*Combinatorics, Probability and Computing* (2008) 17, 67–86. © 2007 Cambridge University Press
doi:10.1017/S096354830700867X First published online 20 July 2007 Printed in the United Kingdom

# The Critical Phase for Random Graphs with a Given Degree Sequence

M. KANG[†] and T. G. SEIERSTAD[‡]

Institut für Informatik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
(e-mail: {kang,seiersta}@informatik.hu-berlin.de)

*Received 16 February 2006; revised 1 February 2007*

We consider random graphs with a fixed degree sequence. Molloy and Reed [11, 12] studied how the size of the giant component changes according to degree conditions. They showed that there is a phase transition and investigated the order of components before and after the critical phase. In this paper we study more closely the order of components at the critical phase, using singularity analysis of a generating function for a branching process which models the random graph with a given degree sequence.

## 1. Introduction and main result

### 1.1. Standard random graph model

One of the most interesting results about the standard random graph model, $G(n,p)$, is the *phase transition* studied by Erdős and Rényi [6], where $G(n,p)$ is a random graph on $n$ vertices where every possible edge is included in the graph independently of each other, with probability $p = p(n)$. Suppose that $p = \frac{c}{n}$. Erdős and Rényi [6] proved that if $c < 1$, then *asymptotically almost surely* (a.a.s. for short), which means with probability tending to 1 as $n \to \infty$, all the components in $G(n,p)$ have order $O(\log n)$, whereas if $c > 1$, then a.a.s. there is a unique component with $\Theta(n)$ vertices, while all other components have $O(\log n)$ vertices. In the case $c = 1$, they proved that $G(n,p)$ a.a.s. has components of order at least $n^{2/3}/\omega(n)$ but no components of order greater than $n^{2/3}\omega(n)$ for any function $\omega(n) \to \infty$. Bollobás [3] and Łuczak [8] considered the case that $c \to 1$ in more detail. Let $\lambda$ be such that

$$p = \frac{1}{n} + \frac{\lambda}{n^{4/3}}. \tag{1.1}$$

[†] Research supported by the Deutsche Forschungsgemeinschaft (DFG Pr 296/7-3).
[‡] Supported by the Deutsche Forschungsgemeinschaft within the European graduate program 'Combinatorics, Geometry, and Computation' (no. GRK 588/2).

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

If $\lambda \to -\infty$, then a.a.s. all the components have order $o(n^{2/3})$. If $\lambda \to \infty$, then there is a.a.s. a unique component of order $\gg n^{2/3}$, while all other components have order $o(n^{2/3})$. Furthermore, if we consider $G(n,p)$ as a graph process where edges are added one by one, then the largest component will a.a.s. remain largest until the end of the process.

### 1.2. Random graphs with a given degree sequence

Molloy and Reed [11] showed that a random graph model with a given degree sequence has a similar phase transition. An *asymptotic degree sequence* is a sequence of integer-valued functions $\mathscr{D} = \{d_0(n), d_1(n), \dots\}$, such that $d_i(n) = 0$ for $i \geqslant n$, and $\sum_{i\geqslant 0} d_i(n) = n$. The value $d_i(n)$ denotes the number of vertices of degree $i$ in a graph of order $n$. If $\mathscr{D}$ is an asymptotic degree sequence, we let $\mathscr{D}_n$ be the degree sequence $\{a_1, a_2, \dots, a_n\}$, where $a_j \geqslant a_{j+1}$ for every $j = 1, \dots, n-1$, and $\#\{j | a_j = i\} = d_i(n)$. Let $\Omega(\mathscr{D}_n)$ be the set of graphs on $n$ vertices with degree sequence $\mathscr{D}_n$. An asymptotic degree sequence $\mathscr{D}$ is *feasible* if $\Omega(\mathscr{D}_n) \neq \emptyset$ for all $n \geqslant 1$. All asymptotic degree sequences in this paper are assumed to be feasible.

If $\mathscr{D}$ is an asymptotic degree sequence, we let $G_n = G_n(\mathscr{D})$ be a random graph chosen from the set $\Omega(\mathscr{D}_n)$ uniformly at random.

For $i \geqslant 0$, we let $\lambda_i(n) = d_i(n)/n$. If $\mathscr{D}$ is such that $\lambda_i(n)$ converges to a constant as $n \to \infty$, for every $i \geqslant 0$, then $\mathscr{D}$ is said to be *smooth*, and we define

$$\lambda_i^* = \lim_{n\to\infty} \lambda_i(n).$$

We will assume that the asymptotic degree sequence $\mathscr{D}$ is *well-behaved* in the following sense, which is similar to the definition in [11], although slightly stronger.

**Definition.** An asymptotic degree sequence $\mathscr{D}$ is *well-behaved* if the following conditions are satisfied.

1 $\mathscr{D}$ is feasible and smooth.

2 Let $f(i)$ be a polynomial in $i$ with degree at most 3. Then

(a) $f(i)\lambda_i(n)$ tends uniformly to $f(i)\lambda_i^*$: that is, for all $\varepsilon > 0$, there exists $N$ such that, for all $n \geqslant N$ and for all $i \geqslant 0$,

$$|f(i)\lambda_i(n) - f(i)\lambda_i^*| < \varepsilon.$$

(b) The limit

$$\lim_{n\to\infty} \sum_{i\geqslant 0} f(i)\lambda_i(n)$$

exists, and the sum approaches the limit $L_f(\mathscr{D}) = \sum_{i\geqslant 0} f(i)\lambda_i^*$ uniformly: that is, for all $\varepsilon > 0$, there exist $i^*$ and $N$ such that, for all $n \geqslant N$,

$$\left| \sum_{i=0}^{i^*} f(i)\lambda_i(n) - L_f(\mathscr{D}) \right| < \varepsilon.$$

In [11], the function $Q(\mathscr{D}) = \sum_{i\geqslant 1} i(i-2)\lambda_i^*$ is defined. The following theorem is the main theorem of [11] and says that a *phase transition* occurs when $Q(\mathscr{D}) = 0$.

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

**Theorem 1.1. (Molloy and Reed 1995)** *Let $\mathscr{D}$ be a well-behaved sparse asymptotic degree sequence for which there exists $\varepsilon > 0$ such that, for all $n$ and $i > n^{1/4-\varepsilon}$, $d_i(n) = 0$.*

(a) *If $Q(\mathscr{D}) < 0$, and for some function $0 \leqslant \omega(n) \leqslant n^{1/8-\varepsilon}$, $d_i(n) = 0$ for all $i \geqslant \omega(n)$, then for some constant $R$ dependent on $Q(\mathscr{D})$, $G_n$ a.a.s. has no component with more than $R\omega(n)^2 \log n$ vertices, and a.a.s. has fewer than $2R\omega(n)^2 \log n$ cycles. Also, a.a.s. no component of $G_n$ has more than one cycle.*

(b) *If $Q(\mathscr{D}) > 0$, then there exist constants $\zeta_1, \zeta_2 > 0$ dependent on $\mathscr{D}$, such that $G_n$ a.a.s. has a component with at least $\zeta_1 n$ vertices and $\zeta_2 n$ cycles. Furthermore, if $Q(\mathscr{D})$ is finite, then $G_n$ a.a.s. has exactly one component of size greater than $C \log n$ for some constant $C$ dependent on $\mathscr{D}$.*

To see that it is natural to consider the quantity $Q(\mathscr{D})$, suppose that we start with a randomly chosen vertex $v$ in the graph and want to determine the order of the component it lies in. The vertex $v$ has degree $i$ with probability $\lambda_i(n)$. Then we can expose the component containing $v$ with a branching process, starting with the neighbours of $v$. When a vertex of degree $i$ is exposed, the number of 'unsaturated' vertices increases by $i-2$. The probability that this happens is roughly $\frac{i\lambda_i(n)}{d}$, where $d = \sum_i i\lambda_i(n)$ is the average degree, so the expected increase in the number of unsaturated vertices is $\frac{1}{d}\sum_i i(i-2)\lambda_i(n)$. If this value is negative, then we expect that the branching process will die out rather quickly. If it is positive, then there is a chance that the number of unsaturated vertices will just continue to grow, so that a large component is generated.

### 1.3. Main result

The structure of random graphs with an asymptotic degree sequence $\mathscr{D}$ such that $Q(\mathscr{D}) = 0$ depends on how fast the quantity $\sum_{i\geqslant 1} i(i-2)\lambda_i(n)$ converges to 0. In order to study this, we define a generating function in the variable $x$,

$$Q_n(x) = \sum_{i\geqslant 1} i(i-2)\lambda_i(n)x^i.$$

If $\lambda_i^* > 0$ for some $i \geqslant 3$, then for sufficiently large $n$, $Q_n''(x) = \sum_{i\geqslant 3} i^2(i-1)(i-2)\lambda_i(n)x^{i-2}$ is positive when $x > 0$, and hence the function $Q_n(x)$ is strictly convex on the interval $[0,\infty)$ and therefore has at most two zeros on this interval. Let $\tau_n$ be the largest value such that

$$Q_n(\tau_n) = 0. \tag{1.2}$$

Note that if $\lim_{n\to\infty} \tau_n = 1$, then $Q(\mathscr{D}) = 0$, and further that if $Q(\mathscr{D}) = 0$, then $\lambda_1^* > 0$ if and only if $\lambda_i^* > 0$ for some $i \geqslant 3$. In the rest of the paper we will make the assumptions that $\lim_{n\to\infty} \tau_n = 1$ and $\lambda_1^* > 0$.

We define the generating function

$$\Lambda_n(x) = \sum_{i\geqslant 0} \lambda_i(n)x^i. \tag{1.3}$$

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

Then $Q_n(x)$ can be written as

$$Q_n(x) = x^2\Lambda_n''(x) - x\Lambda_n'(x). \tag{1.4}$$

Note that since $\mathscr{D}$ is well-behaved, $\Lambda'(1)$, $\Lambda''(1)$ and $\Lambda_n'''(1)$ are all bounded as $n \to \infty$. The average degree of $G_n$ equals $\Lambda'(1)$ and is denoted by $d$.

We will prove the following theorem, which is comparable to the study on the critical phase of $G(n,p)$ by Bollobás [3] and Łuczak [8] (see Section 6).

**Theorem 1.2.** *Assume that $\mathscr{D}$ is a well-behaved asymptotic degree sequence, such that for some $\varepsilon > 0$, $d_i(n) = 0$ whenever $i > n^{1/4-\varepsilon}$. Furthermore assume that $\lim_{n\to\infty} \tau_n = 1$ and $\lambda_1^* > 0$. Let*

$$\delta_n = 1 - \tau_n. \tag{1.5}$$

(a) *If $\delta_n n^{1/3} \to -\infty$, then a.a.s. all components in $G_n$ have $o(n^{2/3})$ vertices.*
(b) *There is a constant $c_1$ such that if $\delta_n n^{1/3} \geqslant c_1 \log n$, then a.a.s. $G_n$ has a single component of order $\gg n^{2/3}$, while all other components have order $o(n^{2/3})$.*

Note that $Q(\mathscr{D}) = \lim_{n\to\infty} Q_n(1)$. Since $Q(\mathscr{D})$ is the quantity used in the statement of Theorem 1.1, it may be interesting to see what Theorem 1.2 says about $Q_n(1)$. The quantities $\delta_n$ and $Q_n(1)$ are asymptotically related by

$$Q_n(1) \sim \delta_n Q_n'(1), \tag{1.6}$$

*i.e.*, $\lim_{n\to\infty} \frac{Q_n(1)}{\delta_n Q_n'(1)} = 1$, where $Q_n'(1)$ converges to a positive constant when $n \to \infty$. (This can be shown using Lemma 4.1 later on.) The quantity $Q_n(1)n^{1/3}$ is analogous to the parameter $\lambda$ in the critical phase of $G(n,p)$, which is given in (1.1); see Section 6 for details. Our theorem is weaker than the corresponding theorem for $G(n,p)$, by Łuczak [8], in the sense that we require a lower bound on how quickly $Q_n(1)n^{1/3}$ tends to infinity.

In the theorem we use the quantity $\delta_n$ rather than $Q_n(1)$, since it occurs naturally in the proof, and it therefore simplifies the notation to state the theorem in terms of $\delta_n$.

In order to prove Theorem 1.2, we will model the order of the components by a branching process introduced in Section 3. In Section 4 we study the behaviour of the branching process, using singularity analysis of the probability generating function associated with the branching process. Then the proof of Theorem 1.2 follows in Section 5.

## 2. Random configurations

It is difficult to study random graphs with a given degree sequence directly. Instead it has become customary to take the route via random configurations. The configuration model was introduced by Bender and Canfield [2], and later studied by Bollobás [4] and Wormald [14].

We define a random configuration $\mathscr{C}_n$ with a given degree sequence $\mathscr{D}_n$, as follows. Let $\mathscr{D}_n = \{a_1, \ldots, a_n\}$. We let $v_1, \ldots, v_n$ be vertices, and let $L_n$ be a set consisting of $a_i$ distinct

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

copies of the vertex $v_i$ for $i = 1, \dots, n$. These copies are called *half-edges*. We let $\mathscr{C}_n$ be equal to $L_n$, together with a perfect matching $\mathscr{P}_n$ of $L_n$, chosen uniformly at random.

A random perfect matching can be constructed greedily: at every step we take an arbitrary, unmatched half-edge, and match it with another half-edge chosen uniformly at random from the remaining half-edges. Using this procedure, every perfect matching has the same probability of being generated.

Given a configuration $\mathscr{C}_n$ on $n$ vertices, we let $G_n^*$ be the multigraph obtained by identifying all copies of $v_i$ with one another, for every $i = 1, \dots, n$, and letting the pairs of the perfect matching in $\mathscr{C}_n$ become edges.

Assume that $d_i(n) = 0$ for all $i > n^{1/4-\varepsilon}$ for some $\varepsilon > 0$, and that $\mathscr{D}_n$ otherwise satisfies the conditions of Theorem 1.2. As observed in [11], the main result of [10] implies that $G_n^*$ is a simple graph with probability tending to $e^{-\nu(\mathscr{D}_n)}$, for some $\nu(\mathscr{D}_n) = O(n^{1/2-\varepsilon})$. If $Q_n(1)$ is bounded, then $\nu(\mathscr{D}_n)$ tends to a constant. This implies the following lemma.

**Lemma 2.1.** *If the underlying multigraph $G_n^*$ of a random configuration $\mathscr{C}_n$ with a given degree sequence $\mathscr{D}_n$ meeting the conditions of Theorem 1.2 has a property $P$ a.a.s., and if $\sup_n Q_n(1) < \infty$, then a random graph $G_n$ with the same degree sequence has $P$ a.a.s.*

## 3. Branching process

We use a branching process in order to study the order of the components in the underlying multigraph $G_n^*$ of the random configuration defined in Section 2.

Consider the set $L_n$, and suppose that the perfect matching $\mathscr{P}_n$ has not been chosen yet. We start by choosing a single pair uniformly at random. We want to determine the order of the component containing this edge, and we will do this by *exposing* the pairs of the perfect matching $\mathscr{P}_n$. Suppose that the pair we choose contains the two half-edges $v_1$ and $v_2$. Then we say that $v_1$ and $v_2$ are *exposed*, while all other half-edges are *unexposed*. A vertex $v$ (consisting of one or more half-edges) is *unexposed* if none of its half-edges are exposed, *partially exposed* if some, but not all, of its half-edges are exposed, and *fully exposed* if all its half-edges have been exposed.

The process of exposing the component containing the pair $v_1v_2$ goes on as follows. At every step we choose an unexposed half-edge $w_1$, randomly or otherwise, in any partially exposed vertex, if such a vertex exists. Then we choose another half-edge $w_2$, chosen uniformly at random from all unexposed half-edges in $L_n$ distinct from $w_1$. Then we add the pair $w_1w_2$ to the matching and say that $w_1$ and $w_2$ are exposed.

When there is no partially exposed vertex left in the configuration, we stop the process. The component containing the edge $v_1v_2$ is then fully exposed. The set of exposed vertices forms a connected component in the underlying multigraph.

We will model this process as a branching process, where the particles in the branching process are edges. An edge consists of two half-edges, which we will call the *upper* and *lower* half-edge. In the branching process an edge gets $i$ edges as children, for $i = 0, 1, 2, \dots$, with probability $\frac{(i+1)\lambda_{i+1}(n)}{d}$, where we recall that $d = \sum i\lambda_i(n)$ is the average degree. This equals the probability that a randomly chosen half-edge is a part of a vertex of degree $i+1$. We will interpret the branching process such that the lower half-edge of an edge,

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

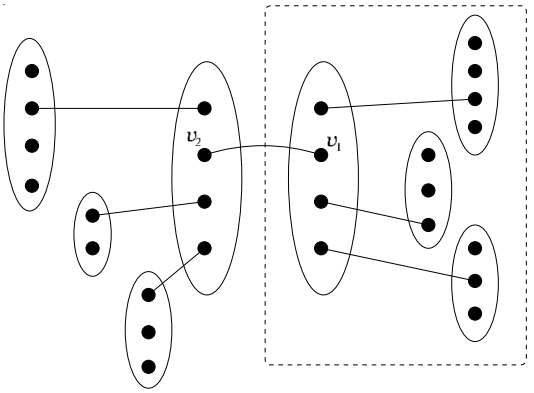



*Figure 1.* The first step of the branching process

together with the upper half-edges of all its children, comprise one vertex in the random graph.

The branching process starts with a single edge, $v_1v_2$, which is special in that we consider both the half-edges $v_1$ and $v_2$ to be lower half-edges. Thus the branching process starts off with two branches which continue independently of each other. Figure 1 shows the situation after the first step of the branching process.

We let $B_n$ be the random variable denoting the number of *vertices* produced in the branching process before it dies out. If the branching process does not die out, but continues forever, then $B_n = \infty$.

Since the branching process starts with two independent branches, it will be convenient to consider the corresponding branching process, which starts with one edge $v_1v_2$, but only lets the branching process continue from one of the half-edges, say $v_1$, as in the boxed part of Figure 1. Let $\beta_n$ be the random variable denoting the number of edges produced in this process, including the original edge. The total number of edges in the original branching process is then $\beta_n^{(1)} + \beta_n^{(2)} - 1$, where $\beta_n^{(1)}$ and $\beta_n^{(2)}$ are independent random variables with the same distribution as $\beta_n$. The number of vertices produced in the process is then $B_n = \beta_n^{(1)} + \beta_n^{(2)}$.

We let $p_n(z)$ be the probability generating function for the number of children of an edge in the branching process. Then

$$p_n(z) = \sum_{i\geqslant 0} p_i(n) z^i = \sum_{i\geqslant 0} \frac{(i+1)\lambda_{i+1}(n) z^i}{\sum_j j\lambda_j(n)} = \frac{\Lambda_n'(z)}{\Lambda_n'(1)}. \tag{3.1}$$

The following is a classical theorem for branching processes. For its proof, see [1].

**Theorem 3.1.** *Let $p(x)$ be the probability generating function for the number of children produced by one particle in a branching process. If the branching process starts with a single particle, then the extinction probability of the branching process is the smallest non-negative root of the equation*

$$p(x) = x.$$

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

*Let $\mu$ be the expected number of children produced by one particle. If $\mu \leqslant 1$, then the extinction probability is* 1 *(except for certain trivial cases). Otherwise, the extinction probability is strictly less than* 1.

The expected number of children of an edge is $p_n'(1)$. It then follows from Theorem 3.1 that if $p_n'(1) \leqslant 1$, then $\mathbb{P}[\beta_n < \infty] = \mathbb{P}[B_n < \infty] = 1$.

Let $e$ be any edge in the branching process. Let $q_k(n)$ be the probability that the total number of descendant edges of $e$, including $e$ itself, is exactly $k$. Let $q_n(z) = \sum q_k(n) z^k$ be the corresponding probability generating function. The probability that $e$ has $i$ children is $p_i(n)$, so

$$q_n(z) = z \sum_{i \geqslant 0} p_i(n) q_n(z)^i = z p_n(q_n(z)). \tag{3.2}$$

Since $B_n = \beta_n^{(1)} + \beta_n^{(2)}$, the probability generating function for $B_n$ is $q_n(z)^2$.

There are two difficulties which are not taken account of by the branching process, compared with $G_n$. The first problem is that in the random graph $G_n$, or more precisely in the underlying multigraph $G_n^*$ of the random configuration, a half-edge may choose to form an edge with a half-edge in a vertex which already contains exposed half-edges, which causes a cycle to be formed. We will show that this happens sufficiently seldom that the branching process is a good enough approximation. The second problem is that when some vertices are already partially exposed, the probability that a new half-edge is in a vertex of degree $i$ generally deviates from $\frac{i\lambda_i(n)}{d}$ in random configurations. However, we will see in Lemma 4.6 that this deviation is also small enough, and that the branching process is a good approximation.

## 4. Analysis of generating functions

In this section we will study the behaviour of the branching process. In particular we want to calculate the probability $\mathbb{P}[B_n \geqslant k]$ for various values of $k$. The event that $B_n \geqslant k$ can happen in two ways: either the branching process dies out after it has produced $k$ vertices, or it continues forever. Hence

$$\mathbb{P}[B_n \geqslant k] = \mathbb{P}[k \leqslant B_n < \infty] + \mathbb{P}[B_n = \infty]. \tag{4.1}$$

In order to calculate this quantity, we will have to study the behaviour of various functions close to 1, and we will use the following observation repeatedly.

**Lemma 4.1.** *Let $f_n(x) = \sum_{i \geqslant 0} \alpha_{ni} x^i$, where $\alpha_{ni}$ are real numbers for $n \geqslant 1$ and $i \geqslant 0$. Assume that $\alpha_{ni} \to \alpha_i^*$ as $n \to \infty$ for constants $\alpha_i^*$ for all $i \geqslant 0$. Let $r(n) = \max(\{i : \alpha_{ni} \neq 0\} \cup \{0\})$. Let $\{a_n\}_{n \geqslant 0}$ and $\{b_n\}_{n \geqslant 0}$ be sequences of real numbers such that $a_n \to 1$ and $b_n \to 0$. Assume that $r(n) b_n = o(1)$. Then, as $n \to \infty$,*

$$f_n(a_n + b_n) = f_n(a_n) + b_n f_n'(a_n) + \frac{1}{2} b_n^2 f_n''(a_n) + O(b_n^3). \tag{4.2}$$

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

**Proof.**

$$\begin{aligned} f_n(a_n + b_n) &= \sum_{i\geqslant 0} \alpha_{ni}(a_n + b_n)^i \\ &= \sum_{i\geqslant 0} \left( \alpha_{ni} \sum_{j=0}^{i} \binom{i}{j} a_n^{i-j} b_n^j \right) \\ &= \sum_{i\geqslant 0} \alpha_{ni} \left( a_n^i + i a_n^{i-1} b_n + \frac{1}{2} i(i-1) a_n^{i-2} b_n^2 + O(b_n^3) \right) \\ &= f_n(a_n) + b_n f_n'(a_n) + \frac{1}{2} b_n^2 f_n''(a_n) + O(b_n^3). \end{aligned}$$

□

We are then able to calculate the second summand of (4.1), the probability that the branching process continues forever, and we find that it is proportional to $\delta_n$ in the supercritical phase, and 0 in the subcritical phase.

**Lemma 4.2.** *If* $\tau_n \downarrow 1$, *then* $\mathbb{P}[B_n = \infty] = 0$. *If* $\tau_n \uparrow 1$. *Then*

$$\mathbb{P}[B_n = \infty] \sim 4\delta_n. \tag{4.3}$$

**Proof.** If $\tau_n \downarrow 1$, the extinction probability is 1 by Theorem 3.1, so assume that $\tau_n \uparrow 1$. We let $\eta_n = \mathbb{P}[\beta_n = \infty]$. Theorem 3.1 implies that $p_n(1-\eta_n) = 1-\eta_n$, so by (3.1) and Lemma 4.1,

$$(1-\eta_n)\Lambda_n'(1) = \Lambda_n'(1-\eta_n) \overset{(4.2)}{=} \Lambda_n'(1) - \eta_n \Lambda_n''(1) + \frac{1}{2}\eta_n^2 \Lambda_n'''(1) + O(\eta_n^3),$$

which gives us

$$\Lambda_n''(1) - \Lambda_n'(1) = \frac{1}{2}\eta_n \Lambda_n'''(1) + O(\eta_n^2). \tag{4.4}$$

The left-hand side of this equation is $Q_n(1)$. Since $Q_n(\tau_n) = 0$ by assumption and $\delta_n \to 0$, we can use Lemma 4.1 to calculate this value. We note that $Q_n(x)$ and $Q_n'(x)$ can be written as

$$\begin{aligned} Q_n(x) &= x^2 \Lambda_n''(x) - x\Lambda_n'(x), \\ Q_n'(x) &= x^2 \Lambda_n'''(x) + x\Lambda_n''(x) - \Lambda_n'(x) \\ &\overset{(1.4)}{=} x^2 \Lambda_n'''(x) + \frac{Q_n(x)}{x}. \end{aligned} \tag{4.5}$$

Thus, by Lemma 4.1 and (1.5),

$$\begin{aligned} Q_n(\tau_n) &= Q_n(1) - \delta_n Q_n'(1) + \delta_n^2 Q_n''(1) + O(\delta_n^3) \\ &\overset{(4.5)}{=} Q_n(1) - \delta_n(\Lambda'''(1) + Q_n(1)) + O(\delta_n^2), \end{aligned}$$

so

$$\begin{aligned} Q_n(1) &= \frac{1}{1-\delta_n}\left(\delta_n \Lambda'''(1) + O(\delta_n^2)\right) \\ &= \delta_n \Lambda'''(1) + O(\delta_n^2). \end{aligned} \tag{4.6}$$

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

Equations (4.4) and (4.6) imply that

$$\delta_n + O(\delta_n^2) = \frac{1}{2}\eta_n + O(\eta_n^2),$$

so

$$\mathbb{P}[\beta_n = \infty] \sim 2\delta_n. \tag{4.7}$$

Since $B_n = \infty$ if and only if $\beta_n^{(1)} = \infty$ or $\beta_n^{(2)} = \infty$, equation (4.3) follows. □

We then have to find $\mathbb{P}[k \leqslant B_n < \infty]$, and for this we will use singularity analysis of generating functions. The key to this approach is the following theorem, which follows from [7, Theorem VI.6]. A function $\phi(w)$ is *periodic* if, for some $d \geqslant 2$ and function $\psi$, we have $\phi(w) = \psi(w^d)$.

**Theorem 4.3.** *Let $\phi(w)$ be a function analytic at 0, having non-negative Taylor coefficients with $\phi(0) \neq 0$, such that there exists a positive solution $\tau$ to the characteristic equation,*

$$\phi(\tau) - \tau\phi'(\tau) = 0,$$

*strictly within the disc of convergence of $\phi$. Let $y(z)$ be the solution analytic at the origin of $y(z) = z\phi(y(z))$. Then $y(z)$ has a dominant singularity at $z = \rho$, where $\rho = \frac{\tau}{\phi(\tau)}$. If furthermore $\phi''(\tau) \neq 0$ and $\phi(w)$ is aperiodic, then the coefficients of $y(z)$ satisfy*

$$[z^i]y(z) \sim c\rho^{-i}i^{-\frac{3}{2}}\big(1 + O(i^{-1})\big),$$

*where $c = \sqrt{\frac{\phi(\tau)}{2\pi\phi''(\tau)}}$.*

We recall that the functions $p_n(x)$ and $q_n(x)$ are related by (3.2). In view of Theorem 4.3, taking $\phi = p_n$ and $y = q_n$, we define $\tau_n$ to be such that

$$p_n(\tau_n) - \tau_n p_n'(\tau_n) = 0,$$

which is equivalent to (1.2). This can also be expressed as

$$\frac{\Lambda_n'(\tau_n)}{\tau_n\Lambda_n''(\tau_n)} = 1. \tag{4.8}$$

We then define

$$\rho_n = \frac{\tau_n}{p_n(\tau_n)},$$

which is the radius of convergence and the location of the dominant singularity of $p_n(x)$. The next lemma gives us a relation between $\rho_n$ and $\delta_n$.

**Lemma 4.4.** *Assume that $\tau_n \to 1$, and let $\rho_n$ and $\delta_n$ be as before. Then*

$$\log \rho_n \sim c_2\delta_n^2,$$

*for a constant $c_2 = \frac{1}{2} + \frac{1}{2}\lim_{n\to\infty} \frac{\Lambda_n'''(1)}{\Lambda_n'(1)} > 0$.*

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

**Proof.** Applying Lemma 4.1 to (1.5), we see that for all suitable functions $f_n(x)$,

$$f_n(\tau_n) = f_n(1) - \delta_n f_n'(1) + \frac{1}{2}\delta_n^2 f_n''(1) + O(\delta_n^3). \tag{4.9}$$

Furthermore, using the equation

$$\log(1+x) = x - \frac{x^2}{2} + O(x^3), \tag{4.10}$$

we get

$$\begin{aligned}
\log f_n(\tau_n) &\overset{(4.9)}{=} \log\left(f_n(1) - \delta_n f_n'(1) + \frac{\delta_n^2}{2} f_n''(1) + O(\delta_n^3)\right) \\
&= \log f_n(1) + \log\left(1 - \frac{\delta_n f_n'(1)}{f_n(1)} + \frac{\delta_n^2 f_n''(1)}{2f_n(1)} + O(\delta_n^3)\right) \\
&\overset{(4.10)}{=} \log f_n(1) - \delta_n \frac{f_n'(1)}{f_n(1)} + \frac{\delta_n^2}{2}\left(\frac{f_n''(1)}{f_n(1)} - \frac{f_n'(1)^2}{f_n(1)^2}\right) + O(\delta_n^3).
\end{aligned} \tag{4.11}$$

The function $\Lambda_n'(x)$ satisfies the conditions of Lemma 4.1. Hence,

$$\begin{aligned}
\log \rho_n &\overset{(4.8)}{=} \log \frac{\tau_n \Lambda_n'(1)}{\Lambda_n'(\tau_n)} \\
&= \log \tau_n + \log \Lambda_n'(1) - \log \Lambda_n'(\tau_n) \\
&\overset{(4.11)}{=} \delta_n\left(\frac{\Lambda_n''(1)}{\Lambda_n'(1)} - 1\right) + \frac{\delta_n^2}{2}\left(\frac{\Lambda_n''(1)^2}{\Lambda_n'(1)^2} - \frac{\Lambda_n'''(1)}{2\Lambda_n'(1)} - 1\right) + O(\delta_n^3).
\end{aligned} \tag{4.12}$$

Because of (4.8) and the fact that $\tau_n \to 1$, we might expect that $\Lambda_n''(1)/\Lambda_n'(1)$ is close to 1. Indeed, applying Lemma 4.1 to the function $x\Lambda_n''(x)$, we get

$$\Lambda_n''(1) = \tau_n \Lambda_n''(\tau_n) + \delta_n(\Lambda_n''(1) + \Lambda_n'''(1)) + O(\delta_n^2),$$

and using instead the function $\Lambda_n'(x)$, we get

$$\Lambda_n'(1) = \Lambda_n'(\tau_n) + \delta_n \Lambda_n''(1) + O(\delta_n^2).$$

Thus

$$\begin{aligned}
\frac{\Lambda_n''(1)}{\Lambda_n'(1)} &= \frac{\tau_n \Lambda_n''(\tau_n) + \delta_n(\Lambda_n''(1) + \Lambda_n'''(1)) + O(\delta_n^2)}{\Lambda_n'(\tau_n) + \delta_n \Lambda_n''(1) + O(\delta_n^2)} \\
&= \frac{1}{\Lambda_n'(\tau_n)} \frac{\tau_n \Lambda_n''(\tau_n) + \delta_n(\Lambda_n''(1) + \Lambda_n'''(1)) + O(\delta_n^2)}{1 + \delta_n \frac{\Lambda_n''(1)}{\Lambda_n'(\tau_n)} + O(\delta_n^2)} \\
&= \frac{1}{\Lambda_n'(\tau_n)} \left(\tau_n \Lambda_n''(\tau_n) + \delta_n(\Lambda_n''(1) + \Lambda_n'''(1)) + O(\delta_n^2)\right)\left(1 - \delta_n \frac{\Lambda_n''(1)}{\Lambda_n'(\tau_n)} + O(\delta_n^2)\right) \\
&= \frac{1}{\Lambda_n'(\tau_n)} \left(\tau_n \Lambda_n''(\tau_n) + \delta_n\left(\Lambda_n''(1) + \Lambda_n'''(1) - \tau_n \Lambda_n''(\tau_n)\frac{\Lambda_n''(1)}{\Lambda_n'(\tau_n)}\right) + O(\delta_n^2)\right).
\end{aligned}$$

Using (4.8), and the fact that $\Lambda_n(\tau_n) = \Lambda_n(1) + O(\delta)$,

$$\begin{aligned}
\frac{\Lambda_n''(1)}{\Lambda_n'(1)} &= 1 + \delta_n\left(\frac{\Lambda_n''(1) + \Lambda_n'''(1)}{\Lambda_n'(\tau_n)} - \frac{\Lambda_n''(1)}{\Lambda_n'(1)}\right) + O(\delta_n^2) \\
&= 1 + \delta_n \frac{\Lambda_n'''(1)}{\Lambda_n'(1)} + O(\delta_n^2),
\end{aligned}$$

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

so by (4.12),

$$\log \rho_n = \frac{\delta_n^2}{2}\left(1 + \frac{\Lambda_n'''(1)}{\Lambda_n'(1)}\right) + O(\delta_n^3). \tag{4.13}$$

□

**Lemma 4.5.** *The probability that one branch of the branching process dies out after producing at least $k$ vertices is*

$$\mathbb{P}[k \leqslant \beta_n < \infty] \sim c_3 e^{-c_2 k \delta_n^2} k^{-1/2}, \tag{4.14}$$

*where $c_3 > 0$ is a constant, and $c_2 > 0$ is as in Lemma 4.4.*

**Proof.** Recall that $q_k(n)$ is the probability that the branching process dies out after precisely $k$ vertices have been produced. According to Theorem 4.3,

$$q_k(n) \sim c\rho_n^{-k} k^{-3/2}(1 + O(k^{-1})),$$

for a constant $c$. Hence

$$\begin{aligned}
\mathbb{P}[k \leqslant \beta_n < \infty] &= \sum_{i \geqslant k} q_i(n) \sim \sum_{i \geqslant k} c\rho_n^{-i} i^{-3/2}(1 + O(i^{-1})) \\
&\sim c \int_k^\infty \rho_n^{-x} x^{-3/2}(1 + O(x^{-1}))dx \\
&\sim c\left[-\rho_n^{-x} x^{-1/2}\left(2 + \frac{4}{3}\log \rho_n + o(x^{-1})\right)\right]_k^\infty \\
&= c\rho_n^{-k} k^{-1/2}\left(2 + \frac{4}{3}\log \rho_n + o(k^{-1})\right).
\end{aligned}$$

Now (4.14) follows from Lemma 4.4. □

Lemma 4.5 tells us the probability that one branch of the branching process dies out after $k$ vertices have been created. The complete branching process has two branches, which produce $\beta_n^{(1)}$ and $\beta_n^{(2)}$ vertices respectively. We have

$$[k \leqslant \beta_n^{(1)} < \infty] \wedge [\beta_n^{(2)} < \infty] \Rightarrow [k \leqslant B_n < \infty]$$

and

$$[k \leqslant B_n < \infty] \Rightarrow [k/2 \leqslant \beta_n^{(1)} < \infty] \vee [k/2 \leqslant \beta_n^{(2)} < \infty].$$

Hence we get the lower bound

$$\begin{aligned}
\mathbb{P}[k \leqslant B_n < \infty] &\overset{(4.7)}{\geqslant} \mathbb{P}[k \leqslant \beta_n^{(1)} < \infty](1 - 2\delta_n) \\
&\sim c_3 e^{-c_2 k \delta_n^2 (1 + o(1))} k^{-1/2},
\end{aligned} \tag{4.15}$$

and the upper bound

$$\begin{aligned}
\mathbb{P}[k \leqslant B_n < \infty] &\leqslant 2\mathbb{P}[k/2 \leqslant \beta_n < \infty] \\
&\leqslant 3c_3 e^{-\frac{1}{2} c_2 k \delta_n^2 (1 + o(1))} k^{-1/2}.
\end{aligned} \tag{4.16}$$

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

Let $C_n$ be the random variable denoting the number of vertices in the component containing a random edge in $G_n^*$. We will show that $C_n$ and $B_n$ behave very similarly.

**Lemma 4.6.** *Let $\alpha_n$ and $\gamma_n$ be such that $k = \gamma_n n^{2/3}$ and $\delta_n = \alpha_n n^{-1/3}$. Suppose that $\gamma_n \ll |\alpha_n|$. Then there are constants $c_4, c_5, c_6, c_7$ such that, for large enough $n$,*

$$\mathbb{P}[C_n \geqslant k] \leqslant c_4 e^{-c_5 k \delta_n^2 (1+o(1))} k^{-1/2} + I_{\delta_n > 0} 4\delta_n (1 + o(1)) \tag{4.17}$$

*and*

$$\mathbb{P}[C_n \geqslant k] \geqslant c_6 e^{-c_7 k \delta_n^2 (1+o(1))} k^{-1/2} + I_{\delta_n > 0} 4\delta_n (1 + o(1)). \tag{4.18}$$

*Let $k_- = \gamma_n^- n^{2/3}$ and $k_+ = \gamma_n^+ n^{2/3}$, where $\gamma_n^- \leqslant \gamma_n^+ \ll \alpha_n$. Then there is a positive constant $c_8$ such that, for large enough $n$,*

$$\mathbb{P}[k_- < C_n < k_+] \leqslant c_8 e^{-c_2 \alpha^2 \gamma_- (1+o(1))} k_-^{-1/2}. \tag{4.19}$$

**Proof.** There are two problems which can cause $B_n$ and $C_n$ to differ. The first is the fact that in the random graph cycles can be formed, whereas this does not happen in the ordinary branching process. The second problem is that in the branching process the probability that a vertex has, say, $i$ children remains the same throughout the process. When exposing the component in the random graph, this is not true, since it depends on how many vertices of degree $i+1$ we have exposed so far. We will show that both of these factors have a negligible effect as long as the number of vertices exposed is not too large.

We first consider the possibility of cycles being formed. Suppose that $k$ vertices have already been exposed in $\mathscr{C}$. We first choose a half-edge $w_1$ in a partially exposed vertex, and then a half-edge $w_2$ uniformly at random from all unexposed half-edges. The probability that $w_2$ is in a partially exposed vertex is then $O\left(\frac{k}{n}\right)$.

On the other hand, let $X_{m,i}$ denote the number of vertices of degree $i$ among the first $m$ vertices exposed. These $m$ vertices are picked at random from the total of $n$ vertices, with every vertex being chosen with probability proportional to its degree. Since we consider the case that $m$ is asymptotically small compared to $n$, the distribution of $X_{m,i}$ approaches a binomial distribution $\mathrm{Bin}\left(m, \frac{i\lambda_i(n)}{d}\right)$ when $n$ tends to infinity. Let $b_m$ be the degree of the $m$th vertex. We obtain the upper bound

$$\begin{aligned}
\mathbb{P}[b_m = i] &= \frac{ni\lambda_i(n) - iX_{m,i}}{dn - \sum_j jX_{mj}} \\
&\leqslant \frac{i\lambda_i(n)}{d} \frac{1}{1 - \frac{1}{dn}\sum_j jX_{m,j}} \\
&= \frac{i\lambda_i(n)}{d}\left(1 + O\left(\frac{k}{n}\right)\right),
\end{aligned}$$

when $m \leqslant k$. The probability that the $m$th vertex exposed has degree $i$ is therefore $(1 + o(1))\frac{i\lambda_i(n)}{d}$ when $m$ is small. Hence,

$$\mathbb{E}[X_{m,i}] = (1 + o(1))\frac{mi\lambda_i(n)}{d},$$

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

while the standard deviation is roughly $\sigma_{m,i} = \sqrt{m\frac{i\lambda_i(n)}{d}\left(1-\frac{i\lambda_i(n)}{d}\right)}$. Using Chernoff bounds, we can show that there is a constant $c$, such that, with probability $1-o(n^{-2})$, we have

$$|X_{m,i} - \mathbb{E}[X_{m,i}]| \leqslant c\sigma_{m,i}\sqrt{\log n}. \tag{4.20}$$

Hence, with probability $1-o(n^{-1})$ (4.20) holds for all $i = 1,\dots,\psi(n)$, where $\psi(n)$ is the maximum degree. We can therefore condition on (4.20) holding, for all $i = 1,\dots,\psi(n)$ and $m = o(n)$.

Let $K_n$ be the set of integers $i$ with $1 \leqslant i \leqslant \psi(n)$ such that $m\frac{i\lambda_i(n)}{d} \geqslant \log^2 n$. By (4.20), for every $i \in K_n$, $X_{m,i} = (1+o(1))\mathbb{E}[X_{m,i}] = (1+o(1))\frac{mi\lambda_i(n)}{d}$, and for every $i \notin K_n$, $X_{m,i} \leqslant c'\log n$ for some constant $c'$. We therefore have

$$\begin{aligned}\sum_i iX_{m,i} &= \sum_{i\in K_n}(1+o(1))\frac{mi^2\lambda_i(n)}{d} + O\left(\sum_{i\notin K_n}\log^2 n\right)\\ &= (1+o(1))\frac{m}{d}\sum_i i^2\lambda_i(n) + O\left(\psi(n)\log^2 n\right)\\ &= (1+o(1))\frac{m}{d}(\Lambda''(1)+\Lambda'(1)).\end{aligned}$$

Then we can also find a lower bound on the probability that the $m$th vertex has degree $i$:

$$\begin{aligned}\mathbb{P}[b_m = i | i \in K_n] &= \frac{ni\lambda_i(n) - iX_{m,i}}{dn - \sum_j jX_{m,j}}\\ &= \frac{i\lambda_i(n)}{d}\frac{1-\frac{(1+o(1))mi}{nd}}{1-\frac{m}{d^2n}(\Lambda''(1)+d)(1+o(1))}\\ &= \frac{i\lambda_i(n)}{d}\left(1+O\left(\frac{k}{n}\right)\right),\end{aligned} \tag{4.21}$$

for $1 \leqslant m \leqslant k$, whereas

$$\begin{aligned}\mathbb{P}[b_m = i | i \notin K_n] &= \frac{ni\lambda_i(n) - iX_{m,i}}{dn - \sum_j jX_{m,j}}\\ &\geqslant \frac{ni\lambda_i(n) - \frac{mi\lambda_i(n)}{d} - c\sqrt{\frac{mi\lambda_i(n)}{d}\log n}}{dn}\\ &\geqslant \frac{ni\lambda_i(n) - c'\log^2 n}{dn}\\ &= \frac{i\lambda_i(n)}{d} - \frac{c'\log^2 n}{dn}.\end{aligned}$$

We will denote the branching process defined in Section 3 by $\mathscr{B}$. Thus we may say that if we use the branching process $\mathscr{B}$ to approximate the component exposure process $\mathscr{C}$, then at every step there is a chance that we choose the 'wrong' degree. We therefore introduce a modified branching process $\overline{\mathscr{B}}$ as follows. In $\mathscr{B}$ an edge gets $i$ children with probability $p_i = \frac{(i+1)\lambda_{i+1}(n)}{d}$. If $i \in K_n$, then an edge in $\overline{\mathscr{B}}$ gets $i$ children with probability $\overline{p}_i = p_i(1+\varepsilon_i)$, where $\varepsilon_i = \varepsilon_i(n)$ are error terms depending on $n$, such that $|\varepsilon_i| \ll |\delta|$. If $i \notin K_n$, then an edge in $\overline{\mathscr{B}}$ gets $i$ children with probability $\overline{p}_i = p_i + \varepsilon_i$, where again $\varepsilon_i = \varepsilon_i(n)$ depends on $n$, and $|\varepsilon_i| \leqslant \frac{c'\log^2 n}{n}$. Note that if $i \notin K_n$, then the error term is not relative to $p_i$. By choosing

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

the functions $\varepsilon_i$ appropriately, we can make sure that the process $\overline{\mathscr{B}}$ becomes either an upper bound or a lower bound of the exposure process $\mathscr{C}$, and we will show that as long as the stated bounds are satisfied, the behaviour of $\overline{\mathscr{B}}$ does not differ much from the behaviour of $\mathscr{B}$. We then obtain the probability generating function for the number of children of an edge in the modified branching process as

$$\begin{aligned}\overline{p}_n(z) &= \sum_i \overline{p}_i z^i = \sum_{i\in K_n} p_i(1+\varepsilon_i)z^i + \sum_{i\notin K_n}(p_i+\varepsilon_i)z^i \\ &= p_n(z) + \sum_{i\in K_n} p_i\varepsilon_i z^i - \sum_{i\notin K_n}\varepsilon_i z^i.\end{aligned}$$

We are interested in the behaviour of $\overline{p}_n(z)$ when $z \in (1-2\delta, 1+2\delta)$. In this interval $z^{\psi(n)} = 1 + o(1)$ since $\psi(n) = o(n^{1/4})$. Hence

$$\left|\sum_{i\in K_n} p_i\varepsilon_i z^i\right| = o(\delta)\left|\sum_{i\in K_n} p_i z^i\right| = o(\delta),$$

and

$$\left|\sum_{i\notin K_n}\varepsilon_i z^i\right| \leqslant (1+o(1))\psi(n)\frac{c'\log n}{dn} = o(n^{-3/4}).$$

We can therefore write $\overline{p}_n(z) = p_n(z) + c_n(z)$, where $c_n(z)$ is a function such that $|c_n(z)| = o(\delta)$ when $z \in (1-2\delta, 1+2\delta)$. Likewise we can show that $|\sum_{i\in K_n} ip_i\varepsilon_i z^{i-1}| = o(\delta_n)$ and $|\sum_{i\notin K_n} i\varepsilon_i z^i| = o(n^{-1/2})$, so that also $|c_n'(z)| = o(\delta)$. We let $\overline{\tau}_n$ be defined such that $\overline{p}_n(\overline{\tau}_n) - \overline{\tau}_n\overline{p}_n'(\overline{\tau}_n) = 0$, and let $\overline{\delta}_n = 1 - \overline{\tau}_n$. Let $\zeta_n = \overline{\tau}_n - \tau_n$. Then

$$\begin{aligned}0 &= \overline{p}_n(\overline{\tau}_n) - \overline{\tau}_n\overline{p}_n'(\overline{\tau}_n) \\ &= \big(p_n(\overline{\tau}_n) + c_n(\overline{\tau}_n)\big) - \overline{\tau}_n\big(p_n'(\overline{\tau}_n) + c_n'(\overline{\tau}_n)\big) \\ &= p_n(\tau_n+\zeta_n) - (\tau_n+\zeta_n)p_n'(\tau_n+\zeta_n) + o(\delta_n) \\ &= p_n(\tau_n) + \zeta_n p_n'(\tau_n) + O(\zeta_n^2) - (\tau_n+\zeta_n)(p_n'(\tau_n) + \zeta_n p_n''(\tau_n) + O(\zeta_n^2)) + o(\delta_n) \\ &= -\zeta_n p_n''(1) + O(\zeta_n^2) + o(\delta_n),\end{aligned}$$

by Lemma 4.1. Hence $\zeta_n = o(\delta_n)$, so $\overline{\delta}_n \sim \delta_n$. Let $\overline{B}_n$ be the random variable denoting the number of vertices generated in $\overline{\mathscr{B}}$, and let $\overline{q}_n(z)$ be the corresponding probability generating function. Then $\overline{q}_n(z)$ is given implicitly by $\overline{q}_n(z) = z\overline{p}_n(\overline{q}_n(z))$. According to Theorem 4.3, the dominant singularity of $\overline{q}_n(z)$ is $\overline{\rho}_n = \frac{\overline{\tau}_n}{\overline{p}_n(\overline{\tau}_n)}$. Lemma 4.4 states that $\log\rho_n \sim c_2\delta_n^2$; we can similarly calculate that $\log\overline{\rho}_n \sim c_2\delta^2$. Hence, by Theorem 4.3,

$$\overline{q}_k \sim c\overline{\rho}_n^{-k}k^{-3/2} = ce^{-c_2k\log\rho_n}k^{-3/2} = ce^{-c_2k\delta_n^2(1+o(1))}k^{-3/2}, \tag{4.22}$$

where $c$ is a positive constant. Furthermore, as in Lemma 4.2, the probability that $\overline{\mathscr{B}}$ does not die out is 0 if $\delta_n < 0$, and if $\delta_n > 0$, then $\mathbb{P}[\overline{B}_n = \infty] \sim 4\overline{\delta}_n \sim 4\delta_n$.

Instead of approximating $\mathscr{C}$ by the branching process $\mathscr{B}$, we will approximate it by $\overline{\mathscr{B}}$, choosing the functions $\varepsilon_i(n)$ appropriately. In $\mathscr{C}$ the probability that the next vertex chosen has degree $i$ depends on the previously exposed vertices, while in $\overline{\mathscr{B}}$, the probabilities are the same all the time. In (4.21) we gave bounds for the deviations of the probabilities between $\mathscr{B}$ and $\mathscr{C}$. In the definition of $\overline{\mathscr{B}}$, we assumed that the deviations $\varepsilon_i$ are asymptotically smaller than $\delta$, while in (4.21) the deviations were found to be $O\left(\frac{k}{n}\right)$. By assumption,

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

$\frac{k}{n} = \gamma_n n^{-1/3} \ll \alpha_n n^{-1/3} = \delta_n$, so in $\overline{\mathscr{B}}$, the error terms $\varepsilon_i$ can be chosen in such a way that $\overline{\mathscr{B}}$ becomes either a lower bound of $\mathscr{C}$, or an upper bound. We can therefore derive (4.17) in the same manner as (4.16), and (4.18) in the same manner as (4.15). As for (4.19), we use (4.22) to obtain

$$\begin{aligned}\mathbb{P}[k_- < C_n < k_+] &= \sum_{k=k_-}^{k_+} \mathbb{P}[C_n = k] \sim \int_{k_-}^{k_+} ce^{-c_2 x\delta^2(1+o(1))} x^{-3/2} dx \\ &\leqslant 2ce^{-c_2 k_- \delta^2(1+o(1))} k_-^{-1/2} \\ &= 2ce^{-c_2\alpha^2\gamma_-(1+o(1))} k_-^{-1/2}.\end{aligned}$$

□

## 5. The phase transition

Using the lemmas of the previous section, we can now prove Theorem 1.2.

### 5.1. The subcritical case

We first consider case (1) of Theorem 1.2, that $\alpha_n = \delta_n n^{1/3} \to -\infty$. We want to show that the largest component in $G_n$ a.a.s. has $o(n^{2/3})$ vertices.

Let $k = cn^{2/3}$ for some constant $c$. Since clearly $c \ll |\alpha_n|$, (4.17) implies that, for large enough $n$,

$$\mathbb{P}[C_n \geqslant k] \leqslant \frac{c_4}{\sqrt{c}} e^{-cc_5\alpha_n^2} n^{-1/3}, \tag{5.1}$$

where $c_4, c_5 > 0$. Let $X_k$ be the number of vertices in components of order greater than $k$, and let $A_k$ be the event that there is a component of order at least $k$. Then, by (5.1) and Markov's inequality,

$$\mathbb{P}[A_k] = \mathbb{P}[X_k \geqslant k] \leqslant \frac{\mathbb{E}[X_k]}{k} \sim \frac{n\mathbb{P}[C_n \geqslant k]}{k} \leqslant \frac{c_4}{c^{3/2}} e^{-cc_5\alpha_n^2} \to 0.$$

There is therefore a.a.s. no component in $G_n^*$ with more than $cn^{2/3}$ vertices, for every positive constant $c$. This, together with Lemma 2.1, completes the proof.

### 5.2. The supercritical case

Now we consider the supercritical phase, when $\alpha_n = \delta_n n^{1/3} \geqslant c_1 \log n$. We call a component *large* if it has $\gg n^{2/3}$ vertices and small if it has $o(n^{2/3})$ vertices. We will prove firstly that a.a.s. every component is either large or small, secondly that there is a.a.s. at least one large component, and thirdly that there is a.a.s. only one large component.

**Lemma 5.1.** *Let $\omega(n)$ be a function which tends to infinity as $n \to \infty$, but such that $\omega(n) = o(\log n)$. There is a constant $c_9$, such that, if $\alpha_n \geqslant c_9 \log n$, then the probability that $G_n^*$ contains a component of order between $k_- = n^{2/3}/\log n$ and $k_+ = n^{2/3}\omega(n)$ is $O(n^{-1})$.*

**Proof.** If $v$ is a vertex, we let $C(v)$ be the component containing $v$. Then, according to (4.19),

$$\begin{aligned}\mathbb{P}[k_- \leqslant |C(v)| \leqslant k_+] &\leqslant c_8 e^{-c_2\alpha_n^2/\log n} k_-^{-1/2} \\ &\leqslant c_8 n^{-c_2 c_9^2} n^{-1/3} \sqrt{\log n}.\end{aligned}$$

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

Let $X$ be the number of vertices contained in components of order between $k_-$ and $k_+$, and let $A$ be the event that there is at least one such component. Then, by Markov's inequality,

$$\mathbb{P}[A] = \mathbb{P}[X \geqslant k_-] \leqslant \frac{\mathbb{E}[X]}{k_-} = \frac{n\mathbb{P}[k_- \leqslant |C(v)| \leqslant k_+]}{k_-} \leqslant c_8 n^{-c_2 c_9^2} \log^{3/2} n.$$

Clearly $c_9$ can be chosen so large that $\mathbb{P}[A] = O(n^{-1})$. □

We assume that $c_1$ in Theorem 1.2 satisfies $c_1 \geqslant c_9$. Thus, by Lemma 5.1 we know that a.a.s. every component in $G_n^*$ is either large or small. We will now show that there is a.a.s. at least one large component in $G_n^*$. We will do this by considering a subgraph $H$ of $G_n^*$, which can also be viewed as a random graph with degree sequence $\mathscr{D}_0'$, where $\mathscr{D}_0'$ is some asymptotic degree sequence.

Let us consider the configuration model $\mathscr{C}_n$. We know that the perfect matching $\mathscr{P}_n$ can be constructed greedily. We will construct it in the following way: for some $m$, we first choose $m$ pairs of half-edges randomly, and label them $e_1, \dots, e_m$. Let $L_0' = L_n \setminus \bigcup_{i=1}^m e_i$ be the set of unmatched half-edges, and let $\mathscr{D}_0'$ be the degree sequence of $L_0'$. Then we choose a random perfect matching, $\mathscr{P}_0'$, of $L_0'$, and let $\mathscr{C}_0'$ be the random configuration consisting of the set $L_0'$ of half-edges and the matching $\mathscr{P}_0'$. Let $H$ be the underlying multigraph of $\mathscr{C}_0'$. Then $H$ is a subgraph of $G_n^*$.

Each of the pairs of half-edges $e_i$, with $1 \leqslant i \leqslant m$ is chosen in the following way. We first choose a half-edge uniformly at random from the set of yet unmatched half-edges in vertices of degree at least 3. The second half-edge is chosen uniformly at random from all unmatched half-edges. This is a valid way to generate the random matching $\mathscr{P}_n$, since, as we stated in Section 2, the first half-edge in every pair can be chosen in an arbitrary manner, as long as the second half-edge is chosen uniformly at random. Moreover, we will make sure that we always have $m = o(n)$, and since, by assumption, $\lambda_i^* > 0$ for some $i \geqslant 3$, the set of vertices of degree at least 3 will not be exhausted.

Recall that $\alpha_n = \delta_n n^{1/3}$. We write $\alpha_n = \alpha(\mathscr{D}_n)$, such that $\alpha$ is a function of the degree sequence $\mathscr{D}_n$. In the process explained in the previous paragraph we do not fix $m$ beforehand, but we choose the pairs $e_1, e_2, \dots$ one by one and remove them from $L_n$. Whenever an edge is removed from $L_n$ in this manner, the value of $Q_n(z)$ decreases. Indeed, by assumption, at least one of the end-vertices of $e_i$, for any $i$, has degree 3 or greater. In the case that the end-vertices have degree 1 and 3 respectively, the value of $Q_n(1)$ decreases by $\frac{2}{n}$. Any other combination of degrees causes $Q_n(1)$ to decrease by a greater amount.

We know from (1.6) (which follows from Lemma 4.1) that $Q_n(1)$ and $\delta$ are proportional. Hence $\delta$ and $\alpha$ similarly decrease whenever a pair of half-edges is removed from $L_n$. It is clear that removing sufficiently many pairs of half-edges will cause $Q_n(1)$, and thereby $\alpha$, to become negative. We will continue to remove pairs of half-edges until the degree sequence $\mathscr{D}'$ of the remaining set is such that $\alpha(\mathscr{D}') \leqslant \log^{3/8} n$. Since the maximum degree is less than $n^{1/4}$ by assumption, the value of $\alpha$ decreases by at most $O(n^{-3/4})$ for every

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

pair removed; hence we will have $\alpha_0' := \alpha(\mathscr{D}') \sim \log^{3/8} n$. The graph $H$ is then a random multigraph with degree sequence $\mathscr{D}_0'$.

Set $\gamma_- = \log^{-9/10} n$ and $\gamma_+ = \log^{2/5} n$, and let $k_- = n^{2/3}\gamma_-$ and $k_+ = n^{2/3}\gamma_+$. We will show that $H$ a.a.s. contains a component of order at least $k_-$. This implies that $G_n^*$ a.a.s. contains a component of order at least $k_-$, and Lemma 5.1 then implies that $G_n^*$ a.a.s. has a component of order asymptotically greater than $n^{2/3}$.

We let $C_0'$ be the random variable denoting the order of the component containing a specified vertex of $H$. We let $A_k'$ be the event that there is at least one component in $H$ of order at least $k$, and we let $X_k'$ be the number of vertices which are contained in such components. Then, using (4.17), we get

$$\mathbb{P}[C_0' \geqslant k_+] \leqslant c_4 e^{-c_5\gamma_+(\alpha_0')^2(1+o(1))}k_+^{-1/2} + 4\alpha_0' n^{-1/3}(1+o(1)) \sim 4\alpha_0' n^{-1/3}.$$

Thus

$$\mathbb{P}[A_{k_+}] = \mathbb{P}[X_{k_+}' \geqslant k_+] \leqslant \frac{n}{k_+}\mathbb{P}[C_0' \geqslant k_+] \leqslant \frac{4\alpha_0'}{\gamma_+} = 4\log^{-1/40} n = o(1).$$

We can therefore condition on $\overline{A_{k_+}}$, namely that $H$ does not contain any components of order $k_+$ or greater.

By (4.18),

$$\begin{aligned}
\mathbb{E}[X_{k_-}'] &= n\mathbb{P}[C_0' \geqslant k_-] \\
&\geqslant n\big(c_6 e^{-c_7\gamma_-(\alpha_0')^2(1+o(1))}k^{-1/2} + 4\alpha_0' n^{-1/3}(1+o(1))\big) \\
&\sim n^{2/3}\big(c_6\gamma_-^{-1/2}e^{-c_7(1+o(1))\log^{-0.15} n} + 4\log^{3/8} n\big) \\
&\sim n^{2/3}\big(c_6\log^{9/20} n + 4\log^{3/8} n\big) \sim c_6 n^{2/3}\log^{9/20} n.
\end{aligned}$$

Suppose that $v$ is a vertex in a component of order at least $k_-$. Since we condition on $\overline{A_{k_+}}$, the component containing $v$ has at most $k_+$ vertices. The expected number of vertices $w \neq v$, such that $w$ is also contained in a component of order at least $k_-$, is therefore bounded by $k_+ + \mathbb{E}[X_{k_-}']$. Since $k_+ \ll \mathbb{E}[X_{k_-}']$,

$$\begin{aligned}
\mathbb{E}[X_{k_-}'(X_{k_-}' - 1)] &\leqslant \mathbb{E}[X_{k_-}']\big(k_+ + \mathbb{E}[X_{k_-}']\big) \\
&= \mathbb{E}[X_{k_-}']^2(1+o(1)),
\end{aligned}$$

so by Chebyshev's inequality, a.a.s. $X_{k_-}' \sim \mathbb{E}[X_{k_-}']$. We conclude that $H$, and thereby $G_n^*$, a.a.s. contains a component of order at least $k_-$. Hence, according to Lemma 5.1, $G_n^*$ a.a.s. contains at least one large component.

It remains to prove that there is just one such component in $G_n^*$. We will do this by deleting a certain number of edges in the same way as in the previous step, but now we fix $m = \lceil n^{2/3}\log n\rceil$. Furthermore, this time we will choose the edges $e_1, \dots, e_m$ uniformly at random. We let as before $L_0' = L \setminus \bigcup_{i=1}^m e_i$, and $\mathscr{P}_0'$ be a random perfect matching of $L_0'$. For $1 \leqslant i \leqslant m$, we let $L_i' = L_{i-1}' \cup e_i$, and let $\mathscr{D}_i'$ be the degree sequence of $L_i'$. Furthermore we let $\mathscr{P}_i'$ be the perfect matching of $L_i'$ consisting of the pairs in $\mathscr{P}_0'$ together with the pairs $e_1, \dots, e_i$. We let $H_i$ be the underlying multigraph of the configuration $\mathscr{C}_i'$, which consists of $L_i'$ and $\mathscr{P}_i'$. Then $H_i$ can be considered a random graph with given degree sequence $\mathscr{D}_i'$, and $G_n^* = H_m$.

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

By assumption $\alpha_n = \alpha(\mathscr{D}'_m) \geqslant c_1 \log n$. We will first show that $\alpha'_0 = \alpha(\mathscr{D}'_0) \asymp \alpha_n$, *i.e.*, $\alpha'_0 = \Theta(\alpha_n)$; $\alpha'_0$ and $\alpha_n$ are of the same order of magnitude. Let $\lambda'_i(n)$ be the proportion of vertices in $H_0$ having degree $i$. Let $Q_n(x) = \sum_i i(i-2)\lambda_i(n)x^i$ and $Q_0(x) = \sum_i i(i-2)\lambda'_i(n)x^i$. Recall that $\tau_n = 1 - \delta_n$, and that $Q_n(\tau_n) = 0$. We let $\tau'_0$ be a number such that $Q_0(\tau'_0) = 0$, and we let $\delta'_0 = 1 - \tau'_0$. We have to show that $\delta'_0 \asymp \delta_n$.

We let $D(x) = Q(x) - Q_0(x) = \sum_i i(i-2)(\lambda_i(n) - \lambda'_i(n))x^i$. Suppose first that the difference between the degree sequence of $H_m = G^*_n$, and the degree sequence of $H_0$ is that precisely one vertex has degree $d$ in $H_0$, but degree $d+1$ in $H_m$, while all other vertices have unchanged degree. (Ignore for a moment that this is impossible.) Then

$$\sum_i i(i-2)(\lambda_i(n) - \lambda'_i(n))\tau_n^i = \frac{1}{n}(-d(d-2)\tau_n^d + (d+1)(d-1)\tau_n^{d+1})$$
$$= \frac{\tau_n^d}{n}(2d - 1 + (1-d^2)\delta_n).$$

Now we remember that the difference between $H_m$ and $H_0$ is more substantial, namely that $m$ edges have been added to $H_0$ to obtain $H_m$. Let $m_i$ be the number of times any vertex gets its degree increased from $i$ to $i+1$ in this process. Then

$$D(\tau_n) = \sum_i i(i-2)(\lambda_i(n) - \lambda'_i(n))\tau_n^i$$
$$= \frac{1}{n}\sum_i (2i - 1 + (1-i^2)\delta_n)m_i\tau_n^i,$$

and $2m = \sum_i m_i$. Since $m = n^{2/3}\log n$, $\tau_n \to 1$ and $\delta_n \to 0$, we see that $D(\tau_n) \sim c_{10}\frac{m}{n} = c_{10}n^{-1/3}\log n$ for a constant $c_{10} > 0$. Hence

$$Q_0(\tau_n) = Q(\tau_n) - D(\tau_n) = -c_{10}n^{-1/3}\log n. \tag{5.2}$$

Let $\Delta\tau = \tau_n - \tau'_0$. By Lemma 4.1,

$$Q_0(\tau_n) = Q_0(\tau'_0) + \Delta\tau Q'_0(\tau'_0) + \Delta\tau^2 Q''_0(\tau'_0) + O(\Delta\tau^3). \tag{5.3}$$

Equations (5.2) and (5.3) imply that

$$\Delta\tau Q'_0(\tau'_0) + O(\Delta\tau^2) = -c_{10}n^{-1/3}\log n,$$

so $\Delta\tau \asymp -n^{-1/3}\log n$. Since the difference between $\tau_n$ and $\tau'_0$ is in the order of $n^{-1/3}\log n$, we get that $\alpha'_0 \sim (c_1 - c_{10})\alpha_n$. We assume that $c_1$ is so large that $c_1 - c_{10} \geqslant c_9$, where $c_9$ is the constant in Lemma 5.1.

Let $\alpha'_i = \alpha(\mathscr{D}'_i)$. Since $\alpha'_{i+1} \geqslant \alpha'_i$, we have $\alpha'_i \geqslant c_9 \log n$ for all $i = 0, \ldots, m$. Thus, by Theorem 5.1, the probability that there is a component with between $n^{2/3}/\log n$ and $n^{2/3}$ in any of the graphs $H_0, \ldots, H_m$ is bounded by $n^{2/3-1}\log n = o(1)$. It follows that every large component in $H_i$ for $i = 1, \ldots, m$ must contain some large component in $H_{i-1}$, and hence every large component in $H_m$ must contain at least one of the large components in $H_0$.

Let $C_1, \ldots, C_l$ be the large components in $H_0$. We must prove that these components a.a.s. are contained in one component in $G^*_n$. Recall that $m$ edges were removed from $G^*_n$ to obtain $H_0$. We will show that for every pair $(C_i, C_j)$ of large components in $H_0$, it is very likely that one of the edges removed from $G^*_n$ joins two vertices in $C_i$ and $C_j$ to each other.

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

Let $E = \{e_1, \ldots, e_m\}$, and let $M$ be the set of vertices in $G_n^*$ which are incident to one of the edges in $E$. Suppose that $v$ is a vertex of degree $i$ in $G_n^*$. The probability that $e_j$ is incident to $v$ is $\frac{i}{dn}$. It follows that the expected number of vertices of degree $i$ which are incident to one of the edges in $E$ is $m\frac{i\lambda_i(n)}{d}$. The expected number of edges in $H_0$ which are incident to a vertex in $M$ is then $\sum_i (i-1)m\frac{i\lambda_i(n)}{d} = \frac{m}{d}\sum_i i(i-1)\lambda_i(n)$. Because $\mathscr{D}$, and therefore also $\mathscr{D}_0$, is well-behaved, the sum tends to a constant, so the expected number of edges in $H_0$ incident to a vertex in $M$ is $c_{11}m$ for some constant $c_{11} > 0$.

Now let $e$ be any edge in $H_0$, and suppose we expose the component containing $e$ as explained in Section 3. At any point in the exposure process the probability that the next vertex is a vertex in $M$ is $\frac{c_{11}m}{dn} \asymp n^{-1/3}\log n$. Let $\omega(n) \to \infty$ be such that each of the components $C_1, \ldots, C_l$ contains at least $n^{2/3}\omega(n)$ vertices. The expected number of vertices in $M$ among the first $n^{2/3}\omega(n)$ vertices exposed in any component is $\Theta(n^{1/3}\omega(n)\log n)$. Furthermore, the distribution of the number of such vertices tends to a binomial distribution $\mathrm{Bin}(n^{2/3}\omega(n), \frac{c_{11}}{d}n^{-1/3}\log n)$ as $n \to \infty$, so we assert that for any $i = 1, \ldots, l$, with probability $1 - o(n^{-1})$, the number of vertices in $M \cap C_i$ is at least $c_{12}n^{1/3}\omega(n)\log n$ for some constant $c_{12} > 0$.

There can be at most $n^{1/3}$ large components in $H_0$, so a.a.s. each of these components has at least $c_{12}n^{1/3}\omega(n)\log n$ vertices from $M$. Consider two components $C_i$ and $C_j$ in $H_0$ with $1 \leqslant i, j \leqslant l$, and let $v$ be a vertex in $M \cap C_i$. The probability that one of the edges in $E$ has $v$ as one endpoint and its other endpoint in $C_j$ is $\frac{|M \cap C_j|}{|M|} \geqslant \frac{c_{12}n^{1/3}\log n}{2n^{2/3}\log n} \asymp n^{-1/3}$. The probability that none of the vertices in $M \cap C_i$ is adjacent to a vertex in $M \cap C_j$ is therefore

$$\left(1 - \frac{1}{n^{1/3}}\right)^{c_{12}n^{1/3}\omega(n)\log n} \sim e^{-c_{12}\omega(n)\log n} = o(n^{-1}).$$

Since $l \leqslant n^{1/3}$, the expected number of pairs of the components $C_i$, $C_j$ with $1 \leqslant i, j \leqslant l$, which are *not* connected by one of the $m$ edges, is then at most $n^{2/3}n^{-1} = o(1)$. Hence, a.a.s. all the components have joined to form a single component. This, together with Lemma 2.1, concludes the proof of Theorem 1.2.

## 6. Concluding remarks

Theorem 1.2 implies that the critical phase of $G_n$ is similar to the critical phase in the standard random graph model, $G(n,p)$. The parameter $\delta_n n^{1/3}$ in Theorem 1.2 plays the same role for $G_n$ as $\lambda$ in (1.1) does for $G(n,p)$. Indeed, the method we have used in this paper also works for $G(n,p)$. Suppose that $p = \frac{c}{n} = \frac{1}{n} + \frac{\varepsilon}{n}$, where $\varepsilon = \varepsilon(n) \to 0$ as $n \to \infty$. It is well known that the components can be modelled using a branching process, where a vertex gets $i$ children with probability $p_i = \frac{e^{-c}c^i}{i!}$. The corresponding probability generating function is $p(x) = \sum p_i x^i = e^{c(x-1)}$. If we let $q(x)$ be the probability generating function for the number of vertices generated before the branching process dies out, we get $q(x) = xp(q(x))$. We let $\tau$ and $\rho$ be as in Theorem 4.3. Then $\tau = \frac{1}{c}$, and $\rho = \frac{1}{c}e^{1-c}$, so

$$\log \rho = \varepsilon - \log(1+\varepsilon) = \varepsilon^2/2 - \varepsilon^3/3 + \cdots.$$

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press

From (4.13) we get

$$\frac{\delta_n^2}{2}\left(1+\frac{\Lambda_n'''(1)}{\Lambda_n'(1)}\right)+O(\delta_n^3)=\log\rho_n=\frac{\varepsilon^2}{2}-\frac{\varepsilon^3}{3}+\cdots.$$

We set $\varepsilon=\lambda n^{-1/3}$ to make it comparable to (1.1). Then $\delta_n$ and $\lambda$ satisfy

$$\delta_n n^{1/3}\asymp \varepsilon n^{1/3}=\lambda.$$

In part (b) of Theorem 1.2 we have a lower bound on $\delta_n n^{1/3}$; however, we believe that it holds whenever $\delta_n n^{1/3}\to\infty$, as for the critical phase in $G(n,p)$.

For further references on the component structure in random graphs with a given degree sequence, see Chung and Lu [5], Łuczak [9] and Newman, Strogatz and Watts [13].

## References

[1] Athreya, K. B. and Ney, P. E. (1972) *Branching Processes*, Vol. 196 of *Die Grundlehren der Mathematischen Wissenschaften*, Springer, New York.
[2] Bender, E. A. and Canfield, E. R. (1978) The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory Ser. A* **24** 296–307.
[3] Bollobás, B. (1984) The evolution of random graphs. *Trans. Amer. Math. Soc.* **286** 257–274.
[4] Bollobás, B. (2001) *Random Graphs*, 2nd edn, Vol. 73 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge.
[5] Chung, F. and Lu, L. (2002) Connected components in random graphs with given expected degree sequences. *Ann. Comb.* **6** 125–145.
[6] Erdős, P. and Rényi, A. (1960) On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** 17–61.
[7] Flajolet, P. and Sedgewick, R. (2006) *Analytic Combinatorics*, book in preparation.
[8] Łuczak, T. (1990) Component behavior near the critical point of the random graph process. *Random Struct. Alg.* **1** 287–310.
[9] Łuczak, T. (1992) Sparse random graphs with a given degree sequence. In *Random Graphs*, Vol. 2 (Poznań, 1989), Wiley–Interscience, pp. 165–182.
[10] McKay, B. D. (1985) Asymptotics for symmetric 0–1 matrices with prescribed row sums. *Ars Combin.* **19** 15–25.
[11] Molloy, M. and Reed, B. (1995) A critical point for random graphs with a given degree sequence. In *Proc. Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, 'Random Graphs '93'* (Poznań, 1993), Vol. 6, pp. 161–179.
[12] Molloy, M. and Reed, B. (1998) The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.* **7** 295–305.
[13] Newman, M. E. J., Strogatz, S. H. and Watts, D. J. (2001) Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64** 026118.
[14] Wormald, N. C. (1978) Some problems in the enumeration of labelled graphs. PhD thesis, Newcastle University.

https://doi.org/10.1017/S096354830700867X Published online by Cambridge University Press