

Random graphs with arbitrary i.i.d. degrees

Remco van der Hofstad*
Gerard Hooghiemstra[†] and Dmitri Znamenski[‡]

February 22, 2005

Abstract

In this paper we study distances and connectivity properties of random graphs with an arbitrary i.i.d. degree sequence. When the tail of the degree distribution is regularly varying with exponent $1 - \tau$ there are three distinct cases: (i) $\tau > 3$, where the degrees have finite variance, (ii) $\tau \in (2, 3)$, where the degrees have infinite variance, but finite mean, and (iii) $\tau \in (1, 2)$, where the degrees have infinite mean. These random graphs can serve as models for complex networks where degree power laws are observed. The distances between pairs of nodes in the three cases mentioned above have been studied in three previous publications, and we survey the results obtained there. Apart from the critical cases $\tau = 1$, $\tau = 2$ and $\tau = 3$, this completes the scaling picture. We explain the results heuristically and describe related work and open problems. We also compare the behavior in this model to Internet data, where a degree power law with exponent $\tau \approx 2.2$ is observed.

Furthermore, in this paper we derive results concerning the connected components and the diameter. We give a criterion when there exists a unique largest connected component of size proportional to the size of the graph, and study sizes of the other connected components. Also, we show that for $\tau \in (2, 3)$, which is most often observed in real networks, the diameter in this model grows much faster than the typical distance between two arbitrary nodes.

1 Introduction

Random graph models for complex networks have received a tremendous amount of attention in the past decade. Measurements have shown that many real networks share two properties. The first fundamental network property is the fact that typical distances between nodes are small. This is called the ‘small world’ phenomenon (see [44]). For example, in Internet, IP-packets cannot use more than a threshold of physical links, and if the distances in the graph would be large, e-mail service would simply break down. Thus, the graph of the Internet has evolved in such a way that typical distances are relatively small, even though the Internet is rather large. The second, maybe more surprising, property of many networks is that the number of nodes with degree k falls off as an inverse power of k . This is called a ‘power law degree sequence’, and resulting graphs often go under the name ‘scale-free graphs’, which refers to the fact that the asymptotics of the degree sequence is independent of the size of the graph. We refer to [4, 38, 42] and the references therein for an introduction to complex networks and many examples where the above two properties hold.

The observation that many real networks have the above two properties have incited a burst of activity in network modelling. Most of the models use random graphs as a way to model the uncertainty and the lack of regularity in real networks. In this paper, we survey some of the

*Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: rhofstad@win.tue.nl

[†]Delft University of Technology, Electrical Engineering, Mathematics and Computer Science, P.O. Box 5031, 2600 GA Delft, The Netherlands. E-mail: G.Hooghiemstra@ewi.tudelft.nl

[‡]EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: znamenski@eurandom.nl

proposals for network models. These models can be divided into two distinct types: ‘static’ models, where we model a graph of a given size as a time snap of a real network, and ‘dynamical’ models, where we model the growth of the network. Static models aim to describe real networks and their topology at a given time instant. Dynamical models aim to *explain* how the networks came to be as they are. Such explanations often focus on the growth of the network as a way to explain the power law degree sequences by means of ‘preferential attachment’ growth rules, where added nodes and edges are more likely to be attached to nodes that already have large degrees. We discuss the history of this problem and the proposed models in more detail in Section 1.7 below.

The current paper reviews recent results on the random fluctuations of the distance between two uniformly chosen nodes in a graph with i.i.d. degrees. This distance is also called the geodesic, and in Internet called the hopcount. The model where the degrees are i.i.d. is sometimes called the *configuration model* (see [38]). See [5, 8, 32] for a general introduction to random graphs, focussing on the random graph $G(N, p)$.

In a paper of the first two authors with Van Mieghem [28], the finite variance case $\tau > 3$ was investigated. In [29] and [30], the cases $\tau \in (2, 3)$ and $\tau \in (1, 2)$ were studied. In this paper, we study properties of the connected components in the random graph, and prove results concerning the scaling of the largest and second largest connected components, as well as the diameter (see Section 1.2 below).

This section is organised as follows. In Section 1.1 we start by introducing the configuration model, and in Section 1.2 review the main results concerning the distances between pairs of nodes. In Section 1.3, we discuss the new results concerning component sizes and diameter of this graph. In Section 1.4, we describe an application of our results to Internet. We describe related work in Section 1.5 and open questions in Section 1.6. Finally, we describe the preferential attachment models in Section 1.7.

1.1 The configuration model

Fix an integer N . Consider an i.i.d. sequence D_1, D_2, \dots, D_N . We will construct an undirected graph with N nodes where node j has degree D_j . We assume that $L_N = \sum_{j=1}^N D_j$ is even. If L_N is odd, then we add a stub to the N^{th} node, so that D_N is increased by 1. This single change will make hardly any difference in what follows, and we will ignore this effect. We will later specify the distribution of D_1 .

To construct the graph, we have N separate nodes and incident to node j , we have D_j stubs. All stubs need to be connected to build the graph. The stubs are numbered in a given order from 1 to L_N . We start by connecting at random the first stub with one of the $L_N - 1$ remaining stubs. Once paired, two stubs form a single edge of the graph. Hence, a stub can be seen as the left or the right half of an edge. We continue the procedure of randomly choosing and pairing the stubs until all stubs are connected. Unfortunately, nodes having self-loops may occur. However, self-loops are scarce when $N \rightarrow \infty$, as shown in [22].

The above model is a variant of the configuration model, which, given a degree sequence, is the random graph with that given degree sequence. The degree sequence of a graph is the vectors of which the k^{th} coordinate equals the fraction of nodes with degree k . In our model, by the law of large numbers, the degree sequence is close to the distribution of the nodal degree D of which D_1, \dots, D_N are i.i.d. copies. We will later comment on differences and similarities between the configuration model with fixed degrees and the configuration model with i.i.d. degrees.

The probability mass function and the distribution function of the nodal degree law are denoted by

$$\mathbb{P}(D_1 = j) = f_j, \quad j = 1, 2, \dots, \quad \text{and} \quad F(x) = \sum_{j=1}^{\lfloor x \rfloor} f_j, \quad (1.1)$$

where $\lfloor x \rfloor$ is the largest integer smaller than or equal to x . We consider distributions of the form

$$1 - F(x) = x^{-\tau+1}L(x), \quad (1.2)$$

where $\tau > 1$ and L is slowly varying at infinity. This means that the random variables D_i obey a power law, and the factor L is meant to generalize the model. We assume the following more specific conditions, splitting between the cases $\tau \in (1, 2)$, $\tau \in (2, 3)$ and $\tau > 3$.

Assumption 1.1 (i) For $\tau \in (1, 2)$, we assume (1.2).

(ii) For $\tau \in (2, 3)$, we assume that there exists $\gamma \in [0, 1)$ and $C > 0$ such that

$$x^{-\tau+1-C(\log x)^{\gamma-1}} \leq 1 - F(x) \leq x^{-\tau+1+C(\log x)^{\gamma-1}}, \quad \text{for large } x. \quad (1.3)$$

(iii) For $\tau > 3$, we assume that there exists a constant $c > 0$ such that

$$1 - F(x) \leq cx^{-\tau+1}, \quad \text{for all } x \geq 1, \quad (1.4)$$

and that $\nu > 1$, where ν is given by

$$\nu = \mathbb{E}[D_1(D_1 - 1)]. \quad (1.5)$$

Distributions satisfying (1.4) include distributions which have a lighter tail than a power law, and (1.4) is only slightly stronger than assuming finite variance. The condition in (1.3) is slightly stronger than (1.2).

1.2 Main results for the graph distance

We define the graph distance H_N between the nodes 1 and 2 as the minimum number of edges that form a path from 1 to 2. By convention, the distance equals ∞ if 1 and 2 are not connected. Observe that the distance between two randomly chosen nodes is equal in distribution to H_N , because the nodes are exchangeable. In order to state the main result concerning H_N , we define the centering constant

$$m_{\tau,N} = \begin{cases} 0 & \text{for } \tau \in (1, 2), \\ 2 \frac{\log \log N}{|\log(\tau-2)|} & \text{for } \tau \in (2, 3), \\ \log_{\nu} N & \text{for } \tau > 3. \end{cases} \quad (1.6)$$

The parameter $m_{\tau,N}$ describes the asymptotic growth of H_N as $N \rightarrow \infty$. A more precise result including the random fluctuations around $m_{\tau,N}$ is formulated in the following theorem.

Theorem 1.2 (The fluctuations of the graph distance) *When Assumption 1.1 holds, then there exist random variables $(R_{a,b})_{a \in (-1,0], b \in \{0,1\}}$ such that, as $N \rightarrow \infty$,*

$$\mathbb{P}\left(H_N = \lfloor m_{\tau,N} \rfloor + l \mid H_N < \infty\right) = \mathbb{P}(R_{a_N, b_N} = l) + o(1), \quad (1.7)$$

where $a_N = \lfloor m_{\tau,N} \rfloor - m_{\tau,N} \in (-1, 0]$ and $b_N = 1$ if $\lfloor m_{\tau,N} \rfloor$ is odd, and 0 otherwise.

Corollary 1.3 (Convergence in distribution along subsequences) *When Assumption 1.1 holds, along sequences N_k where a_{N_k} converges to a and $b_{N_k} = b$, as $k \rightarrow \infty$, the random variables*

$$H_{N_k} - \lfloor m_{\tau,N} \rfloor \mid H_{N_k} < \infty \quad (1.8)$$

converges in distribution to $R_{a,b}$.

Corollary 1.4 (Concentration of the hopcount) *When Assumption 1.1 holds, then the random variables $H_N - m_{\tau,N}$, given that $H_N < \infty$, form a tight sequence, i.e.,*

$$\lim_{K \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{P}\left(|H_N - m_{\tau,N}| \leq K \mid H_N < \infty\right) = 1. \quad (1.9)$$

We next describe the laws of the random variables $(R_{a,b})_{a \in (-1,0], b \in \{0,1\}}$. For this, we need some further notation from branching processes. For $\tau > 2$, we introduce a *delayed* branching process $\{\mathcal{Z}_k\}_{k \geq 1}$, where in the first generation the offspring distribution is chosen according to (1.1) and in the second and further generations the offspring is chosen in accordance to g given by

$$g_j = \frac{(j+1)f_{j+1}}{\mu}, \quad j = 0, 1, \dots, \quad \text{where} \quad \mu = \mathbb{E}[D_1]. \quad (1.10)$$

When $\tau \in (2, 3)$, the branching process $\{\mathcal{Z}_k\}$ has infinite expectation. Under Assumption 1.1, it is proved in [21] that

$$\lim_{n \rightarrow \infty} (\tau - 2)^n \log(\mathcal{Z}_n \vee 1) = Y \quad \text{a.s.}, \quad (1.11)$$

where $x \vee y$ denotes the maximum of x and y .

When $\tau > 3$, the process $\{\mathcal{Z}_k / \mu \nu^{k-1}\}$ is a martingale with uniformly bounded expectation and consequently

$$\lim_{n \rightarrow \infty} \frac{\mathcal{Z}_n}{\mu \nu^{n-1}} = \mathcal{W} \quad \text{a.s.} \quad (1.12)$$

Theorem 1.5 (The limit laws) *When Assumption 1.1 holds, then*

(i) *for $\tau \in (1, 2)$ and for every $a \in (-1, 0], b \in \{0, 1\}$,*

$$\mathbb{P}(R_{a,b} = 2) = 1 - \mathbb{P}(R_{a,b} = 3) = p, \quad (1.13)$$

where $p = p_\tau \in (0, 1)$ is independent of a and b .

(ii) *for $\tau \in (2, 3)$ and for $a \in (-1, 0], b \in \{0, 1\}$,*

$$\mathbb{P}(R_{a,b} > l) = \mathbb{P}\left(\min_{s \in \mathbb{Z}} [(\tau - 2)^{-s} Y^{(1)} + (\tau - 2)^{s - c_l + b} Y^{(2)}] \leq (\tau - 2)^{\lceil (l+b)/2 \rceil + a} \mid Y^{(1)} Y^{(2)} > 0\right),$$

where $c_l = 1$ if l is even, and zero otherwise, and $Y^{(1)}, Y^{(2)}$ are two independent copies of the limit random variable in (1.11).

(iii) *for $\tau > 3$ and for $a \in (-1, 0], b \in \{0, 1\}$,*

$$\mathbb{P}(R_{a,b} > k) = \mathbb{E}\left[\exp\{-\kappa \nu^{a+k} \mathcal{W}^{(1)} \mathcal{W}^{(2)}\} \mid \mathcal{W}^{(1)} \mathcal{W}^{(2)} > 0\right], \quad (1.14)$$

where the limit does not depend on b , $\mathcal{W}^{(1)}$ and $\mathcal{W}^{(2)}$ are two independent copies of the limit variable \mathcal{W} in (1.12) and where $\kappa = \mu(\nu - 1)^{-1}$.

The above results prove that the scaling in these random graphs is quite sensitive to the degree exponent τ . The scaling of the distance between pairs of nodes is proved for all $\tau > 1$, except for the critical cases $\tau = 2$ and $\tau = 3$. The result for $\tau \in (1, 2)$ is proved in [30], the result for $\tau \in (2, 3)$ in [29], and the result for $\tau > 3$ in [28]. For $\tau \in (2, 3)$, the result was formulated slightly differently in [29]. Indeed, in [29], (1.7) was proved with $2 \left\lfloor \frac{\log \log N}{|\log(\tau-2)|} \right\rfloor$ instead of $\lfloor m_{\tau,N} \rfloor = \left\lfloor 2 \frac{\log \log N}{|\log(\tau-2)|} \right\rfloor$ on the right-hand side. Since

$$\left\lfloor 2 \frac{\log \log N}{|\log(\tau-2)|} \right\rfloor - 2 \left\lfloor \frac{\log \log N}{|\log(\tau-2)|} \right\rfloor = b_N, \quad (1.15)$$

this change amounts to replacing l by $l + b_N$.

In Section 2 below, we give a heuristic explanation of the scaling in Theorem 1.2 by comparing the shortest path tree from a given point to a branching process.

1.3 Connected components and diameter of the random graph

In this paper, we also prove results concerning the sizes of the connected components in the random graph, and give bounds on the diameter. These results are stated in this section.

Our main result describes the size of the unique largest connected component and the maximal size of all other connected components. In its statement, we write G for the random graph with degree distribution given by (1.1), and we denote for $\tau > 2$ the survival probability of the delayed branching process $\{\mathcal{Z}_k\}$ described in Section 2 by q . When $1 < \tau < 2$ and hence $\mathbb{E}[D] = \infty$, we define $q = 1$. We define, for $\delta > 0$,

$$\gamma_1^* = \frac{1 + \delta}{\log \mu - \log 2} \quad (\tau > 2), \quad \gamma_2^* = \frac{\tau - 1}{2 - \tau}(1 + \delta) \quad (\tau \in (1, 2)). \quad (1.16)$$

In the sequel we use the abbreviation **whp** to denote that a statement holds with probability $1 - o(1)$ as $N \rightarrow \infty$.

Theorem 1.6 (The giant component) *Fix $\delta > 0$. When Assumption 1.1 holds and $q \in (0, 1]$, then, **whp**, the largest connected component in G has $qN(1 + o(1))$ nodes, and all other connected components have at most γ_2^* nodes when $\tau \in (1, 2)$, and at most $\gamma_1^* \log N$ nodes when $\tau > 2$ and $\mu > 2$.*

In [28], a similar result was proved for the case when $\tau > 3$, using the results of [36, 37], without the assumption that $\mu > 2$. In this case, the main restriction is that $\nu > 1$. The result in [36, 37] can be restated as saying that the largest component is $qN(1 + o(1))$ when $\nu > 1$ and is $o(N)$ when $\nu < 1$. Our result applies in certain cases where the results of [36, 37] do not apply (such as the cases when $\tau \in (2, 3)$ and $\tau \in (1, 2)$), and our proof is relatively simple and yields rather explicit bounds.

The following theorem says that γ_1^* and γ_2^* defined in (1.16) provide quite a sharp estimate for the size of the connected components that are not the largest in the random graph. Define for $f_1 > 0$,

$$\gamma_1^{**} = \frac{1 - \delta}{\log \mu - \log f_1} \quad (\tau > 2), \quad \gamma_2^{**} = \frac{\tau - 1}{2 - \tau}(1 - \delta) \quad (\tau \in (1, 2)). \quad (1.17)$$

Theorem 1.7 (Sizes of non-giant components) (i) *Let $\tau \in (1, 2)$ and $f_1 > 0$. Then, for any $\delta > 0$ and $k \leq \gamma_2^{**}$, and such that $f_k > 0$, **whp** the random graph contains a connected component with $k + 1$ nodes.*

(ii) *Let $\tau > 2$ and $\mu > f_1 > 0$, and assume that*

$$f_k = L_f(k)k^{-\tau}, \quad k \rightarrow \infty, \quad (1.18)$$

*where $L_f(k)$ is a slowly varying function of k . Then, for any $\delta > 0$, and $k = k_N \leq \gamma_1^{**} \log N$ and such that $f_k > 0$, **whp** the random graph contains a connected component with $k + 1$ nodes.*

We present some further results in the more special case when $q = 1$. In this case, either $\mu = \infty$ or $D \geq 2$ a.s. Then, from Theorem 1.6, we have that there exists a unique connected component of size $N - o(N)$ and all other connected components are much smaller. We now investigate when the random graph is **whp** connected. Before stating the theorem, we introduce some notation. Let \mathbf{C}_N denote the number of nodes in the complement of the largest connected component of the random graph.

Theorem 1.8 (Size of complement of giant component) *Let $D \geq 2$ a.s. and $\mu > 2$. Then, there exists a $c < 1$ and $c > 0$ such that*

$$\mathbb{P}(\mathbf{C}_N \geq k) \leq ca^k. \quad (1.19)$$

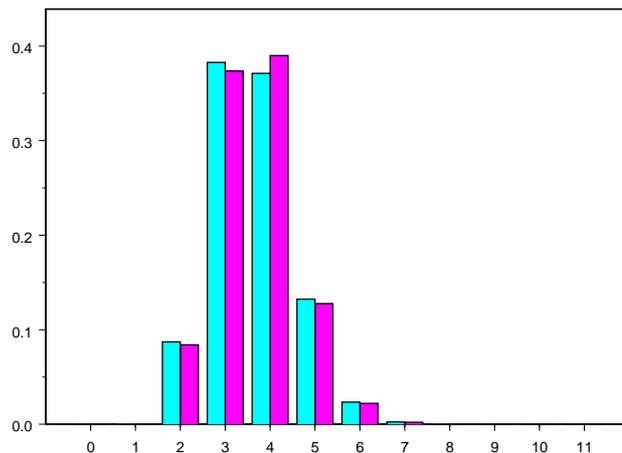


Figure 1: Histograms of the AS-count and graph distance in the configuration model with $N = 10,940$, where the degrees have generating function $f_\tau(s) = \mathbb{E}[s^D] = \frac{(\tau-1)s + (1-s)^{\tau-1} - 1}{\tau-2}$, for which the power law exponent τ takes the value $\tau = 2.25$. The AS-data is lightly shaded, the simulation is darkly shaded.

If, in addition, $\mu = \infty$, then

$$\lim_{N \rightarrow \infty} \mathbb{P}(\mathbf{C}_N = 0) = 1. \quad (1.20)$$

Consequently, in the latter case, the random graph is connected **whp**. The same conclusion holds when L_N/N^2 converges in probability to infinity and the restriction $D \geq 2$ a.s. is dropped.

We close this section by giving lower bounds on the diameter of the graph, which we define as the largest distance between any two nodes that are connected:

Theorem 1.9 (Lower bound on diameter) For $\tau > 2$, assuming that $f_1 > 0$ and $f_2 > 0$, there exists a positive constant α such that **whp** the diameter of G is bounded below by $\alpha \log N$, as $N \rightarrow \infty$.

The result in Theorem 1.9 is most interesting in the case when $\tau \in (2, 3)$. Indeed, by Theorem 1.2, the typical distance in this case is proportional to $\log \log N$, whereas the diameter is bounded below by a constant times $\log N$. The pairs of nodes where the distance is of the order $\log N$ are thus scarce. We comment further on this in Section 1.7.

1.4 Comparison to Internet data

Our work was triggered by the Internet. In a seminal paper [26], Faloutsos *et al.* have shown that the degree distribution in Internet follows a power law with exponent $\tau \approx 2.16 - 2.25$. Thus, the power law random graph with this value of τ can possibly lead to a good Internet model. We study the above version of the configuration model to describe the topology of the Internet at a fixed time instant. In [43], and inspired by the observed power law degree sequence in [26], the degree graph is proposed as a model for the network of *autonomous systems*. In this graph, the nodes are the autonomous systems in the Internet, i.e., the parts of the Internet controlled by a single party (such as a university, company or provider), and the links represent the physical connections between the different autonomous systems. The work of Faloutsos *et al.* in [26] was among others on this graph which at that time had size approximately 10,000.

In [43], it is argued on a qualitative basis that the power law random graph serves as a better model for the Internet topology than the currently used topology generators. Our results can be seen as a step towards the quantitative understanding of whether the AS-count in Internet is described well by the average graph distance in the configuration model. The AS-count gives the number of physical links connecting the various autonomous domains between two randomly chosen nodes in the graph.

To validate the model, we compare a simulation of the distribution of the distance between pairs of nodes in the power law random graph with the same value of N and τ to extensive measurements of the AS-count in Internet. In Figure 1, we see that AS-count in the model with the predicted value of $\tau = 2.25$ and the value of N from the data set fits the data remarkably well.

1.5 Related work for static models

The model studied in this paper with $\tau \in (2, 3)$ is also studied in [41], who prove that **whp**, H_N is less than $m_{\tau,N} + 2\kappa(N)$, where

$$\kappa(N) = \left\lceil \exp\left(\frac{2}{3-\tau}\ell(N)\right) \right\rceil \quad \text{with} \quad \lim_{N \rightarrow \infty} \frac{\ell(N)}{\log \log \log N} = \infty. \quad (1.21)$$

At approximately the same moment the $\log \log N$ -scaling result appeared in the physics literature [18], where it was derived in a non-rigorous way.

There is substantial work on random graphs that are, although different from ours, still similar in spirit. In [1], random graphs were considered with a degree sequence that is *precisely* equal to a power law, meaning that the number of nodes with degree k is precisely proportional to $k^{-\tau}$. Aiello *et al.* show that the largest component is of order the size of the graph when $\tau < \tau_0 = 3.47875\dots$, where τ_0 is the solution of $\zeta(\tau-2) - 2\zeta(\tau-1) = 0$ and where ζ is the Riemann Zeta function. When $\tau > \tau_0$, the largest connected component is of smaller order than the size of the graph and more precise bounds are given for the largest connected component. When $\tau \in (1, 2)$, the graph is almost surely connected. The proofs of these facts use couplings with branching processes and strengthen previous results due to Molloy and Reed [36, 37]. For this same model, Dorogovtsev *et al.* [23, 24] investigate the leading asymptotics and the fluctuations around the mean of the distance between arbitrary nodes in the graph from a theoretical physics point of view, using mainly generating function terminology.

A second related model can be found in [15, 16], where edges between nodes i and j are present with probability equal to $w_i w_j / \sum_l w_l$ for some ‘expected degree vector’ $w = (w_1, \dots, w_N)$ for which $\max_i w_i^2 < \sum_l w_l$. Chung and Lu show that when w_i is proportional to $i^{-\frac{1}{\tau-1}}$, then the average distance between pairs of nodes is $\log_\nu N(1 + o(1))$ when $\tau > 3$, and $\frac{\log \log N}{|\log(\tau-2)|}(1 + o(1))$ when $\tau \in (2, 3)$. The difference between this model and ours is that the nodes are not exchangeable in [15], but the observed phenomena are similar. This can be heuristically understood as follows. Firstly, the actual degree vector in [15] should be close to the expected degree vector. Secondly, for the expected degree vector, we can compute that the number of nodes for which the degree is greater than or equal to k equals

$$|\{i : w_i \geq k\}| \approx |\{i : ci^{-\frac{1}{\tau-1}} \geq k\}| = c^{\tau-1} k^{1-\tau},$$

where the proportionality constant c depends on N . Thus, one expects that the number of nodes with degree at least k decreases as $k^{1-\tau}$, similarly as in our model. In [16], Chung and Lu study the sizes of the connected components in the above model. The advantage of this model is that the edges are present *independently*, which makes the resulting graph closer to a traditional random graph. A related model of this form is studied in [40], where the number of edges between i and j is equal to a Poisson random variable with parameter $\Lambda_i \Lambda_j / L_N$, where $\{\Lambda_i\}_{i=1}^N$ are i.i.d. random variables with a power law distribution as in (1.2) for $\tau \in (2, 3)$. In fact, the above model can

be constructed as a *growing* graph, by adding nodes and edges, and removing edges independently with the correct probability. The results proved in [40] are identical to the ones proved in [41].

In [17], a so-called *hybrid* model is studied. In this hybrid model, a disjoint union is taken of a *local* graph and a *global* graph. The local graph is supposed to describe the geometry present in the graph, while the global graph describes the long connections. The global graph is the model in [15, 16], whereas the local graph is such that it models the local structure well. Then, the main results in [17] imply that the hybrid graph inherits the distance and diameter properties from the global graph, whereas it inherits the local structure, such as its clustering behaviour (see also Section 1.6) from the local graph. In some cases, the diameter of the hybrid graph also depends on the so-called *isoperimetric dimension* of the local graph.

Arratia and Liggett [6] study whether *simple* graphs exist with an i.i.d. degree distribution, i.e., graphs without self-loops and multiple edges. It is not hard to see that when $\tau < 2$ this happens with probability 0 (since the largest degree is larger than N). When $\tau > 2$, however, this probability is asymptotic to the probability that the sum of N i.i.d. random variables is even, which is close to $1/2$. When $\tau = 2$, the probability can converge to any element of $[0, \frac{1}{2}]$, depending on the slowly varying function in (1.2). A similar problem is addressed in [22], where various ways how self-loops and multiple edges can be avoided are discussed. Among others, in [22], it is proved that when the degrees are i.i.d. and all self-loops and multiple edges are removed, then the power law degree sequence remains valid.

1.6 Open problems in static models

There are many important questions remaining in the configuration model. For instance, in [28], we have shown that for $\tau > 3$, the largest connected component has size qN , where q is the survival probability of the delayed branching process. All other connected components have size at most $\gamma \log N$, for some $\gamma > 0$. For $\tau \in (2, 3)$, such a result is given in this paper under the extra assumption that $\mu > 2$. It would be of interest to investigate whether the same result holds for $\tau \leq 3$ and general μ when $q > 0$.

A second quantity of interest is the *diameter* of the graph which is important in many applications. For instance, in Internet, a message is killed when the number of hops exceeds a finite threshold. Thus, it would be interesting to investigate how the diameter grows with the size of the graph. The result in Theorem 1.9 is a lower bound in the case when $f_2 > 0$ only, and a better understanding of the diameter is necessary.

Other open problems involve the critical cases $\tau = 2$ and $\tau = 3$. We believe that for these values, there are again different cases depending on the slowly varying function L in (1.2). This, however, is likely to be a hard problem.

In Theorem 1.2, we have given a description of the law of the hopcount H_N when N is large. In practice, often the mean and variance of such random variables are given (rather than their complete distribution), and it would be of interest to prove that

$$\mathbb{E}[H_N | H_N < \infty] = \lfloor m_{\tau, N} \rfloor + \mathbb{E}[R_{a_N, b_N}] + o(1), \quad \text{and} \quad \text{Var}(H_N | H_N < \infty) = \text{Var}(R_{a_N, b_N}) + o(1). \quad (1.22)$$

Our methods do not imply this convergence of means and variances. In [31], the expected value $\mathbb{E}[R_{a, b}] = \mathbb{E}[R_a]$ is numerically approximated when $\tau > 3$, and simulations confirm the above asymptotics of mean and variance.

An important property of the topology of a graph is its *clustering*, which basically describes how likely two nodes that have an edge to a common node are to be connected by an edge. In general, in random graphs, this clustering is much smaller than the clustering in real networks. It would be of interest to investigate graphs with a higher clustering in more detail. The hybrid graphs in [17] are an important step in that direction.

Further interesting problems arise in the different variants of the configuration model. For example, it would be interesting to see whether the results in Theorems 1.2 and 1.5 are also valid

for the configuration model where the number of nodes with degree k is in some sense close to the *deterministic value* Nf_k , or for the ‘expected degree’ model in [15, 16]. Such results would prove that the distances in power law random graphs are not very sensitive to the precise details of the model.

Finally, in the engineering and networking community, the question has arisen whether the methodology in measuring of the degree sequences (for example, `traceroute` in Internet) has an effect on the observation of power laws in real networks. In [20], it is suggested that by performing `traceroute`-like measurements on the degrees in a classical $G(N, p)$ random graph, the degree distribution may *appear* to be of power-law type (with $\tau = 1$), whereas it is well known that the degree distribution in $G(N, p)$ is a binomial distribution. In particular, this observation sheds doubts on the accuracy of the estimates of the powers observed in real networks through extensive measurements. The influence of measurement methodology deserves to be investigated in more detail.

1.7 Growth models and preferential attachment

The models described above are *static*, i.e., the size of the graph is *fixed*, and we have not modeled the *growth* of the graph. As described in the introduction, there is a large body of work investigating *dynamical* models for complex networks, often in the context of the World-Wide Web [14, 34, 35]. In various forms, preferential attachment has been shown to lead to power law degree sequences. Therefore, such models intend to *explain* the occurrence of power law degree sequences in random graphs. See [2, 3, 4, 9, 10, 11, 12, 13, 19, 35] and the references therein. We will now discuss this model, and some of its variations, in some detail.

In the preferential attachment model, nodes with a fixed degree $m \geq 1$ are added sequentially. Their edges are attached to a receiving node with a probability proportionally to the degree of the receiving node, thus favoring nodes with large degrees. For this model, it is shown in [13] that the number of nodes with degree k decays proportionally to k^{-3} . Furthermore, the diameter is of order $\frac{\log N}{\log \log N}$, when $m \geq 2$ [9], giving an upper bound on the typical distances between pairs of nodes in the graph. Couplings to a classical random graph $G(N, p)$ for an appropriately chosen p are given in [11]. See also [10] for a survey.

The Albert-Barabási model was introduced to explain power laws occurring in real networks, but was not seriously considered as a good model for a real network. Since nodes are added with fixed degree m , there is too much regularity in this graph. For example, the total degree is always equal to $2mN$. Therefore, the question of whether power laws occur for more general preferential attachment models is particularly relevant. This problem was taken up by Cooper and Frieze [19], who show that for a very general class of preferential attachment models the degree sequence obeys a power law, and the exponent depends sensitively on the precise way in which nodes are attached. In [12], a model was proposed for a directed graph with different power exponents for the in- and out degrees. The proof in [12] is simpler than the one in [19], and it would be of interest to investigate whether the arising model can serve to describe the Web graph, which is directed.

We finally describe the relation between the configuration model, and the preferential attachment models. In [4, 38, 39], it is suggested that our model is a snapshot of the preferential attachment models, i.e., a realization of the graph growth process at the time instant that the graph has a certain prescribed size. Thus, rather than to describe the growth of the model, we investigate the properties of the model that is present at a given time instant (see e.g., [4, Section VII.D]). We now discuss these relations in some detail.

There are clearly substantial differences between the configuration model and the general preferential attachment models. One difference arises in the study of the diameter. When $\tau > 2$, the diameter in the configuration model is of the order $\log N$ by Theorem 1.9. In the Albert-Barabási model, on the other hand, the diameter grows proportionally to $\frac{\log N}{\log \log N}$. Thus, at least at the level of diameters, the two models have different behaviour. In order to understand the similarities

and differences between the two models, it would be useful to study the hopcount distribution in preferential attachment models, and to compare this law to the law for the configuration model described in Section 1.2. Also from a practical point of view, the typical distances between pairs of nodes in the network is relevant, since such distances can often be measured rather simply (e.g., by `traceroute` in Internet).

1.8 Organization of the paper

The paper is organized as follows. We first give heuristics for the scaling of the hopcount in Theorems 1.2 and 1.5 by comparing the growth of the shortest path graph from any node to a branching process in Section 2. In Section 3, we prove Theorem 1.8, and in Section 4, we prove Theorem 1.6. Finally, in Section 5, we prove the lower bounds on the second largest cluster in Theorem 1.7 and on the diameter in Theorem 1.9.

2 Heuristic explanations and connections to branching processes

In this section, we present a heuristic explanation of Theorems 1.2 and 1.5. There are two basic ingredients underlying these results. The first one is that for two disjoint sets of stubs of sizes n and m out of a total of L , the probability that none of the stubs in the first set is attached to a stub in the second set, is approximately equal to

$$\prod_{i=0}^{n-1} \left(1 - \frac{m}{L - n - 2i}\right). \quad (2.1)$$

In fact, the product in (2.1) is precisely equal to the probability that none of the n stubs in the first set of stubs is attached to a stub in the second set, given that no two stubs in the first set are attached to one another. When $n = o(L)$, $L \rightarrow \infty$, however, these two probabilities are asymptotically equal. We approximate (2.1) further as

$$\prod_{i=0}^{n-1} \left(1 - \frac{m}{L - n - 2i}\right) \approx \exp \left\{ \sum_{i=0}^{n-1} \log \left(1 - \frac{m}{L} \left(1 + \frac{n + 2i}{L}\right)\right) \right\} \approx e^{-\frac{mn}{L}}, \quad (2.2)$$

where the approximation is valid as long as $nm(n + m) = o(L^2)$, when $L \rightarrow \infty$.

The *shortest path tree* (SPG) from node 1 is the random graph as observed from node 1, and consists of the shortest paths between node 1 and all other nodes $\{2, \dots, N\}$. We define the SPG from node 2 in a similar fashion. We apply the above heuristic asymptotics to the growth of the SPG. Let $Z_j^{(1)}$ denote the number of stubs that are attached to nodes precisely $j - 1$ steps away from node 1, and similarly for $Z_j^{(2)}$. We then apply (2.2) to $n = Z_j^{(1)}$, $m = Z_j^{(2)}$ and $L = L_N$. Let $\mathbb{Q}_Z^{(k,l)}$ be the conditional probabilities given $\{Z_s^{(1)}\}_{s=1}^k$ and $\{Z_s^{(2)}\}_{s=1}^l$. For $l = 0$, we only condition on $\{Z_s^{(1)}\}_{s=1}^k$. For $j \geq 1$, we have the multiplication rule (see [28, Lemma 4.1]),

$$\mathbb{P}(H_N > j) = \mathbb{E} \left[\prod_{i=2}^{j+1} \mathbb{Q}_Z^{(\lceil i/2 \rceil, \lfloor i/2 \rfloor)}(H_N > i - 1 | H_N > i - 2) \right]. \quad (2.3)$$

where $\lceil x \rceil$ is the smallest integer greater than or equal to x and $\lfloor x \rfloor$ the largest integer smaller than or equal to x . Now we approximate

$$\mathbb{Q}_Z^{(\lceil i/2 \rceil, \lfloor i/2 \rfloor)}(H_N > i - 1 | H_N > i - 2) \approx \exp \left\{ -\frac{Z_{\lceil i/2 \rceil}^{(1)} Z_{\lfloor i/2 \rfloor}^{(2)}}{L_N} \right\}. \quad (2.4)$$

This asymptotic identity follows because the event $\{H_N > i - 1 | H_N > i - 2\}$ occurs precisely when none of the stubs $Z_{\lfloor i/2 \rfloor}^{(1)}$ attaches to one of those of $Z_{\lfloor i/2 \rfloor}^{(2)}$. A typical value of the hopcount H_N is the value j for which

$$\frac{1}{L_N} \sum_{i=2}^{j+1} Z_{\lfloor i/2 \rfloor}^{(1)} Z_{\lfloor i/2 \rfloor}^{(2)} \approx 1.$$

This is the first ingredient of the heuristic.

The second ingredient is the connection to branching processes. Given any node i and a stub attached to this node, we attach the stub to a second stub to create an edge of the graph. This chosen stub is attached to a certain node, and we wish to investigate how many further stubs this node has. The conditional probability that this number of stubs, excluding the stub that is attached to the stub at node i , equals n given D_1, \dots, D_N , is approximately equal to the probability that a random stub from all $L_N = D_1 + \dots + D_N$ stubs is attached to a node with in total $n + 1$ stubs. Since there are precisely $\sum_{j=1}^N (n + 1) I[D_j = n + 1]$ stubs that belong to a node with degree $n + 1$, we find for the latter probability

$$g_n^{(N)} = \frac{n + 1}{L_N} \sum_{j=1}^N I[D_j = n + 1], \quad (2.5)$$

where $I[A]$ denotes the indicator function of the event A . The above formula comes from sampling with replacement, whereas in the SPG the sampling is performed without replacement. Now, as we grow the SPGs from nodes 1 and 2, of course the number of stubs that can still be chosen decreases. However, when the size of both SPGs is much smaller than N , for instance at most \sqrt{N} , or slightly bigger, this dependence can be neglected, and it is as if we choose each time *independently* and with replacement. Thus, the growth of the SPGs is closely related to a *branching process* with offspring distribution $\{g_n^{(N)}\}$.

When $\tau > 2$, using the strong law of large numbers for $N \rightarrow \infty$,

$$\frac{L_N}{N} \rightarrow \mu = \mathbb{E}[D], \quad \text{and} \quad \frac{1}{N} \sum_{j=1}^N I[D_j = n + 1] \rightarrow f_{n+1} = \mathbb{P}(D = n + 1),$$

so that, almost surely,

$$g_n^{(N)} \rightarrow \frac{(n + 1)f_{n+1}}{\mu} \equiv g_n, \quad N \rightarrow \infty. \quad (2.6)$$

Therefore, the growth of the shortest path tree should be well described by a branching process with offspring distribution $\{g_n\}$, and we come to the question what is a typical value of j for which

$$\sum_{i=2}^{j+1} Z_{\lfloor i/2 \rfloor}^{(1)} Z_{\lfloor i/2 \rfloor}^{(2)} = L_n \approx \mu N, \quad (2.7)$$

where $\{Z_j^{(1)}\}$ and $\{Z_j^{(2)}\}$ denote two independent copies of a delayed branching process with offspring distribution $\{f_n\}$, $f_n = \mathbb{P}(D = n)$, $n = 1, 2, \dots$, in the first generation and offspring distribution $\{g_n\}$ in all further generations.

To answer this question, we need to make separate arguments depending on the value of τ . When $\tau > 3$, then $\nu = \sum_n n g_n < \infty$. Assume also that $\nu > 1$, so that the branching process is supercritical. In this case, the delayed branching process $Z_j / \mu \nu^{j-1}$ converges almost surely to a random variable \mathcal{W} . Hence for the two *independent* branching processes $\{Z_j^{(i)}\}$, $i = 1, 2$, that locally describe the number of stubs attached to nodes on distance $j - 1$, we find that for $j \rightarrow \infty$,

$$Z_j^{(i)} \sim \mu \nu^{j-1} \mathcal{W}^{(i)}. \quad (2.8)$$

This explains why the average value of $Z_j^{(i)}$ grows like $\mathbb{E}[Z_j^{(i)}] = \mu\nu^{j-1} = \mu \exp((j-1)\log\nu)$, that is exponential in j for $\nu > 1$, so that a typical value of j for which (2.7) holds satisfies

$$\mu \cdot \nu^{j-1} = N, \quad \text{or} \quad j = \log_\nu(N/\mu) + 1.$$

We can extend this argument to describe the fluctuation around the asymptotic mean. Since (2.8) describes the fluctuations of $Z_j^{(i)}$ around the mean value $\mu\nu^{j-1}$, we are able to describe the random fluctuations of H_N around $\log_\nu N$.

When $\tau \in (2, 3)$, the branching processes $\{\mathcal{Z}_j^{(1)}\}$ and $\{\mathcal{Z}_j^{(2)}\}$ are well-defined, but they have infinite mean. Under certain conditions on the underlying offspring distribution, Davies [21] proves for this case that $(\tau-2)^j \log(1 + \mathcal{Z}_j)$ converges almost surely, as $j \rightarrow \infty$, to some random variable Y . Moreover, $\mathbb{P}(Y = 0) = 1 - q$, the extinction probability of $\{\mathcal{Z}_j\}_{j=0}^\infty$. Therefore, also $(\tau-2)^j \log(\mathcal{Z}_j \vee 1)$ converges almost surely to Y .

Since $\tau > 2$, we still have that $L_N \approx \mu N$, so that the typical value of j for which (2.7) holds satisfies

$$\mathcal{Z}_{\lceil (j+1)/2 \rceil}^{(1)} \mathcal{Z}_{\lfloor (j+1)/2 \rfloor}^{(2)} \approx \mu N, \quad \text{or} \quad \log(\mathcal{Z}_{\lceil (j+1)/2 \rceil}^{(1)} \vee 1) + \log(\mathcal{Z}_{\lfloor (j+1)/2 \rfloor}^{(2)} \vee 1) \approx \log N.$$

This indicates that the typical value of j is of order

$$m_{\tau, N} = 2 \frac{\log \log N}{|\log(\tau - 2)|}, \quad (2.9)$$

as formulated in Theorem 1.2(ii), since for some $c \in (0, 1)$

$$\log(\mathcal{Z}_{\lceil (j+1)/2 \rceil}^{(1)} \vee 1) \approx c \log N, \quad \log(\mathcal{Z}_{\lfloor (j+1)/2 \rfloor}^{(2)} \vee 1) \approx (1 - c) \log N$$

so that $(j+1)/2 = \log(c \log N) / |\log(\tau - 2)|$, which induces leading order of $m_{\tau, N}$ stated above. Again we stress that, since Davies result [21] describes a distributional limit, we are able to describe the random fluctuations of H_N around $m_{\tau, N}$.

When $\tau \in (1, 2)$, the sequence g_n is not defined since $\mathbb{P}(L_N/N \rightarrow \infty) = 1$, as $N \rightarrow \infty$. Thus the limiting branching process is not defined. In this case, L_N is the i.i.d. sum of N random variables D_1, D_2, \dots, D_N , with infinite mean. From extreme value theory, it is well known that then the bulk of the contribution to L_N comes from a *finite* number of nodes which have giant degrees (the so-called *giant nodes*). Since these giant nodes have degree roughly $N^{1/(\tau-1)}$, which is much larger than \sqrt{N} , they are all connected to each other, thus forming a complete graph of giant nodes. Each stub of node 1 or node 2 is with probability close to 1 attached to a stub of some giant node, and therefore, the distance between any two nodes is, **whp**, at most 3. In fact, this distance equals 2 precisely when the two nodes are attached to the *same* giant node, and is 3 otherwise.

3 Connectivity properties of the random graph

Let \mathbf{C}_N denote the number of nodes in the complement to the largest connected component of the random graph. In the sections below, we frequently use the notation \mathbb{P}_N to denote the law of the random graph when we condition on D_1, \dots, D_N .

Proposition 3.1 *Let $r \in \{1, 2\}$, and assume that $\mathbb{P}(D_1 \geq r) = 1$. Then, for any $1 \leq s < N/2$,*

$$\mathbb{P}_N(\mathbf{C}_N \geq s) \leq 2 \sum_{j=s}^{N-s} \left(\frac{2N^{\frac{2}{r}}}{L_N} \right)^{\lceil jr/2 \rceil}. \quad (3.1)$$

We first show that Theorem 1.8 is an immediate consequence of Proposition 3.1:

Proof of Theorem 1.8. Consider first the case $2 < \mu < \infty$, and denote $\mu_N = L_N/N$. Then we apply Proposition 3.1 with $r = 2$ to obtain that the expectation of r.h.s. of (3.1) is at most

$$\begin{aligned} & \mathbb{E} \left(2 \sum_{j=s}^{N-s} \left(\frac{2}{\mu_N} \right)^j I[L_N \geq N(1 + \mu/2)] \right) + \mathbb{P}(\mu_N < (1 + \mu/2)) \\ & \leq \frac{2 \left(\frac{2}{1 + \mu/2} \right)^s}{1 - \frac{2}{1 + \mu/2}} + \mathbb{P}(\mu_N < (1 + \mu/2)) \leq \frac{2(2+\mu)}{\mu-2} \left(\frac{4}{2+\mu} \right)^s + e^{-IN}, \end{aligned} \quad (3.2)$$

for large enough N , because of the Large Deviation Principle, and where I is the exponential rate of the event $\mu_N < (1 + \mu/2)$, which is strictly positive since $(1 + \mu/2) < \mu$. This shows (1.19) for $2 < \mu < \infty$.

Consider next the case $\mu = \infty$. Then, for any $\varepsilon > 0$ and large enough N ,

$$\mathbb{P} \left(\frac{2N}{L_N} \leq \varepsilon \right) \geq 1 - \varepsilon.$$

Hence, the probability that the random graph is disconnected, or $\mathbb{P}(\mathbf{C}_N \geq 1)$, is due to (3.1) at most

$$\mathbb{E} \left(2 \sum_{j=1}^{N-1} \left(\frac{2N}{L_N} \right)^j I[2N \leq \varepsilon L_N] \right) + \mathbb{P} \left(\frac{2N}{L_N} > \varepsilon \right) \leq 2 \sum_{j=1}^{N-1} \varepsilon^j + \varepsilon < \frac{2\varepsilon}{1 - \varepsilon} + \varepsilon. \quad (3.3)$$

Since $\varepsilon > 0$ can be chosen arbitrarily small the probability that the graph is disconnected tends to zero, as N tends to infinity.

We complete the proof with the case that $L_N/N^2 \rightarrow \infty$, in probability or equivalently $\tau \in (1, \frac{3}{2})$. In this case, we take $r = s = 1$, and we use the assumption that $N^2/L_N \rightarrow 0$ in probability as $N \rightarrow \infty$, to see that

$$\mathbb{P}(\mathbf{C}_N \geq 1) \leq \mathbb{P}(N^2/L_N > \varepsilon) + 2 \sum_{j=1}^{N-1} (2\varepsilon)^{\lceil j/2 \rceil} \leq 9\varepsilon, \quad (3.4)$$

when $\varepsilon > 0$ is sufficiently small. □

Proof of Proposition 3.1. For any $s < N/2$ we estimate the probability that $\mathbf{C}_N \geq s$. If $\mathbf{C}_N \geq s$, then there exist two disjoint groups of nodes, one with $s \leq j \leq N - s$ nodes and another with $N - j$, such that all stubs of the first group pair within the first group and all stubs of the second group pair within the second. Indeed, when the largest component has size at least s , then this is true with j the size of the largest component, and when the largest component has size at most s , then we can subdivide the components into two groups, where each group has size at least s .

Let i_1, \dots, i_j be the nodes and $A_j = \sum_{i=1}^j D_{i_j}$ be the total number of stubs in the first group, and j_1, \dots, j_{N-j} and B_{N-j} the nodes and number of stubs in the second group. We remark that $A_j + B_{N-j} = L_N$, and since the groups are not connected, both A_j and B_{N-j} are even. The \mathbb{P}_N -probability that the groups are not connected is then, for each choice i_1, \dots, i_j ,

$$\prod_{n=0}^{\frac{A_j}{2}-1} \frac{A_j - 2n - 1}{L_N - 2n - 1} = \frac{\prod_{n=0}^{\frac{A_j}{2}-1} (A_j - 2n - 1)}{\prod_{n=0}^{\frac{A_j}{2}-1} (L_N - 2n - 1)} = \prod_{m=0}^{\frac{A_j}{2}-1} \frac{2m + 1}{L_N - 2m - 1}. \quad (3.5)$$

By symmetry between A_j and B_{N-j} , we also obtain that this probability is equal to

$$\prod_{m=0}^{\frac{B_{N-j}-1}{2}} \frac{2m+1}{L_N-2m-1}. \quad (3.6)$$

Observe that for integers $j \geq 0$, the map

$$j \mapsto \prod_{m=0}^j \frac{2m+1}{L_N-2m-1} \text{ is decreasing for } j \leq \frac{L_N}{4} - \frac{1}{2}. \quad (3.7)$$

Suppose that $A_j \leq L_N/2 - 1$. Then we use (3.5). Due to $\mathbb{P}(D_1 \geq r) = 1$, and since A_j is even we have $\lceil jr/2 \rceil \leq A_j/2 \leq (L_N/2 - 1)/2$ a.s., and hence, by (3.7), the final expression in (3.5) is at most

$$\prod_{m=0}^{\lceil jr/2 \rceil - 1} \frac{2m+1}{L_N-2m-1}.$$

Suppose that $A_j > L_N/2 + 1$. Since $A_j + B_{N-j} = L_N$ and $\mathbb{P}(D_1 \geq r) = 1$, we have then that $\lceil (N-j)r/2 \rceil \leq B_{N-j}/2 \leq (L_N/2 - 1)/2$ a.s., and we estimate (3.6), by (3.7), by

$$\prod_{m=0}^{\lceil (N-j)r/2 \rceil - 1} \frac{2m+1}{L_N-2m-1}.$$

Hence, the \mathbb{P}_N -probability that the two groups of nodes are not connected is at most

$$\begin{aligned} & \prod_{m=0}^{\lceil jr/2 \rceil - 1} \frac{2m+1}{L_N-2m-1} I[A_j \leq L_N/2 - 1] + \prod_{m=0}^{\lceil (N-j)r/2 \rceil - 1} \frac{2m+1}{L_N-2m-1} I[A_j > L_N/2 - 1] \\ & \leq \prod_{m=0}^{\lceil jr/2 \rceil - 1} \frac{2m+1}{L_N-2m-1} + \prod_{m=0}^{\lceil (N-j)r/2 \rceil - 1} \frac{2m+1}{L_N-2m-1}. \end{aligned} \quad (3.8)$$

For $1 \leq j \leq N-1$, we have at most $\binom{N}{j}$ ways to choose j nodes i_1, \dots, i_j . Hence, the \mathbb{P}_N -probability that $\mathbf{C}_N \geq s$ is at most

$$\begin{aligned} & \sum_{j=s}^{N-s} \frac{N!}{j!(N-j)!} \left(\prod_{m=0}^{\lceil jr/2 \rceil - 1} \frac{2m+1}{L_N-2m-1} + \prod_{m=0}^{\lceil (N-j)r/2 \rceil - 1} \frac{2m+1}{L_N-2m-1} \right) \\ & \leq 2 \sum_{j=s}^{N-s} \left(\prod_{m=\lceil jr/2 \rceil}^{j-1} \frac{N-m}{m+1} \right) \left(\prod_{m=0}^{\lceil jr/2 \rceil - 1} \frac{(N-m)(2m+1)}{(m+1)(L_N-2m-1)} \right). \end{aligned} \quad (3.9)$$

We will make use of the following lemma:

Lemma 3.2 For any $1 \leq k \leq N-1$,

$$\prod_{m=0}^{k-1} \frac{(N-m)(2m+1)}{(m+1)(2N-2m-1)} \leq 1. \quad (3.10)$$

Proof. Let, for $0 \leq m \leq N-1$,

$$h(m) = \frac{(N-m)(2m+1)}{(m+1)(2N-2m-1)}.$$

Then

$$\begin{aligned} h(m) &\leq 1, & \text{if } m \leq (N-1)/2, \\ h(m)h(N-m-1) &= 1, & \text{for all } 0 \leq m < N, \\ h((N-1)/2) &= 1, & \text{if } N \text{ is odd.} \end{aligned}$$

Hence, (3.10) is trivial for $k \leq (N-1)/2 + 1$. If $k > (N-1)/2 + 1$ then $N-k < (N-1)/2$ and

$$\left(\prod_{m=0}^{k-1} h(m) \right)^2 \leq \left(\prod_{m=N-k}^{k-1} h(m) \right)^2 = \prod_{m=N-k}^{k-1} h(m)h(N-m-1) = 1.$$

Thus we have (3.10) for all $1 \leq k \leq N-1$. \square

We first complete the proof of Proposition 3.1 when $r = 2$. In this case, due to (3.10), the l.h.s. of (3.9) is at most

$$2 \sum_{j=s}^{N-s} \prod_{m=0}^{j-1} \frac{(N-m)(2m+1)}{(m+1)(L_N-2m-1)} \leq 2 \sum_{j=s}^{N-s} \prod_{m=0}^{j-1} \frac{2N-2m-1}{L_N-2m-1} \leq 2 \sum_{j=s}^{N-s} \left(\frac{2N}{L_N} \right)^j, \quad (3.11)$$

since $L_N \geq 2N$ a.s. This completes the proof when $r \geq 2$.

When $r = 1$, we rewrite r.h.s. of (3.9) with $s = 1$ as

$$2 \sum_{j=1}^{N-1} \left(\prod_{m=\lceil j/2 \rceil}^{j-1} \frac{N-m}{m+1} \right) \left(\prod_{m=0}^{\lceil j/2 \rceil - 1} \frac{(N-m)(2m+1)}{(m+1)(L_N-2m-1)} \right), \quad (3.12)$$

which, due to (3.10), is at most

$$2 \sum_{j=1}^{N-1} \left(\prod_{m=\lceil j/2 \rceil}^{j-1} \frac{N-m}{m+1} \right) \left(\prod_{m=0}^{\lceil j/2 \rceil - 1} \frac{2N-2m-1}{L_N-2m-1} \right) \leq 2 \sum_{j=1}^{N-1} N^{\lfloor j/2 \rfloor} \left(\frac{2N}{L_N} \right)^{\lceil j/2 \rceil} \leq 2 \sum_{j=1}^{N-1} \left(\frac{2N^2}{L_N} \right)^{\lceil j/2 \rceil}.$$

This completes the proof of Proposition 3.1. \square

4 On the connected component sizes

In this section, we investigate the largest connected components in more detail and prove Theorem 1.6. Before stating the result, we give some definitions. For $\delta, \varepsilon > 0$, we define $\gamma_N = \gamma_N(\delta, \varepsilon)$ by

$$\gamma_N = \frac{1 + \delta}{\log \mu_N - \log 2 - \frac{\varepsilon}{1-2\varepsilon}} \log N. \quad (4.1)$$

Note that γ_N is a random variable. We also define a deterministic version of γ_N in the following way:

$$\bar{\gamma}_N = \frac{1 + \delta}{\log \underline{\mu}_N - \log 2 - \frac{\varepsilon}{1-2\varepsilon}} \log N, \quad (4.2)$$

where $\underline{\mu}_N$ is any deterministic sequence for which

$$\mathbb{P}(\mu_N \geq \underline{\mu}_N) = 1 - o(1). \quad (4.3)$$

We start by formulating a version of Theorem 1.6 that is valid under the \mathbb{P}_N -probability. Before stating this theorem, we need a number of assumptions. Define

$$q_N = \frac{1}{N} \sum_{i=1}^N \mathbb{P}_N(|\mathcal{C}_i| \geq \gamma_N), \quad (4.4)$$

where \mathcal{C}_i is the connected component that contains i and $|\mathcal{C}|$ denotes the number of nodes in $\mathcal{C} \subseteq \{1, \dots, N\}$. We assume that

$$\frac{1}{N(N-1)} \sum_{i \neq j} \mathbb{P}_N(i, j \text{ connected}) = q_N^2(1 + o(1)). \quad (4.5)$$

Note that (4.5) is an assumption involving the \mathbb{P}_N -probability. We can interpret (4.5) as saying that, under \mathbb{P}_N , a large proportion of nodes i, j for which the connected component consists of more than γ_N nodes, are connected.

Theorem 4.1 *Assume that (4.5) holds and that $q_N \geq \varepsilon$ as $N \rightarrow \infty$, for some $\varepsilon > 0$. Then, **whp**, the largest connected component in G has $q_N N(1 + o(1))$ nodes, and all other connected components have at most γ_N nodes. Moreover, **whp**, the largest connected component has in between $q_N N(1 \pm \omega_N \sqrt{\frac{\gamma_N}{L_N}})$ nodes for any $\omega_N \rightarrow \infty$.*

The remainder of this section is organized as follows. In Section 4.1, we prove Theorem 4.1. In Section 4.2, we use a modification of the proof of Theorem 4.1 to prove Theorem 1.6.

4.1 Connected components under \mathbb{P}_N

We start with a proposition that shows that the connected components, measured in terms of their number of edges, are either quite small, i.e., less than γ_N , or very large, i.e., a positive fraction of the total number of edges.

Proposition 4.2 *Fix $\delta > 0$, assume that $\mu_N > 2$, and let $0 < \varepsilon < \frac{1}{10}$ be such that $\log \mu_N - \log 2 - \frac{\varepsilon}{1-2\varepsilon} > 0$. Then, the \mathbb{P}_N -probability that there exists a connected component with in between γ_N and εL_N edges is bounded by $N^{-\delta}$.*

Consequently, when $\mu > 2$ or $\mu = \infty$, the \mathbb{P} -probability that there exists a connected component with in between $\bar{\gamma}_N$ and εL_N edges converges to 0, as $N \rightarrow \infty$.

Of course, we expect that there is a unique such large connected component, and this is what we will prove later on. Note that when $\tau \in (1, 2)$, then, with large probability, $\mu_N = L_N/N \geq N^\eta$, for some $\eta > 0$. In this case, we even have that γ_N is uniformly bounded in N , and thus, the connected components that do not contain a positive fraction of the edges are uniformly bounded.

Proof of Proposition 4.2. We start with the first claim. We adapt the proof of Proposition 3.1. Denote by k the number of edges in the connected component with in between γ_N and εL_N edges. Then, we must have that all the $2k$ stubs are connected to each other, i.e., they are not connected to stubs not in the k edges. Thus, this probability is bounded by

$$\prod_{n=0}^{k-1} \frac{2k - 2n - 1}{L_N - 2n - 1} = \prod_{m=0}^{k-1} \frac{2m + 1}{L_N - 2m - 1},$$

since we ignore the fact that the component needs to be connected. We first prove the statement for $k \leq (\frac{N}{2} - 1) \wedge \varepsilon L_N$, and in a second step prove the statement for $(\frac{N}{2} - 1) \wedge \varepsilon L_N < k \leq \varepsilon L_N$. We abbreviate $A_N = (\frac{N}{2} - 1) \wedge \varepsilon L_N$.

Denote the number of nodes in the connected component by l . Note that when a component consisting of k edges is connected, then $l \leq k + 1$. Therefore, the total number of ways in which we can choose these l nodes is at most $\binom{N}{l} \leq \binom{N}{k+1}$, when $k \leq \frac{N}{2} - 1$. Thus, the \mathbb{P}_N -probability

that there exists a component with in between γ_N and A_N edges is bounded by

$$\begin{aligned} \sum_{k=\gamma_N}^{A_N} \binom{N}{k+1} \prod_{m=0}^{k-1} \frac{2m+1}{L_N-2m-1} &\leq \sum_{k=\gamma_N}^{A_N} \frac{N!}{(N-k-1)!} 2^k \prod_{m=0}^{k-1} (L_N-2m-1)^{-1} \\ &\leq N \sum_{k=\gamma_N}^{A_N} \left(\frac{2}{\mu_N}\right)^k \prod_{m=0}^{k-1} \left(1 + \frac{2m+1}{L_N-2m-1}\right). \end{aligned} \quad (4.6)$$

Next, we use that $1+x \leq e^x$ for $x \geq 0$, to obtain that the probability that there exists a component with in between γ_N and A_N edges is bounded by

$$\begin{aligned} N \sum_{k=\gamma_N}^{A_N} \left(\frac{2}{\mu_N}\right)^k e^{\sum_{m=0}^{k-1} \frac{2m+1}{L_N-2m-1}} &\leq N \sum_{k=\gamma_N}^{A_N} \left(\frac{2}{\mu_N}\right)^k e^{\frac{k^2}{L_N-2\varepsilon L_N}} \\ &\leq N \sum_{k=\gamma_N}^{A_N} \left(\frac{2}{\mu_N}\right)^k e^{k \frac{\varepsilon}{1-2\varepsilon}} \leq \eta^{-1} N \left(\frac{2}{\mu_N} e^{\frac{\varepsilon}{1-2\varepsilon}}\right)^{\gamma_N}, \end{aligned} \quad (4.7)$$

provided that $\frac{2}{\mu_N} e^{\frac{\varepsilon}{1-2\varepsilon}} < 1 - \eta$. The right-hand side of (4.7) is bounded by $\eta^{-1} N^{-\delta}$ for the choice of γ_N in (4.1).

We complete the proof by dealing with the range of k 's such that $A_N < k \leq \varepsilon L_N$. In this case, we must have that $A_N = \frac{N}{2} - 1$, otherwise there is nothing to prove, so that $k \geq \frac{N}{2}$. Then, we bound the total number of ways in which we can choose the $l \leq k$ nodes by 2^N . Since $2^N \leq 2^{2k}$, for all $k \geq N/2$, we arrive at the fact that the probability that there exists a connected component with in between $\frac{N}{2}$ and εL_N edges is bounded by

$$\sum_{k=\frac{N}{2}}^{\varepsilon L_N} 2^N \prod_{m=0}^{k-1} \frac{2m+1}{L_N-2m-1} \leq \sum_{k=\frac{N}{2}}^{\varepsilon L_N} 2^{2k} \left(\frac{2\varepsilon}{1-2\varepsilon}\right)^k \leq \frac{1-2\varepsilon}{1-10\varepsilon} \left(\frac{8\varepsilon}{1-2\varepsilon}\right)^{N/2}, \quad (4.8)$$

which is exponentially small in N when $\varepsilon < \frac{1}{10}$. Thus, this probability is certainly bounded above by $N^{-\delta}$.

For the bound on the \mathbb{P} -probability that there exists a connected component with in between $\bar{\gamma}_N$ and εL_N edges, we denote by $F(k, l)$ the event that there exists a connected component with in between k and l edges. Then we can bound

$$\mathbb{P}(F(\bar{\gamma}_N, \varepsilon L_N)) \leq \mathbb{P}(\mu_N \leq \underline{\mu}_N) + \mathbb{E}[I[\mu_N \geq \underline{\mu}_N] \mathbb{P}_N(F(\gamma_N, \varepsilon L_N))], \quad (4.9)$$

where we use that $\bar{\gamma}_N \geq \gamma_N$ when $\mu_N \geq \underline{\mu}_N$. The first term is $o(1)$ by (4.3), the second term is small by the estimate $\mathbb{P}_N(F(\gamma_N, \varepsilon L_N)) \leq N^{-\delta}$ proved above. \square

We next present a lemma that will be used in the proof of Theorem 4.1.

Lemma 4.3 *Fix $\varepsilon > 0$ and $0 < \eta < \varepsilon$ sufficiently small. Then, when $L_N \geq 2N$, the \mathbb{P}_N -probability that there exists a connected component with at most ηN nodes, and in between εL_N and $(1-\varepsilon)L_N$ stubs, is exponentially small in N . Consequently, the same estimate is true for the \mathbb{P} -probability of this event when $\mathbb{P}(L_N \leq 2N)$ is exponentially small.*

Proof. Fix $\eta > 0$. Again denote by k the number of edges in an arbitrary connected component satisfying $\frac{\varepsilon}{2} L_N \leq k \leq \frac{(1-\varepsilon)}{2} L_N$. Then, we must have that all the $2k$ stubs are connected to each other, i.e., they are not connected to stubs not in the k edges. Thus, the \mathbb{P}_N -probability of this event is bounded by

$$\prod_{n=0}^{k-1} \frac{2k-2n-1}{L_N-2n-1} \leq \prod_{n=0}^{k-1} \frac{2k-2n}{L_N-2n} = \prod_{n=0}^{k-1} \frac{k-n}{\frac{L_N}{2}-n} = \binom{\frac{L_N}{2}}{k}^{-1}.$$

We next use that the number ways of choosing at most ηN nodes is bounded from above by

$$\sum_{j=0}^{\eta N} \binom{N}{j} \leq N \binom{N}{\eta N}. \quad (4.10)$$

Therefore, the \mathbb{P}_N -probability that there exists a connected component with in between εL_N and $(1 - \varepsilon)L_N$ stubs and at most ηN nodes is bounded by

$$\begin{aligned} & N \binom{N}{\eta N} \sum_{k=\frac{\varepsilon}{2}L_N}^{\frac{(1-\varepsilon)}{2}L_N} \mathbb{P}_N(2k \text{ stubs are connected to each other}) \\ & \leq N \sum_{k=\frac{\varepsilon}{2}L_N}^{\frac{(1-\varepsilon)}{2}L_N} \frac{\binom{N}{\eta N}}{\binom{\frac{L_N}{2}}{k}} \leq N L_N \frac{\binom{N}{\eta N}}{\binom{\frac{L_N}{2}}{\frac{\varepsilon}{2}L_N}} = N L_N e^{c_\eta N(1+o(1))} e^{-c_\varepsilon \frac{L_N}{2}(1+o(1))}, \end{aligned} \quad (4.11)$$

where we use that for η small, $\binom{N}{\eta N} = e^{c_\eta N(1+o(1))}$, where $c_\eta \downarrow 0$ as $\eta \downarrow 0$. Therefore, using that $L_N \geq 2N$, it suffices to take $\eta > 0$ so small that $c_\eta N < (c_\varepsilon - \delta) \frac{L_N}{2}$, for some $\delta > 0$ sufficiently small, to see that this probability is exponentially small in N . This completes the proof of Lemma 4.3. \square

We are now ready for the proof of Theorem 4.1:

Proof of Theorem 4.1. Fix $\delta > 0$ and γ_N as in (4.1). Recall that \mathcal{C}_i denotes the connected component that i belongs to. We define the random variable X_N by

$$X_N = \sum_{i=1}^N I[|\mathcal{C}_i| \geq \gamma_N], \quad (4.12)$$

so that X_N equals the total number of nodes in connected components of size at least γ_N . By (4.4), conditionally on D_1, \dots, D_N ,

$$\mathbb{E}_N[X_N] = \sum_{i=1}^N \mathbb{P}_N(|\mathcal{C}_i| \geq \gamma_N) = N q_N, \quad (4.13)$$

where \mathbb{E}_N is the expected value w.r.t. \mathbb{P}_N . We first prove that the variance of X_N under the law \mathbb{P}_N is small, so that X_N is with high probability close to $q_N N$:

Lemma 4.4

$$\text{Var}_N(X_N) = q_N(1 - q_N)N + O\left(\frac{\gamma_N N^2}{L_N}\right), \quad (4.14)$$

where Var_N denotes the variance under \mathbb{P}_N .

Proof. We first note that $\text{Var}_N(X_N) = \text{Var}(N - X_N) = \text{Var}_N(Y_N)$, where

$$Y_N = \sum_{i=1}^N I[|\mathcal{C}_i| < \gamma_N]. \quad (4.15)$$

Therefore,

$$\begin{aligned} \text{Var}_N(Y_N) &= \sum_{i,j} \mathbb{P}_N(|\mathcal{C}_i| < \gamma_N, |\mathcal{C}_j| < \gamma_N) - [N(1 - q_N)]^2 \\ &= \sum_{i \neq j} \mathbb{P}_N(|\mathcal{C}_i| < \gamma_N, |\mathcal{C}_j| < \gamma_N) + N(1 - q_N) - N^2(1 - q_N)^2. \end{aligned} \quad (4.16)$$

For the first term, we use the coupling in [28, Section 3] to obtain that

$$\mathbb{P}_N(|\mathcal{C}_i| < \gamma_N) = \mathbb{P}_N\left(\sum_l \hat{Z}_l^{(i,N)} < \gamma_N\right) + O\left(\frac{\gamma_N}{L_N}\right), \quad (4.17)$$

where $\hat{Z}_j^{(i,N)}$ is a delayed branching process with offspring distribution

$$g_n^{(N)} = \frac{n+1}{L_N} \sum_{j=1}^N I[D_j = n+1], \quad (4.18)$$

with $\hat{Z}_1^{(i,N)} = D_i$, the degree of node i . The coupling is described in full detail in [28, Section 3], whereas the bound in (4.17) follows from the proof of [28, Lemma A.2.2], which holds under the \mathbb{P}_N -probability and is therefore true for any degree sequence.

Moreover, for $i \neq j$, it is described in [28, Section 3] that we can couple $|\mathcal{C}_i|$ and $|\mathcal{C}_j|$ simultaneously to two *independent* branching processes to obtain

$$\mathbb{P}_N(|\mathcal{C}_i| < \gamma_N, |\mathcal{C}_j| < \gamma_N) = \mathbb{P}_N\left(\sum_l \hat{Z}_l^{(i,N)} < \gamma_N\right) \mathbb{P}_N\left(\sum_l \hat{Z}_l^{(j,N)} < \gamma_N\right) + O\left(\frac{\gamma_N}{L_N}\right). \quad (4.19)$$

Therefore,

$$\begin{aligned} \sum_{i \neq j} \mathbb{P}_N(|\mathcal{C}_i| < \gamma_N, |\mathcal{C}_j| < \gamma_N) &= \sum_{i \neq j} \left[\mathbb{P}_N\left(\sum_l \hat{Z}_l^{(i,N)} < \gamma_N\right) \mathbb{P}_N\left(\sum_l \hat{Z}_l^{(j,N)} < \gamma_N\right) + O\left(\frac{\gamma_N}{L_N}\right) \right] \\ &= \left(\sum_{i=1}^N \mathbb{P}_N\left(\sum_l \hat{Z}_l^{(i,N)} < \gamma_N\right) \right)^2 + O\left(\frac{\gamma_N N^2}{L_N}\right) \\ &= (1 - q_N)^2 N^2 + O\left(\frac{\gamma_N N^2}{L_N}\right), \end{aligned} \quad (4.20)$$

using (4.4) and (4.18). So, substituting (4.20) into (4.16),

$$\text{Var}_N(X_N) = \text{Var}_N(Y_N) = q_N(1 - q_N)N + O\left(\frac{\gamma_N N^2}{L_N}\right). \quad (4.21)$$

□

We continue with the proof of Theorem 4.1, which is a consequence of the following proposition. This proposition will also be used to prove Theorem 1.6 below. In its statement, we let \mathbb{Q} be a probability distribution, which we will take to be \mathbb{P}_N in the proof of Theorem 4.1 and \mathbb{P} in the proof of Theorem 1.6. Let $\gamma_N^* = \gamma_N$ when $\mathbb{Q} = \mathbb{P}_N$ and $\gamma_N^* = \bar{\gamma}_N$ when $\mathbb{Q} = \mathbb{P}$, for any $\delta > 0$. Furthermore, we take $X_N = \sum_{i=1}^N I[|\mathcal{C}_i| \geq \gamma_N^*]$.

Proposition 4.5 *Let $\mathbb{Q} = \mathbb{P}$ or $\mathbb{Q} = \mathbb{P}_N$. Suppose that $L_N \geq 2N$, $\text{Var}_{\mathbb{Q}}(X_N) \leq B_N$, and*

$$\mathbb{E}_{\mathbb{Q}}[X_N] = q_N N, \quad \sum_{i,j} \mathbb{Q}(i, j \text{ connected}) = q_N^2 N^2 (1 + o(1)), \quad (4.22)$$

where $q_N \geq \varepsilon$ for some $\varepsilon > 0$, then

- (i) **whp** the second largest component has at most γ_N^* nodes;
- (ii) **whp** the largest connected component has in between $q_N N \pm \omega_N \sqrt{B_N}$ nodes for any $\omega_N \rightarrow \infty$.

To prove Theorem 4.1, we use the above with $\mathbb{Q} = \mathbb{P}_N$ and $B_N = C \frac{\gamma_N N^2}{L_N}$.

Proof. We define the event

$$E_N = \{|X_N - q_N N| \leq \omega_N \sqrt{B_N}\}, \quad (4.23)$$

then, by the Chebycheff inequality,

$$\mathbb{Q}(E_N^c) \leq \left(\omega_N \sqrt{B_N}\right)^{-2} \text{Var}_{\mathbb{Q}}(X_N) \leq O(\omega_N^{-2}) = o(1). \quad (4.24)$$

We write $\mathcal{C}_{(1)}, \mathcal{C}_{(2)}, \dots$ for the connected components ordered according to their sizes, so that $\mathcal{C}_{(1)} \geq \mathcal{C}_{(2)} \geq \dots$ and $\mathcal{C}_{(i)}$ and $\mathcal{C}_{(j)}$ are disjoint for $i \neq j$. Then we clearly have that

$$\sum_{i,j} \mathbb{Q}(i, j \text{ connected}) = \sum_{i,j} \mathbb{Q}\left(\bigcup_l \{i, j \in \mathcal{C}_{(l)}\}\right) = \sum_l \sum_{i,j} \mathbb{Q}(i, j \in \mathcal{C}_{(l)}) = \sum_l \mathbb{E}_{\mathbb{Q}}[|\mathcal{C}_{(l)}|^2]. \quad (4.25)$$

By (4.22),

$$\sum_l \mathbb{E}_{\mathbb{Q}}[|\mathcal{C}_{(l)}|^2] = q_N^2 N^2 (1 + o(1)). \quad (4.26)$$

Furthermore,

$$\sum_l \mathbb{E}_{\mathbb{Q}}[|\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| < \gamma_N^*]] \leq \gamma_N^* \sum_l \mathbb{E}_{\mathbb{Q}}[|\mathcal{C}_{(l)}| I[|\mathcal{C}_{(l)}| < \gamma_N^*]] \leq \gamma_N^* N. \quad (4.27)$$

Therefore, since $\gamma_N^* = O(\log N)$ and $q_N \geq \varepsilon$, we obtain that

$$\sum_l \mathbb{E}_{\mathbb{Q}}[|\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*]] = q_N^2 N^2 (1 + o(1)). \quad (4.28)$$

By (4.24), we thus also have that

$$\mathbb{E}_{\mathbb{Q}}\left[\sum_l |\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] I[E_N]\right] = q_N^2 N^2 (1 + o(1)). \quad (4.29)$$

We will now prove that

$$\mathbb{Q}(|\mathcal{C}_{(2)}| \geq \gamma_N^*) = o(1). \quad (4.30)$$

This proceeds in two key steps. We first show that for some $\eta > 0$ sufficiently small

$$\mathbb{Q}(|\mathcal{C}_{(2)}| \geq \gamma_N^*) = \mathbb{Q}(|\mathcal{C}_{(2)}| > \eta N) + o(1), \quad (4.31)$$

and then that the assumption that

$$\limsup_{N \rightarrow \infty} \mathbb{Q}(|\mathcal{C}_{(2)}| > \eta N) = a > 0, \quad (4.32)$$

leads to a contradiction. Together we obtain (4.30). We start by proving (4.31). We note that we only need to prove that $\mathbb{Q}(|\mathcal{C}_{(2)}| \geq \gamma_N^*)$ is less than or equal to the right-hand side of (4.31), since the other bound is trivial (even with $o(1)$ replaced by 0).

To prove (4.31), we split

$$\mathbb{Q}(|\mathcal{C}_{(i)}| \geq \gamma_N^*) = \mathbb{Q}(|\mathcal{C}_{(i)}| \geq \gamma_N^*, |\mathcal{C}_{(i)}|_b \leq \varepsilon L_N) + \mathbb{Q}(|\mathcal{C}_{(i)}| \geq \gamma_N^*, |\mathcal{C}_{(i)}|_b > \varepsilon L_N), \quad (4.33)$$

where $|\mathcal{C}|_b$ denotes the number of edges in \mathcal{C} . Since $|\mathcal{C}|_b \geq |\mathcal{C}| - 1$, for any connected \mathcal{C} , by Proposition 4.2, for any $\delta > 0$, and for $i = 1, 2$, since $\gamma_N^* = \gamma_N$ or $\gamma_N^* = \bar{\gamma}_N$, where **whp** $\bar{\gamma}_N \geq \gamma_N$, we obtain

$$\mathbb{Q}(|\mathcal{C}_{(i)}| \geq \gamma_N^*, |\mathcal{C}_{(i)}|_b \leq \varepsilon L_N) \leq \mathbb{Q}(\gamma_N^* \leq |\mathcal{C}_{(i)}|_b \leq \varepsilon L_N) = o(1), \quad (4.34)$$

so that

$$\mathbb{Q}(|\mathcal{C}_{(2)}| \geq \gamma_N^*) = \mathbb{Q}(|\mathcal{C}_{(2)}| \geq \gamma_N^*, |\mathcal{C}_{(1)}|_b > \varepsilon L_N, |\mathcal{C}_{(2)}|_b > \varepsilon L_N) + o(1). \quad (4.35)$$

By Lemma 4.3, we further have that for $\eta > 0$ sufficiently small

$$\mathbb{Q}(|\mathcal{C}_{(2)}| \leq \eta N, |\mathcal{C}_{(1)}|_b > \varepsilon L_N, |\mathcal{C}_{(2)}|_b > \varepsilon L_N) = o(1), \quad (4.36)$$

Therefore,

$$\mathbb{Q}(|\mathcal{C}_{(2)}| \geq \gamma_N^*) = \mathbb{Q}(|\mathcal{C}_{(2)}| > \eta N, |\mathcal{C}_{(1)}|_b > \varepsilon L_N, |\mathcal{C}_{(2)}|_b > \varepsilon L_N) + o(1) \leq \mathbb{Q}(|\mathcal{C}_{(2)}| > \eta N) + o(1). \quad (4.37)$$

This proves (4.31).

We next prove that (4.32) is in contradiction with (4.29). Observe that on E_N ,

$$q_N N(1 + o(1)) = X_N = \sum_{i=1}^N I[|\mathcal{C}_i| \geq \gamma_N^*] = \sum_l |\mathcal{C}_{(l)}| I[|\mathcal{C}_{(l)}| \geq \gamma_N^*], \quad (4.38)$$

and that we can bound

$$\sum_l |\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] \leq |\mathcal{C}_{(2)}|^2 + \left(\sum_{l \neq 2} |\mathcal{C}_{(l)}| I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] \right)^2 = |\mathcal{C}_{(2)}|^2 + (X_N^* - |\mathcal{C}_{(2)}|)^2. \quad (4.39)$$

We split the involved expectation by intersecting with the event $\{|\mathcal{C}_{(2)}| > \eta N\}$ and its complement:

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} \left[\sum_l |\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] I[E_N] \right] &= \mathbb{E}_{\mathbb{Q}} \left[\sum_l |\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] I[E_N \cap \{|\mathcal{C}_{(2)}| > \eta N\}] \right] \\ &\quad + \mathbb{E}_{\mathbb{Q}} \left[\sum_l |\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] I[E_N \cap \{|\mathcal{C}_{(2)}| \leq \eta N\}] \right] \end{aligned} \quad (4.40)$$

We next use a simple argument from analysis. Note that the function $x \mapsto x^2 + (y-x)^2$ for $x \leq y/2$ and $x \geq \eta N$ is maximal for $x = \eta N$. We apply this inequality on the bound in (4.39), with $x = |\mathcal{C}_{(2)}|$ and $y = X_N \geq 2|\mathcal{C}_{(2)}| = 2x$, so that,

$$\mathbb{E}_{\mathbb{Q}} \left[\sum_l |\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] I[E_N \cap \{|\mathcal{C}_{(2)}| > \eta N\}] \right] \leq \mathbb{Q}(|\mathcal{C}_{(2)}| > \eta N) (\eta^2 + (q_N - \eta)^2) N^2 (1 + o(1)). \quad (4.41)$$

On the other hand,

$$\sum_l |\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] \leq \left(\sum_l |\mathcal{C}_{(l)}| I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] \right)^2 = q_N^2 N^2 (1 + o(1)), \quad (4.42)$$

implying that

$$\mathbb{E}_{\mathbb{Q}} \left[\sum_l |\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] I[E_N \cap \{|\mathcal{C}_{(2)}| \leq \eta N\}] \right] \leq \mathbb{Q}(\gamma_N^* \leq |\mathcal{C}_{(2)}| \leq \eta N) q_N^2 N^2 (1 + o(1)). \quad (4.43)$$

Together, (4.40), (4.41) and (4.43) yield

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} \left[\sum_l |\mathcal{C}_{(l)}|^2 I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] I[E_N] \right] & \leq \left[\mathbb{Q}(|\mathcal{C}_{(2)}| > \eta N) (\eta^2 + (q_N - \eta)^2) N^2 + \mathbb{Q}(\gamma_N^* \leq |\mathcal{C}_{(2)}| \leq \eta N) q_N^2 N^2 \right] (1 + o(1)), \end{aligned} \quad (4.44)$$

so that the assumption that $\limsup_{N \rightarrow \infty} \mathbb{Q}(|\mathcal{C}_{(2)}| > \eta N) \geq a > 0$ is in contradiction with (4.29), because assuming both (4.29) and (4.44) implies that $\eta \geq q$. This proves that the assumption in (4.32) is false, and we conclude that

$$\mathbb{Q}(|\mathcal{C}_{(2)}| \geq \gamma_N^*) = o(1), \quad (4.45)$$

which proves the claim for the second largest connected component.

We now prove that **whp** the largest component has size $q_N N \pm \omega_N \sqrt{B_N}$ for any $\omega_N \rightarrow \infty$. For this, we note that on the event that the second largest component has size less than or equal to γ_N^* , we have that

$$X_N = \sum_l |\mathcal{C}_{(l)}| I[|\mathcal{C}_{(l)}| \geq \gamma_N^*] = |\mathcal{C}_{(1)}|. \quad (4.46)$$

By (4.24) and (4.13), we thus obtain that

$$\mathbb{P}_N\left(|\mathcal{C}_{(1)}| - q_N N \geq \omega_N \sqrt{B_N}\right) \leq \mathbb{P}_N(|\mathcal{C}_{(2)}| \geq \gamma_N^*) + \mathbb{P}(E_N^c) = o(1). \quad (4.47)$$

This completes the proof of Proposition 4.5. \square

The proof of Theorem 4.1 follows from Proposition 4.5, once we notice that the assumptions in (4.22) are implied by (4.5). \square

4.2 Proof of Theorem 1.6

We use the results in [28] for $\tau > 3$, [29] for $\tau \in (2, 3)$ and [30] for $\tau \in (1, 2)$. We use Proposition 4.5 with $\mathbb{Q} = \mathbb{P}$ and $\gamma_N^* = \bar{\gamma}_N$ defined in (4.2). Because the nodes are exchangeable,

$$\mathbb{E}[X_N] = N\mathbb{P}(|\mathcal{C}_1| > \bar{\gamma}_N), \quad (4.48)$$

and this identifies $q_N = \mathbb{P}(|\mathcal{C}_1| > \bar{\gamma}_N)$. We next note that

$$\sum_{i,j} \mathbb{P}(i, j \text{ connected}) = N(N-1)\mathbb{P}(1, 2 \text{ connected}) + N, \quad (4.49)$$

and in [28], [29], [30] it is shown that

$$\mathbb{P}(1, 2 \text{ connected}) = q^2(1 + o(1)), \quad (4.50)$$

where $q = 1$ when $\tau \in (1, 2)$ and q the survival probability of the delayed branching process \mathcal{Z} when $\tau \in (2, 3)$ or $\tau > 3$. Therefore, in order to use Proposition 4.5, it remains to show that $q_N = q + o(1)$, and to give a bound on $\text{Var}(X_N)$, since $L_N \geq 2N$ follows from $\mu > 2$, when $\tau > 2$ or is immediate from $\tau \in (1, 2)$. This is done in Lemmas 4.6 and Lemma 4.7 below.

Lemma 4.6 $q_N = q + o(1)$.

Proof. We use the coupling described in [28, Section 3]. We have that

$$q_N = \mathbb{P}(|\mathcal{C}_1| > \bar{\gamma}_N) = 1 - \mathbb{P}(|\mathcal{C}_1| \leq \bar{\gamma}_N). \quad (4.51)$$

Using (4.17), we obtain

$$\mathbb{P}(|\mathcal{C}_1| \leq \bar{\gamma}_N) = \mathbb{E}\left[\mathbb{P}_N\left(\sum_l \hat{Z}_l^{(1,N)} \leq \bar{\gamma}_N\right)\right] + O\left(\frac{\bar{\gamma}_N}{L_N}\right). \quad (4.52)$$

Therefore,

$$q_N = 1 - \mathbb{E}\left[\mathbb{P}_N\left(\sum_l \hat{Z}_l^{(1,N)} \leq \bar{\gamma}_N\right)\right] + O\left(\frac{\bar{\gamma}_N}{L_N}\right). \quad (4.53)$$

We start with $\tau \in (1, 2)$. We note that

$$\mathbb{P}_N\left(\sum_l \hat{Z}_l^{(1,N)} \leq \bar{\gamma}_N\right) \leq \mathbb{P}_N(\hat{Z}_1^{(1,N)} \leq \bar{\gamma}_N) \leq \sum_{n=1}^{\bar{\gamma}_N} g_n^{(N)} = \sum_{i=1}^N \frac{D_i}{L_N} I[D_i \leq \bar{\gamma}_N + 1] \leq \frac{(\bar{\gamma}_N + 1)N}{L_N}. \quad (4.54)$$

Therefore, $q_N = 1 - O(\frac{\bar{\gamma}_N N}{L_N}) = 1 - o(1)$, **whp** when $\tau \in (1, 2)$.

We next turn to $\tau \in (2, 3)$ and $\tau > 3$, which we treat simultaneously. For this, we use that we can prove by coupling (see [28, Section 3]) that

$$\mathbb{P}_N(\sum_l \hat{Z}_l^{(1,N)} \leq \bar{\gamma}_N) = \mathbb{P}_N(\sum_l \mathcal{Z}_l^{(1,N)} \leq \bar{\gamma}_N) + O(\bar{\gamma}_N p_N) = \mathbb{P}(\sum_l \mathcal{Z}_l^{(1,N)} \leq \bar{\gamma}_N) + O(\bar{\gamma}_N p_N), \quad (4.55)$$

where p_N is the total variation distance between $\{g_n^{(N)}\}$ and $\{g_n\}$ given by

$$p_N = \frac{1}{2} \sum_{n=0}^{\infty} |g_n^{(N)} - g_n|. \quad (4.56)$$

In [28, Proposition 3.4], it is shown that for $\tau > 3$, and some $\alpha_2, \beta_2 > 0$,

$$\mathbb{P}(p_N > N^{-\alpha_2}) \leq N^{-\beta_2}. \quad (4.57)$$

In Remark A.1.3, the same conclusion is derived for $\tau \in (2, 3)$. Therefore,

$$q_N = 1 - \mathbb{P}(\sum_l \mathcal{Z}_l^{(1,N)} \leq \bar{\gamma}_N) + O(\frac{\bar{\gamma}_N}{L_N}) + O(N^{-\beta_2}) + O(\bar{\gamma}_N N^{-\alpha_2}), \quad (4.58)$$

so that, in turn,

$$q_N = q - \mathbb{P}(\bar{\gamma}_N < \sum_l \mathcal{Z}_l^{(1,N)} < \infty) + o(1). \quad (4.59)$$

We have that

$$\mathbb{P}(\bar{\gamma}_N < \sum_l \mathcal{Z}_l^{(1,N)} < \infty) = (1 - q) \mathbb{P}(\sum_l \mathcal{Z}_l^{(1,N)} > \bar{\gamma}_N | \text{extinction}). \quad (4.60)$$

A super-critical branching process conditioned on extinction is a branching process with law

$$g_n^* = (1 - q)^{n-1} g_n \quad (n \geq 1), \quad g_0^* = 1 - \sum_{n \geq 1} g_n^*. \quad (4.61)$$

Indeed, if F_n is the event that $\mathcal{Z}_l^{(1,N)}$ has n children in the first generation, then

$$\mathbb{P}(\mathcal{Z}_l^{(1,N)} \text{ dies out}, F_n) = g_n \mathbb{P}(n \text{ copies of } \mathcal{Z}_l^{(1,N)} \text{ die out}) = (1 - q)^n g_n. \quad (4.62)$$

It is not hard to see that g^* is subcritical, and it clearly has finite mean. Therefore, in particular, the total progeny has finite mean (in fact, even exponential tails), so that by the Markov inequality

$$\mathbb{P}(\sum_l \mathcal{Z}_l^{(1,N)} > \bar{\gamma}_N | \text{extinction}) \leq \bar{\gamma}_N^{-1} \mathbb{E}[\sum_l \mathcal{Z}_l^{(1,N)} | \text{extinction}] = C \bar{\gamma}_N^{-1} = o(1). \quad (4.63)$$

This completes the proof of Lemma 4.6. \square

We must also identify the bound B_N on the variance of X_N :

Lemma 4.7 *There exists $\alpha > 0$ such that*

$$\text{Var}(X_N) = O(N^{2-\alpha}). \quad (4.64)$$

Proof. We follow the proof of Lemma 4.4. We rewrite

$$\text{Var}(X_N) = \text{Var}(Y_N) = \mathbb{E}(\text{Var}_N(Y_N)) + \mathbb{E}(\mathbb{E}_N[Y_N]^2) - \mathbb{E}[Y_N]^2. \quad (4.65)$$

By Lemma 4.4, $\mathbb{E}(\text{Var}_N(Y_N))$ is certainly bounded by $O(N^{2-\alpha})$, and we are left to bound the second term. We start with $\tau \in (1, 2)$. We use (4.54) to see that **whp** $\mathbb{P}_N(\hat{Z}_1^{(j,N)} \leq \bar{\gamma}_N) \leq N^{-\alpha_2}$. Therefore,

$$0 \leq \mathbb{E}(\mathbb{E}_N[Y_N]^2) - \mathbb{E}[Y_N]^2 \leq \mathbb{E}(\mathbb{E}_N[Y_N]^2) \leq N^{2-2\alpha_2}. \quad (4.66)$$

Therefore, (4.64) holds with $\alpha = 2\alpha_2$. We next turn to $\tau \in (2, 3)$ and $\tau > 3$, which we treat simultaneously. For this, we use (4.55) and (4.57). Then, (4.64) holds with $\alpha = \alpha_2 \wedge \beta_2$. \square

This concludes the proof of Theorem 1.6. We even obtain an improvement, since $\sqrt{B_N} \leq N^{1-\alpha'}$ for any $\alpha' < \alpha$, so that **whp** the largest cluster is in between $q_N N \pm \omega_N N^{1-\alpha'}$. \square

5 Lower bounds on connected components and diameter

The proof of Theorem 1.7 is based on the following lemma:

Lemma 5.1 *Let $f_1 > 0$. Suppose further that for some $k = k_N \leq C \log N$, where $C > 0$, and some $0 < \delta < 1/6$, **whp***

$$N f_k \left(\frac{f_1(1-\delta)}{\mu_N} \right)^k \rightarrow \infty. \quad (5.1)$$

*Then **whp** the random graph contains a connected component with $k + 1$ nodes.*

Proof. Take k such that $f_k > 0$ and consider the star-like connected component, with one node of degree k at the center and k nodes of degree 1 at the ends (see Figure 2).

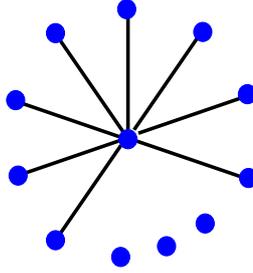


Figure 2: A star-like connected component with $k + 1$ nodes.

We will show that if the condition of the lemma holds, then the random graph contains the above connected component **whp**, uniformly in N .

The idea behind the proof is the following. There are **whp** at least $f_1(1-\delta)N$ nodes of degree one. Hence the probability that we connect a node of degree k to k nodes of degree 1 is at least

$$\left(\frac{f_1(1-\delta)N}{L_N} \right)^k = \left(\frac{f_1(1-\delta)}{\mu_N} \right)^k. \quad (5.2)$$

Since, **whp** there are about $f_k N$ nodes of degree k , we have, about $f_k N$ trials to make a k -star component. The mean number of successful trials is then about

$$N f_k \left(\frac{f_1(1-\delta)}{\mu_N} \right)^k \rightarrow \infty,$$

by condition (5.1). Hence, **whp** we expect to make at least one successful trial, and find a k -star component.

We now give the proof in detail. First of all we define a procedure which determines the existence of k -stars in the random graph. Consider the process of pairing stubs in the graph. We are free to choose the order in which we pair them. Let $D_{i_1}, \dots, D_{i_{\ell_N}}$, where we abbreviate $\ell_N = N(k)$, as the number of the nodes with degree k , which we call k -nodes for brevity. We pair the stubs in the following order.

Let $S_{(1)} = S_{1,1}, \dots, S_{1,k}$ be the stubs of node $i_1^* = i_1$. We first pair $S_{1,1}$. If it is paired with a stub of a node of degree 1, then we call this *pairing successful* and consider the pairing of $S_{1,2}$. If $S_{1,2}$ is paired with a stub of a node of degree 1, then we call the second pairing successful and consider the pairing of $S_{1,3}$, and so on until the first moment when one of the two following things happens. The first case is that all stubs of node i_1 are paired with nodes of degree 1. Then we observe a k -star component, we call the first *trial successful* and stop. The second case is that we come to $l < k$ such that the l^{th} pairing is unsuccessful, i.e., $S_{1,l}$ is not paired with a node of degree 1. Then we call the whole trial unsuccessful and stop with pairing of the stubs in $S_{(1)}$. In the later

case it is possible that $S_{1,l}$ is paired with another node in i_2, \dots, i_{ℓ_N} . Such node can not turn into a k -star anymore, we call this node, as well as node i_1 , *used* and discard them in the procedure.

We define our successive trials inductively. For any $m \geq 2$, let i_m^* be the first unused node in the sequence i_1, \dots, i_{ℓ_N} . Then, for i_m^* we use the same procedure as with i_1^* to determine whether the m^{th} trial is successful or not. If the trial is not successful, then node i_m^* becomes used, and if the corresponding unsuccessful pairing involves another unused k -node, then we also call this node used. We always discard all used nodes from the procedure.

We repeat these trials until we find a successful k -node or until all k -nodes are used. The \mathbb{P}_N -probability that the i^{th} trial is successful is

$$\prod_{j=0}^{k-1} \frac{N(1) - L_N(1, i) - j}{L_N - ki - 2j - 1} I[N(k) \geq i] I[N(1) - L_N(1, i) \geq k], \quad (5.3)$$

where $N(1) - L_N(1, i)$ is the remaining number of free stubs of the degree 1 nodes up to the moment of the i^{th} trial. Let $\tau_N(k)$ be the number of trials. Since at every unsuccessful trial the number of used nodes increases by at most two, we have $\tau_N(k) \geq \lfloor N(k)/2 \rfloor$. Instead of using all these trials, we will only use $\lfloor \delta^2 N(k) \rfloor$ of them. Then, after $\lfloor \delta^2 N(k) \rfloor$ trials, there are at least $N(1) - k \lfloor \delta^2 N(k) \rfloor$ remaining free stubs attached to nodes of degree 1. Hence, for $i \leq \lfloor \delta^2 N(k) \rfloor$, (5.3) is at least

$$\left(\frac{N(1) - \delta^2 k N(k)}{L_N} \right)^k I[N(k) \geq i] I[N(1) - \delta^2 k N(k) \geq 0], \quad (5.4)$$

where we use that if $N(1) - \delta^2 k N(k) \geq 0$, then, for all $i \leq \lfloor \delta^2 N(k) \rfloor - 1$, we have that $N(1) - L_N(1, i) \geq k$.

Then, the \mathbb{P}_N -probability that all $\lfloor \delta N(k) \rfloor$ trials are unsuccessful is at most

$$\prod_{i=0}^{\lfloor \delta^2 N(k) \rfloor - 1} \left(1 - \left(\frac{N(1) - \delta^2 k N(k)}{L_N} \right)^k \right) I[N(1) - \delta^2 k N(k) \geq 0]. \quad (5.5)$$

If we show that **whp**

$$\begin{aligned} (i) \quad & N(1) - \delta^2 k N(k) > f_1(1 - \delta)N, \\ (ii) \quad & N(k) \geq N f_k / 2, \end{aligned} \quad (5.6)$$

then the probability that none of the trials is successful is at most

$$(e^{-1} + o(1))^{\lfloor \delta^2 N(k) \rfloor} \left(\frac{N(1) - \delta^2 k N(k)}{L_N} \right)^k + o(1) \leq (e^{-1} + o(1))^{\frac{N \delta^2 f_k}{2} \left(\frac{f_1(1-\delta)}{\mu_N} \right)^k} + o(1) = o(1), \quad (5.7)$$

whp uniformly in N , due to (5.1). Hence, we are done if we prove (5.6).

For $\tau \in (1, 2)$, we have $f_k > 0$, for some $k \leq \gamma_2^{**}$. It follows that k is bounded and so by the law of large numbers we have **whp**,

$$(1 - \delta)N f_k \leq N(k) \leq (1 + \delta)N f_k. \quad (5.8)$$

Hence, part (5.6(ii)) is clear from the lower bound in (5.8) when $\delta \leq 1/2$. For part (5.6(i)), we use (5.8) together with the similar bound that **whp**

$$(1 - \delta^2)N f_1 \leq N(1) \leq (1 + \delta^2)N f_1. \quad (5.9)$$

Then

$$N(1) - \delta^2 k N(k) \geq (1 - \delta^2)N f_1 - \delta^2(1 + \delta)N k f_k = N[f_1 - \delta^2(f_1 + (1 + \delta)k f_k)] \geq N f_1(1 - \delta), \quad (5.10)$$

when δ is sufficiently small, since k is fixed. This completes the proof of (5.6) when $\tau \in (1, 2)$.

We turn to the case $\tau > 2$. In this case $k = k_N \leq C \log N$ and hence the law of large numbers does not apply. Instead, we use [33], which states that a binomial random variable X with mean μ , satisfies

$$\mathbb{P}(|X - \mu| \geq t) \leq 2e^{-\frac{t^2}{2(\mu+t/3)}}, \quad (5.11)$$

for all $t > 0$. We apply the above result with $X = N(k)$, $\mu = Nf_k$ and $t = \delta Nf_k/2$. Then we obtain for large enough N ,

$$\mathbb{P}_N(|N(k) - Nf_k| \geq \delta f_k N) \leq 2e^{-\frac{\delta^2 N f_k}{2(1+\delta/6)}} = o(1), \quad (5.12)$$

uniformly in N and $k = k_N \leq C \log N$. This yields (5.8) and hence (5.6(ii)), **whp**, when $\delta \leq 1/2$. Furthermore, for $\tau > 2$ we obtain because the expectation $\sum j f_j < \infty$ that $k f_k = o(1)$ when $k \rightarrow \infty$. Hence, **whp** for large enough N ,

$$N(1) - \delta^2 k N(k) \geq (1 - \delta^2) f_1 N - \delta^2 (1 + \delta) k f_k N/2 \geq (1 - \delta) f_1 N,$$

uniformly in N and in $k = k_N \leq C \log N$, for $\delta > 0$ small enough, so that (5.6(i)) holds. This completes the proof of (5.6) for $\tau > 2$. □

Proof of Theorem 1.7(i). We check the conditions of Lemma 5.1 for $\tau \in (1, 2)$. Firstly, since k bounded by some constant, the condition $k = k_N \leq C \log N$, for some $C > 0$ of Lemma 5.1 is trivially fulfilled. Secondly, we rewrite the expression in the l.h.s. of (5.1) as

$$f_k N \left(\frac{f_1(1 - \delta)}{\mu_N} \right)^k = f_k (f_1(1 - \delta))^k e^{\log N - k \log(\mu_N)} \quad (5.13)$$

Fix $0 < \delta' < \delta < 1/6$, and let ε be arbitrary small positive number. Since $L_N = D_1 + \dots + D_N$ where D_i is in the domain of attraction of a stable law ([27, Corollary 2, XVII.5, p. 578]), we have

$$\mathbb{P} \left(\log \mu_N \leq (1 + \delta') \frac{2 - \tau}{\tau - 1} \log N \right) = \mathbb{P} \left(\mu_N \leq N^{(1 + \delta') \frac{2 - \tau}{\tau - 1}} \right) = \mathbb{P} \left(L_N \leq N^{\frac{1}{\tau - 1} + \frac{\delta'(2 - \tau)}{\tau - 1}} \right) \geq 1 - \varepsilon, \quad (5.14)$$

since $(2 - \tau)/(\tau - 1) > 0$, for $\tau \in (1, 2)$. Thus, we obtain that, with probability at least $1 - \varepsilon$,

$$N f_k \left(\frac{f_1(1 - \delta)}{\mu_N} \right)^k = f_k (f_1(1 - \delta))^k e^{\log N - k \log \mu_N} \geq f_k (f_1(1 - \delta))^k N^{\delta - \delta' + \delta \delta'} \rightarrow \infty, \quad (5.15)$$

for every $k \leq \gamma_2^{**}$, defined in (1.17). Therefore, the conditions of Lemma 5.1 are fulfilled, and Theorem 1.7(i) follows. □

Proof of Theorem 1.7(ii). We again use Lemma 5.1 and check its conditions. Firstly, by the condition of the theorem, k clearly satisfies $k = k_N \leq C \log N$, for some $C > 0$. Secondly, we rewrite the expression in the l.h.s. of (5.1), using (1.18), and with δ replaced by δ' , as

$$N f_k \left(\frac{f_1(1 - \delta')}{\mu_N} \right)^k = L_f(k) e^{-\tau \log k} e^{\log N + k \log(f_1(1 - \delta')/\mu_N)}. \quad (5.16)$$

Since $\tau > 2$, we have by the Law of Large Numbers, $\mu_N \rightarrow \mu$ in probability, as $N \rightarrow \infty$, so that **whp** $\mu_N \leq \mu/(1 - \delta')$. On this event, we then estimate

$$\begin{aligned} \log N + k \log \left(\frac{f_1(1 - \delta')}{\mu_N} \right) &\geq \log N - \gamma_1^{**} \log N \cdot \log \left(\frac{\mu_N}{f_1(1 - \delta')^2} \right) \\ &\geq \log N \left(1 - \frac{1 - \delta}{\log(\mu/f_1)} [\log(\mu/f_1) - 2 \log(1 - \delta')] \right) \geq \frac{\delta}{2} \log N, \end{aligned} \quad (5.17)$$

when $\delta' > 0$ is sufficiently small. Substituting the above inequality in the r.h.s. of (5.16), we obtain from (1.18) that, for sufficiently large N , **whp**

$$N f_k \left(\frac{f_1(1 - \delta')}{\mu_N} \right)^k \geq L_f(k) e^{-\tau \log k} N^{\delta/2} \rightarrow \infty, \quad \text{as } N \rightarrow \infty, \quad (5.18)$$

where we have used that $k = k_N \leq C \log N$ to estimate $e^{-\tau \log k} \geq (\log N)^{-\tau C}$. We conclude that the condition (5.1) is fulfilled with some $\delta' > 0$, and we have proved Theorem 1.7(ii). \square

We close this section by proving the lower bound on the diameter in Theorem 1.9:

Proof of Theorem 1.9. The idea behind the proof is simple. Under the conditions of the theorem, **whp** one can find a path $\Gamma(N)$ in the random graph such that this path consists exclusively of nodes with degree 2 and has length at least $2\alpha \log N$. This implies that the diameter $D(N)$ is at least $\alpha \log N$, since the above path could be a cycle.

Below we define a procedure which proves the existence of such a path. Consider the process of pairing stubs in the graph. We are free to choose the order in which we pair the free stubs, since this order is irrelevant for the distribution of the random graph. Hence, we are allowed to start with pairing the stubs of the nodes with degree 2.

Let $S_N(2) = (i_1, \dots, i_{N(2)}) \in \mathbb{N}^{N(2)}$ be the nodes with degree 2, where we recall that $N(2)$ is the number of such nodes. We will pair the stubs and at the same time define a permutation $\Pi(N) = (i_1^*, \dots, i_{N(2)}^*)$ of $S_N(2)$, and a characteristic $\chi(N) = (\chi_1, \dots, \chi_{N(2)})$ on $\Pi(N)$, where χ_j is either 0 or 1. $\Pi(N)$ and $\chi(N)$ will be defined inductively in such a way that for any node $i_k^* \in \Pi(N)$, $\chi_k = 1$, if and only if node i_k^* is connected to node i_{k+1}^* . Hence, if $\chi(N)$ contains a substring of at least $2\alpha \log N$ ones then the random graph contains the required path $\Gamma(N)$.

We initialize our inductive definition by $i_1^* = i_1$. The node i_1^* has two stubs, we consider the second one and pair it to an arbitrary free stub. If this free stub belongs to another node $j \neq i_1^*$ in $N(2)$ then we choose $i_2^* = j$ and $\chi_1 = 1$, else we choose $i_2^* = i_2$, and $\chi_1 = 0$. Suppose for some $1 < k \leq N(2)$, the sequences (i_1^*, \dots, i_k^*) and $(\chi_1, \dots, \chi_{k-1})$ are defined. If $\chi_{k-1} = 1$, then one stub of i_k^* is paired to a stub of i_{k-1}^* , and another stub of i_k^* is free, else, if $\chi_{k-1} = 0$, node i_k^* has two free stubs. Thus, node i_k^* has at least one free stub. We pair this stub to an arbitrary remaining free stub. If this second stub belongs to node $j \in S_N(2) \setminus \{i_1^*, \dots, i_k^*\}$, then we choose $i_{k+1}^* = j$ and $\chi_k = 1$, else we choose i_{k+1}^* as the first stub in $S_N(2) \setminus \{i_1^*, \dots, i_k^*\}$, and $\chi_k = 0$. Hence, we have defined $\chi_k = 1$, if and only if node i_k^* is connected to node i_{k+1}^* .

We show that **whp** there exists a substring of ones of length at least $2\alpha \log N$ in the first half of χ_N , i.e., in $\chi_{\frac{1}{2}}(N) = (\chi_{i_1^*}, \dots, \chi_{i_{\lfloor N(2)/2 \rfloor}^*})$. For this purpose, we couple the sequence $\chi_{\frac{1}{2}}(N)$ with a sequence $B_{\frac{1}{2}}(N) = (B_k)$, where B_k are i.i.d. Bernoulli random variables taking value 1 with probability $f_2/(4\mu)$, and such that $\chi_{i_k^*} \geq B_k$, for all $k \in \{1, \dots, \lfloor N(2)/2 \rfloor\}$, **whp**. Indeed, for any $1 \leq k \leq \lfloor N(2)/2 \rfloor$, the \mathbb{P}_N -probability that $\chi_{i_k^*} = 1$ is at least

$$\frac{2N(2) - C_N(k)}{L_N - C_N(k)}, \quad (5.19)$$

where as before $N(2)$ is the total number of stubs of the nodes with degree 2, and $C_N(k)$ is the total number of paired stubs after $k + 1$ pairings. By definition of $C_N(k)$, for any $k \leq N(2)/2$, we have

$$C_N(k) = 2(k - 1) + 1 \leq N(2). \quad (5.20)$$

Due to the Law of Large Numbers we also have that **whp**

$$N(2) \geq f_2 N/2, \quad L_N \leq 2\mu N. \quad (5.21)$$

Substitution of (5.20) and (5.21) into (5.19) gives us that the r.h.s. of (5.19) is at least

$$\frac{N(2)}{L_N} \geq \frac{f_2}{4\mu}.$$

Thus, **whp** we can couple $\chi_{\frac{1}{2}}(N)$ with a Bernoulli sequence $B_{\frac{1}{2}}(N)$ of n independent trials with success probability p , where $p = f_2/(4\mu)$ and $n = Nf_2/2$. It is well known that the probability of existence of a run of $2\alpha \log N$ ones converges to one whenever $2\alpha \log N \leq (1 - \delta) \frac{\log n}{|\log p|}$ (see [25]).

We conclude that **whp** the sequence $B_{\frac{1}{2}}(N)$ contains a group (and hence a substring) of $2\alpha \log N$ ones. Since **whp** $\chi_N \geq B_{\frac{1}{2}}(N)$, where the ordering is componentwise, **whp** the sequence χ_N also contains a substring of $2\alpha \log N$ ones, and hence there exists a required path consisting of at least $2\alpha \log N$ nodes with degree 2. Thus, **whp** the diameter is at least $\alpha \log N$, and we have proved the theorem. \square

Acknowledgement

We thank Hongsuda Tangmunarunkit for providing us with the Internet data in Figure 1. The work of RvdH and DZ was supported in part by Netherlands Organisation for Scientific Research (NWO). This work was performed in part at the Mittag-Leffler Institute during a visit of all three authors.

References

- [1] W. Aiello, F. Chung and L. Lu. A random graph model for power law graphs. *Experiment. Math.* **10**, no. 1, 53–66, (2001).
- [2] W. Aiello, F. Chung and L. Lu. Random evolution of massive graphs. In Handbook of Massive Data Sets, J. Abello, P.M. Pardalos and M.G.C. Resende, eds., Kluwer Academic, Dordrecht, 97–122, 2002.
- [3] R. Albert and A.-L. Barabási. Emergence of scaling in random networks. *Science* **286**: 509-512, (1999).
- [4] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47-97, (2002).
- [5] N. Alon and J. Spencer. *The Probabilistic Method. Second Edition* John Wiley and Sons, New York (2000).
- [6] R.A. Arratia and T.M. Liggett. How likely is an i.i.d. degree sequence to be graphical? Preprint (2004).
- [7] A.-L. Barabasi. *Linked, The New Science of Networks*. Perseus Publishing, Cambridge, Massachusetts, (2002).
- [8] B. Bollobás. *Random Graphs*, 2nd edition. Academic Press, (2001).
- [9] B. Bollobás and O. Riordan. The diameter of scale-free random graphs. Preprint, (2003).
- [10] B. Bollobás and O. Riordan. Mathematical results on scale-free random graphs. *Handbook of graphs and networks*, 1–34, Wiley-VCH, Weinheim, (2003).
- [11] B. Bollobás and O. Riordan. Coupling scale-free and classical random graphs. Preprint, (2003).

- [12] B. Bollobás, C. Borgs, J.T. Chayes and O. Riordan. Directed scale-free graphs. Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (Baltimore, MD, 2003), 132–139, ACM, New York, (2003).
- [13] B. Bollobás, O. Riordan, J. Spencer and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, **18** no. 3, 279–290, (2001).
- [14] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, J. Wiener, Graph Structure in the Web. *Computer Networks*, **33**, 309-320, 2000.
- [15] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *PNAS*, **99**(25), 15879–15882, (2002).
- [16] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, **6**, 125–145, (2002).
- [17] F. Chung and L. Lu. The small world phenomenon in hybrid power law graphs. Preprint (2004).
- [18] R. Cohen and S. Havlin, Scale free networks are ultrasmall. *Physical Review Letters* 90, 058701 (2003).
- [19] C. Cooper and A.A. Frieze. A general model of web graphs. *Random Structures Algorithms*, **22** no. 3, 311–335, (2003).
- [20] A. Clauset and C. Moore. Traceroute sampling makes random graphs appear to have power law degree distributions. Preprint (2004).
- [21] P.L. Davies. The simple branching process: A note on convergence when the mean is infinite. *J. Appl. Prob.* **15**, 466-480 (1978).
- [22] M. Deijfen, A. Martin-Löf and T. Britton. Generating simple random graphs with arbitrary degree sequences. In preparation.
- [23] S.N. Dorogovtsev, A.V. Goltsev and J.F.F. Mendes. Pseudofractal scale-free web. *Phys. Rev. E* **65**, 066122, (2002).
- [24] S.N. Dorogovtsev, J.F.F. Mendes and A.N. Samuhkin. Metric structure of random networks. *Nucl. Phys. B.* 653, 307, (2003).
- [25] P. Erdős and A. Rényi. On a new law of large numbers. *J. Analyse Math.* **23**, 103–111 (1970).
- [26] C. Faloutsos, P. Faloutsos and M. Faloutsos. On power-law relationships of the internet topology, *Computer Communications Rev.*, **29**, 251-262, (1999).
- [27] W. Feller. *An Introduction to Probability Theory and Its Applications*, Volume II, 2nd edition, John Wiley and Sons, New York, (1971).
- [28] R. van der Hofstad, G. Hooghiemstra and P. Van Mieghem. Distances in random graphs with finite variance degrees. Preprint (2003). To appear in *Random Structures Algorithms*.
- [29] R. van der Hofstad, G. Hooghiemstra and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. Preprint (2004).
- [30] R. van der Hofstad, G. Hooghiemstra and D. Znamenski. Distances in random graphs with infinite mean degrees. Preprint (2004).

- [31] G. Hooghiemstra and P. Van Mieghem. On the mean distance in scale-free graphs. Preprint (2003).
- [32] S. Janson, T. Luczak and A. Rucinski. *Random Graphs*, John Wiley & Sons, New York, (2000).
- [33] S. Janson. On concentration of probability. Contemporary Combinatorics, ed. B. Bollobás, Bolyai Soc. Math. Stud. **10**, János Bolyai Mathematical Society, Budapest, 289-301, (2002).
- [34] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. Trawling the Web for emerging cyber communities. *Computer Networks*, **31**,(11-16), 1481–1493, (1999).
- [35] R. Kumar, P. Raghavan, S. Rajagopalan D. Sivakumar, A. Tomkins and E. Upfal. Stochastic Models for the Web Graph. *42st Annual IEEE Symposium on Foundations of Computer Science*, 57–65, (2000).
- [36] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, **6**, 161-179, (1995).
- [37] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.*, **7**, 295-305, (1998).
- [38] M.E.J. Newman. The structure and function of complex networks. *SIAM Rev.* **45**, no. 2, 167–256, (2003).
- [39] M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Random graphs with arbitrary degree distribution and their application. *Phys. Rev. E* **64**, 1-17, 026118, (2001).
- [40] I. Norros and H. Reittu. On a conditionally Poissonian graph process. Preprint (2004).
- [41] H. Reittu and I. Norros. On the power law random graph model of massive data networks. *Performance Evaluation*, **55** (1-2), 3-23, (2004).
- [42] S. H. Strogatz. Exploring complex networks. *Nature*, 410(8):268–276, March (2001).
- [43] H Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Network topology generators: Degree-based vs. structural. *ACM Sigcomm'02, Pittsburgh, Pennsylvania, Aug. 19-23, USA*, (2002).
- [44] D. J. Watts. *Small Worlds, The Dynamics of Networks between Order and Randomness*. Princeton University Press, Princeton, New Jersey, (1999).