# Entrywise Bounds for Eigenvectors of Random Graphs

Pradipta Mitra

280 Riverside Avenue
Rutherford, NJ 07070, USA

ppmitra@gmail.com

Submitted: May 23, 2009; Accepted: Oct 22, 2009; Published: Oct 31, 2009
Mathematics Subject Classification: 05C80

**Abstract**

Let $G$ be a graph randomly selected from $\mathbf{G}_{n,p}$, the space of Erdős-Rényi Random graphs with parameters $n$ and $p$, where $p \geqslant \frac{\log^6 n}{n}$. Also, let $A$ be the adjacency matrix of $G$, and $v_1$ be the first eigenvector of $A$. We provide two short proofs of the following statement: For all $i \in [n]$, for some constant $c > 0$

$$\left| v_1(i) - \frac{1}{\sqrt{n}} \right| \leqslant c \frac{1}{\sqrt{n}} \frac{\log n}{\log(np)} \sqrt{\frac{\log n}{np}}$$

with probability $1 - o(1)$. This gives nearly optimal bounds on the entrywise stability of the first eigenvector of (Erdős-Rényi) Random graphs. This question about entrywise bounds was motivated by a problem in unsupervised spectral clustering. We make some progress towards solving that problem.

## 1 Introduction

Spectral graph theory has been extensively used to study properties of graphs, and the results from this theory have found many applications in algorithmic graph theory as well. The study of spectral properties of Random graphs and matrices has been particularly fruitful. Starting from Wigner's celebrated semi-circle law [17], a number of results on Random matrices and Random graphs have been proved (See, for example, [10, 16]).

In this paper, we will deal with the well-known $\mathbf{G}_{n,p}$ model of Erdős-Rényi Random graphs. In this model, a random graph $G$ on $n$ vertices is generated by including each of the possible edges independently with probability $p$. For sake of brevity, in the remainder of the paper we will use the term **Random graph** to mean a graph thus generated. Spectral properties of such Random Graphs have been extensively studied. For example, the well known result by Furedi and Komlos (corrected and improved by Vu) [10, 16] implies that, for sufficiently large $p$, if $A$ is the adjacency matrix of a graph $G \in \mathbf{G}_{n,p}$, then with probability $1 - o(1)$

$$\|A - \mathbb{E}(A)\| \leqslant (2 + o(1))\sqrt{np}$$

Here and later $\|M\|$ denotes the spectral norm of a matrix $M$ and $\mathbb{E}(X)$ is the expectation of the random variable $X$ (in this case, a random matrix).

Instead of bounds on the spectral norm, in this paper we shall study the entrywise perturbation for eigenvectors of Random graphs, i.e. $\|v_1(A) - v_1(\mathbb{E}(A))\|_\infty$ (where $v_1(M)$ is the first eigenvector of a square symmetric matrix $M$). Perturbation of eigenvectors and eigenspaces have been classically studied for unitarily invariant norms, in particular for the spectral and Hilbert-Schmidt norms [3]. Perturbations in the $\|\cdot\|_\infty$ norm has been studied in the Markov Chain literature [14] to investigate stability of steady state distributions, however, the error model in those work do not seem to carry over to random graphs in any useful way.

The bound on $\|A - \mathbb{E}(A)\|$ can be converted to a statement about the relationship between $v_1(A)$ and $v_1(\mathbb{E}(A))$, i.e. one can show that $\|v_1(A) - v_1(\mathbb{E}(A))\|$ is small. This in turn does convert to a statement about $\|\cdot\|_\infty$ norm, but it is much weaker than our bounds. Taking random graphs from the $G_{n,\frac{1}{2}}$ model as an example, on the same scale spectral norm bounds only imply $O(\frac{1}{\sqrt{n}})$ entrywise differences, where as our results show that the differences are no larger than $O(\frac{\sqrt{\log n}}{n})$.

Recently, the *delocalization* property of eigenvectors of Wigner random matrices have been studied [15, 9] (and related papers referenced from both). These very general results imply entrywise *upper* bounds on *all* eigenvectors of $A - \mathbb{E}(A)$. Not quite a bound on the first eigenvector of $A$, the bounds are in addition weaker. One gets an upper bound of $O(\frac{\log^c n}{\sqrt{n}})$ on the absolute value on the entries of the eigenvectors, and no useful lower bound (which is not a weakness, one cannot expect such a bound for higher eigenvectors of random matrices). In addition, these works are not concerned with clustering problems on (generalized) Random graphs – something we explore, as described below.

We study the connection between entry-wise bounds for eigenvectors of Random graphs and the clustering problem on graphs generated by the Planted partition model (See [5, 13]), which is a generalization of the Random graph model. In this probabilistic model, the vertex set of the graph is partitioned into $k$ subsets $T_1, T_2, \ldots, T_k$. The input graph is random generated as follows: For two vertices $u \in T_j$, $v \in T_k$ the edge $(u, v)$ is independently chosen to be present with a probability $P_{jk} = P_{kj}$ and absent otherwise. So instead of a single probability $p$, the probability space is defined by a $k \times k$ matrix $P$. The adjacency matrix $A$ of the graph thus generated is presented as input. The task then is to identify the latent clusters $T_1, T_2, \ldots, T_k$ from $A$.

Generalizations of spectral norm bounds discussed above have been successfully used for analyzing spectral heuristics for the Planted partition model [5, 2, 7, 13]. The basic outline of many of these results is this: First one observes that $\mathbb{E}(A)$ is easy to cluster (by design). Since spectral norm bound imply that $\|A - \mathbb{E}(A)\|$ is small, the eigenvector structure of $A$ is not very dissimilar from that of $\mathbb{E}(A)$. This is then converted to a statement that most vertices of $A$ can be put in the correct cluster by looking at the eigenvectors of $A$. However, the small but non-negligible value of $\|A - \mathbb{E}(A)\|$ implies that some vertices might be misclassified. To rectify this, one uses some sort of "clean-up" scheme which for planted partition models invariably turn out to be combinatorial in nature. Experimental results suggest that such clean-up schemes are unnecessary (for large enough values of entries of $P$: for very small probabilities, "clean-up" schemes cannot be avoided). In [13], McSherry made a related conjecture. Proving such a result will most likely involve proving entrywise bounds for second and lower eigenvectors of adjacency matrix of Planted partition models. We make a step towards resolving these questions through computing entrywise bounds for the second eigenvector in a very simple Planted partition model. We will show that for a simple clustering problem, the second eigenvector obeys the cluster boundaries, thus no cleanup phase is necessary. Though our model requires conditions stronger than the ones used in standard results for spectral clustering, the results are non-trivial in the sense that mere eigenvalue bounds are not enough to prove them.

In Section 2 we present useful notation, the basic Random graph model and statement of the result for Random graphs. In Section 3 we present two proofs. Section 4 shows that out bound is tight for quasi-random graphs. In Section 5 we present the model and results for the Planted partition model.

## 2 Notation and Result

As stated in the introduction, the main object of study in this paper is the $\mathbf{G}_{n,p}$ model of Erdős-Rényi random graphs. $\mathbf{G}_{n,p}$ is a probability space on graphs with $n$ vertices. To get a random element $G$ from this space we select the edges independently, each of its $\binom{n}{2}$ possible edges are selected with probability $p$. Random graphs are widely studied objects [4]. In this paper, we will consider a slightly different model, where in addition to the edges between two different vertices, we also allow self loops, which are selected independently with the same probability $p$. This doesn't change the model or the result appreciably, but allows a cleaner exposition. We will call continue to call this modified model $\mathbf{G}_{n,p}$, and use the notation $G \in \mathbf{G}_{n,p}$ to denote that the graph $G$ is a random element of $\mathbf{G}_{n,p}$.

We will use $A(G)$ to denote the adjacency matrix of the graph $G$, and $A$ when $G$ is clear from the context. We will use $\lambda_i(M)$ and $v_i(M)$ to denote the $i^{th}$ largest (in absolute value) eigenvalue and its corresponding eigenvector of a square symmetric matrix $M$. Also let $\lambda = \lambda_1(A)$. If $A$ is the adjacency matrix of $G \in \mathbf{G}_{n,p}$, then note that $\mathbb{E}(A)$ is the $n \times n$ matrix where every entry is $p$.

For sets $R, S \in V$, $e(R, S)$ is the number of edges between $R$ and $S$. We will use the convenient notation $e(R)$ for $e(R, V)$ (where $V$ is the set of all vertices). For a vertex $v$, we will also use the shorthand $e(v) = e(\{v\})$. For any set of vertices $B$, $N(B)$ denotes the set of its neighbors.

Unless otherwise specified, vectors will be of dimension $n$. For two vectors $u$ and $v$, $(u \cdot v)$ denotes their inner product. The unsubscripted norm $\| \cdot \|$ will denote the usual euclidean norm for vectors and the spectral norm for matrices. For a matrix $M$, $M_{ij}$ denotes the entry in the $i^{th}$ row and the $j^{th}$ column. For a vector $x \in \boldsymbol{R}^n$, let $x(i)$ be its $i^{th}$ entry. Let $x_{\max} = \max_{i\in[n]} x(i)$, $x_{\min} = \min_{i\in[n]} x(i)$ and $\|x\|_\infty = \max_{i\in[n]} |x(i)|$. We will use the symbol $\mathbf{1}$ to mean a vector with all entries equal to 1. Also, $c, c_1, c_2 \ldots$ etc are constants throughout the paper. We will use the phrase "with high probability" to mean with probability $1 - o(1)$.

The following is the main result in our paper. Define $\Delta = 4\sqrt{\frac{\log n}{np}}$.

**Theorem 1.** *Let $G \in \mathbf{G}_{n,p}$ be a random graph and $A$ be its adjacency matrix. Assume $p \geqslant \log^6 n/n$. Then, for all $i \in [n]$*

$$\left| v_1(i) - \frac{1}{\sqrt{n}} \right| \leqslant c \frac{\log n}{\log np} \frac{1}{\sqrt{n}} \Delta$$

*with high probability, for some constant $c$.*

# 3 Proofs

## 3.1 The First Proof

In this section we present the first proof of Theorem 1. We will need the following result about random graphs [10, 16].

**Theorem 2.** *Let $A$ be the adjacency matrix of $G \in \mathbf{G}_{n,p}$, where $p \geqslant \frac{\log^6 n}{n}$. Then with high probability,*

$$\|A - \mathbb{E}(A)\| \leqslant 3\sqrt{np}$$

We will also need the following basic results:

**Lemma 3.** *With probability $1 - \frac{1}{n^2}$, for all vertices $v$ of $G \in \mathbf{G}_{n,p}$*

$$|e(v) - np| \leqslant np\Delta$$

*Proof.* Elementary use of the Chernoff bound. ☐

**Lemma 4.** *Let $G$ be a connected graph on $n$ vertices such that $|e(v) - np| \leqslant np\Delta$. Then $\lambda \geqslant np(1 - \Delta)$ and $\lambda \leqslant np(1 + \Delta)$*

*Proof.* For the lower bound, it suffices to observe that $\|A\mathbf{1}\| \geqslant np(1 - \Delta)\|\mathbf{1}\|$.

Now let $v = v_1(A)$. Assume without loss of generality $v(1) = \max_i v(i)$. Then by definition

$$\begin{aligned} & \lambda v(1) = (Av)(1) = \sum_{j:A_{1j}=1} v(j) \leqslant v(1) \sum_{j:A_{1j}=1} 1 \leqslant v(1)np(1+\Delta) \\ \Rightarrow \quad & \lambda \leqslant np(1+\Delta) \end{aligned}$$

That proves the upper bound. ☐

The following is well known [10]:

**Lemma 5.** *Let $G \in \mathbf{G}_{n,p}$ where $p \geqslant \frac{2\log n}{n}$. Then with high probability, $G$ is connected.*

Let us adopt the notation $u = [a \pm b]_v$ for $b \geqslant 0$ to mean that $u$ is a vector such that $a - b \leqslant u(i) \leqslant a + b$ for all $i$.

**Lemma 6.** *Let $u = [1 \pm 3t\Delta]_v$ for some $\log n \geqslant t \geqslant 0$. Define $u' = \frac{1}{\lambda}Au$. Then*

$$u' = [1 \pm 3(t+1)\Delta]_v \tag{1}$$

*Proof.* For any $i \in [n]$

$$\lambda u'(i) = \sum_{j\in[n]} A_{ij}u(j) = \sum_{j\in N(i)} u(j)$$

We know that $|N(i)| \leqslant d(1+\Delta)$ and $u(j) \leqslant 1 + 3t\Delta$. Hence,

$$\begin{aligned} \lambda u'(i) \quad & \leqslant d(1+\Delta)(1+3t\Delta) \\ & \leqslant d(1+\Delta+3t\Delta+3t\Delta^2) \leqslant d(1+\Delta+3t\Delta+o(\Delta)) \\ \Rightarrow \quad & u'(i) \leqslant \frac{d}{\lambda}(1+\Delta+3t\Delta+o(\Delta)) \end{aligned}$$

The assertion $3t\Delta^2 = o(\Delta)$ follows from the assumed bounds on $t$ and $p$. As $\lambda \geqslant d(1-\Delta)$

$$u'(i) \quad \leqslant \quad \frac{1}{1-\Delta}(1+\Delta+3t\Delta+o(\Delta)) \leqslant 1+2\Delta+3t\Delta+o(\Delta) \leqslant 1+3(t+1)\Delta$$

The lower bound is similar. □

**Lemma 7.** *Let $\mathbf{f} \equiv \frac{1}{\sqrt{n}}\mathbf{1} = \alpha v_1 + \beta v_\perp$ where $v_\perp \perp v_1$ and $\|v_\perp\| = 1$. Then, $\alpha \geqslant (1-2\Delta)$*

*Proof.* By definition, $(\mathbf{f} \cdot v_1) = \alpha$.

We claim that $\alpha > 0$. A version of the Perron-Frobenius Theorem [11] implies that the adjacency matrix of a connected graph will have a eigenvector corresponding to its largest eigenvalue that has non-negative entries. We already know that $G$ is connected (Lemma 5). Now by Theorem 2 and Lemma 4, it is clear that the $\lambda_1(A)$ has multiplicity 1. Hence $v_1$ is non-negative (but of course not all zero). Clearly, $\alpha = (\mathbf{f} \cdot v_1) > 0$, which was the claim.

We know (Theorem 2),

$$\begin{aligned} & \|A - \mathbb{E}(A)\| \leqslant 3\sqrt{np} \\ \Rightarrow \quad & \|\lambda v_1 v_1^T + \sum_{i\geqslant 2} \lambda_i v_i v_i^T - np\mathbf{f}\mathbf{f}^T\| \leqslant 3\sqrt{np} \end{aligned} \tag{2}$$

Now

$$\begin{aligned} & (\lambda v_1 v_1^T + \sum_{i\geqslant 2} \lambda_i v_i v_i^T - np \times \mathbf{f}\mathbf{f}^T)v_1 \\ = \quad & \lambda v_1 - np \times \mathbf{f}(\mathbf{f}\cdot v_1) = \lambda v_1 - \alpha np \times \mathbf{f} \\ = \quad & \lambda v_1 - \alpha np(\alpha v_1 + \beta v_\perp) = (\lambda - \alpha^2 np)v_1 + \alpha\beta np v_\perp \end{aligned}$$

Hence $\|(\lambda v_1 v_1^t + \sum_{i\geqslant 2} \lambda_i v_i v_i^T - np \times \mathbf{ff}^T) v_1\|^2 = (\lambda - \alpha^2 np)^2 + (\alpha\beta np)^2$

Comparing this with Equation (2), we get

$$\begin{aligned} & (\lambda - \alpha^2 np)^2 \leqslant 9np \\ \Rightarrow \quad & \alpha^2 np \geqslant \lambda - 3\sqrt{np} \\ \Rightarrow \quad & \alpha \geqslant \alpha^2 \geqslant \frac{1}{np} np(1-2\Delta) = 1 - 2\Delta \end{aligned}$$

Where $\alpha \geqslant \alpha^2$ follows from $1 \geqslant \alpha > 0$. This proves the Lemma. □

Now we can prove Theorem 1:

*Proof.* Let $l = 9\frac{\log n}{\log np}$. Also let

$$u_t = \frac{1}{\lambda^t} A^t \mathbf{1} \tag{3}$$

for $t \geqslant 0$. Note that $A^0 = I$, the identity matrix. By Lemma 7, we know that $\frac{1}{\sqrt{n}}\mathbf{1} = \alpha v_1 + \beta v_\perp$ where $v_\perp \perp v_1$, $\|v_\perp\| = 1$ and $\alpha \geqslant (1 - 2\Delta)$

By Lemma 6

$$u_l = [1 \pm 3l\Delta]_v \tag{4}$$

Let $v_\perp = \sum_{i\geqslant 2} \gamma_i v_i$. Then

$$\begin{aligned} & \frac{1}{\sqrt{n}} A^l \mathbf{1} = \alpha\lambda^l v_1 + \beta \sum_{i\geqslant 2} \gamma_i \lambda_i^l v_i \\ \Rightarrow \quad & \frac{1}{\sqrt{n}} u_l = \alpha v_1 + x_\epsilon \end{aligned} \tag{5}$$

where $x_\epsilon = \beta \sum_{i\geqslant 2} \gamma_i (\frac{\lambda_i}{\lambda})^l v_i$

Now as $\lambda \geqslant np(1-\Delta)$ and $\lambda_{i\geqslant 2} \leqslant 3\sqrt{np}$, $\left(\frac{\lambda_i}{\lambda}\right)^l \leqslant \left(\frac{4}{\sqrt{np}}\right)^l \leqslant \frac{1}{n^4}$

We can compute a bound on each entry of $x_\epsilon$

$$\|x_\epsilon\|_\infty \leqslant \|x_\epsilon\| \leqslant \frac{1}{n^4}\beta \left\|\sum_{i\geqslant 2} \gamma_i v_i\right\| \leqslant \frac{1}{n^4}\beta\|v_\perp\| \leqslant \frac{1}{n^4}$$

Hence, from Equation (4)—(5)

$$\begin{aligned} \alpha v_1 \quad & = \frac{1}{\sqrt{n}} u_l + \left[\pm\frac{1}{n^4}\right]_v = \frac{1}{\sqrt{n}}[1 \pm 3(l+1)\Delta]_v + \left[\pm\frac{1}{n^4}\right]_v = \frac{1}{\sqrt{n}}[1 \pm 4(l+1)\Delta]_v \\ \Rightarrow v_1 \quad & = \frac{1}{\alpha}\frac{1}{\sqrt{n}}[1 \pm 4(l+1)\Delta]_v = \frac{1}{\sqrt{n}}[1 \pm 6(l+1)\Delta]_v \end{aligned}$$

The last line uses the bound $\alpha \geqslant 1 - 2\Delta$. This completes the proof. □

## 3.2 The Second Proof

This proof is slightly longer, but is more elementary (we don't need to use Theorem 2), and perhaps more intuitive. In addition, the proof technique employed here will be used in the next section on spectral clustering, so it is worth introducing.

We will actually prove a theorem on **Quasi-random** graphs [12]: A graph $G(V, E)$ is $(p, \alpha)$-Quasi-random $(p > 0, \alpha > 0)$ if, for all subsets $R, T \in V$

$$|e(R,T) - prt| \leqslant \alpha\sqrt{rt}$$

where $n = |V|$, $r = |R|$ and $t = |T|$.

We will prove the following Theorem

**Theorem 8.** *Assume $G$ is a connected $(p, 2\sqrt{np})$-Quasi-random graph on $n$ vertices. Let $A$ be the adjacency matrix of $G$. Also assume that $|e(v) - np| \leqslant np\Delta$ (We have already defined $\Delta = 4\sqrt{\frac{\log n}{np}}$). Let $v = \gamma v_1(A)$ where $\gamma$ is chosen such that such that $v_{\max} = 1$. Then*

$$v_{\min} \geqslant 1 - c_2 \frac{\log n}{\log np} \Delta \tag{6}$$

*for some constant $c_2$.*

The following corollary of Theorem 8 implies Theorem 1

**Corollary 9.** *Assume $G \in G_{n,p}$ where $p \geqslant \log^6 n/n$. Let $A$ be the adjacency matrix of $G$. Let $v = \gamma v_1(A)$ where $\gamma$ is chosen such that such that $v_{\max} = 1$. Then*

$$v_{\min} \geqslant 1 - c_2 \frac{\log n}{\log np} \Delta \tag{7}$$

*for some constant $c_2$.*

*Proof.* For $p \geqslant \log^6 n/n$, $G \in G_{n,p}$ is $(p, 2\sqrt{np})$-Quasi-random with high probability. This property can be quickly proven by applying the standard Chernoff bound a few times (See the survey by Krivelevich and Sudakov [12] for a reference). Lemmas 3 and 5 imply that the other assumptions needed for Lemma 8 are satisfied. This completes the proof. $\square$

The intuition behind this proof of Theorem 8 can be demonstrated by the following simple observation. Let $v$ is normalized such that $v(1) = 1$ for vertex 1. Now, the first eigenvalue of $A$ is close to $np$, while vertex 1 has a degree of $np(1 \pm \Delta)$. As $(A \cdot v)(1) \approx npv(1) \approx np$, we need $\sum_{j \in N(1)} v(j) \approx np$ where $N(1)$ is 1's neighborhood set. This means *on average*, $N(1)$ will have weights in the range $1 \pm \Delta$. Our technical lemmas that follow show how this intuition can be shown to be true for not only vertices but sets, and how an absolute (not only average) result can be achieved.

Assume, $v$ is defined as in Theorem 8, and $v(1) = v_{\max} = 1$, without loss of generality. We define a sequence of sets $\{S_t\}$ for $t = 1 \ldots$ in the following way:

$$S_1 = \{1\} \tag{8}$$
$$S_{t+1} = \{i : i \in N(S_t) \text{ and } v(i) \geqslant 1 - c(t+1)\Delta\}, \forall t > 1 \tag{9}$$

Now, we define $n_t$ and $F_t$

- $n_t = |S(t)|$
- $F_t = \sum_{i \in S(t)} v(i)$

Note that $n_1 = 1$ and $F_1 = 1$.

**Lemma 10.** *Let $t'$ be the last index such that $n_{t'} \leqslant \frac{60}{p}$. For all $t \leqslant t'$*

$$n_{t+1} \geqslant \frac{np \times n_t}{72 \log^2 n}$$

*Proof.* Let $N = N(S(t))$. Note that $e(S(t)) = e(S(t), N) \leqslant n_t np(1 + \Delta)$.

The edges from $S(t)$ to its neighbors must provide the multiplicative factor of $\lambda$:

$$\lambda F_t = \sum_{i \in N} v(i) e(i, S(t))$$

Now,

$$\begin{aligned}
\lambda F_t &= \sum_{i \in N} v(i) e(i, S(t)) = \sum_{N - S(t+1)} v(i) e(i, S(t)) + \sum_{S(t+1)} v(i) e(i, S(t)) \\
&\leqslant (1 - c(t+1)\Delta) e(N - S(t+1), S(t)) + e(S(t+1), S(t)) \\
&= (1 - c(t+1)\Delta) e(S(t)) + c(t+1) e(S(t), S(t+1))\Delta \\
&\leqslant (1 - c(t+1)\Delta) e(S(t)) + c(t+1) \left(p n_t n_{t+1} + 2\sqrt{n_t n_{t+1} np}\right) \Delta
\end{aligned}$$

The last line uses the quasi-randomness property. Since $\lambda \geqslant np(1 - \Delta)$ (Lemma 4) and $F_t \geqslant n_t(1 - ct\Delta)$ (by definition)

$$\begin{aligned}
& (1 - c(t+1)\Delta)\, e(S(t)) + c(t+1) \left(p n_t n_{t+1} + 2\sqrt{n_t n_{t+1} np}\right) \Delta \geqslant np(1 - \Delta) F_t \\
\Rightarrow \quad & (1 - c(t+1)\Delta)\, n_t np(1 + \Delta) + c(t+1) \left(p n_t n_{t+1} + 2\sqrt{n_t n_{t+1} np}\right) \Delta \\
& \geqslant (1 - ct\Delta) np(1 - \Delta) n_t \\
\Rightarrow \quad & p n_t n_{t+1} + 2\sqrt{n_t n_{t+1} np} \geqslant \left(1 - \frac{2}{c}\right) \frac{n_t np}{t+1}
\end{aligned}$$

As $n_t \leqslant \frac{60}{p}$ by assumption,

$$60 n_{t+1} + 2\sqrt{n_t n_{t+1} np} \geqslant \left(1 - \frac{2}{c}\right) \frac{n_t np}{t+1}$$

Assuming $c \geqslant 10$, either (or both) of the following is true:

$$60n_{t+1} \geqslant \frac{1}{3}\frac{n_t np}{t+1} \tag{10}$$

$$2\sqrt{n_t n_{t+1} np} \geqslant \frac{1}{3}\frac{n_t np}{t+1} \tag{11}$$

From which we get

$$n_{t+1} \geqslant n_t np \max\left(\frac{1}{36(t+1)^2}, \frac{1}{180(t+1)}\right) \tag{12}$$

Hence as long as $t \leqslant \log n$

$$n_{t+1} \geqslant \frac{1}{72}\frac{n_t np}{\log^2 n} \tag{13}$$

All that remains to show is that $t' \leqslant \log n$.

For this, observe that with the growth rate specified in (13), $n_t$ to be at least as large as $\frac{60}{p}$, $t$ need not be larger than $\log_{(np)^{3/4}} \frac{1}{p} = \frac{4}{3}\frac{\log \frac{1}{p}}{\log np} \leqslant \log n$. Hence $t' \leqslant \log n$. □

The following lemma deals with the case of large sets.

**Lemma 11.** *Let $U$ be a set of vertices, where $u = |U| \geqslant \frac{60}{p}$. Also, assume that $F = \sum_{i \in U} v(i) \geqslant u(1 - \alpha\Delta)$ for some $\alpha \geqslant 1$. Let $W(U) = \{i : i \in N(U)$ and $v(i) \geqslant (1 - (12\alpha + 24)\Delta)\}$. Then $w = |W(U)| > \frac{6n}{10}$.*

*Proof.* Assuming that the claim of the lemma is false, $w \leqslant \frac{6n}{10}$.

We know that $\lambda F = \sum_{i \in N(U)} v(i)e(i,U)$. By the lower bound on $F$, we need

$$\sum_{i \in N(U)} v(i)e(i,U) \geqslant \lambda u(1 - \alpha\Delta) \geqslant npu(1 - (\alpha + 1)\Delta) \tag{14}$$

However, using Quasi-randomness and the fact $u \geqslant \frac{60}{p}$

$$e(W(U), U) \leqslant \frac{6npu}{10} + 2\sqrt{\frac{u6n^2p}{10}} \leqslant \frac{4npu}{5}$$

As $v(i) \leqslant 1$ for all $i$ and $v(i) \leqslant 1 - (12\alpha + 24)\Delta$ for $i \in N(U) - W(U)$, this yields,

$$\begin{aligned}
& \sum_{i \in N(U)} v(i)e(i,U) \\
\leqslant\ & e(W(U), U) \times 1 \\
& +(e(U, N(U)) - W(U)) \times (1 - (12\alpha + 24)\Delta) \\
\leqslant\ & \frac{4npu}{5} + \left(u(np(1+\Delta)) - \frac{4unp}{5}\right)(1 - (12\alpha + 24)\Delta) \\
\leqslant\ & npu(1+\Delta) - \frac{npu}{6}(12\alpha + 24)\Delta \\
\leqslant\ & npu(1 - (2\alpha + 2)\Delta)
\end{aligned}$$

This contradicts Equation 14. □

The following two lemmas follow along the same lines as Lemmas 10 and 11, respectively. For these lemmas, we assume $v(1) = v_{\min} = b > 0$ and define $S_t$ and $n_t$ analogously:

- $S_1 = \{1\}$
- $S_{t+1} = \{i : i \in N(S_t) \text{ and } v(i) \leqslant b(1 + c(t+1)\Delta)\}$

And,

- $n_t = |S(t)|$
- $F_t = \sum_{i \in S(t)} v(i)$

Note that $n_1 = 1$ and $F_1 = b$.

**Lemma 12.** *Let $t'$ be the last index such that $n_{t'} \leqslant \frac{60}{p}$. For all $t \leqslant t' + 1$*

$$n_{t+1} \geqslant \frac{n_t np}{9 \log^2 n}$$

**Lemma 13.** *Let $U \subset V$, where $u = |U| \geqslant \frac{60}{p}$. Also, assume that $F(U) = \sum_{i \in U} v(i) \leqslant ub(1 + \alpha\Delta)$ for some $\alpha \geqslant 1$. Let $W(U) = \{i : i \in N(U) \wedge v(i) \leqslant b(1 + (12\alpha + 24)\Delta)\}$. Then $w = |W(U)| > \frac{6n}{10}$.*

### Proof of Theorem 8

*Proof.* Let us consider $S_t$ (as defined in Eqns 8 and 9) for the first $t$ such that $n_t \geqslant \frac{60}{p}$. From Lemma 10, $F_t \geqslant n_t(1 - c\frac{\log n}{\log n}\Delta)$. As $n_t \geqslant \frac{60}{p}$, we can invoke Lemma 11 with $U = S_t$. This gives us a a set $W$ with $|W| > \frac{n}{2}$, such that for every $i \in W, v(i) \geqslant 1 - \beta\Delta$, where $\beta = c_1 \frac{\log n}{\log np}$ for some constant $c_1$.

A similar argument can be put forward using Lemmas 12 and 13. So, for another set $Y$, where $|Y| > \frac{n}{2}$, $v(i) \leqslant b(1 + \beta\Delta)$ for each $i \in Y$. Using the pigeonhole principle to observe that $X$ and $Y$ must intersect, we can conclude that

$$\begin{aligned} & b(1 + \beta\Delta) \geqslant 1 - \beta\Delta \\ \Rightarrow \quad & b \geqslant 1 - 3\beta\Delta \end{aligned}$$

This completes the proof of Theorem 8. □

# 4 Tightness

For general Quasi-random graphs, we will show that our bound is tight up to a constant factor.

We prove the following:

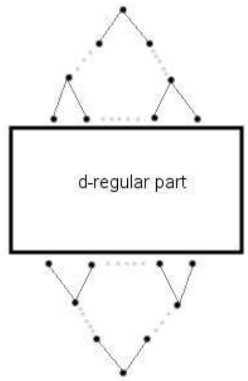


Figure 1: Construction of a Quasi-random graph depicting tightness of the bound.

**Theorem 14.** *For any large enough* $n$*, and any* $\sqrt{n} \geqslant d \geqslant \log^6 n$ *there exists a* $(\frac{d}{n}, 27\sqrt{d})$*-quasirandom graph on* $n$ *nodes so that each vertex has a degree in the range* $d(1 \pm 2\sqrt{\frac{\log n}{d}})$ *and*

$$v_{\max} - v_{\min} \geqslant \frac{\log n}{\log d}\sqrt{\frac{\log n}{d}} \tag{15}$$

*where* $v$ *is the largest eigenvector of the adjacency matrix* $A$ *of the graph.*

*Proof.* Given large enough $n$ and $d$, define $l = \lceil \frac{\log n}{10 \log d} \rceil$ and $\epsilon = \sqrt{\frac{\log n}{d}}$. We construct the Quasi-random graph as follows:

**Construction** Starting from a single node as root, construct a $d(1+\epsilon)$ degree complete tree $T_1$ of depth $l$. Construct another complete tree $T_2$ of same depth, but with degree $d(1-\epsilon)$. Define $L(T)$ to be the set of leaves of a tree $T$. Note that $V(T_1) = d(1+\epsilon)^l = O(n^{1/5})$ and $V(T_2) = d(1-\epsilon)^l = O(n^{1/5})$.

Now let $M$ be a set of $m = n - |V(T_1)| - |V(T_2)|$ new nodes. Set $Q = L(T_1) \cup L(T_2) \cup M$ and construct a $d$-regular expander on $Q$ (by, for example, generating a random $d$-regular graph on them). Expanders are Quasi-random [12], in particular, the subgraph constructed on $q = |Q|$ vertices is $(\frac{d}{q}, 2\sqrt{d})$-Quasi-random.

Now let $G$ be the graph on vertex set $V = M \cup T_1 \cup T_2$ of size $n$ and containing all edges in the two trees and the expander on $Q$.

We claim:

1. $G$ is $(\frac{d}{n}, 27\sqrt{d})$-quasirandom and has vertex degree in the range $d(1 \pm 2\epsilon)$.

2. Let $v = \gamma v_1(A)$ such that $v_{\max} = 1$. Then, $v_{\min} \leqslant 1 - \frac{1}{2} l \frac{\epsilon}{1+\epsilon}$

Let us prove the first claim. The claim about degrees is clear from the construction. Proving the quasi-randomness property is a matter of checking the property for each possible pair of vertex sets. We do a case by case analysis below and then will finally combine the cases to come up with a unified bound.

**Case 1:** Let $R, S \in Q$ and define $q = |Q|$. As stated above, the subgraph on $Q$ is quasirandom, hence

$$
\begin{aligned}
& |e(R,S) - \frac{d|R||S|}{q}| \leqslant 2\sqrt{d|R||S|} \\
\Rightarrow \quad & |e(R,S) - \frac{d|R||S|}{n}| \leqslant 2\sqrt{d|R||S|} + \frac{d|R||S|(n-q)}{nq} \\
& \leqslant 3\sqrt{d|R||S|}
\end{aligned}
$$

The last inequality requires $\frac{d|R||S|(n-q)}{nq} \leqslant \sqrt{d|R||S|}$ which follows easily from $n - q = O(n^{1/5})$.

**Case 2:** Now let $X, Y \in T_1$, define $x = |X|, y = |Y|$ and assume without loss of generality that $x \leqslant y$.

We claim, $e(X,Y) \geqslant \frac{dxy}{n} - 2\sqrt{dxy}$. As, $x, y \leqslant O(n^{1/5})$, $\frac{dxy}{n} < 2\sqrt{dxy}$. So, the bound is trivially true.

Next, we claim $e(X,Y) \leqslant \frac{dxy}{n} + 2\sqrt{dxy}$. We analyze two cases:

- if $x < \frac{y}{d}$. In this case

$$
\begin{aligned}
e(X,Y) &\leqslant x(d(1+\epsilon)+1) \leqslant 2\sqrt{dx}\sqrt{dx} \\
&\leqslant 2\sqrt{dx}\sqrt{y} \leqslant 2\sqrt{dxy}
\end{aligned}
$$

  We use the assumption on $x$ in the last inequality.

- if $x \geqslant \frac{y}{d}$. As $T_1$ is a tree, $e(X,Y) \leqslant x + y \leqslant 2y$ Now,

$$
\begin{aligned}
e(X,Y) &\leqslant 2y = 2\sqrt{y}\sqrt{y} \\
&\leqslant 2\sqrt{y}\sqrt{dx} \leqslant 2\sqrt{dxy}
\end{aligned}
$$

**Case 3:** Let $X \in Q$ and $Y \in T_1$. First, we claim $e(X,Y) \geqslant \frac{dxy}{n} - 2\sqrt{dxy}$. As $d \leqslant \sqrt{n}$ and $y = O(n^{1/5})$, $\frac{dxy}{n} < 2\sqrt{dxy}$. So, the bound is trivially true.

For the other case, note that the only edges from $T_1$ to $Q$ will involve $L(T_1)$, hence the same arguments as in **Case 1** will suffice to prove the claim.

Similar bounds will work for sets involving $T_2$.

To prove the bound for any two sets $S, R \subset V$, assume that for any set $W$, $W_1 = W \cap T_1$, $W_2 = W \cap T_2$ and $W_3 = W \cap Q$. Now,

$$
\begin{aligned}
\left|e(S,T) - \frac{dsr}{n}\right| &\leqslant \sum_{i,j \in \{1,2,3\}} \left|e(S_i, R_j) - \frac{d}{n} s_i r_j\right| \\
\leqslant \sum_{i,j \in \{1,2,3\}} 3\sqrt{ds_i r_j} \leqslant 27\sqrt{dsr}
\end{aligned}
$$

Hence the bound.

Now we prove the second part of the claim. First consider the case where $\lambda_1 \leqslant d$. For this case, assume that the root of $T_1$ is vertex 1. Let, $u = \gamma_1 v_1(A)$ such that $u(1) = 1$. Note that $|\gamma_1| \geqslant |\gamma|$. By Lemma 15 (which we state and prove later), at level $l$ of $T_1$, there is a vertex $j$ for which

$$u(j) \leqslant 1 - \frac{1}{2} l \frac{\epsilon}{1+\epsilon}$$

Since $|\gamma_1| \geqslant |\gamma|$, we get $v_{\min} \leqslant u(j)$. This proves the claim.

Now, if $\lambda_1 \geqslant d$, we can use a similar argument on $T_2$ and prove that if $v_x = 1$ (where $x$ is the root of $T_2$) then there exists a vertex $j$ at level $l$ of $T_2$ such that $v_j \geqslant 1 + \frac{1}{2} l \frac{\epsilon}{1+\epsilon}$. This proves the claim, from which the Theorem 14 follows once we plug in values of $\epsilon$ and $l$. $\square$

**Lemma 15.** *Consider a graph constructed as in the proof of Theorem 14. Assume that $\lambda_1 \leqslant d$ and that the root of $T_1$ is vertex 1. Let, $u = \gamma_1 v_1(A)$ such that $u(1) = 1$. Then, for all $r \leqslant l$ ($l$ is defined in the proof of Theorem 14), there is a vertex $j$ at level $r$ of $T_1$ for which*

$$u(j) \leqslant 1 - \frac{1}{2} r \frac{\epsilon}{1+\epsilon}$$

.

*Proof.* We prove the bound inductively. The claim is trivially true at level $r = 0$. Assume the hypotheis is true at level $r < l$, and that $u(j) \leqslant 1 - \frac{1}{2} r \frac{\epsilon}{1+\epsilon}$ for some vertex $j$ at level $r$. Assume $p$ is $j$ parent in $T_1$, if it has one. Now,

$$\sum_{i \in N(j)} u(i) \leqslant d u_j$$
$$\Rightarrow \sum_{i \in N(j) - \{p\}} u(i) \leqslant (d-1) u_j$$

Since $|N(j) - \{p\}| = d(1+\epsilon)$, by a basic averaging argument, for some $k \in N(i)$ ,

$$\begin{aligned} u(k) \quad &\leqslant \quad \frac{(d-1)u(j)}{d(1+\epsilon)} \leqslant (1 - \frac{1}{2} r \frac{\epsilon}{1+\epsilon})(1 - \frac{1}{d})(1 - \epsilon + \epsilon^2) \\ &\leqslant (1 - \frac{1}{2} r \frac{\epsilon}{1+\epsilon})(1-\epsilon) \leqslant 1 - \frac{1}{2} r \frac{\epsilon}{1+\epsilon} - \epsilon + \frac{r\epsilon^2}{1+\epsilon} \leqslant 1 - \frac{1}{2}(r+1) \frac{\epsilon}{1+\epsilon} \end{aligned}$$

Provided $\epsilon \geqslant r\epsilon^2$ which is true for our construction. $\square$

# 5 Application to clustering

In this section we present a result on spectral clustering using our approach. We will show that a very simple algorithm (Algorithm 1) manages to bi-partition a graph randomly generated from a planted partition model with two clusters. Our model will be a particularly simple instance of the Planted partition model, and the conditions we assume are much stronger than those needed for standard spectral clustering algorithms [13]. The interest, thus, lies in the simplicity of the algorithm, not the generality of the result.

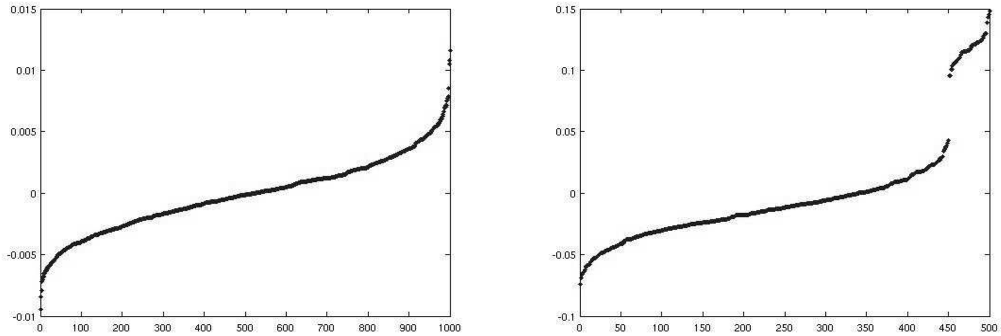


Figure 2: a) A sorted plot of $v_1(A) - v_1(\mathbb{E}A)$ where $A$ was computer generated according to $G_{1000,0.1}$. b) A sorted plot of the second eigenvector of the planted clique problem, where a clique of size 50 is embedded in a 500 node graph. The graph is generated by selecting every edge with $p = \frac{1}{2}$. The largest 50 entries correspond to the clique.

---

**Algorithm 1** Threshold$(A, n)$

---

$\{A$ is the adjacency matrix of the input graph $G$, $n$ is the number of vertices$\}$
Find $v = v_2(A)$, the second eigenvector of $A$
Let $L = \{i : v(i) < 0\}$
Return $L$ and $[n] - L$

---

## 5.1 Model

The input to the algorithm is a graph $G(V, E)$, which has two clusters $T_1$ and $T_2$ such that $T_1 \cup T_2 = V$. Assume $|T_a| = n$ for $a = 1, 2$. The adjacency matrix $A$ is generated thus: For $a, b \in \{1, 2\}$ there are probabilities $p_{ab}(= p_{ba})$ such that if $r \in T_a$ and $s \in T_b$ then $A_{sr} = A_{rs} = 1$ with probability $p_{ab}$ and 0 otherwise.

Assume

$$p_{aa} = \frac{1}{\sqrt{n}} \tag{16}$$

$$p_{ab} \leqslant \frac{1}{\sqrt{n}} - c_1\sqrt{\frac{\log n}{n^{7/6}}} \tag{17}$$

For some large enough constant $c_1$. Let $d_1$ be the expected number of edges of a vertex to vertices in its own cluster and $d_2$ be the expected number of edges from a vertex to vertices in the other cluster. From 16 and 17, $d_1 = \sqrt{n}$ and $d_2 \leqslant \sqrt{n} - c_1\sqrt{n^{5/6}\log n}$, which implies the following **separation condition**:

$$d_1 - d_2 \geqslant c_1\sqrt{n^{5/6}\log n} \tag{18}$$

We prove:

**Theorem 16.** *With high probability, Given a graph $G$ generated from the Planted partition model described above, Algorithm 1 outputs a bi-partitioning of the graph that agrees with the underlying clusters $T_1$ and $T_2$.*

## 5.2 Related work

Clustering problems on probabilistic models have a long history (see references in [13]). Algorithms based on the spectrum of graphs have been considered for both discrete and continuous models by a number of papers (e.g. [1, 13]). A major part in all these papers involves dealing with the "error" $\|A - \mathbb{E}A\|$. Some sort of clean-up method is employed in all work of this kind. These clean-up techniques are usually simple for continuous distributions, but often quite complicated for discrete distributions. In [13], for example, it was shown that a spectral projection based algorithm followed by a combinatorial cross-training succeeded in clustering the graph. The model presented in that paper is quite general, and works under essentially tight separation conditions. For the simplified model presented here, standard algorithms ([13, 6]) successfully partitions the graph as long as

$$d_1 - d_2 \geqslant c\sqrt{n^{1/2}\log n}$$

where $c$ is some constant. Comparing this condition with condition 18 reveals that the latter is a much stronger assumption, as we have mentioned before.

Experiments with Random graphs as well (for example, Fig 2(b)) indicate that these clean-up techniques might be unnecessary for large enough values of edge probabilities (for very small values, they are unavoidable). We show that this is indeed the case for our (simpler) model. Apart from simplicity of the algorithm involved, we believe the question of how the spectral error is distributed is important in extending spectral methods to more complex models.

## 5.3 Proof

The following two Lemmas can be proven using standard techniques in Spectral Clustering literature [2, 6]

**Lemma 17.** $\lambda_2 = \lambda_2(A) \geqslant 0.99(d_1 - d_2)$.

**Lemma 18.** *Let $v = v_2(A)$. Then $v = u + w$ such that*

$$u(i) = sign(a)\frac{1}{\sqrt{2n}} \qquad \forall i \in T_a$$

*and $\|w\| \leqslant \frac{5}{c_1 n^{1/6}\sqrt{\log n}}$, where $sign(1) = 1$ and $sign(2) = -1$.*

Note at this point that the value of $\|w\|$ is enough induce errors (in fact many of them) of the order $\sqrt{\frac{2}{n}}$ in $v$, which is all that is necessary to cause Algorithm 1 to fail, and it is at this point that clean-up phases are necessary. We will show that this doesn't happen. Our idea is to use a analysis of neighborhood sets of a vertex $s$ to show that $\|w\|$ cannot be distributed in an arbitrary fashion.

We will need the following proposition, easily proved from the relation between $l_1$ and $l_2$ norms.

**Proposition 19.** *Consider any subset $S \subset [2n]$. Then*

$$\sum_S |w(i)| \leqslant \sqrt{|S|}\|w\|$$

Here is the main lemma that implies Theorem 16 directly:

**Lemma 20.** *For all $s \in T_1, v(s) > \frac{4}{5\sqrt{2n}}$ and for all $s \in T_2, v(s) < -\frac{4}{5\sqrt{2n}}$.*

*Proof.* The claims for $T_1$ and $T_2$ are symmetric, hence we will only prove the first claim. Let $s \in T_1$. We use the following notation $N_a = \{T_a \cap N(s)\}$ and $N_{ab} = \{T_a \cap N(N_b(s))\}$. Assume $e_i(S)$ is the number of neighbors node $i$ has in set $S$. Then,

$$\begin{aligned} v(s) &= \frac{1}{\lambda_2}\left(\sum_{N_1} v(i) + \sum_{N_2} v(i)\right) \\ &= \frac{1}{\lambda_2^2}\left(\sum_{N_{11}} v(i)e_i(N_1) + \sum_{N_{21}} v(i)e_i(N_1) + \sum_{N_{12}} v(i)e_i(N_2) + \sum_{N_{22}} v(i)e_i(N_2)\right) \end{aligned}$$

Now we claim:

**Claim 21.**

$$\begin{aligned} &\sum_{N_{11}} v(i)e_i(N_1) + \sum_{N_{12}} v(i)e_i(N_2) \\ &+\sum_{N_{21}} v(i)e_i(N_1) + \sum_{N_{22}} v(i)e_i(N_2) > \frac{4}{5\sqrt{2n}}(d_1 - d_2)^2 \end{aligned}$$

This claim would prove the Lemma as it would show

$$v(s) > \frac{1}{\lambda_2^2}\frac{4}{5\sqrt{2n}}(d_1 - d_2)^2 \geqslant 0.8\frac{1}{\sqrt{2n}}$$

The last inequality follows from Lemma 17.

Let's prove the claim. First, for any $a, b \in \{1, 2\}$

$$\begin{aligned} \sum_{N_{ab}} v(i)e_i(N_b) &= \sum_{N_{ab}} u(i)e_i(N_b) + \sum_{N_{ab}} w(i)e_i(N_b) \\ &= sign(b)\frac{1}{\sqrt{2n}}e(N_b, N_{ab}) + \sum_{N_{ab}} w(i)e_i(N_b) \\ \Rightarrow \sum_{a,b}\sum_{N_{ab}} v(i)e_i(N_b) &= e(N_1, N_{11}) + e(N_2, N_{12}) \\ -e(N_2, N_{22}) - e(N_1, N_{21}) &+ \sum_{a,b}\sum_{N_{ab}} w(i)e_i(N_b) \end{aligned} \tag{19}$$

Now, since $s \in T_1, N_1 \geqslant d_1 - 4\sqrt{d_1 \log n}$ and $e(N_1, N_{11}) \geqslant (d_1 - 4\sqrt{d_1 \log n})^2$. Using similar bounds for $N_{12}, N_{21}$ and $N_{22}$

$$\begin{aligned} & e(N_1, N_{11}) + e(N_2, N_{12}) - e(N_2, N_{22}) - e(N_1, N_{21}) \\ \geqslant \; & (d_1 - 4\sqrt{d_1 \log n})^2 + (d_2 - 4\sqrt{d_2 \log n})^2 - 2(d_1 + 4\sqrt{d_1 \log n})(d_2 + 4\sqrt{d_2 \log n}) \\ \geqslant \; & d_1^2 + d_2^2 - 2d_1 d_2 - \Theta(d_1^{3/2}\sqrt{\log n}) \end{aligned}$$

Then,

$$e(N_1, N_{11}) + e(N_2, N_{12}) - e(N_2, N_{22}) - e(N_1, N_{21}) \geqslant 0.95(d_1 - d_2)^2 \tag{20}$$

Now we need to bound $\sum_{ab} \sum_{N_{ab}} |w(i)e_i(N_b)|$. The four terms in the sum are of the same order hence we will only bound one of them. We claim,

$$|\sum_{N_{11}} w(i)e_i(N_1)| \leqslant \frac{4}{c_1} n^{1/3}\sqrt{\log n} \leqslant \frac{1}{50}\frac{1}{\sqrt{2n}}(d_1 - d_2)^2 \tag{21}$$

Again we start with $e_i(N_1) \leqslant d_1(1 + 4\sqrt{\frac{\log n}{d_1}})$. Then,

$$\begin{aligned} \sum_{N_{11}} |w(i)e_i(N_1)| & \leqslant \sum_{t=1 \ldots \log 2d_1} \sum_{i : e_i \geqslant 2^{t-1}} 2^t |w(i)| \\ & \leqslant \log n \max_{t \leqslant \log d_1} 2^t \sum_{i : e_i(N_1) \geqslant 2^{t-1}} |w(i)| \end{aligned}$$

The problem here is that, conceivably, for a large number of vertices $i$, $w(i)$ is large and so is $e_i(N_1)$, thus amplifying the effect of $w$. What we will show is that $e_i(N_1)$ can be large only for a small number of vertices, thus disallowing this effect.

Let us bound for any $t \geqslant 1$ the value of

$$2^t \sum_{i : e_i(N_1) \geqslant 2^{t-1}} |w(i)|$$

For any $f$ define

- $M(f) = \{i : e(i, N_1) \geqslant f\}$
- $m(f) = |M(f)|$

Then setting $f = 2^t$, by Proposition 19 we can write

$$2^t \sum_{i : e_i \geqslant 2^{t-1}} w(i) = f \sum_{M(f/2)} w(i) \leqslant \|w\|\sqrt{m(f/2)}f$$

By the definition of $M(f)$, $e(M(f), N_1) \geqslant f m(f)$. Now by quasi-randomness

$$m(f) + 2\sqrt{n}\sqrt{m(f)} \geqslant f m(f)$$

Then, it is easy to see that $\sqrt{m(f)}f \leqslant 4\sqrt{n}$, the main point being that $f$ doesn't appear on the right hand side.

Therefore $\sum_{i : e_i \geqslant 2^{t-1}} 2^t w_i \leqslant \|w\| 4\sqrt{n} \leqslant \frac{4}{c_1}\frac{n^{1/3}}{\sqrt{\log n}}$. Eqn 21 now follows by simple manipulation and by assuming an appropriate value of $c_1$. Comparing Eqn 20 and 21, Claim 21 follows. □

## References


[1] D. Achlioptas and F. McSherry. On spectral learning of mixtures of distributions. In *Conference on Learning Theory (COLT)*, pages 458-469, 2005.

[2] N. Alon, M. Krivelevich and B. Sudakov. Finding a large hidden clique in a random graph. In *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 594-598, 1998.

[3] R. Bhatia. *Matrix Analysis*. Springer-Verlag, New York, 1997.

[4] B. Bollobas. *Random Graphs*. Cambridge University Press, 2nd Edition, 2001.

[5] R. Boppana. Eigenvalues and graph bisection: an average case analysis. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 280-285, 1987.

[6] A. Dasgupta, J. Hopcroft, R. Kannan, and P. Mitra. Spectral clustering by recursive bi-partitioning. In *European Symposium on Algorithms (ESA)*, pages 256-267, 2006.

[7] M. E. Dyer and A. M. Frieze. The solution of some random NP-hard problems in polynomial expected time. *J. Algorithms*, 10(4):451-489, 1989.

[8] P. Erdős and A. Rényi (1960). The Evolution of Random Graphs. *Publ. Math. Inst. Hung. Acad. Sci*, Vol. 5 (1960), pages 17-61.

[9] L. Erdős, B. Schlein, and H. Yau. Local Semicircle Law and Complete Delocalization for Wigner Random Matrices. *Communications in Mathematical Physics*. Vol. 287, Number 2 (2009), pages 641-655.

[10] Z. Füredi and J. Kómlos. The eigenvalues of random symmetric matrices. *Combinatorica*, 1:233-241, 1981.

[11] R. Horn and C. Johnson. *Matrix Analysis*, Cambridge University Press, 1990.

[12] M. Krivelevich and B. Sudakov. *Pseudo-random graphs*, In: More sets, graphs and numbers, E. Gyori, G. O. H. Katona, L. Lovasz, Eds., Bolyai Soc. Math. Studies Vol. 15, 199-262.

[13] F. McSherry. Spectral partitioning of random graphs. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 529-537, 2001.

[14] C. OCinneide. Entrywise perturbation theory and error analysis for markov chains. *Numer. Math.*, 65:109-120, 1993.

[15] T. Tao and V. Vu, Random matrices: Universality of the local eigenvalue statistics, *submitted*

[16] V. Vu. Spectral norm of random matrices. In *ACM Symposium on Theory of computing (STOC)*, pages 619-626, 2005.

[17] E. Wigner. Characteristic vectors of bordered matrices with innite dimensions. *Annals of Mathematics*, pages 548-564, 1955.