Discrete Mathematics and Theoretical Computer Science **AC**, 2003, 95–104

Rigorous Result for the CHKNS Random Graph Model

Rick Durrett

Cornell University, Department of Mathematics. Mallott Hall, Ithaca, NY 14853, U.S.A.
rtd1@cornell.edu

We study the phase transition in a random graph in which vertices and edges are added at constant rates. Two recent papers in Physical Review E by Callaway, Hopcroft, Kleinberg, Newman, and Strogatz, and Dorogovstev, Mendes, and Samukhin have computed the critical value of this model, shown that the fraction of vertices in finite clusters is infinitely differentiable at the critical value, and that in the subcritical phase the cluster size distribution has a polynomial decay rate with a continuously varying power. Here we sketch rigorous proofs for the first and third results and a new estimates about connectivity probabilities at the critical value.

Keywords: random graph, clusterization, Brownian motion, singularity analysis

1 Introduction

In the last few years, physicists, mathematicians, and computer scientists, motivated by the world wide web (Albert, Jeong, and Barbási 1999, Huberman and Adamic 1999), metabolic networks (Jeong et al. 2000), and other complex structures (for a survey see Strogatz 2001), have begun to investigate the difference between static random graphs and networks in which the node set grows and connections are added over time. Barbási and Albert (1999) considered a model in which new vertices are attached preferentially to already well connected sites and found a power-law distribution for vertex degrees. Callaway, Hopcroft, Kleinberg, Newman, and Strogatz (2001) studied the following model without preferential attachment. At each time a vertex is added to the graph. Number the vertices $1,2,\ldots,n$ in the order they were added. For $k \geq 2$ after the kth vertex is added we add a number of edges with mean δ. The edges are drawn with replacement from the $\binom{k}{2}$ possible edges.

In the original CHKNS model the number of edges was 1 with probability δ, and 0 otherwise. Here, we will primarily study the situation in which a Poisson mean δ number of vertices are added at each step. We prefer this version since in the Poisson case if we let $A_{i,j,k}$ be the event no (i,j) edge is added at time k then $P(A_{i,j,k}) = \exp\left(-\delta/\binom{k}{2}\right)$ for and $i < j \leq k$ and these events are independent.

$$P(\cap_{k=j}^{n} A_{i,j,k}) = \prod_{k=j}^{n} \exp\left(-\frac{2\delta}{k(k-1)}\right) = \exp\left(-2\delta\left(\frac{1}{j-1} - \frac{1}{n}\right)\right) \qquad \#1.$$

The last formula is somewhat ugly, so we will also consider two approximations

$$\approx 1 - 2\delta\left(\frac{1}{j} - \frac{1}{n}\right) \qquad \#2$$

$$\approx 1 - \frac{2\delta}{j} \qquad \#3$$

We will refer to these three models by their numbers, and the original CHKNS model as #0. The second approximation is not as innocent as it looks. If we let $\mathcal{E}_n$ be the number of edges then using the definition of the model and $0 \le e^{-x} - (1-x) \le x^2$ for $0 < x < 1$ we see that

$$E\mathcal{E}_n \approx \begin{cases} \delta n & \#1 \\ \delta(n + O(\log n)) & \#2 \\ 2\delta n & \#3. \end{cases}$$

It turns out however that despite having twice as many edges the connectivity properties of model #3 is almost the same as that of models #1 and #2. See Theorem 3 below. To prepare for an intuitive explanation we will give later, recall that the random graph model of Erdős and Rényi (see Bollobás 1985 for a comprehensive survey) has an edge from i to j with probability $p_{i,j} = \lambda/n$, and note that model #3 corresponds roughly to model #2 plus an independent copy of an Erdős–Rényi random graph with $\lambda = 2\delta$.

The CHKNS analysis of their model begins by examining $N_k(t) =$ the expected number of components of size k at time t. Ignoring terms of $O(1/t^2)$, which correspond to differences between t and $t+1$ in the denominator or picking the same cluster twice:

$$N_1(t+1) = N_1(t) + 1 - 2\delta\frac{N_1(t)}{t}$$

$$N_k(t+1) = N_k(t) - 2\delta\frac{kN_k(t)}{t} + \delta\sum_{j=1}^{k-1}\frac{jN_j(t)}{t}\cdot\frac{(k-j)N_{k-j}(t)}{t}$$

To explain the first equation, note that at each discrete time t one new vertex is added, and a given isolated vertex becomes the endpoint of an added edge with probability $\approx 2\delta/t$. For the second equation, note that the probability an edge connects to a given cluster of size is $\approx 2\delta k/t$, while the second term corresponds to mergers of clusters of size j and $k-j$. There is no factor of 2 in the last term since we sum from 1 to $k-1$.

CHKNS stated the following result without proof. However, due to the triangular nature of the coupled differential equations, it is not difficult to use a little undergraduate analysis and induction to show:

Theorem 1. *For model #0 or #1, as $t \to \infty$, $N_k(t)/t \to a_k$ where $a_1 = 1/(1+2\delta)$ and*

$$a_k = \frac{\delta}{1+2\delta k}\sum_{j=1}^{k-1} ja_j \cdot (k-j)a_{k-j}.$$

To solve for the a_k, which gives the limiting number of clusters of size k per site, CHKNS used generating functions. Let $h(x) = \sum_{k=1}^{\infty} x^k a_k$ and $g(x) = \sum_{k=1}^{\infty} x^k k a_k$. Multiplying the equations in Theorem 1 by $(1+2\delta k)x^k$ and summing gives

$$h(x) + 2\delta g(x) = x + \delta g^2(x).$$

Since $h'(x) = g(x)/x$ differentiation gives $g(x)/x + 2\delta g'(x) = 1 + 2\delta g(x) g'(x)$. Rearranging we have

$$(\star). \qquad g'(x) = \frac{1}{2\delta x} \cdot \frac{x - g(x)}{1 - g(x)}$$

Let $b_k = ka_k$ be the fraction of vertices that belong to clusters of size k. $g(1) = \sum_{k=1}^{\infty} b_k$ gives the fraction of vertices that belong to finite components. $1 - g(1)$ gives the fraction of sites that belong to clusters whose size grows in time. Even though it is not known that the missing mass in the limit belongs to a single cluster, it is common to call $1 - g(1)$ the fraction of sites that belong to the giant component. The next result gives the mean size of finite components.

Lemma 1. *(i) If* $g(1) < 1$ *then* $\sum_{k=1}^{\infty} kb_k = g'(1) = 1/2\delta$.
(ii) If $g(1) = 1$ *then* $g'(1) = (1 - \sqrt{1 - 8\delta})/4\delta$.

Proof. The first conclusion is immediate from $(\star)$. If $g(1) = 1$, L'Hôpital's rule implies

$$2\delta g'(1) = \lim_{x \to 1} \frac{x - g(x)}{1 - g(x)} = \lim_{x \to 1} \frac{1 - g'(x)}{-g'(x)}$$

which gives $2\delta (g'(1))^2 - g'(1) + 1 = 0$. This solution of this quadratic equation indicated in (ii) is the one that tends to 1 as $\delta \to 0$. □

Theorem 2 (Theoem CHKNS). *The critical value* $\delta_c = \sup\{\delta : g(1) = 1\} = 1/8$ *and hence*

$$\sum_k kb_k = \begin{cases} (1 - \sqrt{1 - 8\delta})/4\delta & \delta \le 1/8 \\ 1/2\delta & \delta > 1/8 \end{cases}$$

Remarks. (a) Note that this implies that the mean cluster size $g'(1)$ is always finite but is discontinuous at $\delta = 1/8$ since the value there is 2 but the limit for $\delta \downarrow 1/8$ is 4.
(b) We left a few letters out of the word Theorem to indicate that a few things were left out of the proof. Borrowing a phrase my graduate school roommate Danny Solow invented to vent his frustrations at steps in the proofs in his analysis class appearing as if by magic, we call the nonrigorous proof a

Poof. The formula for the derivative of the real valued function g becomes complex for $\delta > 1/8$ so we must have $\delta_c \le 1$. To argue the other direction, CHKNS note that mean cluster size $g'(1)$ is in general non-analytic only at the critical value and $(1 - \sqrt{1 - 8\delta})/4\delta$ is analytic for $\delta < 1/8$. If you are curious about their exact words, see the paragraph above (17) in their paper. □

2 Rigorous derivation of the critical value

While the reasoning in the poof is not rigorous, our next result shows that the conclusion is correct. In contrast to the situation with ordinary percolation on the square lattice where Kesten (1980) proved the answer was correct nearly twenty year after physicists had guessed it, this time the rigorous answer predates the question by more than 10 years.

Theorem 3. *In models #1, #2, or #3, the critical value* $\delta_c = 1/8$.

To give the promised intuitive explanation of the equality of critical values, recall our remark that model #3 is roughly model #2 plus an independent copy of an Erdős–Rényi random graph with $\lambda = 2\delta$ but for the ER model, cluster sizes have exponentially decaying tails for $\lambda < 1$.

We begin by describing earlier work on the random graph model on $\{1,2,3,\ldots\}$ with $p_{i,j} = \lambda/(i \vee j)$. Kalikow and Weiss (1988) showed that the probability G is connected (ALL vertices in ONE component) is either 0 or 1, and that $1/4 \le \lambda_c \le 1$. They conjectured $\lambda_c = 1$ but Shepp (1989) proved $\lambda_c = 1/4$. To connect with the answer in Theorem 3, note that $\lambda = 2\delta$. Durrett and Kesten (1990) proved a result for a general class of $p_{i,j} = h(i,j)$ that are homogeneous of degree -1, i.e., $h(ci,cj) = c^{-1}h(i,j)$. It is their methods that we will use to prove the result.

Proof of $\delta_c \ge 1/8$. We prove the upper bound for the largest model, #3. An easy comparison shows that the mean size of the cluster containing a given point i is bounded above by the expected value of the total progeny of a discrete time multi-type branching process in which a particle of type j gives birth to one offspring of type k with probability $p_{j,k}$ (with $p_{j,j} = 0$) and the different types of births are independent.

To explain why we expect this comparison to be accurate, we note that in the Erdős–Rényi random graph with $p_{j,k} = \lambda/n$, the upper bound is an ordinary branching process with a Poisson mean λ offspring distribution so we get the correct lower bound $\lambda_c \ge 1$. When $p_{j,k} = 2\delta/(j \vee k)$, the mean of the total progeny starting from one of type i is $\sum_{m=0}^{\infty}\sum_j p^m_{i,j}$, which will be finite if and only if the spectral radius $\rho(p_{i,j}) < 1$. By the Perron–Frobenius theory of positive matrices, ρ is an eigenvalue with positive eigenvector.

Following Shepp (1989) we now make a good guess at this eigenvector.

$$\begin{aligned}
\sum_{j=1}^{n} \frac{1}{i \vee j} \cdot \frac{1}{j^{1/2}} &= \frac{1}{i}\sum_{j=1}^{i-1} \frac{1}{j^{1/2}} + \sum_{j=i+1}^{n} \frac{1}{j^{3/2}} \le \frac{1}{i}\left(1 + \int_1^i \frac{1}{x^{1/2}}\,dx\right) + \int_i^n \frac{1}{x^{3/2}}\,dx \\
&= \frac{1}{i}(1 + 2i^{1/2} - 2) + 2(i^{-1/2} - n^{-1/2}) \le \frac{4}{i^{1/2}}
\end{aligned}$$

This implies $\sum_j i^{1/2} p_{i,j} j^{-1/2} \le 8\delta$ so if we let $b_{n,k}$ be the expected fraction of vertices in clusters of size k in the model on n vertices, and $|\mathcal{C}_i|$ be the size of the cluster $\mathcal{C}_i$ that contains i,

$$\begin{aligned}
\sum_k k b_{n,k} &= \frac{1}{n}\sum_{i=1}^{n} E|\mathcal{C}_i| \le \frac{1}{n}\sum_m \sum_{i,j} p^m_{i,j} \\
&\le \frac{2}{n}\sum_{m=0}^{\infty}\sum_{i \ge j} i^{1/2} p^m_{i,j} j^{-1/2} \le 2\sum_{m=0}^{\infty}(8\delta)^m \le \frac{2}{1-8\delta}
\end{aligned}$$

which completes the proof of the lower bound. □

Proof of $\delta_c \le 1/8$. In this case we need to consider the smallest model, so we set:

$$Q(i,j) = \frac{1}{i \vee j} - \frac{1}{n} \qquad \text{when } K \le i, j \le n.$$

For those who might expect to see some -1's in the denominator, we observe that they can be eliminated by shifting our index set. By the variational characterization of the largest eigenvalue

$$\rho(Q) \ge \left(\sum_{i=1}^{n} v_j^2 \right)^{-1} v^T Q v .$$

Again we take $v_j = 1/\sqrt{j}$.

$$v^T Q v = 2 \sum_{i=K}^{n} \sum_{j=i+1}^{n} \frac{1}{i^{1/2}} \frac{1}{j^{3/2}} - \frac{1}{n} \left(\sum_{j=K}^{n} \frac{1}{j^{1/2}} \right)^2 .$$

The second term is ≥ -4. Bounding sums below by integrals the first is

$$\ge 2 \sum_{i=K}^{n} 2(i+1)^{-1} - 2i^{-1/2}(n+1)^{-1/2} \ge 4 \sum_{i=K}^{n} (i+1)^{-1} - 4 .$$

This implies $\rho(Q) \ge \left(4 \sum_{i=K}^{n} (i+1)^{-1} - 8\right) / \sum_{i=K}^{n} i^{-1}$.

Letting $q(i,j) = 2\delta \left(\frac{1}{i \vee j} - \frac{1}{n} \right)$ for $K \le i, j < KN \le n$ we have

$$\rho(q) \ge 8\delta \frac{\log N - 3}{\log N} .$$

If $8\delta = 1 + 4\varepsilon > 1$ and $N = e^{12+(3/\varepsilon)}$ we have $\rho(q) \ge 1 + 3\varepsilon$ for all $K \ge 1$ and the desired result follows from

(2.16) in Durrett and Kesten (1990). *Consider the q random graph in $[K, NK)$. There are positive constants γ and β so that if $K \ge K_0$ then with probability at least β, K belongs to a component with at least γNK vertices.*

The proof of (2.16) has two steps:

(i) Let $M = 1 + (1/\varepsilon)$, $L = K/M$, subdivide $[K, KN)$ into intervals $[K + (m-1)L, K + mL)$ for $1 \le m \le MN$ to define a multitype branching process with MN types with spectral radius of the mean matrix $\ge 1 + 2\varepsilon$ if K is large.

(ii) Argue that until some interval has more than a fraction ε of its sites occupied the percolation process dominates a branching process with spectral radius $1 + \varepsilon$, so the percolation process will be terminated by this condition with probability $\ge \beta > 0$. For further details see Durrett and Kesten (1990). □

The proof of (2.16) gives a very tiny bound on the fraction of vertices in the large component

$$\gamma = \frac{\varepsilon}{MN} = C\varepsilon^2 e^{-3/\varepsilon} .$$

However it turns out that this estimate is not too bad. By numerically solving ($\star$), CHKNS showed $1 - g(1) \approx \exp(\alpha(\delta - \delta_c)^{-1/2})$. Inspired by their conjecture Dorogovstev, Mendes, and Samukhin (2001) showed:

Theorem 4 (Theoem DMS). *As* $\delta \downarrow 1/8$,

$$S \equiv 1 - g(1) \approx c \exp(-\pi/\sqrt{8\delta - 1}).$$

Note that this implies that the percolation probability S is infinitely differentiable at the critical value, in contrast to the situation for the Erdős–Rényi model and for percolation on $\mathbb{Z}^d$, in which $S \sim (\delta - \delta_c)^\beta$ as $\delta \downarrow \delta_c$ with $\beta \le 1$. See Bollobás (1985) or J.T. Chayes and L. Chayes (1986).

Poof. To derive this result DMS change variables $u(\xi) = 1 - g(1-\xi)$ in $(\star)$ to get

$$u'(\xi) = \frac{1}{2\delta(1-\xi)} \cdot \frac{u(\xi) - \xi}{u(\xi)}.$$

They discard the $1 - \xi$ in the denominator (without any justification or apparent guilt at doing so) and note that the solution to the differential equation is the solution of the following transcendental equation

$$-\frac{1}{\sqrt{8\delta - 1}} \arctan\left(\frac{4\delta[u(\xi)/\xi] - 1}{\sqrt{8\delta - 1}}\right) - \ln\sqrt{\xi^2 - u(\xi)\xi + 2\delta u^2(\xi)}$$
$$= -\frac{\pi/2}{\sqrt{8\delta - 1}} - \ln\sqrt{2\delta} - \ln S$$

This formula is not easy (for me at least) to guess but with patience is not hard to verify. Once this is done, the remainder of the proof is fairly routine asymptotic analysis. □

3 Results at the critical value

It would indeed be a thankless job to fill in details in steps in the last proof that DMS didn't feel the need to justify, so we turn now to an analysis of our model(s) at the critical value. Yu Zhang (1991) studied the percolation process with $p_{i,j} = (1/4)/(i \vee j)$ on $\{1, 2, \ldots\}$ in his Ph.D. thesis at Cornell written under the direction of Harry Kesten. This is a rigorous result, so we modify our naming convention accordingly.

Theorem 5. *If* $i < j$ *and* $i \ge \log^{6+\delta} j$ *then*

$$\frac{c_1 \log(i+1)}{\sqrt{ij}} \le P(i \to j) \le \frac{c_2 \log(i+1)}{\sqrt{ij}}.$$

By adapting Zhang's method we can prove a similar result for model #3:

Theorem 6. *If* $i < j$ *then* $P(i \to j) \le (3/8)\Gamma^n_{i,j}$.
If (i) $(2/\varepsilon)^3 \log^6(4/\varepsilon) \le i < j \le n^{1-\varepsilon}$, $(\varepsilon \le \varepsilon_0)$ *or* (ii) $(\log n)^3 \le i < j \le n$, $(n \ge n_0)$, *then*

$$P(i \to j) \ge c\Gamma^n_{i,j} \quad \textit{where} \quad \Gamma^n_{i,j} = \frac{(\log i + 2)(\log n - \log j + 2)}{(\log n + 4)}.$$

Remarks. (a) The proof of Theorem 6 is a refinement of the proof of Theorem 5. In particular our version of Lemma 2 below allows the condition $i \geq \log^{6+\delta} j$ to be removed from Zhang's result. Unfortunately, the fact that $\Gamma^n_{i,j}$ vanishes at $j = n$ forces us to keep one of the points away from the boundary. (b) From the upper bound in Theorem 6 and some routine summation it follows that

$$\frac{1}{n}\sum_{i=1}^{n} E|C_i| \leq 2\sum_{i<j} P(i \to j) \leq 6\,.$$

This shows that the expected cluster size is finite at the critical value. This upper bound is only 3 times the exact value of 2 given in Theoem 2.

Proof. The expected number of self-avoiding paths from i to j is

$$EV_{i,j} = \sum_{m=0}^{\infty}\sum_{*} h(i,z_1)h(z_1,z_2)\cdots h(z_m,j)$$

where $h(x,y) = (1/4)/(x \vee y)$ and the starred sum is over all self-avoiding paths. The sum restricted to paths with all $z_i \geq 2$ has

$$\Sigma^1_{i,j} \leq \sum_{m=0}^{\infty}\int_1^n dx_1 \cdots \int_1^n dx_m h(i,x_1)h(x_1,x_2)\cdots h(x_m,j)\,.$$

Introducing

$$\pi(u,v) = e^{u/2}h(u,v)e^{v/2} = \begin{cases} (1/4)e^{(u-v)/2} & u \leq v \\ (1/4)e^{(v-u)/2} & u \geq v\,. \end{cases}$$

and setting $\log x_i = y_i$, $dx_i = e^{y_i}\,dy_i$ we have

$$\Sigma^1_{i,j} \leq \frac{1}{\sqrt{ij}} G_{0,\log n}(\log i, \log j)$$

where G is the Green's function for the bilateral exponential random walk killed when it exits $[0, \log n]$.

Suppose the jump distribution is $(\lambda/2)e^{-\lambda|z|}$. Since boundary overshoots are exponential, a standard martingale calculation applied at the exit time from (u,v) shows

$$P_x(T_{(-\infty,u]} < T_{[v,\infty)}) = \frac{(v+1/\lambda)-x}{(v+1/\lambda)-(u-1/\lambda)}$$

the exit probability for Brownian motion from the interval $(u-1/\lambda, v+1\lambda)$. Using this formula and standard reasoning about hitting times, one can show that for the case $\lambda = 1/2$.

$$G_{K,L}(x,z) = \begin{cases} \frac{1}{4}\cdot\frac{(L-x+2)(z-K+2)}{L-K+4} & z \leq x\,, \\ \frac{1}{4}\cdot\frac{(L-z+2)(x-K+2)}{L-K+4} & z \geq x\,. \end{cases}$$

If we discard the $+2$'s and $+4$'s this is exactly the formula for the Green's function of $\sqrt{8}B_t$. Taking $x = \log i$, $z = \log j$ and bounding the paths that visit 1 by $\Sigma^1_{i,1}\cdot\Sigma^1_{1,j}$, the upper bound follows.

To get a lower bound we have to remove the terms from the sum that visit a site more than once. A somewhat lengthy calculation gives:

Lemma 2. *If* $\log(\kappa-1)\geq 6$ *then for* $\kappa^2\leq i<j\leq n$ *we have*

$$EV_{i,j}\geq \frac{1}{8\sqrt{ij}}\left[\frac{(\log i+2)(\log n-\log j+2)}{\log n+4}-\frac{(\log\kappa)^3}{\kappa-1}\right].$$

If conditions (i) or (ii) in Theorem 6 hold then the second term is at most half the first one so $EV_{i,j}\geq (1/16)\Gamma^n_{i,j}$.

By using Zhang's (1991) argument one can show

Lemma 3. $EV^2_{i,j}\leq C\Gamma^n_{i,j}$

Combining the last two lemmas we have $EV^2_{i,j}\leq CEV_{i,j}$. The Cauchy–Schwarz inequality implies

$$EV_{i,j}=E(V_{i,j}1_{\{V_{i,j}>0\}})\leq\sqrt{E(V^2_{i,j})P(V_{i,j}>0)}\,.$$

Rearranging gives $P(i\to j)=P(V_{i,j}>0)\geq EV_{i,j}/C$ which gives the lower bound. □

4 The subcritical case

It is straightforward to generalize the proof of the upper bound in Theorem 7 to show that when $\delta<1/8$,

$$P(i\to j)\leq\frac{1}{\sqrt{ij}}G^{8\delta}(i,j)$$

where $G^{8\delta}$ is the Green's function for the bilateral exponent on **R** killed on each step with probability $1-8\delta$. (One can get lower bounds but they are even worse than our results for the critical case.) Using Fourier transforms one can easily compute $G^{8\delta}(x,y)=\dfrac{2\delta}{\sqrt{1-8\delta}}e^{-r|x-y|}$ where $r=\sqrt{1-8\delta}/2$ which gives for $i<j$: $P(i\to j)\leq\dfrac{c}{i^{1/2-r}j^{1/2+r}}$. Setting $i=1$ and summing over $1<j\leq n$ and doing the same thing to the upper bound in Theorem 7 gives

$$E|\mathcal{C}_1|\leq\begin{cases} cn^{(1-\sqrt{1-8\delta})/2} & 0<\delta<1/8\,,\\ cn^{1/2}/\log n & \delta=1/8\,.\end{cases}$$

I have tried a number of techniques to bound the variance of $|\mathcal{C}_1|$. Thus I leave to the reader to consider:

Problem. *Is* $|\mathcal{C}_1|=O(E|\mathcal{C}_1|)$*?*

Of course making this statement precise is part of the problem. Note that $p_{1,j}=2\delta/j$ for $1\leq j\leq n$ so with high probability the number of edges incident to 1 will be $2\delta\log n+O(\sqrt{\log n})$. Given this it seems likely that the component containing 1 will with high probability be the largest component, but this also needs to be proved.

Dorogovstev, Mendes, and Samukhin (2001) studied the preferential attachment model in which one new vertex and an average of δ edges were added at each time and the probability of an edge from i to j is proportional to $(d_i+\alpha)(d_j+\alpha)$ where d_k is the degree of k. The CHKNS model arises as the limit $\alpha\to\infty$. Taking this limit of the DMS results suggests that the probability a randomly chosen vertex belongs to a cluster of size k has

$$b_k\sim\frac{2}{k^2\ln k}\quad\text{if }\delta=1/8.$$

In the subcritical regime one has (see their (B16) and (B17) and not (21) which is wrong)

$$b_k \sim C_\delta k^{-2/(1-\sqrt{1-8\delta})} \quad \text{if } \delta < 1/8\,.$$

As the next result shows, once again the physicists are right.

Theorem 7. *The formulas for b_k hold for model #0 and #1.*

Proof. As in the first steps of the poof of Theoem 4, we let $u(y) = 1 - g(1-y)$ and $u(y) = y(u'(0) - v(y))$

$$u'(y) = \frac{1}{2\delta(1-y)} \cdot \frac{u(y)-y}{u(y)}\,.$$

Unfortunately, we cannot give ourselves the luxury of discarding the $1-y$ (we could for $\delta > 1/9$ but not for later stages of the argument), so when we plug in the formulas above to get

$$\begin{aligned} v'(y) &= \frac{(1-4\delta u'(0))v(y) + 2\delta v(y)^2}{2\delta y(1-y)(u'(0)-v(y))} + \frac{1}{1-y}(u'(0)-v(y)) \\ &\approx \frac{av(y)}{y} + \frac{v(y)^2}{u'(0)y} + u'(0) \end{aligned}$$

where the constant

$$a = \frac{1-4\delta u'(0)}{2\delta u'(0)} = \begin{cases} 0 & \text{if} \quad \delta = 1/8\,, \\ <1 & \text{if} \quad \delta > 1/9\,. \end{cases}$$

Asymptotic analysis of the differential equation implies

$$v(y) \sim \begin{cases} u'(0)/(-\log y) & \delta = 1/8\,, \\ y^a & \delta \in (1/9, 1/8)\,, \\ u'(0)y(-\log y) & \delta = 1/9\,, \\ cy & \delta < 1/9\,. \end{cases}$$

where $c = 2\delta u'(0)^2/(1-6\delta u'(0))$. From this it follows that

$$\sum_k kb_k(1-(1-y)^{k-1}) = g'(1) - g'(1-y) = yv'(y) + v(y) \sim \begin{cases} 2/\log(1/y) & \delta = 1/8 \\ (1+a)y^a & 1/9 < \delta < 1/8 \end{cases}$$

To check the guesses we note $\displaystyle\sum_{k>1/y} \frac{1}{k(\log k)^2} \approx \frac{1}{\log(1/y)}$ and $\displaystyle\sum_{k>1/y} k^{-\rho+1} \approx (1/y)^{2-\rho}$ so $\rho = a+2 = 1/2\delta u'(0) = 2/(1-\sqrt{1-8\delta})$.

When $\delta < 1/9$, u has two continuous derivatives, so we have taken away more smooth terms to find the singular part. In general if $k < \frac{1}{2\delta u'(0)} - 1 \le k+1$ we can write (recall $u(y) = 1 - g(1-y)$)

$$u(y) = -\sum_{i=1}^{k} c_i(-y)^i + (-y)^k v(y)$$

and analyze $v(y)$ as before. Results of Flajolet and Odlyzko (1990) then allow us to get the desired asymptotics. □

References

[1] Albert, R., Jeong, H., and Barbási A.L. (1999) Diameter of the world-wide web. *Nature*, **401**, 130–131

[2] Barbási A.L., and Albert, R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512

[3] Bollobás, B. (1985) *Random Graphs.* Academic Press, New York

[4] Callaway, D.S., Hopcroft, J.H., Kleinberg, J.M., Newman, M.E.J., and Strogatz, S.H. (2001) Are randomly grown graphs really random? *Physical Review E*, **64**, Paper 041902

[5] Chayes, J.T., and Chayes, L. (1986) inequality for the infinite-cluster density in Bernoulli percolation. *Phy. Rev. Letters.* **56**, 1619–1622

[6] Dorogovstev, S.N., Mends, J.F.F., and Samukhin, A.N. (2001) Anomalous percolation of growing networks. *Physical Review E*, **64**, Paper 066110

[7] Durrett, R., and Kesten, H. (1990) The critical parameter for connectedness of some random graphs. Pages 161–176 in *A Tribute to P. Erdős.* Edited by A. Baker, B. Bollobás, and A. Hajnal. Cambridge U. Press

[8] Flajolet, P. and Odlyzko, A. (1990) Singularity analysis of generating functions. *SIAM J. Discrete Math.* **3**, 216–240

[9] Huberman, B.A., and Adamic, L.A. (1999) Growth dynamics of the world-wide web. *Nature*, **401**, 131

[10] Jeong, H., Tombor, B., Albert, R., Oltval, Z.N., and Barbási, A.L. (2000) The large-scale organization of metabolic networks. *Nature*, **407**, 651–654

[11] Kalikow, S., and Weiss, B. (1988) When are random graphs connected? *Israel J. Math.* **62**, 257–268

[12] Kesten, H. (1980) The critical value for bond percolation on the square lattice equals 1/2. *Commun. Math. Physics.* **74**, 41–59

[13] Shepp, L.A. (1989) Connectedness of certain random graphs. *Israel J. Math.* **67**, 23–33

[14] Strogatz, S. (2001) Exploring complex networks. *Nature*, **410**, 268–276

[15] Zhang, Y. (1991) A power law for connectedness of some random graphs at the critical point. *Random Structures and Algorithms*, **1**, 101–119