Combinatorics, Probability and Computing (2005) 14, 897–938. © 2005 Cambridge University Press
doi:10.1017/S096354830500708X Printed in the United Kingdom

The Small Giant Component in Scale-Free Random Graphs

OLIVER RIORDAN†

Department of Pure Mathematics and Mathematical Statistics,
University of Cambridge, Cambridge CB3 0WB, UK
and
Trinity College, Cambridge CB2 1TQ, UK
(e-mail: o.m.riordan@dpmms.cam.ac.uk)

Received 12 March 2004; revised 29 October 2004

For Béla Bollobás on his 60th birthday

Building on the methods developed in joint work with Béla Bollobás and Svante Janson, we study the phase transition in four 'scale-free' random graph models, obtaining upper and lower bounds on the size of the giant component when there is one. In particular, we determine the extremely slow rate of growth of the giant component just above the phase transition. We greatly reduce the significant gaps between the existing upper and lower bounds, giving bounds that match to within a factor $1 + o(1)$ in the exponent.

In all cases the method used is to couple the neighbourhood expansion process in the graph on n vertices with a continuous-type branching process that is independent of n. It can be shown (requiring some separate argument for each case) that with probability tending to 1 as $n \to \infty$ the size of the giant component divided by n is within $o(1)$ of the survival probability σ of the branching process. This survival probability is given in terms of the maximal solution ϕ to certain non-linear integral equations, which can be written in the form $\phi = \mathbf{F}(\phi)$ for a certain operator $\mathbf{F}$. Upper and lower bounds are found by constructing trial functions ϕ_0, ϕ_1 with $\mathbf{F}(\phi_0) \leqslant \phi_0$ and $\mathbf{F}(\phi_1) \geqslant \phi_1$ holding pointwise; basic properties of branching processes then imply that $\phi_1 \leqslant \phi \leqslant \phi_0$, giving upper and lower bounds on σ.

1. Introduction

In the last few years there has been much interest in new random graph models motivated by the empirical and heuristic study of large-scale real-world graphs such as the world-wide web. For extensive surveys of this area, see [2, 15]. To date, most of the work on these models has been experimental or heuristic, but there is an increasing body of precise mathematical work. For a partial survey of this see [7]. One of the questions frequently asked about random graphs in general, and these new models in particular, is the following. How robust is the graph under random failures? More precisely, if vertices

† Royal Society Research Fellow.

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

or edges of the n-vertex graph are deleted independently with a certain probability $1-p$, how large is the largest component in what remains? In particular, for which p is there a *giant component*, *i.e.*, a component of order $\Theta(n)$? If there is a critical probability p_0 above which such a giant component appears, how large is the giant component when $p = p_0 + \varepsilon$? These are the standard finite versions of the basic questions in percolation theory; of course, here the medium in which percolation is being studied is itself random. Our aim is to answer the last question very precisely for certain models, greatly improving the existing bounds.

Probably the most studied of the new random graph models mentioned above is the Barabási–Albert or BA 'model' of [4]. Given an integer $m \geqslant 1$, the BA model describes a graph which grows by adding one vertex at a time, the new vertex sending out m edges to earlier vertices, chosen according to the 'preferential attachment' rule: the probability that an old vertex is chosen to receive an edge from the new vertex is proportional to its degree. Unfortunately, the description given by Barabási and Albert in [4] (repeated in [2]) is incomplete, and also inconsistent; see [9] for the details, or [7] for a fuller discussion. In fact, it is not obvious how to define a simple graph that has exactly the desired properties. In order to prove results, one must first have a well-defined model, and this is why the *linearized chord diagram* or LCD model was introduced in [9]. This is a precise way of defining a multi-graph (which will have loops and multiple edges, but only very few) that fits the BA description. The LCD model also has the very useful property that it has a static description: there is a way of constructing the n-vertex graph directly, rather than by adding one vertex at a time.

Typical graphs generated by the vague BA model or precise LCD model are 'scale-free' in the sense that their degree sequences follow a power-law (see [4, 11]); this perhaps unfortunate terminology is widely used in this area. The robustness of these graphs has been studied experimentally in [3], heuristically in [13, 14] and rigorously in [8]. In the last paper it is proved that for $m \geqslant 2$, if a constant fraction $p > 0$ of the edges or vertices is retained, then there is always a giant component, but for small p it is very small; the fraction of the vertices that it contains is between $\exp(-\Theta(p^{-2}))$ and $\exp(-\Theta(p^{-1}))$ as $p \to 0$. Here we show that the larger of these two bounds is correct, and find the constant in the exponent in terms of m.

In this way, and in many other ways, the LCD model behaves very differently from a classical random graph $G(n, c/n)$ with a comparable number of edges, which has a giant component if and only if c exceeds a critical value, $c_0 = 1$. One can ask whether this is due to the growth in time, or to preferential attachment. To answer this it makes sense to study graphs which grow in time, with each new vertex joining to a set of m, say, earlier vertices chosen uniformly at random. This model, the *uniformly grown m-out random graph*, studied in [10], is extremely natural in its own right as a model of graphs which grow in time, but where *a priori* we know nothing else about the graph than that it has a number of edges linear in the number of vertices. Even more natural is a slight modification where edges are present independently, the *uniformly grown random graph* or c/j-graph, studied in [6], which was introduced much earlier by Dubins in 1984. See the next section for precise definitions and results, as well as further background on the specific models. One can argue that these models are 'scale-free' in a very different and more natural sense

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

than the BA/LCD model. In any case, we shall see in the next section that the models considered here are related closely enough to justify studying them together.

In [6] and [10] it was shown that the uniformly grown graphs described above are in one sense more like $G(n,c/n)$ than the LCD model: there is a threshold in terms of the parameter c (or the edge-retention probability p) such that a giant component appears only above this threshold. (The result for the c/j-graph essentially dates back to 1989; see [19, 21] for the related earlier results, and [17, 6] for a discussion of their relevance.) However, in [6, 10] it is shown that, above the threshold, the giant component emerges very slowly indeed: for $c = c_0 + \varepsilon$, where c_0 is the critical value of the parameter, upper and lower bounds are given on the size of the giant component which are of the form $\exp(-\Theta(\varepsilon^{-1/2}))$. Here we shall give much more precise bounds for a common generalization of these two models, showing that the lower bounds in [6, 10] are essentially correct, and finding the constant in the exponent exactly.

The questions studied here are all of the form: What is the size of the giant component (if any) in a certain inhomogeneous random graph? Such questions have been studied previously in various cases, generally requiring independence between the edges, which we do not have for two of the models studied here. For example, Söderberg [22, 23, 24] considers a general model with finitely many types of vertices, where the edge probabilities depend on the vertex types, giving formulae for the critical point and giant component size without proof. This model does not cover cases such as that studied here, where the space of types is infinite and the ratios between edge probabilities are unbounded. The case of a rather different growing network model (that nevertheless essentially includes the c/j-graph) has been studied heuristically by Dorogovtsev, Mendes and Samukhin [16]; we return to this in the next section.

Of the models we study, perhaps the most interesting and important is the LCD model. This is the prime example of a properly defined model fitting the vague Barabási–Albert description; this BA 'model' is the model in which there is most current interest. On the other hand, the percolation behaviour near the phase transition is most interesting for the uniformly grown models; the very slow emergence of a giant component above some nonzero threshold (an 'infinite order phase transition') is very striking, and it is surprising that accurate bounds on the size of this giant component can be obtained.

2. Models and results

Throughout this paper we shall write $C_1(G)$ for the maximal order (number of vertices) of a component of a graph G. We say that an event holds *with high probability* (w.h.p.) if it holds with probability tending to 1 as $n \to \infty$, with all other parameters fixed.

2.1. The LCD model

The first model we shall consider is the LCD model of [9], a precise version of the vague Barabási–Albert 'growth with preferential attachment' model [4]. The LCD (multi-)graph $G_m^{(n)}$ on n vertices, where $m \geqslant 1$ is a fixed integer parameter, may be defined as follows. Start with $G_m^{(0)}$ the empty graph with no vertices, or with $G_m^{(1)}$ the graph with one vertex, 1, and m loops. Given $G_m^{(n-1)}$, construct $G_m^{(n)}$ by adding a single new vertex n, and then

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

successively adding m edges $e_1, \dots, e_m$ incident with n. The other end of each e_i is a vertex chosen from $\{1, 2, \dots, n\}$ with probability proportional to its degree, where we count $e_1, \dots, e_{i-1}$ as already contributing to the degrees of their endpoints, and furthermore count e_i as contributing one to the degree of n. $G_m^{(n)}$ contains some loops and multiple edges, but not very many; their presence will not affect our results.

At first sight it is not clear why one should take exactly this definition for $G_m^{(n)}$, which is expanded on in [9]. The answer is that $G_m^{(n)}$ defined in this way has a static description derived via random pairings, or linearized chord diagrams; see [9]. This static description seems to be needed to prove all but the simplest results about $G_m^{(n)}$.

As mentioned in the introduction, the robustness of $G_m^{(n)}$ was considered in [8]. Given $0 < p < 1$, let $G_m^{(n)}(p)$ be the graph formed from $G_m^{(n)}$ by keeping each vertex with probability p, independently of the other vertices and of $G_m^{(n)}$. In [8] it was shown that if $m \geqslant 2$ is fixed and p is small, then w.h.p.

$$\exp(-\Theta(p^{-2})) \leqslant C_1(G_m^{(n)}(p))/n \leqslant \exp(-\Theta(p^{-1})).$$

Here we show that the upper bound is correct, and give a much more precise form of the bound.

Theorem 2.1. *Let $m \geqslant 2$ and $0 < p < 1$ be fixed, and let $G_m^{(n)}(p)$ be the graph formed from $G_m^{(n)}$ by keeping vertices independently with probability p. There is a constant $\lambda_m(p)$ such that*

$$C_1(G_m^{(n)}(p)) = (\lambda_m(p) + o(1))n$$

holds w.h.p. as $n \to \infty$. Furthermore, as $p \to 0$ with m fixed, we have

$$\Omega\left(p^2\left(\frac{m-1}{m+1}\right)^{\frac{1}{2p}}\right) = \lambda_m(p) = O\left(\left(\frac{m-1}{m+1}\right)^{\frac{1}{2p}}\right). \tag{2.1}$$

In particular, as $p \to 0$ with m fixed,

$$\lambda_m(p) = \exp\left(-\frac{c_m}{p} + O(\log(1/p))\right) = \exp\left(-\frac{c_m + o(1)}{p}\right), \tag{2.2}$$

where

$$c_m = \frac{1}{2}\log\left(\frac{m+1}{m-1}\right) = \frac{1}{m} + \frac{1}{3m^3} + O(m^{-5}). \tag{2.3}$$

There will be essentially no difference if we delete edges rather than vertices: the only change is that our bounds on the size of the largest component increase by a factor $1/p$. This affects the bounds (2.1) on $\lambda_m(p)$, but not the asymptotic form (2.2). A similar comment applies to all our results.

The restriction to $m \geqslant 2$ above is essential (a fact missed by some heuristic approaches). For $m = 1$ the model gives a forest with loops, and, as noted in [8], for *any* $p < 1$ w.h.p. there is no giant component.

2.2. The $c/\sqrt{ij}$-graph

It is easy to check that in the j-vertex LCD graph, the expected degree of vertex $i < j$ is $m\sqrt{j/i}(1 + O(1/i))$. It follows from the preferential attachment rule that in $G_m^{(n)}$ the

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

probability of the edge ij, $i < j$, is $\frac{m}{2\sqrt{ij}}(1 + O(1/i))$. This suggests that we should compare $G_m^{(n)}$ with the $c/\sqrt{ij}$*-graph*, *i.e.*, the graph on $\{1, 2, \ldots, n\}$ in which each edge ij is present independently, with probability $c/\sqrt{ij}$. Here $c > 0$ is a parameter, and if $c > \sqrt{2}$ we should write $\min\{1, c/\sqrt{ij}\}$ for the edge probabilities. We shall not do so because our interest is in c tending to zero, and in any case all our arguments do not depend on the probabilities of the first $O(1)$ edges.

The $c/\sqrt{ij}$-graph $G^{(n)}_{1/\sqrt{ij}}(c)$ is very much easier to study that $G_m^{(n)}$, due to the independence between edges, and because the form of the 'kernel' $1/\sqrt{ij}$ makes summing over paths very simple. We shall prove the following rather precise result concerning the largest component in $G^{(n)}_{1/\sqrt{ij}}(c)$.

Theorem 2.2. *There is a function $f(c)$ such that for $c > 0$ fixed, w.h.p. the largest component of $G^{(n)}_{1/\sqrt{ij}}(c)$ has order $(f(c) + o(1))n$ as $n \to \infty$. Furthermore,*

$$f(c) = 2e^{1-\gamma} \exp\left(-\frac{1}{2c}\right)\left(1 + O\left(c^{-1}e^{-1/(2c)}\right)\right), \tag{2.4}$$

as $c \to 0$. In particular,

$$f(c) \sim 2e^{1-\gamma} \exp(-1/(2c)),$$

where $\gamma = 0.57721566\ldots$ is Euler's constant.

This result shows that the LCD graph does not behave as if the edges were independent, but that in a certain numerical sense, the difference is fairly small when m is large. For comparison, the correct normalization is to take $c = pm/2$, so that $G_m^{(n)}(p)$ and $G^{(n)}_{1/\sqrt{ij}}(c)$ have (approximately) the same individual edge probabilities. If edges were independent in the LCD graph, Theorem 2.2 shows that the constant c_m given by (2.3) would be equal to $1/m$. As we can see this is not true, but the difference in the constant in the exponent shrinks surprisingly quickly as m becomes large. Note, however, that small m is the case most often considered. This kind of behaviour has been seen before, in a much simpler context: it was shown in [7] that the number of triangles (or l-cycles) in the LCD graph is almost the same as if edges were independent – there is an extra factor $1 + O(1/m)$. Such results are not so surprising in the light of the equivalent description of the LCD graph given in [9]. There it is shown that, roughly speaking, we may construct the LCD graph by taking certain random variables $X(i)$ for each vertex i, whose distribution is that of a sum of m exponentials, and then putting in each edge ij, $i < j$, with probability $X(i)/\sqrt{ij}$. (There is still some dependence, due to the fixed out degree.) Typically, the $X(i)$ do not vary by more than a factor $1 + O(1/m)$.

2.3. The uniformly grown *m*-out random graph

In the light of the very substantial differences between the LCD/BA graph and a classical random graph (for example, $G(n, c/n)$, or an m-out random graph), it makes sense to ask whether it is growth or preferential attachment that causes these differences. This has lead several people to consider the following model $H_m^{(n)}$, the *uniformly grown m-out random graph*. This is the graph on $\{1, 2, \ldots, n\}$ in which each vertex $i \geqslant m + 1$ is joined to a set of

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

m vertices from $\{1,2,\dots,i-1\}$, chosen uniformly at random from all such sets, the choices being independent over different i. Such a model was considered by Barabási and Albert in their original paper [4], where they noted that it does not have a power-law degree sequence. In fact, the degree sequence of $H_m^{(n)}$ is not that far from a classical random graph with constant average degree. For example, the maximum degree is $\Theta(\log n)$. It is thus not surprising that $H_m^{(n)}$ does display a phase transition: let $H_m^{(n)}(p)$ be the random subgraph of $H_m^{(n)}$ obtained by keeping each edge (or vertex) independently with probability p. It was shown in [8] that for $p < p_l$ there is no giant component, while for $p > p_u$ there is, with $p_l, p_u = \frac{1}{4m}(1 + O(1/m))$. The critical probability $p_c = p_u$ was found in [10], building on the work in [6], and upper and lower bounds on the size of the giant component were given for $p = p_c + \varepsilon$, matching up to a factor π in the exponent. Here we shall show that the lower bound from [10] is correct, and give a rather more precise form.

Theorem 2.3. *Let $m \geqslant 2$ and $0 < p < 1$ be fixed, and set*

$$p_c = p_c(m) = \frac{1}{2}\left(1 - \sqrt{\frac{m-1}{m}}\right) = \frac{1}{4m} + \frac{1}{16m^2} + O(m^{-3}). \tag{2.5}$$

Let $H_m^{(n)}(p)$ be the random subgraph of the uniformly grown m-out random graph $H_m^{(n)}$ obtained by retaining each edge (or each vertex) independently with probability p. There exists a function $f_m(p)$ of m and p, not depending on n, such that $C_1(H_m^{(n)}(p)) = (f_m(p) + o(1))n$ holds w.h.p. as $n \to \infty$, with $f_m(p) = 0$ for $p \leqslant p_c$, and

$$f_m(p_c + \varepsilon) = \exp\left(-\frac{\pi}{2(m(m-1))^{1/4}}\frac{1}{\sqrt{\varepsilon}} + O(\log(1/\varepsilon))\right) \tag{2.6}$$

as $\varepsilon \to 0$ from above with m fixed.

Again, it does not matter whether it is edges or vertices that we retain independently with probability p; the precise value of $f_m(p)$ changes by a factor of p, but this does not affect (2.6). Also, the restriction to $m \geqslant 2$ is again essential; otherwise there is never a giant component. Although we have written the result in a slightly different way, the constant in the exponent is the same as that given by the lower bound in [10].

2.4. The uniformly grown or c/j-graph

The model described in the previous subsection represents a graph which grows in time, each new vertex sending edges to old vertices chosen uniformly at random. While always sending out the same number m of edges is one natural choice, it is perhaps not the most natural. One could consider instead sending out (essentially) a Poisson number of edges with a certain mean c. This has the advantage that different edges are then present independently, leading to the *uniformly grown random graph* or *c/j-graph* $G_{1/j}^{(n)}(c)$, the graph on $\{1,2,\dots n\}$ in which edges are present independently, and each edge ij, $i < j$, has probability c/j. (Or $\min\{1, c/j\}$ if $c > 2$; again this will not be relevant here.) It might be more natural to use $j-1$ instead of j so that the expected out-degrees are all exactly c; such minor changes to the model will not matter here, as can be seen from our general result in Section 8.

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

The claim that such a growing random graph model is very natural is supported by its independent introduction in several places. Callaway, Hopcroft, Kleinberg, Newman and Strogatz [12] introduced a (more or less) equivalent model, the CHKNS model, in 2001, giving heuristic arguments that there is a phase transition at a parameter equivalent to $c = 1/4$, and heuristic and numerical evidence for an infinite order phase transition. In fact, the c/j-graph is much older; it was proposed by Dubins in 1984 (see [19, 21]) who asked (in an essentially equivalent infinite form) the question: What is the critical value of c for percolation? The answer, $c = 1/4$, was proved by Kalikow and Weiss [19] who gave the lower bound, and Shepp [21] who gave the upper bound. A more general result was proved by Durrett and Kesten [18]. As pointed out in [6], and also independently by Durrett [17], these results carry over to the CHKNS model. The question of the rate of emergence of the giant component in the c/j-graph was treated rigorously in [6]; as in [10], upper and lower bounds matching up to a factor π in the exponent were proved, showing that the transition is indeed infinite order. This work predates that in [10] on the more complicated m-out version, which builds very much on the work in [6].

Here we again close the gap, proving the following result.

Theorem 2.4. *There is a function $f(c)$ such that for $c > 0$ fixed, $C_1(G^{(n)}_{1/j}(c)) = (f(c) + o(1))n$ holds w.h.p. as $n \to \infty$. Furthermore, $f(c) = 0$ for $c \leqslant 1/4$, and*

$$f(1/4 + \varepsilon) = \exp\left(-\frac{\pi}{2}\frac{1}{\sqrt{\varepsilon}} + O(\log(1/\varepsilon))\right) \tag{2.7}$$

as $\varepsilon \to 0$ from above.

Again this result carries over to the CHKNS model, taking $c = 2\delta$, where δ is the edge-density parameter in [12]. The argument is exactly the same as in [6], so we do not repeat it here.

For comparison between Theorems 2.3 and 2.4, note that the individual edge probabilities in $H^{(n)}_m(p)$ and $G^{(n)}_{1/j}(mp)$ are (essentially) the same. Thus, if edges in $H^{(n)}_m(p)$ behaved exactly as if independent, we would have $p_c(m) = 1/(4m)$ in place of (2.5), and the constant in the exponent in (2.6) would increase by a factor $((m-1)/m)^{1/4}$, noting that $p = p_c + \varepsilon$ in Theorem 2.3 corresponds to $c = 1/4 + m\varepsilon$ in Theorem 2.4. Hence, we see that the dependence between edges going out from a vertex to earlier vertices, due to the fixed out-degree, affects the critical probability and rate of emergence of the giant component by a factor $1 + O(m^{-1})$. Again this is in some sense a small change; however, one might expect dependence that is large enough to shift the critical probability to radically change the size of the extremely small giant component just above that probability. This does not happen here.

Dorogovtsev, Mendes and Samukhin [16] found by 'rigorous' methods the critical probability and size of the giant component in a family of growing scale-free random graphs. One special case is exactly the CHKNS model. The results of [16] in this case are consistent with, and much more precise than, Theorem 2.4. However, their arguments are not a proof; various possible problems to do with limits are not addressed, for example,

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

including the possible existence of $\Theta(n)$ vertices in 'intermediate' sized components. It may well be, however, that their (difficult) arguments can be made rigorous with further work, perhaps using as a starting point the branching process approximation proved here (see Section 3).

The range of growing models considered in [16] includes graphs with power-law degree distributions, with variable exponents. However, the model is very different from $G_m^{(n)}(p)$; not only are edges added at each step between two random old vertices, but, more importantly, preferential attachment is based on degree *in the final graph*. In contrast, percolation in a preferential attachment graph means growth with preferential attachment, followed by the deletion of some edges; in [16] a random subgraph of a graph generated in some way is never considered. This makes a huge difference to the behaviour of the giant component (except in one case, namely uniform attachment). Setting the parameters of their model to produce a sparse graph with the same degree exponent as $G_m^{(n)}$ or $G_m^{(n)}(p)$ (namely, -3), Dorogovtsev, Mendes and Samukhin [16] find a component containing a constant fraction of the edges, rather than the fraction $\exp(-\Theta(1/p))$ obtained for $G_m^{(n)}(p)$. The heuristics of [16] are not obviously applicable to $G_m^{(n)}(p)$ and related graphs.

2.5. Further models

Our results about the uniformly grown random graph and uniformly grown m-out random graph will be proved by considering a common generalization we call the *uniformly grown Z-out random graph*. Roughly speaking, this is the graph in which each vertex i chooses independently a number Z_i of edges to send to earlier vertices, which are then chosen uniformly at random. Each Z_i has (roughly) the distribution of Z. Precise definitions and a result generalizing Theorems 2.3 and 2.4 are given in Section 8. The case where $Z = m$ is constant is $H_m^{(n)}$, the case where Z is Poisson with mean c is $G_{1/j}^{(n)}(c)$. This graph (called X-out rather than Z-out) was introduced in [10] but with stronger assumptions than we shall make here. Again our result (Theorem 8.1) gives much better bounds. As in [10], it turns out that it is the first two moments of Z that matter.

Having noted that the uniformly grown Z-out graph is much easier to handle when Z has a Poisson distribution, so edges are independent, it is natural to ask whether the LCD model can be modified in this way. The answer is, of course, yes; taking the definition for the LCD model given at the start of this section, one can certainly add a random number of edges for each new vertex rather than just m. Such a model could be analysed using the methods of [9]; just as $G_m^{(n)}$ is constructed by first constructing $G_1^{(n)}$ (via random pairings of random points in $[0,1]$, to achieve some independence), and then merging vertices m at a time, a generalization could be constructed from $G_1^{(n)}$ by merging vertices in groups of random sizes. However, the gain is not so great as for the uniformly grown case – we will not achieve independence between edges. The construction outlined above shows that the positive correlation between edges ij, ik, $i < j < k$ will remain (and is likely to be of the form $1 + O(1/\mathbf{E}(Z))$). In fact, it seems likely that the very close (second-order) agreement between the constant c_m appearing in (2.3) for the LCD graph $G_m^{(n)}$ and the value $1/m$ for the related graph with independence arises from the near cancellation of two effects: in $G_m^{(n)}$ there is positive correlation by a factor

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

of roughly $1 + 1/m$ between edges ij, ik, $i < j < k$, arising from preferential attachment, and negative correlation with correlation factor of roughly $1 - 1/m$ between edges ik, jk, $i < j < k$, resulting from the out-degrees being fixed at m. It seems likely that in the Z-out version of the LCD graph, where Z is Poisson with mean m, as only the positive correlation effect would remain, the equivalent of (2.3) would be $c'_m = 1/m - \Theta(1/m^2)$. More precisely, perhaps

$$c'_m = \frac{1}{m} - \frac{1}{3m^2} + O(m^{-3})$$

would hold. (This is based on solving (7.3) with m in place of $m - 1$ and writing $(p'/2)\log(\rho)$ as a series in m.) It is possible that the methods used in [9], [8] and here could be used to prove such a result, but any proof is not likely to be short!

Finally, in [18], Durrett and Kesten consider an inhomogeneous random graph generalizing the c/j graph: edges are present independently, and the probability that the edge ij, $i < j$, is present is given by $f(i/j)/j$ for some function f. Working in a slightly different context (asking whether the infinite graph is almost surely connected), they give a criterion in terms of f that determines whether or not there is a giant component. It would be interesting to see whether the corresponding phase transition always has infinite order, perhaps with growth of the form $\exp(-\Theta(\varepsilon^{-1/2}))$, where the constant depends on f.

3. The method

We shall use a single basic method to analyse all the models we consider. This is to compare (in fact couple) the neighbourhood expansion process in the random graph G with a certain branching process. Here the neighbourhood expansion process is the random process $(\Gamma_k) = (\Gamma_0, \Gamma_1, \Gamma_2, \dots)$, where Γ_0 consists of a single uniformly chosen random vertex v_0 of G, and Γ_k is the set of vertices at graph distance exactly k from v_0. In all cases, it is not just the number of vertices in each Γ_k that matters, but also where in the graph they are, and often how they were reached. Thus the branching process $(X_k) = (X_0, X_1, \dots)$ we compare with will consist of sets of points x each of which has a certain *type* $\tau(x)$ associated to it. In the simplest case these types will be real numbers in $(0, 1]$. The distribution of the number and types of the children of a point in the branching process will depend on its type.

This coupling result needs to some extent separate arguments in each case. It has already been proved in [8] for the LCD model. After proving a suitable coupling result, there is still much work to be done: having checked that almost all vertices not in small components are in a single giant component, we obtain the relative size of the giant component as the survival probability of the branching process, which in turn is given by the solution to certain non-linear integral equations. At first sight it is not clear how to accurately bound the solution to these equations. It turns out that splitting off early vertices, which are likely to have large degrees, and solving exactly the linearized version of the equations for the resulting process (with a slightly modified value of the edge-probability parameter) gives good results.

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

A consequence of the method used here is that any potential improvement of the bounds in any of the results requires no further combinatorial work. For fixed parameters, the limit of the normalized size of the giant component is given exactly by the survival probability of a certain branching process, and hence by the solution to certain equations. The task of obtaining better bounds is exactly the task of obtaining better estimates on the solutions to these equations.

Another consequence is that it does not matter whether it is edges or vertices that we retain independently with probability p; in the branching process approximation to the neighbourhood expansion process we only consider components that are trees. For these, it makes no difference whether it is each 'outward' edge or the vertex at its 'far end' that is retained with probability p. The expected number of vertices in small components, and hence the size of the giant component, will (in the $n \to \infty$ limit) be exactly a factor of p smaller for vertex deletion, due to the reduced number of starting vertices for the neighbourhood expansion. Of course we could also consider deletion of both edges and vertices.

The rest of the paper is organized as follows. In the next section we consider the simplest model, the $c/\sqrt{ij}$-graph. We shall show that for this model good bounds on the giant component can be obtained directly, by counting paths, rather than by coupling with a branching process. This method will give some intuition into the arguments used later: the way the path counting is done corresponds to the separate consideration of early vertices described above.

In Section 5 we describe precisely the type of branching process we consider, and state and prove some very simple results about such branching processes that we will use to estimate the survival probability. In Section 6 we return to the $c/\sqrt{ij}$-graph and apply the branching process method to prove Theorem 2.2, obtaining much more precise results than by path counting. The analysis of the branching process in this case is very simple and does not require any tricks. The coupling argument is also relatively simple; although this can be viewed as a greatly simplified version of the argument in [8] we give a complete proof. In Section 7 we turn to the LCD model, proving Theorem 2.1; here the relevant coupling result has already been proved, but the analysis of the branching process is more difficult. Finally, in Section 8 we turn to the uniformly grown Z-out model, proving a common generalization of Theorems 2.3 and 2.4.

In all cases our aim is to find the normalized size of the giant component. For this, the branching process method works very well. One might want more information, such as the approximate size of the second largest component above the critical probability, and of the largest component below. It is not clear whether good bounds can easily be obtained by the methods used here. One might guess that the answer to both questions would always be $O(\log n)$, on general grounds based on the behaviour of branching processes. But this is not true: it was shown in [6] by a direct method that the subcritical c/j-graph always has a component of order $n^{\Omega(1)}$.

Throughout the rest of the paper we shall always assume that n, the number of vertices, is at least some sufficiently large constant n_0, which may depend on the parameters (*e.g.*, m, c) of the result we are trying to prove. Thus we shall use inequalities such as $m \log n \leqslant n^{1/10}$ without further comment. We also ignore probability zero events. In particular, we shall

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

not be consistent about the inclusion or exclusion of endpoints when considering whether a continuously distributed random quantity lies in some interval.

4. The $c/\sqrt{ij}$-graph: direct analysis

We start by considering a simple model with edges present independently, motivated by the BA or LCD growth with preferential attachment model. This model is much easier to analyse than the LCD model, due to the independence between edges. In fact, even among models with independence, this model is one of the easiest to analyse, due to the form of the kernel, $c/\sqrt{ij}$. This makes counting paths and finding the eigenfunction of the relevant branching process much simpler than for the uniformly grown random graph considered in [6], with kernel $c/\max\{i,j\}$.

Let $G = G^{(n)}_{1/\sqrt{ij}}(c)$ be the graph on $\{1,2,\dots,n\}$ in which each pair $\{i,j\}$ of vertices is joined independently with probability $c/\sqrt{ij}$. For now, we shall prove a much weaker result than Theorem 2.2, namely, that

$$f_1(c)n \leqslant C_1\big(G^{(n)}_{1/\sqrt{ij}}(c)\big) \leqslant f_2(c)n \tag{4.1}$$

holds w.h.p. for functions $f_1(c)$, $f_2(c)$ satisfying

$$f_1(c), f_2(c) = \exp\left(-\frac{1+o(1)}{2c}\right)$$

as $c \to 0$. The proof of this weaker result will be much simpler.

Proof of (4.1). We start with the upper bound. Fix $\eta > 0$ and set a 'cutoff' $\rho = \exp(-(1-\eta)/c)$. We call a vertex i of G *early* if $i \leqslant \rho n$, and *late* otherwise. It is easy to check that any component of order $\Theta(n)$ is very likely to contain an early vertex: otherwise, the subgraph of G induced by the late vertices contains a component of order $\Theta(n)$. But each individual early vertex has positive probability of being joined to such a component, and the probability that no early vertex joins to it is exponentially small.

We bound the expected number of paths from late to early vertices. More precisely, let us call a path $x_0x_1\cdots x_l$ a *late-early path* if $x_0,\dots,x_{l-1}$ are late vertices and x_l is an early vertex. The expected number N_l of late-early paths of length l is at most

$$\sum_{\rho n < x_0 \leqslant n} \cdots \sum_{\rho n < x_{l-1} \leqslant n} \sum_{0 < x_l \leqslant \rho n} \prod_{j=0}^{l-1} \frac{c}{\sqrt{x_j x_{j+1}}}.$$

In fact, N_l is equal to the sum above when the summation variables are restricted to be distinct. Assuming that ρn is an integer (which we can ensure by adjusting η slightly as n varies), we obtain

$$\begin{aligned} N_l &\leqslant \int_{\rho n < x_0 \leqslant n} \cdots \int_{\rho n < x_{l-1} \leqslant n} \int_{0 < x_l \leqslant \rho n} \prod_{j=0}^{l-1} \frac{c}{\sqrt{x_j x_{j+1}}}\, \mathrm{d}x_l\, \mathrm{d}x_{l-1} \cdots \mathrm{d}x_0 \\ &= c^l n \int_{\rho < x_0 \leqslant 1} \cdots \int_{\rho < x_{l-1} \leqslant 1} \int_{0 < x_l \leqslant \rho} \frac{1}{\sqrt{x_0 x_l}} \prod_{j=1}^{l-1} \frac{1}{x_j}\, \mathrm{d}x_l\, \mathrm{d}x_{l-1} \cdots \mathrm{d}x_0 \end{aligned}$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

$$= c^l n\left(\int_{x_0=\rho}^1 x_0^{-1/2}\,\mathrm{d}x_0\right)\left(\int_{x_l=0}^{\rho} x_l^{-1/2}\,\mathrm{d}x_l\right)\left(\int_{x=\rho}^1 x^{-1}\,\mathrm{d}x\right)^{l-1}$$
$$= 4c^l n\big(1-\rho^{1/2}\big)\rho^{1/2}(\log(1/\rho))^{l-1}$$
$$\leqslant 4cn\rho^{1/2}(c\log(1/\rho))^{l-1}.$$

By our choice of ρ, we have $c\log(1/\rho) = 1-\eta < 1$. It follows that the expectation of the total number of late-early paths is at most

$$\sum_{l\geqslant 1} N_l \leqslant 4c\eta^{-1}\rho^{1/2}n \leqslant \exp(-(1/2-\eta)/c)n \tag{4.2}$$

if η is fixed and c is small enough. In fact, arguing as at the end of Section 6 of [6], we can obtain the same formula (4.2) as an upper bound on the number of late-early paths holding w.h.p., rather than just in expectation. Since w.h.p. any giant component contains an early vertex, w.h.p. the size of the giant component is at most the number ρn of early vertices plus the number of late-early paths, so for small enough c, w.h.p.

$$C_1(G) \leqslant \rho n + \exp(-(1/2-\eta)/c)n \leqslant \exp(-(1/2-2\eta)/c)n.$$

As $\eta > 0$ was arbitrary, this proves the upper bound in (4.1).

We remark that a more complicated version of the argument above appeared in Section 6 of [6]; the argument for the apparently rather distantly related uniformly grown random graph considered there is more complicated, but proceeded by comparison to a path counting argument in G.

The lower bound is an easy application of Theorem 10 of [6]. Again we take a cutoff, this time $\rho = \exp(-(1+\eta)/c)$, and consider the integral operator on $C[\rho,1]$ with kernel $\kappa(x,y) = c/\sqrt{xy}$. Since

$$\int_{y=\rho}^1 \frac{1}{\sqrt{xy}}\frac{1}{\sqrt{y}}\mathrm{d}y = \log(1/\rho)\frac{1}{\sqrt{x}},$$

this has an eigenfunction $\phi(x) = x^{-1/2}$ on $[\rho,1]$, with eigenvalue $c\log(1/\rho) = (1+\eta)$. Viewing κ as the kernel corresponding to an induced subgraph of G, we may thus apply Theorem 10 of [6], with $c = 1/(1+\eta)$ and $\gamma = \eta$. As in [6], there is a very minor technical point concerning the scaling, as the theorem we apply concerns a continuous kernel defined on $[0,1]^2$ rather than $[\rho,1]^2$; it turns out that rescaling does not introduce an extra factor. We conclude that w.h.p. the induced subgraph of G on the vertices $\{\rho n,\dots,n\}$ has a component of order at least

$$\frac{\eta}{1+\eta}n\int_\rho^1 \phi(x)\mathrm{d}x/\sup_{\rho\leqslant x\leqslant 1}\phi(x).$$

Since $\int_\rho^1 \phi(x)\mathrm{d}x \to 2$ as $c\to 0$, while

$$\sup_{\rho\leqslant x\leqslant 1}\phi(x) = \phi(\rho) = \rho^{-1/2} = \exp((1+\eta)/(2c)),$$

this shows that w.h.p. we have

$$C_1(G) \geqslant (2+o(1))\eta\exp(-(1+\eta)/(2c))n \geqslant \exp(-(1+2\eta)/(2c))n$$

if c is small enough. Since $\eta > 0$ was arbitrary, this completes the proof. □

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

We shall return to this model after introducing the basic method we use to analyse all the models considered here. However, the result and proof above illustrate a fact that we shall use later in the analysis of the branching process itself: starting from a supercritical random graph, or branching process, good bounds on the size of the giant component can often be obtained by dealing separately with the few 'early' vertices of highest degree. This turns out to apply to all the models considered here.

5. Branching process background

In this section we state and prove some simple results about the survival probability of certain continuous-type branching processes, in which each point of the process has a *type*, here a real number (or a real number and an integer) associated to it, which affects the distribution of the number and type of its offspring.

Almost certainly, all results we need are special cases of standard results in the theory of branching processes. However, because our particular processes are not of a very standard type, it is likely that any result of sufficient generality to include the cases we need would actually have much more generality than we would like to go into here. Also, it would be misleading to give the impression that we need anything deep from the theory of branching processes: all we need are some very basic first results, which are intuitive and easy to prove. Thus the presentation here will be self-contained. Our aim is to prove only what we need, so we shall solve all problems of analysis (existence of integrals/limits, *etc.*) in the easiest way; rather than aiming to keep generality, we shall impose much stronger conditions than those likely to be really needed whenever it is convenient to do so.

Our branching processes will be associated with a certain *type space* $\mathscr{T}$. It will be convenient to think of $\mathscr{T}$ as the half-open interval $(0,1]$; in all the examples $\mathscr{T}$ will be either $(0,1]$, or $(0,1]\times\{1,2,\dots r\}$ for $r=2$ or $r=3$. We shall write our branching process as $(X_k)=(X_0,X_1,X_2,\dots)$. Each generation X_k consists of a set of *points* x, each of which has a *type* $\tau(x)\in\mathscr{T}$ associated to it. The first generation X_0 will consist of a single point x_0 whose type is chosen according to some distribution. Each $x\in X_k$ with type $\tau(x)=\alpha$ will give rise independently to a set of children $y_1,\dots,y_n$, $0\leqslant n<\infty$, with types $\beta_i=\tau(y_i)$, where n has a certain distribution depending on α, and, given n, the vector $(\beta_1,\dots,\beta_n)$ has a certain continuous distribution, also depending on α. The next generation X_{k+1} will be the disjoint union of the sets of children of all $x\in X_k$.

In general there is a probability density

$$f_n(\alpha,\beta_1,\dots,\beta_n)\,\mathrm{d}\beta_1\cdots\mathrm{d}\beta_n$$

that a point x with $\tau(x)=\alpha$ has exactly n children and that their types lie in small intervals around $\beta_1,\dots,\beta_n$. We shall write $X_1(\alpha)$ for the distribution of the number and types of the children of a point x with type $\tau(x)=\alpha$. The simplest example is the Poisson case, where x gives rise to children whose types β are distributed as a Poisson process with a certain density $g(\alpha,\beta)$. In this case

$$f_n(\alpha,\beta_1,\dots,\beta_n)=\exp\biggl(-\int_{\mathscr{T}}g(\alpha,\beta)\,\mathrm{d}\beta\biggr)\frac{1}{n!}\prod_{i=1}^{n}g(\alpha,\beta_i). \tag{5.1}$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

(Here we treat $(\beta_1,\dots,\beta_n)$ as an n-tuple; of course, it is only the set $\{\beta_1,\dots,\beta_n\}$ that matters in the end.)

Another example is the binomial case, where the number of children of x has a binomial distribution whose parameters depend on α, and the type of each child is independently distributed according to some density depending on α but not the number of children. Two of our examples will involve a mixture of the Poisson and binomial cases.

We shall make the following assumptions about $X_1(\alpha)$. Firstly, we shall assume that $X_1(\alpha)$ is continuous in the following strong sense: for every α and every $\varepsilon > 0$, there is a $\delta > 0$ such that if $|\alpha' - \alpha| < \delta$ then $X_1(\alpha)$ and $X_1(\alpha')$ can be coupled so as to agree exactly (same number of children, and exactly the same types) with probability at least $1 - \varepsilon$. (In the more general case $\mathscr{T} = (0,1] \times \{1,2,\dots r\}$, the condition $|\alpha' - \alpha| < \delta$ of course means that the discrete parts of α, α' coincide, and the continuous parts are close.)

Secondly, we shall assume that the branching process is connected, in the sense that there is some $k > 0$ such that, for any $\alpha, \beta \in \mathscr{T}$ and any neighbourhood N_β of β, starting from $X_0 = \{x_0\}$ with $\tau(x_0) = \alpha$ there is a positive probability that X_k contains some y with $\tau(y) \in N_\beta$. Roughly speaking, this condition says that a point of any type can have descendants of any other type. (The weaker assumption where k is allowed to depend on α and β should suffice.)

Finally, we assume that there is some $\alpha \in \mathscr{T}$ such that a point of type α has positive probability of having at least two children.

Let us say that an $x \in X_k$ *survives* if x has descendants in all later generations. Let $S_\infty(\alpha)$ be the probability that x survives, given that $\tau(x) = \alpha$. Consider the following operator $\mathbf{F}$ which takes as input a (measurable) function $\phi : \mathscr{T} \to [0,1]$ and returns another such function:

$$\mathbf{F}(\phi)(\alpha) = \sum_{n \geqslant 0} \int_{\beta_1,\dots,\beta_n \in \mathscr{T}} f_n(\alpha, \beta_1, \dots, \beta_n) \left(1 - \prod_{i=1}^{n} (1 - \phi(\beta_i))\right) \mathrm{d}\beta_n \cdots \mathrm{d}\beta_1.$$

This has the following interpretation. Start with a point x_0 with $\tau(x_0) = \alpha$, and generate its children. To each child with type β, assign 'success' independently with probability $\phi(\beta)$. Then x_0 has at least one successful child with probability $\mathbf{F}(\phi)(\alpha)$. In particular, we may take $\phi = S_\infty$, assigning 'success' to each point that survives. In this case, the probability that x_0 has at least one successful child is exactly the probability that x_0 survives, so

$$S_\infty = \mathbf{F}(S_\infty).$$

Let us note the following properties of $\mathbf{F}$. Whatever the function $\phi : \mathscr{T} \to [0,1]$, the function $\mathbf{F}(\phi)$ is continuous on $\mathscr{T}$. This follows from our strong continuity assumption: for α and α' close, we can couple the two sets of children to agree with high probability; when they disagree, we use only the fact that ϕ takes values in $[0,1]$ to bound the difference between $\mathbf{F}(\phi)(\alpha)$ and $\mathbf{F}(\phi)(\alpha')$. Also, $\mathbf{F}(\phi)$ clearly takes values in $[0,1]$. Furthermore, $\mathbf{F}$ is monotone: if $\phi_1 \geqslant \phi_2$ holds pointwise, then so does $\mathbf{F}(\phi_1) \geqslant \mathbf{F}(\phi_2)$.

Finally, $\mathbf{F}$ is *superlinear* in the sense that given $\phi : \mathscr{T} \to [0,1]$ and a real number λ, $0 < \lambda < 1$, we have

$$\mathbf{F}(\lambda\phi) \geqslant \lambda \mathbf{F}(\phi). \tag{5.2}$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

(Perhaps sublinear would be the more normal term, but working on $[0,1]$, superlinear seems more natural.) To see this, note that $\mathbf{F}(\lambda\phi)(\alpha)$ has the following interpretation. Start with x_0 with $\tau(x_0) = \alpha$. Generate the children of x_0. Assign to each child having type β 'initial success' with probability $\phi(\beta)$, independently of the other children. Then pick a subset of the initially successful children to be 'successful', selecting each independently with probability λ. $\mathbf{F}(\lambda\phi)$ is the probability that at least one child is successful. In contrast, $\lambda\mathbf{F}(\phi)$ is just λ times the probability that at least one child is initially successful. If p_k is the probability that exactly k children are initially successful, then we have

$$\mathbf{F}(\lambda\phi)(\alpha) = \sum_{k\geqslant 1}(1-(1-\lambda)^k)p_k \geqslant \sum_{k\geqslant 1}\lambda p_k = \lambda\mathbf{F}(\phi)(\alpha),$$

proving (5.2). Note also that if $\phi(\beta) > 0$ for all β, then we have strict inequality in (5.2) for some α, in particular for any α for which x_0 has positive probability of having at least two children.

We have noted already that the survival probability $S_\infty(\alpha)$ solves $S_\infty = \mathbf{F}(S_\infty)$. In fact, it is easy to see that S_∞ is the pointwise supremum of all solutions S to $S = \mathbf{F}(S)$. Taking S_k to be the probability of having descendants in at least the next k generations, we see from the interpretation of $\mathbf{F}$ that $S_{k+1} = \mathbf{F}(S_k)$. In particular, $S_k = \mathbf{F}^k(1)$, the result of applying $\mathbf{F}$ k times to the constant function taking value one. But $S_\infty = \inf_k S_k$, so we have

$$S_\infty(\alpha) = \lim_{k\to\infty}\mathbf{F}^k(1)(\alpha),$$

where the limit is of a decreasing sequence. If S is any function satisfying $S = \mathbf{F}(S)$, then $S \leqslant 1$. From monotonicity of $\mathbf{F}$ we thus have $S = \mathbf{F}^k(S) \leqslant \mathbf{F}^k(1)$ holding pointwise for all k, and $S \leqslant S_\infty$ follows.

Actually, there is at most one nonzero solution to $S = \mathbf{F}(S)$; as pointed out by James Norris [20], this can be shown by looking at the ratio of two solutions. Suppose that S is a solution of $S = \mathbf{F}(S)$ which is not identically zero or equal to S_∞. We may assume that S_∞ is not identically zero, as $S \leqslant S_\infty$. Note that S is continuous, as $S = \mathbf{F}(S)$, and $\mathbf{F}(\phi)$ is always continuous. From connectedness, using $S = \mathbf{F}^k(S)$ for all k, we see that $S(\alpha) > 0$ holds pointwise. (S is nonzero at some β, and therefore in a neighbourhood of that β. Now apply connectedness to show that $S(\alpha) = \mathbf{F}^k(S)(\alpha)$ is positive.) Let $\lambda = \inf S(\alpha)/S_\infty(\alpha)$. Note that $\lambda < 1$ as $S \neq S_\infty$, and $S_\infty \geqslant S$ holds pointwise.

We claim that the infimum above is attained at some α_0; this will require an extra assumption not so far stated. If $\mathscr{T}$ were compact there would be no problem, as S and S_∞ are continuous. We shall use the following additional property of $X_1(\alpha)$: there is some $k_0 \geqslant 1$ such that, given any neighbourhood N_β of some $\beta \in \mathscr{T}$ and any $M > 0$, as $\alpha \to 0$ (or $\alpha \to (0,i)$ in the more general case), the probability that x_0 with $\tau(x_0) = \alpha$ has at least M k_0th generation offspring with types in N_β tends to one. Roughly speaking, this just says that points with very small α have very many children, the expected number (in any interval) tending to infinity. This assumption will hold with $k_0 = 2$ in all cases we consider. This assumption implies that if $\phi : \mathscr{T} \to [0,1]$ is any continuous function not identically zero, then

$$\mathbf{F}^{k_0}(\phi)(\alpha) \to 1 \tag{5.3}$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

as $\alpha \to 0$. (Or as $\alpha \to (0, i)$.) Using (5.3), and the fact that $S = \mathbf{F}^{k_0}(S)$ and $S_\infty = \mathbf{F}^{k_0}(S_\infty)$, we see that $S(\alpha)/S_\infty(\alpha)$ tends to $1/1 = 1$ as $\alpha \to 0$. From this and continuity it follows that the infimum $\lambda = \inf_\alpha S(\alpha)/S_\infty(\alpha) < 1$ must be attained at some $\alpha_0 \in \mathscr{T}$. But now as $S(\alpha_0) > 0$ we have $\lambda > 0$. We have $S(\alpha) \geqslant \lambda S_\infty(\alpha)$ for every α by definition of λ, so from (5.2) and monotonicity of $\mathbf{F}$,

$$S(\alpha_0) = \mathbf{F}(S)(\alpha_0) \geqslant \mathbf{F}(\lambda S_\infty)(\alpha_0) \geqslant \lambda \mathbf{F}(S_\infty)(\alpha_0) = \lambda S_\infty(\alpha_0). \tag{5.4}$$

As we have equality between the first and last terms, equality holds throughout. In particular, $\mathbf{F}(S)(\alpha_0) = \mathbf{F}(\lambda S_\infty)(\alpha_0)$. Since $\mathbf{F}$ is monotone, $S \geqslant \lambda S_\infty$, and S and S_∞ are continuous, it follows that $S(\alpha_1) = \lambda S_\infty(\alpha_1)$ for any α_1 that can be reached from α_0. (That is, for any α_1 such that for any neighbourhood N of α_1 the probability given that $\tau(x_0) = \alpha_0$ that x_0 has a child with type in N is positive.) But then the argument giving (5.4) holds with α_1 in place of α_0. Repeating, and using our connectivity assumption, we see that equality holds in (5.4) if α_0 is replaced by any value α. But this contradicts the fact that strict inequality holds in (5.2) for some α, proving the claim that any solution to $S = \mathbf{F}(S)$ is either S_∞ or identically zero. (Of course, S_∞ may itself be identically zero, if the branching process is subcritical.)

In general, we cannot calculate S_∞; the form of the equation $S = \mathbf{F}(S)$ will be too complicated. What we shall actually use is two simple facts giving bounds on S_∞ in terms of trial functions ϕ satisfying certain conditions.

Fact 1. If $\phi : \mathscr{T} \to [0, 1]$ satisfies $\mathbf{F}(\phi) \geqslant \phi$ pointwise, then $S_\infty \geqslant \phi$ holds pointwise.

This is very easy to see: set $\phi_k = \mathbf{F}^k(\phi)$. Then by assumption $\phi_1 \geqslant \phi_0$. By induction and monotonicity of $\mathbf{F}$, $\phi_{k+1} \geqslant \phi_k$. Thus the pointwise limit $\phi_\infty = \lim_{k\to\infty} \phi_k$ exists, and by monotone convergence of the integral involved,

$$\mathbf{F}(\phi_\infty) = \mathbf{F}(\lim_{k\to\infty} \phi_k) = \lim_{k\to\infty} \mathbf{F}(\phi_k) = \lim_{k\to\infty} \phi_{k+1} = \phi_\infty$$

holds pointwise. Thus ϕ_∞ solves $\mathbf{F}(\phi_\infty) = \phi_\infty$, and as S_∞ is the supremum of all such solutions, $S_\infty \geqslant \phi_\infty \geqslant \phi$.

Fact 2. Suppose that $\phi : \mathscr{T} \to [0, 1]$ satisfies $\mathbf{F}(\phi) \leqslant \phi$ pointwise, and that $\mathbf{F}(\phi)$ is not identically zero. Then $S_\infty \leqslant \phi$ holds pointwise.

The assumption that $\mathbf{F}(\phi)$ is not identically zero is only required to avoid trivialities such as ϕ almost everywhere zero. Recall that $\mathbf{F}(\phi)$ is continuous. As $\mathbf{F}(\phi)$ is not identically zero, it follows from the connectivity assumption that for some k, $\mathbf{F}^k(\phi)$ is nowhere zero. We may and shall assume that $k \geqslant k_0$, the quantity appearing in (5.3) and the preceding assumption, which we use again here. Note that $\mathbf{F}^k(\phi) \leqslant \phi$ holds pointwise, applying monotonicity of $\mathbf{F}$ $k-1$ times to $\mathbf{F}(\phi) \leqslant \phi$. Hence it suffices to show that $S_\infty \leqslant \mathbf{F}^k(\phi)$. Suppose not, and set $\lambda = \inf_{\alpha\in\mathscr{T}} \mathbf{F}^k(\phi)(\alpha)/S_\infty(\alpha)$, so $\lambda < 1$. This infimum is attained at some $\alpha_0 \in \mathscr{T}$. (Again we use continuity of $\mathbf{F}^k(\phi)$ and S_∞, together with (5.3).) Since $\mathbf{F}^k(\phi)$ is nowhere zero, $\lambda > 0$. Now $\mathbf{F}^k(\phi) \geqslant \lambda S_\infty$ holds pointwise by definition of λ. Thus by monotonicity of $\mathbf{F}$, by the assumption $\phi \geqslant \mathbf{F}(\phi)$, and by (5.2),

$$\mathbf{F}^k(\phi) \geqslant \mathbf{F}^{k+1}(\phi) \geqslant \mathbf{F}(\lambda S_\infty) \geqslant \lambda \mathbf{F}(S_\infty) = \lambda S_\infty.$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

At α_0 we have equality between the first and last terms, and hence throughout. Again it follows that $\mathbf{F}^k(\phi) = \lambda S_\infty$ holds pointwise, and hence that we have equality throughout for all α, not just at α_0. But this contradicts the fact that strict inequality in (5.2) always occurs for some α.

6. The $c/\sqrt{ij}$ model revisited

Let us return to the model considered in Section 4, this time using the branching process for more precise analysis. For this section, let $G = G^{(n)}_{1/\sqrt{ij}}(c)$ be the graph on $[n] = \{1, 2, \dots, n\}$ in which each pair of vertices $\{i, j\}$ is joined with probability $c/\sqrt{ij}$, independently of all other pairs.

Our aim is to prove Theorem 2.2, showing that w.h.p. the largest component of G has order $(f(c) + o(1))n$, where $f(c)$ satisfies (2.4). This is a much more precise result than that proved in Section 4.

Let $(\Gamma_k) = (\Gamma_0, \Gamma_1, \Gamma_2, \dots)$ be the neighbourhood expansion process in the graph G, defined as follows: $\Gamma_0 = \{v_0\}$, where v_0 is a vertex of G chosen uniformly at random. For $k \geqslant 0$, set $N_k = \bigcup_{r \leqslant k} \Gamma_r$, and let Γ_{k+1} be the set of vertices of $G \setminus N_k$ adjacent in G to some vertex of Γ_k. In other words, for all k, let Γ_k be the set of vertices at graph distance exactly k from v_0.

6.1. Coupling

We aim to couple (Γ_k) with a branching process $(X_k) = (X_0, X_1, \dots)$ of the general type described in Section 5. Here each generation X_k consists of a set of points x each having a type $\alpha(x) \in (0, 1]$ associated with it. Each such point x gives rise to children y in the next generation whose types $\alpha(y)$ are distributed as a Poisson process on $(0, 1]$ with density $c/\sqrt{\alpha(x)\alpha(y)}$. In particular, the number of children is Poisson with mean $\int_0^1 c/\sqrt{\alpha(x)\beta}\,\mathrm{d}\beta = 2c/\sqrt{\alpha(x)}$. The first generation X_0 consists of a single point x_0 with $\alpha(x_0)$ uniformly distributed on $(0, 1]$.

For a vertex v of G, set $\alpha(v) = v/n$. The aim is to couple the two processes so that in each generation the number of vertices/points coincides, and so do their α values. The intuition is very simple: the chance that a vertex i of G with $\alpha(i) = i/n = \alpha$ has a neighbour j with $\alpha(j) = j/n$ in a certain small interval $[\beta, \beta + \mathrm{d}\beta]$ is roughly

$$\sum_{j \in [n\beta, n(\beta+\mathrm{d}\beta)]} \frac{c}{\sqrt{ij}} \sim \sum_{j \in [n\beta, n(\beta+\mathrm{d}\beta)]} \frac{c}{n\sqrt{\alpha\beta}} \sim \frac{c}{\sqrt{\alpha\beta}}\,\mathrm{d}\beta,$$

which is the probability that a point of type α has a child with type in $[\beta, \beta + \mathrm{d}\beta]$.

In this case, the proof of a suitable coupling result is relatively straightforward and perhaps rather tedious, but we shall write it out in full as this is the basic method underlying all the results of this paper, and this model is much the simplest case.

Lemma 6.1. *Let $c > 0$ be fixed. For each n the processes (Γ_k) and (X_k) can be coupled so that with probability $1 - o(n^{-1/100})$ we have*

$$|\Gamma_k| = |X_k|$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

for $0 \leqslant k \leqslant K$, *and either* $|\Gamma_K| = |X_K| = 0$ *or*

$$\sum_{k=0}^{K+1} |\Gamma_k|, \quad \sum_{k=0}^{K+1} |X_k| \geqslant n^{1/100}. \tag{6.1}$$

In other words, unless a certain event of small probability holds, as far as size is concerned the two processes behave in exactly the same way in the two models until their total size reaches at least $n^{1/100}$ or they die out. A corresponding comparison result is proved in [8] for the much more complicated LCD model, and an appropriate branching process. There almost all the difficulties are due to the lack of independence in the model.

Proof. We shall write ε for $1/100$ throughout, emphasizing that ε is a small positive constant appearing as the exponent of n in the result, whose value we have not attempted to optimize. We shall construct our coupling so that the α values of corresponding points/vertices also agree. Exact agreement is not possible, so we shall ask for agreement only of rounded α values, $\tilde{\alpha}(v) = n^{-1/2}\lceil n^{1/2}\alpha(v)\rceil$, where v is either a vertex of G or a point of some X_k.

We shall inductively couple Γ_k and X_k, $k = 0, 1, 2, \ldots$, until both processes die out, both processes become large (as in the statement of the lemma), or some bad event of small probability occurs. Our coupling will be such that there is a bijection from Γ_k to X_k so that if v maps to x then $\tilde{\alpha}(v) = \tilde{\alpha}(x)$. There is no problem coupling $\Gamma_0 = \{v_0\}$ and $X_0 = \{x_0\}$ in this way, as $\alpha(x_0)$ is by definition uniform on $(0, 1]$, while $\alpha(v_0)$ is uniform on $\{1/n, 2/n, \ldots, n/n\}$, so the rounded values can be coupled to agree with probability $1 - O(n^{-1/2})$.

Let us suppose that we have constructed $\Gamma_0, \ldots, \Gamma_k$ and $X_0, \ldots, X_k$ coupled as above. We aim to extend the coupling to Γ_{k+1} and X_{k+1}. Note that so far we have only 'looked at' potential edges of the graph G that are induced by $N_k = \Gamma_0 \cup \cdots \cup \Gamma_k$, or are incident with N_{k-1}. Hence, conditional on $\Gamma_0, \ldots, \Gamma_k$, each edge from Γ_k to $G \setminus N_k$ is present independently with its original unconditioned probability. Note that from the stopping rule we do not need to consider too many vertices of G. In particular, there is no problem allowing a small ($o(n^{-2\varepsilon})$) failure probability in our coupling for every vertex $v \in N_k$ considered.

We claim that we may assume that $\tilde{\alpha}(v) \geqslant n^{-1/4}$ holds for every $v \in \Gamma_k$, and thus that $\tilde{\alpha}(x) \geqslant n^{-1/4}$ holds for every $x \in X_k$. Suppose not. Then there is a $v \in \Gamma_k$ with $v < n^{3/4} + O(n^{1/2})$. (The $O(n^{1/2})$ allows for the rounding.) Such a vertex v is joined to each $j \in V(G) \setminus N_k$ with $j \geqslant n/2$ independently, with probability $c/\sqrt{vj} \geqslant cn^{-7/8}/2$. But, from our stopping rule, $|N_k| \leqslant n^{\varepsilon}$, so there are at least $n/3$ candidates for the vertex j, and the number of neighbours of v in Γ_{k+1} stochastically dominates a Binomial $\mathrm{Bi}(n/3, cn^{-7/8}/2)$ distribution with mean $cn^{1/8}/6$. As this mean is much larger than n^{ε}, with very high probability, certainly at least $1 - o(n^{-2\varepsilon})$, say, v has at least n^{ε} neighbours in Γ_{k+1}, and in particular, $|\Gamma_{k+1}| \geqslant n^{\varepsilon}$. Similarly, but more simply, with very high probability the corresponding point $x \in X_k$ with $\tilde{\alpha}(x) \geqslant n^{-1/4}$ has at least n^{ε} children, so $|X_{k+1}| \geqslant n^{\varepsilon}$. Thus, we may abandon the coupling at this stage: our stopping rule (6.1) is satisfied for $K = k$ with sufficiently high probability.

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

We assume from now on that the $\tilde{\alpha}(.)$ values for $v \in \Gamma_k$ or $x \in X_k$ are all at least $n^{-1/4}$. Let us consider the vertices $v \in \Gamma_k$ one by one, coupling their contributions to Γ_{k+1} with the children of the corresponding $x \in X_k$. When we come to a particular vertex v, let S be the set of neighbours in Γ_{k+1} of vertices $v' \in \Gamma_k$ previously considered, noting that conditional on everything so far, each vertex w of $V(G) \setminus (N_k \cup S)$ is joined to v independently with probability $c/\sqrt{vw}$. We may assume that $|S| \leqslant n^{\varepsilon}$; otherwise, unless our coupling has already broken down, both Γ_{k+1} and X_{k+1} have cardinality at least n^{ε}, and our stopping rule (6.1) allows us to abandon the coupling. Consider any possible value β of $\tilde{\alpha}(.)$, with $\beta \geqslant n^{-3/8}$. In $V(G)$, there are $n^{1/2} + O(1)$ vertices j with $\tilde{\alpha}(j) = \beta$. Hence, in $V(G) \setminus (N_k \cup S)$ there are $n^{1/2} + O(n^{\varepsilon})$ such vertices. Let us write B for the set of these vertices. Each $j \in B$ is joined to v independently with probability

$$c/\sqrt{vj} = \frac{c}{n\sqrt{\alpha(v)\alpha(j)}} = \frac{c}{n\sqrt{\tilde{\alpha}(v)\beta}}\bigl(1 + O\bigl(n^{-1/8}\bigr)\bigr).$$

For the last step note that for any w the absolute difference between $\alpha(w)$ and $\tilde{\alpha}(w)$ is at most $n^{-1/2}$, and we are assuming that $\tilde{\alpha}(v) > n^{-1/4}$ and $\beta > n^{-3/8}$. Since the individual edge probabilities above are very small, the number of $j \in B$ joined to v is very close in distribution to a Poisson distribution with mean

$$\mu = \sum_{j \in B} \frac{c}{\sqrt{vj}} = \mu_0\bigl(1 + O\bigl(n^{-1/8}\bigr)\bigr),$$

where

$$\mu_0 = \frac{c}{n^{1/2}\sqrt{\tilde{\alpha}(v)\beta}}.$$

In fact, more simply, a direct calculation shows that the probability that v has at least two neighbours in B is at most $\mu^2/2 = O(\mu_0^2)$, while the probability that v has exactly one neighbour in B is $\mu(1 - O(\mu)) = \mu_0 + O(\mu_0^2)$. Hence, the probability that v has no neighbours in B is $1 - \mu_0 + O(\mu_0^2)$. Note that from our bounds on $\tilde{\alpha}(v)$ and β, we have $\mu_0 \leqslant 1/\sqrt{n^{1-1/4-3/8}} = n^{-3/16}$. Also, in the continuous branching process, the number of children y of the point x corresponding to v having $\tilde{\alpha}(y) = \beta$ has a Poisson distribution with mean $\mu' = \mu_0(1 + O(n^{-1/2}))$. We can couple this number with the number of neighbours of v in G so that with probability $1 - O(\mu_0^2 + \mu_0 n^{-1/2}) = 1 - O(\mu_0 n^{-3/16})$ the two numbers agree (and are either 0 or 1). Using independence in both models, we may construct this coupling simultaneously for all possible values $\beta > n^{-3/8}$, with total error probability $O(n^{-3/16} \sum_{\beta} \mu_0(\beta))$. Since

$$\sum_{\beta} \mu_0(\beta) \sim \int_{\beta = n^{-3/8}}^{1} \frac{c}{\sqrt{\tilde{\alpha}(v)\beta}}\,\mathrm{d}\beta < \frac{2c}{\sqrt{\tilde{\alpha}(v)}} \leqslant 2cn^{1/8},$$

this total error probability is $O(n^{-1/16})$ and thus $o(n^{-2\varepsilon})$. As noted earlier, such an error probability per vertex considered is acceptable.

At this point we have coupled the neighbours j of v with $\tilde{\alpha}(j) > n^{-3/8}$ with the children y of x with $\tilde{\alpha}(y) > n^{-3/8}$ so that (with high enough probability) there is a bijection between them preserving $\tilde{\alpha}(.)$ values. To complete the proof we must consider neighbours with $\tilde{\alpha}(j) \leqslant n^{-3/8}$. But the expected number of such neighbours is at most

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

$\sum_{j \leqslant n^{5/8}} c/\sqrt{vj} \leqslant 2cn^{5/16}/\sqrt{v}$. As $\tilde{\alpha}(v) \geqslant n^{-1/4}$, so $v \geqslant n^{3/4} - O(n^{1/2})$, this expectation is at most $3cn^{-1/16} = o(n^{-2\varepsilon})$, so the probability that v has such a neighbour is $o(n^{-2\varepsilon})$. A similar argument applies to x, so we see that (with high enough probability) we have already coupled all neighbours of v in the next generation with all children of x.

Following the argument above for each $v \in \Gamma_k$ in turn, we see that we can construct the required coupling between Γ_{k+1} and X_{k+1} unless either some bad event with probability $o(|\Gamma_k| n^{-2\varepsilon})$ occurs, or both Γ_{k+1} and X_{k+1} have cardinality at least $n^{1/\varepsilon}$, in which case (6.1) holds. Since we always stop before considering more than n^{ε} vertices, we can construct the required coupling with probability $1 - o(n^{\varepsilon})$, completing the proof of the lemma. □

As a simple corollary, Lemma 6.1 gives good bounds on the number of vertices of G in small components. Let us write $\#_r(G)$ for the number of vertices of G in components of order r, and let us write s_r for the probability that the branching process (X_k) has total size $|(X_k)| = \sum_{k \geqslant 0} |X_k|$ equal to r. Note that s_r is independent of n, and that $\sum_{r \geqslant 0} s_r = 1 - \sigma$, where σ is the survival probability of the branching process, *i.e.*, $\sigma = \Pr(|(X_k)| = \infty)$.

Corollary 6.2. *Let $c > 0$ be fixed. There is a positive constant ε (independent of c) such that with probability $1 - o(n^{-1})$ we have*

$$|\#_r(G) - s_r n| \leqslant n^{1-2\varepsilon} \tag{6.2}$$

for all r in the range $0 \leqslant r \leqslant n^{\varepsilon}$.

Proof. We shall take $\varepsilon = 1/1000$, and use Lemma 6.1. Note that $\mathbf{E}(\#_r(G))$ is exactly n times the probability q_r that a random vertex of G lies in a component of order r. From Lemma 6.1, for $r \leqslant n^{\varepsilon}$ we have $|s_r - q_r| = o(n^{-10\varepsilon})$. (Note that we are using $\varepsilon = 1/1000$ here, compared to $1/100$ in the (statement and) proof of Lemma 6.1.) Hence,

$$|\mathbf{E}(\#_r(G)) - s_r n| \leqslant n^{1-3\varepsilon} \tag{6.3}$$

certainly holds for $0 \leqslant r \leqslant n^{\varepsilon}$.

It remains to show concentration of $\#_r(G)$ about its mean. We can do this using standard martingale methods. There is a slight annoyance due to the unbounded (and potentially very large) degrees in G. Perhaps the simplest way to deal with this is to avoid large degrees altogether, as follows. Let the *down-degree* of a vertex i be the number of $j < i$ to which i is adjacent. Note that the down-degree of i in G has expectation $\sum_{j<i} c/\sqrt{ij} \leqslant 2c$. From the independence of the edges, it is easy to check that the chance that a given vertex i has down-degree in G at least $(\log n)^2$, say, is $o(n^{-2})$. Thus with probability $1 - o(n^{-1})$ every vertex of G has down-degree at most $(\log n)^2$.

Let H be the graph obtained from G by removing all edges ij where $j < i$ and i has down-degree at least $(\log n)^2$ in G. Then by construction every vertex of H has down-degree at most $(\log n)^2$. Also, G and H coincide exactly with probability $1 - o(n^{-1})$. Hence $\#_r(G)$ and $\#_r(H)$ coincide with probability $1 - o(n^{-1})$, and $\mathbf{E}(\#_r(G)) = \mathbf{E}(\#_r(H)) + o(1)$. Thus it suffices to prove that $\#_r(H)$ is concentrated about its mean.

Fix $0 \leqslant r \leqslant n^{\varepsilon}$. The graph H can be constructed in n independent steps, at each step deciding the down-neighbourhood of a vertex i. Changing the random choice made at one

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

such step modifies the final graph H in at most $2(\log n)^2$ edges, and thus changes $\#_r(H)$ by, crudely, at most $M = 4r(\log n)^2 = o(n^{2\varepsilon})$. (Adding or deleting a single edge in any graph changes the number of vertices in components of order r by at most $2r$.) It follows from standard martingale inequalities (for example, the Hoeffding–Azuma inequality) that for any $t > 0$ the probability that $\#_r(H)$ deviates from its mean by more than t is at most $2\exp(-t^2/(2nM^2))$. Taking $t = n^{3/4}$ and recalling that $M = o(n^{2\varepsilon})$ this probability is exponentially small in some power of n, and certainly at most n^{-2}, say. Summing over r, we see that with probability $1 - o(n^{-1})$ we have $|\#_r(H) - \mathbf{E}(\#_r(H))| \leqslant n^{3/4}$ for all r, $0 \leqslant r \leqslant n^{\varepsilon}$. Combined with (6.3) and the very close agreement between $\#_r(H)$ and $\#_r(G)$ this proves the corollary. □

Having obtained good bounds on the number of vertices in small components, we would now like to study the order of the giant component. Let $\sigma = \sigma(c)$ be the survival probability of the branching process (X_k), *i.e.*, the probability that $|(X_k)| = \infty$.

Lemma 6.3. *Let $c > 0$ be fixed. Then*

$$C_1(G) = (\sigma(c) + o(1))n$$

holds w.h.p. as $n \to \infty$.

Proof. It only remains to show that w.h.p. almost all vertices not in small components are in the giant component. We could prove this as in [9, 8], by running the neighbourhood expansion argument further. However, there is a much easier way, namely adding vertices to the graph, as in [10]. This method will be applicable to the other models we consider as well.

Let $\varepsilon = 1/1000$ be the constant appearing in Corollary 6.2, and set $L = (\log n)^2/2$. (Here we could take $L = n^{\varepsilon}/2$, but later we shall need to work with smaller L.)

From Corollary 6.2, w.h.p. the number of vertices of G in *small* components, *i.e.*, components of order at most L, is $(\sum_{r \leqslant L} s_r + o(1))n$. Here we use the fact that (6.2) holds with probability $o(L^{-1})$ for each $r \leqslant L$. Now s_r is independent of n, and $\sum_{r \geqslant 0} s_r = 1 - \sigma(c)$. Thus $\sum_{r \leqslant L} s_r = 1 - \sigma(c) + o(1)$, using only $L \to \infty$. Hence, w.h.p. G has $(\sigma(c) + o(1))n$ vertices in *large* components, *i.e.*, components of order greater than L. We shall assume that $\sigma(c) > 0$, as otherwise there is nothing left to prove.

Let $\eta = \eta(n) = (\log n)^{-1/2}$. Let G_t be the subgraph of G induced by the first t vertices, where t runs from $n' = \lfloor (1-\eta)n \rfloor$ to n. Note that $G' = G_{n'}$ has exactly the distribution of $G^{(n')}_{1/\sqrt{ij}}(c)$. Hence, using Corollary 6.2 as above, noting that $L \leqslant (n')^{\varepsilon}$, the number of vertices of G' in large components is w.h.p. $(\sigma(c) + o(1))n' = (\sigma(c) + o(1))n$. Let $C_1, \ldots, C_r$ be the large components of G'. It suffices to show that in G, w.h.p. all the C_i lie within a single component; this large component then contains $(\sigma(c) + o(1))n$ vertices, as required.

We shall condition on G', assuming, as we may, that G' contains at least $\sigma(c)/2$ vertices in large components, *i.e.*, in $\bigcup C_r$. We shall use only one property of the process $(G' = G_{n'}, G_{n'+1}, \ldots, G_{n-1}, G_n = G)$: we claim that there is a $\gamma > 0$ (depending on c but not on n) such that, conditioning on G_{t-1} and fixing two disjoint sets A, B of vertices of G_{t-1}, the probability that the new vertex t of G_t is adjacent to some $a \in A$ and some $b \in B$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

is at least $\gamma|A||B|/n^2$. For the model we consider here this is very easy to see: the edges between vertex t and G_{t-1} are independent of G_{t-1}, and each $a \in A$ has independently probability $c/\sqrt{at} \geqslant c/n$ of being adjacent to t. It follows that t sends an edge to A with probability at least $1-(1-c/n)^{|A|} = \Theta(|A|/n)$, using $c = \Theta(1)$ and $|A| \leqslant n$. Since t sends an edge independently to B with probability $\Theta(|B|/n)$ this proves the claim.

It remains to prove that w.h.p. the C_i all lie in a single component of G. Suppose not. Then there is a partition of $[r] = \{1,2,\dots,r\}$ into (D,D') such that there is no path from $C_D = \bigcup_{i\in D} C_i$ to $C_{D'}$ in G. In particular, no vertex t, $n' < t \leqslant n$, sends edges to both C_D and $C_{D'}$. Let us call such a partition *bad*.

Fix a partition D, D'. Without loss of generality we may take $d = |D| \leqslant |D'|$. Note that as $|D|, |D'| \geqslant d$ and the C_i are all large we have $|C_D|, |C_{D'}| \geqslant dL$. Also, $C_D \cup C_{D'}$ contains all $\geqslant \sigma(c)n/2$ vertices of G' in large components, so at least one of C_D and $C_{D'}$ has cardinality at least $\sigma(c)n/4$. It follows that $|C_D||C_{D'}| \geqslant dL\sigma(c)n/4$. Hence, for each t, $n' < t \leqslant n$, the probability that the new vertex t joins to both C_D and $C_{D'}$ is at least $\gamma dL\sigma(c)/(4n)$. But, as this lower bound on the probability holds for each t conditional on G_{t-1}, the probability that no t, $n' < t \leqslant n$, joins to both C_D and $C_{D'}$ is at most

$$(1-\gamma dL\sigma(c)/(4n))^{n-n'} \leqslant \exp(-\gamma dL\sigma(c)(n-n')/(4n)) \leqslant \exp(-\gamma L\sigma(c)\eta/4)^d.$$

But the number of partitions with $|D| = d$ is at most $\binom{r}{d} \leqslant r^d$. Hence the probability that there is some bad partition is at most

$$\sum_{d=1}^{r/2} \big(r\exp(-\gamma L\sigma(c)\eta/4)\big)^d.$$

Now $L = \Theta((\log n)^2)$, while γ and $\sigma(c)$ are constant, and $\eta = (\log n)^{-1/2}$. Hence,

$$r\exp(-\gamma L\sigma(c)\eta/4) = r\exp(-\Theta((\log n)^{3/2})) = o(rn^{-100}) = o(1),$$

and the sum above is $o(1)$, completing the proof of the lemma. □

The argument above is similar to that used in [10] for a different model. We have written the argument in a cleaner and more general form so that we can apply it to all the models we consider.

6.2. Analysis of the branching process

We have now found the size of the giant component of the $c/\sqrt{ij}$-graph G in terms of $\sigma = \sigma(c)$, the survival probability of a certain branching process. It remains to calculate this survival probability. Of all the branching processes we consider, this is the only one where this can be done exactly (in terms of the inverse of a standard function). In this case we shall not need to use the two facts given at the end of Section 5.

Let us recall that we are considering the continuous type branching process $(X_k) = (X_0, X_1, \dots)$ where each $x \in X_k$ has a type $\alpha(x) \in (0,1]$. The initial condition is that $X_0 = \{x_0\}$ where $\alpha(x_0)$ is uniformly distributed on $(0,1]$. Each $x \in X_k$ gives rise independently to children $y \in X_{k+1}$ according to a Poisson process, so that the values $\alpha(y)$ associated to these children form a Poisson process on $(0,1]$ with density $c/\sqrt{\alpha(x)\alpha(y)}$.

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

As in Section 5, we say that an $x \in X_k$ survives if x has descendants in all later generations. Note that x_0 survives if and only if at least one of its offspring does. Let $S_\infty(\alpha) = S_\infty(\alpha, c)$ be the probability that x_0 survives given that $\alpha(x_0) = \alpha$. Then, given that $\alpha(x_0) = \alpha$, the expected number of surviving children of x_0 is exactly

$$\mu = \int_{\beta=0}^{1} \frac{c}{\sqrt{\alpha\beta}} S_\infty(\beta)\, \mathrm{d}\beta,$$

since each child y with $\alpha(y) = \beta$ survives with probability $S_\infty(\beta)$. Now, from the Poisson nature of the process, given $\alpha(x_0)$, the number of surviving children of x_0 has a Poisson distribution. Since this number has expectation μ, the conditional probability that x_0 survives is thus $1 - \exp(-\mu)$. In other words, for $0 < \alpha \leqslant 1$ we have

$$S_\infty(\alpha) = 1 - \exp\left(- \int_{\beta=0}^{1} \frac{c S_\infty(\beta)}{\sqrt{\alpha\beta}}\, \mathrm{d}\beta \right).$$

This equation is exactly the equation $S_\infty = \mathbf{F}(S_\infty)$ of Section 5, and we could have derived it using the Poisson density (5.1) with $g(\alpha, \beta) = c/\sqrt{\alpha\beta}$. However, the direct argument given above is probably clearer.

Let us write $A = A(c)$ for $c \int_{\beta=0}^{1} S_\infty(\beta)/\sqrt{\beta}\, \mathrm{d}\beta$. Note that as $S_\infty \leqslant 1$ we have $A \leqslant 2c$. Now

$$S_\infty(\alpha) = 1 - \exp(-A/\sqrt{\alpha}),$$

and it only remains to determine A. However, by definition of A we have

$$A = c \int_{\beta=0}^{1} \frac{S_\infty(\beta)}{\sqrt{\beta}}\, \mathrm{d}\beta = c \int_{\beta=0}^{1} \frac{1}{\sqrt{\beta}} \left(1 - e^{-A/\sqrt{\beta}}\right) \mathrm{d}\beta.$$

Let us write $g(x)$ for the function

$$g(x) = \int_{\beta=0}^{1} \frac{1}{\sqrt{\beta}} \left(1 - e^{-x/\sqrt{\beta}}\right) \mathrm{d}\beta.$$

Then it is easy to see that $g(x) \sim 2x \log(1/x)$ as $x \to 0$. Now the relationship between A and c is exactly that $A = cg(A)$, *i.e.*, that $g(A)/A = 1/c$. As $c \to 0$ we certainly have $A \to 0$ (recall that $A \leqslant 2c$). Thus it follows that $2\log(1/A) \sim 1/c$, *i.e.*, that

$$A = \exp\left(- \frac{1 + o(1)}{2c} \right).$$

Actually, we can obtain much more precise results, by writing $g(x)$ in terms of the *exponential integral*

$$E_1(x) = \int_{t=1}^{\infty} \exp(-xt)/t\, \mathrm{d}t.$$

Here we use the notation of [1, equation 5.1.4, p. 228]. Substituting $\beta = t^{-2}$ and integrating by parts shows that

$$g(x) = 2 - 2e^{-x} + 2xE_1(x).$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

Now from [1, equation 5.1.11], we have

$$E_1(x) = -\gamma - \log(x) - \sum_{k=1}^{\infty} (-1)^k \frac{x^k}{k.k!}$$

where $\gamma = 0.57721566\ldots$ is Euler's constant. It follows that

$$g(x)/x = 2\log(1/x) + 2(1-\gamma) + x - x^2/6 + x^3/36 + O(x^4).$$

Finally, using $g(A)/A = 1/c$ we see that

$$A(c) = e^{1-\gamma} \exp\left(-\frac{1}{2c}\right)\left(1 + \Theta\left(e^{-1/(2c)}\right)\right)$$

as $c \to 0$, where we could continue the expansion in terms of $\exp(-1/(2c))$ if we liked.

It remains to calculate $\sigma(c)$. By the definition of the process this is given by

$$\int_{\alpha=0}^{1} S_\infty(\alpha)\,\mathrm{d}\alpha = \int_{\alpha=0}^{1} (1 - \exp(-A/\sqrt{\alpha}))\,\mathrm{d}\alpha.$$

Another reduction to $E_1(x)$ shows that this is $2A + O(A^2\log(1/A))$. (We could obtain further terms in the series, but there seems little point.) Putting this all together we see that as $c \to 0$,

$$\sigma(c) = 2e^{1-\gamma} \exp\left(-\frac{1}{2c}\right)\left(1 + O\left(c^{-1}e^{-1/(2c)}\right)\right).$$

Combined with Lemma 6.3, this gives the size of the giant component in the $c/\sqrt{ij}$-graph, completing the proof of Theorem 2.2.

7. The LCD model

In this section we turn to one of the most basic and most studied scale-free graph models, the 'growth with preferential attachment' model introduced in vague terms by Barabási and Albert in [4] and made precise as the LCD model $G_m^{(n)}$ in [9]. A precise definition of the model was given in Section 2. This model is the hardest to analyse of all those considered here, owing to the very important effect of dependence between edges. After all, the preferential attachment rule is what gives rise to the unbalanced power-law degree sequence, and this rule is based exactly on dependence between edges. Fortunately, most of the work needed has already been done in [8]; a nontrivial amount of calculation remains, however.

Our aim now is to prove Theorem 2.1. For convenience, we recall the statement: let $m \geqslant 2$ and $0 < p < 1$ be fixed, and let $G_m^{(n)}(p)$ be the graph formed from $G_m^{(n)}$ by keeping vertices independently with probability p. Then

$$C_1\left(G_m^{(n)}(p)\right) = (\lambda_m(p) + o(1))n$$

holds w.h.p. as $n \to \infty$ for some constant $\lambda_m(p) > 0$. Furthermore, as $p \to 0$ with m fixed, $\lambda_m(p)$ satisfies the bounds (2.1).

At first sight it is not at all clear how to compare the neighbourhood expansion process in $G_m^{(n)}(p)$ with a branching process. In the case of the $c/\sqrt{ij}$-graph, for example,

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

the independence between edges meant that in the neighbourhood expansion process different vertices $v \in \Gamma_k$ at distance exactly k from the random starting vertex v_0 give rise to neighbours at distance $k+1$ essentially independently of each other. (The only dependence is due to the initially small chance that some w can be reached from v_0 by more than one path of length $k+1$.) Hence it is clear that the neighbourhood expansion process should behave like a branching process, at least initially. Here, however, the presence of an edge ij, $i<j$, makes it likely that i has relatively large degree (as if it has large degree, the edge is more likely), and hence makes it more likely that any given other possible edge ik, $k>i$, is present. (The correlation factor between a pair of edges of this form turns out to be approximately $(m+1)/m$; see [7].) It turns out, however, that partial independence can be re-introduced by introducing auxiliary variables associated to each vertex. This is shown in [9], and relies heavily on the static alternative description of $G_m^{(n)}$. In conclusion, after much work it is shown in [8], using the methods of [9], that the neighbourhood expansion process in $G_m^{(n)}(p)$ can be coupled with a certain branching process, although in a rather indirect way.

Let $0<p<1$ and $0 \leqslant \rho < 1$ be fixed; we shall consider a branching process corresponding to the neighbourhood expansion process in the subgraph of $G_m^{(n)}(p)$ induced by the vertices $\{\lceil \rho n \rceil, \dots, n-1, n\}$. The coupling result we shall quote from [8] applies without a cutoff, *i.e.*, with $\rho = 0$; this is all we need to obtain the size of the giant component. However, in working with the $\rho = 0$ branching process, it will be convenient to also consider a $\rho > 0$ branching process.

Let $L(\alpha)$, $R(\alpha) : [\rho, 1] \to [0, \infty)$ be the maximal solution to

$$\begin{aligned} L(\alpha) &= \frac{p}{2\sqrt{\alpha}} \int_{\beta=\rho}^{\alpha} \frac{1}{\sqrt{\beta}} \left(1 - \frac{(1-L(\beta))^m}{(1+R(\beta))^{m+1}}\right) \mathrm{d}\beta, \\ R(\alpha) &= \frac{p}{2\sqrt{\alpha}} \int_{\beta=\alpha}^{1} \frac{1}{\sqrt{\beta}} \left(1 - \frac{(1-L(\beta))^{m-1}}{(1+R(\beta))^{m}}\right) \mathrm{d}\beta, \end{aligned} \tag{7.1}$$

and let

$$\sigma(p,\rho) = p \int_{\alpha=\rho}^{1} \left(1 - \frac{(1-L(\alpha))^m}{(1+R(\alpha))^m}\right) \mathrm{d}\alpha. \tag{7.2}$$

When $\rho = 0$, $R(0)$ is not defined, and $R(\alpha) \to \infty$ as $\alpha \to 0$. In this case we work with functions L, R on $(0,1]$. Theorem 3.1 of [8] states that in the graph $G_m^{(n)}(p)$ obtained from the LCD graph $G_m^{(n)}$ by keeping vertices independently with probability p, there is w.h.p. a unique giant component of size $(\lambda_m(p) + o(1))n$. Examining the proof in [8], the value of $\lambda_m(p)$ is given by $\sigma(p,0)$. For edge deletion rather than vertex deletion the proof and result are the same, except that the initial factor of p in (7.2) is absent. Upper and lower bounds on $\lambda_m(p) = \sigma(p,0)$ are given in [8]; here we shall use the exact solution of the linearized form of (7.1), with $\rho > 0$, to obtain much better bounds on $\sigma(p,0)$, matching up to a factor $1 + o(1)$ in the exponent.

The description of the branching process, proof of the required coupling, and derivation of the survival probability equations (7.1) are all given in [8], so there is no need to repeat them here. However, we would like to use the two facts proved at the end of Section 5. We shall write (7.1) in the form $(L,R) = \mathbf{F}_\rho((L,R))$, where $\mathbf{F}_\rho$ is the operator defined by

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

the right-hand side of (7.1). Unfortunately, $\mathbf{F}_\rho$ does not have quite the right form for the operator $\mathbf{F}$ of Section 5. Indeed, L and R are not exactly the survival probabilities of particular vertices. However, L and R can be viewed as transformed survival probabilities in a certain modified branching process that survives if and only if the (branching process associated naturally to the) neighbourhood expansion process does. This will involve a distinction between vertices *reached from the left*, *i.e.*, $v \in \Gamma_k$ adjacent to $w \in \Gamma_{k-1}$ with $w < v$, or *reached from the right* ($w > v$). This distinction occurs again in Section 8; in $G_m^{(n)}$ there there are additional complications.

The transformation described above is simple, but cannot be written concisely without reference to the details of the branching process in [8]. In outline, we consider a triple of survival probabilities $(S_1(\alpha), S_2(\alpha), S_3(\alpha))$; these will represent respectively the probabilities of: (1) the survival of the vertex at the end of an outgoing leftwards edge from a vertex x with $\alpha(x) = \alpha$, and (2)/(3) the survival of *some* vertex to the right of x joined directly to x, in the cases that x itself was reached from the left (S_2) and x reached from the right (S_3). There is an operator $\mathbf{F}'$ of the form in Section 5 such that $(S_1, S_2, S_3) = \mathbf{F}'((S_1, S_2, S_3))$. It is shown in [8] that, from the form of the process, if $(\psi_1, \psi_2, \psi_3) = \mathbf{F}'((\phi_1, \phi_2, \phi_3))$ then $\log((1-\psi_3)/(1-\psi_2)) = (m+1)/m$. In the arguments of Section 5 we need only work with functions of the form $\mathbf{F}'(.)$, so we can consider instead the corresponding $\mathbf{F}''((S_1, S_2))$. The monotone transformation $S_1 = L$, $S_2 = 1-(1+R)^{-m}$ transforms $\mathbf{F}''$ into $\mathbf{F}_\rho$.

The key point is that, applying the results of Section 5 to $\mathbf{F}''$, it follows that trial functions (L', R') mapping under $\mathbf{F}_\rho$ to functions above/below themselves give lower/upper bounds on the maximal solution (L, R) to (7.1).

Proof of Theorem 2.1. Consider the following equations, obtained by linearizing (7.1):

$$\begin{aligned} \tilde{L}(\alpha) &= \frac{p'}{2\sqrt{\alpha}} \int_{\beta=\rho}^{\alpha} \frac{1}{\sqrt{\beta}} \big(m\tilde{L}(\beta) + (m+1)\tilde{R}(\beta)\big)\, \mathrm{d}\beta, \\ \tilde{R}(\alpha) &= \frac{p'}{2\sqrt{\alpha}} \int_{\beta=\alpha}^{1} \frac{1}{\sqrt{\beta}} \big((m-1)\tilde{L}(\beta) + m\tilde{R}(\beta)\big)\, \mathrm{d}\beta. \end{aligned} \tag{7.3}$$

We have replaced p by p' since, given p, we shall solve the linearized equations for a slightly different value p' to obtain bounds on the solutions to the non-linear equations. The equations (7.3) can be solved by multiplying through by $\sqrt{\alpha}$ and differentiating. We claim that the functions

$$\begin{aligned} \tilde{L}(\alpha) &= \alpha^{-1/2}\left(\frac{m+1}{m-1}\alpha^{p'/2} - \alpha^{-p'/2}\right) \\ \tilde{R}(\alpha) &= \alpha^{-1/2}\big(\alpha^{-p'/2} - \alpha^{p'/2}\big) \end{aligned}$$

solve (7.3), provided

$$\rho = \left(\frac{m-1}{m+1}\right)^{1/p'}. \tag{7.4}$$

This is easy to verify by substituting back. We shall take (7.4) as the definition of ρ from now on. Note that $\tilde{L}$ and $\tilde{R}$ are nonnegative on $[\rho, 1]$, and are pointwise bounded above by $\frac{10}{m}\alpha^{-1/2}$, say. Indeed, $\alpha^{1/2}\tilde{L}(\alpha)$ is increasing, and hence bounded by $\tilde{L}(1) = 2/(m-1)$,

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

while $\alpha^{1/2}\tilde{R}(\alpha)$ is decreasing, and hence bounded by $\rho^{1/2}\tilde{R}(\rho) = \sqrt{\frac{m+1}{m-1}} - \sqrt{\frac{m-1}{m+1}}$. Note also that for m fixed and p' not too large,

$$\int_{\alpha=\rho}^{1} \tilde{L}(\alpha)\,\mathrm{d}\alpha = \Omega(1), \tag{7.5}$$

due to the contribution from α between $1/2$ and 1, say.

Throughout we shall consider a function $\varepsilon = \varepsilon(p)$ tending to zero at some suitable rate as $p \to 0$. In the end it turns out that we shall take $\varepsilon = p$, but the roles played by ε and p are rather different, so we preserve the separate notation.

To obtain a lower bound on $\lambda_m(p) = \sigma(p, 0)$, set $p' = p(1-\varepsilon)$. Set

$$L_0(\alpha) = \varepsilon\rho^{1/2}\tilde{L}(\alpha)/100, \qquad R_0(\alpha) = \varepsilon\rho^{1/2}\tilde{R}(\alpha)/100.$$

Then L_0 and R_0 solve (7.3) and are bounded above by $\varepsilon/(10m)$. Let us write (7.1) in the form $(L, R) = \mathbf{F}_\rho((L, R))$, where $\mathbf{F}_\rho$ is an operator taking a pair of functions on $[\rho, 1]$ to another such pair. When ε is small enough, since L_0, R_0 are bounded by $\varepsilon/(10m)$, when evaluating $\mathbf{F}_\rho((L_0, R_0))$ the linear approximation of the integrand is correct to within a factor $1 \pm \varepsilon/2$, say. Hence, since (L_0, R_0) solves (7.3) and $p = p'/(1-\varepsilon)$, for p small enough the two functions given by $\mathbf{F}_\rho((L_0, R_0))$ are pointwise at least L_0 and R_0. It follows from the first fact proved at the end of Section 5, or simply by iterating $\mathbf{F}_\rho$, starting from (L_0, R_0), to obtain a solution to $(L, R) = \mathbf{F}_\rho((L, R))$, that L_0, R_0 are pointwise lower bounds for the maximal solution L, R to $(L, R) = \mathbf{F}_\rho((L, R))$. Hence,

$$\lambda_m(p) = \sigma(p, 0) \geqslant \sigma(p, \rho) \geqslant p\int_{\alpha=\rho}^{1} \left(1 - \frac{(1 - L_0(\alpha))^m}{(1 + R_0(\alpha))^m}\right)\mathrm{d}\alpha.$$

Again, the linearized form of the above integral is a good approximation. Using (7.5) we see that this integral is $\Omega(\varepsilon\rho^{1/2})$, so $\lambda_m(p) = \Omega(p\varepsilon\rho^{1/2}) = \Omega(p^2\rho^{1/2})$, recalling that $\varepsilon = p$. Now, as $p' = p(1-\varepsilon) = p(1-p)$,

$$\rho^{1/2} = \exp\left(\frac{\log((m-1)/(m+1))}{2p}(1-p)^{-1}\right) = \Theta\left(\left(\frac{m-1}{m+1}\right)^{\frac{1}{2p}}\right)$$

as $p \to 0$, so our bound $\lambda_m(p) = \Omega(p^2\rho^{1/2})$ gives exactly the lower bound claimed in (2.1).

For the upper bound, we consider $p' = p(1+\varepsilon)$, defining $\tilde{L}$, $\tilde{R}$ and ρ as above. Again we shall take $\varepsilon = p$. Set $\tau = \rho(1+\varepsilon)^{2/p'}$, noting that $\tau = \Theta(\rho)$. Recall from (7.4) that $\rho^{p'/2} = \sqrt{\frac{m-1}{m+1}}$, so

$$\tau^{p'/2} = (1+\varepsilon)\sqrt{\frac{m-1}{m+1}},$$

and

$$\left(\frac{m+1}{m-1}\tau^{p'/2} - \tau^{-p'/2}\right) = \left((1+\varepsilon) - \frac{1}{1+\varepsilon}\right)\sqrt{\frac{m+1}{m-1}} = \Theta(\varepsilon).$$

Hence, for $\alpha \geqslant \tau$ we have

$$\alpha^{1/2}\tilde{L}(\alpha) \geqslant \tau^{1/2}\tilde{L}(\tau) = \Omega(\varepsilon). \tag{7.6}$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

Set

$$L_1(\alpha) = \begin{cases} f\tilde{L}(\alpha) & \alpha \geqslant \tau \\ p & \alpha < \tau, \end{cases}$$

$$R_1(\alpha) = \begin{cases} f\tilde{R}(\alpha) & \alpha \geqslant \tau \\ p/\sqrt{\alpha} & \alpha < \tau, \end{cases}$$

where $f > 0$ is a scale-factor to be chosen later, and define L_2, R_2 by $(L_2, R_2) = \mathbf{F}_0((L_1, R_1))$. Thus,

$$L_2(\alpha) = \frac{p}{2\sqrt{\alpha}} \int_{\beta=0}^{\alpha} \frac{1}{\sqrt{\beta}} \left(1 - \frac{(1 - L_1(\beta))^m}{(1 + R_1(\beta))^{m+1}}\right) \mathrm{d}\beta, \tag{7.7}$$

and

$$R_2(\alpha) = \frac{p}{2\sqrt{\alpha}} \int_{\beta=\alpha}^{1} \frac{1}{\sqrt{\beta}} \left(1 - \frac{(1 - L_1(\beta))^{m-1}}{(1 + R_1(\beta))^{m}}\right) \mathrm{d}\beta.$$

We claim that $L_2(\alpha) \leqslant L_1(\alpha)$ and $R_2(\alpha) \leqslant R_1(\alpha)$ hold for every α, provided p is small enough.

For $\alpha < \tau$ the claim is trivial: the large brackets inside the integrals above are each at most 1, so

$$L_2(\alpha) \leqslant \frac{p}{2\sqrt{\alpha}} \int_{\beta=0}^{\alpha} \frac{1}{\sqrt{\beta}} \, \mathrm{d}\beta = p = L_1(\alpha),$$

and

$$R_2(\alpha) \leqslant \frac{p}{2\sqrt{\alpha}} \int_{\beta=\alpha}^{1} \frac{1}{\sqrt{\beta}} \, \mathrm{d}\beta = p\frac{2 - 2\sqrt{\alpha}}{2\sqrt{\alpha}} \leqslant \frac{p}{\sqrt{\alpha}} = R_1(\alpha).$$

Suppose now that $\alpha \geqslant \tau$. The integral defining $R_2(\alpha)$ only involves values $L_1(\beta)$, $R_1(\beta)$ for which $\beta \geqslant \alpha \geqslant \tau$. In this range $L_1 = f\tilde{L}$ and $R_1 = f\tilde{R}$. Now the linearized form of the right-hand side of (7.1) is always an upper bound on the original form, as $1 - (1 - x)^i(1 + y)^{-j} < ix + jy$ for positive real numbers x, y and for $i, j \geqslant 1$. Hence,

$$R_2(\alpha) \leqslant \frac{p}{2\sqrt{\alpha}} \int_{\beta=\alpha}^{1} \frac{1}{\sqrt{\beta}} ((m-1)f\tilde{L}(\beta) + mf\tilde{R}(\beta)) \, \mathrm{d}\beta.$$

Since $\tilde{L}$ and $\tilde{R}$ solve (7.3), it follows that

$$R_2(\alpha) \leqslant (p/p')f\tilde{R}(\alpha) \leqslant f\tilde{R}(\alpha) = R_1(\alpha).$$

Finally, we must estimate $L_2(\alpha)$ for $\alpha \geqslant \tau$. We shall split the integral on the right-hand side of (7.7) into two parts, bounding the contribution from $\beta < \tau$ using only that the large bracket involved is at most 1, and the contribution from $\beta \geqslant \tau$ as above:

$$L_2(\alpha) = \frac{p}{2\sqrt{\alpha}} \int_{\beta=0}^{\tau} \frac{1}{\sqrt{\beta}} \left(1 - \frac{(1 - L_1(\beta))^m}{(1 + R_1(\beta))^{m+1}}\right) \mathrm{d}\beta$$
$$+ \frac{p}{2\sqrt{\alpha}} \int_{\beta=\tau}^{\alpha} \frac{1}{\sqrt{\beta}} \left(1 - \frac{(1 - L_1(\beta))^m}{(1 + R_1(\beta))^{m+1}}\right) \mathrm{d}\beta$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

$$\begin{aligned}
&\leqslant \frac{p}{2\sqrt{\alpha}} 2\sqrt{\tau} + \frac{p}{2\sqrt{\alpha}} \int_{\beta=\tau}^{\alpha} \frac{1}{\sqrt{\beta}} (mL_1(\beta) + (m+1)R_1(\beta))\,\mathrm{d}\beta \\
&\leqslant p\sqrt{\frac{\tau}{\alpha}} + \frac{p}{p'} f\tilde{L}(\alpha) \\
&\leqslant f\tilde{L}(\alpha) + p\sqrt{\frac{\tau}{\alpha}} - \frac{\varepsilon f}{2}\tilde{L}(\alpha),
\end{aligned}$$

using $p' = p(1+\varepsilon)$. The final bound will be at most $f\tilde{L}(\alpha) = L_1(\alpha)$ as long as we choose f large enough so that $2p\sqrt{\tau} < \varepsilon f\tilde{L}(\alpha)\sqrt{\alpha}$ holds for all $\alpha > \tau$. Using (7.6) and recalling that $\varepsilon = p$, we see that it suffices to take $f = A\sqrt{\tau}/p$ for some constant A.

For this value of f we have shown that our trial functions L_1, R_1 iterate under $\mathbf{F}_0$ to functions below themselves. Using the second fact proved at the end of Section 5, which applies by the remarks made before the start of this proof, it follows that L and R, the maximal solutions to $(L, R) = \mathbf{F}_0((L, R))$, are bounded above by L_1 and R_1. Thus from (7.2) we have

$$\begin{aligned}
\lambda_m(p) = \sigma(p, 0) &\leqslant p\int_{\alpha=0}^{1} \left(1 - \frac{(1 - L_1(\alpha))^m}{(1 + R_1(\alpha))^m}\right)\mathrm{d}\alpha \\
&\leqslant p\tau + p\int_{\alpha=\tau}^{1} (mL_1(\alpha) + mR_1(\alpha))\,\mathrm{d}\alpha.
\end{aligned}$$

Here the last inequality is obtained by splitting the integral at τ, using 1 as an upper bound on the integrand for $\alpha < \tau$, and the linear form as an upper bound for $\alpha > \tau$. As $\tilde{L}$ and $\tilde{R}$ are bounded by $C\alpha^{-1/2}$ for some universal constant, the final integral is bounded by $2mCf$ times $\int_0^1 \alpha^{-1/2}\,\mathrm{d}\alpha = 2$, so

$$\lambda_m(p) \leqslant p\tau + O(pf) = p\tau + O(\sqrt{\tau}).$$

The second term dominates, and, as before, we have

$$\sqrt{\tau} = \Theta(\sqrt{\rho}) = \Theta\left(\left(\frac{m-1}{m+1}\right)^{\frac{1}{2p}}\right),$$

giving the upper bound stated in (2.1) and completing the proof of Theorem 2.1. □

The proof above consists of one way of bounding the solution to certain rather unpleasant equations (7.1). It may be that there is a better way of bounding the solution, eliminating the gap of a factor p^2 between the upper and lower bounds. Although in one sense rather large, in another sense this gap is rather small: not only have we shown that $\lambda_m(p) = \exp(-\Theta(1/p))$, improving the lower bound of $\exp(-\Theta(1/p^2))$ given in [8], but we have found the constant in the exponent to within $o(1)$.

The calculation used in the proof is in some sense a version of the direct argument given for the $c/\sqrt{ij}$ model in Section 4. Indeed, our method of obtaining an upper bound on the size of the giant component involved removing a small number of 'early' vertices (with $i \leqslant \rho n$), so that the graph on the remaining 'late' vertices was subcritical. The final bound on $L_2(\alpha)$ above is really a way of estimating the total number of paths from late to early vertices, using the fact that the branching process restricted to late vertices is

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

subcritical, and that there are very few early vertices. This is analogous to the upper bound on the giant component of the $c/\sqrt{ij}$ graph proved by directly estimating the number of late-early paths in Section 4. For the lower bound, we again removed some early vertices, even though this decreases the size of the giant component. The gain was that the resulting branching process, which is still supercritical, has an eigenfunction that is not too unbalanced (does not become too large as α becomes small). This means that a not too small multiple of the eigenfunction gives a lower bound on the survival probability. Again this method is analogous to, though much more precise than, the method used in Section 4, or in [6] for a different model.

8. The uniformly grown Z-out random graph

We now turn to graphs that grow by adding one vertex at a time, each new vertex sending edges to some number of old vertices that are chosen uniformly at random, considering a common generalization of the uniformly grown m-out random graph and the c/j-graph, as in [10]. The argument for the c/j-graph would be considerably simpler than for this general case, but it turns out that the actual calculations come out more or less the same in the general case.

8.1. Precise definitions and conditions

In Section 2 we introduced the uniformly grown Z-out graph in a rather vague way, leaving precise details and the conditions we shall need for our theorem to this section. The definition will be essentially as in [10] (where X was used instead of Z – here that would clash with our branching process notation), except that we allow rather greater generality.

Let Z be a nonnegative integer-valued random variable; the main examples to bear in mind are the cases where Z is constant, and where Z has a Poisson distribution. Let Z_i, $i = 1, 2, \ldots$, be independent random variables with Z_i taking values in $\{0, 1, \ldots, i-1\}$, and Z_i converging to Z in distribution as $i \to \infty$. (We shall make additional assumptions later.) We say that the random graph process $(G_Z^{(n)})_{n \geqslant 0}$ is a *uniformly growing Z-out process* if $G_Z^{(0)}$ is the empty graph with no vertices, and $G_Z^{(n)}$ is obtained from $G_Z^{(n-1)}$ by adding a new vertex n and joining it to a uniformly chosen set of Z_n vertices from $\{1, 2, \ldots, n-1\}$, this choice being independent of $G_Z^{(n-1)}$. We say that the graph $G_Z^{(n)}$ defined in this way is a *uniformly grown Z-out graph*. Note that two uniformly grown Z-out graphs need not have exactly the same distribution, as the only condition we have so far imposed on the Z_i is that they converge to Z.

Given a $G_Z^{(n)}$ as above, we wish to study the size of the largest component of $G_Z^{(n)}(p)$, the random subgraph of $G_Z^{(n)}$ obtained by retaining each edge independently with probability p. (We could instead keep vertices independently, but as usual it makes no difference in the end; the argument for the vertex-deleted subgraph is notationally more complex but the final bounds differ by exactly a factor of p.) Here we may take $p = 1$, as $G_Z^{(n)}(p)$ has the same distribution as $G_{Z(p)}^{(n)}$, where $Z(p)$ is the random variable defined by the number of successes in Z trials, each of which succeeds with probability p, independently of Z and of all other trials. More precisely, we must replace each Z_i with a corresponding $Z_i(p)$.

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

Throughout we assume that $\Pr(Z \geqslant 2) > 0$, and that the first two moments of Z are finite. We write c for $\mathbf{E}(Z)$ for consistency with earlier work. We also write $r = r(Z)$ for $\mathbf{E}(Z(Z-1))/\mathbf{E}(Z)^2$, noting that $r > 0$, and that $r = 1$ if Z has a Poisson distribution. If $Z(p)$ is defined from Z as above, then $\mathbf{E}(Z(p)) = p\mathbf{E}(Z)$ and $r(Z(p)) = r(Z)$. Indeed, the kth factorial moment of $Z(p)$ is p^k times the corresponding moment of Z; compare the expected number of k-element subsets of a set of size Z and one of size $Z(p)$.

We impose the following additional conditions on Z and the Z_i, which may or may not be needed, but are used in our proof. First, we assume that $\mathbf{E}(Z^3)$ is finite. In fact, as we shall consider asymptotics as the distribution of Z varies, we assume that $\mathbf{E}(Z^3) = O(1)$, *i.e.*, that as we vary Z, the third moment is bounded by some absolute constant. This is only used at one point in the proof; see (8.13). Secondly, we assume that

$$d_{TV}(Z_i, Z) = O((\log i)^{-8}) \tag{8.1}$$

as $i \to \infty$. Here d_{TV} is the total variation distance, namely half the sum of the absolute value of the difference of the probabilities of taking certain values. In (8.1), an exponent of $-2+\varepsilon$ would perhaps suffice if we were more careful in the proof, but this would lead to additional complications. Finally, we assume that $\mathbf{E}(Z_i^2)$ is bounded above by some constant independent of i.

Theorem 8.1. *Let Z and Z_i, $i = 1, 2, \ldots$, be random variables satisfying the conditions above, and set $c = \mathbf{E}(Z)$ and*

$$r = r(Z) = \frac{\mathbf{E}(Z(Z-1))}{\mathbf{E}(Z)^2}.$$

Let $G_Z^{(n)}$ be the uniformly grown Z-out graph defined by the Z_i. Then there is a constant $\lambda(Z)$ such that

$$C_1\big(G_Z^{(n)}\big) = (\lambda(Z) + o(1))n$$

holds w.h.p. as $n \to \infty$. Furthermore, writing

$$c_0 = \frac{1}{2}\frac{1-\sqrt{r}}{1-r}$$

if $r \neq 1$ and $c_0 = 1/4$ if $r = 1$, we have $\lambda(Z) = 0$ if $c \leqslant c_0$, and

$$\lambda(Z) = \Omega\left(\varepsilon^{3/2}\exp\left(-\frac{\pi}{2r^{1/4}}\varepsilon^{-1/2}\right)\right) \tag{8.2}$$

and

$$\lambda(Z) = O\left(\varepsilon^{-3/2}\exp\left(-\frac{\pi}{2r^{1/4}}\varepsilon^{-1/2}\right)\right) \tag{8.3}$$

whenever $c = c_0 + \varepsilon$ with $\varepsilon \to 0$ from above and $r(Z) = r$ is fixed.

Thus, while the exact normalized size of the giant component does depend on the whole distribution Z, the existence of such a giant component, and its asymptotic size when small, are determined by the first two moments of Z only.

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

Taking Z to be Poisson with mean c, and Z_i to be Binomial with parameters $i-1$ and $c/(i-1)$, Theorem 8.1 immediately implies Theorem 2.4.

For Theorem 2.3, note that all our conditions on Z and the Z_i are preserved if we pass to $Z(p)$ and $Z_i(p)$ for $0 < p < 1$. In particular, starting from a Z with mean μ and $r(Z) = r$ satisfying the conditions above, applying Theorem 8.1 with Z replaced by $Z(p)$ shows that in the random subgraph $G_Z^{(n)}(p)$ of $G_Z^{(n)}$ there is a giant component if and only if $p > p_0$, where

$$p_0 = \mu^{-1} c_0 = \frac{1}{2\mu} \frac{1-\sqrt{r}}{1-r}$$

for $r \neq 1$ and $p_0 = \frac{1}{4\mu}$ if $r = 1$, and that if $p = p_0 + \varepsilon$ then the size of the giant component is w.h.p.

$$n \exp\left(-\left(\frac{\pi}{2r^{1/4}\mu^{1/2}} + o(1) \right) \varepsilon^{-1/2} \right)$$

as $\varepsilon \to 0$ from above. (This time we have written the upper and lower bounds together in a weaker but simpler form.) In particular, taking $Z = m$ to be constant, so $\mu = m$ and $r = (m-1)/m$, and taking $Z_i = m$ for all $i > m$ (and $Z_i = 0$ for $i \leqslant m$), we obtain Theorem 2.3.

8.2. Comparison with a branching process

As in Section 7, or in [8], we shall prove results about the largest component in $G_Z^{(n)}$ by comparison with a branching process. This method was used in [10], but with more restrictive conditions. As in [10], each point in the branching process has a type $(\alpha(x), T(x))$, where $T(x)$ has two possible values, which we shall write here as L for points 'reached from the left' and R for points 'reached from the right'. Each point will have a random number of left- and right-children in the next generation, where the α value for a left/right-child is smaller/larger than the α value of the parent, and all left/right-children are reached from the right/left.

The precise definition is that we start from $X_0 = \{x_0\}$, a single point of type (α, R), where α is uniformly distributed on $(0,1)$. Each point $x \in X_k$ gives rise to children in the next generation as follows: x has a random number N of left-children, of types (β_i, R), $i = 1, \ldots, N$, where the β_i are independently uniformly distributed on $(0, \alpha(x))$. Here N has the distribution of Z if $T(x) = R$, and of $Y - 1$ if $T(x) = L$, where Y is the size-biased version of Z, satisfying $\Pr(Y = t) = t \Pr(Z = t)/\mathbf{E}(Z)$. Note that $\mathbf{E}(Y-1) = \mathbf{E}(Z(Z-1))/\mathbf{E}(Z) = r\mathbf{E}(Z) = rc$. The point x also has a Poisson number of right-children of types (β_i, L), where the β_i values are distributed according to a Poisson process on $(\alpha(x), 1)$ with density $c\beta^{-1}\,\mathrm{d}\beta$. Let (X_k) be the branching process defined in this way.

As before, given a (large enough) value of n, let (Γ_k) be the neighbourhood expansion process in $G_Z^{(n)}$, so $\Gamma_0 = \{v_0\}$, where v_0 is chosen uniformly at random from $\{1, 2, \ldots, n\}$, and Γ_k is the set of vertices at graph distance exactly k from v_0 in $G_Z^{(n)}$. We claim that (X_k) and (Γ_k) can be coupled so that with high probability the sizes of corresponding generations agree until they become 'large' or die out. This time, we shall use a smaller

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

meaning of 'large', namely larger than $M = (\log n)^2$. (Perhaps any function $M(n)$ with $M(n)/\log n \to \infty$ would do if we were more careful in the proof.)

Lemma 8.2. *Let Z and $Z_1, Z_2, \ldots$ be given, satisfying the conditions of Theorem 8.1. For each n the processes (Γ_k) and (X_k) can be coupled so that with probability $1 - O(M^{-2})$ we have*

$$|\Gamma_k| = |X_k|$$

for $0 \leqslant k \leqslant K$, and either $|\Gamma_K| = |X_K| = 0$ or

$$\sum_{k=0}^{K+2} |\Gamma_k|, \quad \sum_{k=0}^{K+2} |X_k| \geqslant M, \tag{8.4}$$

where $M = (\log n)^2$.

Outline proof. The basic method is the same as the proof of Lemma 6.1. Also, a very similar result is proved in [10] though under stronger assumptions. Thus we shall only give an outline proof, concentrating on the difficulties, and explaining why with the smaller choice of M here our assumptions are sufficient.

As usual, we inductively couple the branching and neighbourhood expansion processes so that the number and types of points agree. Agreement in type means two things: firstly, agreement as to whether the point/vertex was reached from the left or the right, where a vertex $v \in \Gamma_k$ is reach from the left/right according to whether its (with high enough probability unique) neighbour $w \in \Gamma_{k-1}$ satisfies $w < v$ or $w > v$. Secondly, agreement between α values after rounding to multiples of $n^{-1/2}$, say, where the α value of a vertex v is v/n.

As before, there are two key points. Firstly, we may ignore any bad event that has probability $O(M^{-3})$ per point/vertex reached in X_k or Γ_k, since we abandon the coupling before considering more than M points/vertices. Secondly, we can ignore points/vertices with small α values, say smaller than $n^{-1/4}$; we shall prove the second claim later. As a consequence we may assume that Z is bounded by M^3, and that the Z_i and Z are identically distributed. To see this, note that as Z has finite mean and variance, the probability that Z exceeds M^3 is $O(M^{-6})$. As Y, the size-biased version of Z, has finite mean, the probability that Y exceeds M^3 is $O(M^{-3})$. Defining Z' to have the distribution of Z conditioned on $Z \leqslant M^3$, and Y' to be the size-biased version of Z', we see that $d_{TV}(Z, Z') = O(M^{-6})$ and $d_{TV}(Y, Y') = O(M^{-3})$. Thus Z and Z', and also Y and Y', may be coupled to agree with probability $1 - O(M^{-3})$. It follows that we can couple the branching processes defined using the distributions Z and Z' so that with probability $1 - O(M^{-2})$ they agree until both become large. Similarly, we can couple $G_Z^{(n)}$ with a graph where the down-degree $Z'(i)$ of each vertex $i \geqslant n^{3/4}$ has distribution Z', and these down-degrees are independent. As $d_{TV}(Z, Z') = O(M^{-6})$, while, from (8.1), $d_{TV}(Z_i, Z) = O(M^{-4})$ for all $i \geqslant n^{3/4}$, we can do this coupling so that the down-neighbourhoods of all but $O(nM^{-4})$ vertices agree; call the remaining vertices *bad*. We claim that it is unlikely that the neighbourhood expansion process reaches a bad vertex before becoming large. If we

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

reach a bad vertex of large degree (larger than M), there is no problem, as the next generation is large by definition, so the stopping condition (8.4) in the lemma applies. (Here it helps that we continue the sums in (8.4) to $k = K + 2$ rather than $k = K + 1$.) For bad vertices of degree at most M it is easy to see that the relative chance of reaching some bad vertex to reaching some good vertex is $O(M)$ times the density of bad vertices, and hence $O(M^{-3})$. Such a probability per vertex reached can safely be ignored.

We now turn to the claim that we can ignore points/vertices with α values less than $n^{-1/4}$. We assume here and from now on, as we may, that Z is bounded by M^3, and that all Z_i, $i \geqslant n^{3/4}$, have exactly the distribution of Z. Now a vertex v, $v \geqslant n^{3/4}$, is very unlikely to have a neighbour w with $w < v/M^6$, say. Indeed, as Z_v is bounded by M^3, and from the uniform distribution of neighbours, this probability is $O(M^{-3})$. The same bound holds conditional on Z_v, so we can apply this to a vertex $v \in \Gamma_k$ whether v was reached from the left or the right. It follows that we may assume as we construct the coupling that no $v \in \Gamma_k$ has a left-neighbour $w < v/M^6$. We may also assume that $v_0 \geqslant n/M^3$. Suppose now that some Γ_k, reached before the stopping condition (8.4) holds, contains a vertex v with $\alpha(v) < n^{-1/4}$, *i.e.*, $v < n^{3/4}$. Then we must have $k \geqslant 3$, and Γ_{k-3} must contain a w with $w < M^{18}n^{3/4} < n^{0.76}$. But any such w is likely to have many right-neighbours in Γ_{k-2}: its expected number of neighbours w' with $n^{0.76} < w' < n^{0.77}$ is $\Omega(\log n)$, and from independence and exponential concentration of the Binomial distribution, it follows that with probability $1 - O(M^{-100})$ the number of actual neighbours w' in this range is at least $\Omega(\log n)$ with a reduced implicit constant. Similarly, with very high probability each of these w' has $\Omega(\log n)$ neighbours w'' in Γ_{k-1} with $n^{0.77} < w'' < n^{0.78}$. After one more step we find with high enough probability $\Omega((\log n)^3) > M$ vertices in Γ_k. Thus the probability that some Γ_k contains a v with $\alpha(v) < n^{-1/4}$ but that $|\bigcup_{t \leqslant k} \Gamma_k| < M$ holds is $O(M^{-2})$, and we may abandon our coupling if this event occurs. A similar but simpler argument applies to the branching process.

From this point the rest of the argument is easy, though not easy to write down precisely. We give only an outline; a more complete argument in a special case is given in [10]. The basic method is as in Section 7, but we must deal with dependence between edges introduced via the Z_is. There are two key observations. Firstly, suppose that $i < j$. Then the conditional probability that $ij \in E(G_Z^{(n)})$ given Z_j is proportional to Z_j. Hence, the conditional distribution of Z_j given that $ij \in E(G_Z^{(n)})$ is exactly the size-biased version Y of Z. Thus, ignoring other complications, when we reach a vertex from the left we expect its number of other left-neighbours to have the distribution of $Y - 1$. The second point is that we stop before reaching too many vertices; when we come to consider some $v \in \Gamma_k$, asking what are its neighbours in Γ_{k+1}, a certain set U of vertices have been 'used up'. This set consists of $\bigcup_{t \leqslant k} \Gamma_k$ together with any vertices of Γ_{k+1} found as neighbours of $w \in \Gamma_k$ considered before v. By the stopping rule, we may assume that $|U| < M$. Conditioning on the part of the graph uncovered so far, all we know about the neighbourhood of v is that it includes a certain $u \in \Gamma_{k-1}$, and otherwise avoids a certain subset of U. Similarly, for any $w \in [n] \setminus U$, all we know about the neighbourhood of w is that it avoids a certain subset of U. Now when we come to consider whether some edge vw is present, with $w < v$, say, we need to know the conditional distribution of Z_v. The key point is that since U is so small, Z_v is never too large, and v is large enough, given any possible value

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

of Z_v the probability that v sends an edge to U is extremely small ($o(n^{-1/2})$, say). Hence the conditional distribution of Z_v is very close to Z or to Y, depending on how v was reached. Similarly, the chance that some $w > v$ sends an edge to v is proportional to the conditional expectation of Z_w, which is very close to its unconditional value. □

As before, this coupling result enables us to compare the size of the giant component (if any) in $G_Z^{(n)}$ with the survival probability $\sigma(Z)$ of the branching process (X_k).

Lemma 8.3. *Let Z and $Z_1, Z_2, \ldots$ be given, satisfying the conditions of Theorem* 8.1. *Then*

$$C_1\big(G_Z^{(n)}\big) = (\sigma(Z) + o(1))n$$

holds w.h.p. as $n \to \infty$.

Proof. Let us write $\#_k(G)$ for the number of vertices of a graph G in components of order k. Writing $M = (\log n)^2$ as before, Lemma 8.2 implies immediately that

$$\mathbf{E}\big(\#_k\big(G_Z^{(n)}\big)\big) = (s_k(Z) + O(M^{-2}))n$$

for $0 \leqslant k \leqslant M$, where s_k is the probability that the branching process has total size k, so $\sigma(Z) = 1 - \sum_{k=0}^{\infty} s_k$. We shall in fact apply the lemma to a slightly modified graph.

As usual, the next step is to show concentration. Again there is a potential problem with large degrees, which can be dealt with in the following global way. Let Z' be Z capped at $B = \lceil M^3 \rceil$, say, so Z' is the smaller of Z and B. Define Z_i' from Z_i similarly. Let $G_{Z'}^{(n)}$ be the uniformly grown Z'-out graph defined using the Z_i'. By construction of Z_i', we have $G_{Z'}^{(n)} \subset G_Z^{(n)}$. Furthermore, if ij, $i < j$, is an edge present in $G_Z^{(n)}$ but not $G_{Z'}^{(n)}$, then j has large degree (at least B) in both graphs. It follows that the only way a vertex v can lie in a component of order $k < B$ in one graph and not in the other is if v is in a component of order k in $G_{Z'}^{(n)}$, and that component meets one of the extra edges of $G_Z^{(n)}$. Hence we have

$$|\#_k\big(G_Z^{(n)}\big) - \#_k\big(G_{Z'}^{(n)}\big)| \leqslant k|E\big(G_Z^{(n)} \setminus G_{Z'}^{(n)}\big)|.$$

Now the right-hand side above is just $k\sum_i (Z_i - Z_i')$. For each i we have

$$\mathbf{E}(Z_i - Z_i') = \sum_{t \geqslant B} \Pr(Z_i = t)(t - B) \leqslant \sum_{t \geqslant B} \Pr(Z_i = t)t^2/B \leqslant \mathbf{E}\big(Z_i^2\big)/B.$$

From our assumptions, the $\mathbf{E}(Z_i^2)$ are uniformly bounded above. Hence,

$$\mathbf{E}\big(|\#_k\big(G_Z^{(n)}\big) - \#_k(G_{Z'}^{(n)})|\big) = O(nk/B) = O(nk/M^3),$$

and the probability that $\sum_{k \leqslant M} |\#_k(G_Z^{(n)}) - \#_k(G_{Z'}^{(n)})|$ exceeds $n/\sqrt{M}$ is $o(1)$.

We know from the proof of Lemma 8.2 that truncating Z at $B \geqslant M^3$ has little effect on the branching process, and in particular that $|s_k(Z') - s_k(Z)| = O(M^{-2})$. Since $G_{Z'}^{(n)}$ satisfies the conditions of Lemma 8.2 we thus have

$$\mathbf{E}\big(\#_k\big(G_{Z'}^{(n)}\big)\big) = (s_k(Z') + O(M^{-2}))n = (s_k(Z) + O(M^{-2}))n$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

for $0 \leqslant k \leqslant M$. Finally, $\#_k(G_{Z'}^{(n)})$ is concentrated about its mean: in constructing the graph $G_{Z'}^{(n)}$ we have n independent steps, the construction of the down-neighbourhoods of the vertices. Changing one such step adds/deletes at most B edges to the final graph, and hence changes $\#_k$ by at most $4kB \leqslant n^{1/10}$. It follows that with probability $1 - O(M^{-2})$ we have $\#_k(G_{Z'}^{(n)})$ within $n^{3/4}$ of its mean, say. Combined with the observations above this shows that

$$\Pr\left(\sum_{k \leqslant M} |\#_k(G_Z^{(n)}) - n s_k(Z)| \geqslant 2n/\sqrt{M}\right) = o(1).$$

It follows that w.h.p. the graph $G_Z^{(n)}$ has $\sigma(Z)n + o(n)$ vertices in 'large' components, where we may take large to mean greater than L for any $L \to \infty$ with $L \leqslant M$.

It remains to show that most (all but $o(n)$) vertices in large components are in a single giant component. The argument is exactly as in the proof of Lemma 6.3: recall that $G_Z^{(n)}$ was defined as one graph in a graph process $(G_Z^{(t)})_{t \geqslant 0}$. We consider the graph process from $t = n'$ to $t = n$, where $n' = \lfloor n(1-\eta) \rfloor$ for $\eta = (\log n)^{-1/2}$, say. Each time we add a new vertex, passing from $G_Z^{(t-1)}$ to $G_Z^{(t)}$, there is a positive probability that we add at least two edges. Thus, given disjoint sets of vertices A and B, there is a probability $\Omega(|A||B|/n^2)$ that the new vertex sends edges to both A and B. This is exactly the single condition needed in the proof of Lemma 6.3 to show that almost all vertices in large components are in a single giant component. □

8.3. Analysis of the branching process

We have shown that the size of the giant component in $G_Z^{(n)}$ is w.h.p. close to n times the survival probability $\sigma(Z)$ of a certain branching process defined above; it remains to estimate $\sigma(Z)$. As before, it will be convenient to modify the branching process, by deleting all points with $\alpha(.)$ values less than ρ. The survival probability $\sigma(Z, \rho)$ of the modified process is of course at most $\sigma(Z) = \sigma(Z, 0)$.

We could write down directly the equation giving the survival probabilities of points in our branching process, but as in Section 7 and in [10] it is more convenient to work with a transformation of these probabilities. In the (possibly) truncated process ($\rho \geqslant 0$), let $L(\alpha)$ be the probability that a given left-child of a point x with $\alpha(x) = \alpha$ survives, and let us write the probability that no right-children of x survive as $\exp(-cR(\alpha))$. (This is natural due to the Poisson form of the distribution of right-children.) Then, as noted in [10] (see equations (21), (22) and the modifications in Section 9), it is easy to check that

$$L(\alpha) = \frac{1}{\alpha} \int_{\beta=\rho}^{\alpha} \left(1 - e^{-cR(\beta)} \sum_{t \geqslant 0} p_t (1 - L(\beta))^t\right) \mathrm{d}\beta, \tag{8.5}$$

and

$$R(\alpha) = \int_{\beta=\alpha}^{1} \frac{1}{\beta} \left(1 - e^{-cR(\beta)} \sum_{t \geqslant 0} q_t (1 - L(\beta))^t\right) \mathrm{d}\beta, \tag{8.6}$$

where $p_t = \Pr(Z = t)$ and $q_t = \Pr(Y - 1 = t)$, where Y is the size-biased version of Z, so

$$q_t = (t+1)p_{t+1} / \sum_i i p_i.$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

Also, recalling that the branching process is started from a single point x_0 with $\alpha(x_0)$ uniformly chosen from $[0,1]$ (and then killed immediately if $\alpha(x_0) < \rho$), the survival probability $\sigma(Z,\rho)$ of the process is given by

$$\sigma(Z,\rho) = \int_{\beta=\rho}^{1} \left(1 - e^{-cR(\beta)} \sum_{t \geqslant 0} p_t (1 - L(\beta))^t\right) \mathrm{d}\beta = L(1). \tag{8.7}$$

For the comparison with [10] note that the factor p present there is absent here, as we have absorbed the edge deletion into the distribution of Z. Note also that although all bounds we shall prove depend only on c and r, $\sigma(Z) = \sigma(Z,0)$ itself depends on the whole distribution of Z.

Proof of Theorem 8.1. To complete the proof of Theorem 8.1 it suffices to prove that the bounds (8.2) and (8.3) claimed for $\lambda(Z)$ hold for $\sigma(Z)$. Note that $\sigma(Z)$ is monotonic in Z, in the sense that if Z_1 stochastically dominates Z_2, then $\sigma(Z_1) \geqslant \sigma(Z_2)$. In fact, corresponding random graphs, and hence the branching processes, can be coupled so that one contains the other. It follows that we may restrict our attention to the case $c > c_0$. Indeed, given Z with $c \leqslant c_0$, and given any $c' = c_0 + \varepsilon > c_0$ with $c' < 1$, it is easy to check that we can construct a distribution Z' with $\mathbf{E}(Z') = c'$ and $r(Z') = r(Z)$. (The idea is to shift some weight from value zero to larger values, using $c' < 1$. If we shift weight only to value 1, we can achieve the required mean without increasing the second factorial moment (SFM) at all. Shifting less weight to very large values, we can achieve a large increase in the SFM. Something in between gives exactly the increase of a factor $(c'/c)^2$ in the SFM required.) Now if we prove the upper bound on $\sigma(Z')$ corresponding to (8.3), then it follows that $\sigma(Z) = 0$, since (8.3) tends to 0 as $\varepsilon \to 0$. Similarly, but more simply, given any $c_2 > c_0$ it suffices to consider $c \leqslant c_2$; for larger values of c we may simply use 1 as an upper bound on $\sigma(Z)$, and consider $Z' = Z(c_2/c)$ to obtain a lower bound.

As before, let us write $\mathbf{F}_\rho$ for the operator defined by the right-hand sides of equations (8.5) and (8.6), so we are looking for the maximal solution to $(L,R) = \mathbf{F}_\rho((L,R))$. To obtain upper and lower bounds, it suffices to find pairs of functions which are mapped under $\mathbf{F}_\rho$ to functions above/below themselves. This is because L and the monotonic transformation $1 - \exp(-cR)$ of R are exactly the survival probabilities of a certain branching process. This process is not that considered in the previous subsection, but is equivalent to it, in that one can be transformed into the other. The transformed process has one point of type 1 for each left-edge from a vertex (or left-child of the untransformed process), and one point of type 2 corresponding to the (possibly empty) set of all right-edges (right-children).

We consider the linearized form of (8.5), (8.6). Note that $\sum t p_t = \mathbf{E}(Z) = c$, while $\sum t q_t = \mathbf{E}(Y-1) = \mathbf{E}(Z(Z-1))/\mathbf{E}(Z) = rc$. Thus the linearized form of these equations is

$$\begin{aligned} \tilde{L}(\alpha) &= \frac{1}{\alpha} \int_{\beta=\rho}^{\alpha} (c' \tilde{L}(\beta) + c' \tilde{R}(\beta))\, \mathrm{d}\beta, \\ \tilde{R}(\alpha) &= \int_{\beta=\alpha}^{1} \frac{1}{\beta} (rc' \tilde{L}(\beta) + c' \tilde{R}(\beta))\, \mathrm{d}\beta. \end{aligned} \tag{8.8}$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

Here we have again replaced c by c' as we shall solve the linearized equations with c' slightly different from c to obtain bounds on the solutions to the real equations (8.5), (8.6).

It will be convenient to think of r as fixed and c as variable. But we shall only consider values of c, c' just above the critical value c_0, which lies strictly between 0 and $1/2$. Thus we have $c, c' = \Theta(1)$, and $1/2 - c' = \Theta(1)$.

As in [10], let

$$\theta = \frac{1}{2}\sqrt{-1 + 4c' - 4(c')^2(1-r)},$$

set

$$A = \frac{\theta}{c'r}, \qquad B = \frac{1/2 - c'}{c'r},$$

and let

$$\tilde{L}(\alpha) = \alpha^{-1/2}(A\cos(\theta \log \alpha) + B\sin(-\theta \log \alpha)),$$

$$\tilde{R}(\alpha) = \alpha^{-1/2}\sin(-\theta \log \alpha).$$

Then, as shown in [10], and as may be verified by substituting back, $\tilde{L}$ and $\tilde{R}$ solve (8.8), provided

$$\rho = \exp\left(-\frac{\pi - \tan^{-1}(A/B)}{\theta}\right).$$

From now on we shall define ρ by this condition.

Let us note the following properties of θ, ρ, A, B, $\tilde{L}$ and $\tilde{R}$. As c' tends to c_0 from above, $\theta \to 0$. (In fact, c_0 is defined by $\theta(c_0) = 0$.) In particular, if $c' = c_0 + \delta^2$, then

$$\theta = r^{1/4}\delta + O(\delta^3). \tag{8.9}$$

This may be checked by differentiating the square of $\theta(c)$ with respect to c. We shall always consider $c' = c_0 + o(1)$, so we have $\theta = o(1)$.

Now $A = \Theta(\theta) = o(1)$ while $B = \Theta(1)$. Thus $\pi - \tan^{-1}(A/B) = \pi - \Theta(\theta)$, and

$$\rho = \Theta(\exp(-\pi/\theta)). \tag{8.10}$$

Turning to $\tilde{L}$ and $\tilde{R}$, note that they are nonnegative on $[\rho, 1]$; indeed, ρ is defined as the largest zero of $\tilde{L}$ in $(0, 1)$. Secondly, $\tilde{R}$ is decreasing; this is easy to see from the fact that $\tilde{R}$ solves (8.8). Thirdly,

$$\tilde{L}(\alpha), \tilde{R}(\alpha) = O\left(\theta\rho^{-1/2}\right) \tag{8.11}$$

uniformly for $\alpha \in [\rho, 1]$. This is clear for $\tilde{R}$, as $\tilde{R}$ is decreasing and

$$\tilde{R}(\rho) = \rho^{-1/2}\sin(-\theta \log \rho) = \rho^{-1/2}\sin(\pi - \tan^{-1}(A/B)) = \rho^{-1/2}\sin(\tan^{-1}(A/B)),$$

and the argument of the arctangent is $\Theta(\theta)$. To see that $\tilde{L}(\alpha) = O(\theta\rho^{-1/2})$, note that the sin term contributing to $\tilde{L}(\alpha)$ is $\Theta(\tilde{R}(\alpha))$. For the cos term note that $A = \Theta(\theta)$ and $|\cos(x)| \leqslant 1$. Finally, we claim that

$$\tilde{L}(e^2\rho) = \Theta\left(\theta\rho^{-1/2}\right). \tag{8.12}$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

This is easy to check, as the derivative of $\alpha^{1/2}\tilde{L}(\alpha)$ with respect to $\log\alpha$ is $\theta(-A\sin(\theta\log\alpha) - B\cos(-\theta\log\alpha))$, which is asymptotically $B\theta$ for $\alpha = e^x\rho$ and x bounded, because $-\theta\log\alpha \sim \pi$. Hence, noting that $\tilde{L}(\rho) = 0$ and taking $x = 2$, at $\alpha = e^2\rho$ we have $\alpha^{1/2}\tilde{L}(\alpha) \sim 2B\theta$.

We are now in a position to prove good upper and lower bounds on $\sigma(Z)$, arguing along the lines used in Section 7.

For the lower bound, given Z with $r(Z) = r$ and $c = \mathbf{E}(Z) = c_0 + \varepsilon$, set $c' = c - \varepsilon^{3/2}$, noting that $c' > c_0$ if ε is small enough. Thus we may write $c' = c_0 + \delta^2$, noting that $\delta = \varepsilon^{1/2}(1 - O(\varepsilon^{1/2}))$. Let θ, ρ, $\tilde{L}$ and $\tilde{R}$ be defined as above, using this value of c', and set

$$L_0(\alpha) = f\tilde{L}(\alpha), \qquad R_0(\alpha) = f\tilde{R}(\alpha)$$

for $\rho \leqslant \alpha \leqslant 1$. We shall take f to be a very small constant η times $\theta^{-1}\varepsilon^{3/2}\rho^{1/2}$. Thus, from (8.11) we have that L_0 and R_0 are bounded on $[\rho, 1]$ by $O(\eta\varepsilon^{3/2})$. It follows that if we consider $\mathbf{F}_\rho((L_0, R_0))$, the linearized forms of the integrands in (8.5), (8.6) are pointwise good approximations to the original forms, and in particular, that the non-linear forms are at least $1 - O(\eta\varepsilon^{3/2})$ times the linearized versions. To see this, note that for $x > 0$ and $0 < y < 1$ we have

$$\begin{aligned} 1 - e^{-x}(1-y)^t &\geqslant 1 - e^{-x}e^{-ty} \\ &= 1 - e^{-(x+ty)} \\ &\geqslant x + ty - (x+ty)^2/2. \end{aligned}$$

Thus, recalling that $\sum_t p_t = 1$ and $\sum_t tp_t = \mathbf{E}(Z) = c$,

$$\begin{aligned} 1 - e^{-cR_0(\beta)}\sum_{t\geqslant 0} p_t(1-L_0(\beta))^t &= \sum_t p_t\big(1 - e^{-cR_0(\beta)}(1-L_0(\beta))^t\big) \\ &\geqslant cR_0(\beta) + cL_0(\beta) - O((R_0(\beta)+L_0(\beta))^2). \end{aligned}$$

In the last step we used $\sum t^2 p_t = \mathbf{E}(Z^2) = rc^2 + c = O(1)$. Setting $(L_0', R_0') = \mathbf{F}_\rho(L_0, R_0)$, and recalling that $L_0, R_0 = O(\eta\varepsilon^{3/2})$, it follows that

$$\begin{aligned} L_0'(\alpha) &\geqslant \big(1 - O\big(\eta\varepsilon^{3/2}\big)\big)\frac{1}{\alpha}\int_{\beta=\rho}^{\alpha}(cL_0(\beta) + cR_0(\beta))\,\mathrm{d}\beta \\ &= \big(1 - O\big(\eta\varepsilon^{3/2}\big)\big)(c/c')L_0(\alpha), \end{aligned}$$

since L_0 and R_0 solve (8.8). A similar argument proves a corresponding inequality for R_0, except that we require the additional condition

$$\sum t^2 q_t = O(1). \tag{8.13}$$

This is ensured by our condition on $\mathbf{E}(Z^3)$, and is the only place this condition is used. Finally, choosing η small enough, the factor $(1 - O(\eta\varepsilon^{3/2}))(c/c')$ is larger than 1, so $\mathbf{F}_\rho((L_0, R_0)) \geqslant (L_0, R_0)$ holds pointwise. Thus L_0 and R_0 are pointwise lower bounds on the maximal solution to $\mathbf{F}_\rho((L, R)) = (L, R)$, and from (8.7),

$$\sigma(Z, \rho) \geqslant \int_{\beta=\rho}^{1}\left(1 - e^{-cR_0(\beta)}\sum_{t\geqslant 0} p_t(1-L_0(\beta))^t\right)\mathrm{d}\beta = L_0'(1) \geqslant L_0(1).$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

As $L_0(1) = f\tilde{L}(1) = fA = \Theta(\theta f)$, and we have chosen f to be a small constant times $\theta^{-1}\varepsilon^{3/2}\rho^{1/2}$, we have

$$\sigma(Z) \geqslant \sigma(Z,\rho) = \Omega(\theta f) = \Omega\big(\varepsilon^{3/2}\rho^{1/2}\big) = \Omega\left(\varepsilon^{3/2}\exp\left(-\frac{\pi}{2\theta}\right)\right),$$

where the last step is from (8.10). Now from (8.9), $\theta = r^{1/4}\delta + O(\delta^3) = r^{1/4}\varepsilon^{1/2} + O(\varepsilon)$. It follows that $\exp(-1/\theta) = \Theta(\exp(-1/(r^{1/4}\varepsilon^{1/2})))$, so we obtain the bound on $\sigma(Z)$ required for (8.2).

The argument for the upper bound is similar; this time there is no problem with the linear approximation, as the integrands in (8.5), (8.6) are always at most their linearized forms. However, we must put back the vertices we deleted. The method is exactly analogous to that used in Section 7.

Given Z with $r(Z) = r$ and $c = \mathbf{E}(Z) = c_0 + \varepsilon$, set $c' = c + \varepsilon^{3/2}$, and define δ, θ, ρ, $\tilde{L}$ and $\tilde{R}$ as above, using this value of c'. Let $\tau = e^2\rho$, set

$$L_1(\alpha) = \begin{cases} f\tilde{L}(\alpha) & \alpha \geqslant \tau \\ 1 & \alpha < \tau, \end{cases}$$

$$R_1(\alpha) = \begin{cases} f\tilde{R}(\alpha) & \alpha \geqslant \tau \\ \log(1/\alpha) & \alpha < \tau, \end{cases}$$

where $f > 0$ is a scale-factor to be chosen later, and define L_2, R_2 by $(L_2, R_2) = \mathbf{F}_0((L_1, R_1))$. We claim that L_2, R_2 are pointwise at most L_1, R_1. This is immediate for $\alpha < \tau$, using only the fact that the integrands in (8.5), (8.6) are at most one and $1/\beta$ respectively. Suppose that $\alpha \geqslant \tau$. As in Section 7, $R_2(\alpha) \leqslant R_1(\alpha)$ is almost immediate: the integral defining $R_2(\alpha)$ only involves values of $L_1(\beta)$ and $R_1(\beta)$ for $\beta \geqslant \alpha$, where $L_1 = f\tilde{L}$ and $R_1 = f\tilde{R}$. The integral is bounded above by its linearized form, which is exactly $f(c/c')\tilde{R}(\alpha)$, as $\tilde{L}$, $\tilde{R}$ solve (8.8). This in turn is at most $R_1(\alpha)$ as $c' > c$.

Turning to L_2, bounding the integrand by 1 for $\beta \leqslant \tau$ and by its linearized form for $\beta \geqslant \tau$, and noting that $\rho < \tau$, for $\alpha \geqslant \tau$ we have

$$\begin{aligned} L_2(\alpha) &\leqslant \frac{\tau}{\alpha} + \frac{1}{\alpha}\int_{\beta=\tau}^{\alpha} (cL_1(\beta) + cR_1(\beta))\,\mathrm{d}\beta \\ &\leqslant \frac{\tau}{\alpha} + \frac{c}{c'} f\tilde{L}(\alpha) \\ &= f\tilde{L}(\alpha) + \frac{\tau}{\alpha} - \Omega\big(\varepsilon^{3/2} f\tilde{L}(\alpha)\big), \end{aligned}$$

using $c' = c + \varepsilon^{3/2}$. This will be at most $f\tilde{L}(\alpha) = L_1(\alpha)$ provided we choose f large enough that

$$C\varepsilon^{3/2} f\tilde{L}(\alpha)\alpha \geqslant \tau \tag{8.14}$$

holds for all $\alpha \geqslant \tau$, where C is the constant implied by the $\Omega(.)$ above. Now $\alpha\tilde{L}(\alpha)$ is increasing in α, as $\tilde{L}$ is given by (8.8), so it suffices to ensure that $C\varepsilon^{3/2} f\tilde{L}(\tau) \geqslant 1$. But $\tau = e^2\rho$, so from (8.12) we have $\tilde{L}(\tau) = \Theta(\theta\rho^{-1/2})$, and we may choose f so that (8.14) is satisfied and at the same time

$$f = O\big(\varepsilon^{-3/2}\theta^{-1}\rho^{1/2}\big).$$

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

With this choice of f we have $\mathbf{F}_0((L_1, R_1)) = (L_2, R_2) \leqslant (L_1, R_1)$ holding pointwise, so by the second fact proved at the end of Section 5, L_1 and R_1 are pointwise upper bounds on the maximal solution (L, R) to $\mathbf{F}_0((L, R)) = (L, R)$. It follows that $\sigma(Z)$ is at most $L_1(1) = f\tilde{L}(1) = \Theta(\theta f)$. Hence

$$\sigma(Z) = O\left(\varepsilon^{-3/2}\rho^{1/2}\right).$$

The same calculation as used at the end of the proof of the lower bound shows that

$$\rho^{1/2} = \Theta\left(\exp\left(-\frac{\pi}{2\theta}\right)\right) = \Theta\left(\exp\left(-\frac{\pi}{2r^{1/4}\varepsilon^{1/2}}\right)\right),$$

completing the proof of the bound (8.3) and hence of Theorem 8.1. □

As noted after the statement of Theorem 8.1, this perhaps rather unwieldy result immediately implies specific results for the uniformly grown m-out random graph and the uniformly grown (or c/j) random graph, namely Theorems 2.3 and 2.4. In addition, Theorem 8.1 gives an explanation of the difference between these two specific results: as well as its mean, which controls the edge density, the second moment of Z is important. This is because the second moment (but no higher moments) appear when we count the expected numbers of paths between certain vertices in the Z-out random graph.

References

[1] Abramowitz, M. and Stegun, I. A., eds (1965) *The Handbook of Mathematical Functions*, Dover, New York. (Ninth printing, 1970.)

[2] Albert, R. and Barabási, A.-L. (2002) Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74** 47–97.

[3] Albert, R., Jeong, H. and Barabási, A.-L. (2000) Error and attack tolerance of complex networks. *Nature* **406** 378–382.

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

[5] Bollobás, B. (2001) *Random Graphs*, second edn, Vol. 73 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge.

[6] Bollobás, B., Janson, S. and Riordan, O. (2005) The phase transition in the uniformly grown random graph has infinite order. *Random Struct. Alg.* **26** 1–36.

[7] Bollobás, B. and Riordan, O. (2002) Mathematical results on scale-free random graphs. In *Handbook of Graphs and Networks* (S. Bornholdt and H. G. Schuster, eds), Wiley-VCH, Weinheim, pp. 1–34.

[8] Bollobás, B. and Riordan, O. (2003) Robustness and vulnerability of scale-free random graphs. *Internet Mathematics* **1** 1–35.

[9] Bollobás, B. and Riordan, O. (2004) The diameter of a scale-free random graph. *Combinatorica* **24** 5–34.

[10] Bollobás, B. and Riordan, O. (2005) Slow emergence of the giant component in the growing m-out graph. *Random Struct. Alg.* **27** 1–24.

[11] Bollobás, B., Riordan, O., Spencer, J. and Tusnády, G. (2001) The degree sequence of a scale-free random graph process. *Random Struct. Alg.* **18** 279–290.

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

[13] Callaway, D. S., Newman, M. E. J., Strogatz, S. H. and Watts, D. J. (2000) Network robustness and fragility: percolation on random graphs. *Phys. Rev. Lett.* **85** 5468–5471.

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press

[14] Cohen, R., Erez, K., ben-Avraham, D. and Havlin, S. (2000) Resilience of the Internet to random breakdowns. *Phys. Rev. Lett.* **85** 4626–4628.
[15] Dorogovtsev, S. N. and Mendes, J. F. F. (2002) Evolution of networks. *Adv. Phys.* **51** 1079.
[16] Dorogovtsev, S. N., Mendes, J. F. F. and Samukhin, A. N. (2001) Anomalous percolation properties of growing networks. *Phys. Rev. E* **64** 066110.
[17] Durrett, R. (2003) Rigorous result for the CHKNS random graph model. In *Proc. Discrete Random Walks 2003* (C. Banderier and C. Krattenthaler, eds), *Discrete Mathematics and Theoretical Computer Science* **AC** 95–104. `http://dmtcs.loria.fr/proceedings/`
[18] Durrett, R. and Kesten, H. (1990) The critical parameter for connectedness of some random graphs. In *A Tribute to Paul Erdős* (A. Baker, B. Bollobás and A. Hajnal, eds), Cambridge University Press, Cambridge, pp. 161–176.
[19] Kalikow, S. and Weiss, B. (1988) When are random graphs connected? *Israel J. Math.* **62** 257–268.
[20] Norris, J. private communication.
[21] Shepp, L. A. (1989) Connectedness of certain random graphs. *Israel J. Math.* **67** 23–33.
[22] Söderberg, B. (2002) General formalism for inhomogeneous random graphs. *Phys. Rev. E* **66** 066121.
[23] Söderberg, B. (2003) Random graphs with hidden colour. *Phys. Rev. E* **68** 015102(R).
[24] Söderberg, B. (2003) Properties of random graphs with hidden colour. *Phys. Rev. E* **68** 026107.

https://doi.org/10.1017/S096354830500708X Published online by Cambridge University Press