A tight bound of the largest eigenvalue of sparse random graph

Tomonori Ando

March 2009, C–262
A tight bound of the largest eigenvalue of sparse random graph

Tomonori Ando
Reserch Report C-262

Abstract

We analyze the largest eigenvalue and eigenvector for the adjacency matrices of sparse random graph. Let \( \lambda_1 \) be the largest eigenvalue of an \( n \)-vertex graph, and \( v_1 \) be its corresponding normalized eigenvector. For graphs of average degree \( d \log n \), where \( d \) is a large enough constant, we show \( \lambda_1 = d \log n + 1 + o(1) \) and \( \langle 1, v_1 \rangle \Delta (1 - \Theta \left( \frac{1}{\log n} \right)) \). It shows a limitation of the existing method of analyzing spectral algorithms for NP-hard problems.

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 largest eigenvalue of its adjacency matrix and a corresponding eigenvector.

Let \( \lambda_1 \geq \cdots \geq \lambda_n \) be eigenvalues of the adjacency matrix of \( G \), and \( \lambda = \max[\lambda_2, |\lambda_n|] \). Note that \( np \) is the average degree of \( G \). Then it is well known that \( \lambda_1 \geq np \) and \( \lambda = \Omega(\sqrt{np}) \). For \( p = \Theta(1) \), it is known that \( \lambda_1 = np + 1 - 2p + \epsilon \) with \( |\epsilon| = O \left( \frac{1}{\sqrt{n}} \right) \). This was shown by Furedi and Komlos [2]. Let \( \Delta \) be the maximum degree of \( G \). Krivelevich and Sudakov show that \( \lambda_1 = (1 + o(1)) \max\{ \sqrt{\Delta}, np \} \) [5]. This result does not depend on size of \( np \). Moreover, for \( np = \Omega(\log n) \) and \( np = O \left( n^{1/2} / (\log n)^{1/2} \right) \), it is shown that \( \lambda = O(\sqrt{np}) \) and \( \lambda_1 = np + \epsilon \) with \( |\epsilon| = O(\sqrt{np}) \) by Feige and Ofek [1]. We extend the \( \lambda_1 = np + 1 - 2p + \epsilon \) result to lower values of \( p \) in the \( G_{n,p} \) model. Let \( v_1 \) be the corresponding normalized eigenvector of \( \lambda_1 \), and define \( v_0 = \frac{1}{\sqrt{n}} \).

Theorem 1 Let \( d \) be a sufficiently large constant, and \( p = \frac{d \log n}{n} \). There exists \( c > 0 \) such that

\[
|\lambda_1 - (d \log n + 1)| \leq O \left( \frac{1}{\sqrt{d \log n}} \right)
\]
with probability at least $1 - \frac{1}{n^c}$.

**Theorem 2**  Let $d$ be a sufficiently large constant, and $p = \frac{d \log n}{n}$. There exists $c > 0$ such that

$$|\langle v_\phi, v_1 \rangle| - \left(1 - \frac{1}{2d \log n}\right) \leq O\left(\frac{1}{(d \log n)^{\frac{3}{2}}}\right)$$

with probability at least $1 - \frac{1}{n^c}$.

By the method of analyzing spectral algorithms in [3, 4], it is shown that those spectral algorithms output the wrong assignment for at most $O\left(\frac{1}{p}\right)$ variables. The method is based on the similarity between $v_\phi$ and $v_1$. Our result $\langle v_\phi, v_1 \rangle = 1 - \Theta\left(\frac{1}{\log n}\right)$ shows that the method in [3, 4] never prove that spectral algorithms output the wrong assignment for at most $o\left(\frac{1}{p}\right)$ variables for $p = \Theta\left(\frac{\log n}{n}\right)$.

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. Let $d$ be sufficiently large constant. We assume that $p = d \log n/n$. Let $G$ be a random graph taken from $G_{n,p}$ and $A$ be its adjacency matrix of $G$. Let $\lambda_1 \geq \ldots \geq \lambda_n$ be the eigenvalues of $A$, and $v_1, \ldots, v_n$ be its corresponding orthonormal eigenvectors. Define $v_\phi = \frac{1}{\sqrt{n}} \mathbf{1}$. For regular graphs $v_1 = v_\phi$. Though $G$ is nearly regular and $v_1$ is close to $v_\phi$, $v_1$ differs from $v_\phi$ slightly.

2 The analysis

To bound $\lambda_1$ and $|\langle v_\phi, v_1 \rangle|$, we use the following three properties of a random graph.

**Lemma 3**

$$\Pr\left[|v_\phi^T A v_\phi - d \log n| \geq \frac{2d \log n}{n^{\frac{1}{4}}}\right] \leq \exp\left(-2\sqrt{n}d \log n\right).$$

**Lemma 4** For every $c > 0$ there exists $k > 0$ such that

$$\Pr\left[\exists i \in \{2, \ldots, n\}, |\lambda_i| > k \sqrt{d \log n}\right] \leq \frac{1}{n^c}.$$
Lemma 5 \( f(G) \) denotes \( \sum_{v \in V} (\text{deg}_v - (n - 1)p)^2 \). Then we have

\[
\Pr \left[ |f(G) - n^2p| > 2n^2 \right] \leq 3 \exp \left( -\frac{n^2}{512} \right).
\]

It is not hard to prove Lemma 3. Since \( \frac{n}{2} v^t \phi A v_\phi \) is equal to the number of edge of \( G \), we can prove by Chernoff bound. Lemma 4 was shown in [1] by Feige and Ofek. They also showed essentially the same Lemma 5.

Since \( v_1, \ldots, v_n \) are orthonormal, there exist \( \alpha_1, \ldots, \alpha_n \) with \( \sum_{i=1}^n \alpha_i^2 = 1 \) such that \( v_\phi = \sum_{i=1}^n \alpha_i v_i \). Furthermore, there exist \( \alpha, \beta \) and \( w \) with \( \alpha^2 + \beta^2 = 1 \), \( w \perp v_\phi \) and \( \|w\| = 1 \) such that \( v_1 = \alpha v_\phi + \beta w \). Note that \( \alpha = \alpha_1 = \langle v_\phi, v_1 \rangle \); hence, we have \( \sum_{i=2}^n \alpha_i^2 = \beta^2 \).

Our goal is to show that \( \beta \) is reasonably close to 0 with high probability, but it is in fact nonnegligible with high probability. We prove this by Lemma 6 and Lemma 9.

Lemma 6 There exists \( c > 0 \) such that

\[
\beta^2 \leq \frac{1}{d \log n} + O \left( \frac{1}{(d \log n)^2} \right)
\]

with probability at least \( 1 - \frac{1}{n^c} \).

Proof Note that

\[
f(G) = n \left\| \left( A - \frac{n-1}{n} d \log n E \right) v_\phi \right\|^2.
\]

Thus from Lemma 5, we have the following with high probability

\[
\left\| \left( A - \frac{n-1}{n} d \log n E \right) v_\phi \right\|^2 \leq d \log n + 2n^{-\frac{1}{4}}.
\]

On the other hand, substituting \( v_\phi = \sum_{i=1}^n \alpha_i v_i \) and using Lemma 4, the left hand side of the above is bounded as following by

\[
\left\| \left( A - \frac{n-1}{n} d \log n E \right) v_\phi \right\|^2 = \sum_{i=1}^n \alpha_i^2 \left( \frac{n-1}{n} d \log n - \lambda_i \right)^2 \\
\geq \sum_{i=2}^n \alpha_i^2 \left( \frac{n-1}{n} d \log n - k \sqrt{d \log n} \right)^2.
\]
Hence, we have the following for some constant \( c_1 > 0 \).

\[
\sum_{i=2}^{n} \alpha_i^2 \leq \frac{d \log n + 2n^{-\frac{1}{4}}}{(\frac{n-1}{n} d \log n - k \sqrt{d \log n})^2} \\
\leq \frac{d \log n + 2n^{-\frac{1}{4}}}{(d \log n - 2k \sqrt{d \log n})^2} \\
= \frac{(d \log n - 4k \sqrt{d \log n} + 4k^2) + 4k \sqrt{d \log n} - 4k^2 + 2n^{-\frac{1}{2}}}{d \log n(\sqrt{d \log n} - 2k)^2} \\
\leq \frac{1}{d \log n} + \frac{c_1 k}{(d \log n)^{\frac{3}{2}}}.
\]

\[\square\]

**Corollary 7** There exists \( c > 0 \) such that

\[
\alpha^2 \geq 1 - \frac{1}{d \log n} - \frac{c_1 k}{(d \log n)^{\frac{3}{2}}},
\]

\[
|\alpha| \geq 1 - \frac{1}{2d \log n} - \frac{2c_1 k}{(d \log n)^{\frac{3}{2}}}
\]

and

\[
|\beta| \leq \frac{1}{\sqrt{d \log n}} + \frac{c_1 k}{2d \log n}
\]

with probability at least \( 1 - \frac{1}{n^c} \).

**Proof** Thus \( \alpha^2 + \beta^2 = 1 \), \( \alpha^2 \) is bounded. If we suppose that the second inequality or the third is not true, the first inequality conflict. \( \square \)

**Lemma 8** There exists \( c > 0 \) such that

\[
\lambda_1 \leq d \log n + 1 + O\left(\frac{1}{\sqrt{d \log n}}\right)
\]

with probability at least \( 1 - \frac{1}{n^c} \).

**Proof** We assume that \( \alpha \) is positive. The negative case can be argued similarly. Let \( w' = Av_{\phi} - d \log n v_{\phi} \). By calculating the inner product of \( Av_{\phi} = d \log n v_{\phi} + w' \) and \( v_1 \), we have

\[
\alpha \lambda_1 = \alpha d \log n + \langle v_1, w' \rangle \\
= \alpha d \log n + \langle \alpha v_{\phi}, w' \rangle + \langle \beta w, w' \rangle \\
\leq \alpha d \log n + \alpha \langle v_{\phi}, w' \rangle + |\beta| \|w\| \|w'\|.
\]

4
By lemma 3, the upper bound of $\langle v\phi, w' \rangle$ is

$$\langle v\phi, w' \rangle = v\phi^t A v\phi - d \log n \leq \frac{d \log n}{n^\frac{1}{4}}.$$  

Thus by using the bound of $|\beta|$ stated in Corollary 7, we have

$$\alpha \lambda_1 \leq ad \log n + \frac{d \log n}{n^\frac{1}{4}} + \left( \frac{1}{\sqrt{d \log n}} + \frac{c_1 k}{2d \log n} \right) \|w'\|. \quad (1)$$

Note here that $w' = Av\phi - d \log n v\phi$. Thus by using Lemma 5, we have

$$\|w'\| = \left\| Av\phi - \frac{n-1}{n} d \log n v\phi - \frac{1}{n} d \log n v\phi \right\| \leq \left\| Av\phi - \frac{n-1}{n} d \log n v\phi \right\| + \left\| \frac{1}{n} d \log n v\phi \right\| \leq \sqrt{d \log n} + n^{-\frac{1}{2}} + \frac{d \log n}{n}.$$

Substituting this to (1), we obtain

$$\alpha \lambda_1 \leq ad \log n + \frac{d \log n}{n^\frac{1}{4}} + \left( \frac{1}{\sqrt{d \log n}} + \frac{c_1 k}{2d \log n} \right) \left( \sqrt{d \log n} + n^{-\frac{1}{2}} + \frac{d \log n}{n} \right) = ad \log n + O \left( \frac{1}{\sqrt{d \log n}} \right) + \left( \frac{1}{\sqrt{d \log n}} + O \left( \frac{1}{d \log n} \right) \right) \left( \sqrt{d \log n} + O(1) \right) = ad \log n + 1 + O \left( \frac{1}{\sqrt{d \log n}} \right).$$

By Corollary 7, the lower bound of $|\alpha|$ is $1 - \frac{1}{d \log n}$. Therefore we have the following for some constant $c_2 > 0$.

$$\lambda_1 \leq d \log n + 1 + \frac{c_2 k}{\sqrt{d \log n}}.$$

□

**Lemma 9** There exists $c > 0$ such that

$$\beta^2 \geq \frac{1}{d \log n} - O \left( \frac{1}{(d \log n)^{\frac{3}{2}}} \right)$$

with probability at least $1 - \frac{1}{n^c}$. 

5
The argument is similar to the proof of Lemma 6. First by Lemma 5, we have

\[ \left\| \left( A - \frac{n-1}{n}d \log nE \right) v_\phi \right\|^2 \geq d \log n - 2n^{-\frac{1}{3}}. \]

Then using Lemma 4, the left hand side of the above is bounded as follows.

\[ \left\| \left( A - \frac{n-1}{n}d \log nE \right) v_\phi \right\|^2 \leq \alpha_1^2 \left( \frac{n-1}{n}d \log n - \lambda_1 \right)^2 + \sum_{i=2}^{n} \alpha_i^2 \left( d \log n + k \sqrt{d \log n} \right)^2. \]

Now using the bound \( \lambda_1 \leq d \log n + 1 + \frac{c_3 k}{\sqrt{d \log n}} \), we have the following for some constant \( c_3 > 0 \).

\[ \sum_{i=2}^{n} \alpha_i^2 \left( d \log n + k \sqrt{d \log n} \right)^2 \geq d \log n - 2n^{-\frac{1}{3}} - \alpha_1^2 \left( \frac{n-1}{n}d \log n - \lambda_1 \right)^2 \]

\[ \sum_{i=2}^{n} \alpha_i^2 \geq \frac{d \log n - 2n^{-\frac{1}{3}} - \alpha_1^2 \left( 1 + \frac{c_3 k}{\sqrt{d \log n}} + \frac{d \log n}{n} \right)^2}{(d \log n + k \sqrt{d \log n})^2} \]

\[ = \frac{d \log n - O(1)}{d \log n(\sqrt{d \log n} + k)^2} \]

\[ = \frac{(d \log n + 2k \sqrt{d \log n} + k^2) - 2k \sqrt{d \log n} - k^2 - O(1)}{d \log n(\sqrt{d \log n} + k)^2} \]

\[ \geq \frac{1}{d \log n} - \frac{c_3 k}{(d \log n)^\frac{3}{2}}. \]

\[ \square \]

**Corollary 10** There exists \( c > 0 \) such that

\[ \alpha^2 \leq 1 - \frac{1}{d \log n} + \frac{c_3 k}{(d \log n)^{\frac{3}{2}}}, \]

\[ |\alpha| \leq 1 - \frac{1}{2d \log n} + \frac{c_3 k}{2(d \log n)^{\frac{3}{2}}} \]

and

\[ \beta \geq \frac{1}{\sqrt{d \log n}} - \frac{c_3 k}{d \log n} \]

with probability at least \( 1 - \frac{1}{n^c} \).

**Proof** The proof of this lemma is similar to Corollary 7. Thus we can use \( 1 - \alpha^2 = (1 + |\alpha|)(1 - |\alpha|) \leq 2(1 - |\alpha|) \), the second inequality can be proved more easy. \( \square \)
Lemma 11  There exists $c > 0$ such that
\[ \lambda_1 \geq d \log n + 1 - O \left( \frac{1}{\sqrt{d \log n}} \right) \]
with probability at least \( 1 - \frac{1}{n^c} \).

**Proof** Let \( w'' \) denote \( Av_\phi - (v_\phi^t Av_\phi)v_\phi \). \( w'' \) and \( v_\phi \) are orthogonal. Thus, by computing \( \|w''\|^2 \), we derive
\[ \|Av_\phi\|^2 = (v_\phi^t Av_\phi)^2 + \|w''\|^2. \] (2)

We estimate the right hand side of the above. \( w' = Av_\phi - d \log nv_\phi \), hence,
\[ w'' + (v_\phi^t Av_\phi - d \log n)v_\phi = w'. \]

By triangle inequality,
\[ \|w''\| + \|(v_\phi^t Av_\phi - d \log n)v_\phi\| \geq \|w'\|. \]

Lemma 5 states \( \|w'\| \geq \sqrt{d \log n - n^{-\frac{1}{3}} - d \log n} \). On the other hand, since Lemma 3, \( \|(v_\phi^t Av_\phi - d \log n)v_\phi\| \leq \frac{2d \log n}{n^\frac{1}{4}} \). By Lemma 3, we have
\[ \|w''\| \geq \sqrt{d \log n - n^{-\frac{1}{3}} - d \log n} \]
\[ \geq \sqrt{d \log n - 3d \log n} \]
\[ \geq \sqrt{d \log n} - \frac{3d \log n}{n^\frac{1}{4}}. \]

Then,
\[ \|w''\|^2 \geq d \log n - \frac{6(d \log n)^2}{n^\frac{1}{4}}. \]

On the other hand, again using Lemma 3, we have \( v_\phi^t Av_\phi \geq d \log n - \frac{2d \log n}{n^\frac{1}{4}} \). Hence we have from (2) that
\[ \|Av_\phi\|^2 \geq (1 - \frac{4}{n^\frac{1}{4}})(d \log n)^2 + d \log n - \frac{6(d \log n)^2}{n^\frac{1}{4}} \]
\[ \geq (d \log n)^2 + d \log n - \frac{10(d \log n)^2}{n^\frac{1}{4}}. \] (3)

Since \( v_\phi = \sum_{i=1}^n \alpha_i v_i \), we have \( \|Av_\phi\|^2 = \sum_{i=1}^n \alpha_i^2 \lambda_i^2 \). By Lemma 4, we have \( \lambda_i \leq k^2 d \log n \) for all \( i \geq 2 \). By Lemma 6, we have \( \sum_{i=2}^n \alpha_i^2 \leq \frac{1}{d \log n} + \frac{c_1 k}{(d \log n)^2} \). So, we have
\[ \|Av_\phi\|^2 = \sum_{i=1}^n \alpha_i^2 \lambda_i^2 \]
\[ \leq \alpha_1^2 \lambda_1^2 + k^2 d \log n \left( \frac{1}{d \log n} + \frac{c_1 k}{(d \log n)^2} \right). \] (4)
Hence we have from (3) and (4) that
\[
\alpha^2 \lambda_1^2 \geq (d \log n)^2 + d \log n - \frac{10(d \log n)^2}{n^4} - k^2 d \log n \left( \frac{1}{d \log n} + \frac{c_1 k}{(d \log n)^{\frac{3}{2}}} \right)
\]
\[
\geq (d \log n)^2 + d \log n - 2k^2.
\]
By Corollary 10, we have
\[
\lambda_1^2 \geq \frac{(d \log n)^2 + d \log n - 2k^2}{1 - \frac{1}{d \log n} + \frac{c_1 k}{(d \log n)^{\frac{3}{2}}}}
\]
\[
= \frac{((d \log n)^2 - d \log n + c_3 k \sqrt{d \log n}) + 2d \log n - c_3 k \sqrt{d \log n} - 2k^2}{1 - \frac{1}{d \log n} + \frac{c_1 k}{(d \log n)^{\frac{3}{2}}}}
\]
\[
= (d \log n)^2 + \left( \frac{2d \log n - 2 + 2c_3 k}{\sqrt{d \log n}} \right) + 2 - \frac{2c_3 k}{\sqrt{d \log n}} - c_3 k \sqrt{d \log n} - 2k^2
\]
\[
= (d \log n)^2 + 2d \log n - \Theta(\sqrt{d \log n}) \Theta(1).
\]
For some constant \( c_4 > 0 \),
\[
\lambda_1^2 \geq (d \log n)^2 + 2d \log n - c_4 k \sqrt{d \log n}.
\]
Therefore,
\[
\lambda_1 \geq d \log n + 1 - \frac{c_4 k}{\sqrt{d \log n}}.
\]
\(\square\)

### 2.1 Proof of Lemma 5

Let \( \text{deg}_i = \sum_{j=1}^{n} a_{i,j} \) and \( f(G) = \sum_{i=1}^{n} (\text{deg}_i - (n - 1)p)^2 \). In this section, we prove \( f(G) \) is close to its expectation. The lemma is proved in Section 5 of [1]. We give a strict proof of the lemma. Let \( d_i^* = \sum_{j=1}^{n} a_{j,i+j} \). Let \( a'_{j,i+j} \) be
\[
a'_{j,i+j} = \begin{cases} 
    a_{j,i+j} & (d_i^* \leq knp) \\
    0 & (d_i^* > knp)
\end{cases}
\]
and \( \text{deg}^*_i = \sum_{j=1}^{n} a'_{i,j} \). Let \( k \) be a fixed value. We assume \( k = \Omega(1) \). Let \( D_i \) be
\[
D_i = \begin{cases} 
    (\text{deg}_i^* - (n - 1)p)^2 & (\text{deg}_i^* \leq k(n - 1)p) \\
    (k(n - 1)p - (n - 1)p)^2 & (\text{deg}_i^* > k(n - 1)p)
\end{cases}
\]
and \( D = \sum_{i=1}^{n} D_i \). Note that if \( \text{deg}_i \leq k(n-1)p \) and \( d^*_i \leq knp \) for any \( i \), \( D \) is equal to \( f(G) \). Therefore, it suffices to show \( D \) is close to its expectation and \( D = f(G) \) with high probability. Let \( X_i = E[|D|d^*_1, \ldots, d^*_n|] \) for \( i = 0, \ldots, \lceil \frac{n-1}{2} \rceil \). \( X_0, \ldots, X_{\lceil \frac{n-1}{2} \rceil} \) are the Doob martingale. To use Azuma-Hoeffding inequality, we calculate upper bound of \( |X_i - X_{i+1}| \). Let \( x_{j,i+1} = |E[D_j|d^*_1, \ldots, d^*_n| - E[D_j|d^*_1, \ldots, d^*_n]|] \). Note that \( |X_i - X_{i+1}| \leq \sum_{j=1}^{n} x_{j,i+1} \). If \( d^*_{i+1} \) is decided, \( D_j \) is changed by only \( a_{j,j+i+1}, a_{j,j-i,j} \). We divide the analysis of \( x_{j,i+1} \) into the three cases, and calculate upper bound of \( x_{j,i+1} \).

- Case 1: \( a_{j,j+i+1} + a_{j-j-i-1,j} = 2 \) by \( d^*_{i+1} \)
  
  Note that
  
  \[
  E[D_j|d^*_1, \ldots, d^*_n] = \sum_{d^*_{i+1}} \Pr[d^*_{i+1}|d^*_1, \ldots, d^*_n] E[D_j|d^*_1, \ldots, d^*_n, d^*_{i+1}]
  \]
  
  \[
  = \sum_{x \in \{0,1\}} \sum_{y \in \{0,1\}} \Pr[a_{j,j+i+1} = x, a_{j-j-i-1,j} = y|d^*_1, \ldots, d^*_n] E[D_j|d^*_1, \ldots, d^*_n, a_{j,j+i+1} = x, a_{j-j-i-1,j} = y]
  \]
  
  \[
  = \sum_{x \in \{0,1\}} \sum_{y \in \{0,1\}} \sum_{z = 1}^{n} \Pr[a_{j,j+i+1} = x, a_{j-j-i-1,j} = y, \text{deg}'_j = x + y + z|d^*_1, \ldots, d^*_n] E[D_j|d^*_1, \ldots, d^*_n, a_{j,j+i+1} = x, a_{j-j-i-1,j} = y, \text{deg}'_j = x + y + z].
  \]
  
  In this case, \( a_{j,j+i+1} + a_{j-j-i-1,j} = 2 \) by \( d^*_{i+1} \). Moreover, \( d^*_k \) and \( d^*_l \) are independent for any \( k \neq l \). Thus, we have for any integer \( z \)
  
  \[
  \sum_{x \in \{0,1\}} \sum_{y \in \{0,1\}} \Pr[a_{j,j+i+1} = x, a_{j-j-i-1,j} = y, \text{deg}'_j = x + y + z|d^*_1, \ldots, d^*_n]
  = \Pr[a_{j,j+i+1} = 1, a_{j-j-i-1,j} = 1, \text{deg}'_j = z + 2|d^*_1, \ldots, d^*_{i+1}].
  \]
  
  Note that for any integer \( x, y, z \)
  
  \[
  E[D_j|d^*_1, \ldots, d^*_n, a_{j,j+i+1} = x, a_{j-j-i-1,j} = y, \text{deg}'_j = x + y + z]
  = \begin{cases} 
  (x + y + z - (n-1)p)^2 & (x + y + z \leq k(n-1)p) \\
  (k(n-1)p - (n-1)p)^2 & (x + y + z > k(n-1)p)
  \end{cases}
  \]
  
  Hence we have
  
  \[
  E[D_j|d^*_1, \ldots, d^*_n, a_{j,j+i+1} = x, a_{j-j-i-1,j} = y, \text{deg}'_j = x + y + z] - E[D_j|d^*_1, \ldots, d^*_n, a_{j,j+i+1} = 1, a_{j-j-i-1,j} = 1, \text{deg}'_j = z + 2] \leq (k(n-1)p - (n-1)p)^2 - (k(n-1)p - 2 - (n-1)p)^2 \leq 4knp.
  \]
Therefore,

\[
E[D_j|d_i^*, \ldots, d_i^n] - E[D_j|d_i^*, \ldots, d_i^{n+1}] = \sum_{x \in \{0,1\}} \sum_{y \in \{0,1\}} \sum_{z=1}^n \Pr[a_{j,j+i+1} = x, a_{j-i-1,j} = y, \text{deg}'_j = x + y + z|d_i^*, \ldots, d_i^n]
\]

\[
\left( E[D_j|d_i^*, \ldots, d_i^n, a_{j,j+i+1} = x, a_{j-i-1,j} = y, \text{deg}'_j = x + y + z] - E[D_j|d_i^*, \ldots, d_i^n, a_{j,j+i+1} = 1, a_{j-i-1,j} = 1, \text{deg}'_j = z + 2] \right) \leq 4knp.
\]

- Case 2: \(a_{j,j+i+1} + a_{j-i-1,j} = 1\) by \(d_i^{n+1}\)

  \(x_{j,i+1} \leq 4knp\) is proved in a similar way to Case 1.

- Case 3: \(a_{j,j+i+1} + a_{j-i-1,j} = 0\) by \(d_i^n\),

  Note that

\[
E[D_j|d_i^*, \ldots, d_i^n] - E[D_j|d_i^*, \ldots, d_i^{n+1}] = \sum_{(x,y) \neq (0,0)} \sum_{z=1}^n \Pr[a_{j,j+i+1} = x, a_{j-i-1,j} = y, \text{deg}'_j = x + y + z|d_i^*, \ldots, d_i^n]
\]

\[
\left( E[D_j|d_i^*, \ldots, d_i^n, a_{j,j+i+1} = x, a_{j-i-1,j} = y, \text{deg}'_j = x + y + z] - E[D_j|d_i^*, \ldots, d_i^n, a_{j,j+i+1} = 0, a_{j-i-1,j} = 0, \text{deg}'_j = z] \right).
\]

Since \(\sum_{(x,y) \neq (0,0)} \Pr[a_{j,j+i+1} = x, a_{j-i-1,j} = y]\) is at most \(2p\), we have \(x_{v,i+1} \leq 8knp^2\).

Case 1 and Case 2 cause at most \(2knp\) times for each \(i\). Hence, we have

\[
|X_i - X_{i+1}| \leq \sum_{j=1}^n \left| E[D_j|d_i^*, \ldots, d_i^n] - E[D_j|d_i^*, \ldots, d_i^{n+1}] \right| \leq 2knp \cdot 4knp + n \cdot 8knp^2 \leq 16k^2n^2p^2.
\]

By Azuma-Hoeffding inequality, we have

\[
\Pr[|X_{\left\lceil \frac{n-1}{2} \right\rceil} - X_0| > \lambda] \leq 2 \exp \left(-\frac{\lambda^2}{2 \cdot \frac{n-1}{2} \cdot 256(kd \log n)^4} \right).
\]

Setting \(\lambda = n^2\), \(k = \frac{n^3}{d \log n}\), the value of \(X_0\) is

\[
X_0 = E[D] \leq E[f(G)] \leq nd \log n
\]

\[
X_0 = E[D] \geq E[f(G)] - n^3 \Pr[D \neq f(G)] \leq nd \log n - 2(d \log n)^2.
\]
Pr\[D \neq f(G)\] is at most $2 \cdot 2^{-\frac{n-1}{2}}$. Thus,

$$\Pr[|D - n d \log n| > 2n^{\frac{3}{2}}] \leq \Pr[|X_n - X_0| > n^{\frac{3}{2}}] \leq 2 \exp\left(-\frac{n^{\frac{3}{10}}}{512}\right).$$

Since $\Pr[D \neq f(G)] \leq 2 \cdot 2^{-\frac{n-1}{2}}$, we have $\Pr\left[|f(G) - n^2 p| > 2n^{\frac{3}{2}}\right] \leq 3 \exp\left(-\frac{n^{\frac{3}{10}}}{512}\right)$.

References


