Spectral Techniques Applied to Sparse Random Graphs

Uriel Feige Eran Ofek

July 28, 2006

Abstract

We analyze the eigenvalue gap for the adjacency matrices of sparse random graphs. Let $\lambda_1 \geq \ldots \geq \lambda_n$ be the eigenvalues of an n-vertex graph, and let $\lambda = \max[\lambda_2, |\lambda_n|]$. Let c be a large enough constant. For graphs of average degree $d = c \log n$ it is well known that $\lambda_1 \geq d$, and we show that $\lambda = O(\sqrt{d})$. For $d = c$ it is no longer true that $\lambda = O(\sqrt{d})$, but we show that by removing a small number of vertices of highest degree in G, one gets a graph G' for which $\lambda = O(\sqrt{d})$. Our proofs are based on the techniques of Kahn and Szemeredi from STOC 1989, who proved similar results for regular graphs. Our results are useful for extending the analysis of certain heuristics to sparser instances of NP-hard problems. We illustrate this by removing some unnecessary logarithmic factors in the density of k-SAT formulas that are refuted by the algorithm of Goerdt and Krivelevich from STACS 2001.

1 Introduction

$G_{n,p}$ is the random graph model in which there is an n-vertex graph, and every edge is included independently with probability p. For such a graph G, we study the eigenvalues of its adjacency matrix. Let $\lambda_1 \geq \ldots \geq \lambda_n$ be its eigenvalues, and $\lambda = \max[\lambda_2, |\lambda_n|]$. We will mostly be interested in establishing a large gap between λ_1 and λ.

Let $d \simeq np$ denote the average degree of G. Then it is well known that $\lambda_1 \geq d$, and that $\lambda \geq \Omega(\sqrt{d})$ (where this lower bound on λ holds when d is not too close to n). Moreover, for a variety of random graph models, it is known that $\lambda = O(\sqrt{d})$. This was shown for $d = \Theta(n)$ by Furedi and Komlos [14]. Their proof, which is based on the so called *trace method*, can be extended in a straightforward way to values of d as low as roughly $(\log n)^6$. We remark that in [4] Boppana claimed that the proof can be extended to $d \simeq \log n$, but did not provide a proof of his claim (see also Section 4.2). For random d-regular graphs (with constant values of d), $\lambda = O(\sqrt{d})$ was established by Friedman, Kahn and Szemeredi [13].

Motivated by some recent applications in the design of algorithms for random instances of NP-hard problems, we extend the $\lambda = O(\sqrt{d})$ result to lower values of p in the $G_{n,p}$ model. Our proofs are based on techniques attributed to Kahn and Szemeredi in [13]. It is natural to expect these techniques to extend to random $G_{n,p}$ graphs with p large enough, because once the average degree d is large enough ($d > c \log n$), the random graph is nearly regular in the sense that all vertices have degree $d(1 \pm \epsilon)$. Still, there are some technical difficulties that need to be taken care of for the proof to go through.

Theorem 1.1. *Let G be a random graph taken from $G_{n,p}$ with $c_0 \frac{\log n}{n} \leq p \leq \frac{n^{1/5}}{10n}$. For every $c > 0$ there exists $c' > 0$ such that $\lambda(G) \leq c' \sqrt{np}$ with probability of at least $1 - 1/n^c$.*

The case of $d < \log n$ is more challenging, because as the average degree decreases the graph becomes less regular. In particular, when d is constant, there are vertices of degree $\Omega(\log n/\log\log n)$. The effect of this is more severe than just requiring a patch in the proof. The statement of the theorem needs to be changed, because λ (and also λ_1 and many other eigenvalues) are in fact much larger than d (in the order of $\sqrt{\log n/\log\log n}$, see [18]).

To cope with the case of constant d, we use an idea that was used by Alon and Kahale in [1]. We remove the vertices of highest degree from G, and analyze the eigenvalues of the remaining graph G'. For G' we can show that the average degree remains roughly d, whereas $\lambda = O(\sqrt{d})$. It turns out that in the applications that interest us, analyzing G' rather than G is good enough.

Theorem 1.2. *Let G be a random graph taken from $G_{n,d/n}$ ($d \geq d_0$ for some fixed constant d_0). Let G' be the subgraph induced by removing from G all the vertices of degree greater than $(1+\epsilon)d$ (where $3\sqrt{\log d/d} < \epsilon < 0.9$). Then $\lambda(G') < c'\sqrt{d}$ with probability of $1 - e^{-\Omega(\epsilon^2 d)}$, where c' is some universal constant.*

We are not aware of proofs of Theorems 1.1 and 1.2 being published elsewhere. Part of the contribution of our work is in explicitly stating and proving these theorems. We note however that our proofs of these theorems are based on known approaches, and are in some sense implicit in the known body of earlier works. That is, experts in using the known proof techniques can probably derive these results on their own, and may well have done so in the past (without explicitly publishing these results). In particular, our analysis is very similar to that given in Section 3 of [1], with some technical differences that are mainly the result of the fact that the model of random graphs studied in [1] was not exactly the $G_{n,p}$ model. Additional discussion of related work appears in Section 4 (and in particular in Section 4.3).

As applications of our results, we consider algorithms for handling random instances of NP-hard problems. Such algorithms often use spectral techniques. The eigenvectors of the adjacency matrix of a graph are known to be useful for finding sparse cuts (that tend to show up in the second eigenvector), dense cuts (that tend to show up in the most negative eigenvector), large independent sets (that tend to show up in the most negative eigenvector), colorings (color classes tend to show up in the most negative eigenvectors), and more. Algorithms that are based on semidefinite programming (such as the ϑ function of Lovasz) are often also implicitly based on spectral analysis, and in fact the computationally heavy semidefinite programming machinery can often be replaced by the computationally lighter eigenvalue computations when these algorithms are specialized to random instances. Our analysis of the eigenvalues of sparse matrices shows that the known spectral algorithms for NP-hard problems work on sparser instances than were known before. To illustrate this we show two examples (see Section 4 for details):

1. Improving the bound of the refutation algorithm of Goerdt and Krivelevich [16] by removing unnecessary logarithmic terms. This gives an algorithm which given a random k-SAT formula (for even k) with n variables and at least $c2^k n^{k/2}$ clauses (where c is a large enough constant), almost surely produces a proof that the formula is unsatisfiable (the bound in [16] is $c(\log(n))^7 2^k n^{k/2}$ clauses, for some constant c).

2. An algorithm which given a random sparse graph taken from $G_{n,\frac{d}{n}}$ almost surely certifies that the maximum cut in the graph bounded by $nd/4 + O(\sqrt{d}n)$ (for large enough constant d).

1.1 Notation

The number of vertices in a graph is denoted by n, and p denotes the probability of an edge in the $G_{n,p}$ model. Observe that the expected degree of a vertex in the $G_{n,p}$ model is $p(n-1)$, and with high probability, this is also roughly the average degree in G. We use d to denote np. For the adjacency matrix of graph G, $\lambda_1 \geq \ldots \geq \lambda_n$ denote its eigenvalues, and $v_1, \ldots, v_n$ denote its corresponding eigenvectors. The maximum of λ_2 and $|\lambda_n|$ is denoted by λ.

Throughout this paper we will use various constants. When the value of a constant has no particular importance we will denote it by $\Theta(1)$. In some cases, constants will have generic names (such as c for positive constants, and ϵ for positive constants that may be thought of as being arbitrarily small), and in other cases we give numerical values to constants (e.g., the value 2 in Lemma 2.1). No serious attempt has been made to optimize these constants. We will often use the terms almost surely and w.h.p. (with high probability) to denote probabilities which tend to 1 as n or some other parameter (like d or c) grows. Upon using these terms we will specify the dependency on the parameters. The term w.e.h.p. (with extremely high probability) will be used for probabilities larger than $1 - e^{-\Omega(n^{\delta})}$ (for some fixed $\delta > 0$). We will use $\|\cdot\|$ to denote the l_2 norm of a vector.

2 Bounding λ for $p \geq c_0 \frac{\log n}{n}$

In this section we prove Theorem 1.1.

Let A be the adjacency matrix of graph G, let $\lambda_1 \geq \ldots \geq \lambda_n$ be its eigenvalues, and let $v_1, \ldots, v_n$ be its corresponding eigenvectors. It is well known that λ ($= \max[\lambda_2, |\lambda_n|]$) is given by the Rayleigh quotient:

$$\lambda = \max_{v \perp v_1} |v^t A v / v^t v| = \max_{\substack{v \perp v_1, \\ \|v\|=1}} |v^t A v|.$$

2.1 Replacing the first eigenvector by the "all ones" vector

To use the Rayleigh quotient above, we should first determine v_1. For regular graphs v_1 is the "all ones" vector $\vec{1}$. However, in our $G_{n,p}$ model G is most likely not regular, in which case $v_1 \neq \frac{\vec{1}}{\sqrt{n}}$. However, as G is nearly regular when $p \geq c_0 \frac{\log n}{n}$, v_1 is close to $\frac{\vec{1}}{\sqrt{n}}$. We show that in our $G_{n,p}$ model, we can replace v_1 by $\frac{\vec{1}}{\sqrt{n}}$ when using the Rayleigh quotient to bound λ from above.

Lemma 2.1. *Let A be the adjacency matrix of a graph G. Let u be an arbitrary unit vector, and let $\tilde{d}$ denote the value of $u^t A u$. If for every vector $w \perp u$ of norm 1 it holds that (1) $|w^t A u| \leq c\sqrt{\tilde{d}}$ and (2) $|w^t A w| \leq c\sqrt{\tilde{d}}$, then $\lambda_1 \leq \tilde{d} + 2c\sqrt{\tilde{d}}$, and $\lambda \leq 2c\sqrt{\tilde{d}}$.*

Proof. First we bound λ_2 using the formula $\lambda_2 = \min_H \max_{\substack{x \in H, \\ x \neq 0}} \frac{x^t A x}{x^t x}$, where H ranges over all subspaces of R^n of co-dimension 1. Choosing H to be the subspace perpendicular to u, and using condition (2) of the lemma, we derive that $\lambda_2 \leq c\sqrt{\tilde{d}}$. To bound λ_1 and λ_n, let v be an arbitrary unit vector. We can write $v = \alpha u + \beta w$ where $w \perp u$ and $\alpha^2 + \beta^2 = 1$.

$$v^t A v = (\alpha u + \beta w)^t A (\alpha u + \beta w) = \alpha^2 u^t A u + 2\alpha\beta w^t A u + \beta^2 w^t A w$$

where in the last equality we used the fact that A is a symmetric matrix. Now we use $u^tAu = \tilde{d} > 0$, $0 \le \alpha, \beta \le 1$ and $\alpha\beta \le 1/2$ (because $\alpha^2 + \beta^2 = 1$), and conditions (1) and (2) of the lemma for $w \perp u$, and conclude that

$$-2c\sqrt{\tilde{d}} \le v^tAv \le \tilde{d} + 2c\sqrt{\tilde{d}}.$$

Substituting for v the various eigenvectors of A, we get that all eigenvalues lie in the range $[-2c\sqrt{\tilde{d}}, \tilde{d} + 2c\sqrt{\tilde{d}}]$, which together with the upper bound on λ_2 proves the lemma. $\square$

In the context of the $G_{n,p}$ model, the vector u from Lemma 2.1 will be given the value $\frac{\vec{1}}{\sqrt{n}}$. In this case, u^tAu is exactly equal to the average degree of the graph G, and hence $\tilde{d}$ is the average degree. The following lemma takes care of condition (1) required in Lemma 2.1.

Lemma 2.2. *Let G be a random graph taken from $G_{n,p}$ where $c_0\frac{\log n}{n} \le p \le \frac{n^{1/5}}{10n}$. Let $\bar{d}$ denote the average degree of G, and let d denote np. Then with probability of at least $1 - e^{-\Omega((nd)^{1/3})}$ the adjacency matrix of G denoted by A has the following property:*

$$\forall x \in R^n, x \perp u \quad |x^tAu| \le 2\ \sqrt{\bar{d}}\ \|x\|.$$

Proof. We use the following notation. The variable δ_v denotes the difference between the degree of vertex v and $\bar{d}$. The vector δ denotes the n-vector whose entries are δ_v.

$$x^tAu = x^t \begin{bmatrix} \bar{d}+\delta_1 \\ \bar{d}+\delta_2 \\ . \\ . \\ \bar{d}+\delta_n \end{bmatrix} \frac{1}{\sqrt{n}} = x^t \begin{bmatrix} \delta_1 \\ \delta_2 \\ . \\ . \\ \delta_n \end{bmatrix} \frac{1}{\sqrt{n}} + dx^tu = x^t \begin{bmatrix} \delta_1 \\ \delta_2 \\ . \\ . \\ \delta_n \end{bmatrix} \frac{1}{\sqrt{n}}, \tag{1}$$

hence $|x^tAu| < \|x\|\|\delta\|/\sqrt{n}$. By Lemma 5.1 with probability of at least $1 - e^{-\Omega((nd)^{1/3})}$ it holds that $\|\delta\| \le \sqrt{2n\bar{d}}$. $\square$

From here on our goal is to prove condition (2) of Lemma 2.1. Namely, to bound the Rayleigh quotient of every vector $x \perp u$. We will use the technique of Kahn and Szemeredi [13] with some modifications.

2.2 Bounding the Rayleigh Quotient of all $x \perp u$

Recall that $u = \frac{\vec{1}}{\sqrt{n}}$. We want to bound x^tAx for every $x \perp u$ of norm 1. The first step is to reduce the set of vectors into a finite, yet exponentially large space.

2.2.1 Reduction to Discrete Space

Let $S = \{v \perp \vec{1}\ :\ \|v\| \le 1\}$. We fix some $0 < \delta < 1$ and use it to define a grid which approximates S (we use $\mathbb{Z}$ to denote the set of integers):

$$T = \{x \in (\frac{\delta}{\sqrt{n}}\mathbb{Z})^n\ :\ \sum_i x_i = 0, \|x\| \le 1\}.$$

It will be convenient to think of δ as the constant $\frac{1}{2}$.

Lemma 2.3. *Every vector $v \in S$ whose norm is less than $1-\delta$ is a convex combination of points from T.*

Proof. Let C be the specific hypercube from the grid $(\frac{\delta}{\sqrt{n}}\mathbb{Z})^n$ in which v lies. We first show that the vertices of C are inside T. The length of C's sides is $\frac{\delta}{\sqrt{n}}$, thus any two points inside C are δ close. We conclude that all points of C are at l_2 distance of at most 1 from the origin, which implies that C's vertices are inside T. Notice that the intersection of C with the surface $\sum_i x_i = 0$ is a polytope P (i.e. it is an intersection of half-spaces), and that v is inside this polytope. To show that v is a convex combination of points from T we will prove that all the vertices of the above mentioned polytope are points of T. We call the points of T *integral* points as in each coordinate they have a multiple of $\frac{\delta}{\sqrt{n}}$.

It is well know that for every vertex w of a given polytope P' in $\mathbb{R}^n$, there exists a weight function $\alpha \in \mathbb{R}^n$ such that the maximum of $< x, \alpha >= \sum_{i=1}^n x_i \alpha_i$ over $x \in P'$ is obtained uniquely at w. Therefore it is enough to show that for any weight function $\alpha \in \mathbb{R}^n$, the maximum of $< x, \alpha >$ over $x \in P$ is obtained at some point of T. Fix any weight function α. Let x be the point inside the polytope P for which the maximum is attained. If x is not integral then there exist indexes $i_1, i_2, \ldots i_l$ which correspond to non integral coordinates. Notice that $l \neq 1$ since otherwise $\sum_i x_i \neq 0$ (follows from integrality argument). Fix two non integral coordinates i_j, i_k. Without loss of generality assume that $x_{i_j}, x_{i_k} > 0$ (a similar argument can be used also in the other 3 cases). By adding γ to x_{i_j} and subtracting γ from x_{i_k}, the constraint $\sum_{i=1}^n x_i = 0$ is maintained and the objective function changes by $\gamma(\alpha_{i_j} - \alpha_{i_k})$. By choosing the sign of γ to be sign of $\alpha_{i_j} - \alpha_{i_k}$ we guarantee that the objective function does not decrease. We then increase $|\gamma|$ (starting from 0) until one of x_{i_j}, x_{i_k} (or both) becomes integral. Iterating this operation decreases the number of non integral coordinates until it reaches 0. We will never be left with only one non integral coordinate since in this case it cannot be that $\sum_{i=1}^n x_i = 0$. □

Claim 2.4. *Let $c \in \mathbb{R}$ be an arbitrary constant. If for every $x, y \in T$ $|x^t Ay| \leq c$, then for every $x \in S$ $|x^t Ax| \leq \frac{c}{(1-\delta)^2}$.*

Proof. Let $x \in S$, define $z = (1-\delta)x$. By Lemma 2.3 z is a convex combination of vectors from T, i.e. $z = \sum_i \alpha_i v_i,\ v_i \in T,\ \sum_i \alpha_i = 1$.

$$|z^t Az| = |(\sum_i \alpha_i v_i) A (\sum_i \alpha_i v_i)| \leq \sum_{i,j} \alpha_i \alpha_j |v_i^t A v_j| \leq c \sum_{i,j} \alpha_i \alpha_j \leq c. \tag{2}$$

It follows that $|x^t Ax| \leq \frac{1}{(1-\delta)^2}\ c$. □

2.2.2 The Proof Structure

Claim 2.4 reduced our problem to proving a similar bound only for pairs of vectors from T. In the remainder of this section we will prove the following theorem:

Theorem 2.5 (Main Theorem). *Let A be the adjacency matrix of a graph taken from $G_{n,p}$, where $c_0 \frac{\log n}{n} \leq p \leq \frac{n^{1/3}}{n(\log n)^{5/3}}$. For every $c > 0$ there exists $c' > 0$ such that with probability of $1 - 1/n^c$ the following holds:*

$$\forall x, y \in T: \qquad |x^t Ay| \leq c' \sqrt{np}.$$

In the remainder of section 2 we will use d to denote to expected degree in the graph, i.e. $d = np$. Fix a particular pair of vectors $x, y \in T$. We use the graph vertices to index the vectors x, y. For example we will use x_u to denote the coordinate of x that corresponds to vertex u. We will bound the sum $\sum_{u,v} x_u A_{u,v} y_v$. The proof has two parts. The first part deals with the contribution (to the sum) of couples (u, v) for which $x_u y_v \leq \frac{\sqrt{d}}{n}$; we call these couples light couples. We will show that the expected contribution (over the choice of a random graph G) of these couples is $O(\sqrt{d})$, and also that with extremely high probability this is the case. Using the union bound (over all choices $x, y \in T$) we deduce that w.e.h.p. these contribute at most $O(\sqrt{d})$. The second part deals with the contribution of heavy couples (a couple is heavy if it is not light). For these couples the combination of union bound and tight concentration does not work. To bound their contribution we rely on the following two properties which a random graph almost surely has:

- **Bounded degree property:** every vertex has a degree bounded by $c_1 d$ (for some $c_1 > 1$).
- **Discrepancy property:** the graph has bounded discrepancy. Roughly speaking it means that for every two subsets $A, B \subset [n]$ the number of edges between A and B does not exceed the expectation by much. Let $e(A, B)$ be the random variable which equals to the number of edges between A and B. We set $\mu(A, B)$ to be $|A||B|\ \frac{d}{n}$ so that it is an upper bound on the expected number of edges between A, B (if $A \cap B = \phi$ then the bound is tight). The graph has bounded discrepancy if for every $A, B \subseteq [n]$ ($|B| \geq |A|$) one of the following holds:
  1. $\frac{e(A,B)}{\mu(A,B)} \leq ec_2$.
  2. $e(A, B) \log \frac{e(A,B)}{\mu(A,B)} \leq c_3 |B| \log \frac{n}{|B|}$ (this bound allows larger deviation for small values of $|B|$).

It is not hard to prove that the bounded degree property holds with probability of $1 - e^{-\Omega(\min\{(c_1-1)^2, c_1 \ln c_1\} d)}$, when $d > c_0 \log n$ (union bound, concentration of binomial distribution around its mean). The discrepancy property holds with probability of $1 - 1/n$ for some choice of universal constants c_2, c_3 (this is proved in Section 2.2.5). We will show that these two properties imply that for every choice of x, y the total contribution of heavy couples of x, y is $O(\sqrt{d})$.

2.2.3 Bounding the Contribution of Light Couples

The following lemma bounds the expectation of the contribution of light couples.

Lemma 2.6. *Fix a pair of vectors $x, y \in T$. Let G be a random graph taken from $G_{n,p}$, whose adjacency matrix we denote by A. Let $d = np$, and $L = \{(u, v) : \ |x_u \ y_v| < \frac{\sqrt{d}}{n}\}$. Then the expectation of $|\sum_L x_u y_v A_{u,v}|$ is bounded by $O(\sqrt{d})$.*

Proof. Since $\sum_u x_u = 0$, $\sum_v y_v = 0$, it follows that $\sum_L x_u y_v + \sum_{\overline{L}} x_u y_v = 0$. The sum $\sum_{u,v \in \overline{L}} x_u y_v$ can be bounded as follows:
let $\{\gamma_i\}$ be the values which appear in the vector x. Let a_i be the number of times that the value γ_i appears in x. We define δ_j and b_j similarly with respect to y.

$$|\sum_{(u,v)\in\overline{L}} x_u y_v| = |\sum_{i,j:\ |\gamma_i \ \delta_j| \geq \frac{\sqrt{d}}{n}} a_i b_j \gamma_i \delta_j| = |\sum_{i,j:\ |\gamma_i \ \delta_j| \geq \frac{\sqrt{d}}{n}} a_i b_j \gamma_i^2 \delta_j^2 \frac{1}{\gamma_i \ \delta_j}|$$

$$\leq \frac{n}{\sqrt{d}} \sum_{i,j} a_i b_j \gamma_i^2 \delta_j^2 = \frac{n}{\sqrt{d}} (\sum_i a_i \gamma_i^2)(\sum_j b_j \beta_j^2) \leq \frac{n}{\sqrt{d}}.$$

It follows that $|\sum_L x_u y_v| \leq \frac{n}{\sqrt{d}}$. Since every edge is chosen with probability $\frac{d}{n}$, the expectation of $\sum_L x_u y_v A_{u,v}$ is bounded by $\sqrt{d}$ in absolute value. □

Fix two vectors $x, y \in T$. As before, L denotes the set of light couples. We denote by X the contribution of the set of light couples (X is a random variable induced by the choice of the random matrix A). We already saw that $|\mathbb{E}[X]| \leq \sqrt{d}$.

Claim 2.7. *For every $c > 0$ there exists a number $K > 0$ such that:*

$$\Pr[|X| > K\sqrt{d}] < e^{-cn}. \tag{3}$$

Proof. We prove a concentration result for every $x, y \in T$. In general, the pair x, y that maximizes the Rayleigh quotient of a matrix satisfies $x = y$. However, as we work with a discrete space, the optimum over all $x, y \in T$ need not satisfy $x = y$, and our proof will not assume $x = y$. We remark that proving a concentration result only for $x = y$ is easier as it can be handled as a sum of independent random variables (since $x_u y_v = x_v y_u$). However if $x \neq y$ then $x_u y_v$ may be different than $x_v y_u$, but their contribution to the sum is correlated as they both depend on the edge between vertices u, v. This adds some complications to the proof.
Denote by $a_{(u,v)} = a_{(v,u)}$ the potential contribution of the edge (u, v) to the light couples, i.e. $a_{(u,v)} = f(x_u y_v) + f(x_v y_u)$, where f is :

$$f(x) = \begin{cases} x & |x| \leq \frac{\sqrt{d}}{n} \\ 0 & \text{otherwise} \end{cases}$$

Notice that $a_{(u,v)}$ can be either positive or negative but bounded by $\frac{2\sqrt{d}}{n}$ in absolute value. Let $X = \sum_{\{u,v\}} x_{(u,v)}$, where

$$x_{(u,v)} = x_{(v,u)} \sim \begin{cases} a_{u,v} & w.p. \;\; p = \frac{d}{n} \\ 0 & w.p. \;\; 1-p \end{cases}$$

As the term $\sum_{\{u,v\}} a^2_{u,v}$ will appear in the following analysis we will first bound it. Notice that $a^2_{u,v} \leq \max\{(x_u y_v + x_v y_u)^2, (x_u y_v)^2 + (x_v y_u)^2\}$: if $f(x_u y_v)$ and $f(x_v y_u)$ have the same sign then $a^2_{u,v} = (x_u y_v + x_v y_u)^2$, otherwise $a^2_{u,v} \leq (x_u y_v)^2 + (x_v y_u)^2$.

$$\sum_{\{u,v\} \in \binom{V}{2}} a^2_{u,v} \leq \frac{1}{2} \sum_{\substack{(u,v)\in \\ V\times V}} a^2_{u,v} \leq \frac{1}{2} \sum_{\substack{(u,v)\in \\ V\times V}} \max\{(x_u y_v + x_v y_u)^2, (x_u y_v)^2 + (x_v y_u)^2\} \leq$$

$$\frac{1}{2} \sum_{\substack{(u,v)\in \\ V\times V}} (x_u^2 y_v^2 + x_v^2 y_u^2) + |\sum_{\substack{(u,v)\in \\ V\times V}} x_u y_u x_v y_v| = (\sum_u x_u^2)(\sum_v y_v^2) + |(\sum_u x_u y_u)(\sum_v x_v y_v)| \leq$$

$$\|x\|^2 \|y\|^2 + < x, y >^2 = 2\|x\|^2 \|y\|^2 \leq 2. \tag{4}$$

For simplicity we will use the indexing $i = 1, 2, \ldots, N = \binom{n}{2}$ instead of $\{u, v\}$. Using the new indexing inequality (4) becomes $\sum_{i=1}^{N} a_i^2 \leq 2$.
We know that $|\mathbb{E}[X]| \leq \sqrt{d}$. Without loss of generality we may assume $\mathbb{E}[X] = \mu \geq 0$ (otherwise we can work with $-X$). We fix $\lambda = \frac{n}{4\sqrt{d}}$ so that $|\lambda a_i| \leq \frac{1}{2}$. We will use the inequality $e^x - 1 \leq x + 2x^2$ (valid for $|x| \leq \frac{1}{2}$, see Lemma 2.8 below), and also the inequality $e^x \geq 1 + x$.

$$\Pr[X > k\mu] = \Pr[e^{\lambda X} > e^{\lambda k \mu}] \leq \frac{\mathbb{E}[e^{\lambda X}]}{e^{\lambda k \mu}}.$$

$$\begin{aligned}
\mathbb{E}[e^{\lambda X}] &= \prod_{i=1}^{N} \mathbb{E}[e^{\lambda x_i}] = \prod_{i=1}^{N} [p e^{\lambda a_i} + (1-p) e^0] \\
&= \prod_{i=1}^{N} [p(e^{\lambda a_i} - 1) + 1] \leq \prod_{i=1}^{N} [p(\lambda a_i + 2(\lambda a_i)^2) + 1] \\
&\leq e^{p\lambda \sum_{i=1}^{N} a_i + 2p\lambda^2 \sum_{i=1}^{N} a_i^2} \leq e^{\lambda\mu + 4p\lambda^2}.
\end{aligned}$$

Thus:

$$\Pr[X > k\mu] \leq e^{\lambda\mu + 4p\lambda^2 - k\lambda\mu} = e^{\lambda(4p\lambda - (k-1)\mu)} = e^{\frac{n}{4\sqrt{d}}(\sqrt{d} - (k-1)\mu)}.$$

Setting $k = 3c\frac{\sqrt{d}}{\mu}$ (for some $c \geq 1$) yields that $\Pr[X > 3c\sqrt{d}] < e^{-cn}$. Applying the same argument to $-X$ yields:

$$\Pr[|X| > 3c\sqrt{d}] \leq 2e^{-cn}.$$

□

Lemma 2.8. *$e^x - 1 \leq x + 2x^2$ for $|x| \leq \frac{1}{2}$.*

Proof.

$$\begin{aligned}
e^x - 1 &= x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots \leq x + x^2 + (|x|^3 + |x|^4 + \ldots) \\
&\leq x + x^2 + |x|x^2(1 + |x| + |x|^2 + \ldots) \leq x + 2x^2.
\end{aligned}$$

□

To bound the contribution of light couples we use the union bound.

Claim 2.9. *The number of vectors in T is bounded by e^{cn} for some c which depends on δ.*

Proof. If we shift the grid (which is used to define T) by $\frac{\delta}{2\sqrt{n}}$ in each direction, then every "integer" point of T is covered by exactly one hypercube of the shifted grid. A hypercube of the shifted grid is called *covering* if it contains an integer point of T. We will bound the number of such covering hypercubes using a volume argument. Each covering hypercube has a volume of $(\frac{\delta}{\sqrt{n}})^n$. All the covering hypercubes are inside a ball of radius $1 + \delta$ from the origin, since all the points in T are at distance ≤ 1 from the origin and every two points in the hypercube are δ close. The volume of the ball is bounded by $\frac{((1+\delta)^2\pi)^{n/2}}{(\frac{1}{2}n+1)!} < \left(\frac{(1+\delta)^2 2e\pi}{n}\right)^{n/2}$. It follows that $|T| \leq \left(\frac{(1+\delta)^2 2e\pi}{n}\right)^{n/2} / (\frac{\delta}{\sqrt{n}})^n \leq (9/\delta)^n = e^{n \ln(9/\delta)}$. □

Corollary 2.10. *With probability of $1 - e^{-(c/3 - \ln(9/\delta))n}$ the light couples contribute at most $c\sqrt{d}$.*

2.2.4 Bounding the Contribution of Heavy Couples

Let $x, y \in T$. We wish to bound the absolute value of:

$$x^t Ay = \sum_{u,v:\ |x_u y_v| \geq \frac{\sqrt{d}}{n}} x_u A_{u,v} y_v.$$

First we define an approximation for x, y, which can be handled more easily. We will handle only couples u, v such that $x_u > 0, y_v > 0$, the other 3 cases can be handled similarly. We will use the following notation (δ was already defined at Section 2.2.1):
$\gamma_i = 2^i$.
$A_i = \{u : \frac{\delta}{\sqrt{n}}\gamma_{i-1} \leq x_u < \frac{\delta}{\sqrt{n}}\gamma_i\},\ a_i = |A_i|$ $\qquad i = 1, 2 \ldots, \lceil \log \sqrt{n}/\delta \rceil$.
$B_j = \{u : \frac{\delta}{\sqrt{n}}\gamma_{j-1} \leq y_u < \frac{\delta}{\sqrt{n}}\gamma_j\},\ b_j = |B_j|$ $\qquad j = 1, 2 \ldots, \lceil \log \sqrt{n}/\delta \rceil$ (i runs between 1 and $\lceil \log \sqrt{n}/\delta \rceil$ because x_u is a multiple of $\frac{\delta}{\sqrt{n}}$ and bounded by 1).
$\mu_{i,j} = a_i b_j \frac{d}{n}$ $\quad$ bounds the expected number of edges between the sets A_i, B_j.
$\lambda_{i,j} = \lambda(A_i, B_j) = e(A_i, B_j)/\mu_{i,j}$.

$$x^t Ay \leq \sum_{\substack{i,j:\\ \gamma_i\gamma_j \geq \sqrt{d}}} e(A_i, B_j)\ \frac{\gamma_i \delta}{\sqrt{n}} \frac{\gamma_j \delta}{\sqrt{n}} = \delta^2 \sum_{\substack{i,j:\\ \gamma_i\gamma_j \geq \sqrt{d}}} \mu_{i,j}\lambda_{i,j}\ \frac{\gamma_i}{\sqrt{n}} \frac{\gamma_j}{\sqrt{n}} = \delta^2 \sum_{\substack{i,j:\\ \gamma_i\gamma_j \geq \sqrt{d}}} a_i b_j \frac{d}{n} \lambda_{i,j}\ \frac{\gamma_i}{\sqrt{n}} \frac{\gamma_j}{\sqrt{n}} =$$

$$\delta^2 \sqrt{d} \sum_{\substack{i,j:\\ \gamma_i\gamma_j \geq \sqrt{d}}} \overbrace{a_i \frac{\gamma_i^2}{n}}^{\alpha_i}\ \overbrace{b_j \frac{\gamma_j^2}{n}}^{\beta_j}\ \overbrace{\frac{\lambda_{i,j}\ \sqrt{d}}{\gamma_i\ \gamma_j}}^{\sigma_{i,j}}. \tag{5}$$

We will assume that the graph has the discrepancy and the bounded degree properties (defined at Section 2.2.2) to prove that:

$$\sum_{i,j} \alpha_i\ \beta_j\ \sigma_{i,j} \leq 2 \max\{ec_2(2/\delta)^4, 8c_1/\delta^2, 8c_3\} = \theta(1) \tag{6}$$

(c_1, c_2, c_3 are from the discrepancy and the bounded degree properties). Notice that $\sum_i \alpha_i \leq 4\|x\|^2/\delta^2 \leq 4/\delta^2$, similarly $\sum_j \beta_j \leq 4\|y\|^2 \leq 4/\delta^2$. We will handle only couples for which $b_j \geq a_i$, the other couples can be handled in a similar way. We partition the heavy couples into a union of 6 sets and prove the bound separately for every set. We proceed by a case analysis:

1. $\sigma_{i,j} \leq 1$: the sum is bounded by $(\sum_j \beta_j)(\sum_i \alpha_i) \leq 16/\delta^4 = (2/\delta)^4$.

2. $\lambda_{i,j} \leq ec_2$ (case 1 of the discrepancy property): since $\gamma_i\gamma_j > \sqrt{d}$ it follows that $\sigma_{i,j} \leq ec_2$ and thus the sum is bounded by $ec_2(2/\delta)^4$.

3. $\gamma_i > \sqrt{d}\ \gamma_j$: the bounded degree property implies that $e_{i,j} \leq a_i\ c_1 d \Rightarrow \lambda_{i,j} \leq \frac{c_1 n}{b_j}$.

$$\text{For fixed } i: \qquad \sum_j \beta_j \sigma_{i,j} = \sum_j b_j \frac{\gamma_j^2}{n} \frac{\lambda_{i,j}\sqrt{d}}{\gamma_i\ \gamma_j} \leq c_1 \sum_j \frac{\sqrt{d}\ \gamma_j}{\gamma_i} \leq 2c_1,$$

the last inequality holds since the sum is geometric and each summand is bounded by 1.

$$\text{It follows that:} \qquad \sum_i \left(\alpha_i \cdot \sum_{j:\gamma_i\gamma_j \geq \sqrt{d}} \beta_j \sigma_{i,j} \right) \leq 2c_1 \sum_i \alpha_i \leq \frac{8c_1}{\delta^2}.$$

4. From here we will use case 2 of the discrepancy property (as $\lambda_{i,j} > ec_2$):

$$e_{i,j} \log \lambda_{i,j} \leq c_3 b_j \log \frac{n}{b_j}.$$

First let us express the equation using the notation α, β, σ introduced in equation (5) :

$$(\lambda_{i,j} a_i b_j \frac{d}{n}) \log \lambda_{i,j} \leq c_3 b_j \log \frac{\gamma_j^2}{\beta_j} \quad (\text{substitute: } e_{i,j} = \lambda_{i,j}\mu_{i,j} \ ; \ \beta_j = \frac{\gamma_j^2 b_j}{n}),$$

$$(\lambda_{i,j}\sqrt{d}\ a_i \frac{1}{n}) \log \lambda_{i,j} \leq c_3 \frac{1}{\sqrt{d}} \log \frac{\gamma_j^2}{\beta_j} \qquad (\text{divide last inequality by } \sqrt{d} b_j),$$

$$\underbrace{\frac{\lambda_{i,j}\sqrt{d}}{\gamma_i\gamma_j}}_{=\sigma_{i,j}} \underbrace{a_i \frac{\gamma_i^2}{n}}_{=\alpha_i} \log \lambda_{i,j} \leq c_3 \frac{\gamma_i}{\gamma_j\sqrt{d}} \log \frac{\gamma_j^2}{\beta_j} \qquad (\text{multiply last inequality by } \frac{\gamma_i}{\gamma_j}),$$

$$\sigma_{i,j}\alpha_i \log \lambda_{i,j} \leq c_3 \frac{\gamma_i}{\gamma_j\sqrt{d}} \left[2\log \gamma_j + \log \frac{1}{\beta_j} \right]. \tag{7}$$

In subcases (a),(b) we will fix j and will bound $\sum_i \sigma_{i,j}\alpha_i$ by a constant. This will be enough as $\sum_j \beta_j$ is bounded by constant. A similar argument with respect to i is used in subcase (c). Each of the following subcases corresponds to one of the terms $\lambda_{i,j}, \gamma_j, \frac{1}{\beta_j}$ being the dominant one out of the three.

(a) $\log \lambda_{i,j} > \frac{1}{4}[2\log \gamma_j + \log \frac{1}{\beta_j}]$:
from (7) it follows that $\sigma_{i,j}\ \alpha_i \leq 4c_3\ \frac{\gamma_i}{\gamma_j\sqrt{d}}$, thus for a fixed j:

$$\sum_i \sigma_{i,j}\ \alpha_i \leq 4c_3\ \sum_i \frac{\gamma_i}{\gamma_j\ \sqrt{d}} \leq 8c_3.$$

The last inequality is due to the fact that $\gamma_i \leq \sqrt{d}\ \gamma_j$ (we are not in case 3), and the sum is geometric.

(b) $2\log \gamma_j \geq \log \frac{1}{\beta_j}$:
Since $\log \lambda_{i,j} \leq \frac{1}{4}[2\log \gamma_j + \log \frac{1}{\beta_j}]$ it follows that $\log \lambda_{i,j} \leq \log \gamma_j$ and thus $\lambda_{i,j} \leq \gamma_j$. Combined with the fact that $1 \leq \sigma_{i,j} = \frac{\lambda_{i,j}}{\gamma_j}\frac{\sqrt{d}}{\gamma_i}$, it follows that $\gamma_i \leq \sqrt{d}$.
As $\log \lambda_{i,j} \geq 1$ (we are not in case 2) and $2\log \gamma_j \geq \log \frac{1}{\beta_j}$, we use inequality (7) to derive:

$$\sigma_{i,j}\ \alpha_i \leq c_3 \frac{\gamma_i}{\gamma_j\sqrt{d}}\ 4\log \gamma_j,$$

and so

$$\sum_i \sigma_{i,j}\ \alpha_i \leq 4c_3\ \sum_i \frac{\gamma_i}{\gamma_j\sqrt{d}} \log \gamma_j \leq 4c_3\ \sum_i \frac{\gamma_i}{\sqrt{d}} \leq 8c_3.$$

(c) $2\log\gamma_j \le \log\frac{1}{\beta_j}$:

we are not in (a) thus $\log\lambda_{i,j} \le \log\frac{1}{\beta_j}$. It follows that $\sigma_{i,j} = \frac{\lambda_{i,j}\ \sqrt{d}}{\gamma_i\ \gamma_j} \le \frac{1}{\beta_j}\frac{\sqrt{d}}{\gamma_i\ \gamma_j}$, thus:

$$\sum_j \sigma_{i,j}\ \beta_j \le \sum_j \frac{\sqrt{d}}{\gamma_i\ \gamma_j} \le 2.$$

The last inequality is valid since the sum is geometric and every summand is bounded by 1 ($\gamma_i\gamma_j \le \sqrt{d}$).

Corollary 2.11. *Let G be a graph which has the discrepancy and bounded degree properties (as stated in Section 2.2.5) with constants c_1, c_2, c_3 respectively. Let A be the adjacency matrix of G. For any two vectors x, y which belong to the grid T (as defined in 2.2.1) it holds that:*

$$x^t Ay \le 8\max\{16ec_2/\delta^2, 8c_1, 8c_3\delta^2\}\sqrt{d}.$$

2.2.5 The Discrepancy Property

In this section we prove that in a random graph almost surely the number of edges between every two subsets A, B is not much higher than the expectation.
Let $A, B \subseteq [n]$. Denote by $e(A,B)$ the number of edges between A and B ($e(A,B)$ is a random variable). We set $\mu(A,B)$ to be $|A||B|\ \frac{d}{n}$ so that it bounds the expected number of edges between A, B. If A, B are not disjoint then $\mu(A,B)$ is larger than the expectation. We show that with probability of at least $1-\frac{1}{n}$ the graph has the following property:
for every $A, B \subseteq [n]$ ($|B| \ge |A|$) one of the following holds:

1. $\frac{e(A,B)}{\mu(A,B)} \le ec_2$.
2. $e(A,B)\log\frac{e(A,B)}{\mu(A,B)} \le c_3|B|\log\frac{n}{|B|}$.

Proof. Let a, b denote the cardinalities of A, B respectively. The possible number of edges between A and B is bounded by ab. We will assume that A, B are disjoint and thus there are exactly ab edges to choose from. The expected number of edges between A and B is $\mu = ab\frac{d}{n}$. The bound we are going to prove is valid also for the case in which A, B are not disjoint, since in this case there are less edges to choose from. The edges are chosen independently. We divide the analysis into two cases:
Case 1: $b \ge \frac{n}{e}$. In this case $\frac{e(A,B)}{\mu(A,B)} \le ec_1$ since $\mu(A,B) = ab\frac{d}{n} \ge \frac{ad}{e}$ but $e(A,B) \le ac_1d$ (we use the bounded degree assumption from Section 2.2.2).

Case 2: We will use the following bound which bounds the probability that the sum of independent indicator variables deviate from their expectation (see for example [3] Appendix A):

$$\Pr[e(A,B) > k\mu] < e^{-(k\ln k)\mu/3} \tag{8}$$

(the last inequality is valid for $k \ge 4$). We want to find a minimal $k \ge 8$ such that the event $e(A,B) \le k\mu$ will hold with high probability for all pairs A, B. Thus, we require:

$$e^{-(k\ln k)\mu/3}\binom{n}{a}\binom{n}{b} \le \frac{1}{n^3}.$$

The term $1/n^3$ is chosen in anticipation of later use of the union bound over all choices of a, b. Simplifying, we get:

$$e^{-(k\ln k)\mu/3} \left(\frac{ne}{a}\right)^a \left(\frac{ne}{b}\right)^b e^{3\log n} \leq e^0,$$

equivalently:

$$(k\ln k)\mu/3 \geq a(1+\log\frac{n}{a}) + b(1+\log\frac{n}{b}) + 3\log n.$$

As in this case $b \leq \frac{n}{e}$ we get that $\log(\frac{n}{b}) \geq 1$. Since $x\log\frac{n}{x}$ is monotone in $[1, \frac{n}{e}]$ it follows that $b\log\frac{n}{b} \geq a\log\frac{n}{a}$ (the monotonicity of $x\log\frac{n}{x}$ is postponed to Lemma 2.12). Thus it is enough to require:

$$(k\log k)\mu/3 \geq 4b\log\frac{n}{b} + 3\log n.$$

Again, since $x\log\frac{n}{x}$ is monotone in $[1, \frac{n}{e}]$ it follows that $b\log\frac{n}{b} \geq \log n$, thus it is enough that:

$$k\log k \geq \frac{21b}{\mu} \log\frac{n}{b}.$$

For a fixed size b, we let k' be smallest number such that $k'\log k' \geq \frac{21b}{\mu} \log\frac{n}{b}$. We then set $k = \max\{k', 4\}$, so we can use equation (8). Using the union bound over the choices of a, b we get that with probability of at least $1 - \frac{1}{n}$ for every choice of A, B $(b \leq n/e)$ the following holds:

$$e(A,B) \leq k\mu(A,B).$$

If $k' \leq 4$ then $k = 4$ and condition (1) holds. If $k' > 4$ then using $k\log k = \frac{21b}{\mu(A,B)} \log\frac{n}{b}$, we have:

$$e(A,B) \leq \frac{21}{\mu(A,B)\log k} |B|\log\frac{n}{|B|} \mu(A,B),$$

equivalently:

$$e(A,B) \log(k) \leq 21|B| \log\frac{n}{|B|}.$$

As k upper bounds $\frac{e(A,B)}{\mu(A,B)}$, we get:

$$e(A,B) \log\left(\frac{e(A,B)}{\mu(A,B)}\right) \leq 21|B| \log\frac{n}{|B|}.$$

□

Remark: We proved that a random graph has the discrepancy property with probability of at least $1 - 1/n$. We can increase this probability to $1 - 1/n^c$ by increasing the constant c_3 of the discrepancy property to be larger than 21.

Lemma 2.12. *The function $x\log\frac{n}{x}$ is increasing in $[1, \frac{n}{e}]$.*

Proof.

$$(x\log\frac{n}{x})' = \log\frac{n}{x} + x\cdot\frac{x}{n}\cdot(-\frac{n}{x^2}) = \log\frac{n}{x} - 1.$$

The derivative is greater than 0 if $\frac{n}{x} > e$, which is true for $x \in [1, \frac{n}{e})$. □

2.3 Proof of theorem 1.1

Proof. There are 4 places in the proof where we used a bound on the probability of a bad event:

- To replace v_1 by the "all ones" vector we used a bound on the norm of the vector δ (Lemma 2.2). The complement event in this case has a sub-exponential small probability in n, d.
- To bound the contribution of light couples we proved a concentration result (Claim 2.7). The probability of the complementary event is exponentially small in n.
- The bounded degree property (Section 2.2.2). The probability of not having this property is only polynomially small (for $d = c \log n$, see proof sketch at 2.2.2).
- The discrepancy property (Section 2.2.5). The probability that the graph does not have this property is polynomially small (see the remark at the end of Section 2.2.5).

The probability of the last two events can be made as small as $1/n^c$ by (linearly) increasing the constants c_1, c_2, c_3 involved in them. The constant c' in the bound $c'\sqrt{np}$ depends linearly on c_1, c_2, c_3 (see equation (6)). $\square$

3 The case $p = \frac{d_0}{n}$ for constant d_0

For this value of p the graph contains many disjoint stars of size $\Omega(\frac{\log n}{\log \log n})$ (an l size star is vertex who has l neighbors which form an independent set). The spectrum of an l star is $\{\sqrt{l-1}, 0, -\sqrt{l-1}\}$. It can then be shown that the graph will have many eigenvalues (both positive and negative) of absolute value at least $\Omega(\sqrt{\frac{\log n}{\log \log n}})$. The proof of Section 2 fails in this case because the bounded degree property of Section 2.2.2 is not valid any more. This property is extensively used in bounding the contribution of the heavy couples. The largest eigenvalues of a $\Omega(\frac{\log n}{\log \log n})$ star are induced by eigenvectors which use heavy couples.

In this section we prove theorem 1.2. We will now make some modifications to the proof of Section 2 so that it applies to the induced subgraph. We call a vertex *bad* if it has a degree greater than $(1+\epsilon)d$; we call an edge *bad* if it contains a bad vertex. The probability for a vertex to be bad is at most $e^{-\epsilon^2 d/3}$. The probability for an edge to be bad, is bounded by $2e^{-\epsilon^2 d/3}$. The fraction of bad vertices is denoted by ε. The expected number of bad vertices is bounded by $ne^{-\epsilon^2 d/3}$. Using Markov's inequality we get that $\varepsilon < e^{-\epsilon^2 d/6}$ with probability of at most $e^{-\epsilon^2 d/6}$. In this section we will use $(1+\epsilon)d$ (for $3\sqrt{\log d/d} < \epsilon < 0.9$) as the degree bound instead of $c_1 d$. We limit $\epsilon < 0.9$ only for convenience. Similar results can be obtained also for $\epsilon \geq 0.9$.

3.1 Fixing the first eigenvector

We need to show a lemma equivalent to Lemma 2.2 for the induced subgraph G'.

Lemma 3.1. *Let G be a random graph in which every edge is chosen with probability of d/n. Let G' be the graph induced by the vertices whose degree is bounded by $(1+\epsilon)d$ (where $3\sqrt{\log d/d} < \epsilon < 0.9$). We denote by A' the $n' \times n'$ adjacency matrix of G' ($n' \leq n$); by $\bar{d}$ we denote the average degree in G'. With probability of $1 - 3e^{-\epsilon^2 d/10}$ the matrix A' has the following property:*

$$\forall x \in R^{n'}, x \perp u \quad |x^t A' u| \leq 3\sqrt{\bar{d}}\, \|x\|$$

Proof. We use $\deg_v^G, \deg_v^{G'}$ to denote the degree of v in G, G' respectively. Denote by $\tilde{d}$ the average degree in G. With probability of at least $1 - 3e^{-\epsilon^2 d/6}$ all the following properties hold: $n' \geq (1 - e^{-\epsilon^2 d/6})n$, $|E(G) \setminus E(G')| < 2e^{-\epsilon^2/6}nd$, both $\tilde{d}, \bar{d}$ are in $(1 \pm 0.1)d$. For each $v \in G'$ denote $\delta'_v = \deg_v^{G'} - \tilde{d}$. The degree of a vertex v in G' is just $\tilde{d} + \delta'_v$.

$$x^t A' u = x^t \begin{bmatrix} \tilde{d} + \delta'_1 \\ \tilde{d} + \delta'_2 \\ . \\ . \\ \tilde{d} + \delta'_{n'} \end{bmatrix} \frac{1}{\sqrt{n'}} = x^t \begin{bmatrix} \delta'_1 \\ \delta'_2 \\ . \\ . \\ \delta'_{n'} \end{bmatrix} \frac{1}{\sqrt{n'}} + d x^t u = x^t \begin{bmatrix} \delta'_1 \\ \delta'_2 \\ . \\ . \\ \delta'_{n'} \end{bmatrix} \frac{1}{\sqrt{n'}} \ ,$$

hence $|x^t A' u| \leq \frac{\|x\| \|\delta'\|}{\sqrt{n'}}$. We will now show that $\|\delta'\|^2 = \sum_{v \in G'} (\deg_v^{G'} - \tilde{d})^2 < 9n'\bar{d}$. By Lemma 5.1 with probability of at least $1 - e^{-\Omega(nd)^{1/3}}$ it holds that

$$\sum_{v \in G'} (\deg_v^G - \tilde{d})^2 \leq 2n\tilde{d}.$$

We claim that if $\Delta \geq \max\{|\deg_v^G - \tilde{d}|, |\deg_v^{G'} - \tilde{d}|\}$ for any $v \in G'$ then

$$|\sum_{v \in G'} (\deg_v^G - \tilde{d})^2 - \sum_{v \in G'} (\deg_v^{G'} - \tilde{d})^2| \leq 2\Delta |E(G) \setminus E(G')|. \tag{9}$$

The explanation is as follows. We remove the edges of $E(G) \setminus E(G')$ from G one by one. This process induces a sequence of graphs $G = G_0, G_1, ..., G_k$ where G_k it the union of G' and a set of isolated vertices $V(G) \setminus V(G')$. For each of the graphs G_i we let $f(G_i) = \sum_{v \in G'} (\deg_v^{G_i} - \tilde{d})^2$ (where $\tilde{d}$ is fixed to be the average degree in $G_0 = G$). Note that the term from (9) that we wish to bound is exactly $|f(G_0) - f(G_k)|$. The difference between G_i and G_{i+1} is some edge (u, v) where $u \in V(G) \setminus V(G')$ and $v \in V(G')$. Using the inequality $|x^2 - (x-1)^2| \leq 2\max\{|x|, |x-1|\}$ valid for any integer x we derive

$$|f(G_i) - f(G_{i+1})| \leq 2\max\{|\deg_v^{G_i} - \tilde{d}|, |\deg_v^{G_{i+1}} - \tilde{d}|\} \leq 2\Delta.$$

Assuming $\tilde{d} \in (1 \pm 0.1)d$ and since the maximum degree in G of a vertex $v \in G'$ is most $(1+\epsilon)d < 2d$ we may use $\Delta = 2d$ as an upper bound. Assuming also $|E(G) \setminus E(G')| \leq 2e^{-\epsilon^2 d/6}nd$ we derive

$$\|\delta\|^2 = \sum_{v \in G'} (\deg_v^{G'} - \tilde{d})^2 \leq 2n\tilde{d} + 4de^{-\epsilon^2 d/6}nd < 9n'\bar{d},$$

where in the last inequality we used $\bar{d}, \tilde{d} \in (1 \pm 0.1)d$, $n' > 0.9n$. $\square$

3.2 Light Couples

Let A' be the minor of A induced by removing the high degree vertices. A' has at least $n' = (1-\varepsilon)n$ vertices ($\varepsilon \leq e^{-\epsilon^2 d/6}$) with probability $1 - e^{-\epsilon^2 d/6}$. In the context of A', the light couples are those whose multiplication is bounded in absolute value by $\frac{\sqrt{d}}{n'}$. We would like to argue that A' has the following property:

There exists a universal constant c such that for every two vectors $x, y \in R^{n'}$ the light couples of x, y contribute at most $c\sqrt{d}$ to $xA'y$. If A' has this property we say that A' is *c-light good.*

Lemma 3.2. *A fixed minor A' of size $n' \geq (1-\varepsilon)n$ is c-light good with probability of at least $1 - e^{(c/4 - \ln(9/\delta))n}$.*

Proof. Let A' be an arbitrary minor of size $n' \geq (1-\varepsilon)n$. Every edge in A' is chosen with probability of $\frac{d}{n} = \frac{d'}{n'}$ ($d \leq d' \leq \frac{d}{1-\varepsilon}$). Since $d' \leq d$ we can apply Corollary 2.10 on A' to derive:

$$\Pr[A' \text{ is not } c\text{-light good}] < e^{-(c/3 + \ln(9/\delta))n'} < e^{-(c/4 + \ln(9/\delta))n}.$$

$\square$

To simultaneously cover all relevant minors we use the union bound. There are at most $\binom{n}{\varepsilon n} \leq 2^n$ relevant minors. By Lemma 3.2:

$$\Pr[\text{all minors of size} \geq (1-\varepsilon)n \text{ are } c\text{-light good}] > 1 - 2^n e^{-(c/4+\ln(9/\delta))n} \geq 1 - e^{-(c/4+\ln(9/\delta)+1)n}.$$

3.3 Heavy Couples

We will prove that the discrepancy property is valid for the subgraph G' induced by removing the high degree vertices, with probability of $1 - \max\{e^{-\epsilon^2 d/6}, 1/n\}$. Without loss of generality we may assume that G' contains the first n' vertices. There are at most εn vertices of high degree (where $\varepsilon \leq e^{-\epsilon^2 d/6}$). Observe that the proof of case (2) of the discrepancy property is valid also when d is a fixed constant. The proof of case (1) relies on the bounded degree property. It follows that for every $A, B \subseteq [n']$ (where $|A| \leq |B|$) one of the following holds:

1. $\frac{e^G(A,B)}{\mu(A,B)} \leq ec_2$.

2. $e^G(A,B) \log \frac{e^G(A,B)}{\mu(A,B)} \leq c_3 |B| \log \frac{n}{|B|}$.

We define $d' = n' \frac{d}{n}$ so that $\frac{d'}{n'} = \frac{d}{n}$; $\mu'(A,B) = \frac{e^{G'}(A,B)}{|A||B|d'/n'}$. Using the facts that $e^{G'}(A,B) = e^G(A,B)$ and that $(1-\varepsilon)n \leq n' \leq n$ we get that for every $A, B \subseteq [n']$ ($|A| \leq |B|$):

1. $\frac{e^{G'}(A,B)}{\mu'(A,B)} \leq ec_2$.

2. $e^{G'}(A,B) \log \frac{e^{G'}(A,B)}{\mu'(A,B)} \leq 2c_3 |B| \ \log \frac{n'}{|B|}$.

Having established the relevant discrepancy properties, the proof of Section 2.2.4 goes through for G' with virtually no change. Note that in several places along this proof we will need to use the fact that the degrees of vertices in the underlying graph are bounded by $(1+\epsilon)d$ (this yields a bound of $8 \max\{16ec_2/\delta^2, 8(1+\epsilon), 16c_3\delta^2\}\sqrt{d}$ on the contribution of the heavy couples).

3.4 Proof of theorem 1.2

To get the bound $c'\sqrt{d}$ we use the union bound over the bad events

Proof. There are 4 places in the proof where we used a bound on the probability of a bad event:

- To replace v_1 by the "all ones" vector we used a bound on the norm of the vector δ (Lemma 3.1). The complement event in this case has a probability bounded by $3e^{-\epsilon^2 d/10}$.

- To bound the contribution of light couples we proved that any big enough minor is c-light good (Section 3.2). The probability of the complementary event is exponentially small in n. We also assumed that the fraction of vertices in $G \setminus G'$ is $\varepsilon \leq e^{-\epsilon^2 d/6}$. This happens with probability of $1 - e^{-\epsilon^2 d/6}$.
- The discrepancy property (Sections 3.3, 2.2.5). The probability that the graph does not have this property is polynomially small in n (see the remark at the end of Section 2.2.5).

The dominant probabilities are of order $e^{-\epsilon^2 d/10}$, thus with probability of $1 - e^{-\Omega(\epsilon^2 d)}$ it holds that $\lambda < c'\sqrt{d}$.

$\square$

3.5 Setting the degree bound to be greater than 1

Although Theorem 1.2 is stated only for the case in which ϵ is at most 0.9 (yielding a degree bound $< 1.9d$), similar result can be obtained also for larger values of ϵ. When ϵ is sufficiently large we get that $\lambda < c'\sqrt{d}$ with probability of $1 - e^{-\Omega(\epsilon d)}$. The constant c' grows linearly with ϵ.

4 Applications

For most graphs (namely, random graphs), $\lambda = O(\sqrt{d})$, where d is the average degree of the graph. In the previous sections we extended the known proofs for this also to the case where d is fairly small. Part of the significance of eigenvalues of graphs is that on the one hand, they can be computed in polynomial time, and on the other hand, their value is influenced by some properties of the graph that are NP-hard to compute. For example, if for some constant $\delta > 0$, a nearly regular graph has either a bisection with more than $(\frac{1}{2} + \delta)\frac{d}{2}n$ edges, or a bisection with less than $(\frac{1}{2} - \delta)\frac{d}{2}n$ edges, or an independent set larger than δn, then λ is no longer of the order of $\sqrt{d}$, but rather in the order of d. Hence having small λ is a certificate for the inexistence of small separators, large cuts, or very large independent sets. Likewise, if one "plants" a small bisection, a large cut, a large independent set or a coloring with a small number or colors in an otherwise random graph, this shows up in the eigenvalues of the graph, and often the corresponding eigenvectors can be used in order to retrieve the planted solution. See for example [4, 1, 2, 5, 12, 20]. Extensions of these spectral techniques to the use of semidefinite programming allows one to find planted solutions in so called *semirandom* graphs [11, 10].

Often, previous work only addressed graphs of fairly large average degree, for which bounds on λ are available through the work of [14]. The purpose of the current work is to show that most such results extend in a straightforward way to graphs with average degree as low as $\log n$, and with some extra work (that results from the need to remove vertices with exceptionally high degree), also to graphs of constant average degree. In fact, already [1] addressed graphs of constant degree (and introduced the idea of removing the vertices with highest degree), but they did not provide explicit bounds on λ, which make their results harder to use in a blackbox manner.

We illustrate our treatment of low degree graphs by two examples. One removes some unnecessary logarithmic terms from a bound of [16]. The other addresses relations between spectral properties and semidefinite programming.

4.1 Refuting Random 4-SAT

Let φ be random 4-SAT formula with m clauses and n variables. We use the random model in which each new clause is chosen independently from all possible ordered clauses. The density of φ is defined as $\frac{m}{n}$. It is known that if the density is at least c (a large enough constant) then almost surely φ is unsatisfiable. From a computational point of view one may ask "for which values of the density can we efficiently find a proof for the unsatisfiability of φ?". In [16] an efficient algorithm is given which almost surely certifies the unsatisfiability of φ when the density is at least $\Omega(n \log^7 n)$. The heart of the algorithm is a reduction which translates satisfiability of a formula φ into the existence of a large independent set in a suitable graph, namely if φ is satisfiable then the graph must have a large independent set. Thus if the graph has no large independent set then this is a proof that φ is unsatisfiable. The reduction is the following: given a formula φ we associate with it two graphs $G_t(\varphi), G_f(\varphi)$. The vertices of $G_t(\varphi)$ are all the ordered pairs of negative literals $(\bar{x_i}, \bar{x_j})$. There are $n(n-1)$ such ordered pairs. We put an edge between the vertices $(\bar{x_i}, \bar{x_j})$ and $(\bar{x_k}, \bar{x_l})$ if the clause $(\bar{x_i}, \bar{x_j}, \bar{x_k}, \bar{x_l})$ appears in φ. The graph $G_f(\varphi)$ is defined in a similar way with respect to the positive literals. Notice that if φ is satisfiable then there exists a satisfying assignment in which either half of the variables are assigned true or half of the variables are assigned false. It then follows that either $G_f(\varphi)$ or $G_t(\varphi)$ contains an independent set of size $\frac{n}{2}(\frac{n}{2}-1)$, because there cannot be an edge between $(\bar{x_i}, \bar{x_j})$ and $(\bar{x_k}, \bar{x_l})$ in G_t for variables x_i, x_j, x_k, x_l, which are assigned true. Hence either G_t or G_f contains an independent set with roughly quarter of the vertices. Each one of $G_t(\varphi), G_f(\varphi)$ is a random graph which almost surely contains $(1+O(\frac{1}{\sqrt{m}}))m/16$ edges. Given that the number of edges in $G_t(\varphi)$ equals m', the edges in $G_t(\varphi)$ are chosen independently with repetitions.

We will address a cleaner version of the reduced problem in which a random graph is taken from $G_{n,p}$ (rather than $G_{n,m}$ which is the distribution of $G_t(\varphi)$ given that the number edges is m). There is a simple reduction between the two versions, we omit the details. The problem is formalized as follows: the input is a random graph taken from $G_{n,p}$. Typically the largest independent set is almost surely of size $O(\frac{n \log d}{d})$, where $d = pn$. The goal of the algorithm is to prove that the input graph has no independent set of size $\frac{n}{4}$. We call an algorithm a *refutation heuristic* for independent set if it has the following properties:

1. For most of the graphs (i.e. almost surely) the algorithm will return a proof that the graph does not have an independent set of size $\frac{n}{4}$.
2. For a small fraction of the graphs the algorithm may return *abort*, which means that the algorithm was unable to find a proof that the graph has no large independent set (possibly because the graph does have a large independent set).

Our goal is to find a refutation heuristic for values of p as low as possible. The density of the formula φ corresponds to p. Reducing p reduces the density of φ, namely if we have a refutation heuristic for low values of p, then we have an algorithm that almost surely refutes a random formula with low density. We will show a refutation heuristic for $p = \frac{d}{n}$, for some large enough constant d. This yields an algorithm that refutes most 4-SAT formulas with at least cn^2 clauses (a simple generalization of it refutes most k-SAT formulas with at least $c2^k n^{k/2}$ clauses).

Lemma 4.1 shows that a regular graph which contains an independent set of size $\frac{n}{4}$ must have an eigenvalue as small as $-\Omega(d)$. On the other hand we know that most of the graphs are nearly regular and have $\lambda < O(\sqrt{d})$. This implies that most of the graphs have a proof for not containing large independent set (since a regular graph with small λ cannot contain big independent set).

Lemma 4.1. *Let G be a d-regular graph which contains an independent set of size αn. Then G has an eigenvalue smaller than $-\alpha d \, \frac{1}{1-\alpha}$.*

Proof. Let S be an independent set of size αn. Consider the vector in which all entries corresponding to S have value of $\alpha - 1$ and all other entries are α. The Rayleigh quotient of this vector is:

$$\frac{2E(S,\bar{S})(\alpha-1)\alpha + (2E - 2E(S,\bar{S}))\alpha^2}{|S|(\alpha-1)^2 + |\bar{S}|\alpha^2} = \frac{2\alpha nd(\alpha-1)\alpha + (nd - 2\alpha nd)\alpha^2}{\alpha n(1-2\alpha+\alpha^2) + (1-\alpha)n\alpha^2} =$$

$$\frac{2d(\alpha-1)\alpha + (d-2\alpha d)\alpha}{(1-2\alpha+\alpha^2)+(1-\alpha)\alpha} = -\alpha d \, \frac{2(1-\alpha) + (-1+2\alpha)}{1-\alpha} = -\alpha d \, \frac{1}{1-\alpha}.$$

□

For large enough c and $p = \frac{c \log n}{n}$, almost surely all degrees are in the interval $(1 \pm \epsilon)d$. A similar calculation to the one in Lemma 4.1 implies that G has eigenvalue smaller than $-\alpha d \, \frac{1-3\epsilon}{1-\alpha}$. This together with the bound $\lambda \leq O(\sqrt{d})$ implies:

Corollary 4.2. *For $p = \frac{c \log n}{n}$ there exists a refutation heuristic for independent set.*

The proof of the last corollary does not hold for $p = \frac{d}{n}$ when d is constant, as the graph becomes irregular. For the case $p = \frac{d}{n}$ we consider the subgraph G' induced by removing the vertices of degrees $\geq (1+\epsilon)d$ (ϵ is fixed to be a small constant, e.g. 0.1). With probability of at least $1 - e^{-\Omega(\epsilon^2 d)}$, it has the following properties (when d is large enough):

1. G' contains at most $(1+\epsilon)nd/2$ edges, and at least $(1 - e^{-\Omega(\epsilon^2 d)})n$ vertices.
2. $|E(G) \setminus E(G')|$ is bounded by $e^{-\Omega(\epsilon^2 d)}(1+\epsilon)nd/2$.
3. $\lambda < c\sqrt{d}$ (by Theorem 1.2).
4. $\|\delta\| \leq c\sqrt{nd}$, where $\delta_v + \tilde{d}$ is the degree of v in G', $\tilde{d}$ is the average degree in G, and the dimension of δ is the number of vertices in G' which we denote by n' (see the proof of Lemma 3.1).
5. The average degree $\tilde{d}$ of G is in $(1 \pm \frac{1}{\sqrt{n}})d$.

These properties can be verified in polynomial time. We will show that these properties imply that there is no independent set of size αn in G, where $\alpha = (\frac{2c}{\epsilon\sqrt{d}})$. Notice that $\|\delta\|_1 \leq \sqrt{n'}\|\delta\| \leq cn\sqrt{d}$. Assume by contradiction that G has an independent set of size αn. It follows that G' has an independent set S of size $(\alpha - e^{-\Omega(\epsilon^2 d)})n \geq (1-\epsilon)\alpha n$ (since $\alpha\epsilon = \frac{2c}{\sqrt{d}} \geq e^{-\Omega(\epsilon^2 d)}$). The number of edges touching S in G' is at least:

$$\sum_{v \in S}(\tilde{d} + \delta_v) - |E(G) \setminus E(G')| \geq |S|\tilde{d} - \|\delta\|_1 - |E(G) \setminus E(G')|$$

$$\geq (1-\epsilon)\alpha nd(1-o(1)) - cn\sqrt{d} - e^{-\Omega(\epsilon^2 d)}(1+\epsilon)nd/2 \geq \alpha nd(1-4\epsilon)$$

(the last inequality is true since $\alpha = \frac{2c}{\epsilon\sqrt{d}}$, $e^{-\Omega(\epsilon^2 d)} < \alpha\epsilon$). Since the number of edges in G' is bounded by $(1+\epsilon)nd/2$, we can use an argument similar to Lemma 4.1 to get that G' has an eigenvalue as small as $-\alpha d \, \frac{1-O(\epsilon)}{1+\alpha}$. This contradicts the fact that λ of G' is bounded by $c\sqrt{d}$.

Corollary 4.3. *There exists an efficient procedure which almost surely certifies that a graph taken from $G_{n,\frac{d}{n}}$ has no independent set larger than $\frac{cn}{\sqrt{d}}$ (where c is a universal constant and d is big enough constant).*

The main idea of the above refutation algorithm is to refute the existence of a large independent set in a random graph taken from $G_{n,p}$. An alternative way to do this is to compute the Lovasz ϑ function, which upper bounds the largest independent set of the graph [19]. In the cases where $c\frac{(\log n)^6}{n} < p < 1 - c\frac{(\log n)^6}{n}$ it is well known that almost surely $c'\sqrt{n(1-p)/p} \le \vartheta(G_{n,p}) \le c''\sqrt{n(1-p)/p}$. This was proved in [17] and it relies on the eigenvalue bounds given in [14]. As we extend the bound on the second largest eigenvalue given in [14] also for $p > c\frac{\log n}{n}$, this implies (together with Lemma 5.1) that the above bounds on the ϑ function (maybe with different constants c', c'') are valid also for $c\frac{\log n}{n} < p < 1 - c\frac{\log n}{n}$. Using Theorem 1.2 and Lemma 5.1 one can further extend the range also for $\frac{c}{n} < p < 1 - \frac{c}{n}$. This can be done by bounding the ϑ function of the subgraph induced by the vertices whose degrees are bounded by $(1+\epsilon)np$ (where ϵ is as in the Theorem). To bound the ϑ of the whole graph we just add the number of vertices whose degree is above $(1+\epsilon)np$, but this number is significantly smaller then ϑ of the induced subgraph. We omit the details.

The fact that random 4-SAT with at least cn^2 clauses can be efficiently refuted is also claimed independently in [7]. The refutation algorithm shown in [7] is different from ours. It is our opinion that our approach is conceptually simpler and computationally less demanding.

4.2 Max-cut and relations with semidefinite programming

Boppana [4] used spectral techniques in order to find planted bisections in random graphs. His approach was based on an eigenvalue optimization problem (which can be solved using semidefinite programming), and in its analysis he used a theorem (Theorem 4.2 in [4]) that for a random graph from $G_{n,p}$ with $p > \log n/n$, almost surely $\lambda = O(\sqrt{pn})$. No proof of this theorem is given in Boppana's paper. However, as later observed in [10], a bound of $\lambda = O(\sqrt{pn \log n})$ suffices in order to prove the correctness of Boppana's algorithm for bisection, and such a bound was indeed proved in [10] (for all values of $p > 1/n$). Our proof of Theorem 1.1 provides the missing proof for Theorem 4.2 in [4].

It turns out that for the specific problem that Boppana was studying (planted bisection in a random graph), there is no need to use the computationally heavy semidefinite machinery. The planted bisection can be found using only eigenvalue and eigenvector computations. However, a semidefinite program essentially equivalent to that used by Boppana was later used by Goemans and Williamson [15] in their approximation algorithm for Max Cut. As was later shown by Zwick [21], the approximation ratio of this semidefinite program approaches 1 as the size of the maximum cut in a graph approaches (from above) half the edges. Feige [9] observes that this implies that the semidefinite program can certify that a random graph in the $G_{n,p}$ model with $p = d/n$ for a large enough constant d, does not have a cut with significantly more than half the edges. Specifically, a random graph on n vertices G almost surely has a maximum cut not bigger than $(1/2 + \epsilon_d)|E(G)|$ (where $\epsilon_d = \frac{O(1)}{\sqrt{d}}$). The semidefinite program will give a value bounded by $(1/2 + \delta(\epsilon_d))|E(G)|$ where $\delta(e_d)$ tends to zero as e_d tends to 0. Motivated by the above, we show how one can certify that a random graph does not have a large cut, without the need to explicitly solve a semidefinite program. An added benefit of the new proof is that it places tighter bounds on the size of the maximum cut in a random graph than those one gets by using the results of [21] (the new proof yields something equivalent to $\delta(\epsilon) = O(\epsilon)$ which is best up to a constant).

Given a graph G with adjacency matrix A_G, consider the following optimization problem. *Find the minimum value of β such that there exists a vector $\alpha = (\alpha_1, ..., \alpha_n)$ for which: $\sum_i \alpha_i = \beta$ and $A^* = A_G + diag(\alpha)$ is a positive semi-definite matrix.*

The above problem can be solved using semidefinite programming. Moreover, it provides an upper bound on the size of the maximum cut in the graph as follows. Let m be the number of edges in G, and $m_1 > m/2$ be the number of edges in the maximum cut. Consider a vector v whose value is $+1$ for vertices on one side of the cut and -1 for vertices on the other side. Then $v^t A_G v = 2m - 4m_1$. On the other hand, $v^t A^* v \geq 0$, because A^* is positive semidefinite. Hence $v^t \mathrm{diag}(\alpha) v \geq 4m_1 - 2m$, implying that $\beta \geq 4m_1 - 2m$, or $m_1 \leq \frac{m}{2} + \frac{\beta}{4}$.

Now for a random graph, we upper bound β without explicitly solving the underlying semidefinite program. Notice that if $\lambda_n > -k$ then $\beta \leq nk$ since $A + kI$ is positive semidefinite. In the case that $p \geq c_0 \frac{\log n}{n}$ the matrix A has $\lambda_n > -c\sqrt{d}$ and thus $\beta < cn\sqrt{d}$. The following lemma gives a similar bound for β also for $p \geq \frac{O(1)}{n}$.

Lemma 4.4. *Let G be a random graph with $p = \frac{d}{n}$ with d a sufficiently large constant, let β be defined as above. Then for some universal constant c, $\beta \leq c\sqrt{d}n$ with probability of $1 - e^{-\Omega(d)}$.*

Proof. To prove that $\beta = O(\sqrt{d}n)$ we decompose A_G into the following matrices: $A_G = A_1 + A_2$. The matrix A_1 contains all the edges belonging to low degree vertices whereas A_2 contains all the edges which contain a high degree vertex (degree larger than $(1+\epsilon)d$, for some small fixed $\epsilon > 0$). It is enough to make each separate matrix positive semidefinite by adding values on the diagonals, and then also A will become positive semidefinite. Notice that A_1 has $\lambda = O(\sqrt{d})$, thus by adding $c\sqrt{d}$ edges on its diagonal we can make it positive semidefinite. The matrix A_2 almost surely contains at most $e^{-\Omega(\epsilon^2 d)}n$ edges, thus we can make it positive semidefinite by adding $e^{-\Omega(\epsilon^2 d)}n$ edges on its diagonal. $\square$

Summing up, to show that a random graph G does not have a cut with more than $dn/4 + O(\sqrt{d}n)$ edges, do the following:
Remove all vertices with of degree larger than $(1+\epsilon)d$, and verify that the number of edges removed is $O(\sqrt{d}n)$. Thereafter, for the remaining graph, compute λ_n and verify that $|\lambda_n| < O(\sqrt{d})$.

We remark here that the bound of $dn/2 + O(\sqrt{d}n)$ is tight up to a constant, since a random graph almost surely has a cut of size $nd/2 + \Omega(\sqrt{d}n)$. One way to show it is to select a random graph and simultaneously build the cut in the following way:
In each step we insert a new vertex to the graph, and randomly select the edges connecting it to the vertices added so far. As we add a vertex to the graph we also decide in which side of the cut to put it. We put the new vertex in the side of the cut where it has the least number neighbors.
To see that this procedure creates a cut of size $nd/2 + \Omega(n\sqrt{d})$ (w.h.p.), we consider the last $n/2$ vertices. The first $n/2$ vertices contribute to the cut at least half of the edges in their induced subgraph. For each new vertex (out of the remaining $n/2$) that we add, the expected difference between the number of its neighbors from the two sides of the cut is $\Omega(\sqrt{d})$. It follows that the expected number of edges in the cut is at least $nd/2 + \Omega(n/\sqrt{d})$.

Independently, in [8] it is shown how to upper bound the value of MAX CUT of a random sparse graph ($G_{n,p}$ for $c/n < p < o(n^{1/2})$) using semidefinite programming (in fact they show a more general result for MAX k-CUT). Specifically, it is shown that almost surely the value of the semidefinite relaxation is bounded by $n^2p/4 + cn^{3/2}p^{1/2}$. The approach in [8] involves solving a semidefinite program relaxation for MAX k-CUT.

4.3 Related work

As stated in the introduction, our main contribution is to show that the known spectral techniques work on sparser (random) instances of NP-hard problems than what was known before. As our approach is to analyze the second largest eigenvalue of a random graph, our algorithms are based on computing the eigenvalues (often, just the most negative eigenvalue) of a single matrix, rather than solving a semidefinite program. For some of the above problems, there are strong concentration results and average case algorithms (e.g. [6], [8]), and these issues are not addressed in the current paper.

5 Deviations from average degree

Lemma 5.1. *Let G be a random graph taken from $G_{n,d/n}$ where $d_0 \le d \le 0.1n^{1/5}$ for some universal constant d_0. Let $\bar{d}$ be the average degree in G and define $f(G) := \sum_{v\in V}(deg_v - \bar{d})^2$. The following concentration result holds:*

$$\Pr[f(G) > 2n\bar{d}] < e^{-\Omega((nd)^{1/3})}.$$

Proof. Denote by m the number of edges in G, note that $\bar{d} = \frac{2m}{n}$. Given that G contains exactly m edges, its distribution is $G_{n,m}$ (uniform distribution over graphs with m edges). We will use the following product measure G_n^m to produce graphs: choose m random edges independently (the resulting graph may contain parallel edges). Note that

$$\Pr_{G_n^m}[G \text{ is simple}] \ge \left(1 - \frac{m}{\binom{n}{2}}\right)^m \ge e^{-2m^2/\binom{n}{2}} \ge e^{-\bar{d}^2}.$$

Moreover, given that G is simple its distribution is $G_{n,m}$. It then follows

$$\begin{aligned}\Pr_{G_{n,m}}[\sum(\deg_v - \bar{d})^2 > 2n\bar{d}] &= \Pr_{G_n^m}[\sum(\deg_v - \bar{d})^2 > 2n\bar{d} \mid G \text{ is simple}] \\ &\le \Pr_{G_n^m}[\sum(\deg_v - \bar{d})^2 > 2n\bar{d}]/e^{-\bar{d}^2}. \end{aligned} \tag{10}$$

Define $f(v) = (\deg_v - \bar{d})^2$ (where $\deg_v, \bar{d}$ relate to some known graph). Extending f also to any graph G or any set of edges E we let $f(G) = f(E) = \sum_v f(v)$. Setting $p = \frac{n-1}{\binom{n}{2}} = \frac{2}{n}$, it holds that $\mathbb{E}[f(v)] = p(1-p)m \le \bar{d}$ (for any v) and thus $\mu = \mathbb{E}[f(G)] \le n\bar{d}$. Proving a concentration result for $f(G)$ may be hard, since adding one edge to G may change the value of f by $\Omega(n)$. For this reason we define

$$\tilde{f}(v) = \min\{f(v), (k\bar{d})^2\}, \quad \text{where } k > 10 \text{ is an integer to be fixed later.}$$

The function $\tilde{f}$ has the following three properties. Its expectation is bounded by μ (because $\tilde{f}(v) \le f(v)$). Given that the maximum degree is bounded by kd, it holds that $\tilde{f}(v) = f(v)$ (for every v). For any two configurations of edges E', E (from G_n^m) that differ in exactly one edge it holds that $|\tilde{f}(E)-\tilde{f}(E')| \leq 2[(kd)^2-(kd-1)^2] \leq 4kd$ (if the different edge is (u,v) then the difference in the sum is only in $f(v), f(u)$ and for each of them the extreme case is when the degree without the extra edge is $kd-1$). Since G is taken from a product measure and $\tilde{f}$ has bounded difference of $\Delta = 4kd$ (i.e. adding one edge to G changes $\tilde{f}(G)$ by at most $4kd$) it holds that for any $\lambda > 0$ (see [3] theorem 4.7.1)

$$\Pr_{G_n^m}[\tilde{f}(G) > \lambda + \mu] < e^{-\frac{\lambda^2}{2m\Delta^2}}, \tag{11}$$

setting $\lambda = n\bar{d}$ and using $\mu \leq n\bar{d}$ we derive

$$\Pr_{G_n^m}[\tilde{f}(G) > 2n\bar{d}] < e^{-\frac{n\bar{d}}{(4k\bar{d})^2}} = e^{-\frac{n}{(4k)^2\bar{d}}}$$

It then follows that

$$\begin{aligned}\Pr_{G_n^m}[f(G) > 2n\bar{d}] &< e^{-\frac{n}{(4k)^2\bar{d}}} + \Pr_{G_n^m}[\text{The maximum degree } > k\bar{d}] \\ &< e^{-\frac{n}{(4k)^2\bar{d}}} + ne^{-k\bar{d}} \overset{k:=n/16\bar{d}^2)^{1/3}}{<} e^{-\frac{1}{3}(n\bar{d})^{1/3}}.\end{aligned}$$

Using inequality (10), we derive

$$\Pr_{G_{n,m}}[f(G) > 2n\bar{d}] < e^{-\frac{1}{3}(n\bar{d})^{1/3}+\bar{d}^2} = \exp(-\tfrac{1}{3}(n\bar{d})^{1/3}(1-\tfrac{3\bar{d}^{5/3}}{n^{1/3}})).$$

Using $d \leq 0.1n^{1/5}$ and $\Pr_{G_{n,d/n}}[d/2 \leq \bar{d} \leq 2d] > 1-e^{-\Omega(n)}$ we derive

$$\Pr_{G_{n,d/n}}[f(G) > 2n\bar{d}] < \exp(-\Omega((nd)^{1/3})).$$

$\square$

Acknowledgements

This work was supported in part by a grant from the G.I.F., the German-Israeli Foundation for Scientific Research and Development.

References

[1] N. Alon and N. Kahale. A spectral technique for coloring random 3-colorable graphs. *SIAM Journal on Computing*, 26(6):1733–1748, 1997.

[2] N. Alon, M. Krivelevich, and B. Sudakov. Finding a large hidden clique in a random graph. *Random Structures and Algorithms*, 13(3-4):457–466, 1988.

[3] N. Alon and J. Spencer. *The Probabilistic Method*, Second edition. John Wiley and Sons, New York, NY, 2000.

[4] R. Boppana. Eigenvalues and graph bisection: An average-case analysis. In *Proceedings of the 28th Annual Symposium on Foundations of Computer Science*, pages 280–285, 1987.

[5] H. Chen and A. Frieze. Coloring bipartite hypergraphs. In *Proceedings of the 5th International Conference on Integer Programming and Combinatorial Optimization*, pages 345–358, 1996.

[6] A. Coja-Oghlan. The Lovász number of random graphs. In *Proceedings of the 7th International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 228–239, 2003.

[7] A. Coja-Oghlan, A. Goerdt, A. Lanka, and F. Schadlich. Certifying unsatisfiability of random 2k-sat formulas using approximation techniques. In *Proceedings of the 14th International Symposium on Fundamentals of Computation Theory*, pages 15–26, 2003.

[8] A. Coja-Oghlan, C. Moore, and V. Sanvalani. Max k-cut and approximating the chromatic number of random graphs. In *Proccedings of the 30th International Colloquium on Automata, Languages and Programming*, pages 200–211, 2003.

[9] U. Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 534–543, 2002.

[10] U. Feige and J. Kilian. Heuristics for semirandom graph problems. *Journal of Computing and System Sciences*, 63(4):639–671, 2001.

[11] U. Feige and R. Krauthgamer. Finding and certifying a large hidden clique in a semirandom graph. *Random Structures and Algorithms*, 16(2):195–208, 2000.

[12] A. Flaxman. A spectral technique for random satisfiable 3CNF formulas. In *Proceedings of the 14th annual ACM-SIAM symposium on Discrete algorithms*, pages 357–363, 2003.

[13] J. Friedman, J. Kahn, and E. Szemeredi. On the second eigenvalue in random regular graphs. In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, pages 587–598, 1989.

[14] Z. Furedi and J. Komlos. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[15] M. Goemans and D. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42:1115–1145, 1995.

[16] A. Goerdt and M. Krivelevich. Efficient recognition of random unsatisfiable k-SAT instances by spectral methods. In *STACS: Annual Symposium on Theoretical Aspects of Computer Science*, pages 294–304, 2001.

[17] Ferenc Juhász. The asymptotic behaviour of the Lovász theta function for random graphs. *Combinatorica*, 2(2):153–155, 1982.

[18] M. Krivelevich and B. Sudakov. The largest eigenvalue of sparse random graphs. *Combinatorics, Probability and Computing*, 12:61–72, 2003.

[19] L. Lovász. On the Shannon capcity of a graph. *IEEE Transactions on Information Theory*, IT-25:1–7, 1979.

[20] F. McSherry. Spectral partitioning of random graphs. In *Proceedings of the 42nd Anual Symposium on Foundations of Computer Science*, pages 529–537, 2001.

[21] U. Zwick. Outward rotations: a tool for rounding solutions of semidefinite programming relaxations, with applications to MAX CUT and other problems. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 679–687, 1999.