Concentration of the spectral norm of Erdős-Rényi random graphs

Gábor Lugosi and Shahar Mendelson and Nikita Zhivotovskiy

We present results on the concentration properties of the spectral norm \( \|A_p\| \) of the adjacency matrix \( A_p \) of an Erdős-Rényi random graph \( G(n, p) \). First we consider the Erdős-Rényi random graph process and prove that \( \|A_p\| \) is uniformly concentrated over the range \( p \in [C \log n / n, 1] \). The analysis is based on delocalization arguments, uniform laws of large numbers, together with the entropy method to prove concentration inequalities. As an application of our techniques we prove sharp sub-Gaussian moment inequalities for \( \|A_p\| \) for all \( p \in [c \log^3 n / n, 1] \) that improve the general bounds of Alon, Krivelevich, and Vu [1] and some of the more recent results of Erdős et al. [11]. Both results are consistent with the asymptotic result of Füredi and Komlós [12] that holds for fixed \( p \) as \( n \to \infty \).

1. Introduction

An Erdős-Rényi random graph \( G(n, p) \), named after the authors of the pioneering work [10], is a graph defined on the vertex set \( [n] = \{1, \ldots, n\} \) in which any two vertices \( i, j \in [n], i \neq j \), are connected by an edge independently, with probability \( p \). Such a random graph is represented by its adjacency matrix \( A_p \). \( A_p \) is a symmetric matrix whose entries are

\[
A_{i,j}^{(p)} = \begin{cases} 
0 & \text{if } i = j \\
\mathbb{I}_{U_{i,j} < p} & \text{if } 1 \leq i < j \leq n \\
\mathbb{I}_{U_{j,i} < p} & \text{if } 1 \leq j < i \leq n 
\end{cases}
\]  

(1.1)

where \( (U_{i,j})_{1 \leq i < j \leq n} \) are independent random variables, uniformly distributed on \([0, 1]\) and \( \mathbb{I} \) stands for the indicator function. We call the family of random matrices \( (A_p)_{p \in [0,1]} \)

---

1 Gábor Lugosi was supported by the Spanish Ministry of Economy and Competitiveness, Grant MTM2015-67304-P and FEDER, EU, by “High-dimensional problems in structured probabilistic models - Ayudas Fundación BBVA a Equipos de Investigación Científica 2017” and by “Google Focused Award Algorithms and Learning for AI”. Shahar Mendelson was supported in part by the Israel Science Foundation. Nikita Zhivotovskiy was supported by RSF grant No. 18-11-00132.

2 Department of Economics and Business, Pompeu Fabra University, Barcelona, Spain, gabor.lugosi@upf.edu

3 ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain

4 Barcelona Graduate School of Economics

5 LPSM, Sorbonne University, and Mathematical Sciences Institute, The Australian National University, shahar.mendelson@upmc.fr

6 This work was prepared while Nikita Zhivotovskiy was a postdoctoral fellow at the department of Mathematics, Technion I.I.T. and researcher at National University Higher School of Economics. Now at Google Research, Brain Team, nikita.zhivotovskiy@phystech.edu
the Erdős-Rényi random graph process.

Spectral properties of adjacency matrices of random graphs have received considerable attention, see Füredi and Komlós [12], Krivelevich and Sudakov [14], Vu [20], Erdős, Knowles, Yau, and Yin [11], Benaych-Georges, Bordenave, and Knowles [4, 5], Jung and Lee [13], Tran, Vu, and Wang [18], among many other papers.

In this paper we are primarily concerned with concentration properties of the spectral norm \( \|A_p\| \) of the adjacency matrix. It follows from a general concentration inequality of Alon, Krivelevich, and Vu [1] for the largest eigenvalue of symmetric random matrices with bounded independent entries that for all \( n \geq 1 \), \( p \in [0,1] \), and \( t > 0 \),

\[
P \{ \|A_p\| - E\|A_p\| > t \} \leq 4e^{-t^2/32},
\]

(1.2)

In particular, \( \text{Var}(\|A_p\|) \leq C \) for a universal constant \( C \). (One may take \( C = 16 \), see [8, Example 3.14].) Krivelevich and Sudakov [14] who studied the asymptotic value of \( E\|A_p\| \) raised the question whether it is possible to improve (1.2). As an application of our techniques we settle this question for non-sparse graphs. Moreover, we strengthen (1.2) in two different ways.

Our main result concerns the uniform concentration of the spectral norm. In particular, first we prove that there exists a universal constant \( C \) such that

\[
E \sup_{p \geq C \log n/n} \|A_p\| - E\|A_p\| \leq C
\]

(see Theorem 1 below). Informally, this result means that as we add new edges in the Erdős-Rényi graph process, the value \( \|A_p\| - E\|A_p\| \) is never greater (up to an absolute constant factor) than the same value calculated for just one concrete random graph \( G(n, \frac{1}{2}) \). The proof of this result is based on an extension of the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality (we refer to [16] for the state-of-the-art form) for particular functions of independent random variables. For the entire range \( p \in [0,1] \), we are able to prove a simple but slightly weaker inequality

\[
E \sup_{p \in [0,1]} \|A_p\| - E\|A_p\| \leq C \sqrt{\log \log n}
\]

for a constant \( C \) (Proposition 1). We also prove the tail bound of the form

\[
P \left\{ \sup_{p \geq C \log n/n} \|A_p\| - E\|A_p\| > t \right\} \leq e^{-t^2/C},
\]

(1.3)

which is a uniform version of the sub-Gaussian inequality (1.2) and has the same form up to absolute constant factors. We leave open the question whether the restriction to the range \( p \in \left[ C \log \frac{n}{n}, 1 \right] \) is necessary for uniform concentration. However, we also discuss very sparse regimes (i.e., when \( p \ll \frac{1}{n} \)).

Note that it follows from the Perron-Frobenius theorem that the spectral norm of \( A_p \) equals the largest eigenvalue of \( A_p \), that is, \( \|A_p\| = \lambda_p \). We use both interchangeably throughout the paper, depending on the particular interpretation that is convenient.
Our proofs hinge crucially on the so-called delocalization property of the eigenvector corresponding to the largest eigenvalue (see Erdős, Knowles, Yau, and Yin [11], Mitra [17]), that is, the fact that the normalized eigenvector corresponding to the largest eigenvalue is close, in a certain sense, to the vector \( (1/\sqrt{n}, \ldots, 1/\sqrt{n}) \). We provide delocalization bounds for the top eigenvector of \( A_p \) tailored to our needs (Lemma 3) and a uniform delocalization inequality (Lemma 4). An important fact is that some known delocalization bounds hold with probability \( 1 - \frac{C}{n^\beta} \) (as in [17]) or with quasi-polynomial probability \( 1 - C \exp(-c(\log n)^\beta) \) (see e.g. [18] or Theorem 2.6 in [11]), where any choice of the parameter \( \beta \) greater than zero is responsible for extra logarithmic factors, making these results not applicable in our case. So, to obtain tight concentration results we prove delocalization bounds which hold with the exponential probability of the form \( 1 - C \exp(-cnp) \) (up to logarithmic factors), which is significantly better in the regime when \( p \gg \log n/n \).

As an application of our techniques, we prove sub-Gaussian inequalities for moments of \( \|A_p\| \) of higher order (up to order approximately \( np \)). The precise statement is given in Theorem 2 in Section 2.2 below. In particular, we show that, for small values of \( p \), \( \|A_p\| \) is significantly more concentrated than what the general bound (1.2) suggests. This technique implies, in particular, that there exists a universal constant \( C \) such that

\[
\text{Var}(\|A_p\|) \leq Cp
\]

for all \( n \) and \( p \geq C \log^3 n/n \). The rest of the paper is organized as follows. In Section 2 we formalize and discuss the results of the paper. The proofs are presented in Section 3.

2. Results

2.1. Uniform concentration for the Erdős-Rényi random graph process

Next, we state our inequalities for the uniform concentration of the spectral norm \( \|A_p\| \)—or, equivalently, for the largest eigenvalue \( \lambda_p \) of the adjacency matrix \( A_p \) defined by (1.1). Our first result shows that

**Theorem 1.** There exists a constant \( C \) such that, for all \( n \),

\[
\mathbb{E} \sup_{p \in \left[ \frac{64 \log n}{n}, 1 \right]} |\lambda_p - \mathbb{E}\lambda_p| \leq C.
\]

Moreover, for all \( t \geq 2C \),

\[
\mathbb{P} \left\{ \sup_{p \in \left[ \frac{64 \log n}{n}, 1 \right]} |\lambda_p - \mathbb{E}\lambda_p| \geq t \right\} \leq \exp(-t^2/128).
\]
For the numerical constant, our proof provides the (surely suboptimal) value \( C = 10^9 \). Our proof is based on the fact that the normalized eigenvector corresponding to the largest eigenvalue of \( A_p \) stays close to the vector \((1/\sqrt{n}, \ldots, 1/\sqrt{n})\). In Lemma 4 we prove an \( \ell_2 \) bound that holds uniformly over intervals of the form \([q, 2q]\) when \( q \in [4 \log n/n, 1/2] \). It is because of the restriction of the range of \( q \) in the uniform delocalization lemma that we need to impose \( p \geq 64 \log n/n \) in Theorem 1. We do not know whether this uniform concentration bound holds over the entire interval \( p \in [0, 1] \). However, we are able to prove the following, only slightly weaker bound.

**Proposition 1.** There exists a constant \( C' \) such that, for all \( n \),

\[
\mathbb{E} \sup_{p \in [0,1]} |\lambda_p - \mathbb{E}\lambda_p| \leq C' \sqrt{\log \log n}.
\]

The proof of Proposition 1 uses direct approximation arguments to handle the interval \( p \in [0, 64 \log n/n] \). In particular, we show that

\[
\mathbb{E} \sup_{p \in [0, 64 \log n/n]} |\lambda_p - \mathbb{E}\lambda_p| \leq 5\sqrt{16 + 2 \log \log n},
\]

which, combined with Theorem 1 implies Proposition 1. As a second extension we consider the sparse regime when \( p \ll \frac{1}{n} \).

**Proposition 2.** Fix \( k \in \mathbb{N}, k \geq 2 \). There is a constant \( C_k \) (its value may be extracted from the proof), which depends only on \( k \) such that

\[
\mathbb{E} \sup_{p \in [0, n^{-k/(k-1)}]} |\lambda_p - \mathbb{E}\lambda_p| \leq C_k.
\]

**Remark.** A simple inspection of the proof of the concentration result of Theorem 1 shows that a tail inequality similar to the second inequality of Theorem 1 holds also for the range \( p \in [0, n^{-k/(k-1)}] \). In this case the constant factors may depend on the choice of \( k \). Similarly, Proposition 1 may be used to prove that the concentration result of Theorem 1 holds for all \( p \in [0, 1] \), but only in the regime \( t \geq 2C' \sqrt{\log \log n} \).

### 2.2. Moment inequalities for the spectral norm

As an application of our techniques we show that typical deviations of \( \|A_p\| \) from its expected value are of the order of \( \sqrt{p} \). This is in accordance with the asymptotic normality theorem of Füredi and Komlós [12]. However, while the result of [12] holds for fixed \( p \) as \( n \to \infty \), the theorem below is non-asymptotic. In particular, it holds for \( p = o(1) \) as long as \( np \) is at least of the order of \( \log^4(n) \). Note that the general non-asymptotic concentration inequality of [1] only implies that typical deviations are \( O(1) \) and the question of possible improvements was raised in [14].
Theorem 2. There exist constants $c, C$ such that for every

$$k \in \left(2, \frac{c \left(\log(np)\right)^2 p(n-1) - \log(24(n-1))}{\log\left(\frac{1}{p}\right) + \log(11^5/4)}\right),$$

it holds

$$\left[\mathbb{E}(\|\mathbf{A}_p\| - \mathbb{E}\|\mathbf{A}_p\|)^k\right]^{1/k} \leq (Ckp)^{1/2}$$

and

$$\left[\mathbb{E}(\|\mathbf{A}_p\| - \mathbb{E}\|\mathbf{A}_p\|)^{-k}\right]^{1/k} \leq (C'kp)^{1/2}.$$

In particular, for some absolute constant $\kappa > 0$ it holds for all $n$ and $p \geq \kappa \log^3(n)/n$,

$$\text{Var}(\|\mathbf{A}_p\|) \leq Cp.$$
random variable that is a function of the independent random variables $X_1, \ldots, X_n$, then for all $k \geq 2$,
\[
\left( \mathbb{E} (Z - \mathbb{E}Z)^k \right)^{1/k} \leq \sqrt{3k} \left( \mathbb{E} \left[ V_+^{k/2} \right] \right)^{1/k},
\]
and
\[
\left( \mathbb{E} (Z - \mathbb{E}Z)^k \right)^{1/k} \leq 4.16k \left( \mathbb{E} \left[ V_+^{k/2} \right] \right)^{1/k} \vee \sqrt{k} \left( \mathbb{E} \left[ M^k \right] \right)^{1/k},
\]
where $\vee$ denotes the maximum and the random variable $V_+$ is defined as
\[
V_+ = \mathbb{E}' \sum_{i=1}^{n} (Z - Z'_i)^2_+.
\]
Here $Z'_i = f(X_1, \ldots, X_{i-1}, X'_i, X_{i+1}, \ldots, X_n)$ with $X'_1, \ldots, X'_n$ being independent copies of $X_1, \ldots, X_n$ and $\mathbb{E}'$ denotes expectation with respect to $X'_1, \ldots, X'_n$. Moreover,
\[
M = \max_i (Z - Z'_i)_+.
\]
Recall also that, by the Efron-Stein inequality (e.g., [8, Theorem 3.1])
\[
\text{Var}(Z) \leq \mathbb{E}V_+.
\]

The proof of Theorem 2 is based on (2.2), applied for the random variable $Z = \|A_p\|$. In order to bound moments of the random variable $V_+$, we use the fact that the eigenvector of $A_p$ corresponding to the largest eigenvalue is near the vector $(1/\sqrt{n}, \ldots, 1/\sqrt{n})$. An elegant way of proving such results appears in Mitra [17]. We follow Mitra’s approach though we need to modify his arguments in order to achieve stronger probabilistic guarantees for weak $\ell_\infty$ delocalization bounds. In Lemma 3 we provide the bound we need for the proof of Theorem 2.

3. Proofs

3.1. Proof of Theorem 1

We begin by noting that, if $p \leq q$, then $A_q$ is element-wise greater than or equal to $A_p$ and therefore $\|A_p\| \leq \|A_q\|$ whenever $p \leq q$. (see Corollary 1.5 in [6]).

We start with a lemma for the expected spectral norm for a sparse Erdős-Rényi graph. Since the largest eigenvalue of the adjacency matrix is always bounded by the maximum degree of the graph, $\mathbb{E}\|A_\frac{k}{2}\|$ is at most of the order $\log n$. The next lemma improves this naive bound to $O(\sqrt{\log n})$. With more work, it is possible to improve the rate to $\sqrt{\frac{\log n}{\log \log n}}$ (see the asymptotic result in [14]). However, this slightly weaker version is sufficient for our purposes.
Lemma 1. For all $n \geq 3$, $$\mathbb{E}\|A_{\frac{1}{n}}\| \leq 173\sqrt{\log n}.$$ 

Proof. First write 

$$\mathbb{E}\|A_{\frac{1}{n}}\| \leq \mathbb{E}\|A_{\frac{1}{n}} - \mathbb{E}A_{\frac{1}{n}}\| + \mathbb{E}\|\mathbb{E}A_{\frac{1}{n}} - \mathbb{E}A_{\frac{1}{n}}\| \leq \mathbb{E}\|A_{\frac{1}{n}} - \mathbb{E}A_{\frac{1}{n}}\| + 1.$$ 

Denote $B = A_{\frac{1}{n}} - \mathbb{E}A_{\frac{1}{n}}$ and let $B'$ be an independent copy of $B$. Denoting by $\mathbb{E}'$ the expectation operator with respect to $B'$, note that $\mathbb{E}'B' = 0$ and therefore, by Jensen’s inequality, 

$$\mathbb{E}\|B\| = \mathbb{E}\|B - \mathbb{E}'B'\| \leq \mathbb{E}\|B - B'\|.$$ 

The matrix $B - B'$ is zero mean, its non-diagonal entries have a symmetric distribution with variance $(2/n)(1 - 1/n)$ and all entries have absolute value bounded by 2. Now, applying Corollary 3.6 of Bandeira and van Handel [3] with $\alpha = 3$, 

$$\mathbb{E}\|B - B'\| \leq e^{\frac{3}{2}}(2\sqrt{2} + 84\sqrt{\log n}) \leq 6 + 166\sqrt{\log n}.$$ 

Thus, 

$$\mathbb{E}\|A_{\frac{1}{n}}\| \leq 7 + 166\sqrt{\log n} \leq 173\sqrt{\log n}.$$ 

The next lemma and the uniform delocalization inequality of Lemma 4 (presented in Section 3.3) are the crucial building blocks of the proof of Theorem 1.

Lemma 2. For all $n$ and $q \in [\log n/n, \frac{1}{2}]$, 

$$\mathbb{P}\left\{ \sup_{p \in [q, 2q]} \|A_p - \mathbb{E}A_p\| > 420\sqrt{nq} \right\} \leq e^{-nq/64}.$$ 

Proof. Using the version [8, Example 3.14] of the concentration bound of Alon, Krivelevich and Vu [1], we have for each fixed $p$ and for all $t > 0$, 

$$\mathbb{P}\left\{ \|A_p - \mathbb{E}A_p\| - \mathbb{E}\|A_p - \mathbb{E}A_p\| > t \right\} \leq e^{-t^2/32}.$$ 

On the other hand, using the same symmetrization trick as in Lemma 1, Corollary 3.6 of Bandeira, van Handel [3] implies that for any $p \geq \log n/n$, 

$$\mathbb{E}\|A_p - \mathbb{E}A_p\| \leq e^{\frac{3}{2}}(2\sqrt{2np} + 84\sqrt{\log n}) \leq 170\sqrt{np}.$$ 

These two results imply 

$$\mathbb{P}\left\{ \|A_p - \mathbb{E}A_p\| > 172\sqrt{np} \right\} \leq e^{-np/8}.$$
Let now \( q \geq \log n/n \) and for \( i = 0, 1, \ldots, \lceil nq \rceil \), define \( p_i = q + i/n \). Then using the triangle inequality, combined with \( A_p - A_{p_i} \stackrel{d}{=} A_{p-p_i} \) for \( p > p_i \) and the monotonicity of the operator norm of the matrix with non-negative entries with respect to each entry,

\[
\sup_{p \in [p_i, p_{i+1}]} (\|A_p - \mathbb{E}A_p\| - \|A_{p_i} - \mathbb{E}A_{p_i}\|) \\
\leq \sup_{p \in [p_i, p_{i+1}]} (\|A_p - A_{p_i}\| + \||\mathbb{E}A_p - \mathbb{E}A_{p_i}||) \\
= \sup_{p \in [p_i, p_{i+1}]} (\|A_p - A_{p_i}\| + \||\mathbb{E}A_{p_i}||) \\
= \|A_{p_{i+1}} - A_{p_i}\| + \|\mathbb{E}A_{1/n}\| \\
\leq \|A_{p_{i+1}} - A_{p_i}\| + 1 \\
\leq \|\mathbb{E}A_{1/n}\| + (\|A_{p_{i+1}} - A_{p_i}\| - \|A_{p_{i+1}} - A_{p_i}\|) + 1 \\
\leq 1 + 173\sqrt{\log n + \sqrt{nq}} \\
\leq 176\sqrt{nq},
\]

with probability at least \( 1 - e^{-nq/32} \), where we used Lemma 1 and (1.2). Thus, by the union bound, with probability at least \( 1 - n^{-nq/32} - n^{-nq/8} \geq 1 - e^{-nq/64} \),

\[
\sup_{p \in [q, 2q]} \|A_p - \mathbb{E}A_p\| \leq \max_{i \in \{0, \ldots, \lceil nq \rceil\}} \|A_{p_i} - \mathbb{E}A_{p_i}\| + 176\sqrt{nq} \\
\leq 172\sqrt{2nq} + 176\sqrt{nq} \\
\leq 420\sqrt{nq},
\]

as desired. \( \square \)

**Proof of Theorem 1.** Denote by \( \bar{1} \in \mathbb{R}^n \) the vector whose components are all equal to 1. Let \( B_2^q = \{ x \in \mathbb{R}^n : \|x\|_2 \leq 1 \} \) be the unit Euclidean ball. Let \( v_p \) denote the first unit eigenvector of \( A_p \). Define the event \( E_1 \) that \( v_p \in \frac{1}{\sqrt{nq}} + \frac{2800}{\sqrt{nq}} B_2^q \) for all \( p \in [64 \log n/n, 1] \).

By Lemma 4 (see Section 3.3 below) the vector \( v_p \) can always be chosen in a way such that for \( n \geq 7 \),

\[
\mathbb{P}(E_1) \geq 1 - 4\sum_{j=0}^{\infty} \exp(-2^j \log n) \geq 1 - 4\sum_{j=0}^{\infty} \left( \frac{1}{n} \right)^{2^j} \geq 1 - \frac{4}{n} \sum_{j=0}^{\infty} \left( \frac{1}{7} \right)^j = 1 - \frac{32}{7n} .
\]

Now define the event \( E_2 \) that for all \( p \in \left[ \frac{64 \log n}{n}, 1 \right] \), \( \|A_p - \mathbb{E}A_p\| \leq 420\sqrt{2nq} \). Similarly to the calculation above, by Lemma 2, \( \mathbb{P}(E_2) \geq 1 - \frac{32}{7n} \).

Denoting by \( S_n^{-1} = \{ x \in \mathbb{R}^n : \|x\|_2 = 1 \} \) the Euclidean unit sphere in \( \mathbb{R}^n \), define

\[
\bar{A}_p = \sup_{x \in S_{n}^{-1}} x^T A_p x 1_{E_1 \cap E_2} \quad \text{and} \quad \bar{A}_p = A_p 1_{E_2} .
\]
Then we may write the decomposition
\[
\tilde{\lambda}_p = \sup_{x \in \mathbb{R}^n + \frac{2896}{\sqrt{np}} B_2^2} x^T A_p x = \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} + 2 \sup_{z \in \frac{2896}{\sqrt{np}} B_2^2} z^T A_p \left( \frac{1}{\sqrt{n}} + \frac{z}{2} \right).
\]

Then
\[
\sup_{p \in \left[ \frac{64}{\log n}, 1 \right]} |\lambda_p - \mathbb{E}\lambda_p| \leq 2 \sup_{p \in \left[ \frac{64}{\log n}, 1 \right]} \left| \sup_{z \in \frac{2896}{\sqrt{np}} B_2^2} \left( z^T A_p \left( \frac{1}{\sqrt{n}} + \frac{z}{2} \right) \right) - \mathbb{E} \sup_{z \in \frac{2896}{\sqrt{np}} B_2^2} \left( z^T A_p \left( \frac{1}{\sqrt{n}} + \frac{z}{2} \right) \right) \right| \mathbb{I}_{E_2}
\]
\[
+ \sup_{p \in \left[ \frac{64}{\log n}, 1 \right]} \left| \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} - \mathbb{E} \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} \right| \mathbb{I}_{E_2}.
\]

For the second term on the right-hand side of (3.2), since \( A_p - \lambda_p = A_p \mathbb{I}_{E_2} \), we have
\[
\mathbb{E} \sup_{p \in \left[ \frac{64}{\log n}, 1 \right]} \left| \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} - \mathbb{E} \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} \right| \leq \mathbb{E} \sup_{p \in [0,1]} \left| \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} - \mathbb{E} \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} \right| + 2n \mathbb{P}(E_2).
\]

Note that \( \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} = (2/n) \sum_{i<j} \mathbb{I}_{U_{i,j} < p} \). Thus, the first term on the right-hand side is just the maximum deviation between the cumulative distribution function of a uniform random variable and its empirical counterpart based on \( \binom{n}{2} \) random samples. This may be bounded by the classical Dvoretzky-Kiefer-Wolfowitz theorem [9]. Indeed, by Massart’s version [16], we have
\[
\mathbb{E} \sup_{p \in \left[ \frac{64}{\log n}, 1 \right]} \left| \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} - \mathbb{E} \frac{1}{\sqrt{n}} A_p \frac{1}{\sqrt{n}} \right| \leq 4 \int_0^\infty \exp(-2t^2)dt = \sqrt{2\pi}.
\]

Thus, the second term on the right-hand side of (3.2) is bounded in expectation by the absolute constant \( \sqrt{2\pi} + \frac{64}{\sqrt{n}} \leq 12 \) since \( \mathbb{P}(E_2) \leq \frac{32}{\sqrt{n}} \).
In order to bound the first term on the right-hand side of \((3.2)\), we write

\[
\sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} \left| \sup_{z \in \mathbb{R}^n} z^T A_p \left( \frac{T}{\sqrt{n}} + \frac{z}{2} \right) - \mathbb{E} \sup_{z \in \mathbb{R}^n} z^T A_p \left( \frac{T}{\sqrt{n}} + \frac{z}{2} \right) \right| 1_{E_2}
\]

\[
\leq \sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} \sup_{z \in \mathbb{R}^n} \left| z^T A_p \left( \frac{T}{\sqrt{n}} + \frac{z}{2} \right) - \mathbb{E} z^T A_p \left( \frac{T}{\sqrt{n}} + \frac{z}{2} \right) \right| 1_{E_2}
\]

\[
\leq \sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} \frac{2896}{\sqrt{np}} \sup_{z \in \mathbb{R}^n} \left| \frac{T}{\sqrt{n}} + \frac{z}{2} \right| \cdot \|A_p - \mathbb{E} A_p\| 1_{E_2}
\]

\[
\leq 2896 \times \left( 594 + \frac{32}{7 \sqrt{64 \log(7)}} \right) \left( 1 + \frac{1448}{\sqrt{64 \log(7)}} \right) \leq 4.5 \times 10^8,
\]

where we used that on the event \(E_2\) we have \(A_p = \overline{A}_p\) and

\[
\|A_p - \mathbb{E} A_p\| \leq \|A_p - \mathbb{E} A_p\| + \mathbb{E} \|A_p 1_{E_2}\| \leq 420 \sqrt{2np} + \frac{32}{7}.
\]

Finally, note that with probability at least \(1 - \frac{64}{n}\) for all \(p \in \left[\frac{64 \log n}{n}, 1\right]\) we have \(\overline{A}_p = \lambda_p\).

Moreover, for all \(p\),

\[
\mathbb{E} \sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} |\lambda_p - \overline{A}_p| \leq \mathbb{E} \sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} \sup_{x \in S^{n-1}} |x^T A_p x 1_{E_1 \cup E_2}| \leq nP(E_1 \cup E_2) \leq \frac{64}{7}.
\]

Thus, \((3.3)\) and the bound \(\mathbb{E} \sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} |\lambda_p - \mathbb{E} \lambda_p| 1_{E_2} \leq 2nP(E_2)\) imply

\[
\mathbb{E} \sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} |\lambda_p - \mathbb{E} \lambda_p| \leq \frac{192}{7} + \mathbb{E} \sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} |\overline{A}_p - \mathbb{E} \overline{A}_p| 1_{E_2} \leq 10^9,
\]

proving the first inequality of the theorem.

To prove the second inequality, we follow the argument of Example 6.8 in \([8]\). Denote

\[
Z = \sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} |\lambda_p - \mathbb{E} \lambda_p|\text{ and } Z'_{i,j} = \sup_{p \in \left[\frac{64 \log n}{n}, 1\right]} |\lambda'_p - \mathbb{E} \lambda_p|\text{ where } \lambda'_p\text{ is the largest eigenvalue of the adjacency matrix } A'_p\text{ of the random graph that is obtained from } A_p\text{ by replacing } U_{i,j}\text{ by an independent copy. Denoting as before the first unit eigenvector of } A_p\text{ by } v_p\text{ and the first unit eigenvector of } A'_p\text{ by } v'_p\text{ and the (random) maximizer of}
\]

Concentration of the spectral norm of Erdős-Rényi random graphs

\[
\sup_{p \in [\frac{64 \log n}{n}, 1]} |v_p^T A_p v_p - \mathbb{E} \lambda_p| \] by \( p^* \), we have

\[
(Z - Z')_+ \leq \left( \sup_{p \in [\frac{64 \log n}{n}, 1]} |v_p^T A_p v_p - \mathbb{E} \lambda_p| - \mathbb{E} \sup_{p \in [\frac{64 \log n}{n}, 1]} |v_p^T A_p v_p - \mathbb{E} \lambda_p| \right) \mathbb{1}_{Z \geq Z'}
\]

\[
\leq |v_p^T A_p v_p - \mathbb{E} \lambda_p| - \mathbb{E} \sup_{p \in [\frac{64 \log n}{n}, 1]} |v_p^T A_p v_p - \mathbb{E} \lambda_p| \mathbb{1}_{Z \geq Z'}
\]

\[
\leq 4|v_p^T (A_p - A'_p) v_p| \mathbb{1}_{Z \geq Z'}
\]

This implies \( \sum_{1 \leq i \leq j \leq n} (Z - Z')_+^2 \leq 16 \). Thus, for any \( t \geq 0 \),

\[
P \left\{ \sup_{p \in [\frac{64 \log n}{n}, 1]} |v_p^T A_p v_p - \mathbb{E} \lambda_p| - \mathbb{E} \sup_{p \in [\frac{64 \log n}{n}, 1]} |v_p^T A_p v_p - \mathbb{E} \lambda_p| \geq t \right\} \leq \exp(-t^2/32)
\]

Using the bound \( \mathbb{E} \sup_{p \in [\frac{64 \log n}{n}, 1]} |v_p^T A_p v_p - \mathbb{E} \lambda_p| \leq 10^9 \), we have for \( t' = t + 10^9 \)

\[
P \left\{ \sup_{p \in [\frac{64 \log n}{n}, 1]} |v_p^T A_p v_p - \mathbb{E} \lambda_p| \geq t' \right\} \leq \exp(-(t' - 10^9)^2/32)
\]

For \( t' \geq 2 \times 10^9 \) the claim follows.

### 3.2. Proof of Theorem 2

Let \( v_p \) denote an eigenvector corresponding to the largest eigenvalue of \( A_p \) such that \( \|v_p\| = 1 \). Recall that \( \kappa = 2 \times 835^2 \) and \( c = 1/9408 \). One of the key elements of the proof is the following new variant of a delocalization inequality of Mitra [17].

**Lemma 3.** Let \( n \geq 7 \) and \( p \geq \kappa \log^3(n)/n \). Let \( v_p \) denote an eigenvector corresponding to the largest eigenvalue \( \lambda_p \) of \( A_p \) with \( \|v_p\|_2 = 1 \). Then, with probability at least

\[
1 - 12(n - 1) \exp \left( -2c \left( \frac{\log(np)}{\log n} \right)^2 (n - 1)p \right)
\]

\[
\|v_p\|_{\infty} \leq \frac{11}{\sqrt{n}}
\]

The lemma is proved in Section 3.3 below. Based on this Lemma, we may prove Theorem 2:
Proof of Theorem 2. We apply (2.2) for the random variable \( Z = \|A_p\| \), as a function of the \( \binom{n}{2} \) independent Bernoulli random variables \( A_{i,j} = A_{i,j}^{(p)} \), \( 1 \leq i < j \leq n \). Let \( E_1 \) denote the event \( \|v_p\|_{\infty} \leq 11/\sqrt{n} \). By Lemma 3,

\[
P\{E_1\} \geq 1 - 12(n - 1) \exp\left(-\frac{1}{4704} \left(\frac{\log(np)}{\log n}\right)^2 (n - 1)p\right).
\]

For \( 1 \leq i < j \leq n \), denote by \( \lambda'_{i,j} \) the largest eigenvalue of the adjacency matrix obtained by replacing \( A_{i,j} \) (and \( A_{j,i} \)) by an independent copy \( A'_{i,j} \) and keeping all other entries unchanged. If the components of the unit eigenvector \( v_p \) (corresponding to the eigenvalue \( \lambda_p \)) are \( (v^1_p, \ldots, v^j_p) \), then

\[
V_p = E' \sum_{i<j} (\lambda_p - \lambda'_{i,j})^2 \leq 4 \sum_{i<j} E' \left[ (v^i_p)^2 (v^j_p)^2 (A_{i,j} - A'_{i,j})^2 \right]
\]

\[
= 4 \sum_{i<j} (v^i_p)^2 (v^j_p)^2 (p + (1 - 2p)A_{i,j})_+.
\]

Since \( (A_{i,j} - A'_{i,j})^2 \leq 1 \) and \( \sum_{i<j} (v^i_p)^2 = 1 \), we always have \( V_p \leq 4 \). On the event \( E_1 \), we have a better control:

\[
V_p \mathbb{1}_{E_1} \leq \frac{4 \cdot 11^4}{n^8} \left( \binom{n}{2} p + (1 - 2p) \sum_{i<j} A_{i,j} \right).
\]

Let \( E_2 \) denote the event that \( \sum_{i<j} A_{i,j} \leq 2E \sum_{i<j} A_{i,j} \leq pn(n-1) \). By Bernstein’s inequality,

\[
P\{E_2\} \geq 1 - \exp\left(-\frac{3p(n-1)}{8} \right),
\]

Then

\[
V_p \mathbb{1}_{E_1 \cap E_2} \leq 11^5 p.
\]

Thus,

\[
E \left[ (V_p)^2 \right] = E \left[ (V_p)^2 \mathbb{1}_{E_1 \cap E_2} \right] + E \left[ (V_p)^2 \mathbb{1}_{E_1^c \cap E_2} \right]
\]

\[
\leq (11^5 p)^{k/2} + 4^{k/2} \left( P\{E_1\} + P\{E_2\} \right)
\]

\[
\leq 2(11^5 p)^{k/2},
\]

whenever \( P\{E_1\} + P\{E_2\} \leq (11^5 p/4)^{k/2} \). This holds whenever

\[
24(n - 1) \exp\left(-\frac{1}{4704} \left(\frac{\log(np)}{\log n}\right)^2 (n - 1)p\right) \leq (11^5 p/4)^{k/2},
\]
Concentration of the spectral norm of Erdős-Rényi random graphs

13

guaranteed by our assumption on \(k\). The proof of the bound for the upper tail follows from (2.2). The bound for the variance follows from the Efron-Stein inequality (2.4).

For the bound for the lower tail we use (2.3). Note that

\[
\max_{i<j}(\lambda_p - \lambda'_{i,j}) + \mathbb{1}_{E_1} \leq 2 \max_{i<j}(v_p^iv_p^j(A_{i,j} - A'_{i,j})) + \mathbb{1}_{E_1} \leq \frac{242}{n},
\]

and therefore

\[
\mathbb{E} \max_{i<j}(v_p^iv_p^j(A_{i,j} - A'_{i,j}))^k + \mathbb{1}_{E_1} \leq \left(\frac{242}{n}\right)^k.
\]

Moreover,

\[
\mathbb{E} \max_{i<j}(2v_p^iv_p^j(A_{i,j} - A'_{i,j}))^k + \mathbb{1}_{E_1} \leq 2^{k+2} \left(\frac{1}{4704} \left(\frac{\log(np)}{\log(n)}\right)^2 (n-1)p\right).
\]

We require

\[
\left(\frac{242}{n}\right)^k \geq 3 \times 2^{k+2} (n-1) \exp\left(-\frac{1}{4704} \left(\frac{\log(np)}{\log(n)}\right)^2 (n-1)p\right)
\]

which holds whenever

\[
k \leq \frac{1}{4704} \left(\frac{\log(np)}{\log(n)}\right)^2 (n-1)p - \log(12(n-1)).
\]

Under this condition

\[
\left(\mathbb{E} \max_{i<j}(v_p^iv_p^j(A_{i,j} - A'_{i,j}))^k + \mathbb{1}_{E_1}\right)^{1/2} \leq \frac{484}{n}.
\]

Under our conditions for \(k\) and \(p\), we have \(k(484/n)^2 \leq 2 \cdot 115^5 p\) and therefore (2.3) implies the last inequality of Theorem 2.

**Remark.** It is tempting to understand if different approaches may lead to a simplified proof of Theorem 2 with the weaker condition of \(p \geq \frac{\log n}{n}\). Perturbation theory based approach has been used by [11] for the analysis of concentration of \(\|A_p\|\) around its expectation. To compare with this paper, in this remark we assume that \(A_p\) is the adjacency matrix of an Erdős-Rényi random graph with loops, that is, all vertices link to themselves, each with probability \(p\). Our results may be adapted to this case in a straightforward manner via minor changes in the constant factors. It can be shown (see formula in (6.17) in Section 6 of [11]) that when \(\|A_p - \mathbb{E}A_p\| < \|A_p\|\),

\[
\|A_p\| = \sum_{j=0}^{\infty} p^j \mathbb{1}^j \left(\frac{A_p - \mathbb{E}A_p}{\lambda_p}\right)^j \mathbb{1},
\]

(3.4)
where $\overline{1} \in \mathbb{R}^n$ is the vector whose components are all equal to 1. Theorem 6.2 in [11] (which is based on a thorough analysis of the sum (3.4)) shows that, for any $\xi \in [2, A_0 \log \log (n)]$, provided that $\frac{p_m}{1-p} \geq C_0^2 \log 4 \xi(n)$, we have, with probability at least $1 - \exp \left( -\nu \log \xi(n) \right)$,

$$\|A_p\| = \mathbb{E}\|A_p\| + \frac{\overline{1}^T (A_p - \mathbb{E}A_p)\overline{1}}{n} + O \left( \frac{\log 4 \xi(n)}{1-p} \sqrt{n} \right),$$

(3.5)

where the constant factors in the $O$-notation may depend on $\xi$, and $\nu, A_0 \geq 10$ are absolute constants. It can be easily seen that, up to an absolute constant factor, this bound implies the variance bound (1.4) but only in the regime $p \geq c_0 \log \frac{\xi(n)}{n}$, where $c_0$ is an absolute constant. Moreover, it appears that the probability with which (3.5) holds is not sufficient to recover Theorem 2 in a straightforward manner. Indeed, we know that (3.5) does not hold on the event $E$ with $\mathbb{P}\{E\} \leq \exp \left( -\nu \log \xi(n) \right)$. Let us consider the moments of $\|A_p\|$ when $E$ holds. It can be shown using (1.2) that for some absolute $C > 0$

$$\mathbb{E}(\|A_p\| - \mathbb{E}\|A_p\|)^k \leq \sqrt{\mathbb{E}(\|A_p\| - \mathbb{E}\|A_p\|)^{2k} \mathbb{P}\{E\}} \leq (Ck)^\frac{k}{2} \exp \left( -\nu \log \xi(n)/2 \right).$$

To get the same bound as in Theorem 2 we need $(Ck)^\frac{k}{2} \exp \left( -\nu \log \xi(n)/2 \right) \leq (C\nu k)^\frac{k}{2}$, which holds when

$$k \leq \frac{2\nu \log \xi(n)}{\log \left( \frac{C}{\nu k} \right)}. $$

The last inequality is more restrictive than what is required in Theorem 2 when $p \geq \frac{c \log^{\xi+2}(n)}{n \log^2 (np)}$ for some absolute constant $c > 0$. To sum up, compared to (3.5) our Theorem 2 has a different proof and provides tighter results in some natural situations.

### 3.3. Delocalization bounds

In this section we prove the “delocalization” inequalities that state that the eigenvector $v_p$ corresponding to the largest eigenvalue of $A_p$ is close to the “uniform” vector $n^{-1/2} \overline{1}$. The following lemma is crucial in the proof of Theorem 1. This proof is based on an argument of Mitra [17]. However, we need to modify it to get uniformity and also significantly better concentration guarantees.

**Lemma 4.** Let $n \geq 7$ and $q \in \left[ \frac{4 \log n}{n}, \frac{1}{2} \right]$. Then, with probability $1 - 4 \exp(-nq/64)$,

$$\sup_{p \in [q, 2q]} \left\| v_p - \frac{\overline{1}}{\sqrt{n}} \right\|_2 \leq \frac{2896}{\sqrt{nq}}.$$
Proof. Since the graph with adjacency matrix $A_q$ is connected with probability at least $1 - (n - 1) \exp(-nq/2)$ (see, e.g., [19, Section 5.3.3]), by monotonicity of the property of connectedness, the same holds simultaneously for all graphs $A_p$ for $p \in [q, 2q]$. Also, by the Perron-Frobenius theorem, if the graph is connected, the direction of $v_p$ can be chosen in a way such that the components of $v_p$ are all nonnegative for all $p \in [q, 2q]$. Moreover, the corresponding eigenvalue $\lambda_p$ has multiplicity one as well as the corresponding eigenspace is one dimensional. For the remainder of the proof we work on the event that corresponding graphs are connected.

Note that there exists a unique vector $v_p^\perp$ with $(v_p^\perp, v_p) = 0$ and $\|v_p^\perp\|_2 = 1$ such that
\[
\frac{1}{\sqrt{n}} = \alpha v_p + \beta v_p^\perp
\] (3.6)
for some $\alpha, \beta \in \mathbb{R}$. By Lemma 2, we have with probability at least $1 - \exp(-nq/64) - (n - 1) \exp(-nq/2)$,
\[
\sup_{p \in [q, 2q]} \|A_p - \mathbb{E}A_p\| \leq 420 \sqrt{nq}.
\]
Notice that $\mathbb{E}A_p = pn \frac{\mathbf{1}^T}{\sqrt{n}} - pI_n$, where $I_n$ is an identity $n \times n$ matrix. Using that $\alpha = \frac{\mathbf{1}^T}{\sqrt{n}}, v_p$,
\[
(A_p - \mathbb{E}A_p)v_p = \lambda_p v_p - pn \frac{\mathbf{1}^T}{\sqrt{n}} v_p + pv_p
\]
\[
= \lambda_p v_p - p\alpha \frac{\mathbf{1}}{\sqrt{n}} + pv_p
\]
\[
= \lambda_p v_p - p\alpha(\alpha v_p + \beta v_p^\perp) + pv_p
\]
\[
= (\lambda_p + p - p\alpha^2)v_p - p\alpha \beta v_p^\perp.
\]
This leads to
\[
(\lambda_p + p - p\alpha^2)^2 \leq 420^2 nq.
\] (3.7)
Since $\alpha \in [0, 1]$, this implies that, with probability at least $1 - \exp(-nq/64) - (n - 1) \exp(-nq/2)$, simultaneously for all $p \in [q, 2q]$
\[
\lambda_p \leq p(n - 1) + 420 \sqrt{nq}.
\] (3.8)
We may get a lower bound for $\lambda_p$ by noting that
\[
\lambda_p \geq \frac{1}{n} \mathbf{1}^T A_p \mathbf{1} = \frac{2}{n} \sum_{i < j} \mathbb{1}_{U_{ij} < p}.
\]
Applying Massart’s version of the Dvoretzky-Kiefer-Wolfowitz theorem [16], we have, for all $t \geq 0$,
\[
P \left\{ \sup_{p \in [0,1]} \left| \frac{2}{n} \sum_{i < j} \mathbb{1}_{U_{ij} < p} - (n - 1)p \right| \geq (n - 1)t \right\} \leq 2 \exp \left( -n(n - 1)t^2 \right).
\]
Choosing $t = \frac{\sqrt{nq}}{n}$, we have, with probability at least $1 - 2 \exp(-nq/2)$, simultaneously for all $p \in [q, 2q]$,

$$\lambda_p \geq p(n - 1) - \sqrt{nq} .$$  \hspace{1cm} (3.9)

This lower bound, together with (3.7) gives

$$\alpha \geq \alpha^2 \geq \frac{\lambda_p + p}{pn} - \frac{420\sqrt{nq}}{pn} \geq 1 - \frac{421}{\sqrt{nq}}$$ \hspace{1cm} (3.10)

with probability at least $1 - \exp(-nq/64) - (n - 1) \exp(-nq/2) - 2 \exp(-nq/2) \geq 1 - 4(n - 1) \exp(-nq/64)$. For the rest of the proof, we denote this event by $E$.

Next, write

$$\left\| \frac{T}{\sqrt{n}} - v_p \right\|_2 \leq \left\| A_p \frac{T}{\sqrt{n}} - v_p \right\|_2 + \left\| A_p \frac{T}{\sqrt{n}} - \frac{T}{\sqrt{n}} \right\|_2 .$$  \hspace{1cm} (3.11)

We analyze both terms on the right-hand side. Observe that $E A_p \frac{T}{\sqrt{n}} = \frac{(n-1)pT}{\sqrt{n}}$. The second term on the right-hand side of (3.11) may be bounded on the event $E$, for all $p \in [q, 2q]$, as

$$\left\| A_p \frac{T}{\sqrt{n}} - \frac{T}{\sqrt{n}} \right\|_2 \leq 1 - \frac{\lambda_p}{\sqrt{n}} \leq \left\| A_p \frac{T}{\sqrt{n}} - \frac{T}{\sqrt{n}} \right\|_2 + \left\| (n - 1)p - \lambda_p \right\|_2 \leq \left\| A_p - \frac{T}{\sqrt{n}} \right\| + \left\| (n - 1)p - \lambda_p \right\| \leq 1640 \sqrt{nq} .$$

Thus, on the event $E$, for all $p \in [q, 2q]$,

$$\left\| \frac{T}{\sqrt{n}} - v_p \right\|_2 \leq \left\| A_p \frac{T}{\sqrt{n}} - v_p \right\|_2 + 1640 \sqrt{nq} .$$

For each $p$, we may write $v_p^i = \sum_{i=2}^{n} \gamma_i v_p^i$, where $v_p^i$ is the $i$-th orthonormal eigenvector of $A_p$. Then

$$A_p \frac{T}{\sqrt{n}} = \alpha v_p + \beta \sum_{i=2}^{n} \gamma_i \lambda_i v_p^i ,$$

where $\lambda_i$ is $i$-th eigenvalue of $A_p$. By the Perron-Frobenius theorem, we have $|\lambda_i| \leq \lambda_p$ for all $i = 2, \ldots, n$. Moreover, from Füredi and Komlós [12, Lemmas 1 and 2], for
all $t \in \mathbb{R}$ we have that $|\lambda_i| \leq \|A_p - \frac{t}{\sqrt{n}} \mathbf{1}_n\|_F$ for $i \geq 2$. Choosing $t = np$ we obtain $|\lambda_i| \leq \|A_p - \mathbb{E}A_p\| + p\|I_n\| \leq 420\sqrt{nq} + p \leq 422\sqrt{nq}$. Thus, using (3.10), on the event $E$,

$$
\left\| \frac{\mathbf{1}}{\sqrt{n}} v_p - \lambda_p \right\|_2 \leq 1 - \alpha + \beta \max_{i \geq 2} \frac{|\lambda_i|}{\lambda_p} + \frac{1640}{\sqrt{nq}} + \frac{2061}{\sqrt{nq}} \leq 2896\sqrt{nq},
$$
as desired. \hfill \square

We close this section by proving the “weak” delocalization bound of Lemma 3.

**Proof of Lemma 3.** We use the notation introduced in the proof of Lemma 4. Here we fix $p \geq \kappa \log^3 n/n$. Fix $\ell \in \mathbb{N}$ and write

$$
\|v_p\|_\infty \leq \left\| \left( \frac{A_p}{\lambda_p} \right)^\ell \frac{\mathbf{1}}{\sqrt{n}} - v_p \right\|_\infty + \left\| \left( \frac{A_p}{\lambda_p} \right)^\ell \frac{\mathbf{1}}{\sqrt{n}} \right\|_\infty. \quad (3.12)
$$

We bound both terms on the right-hand side. We start with the second term and rewrite it as

$$
\left\| \left( \frac{A_p}{\lambda_p} \right)^\ell \frac{\mathbf{1}}{\sqrt{n}} \right\|_\infty = \frac{1}{\sqrt{n}} \left| (n-1)p \right|^{\ell} \left\| \left( \frac{A_p}{(n-1)p} \right)^\ell \right\|_\infty.
$$

Denote by $D_i = \sum_{j=1}^n A_{i,j}$ the degree of vertex $i$. By standard tail bounds for the binomial distribution we have, for a fixed $i$ and $0 \leq \Delta \leq 1$,

$$
P\{D_i < p(n-1) - p(n-1)\Delta\} \leq \exp \left( -\frac{\Delta^2 p(n-1)}{2} \right)
$$

and

$$
P\{D_i > p(n-1) + p(n-1)\Delta\} \leq \exp \left( -\frac{3\Delta^2 p(n-1)}{8} \right).$$

Using the union bound, we have

$$
P\left\{ \max_i |D_i - p(n-1)| > p(n-1)\Delta \right\} \leq 2(n-1) \exp \left( -\frac{3\Delta^2 p(n-1)}{8} \right).$$

We denote the event $\max_i |D_i - p(n-1)| \leq p(n-1)\Delta$ by $E_1$. Observe that when $E_1$ holds we have $D_i \leq p(n-1)(1+\Delta)$ and $D_i \geq p(n-1)(1-\Delta)$ for all $i$.

Assume that $u \in \mathbb{R}^n$ is such that

$$
\|u - \mathbf{1}\|_\infty \leq 2t\Delta \quad (3.13)
$$
for some \( t \leq \ell \). In what follows we choose \( \ell = \left\lfloor \frac{21 \log n}{\log(np)} \right\rfloor \) and \( \Delta = \frac{\log(np)}{12 \log n} \). Observe that \( \ell \Delta \leq \frac{1}{2} \). Since \( t \Delta^2 \leq \ell \Delta^2 \leq \frac{1}{2} \Delta \), we have \( \Delta + 2t \Delta^2 \leq 2\Delta \). Thus, on the event \( E_1 \), using the last inequality together with (3.13),

\[
\left( \frac{A_p}{(n-1)p} u \right)_i \leq p(n-1)(1+\Delta)(1+2t\Delta) = 1+\Delta+2t\Delta+2t\Delta^2 \leq 1+(t+1)\Delta .
\] (3.14)

Now consider the term \( \left\lfloor \frac{n-1}{\lambda_p} \right\rfloor^\ell \). Using (3.9) we have, with probability at least \( 1 - 2 \exp(-np/2) \) (denote the corresponding event by \( E_2 \)),

\[
\left| \frac{(n-1)p}{\lambda_p} \right|^\ell \leq \left( 1 - \frac{1}{\sqrt{p(n-1)}} \right)^{-\ell} .
\]

Since \( \ell \leq \sqrt{p(n-1)} \), we obtain \( \left| \frac{(n-1)p}{\lambda_p} \right|^\ell \leq e \). Thus, applying (3.14) \( \ell \) times for vectors satisfying (3.13), on the event \( E_1 \cap E_2 \), we have, for all \( i \),

\[
\left( \frac{A_p}{\lambda_p} \right)_i^\ell \geq \left( \frac{(n-1)p}{\lambda_p} \right)^\ell (1-2\ell\Delta) \geq 0 .
\] (3.15)

Analogously, applying (3.15) \( \ell \) times, on the event \( E_1 \cap E_2 \), we have

\[
\left( \frac{A_p}{\lambda_p} \right)_i^\ell \geq \left( \frac{(n-1)p}{\lambda_p} \right)^\ell (1-2\ell\Delta) \geq 0 .
\]

Hence, on the event \( E_1 \cap E_2 \),

\[
\left\| \left( \frac{A_p}{\lambda_p} \right)_i^\ell \frac{T}{\sqrt{n}} \right\|_\infty \leq 2e \frac{1}{\sqrt{n}} .
\] (3.16)

Next we bound the first term on the right-hand side of (3.12). Recall that for the decomposition \( T/\sqrt{n} = \alpha v_p + \beta v_p^\perp \) from (3.10) we have \( \alpha \geq 1 - \frac{421}{\sqrt{np}} \) on an event \( E_3 \) of probability at least \( 1 - 4(n-1) \exp(-np/64) \). As before, we may write \( v_p^\perp = \sum_{i=2}^n \gamma_i v_i^p \), where \( v_i^p \) is the \( i \)-th orthonormal eigenvector of \( A_p \). Using \( T/\sqrt{n} = \alpha v_p + \beta v_p^\perp \), we have

\[
\left( \frac{A_p}{\lambda_p} \right)_i^\ell \frac{T}{\sqrt{n}} = \alpha v_p + \beta \sum_{i=2}^n \gamma_i v_i^p \left( \frac{\lambda_i}{\lambda_p} \right)^\ell ,
\]
where $\lambda_i$ is $i$-th eigenvalue of $A_p$. Using Füredi and Komlós [12, Lemmas 1 and 2] once again, for all $t \in \mathbb{R}$ we have that $|\lambda_i| \leq \left| A_p - t\frac{1}{\sqrt{n}} \frac{1}{\sqrt{n}} \right|$ for $i \geq 2$. Choosing $t = np$ we obtain $|\lambda_i| \leq \|A_p - E A_p\| + p\|I_n\| \leq 420\sqrt{np} + p \leq 422\sqrt{np}$ on an event $E_4$ of probability at least $1 - 4(n-1) \exp(-np/64)$. Thus, on $E_4$ we have $\frac{|\lambda_i|}{\lambda_p} \leq \frac{835}{\sqrt{np}}$ for $i \geq 2$, and therefore

$$
\left\| \left( A_p - t \frac{1}{\sqrt{n}} \frac{1}{\sqrt{n}} \right) \frac{\lambda_p}{\lambda_p} \right\|_\infty \leq (1 - \alpha) \|v_p\|_\infty + \beta \max_{i \geq 2} \left( \frac{|\lambda_i|}{\lambda_p} \right)^\ell.
$$

(3.17)

Define $\kappa_1 = \frac{\log(835)}{\log(2 \times 835^2)}$. Observe that $\kappa_1 < \frac{1}{2}$. Using $np \geq 2 \times 835^2 = \kappa$, $\beta \max_{i \geq 2} \left( \frac{|\lambda_i|}{\lambda_p} \right)^\ell \leq \beta \left( \frac{835}{\sqrt{np}} \right)^\ell$

$$
\leq \left( \frac{835}{(np)^{\kappa_1}} \right) \exp \left( (1 - \kappa_1) \log \left( \frac{1}{np} \frac{21 \log n}{\log(np)} \right) \right)
$$

$$
\leq \exp \left( -21 \left( \frac{1}{2} - \kappa_1 \right) \log n \right) \leq \frac{1}{\sqrt{n}},
$$

where we used $\left( \frac{835}{(np)^{\kappa_1}} \right)^\ell \leq 1$ and the inequality $21 \left( \frac{1}{2} - \kappa_1 \right) > \frac{1}{2}$. Finally, on the event $E_1 \cap E_2 \cap E_3 \cap E_4$ of probability at least

$$
1 - 2(n-1) \exp \left( -\frac{3 \Delta^2 p(n-1)}{8} \right) - 2 \exp(-np/2) - 8(n-1) \exp(-np/64)
$$

$$
\geq 12(n-1) \exp \left( -\frac{3 \Delta^2 p(n-1)}{8} \right),
$$

we have, using the decomposition (3.12) combined with (3.16) and (3.17), that

$$
\|v_p\|_\infty \leq \frac{1}{\alpha} \left( 1 + 2e \sqrt{\frac{1}{n}} \right) \leq \frac{1}{1 - \frac{42}{\sqrt{np}}} \left( 1 + 2e \sqrt{\frac{1}{n}} \right) \leq \frac{11}{\sqrt{n}}.
$$

3.4. Proof of Proposition 1

It suffices to prove that

$$
\mathbb{E} \sup_{p \in [0, \frac{64 \log n}{n}]} |\lambda_p - E \lambda_p| \leq 5 \sqrt{16 + 2 \log \log n}.
$$

Observe that

$$
\mathbb{E} \sup_{p \in [0, 1]} |\lambda_p - E \lambda_p| \leq \mathbb{E} \sup_{p \in [0, \frac{64 \log n}{n}]} |\lambda_p - E \lambda_p| + \mathbb{E} \sup_{p \in [\frac{64 \log n}{n}, 1]} |\lambda_p - E \lambda_p|.
$$
Let $p_0, p_1, \ldots, p_M$ be such that $0 = p_0 \leq p_1 \leq \cdots \leq p_M = \frac{64 \log n}{n}$ and $\mathbb{E}(\lambda_{p_j} - \lambda_{p_{j-1}}) = \varepsilon$ for some $\varepsilon > 0$ to be specified later. Such a choice is possible since $\lambda_p$ is nondecreasing in $p$. We have

$$
\varepsilon M = \mathbb{E} \lambda_{p_M} \leq \mathbb{E} \| A_{p_M} - \mathbb{E} A_{p_M} \| + \| \mathbb{E} A_{p_M} \| \leq 170 \sqrt{np} + np \leq 1424 \log n. \quad (3.18)
$$

Denote for $p \in [0, p_M]$ the value $\pi_+[p] = \min\{q \in \{p_0, p_1, \ldots, p_M\} \mid q \geq p\}$ and $\pi_-[p] = \max\{q \in \{p_0, p_1, \ldots, p_M\} \mid p \geq q\}$. We have

$$
\mathbb{E} \sup_{p \in [0, \frac{64 \log n}{n}]} |\lambda_p - \mathbb{E} \lambda_p| \\
= \mathbb{E} \sup_{p \in [0, \frac{64 \log n}{n}]} \max\{\lambda_p - \mathbb{E} \lambda_p, \mathbb{E} \lambda_p - \lambda_p\} \\
\leq \mathbb{E} \sup_{p \in [0, \frac{64 \log n}{n}]} \max\{\lambda_{\pi_+[p]} - \mathbb{E} \lambda_{\pi_+[p]} + \varepsilon, \mathbb{E} \lambda_{\pi_-[p]} - \lambda_{\pi_-[p]} + \varepsilon\} \\
= \varepsilon + \mathbb{E} \sup_{p \in [0, \frac{64 \log n}{n}]} \max\{\lambda_{\pi_+[p]} - \mathbb{E} \lambda_{\pi_+[p]}, \mathbb{E} \lambda_{\pi_-[p]} - \lambda_{\pi_-[p]}\} \\
\leq \varepsilon + \mathbb{E} \sup_{q \in \{p_0, \ldots, p_M\}} |\lambda_q - \mathbb{E} \lambda_q|.
$$

Since for each $p_i$, the random variable $|\lambda_q - \mathbb{E} \lambda_q|$ has sub-Gaussian tails by (1.2), for their maximum we obtain the bound

$$
\mathbb{E} \sup_{q \in \{p_0, \ldots, p_M\}} |\lambda_q - \mathbb{E} \lambda_q| \leq 4 \sqrt{2 \log 2M}.
$$

Finally, using (3.18)

$$
\mathbb{E} \sup_{p \in [0, \frac{64 \log n}{n}]} |\lambda_p - \mathbb{E} \lambda_p| \leq \inf_{\varepsilon > 0} \left( \varepsilon + 4 \sqrt{2 \log(2848 \log n/\varepsilon)} \right) \leq 5 \sqrt{2 \log(2848 \log n)},
$$

as desired.

### 3.5. Proof of Proposition 2

The proof is based on two standard facts that may be found in [2]. For $k \geq 2$, let $T_k$ denote the number of connected components in a random graph $G(n, p)$ that are trees on $k$ vertices. By Cayley’s formula, $\mathbb{E} T_k \leq \binom{n}{k} k^{k-2} p^{k-1}$. Now we estimate the probability that there are trees of size at least $k + 1$. Although the asymptotic behavior of this quantity is well understood, in what follows we need a non-asymptotic upper bound. By Markov’s inequality and standard estimates, this probability is bounded by

$$
\mathbb{P} \left\{ \sum_{k+1}^{\infty} T_k \geq 1 \right\} \leq \sum_{j=k+1}^{\infty} \left( \binom{n}{j} \right) j^{j-2} p^{j-1} \leq \sum_{j=k+1}^{\infty} \left( \frac{en}{j} \right)^j j^{j-2} p^{j-1} = \sum_{j=k+1}^{\infty} \frac{en}{j} \left( enp \right)^{j-1}.
$$
The claim follows by observing that for $k = 3$ the event that the graph is a forest is bounded by $\sum_{k=3}^{\infty} c^k = \frac{c^3}{1-c}$. Finally, by the version [8, Example 3.14] of the concentration bound of Alon, Krivelevich and Vu [1], Lemma 1, and the monotonicity of $\lambda_p$, with probability at least $1 - \frac{2}{n}$ we have $\lambda_{n-k/(k-1)} \leq \lambda_{n-1} \leq (173 + \sqrt{32}) \sqrt{\log n} < 179 \sqrt{\log n}$.

Let $E_1$ denote the event that there are no trees of size greater than $k + 1$, let $E_2$ denote the event that the graph is a forest, and let $E_3$ denote the event that $\lambda_{n-k/(k-1)} < 179 \sqrt{\log n}$. Using Jensen’s inequality and the monotonicity of $\lambda_p$, we have

$$\mathbb{E} \sup_{p \in [0,n^{-k/(k-1)}]} |\lambda_p - \mathbb{E}\lambda_p| \leq 2\mathbb{E} \sup_{p \in [0,n^{-k/(k-1)}]} |\lambda_p| = 2\mathbb{E}\lambda_{n-k/(k-1)}.$$  

Since the largest eigenvalue of a forest consisting of trees of size at most $k$ is bounded by $\sqrt{k-1}$ (see e.g. [2]), we have, by the estimates above,

$$\mathbb{E}\lambda_{n-k/(k-1)} \leq \mathbb{E}\lambda_{n-k/(k-1)} \mathbb{I}_{E_1} + \mathbb{E}\lambda_{n-k/(k-1)} \mathbb{I}_{E_2} + \mathbb{E}\lambda_{n-k/(k-1)} \mathbb{I}_{E_3} \mathbb{I}_{E_1} \mathbb{I}_{E_2} \mathbb{I}_{E_3} + 2n\mathbb{P}\{E_3\}$$

$$\leq \mathbb{E}\lambda_{n-k/(k-1)} \mathbb{I}_{E_1} + \mathbb{E}\lambda_{n-k/(k-1)} \mathbb{I}_{E_2} + 179 \sqrt{\log n} (\mathbb{P}\{E_1\} + \mathbb{P}\{E_2\}) + 2n\mathbb{P}\{E_3\}$$

$$\leq \sqrt{k-1} + 179 \sum_{j=k+1}^{\infty} \frac{en\sqrt{\log n}}{j^2} \left( \frac{e}{n^{1/(k-1)}} \right)^{j-1}$$

$$+ \frac{179 \sqrt{\log n}}{(1 - n^{-1/(k-1)}) n^{3/(k-1)}} + 4$$

$$\leq \sqrt{k-1} + 4 + \frac{179 e^{k+1}}{(k+1)^2} \left( \frac{\sqrt{\log n}}{n^{1/(k-1)}} \right) / \left( 1 - \frac{e}{n^{1/(k-1)}} \right)$$

$$+ \frac{179 \sqrt{\log n}}{(1 - n^{-1/(k-1)}) n^{3/(k-1)}} .$$

The claim follows by observing that for $k \geq 2$, $\frac{\sqrt{\log n}}{n^{1/(k-1)}} \leq \frac{\sqrt{\log n}}{n^{1/(k-1)}} \leq c_k$, where $c_k$ depends only on $k$.

References


