Bernoulli **26**(3), 2020, 2253–2274
https://doi.org/10.3150/19-BEJ1192

Concentration of the spectral norm of Erdős–Rényi random graphs

GÁBOR LUGOSI[1,2,3], SHAHAR MENDELSON[4] and NIKITA ZHIVOTOVSKIY[5]

[1]*Department of Economics and Business, Pompeu Fabra University, Barcelona, Spain. E-mail: gabor.lugosi@upf.edu*
[2]*ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain*
[3]*Barcelona Graduate School of Economics, Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain*
[4]*The Australian National University, ACT 2601, Canberra, Australia. E-mail: shahar.mendelson@anu.edu.au*
[5]*Google Research, Brain Team, Brandschenkestrasse 110, 8002 Zürich, Switzerland E-mail: nikita.zhivotovskiy@phystech.edu*

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 (*Israel J. Math.* **131** (2002) 259–267) and some of the more recent results of Erdős et al. (*Ann. Probab.* **41** (2013) 2279–2375). Both results are consistent with the asymptotic result of Füredi and Komlós (*Combinatorica* **1** (1981) 233–241) that holds for fixed p as $n \to \infty$.

Keywords: concentration; empirical processes; random graphs

1. Introduction

An Erdős–Rényi random graph $G(n,p)$, named after the authors of the pioneering work [11], is a graph defined on the vertex set $[n] = \{1, \dots, n\}$ 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

$$A_{i,j}^{(p)} = \begin{cases} 0 & \text{if } i = j, \\ \mathbb{1}_{U_{i,j} < p} & \text{if } 1 \le i < j \le n, \\ \mathbb{1}_{U_{j,i} < p} & \text{if } 1 \le j < i \le n, \end{cases} \tag{1.1}$$

where $(U_{i,j})_{1 \le i < j \le n}$ are independent random variables, uniformly distributed on $[0,1]$ and $\mathbb{1}$ stands for the indicator function. We call the family of random matrices $(A_p)_{p \in [0,1]}$ 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 [10], Benaych-Georges, Bordenave, and Knowles [3,4], 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 $t > 0$,

$$\mathbb{P}\left\{\left|\|A_p\| - \mathbb{E}\|A_p\|\right| > t\right\} \leq 4e^{-t^2/32}. \tag{1.2}$$

In particular, $\mathrm{Var}(\|A_p\|) \leq C$ 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 $\mathbb{E}\|A_p\|$ 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

$$\mathbb{E} \sup_{p \geq C \log n/n} \left|\|A_p\| - \mathbb{E}\|A_p\|\right| \leq C$$

(see Theorem 1 below). Informally, this result means that as we add new edges in the Erdős–Rényi graph process, the value $|\|A_p\| - \mathbb{E}\|A_p\||$ 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

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

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

$$\mathbb{P}\left\{\sup_{p \geq C \log n/n} \left|\|A_p\| - \mathbb{E}\|A_p\|\right| > t\right\} \leq e^{-t^2/C}, \tag{1.3}$$

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 $p \in [C\frac{\log n}{n}, 1]$ 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 [10], Mitra [17]), that is, the fact that the normalized eigenvector corresponding to the largest eigenvalue is close, in a certain sense, to the vector $(1/\sqrt{n}, \dots, 1/\sqrt{n})$. 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 $1 - \frac{C}{n^\alpha}$ (as in [17]) or with quasi-polynomial probability $1 - C\exp(-c(\log n)^\beta)$ (see, e.g., [18] or Theorem 2.6 in [10]), where any choice of the parameter β 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 $p \gg \frac{\log n}{n}$.

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

$$\operatorname{Var}\bigl(\|A_p\|\bigr) \le Cp \tag{1.4}$$

for all n and $p \ge C\log^3 n/n$. 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 λ_p of the adjacency matrix A_p defined by (1.1). Our first result shows the following theorem.

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

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

Moreover, for all $t \ge 2C$,

$$\mathbb{P}\Bigl\{ \sup_{p\in[\frac{64\log n}{n},1]} |\lambda_p - \mathbb{E}\lambda_p| \ge t \Bigr\} \le \exp\bigl(-t^2/128\bigr).$$

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 $(1/\sqrt{n}, \dots, 1/\sqrt{n})$. In Lemma 4, we prove an ℓ_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 $p \ge 64\log n/n$ 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,*

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

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

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

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

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

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 $t \ge 2C'\sqrt{\log\log n}$.

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 $\log^3(n)$. 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*

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

it holds

$$\left[\mathbb{E}\left(\|A_p\| - \mathbb{E}\|A_p\|\right)_+^k\right]^{1/k} \le (Ckp)^{\frac{1}{2}}$$

and

$$\left[\mathbb{E}\left(\|A_p\| - \mathbb{E}\|A_p\|\right)_-^k\right]^{1/k} \le \left(C'kp\right)^{\frac{1}{2}}.$$

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

$$\mathrm{Var}\left(\|A_p\|\right) \le Cp.$$

It is natural to ask whether the condition $p \ge \kappa \log^3 n/n$[1] is necessary. Although we believe that $\log^3 n$ instead of the lower powers of $\log n$ is only an artifact of our technique, the fact that the inequality $\mathrm{Var}(\|A_p\|) \le Cp$ cannot hold for all values of p is easily seen by taking $p = c/n^2$ for a positive constant c. In this case, the probability that the graph $G(n, p)$ is empty is bounded away from zero. In that case, $\|A_p\| = 0$. On the other hand, with probability bounded away from zero, the graph $G(n, p)$ contains a single edge, in which case $\|A_p\| = 1$. Thus, for $p = c/n^2$, $\mathrm{Var}(\|A_p\|) = \Omega(1)$, showing that the bound (1.2) is sharp in this range. Understanding the concentration properties of $\|A_p\|$ in the range $n^{-2} \ll p \ll \log^3(n)/n$ is an intriguing open question.[2]

The proof of Theorem 2 is presented in Section 3.2. The proof reveals that for the values of the constants one may take $\kappa = 2 \times 835^2$, $C = 966{,}306$, $C' = 1{,}339{,}945$, and $c = 1/9408$. However, these values have not been optimized. In the rest of this discussion, we assume these numerical values.

Using the moment bound with $k = t^2/(2Cp)$, Markov's inequality implies that for all $0 < t \le 2\sqrt{Cc}p\sqrt{n-1}\log(np)/(\log n \log(1/p))$,

$$\mathbb{P}\left\{\left|\|A_p\| - \mathbb{E}\|A_p\|\right| \ge t\right\} \le 2^{-t^2/(2Cp)}. \tag{2.1}$$

This result improves (1.2) in the regime when $t \ll p\sqrt{n}$ with some extra logarithmic factors and may be complemented by (1.2) for the remaining values of t. Moreover, a simple inspection of the proof of Theorem 2 shows that it may be extended in a way such that it is always not worse than the tail of (1.2) for all $t \ge 0$. The proof of this theorem is based on general moment inequalities of Boucheron, Bousquet, Lugosi, and Massart [7] (see also [8], Theorems 15.5 and 15.7) that state that if $Z = f(X_1, \dots, X_n)$ is a real random variable that is a function of the independent random variables $X_1, \dots, X_n$, then for all $k \ge 2$,

$$\left[\mathbb{E}(Z - \mathbb{E}Z)_+^k\right]^{1/k} \le \sqrt{3k}\left(\mathbb{E}\left[{V_+}^{k/2}\right]\right)^{1/k} \tag{2.2}$$

and

$$\left[\mathbb{E}(Z - \mathbb{E}Z)_-^k\right]^{1/k} \le \sqrt{4.16k}\left(\left(\mathbb{E}\left[{V_+}^{k/2}\right]\right)^{1/k} \vee \sqrt{k}\left(\mathbb{E}\left[M^k\right]\right)^{1/k}\right), \tag{2.3}$$

[1] Our analysis implies in fact a slightly better factor $\frac{\log^3 n}{n(\log\log n)^2}$ instead of $\frac{\log^3 n}{n}$.

[2] We refer to the work of Lei [15] for some recent progress in this problem.

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)

$$\mathrm{Var}(Z) \leq \mathbb{E} V_+ . \tag{2.4}$$

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 ℓ_∞ 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 [5]).

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, $\mathbb{E}\|A_{\frac{1}{n}}\|$ 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

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

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,

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

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 [2] with $\alpha = 3$,

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

Thus,

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

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 $q \in [\log n/n, \frac{1}{2}]$,*

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

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 $t > 0$,

$$\mathbb{P}\big\{\|A_p - \mathbb{E}A_p\| - \mathbb{E}\|A_p - \mathbb{E}A_p\| > t\big\} \le e^{-t^2/32}.$$

On the other hand, using the same symmetrization trick as in Lemma 1, Corollary 3.6 of Bandeira, van Handel [2] implies that for any $p \ge \log n/n$,

$$\mathbb{E}\|A_p - \mathbb{E}A_p\| \le e^{\frac{2}{3}}(2\sqrt{2np} + 84\sqrt{\log n}) \le 170\sqrt{np}. \tag{3.1}$$

These two results imply

$$\mathbb{P}\big\{\|A_p - \mathbb{E}A_p\| > 172\sqrt{np}\big\} \le e^{-np/8}.$$

Let now $q \ge \log n/n$ and for $i = 0, 1, \dots, \lceil nq \rceil$, 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,

$$\begin{aligned}
&\sup_{p\in[p_i,p_{i+1}]} \big(\|A_p - \mathbb{E}A_p\| - \|A_{p_i} - \mathbb{E}A_{p_i}\|\big) \\
&\quad\le \sup_{p\in[p_i,p_{i+1}]} \big(\|A_p - A_{p_i}\| + \|\mathbb{E}A_p - \mathbb{E}A_{p_i}\|\big) \\
&\quad= \sup_{p\in[p_i,p_{i+1}]} \big(\|A_p - A_{p_i}\| + \|\mathbb{E}A_{p-p_i}\|\big) \\
&\quad= \|A_{p_{i+1}} - A_{p_i}\| + \|\mathbb{E}A_{1/n}\| \\
&\quad\le \|A_{p_{i+1}} - A_{p_i}\| + 1
\end{aligned}$$

$$\le \mathbb{E}\|A_{1/n}\| + \big(\|A_{p_{i+1}} - A_{p_i}\| - \mathbb{E}\|A_{p_{i+1}} - A_{p_i}\|\big) + 1$$
$$\le 1 + 173\sqrt{\log n} + \sqrt{nq} \le 176\sqrt{nq},$$

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 $1 - nqe^{-nq/32} - nqe^{-nq/8} \ge 1 - e^{-nq/64}$,

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

as desired. □

Proof of Theorem 1. Denote by $\overline{1} \in \mathbb{R}^n$ the vector whose components are all equal to 1. Let $B_2^n = \{x \in \mathbb{R}^n : \|x\|_2 \le 1\}$ be the unit Euclidean ball. Let v_p denote the first unit eigenvector of A_p. Define the event E_1 that $v_p \in \frac{\overline{1}}{\sqrt{n}} + \frac{2896}{\sqrt{np}} B_2^n$ 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 $n \ge 7$,

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

Now define the event E_2 that for all $p \in [\frac{64\log n}{n}, 1]$, $\|A_p - \mathbb{E}A_p\| \le 420\sqrt{2np}$. Similarly to the calculation above, by Lemma 2, $\mathbb{P}\{E_2\} \ge 1 - \frac{32}{7n}$.

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

$$\overline{\lambda}_p = \sup_{x\in S^{n-1}} x^T A_p x \mathbb{1}_{E_1\cap E_2} \quad \text{and} \quad \overline{A}_p = A_p \mathbb{1}_{E_2}.$$

Then we may write the decomposition

$$\overline{\lambda}_p = \sup_{x\in \frac{\overline{1}}{\sqrt{n}} + \frac{2896}{\sqrt{np}} B_2^n} x^T \overline{A}_p x = \frac{\overline{1}^T}{\sqrt{n}} \overline{A}_p \frac{\overline{1}}{\sqrt{n}} + 2 \sup_{z\in \frac{2896}{\sqrt{np}} B_2^n} z^T \overline{A}_p \left(\frac{\overline{1}}{\sqrt{n}} + \frac{z}{2}\right).$$

Then

$$\sup_{p\in[\frac{64\log n}{n},1]} |\overline{\lambda}_p - \mathbb{E}\overline{\lambda}_p| \mathbb{1}_{E_2}$$
$$\le 2 \sup_{p\in[\frac{64\log n}{n},1]} \left| \sup_{z\in \frac{2896}{\sqrt{np}} B_2^n} \left(z^T \overline{A}_p \left(\frac{\overline{1}}{\sqrt{n}} + \frac{z}{2}\right)\right) - \mathbb{E} \sup_{z\in \frac{2896}{\sqrt{np}} B_2^n} \left(z^T \overline{A}_p \left(\frac{\overline{1}}{\sqrt{n}} + \frac{z}{2}\right)\right) \right| \mathbb{1}_{E_2}$$
$$+ \sup_{p\in[\frac{64\log n}{n},1]} \left| \frac{\overline{1}}{\sqrt{n}}^T \overline{A}_p \frac{\overline{1}}{\sqrt{n}} - \mathbb{E}\frac{\overline{1}^T}{\sqrt{n}} \overline{A}_p \frac{\overline{1}}{\sqrt{n}} \right| \mathbb{1}_{E_2}. \tag{3.2}$$

For the second term on the right-hand side of (3.2), since $A_p - \overline{A}_p = A_p \mathbb{1}_{\overline{E}_2}$ we have

$$\begin{aligned}
&\mathbb{E} \sup_{p\in[\frac{64\log n}{n},1]} \left| \frac{\overline{1}}{\sqrt{n}}^T \overline{A}_p \frac{\overline{1}}{\sqrt{n}} - \mathbb{E}\frac{\overline{1}^T}{\sqrt{n}} \overline{A}_p \frac{\overline{1}}{\sqrt{n}} \right| \\
&\quad \le \mathbb{E} \sup_{p\in[\frac{64\log n}{n},1]} \left| \frac{\overline{1}}{\sqrt{n}}^T A_p \frac{\overline{1}}{\sqrt{n}} - \mathbb{E}\frac{\overline{1}^T}{\sqrt{n}} A_p \frac{\overline{1}}{\sqrt{n}} \right| + 2n\mathbb{P}(\overline{E}_2).
\end{aligned}$$

Note that $\frac{\overline{1}}{\sqrt{n}}^T A_p \frac{\overline{1}}{\sqrt{n}} = (2/n)\sum_{i<j} \mathbb{1}_{U_{i,j}<p}$. 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

$$\begin{aligned}
\mathbb{E} \sup_{p\in[\frac{64\log n}{n},1]} \left| \frac{\overline{1}}{\sqrt{n}}^T A_p \frac{\overline{1}}{\sqrt{n}} - \mathbb{E}\frac{\overline{1}^T}{\sqrt{n}} A_p \frac{\overline{1}}{\sqrt{n}} \right| &\le \mathbb{E} \sup_{p\in[0,1]} \left| \frac{\overline{1}}{\sqrt{n}}^T A_p \frac{\overline{1}}{\sqrt{n}} - \mathbb{E}\frac{\overline{1}^T}{\sqrt{n}} A_p \frac{\overline{1}}{\sqrt{n}} \right| \\
&\le 4\int_{t=0}^{\infty} \exp\bigl(-2t^2\bigr)\, dt = \sqrt{2\pi}.
\end{aligned}$$

Thus, the second term on the right-hand side of (3.2) is bounded in expectation by the absolute constant $\sqrt{2\pi} + \frac{64}{7} \le 12$ since $\mathbb{P}(\overline{E}_2) \le \frac{32}{7n}$.

In order to bound the first term on the right-hand side of (3.2), we write

$$\begin{aligned}
&\sup_{p\in[\frac{64\log n}{n},1]} \left| \sup_{z\in\frac{2896}{\sqrt{np}}B_2^n} z^T \overline{A}_p \left(\frac{\overline{1}}{\sqrt{n}} + \frac{z}{2}\right) - \mathbb{E} \sup_{z\in\frac{2896}{\sqrt{np}}B_2^n} z^T \overline{A}_p \left(\frac{\overline{1}}{\sqrt{n}} + \frac{z}{2}\right) \right| \mathbb{1}_{E_2} \\
&\le \sup_{p\in[\frac{64\log n}{n},1]} \sup_{z\in\frac{2896}{\sqrt{np}}B_2^n} \left| z^T \overline{A}_p \left(\frac{\overline{1}}{\sqrt{n}} + \frac{z}{2}\right) - \mathbb{E} z^T \overline{A}_p \left(\frac{\overline{1}}{\sqrt{n}} + \frac{z}{2}\right) \right| \mathbb{1}_{E_2} \\
&\le \sup_{p\in[\frac{64\log n}{n},1]} \frac{2896}{\sqrt{np}} \sup_{z\in\frac{2896}{\sqrt{np}}B_2^n} \left\| \frac{\overline{1}}{\sqrt{n}} + \frac{z}{2} \right\| \cdot \|\overline{A}_p - \mathbb{E}\overline{A}_p\| \mathbb{1}_{E_2} \\
&\le \sup_{p\in[\frac{64\log n}{n},1]} 2896 \times \left(594 + \frac{32}{7\sqrt{np}}\right)\left(1 + \frac{1448}{\sqrt{np}}\right) \\
&\le 2896 \times \left(594 + \frac{32}{7\sqrt{64\log 7}}\right)\left(1 + \frac{1448}{\sqrt{64\log(7)}}\right) \le 4.5 \times 10^8,
\end{aligned}$$

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

$$\|\overline{A}_p - \mathbb{E}\overline{A}_p\| \le \|A_p - \mathbb{E}A_p\| + \mathbb{E}\|A_p \mathbb{1}_{\overline{E}_2}\| \le 420\sqrt{2np} + \frac{32}{7}.$$

Finally, note that with probability at least $1-\frac{64}{7n}$ for all $p \in [\frac{64\log n}{n}, 1]$ we have $\overline{\lambda}_p = \lambda_p$. Moreover, for all p,

$$\mathbb{E} \sup_{p\in[\frac{64\log n}{n},1]} |\lambda_p - \overline{\lambda}_p| \le \mathbb{E} \sup_{p\in[\frac{64\log n}{n},1]} \sup_{x\in S^{n-1}} \left|x^T A_p x \mathbb{1}_{\overline{E_1 \cup E_2}}\right| \le n\mathbb{P}(\overline{E}_1 \cup \overline{E_2}) \le \frac{64}{7}. \tag{3.3}$$

Thus, (3.3) and the bound $\mathbb{E}\sup_{p\in[\frac{64\log n}{n},1]} |\overline{\lambda}_p - \mathbb{E}\overline{\lambda}_p| \mathbb{1}_{\overline{E}_2} \le 2n\mathbb{P}(\overline{E}_2)$ imply

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

proving the first inequality of the theorem.

To prove the second inequality, we follow the argument of Example 6.8 in [8]. Denote $Z = \sup_{p\in[\frac{64\log n}{n},1]} |\lambda_p - \mathbb{E}\lambda_p|$ and $Z'_{i,j} = \sup_{p\in[\frac{64\log n}{n},1]} |\lambda'_p - \mathbb{E}\lambda_p|$ where λ'_p is the largest eigenvalue of the adjacency matrix A'_p of the random graph that is obtained from A_p by replacing $U_{i,j}$ by an independent copy. Denoting as before the first unit eigenvector of A_p by v_p and the first unit eigenvector of A'_p by v'_p and the (random) maximizer of $\sup_{p\in[\frac{64\log n}{n},1]} |v_p^T A_p v_p - \mathbb{E}\lambda_p|$ by p^*, we have

$$\begin{aligned}
(Z - Z'_{i,j})_+ &\le \Big(\sup_{p\in[\frac{64\log n}{n},1]} \left|v_p^T A_p v_p - \mathbb{E}\lambda_p\right| - \sup_{p\in[\frac{64\log n}{n},1]} \left|v_p'^T A'_p v'_p - \mathbb{E}\lambda_p\right| \Big) \mathbb{1}_{Z \ge Z'_{i,j}} \\
&\le \left|v_{p^*}^T A_{p^*} v_{p^*} - \mathbb{E}\lambda_{p^*} - v_{p^*}'^T A'_{p^*} v'_{p^*} + \mathbb{E}\lambda_{p^*}\right| \mathbb{1}_{Z \ge Z'_{i,j}} \\
&\le \left|v_{p^*}^T (A_{p^*} - A'_{p^*}) v_{p^*}\right| \mathbb{1}_{Z \ge Z'_{i,j}} \\
&\le 4\left|v_{p^*}^i v_{p^*}^j\right|.
\end{aligned}$$

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

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

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

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

For $t' \ge 2\times 10^9$ the claim follows. $\square$

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 λ_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,

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

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 λ_p) are $(v_p^1, \dots, v_p^j)$, then

$$\begin{aligned} V_+ &= \mathbb{E}' \sum_{i<j}^{n} (\lambda_p - \lambda'_{i,j})_+^2 \leq 4 \sum_{i<j}^{n} \mathbb{E}'\left[(v_p^i)^2 (v_p^j)^2 (A_{i,j} - A'_{i,j})^2\right] \\ &= 4 \sum_{i<j}^{n} (v_p^i)^2 (v_p^j)^2 \left(p + (1-2p)A_{i,j}\right)_+. \end{aligned}$$

Since $(A_{i,j} - A'_{i,j})^2 \leq 1$ and $\sum_i^n (v_p^i)^2 = 1$, we always have $V_+ \leq 4$. On the event E_1, we have a better control:

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

Let E_2 denote the event that $\sum_{i<j}^n A_{i,j} \leq 2\mathbb{E}\sum_{i<j}^n A_{i,j} \leq pn(n-1)$. By Bernstein's inequality, $\mathbb{P}\{E_2\} \geq 1 - \exp(-\frac{3pn(n-1)}{8})$. Then

$$V_+ \mathbb{1}_{E_1 \cap E_2} \leq 11^5 p.$$

Thus,

$$\begin{aligned}\mathbb{E}\big[(V_+)^{\frac{k}{2}}\big] &= \mathbb{E}\big[(V_+)^{\frac{k}{2}}\mathbb{1}_{E_1\cap E_2}\big] + \mathbb{E}\big[(V_+)^{\frac{k}{2}}(\mathbb{1}_{\overline{E_1}} + \mathbb{1}_{\overline{E_2}})\big] \\ &\le \big(11^5 p\big)^{k/2} + 4^{k/2}\big(\mathbb{P}\{\overline{E_1}\} + \mathbb{P}\{\overline{E_2}\}\big) \\ &\le 2\big(11^5 p\big)^{k/2},\end{aligned}$$

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

$$24(n-1)\exp\left(-\frac{1}{4704}\left(\frac{\log(np)}{\log n}\right)^2(n-1)p\right) \le \big(11^5 p/4\big)^{k/2},$$

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

$$\max_{i<j}\big(\lambda_p - \lambda'_{i,j}\big)_+\mathbb{1}_{E_1} \le 2\max_{i<j}\big(v_p^i v_p^j\big(A_{i,j} - A'_{i,j}\big)\big)_+\mathbb{1}_{E_1} \le \frac{242}{n},$$

and therefore

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

Moreover,

$$\begin{aligned}&\mathbb{E}\max_{i<j}\big(2v_p^i v_p^j\big(A_{i,j} - A'_{i,j}\big)\big)_+^k\mathbb{1}_{\overline{E_1}} \\ &\quad\le 2^k\mathbb{P}\{\overline{E_1}\} \le 3\times 2^{k+2}(n-1)\exp\left(-\frac{1}{4704}\left(\frac{\log(np)}{\log(n)}\right)^2(n-1)p\right).\end{aligned}$$

We require

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

which holds whenever

$$k \le \frac{\frac{1}{4704}\big(\frac{\log(np)}{\log(n)}\big)^2(n-1)p - \log(12(n-1))}{\log(\frac{n}{121})}.$$

Under this condition

$$\left(\mathbb{E}\max_{i<j}\big(v_p^i v_p^j\big(A_{i,j} - A'_{i,j}\big)\big)_+^k\right)^{\frac{1}{k}} \le \frac{484}{n}.$$

Under our conditions for k and p, we have $k(484/n)^2 \le 2\cdot 11^5 p$ and therefore (2.3) implies the last inequality of Theorem 2. $\square$

Remark. It is tempting to understand if different approaches may lead to a simplified proof of Theorem 2 with the weaker condition of $p \ge \frac{\log n}{n}$. Perturbation theory based approach has been used by [10] 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 [10]) that when $\|A_p - \mathbb{E}A_p\| < \|A_p\|$,

$$\|A_p\| = \sum_{j=0}^{\infty} p\overline{1}^T \left(\frac{A_p - \mathbb{E}A_p}{\lambda_p}\right)^j \overline{1}, \tag{3.4}$$

where $\overline{1} \in \mathbb{R}^n$ is the vector whose components are all equal to 1. Theorem 6.2 in [10] (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 $\frac{pn}{1-p} \ge C_0^2 \log^{4\xi}(n)$, we have, with probability at least $1 - \exp(-\nu \log^{\xi}(n))$,

$$\|A_p\| = \mathbb{E}\|A_p\| + \frac{\overline{1}^T (A_p - \mathbb{E}A_p)\overline{1}}{n} + O\left(\frac{\log^{2\xi}(n)}{(1-p)\sqrt{n}}\right), \tag{3.5}$$

where the constant factors in the O-notation may depend on ξ, and $\nu, A_0 \ge 10$ 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 $p \ge \frac{c_0 \log^8(n)}{n}$, 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 $\mathbb{P}\{E\} \le \exp(-\nu \log^{\xi}(n))$. Let us consider the moments of $\|A_p\|$ when E holds. It can be shown using (1.2) that for some absolute $C > 0$

$$\mathbb{E}\big(\|A_p\| - \mathbb{E}\|A_p\|\big)^k \mathbb{1}_E \le \sqrt{\mathbb{E}\big(\|A_p\| - \mathbb{E}\|A_p\|\big)^{2k} \mathbb{P}\{E\}} \le (Ck)^{\frac{k}{2}} \exp\big(-\nu \log^{\xi}(n)/2\big).$$

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

$$k \le \frac{2\nu \log^{\xi}(n)}{\log(\frac{C}{C'p})}.$$

The last inequality is more restrictive than what is required in Theorem 2 when $p \ge \frac{c\nu \log^{\xi+2}(n)}{n \log^2(np)}$ 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 $q \in [\frac{4\log n}{n}, \frac{1}{2}]$. 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 λ_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

$$\overline{1}/\sqrt{n} = \alpha v_p + \beta v_p^{\perp} \tag{3.6}$$

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 $\mathbb{E}A_p = pn\frac{\overline{1}}{\sqrt{n}}\frac{\overline{1}^T}{\sqrt{n}} - pI_n$, where I_n is an identity $n \times n$ matrix. Using that $\alpha = (\frac{\overline{1}}{\sqrt{n}}, v_p)$,

$$\begin{aligned}
(A_p - \mathbb{E}A_p)v_p &= \lambda_p v_p - pn\frac{\overline{1}}{\sqrt{n}}\frac{\overline{1}^T}{\sqrt{n}}v_p + pv_p \\
&= \lambda_p v_p - pn\alpha\frac{\overline{1}}{\sqrt{n}} + pv_p \\
&= \lambda_p v_p - pn\alpha\big(\alpha v_p + \beta v_p^{\perp}\big) + pv_p \\
&= \big(\lambda_p + p - pn\alpha^2\big)v_p - pn\alpha\beta v_p^{\perp}.
\end{aligned}$$

This leads to

$$\big(\lambda_p + p - pn\alpha^2\big)^2 \leq 420^2 nq. \tag{3.7}$$

Since $\alpha \in [0, 1]$, this implies that, with probability at least $1 - \exp(-nq/64) - (n-1) \times \exp(-nq/2)$, simultaneously for all $p \in [q, 2q]$

$$\lambda_p \leq p(n-1) + 420\sqrt{nq}. \tag{3.8}$$

We may get a lower bound for λ_p by noting that

$$\lambda_p \geq \frac{1}{n}\overline{1}^T A_p \overline{1} = \frac{2}{n}\sum_{i<j}^{n} \mathbb{1}_{U_{ij}<p}.$$

Applying Massart's version of the Dvoretzky-Kiefer-Wolfowitz theorem [16], we have, for all $t \geq 0$,

$$\mathbb{P}\left\{\sup_{p\in[0,1]}\left|\frac{2}{n}\sum_{i<j}^{n}\mathbb{1}_{U_{ij}<p} - (n-1)p\right| \geq (n-1)t\right\} \leq 2\exp\bigl(-n(n-1)t^2\bigr).$$

Choosing $t = \frac{\sqrt{nq}}{n-1}$, we have, with probability at least $1 - 2\exp(-nq/2)$, simultaneously for all $p \in [q, 2q]$,

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

This lower bound, together with (3.7) gives

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

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

$$\left\|\frac{\overline{1}}{\sqrt{n}} - v_p\right\|_2 \leq \left\|\frac{A_p}{\lambda_p}\frac{\overline{1}}{\sqrt{n}} - v_p\right\|_2 + \left\|\frac{A_p}{\lambda_p}\frac{\overline{1}}{\sqrt{n}} - \frac{\overline{1}}{\sqrt{n}}\right\|_2. \tag{3.11}$$

We analyze both terms on the right-hand side. Observe that $\mathbb{E}A_p\frac{\overline{1}}{\sqrt{n}} = \frac{(n-1)p\overline{1}}{\sqrt{n}}$. 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

$$\begin{aligned}
\left\|\frac{A_p}{\lambda_p}\frac{\overline{1}}{\sqrt{n}} - \frac{\overline{1}}{\sqrt{n}}\right\|_2 &\leq \frac{1}{\lambda_p}\left\|A_p\frac{\overline{1}}{\sqrt{n}} - \frac{(n-1)p\overline{1}}{\sqrt{n}}\right\|_2 + \frac{1}{\lambda_p}\left\|\frac{((n-1)p - \lambda_p)\overline{1}}{\sqrt{n}}\right\|_2 \\
&= \frac{1}{\lambda_p}\left\|A_p\frac{\overline{1}}{\sqrt{n}} - \mathbb{E}A_p\frac{\overline{1}}{\sqrt{n}}\right\|_2 + \frac{|(n-1)p - \lambda_p|}{\lambda_p} \\
&\leq \frac{\|A_p - \mathbb{E}A_p\| + |(n-1)p - \lambda_p|}{\lambda_p} \\
&\leq \frac{420\sqrt{nq} + 420\sqrt{nq}}{p(n-1) - \sqrt{nq}} \\
&\leq \frac{1640}{\sqrt{nq}}.
\end{aligned}$$

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

$$\left\| \frac{\overline{1}}{\sqrt{n}} - v_p \right\|_2 \le \left\| \frac{A_p}{\lambda_p} \frac{\overline{1}}{\sqrt{n}} - v_p \right\|_2 + \frac{1640}{\sqrt{nq}}.$$

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

$$\frac{A_p}{\lambda_p} \frac{\overline{1}}{\sqrt{n}} = \alpha v_p + \beta \sum_{i=2}^n \frac{\gamma_i \lambda_i v_p^i}{\lambda_p},$$

where λ_i is ith eigenvalue of A_p. By the Perron–Frobenius theorem, we have $|\lambda_i| \le \lambda_p$ for all $i = 2, \dots, n$. Moreover, from Füredi and Komlós [12], Lemmas 1 and 2, for all $t \in \mathbb{R}$ we have that $|\lambda_i| \le \|A_p - t \frac{\overline{1}}{\sqrt{n}} \frac{\overline{1}}{\sqrt{n}}^T\|$ for $i \ge 2$. Choosing $t = np$, we obtain $|\lambda_i| \le \|A_p - \mathbb{E} A_p\| + p\|I_n\| \le 420\sqrt{nq} + p \le 422\sqrt{nq}$. Thus, using (3.10), on the event E,

$$\left\| \frac{\overline{1}}{\sqrt{n}} - v_p \right\|_2 \le 1 - \alpha + \beta \max_{i \ge 2} \frac{|\lambda_i|}{\lambda_p} + \frac{1640}{\sqrt{nq}} \le \frac{2061}{\sqrt{nq}} + \frac{422\sqrt{nq}}{(n-1)p - \sqrt{nq}} \le \frac{2896}{\sqrt{nq}},$$

as desired. □

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 $p \ge \kappa \log^3 n/n$. Fix $\ell \in \mathbb{N}$ and write

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

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

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

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 $0 \le \Delta \le 1$,

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

and

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

Using the union bound, we have

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

We denote the event

$$\max_i \big|D_i - p(n-1)\big| \le p(n-1)\Delta$$

by E_1. Observe that when E_1 holds we have $D_i \le p(n-1)(1+\Delta)$ and $D_i \ge p(n-1)(1-\Delta)$ for all i.

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

$$\|u - \overline{1}\|_\infty \le 2t\Delta \tag{3.13}$$

for some $t \le \ell$. In what follows we choose $\ell = \lfloor \frac{21 \log n}{\log(np)} \rfloor$ and $\Delta = \frac{\log(np)}{42 \log n}$. Observe that $\ell\Delta \le \frac{1}{2}$. Since $t\Delta^2 \le \ell\Delta^2 \le \frac{1}{2}\Delta$, we have $\Delta + 2t\Delta^2 \le 2\Delta$. Thus, on the event E_1, using the last inequality together with (3.13),

$$\left(\frac{A_p}{(n-1)p}u\right)_i \le \frac{p(n-1)(1+\Delta)(1+2t\Delta)}{(n-1)p} = 1+\Delta+2t\Delta+2t\Delta^2 \le 1+2(t+1)\Delta. \tag{3.14}$$

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

$$\left|\frac{(n-1)p}{\lambda_p}\right|^\ell \le \left(1 - \frac{1}{\sqrt{p(n-1)}}\right)^{-\ell}.$$

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

$$\left(\left(\frac{A_p}{\lambda_p}\right)^\ell \overline{1}\right)_i = \left|\frac{(n-1)p}{\lambda_p}\right|^\ell \left(\left(\frac{A_p}{(n-1)p}\right)^\ell \overline{1}\right)_i \le e(1+2\ell\Delta) \le 2e.$$

We may similarly derive a lower bound since, for any vector satisfying (3.13),

$$\left(\frac{A_p}{(n-1)p}u\right)_i \ge \frac{p(n-1)(1-\Delta)(1-2t\Delta)}{(n-1)p} = 1-\Delta-2t\Delta+2t\Delta^2 \ge 1-2(t+1)\Delta. \tag{3.15}$$

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

$$\left(\left(\frac{A_p}{\lambda_p}\right)^\ell \overline{1}\right)_i = \left|\frac{(n-1)p}{\lambda_p}\right|^\ell \left(\left(\frac{A_p}{(n-1)p}\right)^\ell \overline{1}\right)_i \ge \left|\frac{(n-1)p}{\lambda_p}\right|^\ell (1-2\ell\Delta) \ge 0.$$

Hence, on the event $E_1 \cap E_2$,

$$\left\|\left(\frac{A_p}{\lambda_p}\right)^\ell \frac{\overline{1}}{\sqrt{n}}\right\|_\infty \le \frac{2e}{\sqrt{n}}. \tag{3.16}$$

Next, we bound the first term on the right-hand side of (3.12). Recall that for the decomposition $\overline{1}/\sqrt{n} = \alpha v_p + \beta v_p^{\perp}$ 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 $v_p^{\perp} = \sum_{i=2}^{n} \gamma_i v_p^i$, where v_p^i is the ith orthonormal eigenvector of A_p. Using $\overline{1}/\sqrt{n} = \alpha v_p + \beta v_p^{\perp}$, we have

$$\left(\frac{A_p}{\lambda_p}\right)^{\ell} \frac{\overline{1}}{\sqrt{n}} = \alpha v_p + \beta \sum_{i=2}^{n} \gamma_i v_p^i \left(\frac{\lambda_i}{\lambda_p}\right)^{\ell},$$

where λ_i is ith eigenvalue of A_p. Using Füredi and Komlós [12], Lemmas 1 and 2, once again, for all $t \in \mathbb{R}$ we have that $|\lambda_i| \leq \|A_p - t\frac{\overline{1}}{\sqrt{n}}\frac{\overline{1}^T}{\sqrt{n}}\|$ for $i \geq 2$. Choosing $t = np$ we obtain $|\lambda_i| \leq \|A_p - \mathbb{E}A_p\| + p\|I_n\| \leq 420\sqrt{np} + p \leq 422\sqrt{np}$ 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

$$\left\| \left(\frac{A_p}{\lambda_p}\right)^{\ell} \frac{\overline{1}}{\sqrt{n}} - v_p \right\|_{\infty} \leq (1-\alpha)\|v_p\|_{\infty} + \beta \max_{i\geq 2}\left(\frac{|\lambda_i|}{\lambda_p}\right)^{\ell}. \tag{3.17}$$

Define $\kappa_1 = \frac{\log(835)}{\log(2\times 835^2)}$. Observe that $\kappa_1 < \frac{1}{2}$. Using $np \geq 2 \times 835^2 = \kappa$,

$$\begin{aligned}
\beta \max_{i\geq 2}\left(\frac{|\lambda_i|}{\lambda_p}\right)^{\ell} &\leq \beta\left(\frac{835}{\sqrt{np}}\right)^{\ell} \\
&\leq \left(\frac{835}{(np)^{\kappa_1}}\right)^{\ell} \exp\left(\left(\frac{1}{2} - \kappa_1\right)\log\left(\frac{1}{np}\right)\frac{21\log n}{\log(np)}\right) \\
&\leq \exp\left(-21\left(\frac{1}{2} - \kappa_1\right)\log n\right) \leq \frac{1}{\sqrt{n}},
\end{aligned}$$

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

$$\begin{aligned}
&1 - 2(n-1)\exp\left(-\frac{3\Delta^2 p(n-1)}{8}\right) - 2\exp(-np/2) - 8(n-1)\exp(-np/64) \\
&\quad \geq 1 - 12(n-1)\exp\left(-\frac{3\Delta^2 p(n-1)}{8}\right),
\end{aligned}$$

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

$$\|v_p\|_{\infty} \leq \frac{1}{\alpha}\left(\frac{1+2e}{\sqrt{n}}\right) \leq \frac{1}{1 - \frac{421}{\sqrt{np}}}\left(\frac{1+2e}{\sqrt{n}}\right) \leq \frac{11}{\sqrt{n}}. \qquad \square$$

3.4. Proof of Proposition 1

It suffices to prove that

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

Observe that

$$\mathbb{E} \sup_{p\in[0,1]} |\lambda_p - \mathbb{E}\lambda_p| \le \mathbb{E} \sup_{p\in[0,\frac{64\log n}{n}]} |\lambda_p - \mathbb{E}\lambda_p| + \mathbb{E} \sup_{p\in[\frac{64\log n}{n},1]} |\lambda_p - \mathbb{E}\lambda_p|.$$

Let $p_0, p_1, \ldots, p_M$ be such that $0 = p_0 \le p_1 \le \cdots \le p_M = \frac{64\log n}{n}$ 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 λ_p is nondecreasing in p. We have

$$\varepsilon M = \mathbb{E}\lambda_{p_M} \le \mathbb{E}\|A_{p_M} - \mathbb{E}A_{p_M}\| + \|\mathbb{E}A_{p_M}\| \le 170\sqrt{np_M} + np_M \le 1424\log n. \tag{3.18}$$

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

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

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

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

Finally, using (3.18)

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

as desired.

3.5. Proof of Proposition 2

The proof is based on two standard facts that may be found in [6]. For $k \geq 2$, let T_k denote the number of 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

$$\mathbb{P}\left\{\sum_{k+1}^{\infty} T_k \geq 1\right\} \leq \sum_{j=k+1}^{\infty} \binom{n}{j} j^{j-2} p^{j-1} \leq \sum_{j=k+1}^{\infty} \left(\frac{en}{j}\right)^j j^{j-2} p^{j-1} = \sum_{j=k+1}^{\infty} \frac{en}{j^2} (enp)^{j-1}.$$

At the same time, Theorem 5.7 (i) in [6] states that if $p = \frac{c}{n}$ for some $c \in [0, 1)$ then probability that $G(n, p)$ is not 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 monotonicty of λ_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 λ_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., [6]), we have, by the estimates above,

$$\begin{aligned}
\mathbb{E}\lambda_{n^{-k/(k-1)}} &\leq \mathbb{E}\lambda_{n^{-k/(k-1)}} \mathbb{1}_{E_1 \cap E_2} + \mathbb{E}\lambda_{n^{-k/(k-1)}} \mathbb{1}_{E_3} (\mathbb{1}_{\overline{E_1}} + \mathbb{1}_{\overline{E_2}}) + 2n\mathbb{P}\{\overline{E_3}\} \\
&\leq \mathbb{E}\lambda_{n^{-k/(k-1)}} \mathbb{1}_{E_1 \cap E_2} + 179\sqrt{\log n}\big(\mathbb{P}\{\overline{E_1}\} + \mathbb{P}\{\overline{E_2}\}\big) + 2n\mathbb{P}\{\overline{E_3}\} \\
&\leq \sqrt{k-1} + 179 \sum_{j=k+1}^{\infty} \frac{en\sqrt{\log n}}{j^2} \left(\frac{e}{n^{1/(k-1)}}\right)^{j-1} \\
&\quad + \frac{179\sqrt{\log n}}{(1 - n^{-1/(k-1)})n^{3/(k-1)}} + 4 \\
&\leq \sqrt{k-1} + 4 + \frac{179 e^{k+1}}{(k+1)^2} \left(\frac{\sqrt{\log n}}{n^{1/(k-1)}}\right) \Big/ \left(1 - \frac{e}{n^{1/(k-1)}}\right) \\
&\quad + \frac{179\sqrt{\log n}}{(1 - n^{-1/(k-1)})n^{3/(k-1)}}.
\end{aligned}$$

The claim follows by observing that for $k \geq 2$, $\frac{\sqrt{\log n}}{n^{3/(k-1)}} \leq \frac{\sqrt{\log n}}{n^{1/(k-1)}} \leq c_k$, where c_k depends only on k.

Acknowledgements

G. Lugosi was supported by the Spanish Ministry of Economy and Competitiveness, Grant PGC2018-101643-B-I00; High-dimensional problems in structured probabilistic models - Ayudas Fundación BBVA a Equipos de Investigación Científica 2017; and Google Focused Award Algorithms and Learning for AI. S. Mendelson was supported in part by the Israel Science Foundation. N. Zhivotovskiy was supported by RSF Grant 18-11-00132.

This work was done while N. Zhivotovskiy was a postdoctoral fellow at the Department of Mathematics, Technion I.I.T. and researcher at National University Higher School of Economics.

References

[1] Alon, N., Krivelevich, M. and Vu, V.H. (2002). On the concentration of eigenvalues of random symmetric matrices. *Israel J. Math.* **131** 259–267. MR1942311 https://doi.org/10.1007/BF02785860

[2] Bandeira, A.S. and van Handel, R. (2016). Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *Ann. Probab.* **44** 2479–2506. MR3531673 https://doi.org/10.1214/15-AOP1025

[3] Benaych-Georges, F., Bordenave, C. and Knowles, A. (2017). Spectral radii of sparse random matrices. arXiv:1704.02945.

[4] Benaych-Georges, F., Bordenave, C. and Knowles, A. (2019). Largest eigenvalues of sparse inhomogeneous Erdős–Rényi graphs. *Ann. Probab.* **47** 1653–1676. MR3945756 https://doi.org/10.1214/18-AOP1293

[5] Berman, A. and Plemmons, R.J. (1994). *Nonnegative Matrices in the Mathematical Sciences*. *Classics in Applied Mathematics* **9**. Philadelphia, PA: SIAM. MR1298430 https://doi.org/10.1137/1.9781611971262

[6] Bollobás, B. (2001). *Random Graphs*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **73**. Cambridge: Cambridge Univ. Press. MR1864966 https://doi.org/10.1017/CBO9780511814068

[7] Boucheron, S., Bousquet, O., Lugosi, G. and Massart, P. (2005). Moment inequalities for functions of independent random variables. *Ann. Probab.* **33** 514–560. MR2123200 https://doi.org/10.1214/009117904000000856

[8] Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities*: *A Nonasymptotic Theory of Independence*. Oxford: Oxford Univ. Press. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001

[9] Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Stat.* **27** 642–669. MR0083864 https://doi.org/10.1214/aoms/1177728174

[10] Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2013). Spectral statistics of Erdős–Rényi graphs I: Local semicircle law. *Ann. Probab.* **41** 2279–2375. MR3098073 https://doi.org/10.1214/11-AOP734

[11] Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Magy. Tud. Akad. Mat. Kut. Intéz. Közl.* **5** 17–61. MR0125031

[12] Füredi, Z. and Komlós, J. (1981). The eigenvalues of random symmetric matrices. *Combinatorica* **1** 233–241. MR0637828 https://doi.org/10.1007/BF02579329

[13] Jung, P. and Lee, J. (2018). Delocalization and limiting spectral distribution of Erdős–Rényi graphs with constant expected degree. *Electron. Commun. Probab.* **23** Art. ID 92. MR3896830 https://doi.org/10.1214/18-ECP198

[14] Krivelevich, M. and Sudakov, B. (2003). The largest eigenvalue of sparse random graphs. *Combin. Probab. Comput.* **12** 61–72. MR1967486 https://doi.org/10.1017/S0963548302005424

[15] Lei, L. (2019). Unified $\ell_{2\to\infty}$ eigenspace perturbation theory for symmetric random matrices. arXiv:1909.04798.

[16] Massart, P. (1990). The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. *Ann. Probab.* **18** 1269–1283. MR1062069

[17] Mitra, P. (2009). Entrywise bounds for eigenvectors of random graphs. *Electron. J. Combin.* **16** Art. ID 131. MR2558268

[18] Tran, L.V., Vu, V.H. and Wang, K. (2013). Sparse random graphs: Eigenvalues and eigenvectors. *Random Structures Algorithms* **42** 110–134. MR2999215 https://doi.org/10.1002/rsa.20406

[19] Tropp, J.A. (2015). An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.* **8** 1–230. https://doi.org/10.1561/2200000048

[20] Vu, V.H. (2005). Spectral norm of random matrices. In *STOC'05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing* 423–430. New York: ACM. MR2181644 https://doi.org/10.1145/1060590.1060654

Received March 2019 and revised October 2019