Combinatorics, Probability and Computing (2011) 20, 131–154. © Cambridge University Press 2010
doi:10.1017/S0963548310000180

The Largest Component in Subcritical Inhomogeneous Random Graphs

TATYANA S. TUROVA[†]

Mathematical Center, University of Lund, Box 118, Lund S-221 00, Sweden
(e-mail: tatyana@maths.lth.se)

Received 13 November 2008; revised 4 May 2010; first published online 9 June 2010

We study the 'rank 1 case' of the inhomogeneous random graph model. In the subcritical case we derive an exact formula for the asymptotic size of the largest connected component scaled to $\log n$. This result complements the corresponding known result in the supercritical case. We provide some examples of applications of the derived formula.

1. Introduction

1.1. Inhomogeneous random graphs

We shall study a subclass of the general inhomogeneous random graph model $G^{\mathcal{V}}(n,\kappa)$ with vertex space

$$\mathcal{V} = \big(\mathcal{S}, \mu, (x_1^{(n)}, \dots, x_n^{(n)})_{n\geqslant 1}\big),$$

as was introduced by Bollobás, Janson and Riordan in [1]. Here $\mathcal{S}$ is a separable metric space and μ is a Borel probability measure on $\mathcal{S}$. No relationship is assumed between $x_i^{(n)}$ and $x_i^{(n')}$. To simplify notation we shall from now on write $(x_1,\dots,x_n) = (x_1^{(n)},\dots,x_n^{(n)})$. We begin by recalling some basic definitions and assumptions from [1]. For each n let $(x_1,\dots,x_n)$ be a deterministic or random sequence of points in $\mathcal{S}$, such that, for any μ-continuity set $A \subseteq \mathcal{S}$,

$$\frac{\#\{i : x_i \in A\}}{n} \xrightarrow{P} \mu(A), \tag{1.1}$$

where $\xrightarrow{P}$ denotes convergence in probability. Given the sequence $x_1,\dots,x_n$, we let $G^{\mathcal{V}}(n,\kappa)$ be the random graph on $\{1,\dots,n\}$, such that any two vertices i and j are connected by an edge independently of the others with probability

$$p_{ij}(n) = \min\{\kappa(x_i,x_j)/n, 1\}, \tag{1.2}$$

[†] Research supported by the Swedish Natural Science Research Council.

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

where κ is a symmetric non-negative measurable kernel defined on $\mathcal{S} \times \mathcal{S}$. Furthermore, we assume that the kernel κ is *graphical* on $\mathcal{V}$, which means that

(i) κ is continuous a.e. on $\mathcal{S} \times \mathcal{S}$;
(ii) $\kappa \in L^1(\mathcal{S} \times \mathcal{S}, \mu \times \mu)$;
(iii)

$$\frac{1}{n}\mathbb{E}e\big(G^{\mathcal{V}}(n,\kappa)\big) \to \frac{1}{2}\int_{\mathcal{S}^2} \kappa(x,y)d\mu(x)d\mu(y),$$

where $e(G)$ denotes the number of edges in a graph G.

Let $C_1(G)$ denote the size of the largest connected component in a graph G. Recall that in the case of a homogeneous random graph $G_{n,p}$ with $p = c/n$, Erdős and Rényi [4] derived the following convergence in probability result, along with even more precise ones: if $c < 1$ then

$$\frac{C_1(G_{n,c/n})}{\log n} \xrightarrow{P} \frac{1}{c - 1 + |\log c|}, \tag{1.3}$$

while if $c > 1$ then

$$\frac{C_1(G_{n,c/n})}{n} \xrightarrow{P} \beta, \tag{1.4}$$

where β is the largest positive root of $\beta = 1 - e^{-c\beta}$.

Karp [7] observed that a random graph can be naturally related to a certain branching process underlying the algorithm to find the connected components in the graph. This idea was extended in [9] for certain inhomogeneous graph models, where a multi-type branching process was introduced to study connectivity. It was not until the work of Bollobás, Janson and Riordan [1] that a comprehensive theory of inhomogeneous random graphs was developed, which provided a unified approach to many models studied previously.

Karp showed in [7] that the constant β in (1.4) equals the survival probability of the associated branching process. Here is the corresponding result that was established by Bollobás, Janson and Riordan [1] for the general model $G^{\mathcal{V}}(n,\kappa)$ described above. Assume, in addition, that κ is irreducible, which means (see [1, Definition 2.10])

$$A \subseteq \mathcal{S} \text{ and } \kappa = 0 \text{ a.e. on } A \times (\mathcal{S} \setminus A) \text{ implies } \mu(A) = 0 \text{ or } \mu(\mathcal{S} \setminus A) = 0.$$

Then Theorem 3.1 from [1] states that

$$\frac{C_1\big(G^{\mathcal{V}}(n,\kappa)\big)}{n} \xrightarrow{P} \rho_\kappa := \int_{\mathcal{S}} \rho_\kappa(x)d\mu(x), \tag{1.5}$$

where $\rho_\kappa(x)$ is the survival probability of the multi-type Galton–Watson process $B_\kappa(x)$ defined as follows. The type space of $B_\kappa(x)$ is $\mathcal{S}$, and initially there is a single particle of type $x \in \mathcal{S}$. Then, at any step, a particle of type $x \in \mathcal{S}$ is replaced in the next generation by a set of particles whose types are distributed as a Poisson process on $\mathcal{S}$ with intensity measure $\kappa(x,y)d\mu(y)$. Bollobás, Janson and Riordan [1] also proved that the function ρ_κ is the maximum solution to

$$\rho_\kappa(x) = 1 - e^{-\int_{\mathcal{S}} \kappa(x,y)\rho_\kappa(y)d\mu(y)}.$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Whether ρ_κ is zero or strictly positive depends only on the norm of the integral operator T_κ defined by

$$(T_\kappa f)(x) = \int_{\mathcal{S}} \kappa(x,y) f(y) d\mu(y) \tag{1.6}$$

with norm

$$\|T_\kappa\| = \sup\{\|T_\kappa f\|_2 : f \geqslant 0, \|f\|_2 \leqslant 1\}.$$

Bollobás, Janson and Riordan [1, Theorem 3.1] proved that the survival probability

$$\rho_\kappa \begin{cases} > 0 & \text{if } \|T_\kappa\| > 1, \\ = 0 & \text{if } \|T_\kappa\| \leqslant 1. \end{cases} \tag{1.7}$$

Whereas (1.5) describes rather well the size of the largest connected component above the phase transition, *i.e.*, when $\|T_\kappa\| > 1$, all that we can infer from (1.5) when $\|T_\kappa\| \leqslant 1$ is that $C_1\big(G^{\mathcal{V}}(n,\kappa)\big) = o_P(n)$. Only under the additional assumption that

$$\sup_{x,y} \kappa(x,y) < \infty \tag{1.8}$$

does Theorem 3.12 in [1] establish in the case $\|T_\kappa\| < 1$ that w.h.p.

$$C_1\big(G^{\mathcal{V}}(n,\kappa)\big) = O(\log n), \tag{1.9}$$

where w.h.p. means 'with high probability', *i.e.*, with probability tending to one as $n \to \infty$.

1.2. Main results

Our aim here is to describe the asymptotic behaviour of the size of the largest component scaled to $\log n$ (as in (1.3)) for the inhomogeneous random graph model in the case $\|T_\kappa\| < 1$. We show that this is also directly related to the parameters of the corresponding branching process B_κ.

Assume from now on that $\mathcal{S} \subseteq \{1,2,\ldots\}$ is finite or countable, μ is a probability measure on $\mathcal{S}$, and the kernel κ on $\mathcal{S} \times \mathcal{S}$ has the form

$$\kappa(x,y) = c\psi(x)\psi(y), \tag{1.10}$$

where c is a positive constant, and ψ is a strictly positive function on $\mathcal{S}$ such that

$$\sum_{x \in \mathcal{S}} \psi(x)\mu(x) < \infty.$$

In this case the operator T_κ defined in (1.6) has rank 1, and

$$\|T_\kappa\| = c \sum_{x \in \mathcal{S}} \psi^2(x)\mu(x). \tag{1.11}$$

We consider the graph $G^{\mathcal{V}}(n,\kappa)$ on a vertex space $\mathcal{V}$ which satisfies condition (1.1). Given $x_1,\ldots,x_n$, each possible edge ij, $i \neq j$, $1 \leqslant i,j \leqslant n$, is present with probability (1.2), *i.e.*,

$$p_{ij}(n) = \min\left\{\frac{c\psi(x_i)\psi(x_j)}{n}, 1\right\}. \tag{1.12}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Whether or not different edges are present are independent events. We shall be imposing some additional conditions on our model.

Assumption 1.1. *There is a monotone non-decreasing function* $\psi_0 : \mathcal{S} \to \mathbb{R}_+$, *and positive constants* $A_1 \leqslant A_2$ *such that*

$$A_1\psi_0(x) \leqslant \psi(x) \leqslant A_2\psi_0(x), \tag{1.13}$$

for all $x \in \mathcal{S}$, *and*

$$\sum_{x\in\mathcal{S}} e^{a\psi(x)}\mu(x) < \infty \tag{1.14}$$

for some positive a.

Assumption 1.2. *For all* $\varepsilon > 0$ *and all* $q > 0$,

$$\mathbb{P}\left\{\frac{\#\{i : x_i = x\}}{n} - \mu(x) \leqslant \varepsilon e^{q\psi(x)}\mu(x) \quad \textit{for all} \quad x \in \mathcal{S}\right\} \to 1 \tag{1.15}$$

as $n \to \infty$.

We shall provide below an example of a model with infinite $\mathcal{S}$ which satisfies Assumptions 1.1 and 1.2.

Remarks. **(1)** If $\mathcal{S}$ is finite, Assumption 1.1 trivially holds, and convergence (1.15) follows by (1.1).

(2) If $\mathcal{S}$ is infinite then convergence (1.15) holds (under condition (1.1)) when, for example, $\sum_{x\in\mathcal{S}} e^{-q\psi(x)} < \infty$ for all $q > 0$ and

$$\mathbb{E}\,\#\{i : x_i = x\} \leqslant (1 + o(1))n\mu(x),$$

where $o(1) \to 0$ as $n \to \infty$ uniformly in x.

(3) Roughly speaking, condition (1.15) together with (1.14) rules out w.h.p. the presence of vertices of too large type in $G^{\mathcal{V}}(n, \kappa)$. More precisely, in order to have vertices of (large) type x in the graph $G^{\mathcal{V}}(n, \kappa)$, condition (1.15) requires $n \geqslant \left((\varepsilon e^{q\psi(x)} + 1)\mu(x)\right)^{-1}$.

Let $B(x)$ be the branching process defined as above. It starts with one particle of type $x \in \mathcal{S}$, and then, at any step, a particle of type $x \in \mathcal{S}$ produces a $\mathrm{Po}\big(c\psi(x)\psi(y)\mu(y)\big)$ number of offspring of each type $y \in S$. Let $\mathcal{X}(x)$ denote the size of the total progeny of $B(x)$, and set

$$r(c) = \sup\left\{z \geqslant 1 : \sum_{x\in\mathcal{S}} \mu(x)\psi(x)\mathbb{E}z^{\mathcal{X}(x)} < \infty\right\}. \tag{1.16}$$

Note that $r(c)$ is well-defined since the sum is finite for $z = 1$ by (1.14). Also $r(c) < \infty$, since $\mathcal{X}(x)$ for any $x \in \mathcal{S}$ stochastically dominates the progeny of a single-type branching process with the distribution of the offspring $\mathrm{Po}(c \min_{x\in\mathcal{S}} \psi(x) \sum_{y\in\mathcal{S}} \psi(y)\mu(y))$.

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Theorem 1.3. *Let* $\kappa(x,y) = c\psi(x)\psi(y)$ *and set*

$$c^{cr} := \left(\sum_{x\in\mathcal{S}} \psi^2(x)\mu(x)\right)^{-1}. \tag{1.17}$$

Under Assumptions 1.1 and 1.2 we have

$$\frac{C_1\big(G^{\mathcal{V}}(n,\kappa)\big)}{\log n} \xrightarrow{P} \frac{1}{\log r(c)} \tag{1.18}$$

as $n\to\infty$. *Furthermore,* $r(c)$ *is a continuous function such that*

$$r(c)\begin{cases} > 1 & \text{if } c < c^{cr}, \\ = 1 & \text{if } c \geqslant c^{cr}. \end{cases} \tag{1.19}$$

Observe that due to (1.11) we have

$$\|T_\kappa\| < 1 \quad \Leftrightarrow \quad c < c^{cr}. \tag{1.20}$$

Hence, the statement of Theorem 1.3 is exactly complementary to (1.5) (under the conditions of Theorem 1.3), since

$$r(c) > 1 \ \text{ implies } \ \rho_\kappa = 0,$$

whereas $\rho_\kappa > 0$ implies $r(c) = 1$. Note, however, that when $c = c^{cr}$ then both $r(c) = 1$ and $\rho_\kappa = 0$, and neither of statements (1.5) or (1.18) provides substantial information.

It is possible that one can extend Theorem 1.3 to the case of non-countable $\mathcal{S}$ under similar assumptions, by replacing the sum by an integral with respect to μ. One may conjecture as well that a statement similar to (1.18) remains true without the rank 1 assumption.

The rank 1 case has proved to be versatile for applications. One may interpret $\psi(x)$ as the 'activity' of a vertex of type x. One particularly common choice of ψ is $\psi(x) = x$ on $\mathcal{S} = \{1,2,\ldots\}$. Here 'type x' can represent the degree of a node as in the study of Britton, Deijfen and Martin-Löf [2], or the size of a macro-vertex as defined in [10] (see also [1, Section 16.4] for other examples). A special feature of the rank 1 case is that it permits one to compute $r(c)$ in a rather closed form, as we shall see below.

Theorem 1.4. *Assume condition* (1.14) *is satisfied, and* $\inf_{x\in\mathcal{S}} \psi(x) > 0$. *Let* X *be a random variable with distribution* μ *on* $\mathcal{S}$, *and define a positive constant* M *by*

$$M = \sum_{k\in\mathcal{S}} \psi(k)\mu(k) = \mathbb{E}\psi(X).$$

If $c < c^{cr}$, *there exists a unique* $y > 1$ *which satisfies*

$$y = \frac{1}{cM}\,\frac{\mathbb{E}\psi(X)\exp\{cM(y-1)\psi(X)\}}{\mathbb{E}\psi^2(X)\exp\{cM(y-1)\psi(X)\}}, \tag{1.21}$$

and then

$$r(c) = \frac{1}{c\mathbb{E}\psi^2(X)\exp\{cM(y-1)\psi(X)\}}. \tag{1.22}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Remark. The conditions of Theorem 1.4 are satisfied under the assumptions of Theorem 1.3.

Note that Theorems 1.3 and 1.4 immediately yield (1.3). Indeed, in the case of a homogeneous model $G_{n,c/n}$ we have $\psi(X) \equiv 1$ in Theorem 1.4, trivially implying $y = 1/c$, which together with (1.22) and (1.18) gives (1.3). The result (1.18) is new for the inhomogeneous random graphs.

The recent study [10] indicated that an upper bound found there for the largest component in the subcritical case of a certain inhomogeneous random graph model should be the optimal one. We shall show here that the conjecture of [10] is a simple corollary of the following modification of Theorem 1.3.

Let L denote a component of the graph $G^{\mathcal{V}}(n,\kappa)$ defined above with κ as in (1.10). In some applications one studies a function of L of the form $\sum_{i\in L}\psi(x_i)$. Sometimes it is natural to call $\psi(x_i)$ the 'activity of vertex i'. In this case, $\sum_{i\in L}\psi(x_i)$ is viewed as the 'activity of the component L'. A similar characteristic of a graph, called the 'volume', has been treated by Chung and Lu [5].

Consider again the branching processes $B(x)$ defined above, which starts with one particle of type $x \in S$. Let $\widetilde{\mathcal{X}}(x)$ denote the set of all offspring of the branching processes $B(x)$. (Recall that previously we considered $\mathcal{X}(x) = |\widetilde{\mathcal{X}}(x)|$.) Let $\Phi(x)$ be the sum of the 'activities' of all offspring of the branching processes $B(x)$, including the initial particle, that is,

$$\Phi(x) = \sum_{i\in\widetilde{\mathcal{X}}(x)} \psi(v_i),$$

where $v_i \in \mathcal{S}$ is the type of particle i. We see that $\Phi(x)$ is equal in distribution to

$$\psi(x) + \sum_{y\in\mathcal{S}} \sum_{i=1}^{N_x(y)} \Phi_i(y), \tag{1.23}$$

where $N_x(y) \in \mathrm{Po}(c\psi(x)\psi(y)\mu(y))$ are independent for different x and y, and the random variables $\Phi_i(x), i \geqslant 1$, are independent and identically distributed as $\Phi(x)$, and are also independent for different values of x. Define also, similarly to (1.16),

$$\alpha(c) = \sup\left\{ z \geqslant 1 : \sum_{x\in\mathcal{S}} \mu(x)\psi(x)\mathbb{E}z^{\Phi(x)} < \infty \right\}. \tag{1.24}$$

Note that the sum is finite for $z = 1$ by (1.14), and also $\alpha(c) < \infty$, since

$$\Phi(x) \geqslant |\widetilde{\mathcal{X}}(x)| \min_{x\in\mathcal{S}} \psi(x) = \mathcal{X}(x) \min_{x\in\mathcal{S}} \psi(x),$$

and $r(c) < \infty$.

Theorem 1.5. *Let $\mathcal{L}$ denote the set of all connected components in $G^{\mathcal{V}}(n,\kappa)$. Under the conditions of Theorem 1.3 we have*

$$\frac{\max_{L\in\mathcal{L}} \sum_{i\in L} \psi(x_i)}{\log n} \xrightarrow{P} \frac{1}{\log \alpha(c)} \tag{1.25}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

as $n \to \infty$. Furthermore, $\alpha(c)$ is a continuous function such that

$$\alpha(c) \begin{cases} > 1 & \text{if } c < c^{cr}, \\ = 1 & \text{if } c \geqslant c^{cr}. \end{cases} \tag{1.26}$$

One can also find a formula for $\alpha(c)$ similar to that in Theorem 1.4.

1.3. Example

Let $G_N(p,c)$ be a graph with the set of vertices $B(N) := \{-N,\dots,N\}^d$ in $\mathbb{Z}^d$, $d \geqslant 1$, having two types of edges: the short-range edges, which connect each pair u and v with $|u-v|=1$ independently with probability p; and the long-range edges, which connect any pair of vertices independently with probability $c/|B(N)|$. By this definition, there can be zero, one or two edges between any two vertices in the graph $G_N(p,c)$. In the last case the edges are of different types. Assume that $0 \leqslant p < p_c(d)$, where $p_c(d)$ is the percolation threshold in dimension d. As shown in [10], this graph is naturally related to the rank 1 case described above. Consider the subgraph of $G_N(p,c)$ induced by the short-range edges only, which forms a purely bond percolation model. Let K_N denote the number of connected (by the short-range edges only) components in this subgraph, and let $\mathbf{X} = \{X_1, X_2, \dots, X_{K_N}\}$ denote the collection of all such components $X_i \subseteq B(N)$. Let C denote the open cluster containing the origin in the bond percolation model on $\mathbb{Z}^d$ where the probability of an edge is p. Recall that

$$\frac{K_N}{|B(N)|} \to \mathbb{E}\frac{1}{|C|} \tag{1.27}$$

a.s. and in L^1 as $N \to \infty$ (see, *e.g.*, [6]). Call each set X_i a *macro-vertex* i of type $|X_i|$.

Now, given a collection of clusters $\mathbf{X}$, introduce another graph $\widetilde{G}_N(\mathbf{X},p,c)$, whose vertices are the macro-vertices $\{1,2,\dots,K_N\}$. The probability that two (macro-)vertices i and j with $|X_i| = x_i$ and $|X_j| = x_j$ are connected is derived from the original model $G_N(p,c)$, that is,

$$p_{ij}(N) = 1 - \left(1 - \frac{c}{|B(N)|}\right)^{x_i x_j} =: \frac{\kappa_N(x_i,x_j)}{K_N}, \tag{1.28}$$

where K_N is random, and thus κ_N is random too. Clearly, the size of the largest connected component in $G_N(p,c)$ has the following representation:

$$C_1\big(G_N(p,c)\big) = \max_L \sum_{i \in L} |X_i|, \tag{1.29}$$

where the maximum is taken over all connected components L in $\widetilde{G}_N(\mathbf{X},p,c)$.

By definition (1.28) we have

$$\kappa_N(x,y) = cxy\frac{K_N}{|B(N)|} + O\left(c^2x^2y^2\frac{K_N}{|B(N)|^2}\right), \tag{1.30}$$

where the last term is uniformly small in x and y not too large. Then convergence (1.27) implies

$$\kappa_N(x,y) \xrightarrow{\text{a.s.}} c\mathbb{E}(|C|^{-1})xy =: \kappa(x,y), \tag{1.31}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

and also

$$\frac{\#\{1 \leqslant i \leqslant K_N : |X_i| = k\}}{K_N} \xrightarrow{\text{a.s.}} \mu(k),$$

where

$$\mu(k) = \frac{1}{\mathbb{E}(|C|^{-1})} \frac{\mathbb{P}\{|C| = k\}}{k}. \tag{1.32}$$

It follows from general results on percolation theory that the distribution of the cluster size $|C|$ for $p < p_c$ has an exponentially decaying tail. Therefore Assumption 1.1 is satisfied for $\psi(x) = x$, $\mathcal{S} = \{1, 2, \ldots\}$, and μ defined in (1.32). Furthermore, it follows from results in [10] that, for any $\varepsilon > 0$ and any $q > 0$,

$$\mathbb{P}\left\{ \frac{\#\{1 \leqslant i \leqslant K_N : X_i = x\}}{K_N} - \mu(x) \leqslant \varepsilon e^{q\psi(x)} \mu(x), \quad \text{for all} \quad x \in \mathcal{S} \right\} \to 1 \tag{1.33}$$

as $N \to \infty$, as Assumption 1.2 requires. We still cannot apply Theorems 1.3 and 1.5 directly to the model $\widetilde{G}_N(\mathbf{X}, p, c)$, since here the kernel κ_N is random and depends on N. We shall use a domination argument as follows. Observe that, due to (1.33), (1.27) and the exponential decay of μ, we have w.h.p. $\max_i |X_i| < (\log K_N)^2$. Also, by (1.30) and (1.27), for any $0 < \delta < 1$ we have w.h.p.

$$(1 - \delta)\kappa(x, y) \leqslant \kappa_N(x, y) \leqslant (1 + \delta)\kappa(x, y) \quad \text{for all} \quad x, y < (\log K_N)^2, \tag{1.34}$$

where $\kappa(x, y)$ is defined in (1.31) Hence, given $\mathbf{X}$ with $K_N = n$ and $\max_i |X_i| < (\log n)^2$, the size $C_1\big(\widetilde{G}_N(\mathbf{X}, p, c)\big)$ is stochastically dominated by the size of the largest component of the inhomogeneous random graph $G^{\mathcal{V}}(n, (1 + \delta)\kappa)$ with the vertex space

$$\mathcal{V} = \big(\mathcal{S}, \mu, (X_1, \ldots, X_n)\big), \tag{1.35}$$

where μ is defined by (1.32) and $\mathcal{S} = \{1, 2, \ldots\}$. In the same way, w.h.p. $C_1\big(\widetilde{G}_N(\mathbf{X}, p, c)\big)$ stochastically dominates $C_1\big(G^{\mathcal{V}}(n, (1 - \delta)\kappa)\big)$. Similar domination results hold for the sum of the types of the vertices in the largest connected component (see formula (1.29)).

Now Theorems 1.3 and 1.5 are applicable to the models $G^{\mathcal{V}}(n, (1 \pm \delta)\kappa)$. Note only that, instead of the sequence $\{n \geqslant 1\}$ as in Theorem 1.3 or Theorem 1.5, one can use $\{K_N, N \geqslant 1\}$, which is unbounded a.s. by (1.27).

By (1.11),

$$\|T_\kappa\| = c\mathbb{E}(|C|^{-1}) \sum_{x \in \mathcal{S}} x^2 \mu(x) = c\mathbb{E}(|C|^{-1}) \sum_{x \in \mathcal{S}} x^2 \frac{1}{\mathbb{E}(|C|^{-1})} \frac{\mathbb{P}\{|C| = x\}}{x} = c\mathbb{E}|C|.$$

Hence, here $\|T_\kappa\| < 1$ if and only if $c < \frac{1}{\mathbb{E}|C|}$. Note that $\lim_{\delta \to 0} \|T_{(1 \pm \delta)\kappa}\| = \|T_\kappa\|$.

Now applying Theorem 1.5 to the models $G^{\mathcal{V}}(n, (1 - \delta)\kappa)$ and $G^{\mathcal{V}}(n, (1 + \delta)\kappa)$, and using the continuity of the function α, we readily get the result for the $G_N(p, c)$ model.

Corollary 1.6. *Assume that $d \geqslant 1$ and $0 \leqslant p < p_c(d)$. Then*

$$\frac{C_1\big(G_N(p, c)\big)}{\log |B(N)|} \xrightarrow{P} \frac{1}{\log \gamma(p, c)}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

as $N \to \infty$, *where* $\gamma(p,c) = \alpha(c\mathbb{E}(|C|^{-1}))$ *is defined by* (1.24) *with* $\psi(x) = x$ *and* μ *defined in* (1.32). *Furthermore,*

$$\gamma(p,c)\begin{cases} > 1 & \text{if } c < \frac{1}{\mathbb{E}|C|}, \\ = 1 & \text{if } c \geqslant \frac{1}{\mathbb{E}|C|}. \end{cases} \tag{1.36}$$

This result was conjectured in [10], where it was proved that, for any $\varepsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left\{ \frac{C_1\big(G_N(p,c)\big)}{\log|B(N)|} > \frac{1}{\log\gamma(p,c)} + \varepsilon \right\} = 0.$$

We also refer to [10] for a more exact description of $\gamma(p,c)$, which is derived similarly to (1.22).

2. Branching processes associated with graphs

2.1. The generating function for the progeny of a branching process

Recall that $\mathcal{X}(x)$ denotes the total number of the particles including the initial one, produced by the branching process $B(x)$, and that $\Phi(x)$ denotes the total activity as defined in (1.23). For $z \geqslant 1$, set

$$h_z(x) = \mathbb{E}z^{\mathcal{X}(x)}, \quad g_z(x) = \mathbb{E}z^{\Phi(x)}.$$

Define also

$$H_z = \sum_{x\in\mathcal{S}} \psi(x)\mu(x)h_z(x). \tag{2.1}$$

Then we rewrite (1.16) as

$$r(c) = \sup\{z \geqslant 1 : H_z < \infty\}.$$

First we shall prove the following lemma, which in particular yields (1.19) and (1.26).

Lemma 2.1. *Let* μ *be a probability measure on* $\mathcal{S}$, *and let* ψ *be a function on* $\mathcal{S}$ *that is positive and satisfies* (1.14). *Write* (*as in* (1.17))

$$c^{cr} := \left(\sum_{x\in\mathcal{S}} \psi^2(x)\mu(x)\right)^{-1}.$$

Then

(I)
$$r(c)\begin{cases} > 1 & \text{if } c < c^{cr}, \\ = 1 & \text{if } c \geqslant c^{cr}; \end{cases} \tag{2.2}$$

(II)
$$\alpha(c)\begin{cases} > 1 & \text{if } c < c^{cr}, \\ = 1 & \text{if } c \geqslant c^{cr}; \end{cases} \tag{2.3}$$

(III) *for all* $x \in S$,
$$\sup\{z \geqslant 1 : h_z(x) < \infty\} = r(c), \tag{2.4}$$

and
$$\sup\{z \geqslant 1 : g_z(x) < \infty\} = \alpha(c). \tag{2.5}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Proof. Note that for all $z \geqslant 1$ the function $h_z(k)$ (as a generating function for a branching process) is the minimal solution to the following equation:

$$h_z(k) = z \exp\left\{\sum_{x\in\mathcal{S}} \kappa(k,x)\mu(x)(h_z(x)-1)\right\}$$
$$= z \exp\left\{c\psi(k)\sum_{x\in\mathcal{S}} \psi(x)\mu(x)\big(h_z(x)-1\big)\right\} \tag{2.6}$$

(see, *e.g.*, [1] for similar results). Let X denote a random variable on $\mathcal{S}$ with distribution μ. Then, from the definition (2.1) of H_z it follows that

$$H_1 = \sum_{x\in\mathcal{S}} \psi(x)\mu(x) = \mathbb{E}\psi(X) =: M,$$

and we can rewrite formula (2.6) as

$$h_z(k) = z \exp\{c\psi(k)(H_z - M)\}. \tag{2.7}$$

Multiplying both sides by $\psi(k)\mu(k)$ and summing up over k, we find for all $1 \leqslant z < r(c)$

$$H_z = \sum_{k\in\mathcal{S}} \psi(k)\mu(k) z \exp\{c\psi(k)(H_z - M)\}$$
$$= z\mathbb{E}\psi(X)\exp\{c\psi(X)(H_z - M)\}. \tag{2.8}$$

Clearly, H_z is an increasing function of z. Hence, H_z is the minimal solution $H_z \geqslant H_1 = M$ to the equation (2.8). Note that equation (2.8) has a finite solution for some $z > 1$ if and only if the equation

$$y = \frac{1}{M} z\mathbb{E}\psi(X)\exp\{c\psi(X)M(y-1)\} \tag{2.9}$$

(for the same value z) has a root $y > 1$. In other words, if equation (2.9) has a finite solution for some $z > 1$ then H_z is finite.

Consider the function on the right in (2.9):

$$f(y,z) := zf(y) := z\frac{1}{M}\mathbb{E}\psi(X)\exp\{c\psi(X)M(y-1)\}. \tag{2.10}$$

It is easy to see, by taking into account assumption (1.14), that $f(y)$ is finite, at least in some neighbourhood $(1, 1+\varepsilon_1)$ of $y = 1$. Furthermore, wherever $f(y,z)$ is defined (and finite), it is increasing in both variables, it has all second-order derivatives and $\frac{\partial^2}{\partial y^2} f(y,z) > 0$. Now compute

$$\frac{\partial}{\partial y} f(y,z)|_{y=1,z=1} = c\mathbb{E}\psi^2(X) = \frac{c}{c^{cr}}. \tag{2.11}$$

Hence, if $c \geqslant c^{cr}$, then for any $z > 1$ we have

$$\frac{\partial}{\partial y} f(y,z)|_{y=1} > 1,$$

which together with the facts that $f(1,z) = z$ and $\frac{\partial^2}{\partial y^2} f(y,z) > 0$ implies

$$f(y,z) > y$$

for all $y \geqslant 1$, and therefore there is no solution $y \geqslant 1$ to (2.9).

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

On the other hand, if $c < c^{cr}$ we have

$$\frac{\partial}{\partial y} f(y,1)|_{y=1} < 1.$$

Note that for any fixed $z > 1$ the function $f(y,z)$ grows exponentially in y. Therefore there exists a $z_0 > 1$ such that, for every $1 \leqslant z \leqslant z_0$, there is a finite solution $y \geqslant 1$ to (2.9). This, in turn, implies the existence of $z > 1$ for which (2.8) has a finite solution H_z. This proves (2.2). Statement (2.4) follows immediately by (2.7).

Next we exploit the relation (1.23) to verify that, for each $z \geqslant 1$, the function $g_z(k)$ is the minimal solution $g_z(k) \geqslant 1$ to the equation

$$g_z(k) = z^{\psi(k)} \exp\left\{ c\psi(k) \sum_{x \in \mathcal{S}} \psi(x)\mu(x)\big(g_z(x) - 1\big) \right\}.$$

Rewrite the last formula as follows:

$$g_z(k) = z^{\psi(k)} \exp\{c\psi(k)(G_z - M)\}, \tag{2.12}$$

where $M = \sum_{x \in \mathcal{S}} \psi(x)\mu(x) = \mathbb{E}\psi(X)$ and

$$G_z = \sum_{x \in \mathcal{S}} \psi(x)\mu(x) g_z(x).$$

Note that for all $z \geqslant 1$

$$G_z \geqslant G_1 = \mathbb{E}\psi(X) = M.$$

Multiplying both sides of (2.12) by $\psi(k)\mu(k)$ and summing up over k, we find that for all $1 \leqslant z < \alpha(c)$, G_z is the minimal solution $G_z \geqslant M$ to

$$\begin{aligned} G_z &= \sum_{k \in \mathcal{S}} \psi(k)\mu(k) z^{\psi(k)} \exp\{c\psi(k)(G_z - M)\} \\ &= \mathbb{E} z^{\psi(X)} \psi(X) \exp\{c\psi(X)(G_z - M)\}. \end{aligned} \tag{2.13}$$

Then G_z is finite for some $z > 1$ if and only if equation

$$y = \frac{1}{M} \mathbb{E} z^{\psi(X)} \psi(X) \exp\{cM\psi(X)(y-1)\} =: F(y,z)$$

has a finite solution $y > 1$.

Observe that the function $F(y,z)$ has essentially the same properties as the function $f(y,z)$, namely, for all (y,z) in some neighbourhood $(1, 1+\varepsilon_1') \times (1, 1+\varepsilon_2')$ of the point $(1,1)$ it is increasing in both variables, it has all second-order derivatives and $\frac{\partial^2}{\partial y^2} F(y,z) > 0$. Furthermore, again as in (2.11), we get

$$\frac{\partial}{\partial y} F(y,z)|_{y=1,z=1} = c\mathbb{E}\psi^2(X) = \frac{c}{c^{cr}}.$$

Therefore the rest of the proof of (2.3) and (2.5) is identical to that of (2.2) and (2.4). □

2.2. A single-type branching process associated with the multi-type one

Here we derive some useful relations between the multi-type branching process $B(x)$ and a certain homogeneous Galton–Watson process.

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Let $\widetilde{X}$ be a random variable with distribution

$$\mathbb{P}\{\widetilde{X}=x\}=\frac{\psi(x)\mu(x)}{M},\quad x\in\mathcal{S}, \tag{2.14}$$

where, as before, $M=\mathbb{E}\psi(X)$. (This distribution is *size-biased* with ψ as the size function.) We also define a random variable Y with a mixed Poisson distribution so that, conditionally on $\widetilde{X}=x$,

$$Y\ |_{\widetilde{X}=x}\in \mathrm{Po}(\psi(x)cM),\quad x\in\mathcal{S}. \tag{2.15}$$

Now let $\widetilde{\mathcal{X}}$ denote the total progeny of a branching process with a single ancestor, where each individual produces a number of offspring that is an independent copy of the random variable Y. It is straightforward to derive that the generating function

$$\widetilde{h}_z:=\mathbb{E}z^{\widetilde{\mathcal{X}}},\quad z\geqslant 1, \tag{2.16}$$

is the minimal solution $\widetilde{h}_z\geqslant 1$ to

$$\widetilde{h}_z=z\mathbb{E}e^{\psi(\widetilde{X})cM(\widetilde{h}_z-1)}. \tag{2.17}$$

Note that in the last equation on the right, the randomness of the function under the sign of expectation comes only from $\widetilde{X}$ defined by (2.14), while the expectation in (2.16) is over the entire branching process.

Rewriting equation (2.8) with the help of $\widetilde{X}$, we get that H_z is the minimal solution $H_z\geqslant M$ to

$$H_z=zM\mathbb{E}e^{c\psi(\widetilde{X})(H_z-M)},$$

which together with (2.17) says that

$$H_z/M=\widetilde{h}_z,\quad z\geqslant 1. \tag{2.18}$$

In particular, this also implies that

$$r(c)=\sup\{z\geqslant 1: H_z<\infty\}=\sup\{z\geqslant 1:\widetilde{h}_z<\infty\}. \tag{2.19}$$

Recall that by (2.1),

$$H_z=\sum_{x\in\mathcal{S}}\psi(x)\mu(x)\mathbb{E}z^{\mathcal{X}(x)}=M\mathbb{E}z^{\mathcal{X}(\widetilde{X})},$$

where $\mathcal{X}(\widetilde{X})$ is defined conditionally on $\widetilde{X}=x$ simply as $\mathcal{X}(x)$. Combining this with (2.18), we finally get

$$\mathbb{E}z^{\mathcal{X}(\widetilde{X})}=\widetilde{h}_z,\quad z\geqslant 1. \tag{2.20}$$

By the Uniqueness Theorem this implies

$$\mathbb{P}\{\mathcal{X}(\widetilde{X})=k\}=\mathbb{P}\{\widetilde{\mathcal{X}}=k\} \tag{2.21}$$

for all $k\geqslant 1$.

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Lemma 2.2. *Let $\widetilde{X}$ be a random variable with distribution defined in* (2.14). *Assume that* (1.14) *holds and* $\inf_{x\in\mathcal{S}}\psi(x)>0$. *Then, for all* $c<c^{cr}$,

$$\lim_{k\to\infty}\big(\mathbb{P}\{\mathcal{X}(\widetilde{X})=k\}\big)^{\frac{1}{k}}=\lim_{k\to\infty}\big(\mathbb{P}\{\widetilde{\mathcal{X}}=k\}\big)^{\frac{1}{k}}=\frac{1}{r(c)}. \tag{2.22}$$

Proof. We shall first establish the existence of the limit in the middle in (2.22). Then the first equality in (2.22) will follow by (2.21), while the second equality in (2.22) will follow by (2.19), which tells us that the radius of convergence of the power series $\widetilde{h}_z=\mathbb{E}z^{\widetilde{\mathcal{X}}}$ is $r(c)$, so

$$\limsup_{k\to\infty}\big(\mathbb{P}\{\widetilde{\mathcal{X}}=k\}\big)^{\frac{1}{k}}=\frac{1}{r(c)}. \tag{2.23}$$

By the formula derived by Otter [8] for the progeny of the branching process,

$$\mathbb{P}\{\widetilde{\mathcal{X}}=k\}=\frac{1}{k}\mathbb{P}\bigg\{\sum_{i=1}^{k}Y_i=k-1\bigg\}, \tag{2.24}$$

where Y_i, $i\geqslant 1$, are i.i.d. copies of Y. Taking into account that the conditional distribution of Y is Poisson (2.15), it is straightforward to derive

$$\mathbb{P}\{\widetilde{\mathcal{X}}=k\}=\frac{1}{k}\mathbb{E}\frac{S_k^{k-1}}{(k-1)!}e^{-S_k}, \tag{2.25}$$

where

$$S_k:=\sum_{i=1}^{k}cM\psi(\widetilde{X}_i)$$

with $\widetilde{X}_1,\dots,\widetilde{X}_k$ being i.i.d. copies of $\widetilde{X}$. Now write

$$N_k=\frac{1}{(k-1)!}S_k^{k-1}e^{-S_k},$$

so that (2.25) gives us

$$k\mathbb{P}\{\widetilde{\mathcal{X}}=k\}=\mathbb{E}N_k=:a_k. \tag{2.26}$$

It remains to show that the limit $\lim_{k\to\infty}a_k^{1/k}$ exists. This together with formula (2.26) will yield the existence of the limit in the middle of (2.22).

We shall use a super-multiplicativity property of a_k. First consider S_k. Write $Z_i=cM\psi(\widetilde{X}_i)$, $i=1,2,\dots$, which are i.i.d. by their definition. Then, for all $n,m\geqslant 1$,

$$S_{n+m}=\sum_{i=1}^{n+m}Z_i=S_n+\sum_{i=n+1}^{n+m}Z_i=:S_n+T,$$

and since each Z_i is positive, we have

$$S_{n+m}^{n+m-1}=\sum_{k=0}^{n+m-1}\binom{n+m-1}{k}S_n^kT^{n+m-1-k}\geqslant\binom{n+m-1}{n-1}S_n^{n-1}T^m.$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Rearranging the above gives us

$$\frac{S_{n+m}^{n+m-1}}{(n+m-1)!} \geqslant \frac{S_n^{n-1}}{(n-1)!}\frac{T^{m-1}}{(m-1)!}\frac{T}{m}.$$

By the assumptions there exists a positive b such that $\psi(x) \geqslant b > 0$ for all $x \in \mathcal{S}$. Thus $T \geqslant mcMb$, and multiplying both sides by $e^{-S_{n+m}} = e^{-S_n}e^{-T}$, we obtain

$$N_{n+m} \geqslant bcM\left(\frac{S_n^{n-1}}{(n-1)!}e^{-S_n}\right)\left(\frac{T^{m-1}}{(m-1)!}e^{-T}\right).$$

Since T has the distribution of S_m, and T and S_n are independent, taking the expectation on both sides yields

$$\mathbb{E}N_{n+m} \geqslant bcM\mathbb{E}N_n\mathbb{E}N_m,$$

that is,

$$\log a_{n+m} \geqslant \log a_n + \log a_m + \log(bcM).$$

Then, by the result of de Bruijn and Erdős [3, Theorem 23], the limit

$$\lim_{k\to\infty}\frac{\log a_k}{k}$$

exists, and, moreover, it cannot be $+\infty$ since $\log a_k = \log\big(k\mathbb{P}\{\widetilde{\mathcal{X}} = k\}\big) \leqslant \log k$. This implies the existence of $\lim_{k\to\infty} a_k^{1/k}$ and completes the proof of Lemma 2.2. □

Corollary 2.3. *Under the assumptions of Lemma 2.2, for any $x \in \mathcal{S}$,*

$$\lim_{k\to\infty}\big(\mathbb{P}\{\mathcal{X}(x) = k\}\big)^{\frac{1}{k}} = \frac{1}{r(c)}. \tag{2.27}$$

Proof. Fix $x \in \mathcal{S}$ arbitrarily. Note that

$$\mathbb{P}\{\mathcal{X}(\widetilde{X}) = k\} \geqslant \mathbb{P}\{\mathcal{X}(x) = k\}\mathbb{P}\{\widetilde{X} = x\}, \tag{2.28}$$

where $\mathbb{P}\{\widetilde{X} = x\}$ is positive by (2.14) and the assumptions. Also, with positive probability $\pi = \psi(x)cMe^{-\psi(x)cM}$, the initial individual of the process $B(x)$ has exactly one child, in which case the type of this child has distribution $\widetilde{X}$, so that for all $k > 0$

$$\mathbb{P}\{\mathcal{X}(x) = k\} \geqslant \pi\mathbb{P}\{\mathcal{X}(\widetilde{X}) = k-1\}. \tag{2.29}$$

Taking the kth root of inequalities (2.29) and (2.28), and applying Lemma 2.2, we readily get (2.27). □

3. Proofs of main results

3.1. Proof of Theorem 1.4

We shall find $r(c) = z_0$ as the (unique!) value for which function y is tangent to $f(y, z_0)$ (see (2.10) and (2.9)) for some $y \geqslant 1$.

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

As noted above, H_z is finite if and only if there is a solution $y > 1$ to $y = zf(y)$. Therefore $z_0 = \max_{y \geqslant 1} y/f(y)$. Observe that this is well-defined, since the function $f(y)$ grows exponentially in y. Let y_0 be the point of maximum, *i.e.*,

$$\left(\frac{y}{f(y)}\right)' |_{y=y_0} = 0.$$

This implies that y_0 is the unique (again, due to the convexity of $f(y)$) solution to

$$y_0 = \frac{f(y_0)}{f'(y_0)} = \frac{1}{cM} \frac{\mathbb{E}\psi(X)\exp\{c\psi(X)M(y_0-1)\}}{\mathbb{E}\psi^2(X)\exp\{c\psi(X)M(y_0-1)\}}, \tag{3.1}$$

and then

$$r(c) = z_0 = \frac{y_0}{f(y_0)} = \frac{1}{f'(y_0)} = \frac{1}{c\mathbb{E}\psi^2(X)\exp\{c\psi(X)M(y_0-1)\}}. \tag{3.2}$$

This proves Theorem 1.4. □

Corollary 3.1. *Under the conditions of Theorem 1.4, the function* $r(c)$ *defined in* (3.2) *is continuous in* $c \leqslant c^{cr}$.

Proof. Taking into account the dependence on c, let us write equation (3.1) as $y_0 = F_c(y_0)$, where $F_c(y) := \frac{f(y)}{f'(y)}$. Observe that the function $F_c(y)$, as well as all its partial derivatives in c and y up to second order, are continuous in both arguments as long as they are finite. Furthermore, it is straightforward to check that

$$\frac{\partial}{\partial y} F_c(y) = 1 - \frac{f(y)f''(y)}{(f'(y))^2} < 1 \tag{3.3}$$

for all (y, c) where defined, since both f and f'' are positive. Hence, by the Implicit Function Theorem, $y_0(c)$ is a continuous function. Thus function $r(c)$ is continuous as a superposition of continuous functions. □

3.2. Proof of Theorem 1.3

3.2.1. The upper bound. First we shall prove that if $c < c^{cr}$, then for any $\delta > 0$

$$\mathbb{P}\left\{C_1\left(G^{\mathcal{V}}(n,\kappa)\right) > \left(\frac{1}{\log r(c)} + \delta\right)\log n\right\} \to 0 \tag{3.4}$$

as $n \to \infty$.

Let a be a constant for which condition (1.14) holds. For any

$$0 \leqslant q < a/2, \tag{3.5}$$

define an auxiliary probability measure on $\mathcal{S}$ by

$$\mu_q(k) = m_q e^{q\psi(k)}\mu(k), \quad k \in \mathcal{S}, \tag{3.6}$$

where the normalizing constant

$$m_q := \left(\sum_{x\in\mathcal{S}} e^{q\psi(k)}\mu(k)\right)^{-1} > 0.$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Note that $\mu_0(k) = \mu(k)$ for all $k \in \mathcal{S}$, and m_q is continuous on $[0, a/2]$ with $m_0 = 1$. Fix $\varepsilon > 0$ and $0 < q < a/2$ arbitrarily and define an event

$$\mathcal{B}_n = \left\{ \frac{\#\{1 \leqslant i \leqslant n : x_i = k\}}{n} - \mu(k) \leqslant \varepsilon \mu_q(k), \quad \text{for all} \ \ k \in \mathcal{S} \right\}. \tag{3.7}$$

By assumption (1.15) we have

$$\mathbb{P}\{\mathcal{B}_n\} = 1 - o(1). \tag{3.8}$$

Let us fix $\mathbf{x} = (x_1, \dots, x_n) \in \mathcal{B}_n$ and consider a graph $G^{\mathcal{V}}(n, \kappa)$ given this sequence. Recall the usual algorithm of finding a connected component in a random graph. Take any vertex $1 \leqslant i \leqslant n$ to be the root. Find all the vertices connected to this vertex i in the graph $G^{\mathcal{V}}(n, \kappa)$, and then mark i as 'saturated'. Then, for each non-saturated revealed vertex j we find all the vertices connected to it but which have not been revealed previously, and then mark j as 'saturated'. We continue this process until we end up with a tree of saturated vertices. Let $\tau_n(i)$ denote the set of vertices in the tree constructed according to this algorithm with the root at a vertex i.

Letting $\mathbb{P}_{\mathbf{x}}$ denote the conditional probability given a sequence $\mathbf{x}$, by the above construction we have for all positive ω

$$\mathbb{P}_{\mathbf{x}}\{C_1\big(G^{\mathcal{V}}(n, \kappa)\big) > \omega\} = \mathbb{P}_{\mathbf{x}}\Big\{ \max_{1 \leqslant i \leqslant n} |\tau_n(i)| > \omega \Big\}, \tag{3.9}$$

where $|\tau_n(i)|$ denotes the number of vertices in $\tau_n(i)$. This together with (3.8) gives us

$$\begin{aligned} \mathbb{P}\{C_1\big(G^{\mathcal{V}}(n, \kappa)\big) > \omega\} &\leqslant \sup_{\mathbf{x} \in \mathcal{B}_n} \mathbb{P}_{\mathbf{x}}\Big\{ \max_{1 \leqslant i \leqslant n} |\tau_n(i)| > \omega \Big\} + o(1) \\ &\leqslant \sup_{\mathbf{x} \in \mathcal{B}_n} \sum_{i=1}^{n} \mathbb{P}_{\mathbf{x}}\{|\tau_n(i)| > \omega\} + o(1). \end{aligned} \tag{3.10}$$

To approximate the distribution of $|\tau_n(i)|$ we shall use the following branching processes. Let $B_{c,q}$ be a process defined similarly to B, but with the distribution of the offspring

$$\mathrm{Po}\big(c\psi(x)\psi(y)\mu_q(y)\big)$$

instead of $\mathrm{Po}\big(c\psi(x)\psi(y)\mu(y)\big)$. Note that $B_{c,0}$ is defined exactly like B.

Further, let $\mathcal{X}^{c,q}(k)$ denote the total number of the particles (including the initial one) produced by the branching process $B_{c,q}$ starting with a single particle of type k.

Proposition 3.2. *For any $c < c^{cr}$ one can find $q > 0$ and $c' > c$ arbitrarily close to 0 and c, respectively, such that for all large n*

$$\mathbb{P}_{\mathbf{x}}\{|\tau_n(i)| > \omega\} \leqslant \mathbb{P}\{\mathcal{X}^{c',q}(x_i) > \omega\} \tag{3.11}$$

for all $\mathbf{x} = (x_1, \dots, x_n) \in \mathcal{B}_n$, $1 \leqslant i \leqslant n$, and $\omega > 0$.

Proof. Observe that at each step of the exploration algorithm which defines τ_n, the number of the type y offspring of a particle of type x has a binomial distribution $\mathrm{Bin}(N'_y, p_{xy}(n))$ where N'_y is the number of the remaining vertices of type y, and, with a

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

slight abuse of the notation (1.12), we write here

$$p_{xy}(n) = \min\left\{c\frac{\psi(x)\psi(y)}{n}, 1\right\}. \tag{3.12}$$

Recall the well-known fact that a binomial $\mathrm{Bin}(n,p)$ distribution is dominated by a Poisson distribution $\mathrm{Po}(-n\log(1-p))$. Hence, if $Y_{n,p} \in \mathrm{Bin}(n,p)$ and $Z_\lambda \in \mathrm{Po}(\lambda)$, we have for all $k \geqslant 0$

$$\mathbb{P}\{Y_{n,p} \geqslant k\} \leqslant \mathbb{P}\{Z_{-n\log(1-p)} \geqslant k\}. \tag{3.13}$$

Conditionally on $\mathbf{x} \in \mathcal{B}_n$ we have

$$N'_y \leqslant \#\{1 \leqslant i \leqslant n : x_i = y\} \leqslant n(\mu(y) + \varepsilon\mu_q(y)) \tag{3.14}$$

for each $y \in \mathcal{S}$ (the constants q and ε are from the definition (3.7) of $\mathcal{B}_n$). The last inequality implies that, for any y such that

$$\#\{1 \leqslant i \leqslant n : x_i = y\} > 0,$$

we have

$$n(\mu(y) + \varepsilon\mu_q(y)) \geqslant 1. \tag{3.15}$$

By assumptions (1.14), (1.13) and (3.5), we have for all large y

$$\mu(y) + \varepsilon\mu_q(y) \leqslant e^{-a\psi(y)} + \varepsilon m_q e^{(q-a)\psi(y)} \leqslant be^{(q-a)A_1\psi_0(y)} \leqslant be^{-aA_1\psi_0(y)/2},$$

where $0 < b < 1 + 2\varepsilon$ for all small $q > 0$. Combining this with (3.15), we obtain for all large n and y that if $\#\{1 \leqslant i \leqslant n : x_i = y\} > 0$ then

$$\frac{1}{n} \leqslant \mu(y) + \varepsilon\mu_q(y) \leqslant be^{-aA_1\psi_0(y)/2},$$

which yields

$$\psi_0(y) \leqslant A_3 \log n$$

for some constant A_3. Therefore, conditionally on $\mathbf{x} \in \mathcal{B}_n$, for all x and y such that $\#\{1 \leqslant i \leqslant n : x_i = x\} > 0$ and $\#\{1 \leqslant i \leqslant n : x_i = y\} > 0$, we have

$$p_{xy}(n) \leqslant c\frac{(A_3 \log n)^2}{n}. \tag{3.16}$$

This and (3.14), together with the continuity of m_q, allow us for any fixed positive ε_1 to choose ε and q in (3.7), so that we get the inequalities

$$\begin{aligned} -N'_y \log(1 - p_{xy}(n)) &\leqslant \big(\mu(y) + \varepsilon\mu_q(y)\big)n|\log(1 - p_{xy}(n))| \\ &\leqslant (1+\varepsilon_1)\mu_q(y)c\psi(x)\psi(y) =: \mu_q(y)c'\psi(x)\psi(y) \end{aligned} \tag{3.17}$$

holding for all large n. In other words, for any $q > 0$ and $c' > c$ sufficiently close to 0 and c, respectively, the bound (3.17) holds for all large n.

Now, due to (3.13) and (3.17),

$$\begin{aligned} \mathbb{P}\{Y_{N'_y, p_{xy}(n)} \geqslant k\} &\leqslant \mathbb{P}\{Z_{-N'_y \log(1-p_{xy}(n))} \geqslant k\} \\ &\leqslant \mathbb{P}\{Z_{\mu_q(y)c'\psi(x)\psi(y)} \geqslant k\}. \end{aligned} \tag{3.18}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

Hence, if at each step of the exploration algorithm we replace the $\mathrm{Bin}(N'_y, p_{xy}(n))$ variable with the $\mathrm{Po}\big(\mu_q(y)c'\psi(x)\psi(y)\big)$ one, we obtain the statement (3.11). This proves Proposition 3.2. □

Substituting (3.11) into (3.10), we derive

$$\begin{aligned}\mathbb{P}\{C_1\big(G^{\mathcal{V}}(n,\kappa)\big) > \omega\} &\leqslant \sup_{\mathbf{x}\in\mathcal{B}_n}\sum_{i=1}^{n}\mathbb{P}\{\mathcal{X}^{c',q}(x_i) > \omega\} + o(1)\\ &\leqslant n\sum_{k\in\mathcal{S}}\big(\mu(k)+\varepsilon\mu_q(k)\big)\mathbb{P}\{\mathcal{X}^{c',q}(k) > \omega\} + o(1)\\ &\leqslant b_1 n\sum_{k\in\mathcal{S}}\mu_q(k)\mathbb{P}\{\mathcal{X}^{c',q}(k) > \omega\} + o(1) \qquad (3.19)\end{aligned}$$

as $n\to\infty$, where b_1 is some positive constant. Then, using Markov's inequality,

$$\mathbb{P}\{\mathcal{X}^{c',q}(k) > \omega\} \leqslant z^{-\omega}\mathbb{E}z^{\mathcal{X}^{c',q}(k)}, \qquad (3.20)$$

we obtain from (3.19) for all $z \geqslant 1$

$$\mathbb{P}\{C_1\big(G^{\mathcal{V}}(n,\kappa)\big) > \omega\} \leqslant b_1 n z^{-\omega}\sum_{k\in\mathcal{S}}\mu_q(k)\mathbb{E}z^{\mathcal{X}^{c',q}(k)} + o(1). \qquad (3.21)$$

Define

$$r_q(c') := \sup\left\{z \geqslant 1 : \sum_{x\in\mathcal{S}}\psi(x)\mu_q(x)\mathbb{E}z^{\mathcal{X}^{c',q}(x)} < \infty\right\}.$$

We shall argue that $r_q(c') > 1$ for some $q > 0$ and $c' > c$, and then, for all $1 < z < r_q(c')$, we will have in (3.21)

$$\sum_{k\in\mathcal{S}}\mu_q(k)\mathbb{E}z^{\mathcal{X}^{c',q}(k)} < \infty. \qquad (3.22)$$

Since ψ and μ_q for some positive q satisfy the conditions of Theorem 1.4, the function $r_q(c)$ is continuous in c by Corollary 3.1 (and Lemma 2.1). Continuity of $r_q(c)$ in q in some neighbourhood of $q = 0$ follows by a similar argument. Indeed, $r_q(c)$ is defined by formulae (3.1) and (3.2), where X is replaced by a random variable X_q with distribution μ_q; in other words, $f(y)$ in these formulae is replaced by

$$f_q(y) := \mathbb{E}\psi(X_q)\exp\{c\psi(X_q)M(y-1)\}.$$

Using the definition (3.6) of μ_q it is straightforward to check that the Implicit Function Theorem is applicable here in the same fashion as in the proof of Corollary 3.1, which gives us continuity of $r_q(c)$ in q in some neighbourhood of $q = 0$.

Since $r_0(c) = r(c) > 1$ for $c < c^{cr}$ by Lemma 2.1, the continuity of $r_q(c')$ allows us for any $\delta > 0$ to choose a small $\delta' > 0$ and (q, c') close to $(0, c)$ so that (3.22) holds with

$$z = r_q(c') - \delta' > 1, \qquad (3.23)$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

and also

$$\left(\frac{1}{\log r(c)}+\delta\right)\log(r_q(c')-\delta')>1. \tag{3.24}$$

Now, setting $\omega=\left(\frac{1}{\log r(c)}+\delta\right)\log n$ and $z=r_q(c')-\delta'$ in (3.21), we derive with the aid of (3.22) that

$$\begin{aligned}\mathbb{P}\left\{C_1\big(G^{\mathcal{V}}(n,\kappa)\big)>\left(\frac{1}{\log r(c)}+\delta\right)\log n\right\}&\leqslant b_2 n z^{-\omega}+o(1)\\ &=b_2 n\exp\left\{-\log(r_q(c')-\delta')\left(\frac{1}{\log r(c)}+\delta\right)\log n\right\}+o(1),\end{aligned} \tag{3.25}$$

where b_2 is some finite positive constant. This together with (3.24) implies statement (3.4). □

3.2.2. The lower bound. Here we show that if $c<c^{cr}$, then for any $\delta>0$

$$\lim_{n\to\infty}\mathbb{P}\left\{C_1\big(G^{\mathcal{V}}(n,\kappa)\big)<\left(\frac{1}{\log r(c)}-\delta\right)\log n\right\}=0. \tag{3.26}$$

Fix any small positive δ and denote

$$\omega=\left(\frac{1}{\log r(c)}-\delta\right)\log n, \tag{3.27}$$

$$N=N(n)=\frac{n}{\omega^2}. \tag{3.28}$$

Also, for an arbitrary fixed finite $D\in\mathcal{S}$ and $\varepsilon_1>0$, introduce the event

$$\mathcal{A}_n=\left\{\frac{\#\{x_i:x_i=y\}}{n}-\mu(y)\geqslant-\varepsilon_1\mu(y)\quad\text{for all}\quad\{y\in\mathcal{S}:\ y\leqslant D\}\right\}.$$

Observe that by (1.1)

$$\mathbb{P}\{\mathcal{A}_n\}=1+o(1). \tag{3.29}$$

For the moment, let us fix $\mathbf{x}=(x_1,\dots,x_n)\in\mathcal{A}_n$ and consider the graph $G^{\mathcal{V}}(n,\kappa)$ given this sequence. We let $\mathbb{P}_{\mathbf{x}}(\cdot)$ denote the corresponding conditional probability. We shall reveal successively the connected components in the following way. Let V_1 be an arbitrary fixed vertex in $\{1,\dots,n\}$, and let $L_1=\tau_n(V_1)$ be the set of vertices in the connected component containing the vertex V_1.

For any set $U\subset\{1,\dots,n\}$ of (previously used) vertices let $\tau_n^U(\cdot)$ denote the set of vertices of the tree constructed in the same way as $\tau_n(\cdot)$ but on $\{1,\dots,n\}\setminus U$ instead of $\{1,\dots,n\}$. In particular, with this notation $\tau_n^{\emptyset}(\cdot)=\tau_n(\cdot)$.

Given constructed components $L_1,\dots,L_k$ for $k\geqslant1$, pick $V_{k+1}\in\{1,\dots,n\}\setminus\cup_{i=1}^k L_i$ according to any fixed rule, and set $L_{k+1}=\tau_n^{\cup_{i=1}^k L_i}(V_{k+1})$. If $\{1,\dots,n\}\setminus\cup_{i=1}^k L_i=\emptyset$, we define $L_{k+1}=\emptyset$. Then, due to (3.29) we have

$$\mathbb{P}\{C_1\big(G^{\mathcal{V}}(n,\kappa)\big)<\omega\}\leqslant\sup_{\mathbf{x}\in\mathcal{A}_n}\mathbb{P}_{\mathbf{x}}\left\{\max_{1\leqslant i\leqslant[N]+1}|L_i|<\omega\right\}+o(1) \tag{3.30}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

as $n \to \infty$, where

$$\mathbb{P}_{\mathbf{x}}\Big\{\max_{1\leqslant i\leqslant [N]+1}|L_i|<\omega\Big\} = \mathbb{P}_{\mathbf{x}}\{|L_1|<\omega\}\prod_{i=1}^{[N]}\mathbb{P}_{\mathbf{x}}\{|L_{i+1}|<\omega \mid |L_1|<\omega,\dots,|L_i|<\omega\}$$
$$\leqslant \prod_{i=1}^{[N]}\max_{(U,V_i):\ |U|\leqslant N\omega, V_i\notin U}\mathbb{P}_{\mathbf{x}}\{|\tau_n^U(V_i)|<\omega\}. \tag{3.31}$$

To approximate the distribution of $|\tau_n^U(i)|$ we introduce another branching process. First define for any value $D \in \mathcal{S}$ an auxiliary probability measure

$$\hat{\mu}_D(y) = \begin{cases} M_D^{-1}\mu(y) & \text{if } y \leqslant D, \\ 0 & \text{otherwise,} \end{cases} \tag{3.32}$$

where $M_D := \sum_{y\leqslant D}\mu(y)$ is a normalizing constant. Then, for any positive c and D let $\hat{B}_{c,D}$ be a process defined similarly to B, but for which the distribution of the offspring is given by

$$\mathrm{Po}\big(c\psi(x)\psi(y)\hat{\mu}_D(y)\big) \tag{3.33}$$

instead of $\mathrm{Po}\big(c\psi(x)\psi(y)\mu(y)\big)$. Note that $\hat{B}_{c,\infty}$ is defined exactly like B. Let $\hat{\mathcal{X}}^{c,D}(x)$ denote the total number of the particles (including the initial one) produced by the branching process $\hat{B}_{c,D}$ starting with a single particle of type x.

Lemma 3.3. *For any $c' < c$, there is a finite D_0 such that for all $D > D_0$ we have*

$$\mathbb{P}_{\mathbf{x}}\{|\tau_n^U(i)|<\omega\} \leqslant \left(1+\frac{b}{n^2}\right)^{n\omega}\mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x_i)<\omega\} \tag{3.34}$$

for all large n, $\mathbf{x} = (x_1,\dots,x_n) \in \mathcal{A}_n$, $1 \leqslant i \leqslant n$, and $U \subset \{1,\dots,n\}$ with $|U| \leqslant N\omega$, where $b = b(c,D)$ is some positive constant independent of c', i and $\mathbf{x}$ (ω and N are defined by (3.27) *and* (3.28)).

Proof. At each step of the exploration algorithm which defines τ_n^U, the number of type y offspring of a particle of type x has a binomial distribution $\mathrm{Bin}(N'_y, p_{xy}(n))$ where N'_y is the number of remaining vertices of type y, and $p_{xy}(n)$ is again as in (3.12). Note that for all $x, y \leqslant D$ and all large n we have, by assumption (1.13), and definition (3.12),

$$p_{xy}(n) = c\frac{\psi(x)\psi(y)}{n} \leqslant c\frac{A_2^2\psi_0^2(D)}{n}. \tag{3.35}$$

Here we shall exploit another relation between the binomial and the Poisson distributions. Let $Y_{n,p} \in \mathrm{Bin}(n,p)$ and $Z_\lambda \in \mathrm{Po}(\lambda)$. It is straightforward to derive from the formulae for the corresponding probabilities that, for all $0 < p < 1/4$ and $0 \leqslant k \leqslant n$,

$$\mathbb{P}\{Y_{n,p}=k\} = \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = \frac{n!}{n^k(n-k)!}\big((1-p)e^{\frac{p}{1-p}}\big)^n e^{-n\frac{p}{1-p}}\frac{\big(n\frac{p}{1-p}\big)^k}{k!}$$
$$\leqslant (1+\gamma p^2)^n\,\mathbb{P}\{Z_{n\frac{p}{1-p}}=k\}, \tag{3.36}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

where γ is some positive constant (independent of n, k and p). In particular, note that (3.36) trivially holds for all $k > n$.

Claim. Let $\mathbf{x} \in \mathcal{A}_n$. Then, for any fixed $\varepsilon_2 > \varepsilon_1$, for all $1 \leqslant y \leqslant D$, $k \leqslant N$ and for all L_i with $|L_i| \leqslant \omega$, $1 \leqslant i \leqslant N$, we have

$$N_y^k := \#\{i \in \{1,\dots,n\} \setminus \cup_{i=1}^k L_i : x_i = y\} \geqslant (1-\varepsilon_2)\mu(y)n \tag{3.37}$$

for all large n.

Proof of Claim. Recall that for any $\mathbf{x} \in \mathcal{A}_n$,

$$N_y := \#\{i \in \{1,\dots,n\} : x_i = y\} \geqslant (1-\varepsilon_1)\mu(y)n \tag{3.38}$$

for all $y < D$. Note that by deleting an arbitrary set U with $|U| \leqslant N\omega$ from $\{1,\dots,n\}$, we may delete at most $N\omega$ vertices of type y. Hence, the number N_y^k of the remaining vertices of type y for any $k \leqslant N$ is bounded from below as follows: we have

$$N_y^k \geqslant N_y - \omega - N\omega,$$

and thus by (3.38)

$$N_y^k \geqslant n(1-\varepsilon_1)\mu(y) - \omega - N\omega$$

for all $y < D$. Taking into account definitions (3.28) and (3.27), we obtain that for any $\varepsilon_2 > \varepsilon_1$ and $D > 0$

$$N_y^k \geqslant (1-\varepsilon_2)\mu(y)n$$

for all $y \leqslant D$ and all large n. This proves the Claim. □

By (3.37) and (3.35), at any step of the exploration algorithm we have

$$N'_y \frac{p_{xy}(n)}{1-p_{xy}(n)} \geqslant \mu(y)(1-\varepsilon_2)c\psi(x)\psi(y) \tag{3.39}$$

for any fixed $\varepsilon_2 > \varepsilon_1, D$, any $y \leqslant D$, and all large n. Now, with the help of (3.32) we rewrite (3.39) as follows:

$$N'_y \frac{p_{xy}(n)}{1-p_{xy}(n)} \geqslant \hat{\mu}_D(y)M_D(1-\varepsilon_2)c\psi(x)\psi(y) = \hat{\mu}_D(y)c'\psi(x)\psi(y) \tag{3.40}$$

for any fixed $\varepsilon_2 > \varepsilon_1, D$, all large n and any $x, y \in \mathcal{S}$, where

$$c' = M_D(1-\varepsilon_2)c.$$

Recall that $\lim_{D\to\infty} M_D \uparrow 1$. Therefore, choosing appropriate constants D and $\varepsilon_2 > \varepsilon_1$, we can make c' arbitrarily close to c.

Now using the relation (3.36) between the Poisson and the binomial distributions, we get for all k

$$\begin{aligned}\mathbb{P}\{Y_{N'_y,p_{xy}(n)} = k\} &\leqslant (1+\gamma p_{xy}^2(n))^{N'_y}\ \mathbb{P}\Big\{Z_{N'_y \frac{p_{xy}(n)}{1-p_{xy}(n)}} = k\Big\} \\ &\leqslant (1+\gamma p_{xy}^2(n))^{n}\ \mathbb{P}\Big\{Z_{N'_y \frac{p_{xy}(n)}{1-p_{xy}(n)}} = k\Big\},\end{aligned} \tag{3.41}$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

where $Z_{N'_y \frac{p_{xy}(n)}{1-p_{xy}(n)}}$ stochastically dominates $Z_{\hat{\mu}_D(y)c'\psi(x)\psi(y)}$ due to (3.40). This implies that if at each of at most ω steps of the exploration algorithm we replace the $\mathrm{Bin}(N'_y, p_{xy}(n))$ variable with the

$$\mathrm{Po}\big(\hat{\mu}_D(y)c'\psi(x)\psi(y)\big)$$

one, we arrive at the following bound using the branching process $\hat{B}_{c,D}$ and bound (3.35):

$$\mathbb{P}_{\mathbf{x}}\{|\tau_n^U(i)| < \omega\} \leqslant \left(1 + \gamma c^2 A_2^4 \frac{\psi_0(D)^4}{n^2}\right)^{n\omega} \mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x_i) < \omega\} \tag{3.42}$$

for all large n. This yields the statement of Lemma 3.3. □

Let $x^0 \in \mathcal{S}$ be such that

$$\psi(x^0) = \min_{x\in\mathcal{S}} \psi(x);$$

the existence of x^0 follows by the assumption (1.13). Then the distribution of $\hat{\mathcal{X}}^{c',D}(x^0)$ is stochastically dominated by $\hat{\mathcal{X}}^{c',D}(x_i)$ with any $x_i \in \mathcal{S}$, due to the Poisson law (3.33) of the offspring distribution. Hence, for all $x_i \in \mathcal{S}$,

$$\mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x_i) < \omega\} \leqslant \mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x^0) < \omega\}.$$

This allows us to derive from Lemma 3.3 a bound that is uniform in i and $\mathbf{x}$, namely

$$\mathbb{P}_{\mathbf{x}}\{|\tau_n^U(i)| < \omega\} \leqslant \left(1 + \frac{b}{n^2}\right)^{n\omega} \mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x^0) < \omega\},$$

which together with (3.31) implies that for all large n

$$\mathbb{P}_{\mathbf{x}}\Big\{\max_{1\leqslant i\leqslant [N]+1} |L_i| < \omega\Big\} \leqslant \left(\left(1 + \frac{b}{n^2}\right)^{n\omega} \mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x^0) < \omega\}\right)^{[N]}.$$

Substituting this into (3.30), we get by taking into account the notation (3.27) and (3.28) that

$$\begin{aligned}\mathbb{P}\{C_1\big(G^{\mathcal{V}}(n,\kappa)\big) < \omega\} &\leqslant \left(1 + \frac{b}{n^2}\right)^{n^2/\omega} \mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x^0) < \omega\}^{n/\omega^2} + o(1)\\ &\leqslant e^{b_1/\log n}\left(1 - \mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x^0) \geqslant \omega\}\right)^{n/\omega^2} + o(1)\end{aligned} \tag{3.43}$$

as $n \to \infty$, where $b_1 = b_1(c, D, \delta)$ is a positive constant.

Assume that $D > x^0$. Consider

$$\hat{r}(c, D) := \sup\{z \geqslant 1 : \mathbb{E}z^{\hat{\mathcal{X}}^{c,D}(x)} < \infty\}.$$

Since $\hat{\mu}_D$ and ψ satisfy the conditions of Lemma 2.1 and Theorem 1.4, the function $\hat{r}(c, D)$ can be derived from (3.1) and (3.2) by replacing the distribution μ by $\hat{\mu}_D$. Again, the Implicit Function Theorem allows us to obtain (as for the proof of Corollary 3.1)

$$\lim_{D\to\infty} \hat{r}(c, D) = r(c)$$

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

for all $c < c^{cr}$. Also $\hat{r}(c, D)$ is continuous in c by Corollary 3.1. Hence we can find for any given $\delta_1 > 0$ a large constant D and $c' < c$ sufficiently close to c, such that

$$\hat{r}(c', D) < r(c) + \delta_1/2. \tag{3.44}$$

Finally, we observe that by Corollary 2.3

$$\lim_{k\to\infty} \big(\mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x^0) = k\}\big)^{\frac{1}{k}} = \frac{1}{r(c', D)},$$

which implies that there is a positive constant $A = A(\delta_1) < \infty$ such that, for all large ω,

$$\mathbb{P}\{\hat{\mathcal{X}}^{c',D}(x^0) \geqslant \omega\} \geqslant A(\hat{r}(c', D) + \delta_1/2)^{-\omega} \geqslant A(r(c) + \delta_1)^{-\omega}, \tag{3.45}$$

where the last bound is due to (3.44). This allows us to derive from (3.43) that, for any $\delta_1 > 0$ and some positive A,

$$\mathbb{P}\{C_1\big(G^{\mathcal{V}}(n, \kappa)\big) < \omega\} \leqslant e^{b_1/\log n}\big(1 - A(r(c) + \delta_1)^{-\omega}\big)^{n/\omega^2} + o(1).$$

Substituting here the formula (3.27) for ω, and setting $\alpha = \big(\frac{1}{\log r(c)} - \delta\big)^{-2}$, we get

$$\mathbb{P}\{C_1\big(G^{\mathcal{V}}(n, \kappa)\big) < \omega\} \leqslant e^{b_1 \log^3 n}\big(1 - A(r(c) + \delta_1)^{-(\frac{1}{\log r(c)} - \delta)\log n}\big)^{\alpha n/\log^2 n} + o(1). \tag{3.46}$$

Now, for any $\delta > 0$ we choose a positive δ_1 so that

$$\gamma_1 := \left(\frac{1}{\log r(c)} - \delta\right)\log\big(r(c) + \delta_1\big) < 1.$$

Then (3.46) becomes

$$\mathbb{P}\left\{C_1\big(G^{\mathcal{V}}(n, \kappa)\big) < \left(\frac{1}{\log r(c)} - \delta\right)\log n\right\} \leqslant e^{b_1/\log n}\left(1 - \frac{A}{n^{\gamma_1}}\right)^{\alpha^2 n/\log^2 n} + o(1), \tag{3.47}$$

where the right-hand side goes to zero as $n \to \infty$. This completes the proof of (3.26). □

3.2.3. End of the proof of Theorem 1.3. The statements (3.26) and (3.4) yield assertion (1.18) of Theorem 1.3 in the case $c < c^{cr}$.

Recall that statement (1.19) of Theorem 1.3 is proved by Lemma 2.1, which also proves continuity of $r(c)$ for $c \geqslant c^{cr}$. Continuity of $r(c)$ for $c \leqslant c^{cr}$ is established in Corollary 3.1.

If $c \geqslant c^{cr}$, then for any $0 < \alpha < \frac{c^{cr}}{c} \leqslant 1$ the size $C_1\big(G^{\mathcal{V}}(n, \kappa)\big)$ stochastically dominates $C_1\big(G^{\mathcal{V}}(n, \alpha\kappa)\big)$, which by the previous part converges in probability to

$$\frac{1}{\log r(\alpha c)}.$$

Since $r(\alpha c) \to 1$ as $\alpha \to \frac{c^{cr}}{c}$, and $r(c) = 1$ for all $c \geqslant c^{cr}$, statement (1.18) of Theorem 1.3 in the case $c \geqslant c^{cr}$ follows. This completes the proof of Theorem 1.3. □

3.3. Proof of Theorem 1.5

The proof of Theorem 1.5 is nearly the same as that of Theorem 1.3. The only difference is that the random variables $|\tau_n(i)|$, $1 \leqslant i \leqslant n$, used in the proof of Theorem 1.3 should be

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press

replaced by

$$\Psi_n(i) := \sum_{v \in \tau_n(i)} \psi(x_v),$$

while $\mathcal{X}^{c,q}$ and $\hat{\mathcal{X}}^{c,D}$ should be replaced by $\Phi^{c,q}$ and $\hat{\Phi}^{c,D}$, which denote the activity of the total progeny of the branching processes $B_{c,q}$ and $\hat{B}^{c,D}$, respectively (see definition of Φ and (1.23)). Then, due to the properties of α, namely, (2.5) and (2.3), given in Lemma 2.1, the proof of Theorem 1.5 follows exactly the same lines as the proof of Theorem 1.3. □

Acknowledgement

The author thanks A. Martin-Löf and D. Mason for helpful discussions.

The author is very grateful to the referee, who not only gave many valuable remarks which improved the proofs and the presentation, but also generously communicated Corollary 2.3.

References

[1] Bollobás, B., Janson, S. and Riordan, O. (2007) The phase transition in inhomogeneous random graphs. *Random Struct. Alg.* **31** 3–122.
[2] Britton, T., Deijfen, M. and Martin-Löf, A. (2006) Generating simple random graphs with prescribed degree distribution. *J. Statist. Phys.* **124** 1377–1397.
[3] de Bruijn, N. G. and Erdős, P. (1951) Some linear and some quadratic recursion formulas I. *Nederl. Akad. Wetensch. Proc. Ser. A.* **54** (*Indagationes Math.* **13**) 374–382.
See also `http://www.renyi.hu/~p_erdos/1952-10.pdf`
[4] Erdős, P. and Rényi, A. (1960) On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5** 17–61.
[5] Chung, F. and Lu, L. (2006) The volume of the giant component of a random graph with given expected degree. *SIAM J. Discrete Math.* **20** 395–411.
[6] Grimmett, G. (1999) *Percolation*, Springer, Berlin.
[7] Karp, R. M. (1990) The transitive closure of a random digraph. *Random Struct. Alg.* **1** 73–93.
[8] Otter, R. (1949) The multiplicative process. *Ann. Math. Statist.* **20** 206–224.
[9] Turova, T. S. (2003) Long paths and cycles in the dynamical graphs. *J. Statist. Phys.* **110** 385–417.
[10] Turova, T. S. and Vallier, T. Merging percolation on Z^d and classical random graphs: Phase transition. *Random Struct. Alg.* **36** 185–217.

https://doi.org/10.1017/S0963548310000180 Published online by Cambridge University Press