

NONBACKTRACKING SPECTRUM OF RANDOM GRAPHS: COMMUNITY DETECTION AND NONREGULAR RAMANUJAN GRAPHS

BY CHARLES BORDENAVE, MARC LELARGE AND LAURENT MASSOULIÉ
CNRS-Université Toulouse III, INRIA-ENS and MSR-INRIA Joint Center

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

CONTENTS

1. Introduction	2
Community detection	3
Ihara zeta function	4
Weak Ramanujan property	4
Organization	5
2. Preliminaries on nonbacktracking matrices	5
2.1. Oriented path symmetry	6
2.2. Chung, Cheeger and Alon–Boppana inequalities for nonbacktracking matrices	7
3. Main results	10
3.1. Erdős–Rényi graphs	10
3.2. Stochastic block model	11
3.3. Notation	14
4. Algebraic tools: Perturbation of eigenvalues and eigenvectors	14
5. Erdős–Rényi graph: Proof strategy for Theorem 3	22
5.1. Proof of Theorem 3	22
5.2. Matrix expansion for B^ℓ	23
5.3. Norm bounds	26
5.4. Proof of Proposition 11	26
6. Proof of Proposition 14: Path count combinatorics	27
6.1. Bound (29) on $\ \Delta^{(k)}\ $	27

Received April 2015; revised August 2016.

MSC2010 subject classifications. Primary 05C80, 60B20; secondary 62M15, 60J85.

Key words and phrases. Random graphs, nonbacktracking matrix, spectral gap, community detection.

6.2. Bound (30) on $\ \Delta^{(k)}\chi\ $	30
6.3. Bound (31) on $\ R_k^{(\ell)}\ $	30
6.4. Bound (32) on $\ B^{(k)}\ $	33
6.5. Bound (33) on $\ S_k^{(\ell)}\ $	34
7. Stochastic block model: Proof of Theorem 4	34
8. Controls on the growth of Poisson multitype branching processes	36
8.1. Two theorems of Kesten and Stigum	37
8.2. Quantitative versions of the Kesten–Stigum theorems	38
8.3. A cross-generation functional	42
8.4. Decorrelation in homogeneous Galton–Watson branching processes	46
9. Local structure of random graphs	47
9.1. Coupling	47
9.2. Geometric growth of linear functions of nonbacktracking walks	51
9.3. Laws of large numbers for local functions	53
9.3.1. Weak laws of large numbers for local functionals: Convergence speed	53
9.3.2. Law of large numbers for specific local functions	55
9.4. Proof of Proposition 19	59
10. Norm of nonbacktracking matrices	62
10.1. Decomposition of B^ℓ	62
10.2. Proof of Proposition 20	63
11. Stochastic block model: Proof of Theorem 5	64
Acknowledgment	69
References	69

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

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

where for any $e = (u, v) \in \vec{E}$, we set $e_1 = u$, $e_2 = v$, $e^{-1} = (v, u)$. This matrix was introduced by Hashimoto [12]. A nonbacktracking walk is a directed path of directed edges of G such that no edge is the inverse of its preceding edge. It is easily seen that for any $k \geq 1$, B_{ef}^k counts the number of nonbacktracking walks of $k + 1$ edges on G starting with e and ending with f .

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

The random graphs we consider are drawn according to the stochastic block model, a generalization of Erdős–Rényi graphs due to Holland et al. [13]. In this model, nodes $v \in V$ are partitioned into r groups. An edge between two nodes u, v

is drawn with probability $W_{\sigma(u),\sigma(v)}/n$, where $\sigma(u) \in [r]$ denotes the group node u belongs to. Thus when the $r \times r$ matrix W is fixed, the expected node degrees remain of order 1 as $n \rightarrow \infty$. We focus moreover on instances where the fraction of nodes in group i converges to a limit $\pi(i)$ as $n \rightarrow \infty$.

An informal statement of our results for eigenvalues is as follows. Let G be drawn according to a stochastic block model with fixed number r of node groups such that all nodes have same fixed expected degree $\alpha > 1$. Let μ_1, \dots, μ_r denote the leading eigenvalues of the expected adjacency matrix $\bar{A} := \mathbb{E}(A)$, ordered so that $\mu_1 = \alpha \geq |\mu_2| \geq \dots \geq |\mu_r|$. Let $r_0 \leq r$ be such that $|\mu_{r_0+1}| \leq \sqrt{\alpha} < |\mu_{r_0}|$. Then the r_0 leading eigenvalues of B are asymptotic to μ_1, \dots, μ_{r_0} , the remaining eigenvalues λ satisfying $|\lambda| \leq (1 + o(1))\sqrt{\alpha}$.

Community detection. Our primary motivation stems from the problem of *community detection*, namely: how to estimate a clustering of graph nodes into groups close to the underlying blocks, based on the observation of such a random graph G ? Decelle et al. [8] conjectured a phase transition phenomenon on detectability, namely: the underlying block structure could be detected if and only if $|\mu_2| > \sqrt{\alpha}$.

In the case of two communities with roughly equal sizes [$\pi(1) = \pi(2) = 1/2$] and symmetric matrix W , the negative part (impossibility of detection when $|\mu_2| \leq \sqrt{\alpha}$) was proven by Mossel et al. [25]. As for the positive part (feasibility of detection when $|\mu_2| > \sqrt{\alpha}$), it was conjectured in [8] that the so-called belief propagation algorithm would succeed. Krzakala et al. [20] then made their so-called “spectral redemption conjecture” according to which detection based on the second eigenvector of the nonbacktracking matrix B would succeed.

Recently, a variant of the spectral redemption conjecture was proven by Masoulié [23]: the spectrum of a matrix counting *self-avoiding paths* in G allows us to detect communities through thresholding of the second eigenvector. More recently and independently of [23], an alternative proof of the positive part of the conjecture in [8] was given by Mossel et al. [24], based on an elaborate construction involving countings of nonbacktracking paths in G .

The two approaches of [23] and [24] to proving the positive part of the conjecture in [8], while both relying ultimately on properties of specific path counts, differ however in the following respects. The method in [23] relies on a clear spectral separation property but its implementation is computationally expensive, as the counts of self-avoiding walks it relies upon take super-linear (though polynomial) time. The method in [24] is computationally efficient as it runs in quasilinear time, but the proof does not establish a spectral separation property. The other two methods conjectured to achieve successful reconstruction, namely belief propagation and analysis of nonbacktracking spectrum, are computationally efficient and they are motivated by a clear intuition as described in the spectral redemption conjecture.

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

Ihara zeta function. Hashimoto introduced the matrix B in the context of the Ihara zeta function [12]. We have the identity

$$\det(I - zB) = \frac{1}{\zeta_G(z)},$$

where ζ_G is the Ihara zeta function of the graph; refer to [16, 19, 29, 30]. It follows that the poles of the Ihara zeta function are the reciprocal of the eigenvalues of B . Our main results have thus consequences on the localization of the poles of the zeta function of random graphs.

Weak Ramanujan property. Our result also has an interpretation from the standpoint of Ramanujan graphs, introduced by Lubotzky et al. [22] (see Murty [26] for a recent survey). These are by definition k -regular graphs such that the second largest modulus of its eigenvalues is at most $2\sqrt{k-1}$. By a result of Alon and Boppana (see [27]) for fixed k , k -regular graphs on n nodes must have their second largest eigenvalue at least $2\sqrt{k-1} - o(1)$ as $n \rightarrow \infty$. Hence, Ramanujan graphs are regular graphs with maximal spectral gap between the first and second eigenvalue moduli. A celebrated result of Friedman [9] states that random k -regular graphs achieve this lower bound with high probability as their number of nodes n goes to infinity.

Lubotzky [21] has proposed an extension of the definition of Ramanujan graphs to the nonregular case. Specifically, G is Ramanujan according to this definition if and only if

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

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

Using the analogy between the Ihara zeta function and the Riemann zeta function, Stark and Terras (see [16]) have defined a graph to satisfy the *graph Riemann hypothesis* if its nonbacktracking matrix B has no eigenvalues λ such that $|\lambda| \in (\sqrt{\rho_B}, \rho_B)$, where ρ_B is the Perron–Frobenius eigenvalue of B . Interestingly, a regular graph G is Ramanujan if and only if it satisfies the graph Riemann hypothesis (see [26] and [16]). Thus, the graph Riemann hypothesis can also be seen as a generalization of the notion of Ramanujan graphs to the nonregular case, phrased in terms of nonbacktracking spectra rather than on spectra of universal covers as in the definition of Lubotzky [21].

Our results imply that for fixed $\alpha > 1$, Erdős–Rényi graphs $\mathcal{G}(n, \alpha/n)$ have an associated nonbacktracking matrix B such that $\rho_B \sim \alpha$ and all its other eigenvalues λ verify $|\lambda| \leq \sqrt{\alpha} + o(1)$ with high probability as $n \rightarrow \infty$. In this sense,

Erdős–Rényi graphs asymptotically satisfy the graph Riemann hypothesis, which itself is a plausible extension of the notion of Ramanujan graphs to the nonregular case. This may be seen as an analogue of Friedman’s theorem [9] in the context of Erdős–Rényi graphs. Similarly, for the stochastic block model, our main result is analogous to recent results on the eigenvalues of random n -lifts of base graphs; see [6, 10]. Interestingly, in [6], the methods developed in the present paper are adapted to lead to a new and simpler proof of Friedman’s theorem and its extensions to random n -lifts. The random graphs studied here will require a more delicate analysis.

Organization. The paper is organized as follows. We start in Section 2 with preliminaries on nonbacktracking matrices. We state our main results in Section 3, namely Theorems 3 and 4 characterizing properties of nonbacktracking spectra of Erdős–Rényi graphs and stochastic block models, respectively, and Theorem 5 establishing their consequence for community detection.

In Section 4, we provide the algebraic ingredients we shall need. Specifically, we establish general bounds on the perturbation of eigenvalues and eigenvectors of not necessarily symmetric matrices that elaborate on the Bauer–Fike theorem.

In Section 5, we give the proof architecture for Theorem 3 on the nonbacktracking spectrum of Erdős–Rényi graphs, and detail a central argument of combinatorial nature, namely a representation of the nonbacktracking matrix B raised to some power ℓ as a sum of products of matrices, elaborating on a technique introduced in [23].

In Section 6, we detail another combinatorial argument needed in our proof, namely we establish bounds on the norms of the various matrices involved in the previous matrix expansion. The latter norm bounds are established by the trace method, adapting arguments due to Füredi and Komlós [11].

Section 7 gives the proof strategy for Theorem 4 on nonbacktracking spectra of stochastic block models.

In Section 8, we establish convergence results on multitype branching processes that extend results of Kesten and Stigum [17, 18]. These are then used in Section 9 to characterize the local structure of the random graphs under study. Specifically, we relate the local statistics of stochastic block models to branching processes via coupling, and then establish weak laws of large numbers on these local statistics. These probabilistic arguments together with the previous algebraic and combinatorial arguments allow us to conclude the proofs of Theorems 3 and 4. Section 11 contains the proof of Theorem 5 on community detection.

2. Preliminaries on nonbacktracking matrices. In this section, we explain how the singular value decomposition of B^ℓ for ℓ large can be used to study the eigenvalues and eigenvectors of B . We also comment on analogues of some classical inequalities known for adjacency or Laplacian matrices of regular graphs.

We set $m = |\vec{E}|$. A priori, B is not a normal matrix. We are interested in its eigenvalues $\lambda_1(B), \dots, \lambda_m(B)$ ordered nonincreasingly, $|\lambda_1(B)| \geq \dots \geq |\lambda_m(B)|$. The Perron–Frobenius theorem implies notably that $\lambda_1(B)$ is a nonnegative real. If G is connected, $\lambda_1(B)$ is equal to the growth rate of the universal cover of G (see Angel, Friedman and Hoory [1]).

2.1. *Oriented path symmetry.* An important remark is that despite B not being a normal matrix, it contains some symmetry. More precisely, we observe that $(B^*)_{ef} = B_{fe} = B_{e^{-1}f^{-1}}$. Introduce for all $x \in \mathbb{R}^{\vec{E}}$ the notation

$$\check{x}(e) = x(e^{-1}), \quad e \in \vec{E}.$$

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

$$(1) \quad \langle y, B^k x \rangle = \langle B^k \check{y}, \check{x} \rangle.$$

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

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

Hence, $B^k P$ is a symmetric matrix (in mathematical physics, this type of symmetry is called PT-invariance, PT standing for parity-time). If $(\sigma_{j,k})$, $1 \leq j \leq m$, are the eigenvalues of $B^k P$ and $(x_{j,k})$, $1 \leq j \leq m$, is an orthonormal basis of eigenvectors, we deduce that

$$(2) \quad B^k = \sum_{j=1}^m \sigma_{j,k} x_{j,k} \check{x}_{j,k}^*.$$

We order the eigenvalues,

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

From the Perron–Frobenius theorem, $x_{1,k}$ can be chosen to have nonnegative entries. Since P is an orthogonal matrix, $(\check{x}_{j,k})$, $1 \leq j \leq m$, is also an orthonormal basis of $\mathbb{R}^{\vec{E}}$. In particular, (2) gives the singular value decomposition of B^k . Indeed, if $s_{j,k} = |\sigma_{j,k}|$ and $y_{j,k} = \text{sign}(\sigma_{j,k}) \check{x}_{j,k}$, we get

$$(3) \quad B^k = \sum_{j=1}^m s_{j,k} x_{j,k} y_{j,k}^*.$$

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

For example, for $k = 1$, it is a simple exercise to compute $(\sigma_{j,1})_{1 \leq j \leq m}$. We find that the eigenvalues of BP are $(\deg(v) - 1)$, $1 \leq v \leq n$, and -1 with multiplicity $m - n$. In particular, the singular values of B contain only information on the degree sequence of the underlying graph G as noted in [20].

For large k however, we may expect that the decomposition (2) carries more structural information on the graph (this will be further discussed in Section 2.2 below). This will be the underlying principle in the proof of our main results. For the moment, we simply note the following. Assume that B is irreducible. From the Perron–Frobenius theorem, if ξ is the Perron eigenvector of B , $\|\xi\| = 1$, then for any n fixed,

$$(4) \quad \lambda_1(B) = \lim_{k \rightarrow \infty} \sigma_{1,k}^{1/k} \quad \text{and} \quad \lim_{k \rightarrow \infty} \|x_{1,k} - \xi\| = 0.$$

A quantitative version of the above limits will be given in the forthcoming Proposition 7. Another consequence of (2) is that, for $i \neq j$, $x_{i,k}$ and $\check{x}_{j,k}$ should be nearly orthogonal if these vectors converge as $k \rightarrow \infty$. Indeed, a heuristic computation gives

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

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

2.2. Chung, Cheeger and Alon–Boppana inequalities for nonbacktracking matrices. The aim of this subsection is to advocate the use of nonbacktracking matrices. Here, we discuss briefly candidate counterparts for irregular graphs of inequalities that are classical in the context of regular graphs. This subsection will not be used in the proof of our main results, it may be skipped.

The diameter bound of Chung [7] gives an upper bound on the diameter of a regular graph in terms of its spectral gap. The following lemma, expressed in terms of the decomposition (3) of B^k , is an analogue.

LEMMA 1. *Let $e = (u, u')$, $f = (v, v') \in \vec{E}$ be such that*

$$x_{1,k}(e)x_{1,k}(f) > s_{2,k}/s_{1,k}.$$

Then the graph distance between u and v is at most $k + 1$.

PROOF. Observe that if $B_{ef}^k > 0$ then u and v are at most at distance $k + 1$. On the other hand from (3),

$$B_{ef}^k - s_{1,k}x_{1,k}(e)x_{1,k}(f) = \sum_{j=2}^m \sigma_{j,k}x_{j,k}(e)x_{j,k}(f)$$

has absolute value at most $s_{2,k}$ from the Cauchy–Schwarz inequality and the orthonormality of the $x_{j,k}$, $1 \leq j \leq m$. Thus, finally, $B_{ef}^k \geq s_{1,k}x_{1,k}(e)x_{1,k}(f) - s_{2,k}$. \square

Cheeger-type inequalities connect the expansion ratio (isoperimetry) of the graph and its spectral gap; for a survey, see [14]. For a subset $X \subset \vec{E}$ of edges, we measure its volume by

$$V_k(X) = \sum_{e \in X} x_{1,k}^2(e).$$

By construction, our volume is normalized with $V_k(\vec{E}) = 1$. We say that $X \subset \vec{E}$ is edge-symmetric if $\check{X} = X$ where $\check{X} = \{e^{-1}, e \in X\}$. For example, the set of edges adjacent to a given subset of vertices is edge-symmetric. If X, Y are edge-symmetric, we define

$$E_k(X, Y) = \sum_{e \in X, f \in Y} x_{1,k}(e)x_{1,k}(f)B_{ef}^k.$$

Since B_{ef}^k is the number of nonbacktracking walks of length $k + 1$ starting with e and ending with f , $E_k(X, Y)$ measures a kind of conductance between X and Y with a proximity range of radius $k + 1$. If $X^c = \vec{E} \setminus X$, the scalar

$$\Sigma_k(X) = \Sigma_k(X^c) = E_k(X, X^c),$$

can be thought of as the outer surface of a set X . The k th expansion ratio of G is then defined as

$$h_k = \min_{X \subset \vec{E}, \check{X} = X} \frac{\Sigma_k(X)}{V_k(X) \wedge V_k(X^c)}.$$

In (2), after reordering the eigenvalues of $B^k P$ as $\sigma_{1,k} \geq \sigma_{2,k} \geq \dots \geq \sigma_{m,k}$, $\sigma_{1,k} - \sigma_{2,k}$ plays the role of the spectral gap in the classical Cheeger inequality. With this new convention, the following lemma is the analogue of the easy part of Cheeger's inequality for graphs.

LEMMA 2.

$$\sigma_{1,k} - \sigma_{2,k} \leq 2h_k.$$

PROOF. The argument is standard. For simplicity, we drop the index k . From the Courant–Fisher min-max theorem, we have

$$\sigma_2 = \max_{\langle x, x_1 \rangle = 0} \frac{\langle x, B^k P x \rangle}{\|x\|^2} = \max_{\langle x, x_1 \rangle = 0} \frac{\langle x, B^k \check{x} \rangle}{\|x\|^2}.$$

Let $X \subset \vec{E}$ be edge-symmetric. We set

$$x(e) = \frac{x_1(e)}{V(X)} \mathbb{1}(e \in X) - \frac{x_1(e)}{V(X^c)} \mathbb{1}(e \in X^c).$$

By construction $\langle x, x_1 \rangle = 0$ and $\|x\|^2 = 1/V(X) + 1/V(X^c)$. Hence,

$$\left(\frac{1}{V(X)} + \frac{1}{V(X^c)} \right) \sigma_2 \geq \langle x, B^k \check{x} \rangle = \sum_{e,f} B_{ef}^k x(e) \check{x}(f).$$

However, using the edge-symmetry of X and X^c ,

$$\begin{aligned} \langle x, B^k \check{x} \rangle &= \frac{1}{V(X)^2} \sum_{e,f \in X} B_{ef}^k x_1(e) \check{x}_1(f) \\ &\quad + \frac{1}{V(X^c)^2} \sum_{e,f \in X^c} B_{ef}^k x_1(e) \check{x}_1(f) - \frac{2\Sigma(X)}{V(X)V(X^c)}. \end{aligned}$$

Also, using the singular value equation $B^k \check{x}_1 = \sigma_1 x_1$,

$$\sum_{e,f \in X} B_{ef}^k x_1(e) \check{x}_1(f) = \sum_{e \in X, f \in \bar{E}} B_{ef}^k x_1(e) \check{x}_1(f) - \Sigma(X) = \sigma_1 V(X) - \Sigma(X),$$

and similarly for X^c . So finally

$$\left(\frac{1}{V(X)} + \frac{1}{V(X^c)} \right) \sigma_2 \geq \left(\frac{1}{V(X)} + \frac{1}{V(X^c)} \right) \sigma_1 - \left(\frac{1}{V(X)} + \frac{1}{V(X^c)} \right)^2 \Sigma(X).$$

Since, for $x, x' > 0$, $1/x + 1/x' \leq 2/(x \wedge x')$, it concludes the proof. \square

The Alon–Boppana theorem gives a lower bound on the second largest eigenvalue of the adjacency matrix of a regular graph (see [14, 27]). We conclude this paragraph with an elementary bound of this type. We introduce for $e \in \bar{E}$,

$$S_k(e) = \|B^k P \delta_e\|_1 \quad \text{where } \delta_e(f) = \mathbb{1}(e = f).$$

In words, $S_k(e)$ is the number of nonbacktracking walks of length $k + 1$ starting with e . As already pointed, if B is irreducible, the Perron eigenvalue is the growth rate of the universal cover of the graph: for any $e \in \bar{E}$,

$$\lambda_1(B) = \lim_{k \rightarrow \infty} s_{1,k}^{1/k} = \lim_{k \rightarrow \infty} S_k(e)^{1/k}.$$

We observe that

$$s_{1,k}^2 + (m-1)s_{2,k}^2 \geq \text{tr}(B^k B^{*k}) \geq \sum_{e \in \bar{E}} S_k(e)$$

where B^* denotes the transpose of B . Hence, we find

$$(5) \quad s_{2,k}^2 \geq \frac{1}{m} \sum_{e \in \bar{E}} S_k(e) - \frac{s_{1,k}^2}{m}.$$

This last crude inequality gives a lower bound on the second largest singular value of B^k .

3. Main results. We now state our results on the nonbacktracking spectra of Erdős–Rényi graphs first, and stochastic block models next.

3.1. *Erdős–Rényi graphs.* Let the vector χ on $\mathbb{R}^{\vec{E}}$ be defined as

$$\chi(e) = 1, \quad e \in \vec{E}.$$

The Euclidean norm of a vector $x \in \mathbb{R}^d$ will be denoted by $\|x\|$. We have the following theorem.

THEOREM 3. *Let G be an Erdős–Rényi graph with parameters $(n, \alpha/n)$ for some fixed parameter $\alpha > 1$. Then, with probability tending to 1 as $n \rightarrow \infty$, the eigenvalues $\lambda_i(B)$ of its nonbacktracking matrix B satisfy*

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

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

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

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

Theorem 3 is illustrated by Figure 1. We conjecture that the lower bound $|\lambda_2(B)| \geq \sqrt{\alpha} - o(1)$ holds, it is reasonable in view of Figure 1. We shall prove a weaker lower bound; see forthcoming Remark 12. It is also an interesting open problem to study the convergence of the empirical distribution of the eigenvalues of B .

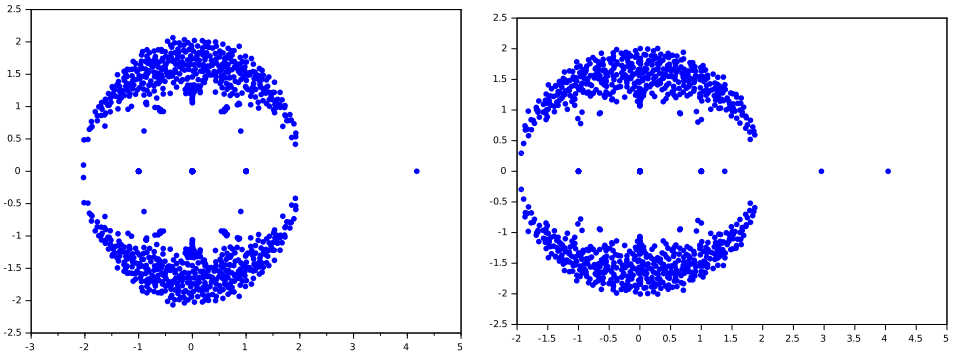


FIG. 1. *Left: eigenvalues of B for a realization of an Erdős–Rényi graph with parameters $(n, \alpha/n)$ with $n = 500$, $\alpha = 4$. Right: eigenvalues of B for Example 2 with $n = 500$, $r = 2$, $a = 7$, $b = 1$.*

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

$$(6) \quad \pi_n(i) := \frac{1}{n} \sum_{v=1}^n \mathbb{1}(\sigma_n(v) = i) = \pi(i) + o(1),$$

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

Given a symmetric matrix $W \in M_r(\mathbb{R}_+)$ we assume that there is an edge between vertices u and v independently with probability

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

We set $\Pi = \text{diag}(\pi(1), \dots, \pi(r))$ and introduce the mean progeny matrix $M = \Pi W$ (the branching process terminology will be clear in Section 8). Note that the eigenvalues of M are the same as the ones of the symmetric matrix $S = \Pi^{1/2} W \Pi^{1/2}$ and in particular are real-valued. They are also the same as the nonzero eigenvalues of the expected adjacency matrix $\bar{A} := \mathbb{E}(A)$ conditioned on the vertex types. We denote them by μ_k and order them by their absolute value,

$$|\mu_r| \leq \dots \leq |\mu_2| \leq \mu_1.$$

We shall make the following assumptions:

$$(7) \quad \mu_1 > 1 \text{ and } M \text{ is positively regular,}$$

that is, for some integer $k \geq 1$, M^k has positive coefficients. In particular, μ_1 is the Perron–Frobenius eigenvalue. It implies notably that for all $i \in [r]$, $\pi(i) > 0$. We define r_0 by

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

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

$$(8) \quad \langle \phi_i, \psi_j \rangle = \delta_{ij} \quad \text{and} \quad \langle \phi_i, \phi_j \rangle_\pi = \delta_{ij},$$

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

$$(9) \quad M = \sum_{k=1}^r \mu_k \psi_k \phi_k^* \quad \text{and} \quad W = \sum_{k=1}^r \mu_k \phi_k \phi_k^*,$$

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

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

$$(10) \quad \alpha = \sum_{i=1}^r \pi(i) W_{ij} = \sum_{i=1}^r M_{ij} \quad \text{for all } j \in [r].$$

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

$$(11) \quad \mu_1 = \alpha > 1, \quad \phi_1 = \mathbb{1} \quad \text{and} \quad \psi_1 = \pi^*.$$

We will also assume that a quantitative version of (6) holds, namely that for some $\gamma \in (0, 1]$,

$$(12) \quad \|\pi - \pi_n\|_\infty = \max_{i \in [r]} |\pi(i) - \pi_n(i)| = O(n^{-\gamma}).$$

The random graph G is usually called the stochastic block model (SBM) or inhomogeneous random graph; see Bollobás, Janson and Riordan [5] and Holland, Laskey and Leinhardt [13]. A popular case is when the map σ is itself random and $\sigma(v)$ are i.i.d. with distribution $(\pi(1), \dots, \pi(r))$. In this case, with probability one, condition (12) is met for any $\gamma < 1/2$.

EXAMPLE 1. If $r = 2$, then we have $\pi(1) = 1 - \pi(2)$. Under condition (10), we have $W_{22} = (\pi(1)W_{11} + (1 - 2\pi(1))W_{12})/(1 - \pi(1))$ so that $\mu_1 = \alpha = \pi(1)W_{11} + (1 - \pi(1))W_{12}$ and $\mu_2 = \pi(1)(W_{11} - W_{12})$.

EXAMPLE 2. If $r \geq 2$, $\pi(i) = 1/r$ and $W_{ii} = a \neq b = W_{ij}$ for all $i \neq j$ so that condition (10) is satisfied. We have $\mu_1 = \alpha = (a + (r - 1)b)/r$ and $\mu_2 = \dots = \mu_r = (a - b)/r$.

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

$$(13) \quad \chi_k(e) = \phi_k(\sigma(e_2)) \quad \text{for all } e \in \vec{E}.$$

In particular, $\chi_1 = \chi$. Our main result is the following generalization of Theorem 3.

THEOREM 4. *Let G be an SBM as above such that hypotheses (7), (10), (12) hold. Then with probability tending to 1 as $n \rightarrow \infty$,*

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

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

$$(14) \quad \frac{B^\ell B^{*\ell} \check{\chi}_k}{\|B^\ell B^{*\ell} \check{\chi}_k\|},$$

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

It follows from this result that a nontrivial estimation of the node types $\sigma(v)$ is feasible on the basis of the eigenvectors $\{\xi_k\}_{2 \leq k \leq r_0}$ provided $r_0 > 1$. More precisely, for vertex type estimators $\hat{\sigma}(v) : [n] \rightarrow [r]$ based on the observed random graph G , following Decelle et al. [8], define the overlap $\text{ov}(\hat{\sigma}, \sigma)$ as the minimum over permutations $p : [r] \rightarrow [r]$ of the quantity

$$(15) \quad \frac{1}{n} \sum_{v=1}^n \mathbb{1}_{\hat{\sigma}(v)=p \circ \sigma(v)} - \max_{k \in [r]} \pi(k).$$

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

As conjectured in [8] and proven in [24], in the setup of Example 2 with $r = 2$, the best possible overlap is $o(1)$ with high probability when $r_0 = 1$, that is, when $\mu_2 \leq \sqrt{\mu_1}$. Conversely, adapting the argument in [23], when $r_0 > 1$, we have the following.

THEOREM 5. *Let G be an SBM as above such that hypotheses (7), (10), (12) hold. Assume further that $\pi(i) \equiv 1/r$, that $r_0 > 1$ and that for some $k \in \{2, \dots, r_0\}$, μ_k is a simple eigenvalue of M . Let $\xi_k \in \mathbb{R}^{\bar{E}}$ be a normalized eigenvector of B associated with $\lambda_k(B)$.*

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

The reason for the existence of the signing $\omega \in \{-1, 1\}^V$ in the above statement is that we do not know a priori whether the vector ξ_k or $-\xi_k$ is asymptotically close to (14). In the simplest case, we will be able to estimate this sign and the vector ω will be equal to $-\mathbb{1}$ or $\mathbb{1}$ and $I^+ = \{i \in [r] : \phi_k(i) > 0\}$, $I^- = [r] \setminus I^+$.

We now explain the main differences with [23]. While [23] constructs a symmetric matrix counting self-avoiding paths of length ℓ of order $\log n$ and studies its spectral decomposition, we analyze the nonbacktracking matrix raised to the power ℓ . The fact that the nonbacktracking matrix is nonnormal complicates significantly the analysis and requires new algebraic ingredients provided in Section 4. We establish general bounds on the perturbation of eigenvalues and eigenvectors of not necessarily symmetric matrices that elaborate on the Bauer–Fike theorem. The proof then requires path counting combinatorial arguments similar

to [23], except that we are dealing here with nonbacktracking paths instead of self-avoiding paths (see Section 6). As in [23], we provide a local analysis of the graph and we make it in a systematic way. In particular, the law of large numbers for local functional given in Section 9.3 might be of independent interest. Moreover, due to the nonnormality of the nonbacktracking matrix, we need to get estimates of cross-generational functionals which were not required in [23].

3.3. Notation. We say that a sequence of events E_n holds *with high probability*, abbreviated w.h.p., if $\lim_{n \rightarrow \infty} \mathbb{P}(E_n) = 1$. The operator norm of a square matrix C is denoted by

$$\|C\| = \sup_{x \neq 0} \frac{\|Cx\|}{\|x\|}.$$

We denote by C^* the transpose of C .

Given a (nonoriented) graph $G = (V, E)$, we denote by $\gamma = (\gamma_0, \dots, \gamma_k)$ a walk of length k where each $\gamma_i \in V$ and $\{\gamma_i, \gamma_{i+1}\} \in E$ for all $i \in \{0, \dots, k-1\}$. We also denote the concatenation of two walks γ and γ' by (γ, γ') . A walk is nonbacktracking if for all $i \in \{0, \dots, k-2\}$, $\gamma_i \neq \gamma_{i+2}$. A walk contains a cycle if there exists $i \neq j$ with $\gamma_i = \gamma_j$.

For convenience, we will extend matrices and vectors from $\mathbb{R}^{\vec{E}}$ with $\vec{E} = \{(u, v) : \{u, v\} \in E\}$ to $\mathbb{R}^{\vec{E}(V)}$ where $\vec{E}(V) = \{(u, v) : u \neq v \in V\}$ is the set of directed edges of the complete graph. The vector χ is defined by $\chi(e) = 1$ for all $e \in \vec{E}(V)$. Recall that for $x \in \mathbb{R}^{\vec{E}(V)}$, we defined

$$\check{x}(e) = x(e^{-1}), \quad e \in \vec{E}(V).$$

4. Algebraic tools: Perturbation of eigenvalues and eigenvectors. One main tool in our analysis is the Bauer–Fike theorem on the perturbation of eigenvalues of matrices. The form given below elaborates on the usual statement of the theorem which in general omits the second half.

THEOREM 6 (Bauer–Fike theorem; see [4], Theorem VI.5.1). *Let D be a diagonalizable matrix such that for some invertible matrix V and diagonal matrix Λ one has $D = V^{-1}\Lambda V$. Let E be a perturbation matrix.*

Then any eigenvalue μ of $D + E$ verifies

$$(16) \quad \min_i |\mu - \lambda_i| \leq \|E\| \cdot \|V\| \cdot \|V^{-1}\|,$$

where λ_i is the i th diagonal entry of Λ .

Denote by R the right-hand side of (16) and $C_i := \mathcal{B}(\lambda_i, R)$ the ball centered at λ_i with radius R . Let \mathcal{I} be a set of indices such that

$$(17) \quad \left(\bigcup_{i \in \mathcal{I}} C_i \right) \cap \left(\bigcup_{i \notin \mathcal{I}} C_i \right) = \emptyset.$$

Then the number of eigenvalues of $D + E$ in $\bigcup_{i \in \mathcal{I}} C_i$ is exactly $|\mathcal{I}|$.

The following proposition on perturbation of rank one matrices will be a basic ingredient to deduce from expressions like (2) quantitative versions of (4). It relies on the stability of eigenvalues and eigenvectors of matrices with a well separated spectrum and which are not too far from being normal (A is normal if and only if $A^*A = AA^*$; see [15]). In this setting, high powers of the matrix can be used to give good estimates on the largest eigenvalues and their eigenvectors.

PROPOSITION 7. *Let $A \in M_n(\mathbb{R})$, $\ell', \ell \geq 1$ be mutually prime integers, $\theta \in \mathbb{R} \setminus \{0\}$, and $c_0, c_1 > 0$ such that for any $k \in \{\ell, \ell'\}$, for some $x_k, y_k \in \mathbb{R}^n$, $R_k \in M_n(\mathbb{R})$,*

$$A^k = \theta^k x_k y_k^* + R_k,$$

with $\langle y_k, x_k \rangle \geq c_0$, $\|x_k\| \|y_k\| \leq c_1$ and

$$\|R_k\| < \frac{c_0^2}{2(\ell \vee \ell') c_1} |\theta|^k.$$

Let (λ_i) , $1 \leq i \leq n$ be the eigenvalues of A , with $|\lambda_n| \leq \dots \leq |\lambda_1|$. Then:

(i) λ_1 has multiplicity one and we have

$$|\lambda_1 - \theta| \leq C|\theta|/\ell \quad \text{and} \quad \text{for } i \geq 2, \quad |\lambda_i| \leq \left(\frac{2c_1}{c_0}\right)^{1/\ell} \|R_\ell\|^{1/\ell},$$

with $C = \pi/2 + 2\sqrt{c_1 \vee 1} \log(2(c_1 \vee c_0^{-1}))$.

(ii) There exists a unit eigenvector ψ of A with eigenvalue λ_1 such that

$$\left\| \psi - \frac{x_\ell}{\|x_\ell\|} \right\| \leq 8c_0^{-1} \|R_\ell\| |\theta|^{-\ell}.$$

The condition on the scalar product $\langle x_k, y_k \rangle$ implies that the left and right eigenvectors of $\theta^k x_k y_k^*$ associated with the eigenvalue θ^k are not orthogonal. The two mutually prime numbers ℓ and ℓ' will be used to lift some possible ambiguities on the phase of the eigenvalue λ_1 (to the best of our knowledge, this idea is new). The Bauer–Fike theorem and its variants are commonly used in polynomial eigenvalues problems; see, for example, [28] and references therein.

PROOF OF PROPOSITION 7. We can assume without loss of generality that $\theta = 1$. We fix $k \in \{\ell, \ell'\}$ and let $\tilde{x} = x_k/\|x_k\|$, $\tilde{y} = y_k/\|y_k\|$, $\sigma = \|x_k\| \|y_k\|$, $\nu = \langle y_k, x_k \rangle$. We have

$$A^k = \sigma \tilde{x} \tilde{y}^* + R_k = S + R_k.$$

The first step of the proof is to apply the Bauer–Fike theorem to A^k to obtain an estimate of λ_1^k . To this end, we write $S = UDU^{-1}$, where $D = \text{diag}(\nu, 0, \dots, 0)$, $U = (\tilde{x}, f_2, \dots, f_n)$ with $f_1 = \tilde{y}$ and $(f_i)_{i \geq 2}$ is an orthogonal basis of \tilde{y}^\perp (we will

soon check that U is indeed invertible). We can also assume that $\tilde{x} \in \text{span}(f_1, f_2)$. Obviously, the eigenvalues of A^k coincide with the eigenvalues of $D + U^{-1}R_kU$. We have to compute the condition number of U :

$$\kappa(U) = \|U\| \|U^{-1}\|.$$

We consider the unitary matrix $V = (f_1, \dots, f_n)$. Let $a, b \in \mathbb{C}$ be such that $\tilde{x} = af_1 + bf_2$. We find

$$V^*U = \begin{pmatrix} W & 0 \\ 0 & I_{n-2} \end{pmatrix} \quad \text{with } W = \begin{pmatrix} a & 0 \\ b & 1 \end{pmatrix}.$$

In particular,

$$\kappa(U) = \kappa(V^*U) = \kappa(W).$$

As $|a|^2 + |b|^2 = 1$ we have

$$W^*W = \begin{pmatrix} 1 & \bar{b} \\ b & 1 \end{pmatrix}.$$

The eigenvalues of WW^* are $1 \pm |b|$. We deduce that

$$\kappa(U) = \sqrt{\frac{1+|b|}{1-|b|}}.$$

Now, by assumption, $|b| = \sqrt{1-|a|^2} = \sqrt{1-|\langle \tilde{x}, \tilde{y} \rangle|^2} \leq \sqrt{1-c_0^2c_1^{-2}} \leq 1 - c_0^2c_1^{-2}/2$. We obtain that

$$\kappa(U) \leq \kappa = 2c_1c_0^{-1}.$$

Notice that by assumption $\ell \vee \ell' \geq 2$ and $2\kappa \|R_k\| < c_0 \leq |v|$. An application of Theorem 6 to $D + U^{-1}R_kU$ then implies that there is a unique eigenvalue of A^k in the ball $\{z \in \mathbb{C} : |v-z| \leq \kappa \|R_k\|\}$ and all the other eigenvalues lie in the disjoint domain $\{z \in \mathbb{C} : |z| \leq \kappa \|R_k\|\}$. Consequently,

$$(18) \quad |\lambda_1^k - v| \leq \kappa \|R_k\| \quad \text{and} \quad \text{for } i \geq 2, \quad |\lambda_i^k| \leq \kappa \|R_k\|.$$

In particular, the eigenvalue λ_1^k has multiplicity one in A^k , and thus λ_1 has multiplicity one in A .

In the second step of the proof, we deduce from what precedes an estimate on λ_1 : we now bound the difference between λ_1 and $\theta = 1$. First, by assumption,

$$(19) \quad c_0 \leq v \leq c_1.$$

From (18), we deduce that $|\lambda_1^k - v| \leq c_0/2$ and hence $c_0/2 \leq |\lambda_1^k| \leq 2c_1$. Since for all $x \in \mathbb{R}$, $|e^x - 1| \leq |x|e^{x+}$, we get

$$|\lambda_1 - 1| \leq |(2c_1)^{1/k} - 1| \vee |(c_0/2)^{1/k} - 1| \leq c_2/k,$$

with $c_2 = \sqrt{(2c_1) \vee 1} \log(2(c_1 \vee c_0^{-1}))$.

We now control the argument $\omega \in (-\pi, \pi]$ of $\lambda_1 = |\lambda_1| \exp(i\omega)$. By assumption, $\nu > c_0 \geq 0$, hence by (18) the real part of λ_1^k is positive. There thus exist an integer $q \in \mathbb{Z}$ and some $\varepsilon \in (-1/2, 1/2)$ such that

$$k\omega = 2q\pi + \varepsilon\pi.$$

Since $|x| \leq (\pi/2)|\sin(x)|$ for $x \in [-\pi/2, \pi/2]$, we obtain from (18) that

$$\kappa \|R_k\| \geq |\lambda_1|^k |\sin(\varepsilon\pi)| \geq |\lambda_1|^k 2|\varepsilon| \geq c_0|\varepsilon|,$$

so that $|\varepsilon| \leq (\kappa/c_0)\|R_k\| = c_1 c_0^{-1} \|R_k\|$. The above holds for $k \in \{\ell, \ell'\}$. Hence, with the notation $\varepsilon = \varepsilon(\ell)$, $\varepsilon' = \varepsilon(\ell')$, $q = q(\ell)$, $q' = q(\ell')$ we have

$$\omega = \frac{2q\pi}{\ell} + \frac{\varepsilon\pi}{\ell} = \frac{2q'\pi}{\ell'} + \frac{\varepsilon'\pi}{\ell'},$$

so that

$$q\ell' - q'\ell = \frac{\varepsilon'\ell - \varepsilon\ell'}{2}.$$

Using the assumption $\|R_k\| < c_0^2/(2c_1(\ell \vee \ell'))$, we find that both $\varepsilon'\ell$ and $\varepsilon\ell'$ have modulus strictly less than 1, and hence so does the right-hand side of the last display. It follows that $q\ell' = q'\ell$. Since ℓ and ℓ' are mutually prime, we deduce that ℓ divides q and ℓ' divides q' , so that modulo 2π , $\omega = \varepsilon\pi/\ell = \varepsilon'\pi/\ell'$. Thus for $k \in \{\ell, \ell'\}$,

$$|\lambda_1 - 1| \leq |\lambda_1| |e^{i\pi\varepsilon/k} - 1| + ||\lambda_1| - 1| \leq \left(1 + \frac{c_2}{k}\right) \frac{\pi|\varepsilon|}{k} + \frac{c_2}{k}.$$

As $\varepsilon < 1/k$, the right-hand side of the above is no larger than $k^{-1}(2c_2 + \pi/2)$ if $k \geq 2$, which must hold for some $k \in \{\ell, \ell'\}$. This completes the proof of the claim of Proposition 7 on eigenvalues.

The final step of the proof gives an estimate on the leading eigenvector using the eigenvector equation of A^k . Consider now a normed eigenvector z of A associated with λ_1 , which for fixed $k \in \{\ell, \ell'\}$ admits the orthogonal decomposition $z = z_0 + z^\perp$, where $z_0 \in \text{span}(\tilde{y}_k, \tilde{x}_k)$. Applying A^k , we obtain

$$\lambda_1^k z = \sigma \tilde{x}_k \tilde{y}_k^* z_0 + R_k z.$$

Projecting onto $\{\tilde{y}_k, \tilde{x}_k\}^\perp$ yields $|\lambda_1|^k \|z^\perp\| \leq \|R_k\|$. Using the bound $|\lambda_1|^k \geq c_0/2$ gives

$$\|z^\perp\| \leq \frac{2}{c_0} \|R_k\|.$$

Projecting onto $\text{span}(\tilde{y}_k, \tilde{x}_k)$ yields

$$\|\lambda_1^k z_0 - \sigma \tilde{x}_k \tilde{y}_k^* z_0\| \leq \|R_k\|.$$

This implies

$$\left\| \frac{z_0}{\|z_0\|} - c\tilde{x}_k \right\| \leq \frac{1}{|\lambda_1|^k \|z_0\|} \|R_k\| \leq \frac{2\|R_k\|}{c_0\|z_0\|},$$

where c is some scalar and we have used $|\lambda_1|^k \geq c_0/2$. We now use the following general inequality:

$$(20) \quad \text{if } \|u\| = \|v\| = 1 \text{ and for some } t \geq 0, \|u - tv\| \leq \varepsilon \text{ then } \|u - v\| \leq 2\varepsilon.$$

We deduce that

$$\left\| \frac{z_0}{\|z_0\|} - \tilde{x}_k \right\| \leq \frac{4\|R_k\|}{c_0\|z_0\|}.$$

From the triangle inequality, $\|z_0\| \geq 1 - \|z^\perp\|$. We then have

$$\|z - \tilde{x}_k\| \leq \|z^\perp\| + \|z_0 - \|z_0\|\tilde{x}_k\| + 1 - \|z_0\| \leq \frac{8\|R_k\|}{c_0}.$$

Finally, since $\lambda = \lambda_1^k$ has multiplicity one in A^k , the eigenspace of λ_1 for A coincides with the eigenspace of λ_1^k for A^k . This completes the proof of Proposition 7. \square

We now provide an extension of Proposition 7 to arbitrary rank tailored to our future needs. For $x = (x_1, \dots, x_n) \in \mathbb{C}^n$, the multiplicity of $z \in \mathbb{C}$ in x is defined as $\sum_i \mathbb{1}(x_i = z)$, the number of coordinates of x equal to z .

PROPOSITION 8. *Let $A \in M_n(\mathbb{R})$, $\ell' < \ell < 2\ell'$ be mutually prime odd integers, $\theta = (\theta_1, \dots, \theta_r) \in (\mathbb{R} \setminus \{0\})^r$ such that for any $k \in \{\ell, \ell'\}$, for some vectors $x_{k,1}, y_{k,1}, \dots, x_{k,r}, y_{k,r} \in \mathbb{R}^n$ and some matrix $R_k \in M_n(\mathbb{R})$,*

$$A^k = \sum_{j=1}^r \theta_j^k x_{k,j} y_{k,j}^* + R_k.$$

Assume there exist $c_0, c_1 > 0$ such that for all $i \neq j \in [r]$, $\langle y_{k,j}, x_{k,j} \rangle \geq c_0$, $\|x_{k,j}\| \|y_{k,j}\| \leq c_1$, $\langle x_{k,j}, y_{k,i} \rangle = \langle x_{k,j}, x_{k,i} \rangle = \langle y_{k,j}, y_{k,i} \rangle = 0$ and

$$\|R_k\| < \left(\frac{c_0(c_0\gamma^k - c_1)_+}{4c_1} \wedge \frac{c_0^2}{2(\ell \vee \ell')c_1} \right) \vartheta^k,$$

where $\vartheta = \min_i |\theta_i|$, $\gamma = \min\{\theta_i/\theta_j : \theta_i > \theta_j > 0 \text{ or } \theta_i < \theta_j < 0\}$ (the minimum over the empty set being $+\infty$). Let (λ_i) , $1 \leq i \leq n$, be the eigenvalues of A with $|\lambda_n| \leq \dots \leq |\lambda_1|$. Then there exists a permutation $\sigma \in S_r$ such that for all $i \in [r]$,

$$|\lambda_i - \theta_{\sigma(i)}| \leq \frac{C|\theta_{\sigma(i)}|}{\ell} \quad \text{and for } i \geq r+1, \quad |\lambda_i| \leq \left(\frac{2c_1}{c_0} \right)^{1/\ell} \|R_\ell\|^{1/\ell},$$

with $C = \pi/2 + 2\sqrt{c_1 \vee 1} \log(2(c_1 \vee c_0^{-1}))$. Moreover, if $\theta_{\sigma(i)}$ has multiplicity one in θ , λ_i is a simple eigenvalue and there exists a unit eigenvector ψ_i of A with eigenvalue λ_i such that

$$\left\| \psi_i - \frac{x_{\ell, \sigma(i)}}{\|x_{\ell, \sigma(i)}\|} \right\| \leq C' \|R_\ell\| |\vartheta|^{-\ell},$$

with $C' = 24c_1 c_0^{-1} / (1 \wedge (c_0 \gamma^k - c_1)_+ \wedge c_0)$.

PROOF. We may assume that $\vartheta = 1$. Fix $k \in \{\ell, \ell'\}$ and let $\tilde{x}_j = x_{k,j} / \|x_{k,j}\|$, $\tilde{y}_j = y_{k,j} / \|y_{k,j}\|$, $\sigma_j = \theta_j^k \|x_{k,j}\| \|y_{k,j}\|$, $v_j = \theta_j^k \langle y_{k,j}, x_{k,j} \rangle = \sigma_j \langle \tilde{y}_{k,j}, \tilde{x}_{k,j} \rangle$. Let $H_j = \text{span}(x_{k,j}, y_{k,j})$. By assumption, the vector spaces H_j , $1 \leq j \leq n$, are orthogonal. For ease of notation, let us assume for all $j \in [r]$, H_j has dimension 2 (the case where $x_{k,j}$ and $y_{k,j}$ are colinear is identical). We consider an orthonormal basis (f_1, \dots, f_n) of \mathbb{C}^n , such that $\text{span}(f_{2j-1}, f_{2j}) = H_j$, $f_{2j-1} = \tilde{y}_j$. We have

$$A^k = \sum_{j=1}^r \sigma_j \tilde{x}_j \tilde{y}_j^* + R_k = U D U^{-1} + R_k,$$

where $D = \text{diag}(v_1, 0, v_2, 0, \dots, v_r, 0, \dots, 0)$, $U = (\tilde{x}_1, f_2, \tilde{x}_2, f_4, \dots, \tilde{x}_r, f_{2r}, \dots, f_n)$ (provided that U is indeed invertible). Arguing as in the proof of Proposition 7, denote by V the unitary matrix $V = (f_1, \dots, f_n)$ and decompose \tilde{x}_j as $\tilde{x}_j = a_j f_{2j-1} + b_j f_{2j}$. Then $V^* U$ has a block diagonal structure with blocks W_j , $1 \leq j \leq r$, and I_{n-2r} , where

$$W_j = \begin{pmatrix} a_j & 0 \\ b_j & 1 \end{pmatrix}.$$

We find, as in Proposition 7,

$$\kappa(U) = \max_j \kappa(W_j) \leq \kappa = 2c_1 c_0^{-1}.$$

Now, by assumption, $2\kappa \|R_k\| < c_0 \wedge (c_0 \gamma^k - c_1)_+$ is less than the minimal distance between the distinct eigenvalues of D . We deduce from Theorem 6 applied to $D + U^{-1} R_k U$ that there is a permutation $s \in S_r$ such that

$$(21) \quad |\lambda_i^k - v_{s(i)}| \leq \kappa \|R_k\| \quad \text{and} \quad \text{for } i \geq r+1, \quad |\lambda_i|^k \leq \kappa \|R_k\|.$$

Importantly, we claim that the permutation $s = s_k \in S_r$ is such that for $1 \leq i \leq r$, $\theta_{s_\ell(i)} = \theta_{s_{\ell'}(i)}$. Indeed, we first observe that the assumptions $\gamma^k c_0 > c_1$ and k odd imply that $v_i = v_j$ is equivalent to $\theta_i = \theta_j$. Assume first for simplicity that all θ_i are positive and let m_1 be the multiplicity in θ of $t_1 = \max_i \theta_i$. Then the m_1 eigenvalues such that $|\lambda_i^k - v_j| \leq \kappa \|R_k\|$ for some j such that $\theta_j = t_1$ are precisely the m_1 largest eigenvalues of A . If $m_1 < r$, we may then repeat the same argument for the second largest value of the set $\{\theta_1, \dots, \theta_r\}$. By iteration, we deduce

the claimed statement when all θ_i have the same sign. In the general case, we notice that if $|\lambda_i^{\ell'} - \nu_j| \leq \varepsilon = \kappa \|R_{\ell'}\|$ with $\theta_j > 0$, then $\nu_j > c_0$, $|\sin(\lambda_i^{\ell'})| \leq \varepsilon/c_0$, $|\arg(\lambda_i^{\ell'})| \leq \pi\varepsilon/(2c_0)$ and $|\arg(\lambda_i^{\ell})| \leq \pi\varepsilon\ell/(2c_0\ell') \leq \pi/2$ (we use here the assumption $\ell < 2\ell'$). It follows that if $|\lambda_i^{\ell'} - \nu_j| \leq \kappa \|R_{\ell'}\|$ with $\theta_j > 0$ then we cannot have $|\lambda_i^{\ell} - \nu_{j'}| \leq \kappa \|R_{\ell}\|$ with $\theta_{j'} < 0$. We may thus repeat the previous argument by considering the largest eigenvalues of A with positive real part and the largest eigenvalues of A with negative real part separately.

We now bound the difference between λ_i and $\theta_{s(i)}$. By assumption, $c_0|\theta_j| \leq |\nu_j| \leq c_1|\theta_j|$. Hence, arguing as in the proof of Proposition 7,

$$||\lambda_i| - |\theta_{s(i)}|| \leq c_2|\theta_{s(i)}|/k,$$

with $c_2 = \sqrt{c_1 \vee 1} \log(2(c_1 \vee c_0^{-1}))$. We next control the argument $\omega_i \in (-\pi, \pi]$ of $\lambda_i = |\lambda_i| \operatorname{sign}(\theta_{\sigma(i)}) e^{i\omega}$. Arguing as in the proof of Proposition 7, we get for $p \in \mathbb{Z}$, $|\omega - 2p\pi| \leq \pi|\varepsilon|/k$ and we may conclude the proof of the claim of Proposition 8 on eigenvalues as in Proposition 7.

It now remains to control the eigenvector of λ_i such that $\theta_{\sigma(i)}$ has multiplicity one. First, from (21), λ_i is a simple eigenvalue of A . Let z be a corresponding normed eigenvector of A . Applying A^k yields

$$(22) \quad \lambda_i^k z = \sum_{j \in [r]} \sigma_j(\tilde{y}_j^* z) \tilde{x}_j + R_k z.$$

Applying A^k once more to (22) yields

$$\lambda_i^{2k} z = \lambda_i^k R_k z + \sum_{j \in [r]} \sigma_j[\sigma_j(\tilde{y}_j^* z)(\tilde{y}_j^* \tilde{x}_j) + \tilde{y}_j^* R_k z] \tilde{x}_j.$$

Multiplying (22) by λ_i^k and subtracting it to the previous display yields

$$\sum_{j \in [r]} \sigma_j[(\lambda_i^k - \nu_j) \tilde{y}_j^* z - \tilde{y}_j^* R_k z] \tilde{x}_j = 0.$$

Thus, for all $j \in [r]$,

$$(23) \quad (\lambda_i^k - \nu_j) \tilde{y}_j^* z - \tilde{y}_j^* R_k z = 0.$$

Now for $j \neq s(i)$, from (21), we have

$$|\lambda_i^k - \nu_j| \geq |\nu_{s(i)} - \nu_j| - |\lambda_i^k - \nu_{s(i)}| \geq \frac{1}{2}((c_0\gamma^k - c_1)_+ \wedge c_0).$$

It follows that

$$\left(\sum_{j \neq s(i)} |\tilde{y}_j^* z|^2 \right)^{1/2} \leq \frac{2\|R_k\|}{(c_0\gamma^k - c_1)_+ \wedge c_0}.$$

Moreover, this implies upon dividing (23) by λ_i^k :

$$\begin{aligned} \left(\sum_{j \neq s(i)} \left| \frac{\sigma_j}{\lambda_i^k} \tilde{y}_j^* z \right|^2 \right)^{1/2} &\leq \frac{c_1}{c_0} \left(\sum_{j \neq s(i)} \left| \frac{v_j}{\lambda_i^k} \tilde{y}_j^* z \right|^2 \right)^{1/2} \\ &\leq \frac{c_1}{c_0} \left(\frac{2\|R_k\|}{(c_0\gamma^k - c_1)_+ \wedge c_0} + \frac{2}{c_0} \|R_k\| \right), \end{aligned}$$

where we have used the fact that $|\lambda_i^k| \geq c_0/2$. It then follows from (22) that for some constant $c = \sigma_{s(i)}(\tilde{y}_{s(i)}^* z)/\lambda_i^k$,

$$\begin{aligned} \|z - c\tilde{x}_{s(i)}\| &= \left\| \sum_{j \neq s(i)} \frac{\sigma_j}{\lambda_i^k} (\tilde{y}_j^* z) \tilde{x}_j + \lambda_i^{-k} R_k z \right\| \\ &\leq \frac{c_1}{c_0} \left(\frac{2\|R_k\|}{(c_0\gamma^k - c_1)_+ \wedge c_0} + \frac{4}{c_0} \|R_k\| \right). \end{aligned}$$

We then obtain the announced bound on $\|z - \tilde{x}_{s(i)}\|$ by appealing to (20). \square

We conclude this paragraph with an elementary lemma on the Gram–Schmidt orthonormalization process. It will be used to obtain vectors which are exactly orthogonal as in the assumptions of Proposition 8.

LEMMA 9. *Let u_1, \dots, u_k be vectors in \mathbb{C}^n with unit norms such that $|\langle u_i, u_j \rangle| \leq \delta$ for all $i \neq j$. If $\delta < k^{-k}$, then (u_1, \dots, u_k) are linearly independent and, if $(\bar{u}_1, \dots, \bar{u}_k)$ is the Gram–Schmidt orthonormalization process of (u_1, \dots, u_k) , we have for all $j \in [k]$,*

$$\|u_j - \bar{u}_j\| \leq \delta j^j.$$

PROOF. We prove the statement by induction. For $k = 1$, $\bar{u}_1 = u_1$. For $k \geq 1$, we denote by v_{k+1} the orthonormal projection of u_{k+1} on the span of (u_1, \dots, u_k) . We have

$$\|v_{k+1}\|^2 = \sum_{j=1}^k |\langle u_{k+1}, \bar{u}_j \rangle|^2.$$

Now, from the induction hypothesis,

$$\sum_{j=1}^k |\langle u_{k+1}, \bar{u}_j \rangle|^2 \leq 2 \sum_{j=1}^k (|\langle u_{k+1}, u_j \rangle|^2 + \|\bar{u}_j - u_j\|^2) \leq 2k\delta^2(1 + k^{2k}).$$

It is easy to check that $\sqrt{2k(1 + k^{2k})} \leq 2^{-1}(k + 1)^{k+1}$ for all $k \geq 1$. In particular, if $\delta < (k + 1)^{-(k+1)}$, $v_{k+1} \neq u_{k+1}$ and then from (20), $\|u_{k+1} - \bar{u}_{k+1}\| \leq 2\|v_{k+1}\| \leq \delta(k + 1)^{k+1}$. \square

5. Erdős–Rényi graph: Proof strategy for Theorem 3. In what follows, we consider a sequence $\ell = \ell(n) \sim \kappa \log_\alpha n$ for some $\kappa \in (0, 1/6)$ as in Theorem 3. Note that the diameter of (the giant component of) an Erdős–Rényi graph with average degree α is larger than $\log_\alpha n$. Hence, for this choice of ℓ , by looking around a typical vertex up to depth ℓ , we are much lower than the diameter but we see some cycles. Indeed, it is clear that we need to take ℓ diverging with n as the local structure of a random graph up to a fixed depth converges to a tree which has very different spectral properties. It turns out that for $\ell = \ell(n) \sim \kappa \log_\alpha n$, we are still able to analyze the local structure of the random graph while we capture the spectral property of B .

5.1. *Proof of Theorem 3.* Let

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

and

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

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

PROPOSITION 10. *For some $c_1, c_0 > 0$, w.h.p.*

$$\langle \zeta, \check{\varphi} \rangle \geq c_0 \quad \text{and} \quad c_0 \alpha^\ell \leq \theta \leq c_1 \alpha^\ell.$$

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

$$\sup_{x: \langle x, \check{\varphi} \rangle = 0, \|x\| = 1} \|B^\ell x\| \leq (\log n)^c \alpha^{\ell/2}.$$

Let us check that the last two Propositions 10 and 11 imply Theorem 3. Let $R = B^\ell - \theta \zeta \check{\varphi}^*$ and $y \in \mathbb{R}^{\bar{E}}$ with $\|y\| = 1$. We write $y = s \check{\varphi} + x$ with $x \in \check{\varphi}^\perp$ and $s \in \mathbb{R}$. We find

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

Hence, Proposition 11 implies that, w.h.p.,

$$(24) \quad \|R\| \leq (\log n)^c \alpha^{\ell/2}.$$

We may now apply Proposition 7. If $\lambda_i = \lambda_i(B)$, we find that, w.h.p.,

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

and the normalized Perron eigenvector ξ of B satisfies, w.h.p.

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

This concludes the proof of Theorem 3.

REMARK 12. We note that from [15], Theorem 3.3.16, we get that in (3)

$$|s_{1,\ell} - \theta| \leq \|R\| \quad \text{and} \quad s_{2,\ell} \leq \|R\|.$$

Hence, from (24) w.h.p.,

$$s_{1,\ell} = O(\alpha^\ell) \quad \text{and} \quad s_{2,\ell} = O((\log n)^c \alpha^{\ell/2}).$$

On the other hand, (5) implies that the above upper bound on $s_{2,\ell}$ is also a lower bound up to the logarithmic factors, more precisely, w.h.p., $s_{2,\ell} \geq c_0 \alpha^{\ell/2}$ for some $c_0 > 0$ (it follows from the proof of the forthcoming Theorem 37). Therefore, the naive lower bound on $s_{2,\ell}$ in (5) is asymptotically tight and Propositions 10–11 may be interpreted as a weak Ramanujan property for Erdős–Rényi random graphs.

Proposition 10 will follow from a *local analysis*. Namely, the statistics of node neighborhoods up to distance ℓ in the original random graph will be related by coupling to a Galton–Watson branching process; relevant properties of the corresponding Galton–Watson process will be established; finally, we shall deduce weak laws of large numbers for the ℓ -neighborhoods of the random graph from the estimations performed on the branching process combined with some asymptotic decorrelation property between distinct node neighborhoods. This is done in Section 9 where Proposition 19, which contains Proposition 10, is proven.

The proof of Proposition 11 relies crucially on a matrix expansion given in Proposition 13, which extends the argument introduced in [23] for matrices counting self-avoiding walks to the present setup where nonbacktracking walks instead are considered. We now introduce some notation to state it.

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

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

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

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

We associate to each walk $\gamma = (\gamma_0, \dots, \gamma_k)$, a graph $G(\gamma) = (V(\gamma), E(\gamma))$ with vertex set $V(\gamma) = \{\gamma_i, 0 \leq i \leq k\}$ and edge set $E(\gamma)$ the set of distinct visited edges $\{\gamma_i, \gamma_{i+1}\}, 0 \leq i \leq k-1$. Following [24], we say that a graph H is *tangle-free* (or ℓ -tangle-free to make the dependence in ℓ explicit) if every neighborhood

of radius ℓ in H contains at most one cycle. Otherwise, H is said to be tangled. We say that γ is tangle-free or tangled if $G(\gamma)$ is. Obviously, if G is tangle-free and $1 \leq k \leq \ell$ then $B^k = B^{(k)}$, where

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

and F_{ef}^{k+1} is the subset of tangle-free paths in Γ_{ef}^{k+1} . For $u \neq v$, we set

$$\underline{A}_{uv} = A_{uv} - \frac{\alpha}{n}.$$

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

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

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

$$(25) \quad \Delta_{ef}^{(0)} = \mathbb{1}(e = f) \underline{A}_e \quad \text{and} \quad B_{ef}^{(0)} = \mathbb{1}(e = f) A_e.$$

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

$$(26) \quad B_{ef}^{(\ell)} = \Delta_{ef}^{(\ell)} + \sum_{t=0}^{\ell} \sum_{\gamma \in F_{ef}^{\ell+1}} \prod_{s=0}^{t-1} \underline{A}_{\gamma_s \gamma_{s+1}} \left(\frac{\alpha}{n}\right) \prod_{s=t+1}^{\ell} A_{\gamma_s \gamma_{s+1}}.$$

Indeed,

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

For $0 \leq t \leq \ell$, we define $R_t^{(\ell)}$ via

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

where for $1 \leq t \leq \ell - 1$, $F_{t,ef}^{\ell+1} \subset \Gamma_{ef}^{\ell+1}$ is the set of nonbacktracking tangled paths $\gamma = (\gamma_0, \dots, \gamma_{\ell+1}) = (\gamma', \gamma'') \in \Gamma_{ef}^{\ell+1}$ with $\gamma' = (\gamma_0, \dots, \gamma_t) \in F_{eg}^t$, $\gamma'' = (\gamma_{t+1}, \dots, \gamma_{\ell+1}) \in F_{g'f}^{\ell-t}$ for some $g, g' \in \vec{E}(V)$. Note that since γ is nonbacktracking, we have $g \neq g'$. For $t = 0$, $F_{0,ef}^{\ell+1}$ is the set of nonbacktracking

tangled paths $\gamma = (\gamma', \gamma'')$ with $\gamma' = e_1$, $\gamma'' = (\gamma_1, \dots, \gamma_{\ell+1}) \in F_{g'f}^\ell$ for some $g' \in \bar{E}(V)$ (necessarily $\gamma_1 = g'_1 = e_2$ and $\gamma_2 = g'_2 \neq e_1$). Similarly, for $t = \ell$, $F_{\ell,ef}^{\ell+1}$ is the set of nonbacktracking tangled paths $\gamma = (\gamma_0, \dots, \gamma_{\ell+1}) = (\gamma', \gamma'')$ with $\gamma'' = f_2$, $\gamma' = (\gamma_0, \dots, \gamma_\ell) \in F_{eg}^\ell$ for some $g \in \bar{E}(V)$ (necessarily $\gamma_\ell = g_2 = f_1$ and $\gamma_{\ell-1} = g_1 \neq f_2$).

We denote by K the nonbacktracking matrix of the complete graph on V . We define

$$L = K^2 - \chi\chi^*$$

(L is nearly the orthogonal projection of K^2 on χ^\perp). We further denote for $1 \leq t \leq \ell - 1$

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

We then have the following.

PROPOSITION 13. *With the above notation, matrix $B^{(\ell)}$ admits the following expansion:*

$$(27) \quad \begin{aligned} B^{(\ell)} &= \Delta^{(\ell)} + \frac{\alpha}{n} K B^{(\ell-1)} + \frac{\alpha}{n} \sum_{t=1}^{\ell-1} \Delta^{(t-1)} K^2 B^{(\ell-t-1)} \\ &\quad + \frac{\alpha}{n} \Delta^{(\ell-1)} K - \frac{\alpha}{n} \sum_{t=0}^{\ell} R_t^{(\ell)}. \end{aligned}$$

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

$$(28) \quad \begin{aligned} \|B^\ell x\| &\leq \|\Delta^{(\ell)}\| + \frac{\alpha}{n} \|K B^{(\ell-1)}\| + \frac{\alpha}{n} \sum_{t=1}^{\ell-1} \|\Delta^{(t-1)} \chi\| \|\langle \chi, B^{(\ell-t-1)} x \rangle\| \\ &\quad + \frac{\alpha}{n} \sum_{t=1}^{\ell-1} \|S_t^{(\ell)}\| + \alpha \|\Delta^{(\ell-1)}\| + \frac{\alpha}{n} \sum_{t=0}^{\ell} \|R_t^{(\ell)}\|. \end{aligned}$$

PROOF. Equation (27) readily follows by adding and subtracting $\frac{\alpha}{n} R_t^{(\ell)}$ to the t th term of the summation in (26) and noticing that this term plus $\frac{\alpha}{n} R_t^{(\ell)}$ factorizes into a matrix product. More precisely, for $1 \leq t \leq \ell - 1$, we have

$$\begin{aligned} (\Delta^{(t-1)} K^2 B^{(\ell-t-1)})_{ef} &= \sum_{g \neq g' \in \bar{E}(V)} \left(\sum_{\gamma \in F_{eg}^t} \prod_{s=0}^{t-1} A_{\gamma_s \gamma_{s+1}} \right) \left(\sum_{\gamma \in F_{g'f}^{\ell-t}} \prod_{s=0}^{\ell-t-1} A_{\gamma_s \gamma_{s+1}} \right) \\ &= (R_t^{(\ell)})_{ef}, \end{aligned}$$

hence we have $\frac{\alpha}{n} \sum_{t=1}^{\ell-1} \Delta^{(t-1)} K^2 B^{(\ell-t-1)} = \frac{\alpha}{n} \sum_{t=1}^{\ell-1} R_t^{(\ell)}$. Similarly, we have $R_0^{(\ell)} = K B^{(\ell-1)}$ and $R_\ell^{(\ell)} = \Delta^{(\ell-1)} K$.

Inequality (28) follows from (27) by noting that $B^\ell = B^{(\ell)}$ as G is tangle free, decomposing K^2 into $L + \chi \chi^*$, and finally using the fact that $\|K\| \leq n$. \square

5.3. *Norm bounds.* The following proposition will be established in Section 6 using path counting combinatorial arguments.

PROPOSITION 14. *Let $\ell \sim \kappa \log_\alpha n$ with $\kappa \in (0, 1/6)$. With high probability, the following norm bounds hold for all $k, 0 \leq k \leq \ell$:*

$$(29) \quad \|\Delta^{(k)}\| \leq (\log n)^{10} \alpha^{k/2},$$

$$(30) \quad \|\Delta^{(k)} \chi\| \leq (\log n)^5 \alpha^{k/2} \sqrt{n},$$

$$(31) \quad \|R_k^{(\ell)}\| \leq (\log n)^{25} \alpha^{\ell-k/2},$$

$$(32) \quad \|B^{(k)}\| \leq (\log n)^{10} \alpha^k \quad \text{and} \quad \|K B^{(k)}\| \leq \sqrt{n} (\log n)^{10} \alpha^k,$$

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

$$(33) \quad \|S_k^{(\ell)}\| \leq \sqrt{n} (\log n)^{20} \alpha^{\ell-k/2}.$$

5.4. *Proof of Proposition 11.* Together with Propositions 13 and 14, we shall also need the next two results, established by local analysis in Section 9. In particular, the forthcoming Lemma 30 implies the following.

LEMMA 15. *For $\ell \sim \kappa \log_\alpha n$ with $\kappa < 1/2$, w.h.p. the random graph G is ℓ -tangle-free.*

For the Erdős–Rényi graph, Corollary 34 states the following.

PROPOSITION 16. *For $\ell \sim \kappa \log_\alpha n$ with $\kappa < 1/2$, w.h.p., for any $0 \leq t \leq \ell - 1$, it holds that*

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

We now have all the ingredients necessary to prove Proposition 11. In view of Lemma 15, we may use the bound (28) of Proposition 13 and take the supremum over of all $x, \|x\| = 1, \langle \check{\varphi}, x \rangle = \langle \chi, B^\ell x \rangle = 0$. By the norm bounds (29)–(31)–(33) of Proposition 14, w.h.p.,

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

Also, from (1), since $\check{\chi} = \chi$,

$$\begin{aligned} \sup_{\|x\|=1, \langle \chi, B^\ell x \rangle = 0} |\langle \chi, B^{(t)} x \rangle| &= \sup_{\|x\|=1, \langle \chi, B^\ell \check{x} \rangle = 0} |\langle \chi, B^{(t)} \check{x} \rangle| \\ &= \sup_{\|x\|=1, \langle B^\ell \chi, x \rangle = 0} |\langle B^{(t)} \chi, x \rangle|. \end{aligned}$$

Hence, from Proposition 16 and norm bound (30), w.h.p.,

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

Hence, w.h.p.,

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

It remains to use norm bound (32) to deal with the term $\|K B^{(\ell-1)}\|/n$ in (28) to conclude the proof of Proposition 11.

6. Proof of Proposition 14: Path count combinatorics. In this section, we use the method of moments to prove the norm upper bounds stated in Proposition 14. Recall that $\ell \sim \kappa \log_\alpha n$ with $\kappa > 0$. All our constants will depend implicitly on κ . We will use a version of the moment method by taking the m th moments with m of the order $\log n / \log \log n$ and then applying Markov's inequality. The particular choice of $\log n / \log \log n$ is mainly technical as we will have error terms of the order $(\ell m)^m$ and we want to bound them by a power of n .

6.1. *Bound (29) on $\|\Delta^{(k)}\|$.* The proof will use a version of the trace method. For $n \geq 3$, we set

$$(34) \quad m = \left\lfloor \frac{\log n}{13 \log(\log n)} \right\rfloor.$$

The symmetry (1) implies that $\Delta_{ef}^{(k)} = \Delta_{f^{-1}e^{-1}}^{(k)}$. With the convention that $e(2m+1) = e(1)$, we get

$$\begin{aligned} \|\Delta^{(k-1)}\|^{2m} &= \|\Delta^{(k-1)} \Delta^{(k-1)*}\|^m \leq \text{tr} \{ (\Delta^{(k-1)} \Delta^{(k-1)*})^m \} \\ &= \sum_{e(1), \dots, e(2m)} \prod_{i=1}^m (\Delta^{(k-1)})_{e(2i-1), e(2i)} (\Delta^{(k-1)})_{e(2i+1), e(2i)} \\ (35) \quad &= \sum_{e(1), \dots, e(2m)} \prod_{i=1}^m (\Delta^{(k-1)})_{e(2i-1), e(2i)} (\Delta^{(k-1)})_{e^{-1}(2i)e^{-1}(2i+1)} \\ &= \sum_{\gamma \in W_{k,m}} \prod_{i=1}^{2m} \prod_{s=1}^k \underline{A}_{\gamma_{i,s-1} \gamma_{i,s}}, \end{aligned}$$

where $W_{k,m}$ is the set of sequences of paths $\gamma = (\gamma_1, \dots, \gamma_{2m})$ such that $\gamma_i = (\gamma_{i,0}, \dots, \gamma_{i,k}) \in V^{k+1}$ is nonbacktracking tangle-free of length k and for all $i = 1, \dots, 2m$,

$$(\gamma_{i,k-1}, \gamma_{i,k}) = (\gamma_{i+1,1}, \gamma_{i+1,0}),$$

with the convention that $\gamma_{2m+1} = \gamma_1$.

We take expectations in (35) and use independence of the edges A_{xy} together with $\mathbb{E}A_{xy} = 0$. We find

$$(36) \quad \mathbb{E} \|\Delta^{(k-1)}\|^{2m} \leq \sum_{\gamma \in W'_{k,m}} \mathbb{E} \prod_{i=1}^{2m} \prod_{s=1}^k \underline{A}_{\gamma_{i,s-1}\gamma_{i,s}},$$

where $W'_{k,m}$ is the subset of $W_{k,m}$ where each nonoriented edge is visited at least twice. For each $\gamma \in W_{k,m}$ we associate the graph $G(\gamma) = (V(\gamma), E(\gamma))$ of visited vertices and edges. We set

$$v(\gamma) = |V(\gamma)| \quad \text{and} \quad e(\gamma) = |E(\gamma)|.$$

We say that a path γ is canonical if $V(\gamma) = \{1, \dots, v(\gamma)\}$ and the vertices are first visited in order. $\mathcal{W}_{k,m}$ will denote the set of canonical paths in $W_{k,m}$. Each canonical path is isomorphic to $\binom{n}{v(\gamma)} v(\gamma)!$ paths in $W_{k,m}$. We also have the following.

LEMMA 17 (Enumeration of canonical paths). *Let $\mathcal{W}_{k,m}(v, e)$ be the set of canonical paths with $v(\gamma) = v$ and $e(\gamma) = e$. We have*

$$|\mathcal{W}_{k,m}(v, e)| \leq k^{2m} (2km)^{6m(e-v+1)}.$$

PROOF. In order to upper bound $|\mathcal{W}_{k,m}(v, e)|$, we need to find an injective way to encode the canonical paths $x \in \mathcal{W}_{k,m}(v, e)$.

Let $x = (x_{i,t})_{1 \leq i \leq 2m, 0 \leq t \leq k} \in \mathcal{W}_{k,m}(v, e)$. We set $y_{i,t} = \{x_{i,t}, x_{i,t+1}\}$, $y_{i,t}$ will be called an edge of x . We explore the sequence $(x_{i,t})$ in lexicographic order denoted by \preceq [i.e., $(i, t) \preceq (i+1, t')$ and $(i, t) \preceq (i, t+1)$]. We think of the index (i, t) as a time. For $0 \leq t \leq k-1$, we say that (i, t) is a *first time*, if $x_{i,t+1}$ has not been seen before [i.e., $x_{i,t+1} \neq x_{i',t'}$ for all $(i', t') \preceq (i, t)$]. If (i, t) is a first time, the edge $y_{i,t}$ is called a *tree edge*. By construction, the set of tree edges forms a tree T with vertex set $\{1, \dots, v\}$. The edges which are not in T are called the *excess edges* of x . Since T has $v-1$ edges, it follows that the cardinal of excess edges is $\epsilon = e - v + 1$.

We build a first encoding of $\mathcal{W}_{k,m}(v, e)$. If (i, t) is not a first time, we say that (i, t) is an *important time* and we mark the time (i, t) by the vector $(x_{i,t+1}, x_{i,\tau})$, where (i, τ) is the next time that $y_{i,\tau}$ will not be a tree edge (by convention $\tau = k$ if $x_{i,s}$ remains on the tree for all $t+1 \leq s \leq k$). Since there is a unique nonbacktracking path between two vertices of a tree, we can reconstruct $x \in \mathcal{W}_{k,m}$ from the position of the important times and their mark. It gives rise to our first encoding.

The main issue with this encoding is that the number of important times could be large. We have however not used so far the hypothesis that each path x_i is tangle free. To this end, we are going to partition important times into three categories, *short cycling*, *long cycling* and *superfluous* times. First, consider the case where the i th path x_i contains a cycle. In this case, the first time (i, t) such that $x_{i,t+1} \in \{x_{i,0}, \dots, x_{i,t}\}$ is called a short cycling time. Let $0 \leq \sigma \leq t$ be such that $x_{i,t+1} = x_{i,\sigma}$. By the assumption of tangle-freeness, $C := (x_{i,\sigma}, \dots, x_{i,t+1})$ is the only cycle visited by x_i . We denote by (i, τ) the first time after (i, t) that $y_{i,\tau}$ is not an edge of C (by convention $\tau = k$ if x_i remains on C). We add the extra mark τ to the short cycling time. Important times (i, t) with $1 \leq t < \sigma$ or $\tau < t \leq k$ are called long cycling times. The other important times are called superfluous. The key observation is that if x_i contains a cycle, the number of long cycling times (i, t) is bounded by $\epsilon - 1$ (since there is at most one cycle, no edge of x can be seen twice outside those of C , the -1 coming from the fact that the short cycling time is an excess edge). Now consider the case where the i th path does not contain a cycle, then all important times are called long cycling times and their number is bounded by ϵ .

We now have our second encoding. We can reconstruct x from the positions of the long cycling and the short cycling times and their marks. For each $1 \leq i \leq 2m$, there are at most 1 short cycling time and $\epsilon - 1$ long cycling times within x_i if x_i contains a cycle and 0 short cycling time and ϵ long cycling times if x_i does not contain a cycle. There are at most $k^{2m\epsilon}$ ways to position them (in time). There are at most v^2 different possible marks for a long cycling time and $v^2 k$ possible marks for a short cycling time. We deduce that

$$|\mathcal{W}_{k,m}(v, e)| \leq k^{2m\epsilon} (v^2 k)^{2m} (v^2)^{2m(\epsilon-1)}.$$

We use $v \leq 2km$ to obtain the announced bound. \square

PROOF OF PROPOSITION 14, NORM BOUND (29). From (36) and Markov's inequality, it suffices to prove that

$$(37) \quad S = \sum_{\gamma \in W'_{k,m}} \mathbb{E} \prod_{i=1}^{2m} \prod_{s=1}^k \underline{A}_{\gamma_{i,s-1}\gamma_{i,s}} \leq (C \log n)^{16m} \alpha^{km}.$$

Observe that if $\gamma \in W'_{k,m}$, $v(\gamma) - 1 \leq e(\gamma) \leq km$ because each edge is visited at least twice and $v(\gamma) \geq 3$. As $\binom{n}{v(\gamma)} v(\gamma)! < n^{v(\gamma)}$, any $\gamma \in \mathcal{W}_{k,m}$ is isomorphic to fewer than $n^{v(\gamma)}$ elements in $W_{k,m}$. Also, from the independence of the edges and $\mathbb{E} \underline{A}_{xy}^p \leq \alpha/n$ for integer $p \geq 2$, we get that

$$(38) \quad \mathbb{E} \prod_{i=1}^{2m} \prod_{s=1}^k \underline{A}_{\gamma_{i,s-1}\gamma_{i,s}} \leq \left(\frac{\alpha}{n}\right)^{e(\gamma)}.$$

Hence, using Lemma 17, we obtain for $k \leq \ell$ (remember that $\alpha > 1$ so that $\alpha^e \leq \alpha^{km}$)

$$\begin{aligned}
 (39) \quad S &\leq \sum_{v=3}^{km+1} \sum_{e=v-1}^{km} |\mathcal{W}_{k,m}(v, e)| \left(\frac{\alpha}{n}\right)^e n^v \\
 &\leq n\alpha^{km} \sum_{v=3}^{km+1} \sum_{e=v-1}^{km} k^{2m} (2km)^{6m(e-v+1)} n^{v-e-1} \\
 &\leq n\alpha^{km} \ell^{2m} (\ell m) \sum_{s=0}^{\infty} \left(\frac{(2\ell m)^{6m}}{n}\right)^s.
 \end{aligned}$$

For our choice of m in (34), we have, for n large enough,

$$n^{1/(2m)} = o(\log n)^7, \quad \ell m = o(\log n)^2 \quad \text{and} \quad (2\ell m)^{6m} \leq n^{12/13}.$$

In particular, the above geometric series converges and (37) follows. \square

6.2. Bound (30) on $\|\Delta^{(k)}\chi\|$.

PROOF. The bound (30) on $\|\Delta^{(k)}\chi\|$ that we will now establish improves, by a factor \sqrt{n} on the trivial estimate $\|\Delta^{(k)}\chi\| \leq \|\chi\| \|\Delta^{(k)}\|$. Its proof parallels the argument used to show (29). We have

$$\begin{aligned}
 \mathbb{E}\|\Delta^{(k-1)}\chi\|^2 &= \mathbb{E} \sum_{e,f,g} \Delta_{ef}^{(k-1)} \Delta_{eg}^{(k-1)} \\
 &= \sum_{\gamma \in W''_{k,1}} \mathbb{E} \prod_{i=1}^2 \prod_{s=1}^k \mathbb{A}_{\gamma_{i,s-1}, \gamma_{i,s}},
 \end{aligned}$$

where $W''_{k,1}$ is the set of pairs of paths (γ_1, γ_2) such $\gamma_i = (\gamma_{i,0}, \dots, \gamma_{i,k})$ is nonbacktracking and $(\gamma_{1,k-1}, \gamma_{1,k}) = (\gamma_{2,1}, \gamma_{2,0})$ and each edge is visited at least twice. The only difference with $W'_{k,1}$ defined above is that we do not require that $(\gamma_{1,0}, \gamma_{1,1}) = (\gamma_{2,k}, \gamma_{2,k-1})$. However, this last condition $(\gamma_{1,0}, \gamma_{1,1}) = (\gamma_{2m,k}, \gamma_{2m,k-1})$ was not used in the proof of Lemma 17. It follows that the set of canonical paths in $W''_{k,1}$ with v distinct vertices and e distinct edges has cardinal bounded by $k^2(2k)^{6(e-v+1)}$. Since the paths are connected and each edge appears at least twice, we have $v-1 \leq e \leq k$. As in the proof of (29), we get from (39) with $m=1$

$$\mathbb{E}\|\Delta^{(k-1)}\chi\|^2 \leq Cn\alpha^k (\log n)^3.$$

We conclude with Markov's inequality and the union bound. \square

6.3. Bound (31) on $\|R_k^{(\ell)}\|$. For $n \geq 3$, we set

$$(40) \quad m = \left\lfloor \frac{\log n}{25 \log(\log n)} \right\rfloor.$$

For $0 \leq k \leq \ell - 1$, we have that

$$(41) \quad \begin{aligned} \|R_k^{(\ell-1)}\|^{2m} &\leq \text{tr} \left\{ (R_k^{(\ell-1)} R_k^{(\ell-1)*})^m \right\} \\ &= \sum_{\gamma \in T'_{\ell,m,k}} \prod_{i=1}^{2m} \prod_{s=1}^k \underline{A}_{\gamma_{i,s-1}\gamma_{i,s}} \prod_{s=k+2}^{\ell} A_{\gamma_{i,s-1}\gamma_{i,s}}, \end{aligned}$$

where $T'_{\ell,m,k}$ is the set of sequences of paths $\gamma = (\gamma_1, \dots, \gamma_{2m})$ such that $\gamma_i^1 = (\gamma_{i,0}, \dots, \gamma_{i,k})$ and $\gamma_i^2 = (\gamma_{i,k+1}, \dots, \gamma_{i,\ell})$ are nonbacktracking tangle-free, $\gamma_i = (\gamma_i^1, \gamma_i^2)$ is nonbacktracking tangled and for all odd $i \in \{1, \dots, 2m\}$,

$$(\gamma_{i,0}, \gamma_{i,1}) = (\gamma_{i-1,0}, \gamma_{i-1,1}) \quad \text{and} \quad (\gamma_{i,\ell-1}, \gamma_{i,\ell}) = (\gamma_{i+1,\ell-1}, \gamma_{i+1,\ell}),$$

with the convention that $\gamma_0 = \gamma_{2m}$.

We define $\tilde{G}(\gamma) = (V(\gamma), \tilde{E}(\gamma))$ as the union of the graph $G(\gamma_i^z)$, $1 \leq i \leq 2m$, $z \in \{1, 2\}$. Note that the edges $(\gamma_{i,k}, \gamma_{i,k+1})$ are not taken into account in $\tilde{G}(\gamma)$. As usual, we set $v(\gamma) = |V(\gamma)|$ and $\tilde{e}(\gamma) = |\tilde{E}(\gamma)| \geq v(\gamma)$. Since γ_i is tangled either (a) $G(\gamma_i^1) \cup G(\gamma_i^2)$ contains a cycle and is connected or (b) both $G(\gamma_i^1)$ and $G(\gamma_i^2)$ contain a cycle. In particular, all connected components of $\tilde{G}(\gamma)$ contain a cycle and it follows that

$$v(\gamma) \leq \tilde{e}(\gamma).$$

Taking the expectation in (41), we find that

$$(42) \quad \mathbb{E} \|R_k^{(\ell-1)}\|^{2m} \leq \sum_{\gamma \in T_{\ell,m,k}} \mathbb{E} \prod_{i=1}^{2m} \prod_{s=1}^k \underline{A}_{\gamma_{i,s-1}\gamma_{i,s}} \prod_{s=k+2}^{\ell} A_{\gamma_{i,s-1}\gamma_{i,s}},$$

where $T_{\ell,m,k}$ is the subset of $\gamma \in T'_{\ell,m,k}$ such that

$$(43) \quad v(\gamma) \leq \tilde{e}(\gamma) \leq km + 2m(\ell - 1 - k) = m(2\ell - 2 - k).$$

Indeed, for the contribution of a given γ in (42) to be nonzero, each pair $\{\gamma_{i,s-1}, \gamma_{i,s}\}$, $1 \leq i \leq 2m$, $1 \leq s \leq k$, should appear at least twice in the sequence of the $2(\ell - 1)m$ pairs $\{\gamma_{i,s-1}, \gamma_{i,s}\}$, $s \neq k + 1$.

LEMMA 18 (Enumeration of canonical tangled paths). *Let $\mathcal{T}_{\ell,m,k}(v, e)$ be the set of canonical paths in $T_{\ell,m,k}$ with $v(\gamma) = v$ and $\tilde{e}(\gamma) = e$. We have*

$$|\mathcal{T}_{\ell,m,k}(v, e)| \leq (4\ell m)^{12m(e-v+1)+8m}.$$

PROOF. We will adapt the proof of Lemma 17 and use the same terminology. We start by reordering $\gamma \in \mathcal{T}_{\ell,m,k}$ into a new sequence which preserves as much as possible the connectivity of the path. First, we reorder $\gamma = (\gamma_1, \dots, \gamma_{2m})$ into $\hat{\gamma} = (\hat{\gamma}_1, \dots, \hat{\gamma}_{2m})$ by setting for i odd, $\hat{\gamma}_i = \gamma_i$ and for i even, $\hat{\gamma}_{i,t} = \gamma_{i,\ell-t}$. Also, for i odd, we set $k_i = k$ and for i even $k_i = \ell - k - 1$. Finally, we write $\hat{\gamma}_i =$

$(\hat{\gamma}'_i, \hat{\gamma}''_i)$ with $\hat{\gamma}'_i = (\hat{\gamma}_{i,0}, \dots, \hat{\gamma}_{i,k_i})$ and $\hat{\gamma}''_i = (\hat{\gamma}_{i,k_i+1}, \dots, \hat{\gamma}_{i,\ell})$. To each i , we say that γ_i is connected or disconnected whether $G(\hat{\gamma}'_i)$ intersects the graph $H_i = \bigcup_{j < i} G(\hat{\gamma}'_j) \cup G(\hat{\gamma}''_i)$ or not. If γ_i is disconnected, we define for $0 \leq t \leq \ell$, $x_{i,t} = \hat{\gamma}_{i,t}$. If γ_i is connected, for $0 \leq t \leq k_i$, we set $x_{i,t} = \hat{\gamma}_{i,t}$, and if $q_i > k_i$ is the first time such that $\hat{\gamma}_{i,q_i} \in H_i$, we set for $k_i + 1 \leq t \leq q_i$, $x_{i,t} = \hat{\gamma}_{i,q_i+k_i+1-t}$ and for $q_i + 1 \leq t \leq \ell$, $x_{i,t} = \hat{\gamma}_{i,t}$. We then explore the sequence $(x_{i,t})$ in lexicographic order and set $y_{i,t} = \{x_{i,t}, x_{i,t+1}\}$. The definition of first time, tree edge and excess edge carry over, that is, $(i, t) \neq (i, k_i)$ is a first time if the end vertex of $y_{i,t}$, that is, $x_{i,t+1}$ has not been seen before. When γ_i is connected, we add the extra mark $(q_i, \hat{\gamma}_{i,q_i})$, if γ_i is disconnected this extra mark is set to 0. With our ordering, all vertices of $V(\gamma) \setminus \{1\}$ will have an associated tree edge, at the exception of x_{i,k_i+1} when γ_i is disconnected. If δ is the number of disconnected γ_i 's, we deduce that there are $\delta + e - v + 1$ excess edges. Note, however, that there are at most $\epsilon = e - v + 1$ excess edges in each connected component of $G(\gamma)$.

We may now repeat the proof of Lemma 17. The main difference is that, for each i , we use that $\hat{\gamma}'_i$ and $\hat{\gamma}''_i$ are tangle free, it gives short cycling times and long cycling times for both $\hat{\gamma}'_i$ and $\hat{\gamma}''_i$. For each i , there are at most 2 short cycling times and $2(\epsilon - 1)$ long cycling times. Since there are at most $\ell^{4m\epsilon}$ ways to position these cycling times, we arrive at

$$|\mathcal{T}_{\ell,m,k}(v, e)| \leq (2\ell v)^{2m} \ell^{4m\epsilon} (v^2 \ell)^{4m} (v^2)^{4m(\epsilon-1)},$$

where the factor $(2\ell v)^{2m}$ accounts for the extra mark. Using $v \leq 2\ell m$, we obtain the claimed statement. \square

PROOF OF BOUND (31). From (42), it suffices to prove that

$$(44) \quad S = \sum_{\gamma \in \mathcal{T}_{\ell,m,k}} \mathbb{E} \prod_{i=1}^{2m} \prod_{s=1}^k \underline{A}_{\gamma_{i,s-1}\gamma_{i,s}} \prod_{s=k+2}^{\ell} A_{\gamma_{i,s-1}\gamma_{i,s}} \leq (C \log n)^{24m} \alpha^{(2\ell-k)m}.$$

As in (38), we find

$$(45) \quad \mathbb{E} \prod_{i=1}^{2m} \prod_{s=1}^k \underline{A}_{\gamma_{i,s-1}\gamma_{i,s}} \prod_{s=k+2}^{\ell} A_{\gamma_{i,s-1}\gamma_{i,s}} \leq \left(\frac{\alpha}{n}\right)^{e(\gamma)}.$$

From (43) and Lemma 18, we obtain

$$\begin{aligned} S &\leq \sum_{v=1}^{m(2\ell-2-k)} \sum_{e=v}^{m(2\ell-2-k)} |\mathcal{T}_{\ell,m,k}(v, e)| \left(\frac{\alpha}{n}\right)^e n^v \\ &\leq \alpha^{(2\ell-k)m} \sum_{v=1}^{m(2\ell-2-k)} \sum_{e=v}^{\infty} (4\ell m)^{12m(e-v)+20m} n^{v-e} \\ &\leq \alpha^{(2\ell-k)m} (4\ell m)^{20m} (2\ell m) \sum_{s=0}^{\infty} \left(\frac{(4\ell m)^{12m}}{n}\right)^s. \end{aligned}$$

For our choice of m in (40), we have $\ell m = o(\log n)^2$ and $(4\ell m)^{12m} \leq n^{24/25}$ and (44) follows. \square

6.4. *Bound (32) on $\|B^{(k)}\|$.* Since $\|K\|$ is of order n , we observe that the second statement in (32) improves by a factor \sqrt{n} the crude bound $\|KB^{(k)}\| \leq \|K\| \|B^{(k)}\|$.

PROOF. We only prove the second statement. The first statement is proved similarly. The argument again parallels that used to show (29). We take m as in (34). We have

$$(46) \quad \begin{aligned} \|KB^{(k-2)}\|^{2m} &\leq \text{tr} \{ ((KB^{(k-2)})(KB^{(k-2)})^*)^m \} \\ &= \sum_{\gamma \in \overline{W}_{k,m}} \prod_{i=1}^m \prod_{s=2}^k A_{\gamma_{2i-1,s-1}\gamma_{2i-1,s}} \prod_{s=1}^{k-1} A_{\gamma_{2i,s-1}\gamma_{2i,s}}, \end{aligned}$$

where $\overline{W}_{k,m}$ is the set of sequences of paths $\gamma = (\gamma_1, \dots, \gamma_{2m})$ such that $\gamma_i = (\gamma_{i,0}, \dots, \gamma_{i,k}) \in V^{k+1}$ are nonbacktracking paths of length $k+1$ with the following constraints: $(\gamma_{2i-1,1}, \dots, \gamma_{2i-1,k})$ and $(\gamma_{2i,0}, \dots, \gamma_{2i,k-1})$ are tangle-free paths of length k , for all $i = 1, \dots, m$, and such that for all $i = 1, \dots, 2m$,

$$(\gamma_{i,k-1}, \gamma_{i,k}) = (\gamma_{i+1,1}, \gamma_{i+1,0}),$$

with the convention that $\gamma_{2m+1} = \gamma_1$. From (46) and Markov's inequality, it suffices to prove that

$$(47) \quad S = \sum_{\gamma \in \overline{W}_{k,m}} \mathbb{E} \prod_{i=1}^m \prod_{s=2}^k A_{\gamma_{2i-1,s-1}\gamma_{2i-1,s}} \prod_{s=1}^{k-1} A_{\gamma_{2i,s-1}\gamma_{2i,s}} \leq (C \log n)^{16m} n^m \alpha^{2km}.$$

Note that $W_{k,m} \subset \overline{W}_{k,m}$ and for a path $\gamma \in \overline{W}_{k,m} \setminus W_{k,m}$, there must exist $i \in \{1, \dots, 2m\}$, such that γ_i is tangled. If i is odd, this can only happen if $\gamma_{i,0} \in \{\gamma_{i,2}, \dots, \gamma_{i,k}\}$, hence there are only $k-1$ possible choices for $\gamma_{i,0}$. Similarly, if i is even and γ_i is tangled, then $\gamma_{i,k} \in \{\gamma_{i,0}, \dots, \gamma_{i,k-2}\}$. Hence, to each path γ in $\overline{W}_{k,m}$ we can associate a path γ' in $W_{k,m}$ such that $(\gamma_{2i-1,1}, \dots, \gamma_{2i-1,k}) = (\gamma'_{2i-1,1}, \dots, \gamma'_{2i-1,k})$ and $(\gamma_{2i,0}, \dots, \gamma_{2i,k-1}) = (\gamma'_{2i,0}, \dots, \gamma'_{2i,k-1})$ and by previous argument, there are at most $(2k)^m$ paths in $\overline{W}_{k,m}$ which are associated with the same path γ' in $W_{k,m}$. Hence, we have, for $k \leq \ell$,

$$S \leq (2\ell)^m \sum_{\gamma \in W_{k,m}} \mathbb{E} \prod_{i=1}^m \prod_{s=2}^k A_{\gamma_{2i-1,s-1}\gamma_{2i-1,s}} \prod_{s=1}^{k-1} A_{\gamma_{2i,s-1}\gamma_{2i,s}}.$$

If $\gamma \in W_{k,m}$ is a canonical element of $W_{k,m}$, then $v(\gamma) - 1 \leq e(\gamma) \leq 2km$ and $v(\gamma) \geq 3$. Also, any $\gamma \in W_{k,m}$ is isomorphic to fewer than $n^{v(\gamma)}$ elements in $W_{k,m}$.

Moreover, we have that

$$(48) \quad \mathbb{E} \prod_{i=1}^m \prod_{s=2}^k A_{\gamma_{2i-1,s-1}\gamma_{2i-1,s}} \prod_{s=1}^{k-1} A_{\gamma_{2i,s-1}\gamma_{2i,s}} \leq \left(\frac{\alpha}{n}\right)^{e(\gamma)-m},$$

indeed, for any $p \geq 1$, $\mathbb{E}A_{u,v}^p \leq \alpha/n$ and, since $(\gamma_{2i+1,0}, \gamma_{2i+1,1}) = (\gamma_{2i,k}, \gamma_{2i,k-1})$ at most m distinct edges are covered by the union of $\{\gamma_{2i-1,0}, \gamma_{2i-1,1}\}$ and $\{\gamma_{2i,k-1}, \gamma_{2i,k}\}$. Hence, using Lemma 17, we obtain

$$\begin{aligned} S &\leq (2\ell)^m \left(\frac{n}{\alpha}\right)^m \sum_{v=3}^{km+1} \sum_{e=v-1}^{2km} |\mathcal{W}_{k,m}(v, e)| \left(\frac{\alpha}{n}\right)^e n^v \\ &\leq (2\ell)^m n^{m+1} \alpha^{(2k-1)m} \sum_{v=3}^{km+1} \sum_{e=v-1}^{2km} k^{2m} (2km)^{6m(e-v+1)} n^{v-e-1} \\ &\leq n^{m+1} \alpha^{(2k-1)m} 2^m (\ell)^{3m} (\ell m) \sum_{s=0}^{\infty} \left(\frac{(2\ell m)^{6m}}{n}\right)^s. \end{aligned}$$

For our choice of m in (34), the above geometric series converges and (47) follows. \square

6.5. Bound (33) on $\|S_k^{(\ell)}\|$.

PROOF. Observe that $L_{ef} = 0$ unless $e = f$, $K_{ef} = 1$, $K_{f^{-1}e} = 1$ or $K_{ef^{-1}} = 1$ in which cases $L_{ef} = -1$. We may thus decompose

$$L = -I - K',$$

where I is the identity, and the nonzero entries of K' are equal 1 and are the pairs (e, f) such that $K_{ef} = 1$, $K_{f^{-1}e} = 1$ or $K_{ef^{-1}} = 1$. Thus,

$$\|S_k^{(\ell)}\| \leq \|\Delta^{(k-1)}\| \|B^{(\ell-k-1)}\| + \|\Delta^{(\ell-1)} K'\| \|B^{(\ell-k-1)}\|.$$

Bounds (29)–(32) imply that the first term has a smaller order than the intended bound (33). Hence, we only need to bound the last term. We use again the method of moments. We observe that $K'_{ef} \leq K_{ef} + (PK)_{ef} + (KP)_{ef}$. A straightforward adaptation of the proof of bound (32) shows that w.h.p., for any $1 \leq k \leq \ell - 1$,

$$\|\Delta^{(k-1)} K'\| \leq \sqrt{n} (\log n)^{10} \alpha^{k/2},$$

which completes the proof. \square

7. Stochastic block model: Proof of Theorem 4. In this section, we give the strategy of proof for Theorem 4. Let $\ell = \ell(n) \sim \kappa \log_{\alpha} n$ for some $\kappa \in (0, \gamma/6)$ as in Theorem 4. Recalling the definition (13) of vector χ_k , we further introduce for all $k \in [r]$

$$(49) \quad \varphi_k = \frac{B^{\ell} \chi_k}{\|B^{\ell} \chi_k\|}, \quad \theta_k = \|B^{\ell} \check{\varphi}_k\|,$$

and

$$\zeta_k = \frac{B^\ell \check{\varphi}_k}{\theta_k} = \frac{B^\ell B^{*\ell} \check{\chi}_k}{\|B^\ell B^{*\ell} \check{\chi}_k\|}$$

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

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

We then have the following.

PROPOSITION 19. *For some $b, c > 0$, w.h.p.:*

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

Proposition 19 will follow from the local analysis done in Section 9. The next proposition will be established in Section 10 using a matrix expansion together with norm bounds derived by combinatorial arguments parallel to the proof of Proposition 11 for the Erdős–Rényi graph.

PROPOSITION 20. *For some $c > 0$, w.h.p.,*

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

We now check that the two preceding propositions imply Theorem 4. We consider $(\bar{\varphi}_1, \dots, \bar{\varphi}_{r'})$ obtained by the Gram–Schmidt orthonormalization of $(\check{\varphi}_1, \dots, \check{\varphi}_r)$. By Lemma 9 and Proposition 19(iv), w.h.p. $r' = r$ and for all $k \in [r]$,

$$(50) \quad \|\check{\varphi}_k - \bar{\varphi}_k\| = O((\log n)^c \alpha^{3\ell/2} n^{-\gamma/2}).$$

Similarly, for $k \in [r_0]$, we denote by $\tilde{\zeta}_k$ the orthogonal projection of ζ_k on the orthogonal of the vector space spanned by $\bar{\varphi}_j$, $j \in [r_0]$, $j \neq k$ and $\tilde{\zeta}_j$, $j < k$. We set $\bar{\zeta}_k = \tilde{\zeta}_k / \|\tilde{\zeta}_k\|$. From Proposition 19(v)–(vi), we find w.h.p. for $k \in [r_0]$,

$$(51) \quad \|\zeta_k - \bar{\zeta}_k\| = O((\log n)^c \alpha^{5\ell/2} n^{-\gamma/2}).$$

We then set

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

Since $\|\check{\varphi}_k - \bar{\varphi}_k\| = o(1)$, from Proposition 19(i)–(iii), we find by induction on $k \in [r]$, w.h.p. for all $k \in [r]$,

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

Consequently, from Proposition 20, we have, w.h.p.,

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

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

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

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

$$(52) \quad \|B^\ell P_0 - D_0\| = O(\alpha^{\ell/2}).$$

We also set $D_1 = B^\ell P_1$ where P_1 is the orthogonal projection of $H_1 = \text{span}(\bar{\varphi}_k, k \in [r] \setminus [r_0])$ and $C = B^\ell - D_0 - D_1$. Arguing similarly, from Proposition 19(iii), w.h.p., for $k \in [r] \setminus [r_0]$,

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

Hence,

$$(53) \quad \|D_1\| = O((\log n)^c \alpha^{\ell/2}).$$

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

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

Hence, Proposition 20 and (52)–(53) imply that w.h.p.:

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

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

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

We are now in position to apply Proposition 8. From (51), the statement of Proposition 19(ii) also holds with ζ_k replaced by $\bar{\zeta}_k$. It readily implies Theorem 4.

8. Controls on the growth of Poisson multitype branching processes. In this section, we derive results for multitype Galton–Watson branching processes with Poisson offspring that will be crucial for the local analysis of Section 9. We refer to Section 3.2 for the notation used below.

8.1. *Two theorems of Kesten and Stigum.* We consider a multitype branching process where a particle of type $j \in [r]$ has a $\text{Poi}(M_{ij})$ number of children with type i . We denote by $Z_t = (Z_t(1), \dots, Z_t(r))$ the population at generation t , where $Z_t(i)$ is the number of particles at generation t with type i . We denote by \mathcal{F}_t the natural filtration associated to Z_t . Following Kesten and Stigum [17, 18], we have the following statement.

THEOREM 21. *For any $k \in [r_0]$,*

$$X_k(t) = \frac{\langle \phi_k, Z_t \rangle}{\mu_k^t} - \langle \phi_k, Z_0 \rangle$$

is an \mathcal{F}_t -martingale converging a.s. and in L^2 such that for some $C > 0$ and all $t \geq 0$, $\mathbb{E}X_k(t) = 0$ and $\mathbb{E}[X_k^2(t)|Z_0] \leq C\|Z_0\|_1$.

PROOF. We include the proof for later use. For $0 \leq s < t$, we have

$$Z_t - M^{t-s}Z_s = \sum_{u=s}^{t-1} M^{t-u-1}(Z_{u+1} - MZ_u),$$

so that, as $\phi_k^*M = \mu_k\phi_k^*$,

$$(54) \quad \frac{\langle \phi_k, Z_t \rangle}{\mu_k^t} = \frac{\langle \phi_k, Z_s \rangle}{\mu_k^s} + \sum_{u=s}^{t-1} \frac{\langle \phi_k, (Z_{u+1} - MZ_u) \rangle}{\mu_k^{u+1}}.$$

It follows easily that $(X_k(t))$ is an \mathcal{F}_t -martingale with mean 0. From Doob's martingale convergence theorem, the statement will follow if we prove that for some $C > 0$ and all integer $t \geq 0$,

$$\mathbb{E}[X_k^2(t)|Z_0] \leq C\|Z_0\|_1 = C\langle \mathbb{1}, Z_0 \rangle.$$

To this end, we denote by $Z_{s+1}(i, j)$ the number of individuals of type i in the $(s+1)$ th generation which descend from a particle of type j in the s th generation. Thus, $\sum_{j \in [r]} Z_{s+1}(i, j) = Z_{s+1}(i)$. We then have

$$(55) \quad \begin{aligned} \mathbb{E}[\|Z_{s+1} - MZ_s\|_2^2 | Z_s] &= \sum_{i \in [r]} \mathbb{E}\left[\left(Z_{s+1}(i) - \sum_{j \in [r]} M_{ij}Z_s(j)\right)^2 \middle| Z_s\right] \\ &= \sum_{i, j \in [r]} \mathbb{E}[(Z_{s+1}(i, j) - M_{ij}Z_s(j))^2 | Z_s(j)] \\ &= \sum_{i, j \in [r]} M_{ij}Z_s(j) \\ &= \langle \mathbb{1}, MZ_s \rangle, \end{aligned}$$

where in the penultimate equality we used the fact that the variance of a Poisson random variable equals its mean. It follows that

$$\mathbb{E}[\|Z_{s+1} - MZ_s\|^2 | Z_0] = \langle \mathbb{1}, M^{s+1} Z_0 \rangle.$$

Recall $\mathbb{1}$ is a left eigenvector of M : $\mathbb{1}^* M^{s+1} = \mu_1^{s+1} \mathbb{1}^*$. We thus obtain

$$\mathbb{E}[\|Z_{s+1} - MZ_s\|_2^2 | Z_0] = \langle \mathbb{1}, Z_0 \rangle \mu_1^{s+1}.$$

Hence finally,

$$\begin{aligned} \mathbb{E}[X_k^2(t) | Z_0] &= \sum_{s=0}^{t-1} \frac{\mathbb{E}[\langle \phi_k, (Z_{s+1} - MZ_s) \rangle^2 | Z_0]}{\mu_k^{2(s+1)}} \\ &\leq \sum_{s=0}^{t-1} \frac{\|\phi_k\|_2^2 \mathbb{E}[\|Z_{s+1} - MZ_s\|_2^2 | Z_0]}{\mu_k^{2(s+1)}} \\ &\leq \langle \mathbb{1}, Z_0 \rangle \sum_{s=0}^{t-1} \left(\frac{\mu_1}{\mu_k^2} \right)^{s+1}. \end{aligned}$$

Since $\mu_k^2 > \mu_1$ the above series is convergent. \square

We also need to control the behavior of $\langle \phi_k, Z_t \rangle$ for $k \in [r] \setminus [r_0]$. The next result is contained in Kesten and Stigum [17], Theorem 2.4.

THEOREM 22. *Assume $Z_0 = x$. For $k \in [r] \setminus [r_0]$ define*

$$X_k(t) = \begin{cases} \frac{\langle \phi_k, Z_t \rangle}{\mu_1^{t/2}} & \text{if } \mu_k^2 < \mu_1, \\ \frac{\langle \phi_k, Z_t \rangle}{\mu_1^{t/2} t^{1/2}} & \text{if } \mu_k^2 = \mu_1. \end{cases}$$

Then $X_k(t)$ converges weakly to a random variable X_k with finite positive variance.

Note that Theorem 2.4 in [17] expresses X_k as a mixture of Gaussian variables. The normalization in the case $\mu_k^2 = \mu_1$ comes from the fact that M is diagonalizable, and hence all its Jordan blocks are of size 1.

8.2. Quantitative versions of the Kesten–Stigum theorems. We will also need probabilistic bounds on the growth of the total population at generation t defined as

$$S_t = \|Z_t\|_1 = \langle \phi_1, Z_t \rangle.$$

We observe that (10) implies that S_t itself is a Galton–Walton branching process with offspring distribution $\text{Poi}(\mu_1)$.

LEMMA 23. *Assume $S_0 = 1$. There exist $c_0, c_1 > 0$ such that, for all $s \geq 0$,*

$$\mathbb{P}(\forall k \geq 1, S_k \leq s\mu_1^k) \geq 1 - c_1 e^{-c_0 s}.$$

PROOF. For $k \geq 1$, we set

$$\varepsilon_k = \mu_1^{-k/2} \sqrt{k} \quad \text{and} \quad f_k = \prod_{\ell=1}^k (1 + \varepsilon_\ell).$$

It is straightforward to check that f_k converges, hence there exist constants $c_0, c_1 > 0$ such that for all $k \geq 1$,

$$(56) \quad c_0 \leq f_k \leq c_1 \quad \text{and} \quad \varepsilon_k \leq c_1.$$

Using Chernov bound, if Y_i are i.i.d. $\text{Poi}(\mu_1)$ variables then, for any integer $\ell \geq 1$ and positive real $s > 1$,

$$(57) \quad \mathbb{P}\left(\sum_{i=1}^{\ell} Y_i \geq \ell \mu_1 s\right) \leq e^{-\ell \mu_1 \gamma(s)},$$

where we have set $\gamma(s) = s \log s - s + 1$. In particular, on the event $\{S_k \leq s f_k \mu_1^k\} \in \mathcal{F}_k$, we have

$$\mathbb{P}(S_{k+1} > s f_{k+1} \mu_1^{k+1} | \mathcal{F}_k) \leq e^{-s \mu_1^{k+1} f_k \gamma(1 + \varepsilon_{k+1})} \leq e^{-c'_0 s \mu_1^{k+1} \varepsilon_{k+1}^2} = e^{-c'_0 (k+1)s},$$

where we have used the existence of some $\theta > 0$ such that for $x \in [0, c_1]$, one has $\gamma(1+x) \geq \theta x^2$. Finally, by our choice of ε_k and (56), if $s \geq \max(1/c'_0, 1/c_1)$,

$$\mathbb{P}(\exists k : S_k > s c_1 \mu_1^{k+1}) \leq \sum_{\ell=1}^k e^{-c'_0 s \ell} \leq \frac{e^{-c'_0 s}}{1 - e^{-c'_0 s}}.$$

Hence, we deduce the statement of the lemma for some (suitably redefined) constants $c_0, c_1 > 0$. \square

A key ingredient in the subsequent analysis will be the following result, which bounds by how much the growth of processes $s \rightarrow \langle \phi_k, Z_s \rangle$ deviates from a purely deterministic exponential growth.

THEOREM 24. *Let $\beta > 0$ and $Z_0 = x \in \mathbb{N}^r$ be fixed. There exists $C = C(x, \beta) > 0$ such that with probability at least $1 - n^{-\beta}$, for all $k \in [r_0]$, all $s, t \geq 0$, with $0 \leq s < t$,*

$$|\langle \phi_k, Z_s \rangle - \mu_k^{s-t} \langle \phi_k, Z_t \rangle| \leq C(s+1) \mu_1^{s/2} (\log n)^{3/2},$$

and for all $k \in [r] \setminus [r_0]$, all $t \geq 0$,

$$|\langle \phi_k, Z_t \rangle| \leq C(t+1)^2 \mu_1^{t/2} (\log n)^{3/2}.$$

Finally, for all $k \in [r] \setminus [r_0]$, all $t \geq 0$, $\mathbb{E}|\langle \phi_k, Z_t \rangle|^2 \leq C(t+1)^3 \mu_1^t$.

PROOF. We start with classical tail bounds for $Y \stackrel{d}{=} \text{Poi}(\lambda)$. From (57) for $s > 0$,

$$\mathbb{P}(Y - \lambda > \lambda s) \leq e^{-\lambda\gamma(1+s)},$$

with $\gamma(s) = s \log s + 1 - s$. Similarly, for $s < 1$ one has

$$\mathbb{P}(Y - \lambda < -\lambda s) \leq e^{-\lambda\gamma(1-s)},$$

where by convention $\gamma(x) = +\infty$ for $x \leq 0$. Let $\delta(x) := \gamma(1-x) \wedge \gamma(1+x)$. Then for any $s \geq 0$,

$$\mathbb{P}(|Y - \lambda| > \lambda s) \leq 2e^{-\lambda\delta(s)}.$$

In particular, for any $i \in [r]$, letting $y := MZ_t$, we have, if $Z_t \neq 0$,

$$\mathbb{P}(|Z_{t+1}(i) - y(i)| > s \|y\|_1^{1/2} | \mathcal{F}_t) \leq 2e^{-y(i)\delta(s\|y\|_1^{1/2}/y(i))}.$$

Consider first the case where $s\|y\|_1^{1/2} \leq y(i)$. As there exists $\theta > 0$ such that for all $x \in [0, 1]$, $\delta(x) \geq \theta x^2$, we get

$$\mathbb{P}(|Z_{t+1}(i) - y(i)| > s\|y\|_1^{1/2} | \mathcal{F}_t) \leq 2e^{-\frac{\theta s^2 \|y\|_1}{y(i)}} \leq 2e^{-\theta s^2}.$$

Consider now the case where $s\|y\|_1^{1/2} > y(i)$. As there exists $\theta' > 0$ such that, for all $x \geq 1$, $\delta(x) \geq \theta' x$, we get

$$\mathbb{P}(|Z_{t+1}(i) - y(i)| > s\|y\|_1^{1/2} | \mathcal{F}_t) \leq 2e^{-\theta' s\|y\|_1^{1/2}} \leq 2e^{-\theta' \sqrt{\mu_1} s},$$

since $Z_t \neq 0$ implies that $\|y\|_1 \geq \mu_1$ from (10). Thus, there exists some $c_0 > 0$ such that, for any $s \geq 0$,

$$\begin{aligned} \mathbb{P}(\|Z_{t+1} - MZ_t\|_2 > s\|Z_t\|_1^{1/2} | \mathcal{F}_t) &\leq \sum_{i=1}^r \mathbb{P}\left(|Z_{t+1}(i) - y(i)| > \frac{s\|Z_t\|_1^{1/2}}{\sqrt{r}} \mid \mathcal{F}_t\right) \\ &\leq 2re^{-c_0(s \wedge s^2)}. \end{aligned}$$

If $Z_t = 0$, then $Z_{t+1} = 0$ and the same bound trivially holds. We thus obtain the existence of constants $c_0, c_1 > 0$ such that, for any $u \geq 1$,

$$\begin{aligned} (58) \quad &\mathbb{P}(\forall t \geq 0, \|Z_{t+1} - MZ_t\|_2 \leq u(t+1) \log n \|Z_t\|_1^{1/2}) \\ &\geq 1 - \sum_{t \geq 1} 2re^{-c_0 u t \log n} \geq 1 - c_1 n^{-c_0 u}. \end{aligned}$$

Now, from (54), for any $s, 0 \leq s \leq t$,

$$|\langle \phi_k, Z_s \rangle - \mu_k^{s-t} \langle \phi_k, Z_t \rangle| \leq \mu_k^s \sum_{h=s}^{t-1} \frac{\|\phi_k\|_2 \|Z_{h+1} - MZ_h\|_2}{\mu_k^{h+1}}.$$

From equation (58) and Lemma 23, for C large enough, with probability at least $1 - n^{-\beta}$, we have for all $h \geq 0$ that $\|Z_{h+1} - MZ_h\|_2 \leq C(\log n)(h+1)\|Z_h\|_1^{1/2}$ and $\|Z_h\|_1 \leq C(\log n)\mu_1^h$. On this event, we get, for $k \in [r_0]$,

$$\begin{aligned} |\langle \phi_k, Z_s \rangle - \mu_k^{s-t} \langle \phi_k, Z_t \rangle| &\leq C'(\log n)^{3/2} \mu_k^s \sum_{h=s}^{t-1} (h+1) \left(\frac{\sqrt{\mu_1}}{\mu_k} \right)^h \\ &\leq C''(\log n)^{3/2} (s+1) \mu_1^{s/2}, \end{aligned}$$

where at the last line, we used that $\mu_k^2 > \mu_1$ and $\sum_{h \geq s} ha^h \leq c(a)sa^s$ for $0 < a < 1$. Similarly, on the same event, for $k \in [r] \setminus [r_0]$, from (54), for $t \geq 1$ and $s = 0$,

$$\begin{aligned} |\langle \phi_k, Z_t \rangle - \mu_k^t \langle \phi_k, Z_0 \rangle| &\leq \mu_k^t \sum_{u=0}^{t-1} \frac{\|\phi_k\|_2 \|Z_{u+1} - MZ_u\|_2}{\mu_k^{u+1}} \\ &\leq C'(\log n)^{3/2} \mu_k^t \sum_{u=0}^{t-1} (u+1) \left(\frac{\sqrt{\mu_1}}{\mu_k} \right)^u. \end{aligned}$$

Using now $\mu_k^2 \leq \mu_1$, we have $\sum_{u=0}^{t-1} (u+1) \left(\frac{\sqrt{\mu_1}}{\mu_k} \right)^u = O(t^2(\sqrt{\mu_1}/\mu_k)^t)$.

For the last result, we define

$$U = \sup_{t \geq 0} \frac{\|Z_{t+1} - MZ_t\|_2}{(t+1)\|Z_t\|_1^{1/2}}.$$

From (58) (with $n = 2$), for any $p \geq 1$, $\mathbb{E}U^p = O(1)$. We obtain from (54) and the Cauchy–Schwarz inequality,

$$\begin{aligned} \mathbb{E}|\langle \phi_k, Z_t \rangle - \mu_k^t \langle \phi_k, Z_0 \rangle|^2 &\leq \mu_k^{2t} \sum_{s=0}^{t-1} \mathbb{E} \frac{\|\phi_k\|_2^2 \|Z_{s+1} - MZ_s\|_2^2}{\mu_k^{2(s+1)}} \\ &\leq \mu_k^{2t} \sum_{s=0}^{t-1} \mathbb{E} \frac{U^2 (s+1)^2 \|Z_s\|_1}{\mu_k^{2(s+1)}} \\ &\leq t^2 \mu_k^{2t} \sqrt{\mathbb{E}U^4} \sum_{s=0}^{t-1} \frac{\sqrt{\mathbb{E}\|Z_s\|_1^2}}{\mu_k^{2(s+1)}} \\ &= O(t^3 \mu_1^t), \end{aligned}$$

where for the last equality, we used the fact that $\mathbb{E}\|Z_s\|_1^2 = O(\mu_1^{2s})$ which follows from Theorem 21 with $k = 1$ (recall that $\phi_1 = \mathbb{1}$), and the bound of $O(t\mu_1^t/\mu_k^{2t})$ on the sum, which holds for $k \notin [r_0]$. \square

8.3. *A cross-generation functional.* For the subsequent analysis, in order to control the law of the candidate eigenvectors $B^\ell B^{*\ell} \check{\chi}_k$, we also need to consider a functional of the multitype branching process which depends on particles in more than one generation. More precisely, assuming that $\|Z_0\|_1 = 1$, we denote by V the particles of the random tree and $o \in V$ the starting particle. Particle $v \in V$ has type $\sigma(v) \in [r]$ and generation $|v|$ from $o \in V$. For $v \in V$ and integer $t \geq 0$, let Y_t^v denote the set of particles of generation t from v in the subtree of particles with common ancestor $v \in V$. Finally, $Z_t^v = (Z_t^v(1), \dots, Z_t^v(r))$ is the vector of population at generation t from v , that is, $Z_t^v(i) = \sum_{u \in Y_t^v} \mathbb{1}(\sigma(u) = i)$. We set

$$S_t^v = \|Z_t^v\|_1 = \langle \phi_1, Z_t^v \rangle.$$

With our previous notation, $Z_t^o = Z_t$, $S_t^o = S_t$. We fix an integer $k \in [r]$, $\ell \geq 1$ and set

$$(59) \quad Q_{k,\ell} = \sum_{(u_0, \dots, u_{2\ell+1}) \in \mathcal{P}_{2\ell+1}} \phi_k(\sigma(u_{2\ell+1})),$$

where the sum is over $(u_0, \dots, u_{2\ell+1}) \in \mathcal{P}_{2\ell+1}$, the set of paths in the tree starting from $u_0 = o$ of length $2\ell + 1$ with (u_0, \dots, u_ℓ) and $(u_\ell, \dots, u_{2\ell+1})$ nonbacktracking and $u_{\ell-1} = u_{\ell+1}$ [i.e., $(u_0, \dots, u_{2\ell+1})$ backtracks exactly once at the $(\ell + 1)$ th step].

The following alternative representation of $Q_{k,\ell}$ will prove useful. By distinguishing paths $(u_0, \dots, u_{2\ell+1})$ according to the smallest depth $t \in \{0, \dots, \ell - 1\}$ to which they climb back after visiting $u_{\ell+1}$ and the node $u_{2\ell-t}$ they then visit at level t we have that

$$(60) \quad Q_{k,\ell} = \sum_{t=0}^{\ell-1} \sum_{u \in Y_t^o} L_{k,\ell}^u,$$

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

$$L_{k,\ell}^u = \sum_{w \in Y_1^u} S_{\ell-t-1}^w \left(\sum_{v \in Y_1^u \setminus \{w\}} \langle \phi_k, Z_t^v \rangle \right).$$

We then have the following.

THEOREM 25. *Assume $Z_0 = \delta_x$. For $k \in [r_0]$, $Q_{k,\ell}/\mu_k^{2\ell}$ converges in L^2 as ℓ tends to infinity to a random variable with mean $\mu_k \phi_k(x)/(\mu_k^2/\alpha - 1)$. For $k \in [r] \setminus [r_0]$, there exists a constant C such that $\mathbb{E} Q_{k,\ell}^2 \leq C \alpha^{2\ell} \ell^5$.*

PROOF. Let \mathcal{F}_t be the filtration generated by (Z_0, \dots, Z_t) . The variables $(L_{k,\ell}^u, u \in Y_t^o)$ are independent given \mathcal{F}_t . We will show that the sum (60) concentrates around its mean. Let us first compute the mean of $L_{k,\ell}^u$ for $u \in Y_t^o$. We

use the fact, that given \mathcal{F}_{t+1} and $v \neq w \in Y_{t+1}^o$, Z_t^v and $S_{\ell-t-1}^w$ are independent. Hence, we have with the short-hand notation $\mathbb{E}_{\mathcal{F}_t} = \mathbb{E}(\cdot | \mathcal{F}_t)$:

$$\mathbb{E}_{\mathcal{F}_t} L_{k,\ell}^u = \mathbb{E}_{\mathcal{F}_t} \sum_{(v,w) \in Y_1^u, v \neq w} \mathbb{E}_{\mathcal{F}_{t+1}}(\phi_k, Z_t^v) \mathbb{E}_{\mathcal{F}_{t+1}} S_{\ell-t-1}^w.$$

By assumption (10), $\mathbb{E}_{\mathcal{F}_{t+1}} S_{\ell-t-1}^w = \alpha^{\ell-t-1}$. Moreover, we have $\mathbb{E}_{\mathcal{F}_{t+1}}(\phi_k, Z_t^v) = \mu_k^t \langle \phi_k, Z_0^v \rangle$ so that

$$\begin{aligned} \mathbb{E}_{\mathcal{F}_t} L_{k,\ell}^u &= \alpha^{\ell-t-1} \mu_k^t \mathbb{E}_{\mathcal{F}_t} \left((|Y_1^u| - 1) \sum_{v \in Y_1^u} \langle \phi_k, Z_0^v \rangle \right) \\ (61) \quad &= \alpha^{\ell-t-1} \mu_k^t \mathbb{E}_{\mathcal{F}_t} \left((|Y_1^u| - 1) |Y_1^u| \sum_{i \in [r]} \phi_k(i) \frac{M_{i,\sigma(u)}}{\alpha} \right) \\ &= \mu_k^{t+1} \alpha^{\ell-t} \langle \phi_k, Z_0^u \rangle \end{aligned}$$

and

$$(62) \quad \mathbb{E}_{\mathcal{F}_t} \sum_{u \in Y_t^o} L_{k,\ell}^u = \mu_k^{t+1} \alpha^{\ell-t} \langle \phi_k, Z_t \rangle = \mu_k^{2t+1} \alpha^{\ell-t} Y_k(t),$$

where $Y_k(t) = X_k(t) + \langle \phi_k, Z_0 \rangle$ and X_k is the centered martingale defined in Theorem 21.

We now prove the statements of the theorem for $k \in [r_0]$. We find similarly

$$\begin{aligned} \text{Var}_{\mathcal{F}_t}(L_{k,\ell}^u) &= \mathbb{E}_{\mathcal{F}_t} (L_{k,\ell}^u - \mathbb{E}_{\mathcal{F}_t} L_{k,\ell}^u)^2 \\ &\leq \mathbb{E}_{\mathcal{F}_t} (L_{k,\ell}^u)^2 \\ &= \mathbb{E}_{\mathcal{F}_t} \left(\sum_{v \neq w \in Y_1^u} S_{\ell-t-1}^w \langle \phi_k, Z_t^v \rangle \right)^2 \\ &\leq C \mathbb{E}_* \langle \phi_k, Z_t \rangle^2 \mathbb{E}_* S_{\ell-t-1}^2, \end{aligned}$$

where $\mathbb{E}_*(\cdot) = \max_{i \in [r]} \mathbb{E}(\cdot | Z_0 = \delta_i)$ and constant C can be taken equal to $\mathbb{E}_* |Y_1^o|^4$. For $k \in [r_0]$, we deduce from Theorem 21 that, for some new $C > 0$,

$$(63) \quad \text{Var}_{\mathcal{F}_t} \left(\sum_{u \in Y_t^o} L_{k,\ell}^u \right) = \sum_{u \in Y_t^o} \text{Var}_{\mathcal{F}_t}(L_{k,\ell}^u) \leq C \mu_k^{2t} \alpha^{2(\ell-t)} S_t.$$

We now define

$$(64) \quad \bar{Q}_{k,\ell} = \sum_{t=0}^{\ell-1} \mathbb{E}_{\mathcal{F}_t} \sum_{u \in Y_t^o} L_{k,\ell}^u = \sum_{t=0}^{\ell-1} \mu_k^{2t+1} \alpha^{\ell-t} Y_k(t).$$

Since $k \in [r_0]$, $\rho_k := \mu_k^2/\alpha > 1$. We write

$$\frac{\bar{Q}_{k,\ell}}{\mu_k^{2\ell}} = \mu_k \sum_{t=0}^{\ell-1} \rho_k^{t-\ell} Y_k(t).$$

From Theorem 21, $\bar{Q}_{k,\ell}/\mu_k^{2\ell}$ converges a.s. to $\mu_k Y_k(\infty)/(\rho_k - 1)$ where $Y_k(\infty) = X_k(\infty) + \langle \phi_k, Z_0 \rangle$ and $X_k(\infty)$ is the limit of the martingale defined in Theorem 21. Moreover, $\bar{Q}_{k,\ell}/\mu_k^{2\ell}$ also converges in L^2 . Indeed, we find easily from the Cauchy–Schwarz inequality

$$\begin{aligned} & \mathbb{E} \left(\frac{\bar{Q}_{k,\ell}}{\mu_k^{2\ell}} - \frac{\mu_k Y_k(\infty)}{\rho_k - 1} (1 - \rho_k^{-\ell}) \right)^2 \\ &= \mu_k^2 \mathbb{E} \left(\sum_{t=0}^{\ell-1} \rho_k^{t-\ell} (Y_k(t) - Y_k(\infty)) \right)^2 \\ &\leq \mu_k^2 \left(\sum_{t=0}^{\ell-1} \rho_k^{t-\ell} \right) \left(\sum_{t=0}^{\ell-1} \mathbb{E} (Y_k(t) - Y_k(\infty))^2 \rho_k^{t-\ell} \right). \end{aligned}$$

Since $\rho_k > 1$, the first term of the above expression is of order $O(1)$. For the second term, from Theorem 21, for any $\varepsilon > 0$, there is t_0 such that for all $t \geq t_0$, $\mathbb{E} (Y_k(t) - Y_k(\infty))^2 \leq \varepsilon$. We find that the second term is $O(\varepsilon + \rho_k^{t_0-\ell}) = o(1)$. It proves that $\bar{Q}_{k,\ell}/\mu_k^{2\ell}$ converges in L^2 .

We now check that $Q_{k,\ell}$ and $\bar{Q}_{k,\ell}$ are close in L^2 for $k \in [r_0]$. For a real random variable Z , set $\|Z\|_2 = \sqrt{\mathbb{E}Z^2}$. From (60)–(63) and the triangle inequality, we get

$$\begin{aligned} \|Q_{k,\ell} - \bar{Q}_{k,\ell}\|_2 &\leq \sum_{t=0}^{\ell-1} \left\| \sum_{u \in Y_t^o} L_{k,\ell}^u - \mathbb{E}_{\mathcal{F}_t} \sum_{u \in Y_t^o} L_{k,\ell}^u \right\|_2 \\ &= \sum_{t=0}^{\ell-1} \left\| \left(\text{Var}_{\mathcal{F}_t} \left(\sum_{u \in Y_t^o} L_{k,\ell}^u \right) \right)^{1/2} \right\|_2 \\ &\leq C \sum_{t=0}^{\ell} \mu_k^t \alpha^{\ell-t} \|\sqrt{S_t}\|_2 \\ &= O(\mu_k^\ell \alpha^{\ell/2}) \\ &= o(\mu_k^{2\ell}), \end{aligned}$$

where at the last line, we have used Lemma 23 and $k \in [r_0]$. It follows that $\|(Q_{k,\ell} - \bar{Q}_{k,\ell})/\mu_k^{2\ell}\|_2$ goes to 0 and it concludes the statements of the theorem for $k \in [r_0]$.

For $k \notin [r_0]$, we note that $\mathbb{E}Z^2 \leq \mathbb{E}(\mathbb{E}_Y Z)^2 + \mathbb{E}\text{Var}_Y(Z)$ so that $\|Z\|_2 \leq \|\mathbb{E}_Y Z\|_2 + \|\text{Var}_Y(Z)\|_2^{1/2}$. From (60) and the triangle inequality, we get

$$(65) \quad \begin{aligned} \|Q_{k,\ell}\|_2 &\leq \sum_{t=0}^{\ell-1} \left\| \sum_{u \in Y_t^o} L_{k,\ell}^u \right\|_2 \\ &\leq \sum_{t=0}^{\ell-1} \left\| \mathbb{E}_{\mathcal{F}_t} \left(\sum_{u \in Y_t^o} L_{k,\ell}^u \right) \right\|_2 + \left\| \left(\text{Var}_{\mathcal{F}_t} \left(\sum_{u \in Y_t^o} L_{k,\ell}^u \right) \right)^{1/2} \right\|_2. \end{aligned}$$

The last statement of Theorem 24 and (63) gives

$$\text{Var}_{\mathcal{F}_t} \left(\sum_{u \in Y_t^o} L_{k,\ell}^u \right) = O(S_t \alpha^{2(\ell-t)} \alpha^t t^3).$$

We deduce from (65) and (62)

$$\|Q_{k,\ell}\|_2 \leq C \sum_{t=0}^{\ell-1} (\mu_k^t \alpha^{\ell-t} \|\langle \phi_k, Z_t \rangle\|_2 + \alpha^{\ell-t} \alpha^{t/2} t^{3/2} \|\sqrt{S_t}\|_2) = O(\alpha^\ell \ell^{5/2}).$$

This completes the proof. \square

We finish this section with a rough bound on $Q_{k,\ell}$.

LEMMA 26. *For any $p \geq 1$, there exists a constant $C = C(p, \alpha) > 0$ such that for any $k \in [r]$,*

$$\mathbb{E}|Q_{k,\ell}|^p \leq C \alpha^{2p\ell}.$$

PROOF. We use the notation of Theorem 25. First, from Lemma 23, for any $p \geq 1$, $\mathbb{E}S_t^p \leq C \alpha^{tp}$. In particular, for any $v \in Y_{t+1}^o$ and $k \in [r]$,

$$\mathbb{E}_{\mathcal{F}_{t+1}} |\langle \phi_k, Z_t^v \rangle|^p = O(\mathbb{E}_{\mathcal{F}_{t+1}} (S_t^v)^p) = O(\alpha^{tp}).$$

We use twice the bound $|\sum_{i=1}^n x_i|^p \leq n^{p-1} \sum_{i=1}^n |x_i|^p$. We find

$$\begin{aligned} \mathbb{E}_{\mathcal{F}_t} \left| \sum_{u \in Y_t^o} L_{k,\ell}^u \right|^p &\leq S_t^{p-1} \sum_{u \in Y_t^o} \mathbb{E}_{\mathcal{F}_t} |L_{k,\ell}^u|^p \\ &\leq S_t^{p-1} \sum_{u \in Y_t^o} \mathbb{E}_{\mathcal{F}_t} \mathbb{E}_{\mathcal{F}_{t+1}} \left| \sum_{(v,w) \in Y_1^u, v \neq w} \langle \phi_k, Z_t^v \rangle S_{\ell-t-1}^w \right|^p \\ &\leq S_t^{p-1} \sum_{u \in Y_t^o} \mathbb{E}_{\mathcal{F}_t} (S_1^u)^{2(p-1)} \sum_{(v,w) \in Y_1^u, v \neq w} C \alpha^{tp} \alpha^{(\ell-t-1)p} \\ &\leq S_t^{p-1} \sum_{u \in Y_t^o} C \alpha^{(\ell-1)p} \mathbb{E}_{\mathcal{F}_t} (S_1^u)^{2p} \\ &\leq C' \alpha^{\ell p} S_t^p, \end{aligned}$$

for some new constant C' depending on α and p . We deduce that for some new $C > 0$,

$$\mathbb{E} \left| \sum_{u \in Y_t^o} L_{k,\ell}^u \right|^p \leq C^p \alpha^{(\ell+t)p}.$$

For a real random variable Z , set $\|Z\|_p = (\mathbb{E}Z^p)^{1/p}$. Using (60) and the triangle inequality we get

$$\|Q_{k,\ell}\|_p \leq \sum_{t=0}^{\ell-1} \left\| \sum_{u \in Y_t^o} L_{k,\ell}^u \right\|_p \leq C \sum_{t=0}^{\ell-1} \alpha^{(\ell+t)} = O(\alpha^{2\ell}). \quad \square$$

8.4. *Decorrelation in homogeneous Galton–Watson branching processes.* We now establish that the variables $Q_{k,\ell}$ and $Q_{j,\ell}$ are uncorrelated when $k \neq j$. To this end, we need the following lemma.

LEMMA 27. *Assume that the type $\sigma(o)$ at the root node o is distributed according to the stationary distribution π . Conditionally, on the branching tree \mathcal{T} , the process of types $\sigma(u)$ attached to the vertices of the tree is a Markov random field. For any two neighbor nodes u, v of \mathcal{T} and any $i, j \in [r]$, one has the following transition probabilities*

$$\mathbb{P}(\sigma(u) = i | \sigma(v) = j, \mathcal{T}) = \frac{1}{\alpha} M_{ij}.$$

For any two (possibly equal) nodes u, v of \mathcal{T} , any $k, j \in [r]$, $k \neq j$, one has

$$(66) \quad \mathbb{E}(\phi_k(\sigma(u))\phi_j(\sigma(v)) | \mathcal{T}) = 0.$$

PROOF. By standard properties of independent Poisson random variables, conditionally on the type $\sigma(o)$ and on the number of children of the root o , the types of each of the children of the root are i.i.d., distributed according to $M_{\cdot\sigma(o)}/\alpha$. Moreover, π is the stationary distribution for this transition kernel, which is reversible, as follows from the relation $M = \Pi W$ and the facts that W is symmetric together with the assumption (10) that the column sums of M all coincide with α . The Markov random field property and the expression of the transition kernel follow by iterating this argument.

We now evaluate the conditional expectation in (66). Let $u_1 = u, \dots, u_t = v$ denote the unique path in \mathcal{T} connecting nodes u and v . Let \mathcal{F}_s denote the σ -field generated by \mathcal{T} and the type variables $\sigma(u_1), \dots, \sigma(u_s)$. We then have by the Markov random field property

$$\mathbb{E}(\phi_j(\sigma(u_{s+1})) | \mathcal{F}_s) = \sum_{i \in [r]} \frac{1}{\alpha} M_{i\sigma(u_s)} \phi_j(i) = \frac{\mu_j}{\alpha} \phi_j(\sigma(u_s)),$$

where we used the fact that ϕ_j is a left-eigenvector of M associated with eigenvalue μ_j . Thus,

$$\begin{aligned} \mathbb{E}(\phi_k(\sigma(u))\phi_j(\sigma(v))|\mathcal{T}) &= \left(\frac{\mu_j}{\alpha}\right)^{t-1} \mathbb{E}(\phi_k(\sigma(u))\phi_j(\sigma(u))|\mathcal{T}) \\ &= \left(\frac{\mu_j}{\alpha}\right)^{t-1} \sum_{i \in [r]} \pi_i \phi_k(i)\phi_j(i) \\ &= 0, \end{aligned}$$

where the last equality follows from π -orthogonality (8) between vectors ϕ_k and ϕ_j for $j \neq k$. \square

We now show the following.

THEOREM 28. *Let $j \neq k \in [r]$ and $Z_0 = \delta_\iota$ where ι has distribution $(\pi(1), \dots, \pi(r))$. Then, for any $\ell \geq 0$,*

$$\mathbb{E}Q_{k,\ell}Q_{j,\ell} = 0.$$

PROOF. Write $Q_{k,\ell}$ as

$$Q_{k,\ell} = \sum_{(v,w) \in \mathcal{P}(\mathcal{T})} \phi_k(\sigma(w)),$$

where the sum extends over a set $\mathcal{P}(\mathcal{T})$ of node pairs (v, w) that depends only on the tree \mathcal{T} . Using the analogue expression for $Q_{j,\ell}$, one obtains

$$\begin{aligned} \mathbb{E}(Q_{k,\ell}Q_{j,\ell}|\mathcal{T}) &= \sum_{(v,w) \in \mathcal{P}(\mathcal{T})} \sum_{(v',w') \in \mathcal{P}(\mathcal{T})} \mathbb{E}(\phi_k(\sigma(w))\phi_j(\sigma(w'))|\mathcal{T}) \\ &= 0 \end{aligned}$$

by Lemma 27, equation (66). This completes the proof. \square

9. Local structure of random graphs. We now derive the necessary controls on the local structure of the SBM random graphs under consideration. Coupling results will allow to bound the deviation of their local structure from branching processes. Asymptotic independence between local neighborhoods of distinct nodes will then be used to establish weak laws of large numbers.

9.1. *Coupling.* For $e \in \vec{E}(V)$ and $f \in \vec{E}$, we define the ‘‘oriented’’ distance

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

where the minimum is taken over all self-avoiding paths $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_{\ell+1})$ in G such that $(\gamma_0, \gamma_1) = e$, $(\gamma_\ell, \gamma_{\ell+1}) = f$ and for all $1 \leq k \leq \ell$, $\{\gamma_k, \gamma_{k+1}\} \in E$.

Recall that $\vec{E}(V)$ is the set of oriented edges of the complete graph whereas \vec{E} is the set of oriented edges of G . In particular, we do not require that $e \in \vec{E}$. Observe that \vec{d} is not symmetric; we have instead $\vec{d}(e, f) = \vec{d}(f^{-1}, e^{-1})$.

Then, for integer $t \geq 0$, we introduce the vector $Y_t(e) = (Y_t(e)(i))_{i \in [r]}$ where, for $i \in [r]$,

$$(67) \quad Y_t(e)(i) = |\{f \in \vec{E} : \vec{d}(e, f) = t, \sigma(f_2) = i\}|.$$

We also set

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

The vector $Y_t(e)$ counts the types at oriented distance t from e .

We shall denote by $S_t(v)$ the set of vertices at distance t from v . We introduce

$$(68) \quad \begin{aligned} n(i) &= \sum_{v=1}^n \mathbb{1}(\sigma(v) = i), & \pi_n(i) &= \frac{n(i)}{n}, \\ \alpha_n(i) &= \sum_{j=1}^r \pi_n(j) W_{ij}, & \bar{\alpha}_n &= \max_{i \in [r]} \alpha_n(i) = \alpha + O(n^{-\gamma}), \end{aligned}$$

where at the last line we have used Assumptions (10)–(12). Central to our local study is the classical exploration process of the neighborhood of v which starts with $A_0 = \{v\}$ and at stage $t \geq 0$, if A_t is not empty, takes a vertex in A_t at minimal distance from v , say v_t , reveals its neighbors, say N_{t+1} , in $[n] \setminus D_t$ where $D_t = \bigcup_{0 \leq s \leq t} A_s$ the set of discovered vertices at time t , and updates $A_{t+1} = (A_t \cup N_{t+1}) \setminus \{v_t\}$. We will denote by \mathcal{F}_t the filtration generated by (A_0, \dots, A_t) . We start by establishing a rough bound on the growth of S_t .

LEMMA 29. *There exist $c_0, c_1 > 0$ such that, for all $s \geq 0$ and for any $w \in [n] \cup \vec{E}(V)$,*

$$\mathbb{P}(\forall t \geq 0 : S_t(w) \leq s \bar{\alpha}_n^t) \geq 1 - c_1 e^{-c_0 s}.$$

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

$$\mathbb{E} \max_{v \in [n], t \geq 0} \left(\frac{S_t(v)}{\bar{\alpha}_n^t} \right)^p \leq c(\log n)^p.$$

PROOF. Recall that $\mathbb{E}|X|^p = p \int_0^\infty x^{p-1} \mathbb{P}(|X| \geq t) dt$ and $\mathbb{P}(\max_v X_v \geq t) \leq 1 \wedge \sum_v \mathbb{P}(X_v \geq t)$. Then the second statement is a direct consequence of the first statement.

To prove the first statement, observe that, in the exploration process, given \mathcal{F}_t , if v_t has type j , the number of neighbors of v_t in $[n] \setminus D_t$ is upper bounded stochastically by

$$V_j = \sum_{i=1}^r V_{ij},$$

where $V_{ij} \stackrel{d}{=} \text{Bin}(n(i), W_{ij}/n) = \text{Bin}(n(i), \pi_n(i)W_{ij}/n(i))$ are independent. In particular, for any $\theta \geq 0$, using $1 + t \leq e^t$, we find

$$\mathbb{E}e^{\theta V_j} = \prod_{i=1}^r \mathbb{E} \left(1 - \frac{\pi_n(i)W_{ij}}{n(i)} + \frac{\pi_n(i)W_{ij}}{n(i)} e^{\theta} \right)^{n(i)} \leq e^{-\alpha_n(j) + \alpha_n(j)e^{\theta}} \leq e^{-\bar{\alpha}_n + \bar{\alpha}_n e^{\theta}}.$$

For any $j \in [r]$, we have thus bounded the characteristic function of V_j for $\theta \geq 0$ by the characteristic function of a $\text{Poi}(\bar{\alpha}_n)$ variable. It remains finally to repeat the proof of Lemma 23 from (57) with μ_1 replaced by $\bar{\alpha}_n$. \square

We now check that the random graph G is locally tree-like. For $v \in [n]$ and integer $h \geq 0$, we denote by $(G, v)_h$ the rooted subgraph of G rooted at v , spanned by the vertices at distance at most h from v . If $e = (u, v) \in \bar{E}(V)$, we set $(G, e)_h = (G', v)_h$ where G' is the graph G with the edge $\{u, v\}$ removed (if it was present in G).

LEMMA 30. *Let $\ell \sim \kappa \log_{\alpha} n$ with $\kappa < 1/2$. Then, w.h.p. the random graph G is ℓ -tangle-free and w.h.p. there are fewer than $\alpha^{2\ell} \log n$ vertices whose ℓ -neighborhood contains a cycle.*

PROOF. We start by proving the second statement. Let τ be defined as the first time t at which all nodes at distance ℓ or less from v have been discovered. It is clearly a stopping time for the filtration \mathcal{F}_t . By construction, given \mathcal{F}_{τ} , the set of discovered edges in $V(G, v)_{\ell}$ builds a spanning tree of $V(G, v)_{\ell}$. Also, given \mathcal{F}_{τ} , the number of undiscovered edges between two vertices in $V(G, v)_{\ell}$ is stochastically upper bounded by $\text{Bin}(m, a/n)$ where $m = |V(G, v)_{\ell}| = S_{\ell}(v)$ and $a = \max_{i,j} W(i, j)$. It follows from Lemma 29 that, for some $c > 0$,

$$(69) \quad \mathbb{P}((G, v)_{\ell} \text{ is not a tree}) \leq \frac{a \mathbb{E} S_{\ell}(v)}{n} \leq \frac{c \alpha^{\ell}}{n}.$$

Hence, from Markov's inequality,

$$\mathbb{P} \left(\sum_v \mathbb{1}((G, v)_{\ell} \text{ is not a tree}) \geq \alpha^{\ell} \log n \right) \leq \frac{c}{\log n}.$$

The second statement follows.

We now turn to the first statement. First, recall that the probability that $\text{Bin}(m, q)$ is not in $\{0, 1\}$ is at most $q^2 m(m-1) \leq q^2 m^2$. Also, if G is ℓ -tangled, then there exists $v \in [n]$ such that $V(G, v)_{\ell}$ has at least two undiscovered edges. In particular, from the union bound,

$$\mathbb{P}(G \text{ is } \ell\text{-tangled}) \leq \sum_{v=1}^n \frac{a^2 \mathbb{E} S_{\ell}(v)^2}{n^2} \leq \frac{c \alpha^{2\ell}}{n} = o(1),$$

where $c > 0$ and we have used again Lemma 29. \square

We conclude this subsection with a coupling of the process $Y_t(e)$ and a multi-type Galton–Watson tree. Recall that for probability measures P, Q on a countable set \mathcal{X} , the total variation distance is given by

$$d_{\text{TV}}(P, Q) = \frac{1}{2} \sum_{x \in \mathcal{X}} |P(x) - Q(x)| = \min \mathbb{P}(X \neq Y),$$

where the minimum is over all coupling (X, Y) such that $X \stackrel{d}{\sim} P, Y \stackrel{d}{\sim} Q$.

PROPOSITION 31. *Let $\ell \sim \kappa \log_\alpha n$ with $0 \leq \kappa < 1/2$ and $e = (u, v) \in \bar{E}(V)$. Let (T, o) be the random rooted tree associated to the Galton–Watson branching process defined in Section 8 started from $Z_0 = \delta_{\sigma(v)}$. The total variation distance between the law of $(G, e)_\ell$ and $(T, o)_\ell$ goes to 0 as $O((\log n)\alpha^\ell n^{-\gamma \wedge (1-\kappa)})$. The same holds with $(G, e)_\ell$ replaced by $(G, v)_\ell$.*

PROOF. We prove the first statement; the proof of the second statement is identical [see comment below (70)]. If $e = (u, v)$ and $G' = G \setminus \{u, v\}$, we consider the filtration \mathcal{F}_t associated to the exploration process of (G', v) . We let τ be the stopping time where all vertices of $(G, e)_\ell$ have been revealed. We set $y_0 = \delta_v$ and at step $t \geq 0$, we denote by $y_{t+1} = (y_{t+1}(1), \dots, y_{t+1}(r))$ the number of discovered neighbors of v_t in $[n] \setminus D_t$ of each type. If $\sigma(v_t) = j$ then, given \mathcal{F}_t , the variables $(y_{t+1}(i))_{i \in [r]}$ are independent and $y_{t+1}(i)$ has distribution $\text{Bin}(n_t(i), W_{ij}/n)$ where

$$(70) \quad n_t(i) = n(i) - \sum_{s=0}^t y_s(i) - \mathbb{1}(t=0, \sigma(u)=i)$$

[the last term comes from the difference between G and G' ; this term is not present in the case of the second statement on $(G, v)_\ell$]. We perform the same exploration process on (T, o) , that is a breath-first search of the tree, we discover at each step the offsprings say $x_{t+1} = (x_{t+1}(1), \dots, x_{t+1}(r))$ of the active vertex v_t . In particular, if v_t has type j then the variables $(x_{t+1}(i))_{i \in [r]}$ are conditionally independent and $x_{t+1}(i)$ has distribution $\text{Poi}(\pi(i)W_{ij})$. To couple the two processes, we shall use the following classical bounds (see, e.g., [2]):

$$(71) \quad d_{\text{TV}}\left(\text{Bin}\left(m, \frac{\lambda}{m}\right), \text{Poi}(\lambda)\right) \leq \frac{\lambda}{m} \quad \text{and} \quad d_{\text{TV}}(\text{Poi}(\lambda), \text{Poi}(\lambda')) \leq |\lambda - \lambda'|.$$

For $0 \leq t \leq \tau$, define the event $\Omega_t = \{|D_t| \leq c\alpha^\ell \log n\} \in \mathcal{F}_t$, where D_t is the set of discovered vertices. By Lemma 29, for c large enough, $\tau \leq c\alpha^\ell \log n$ and Ω_τ holds with probability larger than $1 - 1/n$. Also, by (69), with probability at least $1 - c\alpha^\ell/n$, $(G, e)_\ell$ is a tree. It follows that by iteration, it is enough to check that, if Ω_t holds, there exists $C > 0$ such that

$$(72) \quad d_{\text{TV}}(P_{t+1}, Q_{t+1}) \leq Cn^{-\gamma \wedge (1-\kappa)},$$

where P_{t+1} is the distribution of y_{t+1} under $\mathbb{P}(\cdot|\mathcal{F}_t)$ and Q_{t+1} has the law of $x_{t+1} = (x_1, \dots, x_r)$ where x_i are independent with distribution $\text{Poi}(\pi(i)W_{ij})$, with $j = \sigma(v_t)$. However, from (71) and the triangle inequality, we have

$$\begin{aligned} d_{\text{TV}}(P_{t+1}, Q_{t+1}) &\leq d_{\text{TV}}\left(P_{t+1}, \bigotimes_{i \in [r]} \text{Poi}\left(\frac{n_t(i)W_{ij}}{n}\right)\right) \\ &\quad + d_{\text{TV}}\left(\bigotimes_{i \in [r]} \text{Poi}\left(\frac{n_t(i)W_{ij}}{n}\right), Q_{t+1}\right) \\ &\leq \sum_{i=1}^r \left(\frac{W_{ij}}{n} + W_{ij} \left|\frac{n_t(i)}{n} - \pi(i)\right|\right) \\ &\leq C \frac{\alpha^\ell (\log n)}{n} + \sum_{i=1}^r W_{ij} |\pi_n(i) - \pi(i)|. \end{aligned}$$

From (12), the latter is $O(n^{-\gamma \wedge (1-\kappa)})$. We thus have proved that (72) holds for some new $C > 0$. \square

We will use the following corollary of Proposition 31.

COROLLARY 32. *Let $\ell \sim \kappa \log_\alpha n$ with $0 < \kappa < \gamma \wedge 1/2$. For $e \in \vec{E}(V)$, we define the event $\mathcal{E}(e)$ that for all $0 \leq t < \ell$ and $k \in [r]$: $|\langle \phi_k, Y_t(e) \rangle - \mu_k^{t-\ell} \langle \phi_k, Y_\ell(e) \rangle| \leq (\log n)^4 \alpha^{t/2}$, if $k \in [r_0]$, and $|\langle \phi_k, Y_t(e) \rangle| \leq (\log n)^4 \alpha^{t/2}$, if $k \in [r] \setminus [r_0]$.*

Then w.h.p. the number of edges $e \in \vec{E}$ such that $\mathcal{E}(e)$ does not hold is at most $(\log n)^2 \alpha^\ell n^{1-\gamma}$.

PROOF. First, with exponentially small probability there are fewer than $2\alpha n$ edges in \vec{E} . From the union bound, it is thus enough to prove that for any $e \in \vec{E}(V)$, $\mathbb{P}(\mathcal{E}(e)^c) \leq C(\log n)\alpha^\ell n^{-\gamma}$. To show this, use the coupling result of Proposition 31 to deduce that with probability at least $1 - C(\log n)\alpha^\ell n^{-\gamma}$, the processes $(Y_t(e))_{0 \leq t \leq \ell}$ and $(Z_t)_{0 \leq t \leq \ell}$ coincide. It then remains to use Theorem 24 with $\beta = 1$. \square

9.2. *Geometric growth of linear functions of nonbacktracking walks.* For $k \in [r]$, we recall that

$$\chi_k(e) = \phi_k(\sigma(e_2)).$$

The next proposition asserts that for most $e \in \vec{E}$, $\langle B^t \chi_k, \delta_e \rangle$ grows nearly geometrically in t with rate μ_k up to an error of order $\alpha^{t/2}$.

PROPOSITION 33. *Let $\ell \sim \kappa \log_\alpha n$ with $0 < \kappa < \gamma \wedge 1/2$. There exists a random subset of edges $\vec{E}_\ell \subset \vec{E}$ such that w.h.p. the following hold:*

(i) for all $e \in \vec{E} \setminus \vec{E}_\ell$, $0 \leq t \leq \ell$,

$$\begin{aligned} |\langle B^t \chi_k, \delta_e \rangle - \mu_k^{t-\ell} \langle B^\ell \chi_k, \delta_e \rangle| &\leq (\log n)^4 \alpha^{t/2} && \text{if } k \in [r_0], \\ |\langle B^t \chi_k, \delta_e \rangle| &\leq (\log n)^4 \alpha^{t/2} && \text{if } k \in [r] \setminus [r_0], \end{aligned}$$

(ii) for all $e \in \vec{E}_\ell$, $0 \leq t \leq \ell$ and $k \in [r]$,

$$|\langle B^t \chi_k, \delta_e \rangle| \leq (\log n)^2 \alpha^t,$$

(iii) $|\vec{E}_\ell| \leq (\log n)^3 \alpha^\ell n^{1-\gamma}$.

PROOF. We define \vec{E}_ℓ as the set of oriented edges such that either $(G, e_2)_\ell$ is not a tree or the event $\mathcal{E}(e)$ defined in Corollary 32 does not hold. Then by Lemma 30 and Corollary 32, w.h.p. \vec{E}_ℓ satisfies condition (iii). Moreover, by definition if $(G, e_2)_\ell$ is a tree

$$(73) \quad \langle B^t \chi_k, \delta_e \rangle = \langle \phi_k, Y_t(e) \rangle,$$

and statement (i) follows from Corollary 32. For statement (ii), we simply use that w.h.p. G is tangle free (by Lemma 30), hence, there are at most two nonbacktracking walks of length t from e to any f . We get

$$\langle B^t \chi_k, \delta_e \rangle \leq 2 \|\phi_k\|_\infty S_t(e).$$

However, by Lemma 29 w.h.p. for all $t \geq 0$ and all $e \in \vec{E}$, $|S_t(e)| \leq C(\log n) \alpha^t$. \square

COROLLARY 34. *Let $\ell \sim \kappa \log_\alpha n$ with $0 < \kappa < \gamma/2$. With high probability, for any $0 \leq t \leq \ell - 1$ and $k \in [r]$,*

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

PROOF. We write

$$\langle B^t \chi_k, x \rangle = \sum_{e \in \vec{E}_\ell} x_e \langle B^t \chi_k, \delta_e \rangle + \sum_{e \notin \vec{E}_\ell} x_e \langle B^t \chi_k, \delta_e \rangle = I + J.$$

From the Cauchy–Schwarz inequality, the first term is bounded w.h.p. by

$$\begin{aligned} |I| &\leq (\log n)^2 \alpha^t \sum_{e \in \vec{E}_\ell} |x_e| \leq (\log n)^2 \alpha^t \sqrt{|\vec{E}_\ell|} \\ &\leq (\log n)^4 \alpha^t \alpha^{\ell/2} n^{(1-\gamma)/2} = o(n^{1/2} \alpha^{t/2}), \end{aligned}$$

where we have used that $\alpha^{t/2}\alpha^{\ell/2}n^{-\gamma/2} \leq n^{\kappa-\gamma/2+o(1)}$ and $\kappa < \gamma/2$. For the second term, if $k \in [r_0]$, using $\langle B^\ell \chi_k, x \rangle = 0$, we get similarly w.h.p.

$$\begin{aligned} |J| &\leq \mu_k^{t-\ell} \sum_{e \in \vec{E}_\ell} |x_e| |\langle B^\ell \chi_k, \delta_e \rangle| + \sum_{e \notin \vec{E}_\ell} |x_e| |\langle B^t \chi_k, \delta_e \rangle - \mu_k^{t-\ell} \langle B^\ell \chi_k, \delta_e \rangle| \\ &\leq (\log n)^2 \alpha^{t-\ell} \alpha^\ell \alpha^{\ell/2} n^{(1-\gamma)/2} + (\log n)^4 \sqrt{|\vec{E}|} \alpha^{t/2} \\ &\leq (\log n)^2 n^{1/2} \alpha^{t/2} \alpha^\ell n^{-\gamma/2} + C (\log n)^4 n^{1/2} \alpha^{t/2}. \end{aligned}$$

Finally, if $k \in [r] \setminus [r_0]$, we simply write w.h.p. $|J| \leq \sum_{e \notin \vec{E}_\ell} |x_e| |\langle B^t \chi_k, \delta_e \rangle| \leq (\log n)^4 n^{1/2} \alpha^{t/2}$. \square

9.3. Laws of large numbers for local functions. We first prove weak laws of large numbers for general local functionals of SBM random graphs that will then be applied to specific functionals of interest.

9.3.1. Weak laws of large numbers for local functionals: Convergence speed. We start with a general variance bound for local functions of an inhomogeneous random graph. A colored graph is a graph $G = (V, E)$ with a map $\sigma : V \rightarrow [r]$. We denote by \mathcal{G}^* the set of rooted colored graphs, that is, the set of pairs (G, o) formed by a colored graph G and a distinguished vertex $o \in V$. We shall say that a function τ from \mathcal{G}^* to \mathbb{R} is ℓ -local, if $\tau(G, o)$ is only function of $(G, o)_\ell$.

PROPOSITION 35. *There exists $c > 0$ such that if $\tau, \varphi : \mathcal{G}^* \rightarrow \mathbb{R}$ are ℓ -local, $|\tau(G, o)| \leq \varphi(G, o)$ and φ is nondecreasing by the addition of edges, then*

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

PROOF. We first bound the expectation of

$$Z = \sum_{u=1}^n \Sigma^2(G, u),$$

where $\Sigma(G, u)$ is defined as the number of vertices at distance ℓ from $u \in V$. By Lemma 29, for any $u \in [n]$,

$$\mathbb{P} \left(\Sigma(G, u) \geq s \frac{\bar{\alpha}_n^{\ell+1} - 1}{\bar{\alpha}_n - 1} \right) \leq c_1 e^{-c_0 s}.$$

Hence, for any $p \geq 1$, for some $c_p > 0$,

$$\mathbb{E} Z^p \leq n^{p-1} \mathbb{E} \sum_{u=1}^n \Sigma^{2p}(G, u) \leq c_p n^p \bar{\alpha}_n^{2\ell p}.$$

Now, for $1 \leq k \leq n$, let $X_k = \{1 \leq v \leq k : \{v, k\} \in E\}$, where E is the edge set of G . The vector (X_1, \dots, X_n) is an independent vector and for some function F ,

$$Y := \sum_{u=1}^n \tau(G, u) = F(X_1, \dots, X_n).$$

We also define G_k as the graph with edge set $\bigcup_{v \neq k} X_v$. We set

$$Y_k = \sum_{u=1}^n \tau(G_k, u).$$

Since τ is ℓ -local, we observe that $\tau(G, u) - \tau(G_k, u)$ can be nonzero only if $u \in V((G, k)_\ell)$ and it is bounded by $\Lambda = 2 \max_{u \in [n]} \varphi(G, u)$. Consequently,

$$\sum_{k=1}^n |Y - Y_k|^2 \leq \sum_{k=1}^n \Sigma^2(G, k) \Lambda^2 = Z \Lambda^2.$$

Finally, we conclude by using Efron–Stein’s inequality: $\text{Var}(Y) \leq \mathbb{E} Z \Lambda^2 \leq \sqrt{\mathbb{E} Z^2} \sqrt{\mathbb{E} \Lambda^4}$. \square

We now apply the above proposition and Proposition 31 to show that the SBM random graph with uniform root selection converges weakly to the multitype Galton–Watson process previously studied. The established convergence implies convergence for the local weak topology of Benjamini and Schramm (see [3]). Crucially, we are able to consider local functions with logarithmic distance parameter ℓ and obtain bounds on the convergence speed.

PROPOSITION 36. *Let $\ell \sim \kappa \log_\alpha n$ with $0 < \kappa < 1/2$. There exists $c > 0$, such that if $\tau, \varphi : \mathcal{G}^* \rightarrow \mathbb{R}$ are ℓ -local, $|\tau(G, o)| \leq \varphi(G, o)$ and φ is nondecreasing by the addition of edges, then if $\mathbb{E} \varphi(T, o)$ is finite,*

$$\begin{aligned} & \mathbb{E} \left| \frac{1}{n} \sum_{v=1}^n \tau(G, v) - \mathbb{E} \tau(T, o) \right| \\ & \leq c \frac{\alpha^{\ell/2} \sqrt{\log n}}{n^{\gamma/2}} \left(\left(\mathbb{E} \max_{v \in [n]} \varphi^4(G, v) \right)^{1/4} \vee \left(\mathbb{E} \varphi^2(T, o) \right)^{1/2} \right), \end{aligned}$$

where (T, o) is the random rooted tree associated to the Galton–Watson branching process defined in Section 8 started from $Z_0 = \delta_\iota$ and ι has distribution $(\pi(1), \dots, \pi(r))$.

PROOF. In view of Proposition 35 and Jensen’s inequality, it is sufficient to prove that

$$\begin{aligned} (74) \quad & \left| \frac{1}{n} \sum_{v=1}^n \mathbb{E} \tau(G, v) - \mathbb{E} \tau(T, o) \right| \\ & = O \left(\frac{\alpha^{\ell/2}}{n^\beta} \sqrt{\log n} \left(\max_{v \in [n]} \mathbb{E} \varphi^2(G, v) \vee \mathbb{E} \varphi^2(T, o) \right)^{1/2} \right), \end{aligned}$$

with $\beta = (\gamma \wedge (1 - \kappa))/2$. For $i \in [r]$, we set $\tau_i = \tau(T_i, o)$ where (T_i, o) has distribution the the random tree (T, o) started from $Z_0 = \delta_i$. Let $v \in V$ with $\sigma(v) = i$. We denote by χ_v the indicator function that the coupling of $(G, v)_\ell$ and $(T_i, o)_\ell$ described in Proposition 31 is not successful. We have, from the Cauchy–Schwarz inequality

$$\begin{aligned} |\mathbb{E}\tau(G, v) - \mathbb{E}\tau_i| &= |\mathbb{E}\chi_v \tau(G, v) - \mathbb{E}\chi_v \tau_i| \\ &\leq 2\sqrt{(\mathbb{E}\chi_v)(\mathbb{E}\varphi(G, v)^2 \vee \mathbb{E}\varphi(T_i, o)^2)} \\ &= O(\alpha^{\ell/2} n^{-\beta} \sqrt{\log n} \sqrt{\mathbb{E}\varphi(G, v)^2 \vee \mathbb{E}\varphi(T_i, o)^2}). \end{aligned}$$

Let v_1, \dots, v_r be fixed vertices such that $\sigma(v_i) = i$ [since $\pi(i) > 0$ such v_i exists for n large enough]. We recall that $\mathbb{E}\tau(G, v)$ depends only on $\sigma(v)$. Hence, using (12),

$$\begin{aligned} \mathbb{E} \frac{1}{n} \sum_{v \in V} \tau(G, v) &= \sum_{i=1}^r \frac{n(i)}{n} \mathbb{E}\tau(G, v_i) \\ &= \sum_{i=1}^r \{(\pi(i) + O(n^{-\nu})) \mathbb{E}\tau_i \\ &\quad + O(\alpha^{\ell/2} n^{-\beta} \sqrt{\log n} \sqrt{\mathbb{E}\varphi(G, v_i)^2 \vee \mathbb{E}\varphi(T_i, o)^2})\} \\ &= \mathbb{E}\tau(T, o) + O\left(\alpha^{\ell/2} n^{-\beta} \sqrt{\log n} \sum_{i=1}^r \sqrt{\mathbb{E}\varphi(G, v_i)^2 \vee \mathbb{E}\varphi(T_i, o)^2}\right). \end{aligned}$$

This completes the proof of (74). \square

9.3.2. Law of large numbers for specific local functions. We will now apply Proposition 36 to deduce weak laws of large numbers for expressions closely related to $\langle B^\ell \chi_k, B^\ell \chi_j \rangle$, $\langle B^{2\ell} \chi_k, B^\ell \chi_j \rangle$ and $\langle B^\ell B^{*\ell} \check{\chi}_k, B^\ell B^{*\ell} \check{\chi}_j \rangle$. Recall the definition of $Y_\ell(e)$ in (67).

PROPOSITION 37. *Let $\ell \sim \kappa \log_\alpha n$ with $0 < \kappa < \gamma/4$:*

(i) *For any $k \in [r_0]$, there exists $\rho_k > 0$ such that, in probability,*

$$\frac{1}{\alpha n} \sum_{e \in \vec{E}} \frac{\langle \phi_k, Y_\ell(e) \rangle^2}{\mu_k^{2\ell}} \rightarrow \rho_k.$$

(ii) For any $k \in [r] \setminus [r_0]$, there exists $\rho_k > 0$ such that w.h.p.

$$\frac{1}{\alpha n} \sum_{e \in \bar{E}} \frac{\langle \phi_k, Y_\ell(e) \rangle^2}{\alpha^\ell} \geq \rho_k.$$

(iii) For any $k \neq j \in [r]$,

$$\mathbb{E} \left| \frac{1}{\alpha n} \sum_{e \in \bar{E}} \langle \phi_k, Y_\ell(e) \rangle \langle \phi_j, Y_\ell(e) \rangle \right| = O(\alpha^{5\ell/2} n^{-\gamma/2} (\log n)^{5/2}).$$

(iv) For any $k \neq j \in [r]$,

$$\mathbb{E} \left| \frac{1}{\alpha n} \sum_{e \in \bar{E}} \langle \phi_k, Y_{2\ell}(e) \rangle \langle \phi_j, Y_\ell(e) \rangle \right| = O(\alpha^{7\ell/2} n^{-\gamma/2} (\log n)^{5/2}).$$

(v) For any $k \in [r_0]$, in probability

$$\frac{1}{\alpha n} \sum_{e \in \bar{E}} \frac{\langle \phi_k, Y_{2\ell}(e) \rangle \langle \phi_k, Y_\ell(e) \rangle}{\mu_k^{3\ell}} \rightarrow \rho_k.$$

PROOF. Let Z_t , $t \geq 0$, be the Galton–Watson branching process defined in Section 8 started from $Z_0 = \delta_t$ and t has distribution $(\pi(1), \dots, \pi(r))$. We denote by (T, o) the associated random rooted tree. If $k \in [r_0]$, by Theorem 21, for some $\rho_k > 0$,

$$\mathbb{E} \langle \phi_k, Z_\ell \rangle^2 \mu_k^{-2\ell} = \rho_k + o(1).$$

We set $\tau(G, v) = \sum_{e \in \bar{E}: e_2=v} \langle \phi_k, Y_\ell(e) \rangle^2 \mu_k^{-2\ell}$. We observe that

$$\sum_{e \in \bar{E}: e_2=v} \langle \phi_k, Y_\ell(e) \rangle^2 \leq \sum_{e \in \bar{E}: e_2=v} S_\ell^2(e) \leq S_{\ell+1}^2(v).$$

We get, using $\mu_k^{-2} < \alpha^{-1}$, that $\tau(G, v) \leq \varphi(G, v) := \alpha^{-\ell} S_{\ell+1}^2(v)$. Also, Lemma 29 implies that $\mathbb{E} \max_v \varphi(G, v)^4 = O((\log n)^8 \alpha^{4\ell})$. The same upper bound holds for $\varphi(T, o)$ by Lemma 23. We deduce from Proposition 36 that

$$\mathbb{E} \left| \frac{1}{\alpha n} \sum_{e \in \bar{E}} \frac{\langle \phi_k, Y_\ell(e) \rangle^2}{\mu_k^{2\ell}} - \frac{\mathbb{E} \langle \phi_k, Z_\ell \rangle^2}{\mu_k^{2\ell}} \right| = O(\alpha^{3\ell/2} (\log n)^{5/2} n^{-\gamma/2}).$$

It proves statement (i) of the proposition.

For statement (ii), we use Theorem 22 instead. We denote by X_k the limit martingale in Theorem 22 when $Z_0 = \delta_t$. If $\mu_k^2 < \mu_1$, we find similarly that for any $\theta > 0$,

$$\frac{1}{\alpha n} \sum_{e \in \bar{E}} \left(\frac{\langle \phi_k, Y_\ell(e) \rangle}{\alpha^{\ell/2}} \right)^2 \wedge \theta$$

converges in L^1 to $\mathbb{E}(|X_k|^2 \wedge \theta)$. From Theorem 22, the latter is positive if θ is large enough. In the case $\mu_k^2 = \mu_1$, by Theorem 22, we need to normalize by $\alpha^{\ell/2} \ell^{1/2}$.

For (iii), we use Lemma 27. We set $\tau(G, v) := \sum_{e \in \vec{E}: e_2=v} \langle \phi_k, Y_\ell(e) \rangle \langle \phi_j, Y_\ell(e) \rangle$. Arguing as above, we have $\tau(G, v) \leq \varphi(G, v) = S_{\ell+1}^2(v)$ and $\mathbb{E} \max_v \varphi(G, v)^4$ is $O((\log n)^8 \alpha^{8\ell})$, by Lemma 29. The same upper bound holds for $\varphi(T, o)$ from Lemma 23. It remains to apply Proposition 36.

For statement (iv), we have $\mathbb{E} \langle \phi_k, Z_{2\ell} \rangle \langle \phi_j, Z_\ell \rangle = \mu_k^\ell \mathbb{E} \langle \phi_k, Z_\ell \rangle \langle \phi_j, Z_\ell \rangle = 0$ by Lemma 27. We set $\tau(G, v) = \sum_{e \in \vec{E}: e_2=v} \langle \phi_k, Y_{2\ell}(e) \rangle \langle \phi_j, Y_\ell(e) \rangle$. We have

$$\tau(G, v) \leq \varphi(G, v) = \sum_{e \in \vec{E}: e_2=v} S_{2\ell}(e) S_\ell(e) \leq S_{2\ell+1}(v) S_{\ell+1}(v).$$

Moreover, $\mathbb{E} \max_v \varphi(G, v)^4 = O((\log n)^8 \alpha^{12\ell})$ by Lemma 29 and the Cauchy–Schwarz inequality. The same upper bound holds for $\varphi(T, o)$ from Lemma 23. It remains to apply Proposition 36.

Finally, for statement (v), $\mathbb{E} \langle \phi_k, Z_{2\ell} \rangle \langle \phi_k, Z_\ell \rangle = \mu_k^\ell \mathbb{E} \langle \phi_k, Z_\ell \rangle^2 = \mu_k^{3\ell} (\rho_k + o(1))$. We use $\mu_k^2 > \alpha$ and then repeat the proof of statement (iv), we get

$$\mathbb{E} \left| \frac{1}{\alpha n} \sum_{e \in \vec{E}} \frac{\langle \phi_k, Y_{2\ell}(e) \rangle \langle \phi_k, Y_\ell(e) \rangle}{\mu_k^{3\ell}} - \rho_k \right| = O(\alpha^{7\ell/2 - 3\ell/2} n^{-\gamma/2} (\log n)^{5/2}).$$

Since $\kappa < \gamma/4$, the right-hand side is $o(1)$. \square

We conclude this subsection by estimates on quantities which are closely related to $B^\ell B^{*\ell} \check{\chi}_k$. For $e \in \vec{E}(V)$, we define for $t \geq 0$, $\mathcal{Y}_t(e) = \{f \in \vec{E} : \vec{d}(e, f) = t\}$. For $k \in [r]$, we set

$$(75) \quad P_{k,\ell}(e) = \sum_{t=0}^{\ell-1} \sum_{f \in \mathcal{Y}_t(e)} L_k(f),$$

where

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

and $\tilde{Y}_t(g)$, $\tilde{S}_{\ell-t-1}(h) = \|\tilde{Y}_{\ell-t-1}(h)\|_1$ are the variables $Y_t(g)$, $S_{\ell-t-1}(h)$ defined on the graph G where all edges in $(G, e_2)_t$ have been removed. In particular, if $(G, e)_{2\ell}$ is a tree, $\tilde{Y}_s(g)$ and $Y_s(g)$ coincide for $s \leq 2\ell - t$.

We also define

$$(76) \quad S_{k,\ell}(e) = S_\ell(e) \phi_k(\sigma(e_1)).$$

As can be seen from (60), when $(G, e_2)_{2\ell}$ is a tree, it follows that

$$(77) \quad B^\ell B^{*\ell} \check{\chi}_k(e) = P_{k,\ell}(e) + S_{k,\ell}(e).$$

This is the reason why controls of quantities $P_{k,\ell}$ and $S_{k,\ell}$, themselves based on the previous branching analysis of quantities $Q_{k,\ell}$ and S_ℓ , respectively, will be instrumental in our analysis of vectors $B^\ell B^{*\ell} \check{\chi}_k$; see Lemma 39 below. We have the following.

PROPOSITION 38. *Let $\ell \sim \kappa \log_\alpha n$ with $0 < \kappa < \gamma/5$:*

(i) *For any $k \in [r_0]$, there exists $\rho'_k > 0$ such that, w.h.p.,*

$$\frac{1}{\alpha n} \sum_{e \in \bar{E}} \frac{P_{k,\ell}^2(e)}{\mu_k^{4\ell}} \rightarrow \rho'_k.$$

(ii) *For any $k \in [r] \setminus [r_0]$, there exists $C_k > 0$ such that, w.h.p.,*

$$\frac{1}{\alpha n} \sum_{e \in \bar{E}} \frac{P_{k,\ell}^2(e)}{\alpha^{2\ell} (\log n)^5} \leq C_k.$$

(iii) *For any $k \neq j \in [r]$, for some $c > 0$,*

$$\mathbb{E} \left| \frac{1}{\alpha n} \sum_{e \in \bar{E}} (P_{k,\ell}(e) + S_{k,\ell}(e))(P_{j,\ell}(e) + S_{j,\ell}(e)) \right| = O(\alpha^{9\ell/2} n^{-\gamma/2} (\log n)^c).$$

PROOF. Let Z_t , $t \geq 0$, be the Galton–Watson branching process defined in Section 8 started from $Z_0 = \delta_t$ and ι has distribution $(\pi(1), \dots, \pi(r))$. We denote by (T, o) the associated random rooted tree. For any $k \in [r_0]$, by Theorem 25, for some positive constant ρ'_k ,

$$(78) \quad \frac{\mathbb{E}[Q_{k,\ell}^2]}{\mu_k^{2\ell}} = \rho'_k + o(1).$$

On the other hand, Theorem 25 also ensures that for some $C_k > 0$, for $k \in [r] \setminus [r_0]$,

$$(79) \quad \frac{\mathbb{E}[Q_{k,\ell}^2]}{\alpha^{2\ell} (\log n)^5} \leq 2C_k.$$

Let us denote here by \mathcal{F}_t the σ -algebra spanned by $(G, e)_t$, given \mathcal{F}_{t+1} and $g \in (G, e)_{t+1}$. By a monotonicity argument, the statement of Lemma 29 applies to variables $(\tilde{S}_s(g), s \geq 0)$. Thus, for any $p \geq 1$, there is a constant $c > 0$ such that for any integer $s \geq 0$, $\mathbb{E}[\tilde{S}_s(g)^p | \mathcal{F}_{t+1}] \leq c \bar{\alpha}_n^{sp}$. We thus have, repeating the proof of Lemma 26, that for any fixed $p \geq 1$,

$$\mathbb{E} P_{k,\ell}(e)^p \leq C(\bar{\alpha}_n)^{2\ell p} = O(\alpha^{2\ell p}).$$

Then the argument in the proof of Proposition 37 can be applied. For statement (i), we let $k \in [r_0]$ and define $\tau(G, v) = \sum_{e \in \bar{E}, e_2=v} P_{k,\ell}^2(e) \mu_k^{-4\ell}$. Let

$$M(v) = \max_{0 \leq t \leq \ell} \max_{u \in (G, v)_t} \max_{s \leq 2\ell - t} (S_s(u) / \alpha^s).$$

Since $\mu_k^2 > \alpha$, we have the rough bound

$$\begin{aligned} \tau(G, v) &\leq \alpha^{-2\ell} \sum_{e \in \vec{E}, e_2=v} \left(\sum_{t=0}^{\ell-1} \sum_{f \in \mathcal{Y}_t(e)} M^2(v) \alpha^{t+1} \alpha^{\ell-t} \right)^2 \\ &= (M^2(v) \alpha)^2 \sum_{e \in \vec{E}, e_2=v} \left(\sum_{t=0}^{\ell-1} \sum_{f \in \mathcal{Y}_t(e)} 1 \right)^2. \end{aligned}$$

Hence, $\tau(G, v) \leq C M(v)^6 \alpha^{2\ell} = \varphi(G, v)$. However, we have by Lemma 29 that $\mathbb{E} \max_v \varphi(G, v)^4 = O((\log n)^{24} \alpha^{8\ell})$. By Lemma 23, the same bound holds for $\varphi(T, o)$. We deduce from Proposition 36 that

$$\mathbb{E} \left| \frac{1}{\alpha n} \sum_{e \in \vec{E}} \frac{P_{k,\ell}^2(e)}{\mu_k^{4\ell}} - \frac{\mathbb{E} Q_{k,\ell}^2}{\mu_k^{4\ell}} \right| = O(\alpha^{5\ell/2} n^{-\gamma/2} (\log n)^c).$$

Since $\kappa < \gamma/5$, the right-hand side goes to 0 and statement (i) follows from (78). Statement (ii) is proved similarly using (79). For statement (iii), we use the above computation, together with Theorem 28 and (8). It gives the claimed bound. \square

9.4. *Proof of Proposition 19.* For the next lemma, recall the definitions (75)–(76) of the vectors $P_{k,\ell}$ and $S_{k,\ell}$ in $\mathbb{R}^{\vec{E}}$. We also introduce the vector in $\mathbb{R}^{\vec{E}}$, for $k \in [r]$,

$$N_{k,\ell}(e) = \langle \phi_k, Y_\ell(e) \rangle.$$

LEMMA 39. *Let $\ell \sim \kappa \log_\alpha n$ with $0 < \kappa < \gamma \wedge 1/2$. Then, w.h.p. $\|B^\ell \chi_k - N_{k,\ell}\| = O((\log n)^{5/2} \alpha^{3\ell/2}) = o(\alpha^{\ell/2} \sqrt{n})$, $\|B^\ell B^{*\ell} \check{\chi}_k - P_{k,\ell} - S_{k,\ell}\| = O((\log n)^4 \alpha^{3\ell})$ and $\|B^\ell B^{*\ell} \check{\chi}_k - P_{k,\ell}\| = O(\alpha^\ell \sqrt{n})$.*

PROOF. Let \vec{E}_ℓ be as in Proposition 33 and let $\vec{E}'_\ell \subset \vec{E}_\ell$ be the subset of edges such that $(G, e_2)_\ell$ is a not tree. From (73), Lemma 29 and Proposition 33 we have, w.h.p.,

$$\begin{aligned} \|N_{k,\ell} - B^\ell \chi_k\|^2 &= \sum_{e \in \vec{E}'_\ell} (\langle \phi_k, Y_\ell(e) \rangle - \langle B^\ell \chi_k, \delta_e \rangle)^2 \\ &\leq 2 \sum_{e \in \vec{E}'_\ell} (S_\ell(e)^2 + \langle B^\ell \chi_k, \delta_e \rangle^2) \\ &= O(|\vec{E}'_\ell| \log(n)^4 \alpha^{2\ell}) = O((\log n)^5 \alpha^{3\ell}), \end{aligned}$$

where at the last line, we have used Lemma 30. Since $\kappa < 1/2$, it proves the first statement.

Similarly, for the second statement, recall as stated in (77) that when $(G, e_2)_{2\ell}$ is a tree, then $B^\ell B^{*\ell} \check{\chi}_k(e) = P_{k,\ell}(e) + S_{k,\ell}(e)$. Let $\vec{E}'_{2\ell} \subset \vec{E}_{2\ell}$ be the subset of edges such that $(G, e_2)_{2\ell}$ is not a tree. If G is 2ℓ -tangle-free, then there are at most two different nonbacktracking paths between two edges. Hence, if $e \in \vec{E}'_{2\ell}$,

$$|B^\ell B^{*\ell} \check{\chi}_k(e)| \leq 2\|\phi_k\|_\infty (P_{1,\ell}(e) + S_\ell(e)) \leq 2(P_{1,\ell}(e) + S_\ell(e)).$$

Now, by Lemma 29, w.h.p. $S_\ell(e) \leq C(\log n)\alpha^\ell$. Moreover, if $M = \max_{v,t \leq \ell} S_t(v)/\alpha^t \leq C \log n$, $P_{1,\ell}(e) \leq \sum_{t=0}^{\ell-1} \sum_{f \in \mathcal{Y}_t(e)} \alpha^{t+1} \alpha^{\ell-t} M^2 \leq M^3 \sum_{t=0}^{\ell-1} \alpha^{t+\ell+1}$. So finally, w.h.p. for all $e \in \vec{E}'_{2\ell}$,

$$|B^\ell B^{*\ell} \check{\chi}_k(e)| = O((\log n)^3 \alpha^{2\ell}).$$

Hence, by Lemma 30, w.h.p.,

$$\|B^\ell B^{*\ell} \check{\chi}_k - P_{k,\ell} - S_{k,\ell}\| = O\left(\sqrt{|E'_{2\ell}|} (\log n)^3 \alpha^{2\ell}\right) = O((\log n)^4 \alpha^{3\ell}).$$

On the other hand, from Proposition 37(i), w.h.p.,

$$(80) \quad \|S_{k,\ell}\| = O(\sqrt{n} \alpha^\ell).$$

The conclusion follows since $\kappa < 1/2$. \square

All ingredients are finally gathered to prove Proposition 19.

PROOF OF PROPOSITION 19. We use the notation of Lemma 39.

Proof of (i). Let $k \in [r_0]$. By definition,

$$\theta_k = \frac{\|B^\ell B^{*\ell} \check{\chi}_k\|}{\|B^\ell \chi_k\|}.$$

From Proposition 37(i) and Proposition 38(i), respectively, for some positive constants c_0, c_1 , w.h.p.,

$$\frac{c_0}{2} \leq \frac{\|N_{k,\ell}\|}{\sqrt{n} \mu_k^\ell} \leq 2c_1 \quad \text{and} \quad \frac{c_0}{2} \leq \frac{\|P_{k,\ell}\|}{\sqrt{n} \mu_k^{2\ell}} \leq 2c_1.$$

It remains to use Lemma 39 and the assumption $\mu_k^2 > \alpha$. We find, w.h.p.,

$$(81) \quad c_0 \leq \frac{\|B^\ell \chi_k\|}{\sqrt{n} \mu_k^\ell} \leq c_1 \quad \text{and} \quad c_0 \leq \frac{\|B^\ell B^{*\ell} \check{\chi}_k\|}{\sqrt{n} \mu_k^{2\ell}} \leq c_1.$$

Proof of (iii). Let $k \in [r] \setminus [r_0]$. From Proposition 37(ii) w.h.p. $\|N_{k,\ell}\| \geq (c_0/2)\sqrt{n}\alpha^{\ell/2}$ and from Proposition 38(ii) w.h.p. $\|P_{k,\ell}\| = O(\sqrt{n}(\log n)^{5/2}\alpha^\ell)$. Using Lemma 39, we find, w.h.p.,

$$(82) \quad c_0 \leq \frac{\|B^\ell \chi_k\|}{\sqrt{n}\alpha^{\ell/2}} \quad \text{and} \quad \frac{\|B^\ell B^{*\ell} \check{\chi}_k\|}{\sqrt{n}(\log n)^{5/2}\alpha^\ell} \leq c_1.$$

Proof of (ii). Let $k \in [r_0]$. Since $Px = \check{x}$ is an isometry,

$$\langle \zeta_k, \check{\varphi}_k \rangle = \frac{\langle B^\ell B^{*\ell} \check{\chi}_k, B^{*\ell} \check{\chi}_k \rangle}{\|B^\ell B^{*\ell} \check{\chi}_k\| \|B^{*\ell} \check{\chi}_k\|} = \frac{\langle B^\ell \chi_k, B^{2\ell} \chi_k \rangle}{\|B^\ell B^{*\ell} \check{\chi}_k\| \|B^\ell \chi_k\|}.$$

In view of (81), it is sufficient to prove that for some $c_0 > 0$, w.h.p. $\langle B^\ell \chi_k, B^{2\ell} \chi_k \rangle > c_0 \mu_k^{3\ell} n$. Note then that

$$(83) \quad \begin{aligned} & |\langle B^\ell \chi_k, B^{2\ell} \chi_k \rangle - \langle N_{k,\ell}, N_{k,2\ell} \rangle| \\ & \leq \|B^\ell \chi_k\| \|B^{2\ell} \chi_k - N_{k,2\ell}\| + \|B^\ell \chi_k - N_{k,\ell}\| \|N_{k,2\ell}\|. \end{aligned}$$

However, from (81) and Lemma 39, we have w.h.p. $\|B^\ell \chi_k\| \|B^{2\ell} \chi_k - N_{k,2\ell}\| = o(\mu_k^{3\ell} n)$. Also, from Proposition 37(i), and Lemma 39, we find w.h.p. $\|B^\ell \chi_k - N_{k,\ell}\| \|N_{k,2\ell}\| = o(\mu_k^{3\ell} n)$. So finally, w.h.p.,

$$|\langle B^\ell \chi_k, B^{2\ell} \chi_k \rangle - \langle N_{k,\ell}, N_{k,2\ell} \rangle| = o(\mu_k^{3\ell} n).$$

On the other hand, by Proposition 37(v), $\langle N_{k,\ell}, N_{k,2\ell} \rangle$ is w.h.p. larger than $c_0 \mu_k^{3\ell} n$ for some $c_0 > 0$. This completes the proof of (ii).

Proof of (iv). Let $\bar{\mu}_k = \mu_k \vee \sqrt{\alpha}$. From (1) and (81)–(82) for $k, j \in [r]$, w.h.p.,

$$|\langle \check{\varphi}_j, \check{\varphi}_k \rangle| = \frac{|\langle B^\ell \chi_j, B^\ell \chi_k \rangle|}{\|B^\ell \chi_j\| \|B^\ell \chi_k\|} \leq \frac{|\langle B^\ell \chi_j, B^\ell \chi_k \rangle|}{c_0^2 n \bar{\mu}_j^\ell \bar{\mu}_k^\ell} \leq \frac{|\langle B^\ell \chi_j, B^\ell \chi_k \rangle|}{c_0^2 n \alpha^\ell}.$$

In addition, equations (81)–(82), Proposition 37(i) and Lemma 39 entail that, w.h.p.,

$$\begin{aligned} |\langle B^\ell \chi_j, B^\ell \chi_k \rangle - \langle N_{j,\ell}, N_{k,\ell} \rangle| & \leq \|B^\ell \chi_j\| \|B^\ell \chi_k - N_{k,\ell}\| + \|B^\ell \chi_j - N_{j,\ell}\| \|N_{k,\ell}\| \\ & = O(\alpha^{5\ell/2} (\log n)^{5/2} \sqrt{n}). \end{aligned}$$

From Proposition 37(iii), we get, if $k \neq j$, that w.h.p.,

$$|\langle \check{\varphi}_j, \check{\varphi}_k \rangle| = O(\alpha^{3\ell/2} n^{-\gamma/2} (\log n)^{5/2}).$$

Proof of (v). Let $k \neq j \in [r_0]$. From (1) and (81), w.h.p.,

$$|\langle \zeta_j, \check{\varphi}_k \rangle| = \frac{|\langle B^\ell \chi_j, B^{2\ell} \chi_k \rangle|}{\|B^\ell B^{*\ell} \check{\chi}_k\| \|B^\ell \chi_k\|} \leq \frac{|\langle B^\ell \chi_j, B^{2\ell} \chi_k \rangle|}{c_0^2 n \alpha^{3\ell/2}}.$$

As in (83), we use $|\langle x, y \rangle - \langle x', y' \rangle| \leq \|x'\| \|y - y'\| + \|y\| \|x - x'\|$. We find from (81), Proposition 37(i) and Lemma 39 that, w.h.p.,

$$|\langle B^\ell \chi_j, B^{2\ell} \chi_k \rangle - \langle N_{j,\ell}, N_{k,2\ell} \rangle| = O(\alpha^{4\ell} \sqrt{n} (\log n)^{5/2}).$$

Also, from Proposition 37(iv), w.h.p. $\langle N_{j,\ell}, N_{k,2\ell} \rangle$ is $O(n^{1-\gamma/2} \alpha^{7\ell/2} (\log n)^{5/2})$. We conclude finally that

$$|\langle \zeta_j, \check{\varphi}_k \rangle| = O(\alpha^{2\ell} n^{-\gamma/2} (\log n)^{5/2}).$$

Proof of (vi). We again use the same argument. Let $k \neq j \in [r_0]$. From (81),

$$|\langle \zeta_k, \zeta_j \rangle| = \frac{|\langle B^\ell B^{*\ell} \check{\chi}_k, B^\ell B^{*\ell} \check{\chi}_j \rangle|}{\|B^\ell B^{*\ell} \check{\chi}_k\| \|B^\ell B^{*\ell} \check{\chi}_j\|} \leq \frac{|\langle B^\ell B^{*\ell} \check{\chi}_k, B^\ell B^{*\ell} \check{\chi}_k \rangle|}{c_0^2 n \alpha^{2\ell}}.$$

Recall that $|\langle x, y \rangle - \langle x', y' \rangle| \leq \|x'\| \|y - y'\| + \|y\| \|x - x'\|$. Then from (81), Proposition 38(i) and Lemma 39, we obtain w.h.p.

$$|\langle B^\ell B^{*\ell} \check{\chi}_k, B^\ell B^{*\ell} \check{\chi}_j \rangle - \langle P_{k,\ell} + S_{k,\ell}, N_{j,\ell} + S_{k,\ell} \rangle| = O((\log n)^4 \alpha^{5\ell} \sqrt{n}).$$

Finally, from Proposition 38(iii), the scalar product $\langle P_{k,\ell} + S_{k,\ell}, N_{j,\ell} + S_{k,\ell} \rangle$ is $O(n^{1-\gamma/2} \alpha^{9\ell/2} (\log n)^{5/2})$. \square

10. Norm of nonbacktracking matrices. In this section, we prove Proposition 20. The argument used for Erdős–Rényi graphs extends rather directly to the stochastic block model.

10.1. *Decomposition of B^ℓ .* In this paragraph, we essentially repeat the argument of Section 5.2. We define, for $u \neq v \in V$, the centered variable,

$$\underline{A}_{uv} = A_{uv} - W_{\sigma(u)\sigma(v)}.$$

We now re-define K as the weighted nonbacktracking matrix on the complete graph on V , for $e, f \in \vec{E}(V)$,

$$K_{ef} = \mathbb{1}(e \rightarrow f) W_{\sigma(e_1)\sigma(e_2)},$$

where $e \rightarrow f$ represents the nonbacktracking property, $e_2 = f_1$ and $e \neq f^{-1}$. We also introduce

$$K_{ef}^{(2)} = \mathbb{1}(e \xrightarrow{2} f) W_{\sigma(e_2)\sigma(f_1)},$$

where $e \xrightarrow{2} f$ means that there is a nonbacktracking path with one intermediate edge between e and f . We define $\Delta^{(\ell)}$, $B^{(\ell)}$ as in Section 5.2. $R_t^{(\ell)}$ is now defined as

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

where the set of paths $F_{t,ef}^{\ell+1}$ is still defined as in Section 5.2. We again use the decomposition

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

to obtain

$$(84) \quad \begin{aligned} B^{(\ell)} &= \Delta^{(\ell)} + \frac{1}{n} K B^{(\ell-1)} + \frac{1}{n} \sum_{t=1}^{\ell-1} \Delta^{(t-1)} K^{(2)} B^{(\ell-t-1)} \\ &\quad + \frac{1}{n} \Delta^{(\ell-1)} K - \frac{\alpha}{n} \sum_{t=0}^{\ell} R_t^{(\ell)}. \end{aligned}$$

We introduce

$$\bar{W} = \sum_k \mu_k \chi_k \check{\chi}_k^* \quad \text{and} \quad L = K^{(2)} - \bar{W}.$$

It now follows from (84) that when G is ℓ -tangle-free,

$$(85) \quad \begin{aligned} \|B^\ell x\| &\leq \|\Delta^{(\ell)}\| + \frac{1}{n} \|K B^{(\ell-1)}\| \\ &\quad + \frac{1}{n} \sum_{j=1}^r \mu_j \sum_{t=1}^{\ell-1} \|\Delta^{(t-1)} \chi_j\| \|\langle \check{\chi}_j, B^{\ell-t-1} x \rangle\| \\ &\quad + \frac{1}{n} \sum_{t=1}^{\ell-1} \|S_t^{(\ell)}\| + \|\Delta^{(\ell-1)}\| + \frac{\alpha}{n} \sum_{t=0}^{\ell} \|R_t^{(\ell)}\|, \end{aligned}$$

where we have again let $S_t^{(\ell)} := \Delta^{(t-1)} L B^{(\ell-t-1)}$ as in Section 5.2. We will now upper bound the above expression over all x such that $\langle \check{\chi}_j, B^\ell x \rangle = 0$.

10.2. Proof of Proposition 20. The proof of Proposition 20 parallels that of Proposition 11.

The main task is to adapt the arguments of Section 6 to bound the norms $\|\Delta^{(t)}\|$, $\|\Delta^{(t)} \chi_k\|$, $\|R_t^{(\ell)}\|$, $\|B^{(t)}\|$, $\|K B^{(t)}\|$, $\|S_t^{(\ell)}\|$. We only highlight the two main differences.

First, the expressions (38)–(45)–(48) now depend on the types of the vertices involved in a path. We treat, for example, the case of (38) needed for Proposition 29. We claim that if $\gamma \in W_{k,m}$ is a canonical path with e edges and v vertices,

$$(86) \quad \frac{1}{n^v} \sum_{\tau} \mathbb{E} \prod_{i=1}^{2m} \prod_{s=1}^k \underline{A}_{\tau(\gamma_{i,s-1}), \tau(\gamma_{i,s})} \leq \left(\frac{\bar{\alpha}_n}{n}\right)^{v-1} \left(\frac{a}{n}\right)^{e-v+1},$$

where the sum is over all injections $\tau : [v] \rightarrow [n]$, $a = \max_{i,j} W_{ij}$ and $\bar{\alpha}_n = \alpha + O(n^{-\gamma})$ is defined in (68). Indeed, we consider a spanning tree of $G(\gamma)$, for the $e - v + 1$ edges not present in the spanning tree, we use the bound, $\mathbb{E} \underline{A}_{uv}^p \leq W_{\sigma(u)\sigma(v)}/n \leq a/n$ for any $p \geq 1$ and $u, v \in [n]$. For the remaining $v - 1$ edges, we take a leaf, say l , of the spanning tree of $G(\gamma)$, and denote its unique neighbor by g . Then the injection $\tau : [v] \rightarrow [n]$ will give a label say $i = \sigma(\tau(g))$

to g and $j = \sigma(\tau(l))$ to l . We use the bound that for any $p \geq 1$ and $i \in [n]$, $\sum_{j=1}^r n(j)W_{ij}/n \leq \bar{\alpha}_n$. Hence, summing over all possible values of $\tau(l)$ while fixing $\tau(q)$, $q \neq l$, gives a factor of at most α_n/n in (86). We then remove l from the spanning tree and repeat this procedure $v - 1$ times; this yields (86).

With (86) in place of (38), using Lemma 17 we then bound S given by (37) as follows:

$$\begin{aligned} S &\leq \sum_{v=3}^{km+1} \sum_{e=v-1}^{km} |\mathcal{W}_{k,m}(v, e)| \left(\frac{\bar{\alpha}_n}{n}\right)^{v-1} \left(\frac{a}{n}\right)^{e-v+1} n^v \\ &\leq n \bar{\alpha}_n^{km} \sum_{v=3}^{km+1} \sum_{e=v-1}^{km} (2km)^{10m(e-v+1)+8m} a^{e-v+1} n^{v-e-1} \\ &\leq n \bar{\alpha}_n^{km} (2\ell m)^{8m} (\ell m) \sum_{s=0}^{\infty} \left(\frac{(2\ell am)^{10m}}{n}\right)^s. \end{aligned}$$

In the range of $k \leq \ell$ and m defined by (34), we have $\bar{\alpha}_n^{km} = \alpha^{km} (1 + o(1))$. We deduce that the bound (29) on $\|\Delta\|$ continues to hold for the stochastic block model. Similarly, by the same adaptation of (45)–(48), we find that bounds (31) and (32) on $\|R_t^{(\ell)}\|$ and $\|B^{(t)}\|$ continue to hold for the stochastic block model.

The second difference lies in the definition of the matrix $L = K^{(2)} - \bar{W}$. For the stochastic block model, from (9), the entry L_{ef} is zero unless $e = f$, $e \rightarrow f$, $f^{-1} \rightarrow e$ or $e \rightarrow f^{-1}$. Moreover, the nonzero entries are bounded by a . Then the argument of the proof of bound (33) carries over easily.

With bounds (29)–(31)–(32)–(33) available for the stochastic block model, the remainder of the proof of Proposition 20 repeats the argument of Section 5.4.

11. Stochastic block model: Proof of Theorem 5. The strategy of proof is based on the following lemma which asserts that the existence of a Boolean function nonconstant over the classes ensures the existence of an estimation with asymptotically positive overlap.

LEMMA 40. *Assume that $\pi(i) \equiv 1/r$ and there exists a function $F : V \rightarrow \{0, 1\}$ of the graph G such that in probability, for any $i \in [r]$,*

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

where $f : [r] \rightarrow [0, 1]$ is not a constant function [there exists (i, j) such that $f(i) \neq f(j)$]. Let (I^+, I^-) be a partition of $[r]$, such that $0 < |I^+| < r$ and

$$(87) \quad \frac{1}{|I^+|} \sum_{i \in I^+} f(i) > \frac{1}{|I^-|} \sum_{i \in I^-} f(i).$$

Then the following estimation procedure yields asymptotically positive overlap with permutation p in (15) equal to the identity: assign to each vertex v a label $\hat{\sigma}(v)$ picked uniformly at random from I^+ if $F(v) = 1$ and from I^- if $F(v) = 0$.

Observe that the existence of a nontrivial partition (I^+, I^-) satisfying (87) is implied by the assumption that f is not constant.

PROOF OF LEMMA 40. Let $j \in I^+$ and $v \in V$ such that $\sigma(v) = j$, then given the realization of the graph, the event $\hat{\sigma}(v) = \sigma(v)$ is equal to $F(v)\varepsilon_v$, where ε_v is an independent Bernoulli $\{0, 1\}$ -random variable with $\mathbb{P}(\varepsilon_v = 1) = 1/|I^+|$. From the law of large numbers we deduce that, in probability,

$$\frac{1}{n} \sum_{v=1}^n \mathbb{1}_{\sigma(v)=j} \mathbb{1}_{\hat{\sigma}(v)=\sigma(v)} \rightarrow \frac{f(j)}{r|I^+|}.$$

Summing over all $j \in I^+$, in probability,

$$\frac{1}{n} \sum_{v=1}^n \mathbb{1}_{\sigma(v) \in I^+} \mathbb{1}_{\hat{\sigma}(v)=\sigma(v)} \rightarrow \frac{f_+}{r},$$

where f_+ is the left-hand side of (87). Similarly, if f_- is the right-hand side of (87), in probability,

$$\frac{1}{n} \sum_{v=1}^n \mathbb{1}_{\sigma(v) \in I^-} \mathbb{1}_{\hat{\sigma}(v)=\sigma(v)} \rightarrow \frac{1 - f_-}{r}.$$

Finally, in probability,

$$\frac{1}{n} \sum_{v=1}^n \mathbb{1}_{\hat{\sigma}(v)=\sigma(v)} - \frac{1}{r} \rightarrow \frac{1}{r}(f_+ + 1 - f_- - 1) = \frac{1}{r}(f_+ - f_-) > 0,$$

where the strict inequality comes from (87). \square

Our aim is now to find a nonconstant function over the classes which depends on the eigenvector ξ_k . To this end, we introduce a new random variable, for $v \in V$,

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

where $P_{k,\ell}$ was defined by (75). Our first lemma is an extension of Proposition 38.

LEMMA 41. Let $\ell \sim \kappa \log_\alpha n$ with $0 < \kappa < \gamma \wedge 1/2$, $k \in [r_0]$ and $i \in [r]$. There exists a random variable $Y_{k,i}$ such that $\mathbb{E}Y_{k,i} = 0$, $\mathbb{E}|Y_{k,i}| < \infty$ and for any continuity point t of the distribution of $|Y_{k,i}|$, in L^2 ,

$$\frac{1}{n} \sum_{v=1}^n \mathbb{1}_{\{\sigma(v)=i\}} \mathbb{1}_{\{|I_{k,\ell}(v)\mu_k^{-2\ell} - \alpha\mu_k\phi_k(i)/(\mu_k^2/\alpha - 1)| \geq t\}} \rightarrow \pi(i)\mathbb{P}(|Y_{k,i}| \geq t).$$

PROOF. Let Z_t , $t \geq 0$, be the Galton–Watson branching process defined in Section 8 started from $Z_0 = \delta_t$ and t has distribution $(\pi(1), \dots, \pi(r))$. We denote by (T, o) the associated random rooted tree. Let D be the number of offspring of the root and for $1 \leq x \leq D$, let $Q_{k,\ell}(x)$ be the random variable $Q_{k,\ell}$ defined on the tree T^x where the subtree attached to x is removed and set

$$J_{k,\ell} = \sum_{x=1}^D Q_{k,\ell}(x).$$

We observe that

$$\begin{aligned} \mathbb{E}J_{k,\ell} &= \sum_{n=0}^{\infty} \frac{\alpha^n e^{-\alpha}}{n!} n \mathbb{E}[Q_{k,\ell} | D = n - 1] \\ &= \sum_{n=1}^{\infty} \frac{\alpha^n e^{-\alpha}}{(n-1)!} \mathbb{E}[Q_{k,\ell} | D = n - 1] = \alpha \mathbb{E}Q_{k,\ell}. \end{aligned}$$

Also, by Theorem 25, the variable $Q_{k,\ell} \mu_k^{-2\ell} - \mu_k \phi_k(i) / (\mu_k^2 / \alpha - 1)$ converges in L^2 to a centered variable X_k satisfying $\mathbb{E}X_k^2 \leq C$. However, the variables $Q_{k,\ell}$ and $J_{k,\ell}$ are closely related, indeed,

$$J_{k,\ell} = (D-2)L_{k,\ell}^o + (D-1) \sum_{t=1}^{\ell} \sum_{u \in Y_t^o} L_{k,\ell}^u = (D-1)Q_{k,\ell} - L_{k,\ell}^o,$$

where $L_{k,\ell}^u$ was defined above (60). Inequality (63) for $t = 0$, shows that $\mathbb{E}|L_{k,\ell}^o|^2 = O(\alpha^{2\ell})$. Hence, $L_{k,\ell}^o / \mu_k^{2\ell}$ converges in L^2 to 0. From Theorem 25, we find that $J_{k,\ell} \mu_k^{-2\ell} - \alpha \mu_k \phi_k(i) / (\mu_k^2 / \alpha - 1)$ converges weakly to a centered variable Y_k satisfying $\mathbb{E}|Y_k| \leq C$. In particular, if t is a continuity point of $|Y_k|$, $\mathbb{1}(\sigma(o) = i) \mathbb{1}(|J_{k,\ell} \mu_k^{-2\ell} - \alpha \mu_k \phi_k(i) / (\mu_k^2 / \alpha - 1)| \geq t)$ converges weakly to $\mathbb{1}(\sigma(o) = i) \mathbb{1}(|Y_k| \geq t)$. It then remains to apply Proposition 36. \square

Our second lemma checks that we may replace $P_{k,\ell}$ in the above statement by the eigenvector ξ_k' properly renormalized to be asymptotically close to (14). More precisely, we set

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

where $s = \sqrt{\alpha \rho_k'}$ and ρ_k' was defined in Proposition 38.

LEMMA 42. *Let $k \in [r_0]$, $i \in [r]$ and $Y_{k,i}$ be as in Lemma 41. For any continuity point t of the distribution of $|Y_{k,i}|$, in L^2 ,*

$$\frac{1}{n} \sum_{v=1}^n \mathbb{1}_{\{\sigma(v)=i\}} \mathbb{1}_{\{|I_k(v) - \alpha \mu_k \phi_k(i) / (\mu_k^2 / \alpha - 1)| \geq t\}} \rightarrow \pi(i) \mathbb{P}(|Y_{k,i}| \geq t).$$

PROOF. From Proposition 38(i), we find that, in probability,

$$\frac{1}{n} \sum_{e \in \bar{E}} \frac{P_{k,\ell}^2}{\mu_k^{4\ell}} \rightarrow s^2.$$

We set $\tilde{\xi}_k(e) = P_{k,\ell}/(\mu_k^{2\ell} s \sqrt{n})$. From Lemma 39, we have, w.h.p.,

$$\|\tilde{\xi}_k - \zeta_k\| = O(\alpha^\ell \mu_k^{-2\ell}) = o(1).$$

Also, from Theorem 4, w.h.p.,

$$\|\zeta_k - \xi'_k\| = o(1).$$

Hence, from the triangle inequality, w.h.p.,

$$\|\tilde{\xi}_k - \xi'_k\| = o(1).$$

We deduce from the Cauchy–Schwarz inequality that, w.h.p.,

$$\frac{1}{n} \sum_{v=1}^n \left| I_k(v) - \frac{I_{k,\ell}(v)}{\mu_k^{2\ell}} \right| \leq \frac{s\sqrt{n}}{n} \sum_{e \in \bar{E}} |\xi'_k(e) - \tilde{\xi}_k(e)| \leq \frac{s\sqrt{n|\bar{E}|}}{n} \|\tilde{\xi}_k - \xi'_k\| = o(1).$$

Since t is a continuity point of $|Y_{k,i}|$, it is then a routine to deduce Lemma 42 from Lemma 41. \square

All ingredients are now gathered to prove Theorem 5. We fix $k \in [r_0]$ as in Theorem 5 and let ξ'_k be as above. We set

$$J^+ := \{i \in [r] : \phi_k(i) > 0\} \quad \text{and} \quad J^- = [r] \setminus J^+.$$

From Lemma 42, there exist random variables X_j , $j \in [r]$ on \mathbb{R} such that $\mathbb{E}X_j = \alpha \mu_k \phi_k(j)/(\mu_k^2/\alpha - 1)$ and the following holds for all $j \in [r]$. With I_k as above, for all $t \in \mathbb{R}$ that is a continuity point of the distribution of X_j , the following convergence in probability holds:

$$(88) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{v=1}^n \mathbb{1}_{\sigma(v)=j} \mathbb{1}_{I_k(v) > t} = \pi(j) \mathbb{P}(X_j > t).$$

Write, for $\varepsilon = \pm$, $\pi_\varepsilon = \sum_{j \in J^\varepsilon} \pi(j)$, $g_\varepsilon = \sum_{j \in J^\varepsilon} \pi(j) \phi_k(j)$. Note that $g_+ > 0$ by definition of J^+ . Also by the orthogonality relation (8) between $\phi_1 = \mathbb{1}$ and ϕ_k we obtain $g_+ + g_- = 0$, so that $g_- < 0$. For $\varepsilon = \pm$, we shall denote by X_ε the random variable obtained as a mixture of the X_j for $j \in J^\varepsilon$, with weights $\pi(j)/\pi_\varepsilon$. Note that X_ε has mean $g_\varepsilon/\pi_\varepsilon$.

We now establish the existence of $t_0 \in \mathbb{R}$ that is a continuity point of the distribution of both X_+ and X_- , and such that

$$(89) \quad \mathbb{P}(X_+ > t_0) > \mathbb{P}(X_- > t_0).$$

To this end, since $\mathbb{E}X_+ > 0$, we write

$$\int_0^{+\infty} \mathbb{P}(X_+ > t) dt > \int_0^{+\infty} \mathbb{P}(-X_+ > t) dt = \int_0^{+\infty} \mathbb{P}(-X_+ \geq t) dt.$$

The same argument yields

$$\int_0^{+\infty} \mathbb{P}(X_- > t) dt < \int_0^{+\infty} \mathbb{P}(-X_- \geq t) dt.$$

Combined, these two inequalities imply

$$\int_0^{+\infty} \{[\mathbb{P}(X_+ > t) - \mathbb{P}(X_- > t)] + [\mathbb{P}(X_+ > -t) - \mathbb{P}(X_- > -t)]\} dt > 0.$$

Thus, there is a subset of \mathbb{R}_+ of positive Lebesgue measure on which either $\mathbb{P}(X_+ > t) > \mathbb{P}(X_- > t)$ or $\mathbb{P}(X_+ > -t) > \mathbb{P}(X_- > -t)$. This implies the existence of a continuity point $t_0 \in \mathbb{R}$ of both X_+ , X_- , $-X_+$ and $-X_-$ such that (89) holds.

We may now come back to the eigenvector ξ_k in Theorem 5. We set $\tau = st_0$ in Theorem 5. For some unknown sign $\omega \in \{-1, 1\}$, we have

$$(90) \quad \xi_k = \omega \xi'_k.$$

Case 1: The sign can be estimated. We first assume that ω is known and $\xi'_k = \xi_k$. We consider the function

$$F(v) = \mathbb{1}_{\{\sum_{e:e_2=v} \xi'_k(e) > \tau/\sqrt{n}\}} = \mathbb{1}_{\{I_k(v) > t_0\}}.$$

From (89), (87) is satisfied with $I^\pm = J^\pm$ and we can apply Lemma 40 to obtain an asymptotically positive overlap. We note that the sign ω is easy to estimate consistently if the random variable X which is the mixture of the X_j with weights $\pi(j) = 1/r$ is not symmetric. Indeed, in this case, for some bounded continuous function f ,

$$\mathbb{E}f(X) = \sum_{j=1}^r \pi(j) \mathbb{E}f(X_j) \neq \sum_{j=1}^r \pi(j) \mathbb{E}f(-X_j) = \mathbb{E}f(-X).$$

Then, from (88), given ω , in probability,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{v=1}^n f(\omega I_k(v)) = \mathbb{E}f(\omega X)$$

takes a different value for $\omega = 1$ and $\omega = -1$.

Case 2: Fully symmetric case. Another simple case is if X defined above is symmetric and $|J^+| = |J^-|$. If this occurs, X_+ and $-X_-$ have the same distribution. We consider the function $F(v) = \mathbb{1}_{\{\sum_{e:e_2=v} \xi_k(e) > \tau/\sqrt{n}\}} = \mathbb{1}_{\{\omega I_k(v) > t_0\}}$ and the estimation where a vertex such that $F(v) = 1$ receives a uniform label in J^+ ,

and otherwise a label uniform in J^- . By Lemma 40 applied to F , if $\omega = 1$, we obtain an positive overlap with $I^\pm = J^\pm$ and the permutation p in (15) equal to the identity. If $\omega = -1$, we obtain an positive overlap with $I^\pm = J^\mp$ and any permutation p in (15) such that $p(J^\pm) = J^\mp$.

Case 3: General case. In the general case, we may use the same idea: apply Lemma 40 for a partition (I^+, I^-) which may depend on ω but such that the cardinal of I^\pm does not. First, from (88), the function $f_1(j) = \mathbb{P}(X_j > t_0)$ is not constant on $[r]$ and there exists j_1 such that

$$f_1(j_1) > \frac{1}{r-1} \sum_{j \neq i_1} f_1(j).$$

We distinguish two subcases. The first case is when the function $f_{-1}(j) = \mathbb{P}(-X_j > t_0)$ is also nonconstant. Then there exists j_{-1} such that

$$f_{-1}(j_{-1}) > \frac{1}{r-1} \sum_{j \neq j_{-1}} f_{-1}(j).$$

We consider the function $F(v) = \mathbb{1}(\sum_{e:e_2=v} \xi_k(e) > \tau/\sqrt{n}) = \mathbb{1}(\omega I_k(v) > t_0)$ and the estimation where $\hat{\sigma}(v) = 1$ if $F(v) = 1$ and $\hat{\sigma}(v)$ uniform on $\{2, \dots, r\}$ otherwise. We apply Lemma 40 to F and the partition $I^+ = \{j_\omega\}$, $I^- = [r] \setminus \{j_\omega\}$. We obtain an asymptotically positive overlap for any permutation p in (15) such that $p(j_\omega) = 1$.

In the other case, f_{-1} is constant and equal to say a . We introduce extra random independent variables $\omega'(v) \in \{-1, 1\}$ i.i.d. such that $\mathbb{P}(\omega'(v) = 1) = \mathbb{P}(\omega'(v) = -1) = 1/2$. We consider the function $F(v) = \mathbb{1}(\omega'(v) \sum_{e:e_2=v} \xi_k(e) > \tau/\sqrt{n})$. Then, by (88), in probability,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{v=1}^n \mathbb{1}_{\sigma(v)=j} F(v) = \frac{\pi(j)}{2} (\mathbb{P}(X_j > t_0) + a).$$

Hence, it follows from (89) that (87) is satisfied with $I^\pm = J^\pm$. We can then apply Lemma 40 to obtain an asymptotically positive overlap.

This completes the proof of Theorem 5.

Acknowledgment. The authors thank Lennart Gulikers for pointing out a mistake in a preliminary version of this paper.

REFERENCES

- [1] ANGEL, O., FRIEDMAN, J. and HOORY, S. (2015). The non-backtracking spectrum of the universal cover of a graph. *Trans. Amer. Math. Soc.* **367** 4287–4318. [MR3324928](#)
- [2] BARBOUR, A. D. and CHEN, L. H. Y., eds. (2005). *An Introduction to Stein's Method. Lecture Notes Series. Institute for Mathematical Sciences. National Univ. Singapore* **4**. Singapore Univ. Press, Singapore.

- [3] BENJAMINI, I. and SCHRAMM, O. (2001). Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6** no. 23, 13 pp. (electronic). [MR1873300](#)
- [4] BHATIA, R. (1997). *Matrix Analysis. Graduate Texts in Mathematics* **169**. Springer, New York. [MR1477662](#)
- [5] BOLLOBÁS, B., JANSON, S. and RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31** 3–122. [MR2337396](#)
- [6] BORDENAVE, C. (2015). A new proof of Friedman’s second eigenvalue theorem and its extension to random lifts. Preprint. Available at [arXiv:1502.04482](#).
- [7] CHUNG, F. R. K. (1989). Diameters and eigenvalues. *J. Amer. Math. Soc.* **2** 187–196. [MR0965008](#)
- [8] DECELLE, A., KRZAKALA, F., MOORE, C. and ZDEBOROVÁ, L. (2011). Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Phys. Rev. E* (3) **84** 066106.
- [9] FRIEDMAN, J. (2008). A proof of Alon’s second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.* **195** viii+100. [MR2437174](#)
- [10] FRIEDMAN, J. and KOHLER, D.-E. (2014). The relativized second eigenvalue conjecture of Alon. Preprint. Available at [arXiv:1403.3462](#).
- [11] FÜREDI, Z. and KOMLÓS, J. (1981). The eigenvalues of random symmetric matrices. *Combinatorica* **1** 233–241. [MR0637828](#)
- [12] HASHIMOTO, K. (1989). Zeta functions of finite graphs and representations of p -adic groups. In *Automorphic Forms and Geometry of Arithmetic Varieties. Adv. Stud. Pure Math.* **15** 211–280. Academic Press, Boston, MA. [MR1040609](#)
- [13] HOLLAND, P. W., LASKEY, K. B. and LEINHARDT, S. (1983). Stochastic blockmodels: First steps. *Soc. Netw.* **5** 109–137. [MR0718088](#)
- [14] HOORY, S., LINIAL, N. and WIGDERSON, A. (2006). Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)* **43** 439–561 (electronic). [MR2247919](#)
- [15] HORN, R. A. and JOHNSON, C. R. (1991). *Topics in Matrix Analysis*. Cambridge Univ. Press, Cambridge. [MR1091716](#)
- [16] HORTON, M. D., STARK, H. M. and TERRAS, A. A. (2006). What are zeta functions of graphs and what are they good for? In *Quantum Graphs and Their Applications. Contemp. Math.* **415** 173–189. Amer. Math. Soc., Providence, RI. [MR2277616](#)
- [17] KESTEN, H. and STIGUM, B. P. (1966). Additional limit theorems for indecomposable multi-dimensional Galton–Watson processes. *Ann. Math. Stat.* **37** 1463–1481. [MR0200979](#)
- [18] KESTEN, H. and STIGUM, B. P. (1966). A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Stat.* **37** 1211–1223. [MR0198552](#)
- [19] KOTANI, M. and SUNADA, T. (2000). Zeta functions of finite graphs. *J. Math. Sci. Univ. Tokyo* **7** 7–25. [MR1749978](#)
- [20] KRZAKALA, F., MOORE, C., MOSSEL, E., NEEMAN, J., SLY, A., ZDEBOROVÁ, L. and ZHANG, P. (2013). Spectral redemption in clustering sparse networks. *Proc. Natl. Acad. Sci. USA* **110** 20935–20940. [MR3174850](#)
- [21] LUBOTZKY, A. (1995). Cayley graphs: Eigenvalues, expanders and random walks. In *Surveys in Combinatorics, 1995 (Stirling). London Mathematical Society Lecture Note Series* **218** 155–189. Cambridge Univ. Press, Cambridge. [MR1358635](#)
- [22] LUBOTZKY, A., PHILLIPS, R. and SARNAK, P. (1988). Ramanujan graphs. *Combinatorica* **8** 261–277. [MR0963118](#)
- [23] MASSOULIÉ, L. (2014). Community detection thresholds and the weak Ramanujan property. In *STOC’14—Proceedings of the 2014 ACM Symposium on Theory of Computing* 694–703. ACM, New York. [MR3238997](#)
- [24] MOSSEL, E., NEEMAN, J. and SLY, A. (2013). A proof of the block model threshold conjecture. Preprint. Available at [arXiv:1311.4115v2](#).

- [25] MOSSEL, E., NEEMAN, J. and SLY, A. (2015). Reconstruction and estimation in the planted partition model. *Probab. Theory Related Fields* **162** 431–461. [MR3383334](#)
- [26] MURTY, M. R. (2003). Ramanujan graphs. *J. Ramanujan Math. Soc.* **18** 33–52. [MR1966527](#)
- [27] NILLI, A. (1991). On the second eigenvalue of a graph. *Discrete Math.* **91** 207–210. [MR1124768](#)
- [28] SHI, X. and WEI, Y. (2012). A sharp version of Bauer–Fike’s theorem. *J. Comput. Appl. Math.* **236** 3218–3227. [MR2912686](#)
- [29] SUNADA, T. (1988). Fundamental groups and Laplacians. In *Geometry and Analysis on Manifolds (Katata/Kyoto, 1987)*. *Lecture Notes in Math.* **1339** 248–277. Springer, Berlin. [MR0961485](#)
- [30] TERRAS, A. (2011). *Zeta Functions of Graphs. A Stroll Through the Garden*. *Cambridge Studies in Advanced Mathematics* **128**. Cambridge Univ. Press, Cambridge. [MR2768284](#)

C. BORDENAVE
INSTITUT DE MATHÉMATIQUES DE TOULOUSE
CNRS & UNIVERSITÉ TOULOUSE III
118 ROUTE DE NARBONNE
31062 TOULOUSE
FRANCE
E-MAIL: charles.bordenave@math.univ-toulouse.fr

M. LELARGE
INRIA & ENS
2 RUE SIMONE IFF
CS 42112
75589 PARIS CEDEX 12
FRANCE
E-MAIL: marc.lelarge@ens.fr

L. MASSOULIÉ
MSR-INRIA JOINT CENTRE
1 RUE HONORÉ D’ESTIENNE D’ORVES
CAMPUS DE L’ÉCOLE POLYTECHNIQUE
91128 PALAISEAU
FRANCE
E-MAIL: laurent.massoulie@inria.fr