Finding Better Bounds for Graphs with Small Independent Sets

Philip Carpenter and Jonathan Kropko*
Department of Mathematics
The Ohio State University
231 West 18th Avenue
Columbus, OH 43210

December 29, 2004

We would like to thank Craig Lennon and Steven J. Miller for their valuable suggestions and guidance throughout this project.

Abstract

We examined random graphs on n vertices with m edges. We are specifically interested in graphs with large girth and chromatic number, which will mean a small independent set, whose size we call p. We used the probabilistic method: by comparing the number of the graphs with the desired property to the total number of graphs, we can get an idea of whether or not it is likely for such a graph to occur when the number of vertices becomes very large. In Sarnak's book *Some Applications of Modular Forms* [Sar] he used the number of edges, m, of the order $n^{1+\epsilon}$ and the size of the independent set, p, of the order $n^{1-\eta}$. He shows that these graphs, which have an independent set of size p, and contain at most n edges in the independent set, are not likely to be randomly generated in the limit as n becomes large, provided that $\epsilon > 2\eta$. We have found that these bounds will work, but are generally wasteful, and the same property can be found using smaller m and p, given $\epsilon > \frac{1}{e}$.

---

*E-mail: `carpenter.339@osu.edu, kropko.4@osu.edu`

1 Sarnak's bounds

Sarnak presents a proof, attributed to Erdos, that a graph can be constructed with chromatic number $\geq w$ and girth $\geq k$. The girth of a graph is the length of the smallest circuit in the graph. As the girth of a graph increases, the density of edges in the graph decreases. The chromatic number, sometimes called the coloring number, represents the minimum number of colors needed to color every vertex in a graph such that adjacent vertices are different colors. The chromatic number is a measure of connectivity in a graph, and it is related to the size of the largest set of vertices in a graph that have no connections with each other. Such a set of vertices is called the maximal independent set. Sarnak describes the following relation in [Sar]:

Given a graph X with n vertices, a chromatic number of $c(X)$ and a maximal independent set of size $i(X)$,

$$c(X) \geq \frac{n}{i(X)}. \tag{1}$$

Graphs with large girth and high chromatic number tend to be well connected with relatively few edges relative to the number of vertices. (Obviously this number we would expect to be greater than the number of vertices, n, and for this to be non trivial, we want it significantly lower than the total number of possible edges, $\binom{n}{2}$). Sarnak shows the existence of such graphs for a large number of vertices, n. In order to maximize the chromatic number, it follows that the maximal independent set must be small. To ensure a large chromatic number, it suffices to have a small independent set relative to the number of vertices. In other words, if $n >> i(X)$ so that $i(X) = o(n)$, then $c(X)$ must be large. We want to show that graphs on n vertices with small independent sets can be constructed randomly, by choosing some number of edges m of the total possible $\binom{n}{2}$edges. Sarnak utilized a probability argument; he wanted to show that the number of unfavorable graphs that can be generated randomly out of n vertices becomes negligible as the number of vertices approaches infinity. Sarnak counts the number of graphs with at least m edges and maximal independent set of size p. A graph is considered unfavorable if it has an independent set that is larger than p vertices because this will not be very well connected.

He shows that the ratio of the number of unfavorable graphs to the total number of graphs approaches 0 as the number of vertices goes to infinity. Sarnak chose m, the number of edges, to be of the order $n^{1+\epsilon}$ and p of the order $n^{1-\eta}$. He gives the condition: $\epsilon > 2\eta$. If this holds then the random graph will have the desired property, namely a large girth and large chromatic number, and less than n edges in the independent set.

Fix a set of size p in a graph with n vertices. The number of graphs with at most n edges in this set is at most[1]

$$\sum_{l=0}^{n} \binom{\binom{p}{2}}{l} \binom{\binom{n}{2} - \binom{p}{2}}{m-l}. \tag{2}$$

Since binomial coefficients increase towards the middle, we have the following result for (2):

$$\begin{aligned}
&< (n+1)\binom{\binom{p}{2}}{n}\binom{\binom{n}{2} - \binom{p}{2}}{m} \\
&< p^{2n}\binom{\binom{n}{2} - \binom{p}{2}}{m} \\
&< \binom{\binom{n}{2}}{m} p^{2n}\left(1 - \frac{\binom{p}{2}}{\binom{n}{2}}\right)^{m} \\
&< \binom{\binom{n}{2}}{m} p^{2n}\left(1 - \left(\frac{p-1}{n-1}\right)^{2}\right)^{m} \\
&< \binom{\binom{n}{2}}{m} p^{2n} e^{-m\left(\frac{p-1}{n-1}\right)^{2}}.
\end{aligned} \tag{3}$$

This last step follows from the fact that $(1 - \frac{s}{t})^t$ converges to e^{-s} from below, as $t \longrightarrow \infty$.

The expression $\binom{\binom{n}{2}}{m} p^{2n} e^{-m\left(\frac{p-1}{n-1}\right)^{2}}$ is for a given choice of p vertices. To account for the total number of choices of p from n vertices we must multiply by $\binom{n}{p}$, then to show this number of graphs is a small percent of all graphs with n vertices and m edges we must divide by the total number of graphs, which is $\binom{\binom{n}{2}}{m}$. Since $\binom{n}{p}$ represents the binomial coefficient, it is less than 2^n.

So the final upper bound that Sarnak provides us is

$$2^n p^{2n} e^{-m(\frac{p-1}{n-1})^2}. \tag{4}$$

The entire expression will converge to zero provided that $\epsilon > 2\eta$. This behavior becomes obvious when looking at the expression in terms of exponentials in

---

[1] In [Sar], there were a few typographical errors. See section 7 for a more detailed account of our revisions. Here we present the corrected version from [DSV]

n:

$$\begin{aligned}
2^n p^{2n} e^{-n^{1+\epsilon}(\frac{n^{2-2\eta}-2n^{1-\eta}+1}{n^2-2n+1})} &\leq \frac{\binom{\binom{n}{2}}{m} 2^n p^{2n} e^{-n^{1+\epsilon-2\eta}}}{\binom{\binom{n}{2}}{m}} \\
&= 2^n p^{2n} e^{-n^{1+\epsilon-2\eta}} \\
&= (2p^2 e^{-n^{\epsilon-2\eta}})^n, \qquad (5)
\end{aligned}$$

which approaches 0 in the limit when $\epsilon > 2\eta$.

2 New parameters for edges and independent sets

Our task was to revise Sarnak's argument and to determine better convergence criteria for m and p. We looked at m as $nf(n)$ and p as $\frac{n}{g(n)}$. Our hope was to find functions of n which are much smaller than the exponentials that Sarnak proposed. The smaller number of edges, m, allows us to create graphs which are more sparse and still adequately connected. Our first thought was to use f and g as functions using $\log n$. It turns out that these replacements can be done easily.

Our first step was to find a convergence condition for the upper bound Sarnak's inequalities gave us. Letting $f = f(n)$ and $g = g(n)$:

$$2^n p^{2n} e^{-m\left(\frac{p-1}{n-1}\right)^2}. \qquad (6)$$

Replacing m and p with nf and $\frac{n}{g}$ respectively, we get

$$S(n) = 2^n \left(\frac{n}{g}\right)^{2n} e^{-nf\left(\frac{p-1}{n-1}\right)^2}. \qquad (7)$$

To determine when $S(n)$ will go to zero we note that when this happens, the logarithm of $S(n)$ will diverge to negative infinity. So we examine $\log S(n)$:

$$\begin{aligned}
\log S(n) &= \log\left(2^n \left(\frac{n}{g}\right)^{2n} e^{-nf\cdot\left(\frac{p-1}{n-1}\right)^2}\right) \\
&= n\log 2 + 2n\log\frac{n}{g} - nf\cdot\left(\frac{p-1}{n-1}\right)^2 \\
&= n\left(\log 2 + 2\log\frac{n}{g} - f\cdot\left(\frac{p-1}{n-1}\right)^2\right). \qquad (8)
\end{aligned}$$

Note that $\frac{p-1}{n-1} > \frac{p-1}{n}$ and so if $\log S(n)$ still diverges to negative infinity, $S(n)$ will certainly still approach zero if we replace $f \cdot \left(\frac{p-1}{n-1}\right)^2$ with $f \cdot \left(\frac{p-1}{n}\right)^2$.

$$\begin{aligned} \log S(n) &= n\left(\log 2 + 2\log\frac{n}{g} - f\left(\frac{(p-1)}{n}\right)^2\right) \\ &= n\left(\log 2 + 2\log\frac{n}{g} - f\left(\frac{1}{g} - \frac{1}{n}\right)^2\right) \\ &= n\left(\log 2 + 2\log\frac{n}{g} - f\frac{1}{g^2} + 2\frac{f}{ng} - \frac{f}{n^2}\right). \end{aligned} \tag{9}$$

In order for $\log S(n)$ to be negative (and hence $S(n)$less than 1), we require:

$$\log 2 + 2\log\frac{n}{g} + 2\frac{f}{ng} \leq \frac{f}{g^2} + \frac{f}{n^2}. \tag{10}$$

We want $nf(n) < \binom{n}{2}$ so that the number of edges is fewer than in the complete graph with n vertices. Accordingly, given $f(n) < n$ we can say that $\log 2 + 2\frac{f(n)}{ng(n)} - \frac{f(n)}{n^2}$ will decrease to a constant and thus will definitely be overcome by the $\log g$ term, but because the $\log g$ will be small compared to the $\log n$ we must require that:

$$2\log n \leq f\frac{1}{g^2}. \tag{11}$$

This inequality has been derived without using any specific assumptions about $f(n)$ and $g(n)$ other than that they are positive, monotonic functions increasing slower than n. These assumptions allow us to choose a wider variety of functions for $f(n)$ and $g(n)$, specifically logarithms and powers of logarithms. For example, it is easily seen that for example $f(n) = 2(\log n)^3$ and $g(n) = \log n$ will satisfy this condition and be less than any n^ϵ for large n.

3 Our bounds

One of our goals was to find a better upper bound than the one Sarnak's inequalities gave us. We can show that this closer upper bound still works if we can derive the same parameters from (30). Sarnak's upper bound was sufficient to see behavior in the limit that $n \to \infty$, but as far as seeing convergence for small numbers, we needed something a little bit smaller, so we went through and tried to cut a little bit of a margin. Sarnak still provided our first step, an expression for the number of unfavorable graphs for a fixed p. Since the binomial coefficients are increasing towards the middle,

$$\sum_{l=0}^{n}\binom{\binom{p}{2}}{l}\binom{\binom{n}{2}-\binom{p}{2}}{m-l} \leq (n+1)\binom{\binom{p}{2}}{n}\binom{\binom{n}{2}-\binom{p}{2}}{m}. \tag{12}$$

Multiplying by the number of ways we can choose a set of size p out of n vertices, we get an expression which is an upper bound for the total number of unfavorable graphs:

$$\binom{n}{p}(n+1)\binom{\binom{p}{2}}{n}\binom{\binom{n}{2}-\binom{p}{2}}{m}. \tag{13}$$

In order to show that the probability of an unfavorable graph being chosen at random is small, we must show that the total number of unfavorable graphs is small when compared to the total number of graphs with m edges on n vertices. It is enough to show that

$$\lim_{n\to\infty}\frac{\binom{n}{p}(n+1)\binom{\binom{p}{2}}{n}\binom{\binom{n}{2}-\binom{p}{2}}{m}}{\binom{\binom{n}{2}}{m}} = 0. \tag{14}$$

Let $a=\binom{n}{2}$ and let $b=\binom{p}{2}$. Define a function $T(n)$:

$$T(n) = \frac{\binom{n}{p}(n+1)\binom{b}{n}\binom{a-b}{m}}{\binom{a}{m}}. \tag{15}$$

Let us expand the binomials in $T(n)$ into their factorial forms:

$$\begin{aligned} T(n) &= \frac{(n+1)n!b!(a-b)!m!(a-m)!}{p!(n-p)!n!(b-n)!m!(a-b-m)!a!} \\ &= \frac{(n+1)b!(a-b)(a-b-1)\cdots(a-b-m+1)}{p!(n-p)!(b-n)!(a)\cdots(a-1)(a-m+1)}. \end{aligned} \tag{16}$$

Now consider the ratio

$$\begin{aligned} & \frac{(a-b)(a-b-1)\cdots(a-b-m+1)}{(a)\cdots(a-1)(a-m+1)} \\ = \; & \left(\frac{a-b}{a}\right)\left(\frac{a-b-1}{a-1}\right)\cdots\left(\frac{a-b-m+1}{a-m+1}\right) \\ < \; & \left(\frac{a-b}{a}\right)^m . \end{aligned} \tag{17}$$

So,

$$T(n) < \frac{(n+1)b!}{p!(n-p)!(b-n)!}\left(\frac{a-b}{a}\right)^m . \tag{18}$$

Note: Stirling's Formula [Kar] will also give an approximation of the factorial function, however it has an error term of order $\frac{1}{n}$, which we avoided because the error is positive and though it will be closer, it would be greater than $n!$, and so would not satisfy $n! \leq (2x\pi)^{\frac{1}{2}} x^x e^{-x}$. For our purposes we only had to improve a small amount on Sarnak's bound and so going through his arguments and being a little bit more careful was satisfactory for our purposes.

Now consider separately

$$\frac{(n+1)b!}{p!(n-p)!(b-n)!}, \tag{19}$$

We can assume that $\frac{(n+1)}{(n-p)!} < 1$, so

$$\begin{aligned} \frac{(n+1)b!}{p!(n-p)!(b-n)!} & < \frac{b!}{p!(b-n)!} \\ & = \frac{(b)(b-1)\cdots(b-n+1)}{p!}. \end{aligned} \tag{20}$$

Proceeding, we will refer back to our definition of b as $\binom{p}{2}$. First, it is clear that $(b)(b-1)\cdots(b-n+1) < b^n$, and since $b = \binom{p}{2} = \frac{p(p-1)}{2}$, we know that $b^n = \frac{p^n(p-1)^n}{2^n}$.

So

$$\begin{aligned} \frac{(n+1)b!}{p!(n-p)!(b-n)!} & < \frac{p^n(p-1)^n}{2^n p!} \\ & = \frac{p^{n-1}(p-1)^{n-1}}{2^n(p-2)!}. \end{aligned} \tag{21}$$

By removing one power of 2 from each even term in $(p-2)!$ and disregarding the odd terms, we arrive at a somewhat closer approximation to the initial function:

$$\frac{p^{n-1}(p-1)^{n-1}}{2^n(p-2)!} < \frac{p^{n-1}(p-1)^{n-1}}{2^{n+\frac{p}{2}-1}} < \frac{p^{2n-2}}{2^{n+\frac{p}{2}-1}} \tag{22}$$

Therefore,

$$\begin{aligned} T(n) &< \frac{p^{2n-2}}{2^{n+\frac{p}{2}-1}}\left(\frac{a-b}{a}\right)^m \\ &= \frac{p^{2n-2}}{2^{n+\frac{p}{2}-1}}\left(1-\frac{b}{a}\right)^m. \end{aligned} \tag{23}$$

The final upper bound that we will use for our purposes is:

$$T(n) < \frac{p^{2n-2}}{2^{n+\frac{p}{2}-1}}e^{-\frac{bm}{a}} \tag{24}$$

This bound will give us significantly quicker convergence because now we are dividing by an exponential of 2 instead of multiplying by it. To show that our bound is consistent with Sarnak's upper bound (and thus we can still look at the same variety of functions for $f(n)$ and $g(n)$), we will check the convergence criteria in a similar way to what we did previously. First convert everything to a power of e:

$$e^{(2n-2)\log p-(n+\frac{p}{2}-1)\log 2-\frac{bm}{a}} \tag{25}$$

Since we want this expression to converge to 0, we want

$$(2n-2)\log p-(n+\frac{p}{2}-1)\log 2-\frac{bm}{a} \to -\infty. \tag{26}$$

The constant coefficients are insignificant in the limit, so we may disregard them. If we overestimate by ignoring $\frac{p}{2}$ and $\log p$ we come to:

$$2n\log p-n\log 2-\frac{bm}{a} \to -\infty. \tag{27}$$

So it is enough that

$$2\log p-\log 2-f\frac{b}{a} < 0. \tag{28}$$

Recall that $a = \binom{n}{2} = \frac{n(n-1)}{2}$ and that $b = \binom{p}{2} = \frac{n}{2g}(\frac{n}{g} - 1)$, so $\frac{b}{a} \geq \frac{1}{g^2} - \frac{1}{ng}$. Since $p = \frac{n}{g}$, we know $\log p = \log n - \log g$. So we want:

$$2 \log n - 2 \log g - \log 2 - \frac{f}{g^2} + \frac{f}{ng} < 0 \tag{29}$$

Once again, $\log g$, $\log 2$, and $\frac{f}{ng}$ will be insignificant in the limit. We arrive at the conclusion that

$$\frac{f}{2g^2} > \log n \tag{30}$$

which is exactly the same condition as we derived from Sarnak's upper bound. Except now, however, we have an upper bound allowing the solutions where f is a small power of log(n), (for example things like: $f = 2(\log n)^{1+2\beta+\delta}$, and $g = (\log n)^{\beta}$, for $\beta, \delta > 0$, using δ as a small margin to fulfill the inequality) to have more rapid convergence in our upper bound, this allows us to use these solutions for significantly smaller graphs. This upper bound is very similar to Sarnak's but slightly modified to achieve a more applicable upper bound using smaller numbers. It was needed because dealing with a smaller number of edges, convergence was too slow using the upper bound used in Sarnak's existence proof. More concisely we were a bit more careful with some terms which were insignificant in the limit but which for relatively small numbers could be important factors.

4 Choosing a better upper bound and number of edges

In this section we will look for ways to improve upon Sarnak's choices of n^{ϵ} and n^{η}, continuing the methods introduced in the previous section. We look specifically at powers of log(n). We have two major concerns when choosing suitable replacements. First we want to make sure that our number of edges, $nf(n)$, is less than the complete graph with n vertices, $\binom{n}{2}$. Secondly we want to use our upper bound to show that this choice will give a favorable probability of getting a graph with high connectivity and small independence set. For this choice to be an improvement, we will either need a choice of $f(n)$ so that it is always smaller than n^{ϵ}, or a range for which our new choice of $f(n)$ is smaller than n^{ϵ}. The second case is easy to come by because we know that a positive power of n in the limit will be far bigger than any power of $\log n$. For some general choice of ϵ we can take any relatively large positive power of $\log n$ and know that n^ϵ will be bigger for some (though possibly very large) value of n. This gives a situation when n^ϵ is better for small n and the power of $\log n$ will be better for large n. The first possible improvement, when a power of $\log n$ is always less than or equal to n^ϵ is the more difficult to choose for a given ϵ. And so we study the following equation:

$$\log^\lambda n \leq n^\epsilon \tag{31}$$

Taking the epsilon root and substituting $\alpha = \frac{\lambda}{\epsilon}$ we look at

$$\log^\alpha n \leq n \tag{32}$$

Let us introduce two functions $h_1(n) = \log^\alpha n$ and $h_2(n) = n$, and talk about the graphs of these functions.

We wish to find a value of α such that $h_1(n)$ is equal and tangent to $h_2(n)$ at one point. Finding this value of α we know that raising the power will increase $h_1(n)$, and the derivative, ensuring that there is some intersection and making that case the same as that previously discussed where there is a range where n^ϵ is a better choice and another range where $(\log n)^\lambda$ is a better choice. If we make α smaller then we know that the value of $h_1(n)$ will be smaller and it's derivative also smaller, and have no intersections with n^ϵ; while this is acceptable it will sacrifice some convergence when used in our upper bound and so we are interested in the largest power α we can choose so that $h_1(n) \leq h_2(n)$. We present the following proof that this largest power of $\log n$ that we can choose is $\alpha = e$.

If the functions $h_1(n) = (\log n)^\alpha$ and n are equal, then

$$n - (\log n)^\alpha = 0 \tag{33}$$

And if the functions are tangent at that same point, then

$$1 - \frac{\alpha(\log n)^{\alpha-1}}{n} = 0. \tag{34}$$

Let us solve this system of equations. First consider

$$\begin{aligned} n - (\log n)^\alpha &= 0 \\ (\log n)^\alpha &= n \\ (\log n)^{\alpha-1} &= n^{\frac{\alpha-1}{\alpha}} \end{aligned} \tag{35}$$

Now from the derivative of (33) we have

$$\begin{aligned} 1 - \frac{\alpha(\log n)^{\alpha-1}}{n} &= 0 \\ \frac{\alpha(\log n)^{\alpha-1}}{n} &= 1 \\ (\log n)^{\alpha-1} &= \frac{n}{\alpha} \end{aligned} \tag{36}$$

So we obtain the following equation:

$$\begin{aligned} n^{\frac{\alpha-1}{\alpha}} &= \frac{n}{\alpha} \\ n^{\alpha-1} &= \frac{n^\alpha}{\alpha^\alpha} \\ \frac{n^\alpha}{\alpha^\alpha} - n^{\alpha-1} &= 0 \\ n^\alpha - \alpha^\alpha n^{\alpha-1} &= 0 \\ n^{\alpha-1}(n - \alpha^\alpha) &= 0 \end{aligned} \tag{37}$$

We disregard the solution with $n = 0$ simply because n is the number of vertices. We have $n = \alpha^\alpha$. Substituting this solution back into (33), we get

$$\begin{aligned} \alpha^\alpha - (\log \alpha^\alpha)^\alpha &= 0 \\ \alpha^\alpha - \alpha^\alpha(\log \alpha)^\alpha &= 0 \\ \alpha^\alpha &= \alpha^\alpha(\log \alpha)^\alpha \\ 1 &= \log \alpha \end{aligned} \tag{38}$$

Therefore $\alpha = e$. So $h_1(n)$ is equal to and tangent to $h_2(n)$ when $n = e^e$ and $\alpha = e$. It still remains to be shown now that $(\log n)^e \leq n$ for all values of $n \geq 1$. Consider the logarithm of this inequality:

$$\begin{aligned} e \log(\log n) &\overset{?}{\leq} \log n \\ \log n - e \log\log n &\overset{?}{\geq} 0 \end{aligned} \tag{39}$$

Taking the derivative:

$$\begin{aligned} \frac{1}{n} - \frac{e}{n \log n} &\overset{?}{\geq} 0 \\ 1 - \frac{e}{\log n} &\overset{?}{\geq} 0 \end{aligned} \tag{40}$$

The derivative is negative when $\log n < e$ and positive when $\log n > e$. Therefore (33), which is the difference between n and $(\log n)^{\alpha}$, is decreasing when $n < e^e$ and increasing when $n > e^e$. Since the difference is decreasing on $(1, e^e)$, $n = e^e$ is the only intersection point on $(1, e^e)$. And since the difference is increasing on (e^e, ∞) it follows that $n = e^e$ is the only point where $(\log n)^{\alpha} = n$. Therefore, for $\alpha = e$, $(\log n)^{\alpha} \leq n$ for all $n \geq 1$.

Replacing α with e in $\alpha = \frac{\lambda}{\epsilon}$ and solving for our power λ:

$$\begin{aligned} \alpha &= \frac{\lambda}{\epsilon} \\ & or \\ \lambda &= e\epsilon \end{aligned} \tag{41}$$

This condition allows us to choose a fractional power of $\log n$ that will always be less than or equal to n^{ϵ} for a given choice of epsilon. Now in order to satisfy our convergence condition,

$$2 \log n \leq f \frac{1}{g^2} \tag{42}$$

we will select $f(n)$ to be of the form $2(\log n)^{1+2\beta+\delta}$ where $g(n)$ is $(\log n)^{\beta}$. $\delta > 0$ is some arbitrarily small parameter so that our condition is satisfied in the limit, and that we are keeping as few edges as possible.

We realize that adding this coefficient of 2 will make our choice of $f(n)$ twice as large, so it will no longer always be less than or equal to n^{ϵ}. It does however present a good way to choose a power of $\log n$ so that it will work better than a given n^{ϵ}. There may exist, then, a range on which one should use n^{ϵ} and another range on which $2(\log n)^{\lambda}$ will be a better option. Consider though, that the choice of $f(n) = n^{\epsilon}$ and $g(n) = n^{\eta}$, using our convergence condition, will generate a favorable probability only when $\frac{n^{\epsilon - 2\eta}}{2} > \log n$.

We should also note however that choosing $\lambda \leq e\epsilon$ will not converge for $\epsilon < \frac{1}{e}$ simply because $\lambda > 1$ is required since $\log n < \frac{f}{2g^2}$. Using our convergence condition, a larger ϵ with a small η should work well, and should overcome $\log n$ more quickly. We expect that smaller choices of ϵ will require a large number of vertices to converge; a power of $\log n$ will again be larger but will also work for smaller numbers of vertices. Unfortunately, we haven't had time to thoroughly investigate this case, and so we must leave that question to future mathematicians.

To examine when $f(n) = n^{\epsilon}$ is a better choice than $2(\log n)^{\lambda}$ we study two conditions. Again using $\alpha = \frac{\lambda}{\epsilon}$,we require that

$$2^{\frac{1}{\epsilon}}(\log n)^{\alpha} \geq n. \tag{43}$$

Letting $\mu = \epsilon - 2\eta$ we require also that

$$\begin{aligned} \frac{n^{\mu}}{2} &\leq \log n \\ n &\geq 2^{\frac{1}{\mu}}(\log n)^{\frac{1}{\mu}} \end{aligned} \tag{44}$$

This gives the equation:

$$2^{\frac{1}{\epsilon}}(\log n)^{\alpha} \geq n \geq 2^{\frac{1}{\mu}}(\log n)^{\frac{1}{\mu}} \tag{45}$$

It is clear to see that if there exists an n that satisfies this, then we can deduce for this n that

$$\begin{aligned} 2^{\frac{1}{\epsilon}}(\log n)^{\alpha} &\geq 2^{\frac{1}{\mu}}(\log n)^{\frac{1}{\mu}} \\ 2^{\frac{1}{\epsilon}-\frac{1}{\mu}} &\geq (\log n)^{\frac{1}{\mu}-\frac{\lambda}{\epsilon}} \\ 2^{\frac{\frac{1}{\epsilon}-\frac{1}{\mu}}{\frac{1}{\mu}-\frac{\lambda}{\epsilon}}} &\geq \log n \\ 2^{\frac{\mu-\epsilon}{\epsilon-\lambda\mu}} &\geq \log n \end{aligned} \tag{46}$$

Knowing $\mu = \epsilon - 2\eta$,

$$2^{\frac{-2\eta}{\epsilon-\lambda\mu}} \geq \log n \tag{47}$$

Now we need an assumption. We will assume that we want n at least e^2, which means that n is roughly 7.39. In order for this above inequality to be satisfied for $n \geq e^2$ we require that,

$$\begin{aligned} \frac{-2\eta}{\epsilon-\lambda\mu} &\geq 1 \\ -2\eta &\geq \epsilon - \mu\lambda \\ \frac{\epsilon+2\eta}{\mu} &\leq \lambda. \end{aligned} \tag{48}$$

Making a substitution $\lambda = 1 + 2\beta + \delta$, our general choice for λ becomes:

$$\begin{aligned} \frac{\epsilon + 2\eta}{\epsilon - 2\eta} &< 1 + 2\beta + \delta && (49) \\ \frac{\epsilon - 2\eta}{\epsilon - 2\eta} + \frac{4\eta}{\epsilon - 2\eta} &< 1 + 2\beta + \delta && (50) \\ \frac{4\eta}{2(\epsilon - 2\eta)} &< \beta + \frac{\delta}{2}. && (51) \end{aligned}$$

Now recalling our definition that δ can be arbitrarily small, we accept $\frac{\delta}{2}$ as negligible and require

$$\frac{2\eta}{\epsilon - 2\eta} \leq \beta. \tag{52}$$

This condition is only for $n > e^2$, though it could be easily modified for other numbers. Although n is not required to be within these ranges, we know that for some range both

$$\begin{aligned} 2^{\frac{1}{\epsilon}}(\log n)^{\alpha} &\geq n \\ 2^{\frac{1}{\mu}}(\log n)^{\frac{1}{\mu}} &\leq n && (53) \end{aligned}$$

are satisfied. So over the range of all numbers n such that $n > e^2$, if both of the inequalities in (53) are satisfied then n^{ϵ} is a better choice for $f(n)$. It is fairly plausible that a range where n^{ϵ} is a better choice for $f(n)$ exists, but (52) only gives us a necessary condition for this existance and not a sufficient one.

Consider the equation $2^{\frac{1}{\epsilon}}(\log n)^{\alpha}$ at $n = e^2$ and suppose

$$2^{\frac{1}{\epsilon}}(\log n)^{\alpha} > e^2. \tag{54}$$

Choosing β accordingly, there exists a range on which n^{ϵ} is a better choice for $f(n)$. So n is smaller than this larger logarithmic expression at e^2, but we know n will be greater in the limit, so a trivial application of the intermediate value theorem can be used to show that n will be between the two logarithmic expressions for some range, which will be the range where a power of n is a better choice for $f(n)$. This simple evaluation can turn our condition into a sufficient condition, but this relationship at $n = e^2$ is not guaranteed for all choices of powers.

5 Examples of smaller choices being sufficiently effective

Now, to support our new upper bound and to illustrate the use of our paper, we will give an example that applies what we have done. We will show how to choose a power of the logarithm so that it will be better than a given n^ϵ.

First choose ϵ (keeping in mind that $\epsilon > \frac{1}{e}$); for our calculation we will use $\epsilon = \frac{1}{2}$. To choose our η, we want to find a choice of $f(n)$ and $g(n)$ so that it always gives better results than our choice of n^ϵ. Recall that we previously calculated that for $(\log n)^\lambda$ to be less than n^ϵ, we wanted $\lambda \leq e\epsilon$, so we choose λ to be $1+2\beta+\delta$, and solve for β, ignoring δ for now. β is approximately $.1795$ at its maximum, so to give $f(n)$ a margin, we choose $\delta = .001$, making $2\beta + \delta = .36$ and call $1 + 2\beta + \delta = 1.36$. So we have chosen $f = 2(\log n)^{1.36}$, and $g = (\log n)^{.1795}$. Now we choose our η according to our condition from the last section,

$$\frac{2\eta}{\epsilon - 2\eta} \leq \beta. \tag{55}$$

Recalling that this condition was derived for when $n^\epsilon < \log^\lambda n$, and now we want our logarithmic power to be better for all n, so we need to switch the direction of the inequality, which gives us:

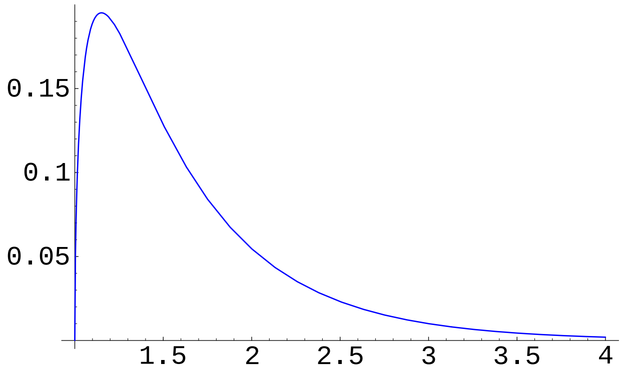
$$\frac{2\eta}{\epsilon - 2\eta} > \beta. \tag{56}$$

Solving for η,

$$\eta \geq \frac{\beta\epsilon}{2(1+\beta)} \tag{57}$$

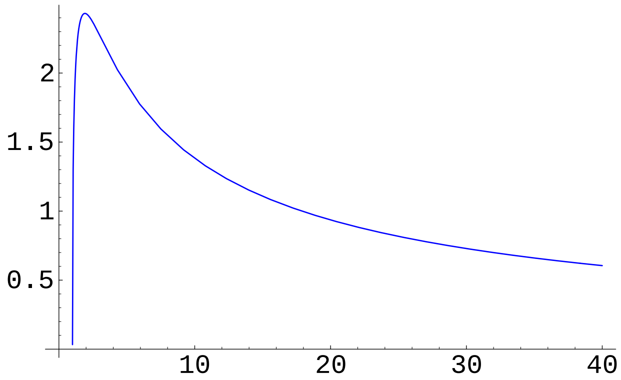
Now plugging in $\beta = .1795$, $\epsilon = .5$ we get that $\eta \geq .03805$ (approximately, this number is rounded up). This means that we can certainly choose a fairly small η so that $\epsilon > 2\eta$, and still be guaranteed that our choice $f = 2(\log n)^{1.36}$ will work better for all n.

Now with the help of Mathematica, we provide the following graphs of what we can expect from choosing such an $f(n)$ and $g(n)$. Here is a graph of the upper bound on the probability of choosing a bad graph:



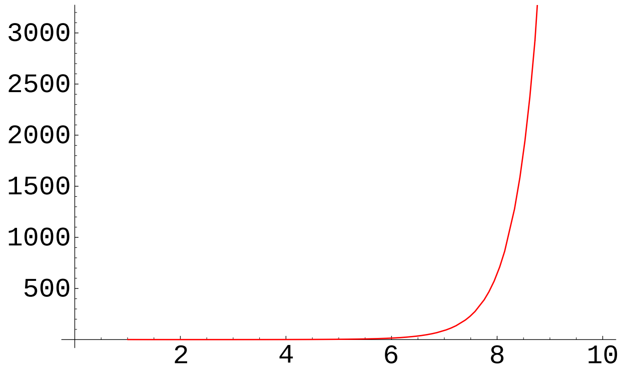
This graph illustrates numerically that the probability of choosing a bad graph becomes very small very quickly.

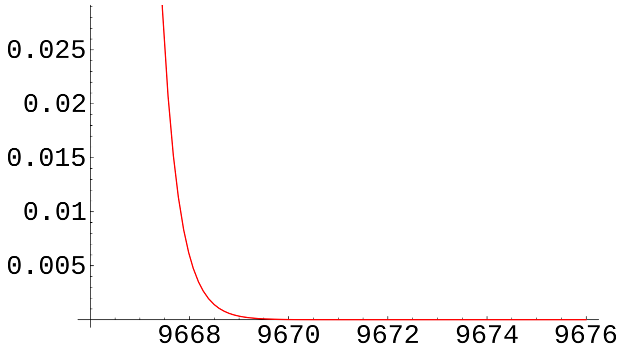
Below is the graph of our other concern; that the number of edges is less than the complete graph. This graph is of the ratio of $\frac{nf}{\binom{n}{2}}$, which we want to be less than 1.



We see from Mathematica that this ratio falls below 1 around x=18. (This is not clear in the graph, necessarily, but running a "For" loop and looking at values at successive integers, we can see that 18 is the first value for which this ratio is less than 1).

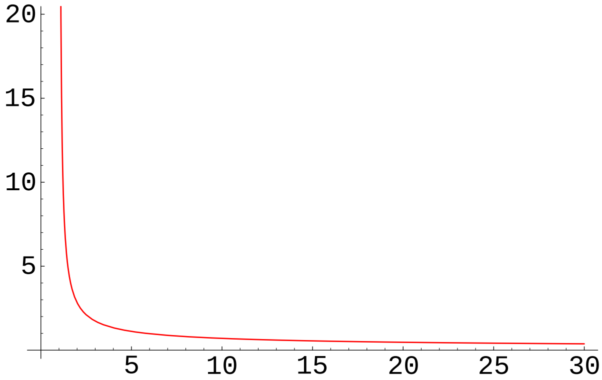
Now we look at the data from the choice of $f(n) = n^{.5}$ and $g(n) = n^{.1}$. Here is a graph of the upper bound on the probability of choosing a bad graph:





This graph illustrates numerically that the probability of choosing a bad graph becomes very large around 4 or 5 (these numbers are actually from a for loop that showed values at successive integers) and becomes small again around 9667. These are in two separate graphs because it's too large a range too see very well.

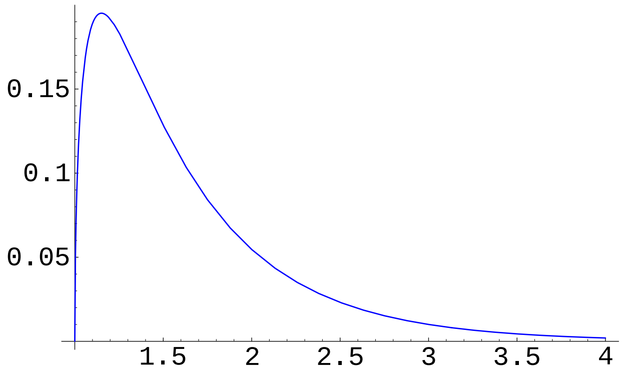
Here we show a graph of the ratio of number of edges, $n^{1+\epsilon}$, to the complete graph, $\binom{n}{2}$.



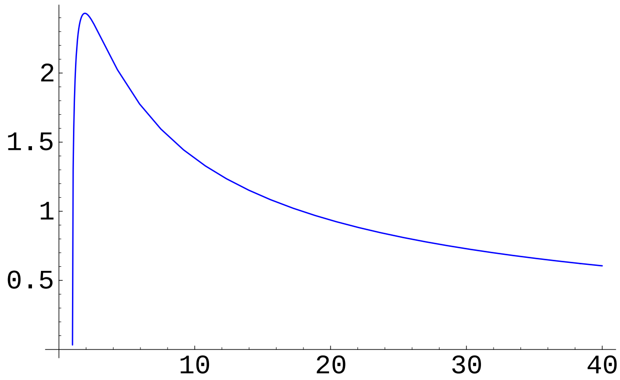
While this is at first fewer edges, the probability generated for the range when $n^{.5}$ is smaller than $2(\log n)^{1.36}$ is not favorable. The number of vertices when $2(\log n)^{1.36}$ becomes smaller than $n^{.5}$ is around 640, however at this point $2(\log n)^{1.36}$ is still giving us a much better probability and so we will continue to use it. This graph is an example of when a power of $\log n$ is always a better choice than a given n^{ϵ}.

Now we present an example of a range with n^{ϵ} being a better choice. We will use the same choices as above, but we will choose η differently. We are choosing: $f(n) = 2(\log n)^{1.36}$, $g(n) = (\log n)^{.1795}$, and $\epsilon = .5$. We calculated earlier that this choice for the power of $\log n$ would always be better as long as $\eta > .03805$. Now we choose $\eta = .03$, and so we may see a range. In fact, again utilizing Mathematica, we can see a range when $n^{.5}$ is clearly a better choice. $n^{.5}$ becomes larger than $2(\log n)^{1.36}$ at $n = 640$. However we see that as we make η smaller, a favorable probability in our upper bound is generated much earlier, around $n = 127$. So in the range $127 \leq n \leq 640$, it is better to use small powers of n. We include the graphics for these as we did previously.

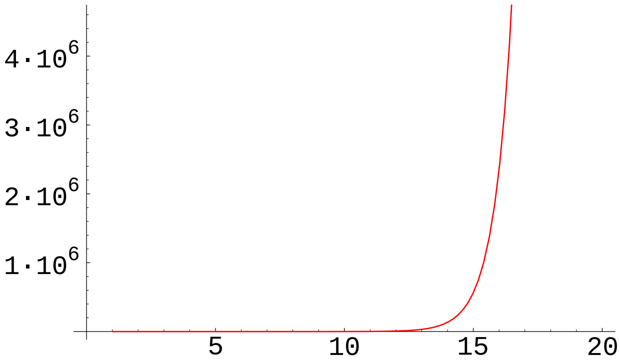
Here is a graph of the upper bound on the probability of choosing a bad graph using $f(n) = 2(\log n)^{1.36}$, $g(n) = (\log n)^{.1795}$ the same as before because we used the same choices :

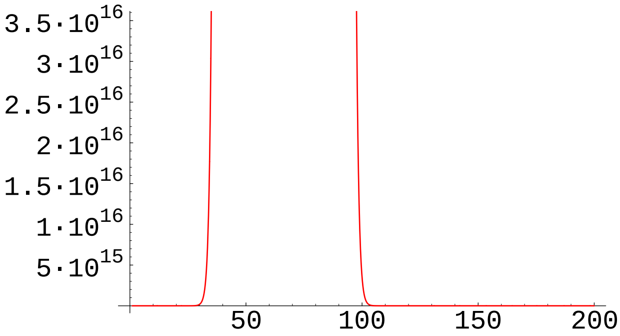


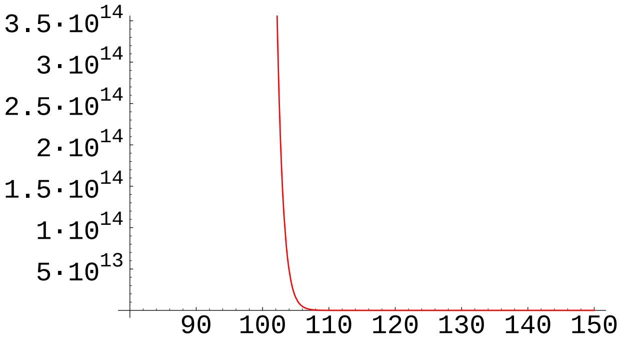
Here is the graph of the percent of the complete graph, again the same choices for $f(n)$, $g(n)$ as above:



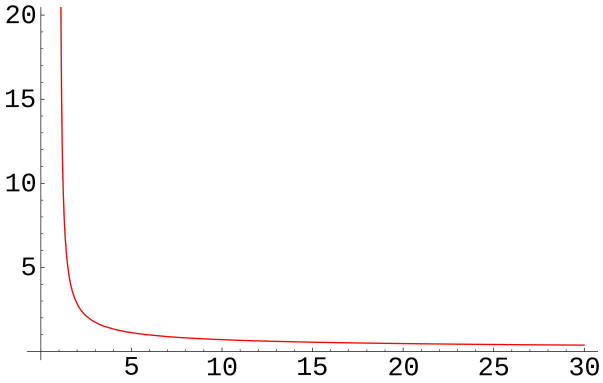
Now we look at the data from the choice of $f(n) = n^{.5}$ and $g(n) = n^{.03}$. Here is a graph of the upper bound on the probability of choosing a bad graph, note this looks similar but becomes favorable a great deal earlier than our previous choice:







Here we show a graph of the ratio of number of edges, $n^{1+\epsilon}$, to the complete graph, $\binom{n}{2}$.



6 Simple exercises

These are some simple exercises which deal with some important aspects of our study. We have included them to help the reader understand our paper.

1. Show that $\lim_{n\to\infty} \frac{\log n}{n^\epsilon} = 0$ for any $\epsilon > 0$.
2. Show that the inequalities in (1) still hold with $m = nf(n)$ and $p = \frac{n}{g(n)}$.
3. Show that in the limit that $n \longrightarrow \infty$, Sarnak's bounds satisfy our inequality: $2\log n \leq f\frac{1}{g^2}$.
4. Prove that

   - $(\log n)^3 > n$ for some $n \geq 1$.
   - $(\log n)^2 < n$ for all n.

   (hint: find where the derivatives are equal if ever. why?)

5. Find expressions for f and g for arbitrary powers of $\log n$. (You may ignore the concern of making nf less than $\binom{n}{2}$)

6. Show that a similar expression for f and g can be derived by substituting p^n for 2^n in equation 2.

7 Errors in *Some Applications of Modular Forms*

As noted in section 1, there were some small mistakes in [Sar]. Here is a list of revisions, along with a small explanation of why these are incorrect, to help the readers follow the book more smoothly.

1. Line 1 of equation (3.1.4)[Sar] has $\binom{\binom{p}{2}}{m}$,where it should read $\binom{\binom{p}{2}}{n}$ because the number of edges that he is choosing in the independent set only goes up to n. And so the number of ways to choose n edges in a set of size p is simply the number of ways to choose n edges out of the number of possible edges, which is $\binom{\binom{p}{2}}{n}$. This mistake becomes critical for convergence, and it can be shown that the final result does not follow with this small typographical error. The proof that this error changes the result is too long to add here but it can be shown using Stirling's approximation. Stirling's approximation formula says that x! asymptotically approaches $c(2x\pi)^{\frac{1}{2}}x^xe^{-x}$. Use this formula to approximate the product in Line 1 of (3.1.4), multiplied by $\frac{\binom{n}{p}}{\binom{\binom{n}{2}}{m}}$. Then take the logarithm and factor into a sum of the exponents, each multiplied by a logarithmic part, and it can be shown from there that all the logarithms are of quantities greater than 1. Some of the exponents go to infinity, and will overpower any constant. What is shown is that in the limit, the quantity from Line 1 of (3.1.4) over $\binom{\binom{n}{2}}{m}$ goes to infinity. Since this ratio can be expressed as an exponential, the power of which is the sum of positive, increasing quantities, it does not give us a useful upper bound.

2. In Line 2 of equation (3.1.4)[Sar] it reads $\binom{\binom{n}{2}-\binom{m}{2}}{m}$ where it should read $\binom{\binom{n}{2}-\binom{p}{2}}{m}$, as would follow from his previous line. This incorrect version

does not make sense because $n < m$ and therefore $\binom{n}{2} < \binom{m}{2}$ which would be choosing m edges from a negative number of possible edges, which is clearly not possible.

3. The third small error is that Sarnak has $\left(1 - \frac{\binom{p}{2}}{\binom{n}{2}}\right) < \left(1 - \frac{p^2}{n^2}\right)$, which implies $\frac{p^2-p}{n^2-n} > \frac{p^2}{n^2}$ from which we get

$$\frac{p-1}{n-1} > \frac{p}{n} \tag{58}$$

$$np - n > pn - p \tag{59}$$

$$n < p. \tag{60}$$

This is not true for our choices of p and n. This mistake however is more of a technicality because the error in this term is very small and vanishes as $n \to \infty$. To show how minor this error was:

$$\frac{p-1}{n-1} > \frac{p}{n} \tag{61}$$

$$\frac{p}{n-1} - \frac{1}{n-1} > \frac{p}{n} \tag{62}$$

It is easy to see that this inequality almost becomes true except for the $\frac{1}{n-1}$, but because this disappears in the limit $n \to \infty$, these two quantities approach equality, though never actually achieving it. Which means that this error is extremely minor and so the arguments that Sarnak makes still work.

References

[DSV] G. Davidoff, P. Sarnak and A. Valette, *Elementary Number Theory, Group Theory, and Ramanujan Graphs*, London Mathematical Society, Student Texts **55**, Cambridge University Press, 2003.

Note to the Reader: This book provided the correct argument which we have used for our upper bound.

[Sar] P. Sarnak *Some Applications of Modular Forms*, Cambridge Trusts in Math. **99**, Cambridge University Press, 1990.

[Kar] Karatsuba Anatolij, *Basic Analytic Number Theory*, Translated by Melvyn B Nathanson. New York: Springer-Verlag,1993.