Distances in random graphs with infinite mean degrees

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

June 2, 2004

Abstract

We study random graphs with an i.i.d. degree sequence of which the tail of the distribution function F is regularly varying with exponent $\tau \in (1, 2)$. Thus, the degrees have infinite mean. Such random graphs can serve as models for complex networks where degree power laws are observed.

The minimal number of edges between two arbitrary nodes, also called the graph distance or the hopcount, in a graph with N nodes is investigated when $N \to \infty$. The paper is part of a sequel of three papers. The other two papers study the case where $\tau \in (2, 3)$, and $\tau \in (3, \infty)$, respectively.

The main result of this paper is that the graph distance converges for $\tau \in (1, 2)$ to a limit random variable with probability mass exclusively on the points 2 and 3. We also consider the case where we condition the degrees to be at most N^α for some $\alpha > 0$. For $\tau^{-1} < \alpha < (\tau-1)^{-1}$, the hopcount converges to 3 in probability, while for $\alpha > (\tau-1)^{-1}$, the hopcount converges to the same limit as for the unconditioned degrees. Our results give convincing asymptotics for the hopcount when the mean degree is infinite, using extreme value theory.

Key words and phrases: Extreme value theorem, Internet, random graphs.
AMS 2000 *Subject classifications: Primary 60G70, Secondary 05C80.*

1 Introduction

The study of complex networks has attracted considerable attention in the past decade. There are numerous examples of complex networks, such as social relations, the World-Wide Web and Internet, co-authorship and citation networks of scientists, etc. The topological structure of networks affects the performance in those networks. For instance, the topology of social networks affects the spread of information and disease (see e.g., [19]), while the performance of traffic in Internet depends heavily on the topology of the Internet.

Measurements on complex networks have shown that many real networks have similar properties. A first example of such a fundamental network property is the fact that typical distances between nodes are small. This is called the 'small world' phenomenon, see the pioneering work of Watts [20] and the references therein. In Internet, for example, e-mail messages cannot use more than a threshold of physical links, and if the distances in Internet 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.

*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`

A second, maybe more surprising, property of many networks is that the number of nodes with degree k falls of as an inverse power of k. This is called a 'power law degree sequence'. In Internet, the power law degree sequence was first observed in [9]. The observation that many real networks have the above 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. See [3, 16] and the references therein for an introduction to complex networks and many examples where the above two properties hold.

The current paper presents a rigorous derivation for the random fluctuations of the distance between two arbitrary nodes (also called the geodesic) in a graph with i.i.d. degrees with *infinite* mean. The model with i.i.d. degrees is a variation of the *configuration model*, which was originally proposed by Newman, Strogatz and Watts [17], where the degrees originate from a given deterministic sequence. The observed power exponents are in the range from $\tau = 1.5$ to $\tau = 3.2$ (see [3, Table II] or [16, Table II]). In a previous paper with Van Mieghem [11], we have investigated the case $\tau > 3$. The case $\tau \in (2, 3)$ was studied in [12]. Here we focus on the case $\tau \in (1, 2)$, and study the typical distances between arbitrary connected nodes. In a forthcoming paper [13], we will survey the results from the different cases for τ, and investigate the connected components of the random graphs.

This section is organized as follows. In Section 1.1 we start by introducing the model, and in Section 1.2 we state our main results. In Section 1.3 we explain heuristically the results obtained. Finally we describe related work in Section 1.4.

1.1 The model

Consider an i.i.d. sequence $D_1, D_2, \ldots, D_N$. Assume that $L_N = \sum_{j=1}^{N} D_j$ is even. When L_N is odd, then we increase the number of stubs of D_N by 1, i.e., we replace D_N by $D_N + 1$. This change will make hardly any difference in what follows, and we will ignore it in the sequel.

We will construct a graph in which node j has degree D_j for all $1 \le j \le N$. We will later specify the distribution of D_j. We start with N separate nodes and incident to node j, we have D_j stubs which still need to be connected to build the graph.

The stubs are numbered in an arbitrary order from 1 to L_N. We now continue by matching at random the first stub with one of the $L_N - 1$ remaining stubs. Once paired, two stubs form an 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 next stub and so on, until all stubs are connected.

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

$$\mathbb{P}(D_1 = j) = f_j, \quad j = 0, 1, 2, \ldots, \quad \text{and} \quad F(x) = \sum_{j=0}^{\lfloor x \rfloor} f_j, \tag{1.1}$$

where $\lfloor x \rfloor$ is the largest integer smaller than or equal to x. Our main assumption will be that

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

where $\tau \in (1, 2)$ and L is slowly varying at infinity. This means that the random variables D_j obey a power law with infinite mean. The factor L is meant to generalize the model.

1.2 Main results

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, where, by convention, this 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.

We will present in this paper two separate theorems for the case $\tau \in (1, 2)$. In the first theorem we take the sequence $D_1, D_2, \ldots, D_N$ an i.i.d. sequence with distribution F, satisfying (1.2), with

$\tau \in (1,2)$. The result is that the graph distance or hopcount converges in distribution to a limit random variable with mass $p = p_\tau$, $1-p$, on the values 2, 3, respectively.

One might argue that including degrees larger than $N-1$ is artificial in a network with N nodes. In fact, in many real networks, the degree is bounded by a physical constant. Therefore we also consider the case where the degrees are conditioned to be smaller than N^α, where α is an arbitrary positive number. Of course, we cannot condition on the degrees to be at most M, where M is fixed and independent on N, since in this case, the degrees are uniformly bounded, and this case is treated in [11]. Therefore, we consider cases where the degrees are conditioned to be at most a given power of N.

The result with conditioned degrees appears in the second theorem below. It turns out that for $\alpha > 1/(\tau-1)$, the conditioning has no influence in the sense that the limit random variable is the same as that for the unconditioned case. This is not so strange, since the maximal degree is of order $N^{1/(\tau-1)}$, so that the conditioning does nothing in this case. However, for $1/\tau < \alpha < 1/(\tau-1)$, the graph distance converges to a degenerate limit random variable with mass 1 on the value 3. It would be of interest to extend the possible conditioning schemes, but we will not elaborate further on it here.

Theorem 1.1 *Fix $\tau \in (1,2)$ in (1.2) and let $D_1, D_2, \ldots, D_N$ denote an i.i.d. sequence with common distribution F. Then,*

$$\lim_{N\to\infty} \mathbb{P}(H_N = 2) = 1 - \lim_{N\to\infty} \mathbb{P}(H_N = 3) = p, \tag{1.3}$$

where $p = p_\tau \in (0,1)$.

In the theorem below we will write $D^{(N)}$ for the random variable D conditioned on $D < N^\alpha$. Thus,

$$\mathbb{P}(D^{(N)} = k) = \frac{f_k}{\mathbb{P}(D < N^\alpha)}, \qquad 0 \le k < N^\alpha. \tag{1.4}$$

Theorem 1.2 *Fix $\tau \in (1,2)$ in (1.2) and let $D_1^{(N)}, D_2^{(N)}, \ldots, D_N^{(N)}$ be a sequence of i.i.d. copies of $D^{(N)}$.*

(i) *If $1/\tau < \alpha < 1/(\tau-1)$, then*

$$\lim_{N\to\infty} \mathbb{P}(H_N = 3) = 1. \tag{1.5}$$

(ii) *If $\alpha > 1/(\tau-1)$, then*

$$\lim_{N\to\infty} \mathbb{P}(H_N = 2) = 1 - \lim_{N\to\infty} \mathbb{P}(H_N = 3) = p, \tag{1.6}$$

where $p = p_\tau \in (0,1)$.

Remark: For $\alpha < 1/\tau$, we have reasons to believe that $\mathbb{P}(H_N > 3)$ remains uniformly positive when $N \to \infty$. A heuristic for this fact is given in the next section.

1.3 Heuristics

When $\tau \in (1,2)$, we consider two different cases. In Theorem 1.1, the degrees are not conditioned, while in Theorem 1.2 we condition on each node having a degree smaller than N^α. We will now give a heuristic explanation of our results.

In two previous papers [11, 12], we have treated the cases $\tau \in (2,3)$ and $\tau > 3$. In these cases, it follows that the probability mass function $\{f_j\}$ introduced in (1.1) has a finite mean, and the the number of nodes on graph distance k from node 1 can be coupled to the number of individuals in the k^{th} generation of a branching process with offspring distribution $\{g_j\}$ given by

$$g_j = \frac{j+1}{\mu} f_j, \tag{1.7}$$

where $\mu = \mathbb{E}[D_1]$. For $\tau \in (1, 2)$, as we are currently investigating, we have $\mu = \infty$, and the branching process used in [11, 12] does not exist.

When we do not condition on D_j being smaller than N^{α}, then L_N is the i.i.d. sum of N random variables $D_1, D_2, \ldots, D_N$, with infinite mean. It is well known that in this case the bulk of the contribution to $L_N \sim N^{1/(\tau-1)}$ comes from a *finite* number of nodes which have giant degrees (the so-called *giant nodes*). A basic fact in the configuration model is that two sets of stubs of sizes n and m are connected with high probability when nm is at least of order L_N. Since the giant nodes have degree roughly $N^{1/(\tau-1)}$, which is much larger than $\sqrt{L_N}$, they are all attached to each other, thus forming a complete graph of giant nodes. Each stub is with probability close to 1 attached to a giant node, and therefore, the distance between any two nodes is, with large probability, at most 3. In fact this distance equals 2 when the two nodes are connected to the *same* giant node, and is 3 otherwise.

When we truncate the distribution as in (1.4), with $\alpha > 1/(\tau - 1)$, we hardly change anything since without truncation with probability $1 - o(1)$ all degrees are below N^{α}. On the other hand, if $\alpha < 1/(\tau - 1)$ when, with truncation, the largest nodes have degree of order N^{α}, and $L_N \sim N^{1+\alpha(2-\tau)}$. Again, the bulk of the total degree L_N comes from nodes with degree of the order N^{α}, so that now these are the giant nodes. Hence, for $1/\tau < \alpha < 1/(\tau - 1)$, the largest nodes have degrees much larger than $\sqrt{L_N}$, and thus, with probability $1-o(1)$, still constitute a complete graph. The number of giant nodes converges to infinity, as $N \to \infty$. Therefore, the probability that two arbitrary nodes are connected to the *same* giant node converges to 0. Therefore, the hopcount equals 3 with probability converging to 1. If $\alpha < 1/\tau$, then the giant nodes no longer constitute a complete graph suggesting that the resulting hopcount can be greater than 3. It remains unclear to us whether the hopcount converges to a *single* value (as in Theorem 1.2), or to more than one possible value (as in Theorem 1.1). We do expect that the hopcount remains uniformly bounded.

The proof in this paper is based on detailed asymptotics of the sum of N i.i.d. random variables with infinite mean, as well as on the scaling of the order statistics of such random variables. The scaling of these order statistics is crucial in the definition of the giant nodes which are described above. The above considerations are the basic idea in the proof of Theorem 1.1. In the proof of Theorem 1.2, we need to investigate what the conditioning does to the scaling of both the total degree L_N, as well as to the largest degrees.

1.4 Related work

The above model is a variation of the configuration model. In the usual configuration model one often starts from a given *deterministic* degree sequence. In our model, the degree sequence is an i.i.d. sequence $D_1, \ldots, D_N$ with distribution equal to a power law. The reason for this choice is that we are interested in models for which all nodes are exchangeable, and this is not the case when the degrees are fixed. The study of this variation of the configuration model was started in [17] for the case $\tau > 3$ and studied by Norros and Reittu [18] in case $\tau \in (2, 3)$.

For a complete survey to complex networks, power law degree sequences and random graph models for such networks, see [3] and [16]. There a heuristic is given why the hopcount scales proportionally to $\log N$, which is originally from [17]. The argument uses a variation of the power law degree model, namely, a model where an exponential cut off is present. An example of such a degree distribution is

$$f_k = Ck^{-\tau}e^{-k/\kappa} \tag{1.8}$$

for some $\kappa > 0$. The size of κ indicates up to what degree the power law still holds, and where the exponential cut off starts to set in. The above model is treated in [11] for any $\kappa < \infty$, but, for $\kappa = \infty$, falls within the regimes where $\tau \in (2, 3)$ in [12] and within the regime in this paper for $\tau \in (1, 2)$. In [17], the authors conclude that since the limit as $\kappa \to \infty$ does not seem to converge, the 'average distance is not well-defined when $\kappa < 3$'. In this paper, as well as in [12], we show that the average distance *is* well-defined, but it scales differently from the case where $\tau > 3$.

In the paper [13], we give a survey to the results for the hopcount in the three different regimes $\tau \in (1,2)$, $\tau \in (2,3)$ and $\tau > 3$. There, we also prove results for the connectivity properties of the random graph in these cases. These results assume that the expected degree is larger than 2. This is always the case when $\tau \in (1,2)$, and stronger results have been shown there. We prove that the largest connected component has size $N(1+o(1))$ with probability converging to one. When $\tau \in (1, \frac{3}{2})$, we can even prove that with large probability, the graph is connected. When $\tau > \frac{3}{2}$, this is not true, and we investigate the structure of the remaining 'dust' that does not belong to the largest connected component. In the analysis, we will make use of the results obtained in this paper for $\tau \in (1,2)$. For instance, it will be crucial that the probability that two arbitrary nodes are connected converges to 1.

There is substantial related work on the configuration model for the case $\tau \in (2,3)$ and $\tau > 3$. References are included in the paper [12] for the case $\tau \in (2,3)$ and in [11] for $\tau > 3$. We again refer to the references in [13] and [3, 16] for more details. The graph distance for $\tau \in (1,2)$, that we study here, has to our best knowledge not been studied before. Values of $\tau \in (1,2)$ have been observed in networks of e-mail messages and networks where the nodes consist of software packages (see [16, Table II]), for which our configuration model with $\tau \in (1,2)$ can possibly give a good model.

In [1], random graphs are 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.* [1] show that the largest connected component is of the order of the size of the graph when $\tau < \tau_0 = 3.47875\ldots$, 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 with high probability connected. The proofs of these facts use couplings with branching processes and strengthen previous results due to Molloy and Reed [14, 15]. See also [1] for a history of the problem and references predating [14, 15]. The problem of distances in the configuration model with $\tau \in (1,2)$ has, up to our best knowledge, not been addressed. See [2] for an introduction to the mathematical results of various models for complex networks (also called massive graphs), as well as a detailed account of the results in [1].

A detailed account for a related model can be found in [6] and [7], where links between nodes i and j are present with probability equal to $w_i w_j / \sum_l w_l$ for some 'expected degree vector' $w = (w_1, \ldots, w_N)$. Chung and Lu [6] show that when w_i is proportional to $i^{-\frac{1}{\tau-1}}$ the average distance between pairs of nodes is $\log_\nu N(1+o(1))$ when $\tau > 3$, and $2\frac{\log\log N}{|\log(\tau-2)|}(1+o(1))$ when $\tau \in (2,3)$. In their model, also $\tau < 1$ is possible, and in this case, similarly to $\tau \in (1, \frac{3}{2})$ in our paper, the graph is connected with high probability.

The difference between this model and ours is that the nodes are not exchangeable in [6], but the observed phenomena are similar. This result can be understood as follows. Firstly, the actual degree vector in [6] should be close to the expected degree vector. Secondly, for the expected degree vector, we can compute that the number of components for which the degree is less than or equal to k equals

$$|\{i : w_i \le k\}| \propto |\{i : i^{-\frac{1}{\tau-1}} \le k\}| \approx k^{-\tau+1}.$$

Thus, one expects that the number of nodes with degree at most k decreases as $k^{-\tau+1}$, similarly as in our model. In [7], Chung and Lu study the sizes of the connected components in the above model. The advantage of working with the 'expected degree model' is that different links are present independently of each other, with makes this model closer to the random graph $G(p,N)$.

1.5 Organization of the paper

The main body of the paper consists of the proofs of Theorem 1.1 in Section 2 and the proof of Theorem 1.2 in Section 3. Both proofs contain a technical lemma and in order to make the argument more transparent, we have postponed the proofs of these lemmas to the appendix. Section 4 contains simulation results and some conclusions. **conclusions: to be added**

2 Proof of Theorem 1.1

In this section, we prove Theorem 1.1, which states that the hopcount between two arbitrary nodes has, with probability $1-o(1)$, a non-trivial distribution on 2 and 3.

We will use an auxiliary lemma, which is a modification of the extreme value theorem for the k largest degrees, $k \in \mathbb{N}$. We introduce

$$D_{(1)} \leq D_{(2)} \leq \cdots \leq D_{(N)},$$

to be the order statistics of $D_1, \ldots, D_N$, so that $D_{(1)} = \min\{D_1, \ldots, D_N\}$, $D_{(2)}$ is the second smallest degree, etc. Let (u_N) be an increasing sequence such that

$$\lim_{N\to\infty} N\left(1 - F(u_N)\right) = 1. \tag{2.1}$$

It is well known that the order statistics of the degrees, as well as the total degree, are governed by u_N in the case that $\tau \in (1,2)$. The following lemma shows this in some detail.

Lemma 2.1 *(a) for any $k \in \mathbb{N}$,*

$$\left(\frac{D_{(N)}}{u_N}, \ldots, \frac{D_{(N-k+1)}}{u_N}\right) \longrightarrow (\xi_1, \ldots, \xi_k), \text{ in distribution, as } N \to \infty, \tag{2.2}$$

where $(\xi_1, \ldots, \xi_k)$ is a random vector with marginals in $(0,\infty)$ and with joint distribution function given, for any tuple $0 < y_k < \cdots < y_1 < \infty$, by

$$\begin{aligned}&\mathbb{P}\left(\xi_1 < y_1, \ldots, \xi_k < y_k\right) \\ &\quad = \sum_{0\leq r_1 \leq \cdots \leq r_k < k} \frac{y_1^{(1-\tau)r_1}}{r_1!} \frac{(y_2^{1-\tau} - y_1^{1-\tau})^{r_2}}{r_2!} \cdots \frac{(y_k^{1-\tau} - y_{k-1}^{1-\tau})^{r_k}}{r_k!} e^{-y_k^{1-\tau}}.\end{aligned} \tag{2.3}$$

Moreover,

$$\xi_k \to 0, \text{ in probability, as } k \to \infty. \tag{2.4}$$

(b)

$$\frac{L_N}{u_N} \longrightarrow \eta, \text{ in distribution, as } N \to \infty,$$

where η is a random variable on $(0,\infty)$.

Proof. For part (a), we take $\rho_i = y_i^{1-\tau}$ and $u_N(\rho_i) = y_i u_N$, $i \in \{1, \ldots, k\}$. Since $u_N = L'(N) N^{\frac{1}{\tau-1}}$ (see e.g., [8]) for some slowly varying function $L'(N)$, it follows from (2.1) that,

$$\lim_{N\to\infty} N\left(1 - F(u_N(\rho_i))\right) = \rho_i, \qquad i \in \{1, \ldots, k\}.$$

Hence by [8, Theorem 4.2.6 and (4.2.4)], we have

$$\begin{aligned}&\lim_{N\to\infty} \mathbb{P}\left(D_{(N)} < u_N(\rho_1), \ldots, D_{(N-k+1)} < u_N(\rho_k)\right) \\ &\quad = \sum_{0\leq r_1 \leq \cdots \leq r_k < k} \frac{\rho_1{}^{r_1}}{r_1!} \frac{(\rho_2 - \rho_1)^{r_2}}{r_2!} \cdots \frac{(\rho_k - \rho_{k-1})^{r_k}}{r_k!} e^{-\rho_k} \\ &\quad = \sum_{0\leq r_1 \leq \cdots \leq r_k < k} \frac{y_1^{(1-\tau)r_1}}{r_1!} \frac{(y_2^{1-\tau} - y_1^{1-\tau})^{r_2}}{r_2!} \cdots \frac{(y_k^{1-\tau} - y_{k-1}^{1-\tau})^{r_k}}{r_k!} e^{-y_k^{1-\tau}}.\end{aligned}$$

We compute now the marginal distribution of ξ_k, for $k \geq 1$. For any $x > 0$, due to (2.3), we have

$$\begin{aligned}
\mathbb{P}(\xi_k < x) &= \mathbb{P}(\xi_1 < \infty, \dots, \xi_{k-1} < \infty, \xi_k < x) \\
&= \lim_{y_1,\dots,y_{k-1}\to\infty} \sum_{0\leq r_1 \leq \dots \leq r_{k-1} \leq r_k < k} e^{-x^{1-\tau}} \prod_{i=1}^{k} \frac{\left(y_i^{(1-\tau)r_i} - y_{i-1}^{(1-\tau)r_{i-1}}\right)}{r_i!} \\
&= \sum_{r=0}^{k-1} \frac{x^{(1-\tau)r}}{r!} e^{-x^{1-\tau}} \to 1, \text{ as } k \to \infty,
\end{aligned}$$

where in the middle expression, we write $y_0 = 0$ and $y_k = x$. Hence we have (2.4).

Part (b) follows since $L_N = D_1 + \cdots + D_N$ is in the domain of attraction of a stable law ([10, Corollary 2, XVII.5, p. 578]). □

We need some additional notation. In this section (Section 2) we define the *giant nodes* as the k_ε largest nodes, i.e., those nodes with degrees $D_{(N)}, \dots, D_{(N-k_\varepsilon+1)}$, where k_ε is some function of ε, to be chosen below. We define

$$A_{\varepsilon, N} = B_{\varepsilon, N} \cap C_{\varepsilon, N} \cap D_{\varepsilon, N}, \tag{2.5}$$

where

(i) $B_{\varepsilon, N}$ is the event that the stubs of node 1 and node 2 are attached exclusively to stubs of giant nodes.

(ii) $C_{\varepsilon, N}$ is the event that any two giant nodes are attached to each other; and

(iii) $D_{\varepsilon, N}$ is defined as

$$D_{\varepsilon, N} = \{D_1 \leq b_{D,\varepsilon} \cap D_2 \leq b_{D,\varepsilon}\},$$

where $b_{D,\varepsilon} = \min\{k : 1 - F(k) < \varepsilon/8\}$, so that $b_{D,\varepsilon} = \varepsilon^{-1/(\tau-1)(1+o(1))}$.

The sets $B_{\varepsilon, N}$ and $C_{\varepsilon, N}$ depend on the integer k_ε, which we will take to be large for ε small, and will be defined now. The choice of the index k_ε is rather technical, and depends on the distributional limits defined in Lemma 2.1. Since L_N/u_N converges in distribution to the random variable η, with support $(0, \infty)$, we can find $a_{\eta,\varepsilon}$, such that

$$\mathbb{P}(L_N < a_{\eta,\varepsilon} u_N) < \varepsilon/36, \quad \forall N. \tag{2.6}$$

This follows since convergence in distribution implies tightness of the sequence L_N/u_N ([4, p. 9]), so that we can find a closed subinterval $[a, b] \subset (0, \infty)$, with

$$\mathbb{P}(L_N/u_N \in [a, b]) > 1 - \varepsilon, \quad \forall N.$$

The definition of $b_{\xi_{k_\varepsilon}}$ is rather involved. It depends on ε, the quantile $b_{D,\varepsilon}$, the value $a_{\eta,\varepsilon}$ defined above and the value of $\tau \in (1, 2)$ and reads

$$b_{\xi_{k_\varepsilon}} = \left(\frac{\varepsilon^2 a_{\eta,\varepsilon}}{2304 b_{D,\varepsilon}}\right)^{\frac{1}{2-\tau}}, \tag{2.7}$$

where the peculiar integer 2304 is the product of 8^2 and 36. Given $b_{\xi_{k_\varepsilon}}$, we take k_ε equal to the minimal value such that

$$\mathbb{P}(\xi_{k_\varepsilon} \geq b_{\xi_{k_\varepsilon}}/2) \leq \varepsilon/72. \tag{2.8}$$

This we can do due to (2.4). We have now defined the constants that we will use in the proof, and we will next show that the probability of $A^c_{\varepsilon, N}$ is at most ε:

Lemma 2.2 *For each $\varepsilon > 0$, there exists N_ε, such that*

$$\mathbb{P}(A^c_{\varepsilon, N}) < \varepsilon, \quad N \geq N_\varepsilon. \tag{2.9}$$

The proof of this Lemma is rather technical and can be found in the appendix. We will now complete the proof of Theorem 1.1 subject to Lemma 2.2.

Proof of Theorem 1.1. The proof consist of two parts. The event $A_{\varepsilon,N}$, implies the event $\{H_N \le 3\}$, so that $\mathbb{P}(A^c_{\varepsilon,N}) < \varepsilon$ induces that $\{H_N \le 3\}$ with probability at least $1-\varepsilon$.

In the first part we show that $\mathbb{P}(\{H_N = 1\} \cap A_{\varepsilon,N}) < \varepsilon$. In the second part we prove that

$$\lim_{N\to\infty} \mathbb{P}(H_N = 2) = \lim_{\varepsilon\to 0}\lim_{N\to\infty} \mathbb{P}(\{H_N = 2\} \cap D_{\varepsilon,N} \,|\, B_{\varepsilon,N}) = p,$$

for the events $B_{\varepsilon,N}, D_{\varepsilon,N}$ defined above and some $0 < p < 1$. Since ε is arbitrary positive, the above statements yield the content of the theorem.

We first prove that $\mathbb{P}(\{H_N = 1\} \cap A_{\varepsilon,N}) < \varepsilon$ for sufficiently large N. The event $\{H_N = 1\}$ occurs *iff* at least one stub of node 1 connects to a stub of node 2. For $j \le D_1$, we denote by $\{[1.j] \to [2]\}$ the event that j-th stub of node 1 connects to a stub of node 2. Then, with $\mathbb{P}_N$ the conditional probability given the degrees $D_1, D_2, \ldots, D_N$,

$$\begin{aligned} \mathbb{P}(H_N = 1,\, A_{\varepsilon,N}) &\le \mathbb{E}\left[\sum_{j=1}^{D_1} \mathbb{P}_N(\{[1.j] \to [2]\},\, A_{\varepsilon,N})\right] \\ &\le \mathbb{E}\left[\sum_{j=1}^{D_1} \frac{D_2}{L_N - 1}\mathbf{1}\{A_{\varepsilon,N}\}\right] \le \frac{b^2_{D,\varepsilon}}{N-1} < \varepsilon, \end{aligned} \tag{2.10}$$

for large enough N, since $L_N \ge N$.

We prove now that $\lim_{N\to\infty} \mathbb{P}(H_N = 2) = p$, for some $0 < p < 1$. Since by definition for any $\varepsilon > 0$,

$$\max\{\mathbb{P}(B^c_{\varepsilon,N}), \mathbb{P}(D^c_{\varepsilon,N})\} \le \mathbb{P}(A^c_{\varepsilon,N}) \le \varepsilon,$$

we have that

$$\begin{aligned} &|\mathbb{P}(H_N = 2) - \mathbb{P}(\{H_N = 2\} \cap D_{\varepsilon,N} \,|\, B_{\varepsilon,N})| \\ &\le \left|\mathbb{P}(H_N = 2)\left(1 - \frac{1}{\mathbb{P}(B_{\varepsilon,N})}\right)\right| + \left|\frac{\mathbb{P}(H_N = 2) - \mathbb{P}(\{H_N = 2\} \cap D_{\varepsilon,N} \cap B_{\varepsilon,N})}{\mathbb{P}(B_{\varepsilon,N})}\right| \\ &\le \frac{2\mathbb{P}(B^c_{\varepsilon,N}) + \mathbb{P}(D^c_{\varepsilon,N})}{\mathbb{P}(B_{\varepsilon,N})} \le \frac{3\varepsilon}{1-\varepsilon}, \end{aligned}$$

uniformly in N, for N sufficiently large. If we show that

$$\lim_{N\to\infty} \mathbb{P}(\{H_N = 2\} \cap D_{\varepsilon,N} \,|\, B_{\varepsilon,N}) = p_{\tau,\varepsilon},$$

then there exists a double limit

$$p_\tau = \lim_{\varepsilon\to 0}\lim_{N\to\infty} \mathbb{P}(\{H_N = 2\} \cap D_{\varepsilon,N} \,|\, B_{\varepsilon,N}) = \lim_{N\to\infty} \mathbb{P}(H_N = 2).$$

Moreover, if we can bound $p_{\tau,\varepsilon}$ from 0 and 1, uniformly in ε, for ε small enough, then we also obtain that $0 < p_\tau < 1$.

First, to prove the existence of

$$\lim_{N\to\infty} \mathbb{P}(\{H_N = 2\} \cap D_{\varepsilon,N} \,|\, B_{\varepsilon,N}) = \lim_{N\to\infty} \mathbb{E}(\mathbb{P}_N(\{H_N = 2\} \cap D_{\varepsilon,N} \,|\, B_{\varepsilon,N})) = p_{\tau,\varepsilon}, \tag{2.11}$$

we will show that $\mathbb{P}_N(\{H_N = 2\} \cap D_{\varepsilon,N} \,|\, B_{\varepsilon,N})$ is a continuous function of the vector

$$\bar{D}_{k_\varepsilon} = \left(\frac{D_{(N)}}{u_N}, \ldots \frac{D_{(N-k_\varepsilon+1)}}{u_N}, \frac{1}{u_N}\right).$$

This vector, due to (2.2), converges in distribution to $(\xi_1, \dots, \xi_{k_\varepsilon}, 0)$. Hence, by the continuous mapping theorem [4, Theorem 5.1, p. 30], we have the existence of the limit (2.11). We now prove the claimed continuity.

We will show an even stronger statement, namely, that $\mathbb{P}_N(\{H_N = 2\} \cap D_{\varepsilon,N} \mid B_{\varepsilon,N})$ is the ratio of two polynomials of the components of the vector $\bar{D}_{k_\varepsilon}$, where the polynomial in the denominator is strictly positive. Indeed, the hopcount between nodes 1 and 2 is 2 iff both nodes are connected to the same giant node. For any $0 \le i \le D_1$, $0 \le j \le D_2$ and $0 \le k < k_\varepsilon$ let $\mathcal{A}_{i,j,k}$ be the event that both the i^{th} stub of node 1 and the j^{th} stub of node 2 are connected to the node with the $(N-k)^{\text{th}}$ largest degree. Then,

$$\mathbb{P}_N(\{H_N = 2\} \cap D_{\varepsilon,N} \mid B_{\varepsilon,N}) = \mathbb{P}_N\left(\bigcup_{i=1}^{D_1} \bigcup_{j=1}^{D_2} \bigcup_{k=0}^{k_\varepsilon - 1} \mathcal{A}_{i,j,k} \,\middle|\, B_{\varepsilon,N} \right),$$

where the r.h.s. can be written by inclusion-exclusion formula, as a linear combination of terms

$$\mathbb{P}_N(\mathcal{A}_{i_1,j_1,k_1} \cap \dots \cap \mathcal{A}_{i_n,j_n,k_n} \mid B_{\varepsilon,N}). \tag{2.12}$$

It is not difficult to see that these probabilities are ratios of polynomials. For example,

$$\begin{aligned} \mathbb{P}_N(\mathcal{A}_{i,j,k} \mid B_{\varepsilon,N}) &= \frac{D_{(N-k)}(D_{(N-k)} - 1)}{(D_{(N-k_\varepsilon+1)} + \dots + D_{(N)})(D_{(N-k_\varepsilon+1)} + \dots + D_{(N)} - 1)} \\ &= \frac{\frac{D_{(N-k)}}{u_N}(\frac{D_{(N-k)}}{u_N} - \frac{1}{u_N})}{(\frac{D_{(N-k_\varepsilon+1)}}{u_N} + \dots + \frac{D_{(N)}}{u_N})(\frac{D_{(N-k_\varepsilon+1)}}{u_N} + \dots + \frac{D_{(N)}}{u_N} - \frac{1}{u_N})}. \end{aligned} \tag{2.13}$$

Similar arguments hold for general terms of the form in (2.12). Hence, $\mathbb{P}_N(\{H_N = 2\} \cap D_{\varepsilon,N} \mid B_{\varepsilon,N})$ itself can be written as a ratio of two polynomials where the polynomial in the denominator is strictly positive. Therefore, the limit (2.11) exists.

We finally bound p_ε from 0 and 1 uniformly in ε, for any $\varepsilon < 1/2$. Since the hopcount between nodes 1 and 2 is 2, given $B_{\varepsilon,N}$, if they are both connected to the node with largest degree,

$$\mathbb{P}(\{H_N = 2\} \cap D_{\varepsilon,N} \mid B_{\varepsilon,N}) \ge \mathbb{E}\left[\mathbb{P}_N(\mathcal{A}_{1,1,(N)} \mid B_{\varepsilon,N})\right],$$

and by (2.13) we have

$$\begin{aligned} p_{\tau,\varepsilon} &= \lim_{N\to\infty} \mathbb{P}(\{H_N = 2\} \cap D_{\varepsilon,N} \mid B_{\varepsilon,N}) \ge \lim_{N\to\infty} \mathbb{E}\left[\frac{D_{(N)}(D_{(N)}-1)}{(D_{(N)} + \dots + D_{(N-k_\varepsilon+1)} - 1)^2}\right] \\ &= \mathbb{E}\left[\left(\frac{\xi_1}{\xi_1 + \dots + \xi_{k_\varepsilon}}\right)^2\right] \ge \mathbb{E}\left[\left(\frac{\xi_1}{\eta}\right)^2\right]. \end{aligned}$$

On the other hand, the hopcount between nodes 1 and 2 is at most 3, given $B_{\varepsilon,N}$ when all stubs of the node 1 are connected to the node with largest degree, and all stubs of the node 2 are connected to the node with the one but largest degree. Hence, for any $\varepsilon < 1/2$ and similarly to (2.13), we have

$$\begin{aligned} p_{\tau,\varepsilon} &= \lim_{N\to\infty} \mathbb{P}(\{H_N = 2\} \cap D_{\varepsilon,N} \mid B_{\varepsilon,N}) \le 1 - \lim_{N\to\infty} \mathbb{P}(\{H_N > 2\} \cap D_{\varepsilon,N} \mid B_{\varepsilon,N}) \\ &\le 1 - \lim_{N\to\infty} \mathbb{P}\left(\{H_N > 2\} \cap D_{\frac{1}{2},N} \mid B_{\varepsilon,N}\right) \\ &\le 1 - \lim_{N\to\infty} \mathbb{E}\left[\left(\prod_{i=0}^{D_1} \frac{D_{(N)} - 2i}{D_{(N)} + \dots + D_{(N-k_\varepsilon+1)} - D_1}\right)\left(\prod_{i=0}^{D_2} \frac{D_{(N-1)} - 2i}{D_{(N)} + \dots + D_{(N-k_\varepsilon+1)} - D_2}\right) \mathbf{1}\{D_{\frac{1}{2},N}\}\right] \\ &\le 1 - \tfrac{1}{2}\mathbb{E}\left[\left(\frac{\xi_1}{\xi_1 + \dots + \xi_{k_\varepsilon}}\right)^{b_{D,1/2}}\left(\frac{\xi_2}{\xi_1 + \dots + \xi_{k_\varepsilon}}\right)^{b_{D,1/2}}\right] \le 1 - \tfrac{1}{2}\mathbb{E}\left[\left(\frac{\xi_1\xi_2}{\xi^2}\right)^{b_{D,1/2}}\right]. \end{aligned}$$

Both the upper and lower bound are strictly positive, independently of ε. Hence, for any $\varepsilon < 1/2$, the quantity $p_{\tau,\varepsilon}$ is bounded away from 0 and 1, where the bounds are *independent of* ε, and thus $0 < p = p_\tau < 1$. □

This completes the proof of Theorem 1.1 subject to Lemma 2.2.

3 Proof of Theorem 1.2

In Theorem 1.2, we consider the hopcount in the configuration model with degrees an i.i.d. sequence with distribution given by (1.4), where D has distribution F satisfying (1.2). We distinguish two cases: (i) $1/\tau < \alpha < 1/(\tau-1)$ and (ii) $\alpha > 1/(\tau-1)$.

We first prove part (ii), which states that the limit distribution of H_N is a mixed distribution with probability mass p on 2 and probability mass $1-p$ on 3, for some $0 < p < 1$. Part (ii) of Theorem 1.2 is almost immediate from Theorem 1.1. As before we denote by $D_1, D_2, \ldots, D_N$ the i.i.d. sequence without conditioning, then $\mathbb{P}(\cup_{i=1}^N \{D_i > N^\alpha\})$, which is the probability that for at least one index $i \in \{1, 2, \ldots, N\}$ the degree D_i exceeds N^α, is bounded by

$$\sum_{i=1}^N \mathbb{P}(D_i > N^\alpha) = N\mathbb{P}(D_1 > N^\alpha) = N^{1+\alpha(1-\tau)}L(N) = N^{-\varepsilon},$$

for some positive ε, because $\alpha > 1/(\tau-1)$. We can therefore couple the i.i.d. sequence $\vec{D}^{(N)} = (D_1^{(N)}, D_2^{(N)}, \ldots, D_N^{(N)})$ to the sequence $\vec{D} = (D_1, D_2, \ldots, D_N)$, where the probability of a miscoupling, i.e., a coupling such that $\vec{D}^{(N)} \neq \vec{D}$, is at most $N^{-\varepsilon}$. Therefore, the result of Theorem 1.1 carries over to case (ii) in Theorem 1.2.

We now turn to case (i) in Theorem 1.2. We must prove that if we condition the degrees to be smaller than N^α with $1/\tau < \alpha < 1/(\tau-1)$, then the graph distance between two arbitrary nodes is 3 with probability equal to $1 - o(1)$. We define

$$L_N^{(N)} = \sum_{n=1}^N D_n^{(N)},$$

to be the total degree of the conditioned model. The steps in the proof Theorem 1.2(i) are identical to those in the proof of Theorem 1.1, however, the details differ considerably. We take an arbitrary $\varepsilon > 0$ and define an event $G_{\varepsilon,N}$ such that $\mathbb{P}(G^c_{\varepsilon,N}) < \varepsilon$, and prove subsequently that for large enough N, $\mathbb{P}(\{H_N = 1\} \cap G_{\varepsilon,N}) < \varepsilon$, $\mathbb{P}(\{H_N = 2\} \cap G_{\varepsilon,N}) < \varepsilon$, and that $\mathbb{P}(\{H_N \leq 3\} \cap G_{\varepsilon,N}) \geq 1-\varepsilon$. The proof that $\mathbb{P}(G^c_{\varepsilon,N}) < \varepsilon$ is quite technical and moved to the appendix. Since $\varepsilon > 0$ is arbitrary, the above statements imply that

$$\lim_{N\to\infty} \mathbb{P}(H_N = 3) = 1.$$

Let $\varepsilon > 0$ be fixed. The event $G_{\varepsilon,N}$ is defined by

$$G_{\varepsilon,N} = \{D_1^{(N)} \leq b_{D,\varepsilon},\ D_2^{(N)} \leq b_{D,\varepsilon},\ L_N^{(N)} \geq L_\varepsilon(N)N^{1+\alpha(2-\tau)}\},$$

where, as before, $b_{D,\varepsilon} = \min\{k : 1 - F(k) < \varepsilon/8\}$ and $L_\varepsilon(N)$ is some function satisfying for each $\delta > 0$ that

$$\liminf_{N\to\infty} N^\delta L_\varepsilon(N) > 1.$$

Lemma 3.1 *For each $\varepsilon > 0$, there exists N_ε, such that, for all $N \geq N_\varepsilon$,*

$$\mathbb{P}(G^c_{\varepsilon,N}) < \varepsilon. \tag{3.1}$$

As explained above, the proof of this lemma is rather technical and can be found in the appendix.

Proof of Theorem 1.2, part (i). We start with the proof that $\mathbb{P}(\{H_N = 1\} \cap G_{\varepsilon,N}) < \varepsilon$, for sufficiently large N. The event $\{H_N = 1\}$ occurs *iff* at least one stub of node 1 connects to a stub of node 2. As before we denote for $j \le D_1^{(N)}$ by $\{[1.j] \to [2]\}$ the event that j-th stub of node 1 attaches to a stub of node 2. Then

$$\begin{aligned}\mathbb{P}\left(\{H_N = 1\} \cap G_{\varepsilon,N}\right) \le & \;\; \mathbb{E}\left[\sum_{j=1}^{D_1^{(N)}} \mathbb{P}_N\left([1.j] \to [2],\, G_{\varepsilon,N}\right)\right] \\ \le & \;\; \mathbb{E}\left[\sum_{j=1}^{D_1^{(N)}} \frac{D_2^{(N)}}{L_N^{(N)} - 1}\mathbf{1}\{G_{\varepsilon,N}\}\right] \le \frac{b_{D,\varepsilon}^2}{N-1} < \varepsilon,\end{aligned} \tag{3.2}$$

for large enough N, since $L_N^{(N)} \ge N$.

We will now bound $\mathbb{P}\left(\{H_N = 2\} \cap G_{\varepsilon,N}\right)$. Note that the event $\{H_N = 2\}$ occurs *iff* there exists node $k \ge 3$ and two stubs of node k such that the first one connects to a stub of node 1 and the second one connects to a stub of node 2. For $k \ge 3$ such that $D_k^{(N)} \ge 2$, $i, j \le D_k^{(N)}$, $i \ne j$, we denote by $\{[k.i] \to [1],\, [k.j] \to [2]\}$ the event that the i-th stub of node k connects to a stub of node 1 and the j-th stub of node k connects to a stub of node 2. Then,

$$\begin{aligned}\mathbb{P}\left(\{H_N = 2\} \cap G_{\varepsilon,\,N}\right) \le & \;\; \mathbb{E}\left[\sum_{k=3}^{N} \sum_{i\ne j}^{D_k^{(N)}} \mathbb{P}_N\left(\{[k.i] \to [1],\, [k.j] \to [2]\} \cap G_{\varepsilon,\,N}\right)\right] \\ \le & \;\; \mathbb{E}\left[\sum_{k=3}^{N} \sum_{i\ne j}^{D_k^{(N)}} \frac{D_1^{(N)}}{L_N^{(N)} - 1}\frac{D_2^{(N)}}{L_N^{(N)} - 3}\mathbf{1}\{G_{\varepsilon,N}\}\right] \\ \le & \;\; N\mathbb{E}\left((D^{(N)})^2\right) \frac{b_{D,\varepsilon}^2}{(L_\varepsilon(N)N^{1+\alpha(2-\tau)}-3)^2}.\end{aligned} \tag{3.3}$$

Observe that

$$\begin{aligned}\mathbb{E}[(D^{(N)})^2] &= \textstyle\sum_{n<N^\alpha}(2n-1)\mathbb{P}(D_1 \ge n | D_1 < N^\alpha) \\ &= \frac{1}{\mathbb{P}(D_1<N^\alpha)} \textstyle\sum_{n<N^\alpha}(2n-1)L(n-1)(n-1)^{1-\tau} = L_4(N)N^{\alpha(3-\tau)},\end{aligned}$$

for some slowly varying function L_4. Substitution of this upper bound in the right-hand side of (3.3) shows that for N large enough,

$$\mathbb{P}\left(\{H_N = 2\} \cap G_{\varepsilon,\,N}\right) \le N L_4(N) N^{\alpha(3-\tau)} \frac{b_{D,\varepsilon}^2}{(L_\varepsilon(N)N^{1+\alpha(2-\tau)} - 3)^2} < \varepsilon,$$

because $\alpha(3-\tau) < 2(1+\alpha(2-\tau))$ when $\alpha < 1/(1-\tau)$.

We will complete the proof by showing that, for sufficiently large N,

$$\mathbb{P}(\{H_N \le 3\} \cap G_{\varepsilon,\,N}) \ge 1 - \varepsilon. \tag{3.4}$$

Let $\beta = (1 + \alpha(4-\tau))/4$. Since $1/\tau < \alpha < 1/(\tau - 1)$, we have

$$\frac{1+\alpha(2-\tau)}{2} < \beta < \alpha. \tag{3.5}$$

In this section we will call a node k, for $1 \le k \le N$, a *giant node* if its degree $D_k^{(N)}$ satisfies

$$N^\beta < D_k^{(N)} < N^\alpha. \tag{3.6}$$

We will show below that with probability close to 1 at least one of the stubs of node 1 and at least one of the stubs of node 2 are connected to stubs of giant nodes, and that any two giant nodes have mutual graph distance 1. This implies that the hopcount between the nodes 1 and 2 is at most 3.

The non-giant nodes, i.e. nodes with degree less than or equal to N^β, are called *normal nodes* . First we will show that the total degree of the normal nodes is negligible with respect to $L_N^{(N)}$. The mean degree of a normal node is

$$\begin{aligned}\mathbb{E}[D^{(N)}\mathbf{1}\{D \le N^\beta\}] &= \sum_{n=1}^{\lfloor N^\beta \rfloor} \mathbb{P}(D \ge n | D < N^\alpha) \\ &= \tfrac{1}{\mathbb{P}(D<N^\alpha)} \sum_{n=1}^{\lfloor N^\beta \rfloor} L(n-1)(n-1)^{1-\tau} = L_5(N)N^{\beta(2-\tau)},\end{aligned}$$

for some slowly varying function L_5. Thus, by the Markov inequality,

$$\mathbb{P}\left(\sum_{i=1}^N D_i^{(N)}\mathbf{1}\{D_i \le N^\beta\} \ge \frac{3}{\varepsilon}L_5(N)N^{1+\beta(2-\tau)}\right) \le \frac{\varepsilon}{3},$$

so that, with probability at least $1-\varepsilon/3$, the fraction of the contribution from normal nodes on $G_{\varepsilon,N}$, for sufficiently large N, is at most

$$\frac{\frac{3}{\varepsilon}L_5(N)N^{1+\beta(2-\tau)}}{L_\varepsilon(N)N^{1+\alpha(2-\tau)}} = \frac{3L_5(N)}{\varepsilon L_\varepsilon(N)}N^{2(\beta-\alpha)}.$$

Since $\beta < \alpha$, for $\tau \in (1,2)$ the above fraction tends to 0, as $N \to \infty$. Thus the total contribution of the normal nodes is negligible with respect to $L_N^{(N)}$ on $G_{\varepsilon,N}$. This implies that for sufficiently large N, with probability at most $1-2\varepsilon/3$, both nodes 1 and 2 are at graph distance 1 from some giant node on $G_{\varepsilon,N}$. It remains to show that any two giant nodes have graph distance 1 with probability at least $1-\varepsilon/3$. For this we need an upper bound of $L_N^{(N)}$. Similarly as above we obtain

$$\mathbb{E}[D_1^{(N)}] = L_6(N)N^{\alpha(2-\tau)},$$

for some slowly varying function L_6. Hence from the Markov inequality and since
$L_N^{(N)} = D_1^{(N)} + \cdots + D_N^{(N)}$,

$$\mathbb{P}\left(L_N^{(N)} > \frac{6}{\varepsilon}L_6(N)N^{1+\alpha(2-\tau)}\right) \le \frac{\varepsilon N L_6(N)N^{\alpha(2-\tau)}}{6L_6(N)N^{1+\alpha(2-\tau)}} \le \frac{\varepsilon}{6}.$$

Since we have at most $N(N-1) < N^2$ pairs of giant nodes, the probability that two of them, say g_1 and g_2, have graph distance greater than 1 is at most (compare (4.5) of [11]),

$$\begin{aligned}&\mathbb{E}\left(N^2 \prod_{j=0}^{\lfloor D_{g_1}/2\rfloor - 1}\left(1 - \frac{D_{g_2}}{L_N^{(N)} - 2i - 1}\right)\mathbf{1}\{L_N^{(N)} \le \frac{6}{\varepsilon}L_6(N)N^{1+\alpha(2-\tau)}\}\right) \\ &+\mathbb{P}\left((L_N^{(N)} > \frac{6}{\varepsilon}L_6(N)N^{1+\alpha(2-\tau)}\right) \le N^2\left(1 - \frac{N^{\lfloor\beta\rfloor}}{\frac{6}{\varepsilon}L_6(N)N^{1+\alpha(2-\tau)}}\right)^{\frac{1}{2}N^\beta} + \frac{\varepsilon}{6} \\ &= N^2\left(e^{-1} + o(1)\right)^{\varepsilon N^{\frac{2\beta-(1+\alpha(2-\tau))}{(12L_6(N))}}} + \frac{\varepsilon}{6} \le \frac{\varepsilon}{3},\end{aligned}$$

for large N, since the exponent grows faster than any power of N. Thus, with probability at least $1-\frac{\varepsilon}{3}$, all giant nodes are on graph distance 1, and thus form a complete graph. This completes the proof. □

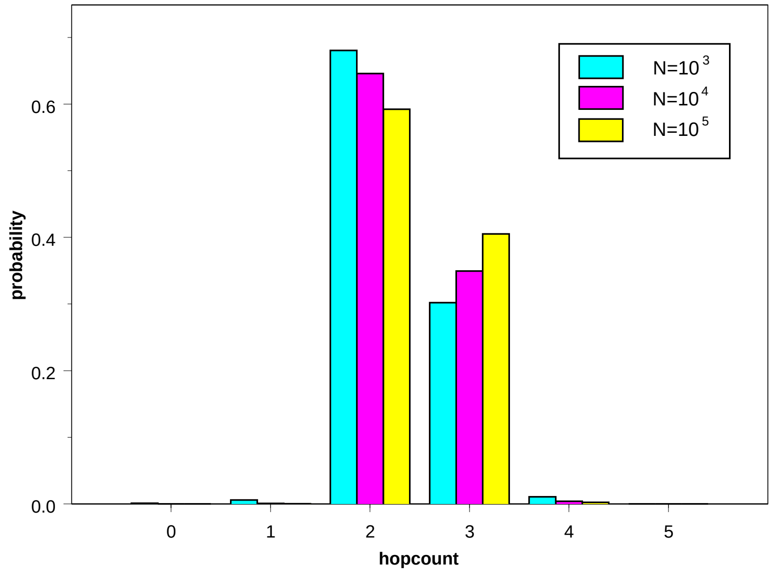

Figure 1: Empirical probability mass function of the hopcount for $\tau = 1.8$ and $N = 10^3, 10^4, 10^5$, for the unconditioned degrees.

4 Simulation and conclusions

To illustrate Theorem 1.1 and Theorem 1.2, we have simulated our random graph with degree distribution $D = \lceil U^{-\frac{1}{\tau-1}} \rceil$, where U is uniformly distributed over $(0, 1)$ and where for $x \in \mathbb{R}$, $\lceil x \rceil$ is the smallest integer greater than or equal to x. Thus,

$$1 - F(k) = \mathbb{P}(U^{-\frac{1}{\tau-1}} > k) = k^{1-\tau}, \quad k = 1, 2, 3, \ldots$$

In Figure 1, we have simulated the graph distance or hopcount with $\tau = 1.8$ and the values of $N = 10^3, 10^4, 10^5$. The histogram is in accordance with Theorem 1.1: for increasing values of N we see that the probability mass is divided over the values $H_N = 2$ and $H_N = 3$, where the probability $\mathbb{P}(H_N = 2)$ converges.

As an illustration of Theorem 1.2, we again take $\tau = 1.8$, but now conditioned the degrees to be less than N, so that $\alpha = 1$. Since in this case $(\tau - 1)^{-1} = \frac{5}{4}$, we expect from Theorem 1.2 case (i), that in the limit the hopcount will concentrate on the value $H_N = 3$. This is indeed the case as is shown in Figure 2

Our results give convincing asymptotics for the hopcount when the mean degree is infinite, using extreme value theory. Some interesting problems remain, such as

(i) Can one compute the exact value of p_τ?

(ii) What is the limit behavior of the hopcount when we condition the degrees on being less than N^α, with $\alpha < 1/\tau$?

Acknowledgement

The work of RvdH and DZ was supported in part by Netherlands Organisation for Scientific Research (NWO).

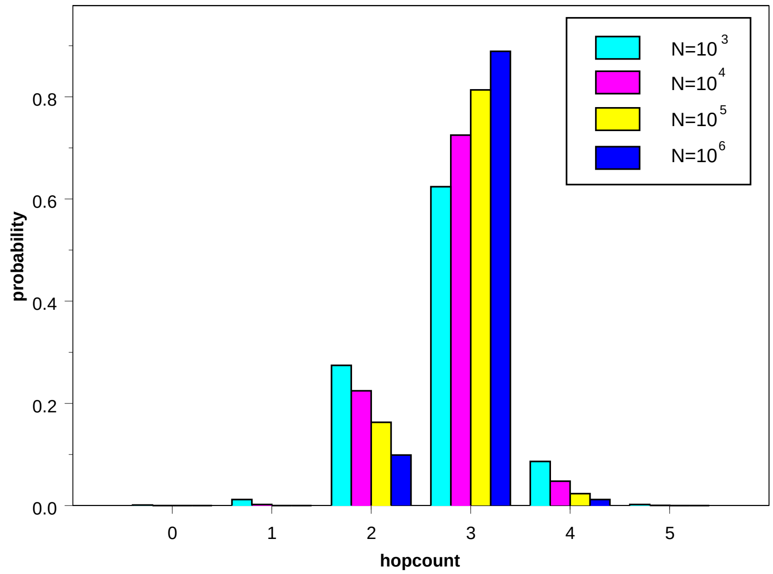

Figure 2: Empirical probability mass function of the hopcount for $\tau = 1.8$ and $N = 10^3, 10^4, 10^5, 10^6$, where the degrees are conditioned to be less than N.

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. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**: 47-97, 2002.

[4] P. Billingsley. *Convergence of Probability Measures.* Wiley, New York, 1968.

[5] B. Bollobás. *Random Graphs.* 2nd edition. Academic Press, 2001.

[6] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. PNAS, **99**(25), 15879–15882, 2002.

[7] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. Annals of Combinatorics, **6**, 125–145, 2002.

[8] P. Embrechts, C. Klüppelberg and T. Mikosch. *Modelling Extremal Events.* Springer Verlag, 1997.

[9] C. Faloutsos, P. Faloutsos and M. Faloutsos. On power-law relationships of the internet topology, Computer Communications Rev., **29**, 251-262, 1999.

[10] W. Feller. *An Introduction to Probability Theory and Its Applications* Volume II, 2nd edition, John Wiley and Sons, New York, 1971.

[11] R. van der Hofstad, G. Hooghiemstra and P. Van Mieghem. Distnaces in random graphs with finite variance degrees. Preprint 2003, submitted to Random Structures and Algorithms.

[12] R. van der Hofstad, G. Hooghiemstra and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. In preparation.

[13] R. van der Hofstad, G. Hooghiemstra and D. Znamenski. Random graphs with arbitrary i.i.d. degrees. Survey paper. In preparation.

[14] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, **6**, 161-179, 1995.

[15] 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.

[16] M.E.J. Newman The structure and function of complex networks. *SIAM Rev.* **45**, no. 2, 167–256, 2003.

[17] M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Random graphs with arbitrary degree distribution and their application. Pys. Rev. E, **64,** 026118, 2000.

[18] H. Reittu and I. Norros. On large random graphs of the "Internet type". Performance Evaluation, **55** (1-2), 3-23, 2004.

[19] S. H. Strogatz. Exploring complex networks. *Nature*, 410(8), 268–276, 2001.

[20] D. J. Watts. *Small Worlds, The Dynamics of Networks between Order and Randomness.* Princeton University Press, Princeton, New Jersey, 1999.

A Appendix.

A.1 Proof of Lemma 2.2

In this section we restate Lemma 2.2 and then give a proof.

Lemma A.1.1 *For each $\varepsilon > 0$, there exists N_ε such that*

$$\mathbb{P}(A^c_{\varepsilon,N}) < \varepsilon, \quad N \geq N_\varepsilon. \tag{A.1.1}$$

Proof. We define the event $E_{\varepsilon,N}$ by

$$\begin{aligned} E_{\varepsilon,N} = \quad & \left\{ \sum_{n=0}^{k_\varepsilon} D_{(N-n)} \leq \frac{\varepsilon}{8b_{D,\varepsilon}} L_N \right\} & (a) \\ & \cap \left\{ D_{(N-k_\varepsilon+1)} \geq a_{\xi_{k_\varepsilon}} u_N \right\} & (b) \\ & \cap \left\{ L_N \leq b_{\eta,\varepsilon} u_N \right\}, & (c) \end{aligned} \tag{A.1.2}$$

where $b_{D,\varepsilon}$ is the ε-quantile of F used in the definition of $D_{\varepsilon,N}$ and where $a_{\xi_{k_\varepsilon}}, b_{\eta,\varepsilon} > 0$ are defined by

$$\mathbb{P}\left(\xi_{k_\varepsilon} < a_{\xi_{k_\varepsilon}}\right) < \varepsilon/24 \quad \text{and} \quad \mathbb{P}(\eta > b_{\eta,\varepsilon}) < \varepsilon/24,$$

respectively. Observe that $a_{\xi_{k_\varepsilon}}$ is a lower quantile of ξ_{k_ε}, whereas $b_{\xi_{k_\varepsilon}}$ defined in (2.7) and (2.8) is an upper quantile of ξ_{k_ε}. Furthermore, $b_{\eta,\varepsilon}$ is an upper quantile of η, whereas $a_{\eta,\varepsilon}$ defined in (2.6) is a lower quantile of η. Since

$$A_{\varepsilon,N} = B_{\varepsilon,N} \cap C_{\varepsilon,N} \cap D_{\varepsilon,N},$$

(see (2.5) and below the proof of Lemma 2.1 for the definition of $A_{\varepsilon,N}$, $B_{\varepsilon,N}$, $C_{\varepsilon,N}$ and $D_{\varepsilon,N}$). we have

$$\mathbb{P}(A^c_{\varepsilon,N}) \leq \mathbb{P}(B^c_{\varepsilon,N} \cap D_{\varepsilon,N} \cap E_{\varepsilon,N}) + \mathbb{P}(C^c_{\varepsilon,N} \cap D_{\varepsilon,N} \cap E_{\varepsilon,N}) + \mathbb{P}(D^c_{\varepsilon,N}) + \mathbb{P}(E^c_{\varepsilon,N}), \tag{A.1.3}$$

and in order to prove the lemma we should show that each of the four terms on the right-hand side is at most $\varepsilon/4$.

Since on $D_{\varepsilon,N}$, the nodes 1 and 2 each have at most $b_{D,\varepsilon}$ stubs, the first term satisfies

$$\mathbb{P}(B^c_{\varepsilon,N}\cap D_{\varepsilon,N}\cap E_{\varepsilon,N})\leq 2b_{D,\varepsilon}\mathbb{E}\left(\frac{1}{L_N}\sum_{n=0}^{k_\varepsilon}D_{(N-n)}\mathbf{1}\{E_{\varepsilon,N}\}\right)\leq \varepsilon/4,$$

due to point (a) of $E_{\varepsilon,N}$. This bounds the first term of (A.1.3).

We turn to the second term of (A.1.3). Recall that $C^c_{\varepsilon,N}$ induces that no stubs of at least two giant nodes are attached to one another. Since we have at most N^2 pairs of giant nodes g_1 and g_2, the items (b), (c) of $E_{\varepsilon,N}$ imply

$$\begin{aligned}\mathbb{P}(C^c_{\varepsilon,N}\cap D_{\varepsilon,N}\cap E_{\varepsilon,N}) &\leq \mathbb{E}\left(N^2\prod_{j=0}^{\lfloor D_{g_1}/2\rfloor-1}\left(1-\frac{D_{g_2}}{L_N-2i-1}\right)\right)\\ &\leq N^2\left(1-\frac{a_{\xi_{k_\varepsilon}}u_N}{b_{\eta,\varepsilon}u_N}\right)^{a_{\xi_{k_\varepsilon}}u_N/2}\\ &= N^2\left(e^{-1}+o(1)\right)^{(a_{\xi_{k_\varepsilon}})^2/(2b_{\eta,\varepsilon})u_N}\\ &= N^2\left(e^{-1}+o(1)\right)^{(a_{\xi_{k_\varepsilon}})^2/(2b_{\eta,\varepsilon})L'(N)N^{\frac{1}{\tau-1}}}\leq \frac{\varepsilon}{4},\end{aligned}$$

for large enough N, since the exponent grows faster than any power of N.

The third term at the r.h.s. of (A.1.3) is at most $\varepsilon/4$, because

$$\mathbb{P}(D^c_{\varepsilon,N})\leq 2\mathbb{P}(D_1>b_{D,\varepsilon})\leq 2\varepsilon/8=\varepsilon/4.$$

It remains to estimate the last term at the r.h.s. of (A.1.3). Clearly,

$$\begin{aligned}\mathbb{P}\left(E^c_{\varepsilon,N}\right)\leq\ &\mathbb{P}\left(\sum_{n=0}^{k_\varepsilon}D_{(N-n)}>\frac{\varepsilon}{8b_{D,\varepsilon}}L_N\right) &&(a)\\ &+\mathbb{P}\left(D_{(N-k_\varepsilon+1)}<a_{\xi_{k_\varepsilon}}u_N\right) &&(b)\\ &+\mathbb{P}\left(L_N>b_{\eta,\varepsilon}u_N\right). &&(c)\end{aligned}\tag{A.1.4}$$

We will consequently show that each term in the above expression is at most $\varepsilon/12$. Let $a_{\eta,\varepsilon}$ and $b_{\xi_{k_\varepsilon}}>0$ be as in (2.6) and (2.7), then we can decompose the first term at the r.h.s. of (A.1.4) as

$$\begin{aligned}\mathbb{P}\left(\sum_{n=0}^{k_\varepsilon}D_{(N-n)}>\frac{\varepsilon}{8b_{D,\varepsilon}}L_N\right) &\leq \mathbb{P}\left(L_N<a_{\eta,\varepsilon}u_N\right)+\mathbb{P}\left(\sum_{n=0}^{k_\varepsilon}D_{(N-n)}>\frac{\varepsilon}{8b_{D,\varepsilon}}a_{\eta,\varepsilon}u_N\right)\\ &\leq \mathbb{P}\left(L_N<a_{\eta,\varepsilon}u_N\right)+\mathbb{P}\left(D_{(N-k_\varepsilon+1)}>b_{\xi_{k_\varepsilon}}u_N\right)\\ &\quad+\mathbb{P}\left(\sum_{i=1}^{N}D_i\mathbf{1}\{D_i<b_{\xi_{k_\varepsilon}}u_N\}>\frac{\varepsilon}{8b_{D,\varepsilon}}a_{\eta,\varepsilon}u_N\right).\end{aligned}\tag{A.1.5}$$

From the Markov inequality,

$$\mathbb{P}\left(\sum_{i=1}^{N}D_i\mathbf{1}\{D_i<b_{\xi_{k_\varepsilon}}u_N\}>\frac{\varepsilon}{8b_{D,\varepsilon}}a_{\eta,\varepsilon}u_N\right)\leq\frac{N\mathbb{E}\left(D_1\mathbf{1}\{D_1<b_{\xi_{k_\varepsilon}}u_N\}\right)}{\frac{\varepsilon}{8b_{D,\varepsilon}}a_{\eta,\varepsilon}u_N}.\tag{A.1.6}$$

Since $1-F(x)$ varies regularly with exponent $\tau-1$ we have, by [10, Theorem 1(b), VIII.9, p.281],

$$\mathbb{E}\left(D_1\mathbf{1}\{D_1<b_{\xi_{k_\varepsilon}}u_N\}\right)=\sum_{k=0}^{\lfloor b_{\xi_{k_\varepsilon}}u_N\rfloor}(1-F(k))\leq 2(2-\tau)b_{\xi_{k_\varepsilon}}u_N\left(1-F(b_{\xi_{k_\varepsilon}}u_N)\right),\tag{A.1.7}$$

for large enough N. Due to (2.1), for large enough N, we have also

$$N\left(1-F(u_N)\right) \leq 2. \tag{A.1.8}$$

Substituting (A.1.7) and (A.1.8) in (A.1.6), we obtain

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{N} D_i \mathbf{1}\{D_i < b_{\xi_{k_\varepsilon}} u_N\} > \frac{\varepsilon}{8b_{D,\varepsilon}} a_{\eta,\varepsilon} u_N\right) &\leq \frac{2N(2-\tau) b_{\xi_{k_\varepsilon}} u_N \left(1-F(b_{\xi_{k_\varepsilon}} u_N)\right)}{\frac{\varepsilon}{8b_{D,\varepsilon}} u_N a_{\eta,\varepsilon}} \\ &\leq \frac{4(2-\tau) b_{\xi_{k_\varepsilon}} \left(1-F(b_{\xi_{k_\varepsilon}} u_N)\right)}{\frac{\varepsilon}{8b_{D,\varepsilon}} a_{\eta,\varepsilon}\left(1-F(u_N)\right)},\end{aligned} \tag{A.1.9}$$

for large enough N. Since $1-F(x)$ varies regularly with exponent $\tau - 1$,

$$\lim_{N\to\infty} \frac{\left(1-F(b_{\xi_{k_\varepsilon}} u_N)\right)}{\left(1-F(u_N)\right)} = \left(b_{\xi_{k_\varepsilon}}\right)^{1-\tau}.$$

Hence the r.h.s. of (A.1.9) is at most

$$\frac{8(2-\tau)\left(b_{\xi_{k_\varepsilon}}\right)^{2-\tau}}{\frac{\varepsilon}{8b_{D,\varepsilon}} a_{\eta,\varepsilon}} = \varepsilon/36,$$

for sufficiently large N, by definition of $b_{\xi_{k_\varepsilon}}$ in (2.7). We now show that the second term of (A.1.5) is at most $\varepsilon/36$. Since $D_{(N-k_\varepsilon+1)}/u_N$ converges in distribution to ξ_{k_ε}, we find from (2.8),

$$\mathbb{P}\left(D_{(N-k_\varepsilon+1)} > b_{\xi_{k_\varepsilon}} u_N\right) \leq \mathbb{P}\left(\xi_{k_\varepsilon} > b_{\xi_{k_\varepsilon}}/2\right) + \varepsilon/72 \leq \varepsilon/36,$$

for large enough N. Similarly, by definition of $a_{\eta,\varepsilon}$, in (2.6), we have

$$\mathbb{P}\left(L_N < a_{\eta,\varepsilon} u_N\right) \leq \varepsilon/36.$$

Thus, the term (A.1.4)(a) is at most $\varepsilon/12$.

The upper bound for (A.1.4)(b), i.e., the bound

$$\mathbb{P}\left(D_{(N-k_\varepsilon+1)} < a_{\xi_{k_\varepsilon}} u_N\right) < \varepsilon/12,$$

is an easy consequence of the distributional convergence of $D_{(N-k_\varepsilon+1)}/u_N$ to ξ_{k_ε} and the definition of $a_{\xi_{k_\varepsilon}}$. Similarly, we obtain the upper bound for the term in (A.1.4)(c), i.e.,

$$\mathbb{P}\left(L_N > b_{\eta,\varepsilon} u_N\right) < \varepsilon/12,$$

from the distributional convergence of L_N/u_N to η and the definition of $b_{\eta,\varepsilon}$.

Thus we have shown that $\mathbb{P}(E^c_{\varepsilon,N}) < \varepsilon/4$. This completes the proof of Lemma 2.2. □

A.2 Proof of Lemma 3.1

In this section we restate Lemma 3.1 and give a proof.

Lemma A.2.1 *For each $\varepsilon > 0$, there exists N_ε such that for all $N \geq N_\varepsilon$,*

$$\mathbb{P}(G^c_{\varepsilon,N}) < \varepsilon. \tag{A.2.1}$$

Proof. Clearly,

$$\mathbb{P}(G^c_{\varepsilon,N}) < 2\mathbb{P}\left(D_1^{(N)} > b_{D,\varepsilon}\right) + \mathbb{P}\left(L_N^{(N)} < L_\varepsilon(N) N^{1+\alpha(2-\tau)}\right). \tag{A.2.2}$$

From the definition of $b_{D,\varepsilon} \in \mathbb{N}$, and the string of inequalities,

$$\mathbb{P}(D_1^{(N)} > b_{D,\varepsilon}) \le \mathbb{P}(D_1 > b_{D,\varepsilon}) < \varepsilon/4,$$

we obtain that the first term on the right-hand side of (A.2.2) is at most $\varepsilon/2$. We show now that the second term on the right-hand side of (A.2.2) is at most $\varepsilon/2$. Since

$$D_i^{(N)} = \sum_{j<N^\alpha} \mathbf{1}\{D_i^{(N)} \ge j\},$$

we obtain, after interchanging the order of the two summations involved, that

$$L_N^{(N)} = \sum_{i=1}^N D_i^{(N)} = \sum_{j=1}^{\lfloor N^\alpha \rfloor -1} \sum_{i=1}^N \mathbf{1}\{D_i^{(N)} \ge j\}. \tag{A.2.3}$$

We would like to approximate

$$\sum_{i=1}^N \mathbf{1}\{D_i^{(N)} \ge j\} \qquad \text{by} \qquad N\mathbb{P}\left(D_1^{(N)} \ge j\right).$$

For this we will use ([5, Theorem 1.7(i), p. 14]) which states that for a binomial random variable X with parameters N and $p \le 1/2$, and a such that

$$a(1-p) \ge 12 \qquad \text{and} \qquad 0 < \frac{a}{Np} \le \frac{1}{12}, \tag{A.2.4}$$

we have

$$\mathbb{P}(|X - Np| \ge a) \le \sqrt{\frac{Np}{a^2}} e^{-\frac{a^2}{3Np}}. \tag{A.2.5}$$

Observe that $p_N(j) = \mathbb{P}\left(D_1^{(N)} \ge j\right)$ is non-increasing in $j \ge 1$ and that $\lim_{N\to\infty} p_N(N-1) = 0$. We will first complete the proof under an extra assumption on $\{p_N(j)\}_{j\ge1}$, and then prove that the assumption indeed holds. The extra assumption reads that for some $1 \le \kappa < k_N < N^\alpha$, all $j \in [\kappa, k_N]$ and N large enough, we have

$$\begin{array}{ll} (a) & p_N(j) < 1/2, \\ (b) & Np_N(j) > 288, \\ (c) & Np_N(j)/432 - \alpha \log N \to \infty, \text{ as } N \to \infty, \\ (d) & p_N(j) > \frac{1}{2}\mathbb{P}\left(D_1 \ge j\right). \end{array} \tag{A.2.6}$$

Then, with $a = Np_N(j)/12$, we have that

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^N \mathbf{1}\{D_i^{(N)} < j\} < Np_N(j)/2\right) &\le \mathbb{P}(|\sum_{i=1}^N \mathbf{1}\{D_i^{(N)} < j\} - Np_N(j))| > a/2)) \\ &\le 12(Np_N(j))^{-1/2} e^{-Np_N(j)/432}, \end{aligned}$$

because from (a) and (b) of (A.2.6), we have $a(1-p) = (Np_N(j)/12)(1-p_N(j)) \geq Np_N(j)/24 > 288/24 = 12$ and $a/Np = 1/12$. Hence, we find

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{j=\kappa}^{k_N}\{\sum_{i=1}^{N}\mathbf{1}\{D_i^{(N)} \geq j\} \geq Np_N(j)/2\}\right) &= 1 - \mathbb{P}\left(\bigcup_{j=\kappa}^{k_N}\{\sum_{i=1}^{N}\mathbf{1}\{D_i^{(N)} < j\} \geq Np_N(j)/2\}\right) \\
&\geq 1 - \sum_{j=\kappa}^{k_N}\mathbb{P}\left(\{\sum_{i=1}^{N}\mathbf{1}\{D_i^{(N)} < j\} \geq Np_N(j)/2\}\right) \\
&\geq 1 - \sum_{j=\kappa}^{k_N} 12(Np_N(j))^{-1/2}\exp(-Np_N(j)/432) \geq 1 - k_N\exp(-Np_N(k_N)/432) \\
&\geq 1 - \exp\{\alpha\log(N) - Np_N(k_N)/432\} \geq 1 - \varepsilon/2, \qquad \text{(A.2.7)}
\end{aligned}$$

for large N, due to (A.2.6(b)) and (A.2.6(c)). Given (A.2.7) and (A.2.6,(d)), we have

$$L_N^{(N)} \geq \sum_{j=\kappa}^{k_N}\sum_{i=1}^{N}\mathbf{1}\{D_i^{(N)} \geq j\} \geq \frac{N}{2}\sum_{j=\kappa}^{k_N} p_N(j) \geq \frac{N}{4}\sum_{j=\kappa}^{k_N}\mathbb{P}(D \geq j) = \frac{N}{4}\sum_{j=\kappa-1}^{k_N-1} j^{1-\tau}L(j), \qquad \text{(A.2.8)}$$

with probability at least $1-\varepsilon/2$.

We will call a function $f(N)$ *slow*, if for each $\delta > 0$, $f(N) > N^{-\delta}$, for large enough N. Observe that for any $a > 0$ and slow f

$$\sum_{j=1}^{k} f(j)j^a = f_1(k)k^{a+1},$$

for some slow function f_1. We further assume that we can take $k_N = L_1(N)N^\alpha$ for some slow function $L_1(N)$, then the r.h.s. of (A.2.8) is $L_\varepsilon(N)N^{1+\alpha(2-\tau)}$, for some slow function $L_\varepsilon(N)$. This would imply that the second term of (A.2.2) is at most $\varepsilon/2$ and thus completes the proof of Lemma A.2.1 subject to (A.2.6) and the fact that $k_N = L_1(N)N^\alpha$ for some slow function $L_1(N)$.

We now specify $1 \leq \kappa < N^\alpha$, check points (A.2.6(a))- (A.2.6(d)) and demonstrate that we can take $k_N = L_1(N)N^\alpha$, for some slow function $L_1(N)$. The only restriction on κ is (A.2.6(a)). Just take κ large enough such that $p_N(\kappa) < 1/2$. Then, since $j \mapsto p_N(j)$ is non-increasing, we obtain that $p_N(j) < 1/2$ for all $j \geq \kappa$. Therefore, (A.2.6(a)) is satisfied. This introduces κ and proves (A.2.6(a)).

We next define k_N. Choose

$$k_N = \max_k\{\mathbb{P}(D_1 > k) > 2(1 - F(\lfloor N^\alpha\rfloor - 1))\}.$$

This definition gives us point (d) of (A.2.6). Indeed for any $j \leq k_N$,

$$\begin{aligned}
p_N(j) = \mathbb{P}\left(D_1^{(N)} \geq j\right) &= \frac{\mathbb{P}\left(D_1 \geq j\right) - \mathbb{P}\left(D_1 \geq N^\alpha\right)}{1 - \mathbb{P}(D_1 \geq N^\alpha)} \\
&> \mathbb{P}\left(D_1 \geq j\right) - \mathbb{P}\left(D_1 \geq N^\alpha\right) \geq \frac{1}{2}\mathbb{P}\left(D_1 \geq j\right). \qquad \text{(A.2.9)}
\end{aligned}$$

Before we check the other items of (A.2.6), we prove that $k_N = L_1(N)N^\alpha$, for some slow function $L_1(N)$. We argue by contradiction. Suppose there exists $\delta > 0$ such that $k_N < N^{(1-\delta)\alpha}$ for all N sufficiently large. Then, by definition of k_N, we have

$$\mathbb{P}(D_1 > k_N + 1) \leq 2\mathbb{P}(D_1 > \lfloor N^\alpha\rfloor) = L(N)N^{\alpha(1-\tau)}.$$

However, when we compute probability on the l.h.s., then we obtain that

$$\mathbb{P}(D_1 > k_N + 1) \geq L_3(N)N^{-(\tau-1)(1-\delta)\alpha},$$

for some slow function $L_3(N)$. This gives a contradiction. Hence, for any $\delta > 0$, $k_N > N^{\alpha(1-\delta)}$, for large enough N. This is equivalent to the statement that $k_N = L_1(N)N^\alpha$ for some slow function $L_1(N)$.

Since $L_1(N)$ is slow, $(A.2.6(\mathrm{c}))$ follows from $(A.2.6(\mathrm{d}))$, since

$$\frac{N}{2}\mathbb{P}\left(D_1 \geq k_N\right) = \frac{N}{2}L(k_N - 1)(k_N - 1)^{1-\tau} = L_2(N)N^{1-\alpha(\tau-1)}, \tag{A.2.10}$$

for some slow function $L_2(N)$. Since $\alpha < 1/(\tau - 1)$, the right-hand side of (A.2.10) increases as a positive power of N, and hence exceeds $\alpha \log(N)$, eventually for sufficiently large N. Finally, (A.2.6(b)) follows from (A.2.6(c)). Thus, we have proved that κ, k_N defined above satisfy the assumptions in (A.2.6) and that $k_N = L_1(N)N^\alpha$ for some slow function $L_1(N)$. Therefore, the second term in (A.2.2) is at most $\varepsilon/2$, and we conclude that $\mathbb{P}(G^c_{\varepsilon,N}) < \varepsilon$. □