ON A CONDITIONALLY POISSONIAN GRAPH PROCESS

ILKKA NORROS [*] [**] AND

HANNU REITTU,[*] [***] *VTT Technical Research Centre of Finland*

Abstract

Random (pseudo)graphs G_N with the following structure are studied: first, independent and identically distributed *capacities* Λ_i are drawn for vertices $i = 1, \ldots, N$; then, each pair of vertices (i, j) is connected, independently of the other pairs, with $E(i, j)$ edges, where $E(i, j)$ has distribution Poisson($\Lambda_i \Lambda_j / \sum_{k=1}^{N} \Lambda_k$). The main result of the paper is that when $P(\Lambda_1 > x) \geq x^{-\tau+1}$, where $\tau \in (2, 3)$, then, asymptotically almost surely, G_N has a giant component, and the distance between two randomly selected vertices of the giant component is less than $(2 + o(N))(\log \log N)/(-\log (\tau - 2))$. It is also shown that the cases $\tau > 3$, $\tau \in (2, 3)$, and $\tau \in (1, 2)$ present three qualitatively different connectivity architectures.

Keywords: Random graph; branching process; giant component; power law

2000 Mathematics Subject Classification: Primary 05C80
Secondary 05C12

1. Introduction

In their classical paper [8], Erdös and Rényi considered a graph process with N vertices that starts with no edges. Edges are then added one by one between randomly selected pairs of vertices. The most exciting phenomenon in this process is the sudden appearance of a giant connected component at a certain critical point. The model of Erdös and Rényi was the starting point of the theory of random graphs, which has grown in many directions and found applications in different parts of science. Mainly, however, random graphs are of purely mathematical interest. They can be applied as models for connection structures that are too complicated to be studied in full detail, and which require a statistical approach if any general characterization and understanding of the object is to be obtained.

One feature that is characteristic of most of the older random graph models is a certain homogeneity, in the sense that the vertices have roughly similar roles in the random network. The degree of a randomly chosen vertex often has an approximately Poissonian distribution. However, many real-world networks – social networks in particular – display a strong heterogeneity that is reflected in a heavy-tailed distribution of vertex degrees. For reviews in this field, see [3] and [14]. Since the 1990s, the Internet has both provided new tools for studying such often huge networks and itself served as a highly interesting object of study.

There are (at least) two types of model that yield highly variable vertex degrees. One class comprises the preferential growth models, where the graph grows from a single initial vertex and

each new vertex tends to get connected to others that are already big, rather than to small ones. A pioneering contribution in this direction was the model proposed by Barabási and Albert [2], whose idea has then been developed further in many modifications. Rigorous results were obtained by Bollobás and Riordan (see [5] and references therein).

Another class consists of models in which the vertex degrees are taken directly as the starting point. Again, there are several variants. Most of them are in fact pseudographs, because multiple edges between the same pairs of vertices and loops from a vertex to itself are not prevented. Molloy and Reed [12], [13] defined the notion of a feasible degree sequence and studied the conditions for a giant component to appear when the graph is randomly chosen among all graphs with that degree sequence. Aiello *et al.* [1] continued in this direction, focusing on power law sequences. Newman *et al.* [15] assumed there to be independent and identically distributed (i.i.d.) random degrees and added one vertex with degree 1 if the sum of the degrees happened to be odd. Their paper focused on the typical distances between vertices belonging to the giant component and showed it to be at most logarithmic in N. Reittu and Norros [16], [17] considered the model of [15] and showed that, for a power-tailed degree distribution with infinite variance, these distances were in fact only of order $\log \log N$. Chung and Lu [7] proved a $\log \log N$ asymptotics in a model that assumed a power law sequence of 'expected degrees' $d_1, \ldots, d_N$, and connected vertices i and j with probability $d_i d_j / \sum_{k=1}^{N} d_k$ (assuming a cut-off $d_i \le$ const. $\sqrt{N}$, guaranteeing that these numbers do not exceed 1). Recently, van der Hofstad *et al.* [19], [20] and Hooghiemstra and van Mieghem [9] have improved the result of [17] considerably by identifying the limit distributions of the distance between two vertices of the giant component in the following three qualitatively different cases of degree distribution: finite variance, finite mean plus infinite variance, and infinite mean.

The present paper introduces a new variant of the latter class of model that is a kind of hybrid of the models of [15] and [7]. The idea is as follows: since in many of the above-mentioned models the number of edges joining two disjoint sets of vertices are approximately Poisson distributed given their degree sums, we propose to define them to be *exactly* Poissonian, conditioned on some mean values. It turns out that this approach yields a very pleasant mathematical structure that, moreover, can be embedded into a simple growth model.

The paper is structured as follows. In Section 2 we present the model, in Section 3 we study a coupling of neighborhood shells to a branching process, and in Section 4 we present the results on the asymptotic architectures.

2. Model definition

Let $\mathbf{\Lambda} = (\Lambda_1, \Lambda_2, \ldots)$ be a sequence of independent, strictly positive random variables with a common distribution function F. As usual, Λ (without a subscript) denotes a generic random variable with distribution function F. We assume throughout that $\mathrm{E}[\Lambda_1] < \infty$ and $\mathrm{P}(\Lambda \ge 1) = 1$ (the latter assumption is not essential but sometimes convenient). In the graphs that we shall define below, Λ_i will represent the mean degree of vertex i. Let us call Λ_i the *capacity* of vertex i.

In the construction that follows, the sequence (Λ_i) can be thought to be drawn first; probabilities defined thereafter are sometimes understood to be conditioned with respect to $\mathbf{\Lambda}$, even without writing this explicitly (when it is clear from the context). The total capacity of the vertices in $\{1, 2, \ldots, N\}$ is denoted by $L_N = \sum_{i=1}^{N} \Lambda_i$. While conditioning on the sequence $\mathbf{\Lambda}$, let us define a sequence of random (pseudo)graphs G_N, $N = 1, 2, \ldots$, as follows (here 'pseudo' means that we allow multiple edges between a pair of vertices, and loops from a

vertex to itself; we suppress the word 'pseudo' from now on). The set of vertices of G_N is $\{1, \ldots, N\}$. Denote by $E_N(i, j)$, $i, j \in \{1, \ldots, N\}$, the number of edges joining vertices i and j in G_N. The sequence of graphs is generated inductively through the following growth process:

Step 1. $E_1(1, 1)$ is a random variable with distribution Poisson(Λ_1).

Step $N + 1$. G_{N+1} is the same as G_N except for the following changes.

(i) Addition of new edges to the new vertex $N + 1$: for $i = 1, \ldots, N + 1$, the number $E_{N+1}(i, N + 1)$ is a Poisson random variable with mean

$$\mathrm{E}[E_{N+1}(i, N + 1)] = \frac{\Lambda_i \Lambda_{N+1}}{L_{N+1}},$$

independent of the previously drawn edges.

(ii) Thinning of the old edges: each edge of G_N is removed independently with probability

$$p_{N+1} := 1 - \frac{L_N}{L_{N+1}}.$$

We denote by $i \leftrightarrow_N j$ the event that vertices i and j are connected with an edge in G_N. If U and V are subsets of $\{1, \ldots, N\}$, $U \leftrightarrow_N V$ denotes the event that some vertex in U is connected to some vertex in V by an edge in G_N.

Denote by $D_N(i)$ the degree of vertex i in G_N:

$$D_N(i) := \sum_{j=1}^{N} E_N(i, j).$$

Thus, in determining the degree of a vertex, we count a loop only once.

Proposition 2.1. *Conditionally on $\boldsymbol{\Lambda}$, the following assertions hold for each N.*

(i) *The distribution of $E_N(i, j)$ is Poisson with mean*

$$\mathrm{E}[E_N(i, j) \mid \boldsymbol{\Lambda}] = \frac{\Lambda_i \Lambda_j}{L_N}.$$

(ii) *The distribution of $D_N(i)$ is* Poisson(Λ_i).

(iii) *The numbers of edges joining different pairs of vertices are independent.*

Proof. The claims hold with $N = 1$. Assume that they hold for a general N, and consider G_{N+1}. Assertion (i) holds by definition when $j = N + 1$. For $i, j \leq N$, note that independent thinning retains Poissonianity. The induction hypothesis thus implies that

$$\begin{aligned}
\mathrm{E}[E_{N+1}(i, j) \mid \boldsymbol{\Lambda}] &= (1 - p_{N+1}) \, \mathrm{E}[E_N(i, j) \mid \boldsymbol{\Lambda}] \\
&= \frac{L_N}{L_{N+1}} \frac{\Lambda_i \Lambda_j}{L_N} = \frac{\Lambda_i \Lambda_j}{L_{N+1}}, \qquad i, j \in \{1, \ldots, N\}.
\end{aligned}$$

Assertion (iii) also follows from the induction hypothesis, since the thinning probability is independent of any edges, and the edges to vertex $N + 1$ are drawn independently of previously

drawn edges and each other. Assertion (ii) follows from assertions (i) and (iii) since a sum of independent Poisson random variables is Poisson and

$$\mathrm{E}[D_N(i) \mid \boldsymbol{\Lambda}] = \sum_{j=1}^{N} \mathrm{E}[E_N(i, j) \mid \boldsymbol{\Lambda}] = \sum_{j=1}^{N} \frac{\Lambda_i \Lambda_j}{L_N} = \Lambda_i.$$

3. Coupling to a branching process

The neighborhood shells around a vertex of a random graph often have a close resemblance to the subsequent generations of a branching process. This section analyses the neighborhood shells in our model, and their coupling to a related branching process.

Assume that the sequence $\boldsymbol{\Lambda}$ has been generated and that probabilities are conditioned on it. For each N, we define a *marked branching process*, i.e. a branching process where each individual is associated with some element of a *mark space*, which in this case is $\{1, \ldots, N\}$, the set of vertices of G_N. More specifically, let us define a process $(Z, \boldsymbol{J}) = (Z_n, (J_{n,i}))$, where Z_n is the size of generation n and $J_{n,i} \in \{1, \ldots, N\}$ is the mark of member i of generation n. We set $Z_0 \equiv 1$ and take $J_{0,1}$ from the uniform distribution uniform$(1, \ldots, N)$. The process then proceeds so that, for each $j = 1, \ldots, N$, an individual bearing mark i gives birth independently to a Poisson$(\Lambda_i \Lambda_j / L_N)$-distributed number of members of the next generation, each bearing mark j.

Now consider the graph G_N, and randomly choose (from uniform distribution) a vertex i_0. Recursively define the neighborhood shells, $\mathcal{N}_k$, of i_0 as follows:

$$\mathcal{N}_0(i_0) = \{i_0\},$$

$$\mathcal{N}_{k+1}(i_0) = \left\{ j \in \left(\bigcup_{l=0}^{k} \mathcal{N}_l(i_0) \right)^{\mathrm{c}} : j \leftrightarrow \mathcal{N}_k(i_0) \right\}.$$

Proposition 3.1. *Let $(Z, \boldsymbol{J})$ be the marked branching process defined above. Define a reduced process by proceeding generation by generation, i.e. in the order $J_{0,1}; J_{1,1}, J_{1,2}, \ldots, J_{1,Z_1}; J_{2,1}, \ldots,$ and pruning (that is, deleting) from $(Z, \boldsymbol{J})$ each individual whose mark has already appeared, together with all its descendants (which are not individually considered for pruning). Denote the resulting finite process by $(\hat{Z}, \hat{\boldsymbol{J}})$, and let $\hat{\mathcal{J}}_k$ be the set of marks in generation k of the reduced process. Then the sequence of the sets $\hat{\mathcal{J}}_k$ has the same distribution as the sequence $\mathcal{N}_k$.*

Proof. We proceed by induction. By definition, i_0 and $J_{0,1}$ have the same (uniform) distribution and $\hat{\mathcal{J}}_0 = \{J_{0,1}\}$. It is clear that the reduced process can also be generated simultaneously with $(Z, \boldsymbol{J})$, by deleting each new individual whose mark has already been seen without letting it reproduce. Let us do this, and assume that generation n has been fully generated and is identical in both processes, and that the same holds for the children of the first i members of generation n. It now suffices to show that we obtain a similarly distributed result for the children of member $i + 1$.

First consider the graph neighborhood process. Here we need only consider vertices that have not yet been reached. With probability $1 - \exp(-\Lambda_{\hat{J}_{n,i+1}} \Lambda_j / L_N)$ there is at least one edge between each such vertex, j, and the present vertex. Moreover, the numbers of edges to different vertices are independent, by assumption. In the branching process, on the other hand, for each $j = 1, \ldots, N$ the corresponding individual produces a Poisson$(\Lambda_{\hat{J}_{n,i+1}} \Lambda_j / L_N)$-distributed

number of offspring each bearing mark j. For different js, these numbers are independent. The pruning then removes all vertices whose marks were generated in previous steps, and also all newly generated duplicates. Clearly, the set of remaining marks is distributed in the same way as the above-described set of new neighbors in the neighborhood process.

We now make a very useful observation: the number and the marks of the children of an individual in $(Z, \boldsymbol{J})$ can be generated independently. This follows from the properties of the Poisson distribution. Denote by q_N the distribution

$$q_N(j) := \frac{\Lambda_j}{L_N}, \qquad j = 1, \dots, N,$$

and introduce another marked branching process, $(\tilde{Z}_n, (\tilde{J}_{n,i}))$. In this process, $\tilde{Z}_0$ and $\tilde{J}_{0,1}$ are defined as before but then we proceed as follows. We define Z_1 to have the distribution Poisson($\Lambda_{J_{0,1}}$). Each individual of a generation greater than or equal to 1 is given a mark J from the distribution q_N, independently of everything that happened before, and then produces a Poisson(Λ_J)-distributed number of children. Thus, after the first generation, the process $(\tilde{Z})$ is a *standard* Galton–Watson process with the mixed Poisson ('doubly stochastic') offspring distribution Poisson(Λ_J), $J \sim q_N$.

Proposition 3.2. *The processes $(Z_n, (J_{n,i}))$ and $(\tilde{Z}_n, (\tilde{J}_{n,i}))$ are stochastically identical. In particular, conditionally on*

$$\sigma\{Z_0, \dots, Z_n; \; J_{k,i}, \; k = 0, \dots, n-1, \; i = 1, \dots, Z_k\},$$

the marks of generation n in $(Z_n, (J_{n,i}))$ are i.i.d. random variables with distribution q_N.

Proof. Fix n, and assume as an induction hypothesis that the nth generation of each process has been generated and that these generations are identical. For each individual $i \in \{1, \dots, Z_n\}$, denote by $N_{i,j}$ the number of its offspring bearing mark j. Furthermore, let $\tilde{N}_{i,j}$ be the number of offspring bearing mark j when the total number of offspring is first drawn from the above-mentioned doubly stochastic distribution and each child is then independently given a mark according to the distribution q_N. It is straightforward to check that $N_{i,j}$ and $\tilde{N}_{i,j}$ are identical in distribution. Since in both processes the reproduction of each individual happens independently of other individuals, the $(n+1)$th generation has the same distribution in both processes. Thus, the assertion follows by induction.

Remark 3.1. In fact, the Poisson distribution allows us to remove the explicit branching altogether: we would obtain the same process by drawing the size of the $(n+1)$th generation directly from the distribution

$$\text{Poisson}\left(\sum_{i=1}^{Z_n} \Lambda_{J_{n,i}}\right)$$

and then giving each individual an independent q_N-distributed mark.

Remark 3.2. Denote by π^* the mixed Poisson distribution Poisson(Γ), where Γ is a random variable with distribution

$$\text{P}(\Gamma \in \mathrm{d}x) = \frac{x\,\text{P}(\Lambda \in \mathrm{d}x)}{\text{E}[\Lambda]}.$$

Obviously, the distribution of Poisson($\Lambda_{J^{(N)}}$), $J^{(N)} \sim q_N$, converges to π^* in probability.

For each N, let $J_1^{(N)}$, $J_2^{(N)}$, ... be i.i.d. random variables with distribution q_N, and let $J_0^{(N)}$ be a random variable with uniform distribution on $\{1, \ldots, N\}$, independent of the previous ones (we usually suppress the superscript '(N)', to simplify the notation). Let

$$\nu \equiv \nu(N) = \inf\{n : J_n = J_m \text{ for some } m < n\},$$
$$\Delta \equiv \Delta(N) = \sup\{n \geq 0 : J_n \text{ is used in the construction of } (\tilde{Z}, \tilde{J})\}. \tag{3.1}$$

Thus, ν is the first time that a vertex is encountered a second time and Δ is the time at which the branching process dies, i.e. the time at which the last survivor is drawn to be childless. Note that 'time' here refers to the index of the sequence J_n, with which we generate the process $(\tilde{Z}, \tilde{J})$.

It is crucial in some of the following proofs that the coupling between the neighborhood shells and the branching process lasts long enough, i.e. that ν is sufficiently large with high probability. The following lemma provides an appropriate estimate in the case of a power-tailed distribution of vertex capacities.

Lemma 3.1. *Either*

(i) *assume that* $\mathrm{E}[\Lambda^2] < \infty$ *and let* $\eta \in (0, \frac{1}{2})$, *or*

(ii) *assume that*

$$\mathrm{P}(\Lambda > x) = L(x)x^{-\tau+1},$$

where $\tau \in (2, 3)$ *and* $L(x)$ *is a slowly varying function (i.e.* $\lim_{N \to \infty} L(cN)/L(N) = 1$ *for every* $c > 0$*), and let* $\eta \in (0, (\tau - 2)/(\tau - 1))$.

Moreover, let $\gamma \equiv \gamma(N)$ *be a natural number such that* $\gamma(N)/N^\eta \to 0$ *as* $N \to \infty$. *Then*

$$\lim_{N \to \infty} \mathrm{P}(\nu > \gamma \mid \Lambda) = 1 \quad \text{almost surely.}$$

Proof. With $V_n = \sum_{i=0}^n \Lambda_{J_i}$, we have

$$\mathrm{P}(\nu \leq \gamma \mid \Lambda) = \sum_{n=1}^{\gamma} \mathrm{P}(\nu = n \mid \Lambda)$$

$$= \sum_{n=1}^{\gamma} \mathrm{E}[\mathrm{P}(\nu = n \mid J_0, \ldots, J_{n-1}; \Lambda) \mid \Lambda]$$

$$= \sum_{n=1}^{\gamma} \mathrm{E}[\mathrm{P}(\nu \geq n \mid J_0, \ldots, J_{n-1}; \Lambda)$$

$$\times \mathrm{P}(\nu = n \mid \nu \geq n; J_0, \ldots, J_{n-1}; \Lambda) \mid \Lambda]$$

$$\leq \sum_{n=1}^{\gamma} \mathrm{E}\left[\frac{V_{n-1}}{L_N} \,\middle|\, \Lambda\right] \leq \gamma \, \mathrm{E}\left[\frac{V_{\gamma-1}}{L_N} \,\middle|\, \Lambda\right]$$

$$\leq \frac{\gamma^2}{L_N} \mathrm{E}[\Lambda_{J_1} \mid \Lambda]$$

$$= \frac{\gamma^2}{L_N^2} \sum_{i=1}^N \Lambda_i^2. \tag{3.2}$$

The last inequality follows from the fact that $E[\Lambda_{J_0} \mid \mathbf{\Lambda}]^2 \le E[\Lambda_{J_0}^2 \mid \mathbf{\Lambda}]$. Under hypothesis (i), the last expression converges to 0 simply by the law of large numbers. Under hypothesis (ii), we choose a number $\alpha \in (1/(2(1-\eta)), (\tau-1)/2)$; since then,

$$P(\Lambda^{2\alpha} > x) = L(x^{1/(2\alpha)})x^{-(\tau-1)/(2\alpha)},$$

we have $E[\Lambda^{2\alpha}] < \infty$. Now, by (3.2) and the subadditivity of x^α, $\alpha < 1$, we find that

$$
\begin{aligned}
P(\nu(N) \le \gamma(N) \mid \mathbf{\Lambda}) &\le \frac{\gamma(N)^2}{L_N^2} \sum_{i=1}^N \Lambda_i^2 \\
&\le \frac{\gamma(N)^2}{L_N^2} \left(\sum_{i=1}^N \Lambda_i^{2\alpha} \right)^{1/\alpha} \\
&= \frac{\gamma(N)^2}{N^{2-1/\alpha}} \cdot \left(\frac{N}{L_N} \right)^2 \cdot \left(\frac{1}{N} \sum_{i=1}^N \Lambda_i^{2\alpha} \right)^{1/\alpha}.
\end{aligned}
$$

The first factor goes to 0 and the two others converge almost surely to positive constants by the law of large numbers.

4. Asymptotic architectures

Let us assume that

$$P(\Lambda > x) = x^{-\tau+1}L(x), \tag{4.1}$$

where $\tau > 1$ and $L(x)$ is a slowly varying function. In this section we show that the three parameter regions $\tau \in (1,2)$, $\tau \in (2,3)$, and $\tau > 3$ correspond to three qualitatively different asymptotic connectivity structures, or *architectures*, of the graphs G_N (we have not considered the boundary cases $\tau = 2$ and $\tau = 3$; however, Bollobás and Riordan [6] already have a result in the case $\tau = 3$ in a related model). Each of these cases is considered in its own subsection. The most intricate case is $\tau \in (2,3)$, where we present a modification of the 'soft hierarchy' proof of [17]. The other cases are much more easy to treat.

4.1. Preliminaries

We are mostly interested in properties that a large graph has *asymptotically almost surely* *(a.a.s.)*. This means the following. Let $A^{(N)}$ be the event that the random graph G_N has a certain property A. We say that A a.a.s. occurs if

$$\lim_{N \to \infty} P(A^{(N)}) = 1.$$

As an example of the use of this notion, we say that the graph process G_N *a.a.s. has a giant component* if there exists a number $a \in (0,1)$ such that the cardinality of a connected component of G_N is a.a.s. greater than aN.

As another example, if $E[\Lambda]$ is finite then the weak law of large numbers states that the total number of edges is a.a.s. approximately proportional to N. This is the context of the following lemma.

Lemma 4.1. *Assume that* $E[\Lambda] < \infty$. *For any* $\eta > 0$,

$$\sum_{i=1}^N \Lambda_i \in [(E[\Lambda] - \eta)N, (E[\Lambda] + \eta)N] \quad a.a.s.$$

It is also useful to note the following obvious facts.

Lemma 4.2. *If* $P(A^{(N)} \mid \Lambda) \to 1$ *in probability, then* $A^{(N)}$ *a.a.s. occurs.*

Lemma 4.3. *Let* $\phi, \psi : \mathbb{N} \to \mathbb{R}$ *be functions such that* $\phi(N) \to \infty$ *and* $\psi(N) \to \infty$, *and assume that the limit*

$$\lim_{N \to \infty} \frac{\psi(N)}{\phi(N)} = a \in [0, \infty]$$

exists. Then

$$\lim_{N \to \infty} \left(1 - \frac{1}{\phi(N)}\right)^{\psi(N)} = \mathrm{e}^{-a} \in [0, 1].$$

We can now turn to the graphs G_N. The following lemma shows how large the aggregated capacities of two sets of vertices must be in order a.a.s. to have a connecting edge between them. By '$\xrightarrow{\mathrm{P}}$' we denote convergence in probability.

Lemma 4.4. *For each* N, *let* $U^{(N)}$ *and* $V^{(N)}$ *be two* $\sigma\{\Lambda_1, \ldots, \Lambda_N\}$-*measurable, disjoint, random sets of vertices. Denote their aggregated capacities by*

$$F^{(N)} = \sum_{i \in U^{(N)}} \Lambda_i, \qquad G^{(N)} = \sum_{j \in V^{(N)}} \Lambda_j.$$

(i) *If* $F^{(N)} G^{(N)} / L_N \xrightarrow{\mathrm{P}} \infty$ *then* $U^{(N)} \leftrightarrow_N V^{(N)}$ *a.a.s.*

(ii) *If* $F^{(N)} G^{(N)} / L_N \xrightarrow{\mathrm{P}} 0$ *then* $U^{(N)} \nleftrightarrow_N V^{(N)}$ *a.a.s.*

(iii) *If* $\mathrm{E}[\Lambda] < \infty$ *then in the previous conditions* L_N *can be replaced by* N.

Proof. Conditioned on Λ, the number of edges joining a vertex in $U^{(N)}$ to a vertex in $V^{(N)}$, say $E_N(U^{(N)}, V^{(N)})$, has conditional distribution

$$\mathcal{L}(E_N(U^{(N)}, V^{(N)}) \mid \Lambda) = \mathrm{Poisson}\left(\frac{F^{(N)} G^{(N)}}{L_N}\right),$$

which implies the first two assertions. If $\mathrm{E}[\Lambda] < \infty$, then $L_N \in [(\mathrm{E}[\Lambda] - \eta)N, (\mathrm{E}[\Lambda] + \eta)N]$ a.a.s. for any positive η (see Lemma 4.1) and $F^{(N)} G^{(N)} / L_N$ converges in probability to ∞ and 0 in the cases (i) and (ii), respectively. It follows that the conditional probability

$$P(U^{(N)} \leftrightarrow_N V^{(N)} \mid \Lambda)$$

converges in probability to 1 and 0 in these respective cases.

4.2. The $\tau > 3$ case: finite variance

Let $\tau > 3$ and assume only that $P(\Lambda > x) < x^{-\tau+1}$ for large x. Thus, distributions with tails lighter than power-tailed are included.

Let X be a random variable with the distribution π^* (see Remark 3.2). Then $\mathrm{E}[X^\alpha]$ is finite for some $\alpha > 1$. If additionally $c := \mathrm{E}[X] > 1$, then a Galton–Watson process $Z_n^{(\pi^*)}$ with offspring distribution π^* has a positive probability of living forever. It is a natural conjecture that the graphs G_N a.a.s. have a giant component if and only if $c > 1$, and that the relative size of this component approaches

$$1 - \sum_{j=1}^{\infty} P(\mathrm{Poisson}(\Lambda) = j) \, P(Z_\infty^{(\pi^*)} = 0)^j.$$

Indeed, a corresponding result holds for the rather similar model considered in [19]. However, we leave this conjecture unproven and just show, by a short argument, that it is not possible to reach a positive fraction of the vertices by many fewer than $\log_c N$ steps. Thus, if there is a giant component, its asymptotic architecture is 'flat' compared with the 'hierarchical' structures found in the two other cases (see Subsections 4.3 and 4.4).

Proposition 4.1. *Let $\eta \in (0, 1)$. The set of vertices reachable from a randomly chosen vertex J_0 of G_N in $\lfloor \eta \log_c N \rfloor$ steps is negligible in the sense that*

$$\frac{1}{N} \sum_{n=0}^{\lfloor \eta \log_c N \rfloor} \mathrm{card}(\mathcal{N}_n(J_0)) \xrightarrow{\mathrm{P}} 0.$$

Here $\lfloor \cdot \rfloor$ denotes the greatest integer less than or equal to its argument.

Proof. Choose the distribution of a random variable $\tilde{X}$ such that (i) $\mathrm{E}[\tilde{X}^\alpha] < \infty$ for some $\alpha > 1$, (ii) $\tilde{c} := \mathrm{E}[\tilde{X}]$ satisfies $\eta \log \tilde{c} < \log c$, and (iii) the events

$$A_N = \{\mathrm{Poisson}(\Lambda_{J^{(N)}}) \leq_{\mathrm{st}} \tilde{X}\}, \qquad J^{(N)} \sim q_N,$$

occur a.a.s. This choice is possible because the mixed-Poisson random variable converges to π^* in distribution.

Fix N, and start the neighborhood process and the coupled marked branching process $(Z, \boldsymbol{J})$ as in Section 3. Given that $Z_1 = D_N(J_0)$, we can write $Z_n = \sum_{i=1}^{Z_1} Z_{n-1}^{(i)}$, $n \geq 1$, where the $Z^{(i)}$ are independent branching processes with $Z_0^{(i)} = 1$. Using a pathwise coupling argument, the processes $Z^{(i)}$ can be respectively majorized on A_N by independent branching processes $\tilde{Z}^{(i)}$ with offspring distribution $\mathcal{L}(\tilde{X})$. By standard results on branching processes, the scaled processes $\tilde{W}_n^{(i)} = \tilde{Z}_n^{(i)}/\tilde{c}^n$ are uniformly integrable martingales with integrable limit variables

$$\lim_{n \to \infty} \frac{\tilde{Z}_n^{(i)}}{\tilde{c}^n} = \tilde{W}_\infty^{(i)}.$$

By a classical martingale result, the suprema $\tilde{W}^{(i)*} = \sup_n \tilde{W}_n^{(i)}$ are almost surely finite as well. Thus, the $\tilde{Z}_n^{(i)}$ cannot grow faster than powers of $\tilde{c}$: $\tilde{Z}_n^{(i)} \leq \tilde{W}^{(i)*} \tilde{c}^n$. With $k = \lfloor \eta \log_c N \rfloor$, on A_N we have

$$\sum_{n=0}^{k} \mathrm{card}(\mathcal{N}_n(J_0)) \leq \sum_{i=1}^{D_N(J_0)} \sum_{n=0}^{k} Z_n^{(i)}$$

$$\leq \sum_{i=1}^{D_N(J_0)} \sum_{n=0}^{k} \tilde{Z}_n^{(i)}$$

$$\leq \sum_{i=1}^{D_N(J_0)} \tilde{W}^{(i)*} \frac{\tilde{c}^{k+1}}{\tilde{c} - 1}$$

$$\leq Y N^{\eta \log \tilde{c} / \log c},$$

where $Y = \sum_{i=1}^{D_N(J_0)} \tilde{W}^{(i)*} \tilde{c}/(\tilde{c}-1)$ is a finite random variable whose distribution is independent of N.

The assertion now follows, since the exponent of N in the last expression is less than 1 and since A_N occurs a.a.s.

4.3. The $\tau < 2$ case: infinite mean

Assume that

$$P(\Lambda > x) = x^{-\tau+1} L(x),$$

where $\tau \in (1, 2)$. As in [18], the asymptotic architecture is the following: there is a completely connected 'core', to which all vertices of positive degree are directly connected. In particular, there is a giant component whose relative size is $P(\text{Poisson}(\Lambda) > 0) = E[1 - \exp(-\Lambda)]$. Here is the exact result.

Theorem 4.1. *Assume that $\tau \in (1, 2)$ and fix an arbitrary number $\delta \in (\frac{1}{2}, 1)$. Then*

(i) *a.a.s., a randomly chosen vertex either has degree 0 or is directly connected to the set*

$$C_\delta := \{i \leq N \colon \Lambda_i > N^{\delta/(\tau-1)}\};$$

(ii) *two randomly chosen vertices of C_δ are a.a.s. connected with an edge.*

Proof. Since $E[\Lambda^p] < \infty$ for $p < \tau - 1$ and $E[\Lambda^p] = \infty$ for $p > \tau - 1$, a generalized law of large numbers [11, Theorem 3.23, p. 51] implies that $L_N / N^{1/p}$ converges almost surely to 0 for $p < \tau - 1$ and to ∞ for $p > \tau - 1$. Thus,

$$L_N \in (N^{(1-\eta)/(\tau-1)}, N^{(1+\eta)/(\tau-1)}) \quad \text{a.a.s.} \tag{4.2}$$

for any $\eta \in (0, 1)$. Define $\Lambda(x) := \Lambda 1_{\{\Lambda \leq x\}}$, where $1_{\{\cdot\}}$ is the indicator function. By Karamata's theorem (see, e.g. [4, Proposition 1.5.8]),

$$E[\Lambda(x)] \sim \frac{L(x)}{2 - \tau} x^{2-\tau}, \qquad E[\Lambda(x)^2] \sim \frac{2L(x)}{3 - \tau} x^{3-\tau}.$$

Using these estimates, it is straightforward to show that

$$\frac{\sum_{i=1}^{N} \Lambda_i 1_{\{\Lambda_i \leq N^{\delta/(\tau-1)}\}}}{N^{1+\delta(2-\tau)/(\tau-1)} L(N^{\delta/(\tau-1)})} \to 2 - \tau, \quad \text{as } N \to \infty,$$

in L^2, and thus also in probability. Now, since $\delta < 1$, we have

$$1 + \frac{\delta(2 - \tau)}{\tau - 1} < \frac{1}{\tau - 1}.$$

Using this together with (4.2), we see that the total capacity of the 'core' C_δ is asymptotically arbitrarily much larger than the capacity of its complement, and the first assertion follows. Since $\delta > \frac{1}{2}$, the second assertion follows from Lemma 4.4 and (4.2).

Corollary 4.1. (Cf. [18].) *If $\tau < 2$ then the distance between two randomly chosen vertices of nonzero degree is a.a.s. either 2 or 3.*

4.4. The $\tau \in (2, 3)$ case: finite mean and infinite variance

In the case $\tau \in (2, 3)$, in which the vertex capacity distribution has a finite mean but an infinite variance, we find the most interesting asymptotic architecture. As in subsection 4.3, a 'core network' plays a crucial role in the connectivity. Now, however, it is not totally connected, but has an internal structure.

Since our result is only an upper bound for the distance between two randomly selected vertices, let us assume for simplicity that

$$P(\Lambda > x) \geq x^{-\tau+1}, \qquad x \geq 1. \tag{4.3}$$

If the distribution of Λ satisfies (4.1), then a slightly larger value of τ satisfies (4.3) for large x, and only the asymptotics plays a role in our proof of the following result.

Theorem 4.2. *Let $\tau \in (2, 3)$ and assume that (4.3) holds. Let*

$$k^* \equiv k^*(N) := \left\lceil \frac{\log \log N}{-\log (\tau - 2)} \right\rceil,$$

where $\lceil \cdot \rceil$ denotes the least integer greater than or equal to its argument. Then

- *the graph G_N a.a.s. has a giant component;*

- *the distance between two randomly chosen vertices of the giant component is almost surely less than $2k^*(N)(1 + o(1))$;*

- *in probability, the relative size of the giant component approaches the value*

$$1 - \sum_{j=1}^{\infty} P(D = j) \, P(Z_\infty^{(\pi^*)} = 0)^j,$$

where D has the conditionally Poissonian distribution of Proposition 2.1 and $Z_n^{(\pi^*)}$ is a Galton–Watson branching process with offspring distribution π^*.

The rest of this section is devoted to the proof of Theorem 4.2.

Let us fix an increasing function $\ell : \mathbb{N} \to \mathbb{R}$ with the following properties:

$$\ell(1) = 1, \qquad \frac{\ell(N)}{\log \log \log N} \to 0 \quad \text{and} \quad \frac{\ell(N)}{\log \log \log \log N} \to \infty \quad \text{as } N \to \infty.$$

We will sometimes write $\log^2 x = \log \log x$, for example.

The following N-dependent definition of a 'small number' turns out to be very useful:

$$\varepsilon(N) = \frac{\ell(N)}{\log N}.$$

In particular, note that, as $N \to \infty$,

$$\varepsilon(N) \to 0 \quad \text{but} \quad N^{\epsilon(N)} = e^{\ell(N)} \to \infty.$$

4.4.1. *Reachability functions and the core.* When $\tau < 3$, vertices with high capacity turn out to play a central role in connectivity. Whereas in the case $\tau < 2$ almost all vertices of the giant component had a connection to the vertex with highest capacity with at most two steps, a less trivial hierarchical structure builds up in the case $\tau \in (2, 3)$. Loosely speaking, a typical vertex of some given large capacity $u = N^\eta$ has edges to vertices with still higher capacities, up to order $N^{\eta/(\tau-2)-\epsilon(N)}$, but not to vertices with capacities higher than $N^{\eta/(\tau-2)+\epsilon(N)}$. We first derive some bounds and limits describing the reachability of large vertices from a vertex of given capacity.

Let

$$U(x) \equiv U^{(N)}(x) = \{i \leq N : \Lambda_i > x\},$$

$$L_N(x) = \sum_{i \in U^{(N)}(x)} \Lambda_i = \sum_{i=1}^{n} \Lambda_i 1_{\{\Lambda_i > x\}},$$

and

$$\rho(x) = \frac{E[\Lambda 1_{\{\Lambda > x\}}]}{E[\Lambda]},$$

and define the *reachability functions*

$$R^{(N)}(x|u) = P(1 \leftrightarrow_N U^{(N)}(x) \mid \Lambda_1 = u),$$

$$R_2^{(N)}(x|v|u) = P(U^{(N)}(x) \leftrightarrow_N \mathcal{N}_1^{(N)}(1) \cap U^{(N)}(v)^c \mid \Lambda_1 = u).$$

Assuming that we start at a vertex with capacity u, $R^{(N)}(x|u)$ is the probability of reaching $U^{(N)}(x)$ in one step, and $R_2^{(N)}(x|v|u)$ is the probability of reaching $U^{(N)}(x)$ in two steps such that the intermediate vertex has a capacity at most v. We skip the proofs of the following limit formulae, since they will be used only in the heuristic discussion below.

Proposition 4.2. *With $u \leq v < x$, the reachability functions have the following limits:*

$$\lim_{N \to \infty} R^{(N)}(x|u) = 1 - \exp(-u\rho(x)), \tag{4.4}$$

$$\lim_{N \to \infty} R_2^{(N)}(x|v|u) = 1 - \exp\left(-\frac{u}{E[\Lambda]} E[\Lambda(1 - \exp(-\rho(x)\Lambda 1_{\{\Lambda \leq v\}}))]\right). \tag{4.5}$$

Before proving Theorem 4.2, we consider some heuristics. With $P(\Lambda > x) = x^{-\tau+1}$, we have $\rho(x) = x^{-\tau+2}$. Using (4.4) as an estimate of $R^{(N)}(x|u)$, we see that the 'upward horizon' of a vertex with capacity u is at the level $x \approx u^{1/(\tau-2)}$. In a graph of size N, the highest capacity present is about $N^{1/(\tau-1)}$ (see Lemma 4.5, below). Descending from this 'big boss', how many distinct 'reachability layers' do we have until we get down to the level of 'small bosses' with capacity $N^{\epsilon(N)}$? The condition

$$N^{(\tau-2)^k/(\tau-1)} = N^{\varepsilon(N)}$$

yields

$$k = \frac{\log \log N}{-\log(\tau-2)} - \frac{\log \log N^{\varepsilon(N)} + \log(\tau-1)}{-\log(\tau-2)},$$

where the second term is asymptotically negligible in comparison with the first. This suggests that the height of the hierarchy within $U^{(N)}(N^{\varepsilon(N)})$ is at most k^*. We call this upper part of the hierarchy the core.

In the case $\tau < 2$, the core network of the asymptotic architecture was just a clique (or at least a 'quasiclique', in the sense that its two randomly chosen vertices were a.a.s. directly connected) to which all other vertices were directly connected. In the present case, the core network is not itself a clique, but contains a clique at the top. Indeed, since $1/(\tau-1) > \frac{1}{2}$, the graph a.a.s. has vertices with capacity higher than $N^{1/2}$, and two random vertices of such degree are a.a.s. directly connected, according to Lemma 4.4.

Formula (4.5) yields an argument suggesting that the upper bound k^* is tight. With $P(\Lambda > x) = x^{-\tau+1}$, we obtain

$$R_2^{(\infty)}(x|u|u) = 1 - \exp\left(-\frac{\tau-2}{3-\tau}u^{4-\tau}x^{-\tau+2} + r(x,u)\right),$$

where $r(x,u)$ satisfies $r(x,u) \to 0$ as $u \to \infty$, and $u\rho(x) \to 0$. A simple computation shows that we reach a higher capacity vertex by going straight to the highest capacity vertices available in each step, rather than by making the first step among all neighbors with at most the same capacity. However, we do not have a full proof that the upper bound k^* could not be improved. It is plausible that the complicated analysis of [20] is transferable to our model, but simply to prove the tightness of k^*, it would be nice to have a simpler argument.

Proof of Theorem 4.2. Denote the vertex with largest capacity by

$$i^*(N) := \min\left\{i \in \{1,\ldots,N\} \colon \Lambda_i = \max_{j\in\{1,\ldots,N\}} \Lambda_j\right\}.$$

Lemma 4.5. *Assume that (4.3) holds. The largest vertex capacity satisfies*

$$P(\Lambda_{i^*(N)} \le N^{\alpha^*(N)}) \le \exp(-e^{(\tau-1)\ell(N)}), \quad \text{where } \alpha^*(N) = \frac{1}{\tau-1} - \varepsilon(N).$$

In particular, $\Lambda_{i^*(N)} > N^{\alpha^*(N)}$ a.a.s.

Proof. We have

$$\begin{aligned}
P(\max \Lambda_i \le N^{\alpha^*(N)}) &= F(N^{\alpha^*(N)})^N \\
&\le (1 - N^{-\alpha^*(N)(\tau-1)})^N \\
&\le \exp(-N^{1-\alpha^*(N)(\tau-1)}) \\
&= \exp(-e^{(\tau-1)\ell(N)}).
\end{aligned}$$

Note that $1/(\tau-1) > \frac{1}{2}$; thus, $\Lambda_{i^*} > \sqrt{N}$ a.a.s.

Lemma 4.6. *Assume that (4.3) holds, and let $\delta(N)$ satisfy*

$$\delta(N) \in (0, \alpha^*).$$

Then

$$P\left(\sum_{i=1}^N \Lambda_i 1_{\{\Lambda_i > N^{\delta(N)}\}} \le \frac{1}{2}N^{1-(\tau-2)\delta(N)}\right) \le \exp\left(-\frac{N}{8}P(\Lambda > N^{\delta(N)})\right).$$

In particular, $\sum_{i=1}^N \Lambda_i 1_{\{\Lambda_i > N^{\delta(N)}\}} > \frac{1}{2}N^{1-(\tau-2)\delta(N)}$ a.a.s.

Proof. Apply the bound (see e.g. [10, Equation (2.6), p. 26])

$$P\left(\text{bin}(n,p) \le \frac{np}{2}\right) \le \exp\left(-\frac{np}{8}\right)$$

to the binomial random variable $\sum_{i=1}^N 1_{\{\Lambda_i \ge N^{\delta(N)}\}}$, and use (4.3).

Recall the number $\alpha^*(N)$ from Lemma 4.5, and recursively define

$$\beta_0(N) = \frac{1 - \alpha^*(N)}{\tau - 2} = \frac{1}{\tau - 1} + \frac{\varepsilon(N)}{\tau - 2},$$

$$\beta_j(N) = (\tau - 2)\beta_{j-1}(N) + \varepsilon(N), \qquad j = 1, 2, \ldots.$$

Then

$$\beta_k(N) = (\tau - 2)^k \beta_0(N) + \varepsilon(N) \sum_{i=0}^{k-1} (\tau - 2)^i. \tag{4.6}$$

For sufficiently large N, we have $\beta_0(N) > \varepsilon(N)/(3 - \tau)$, and the sequence $(\beta_k(N))_{k=0,1,\ldots}$ decreases toward the limit value $\varepsilon(N)/(3 - \tau)$.

Define

$$U_0 := \{i^*(N)\}, \qquad U_j := U(N^{\beta_j(N)}), \quad j = 1, 2, \ldots, \qquad C \equiv C(N) := U_{k^*}.$$

We call $C(N)$ the *core* of the graph G_N. Lemma 4.6 tells us that

$$\sum_{i \in U_j} \Lambda_i \geq \frac{1}{2} N^{1-(\tau-2)\beta_j(N)} \quad \text{a.a.s.} \tag{4.7}$$

for all j (including $j = 0$, since $1 - (\tau - 2)\beta_0(N) = \alpha^*(N)$). From this, we obtain a lower bound of the capacity of the core.

Lemma 4.7. *Assume that (4.3) holds. Then*

$$\sum_{i \in C(N)} \Lambda_i \geq \frac{1}{2} N^{1-\theta\varepsilon(N)}, \quad \text{where } \theta := \frac{(\tau - 2)(4 - \tau)}{3 - \tau}.$$

Proof. From (4.6) with $k = k^*(N)$, and noting that

$$\frac{(\tau - 2)^{k^*(N)}\beta_0(N)}{\varepsilon(N)} \leq \frac{\beta_0(N)}{\ell(N)} \to 0,$$

for sufficiently large N we find that

$$\beta_{k^*(N)}(N) \leq \varepsilon(N)\left(1 + \sum_{i=0}^{\infty}(\tau - 2)^i\right) = \frac{4 - \tau}{3 - \tau}\varepsilon(N). \tag{4.8}$$

The claim now follows from Lemma 4.6.

We next show that a randomly chosen vertex belonging to the core can reach the largest vertex i^* in at most k^* steps. It turns out that an extremely simple search strategy works here: each time, step to the vertex with highest capacity in the neighborhood. For any vertex $i \in C$, define

$$M(i) = \inf\{j \in \{i\} \cup \mathcal{N}_1(i) \colon \Lambda_j = \sup\{\Lambda_k : k \in \{i\} \cup \mathcal{N}_1(i)\}\},$$

where, recall, $\mathcal{N}_1(\cdot)$ denotes the one-step neighborhood, and let

$$K(i) = \inf\{k \colon i \in U_k\}$$

be the 'layer' containing vertex i.

Proposition 4.3. *For each N, let $I_0 \equiv I_0(N) \in C(N)$ be a random vertex of the core such that, conditioned on $\mathbf{\Lambda}$, I_0 is independent of the edges within the core. Define the sequence of vertices*

$$I_{n+1} = M(I_n), \qquad n = 0, 1, 2, \ldots.$$

Then $I_{k^*} = i^*$ a.a.s.

Proof. Let

$$A_n = \{I_n = i^* \text{ or } K(I_{n+1}) < K(I_n)\}, \qquad n = 0, 1, 2, \ldots,$$

$$B' = \left\{ \frac{L_N}{N} \leq \mathrm{E}[\Lambda] + 1 \right\},$$

$$B_k = \left\{ \sum_{j \in U_k} \Lambda_j > \frac{1}{2} N^{1-(\tau-2)\beta_k} \right\}, \qquad k = 0, \ldots, k^* - 1,$$

$$B = B' \cap B_0 \cap \cdots \cap B_{k^*-1}.$$

It is sufficient to show that $\mathrm{P}(A_0 \cap \cdots \cap A_{k^*-1} \mid \mathbf{\Lambda})$ converges to 1 in probability. Using Bonferroni's inequality, write

$$\mathrm{P}(A_0 \cap \cdots \cap A_{k^*-1} \mid \mathbf{\Lambda}) \geq 1_B \left(1 - \sum_{n=0}^{k^*-1} \mathrm{P}(A_n^c \mid \mathbf{\Lambda}) \right),$$

and note that, by Lemmas 4.1, 4.5, and 4.6, $\mathrm{P}(B)$ approaches 1 since

$$\mathrm{P}(B_1^c \cup \cdots \cup B_{k^*-1}^c) \leq (k^* - 1) \mathrm{e}^{-N^{1-(\tau-1)\beta_1}/8} \to 0, \qquad N \to \infty.$$

On B, we have

$$\mathrm{P}(A_n^c \mid \mathbf{\Lambda}) = \mathrm{E}[\mathrm{P}(A_n^c \mid I_n; \mathbf{\Lambda}) \mid \mathbf{\Lambda}]$$

$$= \mathrm{E}\left[1_{\{I_n \neq i^*\}} \exp\left(-\frac{\Lambda_{I_n} \sum_{j \in U_{K(I_n)-1}} \Lambda_j}{L_N} \right) \,\Big|\, \mathbf{\Lambda} \right]$$

$$\leq \mathrm{E}\left[1_{\{I_n \neq i^*\}} \exp\left(-\frac{1}{(\mathrm{E}[\Lambda]+1)N} N^{\beta_{K(I_n)}} \frac{1}{2} N^{1-(\tau-2)\beta_{K(I_n)-1}} \right) \,\Big|\, \mathbf{\Lambda} \right]$$

$$\leq \exp\left(-\frac{\mathrm{e}^{\ell(N)}}{2(\mathrm{E}[\Lambda]+1)} \right).$$

The last expression does not depend on n. It remains to note that, for $c > 0$,

$$k^* \exp(-c\mathrm{e}^{\ell(N)}) = \frac{k^*}{\log^2 N} \exp(\log^3 N - c(\mathrm{e}^{\log^4 N})^{\ell(N)/\log^4 N})$$

$$= \frac{k^*}{\log^2 N} \exp(-\log^3 N(c(\log^3 N)^{\ell(N)/\log^4 N - 1} - 1))$$

$$\to 0$$

by the assumption that $\ell(N)/\log^4 N \to \infty$.

4.4.2. *Connecting arbitrary vertices to the core.* By the definition of $\varepsilon(N)$, we can choose a natural number $\kappa \equiv \kappa(N)$, depending on N, such that

$$\frac{\kappa(N)}{N^{\theta \varepsilon(N)}} \to \infty, \qquad \frac{\kappa(N)}{k^*(N)} \to 0, \qquad \text{as } N \to \infty.$$

To finish the proof of Theorem 4.2, it is now sufficient to prove that, a.a.s., a randomly chosen vertex is connected to the core with the probability given in the theorem, and that the number of steps needed to reach the core is at most κ. This is done using the coupling to a branching process, presented in Section 3. Recall the definition of ν and Δ in (3.1), and let

$$\mu \equiv \mu(N) = \inf\{n \geq 0 \colon J_n \in C(N)\},$$

i.e. the time at which an element of the core is seen in the mark sequence (J_n) for the first time (see Section 3 for our notation).

Lemma 4.8. *A.a.s.,* $\mu(N) \leq \kappa(N)$, *i.e.* $\{J_0, \ldots, J_\kappa\} \cap C(N) \neq \varnothing$.

   Proof. We have

$$P(J_1, \ldots, J_\kappa \in C(N)^c \mid \mathbf{\Lambda}) = \left(1 - \frac{\sum_{i \in C(N)} \Lambda_i}{L_N}\right)^\kappa$$

$$\leq \left(1 - \frac{N}{2L_N} N^{-\theta \varepsilon(N)}\right)^\kappa \quad \text{a.a.s.}$$

The last expression converges to 0 by the law of large numbers, because $\kappa(N) N^{-\theta \varepsilon(N)} \to \infty$ by assumption.

Lemma 4.9. *A.a.s.,* $\mu(N) > \Delta(N)$ *on* $\{\Delta(N) < \infty\}$.

   Proof. Denote by $\mathcal{F}_n$ the σ-algebra generated in the development of the branching process $(\tilde{Z}, \tilde{J})$ up to draw J_n. Since each new individual generates a new copy of the branching process, we have

$$P(\Delta(N) < \infty \mid \mathbf{\Lambda}; \mathcal{F}_{\mu(N)}) 1_{\{\mu(N) \leq \Delta(N)\}}$$

$$\leq P(\Delta_0(N) < \infty \mid \mathbf{\Lambda})^{\text{Poisson}(N^{\beta_{k^*(N)}(N)})} 1_{\{\mu(N) \leq \Delta(N)\}},$$

where Δ_0 refers to the branching process in which the mark of the first individual has, instead of the uniform distribution, the same distribution, q_N, as the others. The right-hand side converges to 0 in probability because $N^{\beta_{k^*(N)}(N)} \to \infty$, and the conditional probability $P(\Delta_0(N) < \infty \mid \mathbf{\Lambda})$ converges in probability to the constant $P(\Delta_0(\infty) < \infty)$, the extinction probability of the Galton–Watson process with the mixed-Poisson offspring distribution π^*.

   Since $\kappa(N)$ is (possibly much) smaller than the power function N^η appearing in Lemma 3.1, we obtain the following picture:

- on $\{\Delta(N) < \infty\}$, we have $\Delta(N) < \mu(N) \leq \kappa(N) < \nu(N)$ a.a.s., i.e. the neighborhood process a.a.s. does not reach the core;

- on $\{\Delta(N) = \infty\}$, we have $\mu(N) \leq \kappa(N) < \nu(N)$ a.a.s., i.e. the neighborhood process a.a.s. reaches the core in at most $\kappa(N)$ steps.

Combining this with Proposition 4.3 completes the proof of Theorem 4.2.

Acknowledgement

References

[1] AIELLO, W., CHUNG, F. AND LU, L. (2000). A random graph model for massive graphs. In *Proc. 32nd Annual ACM Symp. Theory Comput.*, ACM, New York, pp. 171–180.

[2] BARABÁSI, A.-L. AND ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286,** 509–512.

[3] BARABÁSI, A.-L. AND ALBERT, R. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74,** 47–97.

[4] BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1987). *Regular Variation*. Cambridge University Press.

[5] BOLLOBÁS, B. AND RIORDAN, O. (2003). Coupling scale-free and classical random graphs. *Internet Math.* **1,** 215–225.

[6] BOLLOBÁS, B. AND RIORDAN, O. (2004). The diameter of a scale-free random graph. *Combinatorica* **4,** 5–34.

[7] CHUNG, F. AND LU, L. (2003). The average distance in a random graph with given expected degrees (full version). *Internet Math.* **1,** 91–114.

[8] ERDÖS, P. AND RÉNYI, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.* **5,** 17–61.

[9] HOOGHIEMSTRA, G. AND VAN MIEGHEM, P. (2005). On the mean distance in scale-free graphs. *Method. Comput. Appl. Prob.* **7,** 285–306.

[10] JANSON, S., ŁUCZAK, T. AND RUCIŃSKI, A. (2000). *Random Graphs*. John Wiley, New York.

[11] KALLENBERG, O. (1997). *Foundations of Modern Probability*. Springer, New York.

[12] MOLLOY, M. AND REED, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6,** 161–179.

[13] MOLLOY, M. AND REED, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combin. Prob. Comput.* **7,** 295–305.

[14] NEWMAN, M. E. J. (2003). The size of the giant component of a random graph with a given degree sequence. *SIAM Rev.* **45,** 167–256.

[15] NEWMAN, M. E. J., STROGATZ, S. H. AND WATTS, D. J. (2001). Random graphs with arbitrary degree distribution and their applications. *Phys. Rev. E* **64,** 026118, 17pp.

[16] REITTU, H. AND NORROS, I. (2002). On the effect of very large nodes in Internet graphs. In *Globecom'02*, Vol. III (Proc. Global Telecommunications Conf., Taipei, 2002), IEEE, pp. 2624–2628.

[17] REITTU, H. AND NORROS, I. (2004). On the power law random graph model of massive data networks. *Performance Evaluation* **55,** 3–23.

[18] VAN DER ESKER, H., VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND ZNAMENSKI, D. (2004). Distances in random graphs with infinite mean degrees. To appear in *Extremes*.

[19] VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND VAN MIEGHEM, P. (2005). Random graphs with finite variance degrees. *Random Structures Algorithms* **27,** 76–123.

[20] VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND ZNAMENSKI, D. (2005). Distances in random graphs with finite mean and infinite variance degrees. Submitted.