The Effect of Random Edge Removal on Network Degree Sequence

Thomas DuBois[*]
University of Maryland
tdubois@cs.umd.edu

Stephen Eubank[†]
Virginia Tech
seubank@vbi.vt.edu

Aravind Srinivasan[‡]
University of Maryland
srin@cs.umd.edu

Abstract

Many networks arise in a random and distributed fashion, and yet result in having a specific type of degree structure: e.g., the WWW, many social networks, biological networks, etc., exhibit power-law, stretched exponential, or similar degree structures. Much work has examined how a graph's degree-structure influences other graph properties such as connectivity, diameter, etc. Probabilistic edge removal models link failures, information spreading, and processes that consider (random) subgraphs. They also model spreading influence of information as in the independent cascade model [20]. We examine what happens to a graph's degree structure under edge failures where the edges are removed independently with identical probabilities. We start by analyzing the effect of edge failure on the degree sequence for power-law and exponential networks, and improve upon results of Martin, Carr & Faulon and Cooper & Lu; then, using intuition from the power-law case, we derive asymptotic results for almost any degree sequence of interest. Our major result shows a classification of degree sequences which leads to simple rules that give much of the new expected degree sequence after random edge-removal; we also provide associated concentration bounds.

Keywords: random graph, degree distribution

1 Introduction

Many processes can cause the edges of a given graph G to fail independently, say with identical probabilities. This classical idea includes the original random-graph model $G(n, p)$

---

of Erdős & Rényi [18], where the host graph G being the complete graph. This topic has received much attention recently in the case where the host graph G is not complete, but comes from some interesting family of graphs. Motivations for studying random edge-removal are many: e.g., link failure in peer-to-peer networks [3], instant-messaging or the fact that we can only sample (random) subgraphs of massive graphs [10], pruning of relationships in online social networks, and disease propagation or network attacks [25].

Many properties such as diameter, emergence of the giant component, and various spectral parameters (e.g., spectral gap, mixing, expansion) have been studied when edge removal is conducted on graphs from various natural families [10, 18, 9, 8, 1]. Another important parameter of a graph family is its degree sequence. Much work has focused on various properties of random graphs with given (expected) degree sequences [11, 14, 13]. For instance, Chung & Lu [12] show how the size of such a random graph's likely giant-component depends on the average expected degree and the second order average degree.

The key role often played by the degree sequence leads us to the question: what happens to the degree sequence of a graph family when edges are removed independently with identical probabilities? If the initial graph has $d(j)$ vertices of degree j and edges are each removed with probability p, there are essentially $d'(k) \doteq \sum_{j=k}^{\infty} d(j) \cdot \binom{j}{k} \cdot p^{j-k}(1-p)^k$ expected nodes of final degree k. Note that this is not a sum of independent variables, but that the final degree of any node is intertwined with the final degree of its neighbors. We provide a characterization of how this new distribution relates to d for most interesting graph families. We also provide concentration bounds that show that the degree sequence matches the sequence d' with high probability.

Let us start by discussing two basic types of degree sequences. It is known that real-world, large-scale graphs often have a degree structure similar to power-law. These graphs include the World Wide Web [5, 6, 22], internet routers [19], many social networks including scientific co-authorship [4], as well as biological networks such as protein-protein interaction graphs [28]. Some graphs previously identified as being power-law may have degree sequences closer to stretched exponential, log-normal, or other. For a thorough examination of many empirical datasets, see Clauset et al. [15]. We will focus on power-law sequences as a concrete example, and later show that the analysis works for essentially all other interesting sequences as well.

A graph has a power-law structure if there exist parameters α and γ such that for essentially all k the number of nodes of degree k is $\sim \alpha k^{-\gamma}$. To accommodate small deviations, we define a family of power-law-like networks such that for some interval I of degrees, the number of nodes of degree $k \in I$ is within a constant factor of power-law (i.e., $\Theta(\alpha k^{-\gamma})$). What happens to the degree structure of the graph family when edges are removed independently, with some probability p? Does a power-law-like graph remain power-law-like? If so, how do its parameters vary, and how good is the concentration of the number of final nodes with some degree k? What about graphs with exponential degree sequences, where the degree sequence decays exponentially starting at some minimum degree of interest?

Callaway et al. [7] address some of these questions using generating functions to show how random edge removal affects the existence of the giant component for original random

graphs of arbitrary degree distribution. Martin et al. [25] is a more direct predecessor of our work. They considered the special case of $p = 1/2$ and $\gamma = 2$ for power-law graphs, and derived empirical results demonstrating that removing edges with probability $\frac{1}{2}$ from a power-law graph gives an expected new degree structure which is close to power-law. More recently Cooper and Lu [17] have shown that only power-law distributions are "scale-free", in the sense that a random subgraph of a power-law graph is likely to be "scale-free" also. (Technically, the work of [17] is on "site percolation", wherein vertices are removed randomly, while those of [7, 25] and ours are on "bond percolation", wherein edges are removed. Because our bounds for $E[d_k]$ rely on a sum over the expected contribution from each node and every node not removed by site percolation has the same behavior as it does in bond percolation, our qualitative statements about expected degree sequences apply to the bond percolation context as well.)

In Sections 2 we prove that after random edge removal, power-law-like networks retain expected degree sequences that are power-law-like with the same value of γ. We do this through a careful algebraic argument where we prove that the vast majority of nodes with final degree k have initial degrees near $k/(1-p)$ and that for any two initial degrees in this range, they have close enough to the same number of nodes for each. We go on to show in Section 3 how the value p changes the exponent in exponential degree sequences. In both cases we get explicit, analytical bounds by combining a binomial distribution, a step function, and the degree distribution. These results are in contrast with the three works mentioned above, as follows. Given the general generating-function bounds of [7], one needs to evaluate numerically the parameters of interest, in some cases by iterated numerical evaluation. The main results of [25] are largely empirical, and focus on the case where $p = \frac{1}{2}$ and $\gamma = 2$. Finally, the generating function based work of [17] requires that the graph have *bounded* degree D, and that for all $k \le D$, the number of nodes of degree k is *precisely power-law* up to a lower-order term: i.e., $(c + o(1))nk^{-\gamma}$ for some constant c, whereas we allow a multiplicative-constant deviation from strict power-law behavior. (We also improve upon [17] in a tail bound, as described below. Note, however, that given a precise "$(c + o(1))nk^{-\gamma}$" bound as above, the elegant work of [17] shows that only such degree sequences remain purely "scale-free" under edge removal.)

In Section 4 we give a unified view of how degree distributions change with edge removal. Specifically, for any general, non-increasing degree distribution where $d(j)$ nodes have degree j, the degree distribution after edge removal can be classified based on the limit behavior of $d(j)/d(j-1)$. We study three main classes based on whether $\lim_{j \to \infty} d(j)/d(j-1)$ exists and is bounded away from both 0 and 1, approaches 1, or approaches 0. For $d(j)/d(j-1)$ bounded away from 0 and 1, the distribution is exponential and after edge removal the new distribution is a different exponential; for $d(j)/d(j-1)$ approaching 0, the new distribution $d'(k)$ is $\Theta((1-p)^k d(k))$; and for $d(j)/d(j-1)$ approaching 1 quickly enough, the new distribution is $\Theta(d(j))$. This final result means that *for almost any distribution of interest, random edge-removal does not affect the distribution type, only the "scale" of the distribution.* Thus whether a network is power-law or has a similar distribution does not matter (which is useful due to the subtleties in distinguishing such distributions [24, 26]), and the relationship between the expected distribution in

the new network and the original distribution can be easily determined.

The above discussion focuses on the expected number of nodes of degree k in the graph after random edge removal, for an arbitrary but *single* value of k. How well is the final degree sequence concentrated around this target? In Section 5, we demonstrate a constant factor concentration, by a careful grouping of the vertices and through Martingale inequalities. One further way in which we improve upon [17] here is as follows: they require that the sum of the degrees-squared in the original graph be $O(n^{2-\Omega(1)})$ for their tail-bounds to hold. We require something much weaker than the analog of this for bond percolation. In Section 6 we discuss the effect of edge removal on probabilistic degree distributions: in models like those proposed by Chung and Lu [12] or Leskovec et al. [23] edges in a graph are already specified probabilistically. Thus nodes do not have degrees until after a graph is instantiated; instead they have expected degrees. We briefly define what it means for such a graph to have a power-law or exponential distribution, and show what effect edge-removals have on the distributions.

Thus, this work conducts a systematic study of what happens to a fundamental parameter of a graph, its degree sequence, under edge removal. The general flavor of the results obtained is that for essentially all of the interval of degrees of interest, the degree sequence retains its qualitative character, and several quantitative aspects as well.

2 Graphs with Power-Law Degree Sequences

In this section we derive specific bounds for how a power-law-like degree sequence changes through independently random edge removal. Letting $\deg(v)$ denote the degree of a node v, we define a graph with n vertices to be power-law-like with parameters $c_1, c_2, d_{\min}, d_{\max}$, and γ if $\forall i \in [d_{\min}, d_{\max}]$, $|\{v | \deg(v) = i\}| \in [\frac{c_1 n}{i^\gamma}, \frac{c_2 n}{i^\gamma}]$. In doing so, we develop the intuition and details which lead to the general categorization presented later.

Define $\deg(v)$ to be the degree of any vertex v. Given any graph G that has a power-law-like distribution on its node degrees we remove each edge independently with probability p. We show for all degrees k in the range $R = [d_{\min} \cdot (1-p), d_{\max} \cdot (1-p)]$, that there are an expected $\Theta(nk^{-\gamma})$ nodes of degree k in the new graph. The size of the new graph's power-law-like degree range is the original graph's power law range scaled down by a factor of $1-p$. Note that this places an implicit limit on how close p can be to 1 and still have a largely power-law new graph. If $(1-p)(d_{\max} - d_{\min})$ is small then the range of degrees for which the power-law-like property holds will be small as well. This makes intuitive sense because if p is small, only a few edges are removed and the graph should not change much. If p is close to one, then most edges are removed, the graph changes drastically, power-law-like behavior only remains for a small range of nodes.

We define the random variable d_k to be the number of vertices with degree k after edge removal. This leads to the equation:

$$\sum_{j=\max(d_{\min},k)}^{d_{\max}} \binom{j}{k} p^{j-k}(1-p)^k \frac{c_1 n}{j^\gamma} \leq E[d_k] \leq \sum_{j=\max(d_{\min},k)}^{d_{\max}} \binom{j}{k} p^{j-k}(1-p)^k \frac{c_2 n}{j^\gamma}. \quad (1)$$
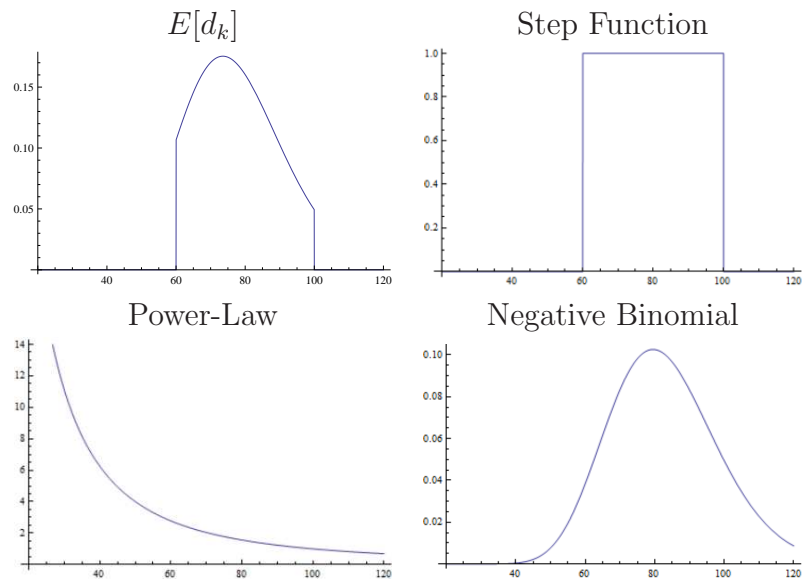
Figure 1: Examples of the three functions which make up g_k for $k = 20, d_{\min} = 60, d_{\max} = 100, c_1 n = 10000, \gamma = 2, p = .75$.

This equation comes from taking a sum over all degrees $j \geq k$ of the number of nodes with degree j times the probability that this node has degree k after the edge removal. If the distribution is exactly power-law ($c_1 = c_2 = c$, $d_{\min} = 1$, and $d_{\max} = \infty$) this equation reduces to $E[d_k] = \sum_{j=k}^{\infty} \binom{j}{k} p^{j-k} (1-p)^k cnj^{-\gamma}$, which is commonly used elsewhere [25, 17, 16]. If $c_1 \neq c_2$ then we can use c_1 throughout and introduce only a constant factor error term of $\frac{c_2}{c_1}$ which can be applied to the upper bound at the end, so for the remainder of our analysis we assume that $c_1 = c_2$.

Based on (1) we define for all $k \in R$ the functions g_k and f_k. Specifically we define $g_k(j)$ for all $j \geq k$ to be the "j term" in the summation for $E[d_k]$:

$$g_k(j) = \binom{j}{k} p^{j-k} (1-p)^k c_1 n j^{-\gamma}. \tag{2}$$

We view $g_k(j)$ as the product of three separate functions. First is a step function which is 1 when $d_{\min} \leq j \leq d_{\max}$ and 0 otherwise. The second is a power-law function giving the number of nodes of degree j: $c_1 n j^{-\gamma}$. Finally we have the function that gives the probability of a node with degree j having final degree k. The type of degree distribution only affects the second of these functions; the other two are degree distribution independent. Examples of the three functions appear in Figure 1.

The intuition behind our method comes in two parts. First, for any given positive σ the third component function (which is almost a negative binomial probability distribution) has the vast majority of its mass (all but an amount exponential in $-\sigma$) within an $O(\sigma \cdot \frac{\sqrt{k}}{1-p})$ range around $j = \frac{k}{1-p}$. And second, that within this range the value of $j^{-\gamma}$ does not change much.

We also define the function $f_k(j)$ for $j \geq k$ to be the ratio of successive terms in the sum $E[d_k]$ such that

$$f_k(j) = g_k(j+1)/g_k(j) = \frac{pj^\gamma}{(j+1-k)(j+1)^{\gamma-1}}. \tag{3}$$

This leads to the following lemma:

Lemma 1 $\forall k > \gamma$, the function $f_k(j)$ is strictly decreasing as j increases.

The proof proceeds by treating f_k as a continuous function, and showing that the derivative is always negative.

Proof of Lemma 1 Taking the derivative of $f_k(j)$ with respect to j yields:

$$
\begin{aligned}
f_k'(j) &= \frac{\gamma p j^{\gamma-1}}{(j+1-k)(j+1)^{\gamma-1}} - \frac{pj^\gamma \left((j+1)^{\gamma-1} + (j+1-k)(\gamma-1)(j+1)^{\gamma-2}\right)}{(j+1-k)^2(j+1)^{2\gamma-2}} \\
&= \frac{\gamma p j^{\gamma-1}(j+1-k)(j+1)^{\gamma-1} - pj^\gamma(j+1)^{\gamma-1} - pj^\gamma(j+1-k)(\gamma-1)(j+1)^{\gamma-2}}{(j+1-k)^2(j+1)^{2\gamma-2}} \\
&= \frac{pj^{\gamma-1}(j+1)^{\gamma-2}}{(j+1-k)^2(j+1)^{2\gamma-2}} \left(\gamma(j+1-k)(j+1) - j(j+1) - j(j+1-k)(\gamma-1)\right).
\end{aligned} \tag{4}
$$

The fractional part of (4) is strictly positive. Therefore $f_k'(j)$ has the same sign as

$$\gamma(j+1-k)(j+1) - j(j+1) - j(j+1-k)(\gamma-1) \leq j\gamma - jk < 0.$$

Since the derivative of $f_k(j)$ with respect to j is always negative when $k > \gamma$, $f_k(j)$ always decreases as j increases. ∎

Using the fact that f_k is strictly decreasing we next find the maximum value of g_k for all $k > \gamma$. When $f_k(j) \geq 1$, g_k increases from j to $j+1$ and when $f_k(j) < 1$, g_k decreases from j to $j+1$. Therefore the maximum term $j_{\max}$ in (1) occurs within one of where $f_k(j)$ crosses from above 1 to below 1. We treat $f_k(j)$ as a continuous function and round to find the desired j:

$$1 = f_k(j) = \frac{pj}{j+1-k} \cdot \frac{j^{\gamma-1}}{(j+1)^{\gamma-1}} \quad \text{i.e., } j = \frac{k-1}{1 - p\left(j/(j+1)\right)^{\gamma-1}}. \tag{5}$$

It immediately follows from (5) that $j_{\max}$ is always less than $\frac{k-1}{1-p}$ for any $\gamma > 1$, which is inherent in a power-law distribution. Furthermore, since $j \geq k$, $\frac{j}{j+1}$ is typically close to 1, thus $j_{\max}$ is typically close to $\frac{k-1}{1-p}$ as well. This observation drives our focus on $j \approx \frac{k}{1-p}$ from here on. Specifically we will use $g_k(\frac{k}{1-p})$ as a proxy for the more elusive $g_k(j_{\max})$.

We proceed by showing bounds on $g_k(\frac{k}{1-p})$, using the error factor $\xi = 1 - O(1/(pk))$.

$$g_k(\frac{k}{1-p}) = \frac{\left(\frac{k}{1-p}\right)!}{k!\left(\frac{pk}{1-p}\right)!} p^{pk/(1-p)}(1-p)^k c_1 n \left(\frac{k}{1-p}\right)^{-\gamma}$$

$$= \frac{\sqrt{2\pi\frac{k}{1-p}} \cdot \left(\frac{k}{1-p}\right)^{\frac{k}{1-p}} e^{\lambda_{k/(1-p)} - \lambda_k - \lambda_{pk/(1-p)}}}{\sqrt{2\pi k} \cdot \sqrt{2\pi\frac{pk}{1-p}} \cdot k^k \left(\frac{pk}{1-p}\right)^{pk/(1-p)}} p^{\frac{pk}{1-p}}(1-p)^k c_1 n \left(\frac{k}{1-p}\right)^{-\gamma} \cdot \xi \quad (6)$$

$$= \frac{(1-p)^\gamma}{\sqrt{2\pi p}} \cdot c_1 n \cdot k^{-\gamma-.5} \cdot \xi \quad (7)$$

where (6) relies on Stirling's approximation for factorials: $n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot e^{\lambda_n}$ where $\lambda_n \in [1/(12n+1), 1/(12n)]$.

Next we use $j = k/(1-p)$ as a starting point and bound how much g_k changes around this point. To that end we derive the bounds $f_k((k+\sigma\sqrt{k})/(1-p)) \leq 1-((1-p)\sigma)/(p\sqrt{k}+\sigma)$ and $g_k(\frac{k+\sigma\sqrt{k}}{1-p})/g_k(\frac{k}{1-p}) \leq e^{p\gamma/(1-p)-|\sigma|/8}$ for both positive and negative σ.

We have already shown that $g_k(\frac{k}{1-p}) \approx \left(\frac{(1-p)^\gamma}{\sqrt{2\pi p}}\right) c_1 n k^{-\gamma-.5}$ with very small error and that $\frac{k}{1-p}$ is close to $j_{\max}$. In this section we show bounds on how quickly g_k and f_k change as j moves away from $\frac{k}{1-p}$.

We start by showing bounds on $f_k(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k})$ and $\frac{g_k(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k})}{g_k(\frac{k}{1-p})}$ for any real (positive or negative) σ such that $\sigma > -p\sqrt{k}$. We need these bounds to prove that the overwhelming amount of the probability weight of nodes with final degree k comes from nodes with initial degrees in the range $\frac{k \pm O(\sqrt{k})}{1-p}$.

First we upper bound $f_k(\frac{k}{1-p} + \frac{\sigma}{1-p}\sqrt{k})$.

$$f_k(\frac{k}{1-p} + \frac{\sigma}{1-p}\sqrt{k}) = \frac{p(\frac{k}{1-p} + \frac{\sigma}{1-p}\sqrt{k})}{\left(\frac{k}{1-p} + \frac{\sigma}{1-p}\sqrt{k} + 1 - k\right)} \frac{(\frac{k}{1-p} + \frac{\sigma}{1-p}\sqrt{k})^{\gamma-1}}{(\frac{k}{1-p} + \frac{\sigma}{1-p}\sqrt{k} + 1)^{\gamma-1}}$$

$$\leq \frac{pk + p\sigma\sqrt{k}}{pk + \sigma\sqrt{k}} = 1 - \frac{(1-p)\sigma}{p\sqrt{k}+\sigma} \quad (8)$$

Next we upper bound $\dfrac{g_k(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k})}{g_k(\frac{k}{1-p})}$ when σ is positive:

$$\frac{g_k(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k})}{g_k(\frac{k}{1-p})} = \frac{\binom{\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k}}{k}(1-p)^k p^{\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k}-k}(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k})^{-\gamma}}{\binom{\frac{k}{1-p}}{k}(1-p)^k p^{\frac{k}{1-p}-k}\frac{k}{1-p}^{-\gamma}}$$

$$= \left(\frac{(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k})!}{(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k}-k)!}\right) \Big/ \left(\frac{\frac{k}{1-p}!}{(\frac{k}{1-p}-k)!}\right) p^{\frac{\sigma}{1-p}\sqrt{k}}\left(\frac{k}{k+\sigma\sqrt{k}}\right)^{\gamma} \quad (9)$$

$$= \prod_{i=1}^{\frac{\sigma}{1-p}\sqrt{k}}\left(\frac{\frac{k}{1-p}+i}{\frac{k}{1-p}-k+i}p\right) \leq \prod_{i=\frac{\sigma}{2-2p}\sqrt{k}+1}^{\frac{\sigma}{1-p}\sqrt{k}}\left(\frac{k+\frac{\sigma}{2}\sqrt{k}}{kp+\frac{\sigma}{2}\sqrt{k}}p\right) \quad (10)$$

$$\leq \left(1-\frac{(1-p)\sigma}{2p\sqrt{k}+\sigma}\right)^{\frac{\sigma}{2-2p}\sqrt{k}} \leq e^{\frac{-\sigma^2}{4p+2\sigma/\sqrt{k}}} \leq e^{\max(\frac{-\sigma^2}{8p},\frac{-\sigma\sqrt{k}}{4})} \leq e^{-\sigma/8}. \quad (11)$$

The bound on $\dfrac{g_k(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k})}{g_k(\frac{k}{1-p})}$ when σ is negative is a little more involved. First, the $\left(\frac{k}{k+\sigma\sqrt{k}}\right)^{\gamma}$ term in (9) is greater than 1 when σ is negative, so it cannot simply be discounted. However it is at most $\left(\frac{k}{k-p\sqrt{k}\sqrt{k}}\right)^{\gamma} = \left(1+\frac{p}{1-p}\right)^{\gamma}$ which is less than $e^{p\gamma/(1-p)}$, which only depends on the fixed parameters p and γ.

The next change from positive to negative σ comes on (10) which is replaced by:

$$\frac{g_k(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k})}{g_k(\frac{k}{1-p})} \leq e^{p\gamma/(1-p)}\prod_{i=\frac{-\sigma}{2-2p}\sqrt{k}+1}^{\frac{-\sigma}{1-p}\sqrt{k}}\left(\frac{pk-i(1-p)}{pk-pi(1-p)}\right)$$

$$\leq e^{p\gamma/(1-p)}\left(1-\frac{\frac{-\sigma}{2}\sqrt{k}(1-p)}{pk-p\frac{-\sigma}{2}\sqrt{k}}\right)^{\frac{-\sigma}{2-2p}\sqrt{k}}$$

$$\leq e^{p\gamma/(1-p)-\frac{\sigma^2}{4p(1+\sigma/(2\sqrt{k}))}} \leq e^{p\gamma/(1-p)-\frac{\sigma^2}{4p}} \text{ if } \sigma > -2\sqrt{k}.$$

While our analysis runs into a discontinuity at $\sigma = -2\sqrt{k}$, this is unimportant. When $\sigma < -p\sqrt{k}$ then $\frac{k}{1-p}+\frac{\sigma}{1-p}\cdot\sqrt{k} < k$. Since there are no nodes with initial degree $< k$ that end up with degree k, in these cases it follows that $g_k(\frac{k}{1-p}+\frac{\sigma}{1-p}\cdot\sqrt{k}) = 0$.

Combining all of these bounds gives that for any meaningful σ,

$$\frac{g_k(\frac{k}{1-p}+\frac{\sigma}{1-p}\sqrt{k})}{g_k(\frac{k}{1-p})} \leq e^{p\gamma/(1-p)-|\sigma|/8}. \quad (12)$$

Since we have a dropoff in $g_k(j)$ from $g_k(\frac{k}{1-p})$ that is exponential in $(j-\frac{k}{1-p})/\sqrt{k}$, almost all (at least $1 - O(\frac{1}{n})$) of the concentration of $E[d_k]$ comes from a $\sqrt{k}\ln n$ region around $g_k(\frac{k}{1-p})$.
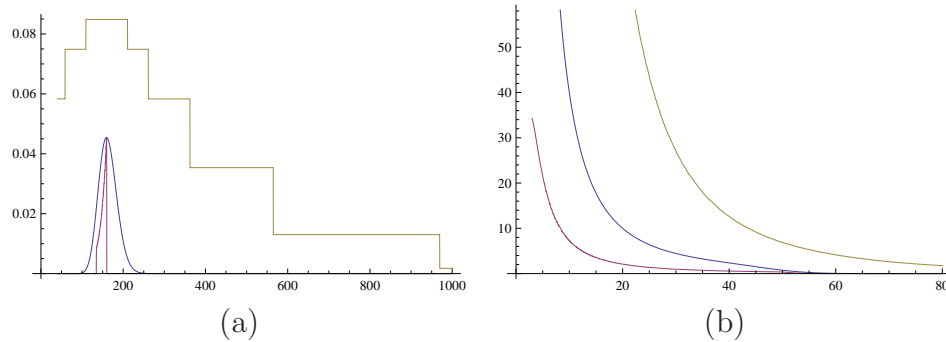
(a) (b)

Figure 2: Figure (a) shows the function g_{40} with $p = .75$, $\gamma = 2$ and $c_1 = 1000$ in blue along with our upper and lower bounds for it. Figure (b) shows $E[d_k]$ for various values of k along with our upper and lower bounds. With these parameters, the lower bound is approximately $E[d_k]/5$ and the upper bound approximately $14E[d_k]$.

From here we use the bounds on $f_k((k + \sigma\sqrt{k})/(1 - p))$ to create a geometric series whose sum is a lower bound for the $O(\sqrt{k})$ terms of g_k leading up to $g_k(k/(1-p))$, which is itself a lower bound on $E[d_k]$. We go on to use the bounds on $g_k(\frac{k+\sigma\sqrt{k}}{1-p})/g_k(\frac{k}{1-p})$ directly to upper bound $E[d_k]$. Figure 2 gives a graphical illustration of our upper and lower bounds. This leads to the following theorem:

Theorem 2 *Suppose we are given a fixed edge removal probability p and a power-law-like graph G given by constants γ, c_1, c_2 and range $d_{\min}, d_{\max}$. The degree distribution of the induced graph where each edge in G is removed with probability p is power-law-like as well. Specifically for all k in $[\frac{d_{\min}}{1-p}, \frac{d_{\max}}{1-p}]$ the term $E[d_k]$ falls between*

$$\left(\left(1 - e^{-\frac{1}{p}}\right)(p - k^{.5})(1 - p)^{\gamma-1}/\sqrt{2\pi p}\right) \cdot \left(c_1 n k^{-\gamma}/\gamma^\gamma\right) \cdot \left(1 - O(1/(12pk))\right)$$

(where the "γ^γ" term is only necessary when $k < \gamma$) and

$$\left(20e^{(p\gamma/(8(1-p)))}(1 - p)^{\gamma-1}/\sqrt{2\pi p}\right) \cdot \left(c_2 n k^{-\gamma}\right).$$

Proof of Theorem 2

Our general technique to lower bound $E[d_k]$ is to lower bound $g_k(j)$ for the $\sqrt{k}$ values of j preceding $\frac{k}{1-p}$. We do this by taking $g_k(\frac{k}{1-p})$ and successively dividing it by the intermediate $\frac{k}{1-p} - j$ values of f_k. From Eq (8) we know that these f_k values are not too large. The sum of these lower bounds, which we show to be $\Theta(c_1 n k^{-\gamma})$, is itself a lower bound on $E[d_k]$. Figure 2 shows our upper and lower bounds for both g_k and $E[d_k]$.

For any k such that $k \geq \left(\frac{1}{p}\right)^2$ and $d_{\max} \geq \frac{k}{1-p}$:

$$E[d_k] = \sum_{i=k}^{d_{\max}} g_k(i) \geq \sum_{i=\frac{k-\sqrt{k}}{1-p}}^{\frac{k}{1-p}} g_k(i) = \sum_{i=\frac{k-\sqrt{k}}{1-p}}^{\frac{k}{1-p}} \left(g_k(\frac{k}{1-p}) \prod_{j=0}^{\frac{k}{1-p}-i-1} \frac{1}{f_k(i+j)} \right)$$

$$\geq \sum_{i=\frac{k-\sqrt{k}}{1-p}}^{\frac{k}{1-p}} \left(g_k(\frac{k}{1-p}) \prod_{j=0}^{\frac{k}{1-p}-i-1} \frac{1}{1+\frac{1-p}{p\sqrt{k}-1}} \right) \quad \text{(from (8) and } f_k \text{ decreasing)}$$

$$\geq g_k(\frac{k}{1-p}) \sum_{i=0}^{\frac{\sqrt{k}}{1-p}-1} \left(1 - \frac{1-p}{p\sqrt{k}-1} \right)^i$$

$$= \left(\frac{\left(1-e^{-\frac{1}{p}}\right)(p-k^{-.5})(1-p)^{\gamma-1}}{\sqrt{2\pi p}} \right) c_1 n k^{-\gamma}. \qquad (13)$$

Note that we could equivalently bound $E[d_k] \geq \sum_{j=\frac{k}{1-p}}^{\frac{k+\sqrt{k}}{1-p}} g_k(j) \geq \Theta(\frac{pk^{-\gamma}}{1-p})$ using nearly identical steps. As long as $d_{\max} - d_{\min} > 2\sqrt{k}$ there are always enough j values either above or below $\frac{k}{1-p}$ to contribute enough probability weight to this lower bound that $E[d_k] \geq \Theta(\frac{pk^{-\gamma}}{1-p})$.

Whenever $k < \gamma$ we lose the guarantee that $f_k(j)$ is monotonically decreasing with j. However, for every term $g_k(j)$:

$$g_k(j) = \binom{j}{k} p^{j-k}(1-p)^k c_1 n j^{-\gamma}$$

$$\geq \binom{j+\gamma}{k+\gamma} p^{j+\gamma-k-\gamma}(1-p)^{k+\gamma} c_1 n (j+\gamma)^{-\gamma}$$

$$\geq g_{k+\gamma}(j+\gamma).$$

Therefore $E[d_k] \geq E[d_{k+\gamma}]$. So if $E[d_{k+\gamma}] \geq \alpha n(k+\gamma)^{-\gamma}$, then $E[d_k]$ also has a power-law expected degree sequence, since $E[d_k] \geq E[d_{k+\gamma}] \geq \alpha n(k+\gamma)^{-\gamma} \geq \frac{\alpha}{\gamma^\gamma} n k^{-\gamma}$.

For the upper bound on $E[d_k]$ we use the fact from (12) that $\forall \sigma$, $g_k(\frac{k\pm\sigma\sqrt{k}}{1-p}) \leq g_k(\frac{k}{1-p}) \cdot e^{p\gamma/(1-p)-\frac{|\sigma|}{8}}$. We group j values based upon which value of σ gives $\frac{k+(2^\sigma-1)\sqrt{k}}{1-p}$ closest to j in the direction towards $\frac{k}{1-p}$. This grouping is evident in the step-like behavior of the upper bound on g_k in Figure 2. We then bound a sum over all of these groups of an upper

bound on the total weight from each group.

$$\sum_{i=k}^{d_{\max}} g_k(i) \leq \sum_{i=-\infty}^{\infty} g_k(i)$$

$$\leq \sum_{\sigma=0}^{\infty} \sum_{i=0}^{\sqrt{k}(2^\sigma)/(1-p)-1} \left(g_k(\frac{k}{1-p} - \sqrt{k} \cdot \frac{2^\sigma - 1}{1-p} - i) + g_k(\frac{k}{1-p} + \sqrt{k} \cdot \frac{2^\sigma - 1}{1-p} + i) \right) \quad (14)$$

$$\leq \frac{\sqrt{k}}{1-p} \cdot \sum_{\sigma=0}^{\infty} 2^\sigma \left(2g_k(\frac{k}{1-p}) e^{\frac{p\gamma/(1-p)-(2^\sigma-1)}{8}} \right) \text{ from (12)}$$

$$= \left(\frac{2e^{p\gamma/(1-p)+1/8}(1-p)^{\gamma-1}}{\sqrt{2\pi p}} \right) c_1 n k^{-\gamma} \sum_{\sigma=0}^{\infty} (2^\sigma) \left(e^{\frac{-2^\sigma}{8}} \right) \cdot (1 - O(1/(12pk))) \quad (15)$$

$$= \left(\frac{20e^{p\gamma/(1-p)+1/8}(1-p)^{\gamma-1}}{\sqrt{2\pi p}} \right) c_1 n k^{-\gamma}, \quad (16)$$

where the slack from (15) comes from Robbins' approximation[27] and (16) follows because the preceding summation is less than 10.

Putting the two bounds from (13) and (16) together completes the proof. ∎

3 Exponential Degree Sequences

We follow up with a similar, though much more straightforward result for graphs with exponential degree sequences. We say a graph has a γ exponential degree sequence if there exist constants c_1, c_2 such that there are between $c_1 n \gamma^k$ and $c_2 n \gamma^k$ nodes of degree k (for the analysis we use a single constant c_1). If each edge is removed with probability p, then the expected number of nodes with final degree k is $E[d_k] = \sum_{j \geq k} c_1 n \gamma^j \binom{j}{k} p^{j-k} (1-p)^k$. In the remainder of this section we prove the following theorem:

Theorem 3 *Given a graph with a γ exponential degree sequence and an edge removal probability p such that $0 \leq p\gamma < 1$, the expected degree sequence for the resulting graph is an $\frac{\gamma(1-p)}{1-p\gamma}$-exponential.*

As with the power-law case, the terms in the summation can be broken into three components: a step function, an exponential function, and a negative binomial. However unlike the power-law case, the exponential function is significant enough that the negative binomial does not dominate it, and the algebra becomes much simpler. We can rewrite

$E[d_k]$ as

$$\sum_{j \geq k} c_1 n \gamma^j \binom{j}{k} p^{j-k} (1-p)^k = c_1 n \left(\frac{\gamma(1-p)}{1-p\gamma} \right)^k \sum_{j \geq k} \binom{j}{k} (p\gamma)^{j-k} (1-p\gamma)^k$$

$$= \frac{c_1 n}{1-p\gamma} \left(\frac{\gamma(1-p)}{1-p\gamma} \right)^k .$$

This shows directly that $E[d_k]$ is exponential in k. Specifically the new graph has a $\frac{\gamma(1-p)}{1-p\gamma}$ exponential degree sequence. Note that this result does not require that $\gamma < 1$, but only requires the weaker condition that $p\gamma < 1$. This restriction is in place because if $p\gamma \geq 1$ then the new negative binomial term is no longer a valid probability distribution.

4 General Principles

So far we have given two specific examples of what happens to a graph's degree distribution when edges are removed uniformly and independently at random. In this section we expand upon those results by developing general asymptotic principles that apply to almost any monotonically non-increasing degree distribution. We look at any degree distribution given in terms of two parts, a unit step function $s(j)$ which is one for all j between $d_{\min}$ and $d_{\max}$ and a density function $d(j)$ giving the number of nodes of degree j if $s(j) = 1$. These correspond to the first and second components of the terms of g_k, where the third component remains the negative binomial function. For now we assume $d_{\min} = 1$ and $d_{\max} = \infty$. For review, $E[d_k] = c_1 n \sum_{j=k}^{\infty} \binom{j}{k} \left(\frac{1-p}{p} \right)^k p^j s(j) d(j)$.

We classify most such distributions into three classes based upon the asymptotic behavior of $\frac{d(j)}{d(j-1)}$ in the following theorem:

Theorem 4 *For all k such that the initial degrees contribution most to nodes of final degree k are not excluded from the graphs:*

- *For exponential degree sequences where $\lim_{j \to \infty} \frac{d(j)}{d(j-1)} = \gamma$ for some $0 < \gamma < 1$, keeping edges with probability pkeep yields an expected new degree sequence of*

$$E[d_k] = \frac{c_1 n}{1-p\gamma} \cdot \left(\frac{\gamma(1-p)}{1-p\gamma} \right)^k .$$

- *For super-exponential degree sequences where $d(j)/d(d-1) = o(\frac{1}{pj})$,*

$$E[d_k] = \Theta((1-p)^k d(k)).$$

- *For sub-exponential degree sequences where $d(j)/d(j-1) = 1 \pm o(\frac{1-p}{\sqrt{j}})$,*

$$E[d_k] = \Theta(d(\frac{k}{1-p})).$$

We derive the individual cases of this theorem in the remainder of this section.

The first class includes all exponential degree distributions. The exponential case is unique in that the p^j term and the $d(j)$ term align to produce a new negative binomial distribution, plus a term exponential in k. The new negative binomial sums to a constant, leaving behind only the new exponential term.

The second class of functions we consider are super-exponential decreasing functions where $\frac{d(j)}{d(j-1)} = o(\frac{1}{pj})$. Here we have functions that decrease significantly faster than any exponential. If a function $d(j)$ decreases fast enough that asymptotically $\frac{d(j)}{d(j-1)} \leq \frac{\epsilon}{pj}$ for some $\epsilon < 1$, then for sufficiently large k,

$$(1-p)^k d(k) \leq E[d_k] = \sum_{j=k}^{\infty} p^{j-k}(1-p)^k d(j)$$

$$\leq \sum_{j=k}^{\infty} (1-p)^k d(k) \left(\frac{\epsilon}{j}\right)^{j-k} = O((1-p)^k d(k)).$$

The degree distribution decreases so rapidly in this case, that the number of nodes with final degree k is dominated by the number of nodes with initial degree k. Thus the distribution stays asymptotically the same.

Finally we consider the case where $\frac{d(j)}{d(j-1)} = 1 \pm o(\frac{1-p}{\sqrt{j}}))$, which includes all polynomial degree distributions, including power law distributions. In these cases, the maximum term in the $E[d_k]$ summation occurs at the highest j such that $\frac{d(j)}{d(j-1)} \cdot p \cdot \frac{j}{j-k} \geq 1$. Since $\frac{d(j)}{d(j-1)} \approx 1$, this occurs near $j = \frac{k}{1-p}$, which maximizes the negative binomial component function. Recall that the vast majority of the weight in a negative binomial occurs within an $O(\frac{\sqrt{k}}{1-p})$ region around its maximum. For some $x, y \in [\frac{k-\sigma\sqrt{k}}{1-p}, \frac{k+\sigma\sqrt{k}}{1-p}]$, $d(j)$ can change by at most a ratio of

$$\frac{d(x)}{d(y)} \leq \prod_{j=(k-\sigma\sqrt{k})/(1-p)+1}^{(k+\sigma\sqrt{k})/(1-p)} \max\left(\frac{d(j)}{d(j-1)}, \frac{d(j-1)}{d(j)}\right),$$

i.e., at most $\left(1 + o((1-p)/(\sqrt{k}))\right)^{2\sigma\sqrt{k}/(1-p)}$ which for sufficiently large k is at most

$$\left(1 + (1-p)/(2\sigma\sqrt{k})\right)^{2\sigma\sqrt{k}/(1-p)} \leq e^{(1-p)/(2\sigma\sqrt{k}) \cdot 2\sigma\sqrt{k}/(1-p)} = e.$$

Since for sufficiently large k, the value of $d(j)$ changes by at most a factor of e within the range for which the negative binomial component is large, we can approximate d as a constant $d(k/(1-p))$ with at most a factor of error e. Thus for this category of $d(j)$ functions, $E[d_k] = \Theta(d(\frac{k}{1-p}))$. This result holds for all power-law distributions, as well as any smooth, piece-wise combination of power-law functions, or any other slowly changing function.

In each of these three cases, the effect of the $s(j)$ step function on the summations is that of a range selector. Any k such that $\frac{k}{1-p}$ falls well within the $s(j) = 1$ region will behave as described above. Any k with $\frac{k}{1-p}$ is sufficiently far away from where $s(j) = 1$ will have very small $E[d_k]$, with $E[d_k]$ transitioning slowly between the two extremes.

5 Bounds on Large Deviations from Expected Degrees

In the previous section, we categorize how the expected degree sequence of a graph changes when its edges are removed at random. In this section we take that one step further and show that with probability at least $1 - \frac{1}{n}$, all relevant degrees k will simultaneously have $\Theta(E[d_k])$ nodes of degree k in the resulting graph as long as a few assumptions are met. The challenge in developing tail bounds comes from the observation that every edge's final degree is correlated with the final degrees of its neighbors. This dependency not only means that Chernoff bounds do not apply, but techniques requiring low degree dependency graphs are not sufficient either. Cooper and Lu [17] use the Azuma-Hoeffding inequality [2, 21] to show for vertex removal that when $\sum_{v \in V} \deg(v)^2 = O(n^{2-\epsilon})$ for some positive ϵ, the probability that the actual number of nodes with degree k after edge removal deviates from $E[d_k]$ by more than a constant factor is something at most $O(n^{-2})$ (they give a more specific bound). We expand upon their technique, apply it to edge and not vertex removal, and get improved results in several ways. In addition to requiring a bound on the sum of the degrees squared, they reach their result by requiring the power law range to include all degrees (thus not allowing graphs with $d_{\min} \neq 0$ or $d_{\max} \neq \infty$) and the number of nodes of degree k to grow linearly with n. We remove these requirements and replace them with more generally applicable requirements. In this section we will prove the following large deviation theorem:

Theorem 5 *Suppose we are given an initial graph such that for some constants c_1, c_2, and with probability at least $1 - O(n^{-4})$ for all relevant k, the nodes with final degree k come from a set of size at most $E[d_k]c_1\sqrt{k \ln n}$ and with maximum degree $c_2 k(1 + \ln n/\sqrt{k})$. If each edge is removed independently with probability p, then for all such k simultaneously, the final number d_k of nodes of degree k will be within $[\frac{E[d_k]}{2}, \frac{3E[d_k]}{2}]$ with probability at least $1 - O(\frac{1}{n})$. For this theorem, we define all relevant k to mean those $k \in [d_{min}(1 - p), d_{\max}(1 - p)]$ such that $E[d_k] \geq 72c_1c_2 k \ln^2 n(\sqrt{k} + \ln n)$ or both $E[d_k] = \Omega(\ln n)$ and $k = \Omega(\ln^4 n)$.*

Remark. If we set $c_1 = 4/(1 - p)$ and $c_2 = 1/(1 - p)$, any power law graph satisfies Theorem 5's requirements. This includes power-law graphs that the Cooper-Lu method does not apply to because of the edge density. To demonstrate how our theorem applies to a concrete example, consider a power-law graph where nodes vary in degree from 1 to $n/10$ with $\gamma = 2 + \epsilon$. The expression $\sum_{i=1}^{n/10} \alpha i^{-2+\epsilon}$ for some α gives the number of nodes in the graph. Solving for α gives $\alpha \approx n/1.6$, and thus for k small enough, the

expected number of nodes with final degree k is $\Theta(n/k^{2+\epsilon})$. Applying Theorem 5 shows us that the actual number of nodes of degree k will be close to $E[d_k]$ roughly as long as either: (a) $E[d_k] \geq 72c_1c_2k\ln^2 n(\sqrt{k} + \ln n)$ which leads to $k = O(n^{1/3}/\ln n)$; or (b) $k = \Omega(\ln^4 n)$, which for reasonably large n is satisfied if $E[d_k] = \Omega(\ln n)$ and (a)'s conditions are not satisfied. ($\Omega(\ln n) = E[d_k]$ is equivalent to $\ln n = O(n/k^{2+\epsilon})$ and thus $k = O((n/\ln n)^{1/(2+\epsilon)})$.) These two cases combine to show that our theorem covers most of the range of degrees k likely to occur in the random graph. The only degrees likely to occur for which our tail bounds do not apply are those where $E[d_k] = o(\ln n)$. In these cases even if we could directly apply Chernoff bounds to the large deviation probabilities, there are simply not enough expected nodes of degree k to drive the probability of a bad event below $1/n$. Thus no scheme can be expected to improve significantly upon our result in this particular example.

For the first step in our proof we take any arbitrary ordering on the edges and consider the random process where the edges are exposed sequentially, each being removed with probability p. For any vertex v, we can use Chernoff bounds to show that at every step in the random process, if v has had x of its edges exposed, then less than $xp - 4\sqrt{x\ln n}$ or more than $xp + 4\sqrt{x\ln n}$ of those edges will have been removed with probability at most $2\exp(-xp(\frac{4\sqrt{x\ln n}}{xp})^2/3) = 2\exp(-(\frac{16\ln n}{3p})) = O(n^{-5})$. If we then take the union bound over all vertices and all time steps we have that with probability at least $1 - O(n^{-3})$ for every vertex v and number of exposed edges x, between $xp - 4\sqrt{x\ln n}$ and $xp + 4\sqrt{x\ln n}$ of those edges will have been removed. This means that for any k, with probability at least $1 - O(n^{-3})$ only those vertices for which the initial degree $\deg(v)$ satisfies

$$\deg(v)(1-p) - 4\sqrt{\deg(v)\ln n} \leq k \leq \deg(v)(1-p) + 4\sqrt{\deg(v)\ln n}$$

can have final degree k. We call this Event (A). Note that this result immediately implies one of the conditions for Theorem 5, namely that with high probability only nodes of initial degree at most $c_2k(1 + \ln n/\sqrt{k})$ can end up with final degree k.

We also rely upon the high probability event that for all relevant k, only nodes from the set S_k of size at most $E[d_k]c_1\sqrt{k}\ln n$ might end up with final degree k (this is a precondition of Theorem 5). By the union bound over all k individually, with probability at least $1 - O(n^{-3})$ this will happen for all k simultaneously. We call this Event (B).

For any relevant k, when $E[d_k] \geq 72c_1c_2k\ln^2 n(\sqrt{k} + \ln n)$, only $|S_k| \leq E[d_k]c_1\sqrt{k}\ln n$ nodes are conditioned to possibly have final degree k, and each of these has at most $c_2k(1+ \ln n/\sqrt{k})$ edges to other nodes in this set, for a total of at most $E[d_k]c_1c_2k^{1.5}\sqrt{\ln n}(1+ \ln n/\sqrt{k})/2$ edges internal to this set. Furthermore, each edge we expose in our random process can affect at most two of these nodes (its endpoints). Define the random variable X to be the expected number of nodes from S_k with final degree k after exposing these

internal edges. We can use Azuma's inequality which says

$$Pr[|X - E[d_k]| \geq E[d_k]/3] \leq 2\exp(-E[d_k]^2/(18 \sum_{\text{internal edges}} 2^2))$$

$$\leq 2\exp(-E[d_k]^2/(36E[d_k]c_1c_2k^{1.5}\sqrt{\ln n}(1 + \ln n/\sqrt{k})))$$

$$\leq 2\exp(-2\ln n) = 2n^{-2}.$$

Therefore, with high probability $X \in [\frac{2}{3} \cdot E[d_k], \frac{4}{3} \cdot E[d_k]]$. After exposing all of the internal edges, those nodes with external edges will not yet have their degrees fixed because their external edges have not been exposed. By definition, each external edge affects at most one relevant node. Therefore these remaining nodes end up with degree k independently from each other, and thus we can apply Chernoff bounds directly to exposing the external edges. Because $X \gg \ln n$, with high probability the final result will be within a small constant factor of X. Specifically, if the final number of nodes of degree k is Y:

$$Pr[|Y - E[d_k]| > E[d_k]/2] \leq 2n^{-2} + Pr[|Y - X| \geq E[d_k]/6|X > 2E[d_k]/3]$$

$$\leq 2n^{-2} + \exp(-2E[d_k]/3(1/4)^2/3)$$

$$\leq 2n^{-2} + \exp(-c_1c_2k^{1.5}\ln^2 n) = (2 + o(1))n^{-2}.$$

For the second case of Theorem 5, when $E[d_k] = \Omega(\ln n)$ and $k = \Omega(\ln^4 n)$, once again for any k we have $E[d_k]c_1\sqrt{k \ln n}$ nodes which might end up with degree k. Each of these nodes has at most a $\binom{k/(1-p)}{k}(1-p)^k p^{k/(1-p)-k} = O(1/\sqrt{k})$ probability of having final degree k. We divide the set S_k of nodes into groups of size at most $k/(2-2p)$ (and at most $k/(4-4p)$ unless there is only one group). We make this partition fairly uniformly, so that within each group G, the total expected number of nodes of degree k within the group, denoted $E[d_k^G]$ is within a constant factor, say 2, of every other group. The balls-and-bins model shows that even if we assign the nodes to groups randomly, this will be satisfied. Thus each group will have an $E[d_k^G]$ such that

$$E[d_k^G] \geq E[d_k] \cdot \frac{k/(4-4p)}{2E[d_k]c_1\sqrt{k \ln n}} \geq \frac{\sqrt{k}}{2c_1(4-4p)\ln n} \geq \Omega(\ln n).$$

Now consider the process where we arbitrarily order the internal edges, exposing them one at a time, and once finished we expose all of the external edges at once. How much can exposing one of the internal edges change $E[d_k^G]$? For any node v at any point in this process the internal edges already exposed are at most half of its total edges. Thus for each internal edge and each of its two endpoints, for some i: $\deg(v)/2 \leq i \leq \deg(v)$ and $i \leq j$, where i is the number of edges left unexposed and j the number left to keep, the effect on the expectation of exposing a single edge one way or the other per endpoint is

at most

$$\left| \binom{i-1}{j-1} p^{i-j}(1-p)^{j-1} - \binom{i-1}{j} p^{i-j-1}(1-p)^{j} \right|$$

$$= \left| \binom{i}{j} p^{i-j}(1-p)^{j} \left(\frac{j}{i(1-p)} - \frac{i-j}{ip} \right) \right| = O(\frac{1}{\sqrt{i}}) \left| \frac{j-i(1-p)}{i(1-p)p} \right|.$$

Because of the conditioning that at any point the number of edges removed r is between $xp - 4\sqrt{x \ln n}$ and $xp + 4\sqrt{x \ln n}$, $(j - i(1-p))$ is at most $8\sqrt{\deg(v) \ln n}$. This follows because $i = \deg(v) - x$ and $j = k - (x - r)$, and thus

$$|j - i(1-p)| = |k - x + r - (\deg(v) - x)(1-p)|$$

$$= |(k - \deg(v)(1-p))| + |(r - xp)|$$

$$\leq 4\sqrt{\deg(v) \ln n} + 4\sqrt{x \ln n} \leq 8\sqrt{\deg(v) \ln n}.$$

Therefore

$$\left| \binom{i-1}{j-1} p^{i-j}(1-p)^{j-1} - \binom{i-1}{j} p^{i-j-1}(1-p)^{j} \right| = O(\frac{\sqrt{\ln n}}{i}) = O(\frac{\sqrt{\ln n}}{k}). \qquad (17)$$

Define X^G and Y^G to be the expected number of nodes in G of degree k after exposing the internal and all edges respectively. If we look at any single such group G in isolation and using the bound from (17) in applying Azuma's inequality, we get

$$Pr[|X^G - E[d_k^G]| \geq E[d_k]/2 \leq 2\exp(-E[d_k^G]^2/(8 \sum_{\text{internal edges}} O(\sqrt{\ln n}/k)^2))$$

$$\leq 2\exp(-E[d_k^G]^2/O(\ln n)) = n^{-\Omega(1)}.$$

After exposing the internal edges, we can expose all of the external edges at random, and apply Chernoff bounds to show that the final Y^G values will be very close to X^G. The analysis details are the same as in the first case of the theorem. For n large enough, $n^{-\Omega(1)}$ plus the large deviation bound from the external edges is a small enough per group, that with probability at most n^{-2} *any* group will have large deviation. As long as no group's Y^G has a large deviation from $E[d_k^G]$, the sum Y cannot have large deviation from $E[d_k]$ either. Since for each k the large deviation probability of Y is at most $O(1/n^2)$. By taking a union bound over all such bad events for k from 0 to n and the event that at least one of Events (A,B) do not hold, we see that the probability that *any* such k has a large deviation is at most $O(1/n)$. This completes the proof of Theorem 5.

6 Edge Removal on Probabilistic Degree Distributions

For completeness we also mention the effect of random edge removal on stochastic graph models. Two such models are the Chung-Lu [12] and Kronecker [23] models. In the

Chung-Lu model, every node v in a bipartite graph is parametrized with an expected degree x_v. The model includes each edge u, v with probability proportional to $x_u \cdot x_v$. The Kronecker graph model starts with a small initiator adjacency matrix with entries between 0 and 1. A Kronecker product of two matrices A and B each of size n by n gives a new n^2 by n^2 matrix where entry i, j is $A_{(\lceil i/n \rceil, \lceil j/n \rceil)} \cdot B_{(i \mod n, j \mod n)}$. A Kronecker power of the matrix gives a probabilistic adjacency matrix from which graphs can be instantiated. In either case, each edge has a probability with which it is independently placed in an instance of the graph. We can sum the probabilities of all the edges incident on a node to get that node's expected degree.

In these, or any other, random graph models, we cannot used the definition of a degree distribution $d(j)$ for integers j giving the number of nodes of degree j. We modify the definition slightly, to say a stochastic graph specification has distribution $d(j)$ if for all j the number of nodes with expected degrees in the range $[j, j+1)$ is $d(j)$. Using this definition $E[d_k]$ is approximately $\sum_{j=\frac{k}{1-p}}^{\frac{k+1}{1-p}} d(j)$. For any $d(j)$ such that $\frac{d(j)}{d(j-1)} = 1 - o(1)$, this yields $E[d_k] \approx \frac{1}{1-p} \cdot d(\frac{j}{1-p})$. For any exponential distribution with $\frac{d(j)}{d(j-1)} = c$ for some constant c, $E[d_k] = \Theta(d(\frac{j}{1-p}))$. Meanwhile for an exponential distribution, we have $\sum_{j=\frac{k}{1-p}}^{\frac{k+1}{1-p}} d(j) \approx d(\frac{k}{1-p}) \sum_{i=0}^{1/(1-p)} \gamma^{-i} = d(k/(1-p)) \frac{1 - \gamma^{-1/(1-p)}}{1 - \gamma^{-1}} \approx d(k/(1-p))$.

In the case of an initial power-law distribution, the new distribution is also power-law, and has the same parameter γ. Interestingly, exponential graphs behave differently for stochastic and discrete models. In the discrete model, edge removal changes the exponential base, while the base remains unchanged for the stochastic models.

7   Conclusion

We examine the effect of random edge removal on a graph's degree sequence. We find that with power-law sequences, or any sequence whose rate of change is slow enough, the new degree sequence is a shifted version of the original, with the new sequence $d'(k) = O(d(\frac{k}{1-p}))$. This immediately implies that for power-law degree sequences, the new distribution remains power-law-like with the original exponent γ. Exponential degree sequences exhibit an interesting mathematical property where the $d(j)$ terms in the summation for $E[d_k]$ combine with the probability terms to create a new exponential with exponent $\gamma \cdot \frac{1-p}{1-p\gamma}$. We then conclude by showing that, for many types of graphs, with high probability the actual degree sequences will be within a small constant factor of the expected sequence. These results provide a useful bridge in settings where a graph property that is tied to the degree sequence may be extended with high probability to an edge-percolated subgraph because of the similarity in degree sequences.

Acknowledgement

References

[1] N. Alon, I. Benjamini, and A. Stacey. Percolation on finite graphs and isoperimetric inequalities. In *Annals of Probability*, volume 32, pages 1727–1745, July 2004.

[2] Kazuoki Azuma. Weighted sums of certain dependent random variables. *Tohoku Mathematical Journal*, 19(3):357–367, 1967.

[3] S. Banerjee, S. Lee, B. Bhattacharjee, and A. Srinivasan. Resilient multicast using overlays. In *IEEE/ACM Transactions on Networking*, pages 237–248, 2006.

[4] A. L. Barabasi, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A*, 311:3, 2002.

[5] Albert-László Barabási and Réka Albert. Emergence of Scaling in Random Networks. *Science*, 286(5439):509–512, 1999.

[6] Albert-László Barabási, Réka Albert, and H. Jeong. Scale-free characteristics of random networks: the topology of the World-Wide Web. *Physica A*, 281:69–77(9), 2000.

[7] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, 85:5468, 2000.

[8] F. Chung. *Spectral Graph Theory*. AMS Publications, 1997.

[9] F. Chung and L. Lu. *Complex Graphs and Networks*. AMS Publications, 2006.

[10] Fan Chung, Paul Horn, and Linyuan Lu. The giant component in a random subgraph of a given graph. In *WAW '09: Proceedings of the 6th International Workshop on Algorithms and Models for the Web-Graph*, pages 38–49, Berlin, Heidelberg, 2009. Springer-Verlag.

[11] Fan Chung and Linyuan Lu. The average distance in random graphs with given expected degrees. In *Proceedings of National Academy of Science*, volume 99, pages 15879–15882, 2002.

[12] Fan Chung and Linyuan Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6:125–145, November 2002.

[13] Fan Chung and Linyuan Lu. The volume of the giant component for a random graph with given expected degrees. *SIAM J. Discrete Math.*, 20(2):395–411, 2006.

[14] Fan Chung, Linyuan Lu, and Van Vu. The spectra of random graphs with given expected degrees. In *Proceedings of National Academy of Science*, volume 100, pages 6313–6318, 2003.

[15] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Rev.*, 51(4):661–703, 2009.

[16] Reuven Cohen, Keren Erez, Daniel Ben-Avraham, and Shlomo Havlin. Resilience of the Internet to Random Breakdowns. *Physical Review Letters*, 85(21):4626–4628, November 2000.

[17] Joshua N. Cooper and Lincoln Lu. Graphs with asymptotically invariant degree sequence under restriction. *Internet Mathematics*, 7(1):67–80, 2011.

[18] Erdős and Rényi. On random graphs. *Publ. Math. Debrecen*, pages 290–297, 1959.

[19] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos. On power-law relationships of the internet topology. In *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*, SIGCOMM '99, pages 251–262, New York, NY, USA, 1999. ACM.

[20] J. Goldenberg, B. Libai, and Muller. Using complex systems analysis to advance marketing theory development. *Academy of Marketing Science Review*, 2001.

[21] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *American Statistical Association*, 58(301):13–30, Mar 1963.

[22] Jon M. Kleinberg, Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew S. Tomkins. The web as a graph: Measurements, models, and methods. In T. Asano, editor, *International Conference on Combinatorics and Computing*, Lecture Notes in Computer Science, pages 1–17. Springer Berlin / Heidelberg, 1999.

[23] Jure Leskovec, Deepayan Chakrabarti, Jon Kleinberg, Christos Faloutsos, and Zoubin Ghahramani. Kronecker graphs: An approach to modeling networks. *J. Mach. Learn. Res.*, 11:985–1042, March 2010.

[24] Y. Malevergne, V. Pisarenko, and D. Sornette. Empirical distributions of stock returns: between the stretched exponential and the power law? *Quantitative Finance*, 5(4):379–401, 2005.

[25] S. Martin, R. D. Carr, and J.-L. Faulon. Random removal of edges from scale free graphs. *Physica A Statistical Mechanics and its Applications*, 371:870–876, November 2006.

[26] M. Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, 1(2):226–251, 2004.

[27] Herbert Robbins. A remark on Stirling's formula. *Amer. Math. Monthly*, 62:26–29, 1955.

[28] Andreas Wagner. How the global structure of protein interaction networks evolves. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1514):457–466, 2003.