Expected values of parameters associated with the minimum rank of a graph*

Tracy Hall[†] Leslie Hogben[‡] Ryan Martin[§] Bryan Shader[¶]

February 16, 2009

Abstract

We investigate the expected value of various graph parameters associated with the minimum rank of a graph, including minimum rank/maximum nullity and related Colin de Verdière-type parameters. Let $G(v,p)$ denote the usual Erdős-Rényi random graph on v vertices with edge probability p. We obtain bounds for the expected value of the random variables $\operatorname{mr}(G(v,p))$, $\mathrm{M}(G(v,p))$, $\nu(G(v,p))$ and $\xi(G(v,p))$, which yield bounds on the average values of these parameters over all labeled graphs of order v.

Keywords. minimum rank, maximum nullity, average minimum rank, average maximum nullity, expected value, Colin de Verdière type parameter, positive semidefinite minimum rank, delta conjecture, rank, matrix, random graph, graph
AMS subject classifications. 05C50, 05C80, 15A03

1 Introduction

The set of $v \times v$ real symmetric matrices will be denoted by $\mathbb{R}^{(v)}$. For $A \in \mathbb{R}^{(v)}$, the *graph* of A, denoted $\mathcal{G}(A)$, is the graph with vertices $\{1,\dots,v\}$ and edges $\{\{i,j\} : \ a_{ij} \neq 0, 1 \leq i < j \leq v\}$. Note that the diagonal of A is ignored in determining $\mathcal{G}(A)$. The *minimum rank* of a graph G on v vertices is

$$\operatorname{mr}(G) = \min\{\operatorname{rank}(A) : A \in \mathbb{R}^{(v)}, \mathcal{G}(A) = G\}.$$

The *maximum nullity* or *maximum corank* of a graph G is

$$\mathrm{M}(G) = \min\{\operatorname{null}(A) : A \in \mathbb{R}^{(v)}, \mathcal{G}(A) = G\}.$$

Note that

$$\operatorname{mr}(G) + \mathrm{M}(G) = v.$$

Here a *graph* is a pair $G = (V(G), E(G))$, where V is the (finite, nonempty) set of vertices and E is the set of edges (an edge is a two-element subset of vertices); what we call a graph is sometimes called a simple undirected graph. We use the notation $v(G) = |V(G)|$ and $e(G) = |E(G)|$.

*Research of Hall, Hogben, Shader partially supported by American Institute of Mathematics SQuaRE, "Minimum Rank of Symmetric Matrices described by a Graph." Martin's research partially supported by NSA Grant H-98230-08-1-0015.

[†]Department of Mathematics, Brigham Young University, Provo UT 84602 (h.tracy@gmail.com).

[‡]Department of Mathematics, Iowa State University, Ames, IA 50011, USA (LHogben@iastate.edu) and American Institute of Mathematics, 360 Portage Ave, Palo Alto, CA 94306, USA (hogben@aimath.org).

[§]Department of Mathematics, Iowa State University, Ames, IA 50011, USA (rymartin@iastate.edu).

[¶]Department of Mathematics, University of Wyoming, Laramie, WY 82071, USA (bshader@uwyo.edu).

The *minimum rank problem* (of a graph, over the real numbers) is to determine $\mathrm{mr}(G)$ for any graph G. See [12] for a survey of known results and discussion of the motivation for the minimum rank problem; an extensive bibliography is also provided there. The minimum rank problem was a focus of the 2006 workshop "Spectra of families of matrices described by graphs, digraphs, and sign patterns" held at the American Institute of Mathematics [2]. One of the questions raised during the workshop was:

Question 1.1. *What is the average minimum rank of a graph on v vertices?*

Formally, we define the *average minimum rank* of graphs of order v to be the sum over all labeled graphs of order v of the minimum ranks of the graphs, divided by the number of (labeled) graphs of order v. That is,

$$\mathrm{amr}(v) = \frac{\sum_{v(G)=v} \mathrm{mr}(G)}{2^{\binom{v}{2}}}.$$

Let $G(v,p)$ denote the Erdős-Rényi random graph on v vertices with edge probability p. That is, every pair of vertices is adjacent, independently, with probability p. Note that $G(v,1/2)$ gives that every labeled v-vertex graph is equally likely (each labeled graph is chosen with probability $2^{-\binom{v}{2}}$), so

$$\mathrm{amr}(v) = E\left[\mathrm{mr}(G(v,1/2))\right].$$

Our goal in this paper is to determine statistics about the random variable $\mathrm{mr}(G(v,p))$ and other related parameters. We highlight the two main results of this paper by focusing on the $p = 1/2$ case:

Theorem 1.2. *Given* $\mathrm{amr}(v) = E\left[\mathrm{mr}(G(v,1/2))\right]$*, then for v sufficiently large,*

1. $|\mathrm{mr}(G(v,1/2)) - \mathrm{amr}(v)| < \sqrt{v \ln \ln v}$ *with probability approaching 1 as $v \to \infty$, and*
2. $0.146907v < \mathrm{amr}(v) < 0.5v + \sqrt{7v \ln v}$.

In general, we show that the random variable $\mathrm{mr}(G(v,p))$ is tightly concentrated around its mean (Section 2), and establish lower and upper bounds for its expected value in Sections 4 and 5. We also establish lower and upper bounds on the Colin de Verdière type parameters $\nu(G)$ and $\xi(G)$, which are related to $\mathrm{M}(G)$, in Section 6, where the definitions of these parameters are given, and Section 7. These bounds are used in Section 8 to establish bounds on the expected value of the random variable $\xi(G(v,p))$. The upper bound on $\xi(G(v,p))$ may lead to a better upper bound on the expected value of $\mathrm{M}(G(v,p))$ and hence a better lower bound on the expected value of $\mathrm{mr}(G(v,p))$.

2 Tight concentration of expected minimum rank

Although we are unable to determine precisely the mean of $\mathrm{mr}(G(v,p))$, in this section we show that this random variable is tightly concentrated around its mean, and thus average minimum rank is tightly concentrated about its mean.

A *martingale* is a sequence of random variables $X_0, \ldots, X_{v-1}$ such that

$$\mathbf{E}[X_{i+1} | X_i, X_{i-1}, \ldots, X_1] = X_i.$$

The martingale we use is the vertex exposure martingale (as described on pages 94-95 of [1]) for the graph parameter $f(G) = \frac{1}{2}\,\mathrm{mr}(G)$ (the factor $\frac{1}{2}$ is needed because deletion of a vertex may change the minimum rank by 2; see Corollary 2.3 below). $G(v,p)$ is sampled to obtain a specific graph H, and X_i is the expected value of the graph parameter $f(G) = \frac{1}{2}\operatorname{mr}(G)$ when the neighbors of vertices $v_1, \dots, v_i$ are known. Since nothing is known for X_0, $X_0 = \mathbf{E}[f(G(v,p))] = \frac{1}{2}\mathbf{E}[\operatorname{mr}(G(v,p))]$. Since the entire graph H is revealed at stage $v-1$, $X_{v-1} = \frac{1}{2}\operatorname{mr}(H)$.

The method for showing tight concentration uses Azuma's inequality for martingales (see Section 7.2 of [1]) and was pioneered by Shamir and Spencer [18]. The following corollary of Azuma's inequality is used.

Theorem 2.1. [1, Corollary 7.2.2] *Let $b = X_0, \dots, X_v$ be a martingale with*

$$|X_{i+1} - X_i| \leq 1$$

for all $0 \leq i \leq v$. Then

$$\Pr[|X_v - b| > \beta\sqrt{v}] < 2e^{-\beta^2/2}.$$

The proof that derives the tight concentration of the chromatic number of the random graph [1, Theorem 7.2.4] from [1, Corollary 7.2.2] via the vertex exposure martingale remains valid for any graph parameter $f(G)$ such that when G and H differ only in the exposure of a single vertex, then $|f(G) - f(H)| \leq 1$.

Theorem 2.2. *Let $p \in (0,1)$. Let f be a graph invariant such that for any graphs G and H, if $x \in V(G) = V(H)$ and $G - x = H - x$, then $|f(G) - f(H)| \leq 1$. Let $\mu = E\left[f(G(v,p))\right]$. Then, for any $\beta > 0$,*

$$\Pr\left[|f(G(v,p)) - \mu| > \beta\sqrt{v-1}\right] < 2e^{-\beta^2/2}.$$

Corollary 2.3. *Let $p \in (0,1)$ be fixed and $\mu = E\left[\operatorname{mr}(G(v,p))\right]$. For any $\beta > 0$,*

$$\Pr\left[|\operatorname{mr}(G(v,p)) - \mu| > 2\beta\sqrt{v-1}\right] < 2e^{-\beta^2/2}.$$

In particular,

$$|\operatorname{mr}(G(v,p)) - \mu| < \sqrt{v \ln\ln v}$$

with probability approaching 1 as $v \to \infty$.

Proof. It is well-known that for any graph G and any vertex $x \in V(G)$, $0 \leq \operatorname{mr}(G) - \operatorname{mr}(G-x) \leq 2$. Thus if $V(H) = V(G)$ and $G - x = H - x$, then $|\operatorname{mr}(G) - \operatorname{mr}(H)| \leq 2$. For the first statement, apply Theorem 2.2 with $f(G) = \frac{1}{2}\operatorname{mr}(G)$. For the second statement, let $\beta = \frac{1}{2}\sqrt{\ln\ln v}$ and conclude

$$\Pr\left[|\operatorname{mr}(G(v,p)) - \mu| > \sqrt{v\ln\ln v}\right] < \left(\frac{1}{\ln v}\right)^{1/8}.$$

□

Note that Corollary 2.3 gives the result in Theorem 1.2(1).

3 Observations on parameters of random graphs

Large deviation bounds easily show that the degree sequence of the random graph is tightly concentrated. In this section, we provide some well-known results that will be used later. The version of the Chernoff-Hoeffding bound that we use is given in [1].

Theorem 3.1. [1, Theorem A.1.16] *Let X_i, $1 \leq i \leq n$, be mutually independent random variables with all $\mathbf{E}[X_i] = 0$ and all $|X_i| \leq 1$. Set $S = X_1 + \cdots + X_n$. Then for any $a > 0$,*

$$\Pr[S > a] < \exp\{-a^2/(2n)\}.$$

It is well-known that Theorem 3.1 can be applied to the number of edges in a random graph:

Theorem 3.2. *Let p be fixed and let G be distributed according to $G(v, p)$. Then,*

$$e(G) \leq p\binom{v}{2} + v\sqrt{2\ln v},$$

with probability at least $1-v^{-2}$. In addition, $e(G) \geq p\binom{v}{2} - v\sqrt{2\ln v}$ with probability at least $1-v^{-2}$.

Proof. Let G be distributed according to the random variable $G(v, p)$. We may regard $\{\{x, y\} \in E(G) : x \neq y\}$ to be $\binom{v}{2}$ mutually independent indicator random variables. Subtract p from each and they become random variables with mean 0 and magnitude at most 1. Using Theorem 3.1, we see that $\Pr\left[e(G) - p\binom{v}{2} > a\right] < \exp\left\{-a^2/\left(2\binom{v}{2}\right)\right\}$.

Choose $a = v\sqrt{2\ln v}$; we see that

$$e(G) - p\binom{v}{2} \leq v\sqrt{2\ln v},$$

with probability at least $1 - v^{-2}$. By multiplying the random variables above by -1, we obtain

$$e(G) - p\binom{v}{2} \geq -v\sqrt{2\ln v},$$

with probability at least $1 - v^{-2}$. $\square$

Theorem 3.1 can also be applied to the neighborhood of each vertex:

Theorem 3.3. *Let p be fixed and let G be distributed according to $G(v, p)$. Then,*

$$pv - \sqrt{6v\ln v} \leq \delta(G) \leq \Delta(G) \leq pv + \sqrt{6v\ln v}$$

with probability at least $1 - 2v^{-2}$.

Proof. Let G be distributed according to the random variable $G(v, p)$. For each $x \in V(G)$, we may regard $\{\{x, y\} \in E(G) : y \neq x\}$ to be $v - 1$ mutually independent indicator random variables. Using Theorem 3.1, we see that $\Pr\left[|\deg(x) - p(v-1)| > a\right] < 2\exp\{-a^2/(2(v-1))\}$.

Thus, the probability that there exists a vertex with degree that deviates by more than a from $p(v - 1)$ is at most

$$v \times 2\exp\{-a^2/(2(v-1))\}.$$

Choose $a = \sqrt{6v\ln v}$ and we see that, simultaneously for all $x \in V(G)$,

$$|\deg(x) - pv| \leq \sqrt{6v\ln v},$$

with probability at least $1 - 2v^{-2}$. $\square$

4 A lower bound for expected minimum rank

In this section we show that if v is sufficiently large, then the expected value of $\mathrm{mr}(G(v,p))$ is at least $c(p)v+o(v)$, where $c(p)$ is the solution to equation (1) below. In the case $p=1/2$, $c(p)\approx 0.1469077$, so the average minimum rank is greater than $0.146907v$ for v sufficiently large.

The zero-pattern $\zeta(\mathbf{x})$ of the real vector $\mathbf{x}=(x_1,\dots,x_\ell)$ is the $(0,*)$-vector obtained from $\mathbf{x}$ by replacing its nonzero entries by $*$. The *support* of the zero pattern $\mathbf{z}=(z_1,\dots,z_\ell)$ is the set $S(\mathbf{z})=\{i : z_i\neq 0\}$. We will modify the proof of Theorem 4.1 from [17] to obtain the following result.

Theorem 4.1. *If $\mathbf{f}(\mathbf{x})=(f_1(\mathbf{x}),f_2(\mathbf{x}),\dots,f_m(\mathbf{x}))$ is an m-tuple of polynomials in n variables over a field F with $m\geq n$, each f_i of degree at most d, then the number of zero-patterns $\mathbf{z}=\zeta(\mathbf{f}(\mathbf{x}))$ with $|S(\mathbf{z})|\leq s$ is at most*

$$\binom{n+sd}{n}.$$

Proof. We follow the proof in [17]. Assume that the m-tuple $\mathbf{f}=(f_1,\dots,f_m)$ of polynomials over field F has the M zero-patterns $\mathbf{z}_1,\dots,\mathbf{z}_M$. Choose $\mathbf{u}_1,\dots,\mathbf{u}_M\in F^n$ such that $\zeta(\mathbf{f}(\mathbf{u}_i))=\mathbf{z}_i$.

Set

$$g_i=\prod_{k\in S(\mathbf{z}_i)} f_k.$$

Note that

$$g_i(\mathbf{u}_j)\neq 0 \qquad \text{if and only if} \qquad S(\mathbf{z}_i)\subseteq S(\mathbf{z}_j).$$

We show that polynomials $g_1,\dots,g_M$ are linearly independent. Assume on the contrary that there is a nontrivial linear combination $\sum_{i=1}^M \beta_i g_i=0$, where each $\beta_i\in F$. Let j be a subscript such that $|S(\mathbf{z}_j)|$ is minimal among the $S(\mathbf{z}_i)$ with $\beta_i\neq 0$, so for every i such that $i\neq j$ and $\beta_i\neq 0$, $S(\mathbf{z}_i)\not\subseteq S(\mathbf{z}_j)$. So substituting $\mathbf{u}_j$ into the linear combination gives $\beta_j g_j(\mathbf{u}_j)=0$, a contradiction.

Thus, $g_1,\dots,g_M$ are linearly independent over F. Each g_i has degree at most sd and the dimension of the space of polynomials of degree $\leq D$ is exactly $\binom{n+D}{n}$. $\square$

By Sylvester's Law of Inertia, every real symmetric $v\times v$ matrix of rank at most r can be expressed in the form X^TD_iX for some i such that $0\leq i\leq r$, where $D_i=\mathrm{diag}(1,\dots,1,-1,\dots,-1)$ is an $r\times r$ diagonal matrix with i diagonal entries equal to 1 and $r-i$ equal to -1 and X is an $r\times v$ real matrix. There are $r+1$ diagonal matrices D_i. Let each entry of X be a variable; the total number of variables is rv and each entry of the matrix X^TD_iX is a polynomial of degree at most 2.

Let $c(p)$ be the solution to

$$\frac{(c+p)^{2c+2p}}{(c)^{2c}(p)^{2p}}p^p(1-p)^{(1-p)}=1 \tag{1}$$

for a fixed value of p. The values of $c(p), 0<p<1$ are graphed in Figure 1.

Theorem 4.2. *Let G be distributed according to $G(v,p)$ for a fixed p, $0<p<1$. For any $c<c(p)$, the expectation $\mathbf{E}[\mathrm{mr}(G)]$ satisfies*

$$\mathbf{E}[\mathrm{mr}(G)]>cv$$

for v sufficiently large.

Furthermore, for any such c, $\Pr[\mathrm{mr}(G(v,p))\leq cv]\to 0$ as $v\to\infty$.

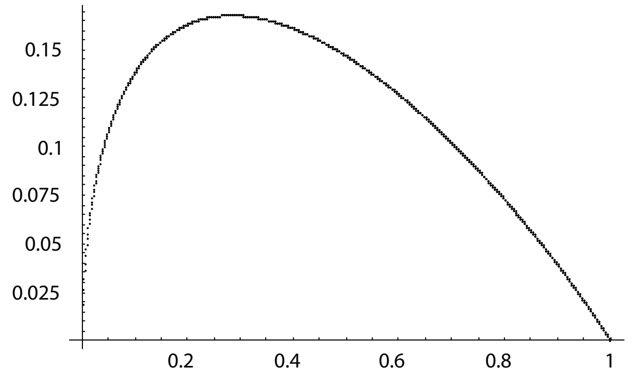

Figure 1: The graph of $c(p)$

Proof. Let G be distributed according to $G(v,p)$. Let $\mathcal{E}$ be the event that $\left|e(G) - p\binom{v}{2}\right| \le v\sqrt{2\ln v}$. By the law of total expectation,

$$\begin{aligned}
\mathbf{E}[\mathrm{mr}(G)] &= \mathbf{E}[\mathrm{mr}(G) \mid \mathcal{E}] \Pr[\mathcal{E}] + \mathbf{E}[\mathrm{mr}(G) \mid \mathcal{E}^c] \Pr[\mathcal{E}^c] \\
&\ge \mathbf{E}[\mathrm{mr}(G) \mid \mathcal{E}] \Pr[\mathcal{E}] \\
&\ge (r+1) \Pr[\mathrm{mr}(G) > r \mid \mathcal{E}] \Pr[\mathcal{E}] \\
&\ge (r+1)\,(1 - \Pr[\mathrm{mr}(G) \le r \mid \mathcal{E}])\,(1 - \Pr[\mathcal{E}^c]) \\
&\ge (r+1) - (r+1)\Pr[\mathrm{mr}(G) \le r \mid \mathcal{E}] - (r+1)\Pr[\mathcal{E}^c] \\
&\ge (r+1) - v\Pr[\mathrm{mr}(G) \le r \mid \mathcal{E}] - v\Pr[\mathcal{E}^c]
\end{aligned}$$

Theorem 3.2 shows that $v\Pr[\mathcal{E}^c] \le v^{-1}$. It remains to bound $\Pr[\mathrm{mr}(G) \le r \mid \mathcal{E}]$.

$$\begin{aligned}
\Pr[\mathrm{mr}(G) \le r \mid \mathcal{E}] &= \sum_{\substack{G:\, v(G)=v,\, \mathrm{mr}(G) \le r \\ \left|e(G) - p\binom{v}{2}\right| \le v\sqrt{2\ln v}}} \Pr[G \in G(v,p)] \\
&= \sum_{\substack{G:\, v(G)=v,\, \mathrm{mr}(G) \le r \\ \left|e(G) - p\binom{v}{2}\right| \le v\sqrt{2\ln v}}} p^{e(G)}(1-p)^{\binom{v}{2} - e(G)} \\
&= \sum_{\substack{G:\, v(G)=v,\, \mathrm{mr}(G) \le r \\ \left|e(G) - p\binom{v}{2}\right| \le v\sqrt{2\ln v}}} \left(\frac{p}{1-p}\right)^{e(G)} (1-p)^{\binom{v}{2}}
\end{aligned}$$

If $p < 1/2$, then we use a lower bound for $e(G)$, given $\mathcal{E}$; if $p > 1/2$, an upper bound. So, we can bound the term inside the summation as

$$\left(\frac{p}{1-p}\right)^{e(G)} (1-p)^{\binom{v}{2}} \le \left(\frac{\max\{p, 1-p\}}{\min\{p, 1-p\}}\right)^{v\sqrt{2\ln v}} \left(p^p (1-p)^{(1-p)}\right)^{\binom{v}{2}}.$$

Hence,

$$\begin{aligned}
&\Pr[\mathrm{mr}(G) \le r \mid \mathcal{E}] \\
&\le \left(\frac{\max\{p, 1-p\}}{\min\{p, 1-p\}}\right)^{v\sqrt{2\ln v}} \left(p^p (1-p)^{(1-p)}\right)^{\binom{v}{2}} \left|\left\{G : v(G) = v, \left|e(G) - p\binom{v}{2}\right| \le v\sqrt{2\ln v}, \mathrm{mr}(G) \le r\right\}\right|.
\end{aligned}$$

The number of v vertex graphs with between $p\binom{v}{2} - v\sqrt{2\ln v}$ and $p\binom{v}{2} + v\sqrt{2\ln v}$ edges and minimum rank at most r is at most the number of $v \times v$ symmetric pattern matrices obtained as $X^T D_i X, i = 0, \ldots, r$ with X an $r \times v$ matrix for which the cardinality of the support of the superdiagonal entries is at most $p\binom{v}{2} + v\sqrt{2\ln v}$. We can apply Theorem 4.1 with $n = rv$, $d = 2$ and $s \leq p\binom{v}{2} + v\sqrt{2\ln v}$. Therefore, because there are $r+1$ diagonal matrices,

$$\begin{aligned} &\Pr[\mathrm{mr}(G) \leq r \mid \mathcal{E}] \\ &\leq \left(\frac{\max\{p, 1-p\}}{\min\{p, 1-p\}}\right)^{v\sqrt{2\ln v}} \left(p^p(1-p)^{(1-p)}\right)^{\binom{v}{2}} (r+1) \binom{rv + 2p\binom{v}{2} + 2v\sqrt{2\ln v}}{rv}. \end{aligned}$$

By Corollary A.2 in Appendix A, for fixed c and p with $r = cv$,

$$\binom{rv + 2p\binom{v}{2} + 2v\sqrt{2\ln v}}{rv} \leq \left((1 + o(1))\left(\frac{(c+p)^{c+p}}{c^c p^p}\right)\right)^{v^2}.$$

Thus

$$\Pr[\mathrm{mr}(G) \leq cv \mid \mathcal{E}] \leq \left((1 + o(1))\frac{(c+p)^{2c+2p}}{(c)^{2c}(p)^{2p}} p^p (1-p)^{(1-p)}\right)^{v^2/2}.$$

As long $c < c(p)$ and v is sufficiently large, the quantity $v \Pr[\mathrm{mr}(G) \leq r \mid \mathcal{E}]$ is less than 1, giving

$$\mathbf{E}[\mathrm{mr}(G)] \geq (r+1) - v\Pr[\mathrm{mr}(G) \leq r \mid \mathcal{E}] - v\Pr[\mathcal{E}^c] > r + 1 - o(1) \geq r.$$

Furthermore, as long as $c < c(p)$, $\Pr[\mathrm{mr}(G) \leq cv \mid \mathcal{E}] \to 0$ as $v \to \infty$, and by Theorem 4.1, $\Pr[\mathcal{E}^c] \to 0$ as $v \to \infty$. Since

$$\Pr[\mathrm{mr}(G) \leq cv] \leq \Pr[\mathrm{mr}(G) \leq cv \mid \mathcal{E}] + \Pr[\mathcal{E}^c],$$

$\Pr[\mathrm{mr}(G) \leq cv] \to 0$ as $v \to \infty$. □

Corollary 4.3. *For v sufficiently large, the average minimum rank over all labeled graphs of order v satisfies*

$$\mathrm{amr}(v) > 0.146907v.$$

Furthermore, if G is chosen at random from all labeled graphs of order v, $\Pr[\mathrm{mr}(G) \leq 0.146907v] \to 0$ *as $v \to \infty$.*

Proof. For $p = 1/2$, $\mathbf{E}[\mathrm{mr}(G)] = \mathrm{amr}(v)$ and $0.146907 < c(p)$. □

Note that Corollary 2.3 gives the lower bound in Theorem 1.2(2). We note further the lack of symmetry with respect to p. The value $c(p)$ approaches zero as p approaches zero, which is not the case with the upper bound, that we describe in the next section.

5 An upper bound for expected minimum rank

In this section we show that if v is sufficiently large, then the expected value of $\mathrm{mr}(G(v,p))$ is at most $(1-p)v + \sqrt{7v\ln v}$. Thus the average minimum rank for graphs of order v is at most $0.5v + \sqrt{7v\ln v}$.

Let $\kappa(G)$ denote the *vertex connectivity* of G. That is, if G is not complete, it is the smallest number k such that there is a set of vertices S, with $|S| = k$, for which $G - S$ is disconnected. By convention, $\kappa(K_v) = v - 1$.

Following the terminology of [14], for a graph G an *orthogonal representation of G of dimension d* is a set of vectors in $\mathbb{R}^d$, one corresponding to each vertex, such that if two vertices are nonadjacent, then their corresponding vectors are orthogonal. Every graph has an orthogonal representation in any dimension by associating the zero vector with every vertex. A *faithful orthogonal representation of G of dimension d* is a set of vectors in $\mathbb{R}^d$, one corresponding to each vertex, such that two (distinct) vertices are nonadjacent if and only if their corresponding vectors are orthogonal. Note that in the minimum rank literature, the term "orthogonal representation" is customarily used for what is here called a faithful orthogonal representation.

The following result of Lovász, Saks and Schrijver [14] (see also the note on errata, [15] Theorem 1.1) is the basis for an upper bound for minimum rank.

Theorem 5.1. [14, Corollary 1.4] *Every graph G on v vertices has a faithful orthogonal representation of dimension $v - \kappa(G)$.*

Let $\mathrm{mr}_+(G)$ denote the minimum rank among all symmetric positive semidefinite matrices A such that $\mathcal{G}(A) = G$, and let $\mathrm{M}_+(G)$ denote the maximum nullity among all such matrices. It is well known (and easy to see) that every faithful orthogonal representation of dimension d gives rise to a positive semidefinite matrix of rank d and vice versa.

Corollary 5.2. *For any graph G on v vertices,*

$$\mathrm{mr}(G) \leq \mathrm{mr}_+(G) \leq v - \kappa(G), \tag{2}$$

or equivalently,

$$\kappa(G) \leq \mathrm{M}_+(G) \leq \mathrm{M}(G). \tag{3}$$

Our proof of the upper bound on the expected value of $\mathrm{mr}(G(v, 1/2))$ uses the bound (2) and the relationship (on average) between the connectivity $\kappa(G)$ and the minimum degree $\delta(G)$. At the AIM workshop [2] it was conjectured that for any graph G, $\delta(G) \leq M(G)$, or equivalently $\mathrm{mr}(G) \leq v(G) - \delta(G)$ [9]. The conjecture was proved for bipartite graphs in [4] but remains open in general. In [14] it is reported that in 1987, Maehara made a conjecture equivalent to $\mathrm{mr}_+(G) \leq v(G) - \delta(G)$, which would imply $\mathrm{mr}(G) \leq v(G) - \delta(G)$.

Theorem 5.3. *Let G be distributed according to $G(v, p)$. For v sufficiently large, the expected value of minimum rank satisfies $\mathbf{E}[\mathrm{mr}(G)] \leq (1 - p)v + \sqrt{7v \ln v}$.*

For v sufficiently large, the average minimum rank over all labeled graphs of order v satisfies

$$\mathrm{amr}(v) \leq 0.5v + \sqrt{7v \ln v}.$$

Proof. In [8] (see also section 7.2 of [7]), Bollobás and Thomason prove that if G is distributed according to $G(v, p)$, then $\Pr[\kappa(G) = \delta(G)] \to 1$ as $v \to \infty$, without any restriction on p. Lemma B.1 in Appendix B shows that for p fixed and v large enough, $\Pr[\kappa(G) < \delta(G)] \leq 4v^{-2}$. Let $\mathcal{E}$ be the event that $\kappa(G) = \delta(G)$ and $\delta(G) \geq pv - \sqrt{6v \ln v}$. For G distributed according to $G(v, p)$, the law of total expectation gives

$$\begin{aligned}
\mathbf{E}[\kappa(G)] &= \mathbf{E}[\kappa(G) \mid \mathcal{E}] \Pr[\mathcal{E}] + \mathbf{E}[\kappa(G) \mid \mathcal{E}^c] \Pr[\mathcal{E}^c] \\
&\geq \left(pv - \sqrt{6v \ln v}\right)(1 - \Pr[\mathcal{E}^c]) \\
&\geq pv - \sqrt{6v \ln v} - v\left(\Pr[\delta(G) < pv - \sqrt{6v \ln v}] + \Pr[\kappa(G) < \delta(G)]\right).
\end{aligned}$$

We use Theorem 3.3 and the result that $v \Pr[\kappa(G) < \delta(G)] \leq 4v^{-1}$ to see that

$$\mathbf{E}\left[\kappa(G)\right] \geq pv - \sqrt{6v \ln v} - 2v^{-1} - 4v^{-1} \geq pv - \sqrt{7v \ln v},$$

for v sufficiently large. By (2), $\mathbf{E}[\mathrm{mr}(G)] \leq (1 - p)v + \sqrt{7v \ln v}$. □

Theorem 5.3 gives the upper bound in Theorem 1.2(2). Note that Theorem 5.3 actually establishes $\mathbf{E}[\mathrm{mr}_+(G)] \le (1-p)v + \sqrt{7v \ln v}$. Since $\mathrm{mr}(G) \le \mathrm{mr}_+(G)$ for any graph G, the lower bound in Theorem 4.2 in Section 4 certainly bounds $\mathbf{E}[\mathrm{mr}_+(G)]$ from below. No improvement in the lower bound results by assuming that the matrices are positive semidefinite (the only change is the replacement of the $r+1$ diagonal matrices $D_i, i = 0, ..., r$ by D_r).

6 A lower bound for $\nu(G)$ and $\xi(G)$

In this section we discuss the the Colin de Verdière type parameters $\nu(G)$ and $\xi(G)$ and show that the vertex connectivity $\kappa(G)$ bounds $\nu(G)$ and $\xi(G)$ from below. This bound has implications for the average value of ν and ξ (see Section 8).

In 1990 Colin de Verdière ([10] in English) introduced the graph parameter μ that is equal to the maximum multiplicity of eigenvalue 0 among all matrices satisfying several conditions including the Strong Arnold Hypothesis (defined below). The parameter μ, which is used to characterize planarity, is the first of several parameters that require the Strong Arnold Hypothesis and bound the maximum nullity from below (called *Colin de Verdière type parameters*). All the Colin de Verdière type parameters we discuss have been shown to be minor monotone.

A *contraction* of G is obtained by identifying two adjacent vertices of G, deleting any loops that arise in this process, and replacing any multiple edges by a single edge. A *minor* of G arises by performing a sequence of deletions of edges, deletions of isolated vertices, and/or contractions of edges. A graph parameter β is *minor monotone* if for any minor G' of G, $\beta(G') \le \beta(G)$.

A symmetric real matrix M is said to satisfy the *Strong Arnold Hypothesis (SAH)* provided there does not exist a nonzero real symmetric matrix X satisfying $AX = 0$, $A \circ X = 0$, and $I \circ X = 0$, where $\circ$ denotes the Hadamard (entrywise) product and I is the identity matrix.

The SAH is equivalent to the requirement that certain manifolds intersect transversally. Specifically, for $A = [a_{ij}] \in \mathbb{R}^{(v)}$ let

$$\mathcal{R}_A = \{B \in \mathbb{R}^{(v)} : \operatorname{rank} B = \operatorname{rank} A\},$$

and

$$\mathcal{S}_A = \{B \in \mathbb{R}^{(v)} : \mathcal{G}(B) = \mathcal{G}(A)\}.$$

Then $\mathcal{R}_A$ and $\mathcal{S}_A$ intersect transversally if and only if A satisfies the SAH (see [13]).

Another minor monotone parameter, introduced by Colin de Verdière in [11], is denoted by $\nu(G)$ and defined to be the maximum nullity among matrices A that satisfy:

1. $\mathcal{G}(A) = G$;

2. A is positive semidefinite;

3. A satisfies the Strong Arnold Hypothesis.

Clearly $\nu(G) \le \mathrm{M}_+(G)$.

The parameter $\xi(G)$ was introduced in [3] as a Colin de Verdière type parameter intended for use in computing maximum nullity and minimum rank, by removing any unnecessary restrictions while preserving minor monotonicity (the SAH seems to be necessary to obtain minor monotonicity). Define $\xi(G)$ to be the maximum multiplicity of 0 as an eigenvalue among matrices $A \in \mathbb{R}^{(v)}$ that satisfy:

- $\mathcal{G}(A) = G$.

- A satisfies the Strong Arnold Hypothesis.

Clearly, $\nu(G) \leq \xi(G) \leq \mathrm{M}(G)$. In [3] it is shown that the parameter $\xi(G)$ is minor monotone.

An orthogonal representation $\mathbf{a}_1, \dots \mathbf{a}_v \in \mathbb{R}^d$ of G is in *general position* if every subset of d vectors is linearly independent.

Theorem 6.1. *For every graph G,*

$$\kappa(G) \leq \nu(G) \leq \xi(G).$$

Proof. The proof in [14] and [15] that there is a faithful orthogonal representation of dimension $d = v - \kappa(G)$ is done by providing an algorithm to construct an orthogonal representation $\mathbf{a}_1, \dots \mathbf{a}_v \in \mathbb{R}^d$ that is in general position with probability 1 and faithful with probability 1. So there is a faithful orthogonal representation $\mathbf{a}_1, \dots \mathbf{a}_v \in \mathbb{R}^d$ in general position.

Let $A = [\mathbf{a}_1, \dots \mathbf{a}_v]$, so $\mathcal{G}(A^T A) = G$, $A^T A$ is positive semidefinite, and $\operatorname{rank} A^T A = d$. To show $\kappa(G) \leq \nu(G)$, it remains to show $A^T A$ satisfies the Strong Arnold Hypothesis. Suppose $A^T AX = 0, (A^T A) \circ X = 0$, and $I \circ X = 0$. Every column of $A^T A$ has at least $\kappa(G) = v - d$ nonzero off-diagonal entries. Fix column k and let the position of the off-diagonal zero entries in column k of $A^T A$ be $i_1, \dots, i_s$; note that $s \leq d - 1$. So every nonzero entry in column k of X must be in one of the positions $i_1, \dots, i_s$. Note that $A^T AX = 0$ implies $AX = 0$. Hence,

$$x_{i_1} \mathbf{a}_{i_1} + \cdots + x_{i_s} \mathbf{a}_{i_s} = 0.$$

Since every subset of at most d columns of A is linearly independent, $x_{i_1} = \cdots = x_{i_s} = 0$, i.e., column k of X is 0. Thus $A^T A$ satisfies the SAH. □

7 An upper bound for $\xi(G)$

In this section, we establish an upper bound on the Colin de Verdière type parameter $\xi(G)$ in terms of the number of edges of the graph.

The following bound on the Colin de Verdière number μ in terms of the number of edges $e(G)$ is given in [16] for any connected graph $G \neq K_{3,3}$:

$$e(G) \geq \frac{\mu(G)(\mu(G) + 1)}{2}.$$

We will show that for any connected graph G,

$$e(G) + b \geq \frac{\xi(G)(\xi(G) + 1)}{2}$$

where $b = 1$ if G is bipartite and $b = 0$ otherwise. This bound has implications for the average value of ξ (see Section 8).

For a manifold $\mathcal{M}$ and matrix $A \in \mathcal{M}$, let $\mathcal{T}_{\mathcal{M}_A}$ be the tangent space in $\mathbb{R}^{(v)}$ to $\mathcal{M}$ at A and let $\mathcal{N}_{\mathcal{M}_A}$ be the normal (orthogonal complement) to $\mathcal{T}_{\mathcal{M}_A}$.

Observation 7.1. [13, p. 9]

1. $\mathcal{T}_{\mathcal{S}_A} = \{B : \forall i \neq j, a_{ij} = 0 \Rightarrow b_{ij} = 0\}$.

2. $\mathcal{N}_{\mathcal{S}_A} = \{X : \forall i, x_{ii} = 0$ *and* $\forall i \neq j, a_{ij} \neq 0 \Rightarrow x_{ij} = 0\}$.

3. $\mathcal{T}_{\mathcal{R}_A} = \{WA + AW^T : W \in \mathbb{R}^{n\times n}\} = \{B \in \mathbb{R}^{(v)} : \mathbf{v}^T B\mathbf{v} = 0\ \forall \mathbf{v} \in \ker A\}$.

4. $\mathcal{N}_{\mathcal{R}_A} = \operatorname{span}(\{\mathbf{v}\mathbf{v}^T : \mathbf{v} \in \ker A\}) = \{X \in \mathbb{R}^{(v)} : AX = 0\}$.

Clearly $\dim \mathcal{T}_{\mathcal{S}_A} = e(G)+v$. These observations can also be used to provide the exact dimension of $\mathcal{N}_{\mathcal{R}_A}$ and thus of $\mathcal{T}_{\mathcal{R}_A}$.

Proposition 7.2. *Let $A \in \mathbb{R}^{(v)}$ and let $\mathbf{u}_1, \ldots, \mathbf{u}_q$ be an orthonormal basis for* $\ker A$. *Then $U = \{\mathbf{u}_i\mathbf{u}_i^T : 1 \le i \le q\} \cup \{\mathbf{u}_i\mathbf{u}_j^T + \mathbf{u}_j\mathbf{u}_i^T : 1 \le i < j \le q\}$ is a basis for* $\operatorname{span}(\{\mathbf{v}\mathbf{v}^T : \mathbf{v} \in \ker A\})$. *Thus* $\dim \mathcal{N}_{\mathcal{R}_A} = \frac{q(q+1)}{2}$.

Proof. Let $\mathcal{N} = \operatorname{span}(\{\mathbf{v}\mathbf{v}^T : \mathbf{v} \in \ker A\})$. Since $\mathbf{u}_i\mathbf{u}_j^T + \mathbf{u}_j\mathbf{u}_i^T = (\mathbf{u}_i+\mathbf{u}_j)(\mathbf{u}_i+\mathbf{u}_j)^T - \mathbf{u}_i\mathbf{u}_i^T - \mathbf{u}_j\mathbf{u}_j^T$, $U \subset \mathcal{N}$.

Show U spans $\mathcal{N}$:

$$\left(\sum_{i=1}^{q} s_i\mathbf{u}_i\right)\left(\sum_{j=1}^{q} s_j\mathbf{u}_i\right)^T = \sum_{i=1}^{q}\sum_{j=1}^{q} s_i s_j \mathbf{u}_i\mathbf{u}_j^T = \sum_{i=1}^{q} s_i^2\mathbf{u}_i\mathbf{u}_i^T + \sum_{1\le i<j\le q} s_i s_j(\mathbf{u}_i\mathbf{u}_j^T + \mathbf{u}_j\mathbf{u}_i^T)$$

Show U is linearly independent: Let $Y = \sum_{i=1}^{q} s_i\mathbf{u}_i\mathbf{u}_i^T + \sum_{1\le i<j\le q} s_{ij}(\mathbf{u}_i\mathbf{u}_j^T + \mathbf{u}_j\mathbf{u}_i^T)$ and suppose $Y = 0$. For any k, $0 = \mathbf{u}_k^T Y\mathbf{u}_k = s_k$ and for $\ell < k$, $0 = \mathbf{u}_\ell Y\mathbf{u}_k = s_{\ell k}$, and U is linearly independent. □

Corollary 7.3. $\dim \mathcal{T}_{\mathcal{R}_A} = v \operatorname{rank} A - \frac{\operatorname{rank} A(\operatorname{rank} A-1)}{2}$.

Proof. Let $\operatorname{rank} A = r$. By Observation 7.1 and Proposition 7.2,

$$\dim \mathcal{T}_{\mathcal{R}_A} = \dim \mathbb{R}^{(v)} - \dim \mathcal{N}_{\mathcal{R}_A} = \frac{v(v+1)}{2} - \frac{(v-r)(v-r+1)}{2}. \quad \square$$

An *optimal matrix for* $\xi(G)$ is a matrix A such that $\mathcal{G}(A) = G, \operatorname{null} A = \xi(A)$, and A has the Strong Arnold Hypothesis.

Theorem 7.4. *Let G be a connected graph.*

$$e(G) + b \ge \frac{\xi(G)(\xi(G)+1)}{2} \tag{4}$$

where $b = 1$ if G is bipartite and every optimal matrix for $\xi(G)$ has zero diagonal, and $b = 0$ otherwise.

Proof. Let A be an optimal matrix for $\xi(G)$, chosen to have at least one nonzero diagonal entry if there is such an optimal matrix. Let $\operatorname{rank} A = r$.

The Strong Arnold Hypothesis for A is $\mathcal{N}_{\mathcal{R}_A} \cap \mathcal{N}_{\mathcal{S}_A} = \{0\}$, which is equivalent by taking orthogonal complements to

$$\mathcal{T}_{\mathcal{R}_A} + \mathcal{T}_{\mathcal{S}_A} = \mathbb{R}^{(v)}$$

Therefore

$$\begin{aligned} \dim \mathcal{T}_{\mathcal{R}_A} + \dim \mathcal{T}_{\mathcal{S}_A} - \dim(\mathcal{T}_{\mathcal{R}_A} \cap \mathcal{T}_{\mathcal{S}_A}) &= \dim \mathbb{R}^{(v)} \\ vr - \frac{r(r-1)}{2} + e(G) + v - \dim(\mathcal{T}_{\mathcal{R}_A} \cap \mathcal{T}_{\mathcal{S}_A}) &= \frac{v(v+1)}{2} \end{aligned}$$

$$\begin{aligned} e(G) &= \frac{v(v+1)}{2} - vr + \frac{r(r-1)}{2} - v + \dim(\mathcal{T}_{\mathcal{R}_A} \cap \mathcal{T}_{\mathcal{S}_A}) \\ &= \frac{1}{2}((v-r)^2 + (v-r)) - v + \dim(\mathcal{T}_{\mathcal{R}_A} \cap \mathcal{T}_{\mathcal{S}_A}) \\ &= \frac{\xi(G)(\xi(G)+1)}{2} - v + \dim(\mathcal{T}_{\mathcal{R}_A} \cap \mathcal{T}_{\mathcal{S}_A}) \end{aligned}$$

Thus

$$\frac{\xi(G)(\xi(G)+1)}{2} = e(G) + v - \dim(\mathcal{T}_{\mathcal{R}_A} \cap \mathcal{T}_{\mathcal{S}_A}).$$

Let $D = \mathrm{diag}(d_1, \ldots, d_n)$ be a diagonal matrix. Then by Observation 7.1.3, $DA + AD \in \mathcal{T}_{\mathcal{R}_A}$. Clearly, $DA + AD \in \mathcal{T}_{\mathcal{S}_A}$, so $DA + AD \in \mathcal{T}_{\mathcal{R}_A} \cap \mathcal{T}_{\mathcal{S}_A}$. Let $\mathbf{e}_k$ be the kth standard basis vector of $\mathbb{R}^v$. Define $D_k = \mathrm{diag}(\mathbf{e}_k)$ and $B_k = D_k A + AD_k$. Note that $(B_k)_{ij} = (\delta_{ki} + \delta_{kj})a_{ij}$, where $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$.

We show first that if $\sum_{k=1}^{v} c_k B_k = 0$ and $c_t = 0$ for some t such that $1 \leq t \leq v$, then $c_k = 0$ for all $1 \leq k \leq v$. For every neighbor y of t,

$$0 = \left(\sum_{k=1}^{v} c_k B_k\right)_{ty} = \sum_{k=1}^{v} c_k(\delta_{kt} + \delta_{ky})a_{ty} = c_y a_{ty}.$$

Since $\{t, y\}$ is an edge of G, $a_{ty} \neq 0$, and so $c_y = 0$. Since G is connected, every vertex can be reached by a path from t, and so $c_1 = \cdots = c_v = 0$.

Since

$$\sum_{k=1}^{v-1} c_k B_k = \sum_{k=1}^{v} c_k B_k \text{ with } c_v = 0,$$

it follows that for every graph G and $\xi(G)$-optimal matrix A (without any assumption about the diagonal), the matrices $B_k, k = 1, \ldots, v-1$, are linearly independent, and thus

$$\dim(\mathcal{T}_{\mathcal{R}_A} \cap \mathcal{T}_{\mathcal{S}_A}) \geq v - 1 \qquad \text{and} \qquad \frac{\xi(G)(\xi(G)+1)}{2} \leq e(G) + 1,$$

Now suppose that A has a nonzero diagonal entry or G is not bipartite. We show that the matrices $B_k, k = 1, \ldots, v$ are linearly independent, so

$$\dim(\mathcal{T}_{\mathcal{R}_A} \cap \mathcal{T}_{\mathcal{S}_A}) \geq v \qquad \text{and} \qquad \frac{\xi(G)(\xi(G)+1)}{2} \leq e(G)$$

Let

$$\sum_{k=1}^{v} c_k B_k = 0.$$

If A has a nonzero diagonal entry a_{tt}, then $0 = (\sum_{k=1}^{v} c_k B_k)_{tt} = 2c_t a_{tt}$, and so $c_t = 0$. If G is not bipartite, there is an odd cycle; without loss of generality let this odd cycle be $(1, \ldots, t)$. Then for $i = 1, \ldots, t-1$,

$$0 = \left(\sum_{k=1}^{v} c_k B_k\right)_{i,i+1} = (c_i + c_{i+1})a_{i,i+1};$$

similarly $0 = (c_t + c_1)a_{t,1}$. Since $\{t, 1\}$ and $\{i, i+1\}, i = 1, \ldots, t-1$ are edges of G,

$$c_i + c_{i+1} = 0, i = 1, \ldots, t-1, \text{ and } c_t + c_1 = 0.$$

By adding equations $(-1)^i(c_i + c_{i+1} = 0), i = 1, \ldots, t-1$ to $c_t + c_1 = 0$, we obtain $2c_t = 0$. □

If G is the disjoint union of its connected components $G_1, \ldots, G_h$, then $\xi(G) = \max_{i=1,\ldots,h}\{\xi(G_i)\}$ [3].

Corollary 7.5. *For every graph G,*

$$\frac{\xi(G)(\xi(G)+1)}{2} \leq e(G) + 1.$$

Example 7.6. The complete bipartite graph $K_{3,3}$ demonstrates that $b = 1$ is sometimes necessary in the bound (4), because $\xi(K_{3,3}) = 4$, so $\frac{\xi(G)(\xi(G)+1)}{2} = 10$, and $e(K_{3,3}) = 9$.

8 Bounds for the expected value of ξ

In this section we show that if v is sufficiently large, then the expected value of $\xi(G(v,p))$ is asymptotically at most $\sqrt{p}v$. It follows that the average value of ξ for graphs of order v is asymptotically at most $\frac{1}{\sqrt{2}}v$.

We will make the notion of asymptotic expected value more precise, both for minimum rank and for ξ.

Define

$$\overline{\mathrm{mr}}(p) = \limsup_{v\to\infty} \frac{\mathbf{E}\left[\mathrm{mr}(G(v,p))\right]}{v}.$$

(This is a careful definition, as the lim sup is almost certainly a limit.) In previous sections we have shown that for $0 < p < 1$,

$$c(p) \leq \overline{\mathrm{mr}}(p) \leq 1 - p.$$

Now define

$$\bar{\xi}(p) = \liminf_{v\to\infty} \frac{\mathbf{E}\left[\xi(G(v,p))\right]}{v}.$$

The quantity $\bar{\xi}(p)$ should be compared to $1 - \overline{\mathrm{mr}}(p)$ rather than $\overline{\mathrm{mr}}(p)$, since $\xi(G)$ measures a nullity rather than a rank.

Our starting point is an immediate consequence of Corollary 7.5.

Corollary 8.1. *For every graph G,*

$$\xi(G) \quad \leq \quad \frac{1}{2}(-1 + \sqrt{9 + 8e(G)}). \tag{5}$$

Corollary 8.2. *For $0 < p < 1$,*

$$p \leq \bar{\xi}(p) \leq \sqrt{p}.$$

Proof. The proof that $p \leq \bar{\xi}(p)$ follows from Theorem 6.1 by exactly the same reasoning that showed that $\overline{\mathrm{mr}}(p) \leq 1 - p$.

From inequality (5), if $e(G) \geq 2$,

$$\xi(G) \leq \sqrt{2e(G)}.$$

For a fixed $\epsilon > 0$, as $v \to \infty$, almost all graphs sampled from $G(v,p)$ satisfy

$$e(G) \leq (1+\epsilon)\frac{p}{2}v(v-1),$$

so almost all graphs satisfy

$$\xi(G) \leq \sqrt{2e(G)} \leq \sqrt{2(1+\epsilon)\frac{p}{2}v(v-1)} \leq \sqrt{1+\epsilon}\sqrt{p}v.$$

This completes the proof of the second inequality $\bar{\xi}(p) \leq \sqrt{p}$. □

Since for every graph G, $\xi(G) \le \mathrm{M}(G)$ and for every $v > 1$ there exists a graph H such that $\xi(H) < \mathrm{M}(H)$, $\mathbf{E}[\xi(G(v,p))]$ is strictly less than $\mathbf{E}[\mathrm{M}(G(v,p))]$ for $v > 1$ and $0 < p < 1$. However it is quite possible that taking the limit gives $\bar{\xi}(p) + \overline{\mathrm{mr}}(p) = 1$, in which case Corollary 8.2 would provide a better asymptotic lower bound for expected minimum rank than that given in Corollary 4.3. The graphs of these bounds are shown in Figure 2.

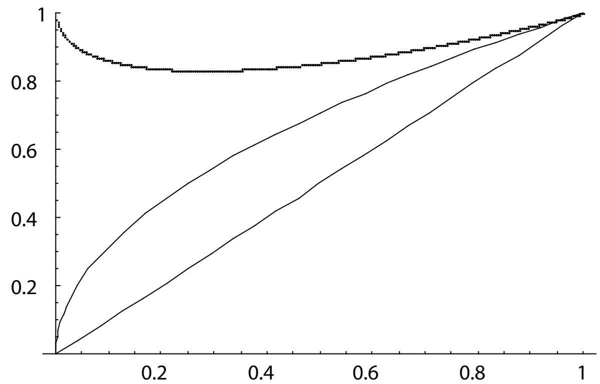

Figure 2: The graphs of $1 - c(p) > \sqrt{p} > p$ for $0 < p < 1$

A Appendix: Estimation of the binomial coefficient

Lemma A.1. *Let N be a positive integer and α, β, γ be real numbers with $\alpha, \beta \in (0,1)$ and $\gamma \in [0,1]$. Then,*

$$\binom{(\alpha+\beta+\gamma)N}{\alpha N} \le E(\alpha,\beta,\gamma,N)\left(\frac{(\alpha+\beta)^{\alpha+\beta}}{\alpha^\alpha \beta^\beta}\right)^N,$$

where

$$E(\alpha,\beta,\gamma,N) = \sqrt{\frac{\alpha+\beta}{2\pi\alpha\beta N}}\exp\left\{\frac{1}{12(\alpha+\beta)N} + \gamma\left(1+\frac{\alpha}{\beta}\right)N\right\}.$$

Proof. We use Stirling's formula as given in [6, page 216]:

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \le n! \le e^{1/(12n)}\sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

From this formula,

$$\begin{aligned}
&\binom{(\alpha+\beta+\gamma)N}{\alpha N} \\
&\le \; e^{1/(12(\alpha+\beta+\gamma)N)}\frac{\sqrt{2\pi(\alpha+\beta+\gamma)N}}{\sqrt{2\pi\alpha N}\sqrt{2\pi(\beta+\gamma)N}}\left(\frac{(\alpha+\beta+\gamma)N}{e}\right)^{(\alpha+\beta+\gamma)N}\left(\frac{e}{\alpha N}\right)^{\alpha N}\left(\frac{e}{(\beta+\gamma)N}\right)^{(\beta+\gamma)N} \\
&\le \; e^{1/(12(\alpha+\beta)N)}\sqrt{\frac{\alpha+\beta+\gamma}{2\pi\alpha(\beta+\gamma)N}}\frac{(\alpha+\beta+\gamma)^{(\alpha+\beta+\gamma)N}}{\alpha^{\alpha N}(\beta+\gamma)^{(\beta+\gamma)N}}
\end{aligned}$$

Since $\frac{\alpha+\beta+\gamma}{\alpha(\beta+\gamma)} \leq \frac{\alpha+\beta}{\alpha\beta}$,

$$\begin{aligned}
&\binom{(\alpha+\beta+\gamma)N}{\alpha N} \\
&\leq \quad e^{1/(12(\alpha+\beta)N)}\sqrt{\frac{\alpha+\beta}{2\pi\alpha\beta N}}\left(\frac{(\alpha+\beta+\gamma)^{(\alpha+\beta+\gamma)}}{\alpha^\alpha(\beta+\gamma)^{(\beta+\gamma)}}\right)^N \\
&\leq \quad e^{1/(12(\alpha+\beta)N)}\sqrt{\frac{\alpha+\beta}{2\pi\alpha\beta N}}\left(\frac{(\alpha+\beta)^{\alpha+\beta}}{\alpha^\alpha\beta^\beta}\right)^N\left(\left(\frac{\alpha+\beta+\gamma}{\alpha+\beta}\right)^{\alpha+\beta}\left(\frac{\beta}{\beta+\gamma}\right)^\beta\left(\frac{\alpha+\beta+\gamma}{\beta+\gamma}\right)^\gamma\right)^N \\
&\leq \quad \left(\frac{(\alpha+\beta)^{\alpha+\beta}}{\alpha^\alpha\beta^\beta}\right)^N\sqrt{\frac{\alpha+\beta}{2\pi\alpha\beta N}}e^{1/(12(\alpha+\beta)N)}\left(\left(1+\frac{\gamma}{\alpha+\beta}\right)^{\alpha+\beta}\left(1+\frac{\alpha}{\beta}\right)^\gamma\right)^N
\end{aligned}$$

Because $1+x \leq e^x$,

$$\binom{(\alpha+\beta+\gamma)N}{\alpha N} \leq \left(\frac{(\alpha+\beta)^{\alpha+\beta}}{\alpha^\alpha\beta^\beta}\right)^N\sqrt{\frac{\alpha+\beta}{2\pi\alpha\beta N}}e^{1/(12(\alpha+\beta)N)}\exp\left\{\gamma N+\gamma\frac{\alpha}{\beta}N\right\} \quad \square$$

Corollary A.2. *Let p, c be fixed and let $r = cv$, with $v \to \infty$.*

$$\binom{rv+2p\binom{v}{2}+2v\sqrt{2\ln v}}{rv} \leq \left((1+o(1))\left(\frac{(c+p)^{c+p}}{c^c p^p}\right)\right)^{v^2}.$$

Proof.

$$\binom{rv+2p\binom{v}{2}+2v\sqrt{2\ln v}}{rv} = \binom{cv^2+pv^2-pv+2v\sqrt{2\ln v}}{cv^2}.$$

Let $N = v^2$, $\alpha = c$, $\beta = p$, and $\gamma = \frac{1}{v^2}(-pv+2v\sqrt{2\ln v})$. With p and c fixed, by Lemma A.1 we see that

$$\binom{rv+2p\binom{v}{2}+2v\sqrt{2\ln v}}{rv} \leq \left((1+o(1))\left(\frac{(c+p)^{c+p}}{c^c p^p}\right)\right)^{v^2}. \quad \square$$

B Appendix: Connectivity is minimum degree

Bollobás and Thomason [8] proved that for $G \sim G(v,p)$, regardless of p, then $\Pr[\kappa(G) < \delta(G)] \to 0$ as $v \to \infty$. Bollobás [5] proved the result for p in a restricted interval, but the statement of his theorem is much more general. For our result, we need to compute the probability that $\kappa(G) = \delta(G)$ where $G \sim G(v,p)$, but only need the result for p fixed.

Lemma B.1. *Let $p \in (0,1)$ be fixed and G be distributed according to $G(v,p)$.*

$$\Pr[\kappa(G) < \delta(G)] \leq 4v^{-2}.$$

Proof. Let $\delta = \delta(G)$. By Theorem 3.2 we see that, with probability at least $1 - v^{-2}$,

$$\delta \leq \frac{2e(G)}{v} \leq \frac{2\left(p\binom{v}{2}+v\sqrt{2\ln v}\right)}{v} = p(v-1)+2\sqrt{2\ln v} \leq pv+2\sqrt{2\ln v}. \tag{6}$$

If $\kappa(G) < \delta$, then there exists a partition $V(G) = V_1 \cup S \cup V_2$ such that $|S| < \delta$, $2 \leq |V_1| \leq |V_2|$ and there is no edge between V_1 and V_2. Let the *closed neighborhood of vertex x* be denoted $N[x]$ and be equal to $\{x\} \cup N(x)$.

We will first show that there is an integer t such that if $|V_1| \geq 2$, then, with high probability, $|V_1| \geq t$ (we will determine the value of t later). Afterwards, we will show that if $|V_1| \geq t$, then $|V_1 \cup S| \geq (v+\delta)/2$, which contradicts the choice of V_1.

Let x_1 and x_2 be distinct vertices in V_1. It is easy to see that

$$\mathbf{E}\left[|N[x_1] \cup N[x_2]|\right] = 2 + (2p - p^2)(v-2).$$

$$\begin{aligned}
&\Pr\left[|N[x_1] \cup N[x_2]| < t+\delta\right] \\
&\quad = \Pr\left[|N[x_1] \cup N[x_2]| - \big(2 + (2p-p^2)(v-2)\big) < (t+\delta) - \big(2 + (2p-p^2)(v-2)\big)\right] \\
&\quad \leq \exp\left\{-\frac{1}{2(v-2)}\left((t+\delta) - 2 - (2p-p^2)(v-2)\right)^2\right\},
\end{aligned}$$

by the negative version of Theorem 3.1. Thus if $t + \delta \leq (2p-p^2)(v-2) + 2 - 2\sqrt{v \ln v}$, then $|V_1 \cup S| \geq t + \delta$ with probability at least $1 - v^{-2}$.

Set $t = (2p-p^2)v - 2\sqrt{v\ln v} - \delta$, so $|V_1 \cup S| \geq t+\delta$ with probability at least $1 - v^{-2}$. Let $\mathcal{E}$ be the event that $\delta \leq pv + 2\sqrt{2\ln v}$. If event $\mathcal{E}$ happens, then as long as $v \geq 2$, our value of t satisfies

$$t \geq (2p-p^2)v - 2\sqrt{v\ln v} - \left(pv + 2\sqrt{2\ln v}\right) \geq p(1-p)v - 4\sqrt{v\ln v}.$$

By (6), with probability at least $1 - v^{-2}$,

$$\left\{\delta \leq pv + 2\sqrt{2\ln v}\right\} \text{ and } \left\{t \geq p(1-p)v - 4\sqrt{v\ln v}\right\}.$$

Now let $x_1, x_2, \ldots, x_t$ be distinct vertices in V_1. Again, it is easy to see that $\mathbf{E}\left[\left|\bigcup_{i=1}^t N[x_i]\right|\right] = t + \left(1 - (1-p)^t\right)(v-t)$. By applying Theorem 3.1, we have

$$\begin{aligned}
&\Pr\left[\left|\bigcup_{i=1}^t N[x_i]\right| < \frac{v+\delta}{2}\right] \\
&\quad = \Pr\left[\left|\bigcup_{i=1}^t N[x_i]\right| - \left(t + \left[1-(1-p)^t\right](v-t)\right) < \frac{v+\delta}{2} - \left(t + \left[1-(1-p)^t\right](v-t)\right)\right] \\
&\quad \leq \exp\left\{-\frac{1}{2(v-t)}\left(\frac{v+\delta}{2} - \left(t + \left[1-(1-p)^t\right](v-t)\right)\right)^2\right\}.
\end{aligned}$$

Since $\Pr[\mathcal{E}] \geq 1 - v^{-2}$,

$$\Pr\left[\left|\bigcup_{i=1}^t N[x_i]\right| < \frac{v+\delta}{2} \,\middle|\, \mathcal{E}\right] \leq \exp\left\{-\frac{1}{2(v-t)}\left(\frac{v+\delta}{2} - \left(t + \left[1-(1-p)^t\right](v-t)\right)\right)^2\right\}(1-v^{-2})^{-1}.$$

Event $\mathcal{E}$ gives that $t \geq p(1-p)v - 4\sqrt{v\ln v}$, which easily implies

$$t \geq \frac{1}{\ln(1-p)} \ln\left(\frac{1-p}{2} - 4\sqrt{\frac{\ln v}{v}}\right)$$

when p is fixed and v is large enough. Thus

$$(1-p)^t \quad \leq \quad \frac{1}{2} - \frac{p}{2} - \frac{\sqrt{2\ln v}}{v} - 2\sqrt{\frac{\ln v}{v}}$$

Since Event $\mathcal{E}$ gives that $\delta \leq pv + 2\sqrt{2\ln v}$,

$$(1-p)^t \quad \leq \quad 1 - \frac{v+\delta}{2v} - \frac{4\sqrt{v\ln v}}{2v}.$$

Then

$$\frac{v+\delta}{2} - \left(t + \left[1-(1-p)^t\right](v-t)\right) \geq \frac{v+\delta}{2} + 2\sqrt{v\ln v},$$

and so

$$\Pr\left[|V_1 \cup S| < \frac{v+\delta}{2} \mid \mathcal{E}\right] \leq v^{-2}(1-v^{-2})^{-1} \leq 2v^{-2}.$$

Summarizing, the assumption that $\kappa(G) < \delta(G)$ leads to the contradiction $|V_1 \cup S| \geq \frac{1}{2}(v+\delta)$ with probability at least $1-4v^{-2}$. $\square$

References

[1] N. Alon and J.H. Spencer, *The Probabilistic Method*, second ed., John Wiley & Sons, New York, 2000.

[2] American Institute of Mathematics workshop "Spectra of Families of Matrices described by Graphs, Digraphs, and Sign Patterns," held October 23-27, 2006 in Palo Alto, CA. Workshop webpage: `http://aimath.org/pastworkshops/matrixspectrum.html`.

[3] F. Barioli, S.M. Fallat, and L. Hogben, A variant on the graph parameters of Colin de Verdière: Implications to the minimum rank of graphs. *Electron. J. Linear Algebra*, **13** (2005), 387–404.

[4] A. Berman, S. Friedland, L. Hogben, U.G. Rothblum and B. Shader, An upper bound for minimum rank of a graph, *Linear Algebra Appl.* **429** (2008), no. 7, 1629-1638.

[5] B. Bollobás, Degree sequences of random graphs, *Discrete Math.* **33** (1981), no. 1, 1–19.

[6] B. Bollobás, *Modern Graph Theory*. Graduate Texts in Mathematics, 184. Springer-Verlag, New York, 1998.

[7] B. Bollobás, *Random Graphs*, second ed., Cambridge Studies in Advanced Mathematics, 73, Cambridge University Press, Cambridge, 2001.

[8] B. Bollobás and A. Thomason, Random graphs of small order, in: *Random Graphs '83 (Poznań, 1983)*, North-Holland Math. Stud., 118, North-Holland, Amsterdam, 1985, pp. 47–97.

[9] R. Brualdi, L. Hogben and B. Shader, AIM Workshop Spectra of Families of Matrices described by Graphs, Digraphs, and Sign Patterns, Final Report: Mathematical Results (revised, March 30, 2007). `http://aimath.org/pastworkshops/matrixspectrumrep.pdf`.

[10] Y. Colin de Verdière, On a new graph invariant and a criterion for planarity, in: *Graph Structure Theory*, Contemporary Mathematics, vol. 147, American Mathematical Society, Providence, 1993, pp. 137–147.

[11] Y. Colin de Verdière, Multplicities of eigenvalues and tree-width graphs, *J. Combin. Theory Ser. B* **74** (1998), no. 2, 121–146.

[12] S. Fallat and L. Hogben, The minimum rank of symmetric matrices described by a graph: A survey, *Linear Algebra Appl.* **426** (2007) 558–582.

[13] H. van der Holst, L. Lovász, and A. Schrijver, The Colin de Verdière graph parameter, in: *Graph Theory and Computational Biology (Balatonlelle, 1996)*, Bolyai Soc. Math. Stud., 7, János Bolyai Math. Soc., Budapest, 1999, pp. 29–85.

[14] L. Lovász, M. Saks and A. Schrijver, Orthogonal representations and connectivity of graphs, *Linear Algebra Appl.* **114/115** (1989), 439–454.

[15] L. Lovász, M. Saks and A. Schrijver, A correction: "Orthogonal representations and connectivity of graphs," *Linear Algebra Appl.* **313** (2000), no. 1-3, 101–105.

[16] R. Pendavingh, On the relation between two minor-monotone graph parameters, *Combinatorica*, **18** (1998), no. 2, 281-292.

[17] L. Rónyai, L. Babai and M.K. Ganapathy, On the number of zero-patterns of a sequence of polynomials, *J. Amer. Math. Soc.* **14** (2001), no. 3, 717–735 (electronic).

[18] E. Shamir and J.H. Spencer, Sharp concentration of the chromatic number in random graphs $G_{n,p}$, *Combinatorica* **7** (1987), no. 1, 121–129.