Electronic Journal of Probability

Vol. 12 (2007), Paper no. 25, pages 703–766.

Journal URL
http://www.math.washington.edu/~ejpecp/

Distances in random graphs with finite mean and infinite variance degrees

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

Abstract

In this paper we study typical distances in random graphs with i.i.d. degrees of which the tail of the common distribution function is regularly varying with exponent $1-\tau$. Depending on the value of the parameter τ we can distinct three 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. The distances between two randomly chosen nodes belonging to the same connected component, for $\tau > 3$ and $\tau \in (1,2)$, have been studied in previous publications, and we survey these results here. When $\tau \in (2,3)$, the graph distance centers around $2\log\log N/|\log(\tau-2)|$. We present a full proof of this result, and study the fluctuations around this asymptotic means, by describing the asymptotic distribution. The results presented here improve upon results of Reittu and Norros, who prove an upper bound only.
The random graphs studied here can serve as models for complex networks where degree power laws are observed; this is illustrated by comparing the typical distance in this model to Internet data, where a degree power law with exponent $\tau \approx 2.2$ is observed for the so-called Autonomous Systems (AS) graph .

Key words: Branching processes, configuration model, coupling, graph distance.

[*]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

[†]Supported in part by Netherlands Organization for Scientific Research (NWO).

[‡]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

AMS 2000 Subject Classification: Primary 05C80; Secondary: 05C12, 60J80.

Submitted to EJP on February 27, 2007, final version accepted April 10, 2007.

1 Introduction

Complex networks are encountered in a wide variety of disciplines. A rough classification has been given by Newman (18) and consists of: (i) Technological networks, e.g. electrical power grids and the Internet, (ii) Information networks, such as the World Wide Web, (iii) Social networks, like collaboration networks and (iv) Biological networks like neural networks and protein interaction networks.

What many of the above examples have in common is that the typical distance between two nodes in these networks are small, a phenomenon that is dubbed the *'small-world'* phenomenon. A second key phenomenon shared by many of those networks is their *'scale-free'* nature; meaning that these networks have so-called power-law degree sequences, i.e., the number of nodes with degree k falls of as an inverse power of k. We refer to (1; 18; 25) and the references therein for a further introduction to complex networks and many more examples where the above two properties hold.

A random graph model where both the above key features are present is the configuration model applied to an i.i.d. sequence of degrees with a power-law degree distribution. In this model we start by sampling the degree sequence from a power law and subsequently connect nodes with the sampled degree purely at random. This model automatically satisfies the power law degree sequence and it is therefore of interest to rigorously derive the typical distances that occur.

Together with two previous papers (10; 14), the current paper describes the random fluctuations of the graph distance between two arbitrary nodes in the configuration model, where the i.i.d. degrees follow a power law of the form

$$\mathbb{P}(D > k) = k^{-\tau+1}L(k),$$

where L denotes a slowly varying function and the exponent τ satisfies $\tau \geq 1$. To obtain a complete picture we include a discussion and a heuristic proof of the results in (10) for $\tau \in [1, 2)$, and those in (14) for $\tau > 3$. However, the main goal of this paper is the complete description, including a full proof of the case where $\tau \in (2, 3)$. Apart from the critical cases $\tau = 2$ and $\tau = 3$, which depend on the behavior of the slowly varying function L (see (10, Section 4.2) when $\tau = 2$), we have thus given a complete analysis for all possible values of $\tau \geq 1$.

This section is organized as follows. In Section 1.1, we start by introducing the model, in Section 1.2 we state our main results. Section 1.3 is devoted to related work, and in Section 1.4, we describe some simulations for a better understanding of our main results. Finally, Section 1.5 describes the organization of the paper.

1.1 Model definition

Fix an integer N. Consider an i.i.d. sequence $D_1, D_2, \ldots, D_N$. 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 increase D_N 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 or half-edges. All stubs need to be connected to build the graph. The stubs are numbered in any 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 (half-edges) 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 $N \to \infty$, as shown in (5).

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 vector of which the k^{th} coordinate equals the proportion of nodes with degree k. In our model, by the law of large numbers, the degree sequence is close to the probability mass function of the nodal degree D of which $D_1, \ldots, D_N$ are independent copies.

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

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

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

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

where $\tau > 1$ and L is slowly varying at infinity. This means that the random variables D_j 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*

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

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

$$1 - F(x) \leq c x^{-\tau+1}, \qquad \textit{for all } x \geq 1, \tag{1.4}$$

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

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

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

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

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

The parameter $m_{\tau,N}$ describes the asymptotic growth of H_N as $N \to \infty$. 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*

(i) *for $\tau \in (1,2)$,*

$$\lim_{N\to\infty} \mathbb{P}(H_N = 2) = 1 - \lim_{N\to\infty} \mathbb{P}(H_N = 3) = p, \tag{1.7}$$

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

(ii) *for $\tau \in (2,3)$ or $\tau > 3$ there exist random variables $(R_{\tau,a})_{a\in(-1,0]}$, such that as $N \to \infty$,*

$$\mathbb{P}\Big(H_N = m_{\tau,N} + l \;\Big|H_N < \infty\Big) = \mathbb{P}(R_{\tau,a_N} = l) + o(1), \tag{1.8}$$

where

$$a_N = \begin{cases} \lfloor \frac{\log\log N}{|\log(\tau-2)|} \rfloor - \frac{\log\log N}{|\log(\tau-2)|}, & \text{for } \tau \in (2,3), \\ \lfloor \log_\nu N \rfloor - \log_\nu N, & \text{for } \tau > 3. \end{cases}$$

We see that for $\tau \in (1,2)$, the limit distribution exists and concentrates on the two points 2 and 3. For $\tau \in (2,3)$ or $\tau > 3$ the limit behavior is more involved. In these cases the limit distribution does not exist, caused by the fact that the correct centering constants, $2\log\log N/(|\log(\tau-2)|)$, for $\tau \in (2,3)$ and $\log_\nu N$, for $\tau > 3$, are in general not integer, whereas H_N is with probability 1 concentrated on the integers. The above theorem claims that for $\tau \in (2,3)$ or $\tau > 3$ and large N, we have $H_N = m_{\tau,N} + O_p(1)$, with $m_{\tau,N}$ specified in (1.6) and where $O_p(1)$ is a random contribution, which is tight on $\mathbb{R}$. The specific form of this random contribution is specified in Theorem 1.5 below.

In Theorem 1.2, we condition on $H_N < \infty$. In the course of the proof, here and in (14), we also investigate the probability of this event, and prove that

$$\mathbb{P}(H_N < \infty) = q^2 + o(1), \tag{1.9}$$

where q is the survival probability of an appropriate branching process.

Corollary 1.3 (Convergence in distribution along subsequences). *For $\tau \in (2,3)$ or $\tau > 3$, and when Assumption 1.1 is fulfilled, we have that, for $k \to \infty$,*

$$H_{N_k} - m_{\tau,N_k} | H_{N_k} < \infty \tag{1.10}$$

converges in distribution to $R_{\tau,a}$, along subsequences N_k where a_{N_k} converges to a.

A simulation for $\tau \in (2,3)$ illustrating the weak convergence in Corollary 1.3 is discussed in Section 1.4.

Corollary 1.4 (Concentration of the hopcount)**.** *For $\tau \in (2,3)$ or $\tau > 3$, and when Assumption 1.1 is fulfilled, we have that the random variables $H_N - m_{\tau,N}$, given that $H_N < \infty$, form a tight sequence, i.e.,*

$$\lim_{K\to\infty} \limsup_{N\to\infty} \mathbb{P}\Big(\big|H_N - m_{\tau,N}\big| \le K \;\Big| H_N < \infty\Big) = 1. \tag{1.11}$$

We next describe the laws of the random variables $(R_{\tau,a})_{a\in(-1,0]}$. For this, we need some further notation from branching processes. For $\tau > 2$, we introduce a *delayed* branching process $\{\mathcal{Z}_k\}_{k\ge1}$, 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

$$g_j = \frac{(j+1)f_{j+1}}{\mu}, \quad j = 0,1,\ldots, \qquad \text{where} \qquad \mu = \mathbb{E}[D_1]. \tag{1.12}$$

When $\tau \in (2,3)$, the branching process $\{\mathcal{Z}_k\}$ has infinite expectation. Under Assumption 1.1, it is proved in (8) that

$$\lim_{n\to\infty} (\tau-2)^n \log(\mathcal{Z}_n \vee 1) = Y, \qquad \text{a.s.}, \tag{1.13}$$

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

When $\tau > 3$, the process $\{\mathcal{Z}_n/\mu\nu^{n-1}\}_{n\ge1}$ is a non-negative martingale and consequently

$$\lim_{n\to\infty} \frac{\mathcal{Z}_n}{\mu\nu^{n-1}} = \mathcal{W}, \quad \text{a.s.} \tag{1.14}$$

The constant q appearing in (1.9) is the survival probability of the branching process $\{\mathcal{Z}_k\}_{k\ge1}$. We can identify the limit laws of $(R_{\tau,a})_{a\in(-1,0]}$ in terms of the limit random variables in (1.13) and (1.14) as follows:

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

(i) *for $\tau \in (2,3)$ and for $a \in (-1,0]$,*

$$\mathbb{P}(R_{\tau,a} > l) = \mathbb{P}\Big(\min_{s\in\mathbb{Z}} \big[(\tau-2)^{-s}Y^{(1)} + (\tau-2)^{s-c_l}Y^{(2)}\big] \le (\tau-2)^{\lceil l/2\rceil + a}\big|Y^{(1)}Y^{(2)} > 0\Big), \tag{1.15}$$

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 (1.13).

(ii) *for $\tau > 3$ and for $a \in (-1,0]$,*

$$\mathbb{P}(R_{\tau,a} > l) = \mathbb{E}\big[\exp\{-\tilde{\kappa}\nu^{a+l}\mathcal{W}^{(1)}\mathcal{W}^{(2)}\}\big|\mathcal{W}^{(1)}\mathcal{W}^{(2)} > 0\big], \tag{1.16}$$

where $\mathcal{W}^{(1)}$ and $\mathcal{W}^{(2)}$ are two independent copies of the limit random variable $\mathcal{W}$ in (1.14) and where $\tilde{\kappa} = \mu(\nu-1)^{-1}$.

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 $\tau \ge 1$, except for the critical cases $\tau = 2$ and $\tau = 3$. The result for $\tau \in (1,2)$, and the case $\tau = 1$, where $H_N \xrightarrow{P} 2$, are both proved in (10), the result for $\tau > 3$ is proved in (14). In Section 2 we will present heuristic proofs for all three cases, and in Section 4 a full proof for the case where

$\tau \in (2,3)$. Theorems 1.2-1.5 quantify the small-world phenomenon for the configuration model, and explicitly divide the scaling of the graph distances into three distinct regimes

In Remarks 4.2 and A.1.5 below, we will explain that our results also apply to the usual configuration model, where the number of nodes with a given degree is deterministic, when we study the graph distance between two *uniformly* chosen nodes, and the degree distribution satisfied certain conditions. For the precise conditions, see Remark A.1.5 below.

1.3 Related work

There are many papers on scale-free graphs and we refer to reviews such as the ones by Albert and Barabási (1), Newman (18) and the recent book by Durrett (9) for an introduction; we refer to (2; 3; 17) for an introduction to classical random graphs.

Papers involving distances for the case where the degree distribution F (see (1.2)), has exponent $\tau \in (2,3)$ are not so wide spread. In this discussion we will focus on the case where $\tau \in (2,3)$. For related work on distances for the cases $\tau \in (1,2)$ and $\tau > 3$ we refer to (10, Section 1.4) and (14, Section 1.4), respectively.

The model investigated in this paper with $\tau \in (2,3)$ was first studied in (21), where it was shown that with probability converging to 1, H_N is less than $m_{\tau,N}(1+o(1))$. We improve the results in (21) by deriving the asymptotic distribution of the random fluctuations of the graph distance around $m_{\tau,N}$. Note that these results are in contrast to (19, Section II.F, below Equation (56)), where it was suggested that if $\tau < 3$, then an exponential cut-off is necessary to make the graph distance between an arbitrary pair of nodes well-defined. The problem of the mean graph distance between an arbitrary pair of nodes was also studied non-rigorously in (7), where also the behavior when $\tau = 3$ and $x \mapsto L(x)$ is the constant function, is included. In the latter case, the graph distance scales like $\frac{\log N}{\log\log N}$. A related model to the one studied here can be found in (20), where a Poissonian graph process is defined by adding and removing edges. In (20), the authors prove similar results as in (21) for this related model. For $\tau \in (2,3)$, in (15), it was further shown that the diameter of the configuration model is bounded below by a constant times $\log N$, when $f_1 + f_2 > 0$, and bounded above by a constant times $\log\log N$, when $f_1 + f_2 = 0$.

A second related model can be found in (6), 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' $w = (w_1, \ldots, w_N)$. It is assumed that $\max_i w_i^2 < \sum_i w_i$, so that $w_i w_j / \sum_l w_l$ are probabilities. In (6), w_i is often taken as $w_i = c i^{-\frac{1}{\tau-1}}$, where c is a function of N proportional to $N^{\frac{1}{\tau-1}}$. In this case, the degrees obey a power law with exponent τ. Chung and Lu (6) show that in this case, the graph distance between two uniformly chosen nodes is with probability converging to 1 proportional to $\log N(1+o(1))$ when $\tau > 3$, and to $2\frac{\log\log N}{|\log(\tau-2)|}(1+o(1))$ when $\tau \in (2,3)$. The difference between this model and ours is that the nodes are not exchangeable in (6), but the observed phenomena are similar. This result can be heuristically 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 nodes for which the degree is at least k equals

$$|\{i : w_i \geq k\}| = |\{i : c i^{-\frac{1}{\tau-1}} \geq k\}| \propto k^{-\tau+1}.$$

Thus, one expects that the number of nodes with degree at least k decreases as $k^{-\tau+1}$, similarly as in our model. The most general version of this model can be found in (4). All these models

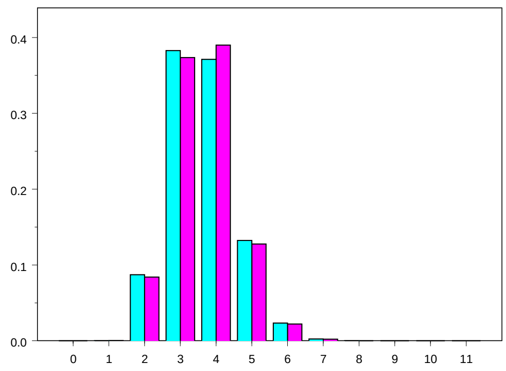

Figure 1: Histograms of the AS-count and graph distance in the configuration model with $N = 10,940$, where the degrees have generating function $f_\tau(s)$ in (1.18), for which the power law exponent τ takes the value $\tau = 2.25$. The AS-data is lightly shaded, the simulation is darkly shaded.

assume some form of (conditional) independence of the edges, which results in asymptotic degree sequences that are given by mixed Poisson distributions (see e.g. (5)). In the configuration model, instead, the *degrees* are independent.

1.4 Demonstration of Corollary 1.3

Our motivation to study the above version of the configuration model is to describe the topology of the Internet at a fixed time instant. In a seminal paper (12), 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. In (24), and inspired by the observed power law degree sequence in (12), the power law random 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 edges represent the physical connections between the different autonomous systems. The work of Faloutsos *et al.* in (12) was among others on this graph which at that time had size approximately 10,000. In (24), 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 graph distance in the configuration model. The AS-count gives the number of physical links connecting the various autonomous domains between two randomly chosen domains. To validate the model studied here, we compare a simulation of the distribution of the distance between pairs of nodes in the configuration model with the same value of N and τ to extensive measurements of the AS-count in Internet. In Figure 1, we see that the graph distance in the model with the predicted value of $\tau = 2.25$ and the value of N from the data set fits the AS-count data remarkably well.

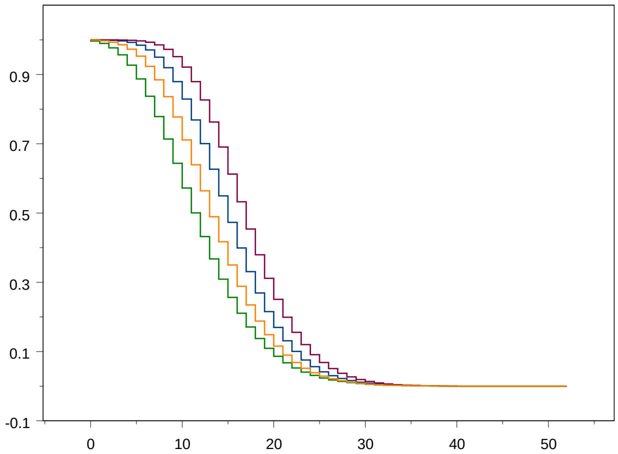

Figure 2: Empirical survival functions of the graph distance for $\tau = 2.8$ and for the four values of N.

Having motivated why we are interested to study distances in the configuration model, we now explain by a simulation the relevance of Theorem 1.2 and Corollary 1.3 for $\tau \in (2,3)$. We have chosen to simulate the distribution (1.12) using the generating function:

$$g_\tau(s) = 1 - (1-s)^{\tau-2}, \qquad \text{for which} \qquad g_j = (-1)^{j-1}\binom{\tau-2}{j} \sim \frac{c}{j^{\tau-1}}, \quad j \to \infty. \tag{1.17}$$

Defining

$$f_\tau(s) = \frac{\tau-1}{\tau-2}s - \frac{1-(1-s)^{\tau-1}}{\tau-2}, \quad \tau \in (2,3), \tag{1.18}$$

it is immediate that

$$g_\tau(s) = \frac{f'_\tau(s)}{f'_\tau(1)}, \qquad \text{so that} \qquad g_j = \frac{(j+1)f_{j+1}}{\mu}.$$

For fixed τ, we can pick different values of the size of the simulated graph, so that for each two simulated values N and M we have $a_N = a_M$, i.e., $N = \lceil M^{(\tau-2)^{-k}} \rceil$, for some integer k. For $\tau = 2.8$, this induces, starting from $M = 1000$, by taking for k the successive values $1, 2, 3$,

$$M = 1,000, \qquad N_1 = 5,624, \qquad N_2 = 48,697, \qquad N_3 = 723,395.$$

According to Corollary 1.3, the survival functions of the hopcount H_N, given by $k \mapsto \mathbb{P}(H_N > k | H_N < \infty)$, and for $N = \lceil M^{(\tau-2)^{-k}} \rceil$, run approximately parallel on distance 2 in the limit for $N \to \infty$, since $m_{\tau,N_k} = m_{\tau,M} + 2k$ for $k = 1, 2, 3$. In Section 3.1 below we will show that the distribution with generating function (1.18) satisfies Assumption 1.1(ii).

1.5 Organization of the paper

The paper is organized as follows. In Section 2 we heuristically explain our results for the three different cases. The relevant literature on branching processes with infinite mean. is reviewed in Section 3, where we also describe the growth of shortest path graphs, and state coupling results needed to prove our main results, Theorems 1.2–1.5 in Section 4. In Section 5, we prove three technical lemmas used in Section 4. We finally prove the coupling results in the Appendix. In the sequel we will write that event $\mathcal{E}$ occurs **whp** for the statement that $\mathbb{P}(\mathcal{E}) = 1 - o(1)$, as the total number of nodes $N \to \infty$.

2 Heuristic explanations of Theorems 1.2 and 1.5

In this section, we present a heuristic explanation of Theorems 1.2 and 1.5.

When $\tau \in (1, 2)$, the total degree L_N is the i.i.d. sum of N random variables $D_1, D_2, \ldots, D_N$, 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 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. For $\tau = 1$ the quotient M_N/L_N, where M_N denotes the maximum of $D_1, D_2, \ldots, D_N$, converges to 1 in probability, and consequently the asymptotic distance is 2 in this case, as basically all nodes are attached to the *unique* giant node. As mentioned before, full proofs of these results can be found in (10).

For $\tau \in (2, 3)$ or $\tau > 3$ there are two basic ingredients underlying the graph distance 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

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

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 $n = o(L)$, $L \to \infty$, however, these two probabilities are asymptotically equal. We approximate (2.1) further as

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

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

The *shortest path graph* (SPG) from node 1 is the union of all shortest paths between node 1 and all other nodes $\{2, \ldots, N\}$. We define the SPG from node 2 in a similar fashion. We apply the above heuristic asymptotics to the growth of the SPG's. Let $Z_j^{(1,N)}$ denote the number of stubs that are attached to nodes precisely $j-1$ steps away from node 1, and similarly for $Z_j^{(2,N)}$. We then apply (2.2) to $n = Z_j^{(1,N)}, m = Z_j^{(2,N)}$ and $L = L_N$. Let $\mathbb{Q}_Z^{(k,l)}$ be the conditional distribution given $\{Z_s^{(1,N)}\}_{s=1}^k$ and $\{Z_s^{(2,N)}\}_{s=1}^l$. For $l = 0$, we only condition on $\{Z_s^{(1,N)}\}_{s=1}^k$. For $j \geq 1$, we have the multiplication rule (see (14, Lemma 4.1)),

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

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 from (2.1) and (2.2) we find,

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

This asymptotic identity follows because the event $\{H_N > i-1 | H_N > i-2\}$ occurs precisely when none of the stubs $Z_{\lceil i/2\rceil}^{(1,N)}$ attaches to one of those of $Z_{\lfloor i/2\rfloor}^{(2,N)}$. Consequently we can approximate

$$\mathbb{P}(H_N > j) \approx \mathbb{E}\left[\exp\left\{-\frac{1}{L_N}\sum_{i=2}^{j+1} Z_{\lceil i/2\rceil}^{(1,N)} Z_{\lfloor i/2\rfloor}^{(2,N)}\right\}\right]. \tag{2.5}$$

A typical value of the hopcount H_N is the value j for which

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

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 (these stubs are called 'brother' stubs of the chosen stub). The conditional probability that this number of 'brother' stubs equals n given $D_1, \ldots, D_N$, is approximately equal to the probability that a random stub from all $L_N = D_1 + \ldots + D_N$ stubs is attached to a node with in total $n+1$ stubs. Since there are precisely $\sum_{j=1}^N (n+1)\mathbf{1}_{\{D_j = n+1\}}$ stubs that belong to a node with degree $n+1$, we find for the latter probability

$$g_n^{(N)} = \frac{n+1}{L_N}\sum_{j=1}^N \mathbf{1}_{\{D_j = n+1\}}, \tag{2.6}$$

where $\mathbf{1}_A$ 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 SPG's from nodes 1 and 2, of course the number of stubs that can still be chosen decreases. However, when the size of both SPG's 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 SPG's is closely related to a *branching process* with offspring distribution $\{g_n^{(N)}\}_{n=1}^\infty$.

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

$$\frac{L_N}{N} \to \mu = \mathbb{E}[D_1], \quad \text{and} \quad \frac{1}{N}\sum_{j=1}^{N} \mathbf{1}_{\{D_j=n+1\}} \to f_{n+1} = \mathbb{P}(D_1 = n+1),$$

so that, almost surely,

$$g_n^{(N)} \to \frac{(n+1)f_{n+1}}{\mu} = g_n, \quad N \to \infty. \tag{2.7}$$

Therefore, the growth of the shortest path graph 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

$$\sum_{i=2}^{j+1} \mathcal{Z}^{(1)}_{\lceil i/2\rceil}\mathcal{Z}^{(2)}_{\lfloor i/2\rfloor} = L_N \approx \mu N, \tag{2.8}$$

where $\{\mathcal{Z}^{(1)}_j\}$ and $\{\mathcal{Z}^{(2)}_j\}$ denote two independent copies of a delayed branching process with offspring distribution $\{f_n\}$, $f_n = \mathbb{P}(D = n)$, $n = 1, 2, \ldots$, 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 $\nu = \sum_{n\geq 1} ng_n < \infty$. Assume also that $\nu > 1$, so that the branching process is supercritical. In this case, the branching process $\mathcal{Z}_j/\mu\nu^{j-1}$ converges almost surely to a random variable $\mathcal{W}$ (see (1.14)). Hence, for the two *independent* branching processes $\{\mathcal{Z}^{(i)}_j\}$, $i = 1, 2$, that locally describe the number of stubs attached to nodes on distance $j-1$, we find that, for $j \to \infty$,

$$\mathcal{Z}^{(i)}_j \sim \mu\nu^{j-1}\mathcal{W}^{(i)}. \tag{2.9}$$

This explains why the average value of $Z^{(i,N)}_j$ grows like $\mu\nu^{j-1} = \mu\exp((j-1)\log\nu)$, that is, exponential in j for $\nu > 1$, so that a typical value of j for which (2.8) 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 (2.9) describes the fluctuations of $\mathcal{Z}^{(i)}_j$ around the mean value $\mu\nu^{j-1}$, we are able to describe the random fluctuations of H_N around $\log_\nu N$. The details of these proofs can be found in (14).

When $\tau \in (2, 3)$, the branching processes $\{\mathcal{Z}^{(1)}_j\}$ and $\{\mathcal{Z}^{(2)}_j\}$ are well-defined, but they have infinite mean. Under certain conditions on the underlying offspring distribution, which are implied by Assumption 1.1(ii), Davies (8) proves for this case that $(\tau-2)^j\log(\mathcal{Z}_j+1)$ converges almost surely, as $j \to \infty$, 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 $L_N \approx \mu N$. Furthermore by the double exponential behavior of $\mathcal{Z}_i$, the size of the left-hand side of (2.8) is equal to the size of the last term, so that the typical value of j for which (2.8) holds satisfies

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

This indicates that the typical value of j is of order

$$2\frac{\log\log N}{|\log(\tau-2)|}, \tag{2.10}$$

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

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

then $(j+1)/2 = \log(c \log N)/|\log(\tau-2)|$, which induces the leading order of $m_{\tau,N}$ defined in (1.6). Again we stress that, since Davies' result (8) describes a distributional limit, we are able to describe the random fluctuations of H_N around $m_{\tau,N}$. The details of the proof are given in Section 4.

3 The growth of the shortest path graph

In this section we describe the growth of the shortest path graph (SPG). This growth relies heavily on branching processes (BP's). We therefore start in Section 3.1 with a short review of the theory of BP's in the case where the expected value (mean) of the offspring distribution is infinite. In Section 3.2, we discuss the coupling between these BP's and the SPG, and in Section 3.3, we give the bounds on the coupling. Throughout the remaining sections of the sequel we will assume that $\tau \in (2,3)$, and that F satisfies Assumption 1.1(ii).

3.1 Review of branching processes with infinite mean

In this review of BP's with infinite mean we follow in particular (8), and also refer the readers to related work in (22; 23), and the references therein.

For the formal definition of the BP we define a double sequence $\{X_{n,i}\}_{n\geq 0, i\geq 1}$ of i.i.d. random variables each with distribution equal to the offspring distribution $\{g_j\}$ given in (1.12) with distribution function $G(x) = \sum_{j=0}^{\lfloor x \rfloor} g_j$. The BP $\{\mathcal{Z}_n\}$ is now defined by $\mathcal{Z}_0 = 1$ and

$$\mathcal{Z}_{n+1} = \sum_{i=1}^{\mathcal{Z}_n} X_{n,i}, \quad n \geq 0.$$

In case of a delayed BP, we let $X_{0,1}$ have probability mass function $\{f_j\}$, independently of $\{X_{n,i}\}_{n\geq 1}$. In this section we restrict to the non-delayed case for simplicity.

We follow Davies in (8), who gives the following sufficient conditions for convergence of $(\tau-2)^n \log(1+\mathcal{Z}_n)$. Davies' main theorem states that if there exists a non-negative, non-increasing function $\gamma(x)$, such that,

(i) $x^{-\zeta-\gamma(x)} \leq 1 - G(x) \leq x^{-\zeta+\gamma(x)}$, for large x and $0 < \zeta < 1$,

(ii) $x^{\gamma(x)}$ is non-decreasing,

(iii) $\int_0^\infty \gamma(e^{e^x})\, dx < \infty$, or, equivalently, $\int_e^\infty \frac{\gamma(y)}{y \log y}\, dy < \infty$,

then $\zeta^n \log(1+\mathcal{Z}_n)$ converges almost surely to a non-degenerate finite random variable Y with $\mathbb{P}(Y=0)$ equal to the extinction probability of $\{\mathcal{Z}_n\}$, whereas $Y|Y>0$ admits a density on $(0,\infty)$. Therefore, also $\zeta^n \log(\mathcal{Z}_n \vee 1)$ converges to Y almost surely.

The conditions of Davies quoted as (i-iii) simplify earlier work by Seneta (23). For example, for $\{g_j\}$ in (1.17), the above is valid with $\zeta = \tau - 2$ and $\gamma(x) = C(\log x)^{-1}$, where C is sufficiently large. We prove in Lemma A.1.1 below that for F as in Assumption 1.1(ii), and G the distribution function of $\{g_j\}$ in (1.12), the conditions (i-iii) are satisfied with $\zeta = \tau - 2$ and $\gamma(x) = C(\log x)^{\gamma-1}$, with $\gamma < 1$.

Let $Y^{(1)}$ and $Y^{(2)}$ be two independent copies of the limit random variable Y. In the course of the proof of Theorem 1.2, for $\tau \in (2, 3)$, we will encounter the random variable $U = \min_{t\in\mathbb{Z}}(\kappa^t Y^{(1)} + \kappa^{c-t} Y^{(2)})$, for some $c \in \{0, 1\}$, and where $\kappa = (\tau - 2)^{-1}$. The proof relies on the fact that, conditionally on $Y^{(1)}Y^{(2)} > 0$, U has a density. The proof of this fact is as follows. The function $(y_1, y_2) \mapsto \min_{t\in\mathbb{Z}}(\kappa^t y_1 + \kappa^{c-t} y_2)$ is discontinuous precisely in the points (y_1, y_2) satisfying $\sqrt{y_2/y_1} = \kappa^{n-\frac{1}{2}c}$, $n \in \mathbb{Z}$, and, conditionally on $Y^{(1)}Y^{(2)} > 0$, the random variables $Y^{(1)}$ and $Y^{(2)}$ are independent continuous random variables. Therefore, conditionally on $Y^{(1)}Y^{(2)} > 0$, the random variable $U = \min_{t\in\mathbb{Z}}(\kappa^t Y^{(1)} + \kappa^{c-t} Y^{(2)})$ has a density.

3.2 Coupling of SPG to BP's

In Section 2, it has been shown informally that the growth of the SPG is closely related to a BP $\{\hat{Z}_k^{(1,N)}\}$ with the *random* offspring distribution $\{g_j^{(N)}\}$ given by (2.6); note that in the notation $\hat{Z}_k^{(1,N)}$ we do include its dependence on N, whereas in (14, Section 3.1) this dependence on N was left out for notational convenience. The presentation in Section 3.2 is virtually identical to the one in (14, Section 3). However, we have decided to include most of this material to keep the paper self-contained.

By the strong law of large numbers,

$$g_j^{(N)} \to (j+1)\mathbb{P}(D_1 = j+1)/\mathbb{E}[D_1] = g_j, \qquad N \to \infty.$$

Therefore, the BP $\{\hat{Z}_k^{(1,N)}\}$, with offspring distribution $\{g_j^{(N)}\}$, is expected to be close to the BP $\{\mathcal{Z}_k^{(1)}\}$ with offspring distribution $\{g_j\}$ given in (1.12). So, in fact, the coupling that we make is two-fold. We first couple the SPG to the $N-$dependent branching process $\{\hat{Z}_k^{(1,N)}\}$, and consecutively we couple $\{\hat{Z}_k^{(1,N)}\}$ to the BP $\{\mathcal{Z}_k^{(1)}\}$. In Section 3.3, we state bounds on these couplings, which allow us to prove Theorems 1.2 and 1.5 of Section 1.2.

The shortest path graph (SPG) from node 1 consists of the shortest paths between node 1 and all other nodes $\{2, \ldots, N\}$. As will be shown below, the SPG is not necessarily a tree because cycles may occur. Recall that two stubs together form an edge. We define $Z_1^{(1,N)} = D_1$ and, for $k \geq 2$, we denote by $Z_k^{(1,N)}$ the number of stubs attached to nodes at distance $k-1$ from node 1, but are not part of an edge connected to a node at distance $k-2$. We refer to such stubs as 'free stubs', since they have not yet been assigned to a second stub to from an edge. Thus, $Z_k^{(1,N)}$ is the number of outgoing stubs from nodes at distance $k-1$ from node 1. By SPG$_{k-1}$ we denote the SPG up to level $k-1$, i.e., up to the moment we have $Z_k^{(1,N)}$ free stubs attached to nodes on distance $k-1$, and no stubs to nodes on distance k. Since we compare $Z_k^{(1,N)}$ to the k^{th} generation of the BP $\hat{Z}_k^{(1,N)}$, we call $Z_k^{(1,N)}$ the stubs of level k.

For the complete description of the SPG $\{Z_k^{(1,N)}\}$, we have introduced the concept of labels in (14, Section 3). These labels illustrate the resemblances and the differences between the SPG $\{Z_k^{(1,N)}\}$ and the BP $\{\hat{Z}_k^{(1,N)}\}$.

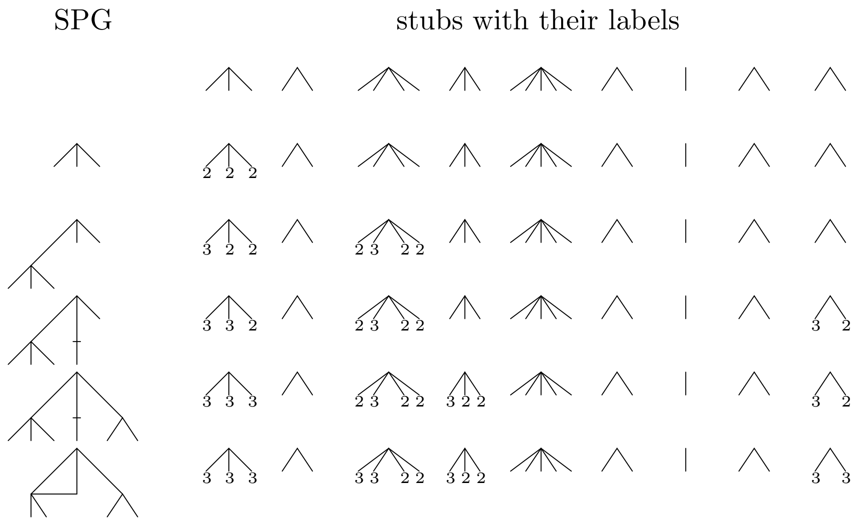

Figure 3: Schematic drawing of the growth of the SPG from the node 1 with $N = 9$ and the updating of the labels. The stubs without a label are understood to have label 1. The first line shows the N different nodes with their attached stubs. Initially, all stubs have label 1. The growth process starts by choosing the first stub of node 1 whose stubs are labelled by 2 as illustrated in the second line, while all the other stubs maintain the label 1. Next, we uniformly choose a stub with label 1 or 2. In the example in line 3, this is the second stub from node 3, whose stubs are labelled by 2 and the second stub by label 3. The left hand side column visualizes the growth of the SPG by the attachment of stub 2 of node 3 to the first stub of node 1. Once an edge is established the paired stubs are labelled 3. In the next step, again a stub is chosen uniformly out of those with label 1 or 2. In the example in line 4, it is the first stub of the last node that will be attached to the second stub of node 1, the next in sequence to be paired. The last line exhibits the result of creating a cycle when the first stub of node 3 is chosen to be attached to the second stub of node 9 (the last node). This process is continued until there are no more stubs with labels 1 or 2. In this example, we have $Z_1^{(1,N)} = 3$ and $Z_2^{(1,N)} = 6$.

Initially, all stubs are labelled 1. At each stage of the growth of the SPG, we draw uniformly at random from all stubs with labels 1 and 2. After each draw we will update the realization of the SPG according to three categories, which will be labelled 1, 2 and 3. At any stage of the generation of the SPG, the labels have the following meaning:

1. Stubs with label 1 are stubs belonging to a node that is not yet attached to the SPG.
2. Stubs with label 2 are attached to the SPG (because the corresponding node has been chosen), but not yet paired with another stub. These are the 'free stubs' mentioned above.
3. Stubs with label 3 in the SPG are paired with another stub to form an edge in the SPG.

The growth process as depicted in Figure 3 starts by labelling all stubs by 1. Then, because we construct the SPG starting from node 1 we relabel the D_1 stubs of node 1 with the label 2. We note that $Z_1^{(1,N)}$ is equal to the number of stubs connected to node 1, and thus $Z_1^{(1,N)} = D_1$. We next identify $Z_j^{(1,N)}$ for $j > 1$. $Z_j^{(1,N)}$ is obtained by sequentially growing the SPG from the free stubs in generation $Z_{j-1}^{(1,N)}$. When all free stubs in generation $j-1$ have chosen their connecting stub, $Z_j^{(1,N)}$ is equal to the number of stubs labelled 2 (i.e., free stubs) attached to the SPG. Note that not necessarily each stub of $Z_{j-1}^{(1,N)}$ contributes to stubs of $Z_j^{(1,N)}$, because a cycle may 'swallow' two free stubs. This is the case when a stub with label 2 is chosen.

After the choice of each stub, we update the labels as follows:

1. If the chosen stub has label 1, we connect the present stub to the chosen stub to form an edge and attach the brother stubs of the chosen stub as children. We update the labels as follows. The present and chosen stub melt together to form an edge and both are assigned label 3. All brother stubs receive label 2.

2. When we choose a stub with label 2, which is already connected to the SPG, a self-loop is created if the chosen stub and present stub are brother stubs. If they are not brother stubs, then a cycle is formed. Neither a self-loop nor a cycle changes the distances to the root in the SPG. The updating of the labels solely consists of changing the label of the present and the chosen stubs from 2 to 3.

The above process stops in the j^{th} generation when there are no more free stubs in generation $j-1$ for the SPG, and then $Z_j^{(1,N)}$ is the number of free stubs at this time. We continue the above process of drawing stubs until there are no more stubs having label 1 or 2, so that all stubs have label 3. Then, the SPG from node 1 is finalized, and we have generated the shortest path graph as seen from node 1. We have thus obtained the structure of the shortest path graph, and know how many nodes there are at a given distance from node 1.

The above construction will be performed identically from node 2, and we denote the number of free stubs in the SPG of node 2 in generation k by $Z_k^{(2,N)}$. This construction is close to being independent, when the generation size is not too large. In particular, it is possible to couple the two SPG growth processes with two independent BP's. This is described in detail in (14, Section 3). We make essential use of the coupling between the SPG's and the BP's, in particular, of (14, Proposition A.3.1) in the appendix. This completes the construction of the SPG's from both node 1 and 2.

3.3 Bounds on the coupling

We now investigate the relationship between the SPG $\{Z_k^{(i,N)}\}$ and the BP $\{\mathcal{Z}_k^{(i)}\}$ with law g. These results are stated in Proposition 3.1, 3.2 and 3.4. In their statement, we write, for $i = 1, 2$,

$$Y_k^{(i,N)} = (\tau-2)^k \log(Z_k^{(i,N)} \vee 1) \qquad \text{and} \qquad Y_k^{(i)} = (\tau-2)^k \log(\mathcal{Z}_k^{(i)} \vee 1), \tag{3.1}$$

where $\{\mathcal{Z}_k^{(1)}\}_{k\geq 1}$ and $\{\mathcal{Z}_k^{(2)}\}_{k\geq 1}$ are two independent delayed BP's with offspring distribution $\{g_j\}$ and where $\mathcal{Z}_1^{(i)}$ has law $\{f_j\}$. Then the following proposition shows that the first levels of the SPG are close to those of the BP:

Proposition 3.1 (Coupling at fixed time). *If F satisfies Assumption 1.1(ii), then for every m fixed, and for $i = 1, 2$, there exist* independent *delayed BP's $\mathcal{Z}^{(1)}, \mathcal{Z}^{(2)}$, such that*

$$\lim_{N\to\infty} \mathbb{P}(Y_m^{(i,N)} = Y_m^{(i)}) = 1. \tag{3.2}$$

In words, Proposition 3.1 states that at any *fixed* time, the SPG's from 1 and 2 can be coupled to two independent BP's with offspring g, in such a way that the probability that the SPG differs from the BP vanishes when $N \to \infty$.

In the statement of the next proposition, we write, for $i = 1, 2$,

$$\begin{aligned} \mathcal{T}_m^{(i,N)} = \mathcal{T}_m^{(i,N)}(\varepsilon) &= \{k > m : \big(Z_m^{(i,N)}\big)^{\kappa^{k-m}} \leq N^{\frac{1-\varepsilon^2}{\tau-1}}\} \\ &= \{k > m : \kappa^k Y_m^{(i,N)} \leq \frac{1-\varepsilon^2}{\tau-1}\log N\}, \end{aligned} \tag{3.3}$$

where we recall that $\kappa = (\tau-2)^{-1}$. We will see that $Z_k^{(i,N)}$ grows super-exponentially with k as long as $k \in \mathcal{T}_m^{(i,N)}$. More precisely, $Z_k^{(i,N)}$ is close to $\big(Z_m^{(i,N)}\big)^{\kappa^{k-m}}$, and thus, $\mathcal{T}_m^{(i,N)}$ can be thought of as the generations for which the generation size is bounded by $N^{\frac{1-\varepsilon^2}{\tau-1}}$. The second main result of the coupling is the following proposition:

Proposition 3.2 (Super-exponential growth with base $Y_m^{(i,N)}$ for large times). *If F satisfies Assumption 1.1(ii), then, for $i = 1, 2$,*

$$(a) \qquad \mathbb{P}\Big(\varepsilon \leq Y_m^{(i,N)} \leq \varepsilon^{-1}, \max_{k\in\mathcal{T}_m^{(i,N)}(\varepsilon)} |Y_k^{(i,N)} - Y_m^{(i,N)}| > \varepsilon^3\Big) = o_{N,m,\varepsilon}(1), \tag{3.4}$$

$$(b) \qquad \mathbb{P}\Big(\varepsilon \leq Y_m^{(i,N)} \leq \varepsilon^{-1}, \exists k \in \mathcal{T}_m^{(i,N)}(\varepsilon) : \, Z_{k-1}^{(i,N)} > Z_k^{(i,N)}\Big) = o_{N,m,\varepsilon}(1), \tag{3.5}$$

$$\mathbb{P}\Big(\varepsilon \leq Y_m^{(i,N)} \leq \varepsilon^{-1}, \exists k \in \mathcal{T}_m^{(i,N)}(\varepsilon) : \, Z_k^{(i,N)} > N^{\frac{1-\varepsilon^4}{\tau-1}}\Big) = o_{N,m,\varepsilon}(1), \tag{3.6}$$

where $o_{N,m,\varepsilon}(1)$ denotes a quantity $\gamma_{N,m,\varepsilon}$ that converges to zero when first $N \to \infty$, then $m \to \infty$ and finally $\varepsilon \downarrow 0$.

Remark 3.3. *Throughout the paper limits will be taken in the above order, i.e., first we send $N \to \infty$, then $m \to \infty$ and finally $\varepsilon \downarrow 0$.*

Proposition 3.2 (a), i.e. (3.4), is the main coupling result used in this paper, and says that as long as $k \in \mathcal{T}_m^{(i,N)}(\varepsilon)$, we have that $Y_k^{(i,N)}$ is close to $Y_m^{(i,N)}$, which, in turn, by Proposition 3.1, is close to $Y_m^{(i)}$. This establishes the coupling between the SPG and the BP. Part (b) is a technical result used in the proof. Equation (3.5) is a convenient result, as it shows that, with high probability, $k \mapsto Z_k^{(i,N)}$ is monotonically increasing. Equation (3.6) shows that with high probability $Z_k^{(i,N)} \leq N^{\frac{1-\varepsilon^4}{\tau-1}}$ for all $k \in \mathcal{T}_m^{(i,N)}(\varepsilon)$, which allows us to bound the number of free stubs in generation sizes that are in $\mathcal{T}_m^{(i,N)}(\varepsilon)$.

We complete this section with a final coupling result, which shows that for the first k which is *not* in $\mathcal{T}_m^{(i,N)}(\varepsilon)$, the SPG has many free stubs:

Proposition 3.4 (Lower bound on $Z_{k+1}^{(i,N)}$ for $k+1 \notin \mathcal{T}_m^{(i,N)}(\varepsilon)$)**.** *Let F satisfy Assumption 1.1(ii). Then,*

$$\mathbb{P}\Big(k \in \mathcal{T}_m^{(i,N)}(\varepsilon), k+1 \notin \mathcal{T}_m^{(i,N)}(\varepsilon), \varepsilon \leq Y_m^{(i,N)} \leq \varepsilon^{-1}, Z_{k+1}^{(i,N)} \leq N^{\frac{1-\varepsilon}{\tau-1}}\Big) = o_{N,m,\varepsilon}(1). \tag{3.7}$$

Propositions 3.1, 3.2 and 3.4 will be proved in the appendix. In Section 4 and 5, we will prove the main results in Theorems 1.2 and 1.5 subject to Propositions 3.1, 3.2 and 3.4.

4 Proof of Theorems 1.2 and 1.5 for $\tau \in (2,3)$

For convenience we combine Theorem 1.2 and Theorem 1.5, in the case that $\tau \in (2,3)$, in a single theorem that we will prove in this section.

Theorem 4.1. *Fix $\tau \in (2,3)$. When Assumption 1.1(ii) holds, then there exist random variables $(R_{\tau,a})_{a\in(-1,0]}$, such that as $N \to \infty$,*

$$\mathbb{P}\Big(H_N = 2\lfloor \frac{\log\log N}{|\log(\tau-2)|}\rfloor + l \;\Big|H_N < \infty\Big) = \mathbb{P}(R_{\tau,a_N} = l) + o(1), \tag{4.1}$$

where $a_N = \lfloor \frac{\log\log N}{|\log(\tau-2)|}\rfloor - \frac{\log\log N}{|\log(\tau-2)|} \in (-1,0]$. The distribution of $(R_{\tau,a})$, for $a \in (-1,0]$, is given by

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

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 (1.13).

4.1 Outline of the proof

We start with an outline of the proof. The proof is divided into several key steps proved in 5 subsections, Sections 4.2 - 4.6.

In the first key step of the proof, in Section 4.2, we split the probability $\mathbb{P}(H_N > k)$ into separate parts depending on the values of $Y_m^{(i,N)} = (\tau-2)^m \log(Z_m^{(i,N)} \vee 1)$. We prove that

$$\mathbb{P}(H_N > k, Y_m^{(1,N)}Y_m^{(2,N)} = 0) = 1 - q_m^2 + o(1), \qquad N \to \infty, \tag{4.2}$$

where $1-q_m$ is the probability that the delayed BP $\{\mathcal{Z}_j^{(1)}\}_{j\geq1}$ dies at or before the m^{th} generation. When m becomes large, then $q_m \uparrow q$, where q equals the survival probability of $\{\mathcal{Z}_j^{(1)}\}_{j\geq1}$. This leaves us to determine the contribution to $\mathbb{P}(H_N > k)$ for the cases where $Y_m^{(1,N)}Y_m^{(2,N)} > 0$. We further show that for m large enough, and on the event that $Y_m^{(i,N)} > 0$, **whp**, $Y_m^{(i,N)} \in [\varepsilon, \varepsilon^{-1}]$, for $i = 1,2$, where $\varepsilon > 0$ is small. We denote the event where $Y_m^{(i,N)} \in [\varepsilon, \varepsilon^{-1}]$, for $i = 1,2$, by $E_{m,N}(\varepsilon)$, and the event where $\max_{k\in\mathcal{T}_m^{(N)}(\varepsilon)} |Y_k^{(i,N)} - Y_m^{(i,N)}| \leq \varepsilon^3$ for $i = 1,2$ by $F_{m,N}(\varepsilon)$. The events $E_{m,N}(\varepsilon)$ and $F_{m,N}(\varepsilon)$ are shown to occur **whp**, for $F_{m,N}(\varepsilon)$ this follows from Proposition 3.2(a).

The second key step in the proof, in Section 4.3, is to obtain an asymptotic formula for $\mathbb{P}(\{H_N > k\} \cap E_{m,N}(\varepsilon))$. Indeed we prove that for $k \geq 2m-1$, and any k_1 with $m \leq k_1 \leq (k-1)/2$,

$$\mathbb{P}(\{H_N > k\} \cap E_{m,N}(\varepsilon)) = \mathbb{E}\Big[\mathbf{1}_{E_{m,N}(\varepsilon)\cap F_{m,N}(\varepsilon)} P_m(k, k_1)\Big] + o_{N,m,\varepsilon}(1), \tag{4.3}$$

where $P_m(k, k_1)$ is a product of conditional probabilities of events of the form $\{H_N > j | H_N > j-1\}$. Basically this follows from the multiplication rule. The identity (4.3) is established in (4.32).

In the third key step, in Section 4.4, we show that, for $k = k_N \to \infty$, the main contribution of the product $P_m(k, k_1)$ appearing on the right side of (4.3) is

$$\exp\Big\{ -\lambda_N \min_{k_1 \in \mathcal{B}_N} \frac{Z^{(1,N)}_{k_1+1} Z^{(2,N)}_{k_N-k_1}}{L_N} \Big\}, \tag{4.4}$$

where $\lambda_N = \lambda_N(k_N)$ is in between $\frac{1}{2}$ and $4k_N$, and where $\mathcal{B}_N = \mathcal{B}_N(\varepsilon, k_N)$ defined in (4.51) is such that $k_1 \in \mathcal{B}_N(\varepsilon, k_N)$ precisely when $k_1 + 1 \in \mathcal{T}^{(1,N)}_m(\varepsilon)$ and $k_N - k_1 \in \mathcal{T}^{(2,N)}_m(\varepsilon)$. Thus, by Proposition 3.2, it implies that **whp**

$$Z^{(1,N)}_{k_1+1} \leq N^{\frac{1-\varepsilon^4}{\tau-1}} \qquad \text{and} \qquad Z^{(2,N)}_{k_N-k_1} \leq N^{\frac{1-\varepsilon^4}{\tau-1}}.$$

In turn, these bounds allow us to use Proposition 3.2(a). Combining (4.3) and (4.4), we establish in Corollary 4.10, that for all l and with

$$k_N = 2\left\lfloor \frac{\log\log N}{|\log(\tau-2)|} \right\rfloor + l, \tag{4.5}$$

we have

$$\begin{aligned}
\mathbb{P}(\{H_N > k_N\} \cap E_{m,N}(\varepsilon)) &= \mathbb{E}\Big[\mathbf{1}_{E_{m,N}(\varepsilon)\cap F_{m,N}(\varepsilon)} \exp\Big\{ -\lambda_N \min_{k_1\in\mathcal{B}_N} \frac{Z^{(1,N)}_{k_1+1} Z^{(2,N)}_{k_N-k_1}}{L_N}\Big\}\Big] + o_{N,m,\varepsilon}(1) \\
&= \mathbb{E}\Big[\mathbf{1}_{E_{m,N}(\varepsilon)\cap F_{m,N}(\varepsilon)} \exp\Big\{ -\lambda_N \min_{k_1\in\mathcal{B}_N} \frac{\exp\{\kappa^{k_1+1} Y^{(1,N)}_{k_1+1} + \kappa^{k_N-k_1} Y^{(2,N)}_{k_N-k_1}\}}{L_N}\Big\}\Big] + o_{N,m,\varepsilon}(1),
\end{aligned} \tag{4.6}$$

where $\kappa = (\tau-2)^{-1} > 1$.

In the final key step, in Sections 4.5 and 4.6, the minimum occurring in (4.6), with the approximations $Y^{(1,N)}_{k_1+1} \approx Y^{(1,N)}_m$ and $Y^{(2,N)}_{k_N-k_1} \approx Y^{(2,N)}_m$, is analyzed. The main idea in this analysis is as follows. With the above approximations, the right side of (4.6) can be rewritten as

$$\mathbb{E}\Big[\mathbf{1}_{E_{m,N}(\varepsilon)\cap F_{m,N}(\varepsilon)} \exp\Big\{ -\lambda_N \exp\Big[\min_{k_1\in\mathcal{B}_N} (\kappa^{k_1+1} Y^{(1,N)}_m + \kappa^{k_N-k_1} Y^{(2,N)}_m) - \log L_N\Big]\Big\}\Big] + o_{N,m,\varepsilon}(1). \tag{4.7}$$

The minimum appearing in the exponent of (4.7) is then rewritten (see (4.73) and (4.75)) as

$$\kappa^{\lceil k_N/2\rceil}\big\{\min_{t\in\mathbb{Z}}(\kappa^t Y^{(1,N)}_m + \kappa^{c_l - t} Y^{(2,N)}_m) - \kappa^{-\lceil k_N/2\rceil}\log L_N\big\}.$$

Since $\kappa^{\lceil k_N/2\rceil} \to \infty$, the latter expression only contributes to (4.7) when

$$\min_{t\in\mathbb{Z}}(\kappa^t Y^{(1,N)}_m + \kappa^{c_l - t} Y^{(2,N)}_m) - \kappa^{-\lceil k_N/2\rceil}\log L_N \leq 0.$$

Here it will become apparent that the bounds $\frac{1}{2} \le \lambda_N(k) \le 4k$ are sufficient. The expectation of the indicator of this event leads to the probability

$$\mathbb{P}\left(\min_{t\in\mathbb{Z}}(\kappa^t Y^{(1)} + \kappa^{c_l - t} Y^{(2)}) \le \kappa^{a_N - \lceil l/2 \rceil}, Y^{(1)} Y^{(2)} > 0\right),$$

with a_N and c_l as defined in Theorem 4.1. We complete the proof by showing that conditioning on the event that 1 and 2 are connected is asymptotically equivalent to conditioning on $Y^{(1)}Y^{(2)} > 0$.

Remark 4.2. *In the course of the proof, we will see that it is not necessary that the degrees of the nodes are i.i.d. In fact, in the proof below, we need that Propositions 3.1–3.4 are valid, as well as that L_N is concentrated around its mean μN. In Remark A.1.5 in the appendix, we will investigate what is needed in the proof of Propositions 3.1– 3.4. In particular, the proof applies also to some instances of the configuration model where the number of nodes with degree k is deterministic for each k, when we investigate the distance between two* uniformly *chosen nodes.*

We now go through the details of the proof.

4.2 A priory bounds on $Y_m^{(i,N)}$

We wish to compute the probability $\mathbb{P}(H_N > k)$. To do so, we split $\mathbb{P}(H_N > k)$ as

$$\mathbb{P}(H_N > k) = \mathbb{P}(H_N > k, Y_m^{(1,N)} Y_m^{(2,N)} = 0) + \mathbb{P}(H_N > k, Y_m^{(1,N)} Y_m^{(2,N)} > 0). \tag{4.8}$$

We will now prove two lemmas, and use these to compute the first term in the right-hand side of (4.8).

Lemma 4.3. *For any m fixed,*

$$\lim_{N\to\infty} \mathbb{P}(Y_m^{(1,N)} Y_m^{(2,N)} = 0) = 1 - q_m^2,$$

where

$$q_m = \mathbb{P}(Y_m^{(1)} > 0).$$

Proof. The proof is immediate from Proposition 3.1 and the independence of $Y_m^{(1)}$ and $Y_m^{(2)}$. □

The following lemma shows that the probability that $H_N \le m$ converges to zero for any *fixed* m:

Lemma 4.4. *For any m fixed,*

$$\lim_{N\to\infty} \mathbb{P}(H_N \le m) = 0.$$

Proof. As observed above Theorem 1.2, by exchangeability of the nodes $\{1, 2, \ldots, N\}$,

$$\mathbb{P}(H_N \le m) = \mathbb{P}(\widetilde{H}_N \le m), \tag{4.9}$$

where $\widetilde{H}_N$ is the hopcount between node 1 and a uniformly chosen node unequal to 1. We split, for any $0 < \delta < 1$,

$$\mathbb{P}(\widetilde{H}_N \le m) = \mathbb{P}(\widetilde{H}_N \le m, \sum_{j\le m} Z_j^{(1,N)} \le N^\delta) + \mathbb{P}(\widetilde{H}_N \le m, \sum_{j\le m} Z_j^{(1,N)} > N^\delta). \tag{4.10}$$

The number of nodes at distance at most m from node 1 is bounded from above by $\sum_{j\le m} Z_j^{(1,N)}$. The event $\{\widetilde{H}_N \le m\}$ can only occur when the end node, which is uniformly chosen in $\{2,\dots,N\}$, is in the SPG of node 1, so that

$$\mathbb{P}\Big(\widetilde{H}_N \le m, \sum_{j\le m} Z_j^{(1,N)} \le N^\delta\Big) \le \frac{N^\delta}{N-1} = o(1), \qquad N\to\infty. \tag{4.11}$$

Therefore, the first term in (4.10) is $o(1)$, as required. We will proceed with the second term in (4.10). By Proposition 3.1, **whp**, we have that $Y_j^{(1,N)} = Y_j^{(1)}$ for all $j \le m$. Therefore, we obtain, because $Y_j^{(1,N)} = Y_j^{(1)}$ implies $Z_j^{(1,N)} = \mathcal{Z}_j^{(1)}$,

$$\mathbb{P}\Big(\widetilde{H}_N \le m, \sum_{j\le m} Z_j^{(1,N)} > N^\delta\Big) \le \mathbb{P}\Big(\sum_{j\le m} Z_j^{(1,N)} > N^\delta\Big) = \mathbb{P}\Big(\sum_{j\le m} \mathcal{Z}_j^{(1)} > N^\delta\Big) + o(1).$$

However, when m is fixed, the random variable $\sum_{j\le m}\mathcal{Z}_j^{(1)}$ is finite with probability 1, and therefore,

$$\lim_{N\to\infty} \mathbb{P}\Big(\sum_{j\le m} Z_j^{(1,N)} > N^\delta\Big) = 0. \tag{4.12}$$

This completes the proof of Lemma 4.4. □

We now use Lemmas 4.3 and 4.4 to compute the first term in (4.8). We split

$$\mathbb{P}(H_N > k, Y_m^{(1,N)}Y_m^{(2,N)} = 0) = \mathbb{P}(Y_m^{(1,N)}Y_m^{(2,N)} = 0) - \mathbb{P}(H_N \le k, Y_m^{(1,N)}Y_m^{(2,N)} = 0). \tag{4.13}$$

By Lemma 4.3, the first term is equal to $1 - q_m^2 + o(1)$. For the second term, we note that when $Y_m^{(1,N)} = 0$ and $H_N < \infty$, then $H_N \le m-1$, so that

$$\mathbb{P}(H_N \le k, Y_m^{(1,N)}Y_m^{(2,N)} = 0) \le \mathbb{P}(H_N \le m-1). \tag{4.14}$$

Using Lemma 4.4, we conclude that

Corollary 4.5. *For every m fixed, and each $k\in\mathbb{N}$, possibly depending on N,*

$$\lim_{N\to\infty} \mathbb{P}(H_N > k, Y_m^{(1,N)}Y_m^{(2,N)} = 0) = 1 - q_m^2.$$

By Corollary 4.5 and (4.8), we are left to compute $\mathbb{P}(H_N > k, Y_m^{(1,N)}Y_m^{(2,N)} > 0)$. We first prove a lemma that shows that if $Y_m^{(1,N)} > 0$, then **whp** $Y_m^{(1,N)} \in [\varepsilon, \varepsilon^{-1}]$:

Lemma 4.6. *For $i = 1, 2$,*

$$\limsup_{\varepsilon\downarrow 0}\limsup_{m\to\infty}\limsup_{N\to\infty} \mathbb{P}(0 < Y_m^{(i,N)} < \varepsilon) = \limsup_{\varepsilon\downarrow 0}\limsup_{m\to\infty}\limsup_{N\to\infty} \mathbb{P}(Y_m^{(i,N)} > \varepsilon^{-1}) = 0.$$

Proof. Fix m, when $N\to\infty$ it follows from Proposition 3.1 that $Y_m^{(i,N)} = Y_m^{(i)}$, **whp**. Thus, we obtain that

$$\limsup_{\varepsilon\downarrow 0}\limsup_{m\to\infty}\limsup_{N\to\infty} \mathbb{P}(0 < Y_m^{(i,N)} < \varepsilon) = \limsup_{\varepsilon\downarrow 0}\limsup_{m\to\infty} \mathbb{P}(0 < Y_m^{(i)} < \varepsilon),$$

and similarly for the second probability. The remainder of the proof of the lemma follows because $Y_m^{(i)} \xrightarrow{d} Y^{(i)}$ as $m \to \infty$, and because conditionally on $Y^{(i)} > 0$ the random variable $Y^{(i)}$ admits a density. $\square$

Write

$$E_{m,N} = E_{m,N}(\varepsilon) = \{Y_m^{(i,N)} \in [\varepsilon, \varepsilon^{-1}], i = 1, 2\}, \tag{4.15}$$

$$F_{m,N} = F_{m,N}(\varepsilon) = \big\{ \max_{k \in \mathcal{T}_m^{(N)}(\varepsilon)} |Y_k^{(i,N)} - Y_m^{(i,N)}| \le \varepsilon^3, i = 1, 2\big\}. \tag{4.16}$$

As a consequence of Lemma 4.6, we obtain that

$$\mathbb{P}(E_{m,N}^c \cap \{Y_m^{(1,N)} Y_m^{(2,N)} > 0\}) = o_{N,m,\varepsilon}(1), \tag{4.17}$$

so that

$$\mathbb{P}(H_N > k, Y_m^{(1,N)} Y_m^{(2,N)} > 0) = \mathbb{P}(\{H_N > k\} \cap E_{m,N}) + o_{N,m,\varepsilon}(1). \tag{4.18}$$

In the sequel, we compute

$$\mathbb{P}(\{H_N > k\} \cap E_{m,N}), \tag{4.19}$$

and often we will make use of the fact that by Proposition 3.2,

$$\mathbb{P}(E_{m,N} \cap F_{m,N}^c) = o_{N,m,\varepsilon}(1). \tag{4.20}$$

4.3 Asymptotics of $\mathbb{P}(\{H_N > k\} \cap E_{m,N})$

We next give a representation of $\mathbb{P}(\{H_N > k\} \cap E_{m,N})$. In order to do so, we write $\mathbb{Q}_Z^{(i,j)}$, where $i, j \ge 0$, for the conditional probability given $\{Z_s^{(1,N)}\}_{s=1}^i$ and $\{Z_s^{(2,N)}\}_{s=1}^j$ (where, for $j = 0$, we condition only on $\{Z_s^{(1,N)}\}_{s=1}^i$), and $\mathbb{E}_Z^{(i,j)}$ for its conditional expectation. Furthermore, we say that a random variable k_1 is *Z_m-measurable* if k_1 is measurable with respect to the σ-algebra generated by $\{Z_s^{(1,N)}\}_{s=1}^m$ and $\{Z_s^{(2,N)}\}_{s=1}^m$. The main rewrite is now in the following lemma:

Lemma 4.7. *For $k \ge 2m - 1$,*

$$\mathbb{P}(\{H_N > k\} \cap E_{m,N}) = \mathbb{E}\Big[\mathbf{1}_{E_{m,N}} \mathbb{Q}_Z^{(m,m)}(H_N > 2m-1) P_m(k, k_1)\Big], \tag{4.21}$$

where, for any Z_m-measurable k_1, with $m \le k_1 \le (k-1)/2$,

$$P_m(k, k_1) = \prod_{i=2m}^{2k_1} \mathbb{Q}_Z^{(\lfloor i/2 \rfloor + 1, \lceil i/2 \rceil)}(H_N > i | H_N > i-1) \tag{4.22}$$
$$\times \prod_{i=1}^{k-2k_1} \mathbb{Q}_Z^{(k_1+1, k_1+i)}(H_N > 2k_1 + i | H_N > 2k_1 + i - 1).$$

Proof. We start by conditioning on $\{Z_s^{(1,N)}\}_{s=1}^m$ and $\{Z_s^{(2,N)}\}_{s=1}^m$, and note that $\mathbf{1}_{E_{m,N}}$ is Z_m-measurable, so that we obtain, for $k \ge 2m - 1$,

$$\mathbb{P}(\{H_N > k\} \cap E_{m,N}) = \mathbb{E}\Big[\mathbf{1}_{E_{m,N}} \mathbb{Q}_Z^{(m,m)}(H_N > k)\Big] \tag{4.23}$$
$$= \mathbb{E}\Big[\mathbf{1}_{E_{m,N}} \mathbb{Q}_Z^{(m,m)}(H_N > 2m-1) \mathbb{Q}_Z^{(m,m)}(H_N > k | H_N > 2m-1)\Big].$$

Moreover, for i, j such that $i + j \leq k$,

$$\begin{aligned}
&\mathbb{Q}_Z^{(i,j)}(H_N > k | H_N > i + j - 1) \qquad (4.24)\\
&= \mathbb{E}_Z^{(i,j)}\big[\mathbb{Q}_Z^{(i,j+1)}(H_N > k | H_N > i + j - 1)\big]\\
&= \mathbb{E}_Z^{(i,j)}\big[\mathbb{Q}_Z^{(i,j+1)}(H_N > i + j | H_N > i + j - 1)\mathbb{Q}_Z^{(i,j+1)}(H_N > k | H_N > i + j)\big],
\end{aligned}$$

and, similarly,

$$\begin{aligned}
&\mathbb{Q}_Z^{(i,j)}(H_N > k | H_N > i + j - 1) \qquad (4.25)\\
&= \mathbb{E}_Z^{(i,j)}\big[\mathbb{Q}_Z^{(i+1,j)}(H_N > i + j | H_N > i + j - 1)\mathbb{Q}_Z^{(i+1,j)}(H_N > k | H_N > i + j)\big].
\end{aligned}$$

In particular, we obtain, for $k > 2m - 1$,

$$\begin{aligned}
\mathbb{Q}_Z^{(m,m)}(H_N > k | H_N > 2m - 1) = \mathbb{E}_Z^{(m,m)}\Big[&\mathbb{Q}_Z^{(m+1,m)}(H_N > 2m | H_N > 2m - 1) \qquad (4.26)\\
&\times \mathbb{Q}_Z^{(m+1,m)}(H_N > k | H_N > 2m)\Big],
\end{aligned}$$

so that, using that $E_{m,N}$ is Z_m-measurable and that $\mathbb{E}[\mathbb{E}_Z^{(m,m)}[X]] = \mathbb{E}[X]$ for any random variable X,

$$\begin{aligned}
&\mathbb{P}(\{H_N > k\} \cap E_{m,N}) \qquad (4.27)\\
&= \mathbb{E}\Big[\mathbf{1}_{E_{m,N}}\mathbb{Q}_Z^{(m,m)}(H_N > 2m - 1)\mathbb{Q}_Z^{(m+1,m)}(H_N > 2m | H_N > 2m - 1)\mathbb{Q}_Z^{(m+1,m)}(H_N > k | H_N > 2m)\Big].
\end{aligned}$$

We now compute the conditional probability by repeatedly applying (4.24) and (4.25), increasing i or j as follows. For $i + j \leq 2k_1$, we will increase i and j in turn by 1, and for $2k_1 < i + j \leq k$, we will only increase the second component j. This leads to

$$\begin{aligned}
\mathbb{Q}_Z^{(m,m)}(H_N > k | H_N > 2m - 1) &= \mathbb{E}_Z^{(m,m)}\Big[\prod_{i=2m}^{2k_1} \mathbb{Q}_Z^{(\lfloor i/2 \rfloor + 1, \lceil i/2 \rceil)}(H_N > i | H_N > i - 1) \qquad (4.28)\\
&\quad \times \prod_{j=1}^{k-2k_1} \mathbb{Q}_Z^{(k_1+1, k_1+j)}(H_N > 2k_1 + j | H_N > 2k_1 + j - 1)\Big]\\
&= \mathbb{E}_Z^{(m,m)}[P_m(k, k_1)],
\end{aligned}$$

were we used that we can move the expectations $\mathbb{E}_Z^{(i,j)}$ outside, as in (4.27), so that these do not appear in the final formula. Therefore, from (4.23), (4.28), and since $\mathbf{1}_{E_{m,N}}$ and $\mathbb{Q}_Z^{(m,m)}(H_N > 2m - 1)$ are Z_m-measurable,

$$\begin{aligned}
\mathbb{P}(\{H_N > k\} \cap E_{m,N}) &= \mathbb{E}\Big[\mathbf{1}_{E_{m,N}}\mathbb{Q}_Z^{(m,m)}(H_N > 2m - 1)\mathbb{E}_Z^{(m,m)}[P_m(k, k_1)]\Big]\\
&= \mathbb{E}\Big[\mathbb{E}_Z^{(m,m)}[\mathbf{1}_{E_{m,N}}\mathbb{Q}_Z^{(m,m)}(H_N > 2m - 1)P_m(k, k_1)]\Big]\\
&= \mathbb{E}\Big[\mathbf{1}_{E_{m,N}}\mathbb{Q}_Z^{(m,m)}(H_N > 2m - 1)P_m(k, k_1)\Big]. \qquad (4.29)
\end{aligned}$$

This proves (4.22). □

We note that we can omit the term $\mathbb{Q}_Z^{(m,m)}(H_N > 2m-1)$ in (4.21) by introducing a small error term. Indeed, we can write

$$\mathbb{Q}_Z^{(m,m)}(H_N > 2m-1) = 1 - \mathbb{Q}_Z^{(m,m)}(H_N \leq 2m-1). \tag{4.30}$$

Bounding $\mathbf{1}_{E_{m,N}} P_m(k,k_1) \leq 1$, the contribution to (4.21) due to the second term in the right-hand side of (4.30) is according to Lemma 4.4 bounded by

$$\mathbb{E}\Big[\mathbb{Q}_Z^{(m,m)}(H_N \leq 2m-1)\Big] = \mathbb{P}(H_N \leq 2m-1) = o_N(1). \tag{4.31}$$

We conclude from (4.20), (4.21), and (4.31), that

$$\mathbb{P}(\{H_N > k\} \cap E_{m,N}) = \mathbb{E}\Big[\mathbf{1}_{E_{m,N}\cap F_{m,N}} P_m(k,k_1)\Big] + o_{N,m,\varepsilon}(1). \tag{4.32}$$

We continue with (4.32) by bounding the conditional probabilities in $P_m(k,k_1)$ defined in (4.22).

Lemma 4.8. *For all integers $i,j \geq 0$,*

$$\exp\left\{-\frac{4Z_{i+1}^{(1,N)} Z_j^{(2,N)}}{L_N}\right\} \leq \mathbb{Q}_Z^{(i+1,j)}(H_N > i+j | H_N > i+j-1) \leq \exp\left\{-\frac{Z_{i+1}^{(1,N)} Z_j^{(2,N)}}{2L_N}\right\}. \tag{4.33}$$

The upper bound is always valid, the lower bound is valid whenever

$$\sum_{s=1}^{i+1} Z_s^{(1,N)} + \sum_{s=1}^{j} Z_s^{(2,N)} \leq \frac{L_N}{4}. \tag{4.34}$$

Proof. We start with the upper bound. We fix two sets of n_1 and n_2 stubs, and will be interested in the probability that none of the n_1 stubs are connected to the n_2 stubs. We order the n_1 stubs in an arbitrary way, and connect the stubs iteratively to other stubs. Note that we must connect at least $\lceil n_1/2 \rceil$ stubs, since any stub that is being connected removes at most 2 stubs from the total of n_1 stubs. The number $n_1/2$ is reached for n_1 even precisely when all the n_1 stubs are connected with each other. Therefore, we obtain that the probability that the n_1 stubs are not connected to the n_2 stubs is bounded from above by

$$\prod_{t=1}^{\lceil n_1/2 \rceil}\Big(1 - \frac{n_2}{L_N - 2t + 1}\Big) \leq \prod_{t=1}^{\lceil n_1/2 \rceil}\Big(1 - \frac{n_2}{L_N}\Big). \tag{4.35}$$

Using the inequality $1 - x \leq e^{-x}$, $x \geq 0$, we obtain that the probability that the n_1 stubs are not connected to the n_2 stubs is bounded from above by

$$e^{-\lceil n_1/2 \rceil \frac{n_2}{L_N}} \leq e^{-\frac{n_1 n_2}{2L_N}}. \tag{4.36}$$

Applying the above bound to $n_1 = Z_{i+1}^{(1,N)}$ and $n_2 = Z_j^{(2,N)}$, and noting that the probability that $H_N > i+j$ given that $H_N > i+j-1$ is bounded from above by the probability that none of the $Z_{i+1}^{(1,N)}$ stubs are connected to the $Z_j^{(2,N)}$ stubs leads to the upper bound in (4.33).

We again fix two sets of n_1 and n_2 stubs, and are again interested in the probability that none of the n_1 stubs are connected to the n_2 stubs. However, now we use these bounds repeatedly, and we assume that in each step there remain to be at least L stubs available. We order the n_1 stubs in an arbitrary way, and connect the stubs iteratively to other stubs. We obtain a lower bound by further requiring that the n_1 stubs do not connect to each other. Therefore, the probability that the n_1 stubs are not connected to the n_2 stubs is bounded below by

$$\prod_{t=1}^{n_1}\Big(1-\frac{n_2}{L-2t+1}\Big). \tag{4.37}$$

When $L-2n_1\geq \frac{L_N}{2}$ and $1\leq t\leq n_1$, we obtain that $1-\frac{n_2}{L-2t+1}\geq 1-\frac{2n_2}{L_N}$. Moreover, when $x\leq \frac{1}{2}$, we have that $1-x\geq e^{-2x}$. Therefore, we obtain that when $L-2n_1\geq \frac{L_N}{2}$ and $n_2\leq \frac{L_N}{4}$, then the probability that the n_1 stubs are not connected to the n_2 stubs when there are still at least L stubs available is bounded below by

$$\prod_{t=1}^{n_1}\Big(1-\frac{n_2}{L-2t+1}\Big)\geq \prod_{t=1}^{n_1} e^{-\frac{4n_2}{L_N}}=e^{-\frac{4n_1n_2}{L_N}}. \tag{4.38}$$

The event $H_N>i+j$ conditionally on $H_N>i+j-1$ precisely occurs when none of the $Z_{i+1}^{(1,N)}$ stubs are connected to the $Z_j^{(2,N)}$ stubs. We will assume that (4.34) holds. We have that $L=L_N-2\sum_{s=1}^{i}Z_s^{(1,N)}-2\sum_{s=1}^{j}Z_s^{(2,N)}$, and $n_1=Z_{i+1}^{(1,N)}$, $n_2=Z_j^{(2,N)}$. Thus, $L-2n_1\geq \frac{L_N}{2}$ happens precisely when

$$L-2n_1=L_N-2\sum_{s=1}^{i+1}Z_s^{(1,N)}-2\sum_{s=1}^{j}Z_s^{(2,N)}\geq \frac{L_N}{2}. \tag{4.39}$$

This follows from the assumed bound in (4.34). Also, when $n_2=Z_j^{(2,N)}$, $n_2\leq \frac{L_N}{4}$ is implied by (4.34). Thus, we are allowed to use the bound in (4.38). This leads to

$$\mathbb{Q}_Z^{(i+1,j)}(H_N>i+j|H_N>i+j-1)\geq \exp\Big\{-\frac{4Z_{i+1}^{(1,N)}Z_j^{(2,N)}}{L_N}\Big\}, \tag{4.40}$$

which completes the proof of Lemma 4.8. □

4.4 The main contribution to $\mathbb{P}(\{H_N>k\}\cap E_{m,N})$

We rewrite the expression in (4.32) in a more convenient form, using Lemma 4.8. We derive an upper and a lower bound. For the upper bound, we bound all terms appearing on the right-hand side of (4.22) by 1, except for the term $\mathbb{Q}_Z^{(k_1+1,k-k_1)}(H_N>k|H_N>k-1)$, which arises when $i=k-2k_1$, in the second product. Using the upper bound in Lemma 4.8, we thus obtain that

$$P_m(k,k_1)\leq \exp\Big\{-\frac{Z_{k_1+1}^{(1,N)}Z_{k-k_1}^{(2,N)}}{2L_N}\Big\}. \tag{4.41}$$

The latter inequality is true for any Z_m-measurable k_1, with $m\leq k_1\leq (k-1)/2$.

To derive the lower bound, we next assume that

$$\sum_{s=1}^{k_1+1}Z_s^{(1,N)}+\sum_{s=1}^{k-k_1}Z_s^{(2,N)}\leq \frac{L_N}{4}, \tag{4.42}$$

so that (4.34) is satisfied for all i in (4.22). We write, recalling (3.3),

$$\mathcal{B}_N^{(1)}(\varepsilon,k)=\Big\{m\le l\le (k-1)/2:\ l+1\in \mathcal{T}_m^{(1,N)}(\varepsilon),\ k-l\in \mathcal{T}_m^{(2,N)}(\varepsilon)\Big\}. \tag{4.43}$$

We restrict ourselves to $k_1\in\mathcal{B}_N^{(1)}(\varepsilon,k)$, if $\mathcal{B}_N^{(1)}(\varepsilon,k)\neq\varnothing$. When $k_1\in\mathcal{B}_N^{(1)}(\varepsilon,k)$, we are allowed to use the bounds in Proposition 3.2. Note that $\{k_1\in\mathcal{B}_N^{(1)}(\varepsilon,k)\}$ is Z_m-measurable. Moreover, it follows from Proposition 3.2 that if $k_1\in\mathcal{B}_N^{(1)}(\varepsilon,k)$, that then, with probability converging to 1 as first $N\to\infty$ and then $m\to\infty$,

$$Z_s^{(1,N)}\le N^{\frac{1-\varepsilon^4}{\tau-1}},\quad \forall m<s\le k_1+1,\qquad\text{and}\qquad Z_s^{(2,N)}\le N^{\frac{1-\varepsilon^4}{\tau-1}},\quad \forall m<s\le k-k_1. \tag{4.44}$$

When $k_1\in\mathcal{B}_N^{(1)}(\varepsilon,k)$, we have

$$\sum_{s=1}^{k_1+1}Z_s^{(1,N)}+\sum_{s=1}^{k-k_1}Z_s^{(2,N)}=kO(N^{\frac{1}{\tau-1}})=o(N)=o(L_N),$$

as long as $k=o(N^{\frac{\tau-2}{\tau-1}})$. Since throughout the paper $k=O(\log\log N)$ (see e.g. Theorem 1.2), and $\frac{\tau-2}{\tau-1}>0$, the Assumption (4.42) will always be fulfilled.

Thus, on the event $E_{m,N}\cap\{k_1\in\mathcal{B}_N^{(1)}(\varepsilon,k)\}$, using (3.5) in Proposition 3.2 and the lower bound in Lemma 4.8, with probability $1-o_{N,m,\varepsilon}(1)$, and for all $i\in\{2m,\dots,2k_1-1\}$,

$$\mathbb{Q}_Z^{(\lfloor i/2\rfloor+1,\lceil i/2\rceil)}(H_N>i|H_N>i-1)\ge \exp\Big\{-\frac{4Z_{\lfloor i/2\rfloor+1}^{(1,N)}Z_{\lceil i/2\rceil}^{(2,N)}}{L_N}\Big\}\ge \exp\Big\{-\frac{4Z_{k_1+1}^{(1,N)}Z_{k-k_1}^{(2,N)}}{L_N}\Big\}, \tag{4.45}$$

and, for $1\le i\le k-2k_1$,

$$\mathbb{Q}_Z^{(k_1+1,k_1+i)}(H_N>2k_1+i|H_N>2k_1+i-1)\ge \exp\Big\{-\frac{4Z_{k_1+1}^{(1,N)}Z_{k_1+i}^{(2,N)}}{L_N}\Big\}\ge \exp\Big\{-\frac{4Z_{k_1+1}^{(1,N)}Z_{k-k_1}^{(2,N)}}{L_N}\Big\}. \tag{4.46}$$

Therefore, by Lemma 4.7, and using the above bounds for each of the in total $k-2m+1$ terms, we obtain that when $k_1\in\mathcal{B}_N^{(1)}(\varepsilon,k)\neq\varnothing$, and with probability $1-o_{N,m,\varepsilon}(1)$,

$$P_m(k,k_1)\ge\left(\exp\Big\{-4\frac{Z_{k_1+1}^{(1,N)}Z_{k-k_1}^{(2,N)}}{L_N}\Big\}\right)^{k-2m+1}\ge \exp\Big\{-4k\frac{Z_{k_1+1}^{(1,N)}Z_{k-k_1}^{(2,N)}}{L_N}\Big\}. \tag{4.47}$$

We next use the symmetry for the nodes 1 and 2. Denote

$$\mathcal{B}_N^{(2)}(\varepsilon,k)=\Big\{m\le l\le (k-1)/2:\ l+1\in \mathcal{T}_m^{(2,N)}(\varepsilon),\ k-l\in \mathcal{T}_m^{(1,N)}(\varepsilon)\Big\}. \tag{4.48}$$

Take $\tilde l=k-l-1$, so that $(k-1)/2\le\tilde l\le k-1-m$, and thus

$$\mathcal{B}_N^{(2)}(\varepsilon,k)=\Big\{(k-1)/2\le \tilde l\le k-1-m:\ \tilde l+1\in \mathcal{T}_m^{(1,N)}(\varepsilon),\ k-\tilde l\in \mathcal{T}_m^{(2,N)}(\varepsilon)\Big\}. \tag{4.49}$$

Then, since the nodes 1 and 2 are exchangeable, we obtain from (4.47), when $k_1\in\mathcal{B}_N^{(2)}(\varepsilon,k)\neq\varnothing$, and with probability $1-o_{N,m,\varepsilon}(1)$,

$$P_m(k,k_1)\ge \exp\Big\{-4k\frac{Z_{k_1+1}^{(1,N)}Z_{k-k_1}^{(2,N)}}{L_N}\Big\}. \tag{4.50}$$

We define $\mathcal{B}_N(\varepsilon,k)=\mathcal{B}_N^{(1)}(\varepsilon,k)\cup\mathcal{B}_N^{(2)}(\varepsilon,k)$, which is equal to

$$\mathcal{B}_N(\varepsilon,k)=\Big\{m\le l\le k-1-m:\ l+1\in\mathcal{T}_m^{(1,N)}(\varepsilon),\ k-l\in\mathcal{T}_m^{(2,N)}(\varepsilon)\Big\}. \tag{4.51}$$

We can summarize the obtained results by writing that with probability $1-o_{N,m,\varepsilon}(1)$, and when $\mathcal{B}_N(\varepsilon,k)\neq\varnothing$, we have

$$P_m(k,k_1)=\exp\big\{-\lambda_N\frac{Z_{k_1+1}^{(1,N)}Z_{k-k_1}^{(2,N)}}{L_N}\big\}, \tag{4.52}$$

for all $k_1\in\mathcal{B}_N(\varepsilon,k)$, where $\lambda_N=\lambda_N(k)$ satisfies

$$\frac{1}{2}\le\lambda_N(k)\le 4k. \tag{4.53}$$

Relation (4.52) is true for any $k_1\in\mathcal{B}_N(\varepsilon,k)$. However, our coupling fails when $Z_{k_1+1}^{(1,N)}$ or $Z_{k-k_1}^{(2,N)}$ grows too large, since we can only couple $Z_j^{(i,N)}$ with $\hat{Z}_j^{(i,N)}$ up to the point where $Z_j^{(i,N)}\le N^{\frac{1-\varepsilon^2}{\tau-1}}$. Therefore, we next take the maximal value over $k_1\in\mathcal{B}_N(\varepsilon,k)$ to arrive at the fact that, with probability $1-o_{N,m,\varepsilon}(1)$, on the event that $\mathcal{B}_N(\varepsilon,k)\neq\varnothing$,

$$P_m(k,k_1)=\max_{k_1\in\mathcal{B}_N(\varepsilon,k)}\exp\big\{-\lambda_N\frac{Z_{k_1+1}^{(1,N)}Z_{k-k_1}^{(2,N)}}{L_N}\big\}=\exp\Big\{-\lambda_N\min_{k_1\in\mathcal{B}_N(\varepsilon,k)}\frac{Z_{k_1+1}^{(1,N)}Z_{k-k_1}^{(2,N)}}{L_N}\Big\}. \tag{4.54}$$

From here on we take $k=k_N$ as in (4.5) with l a fixed integer.

In Section 5, we prove the following lemma that shows that, apart from an event of probability $1-o_{N,m,\varepsilon}(1)$, we may assume that $\mathcal{B}_N(\varepsilon,k_N)\neq\varnothing$:

Lemma 4.9. *For all l, with k_N as in (4.5),*

$$\limsup_{\varepsilon\downarrow 0}\limsup_{m\to\infty}\limsup_{N\to\infty}\mathbb{P}(\{H_N>k_N\}\cap E_{m,N}\cap\{\mathcal{B}_N(\varepsilon,k_N)=\varnothing\})=0.$$

From now on, we will abbreviate $\mathcal{B}_N=\mathcal{B}_N(\varepsilon,k_N)$. Using (4.32), (4.54) and Lemma 4.9, we conclude that,

Corollary 4.10. *For all l, with k_N as in (4.5),*

$$\mathbb{P}\big(\{H_N>k_N\}\cap E_{m,N}\big)=\mathbb{E}\Big[\mathbf{1}_{E_{m,N}\cap F_{m,N}}\exp\Big\{-\lambda_N\min_{k_1\in\mathcal{B}_N}\frac{Z_{k_1+1}^{(1,N)}Z_{k_N-k_1}^{(2,N)}}{L_N}\Big\}\Big]+o_{N,m,\varepsilon}(1),$$

where

$$\frac{1}{2}\le\lambda_N(k_N)\le 4k_N.$$

4.5 Application of the coupling results

In this section, we use the coupling results in Section 3.3. Before doing so, we investigate the minimum of the function $t\mapsto\kappa^t y_1+\kappa^{n-t}y_2$, where the minimum is taken over the discrete set $\{0,1,\dots,n\}$, and where we recall that $\kappa=(\tau-2)^{-1}$.

Lemma 4.11. *Suppose that $y_1 > y_2 > 0$, and $\kappa = (\tau - 2)^{-1} > 1$. Fix an integer n, satisfying $n > \frac{|\log(y_2/y_1)|}{\log \kappa}$, then*

$$t^* = \mathrm{argmin}_{t\in\{1,2,\ldots,n\}} \left(\kappa^t y_1 + \kappa^{n-t} y_2\right) = \mathrm{round}\left(\frac{n}{2} + \frac{\log(y_2/y_1)}{2\log\kappa}\right),$$

where round(x) is x rounded off to the nearest integer. In particular,

$$\max\left\{\frac{\kappa^{t^*} y_1}{\kappa^{n-t^*} y_2}, \frac{\kappa^{n-t^*} y_2}{\kappa^{t^*} y_1}\right\} \le \kappa.$$

Proof. Consider, for real-valued $t \in [0, n]$, the function

$$\psi(t) = \kappa^t y_1 + \kappa^{n-t} y_2.$$

Then,

$$\psi'(t) = (\kappa^t y_1 - \kappa^{n-t} y_2)\log\kappa, \qquad \psi''(t) = (\kappa^t y_1 + \kappa^{n-t} y_2)\log^2\kappa.$$

In particular, $\psi''(t) > 0$, so that the function ψ is strictly convex. The unique minimum of ψ is attained at $\hat{t}$, satisfying $\psi'(\hat{t}) = 0$, i.e.,

$$\hat{t} = \frac{n}{2} + \frac{\log(y_2/y_1)}{2\log\kappa} \in (0, n),$$

because $n > -\log(y_2/y_1)/\log\kappa$. By convexity $t^* = \lfloor \hat{t} \rfloor$ or $t^* = \lceil \hat{t} \rceil$. We will show that $|t^* - \hat{t}| \le \frac{1}{2}$. Put $t_1^* = \lfloor \hat{t} \rfloor$ and $t_2^* = \lceil \hat{t} \rceil$. We have

$$\kappa^{\hat{t}} y_1 = \kappa^{n-\hat{t}} y_2 = \kappa^{\frac{n}{2}} \sqrt{y_1 y_2}. \tag{4.55}$$

Writing $t_i^* = \hat{t} + t_i^* - \hat{t}$, we obtain for $i = 1, 2$,

$$\psi(t_i^*) = \kappa^{\frac{n}{2}} \sqrt{y_1 y_2}\{\kappa^{t_i^* - \hat{t}} + \kappa^{\hat{t} - t_i^*}\}.$$

For $0 < x < 1$, the function $x \mapsto \kappa^x + \kappa^{-x}$ is increasing so $\psi(t_1^*) \le \psi(t_2^*)$ if and only if $\hat{t} - t_1^* \le t_2^* - \hat{t}$, or $\hat{t} - t_1^* \le \frac{1}{2}$, i.e., if $\psi(t_1^*) \le \psi(t_2^*)$ and hence the minimum over the discrete set $\{0, 1, \ldots, n\}$ is attained at t_1^*, then $\hat{t} - t_1^* \le \frac{1}{2}$. On the other hand, if $\psi(t_2^*) \le \psi(t_1^*)$, then by the 'only if' statement we find $t_2^* - \hat{t} \le \frac{1}{2}$. In both cases we have $|t^* - \hat{t}| \le \frac{1}{2}$. Finally, if $t^* = t_1^*$, then we obtain, using (4.55),

$$1 \le \frac{\kappa^{n-t^*} y_2}{\kappa^{t^*} y_1} = \frac{\kappa^{\hat{t} - t_1^*}}{\kappa^{t_1^* - \hat{t}}} = \kappa^{2(\hat{t} - t_1^*)} \le \kappa,$$

while for $t^* = t_2^*$, we obtain $1 \le \frac{\kappa^{t^*} y_1}{\kappa^{n-t^*} y_2} \le \kappa$. $\square$

We continue with our investigation of $\mathbb{P}\big(\{H_N > k_N\} \cap E_{m,N}\big)$. We start from Corollary 4.10, and substitute (3.1) to obtain,

$$\begin{aligned} &\mathbb{P}\big(\{H_N > k_N\} \cap E_{m,N}\big) \\ &= \mathbb{E}\Big[\mathbf{1}_{E_{m,N}\cap F_{m,N}} \exp\Big\{-\lambda_N \exp\Big[\min_{k_1\in\mathcal{B}_N} \big(\kappa^{k_1+1} Y^{(1,N)}_{k_1+1} + \kappa^{k_N-k_1} Y^{(2,N)}_{k_N-k_1}\big) - \log L_N\Big]\Big\}\Big] + o_{N,m,\varepsilon}(1), \end{aligned} \tag{4.56}$$

where we rewrite, using (4.51) and (3.3),

$$\mathcal{B}_N = \left\{ m \leq k_1 \leq k_N - 1 - m : \kappa^{k_1+1} Y_m^{(1,N)} \leq \frac{1-\varepsilon^2}{\tau - 1} \log N, \kappa^{k_N - k_1} Y_m^{(2,N)} \leq \frac{1-\varepsilon^2}{\tau-1} \log N \right\}. \tag{4.57}$$

Moreover, on $F_{m,N}$, we have that $\min_{k_1 \in \mathcal{B}_N} (\kappa^{k_1+1} Y_{k_1+1}^{(1,N)} + \kappa^{k_N-k_1} Y_{k_N-k_1}^{(2,N)})$ is between

$$\min_{k_1 \in \mathcal{B}_N} \left(\kappa^{k_1+1}(Y_m^{(1,N)} - \varepsilon^3) + \kappa^{k_N-k_1}(Y_m^{(2,N)} - \varepsilon^3)\right)$$

and

$$\min_{k_1 \in \mathcal{B}_N} \left(\kappa^{k_1+1}(Y_m^{(1,N)} + \varepsilon^3) + \kappa^{k_N-k_1}(Y_m^{(2,N)} + \varepsilon^3)\right).$$

To abbreviate the notation, we will write, for $i = 1, 2$,

$$Y_{m,+}^{(i,N)} = Y_m^{(i,N)} + \varepsilon^3, \qquad Y_{m,-}^{(i,N)} = Y_m^{(i,N)} - \varepsilon^3. \tag{4.58}$$

Define for $\varepsilon > 0$,

$$H_{m,N} = H_{m,N}(\varepsilon) = \left\{ \min_{0 \leq k_1 \leq k_N - 1} \left(\kappa^{k_1+1} Y_{m,-}^{(1,N)} + \kappa^{k_N-k_1} Y_{m,-}^{(2,N)}\right) \leq (1+\varepsilon^2)\log N \right\}.$$

On the complement $H_{m,N}^c$, the minimum over $0 \leq k_1 \leq k_N - 1$ of $\kappa^{k_1+1} Y_{m,-}^{(1,N)} + \kappa^{k_N-k_1} Y_{m,-}^{(2,N)}$ exceeds $(1+\varepsilon^2)\log N$. Therefore, also the minimum over the set $\mathcal{B}_N$ of $\kappa^{k_1+1} Y_{m,-}^{(1,N)} + \kappa^{k_N-k_1} Y_{m,-}^{(2,N)}$ exceeds $(1+\varepsilon^2)\log N$, so that from (4.56), Lemma 4.8 and Proposition 3.2 and with error at most $o_{N,m,\varepsilon}(1)$,

$$\begin{aligned}
&\mathbb{P}\Big(\{H_N > k_N\} \cap E_{m,N} \cap H_{m,N}^c\Big) \\
&\quad \leq \mathbb{E}\Big[\mathbf{1}_{H_{m,N}^c} \exp\Big\{ -\frac{1}{2} \exp\Big[\min_{k_1 \in \mathcal{B}_N} \left(\kappa^{k_1+1} Y_{k_1+1}^{(1,N)} + \kappa^{k_N-k_1} Y_{k_N-k_1}^{(2,N)}\right) - \log L_N\Big]\Big\}\Big] \\
&\quad \leq \mathbb{E}\Big[\mathbf{1}_{H_{m,N}^c} \exp\Big\{ -\frac{1}{2} \exp\Big[\min_{k_1 \in \mathcal{B}_N} \left(\kappa^{k_1+1} Y_{m,-}^{(1,N)} + \kappa^{k_N-k_1} Y_{m,-}^{(2,N)}\right) - \log L_N\Big]\Big\}\Big] \\
&\quad \leq \mathbb{E}\Big[\exp\Big\{ -\frac{1}{2} \exp\left((1+\varepsilon^2)\log N - \log L_N\right)\Big\}\Big] \leq e^{-\frac{1}{2c} N^{\varepsilon^2}} = o_{N,m,\varepsilon}(1),
\end{aligned} \tag{4.59}$$

because $L_N \leq cN$, **whp**, as $N \to \infty$. Combining (4.59) with (4.18) yields

$$\mathbb{P}(H_N > k, Y_m^{(1,N)} Y_m^{(2,N)} > 0) = \mathbb{P}(\{H_N > k\} \cap E_{m,N} \cap H_{m,N}) + o_{N,m,\varepsilon}(1). \tag{4.60}$$

Therefore, in the remainder of the proof, we assume that $H_{m,N}$ holds.

Lemma 4.12. *With probability exceeding* $1 - o_{N,m,\varepsilon}(1)$,

$$\min_{k_1 \in \mathcal{B}_N} \left(\kappa^{k_1+1} Y_{m,+}^{(1,N)} + \kappa^{k_N-k_1} Y_{m,+}^{(2,N)}\right) = \min_{0 \leq k_1 < k_N} \left(\kappa^{k_1+1} Y_{m,+}^{(1,N)} + \kappa^{k_N-k_1} Y_{m,+}^{(2,N)}\right), \tag{4.61}$$

and

$$\min_{k_1 \in \mathcal{B}_N} \left(\kappa^{k_1+1} Y_{m,-}^{(1,N)} + \kappa^{k_N-k_1} Y_{m,-}^{(2,N)}\right) = \min_{0 \leq k_1 < k_N} \left(\kappa^{k_1+1} Y_{m,-}^{(1,N)} + \kappa^{k_N-k_1} Y_{m,-}^{(2,N)}\right). \tag{4.62}$$

Proof. We start with (4.61), the proof of (4.62) is similar, and, in fact, slightly simpler, and is therefore omitted. To prove (4.61), we use Lemma 4.11, with $n = k_N + 1$, $t = k_1 + 1$, $y_1 = Y^{(1,N)}_{m,+}$ and $y_2 = Y^{(2,N)}_{m,+}$. Let

$$t^* = \mathrm{argmin}_{t\in\{1,2,\ldots,n\}} \left(\kappa^t y_1 + \kappa^{n-t} y_2\right),$$

and assume (without restriction) that $\kappa^{t^*} y_1 \geq \kappa^{n-t^*} y_2$. We have to show that $t^* - 1 \in \mathcal{B}_N$. According to Lemma 4.11,

$$1 \leq \frac{\kappa^{t^*} Y^{(1,N)}_{m,+}}{\kappa^{n-t^*} Y^{(2,N)}_{m,+}} = \frac{\kappa^{t^*} y_1}{\kappa^{n-t^*} y_2} \leq \kappa. \tag{4.63}$$

We define $x = \kappa^{t^*} Y^{(1,N)}_{m,+}$ and $y = \kappa^{n-t^*} Y^{(2,N)}_{m,+}$, so that $x \geq y$. By definition, on $H_{m,N}$,

$$\kappa^{t^*} Y^{(1,N)}_{m,-} + \kappa^{n-t^*} Y^{(2,N)}_{m,-} \leq (1+\varepsilon^2)\log N.$$

Since, on $E_{m,N}$, we have that $Y^{(1,N)}_m \geq \varepsilon$,

$$Y^{(1,N)}_{m,+} = Y^{(1,N)}_{m,-}\left(1 + \frac{2\varepsilon^3}{Y^{(1,N)}_{m,-}}\right) \leq \frac{\varepsilon+\varepsilon^3}{\varepsilon-\varepsilon^3} Y^{(1,N)}_{m,-} = \frac{1+\varepsilon^2}{1-\varepsilon^2} Y^{(1,N)}_{m,-}, \tag{4.64}$$

and likewise for $Y^{(2,N)}_{m,+}$. Therefore, we obtain that on $E_{m,N} \cap H_{m,N}$, and with ε sufficiently small,

$$x + y \leq \frac{1+\varepsilon^2}{1-\varepsilon^2}\left[\kappa^{t^*} Y^{(1,N)}_{m,-} + \kappa^{n-t^*} Y^{(2,N)}_{m,-}\right] \leq \frac{(1+\varepsilon^2)^2}{1-\varepsilon^2}\log N \leq (1+\varepsilon)\log N. \tag{4.65}$$

Moreover, by (4.63), we have that

$$1 \leq \frac{x}{y} \leq \kappa. \tag{4.66}$$

Hence, on $E_{m,N} \cap H_{m,N}$, we have, with $\kappa^{-1} = \tau - 2$,

$$x = \frac{x+y}{1+\frac{y}{x}} \leq (1+\varepsilon)\frac{1}{1+\kappa^{-1}}\log N = \frac{1+\varepsilon}{\tau-1}\log N, \tag{4.67}$$

when $\varepsilon > 0$ is sufficiently small. We claim that if (note the difference with (4.67)),

$$x = \kappa^{t^*} Y^{(1,N)}_{m,+} \leq \frac{1-\varepsilon}{\tau-1}\log N, \tag{4.68}$$

then $k_1^* = t^* - 1 \in \mathcal{B}_N(\varepsilon, k_N)$, so that (4.61) follows. Indeed, we use (4.68) to see that

$$\kappa^{k_1^*+1} Y^{(1,N)}_m = \kappa^{t^*} Y^{(1,N)}_m \leq \kappa^{t^*} Y^{(1,N)}_{m,+} \leq \frac{1-\varepsilon}{\tau-1}\log N, \tag{4.69}$$

so that the first bound in (4.57) is satisfied. The second bound is satisfied, since

$$\kappa^{k_N - k_1^*} Y^{(2,N)}_m = \kappa^{n-t^*} Y^{(2,N)}_m \leq \kappa^{n-t^*} Y^{(2,N)}_{m,+} = y \leq x \leq \frac{1-\varepsilon}{\tau-1}\log N, \tag{4.70}$$

where we have used $n = k_N + 1$, and (4.68). Thus indeed $k_1^* \in \mathcal{B}_N(\varepsilon, k_N)$.

We conclude that, in order to show that (4.61) holds with error at most $o_{N,m,\varepsilon}(1)$, we have to show that the probability of the intersection of the events $\{H_N > k_N\}$ and

$$\mathcal{E}_{m,N} = \mathcal{E}_{m,N}(\varepsilon) = \Big\{\exists t : \frac{1-\varepsilon}{\tau-1}\log N < \kappa^t Y^{(1,N)}_{m,+} \le \frac{1+\varepsilon}{\tau-1}\log N, \\ \kappa^t Y^{(1,N)}_{m,+} + \kappa^{n-t} Y^{(2,N)}_{m,+} \le (1+\varepsilon)\log N\Big\}, \tag{4.71}$$

is of order $o_{N,m,\varepsilon}(1)$. This is contained in Lemma 4.13 below. □

Lemma 4.13. *For k_N as in (4.5),*

$$\limsup_{\varepsilon\downarrow 0}\limsup_{m\to\infty}\limsup_{N\to\infty}\mathbb{P}(E_{m,N}(\varepsilon)\cap\mathcal{E}_{m,N}(\varepsilon)\cap\{H_N > k_N\}) = 0.$$

The proof of Lemma 4.13 is deferred to Section 5.

From (4.56), Lemmas 4.12 and 4.13, we finally arrive at

$$\mathbb{P}\big(\{H_N > k_N\}\cap E_{m,N}\big) \\ \le \mathbb{E}\Big[\mathbf{1}_{E_{m,N}}\exp\Big\{-\lambda_N\exp\Big[\min_{0\le k_1<k_N}\big(\kappa^{k_1+1}Y^{(1,N)}_{m,-} + \kappa^{k_N-k_1}Y^{(2,N)}_{m,-}\big) - \log L_N\Big]\Big\}\Big] + o_{N,m,\varepsilon}(1), \tag{4.72}$$

and at a similar lower bound where $Y^{(i,N)}_{m,-}$ is replaced by $Y^{(i,N)}_{m,+}$. Note that on the right-hand side of (4.72), we have replaced the intersection of $\mathbf{1}_{E_{m,N}\cap F_{m,N}}$ by $\mathbf{1}_{E_{m,N}}$, which is allowed, because of (4.20).

4.6 Evaluating the limit

The final argument starts from (4.72) and the similar lower bound, and consists of letting $N\to\infty$ and then $m\to\infty$. The argument has to be performed with $Y^{(i,N)}_{m,+}$ and $Y^{(i,N)}_{m,-}$ separately, after which we let $\varepsilon\downarrow 0$. Since the precise value of ε plays no role in the derivation, we only give the derivation for $\varepsilon = 0$. Observe that

$$\begin{aligned}&\min_{0\le k_1<k_N}(\kappa^{k_1+1}Y^{(1,N)}_m + \kappa^{k_N-k_1}Y^{(2,N)}_m) - \log L_N \\ &\quad= \kappa^{\lceil k_N/2\rceil}\min_{0\le k_1<k_N}\Big(\kappa^{k_1+1-\lceil k_N/2\rceil}Y^{(1,N)}_m + \kappa^{\lfloor k_N/2\rfloor-k_1}Y^{(2,N)}_m - \kappa^{-\lceil k_N/2\rceil}\log L_N\Big) \\ &\quad= \kappa^{\lceil k_N/2\rceil}\min_{-\lceil k_N/2\rceil+1\le t<\lfloor k_N/2\rfloor+1}(\kappa^t Y^{(1,N)}_m + \kappa^{c_l-t}Y^{(2,N)}_m - \kappa^{-\lceil k_N/2\rceil}\log L_N),\end{aligned} \tag{4.73}$$

where $t = k_1+1-\lceil k_N/2\rceil$, $c_{k_N} = c_l = \lfloor l/2\rfloor - \lceil l/2\rceil + 1 = \mathbf{1}_{\{l \text{ is even}\}}$. We further rewrite, using the definition of a_N in Theorem 4.1,

$$\kappa^{-\lceil k_N/2\rceil}\log L_N = \kappa^{\frac{\log\log N}{\log\kappa}-\lfloor\frac{\log\log N}{\log\kappa}\rfloor-\lceil l/2\rceil}\frac{\log L_N}{\log N} = \kappa^{-a_N-\lceil l/2\rceil}\frac{\log L_N}{\log N}. \tag{4.74}$$

Calculating, for $Y^{(i,N)}_m \in [\varepsilon,\varepsilon^{-1}]$, the minimum of $\kappa^t Y^{(1,N)}_m + \kappa^{c_l-t}Y^{(2,N)}_m$, over all $t\in\mathbb{Z}$, we conclude that the argument of the minimum is contained in the interval $[\frac{1}{2}, \frac{1}{2} + \log(\varepsilon^2)/2\log\kappa]$. Hence from Lemma 4.11, for $N\to\infty$, $n = c_l \in \{0,1\}$ and on the event $E_{m,N}$,

$$\min_{-\lceil k_N/2\rceil+1\le t\le\lfloor k_N/2\rfloor}(\kappa^t Y^{(1,N)}_m + \kappa^{c_l-t}Y^{(2,N)}_m) = \min_{t\in\mathbb{Z}}(\kappa^t Y^{(1,N)}_m + \kappa^{c_l-t}Y^{(2,N)}_m). \tag{4.75}$$

We define

$$W_{m,N}(l) = \min_{t\in\mathbb{Z}}(\kappa^t Y_m^{(1,N)} + \kappa^{c_l-t} Y_m^{(2,N)}) - \kappa^{-a_N-\lceil l/2\rceil}\frac{\log L_N}{\log N}, \tag{4.76}$$

and, similarly we define $W^+_{m,N}$ and $W^-_{m,N}$, by replacing $Y_m^{(i,N)}$ by $Y_{m,+}^{(i,N)}$ and $Y_{m,-}^{(i,N)}$, respectively. The upper and lower bound in (4.72) now yield:

$$\begin{aligned}&\mathbb{E}\Big[\mathbf{1}_{E_{m,N}}\exp\big[-\lambda_N e^{\kappa^{\lceil k_N/2\rceil}W^+_{m,N}(l)}\big]\Big] + o_{N,m,\varepsilon}(1) \\ &\quad\le \mathbb{P}(\{H_N > k_N\}\cap E_{m,N}) \le \mathbb{E}\Big[\mathbf{1}_{E_{m,N}}\exp\big[-\lambda_N e^{\kappa^{\lceil k_N/2\rceil}W^-_{m,N}(l)}\big]\Big] + o_{N,m,\varepsilon}(1).\end{aligned} \tag{4.77}$$

We split

$$\mathbb{E}\Big[\mathbf{1}_{E_{m,N}}\exp\big[-\lambda_N e^{\kappa^{\lceil k_N/2\rceil}W_{m,N}(l)}\big]\Big] = \mathbb{P}(\widetilde{G}_N\cap E_{m,N}) + I_N + J_N + K_N + o_{N,m,\varepsilon}(1), \tag{4.78}$$

where for $\varepsilon > 0$,

$$\widetilde{F}_N = \widetilde{F}_N(l,\varepsilon) = \big\{W_{m,N}(l) > \varepsilon\big\}, \qquad \widetilde{G}_N = \widetilde{G}_N(l,\varepsilon) = \big\{W_{m,N}(l) < -\varepsilon\big\}, \tag{4.79}$$

and where we define

$$I_N = \mathbb{E}\Big[\exp\big[-\lambda_N e^{\kappa^{\lceil k_N/2\rceil}W_{m,N}(l)}\big]\mathbf{1}_{\widetilde{F}_N\cap E_{m,N}}\Big], \tag{4.80}$$

$$J_N = \mathbb{E}\Big[\Big(\exp\big[-\lambda_N e^{\kappa^{\lceil k_N/2\rceil}W_{m,N}(l)}\big] - 1\Big)\mathbf{1}_{\widetilde{G}_N\cap E_{m,N}}\Big], \tag{4.81}$$

$$K_N = \mathbb{E}\Big[\exp\big[-\lambda_N e^{\kappa^{\lceil k_N/2\rceil}W_{m,N}(l)}\big]\mathbf{1}_{\widetilde{F}^c_N\cap\widetilde{G}^c_N\cap E_{m,N}}\Big]. \tag{4.82}$$

The split (4.78) is correct since (using the abbreviation $\exp W$ for $\exp\big[-\lambda_N e^{\kappa^{\lceil k_N/2\rceil}W_{m,N}(l)}\big]$),

$$\begin{aligned}I_N + J_N + K_N &= \mathbb{E}\Big[\mathbf{1}_{E_{m,N}}\big[(\exp W)\{\mathbf{1}_{\widetilde{F}_N} + \mathbf{1}_{\widetilde{G}_N} + \mathbf{1}_{\widetilde{F}^c_N\cap\widetilde{G}^c_N}\} - \mathbf{1}_{\widetilde{G}_N}\big]\Big] \\ &= \mathbb{E}\big[\mathbf{1}_{E_{m,N}}[(\exp W) - \mathbf{1}_{\widetilde{G}_N}]\big] \\ &= \mathbb{E}\big[\mathbf{1}_{E_{m,N}}(\exp W)\big] - \mathbb{P}(E_{m,N}\cap\widetilde{G}_N).\end{aligned} \tag{4.83}$$

Observe that

$$\kappa^{\lceil k_N/2\rceil}W_{m,N}(l)\cdot\mathbf{1}_{\widetilde{F}_N} > \varepsilon\kappa^{\lceil k_N/2\rceil}, \qquad \kappa^{\lceil k_N/2\rceil}W_{m,N}(l)\cdot\mathbf{1}_{\widetilde{G}_N} < -\varepsilon\kappa^{\lceil k_N/2\rceil}. \tag{4.84}$$

We now show that I_N, J_N and K_N are error terms, and then prove convergence of $\mathbb{P}(E_{m,N}\cap\widetilde{G}_N)$. We start by bounding I_N. By the first bound in (4.84), for every $\varepsilon > 0$, and since $\lambda_N \ge \frac{1}{2}$,

$$\limsup_{N\to\infty} I_N \le \limsup_{N\to\infty}\exp\Big\{-\frac{1}{2}\exp\{\kappa^{\lceil k_N/2\rceil}\varepsilon\}\Big\} = 0. \tag{4.85}$$

Similarly, by the second bound in (4.84), for every $\varepsilon > 0$, and since $\lambda_N \le 4k_N$, we can bound J_N as

$$\limsup_{N\to\infty}|J_N| \le \limsup_{N\to\infty}\mathbb{E}\Big[1 - \exp\big\{-4k_N\exp\{-\kappa^{\lceil k_N/2\rceil}\varepsilon\}\big\}\Big] = 0. \tag{4.86}$$

Finally, we bound K_N by

$$K_N \le \mathbb{P}(\widetilde{F}^c_N\cap\widetilde{G}^c_N\cap E_{m,N}), \tag{4.87}$$

and apply the following lemma, whose proof is deferred to Section 5:

Lemma 4.14. *For all l,*

$$\limsup_{\varepsilon\downarrow 0}\limsup_{m\to\infty}\limsup_{N\to\infty}\mathbb{P}\big(\widetilde{F}_N(l,\varepsilon)^c\cap\widetilde{G}_N(l,\varepsilon)^c\cap E_{m,N}(\varepsilon)\big)=0.$$

The conclusion from (4.77)-(4.87) is that:

$$\mathbb{P}(\{H_N>k_N\}\cap E_{m,N})=\mathbb{P}(\widetilde{G}_N\cap E_{m,N})+o_{N,m,\varepsilon}(1). \tag{4.88}$$

To compute the main term $\mathbb{P}(\widetilde{G}_N\cap E_{m,N})$, we define

$$U_l=\min_{t\in\mathbb{Z}}(\kappa^t Y^{(1)}+\kappa^{c_l-t}Y^{(2)}), \tag{4.89}$$

and we will show that

Lemma 4.15.

$$\mathbb{P}(\widetilde{G}_N\cap E_{m,N})=\mathbb{P}\big(U_l-\kappa^{-a_N-\lceil l/2\rceil}<0, Y^{(1)}Y^{(2)}>0\big)+o_{N,m,\varepsilon}(1). \tag{4.90}$$

Proof. From the definition of $\widetilde{G}_N$,

$$\widetilde{G}_N\cap E_{m,N}=\Big\{\min_{t\in\mathbb{Z}}(\kappa^t Y_m^{(1,N)}+\kappa^{c_l-t}Y_m^{(2,N)})-\kappa^{-a_N-\lceil l/2\rceil}\frac{\log U_N}{\log N}<-\varepsilon, Y_m^{(i,N)}\in[\varepsilon,\varepsilon^{-1}]\Big\}. \tag{4.91}$$

By Proposition 3.1 and the fact that $L_N=\mu N(1+o(1))$,

$$\mathbb{P}(\widetilde{G}_N\cap E_{m,N})-\mathbb{P}\Big(\min_{t\in\mathbb{Z}}(\kappa^t Y_m^{(1)}+\kappa^{c_l-t}Y_m^{(2)})-\kappa^{-a_N-\lceil l/2\rceil}<-\varepsilon, Y_m^{(i)}\in[\varepsilon,\varepsilon^{-1}]\Big)=o_{N,m,\varepsilon}(1). \tag{4.92}$$

Since $Y_m^{(i)}$ converges to $Y^{(i)}$ almost surely, as $m\to\infty$, $\sup_{s\geq m}|Y_s^{(i)}-Y^{(i)}|$ converges to 0 a.s. as $m\to\infty$. Therefore,

$$\mathbb{P}(\widetilde{G}_N\cap E_{m,N})-\mathbb{P}\Big(U_l-\kappa^{-a_N-\lceil l/2\rceil}<-\varepsilon, Y^{(i)}\in[\varepsilon,\varepsilon^{-1}]\Big)=o_{N,m,\varepsilon}(1), \tag{4.93}$$

Moreover, since $Y^{(1)}$ has a density on $(0,\infty)$ and an atom at 0 (see (8)),

$$\mathbb{P}(Y^{(1)}\notin[\varepsilon,\varepsilon^{-1}], Y^{(1)}>0)=o(1), \qquad \text{as } \varepsilon\downarrow 0.$$

Recall from Section 3.1 that for any l fixed, and conditionally on $Y^{(1)}Y^{(2)}>0$, the random variable U_l has a density. We denote this density by f_2 and the distribution function by F_2. Also, $\kappa^{-a_N-\lceil l/2\rceil}\in I_l=[\kappa^{-\lceil l/2\rceil},\kappa^{-\lceil l/2\rceil+1}]$. Then,

$$\mathbb{P}\Big(-\varepsilon\leq U_l-\kappa^{-a_N-\lceil l/2\rceil}<0\Big)\leq\sup_{a\in I_l}[F_2(a)-F_2(a-\varepsilon)]. \tag{4.94}$$

The function F_2 is continuous on I_l, so that in fact F_2 is uniformly continuous on I_l, and we conclude that

$$\limsup_{\varepsilon\downarrow 0}\sup_{a\in I_l}[F_2(a)-F_2(a-\varepsilon)]=0. \tag{4.95}$$

This establishes (4.90). □

We summarize the results obtained sofar in the following corollary:

Corollary 4.16. *For all l, with k_N as in (4.5),*

$$\mathbb{P}\big(\{H_N > k_N\} \cap E_{m,N}\big) = q^2\mathbb{P}\Big(U_l \le \kappa^{-a_N - \lceil l/2 \rceil}\Big| Y^{(1)}Y^{(2)} > 0\Big) + o_{N,m,\varepsilon}(1).$$

Proof. By independence of $Y^{(1)}$ and $Y^{(2)}$, we obtain $\mathbb{P}(Y^{(1)}Y^{(2)} > 0) = q^2$. Combining this with (4.88) and (4.90), yields

$$\begin{aligned}\mathbb{P}(\{H_N > k_N\} \cap E_{m,N}) &= \mathbb{P}\Big(U_l - \kappa^{-a_N - \lceil l/2 \rceil} < 0, Y^{(1)}Y^{(2)} > 0\Big) + o_{N,m,\varepsilon}(1) \\ &= q^2\mathbb{P}\Big(U_l - \kappa^{-a_N - \lceil l/2 \rceil} < 0\Big| Y^{(1)}Y^{(2)} > 0\Big) + o_{N,m,\varepsilon}(1). \end{aligned} \tag{4.96}$$

Note that the change from $U_l < \kappa^{-a_N - \lceil l/2 \rceil}$ to $U_l \le \kappa^{-a_N - \lceil l/2 \rceil}$ is allowed because F_2 admits a density. □

We now come to the conclusion of the proof of Theorem 4.1. Corollary 4.16 yields, with $m_{\tau,N} = 2\left\lfloor \frac{\log\log N}{|\log(\tau-2)|}\right\rfloor$, so that $k_N = m_{\tau,N} + l$,

$$\begin{aligned}\mathbb{P}\left(H_N > m_{\tau,N} + l\right) &= \mathbb{P}(\{H_N > k_N\} \cap E_{m,N}) + \mathbb{P}(\{H_N > k_N\} \cap E^c_{m,N}) \\ &= q^2\mathbb{P}\left(U_l \le \kappa^{-a_N - \lceil l/2 \rceil}\Big| Y^{(1)}Y^{(2)} > 0\right) + 1 - q^2 + o_{N,m,\varepsilon}(1),\end{aligned}$$

because

$$\begin{aligned}&\mathbb{P}(\{H_N > k_N\} \cap E^c_{m,N}) \\ &\quad= \mathbb{P}(\{H_N > k_N\} \cap E^c_{m,N} \cap \{Y^{(1,N)}_m Y^{(2,N)}_m = 0\}) + \mathbb{P}(\{H_N > k_N\} \cap E^c_{m,N} \cap \{Y^{(1,N)}_m Y^{(2,N)}_m > 0\}) \\ &\quad= \mathbb{P}(\{H_N > k_N\} \cap \{Y^{(1,N)}_m Y^{(2,N)}_m = 0\}) + \mathbb{P}(\{H_N > k_N\} \cap E^c_{m,N} \cap \{Y^{(1,N)}_m Y^{(2,N)}_m > 0\}) \\ &\quad= \lim_{\varepsilon\downarrow 0}\lim_{m\to\infty}(1 - q_m^2) + o_{N,m,\varepsilon}(1) = (1 - q^2) + o_{N,m,\varepsilon}(1),\end{aligned}$$

where the second equality follows from $\{Y^{(1,N)}_m Y^{(2,N)}_m = 0\} \subseteq E^c_{m,N}$, and the one but final equality from Corollary 4.5 and (4.17), respectively.

Taking complementary events, we obtain,

$$\mathbb{P}\left(H_N \le m_{\tau,N} + l\right) = q^2\mathbb{P}\left(U_l > \kappa^{-a_N - \lceil l/2 \rceil}\big| Y^{(1)}Y^{(2)} > 0\right) + o_{N,m,\varepsilon}(1).$$

Note that in the above equation the terms, except the error $o_{N,m,\varepsilon}(1)$, are independent of m and ε, so that, in fact, we have, for $N \to \infty$,

$$\mathbb{P}\left(H_N \le m_{\tau,N} + l\right) = q^2\mathbb{P}\left(U_l > \kappa^{-a_N - \lceil l/2 \rceil}\big| Y^{(1)}Y^{(2)} > 0\right) + o(1). \tag{4.97}$$

We claim that (4.97) implies that, when $N \to \infty$,

$$\mathbb{P}\left(H_N < \infty\right) = q^2 + o(1). \tag{4.98}$$

Indeed, to see (4.98), we prove upper and lower bounds. For the lower bound, we use that for any $l \in \mathbb{Z}$

$$\mathbb{P}\left(H_N < \infty\right) \ge \mathbb{P}\left(H_N \le k_{\tau,N} + l\right),$$

and let $l \to \infty$ in (4.97), noting that $\kappa^{-a_N - \lceil l/2 \rceil} \to 0$ as $l \to \infty$. For the upper bound, we split

$$\mathbb{P}(H_N < \infty) = \mathbb{P}\left(\{H_N < \infty\} \cap \{Y_m^{(1,N)} Y_m^{(2,N)} = 0\}\right) + \mathbb{P}\left(\{H_N < \infty\} \cap \{Y_m^{(1,N)} Y_m^{(2,N)} > 0\}\right).$$

For $N \to \infty$, the first term is bounded by $\mathbb{P}(H_N \leq m-1) = o(1)$, by Lemma 4.4. Similarly as $N \to \infty$, the second term is bounded from above by, using Proposition 3.1,

$$\mathbb{P}\left(\{H_N < \infty\} \cap \{Y_m^{(1,N)} Y_m^{(2,N)} > 0\}\right) \leq \mathbb{P}\left(Y_m^{(1,N)} Y_m^{(2,N)} > 0\right) = q_m^2 + o(1), \tag{4.99}$$

which converges to q^2 as $m \to \infty$. This proves (4.98). We conclude from (4.97) and (4.98) that for $N \to \infty$,

$$\mathbb{P}\left(H_N \leq m_{\tau,N} + l \,\middle|\, H_N < \infty\right) = \mathbb{P}\left(U_l > \kappa^{-a_N - \lceil l/2 \rceil} \,\middle|\, Y^{(1)} Y^{(2)} > 0\right) + o(1). \tag{4.100}$$

Substituting $\kappa = (\tau - 2)^{-1}$, and taking complements in (4.100) this yields the claims in Theorem 4.1. $\square$

5 Proofs of Lemmas 4.9, 4.13 and 4.14

In this section, we prove the three lemmas used in Section 4. The proofs are similar in nature. Denote

$$\{k \in \partial \mathcal{T}_m^{(i,N)}\} = \{k \in \mathcal{T}_m^{(i,N)}\} \cap \{k+1 \notin \mathcal{T}_m^{(i,N)}\}. \tag{5.1}$$

We will make essential use of the following consequences of Propositions 3.1 and 3.2:

Lemma 5.1. *For any $u > 0$, and $i = 1, 2$,*

$$(i) \quad \mathbb{P}\Big(\{k \in \mathcal{T}_m^{(i,N)}\} \cap E_{m,N} \cap \{Z_k^{(i,N)} \in [N^{u(1-\varepsilon)}, N^{u(1+\varepsilon)}]\}\Big) = o_{N,m,\varepsilon}(1), \tag{5.2}$$

$$(ii) \quad \mathbb{P}\Big(\{k \in \partial \mathcal{T}_m^{(i,N)}\} \cap E_{m,N} \cap \{Z_k^{(i,N)} \leq N^{\frac{1}{\kappa(\tau-1)} + \varepsilon}\}\Big) = o_{N,m,\varepsilon}(1). \tag{5.3}$$

Proof. We start with the proof of (i). In the course of this proof the statement **whp** means that the complement of the involved event has probability $o_{N,m,\varepsilon}(1)$. By Proposition 3.2, we have **whp**, for $k \in \mathcal{T}_m^{(i,N)}$, and on the event $E_{m,N}$, that

$$Y_k^{(i,N)} \leq Y_m^{(i,N)} + \varepsilon^3 \leq Y_m^{(i,N)}(1 + \varepsilon^2), \tag{5.4}$$

where the last inequality follows from $Y_m^{(i,N)} \geq \varepsilon$. Therefore, also

$$Y_m^{(i,N)} \geq Y_k^{(i,N)}(1 - 2\varepsilon^2), \tag{5.5}$$

when ε is so small that $(1 + \varepsilon^2)^{-1} \geq 1 - 2\varepsilon^2$. In a similar way, we conclude that with $k \in \mathcal{T}_m^{(i,N)}$, and on the event $E_{m,N}$,

$$Y_m^{(i,N)} \leq Y_k^{(i,N)}(1 + 2\varepsilon^2). \tag{5.6}$$

Furthermore, the event $Z_k^{(i,N)} \in [N^{u(1-\varepsilon)}, N^{u(1+\varepsilon)}]$ is equivalent to

$$\kappa^{-k} u(1-\varepsilon) \log N \leq Y_k^{(i,N)} \leq \kappa^{-k} u(1+\varepsilon) \log N. \tag{5.7}$$

Therefore, we obtain that, with $u_{k,N} = u\kappa^{-k}\log N$,

$$Y_m^{(i,N)} \leq (1+2\varepsilon^2)(1+\varepsilon)u\kappa^{-k}\log N \leq (1+2\varepsilon)u_{k,N}, \tag{5.8}$$

and similarly,

$$Y_m^{(i,N)} \geq (1-2\varepsilon^2)(1-\varepsilon)u\kappa^{-k}\log N \geq (1-2\varepsilon)u_{k,N}. \tag{5.9}$$

We conclude that **whp** the events $k \in \mathcal{T}_m^{(i,N)}, \varepsilon \leq Y_m^{(i,N)} \leq \varepsilon^{-1}$ and $Z_k^{(i,N)} \in [N^{u(1-\varepsilon)}, N^{u(1+\varepsilon)}]$ imply

$$Y_m^{(i,N)} \in u_{k,N}[1-2\varepsilon, 1+2\varepsilon] \equiv [u_{k,N}(1-2\varepsilon), u_{k,N}(1+2\varepsilon)]. \tag{5.10}$$

Since $\varepsilon \leq Y_m^{(i,N)} \leq \varepsilon^{-1}$, we therefore must also have (when ε is so small that $1-2\varepsilon \geq \frac{1}{2}$),

$$u_{k,N} \in [\frac{\varepsilon}{2}, \frac{2}{\varepsilon}]. \tag{5.11}$$

Therefore,

$$\begin{aligned}&\mathbb{P}\Big(\{k \in \mathcal{T}_m^{(i,N)}\} \cap E_{m,N} \cap \{Z_k^{(i,N)} \in [N^{u(1-\varepsilon)}, N^{u(1+\varepsilon)}]\}\Big) \\ &\qquad \leq \sup_{x \in [\frac{\varepsilon}{2}, \frac{2}{\varepsilon}]} \mathbb{P}\big(Y_m^{(i,N)} \in x[1-2\varepsilon, 1+2\varepsilon]\big) + o_{N,m,\varepsilon}(1).\end{aligned} \tag{5.12}$$

Since, for $N \to \infty$, $Y_m^{(i,N)} = Y_m^{(i)}$ in probability, by Proposition 3.1, we arrive at

$$\begin{aligned}&\limsup_{\varepsilon \downarrow 0} \limsup_{m\to\infty} \limsup_{N\to\infty} \mathbb{P}\big(\{k \in \mathcal{T}_m^{(i,N)}\} \cap E_{m,N} \cap \{Z_k^{(i,N)} \in [N^{u(1-\varepsilon)}, N^{u(1+\varepsilon)}]\}\big) \\ &\qquad \leq \limsup_{\varepsilon \downarrow 0} \limsup_{m\to\infty} \sup_{x \in [\frac{\varepsilon}{2}, \frac{2}{\varepsilon}]} \mathbb{P}\big(Y_m^{(i)} \in x[1-2\varepsilon, 1+2\varepsilon]\big).\end{aligned} \tag{5.13}$$

We next use that $Y_m^{(i)}$ converges to $Y^{(i)}$ almost surely as $m \to \infty$ to arrive at

$$\begin{aligned}&\limsup_{m\to\infty} \limsup_{N\to\infty} \mathbb{P}\big(\{k \in \mathcal{T}_m^{(i,N)}\} \cap E_{m,N} \cap \{Z_k^{(i,N)} \in [N^{u(1-\varepsilon)}, N^{u(1+\varepsilon)}]\}\big) \\ &\qquad \leq \sup_{x \in [\frac{\varepsilon}{2}, \frac{2}{\varepsilon}]} \mathbb{P}\big(Y^{(i)} \in x[1-2\varepsilon, 1+2\varepsilon]\big) \leq \sup_{x>0} [F_1(x(1+2\varepsilon)) - F_1(x(1-2\varepsilon))],\end{aligned} \tag{5.14}$$

where F_1 denotes the distribution function of $Y^{(i)}$. Since F_1 is a proper distribution function on $[0, \infty)$, with an atom at 0 and admitting a density on $(0, \infty)$, we have

$$\limsup_{\varepsilon \downarrow 0} \sup_{x>0} [F_1(x(1+2\varepsilon)) - F_1(x(1-2\varepsilon))] = 0. \tag{5.15}$$

This is immediate from uniform continuity of F_1 on each bounded subinterval of $(0, \infty)$ and by the fact that $F_1(\infty) = 1$. The upper bound (5.14) together with (5.15) completes the proof of the first statement of the lemma.

We turn to the statement (ii). The event that $k \in \partial\mathcal{T}_m^{(i,N)}$ implies

$$Y_m^{(i,N)} \geq \frac{1-\varepsilon^2}{\tau-1}\kappa^{-(k+1)}\log N. \tag{5.16}$$

By (5.6), we can therefore conclude that **whp**, for $k \in \partial\mathcal{T}_m^{(i,N)}$, and on the event $E_{m,N}$, that, for ε so small that $(1-\varepsilon^2)/(1+2\varepsilon^2) \geq 1-\varepsilon$,

$$Y_k^{(i,N)} \geq \frac{1-\varepsilon}{\tau-1}\kappa^{-(k+1)}\log N, \tag{5.17}$$

which is equivalent to

$$Z_k^{(i,N)} \geq N^{\frac{1-\varepsilon}{\kappa(\tau-1)}} \geq N^{\frac{1}{\kappa(\tau-1)}-\varepsilon}. \tag{5.18}$$

Therefore,

$$\limsup_{\varepsilon\downarrow 0}\limsup_{m\to\infty}\limsup_{N\to\infty}\mathbb{P}\big(\{k \in \partial\mathcal{T}_m^{(i,N)}\} \cap E_{m,N} \cap \{Z_k^{(i,N)} \leq N^{\frac{1}{\kappa(\tau-1)}+\varepsilon}\}\big) \tag{5.19}$$
$$\leq \limsup_{\varepsilon\downarrow 0}\limsup_{m\to\infty}\limsup_{N\to\infty}\mathbb{P}\big(\{k \in \mathcal{T}_m^{(i,N)}\} \cap E_{m,N} \cap \{Z_k^{(i,N)} \in [N^{\frac{1}{\kappa(\tau-1)}-\varepsilon}, N^{\frac{1}{\kappa(\tau-1)}+\varepsilon}]\}\big) = 0,$$

which follows from the first statement in Lemma 5.1 with $u = \frac{1}{\kappa(\tau-1)}$. □

Proof of Lemma 4.9. By (4.20), it suffices to prove that

$$\mathbb{P}(\{H_N > k_N\} \cap E_{m,N} \cap F_{m,N} \cap \{\mathcal{B}_N(\varepsilon, k_N) = \varnothing\}) = o_{N,m,\varepsilon}(1), \tag{5.20}$$

which shows that in considering the event $\{H_N > k_N\} \cap E_{m,N} \cap F_{m,N}$, we may assume that $\mathcal{B}_N(\varepsilon, k_N) \neq \varnothing$.

Observe that if $\mathcal{B}_N(\varepsilon, k) = \varnothing$, then $\mathcal{B}_N(\varepsilon, k+1) = \varnothing$. Indeed, if $l \in \mathcal{B}_N(\varepsilon, k+1)$ and $l \neq m$, then $l-1 \in \mathcal{B}_N(\varepsilon, k)$. If, on the other hand, $\mathcal{B}_N(\varepsilon, k+1) = \{m\}$, then also $m \in \mathcal{B}_N(\varepsilon, k)$. We conclude that the random variable,

$$l^* = \sup\{k : \mathcal{B}_N(\varepsilon, k) \neq \varnothing\}. \tag{5.21}$$

is well defined.

Hence $\{\mathcal{B}_N(\varepsilon, k_N) = \varnothing\} = \{k_N \geq l^*+1\}$ and we therefore have

$$\{\mathcal{B}_N(\varepsilon, k_N) = \varnothing\} = \{l^* < k_N\} = \{l^* \leq k_N - 2\} \,\dot\cup\{l^* = k_N - 1\}. \tag{5.22}$$

We deal with each of the two events separately. We start with the first.

Since the sets $\mathcal{B}_N(\varepsilon, k)$ are Z_m-measurable, we obtain, as in (4.32),

$$\mathbb{P}(\{H_N > k_N\} \cap E_{m,N} \cap F_{m,N} \cap \{l^* \leq k_N - 2\}) \leq \mathbb{P}(\{H_N > l^*+2\} \cap E_{m,N} \cap F_{m,N}) \tag{5.23}$$
$$= \mathbb{E}\Big[\mathbf{1}_{E_{m,N}\cap F_{m,N}} P_m(l^*+2, k_1)\Big] + o_{N,m,\varepsilon}(1).$$

We then use (4.41) to bound

$$P_m(l^*+2, k_1) \leq \exp\left\{-\frac{Z_{k_1+1}^{(1,N)} Z_{l^*+2-k_1}^{(2,N)}}{2L_N}\right\}. \tag{5.24}$$

Now, since $\mathcal{B}_N(\varepsilon, l^*) \neq \varnothing$, we can pick k_1 such that $k_1 - 1 \in \mathcal{B}_N(\varepsilon, l^*)$. Since $\mathcal{B}_N(\varepsilon, l^*+1) = \varnothing$, we have $k_1 - 1 \notin \mathcal{B}_N(\varepsilon, l^*+1)$, implying $l^*+1-k_1 \in \mathcal{T}_m^{(2,N)}$ and $l^*+2-k_1 \notin \mathcal{T}_m^{(2,N)}$ so that, by (3.7), $Z_{l^*+2-k_1}^{(2,N)} \geq N^{\frac{1-\varepsilon}{\tau-1}}$.

Similarly, since $k_1 \notin \mathcal{B}_N(\varepsilon, l^*+1)$ we have that $k_1 \in \mathcal{T}_m^{(1,N)}$ and $k_1+1 \notin \mathcal{T}_m^{(1,N)}$, so that, again by (3.7), $Z^{(1,N)}_{k_1+1} \geq N^{\frac{1-\varepsilon}{\tau-1}}$. Therefore, since $L_N \geq N$, **whp**,

$$\frac{Z^{(1,N)}_{k_1+1} Z^{(2,N)}_{l^*+2-k_1}}{L_N} \geq N^{\frac{2(1-\varepsilon)}{\tau-1}-1}, \tag{5.25}$$

and the exponent of N is strictly positive for $\tau \in (2,3)$ and $\varepsilon > 0$ small enough. This bounds the contribution in (5.23) due to $\{l^* \leq k_N - 2\}$.

We proceed with the contribution due to $\{l^* = k_N - 1\}$. In this case, there exists a k_1 with $k_1 - 1 \in \mathcal{B}_N(\varepsilon, k_N - 1)$, so that $k_1 \in \mathcal{T}_m^{(1,N)}$ and $k_N - k_1 \in \mathcal{T}_m^{(2,N)}$. On the other hand, $\mathcal{B}_N(\varepsilon, k_N) = \varnothing$, which together with $k_1 - 1 \in \mathcal{B}_N(\varepsilon, k_N - 1)$ implies that $k_N - k_1 \in \mathcal{T}_m^{(2,N)}$, and $k_N - k_1 + 1 \notin \mathcal{T}_m^{(2,N)}$. Similarly, we obtain that $k_1 \in \mathcal{T}_m^{(1,N)}$ and $k_1 + 1 \notin \mathcal{T}_m^{(1,N)}$. Using Proposition 3.4, we conclude that, **whp**, $Z^{(1,N)}_{k_1+1} \geq N^{\frac{1-\varepsilon}{\tau-1}}$.

We now distinguish two possibilities: (a) $Z^{(2,N)}_{k_N-k_1} \leq N^{\frac{\tau-2}{\tau-1}+\varepsilon}$; and (b) $Z^{(2,N)}_{k_N-k_1} > N^{\frac{\tau-2}{\tau-1}+\varepsilon}$. By (5.3) and the fact that $k_N - k_1 \in \partial\mathcal{T}_m^{(2,N)}$, case (a) has small probability, so we need to investigate case (b) only.

In case (b), we can write

$$\begin{aligned}
&\mathbb{P}\Big(\{H_N > k_N\} \cap E_{m,N} \cap F_{m,N} \cap \{l^* = k_N - 1\} \cap \{Z^{(2,N)}_{k_N-k_1} > N^{\frac{\tau-2}{\tau-1}+\varepsilon}\}\Big) \\
&\quad = \mathbb{E}\Big[\mathbf{1}_{E_{m,N}\cap F_{m,N}\cap\{l^*=k_N-1\}} \mathbf{1}_{\{Z^{(2,N)}_{k_N-k_1} > N^{\frac{\tau-2}{\tau-1}+\varepsilon}\}\cap\{k_1-1\in\mathcal{B}_N(\varepsilon,k_N-1)\}} P_m(k_N, k_1)\Big] + o_{N,m,\varepsilon}(1),
\end{aligned} \tag{5.26}$$

where according to (4.41), we can bound

$$P_m(k_N, k_1) \leq \exp\left\{-\frac{Z^{(1,N)}_{k_1+1} Z^{(2,N)}_{k_N-k_1}}{2L_N}\right\}. \tag{5.27}$$

We note that by Proposition 3.4 and similarly to (5.25),

$$\frac{Z^{(1,N)}_{k_1+1} Z^{(2,N)}_{k_N-k_1}}{L_N} \geq N^{\frac{1-\varepsilon}{\tau-1}} N^{\frac{\tau-2}{\tau-1}+\varepsilon-1} = N^{\left(1-\frac{1}{\tau-1}\right)\varepsilon}, \tag{5.28}$$

and again the exponent is strictly positive, so that, following the arguments in (5.23–5.27), we obtain that also the contribution due to case (b) is small. □

Proof of Lemma 4.13. Recall that we have defined $x = x(t) = \kappa^t Y^{(1,N)}_{m,+}$ and $y = y(t) = \kappa^{n-t} Y^{(2,N)}_{m,+}$, with $n = k_N + 1$, and that $x \geq y$. The event $\mathcal{E}_{m,N}$ in (4.71) is equal to the existence of a t such that,

$$\frac{1-\varepsilon}{\tau-1}\log N \leq x \leq \frac{1+\varepsilon}{\tau-1}\log N, \qquad \text{and} \qquad x + y \leq (1+\varepsilon)\log N. \tag{5.29}$$

Therefore, by (4.66),

$$y \geq \frac{x}{\kappa} \geq (1-\varepsilon)\frac{\tau-2}{\tau-1}\log N. \tag{5.30}$$

On the other hand, by the bounds in (5.29),

$$y \leq (1+\varepsilon)\log N - x \leq (1+\varepsilon)\log N - \frac{1-\varepsilon}{\tau-1}\log N = \big(1+\varepsilon\frac{\tau}{\tau-2}\big)\frac{\tau-2}{\tau-1}\log N. \tag{5.31}$$

Therefore, by multiplying the bounds on x and y, we obtain

$$(1-\varepsilon)^2\frac{\tau-2}{(\tau-1)^2}\log^2 N \leq \kappa^{k_N+1}Y^{(1,N)}_{m,+}Y^{(2,N)}_{m,+} \leq \big(1+\varepsilon\frac{\tau}{\tau-2}\big)(1+\varepsilon)\frac{\tau-2}{(\tau-1)^2}\log^2 N, \tag{5.32}$$

and thus

$$\mathbb{P}(E_{m,N}\cap \mathcal{E}_{m,N}\cap\{H_N > k_N\}) \leq \mathbb{P}\Big((1-\varepsilon)^2 \leq \frac{\kappa^{k_N+1}}{c\log^2 N}Y^{(1,N)}_{m,+}Y^{(2,N)}_{m,+} \leq \big(1+\varepsilon\frac{\tau}{\tau-2}\big)(1+\varepsilon)\Big), \tag{5.33}$$

where we abbreviate $c = \frac{\tau-2}{(\tau-1)^2}$. Since $\frac{\kappa^{k_N+1}}{c\log^2 N}$ is bounded away from 0 and ∞, we conclude that the right-hand side of (5.33) is $o_{N,m,\varepsilon}(1)$, analogously to the final part of the proof of Lemma 5.1(i). $\square$

Proof of Lemma 4.14. We recall that $U_l = \min_{t\in\mathbb{Z}}(\kappa^t Y^{(1)} + \kappa^{c_l-t}Y^{(2)})$, and repeat the arguments leading to (4.92–4.96) to see that, as first $N\to\infty$ and then $m\to\infty$,

$$\begin{aligned}\mathbb{P}(\widetilde{F}^c_N \cap \widetilde{G}^c_N \cap E_{m,N}) &\leq \mathbb{P}\left(-\varepsilon \leq U_l - \kappa^{-a_N-\lceil l/2\rceil} \leq \varepsilon, Y^{(1)}Y^{(2)}>0\right) + o_{N,m}(1) \qquad (5.34)\\ &= q^2\mathbb{P}\left(-\varepsilon \leq U_l - \kappa^{-a_N-\lceil l/2\rceil} \leq \varepsilon \Big| Y^{(1)}Y^{(2)}>0\right) + o_{N,m}(1).\end{aligned}$$

Recall from Section 3.1 that, conditionally on $Y^{(1)}Y^{(2)}>0$, the random variable U_l has a density, and that we denoted the distribution function of U_l given $Y^{(1)}Y^{(2)}>0$ by F_2. Furthermore, $\kappa^{-a_N-\lceil l/2\rceil} \in I_l = [\kappa^{-\lceil l/2\rceil}, \kappa^{-\lceil l/2\rceil+1}]$, so that, uniformly in N,

$$\mathbb{P}\left(-\varepsilon \leq U_l - \kappa^{-a_N-\lceil l/2\rceil} \leq \varepsilon \Big| Y^{(1)}Y^{(2)}>0\right) \leq \sup_{u\in I_l}[F_2(u+\varepsilon) - F_2(u-\varepsilon)] = 0,$$

where the conclusion follows by repeating the argument leading to (5.15). This completes the proof of Lemma 4.14. $\square$

A Proof of Propositions 3.1, 3.2 and 3.4

The appendix is organized as follows. In Section A.1, we prove three lemmas that are used in Section A.2 to prove Proposition 3.1. In Section A.3, we continue with preparations for the proofs of Proposition 3.2 and 3.4. In this section we formulate key Proposition A.3.2, which will be proved in Section A.4. In Section A.5, we end the appendix with the proofs of Proposition 3.2 and 3.4. As in the main body of the paper, we will assume throughout the appendix that $\tau\in(2,3)$, so if we refer to Assumption 1.1, we mean Assumption 1.1(ii).

A.1 Some preparatory lemmas

In order to prove Proposition 3.1, we make essential use of three lemmas, that also play a key role in Section A.4 below. The first of these three lemmas investigates the tail behaviour of $1 - G(x)$ under Assumption 1.1. Recall that G is the distribution function of the probability mass function $\{g_j\}$, defined in (1.12).

Lemma A.1.1. *If F satisfies Assumption 1.1(ii) then there exists $K_\tau > 0$ such that for x large enough*

$$x^{2-\tau-K_\tau\gamma(x)} \leq 1 - G(x) \leq x^{2-\tau+K_\tau\gamma(x)}, \tag{A.1.1}$$

where $\gamma(x) = (\log x)^{\gamma-1}$, $\gamma \in [0, 1)$.

Proof. Using (1.12) we rewrite $1 - G(x)$ as

$$1 - G(x) = \sum_{j=x+1}^{\infty} \frac{(j+1)f_{j+1}}{\mu} = \frac{1}{\mu}\left[(x+2)\left[1 - F(x+1)\right] + \sum_{j=x+2}^{\infty}\left[1 - F(j)\right]\right].$$

Then we use (13, Theorem 1, p. 281), together with the fact that $1 - F(x)$ is regularly varying with exponent $1 - \tau \neq 1$, to deduce that there exists a constant $c = c_\tau > 0$ such that

$$\sum_{j=x+2}^{\infty}\left[1 - F(j)\right] \leq c_\tau(x+2)\left[1 - F(x+2)\right].$$

Hence, if F satisfies Assumption 1.1(ii), then

$$\begin{aligned} 1 - G(x) &\geq \tfrac{1}{\mu}(x+2)\left[1 - F(x+1)\right] \geq x^{2-\tau-K_\tau\gamma(x)}, \\ 1 - G(x) &\leq \tfrac{1}{\mu}(c+1)(x+2)\left[1 - F(x+1)\right] \leq x^{2-\tau+K_\tau\gamma(x)}, \end{aligned}$$

for some $K_\tau > 0$ and large enough x. □

Remark A.1.2. *It follows from Assumption 1.1(ii) and Lemma A.1.1, that for each $\varepsilon > 0$ and sufficiently large x,*

$$\begin{aligned} x^{1-\tau-\varepsilon} &\leq 1 - F(x) \leq x^{1-\tau+\varepsilon}, \qquad (a) \\ x^{2-\tau-\varepsilon} &\leq 1 - G(x) \leq x^{2-\tau+\varepsilon}. \qquad (b) \end{aligned} \tag{A.1.2}$$

We will often use (A.1.2) with ε replaced by ε^6. □

Let us define for $\varepsilon > 0$,

$$\alpha = \frac{1 - \varepsilon^5}{\tau - 1}, \qquad h = \varepsilon^6, \tag{A.1.3}$$

and the auxiliary event F_ε by

$$F_\varepsilon = \{\forall 1 \leq x \leq N^\alpha : |G(x) - G^{(N)}(x)| \leq N^{-h}[1 - G(x)]\}, \tag{A.1.4}$$

where $G^{(N)}$ is the (random) distribution function of $\{g_n^{(N)}\}$, defined in (2.6).

Lemma A.1.3. *For ε small enough, and N sufficiently large,*

$$\mathbb{P}(F_\varepsilon^c) \le N^{-h}. \tag{A.1.5}$$

Proof. First, we rewrite $1 - G^{(N)}(x)$, for $x \in \mathbb{N} \cup \{0\}$, in the following way:

$$\begin{aligned} 1 - G^{(N)}(x) &= \sum_{n=x+1}^{\infty} g_n^{(N)} = \frac{1}{L_N} \sum_{j=1}^{N} \sum_{n=x+1}^{\infty} D_j \mathbf{1}_{\{D_j = n+1\}} = \frac{1}{L_N} \sum_{j=1}^{N} D_j \mathbf{1}_{\{D_j \ge x+2\}} \\ &= \frac{1}{L_N} \sum_{j=1}^{N} \sum_{l=1}^{D_j} \mathbf{1}_{\{D_j \ge x+2\}} = \frac{1}{L_N} \sum_{l=1}^{\infty} \sum_{j=1}^{N} \mathbf{1}_{\{D_j \ge (x+2)\vee l\}}. \end{aligned} \tag{A.1.6}$$

Writing

$$B_y^{(N)} = \sum_{j=1}^{N} \mathbf{1}_{\{D_j \ge y\}}, \tag{A.1.7}$$

we thus end up with

$$1 - G^{(N)}(x) = \frac{1}{L_N} \sum_{l=1}^{\infty} B^{(N)}_{(x+2)\vee l}. \tag{A.1.8}$$

We have a similar expression for $1 - G(x)$ that reads

$$1 - G(x) = \frac{1}{\mu} \sum_{l=1}^{\infty} \mathbb{P}(D_1 \ge (x+2) \vee l). \tag{A.1.9}$$

Therefore, with

$$\beta = \frac{1-h}{\tau-1}, \qquad \text{and} \qquad \chi = \frac{1+2h}{\tau-1},$$

we can write

$$\begin{aligned} [G(x) - G^{(N)}(x)] \; &= \left(\tfrac{N\mu}{L_N} - 1\right)[1 - G(x)] \\ &+ \tfrac{1}{L_N} \textstyle\sum_{l=1}^{N^\beta} \left[B^{(N)}_{(x+2)\vee l} - N\mathbb{P}\big(D_1 \ge (x+2)\vee l\big)\right] \\ &+ \tfrac{1}{L_N} \textstyle\sum_{l=N^\beta+1}^{N^\chi} \left[B^{(N)}_{(x+2)\vee l} - N\mathbb{P}\big(D_1 \ge (x+2)\vee l\big)\right] \\ &+ \tfrac{1}{L_N} \textstyle\sum_{l=N^\chi+1}^{\infty} \left[B^{(N)}_{(x+2)\vee l} - N\mathbb{P}\big(D_1 \ge (x+2)\vee l\big)\right]. \end{aligned} \tag{A.1.10}$$

Hence, for large enough N and $x \le N^\alpha < N^\beta < N^\chi$, we can bound

$$\begin{aligned} R_N(x) \equiv \left|G(x) - G^{(N)}(x)\right| \; &\le \left|\tfrac{N\mu}{L_N} - 1\right| [1 - G(x)] && (a) \\ &+ \tfrac{1}{L_N} \textstyle\sum_{l=1}^{N^\beta} \left|B^{(N)}_{(x+2)\vee l} - N\mathbb{P}\big(D_1 \ge (x+2)\vee l\big)\right| && (b) \\ &+ \tfrac{1}{L_N} \textstyle\sum_{l=N^\beta+1}^{N^\chi} \left|B^{(N)}_{l} - N\mathbb{P}\big(D_1 \ge l\big)\right| && (c) \\ &+ \tfrac{1}{L_N} \textstyle\sum_{l=N^\chi+1}^{\infty} B^{(N)}_{l} && (d) \\ &+ \tfrac{1}{L_N} \textstyle\sum_{l=N^\chi+1}^{\infty} N\mathbb{P}\big(D_1 \ge l\big). && (e) \end{aligned}$$

(A.1.11)

We use (A.1.2(b)) to conclude that, in order to prove $\mathbb{P}(F_\varepsilon^c) \leq N^{-h}$, it suffices to show that

$$\mathbb{P}\left(\bigcup_{1\leq x\leq N^\alpha} \left\{|R_N(x)| > C_g N^{-h} x^{2-\tau-h}\right\}\right) \leq N^{-h}, \tag{A.1.12}$$

for large enough N, and for some C_g, depending on distribution function G. We will define an auxiliary event $A_{N,\varepsilon}$, such that $|R_N(x)|$ is more easy to bound on $A_{N,\varepsilon}$ and such that $\mathbb{P}(A_{N,\varepsilon}^c)$ is sufficiently small. Indeed, we define, with $A = 3(\beta + 2h)$,

$$\begin{aligned}
A_{N,\varepsilon}(a) &= \left\{|\tfrac{N\mu}{L_N} - 1| \leq N^{-3h}\right\}, && (a)\\
A_{N,\varepsilon}(b) &= \{\max_{1\leq j\leq N} D_j \leq N^\chi\}, && (b)\\
A_{N,\varepsilon}(c) &= \bigcap_{1\leq x\leq N^\beta} \left\{|B_x^{(N)} - N\mathbb{P}(D_1 \geq x)| \leq \sqrt{A(\log N)N\mathbb{P}(D_1 \geq x)}\right\}, && (c)
\end{aligned} \tag{A.1.13}$$

and

$$A_{N,\varepsilon} = A_{N,\varepsilon}(a) \cap A_{N,\varepsilon}(b) \cap A_{N,\varepsilon}(c).$$

By intersecting with $A_{N,\varepsilon}$ and its complement, we have

$$\begin{aligned}
&\mathbb{P}\big(\bigcup_{1\leq x\leq N^\alpha} \{|R_N(x)| > C_g N^{-h} x^{2-\tau-h}\}\big)\\
&\qquad \leq \mathbb{P}\left(A_{N,\varepsilon} \cap \left\{\bigcup_{1\leq x\leq N^\alpha} \{|R_N(x)| > C_g N^{-h} x^{2-\tau-h}\}\right\}\right) + \mathbb{P}(A_{N,\varepsilon}^c).
\end{aligned} \tag{A.1.14}$$

We will prove that $\mathbb{P}(A_{N,\varepsilon}^c) \leq N^{-h}$, and that on the event $A_{N,\varepsilon}$, and for each $1 \leq x \leq N^\alpha$, the right-hand side of (A.1.11) can be bounded by $C_g N^{-h} x^{2-\tau-h}$. We start with the latter statement.

Consider the right-hand side of (A.1.11). Clearly, on $A_{N,\varepsilon}(a)$, the first term of $|R_N(x)|$ is bounded by $N^{-3h}[1-G(x)] \leq C_g N^{-3h} x^{2-\tau+h} \leq C_g N^{-h} x^{2-\tau-h}$, where the one but last inequality follows from (A.1.2(b)), and the last since $x \leq N^\alpha < N$ so that $x^{2h} < N^{2h}$. Since for $l > N^\chi$ and each j, $1 \leq j \leq N$, we have that $\{D_j > l\}$ is the empty set on $A_{N,\varepsilon}(b)$, the one but last term of $|R_N(x)|$ vanishes on $A_{N,\varepsilon}(b)$. The last term of $|R_N(x)|$ can, for N large, be bounded, using the inequality $L_N \geq N$ and (A.1.2(a)),

$$\frac{1}{L_N}\sum_{l=N^\chi+1}^{\infty} N\mathbb{P}(D_1 \geq l) \leq \sum_{l=N^\chi+1}^{\infty} l^{1-\tau+h} \quad \leq \frac{N^{\chi(2-\tau+h)}}{\tau-2} < C_g N^{-h+\alpha(2-\tau+h)} \leq C_g N^{-h} x^{2-\tau-h},$$

for all $x \leq N^\alpha$, and where we also used that for ε sufficiently small and $\tau > 2$,

$$\chi(2-\tau+h) < -h + \alpha(2-\tau+h).$$

We bound the third term of $|R_N(x)|$ by

$$\begin{aligned}
\frac{1}{L_N}\sum_{l=N^\beta+1}^{N^\chi} \left|B_l^{(N)} - N\mathbb{P}(D_1 \geq l)\right| &\leq \frac{1}{N}\sum_{l=N^\beta+1}^{N^\chi} [B_l^{(N)} + N\mathbb{P}(D_1 \geq l)]\\
&\leq N^\chi[N^{-1}B_{N^\beta}^{(N)} + \mathbb{P}(D_1 \geq N^\beta)].
\end{aligned} \tag{A.1.15}$$

We note that due to (A.1.2(a)),

$$\mathbb{P}(D_1 \geq N^{\beta}) \geq N^{\beta(1-\tau-h)}, \tag{A.1.16}$$

for large enough N, so that

$$b_{N} = \sqrt{A(\log N)N\mathbb{P}(D_1 \geq N^{\beta})} \leq N\mathbb{P}(D_1 \geq N^{\beta}). \tag{A.1.17}$$

Therefore, on $A_{N,\varepsilon}(c)$, we obtain that

$$B^{(N)}_{N^{\beta}} \leq 2N\mathbb{P}(D_1 \geq N^{\beta}), \tag{A.1.18}$$

for ε small enough and large enough N. Furthermore as $\varepsilon \downarrow 0$,

$$N^{\chi+\beta(1-\tau+h)} < C_g N^{-h+\alpha(2-\tau-h)} \leq C_g N^{-h} x^{2-\tau-h},$$

for $x \leq N^{\alpha}$, $2-\tau-h<0$, because (after multiplying by $\tau-1$ and dividing by ε^5)

$$\chi+\beta(1-\tau+h) < -h+\alpha(2-\tau-h), \qquad \text{or} \qquad \varepsilon(2+2\tau-h) < \tau-2+h,$$

as ε is sufficiently small. Thus, the third term of $|R_{N}(x)|$ satisfies the required bound. We bound the second term of $|R_{N}(x)|$ on $A_{N,\varepsilon}(c)$, using again $L_{N} \geq N$, by

$$\frac{1}{N}\sum_{l=1}^{N^{\beta}} \sqrt{A(\log N)N\mathbb{P}\big(D_1 \geq (x+2)\vee l\big)} = \frac{\sqrt{A\log N}}{\sqrt{N}}\sum_{l=1}^{N^{\beta}} \sqrt{\mathbb{P}\big(D_1 \geq (x+2)\vee l\big)}. \tag{A.1.19}$$

Let c be a constant such that $(\mathbb{P}(D_1 > x))^{\frac{1}{2}} \leq c x^{(1-\tau+h)/2}$, then for all $1 \leq x \leq N^{\alpha}$,

$$\begin{aligned}
\frac{1}{L_{N}}\sum_{l=1}^{N^{\beta}} \big|B^{(N)}_{(x+2)\vee l} - N\mathbb{P}\big(D_1 \geq (x+2)\vee l\big)\big| &\leq \frac{c\sqrt{A\log N}}{\sqrt{N}}\sum_{l=1}^{N^{\beta}} \big((x+2)\vee l\big)^{(1-\tau+h)/2} \\
&\leq \frac{c\sqrt{A\log N}}{\sqrt{N}}\big[x^{(3-\tau+h)/2} + N^{\beta(3-\tau+h)/2}\big] \leq \frac{2c\sqrt{A\log N}}{\sqrt{N}} N^{\beta(3-\tau+h)/2} \\
&\leq N^{h-1/2}N^{\beta(3-\tau+h)/2} < C_g N^{-h}N^{\alpha(2-\tau-h)} \leq C_g N^{-h}x^{2-\tau-h},
\end{aligned} \tag{A.1.20}$$

because

$$h-1/2+\beta(3-\tau+h)/2 < -h+\alpha(2-\tau-h), \qquad \text{or} \qquad h(5\tau-4-h) < 2\varepsilon^5(\tau-2+h),$$

for ε small enough and $\tau \in (2,3)$. We have shown that for $1 \leq x \leq N^{\alpha}$, N sufficiently large, and on the event $A_{N,\varepsilon}$,

$$|R_{N}(x)| \leq C_g N^{-h} x^{2-\tau-h}. \tag{A.1.21}$$

It remains to prove that $\mathbb{P}(A^c_{N,\varepsilon}) \leq N^{-h}$. We use that

$$\mathbb{P}(A^c_{N,\varepsilon}) \leq \mathbb{P}(A_{N,\varepsilon}(a)^c) + \mathbb{P}(A_{N,\varepsilon}(b)^c) + \mathbb{P}(A_{N,\varepsilon}(c)^c), \tag{A.1.22}$$

and we bound each of the three terms separately by $N^{-h/3}$.

Using the Markov inequality followed by the Marcinkiewicz-Zygmund inequality, see e.g. (11, Corollary 8.2 in Section 3), we obtain, with $1 < r < \tau - 1$, and again using that $L_N \geq N$,

$$\begin{aligned}\mathbb{P}(A_{N,\varepsilon}(a)^c) &= \mathbb{P}\Big(\Big|\frac{1}{N}\sum_{j=1}^{N}(D_j-\mu)\Big| > N^{-3h}\cdot L_N/N\Big) \qquad \text{(A.1.23)}\\ &= \mathbb{P}\Big(\big|\sum_{j=1}^{N}(D_j-\mu)\big|^r > (N^{1-3h}\cdot L_N/N)^r\Big) \leq C_r(N^{1-3h})^{-r}N\mathbb{E}[|D_1-\mu|^r] \leq \frac{1}{3}N^{-h},\end{aligned}$$

by choosing h sufficiently small depending on r.

The bound on $\mathbb{P}(A_{N,\varepsilon}(b)^c)$ is a trivial estimate using (A.1.2(a)). Indeed, for N large,

$$\mathbb{P}(A_{N,\varepsilon}(b)^c) = \mathbb{P}\Big(\max_{1\leq j\leq N} D_j > N^{\chi}\Big) \leq N\mathbb{P}(D_1 \geq N^{\chi}) \leq N^{\chi(1-\tau+h)+1} \leq \frac{1}{3}N^{-h}, \qquad \text{(A.1.24)}$$

for small enough ε, because $\tau > 2+h$. For $\mathbb{P}(A_{N,\varepsilon}(c)^c)$, we will use a bound given by Janson (16), which states that for a binomial random variable X with parameters N and p, and all $t > 0$,

$$\mathbb{P}(|X - Np| \geq t) \leq 2\exp\left\{-\frac{t^2}{2(Np+t/3)}\right\}. \qquad \text{(A.1.25)}$$

We will apply (A.1.25) with $t = b_N(x) = \sqrt{A(\log N)N\mathbb{P}(D_1 \geq x)}$, and obtain that uniformly in $x \leq N^{\alpha}$,

$$\begin{aligned}&\mathbb{P}\left(|B_x^{(N)} - N\mathbb{P}(D_1 \geq x)| > b_N(x)\right) \leq 2\exp\left\{-\frac{b_N(x)^2}{2(N\mathbb{P}(D_1\geq x)+b_N(x)/3)}\right\}\\ &\leq 2\exp\left\{-\frac{A\log N}{2(1+\frac{1}{3}\sqrt{A\log N/(N\mathbb{P}(D_1\geq N^{\alpha}))})}\right\} \leq 2N^{-A/3}, \qquad \text{(A.1.26)}\end{aligned}$$

because

$$\frac{\log N}{N\mathbb{P}(D_1 \geq N^{\alpha})} \leq \frac{\log N}{N^{1+\alpha(\tau-1-h)}} \to 0,$$

as $N \to \infty$. Thus, (A.1.26) yields, using $A = 3(\beta+2h)$,

$$\mathbb{P}(A_{N,\varepsilon}(c)^c) \leq \sum_{x=1}^{N^{\beta}}\mathbb{P}\big(|B_x^{(N)} - N\mathbb{P}(D_1\geq x)| > b_N(x)\big) \leq 2N^{\beta-A/3} = 2N^{-2h} \leq \frac{1}{3}N^{-h}. \qquad \text{(A.1.27)}$$

This completes the proof of the lemma. $\square$

For the third lemma we introduce some further notation. For any $x \in \mathbb{N}$, define

$$\hat{S}_x^{(N)} = \sum_{i=1}^{x}\hat{X}_i^{(N)}, \qquad \hat{V}_x^{(N)} = \max_{1\leq i\leq x}\hat{X}_i^{(N)},$$

where $\{\hat{X}_i^{(N)}\}_{i=1}^{x}$ have the same law, say $\hat{H}^{(N)}$, but are *not necessarily* independent.

Lemma A.1.4 (Sums and maxima with law $\hat{H}^{(N)}$ on the good event)**.**

(i) *If* $\hat{H}^{(N)}$ *satisfies*

$$[1-\hat{H}^{(N)}(z)] \leq [1+2N^{-h}][1-G(z)], \quad \forall\, z \leq y, \tag{A.1.28}$$

then for all $x \in \mathbb{N}$*, there exists a constant* b'*, such that:*

$$\mathbb{P}\Big(\hat{S}_x^{(N)} \geq y\Big) \leq b'x[1+2N^{-h}]\big[1-G\big(y\big)\big]. \tag{A.1.29}$$

(ii) *If* $\hat{H}^{(N)}$ *satisfies*

$$[1-\hat{H}^{(N)}(y)] \geq [1-2N^{-h}][1-G(y)], \tag{A.1.30}$$

and $\{\hat{X}_i^{(N)}\}_{i=1}^x$ *are independent, then for all* $x \in \mathbb{N}$*,*

$$\mathbb{P}\left(\hat{V}_x^{(N)} \leq y\right) \leq \Big(1-[1-2N^{-h}][1-G\left(y\right)]\Big)^x. \tag{A.1.31}$$

Proof. We first bound $\mathbb{P}\Big(\hat{S}_x^{(N)} \geq y\Big)$. We write

$$\mathbb{P}\Big(\hat{S}_x^{(N)} \geq y\Big) \leq \mathbb{P}\Big(\hat{S}_x^{(N)} \geq y, \hat{V}_x^{(N)} \leq y\Big) + \mathbb{P}\big(\hat{V}_x^{(N)} > y\big). \tag{A.1.32}$$

Due to (A.1.28), the second term is bounded by

$$x\mathbb{P}\big(\hat{X}_1^{(N)} > y\big) = x\Big[1-\hat{H}^{(N)}\big(y\big)\Big] \leq x[1+2N^{-h}]\Big[1-G\big(y\big)\Big]. \tag{A.1.33}$$

We use the Markov inequality and (A.1.28) to bound the first term on the right-hand side of (A.1.32) by

$$\begin{aligned}\mathbb{P}\Big(\hat{S}_x^{(N)} \geq y, \hat{V}_x^{(N)} \leq y\Big) &\leq \frac{1}{y}\mathbb{E}\left(\hat{S}_x^{(N)}\mathbf{1}_{\{\hat{V}_x^{(N)} \leq y\}}\right) \leq \frac{x}{y}\mathbb{E}\left(\hat{X}_1^{(N)}\mathbf{1}_{\{\hat{X}_1^{(N)} \leq y\}}\right)\\ &\leq \frac{x}{y}\sum_{i=1}^{y}[1-\hat{H}^{(N)}(i)] \leq \frac{x}{y}[1+2N^{-h}]\sum_{i=1}^{y}[1-G(i)].\end{aligned} \tag{A.1.34}$$

For the latter sum, we use (13, Theorem 1(b), p. 281), together with the fact that $1-G(y)$ is regularly varying with exponent $2-\tau \neq 1$, to deduce that there exists a constant c_1 such that

$$\sum_{i=1}^{y}[1-G(i)] \leq c_1 y[1-G(y)]. \tag{A.1.35}$$

Combining (A.1.32), (A.1.33), (A.1.34) and (A.1.35), we conclude that

$$\mathbb{P}\Big(\hat{S}_x^{(N)} \geq y\Big) \leq b'x[1+2N^{-h}]\big[1-G\big(y\big)\big)\big], \tag{A.1.36}$$

where $b' = c_1 + 1$. This completes the proof of Lemma A.1.4(i).

For the proof of (ii), we use independence of $\{\hat{X}_i^{(N)}\}_{i=1}^x$, and condition (A.1.30), to conclude that

$$\mathbb{P}\Big(\hat{V}_x^{(N)} \leq y\Big) = \left(\hat{H}^{(N)}\left(y\right)\right)^x = \Big(1-\Big[1-\hat{H}^{(N)}(y)\Big]\Big)^x \leq \Big(1-[1-2N^{-h}][1-G\left(y\right)]\Big)^x.$$

Hence, (A.1.31) holds. $\square$

Remark A.1.5. *In the proofs in the appendix, we will only use that*

(i) *the event F_ε holds* **whp**;

(ii) *that L_N is concentrated around its mean;*

(iii) *that,* **whp**, *the maximal degree is bounded by N^χ for any $\chi > 1/(\tau-1)$.*

Moreover, the proof of Proposition 3.1 relies on (14, Proposition A.3.1), and in its proof it was further used that

(iv) $p_N \le N^{\alpha_2}$, **whp**, *for any $\alpha_2 > 0$, where p_N is the total variation distance between g and $g^{(N)}$, i.e.,*

$$p_N = \frac{1}{2}\sum_n |g_n - g_n^{(N)}|. \tag{A.1.37}$$

Therefore, if instead of taking the degrees i.i.d. with distribution F, we would take the degrees in an exchangeable way such that the above restrictions hold, then the proof carries on verbatim. In particular, this implies that our results also hold for the usual configuration model, where the degrees are fixed, as long as the above restrictions are satisfied. □

A.2 Proof of Proposition 3.1

The proof makes use of (14, Proposition A.3.1), which proves the statement in Proposition 3.1 under an additional condition.

In order to state this condition let $\{\hat{Z}_j^{(i,N)}\}_{j\ge 1}$, $i = 1, 2$, be two independent copies of the delayed BP, where $\hat{Z}_1^{(i,N)}$ has law $\{f_n\}$ given in (1.1), and where the offspring of any individual in generation j with $j > 1$ has law $\{g_n^{(N)}\}$, where $g_n^{(N)}$ is defined in (2.6). Then, the conclusion of Proposition 3.1 follows from (14, Proposition A.3.1), for any m such that, for any $\eta > 0$, and $i = 1, 2$,

$$\mathbb{P}(\sum_{j=1}^{m} \hat{Z}_j^{(i,N)} \ge N^\eta) = o(1)., \quad N \to \infty \tag{A.2.1}$$

By exchangeability it suffices to prove (A.2.1) for $i = 1$ only, we can therefore simplify notation and write further $\hat{Z}_j^{(N)}$ instead of $\hat{Z}_j^{(i,N)}$. We turn to the proof of (A.2.1).

By Lemma A.1.3 and (A.1.2(b)), respectively, for every $\eta > 0$, there exists a $c_\eta > 0$, such that **whp** for all $x \le N^\alpha$,

$$1 - G^{(N)}(x) \le [1 + 2N^{-h}][1 - G(x)] \le c_\eta x^{2-\tau+\eta}. \tag{A.2.2}$$

We call a generation $j \ge 1$ **good**, when

$$\hat{Z}_j^{(N)} \le \left(\hat{Z}_{j-1}^{(N)} \log N\right)^{\frac{1}{\tau-2-\eta}}, \tag{A.2.3}$$

and **bad** otherwise, where as always $\hat{Z}_0^{(N)} = 1$. We further write

$$H_m = \{\text{generations } 1, \ldots, m \text{ are good}\}. \tag{A.2.4}$$

We will prove that when H_m holds, then $\sum_{j=1}^m \hat{Z}_j^{(N)} \leq N^\eta$. Indeed, when generations $1,\ldots,m$ are all good, then, for all $j \leq m$,

$$\hat{Z}_j^{(N)} \leq (\log N)^{\sum_{i=1}^j (\tau-2-\eta)^{-i}}. \tag{A.2.5}$$

Therefore,

$$\sum_{j=1}^m \hat{Z}_j^{(N)} \leq m(\log N)^{\sum_{i=1}^m (\tau-2-\eta)^{-i}} \leq m(\log N)^{\frac{(\tau-2-\eta)^{-m-2}}{(\tau-2-\eta)^{-1}-1}} \leq N^\eta, \tag{A.2.6}$$

for any $\eta > 0$, when N is sufficiently large. We conclude that

$$\mathbb{P}\Big(\sum_{j=1}^m \hat{Z}_j^{(N)} > N^\eta\Big) \leq \mathbb{P}(H_m^c), \tag{A.2.7}$$

and Proposition 3.1 follows if we show that $\mathbb{P}(H_m^c) = o(1)$. In order to do so, we write

$$\mathbb{P}(H_m^c) = \mathbb{P}(H_1^c) + \sum_{j=1}^{m-1} \mathbb{P}(H_{j+1}^c \cap H_j). \tag{A.2.8}$$

For the first term, we use (A.1.2(a)) to deduce that

$$\mathbb{P}(H_1^c) = \mathbb{P}\big(D_1 > (\log N)^{\frac{1}{\tau-2-\eta}}\big) \leq (\log N)^{-\frac{\tau-1-\eta}{\tau-2-\eta}} \leq (\log N)^{-1}. \tag{A.2.9}$$

For $1 \leq j \leq m$, we have $\hat{Z}_j^{(N)} \leq \sum_{k=1}^m \hat{Z}_k^{(N)}$, and using (A.2.6),

$$\sum_{j=1}^m \hat{Z}_j^{(N)} \leq m(\log N)^{\frac{(\tau-2-\eta)^{-m-2}}{(\tau-2-\eta)^{-1}-1}} \equiv W_N. \tag{A.2.10}$$

Using Lemma A.1.4(i) with $\hat{H}^{(N)} = G^{(N)}$, $x = l$ and $y = v_N(l) = (l\log N)^{\frac{1}{\tau-2-\eta}}$, where (A.1.28) follows from (A.2.2), we obtain that

$$\begin{aligned}\mathbb{P}(H_{j+1}^c \cap H_j) &\leq \sum_{l=1}^{W_N} \mathbb{P}\Big(\hat{Z}_{j+1}^{(N)} \geq v_N(l) \Big| \hat{Z}_j^{(N)} = l\Big)\mathbb{P}(\hat{Z}_j^{(N)} = l) \\ &\leq \max_{1\leq l\leq W_N} \mathbb{P}\Big(\hat{S}_l^{(N)} \geq v_N(l)\Big) \leq b' \max_{1\leq l\leq W_N} l[1+2N^{-h}]\big[1-G(v_N(l))\big]. \end{aligned} \tag{A.2.11}$$

Furthermore by (A.1.2(b)),

$$\max_{1\leq l\leq W_N} l[1-G(v_N(l))] \leq \max_{1\leq l\leq W_N} l v_N(l)^{2-\tau+\eta} = (\log N)^{-1}. \tag{A.2.12}$$

This completes the proof of Proposition 3.1. □

A.3 Some further preparations

Before we can prove Propositions 3.2 and 3.4 , we state a lemma that was proved in (14).

We introduce some notation. Suppose we have L objects divided into N groups of sizes $d_1, \ldots, d_N$, so that $L = \sum_{i=1}^N d_i$. Suppose we draw an object at random. This gives a distribution $g^{(\vec{d})}$, i.e.,

$$g_n^{(\vec{d})} = \frac{1}{L}\sum_{i=1}^N d_i \mathbf{1}_{\{d_i = n+1\}}, \quad n = 0, 1, \ldots \tag{A.3.1}$$

Clearly, $g^{(N)} = g^{(\vec{D})}$, where $\vec{D} = (D_1, \ldots, D_N)$. We further write

$$G^{(\vec{d})}(x) = \sum_{n=0}^{x} g_n^{(\vec{d})}. \tag{A.3.2}$$

We next label M of the L objects in an arbitrary way, and suppose that the distribution $G_M^{(\vec{d})}(x)$ is obtained in a similar way from drawing conditionally on drawing an unlabelled object. More precisely, we remove the labelled objects from all objects thus creating new $d'_1, \ldots, d'_N$, and we let $G_M^{(\vec{d})}(x) = G^{(\vec{d}')}(x)$. Even though this is not indicated, the law $G_M^{(\vec{d})}$ depends on what objects have been labelled.

Lemma A.3.1 below shows that the law $G_M^{(\vec{d})}$ can be stochastically bounded above and below by two specific ways of labeling objects. Before we can state the lemma, we need to describe those specific labellings.

For a vector $\vec{d}$, we denote by $d_{(1)} \leq d_{(2)} \leq \ldots \leq d_{(N)}$ the ordered coordinates. Then the laws $\overline{G}_M^{(\vec{d})}$ and $\underline{G}_M^{(\vec{d})}$, respectively, are defined by successively decreasing $d_{(N)}$ and $d_{(1)}$, respectively, by one. Thus,

$$\begin{aligned} \overline{G}_1^{(\vec{d})}(x) &= \frac{1}{L-1}\sum_{i=1}^{N-1} d_{(i)} \mathbf{1}_{\{d_{(i)} \leq x+1\}} + \frac{d_{(N)}-1}{L-1}\mathbf{1}_{\{d_{(N)}-1 \leq x+1\}}, && \text{(A.3.3)} \\ \underline{G}_1^{(\vec{d})}(x) &= \frac{1}{L-1}\sum_{i=2}^{N} d_{(i)} \mathbf{1}_{\{d_{(i)} \leq x+1\}} + \frac{d_{(1)}-1}{L-1}\mathbf{1}_{\{d_{(1)}-1 \leq x+1\}}. && \text{(A.3.4)} \end{aligned}$$

For $\overline{G}_M^{(\vec{d})}$ and $\underline{G}_M^{(\vec{d})}$, respectively, we perform the above change M times, and after each repetition we reorder the groups. Here we note that when $d_{(N)} = 1$ (in which case $d_i = 1$, for all i), and for $\overline{G}_1^{(\vec{d})}$ we decrease $d_{(N)}$ by one, that we only keep $d_{(1)}, \ldots, d_{(N-1)}$. A similar rule applies when $d_{(1)} = 1$ and for $\underline{G}_1^{(\vec{d})}$ we decrease $d_{(1)}$ by one. Thus, in these cases, the number of groups of objects, indicated by N, is decreased by 1. Applying the above procedure to $\vec{d} = (D_1, \ldots, D_N)$ we obtain that, for all $x \geq 1$,

$$\overline{G}_M^{(N)}(x) \equiv \overline{G}_M^{(\vec{D})}(x) \leq \frac{1}{L_N - M}\sum_{i=1}^N D_i \mathbf{1}_{\{D_i \leq x+1\}} = \frac{L_N}{L_N - M} G^{(N)}(x), \tag{A.3.5}$$

$$\underline{G}_M^{(N)}(x) \equiv \underline{G}_M^{(\vec{D})}(x) \geq \frac{1}{L_N - M}\Big[\sum_{i=1}^N D_i \mathbf{1}_{\{D_i \leq x+1\}} - M\Big] = \frac{1}{L_N - M}\Big[L_N G^{(N)}(x) - M\Big], \tag{A.3.6}$$

where equality is achieved precisely when $D_{(N)} \geq x + M$, and $\#\{i : D_i = 1\} \geq M$, respectively. Finally, for two distribution functions F, G, we write that $F \preceq G$ when $F(x) \geq G(x)$ for all x. Similarly, we write that $X \preceq Y$ when for the distribution functions F_X, F_Y we have that $F_X \preceq F_Y$.

We next prove stochastic bounds on the distribution $G_M^{(\vec{d})}(x)$ that are uniform in the choice of the M labelled objects. The proof of Lemma A.3.1 can be found in (14).

Lemma A.3.1. *For all choices of M labelled objects*

$$\underline{G}_M^{(\vec{d})} \preceq G_M^{(\vec{d})} \preceq \overline{G}_M^{(\vec{d})}. \tag{A.3.7}$$

Moreover, when $X_1, \ldots, X_j$ are draws from $G_{M_1}^{(\vec{d})}, \ldots, G_{M_l}^{(\vec{d})}$, where the only dependence between the X_i resides in the labelled objects, then

$$\sum_{i=1}^{j} \underline{X}_i \preceq \sum_{i=1}^{j} X_i \preceq \sum_{i=1}^{j} \overline{X}_i, \tag{A.3.8}$$

where $\{\underline{X}_i\}_{i=1}^j$ and $\{\overline{X}_i\}_{i=1}^j$, respectively, are i.i.d. copies of $\underline{X}$ and $\overline{X}$ with laws $\underline{G}_M^{(N)}$ and $\overline{G}_M^{(N)}$ for $M = \max_{1 \leq i \leq l} M_i$, respectively.

We will apply Lemma A.3.1 to $G^{(\vec{D})} = G^{(N)}$.

A.3.1 The inductive step

Our key result, which will yield the proofs of Proposition 3.2 and 3.4, is Proposition A.3.2 below. This proposition will be proved in Section A.4. For its formulation we need some more notation. As before we simplify notation and write further on $Z_k^{(N)}$ instead of $Z_k^{(i,N)}$. Similarly, we write $\mathcal{Z}_k$ instead of $\mathcal{Z}_k^{(i)}$ and $\mathcal{T}_m^{(N)}(\varepsilon)$ instead of $\mathcal{T}_m^{(i,N)}(\varepsilon)$. Recall that we have defined previously

$$\kappa = \frac{1}{\tau - 2} > 1 \qquad \text{and} \qquad \alpha = \frac{1 - \varepsilon^5}{\tau - 1}.$$

In the sequel we work with $Y_k^{(N)} > \varepsilon$, for k large enough, i.e., we work with $Z_k^{(N)} > e^{\varepsilon \kappa^k} > 1$, due to definition (3.1). Hence, we can treat these definitions as

$$Y_k^{(N)} = \kappa^{-k} \log(Z_k^{(N)}) \qquad \text{and} \qquad Y_k = \kappa^{-k} \log(\mathcal{Z}_k). \tag{A.3.9}$$

With γ defined in the Assumption 1.1(ii), and $0 < \varepsilon < 3 - \tau$, we take m_ε sufficiently large to have

$$\sum_{k=m_\varepsilon}^{\infty} (\tau - 2 + \varepsilon)^{k(1-\gamma)} \leq \varepsilon^3 \qquad \text{and} \qquad \sum_{k=m_\varepsilon}^{\infty} k^{-2} \leq \varepsilon/2. \tag{A.3.10}$$

For any $m_\varepsilon \leq m < k$, we denote

$$M_k^{(N)} = \sum_{j=1}^{k} Z_j^{(N)}, \quad \text{and} \quad M_k = \sum_{j=1}^{k} \mathcal{Z}_j. \tag{A.3.11}$$

As defined in Section 3 of (14) we speak of free stubs at level l, as the free stubs connected to nodes at distance $l-1$ from the root; the total number of free stubs, obtained immediately after pairing of all stubs at level $l-1$ equals $Z_l^{(N)}$(see also Section 3.2 above). For any $l \geq 1$ and $1 \leq x \leq Z_{l-1}^{(N)}$, let $Z_{x,l}^{(N)}$ denote the number of constructed free stubs at level l after pairing of the first x stubs of $Z_{l-1}^{(N)}$. Note that for $x = Z_{l-1}^{(N)}$, we obtain $Z_{x,l}^{(N)} = Z_l^{(N)}$. For general x, the quantity $Z_{x,l}^{(N)}$ is loosely speaking the sum of the number of children of the first x stubs at level $l-1$, and according to the coupling at fixed times (Proposition 3.1) this number is for fixed l, **whp** equal to the number of children of the first x individuals in generation $l-1$ of the BP $\{\mathcal{Z}_k\}_{k\geq 1}$.

We introduce the event $\hat{F}_{m,k}(\varepsilon)$,

$$\hat{F}_{m,k}(\varepsilon) = \begin{array}{ll} \{k \in \mathcal{T}_m^{(N)}(\varepsilon)\} & (a) \\ \cap\{\forall m < l \leq k-1: \ |Y_l^{(N)} - Y_m^{(N)}| \leq \varepsilon^3\} & (b) \\ \cap\{\varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}\} & (c) \\ \cap\{M_m^{(N)} \leq 2Z_m^{(N)}\}. & (d) \end{array} \tag{A.3.12}$$

We denote by $X_{i,l-1}^{(N)}$ the number of brother stubs of a stub attached to the i^{th} stub of SPG_{l-1}. In the proof of Proposition A.3.2 we compare the quantity $Z_{x,l}^{(N)}$ to the sum $\sum_{i=1}^x X_{i,l-1}^{(N)}$ for part (a) and to $\max_{1\leq i\leq x} X_{i,l-1}^{(N)}$ for part (b). We then couple $X_{i,l-1}^{(N)}$ to $\overline{X}_{i,l-1}^{(N)}$ for part (a) and to $\underline{X}_{i,l-1}^{(N)}$ for part (b). Among other things, the event $\hat{F}_{m,k}(\varepsilon)$ ensures that these couplings hold.

Proposition A.3.2 (Inductive step). *Let F satisfy Assumption 1.1(ii). For $\varepsilon > 0$ sufficiently small and c_γ sufficiently large, there exist a constant $b = b(\tau, \varepsilon) > 0$ such that, for $x = Z_{l-1}^{(N)} \wedge N^{\frac{(1-\varepsilon/2)}{\kappa(\tau-1)}}$,*

$$\mathbb{P}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \geq (l^3 x)^{\kappa + c_\gamma \gamma(x)}\big\}\Big) \leq b l^{-3}, \qquad (a)$$

$$\mathbb{P}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \leq \big(\tfrac{x}{l^3}\big)^{\kappa - c_\gamma \gamma(x)}\big\}\Big) \leq b l^{-3}. \qquad (b)$$

The proof of Proposition A.3.2 is quite technical and is given in Section A.4. In this section we give a short overview of the proof. For $l \geq 1$, let SPG_l denote the shortest path graph containing all nodes on distance $l-1$, and including all stubs at level l, i.e., the moment we have $Z_l^{(N)}$ free stubs at level l. As before, we denote by $X_{i,l-1}^{(N)}$, $i \in \{1, \ldots, x\}$, the number of brother stubs of a stub attached to the i^{th} stub of SPG_{l-1} (see Figure A.3.1).

Because $Z_{x,l}^{(N)}$ is the number of free stubs at level l after the pairing of the first x stubs, one would expect that

$$Z_{x,l}^{(N)} \sim \sum_{i=1}^{x} X_{i,l-1}^{(N)}, \tag{A.3.13}$$

where $\sim$ denotes that we have an uncontrolled error term. Indeed, the intuition behind (A.3.13) is that loops or cycles should be rare for small l. Furthermore, when $M_{l-1}^{(N)}$ is much smaller than N, then the law of $X_{i,l-1}^{(N)}$ should be quite close to the law $G^{(N)}$, which, in turn, by Lemma A.1.3 is close to G. If $X_{i,l-1}^{(N)}$ would have distribution $G(x)$, then we could use the theory of sums of random variables with infinite expectation, as well as extreme value theory, to obtain the inequalities of Proposition A.3.2.

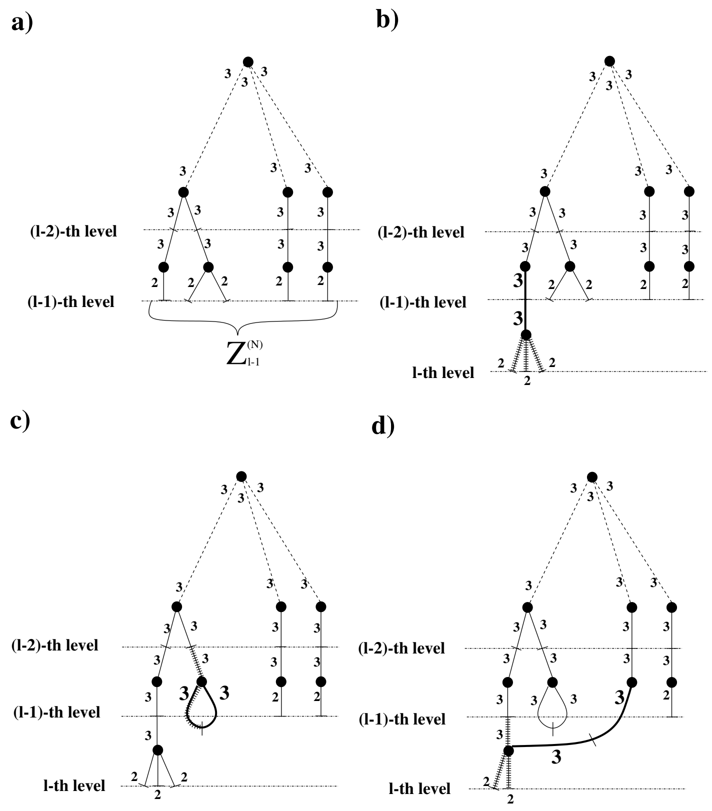

Figure 4: The building of the l^{th} level of SPG. The last paired stubs are marked by thick lines, the brother stubs by dashed lines. In a) the $(l-1)^{\text{st}}$ level is completed, in b) the pairing with a new node is described, in c) the pairing within the $(l-1)^{\text{st}}$ level is described, and in d) the pairing with already existing node at l^{th} level is described.

In order to make the above intuition rigorous, we use upper and lower bounds. We note that the right-hand side of (A.3.13) is a valid upper bound for $Z_{x,l}^{(N)}$. We show below that $X_{i,l-1}^{(N)}$ have the same law, and we wish to apply Lemma A.1.4(i). For this, we need to control the law $X_{i,l-1}^{(N)}$, for which we use Lemma A.3.1 to bound each $X_{i,l-1}^{(N)}$ from above by a random variable with law $\overline{G}_M^{(N)}$. This coupling makes sense only on the *good event* where $\overline{G}_M^{(N)}$ is sufficiently close to G.

For the lower bound, we have to do more work. The basic idea from the theory of sums of random variables with infinite mean is that the sum has the same order as the maximal summand. Therefore, we bound from below

$$Z_{x,l}^{(N)} \geq \underline{Z}_{x,l}^{(N)} - x. \tag{A.3.14}$$

where

$$\underline{Z}_{x,l}^{(N)} = \max_{1 \leq i \leq x} X_{i,l-1}^{(N)}. \tag{A.3.15}$$

However, we will see that this lower bound is only valid when the chosen stub is not part of the shortest path graph up to that point. We show in Lemma A.3.4 below that the chosen stub has label 1 when $\underline{Z}^{(N)}_{x,l} > 2M^{(N)}_{l-1}$. In this case, (A.3.14) follows since the $x-1$ remaining stubs can 'eat up' at most $x-1 \le x$ stubs. To proceed with the lower bound, we bound $(X^{(N)}_{1,l-1}, \ldots, X^{(N)}_{x,l-1})$ stochastically from below, using Lemma A.3.1, by an i.i.d. sequence of random variables with laws $\underline{G}^{(N)}_M$, where M is chosen appropriately and serves as an upper bound on the number of stubs with label 3. Again on the *good event*, $\underline{G}^{(N)}_M$ is sufficiently close to G. Therefore, we are now faced with the problem of studying the maximum of a number of random variables with a law close to G. Here we can use Lemma A.1.4(ii), and we conclude in the proof of Proposition A.3.2(a) that $\underline{Z}^{(N)}_{x,l}$ is to leading order equal to x^{κ}, when $x = Z^{(N)}_{l-1} \wedge N^{\frac{1-\varepsilon/2}{\kappa(\tau-1)}}$. For this choice of x, we also see that $\underline{Z}^{(N)}_{x,l}$ is of bigger order than $M^{(N)}_{l-2}$, so that the basic assumption in the above heuristic is satisfied. This completes the overview of the proof.

We now state and prove the Lemmas A.3.3 and A.3.4. The proof of Proposition A.3.2 then follows in Section A.4. We define the good event mentioned above by

$$F_{\varepsilon,M} = \bigcap_{x=1}^{N^{\alpha}} \Big\{ [1-2N^{-h}][1-G(x)] \le 1 - \overline{G}^{(N)}_M(x) \le 1 - \underline{G}^{(N)}_M(x) \le [1+2N^{-h}][1-G(x)] \Big\}. \tag{A.3.16}$$

The following lemma says that for $M \le N^{\alpha}$, the probability of the good event is close to one.

Lemma A.3.3. *Let F satisfy Assumption 1.1(ii). Then, for $\varepsilon > 0$ sufficiently small,*

$$\mathbb{P}(F^c_{\varepsilon,N^{\alpha}}) \le N^{-h}, \qquad \textit{for large } N.$$

Proof. Due to Lemma A.1.3 it suffices to show that for ε small enough, and N sufficiently we have

$$F^c_{\varepsilon,N^{\alpha}} \subseteq F^c_{\varepsilon}. \tag{A.3.17}$$

We will prove the equivalent statement that

$$F_{\varepsilon} \subseteq F_{\varepsilon,N^{\alpha}}. \tag{A.3.18}$$

It follows from (A.3.5) and (A.3.6) that for every M and x

$$1 - \underline{G}^{(N)}_M(x) \le 1 - G^{(N)}(x) \le 1 - \overline{G}^{(N)}_M(x), \tag{A.3.19}$$

and, in particular, that for $M \le N^{\alpha}$,

$$[1 - \overline{G}^{(N)}_M(x)] - [1 - \underline{G}^{(N)}_M(x)] \le \frac{M}{L_N - M} \le O(N^{\alpha-1}). \tag{A.3.20}$$

Then we use (A.1.2(b)) to obtain that for all $x \le N^{\alpha}$, ε small enough, and N sufficiently large,

$$\begin{aligned} O(N^{\alpha-1}) &\le N^{\alpha-1+h} = N^{\frac{1-\varepsilon^5}{\tau-1}-1+\varepsilon^6} < N^{-2\varepsilon^6} N^{\frac{1-\varepsilon^5}{\tau-1}(2-\tau-\varepsilon^6)} \\ &= N^{-2h} N^{\alpha(2-\tau-h)} \le N^{-2h} x^{2-\tau-h} \le N^{-h}[1-G(x)]. \end{aligned} \tag{A.3.21}$$

Therefore, for $M \le N^{\alpha}$ and with the above choices of ε, α and h, we have, uniformly for $x \le N^{\alpha}$ and on F_{ε},

$$[1 - \underline{G}^{(N)}_M(x)] \le \quad 1 - G^{(N)}(x) + [1 - \overline{G}^{(N)}_M(x)] - [1 - \underline{G}^{(N)}_M(x)] \le [1+2N^{-h}][1-G(x)],$$

$$[1 - \underline{G}^{(N)}_M(x)] \ge \quad 1 - G^{(N)}(x) - [1 - \overline{G}^{(N)}_M(x)] + [1 - \underline{G}^{(N)}_M(x)] \ge [1-2N^{-h}][1-G(x)],$$

i.e., we have (A.3.16), so that indeed $F_\varepsilon \subseteq F_{\varepsilon,N^\alpha}$. □

For the coupling of $X^{(N)}_{i,l-1}$ with the random variables with laws $\underline{G}^{(N)}_M(x)$ and $\overline{G}^{(N)}_M(x)$ we need the following lemma. Recall the definition of $M^{(N)}_l$ given in (A.3.11.

Lemma A.3.4. *For any $l \geq 1$ there are at most $2M^{(N)}_l$ stubs with label 3 in* SPG_{l+1}*, while the number of stubs with label 2 is (by definition) equal to $Z^{(N)}_{l+1}$.*

Proof. The proof is by induction on l. There are $Z^{(N)}_1$ free stubs in SPG_1. Some of these stubs will be paired with stubs with label 2 or 3, others will be paired to stubs with label 1 (see Figure A.3.1). This gives us at most $2Z^{(N)}_1$ stubs with label 3 in SPG_2. This initializes the induction. We next advance the induction. Suppose that for some $l \geq 1$ there are at most $2M^{(N)}_l$ stubs with label 3 in SPG_{l+1}. There are $Z^{(N)}_{l+1}$ free stubs (with label 2) in SPG_{l+1}. Some of these stubs will be paired with stubs with label 2 or 3, others will be linked with stubs with label 1 (again see Figure A.3.1). This gives us at most $2Z^{(N)}_{l+1}$ new stubs with label 3 in SPG_{l+2}. Hence the total number of these stubs is at most $2M^{(N)}_l + 2Z^{(N)}_{l+1} = 2M^{(N)}_{l+1}$. This advances the induction hypothesis, and proves the claim. □

A.4 The proof of Proposition A.3.2

We state and prove some consequences of the event $\hat{F}_{m,k}(\varepsilon)$, defined in (A.3.12). We refer to the outline of the proof of Proposition A.3.2, to explain where we use these consequences.

Lemma A.4.1. *The event $\hat{F}_{m,k}(\varepsilon)$ implies, for sufficiently large N, the following bounds:*

$$\begin{array}{lll} (a) & M^{(N)}_{k-1} < N^{\frac{1-3\varepsilon^4/4}{\kappa(\tau-1)}}, & \\ (b) & \text{for any } \delta > 0,\ N^{-\delta} \leq k^{-3}, & \\ (c) & \kappa^{k-1}(\varepsilon - \varepsilon^3) \leq \log\left(Z^{(N)}_{k-1}\right) \leq \kappa^{k-1}(\varepsilon^{-1} + \varepsilon^3), \quad \text{for} \quad k-1 \geq m, & \\ (d) & M^{(N)}_{k-1} \leq 2Z^{(N)}_{k-1}, \quad \text{for} \quad k-1 \geq m. & \end{array} \tag{A.4.1}$$

Proof. Assume that (A.3.12(a)-(d)) holds. We start by showing (A.4.1(b)), which is evident if we show the following claim:

$$k \leq \frac{\log\left(\frac{1-\varepsilon^2}{\varepsilon(\tau-1)}\log N\right)}{\log \kappa}, \tag{A.4.2}$$

for N large enough. In order to prove (A.4.2), we note that if $k \in \mathcal{T}^{(N)}_m(\varepsilon)$ then, due to definition (3.3),

$$\kappa^{k-m} \leq \frac{1-\varepsilon^2}{\tau-1}\frac{\log N}{\log\left(Z^{(N)}_m\right)} < \frac{1-\varepsilon^2}{\varepsilon(\tau-1)}\kappa^{-m}\log N, \tag{A.4.3}$$

where the latter inequality follows from $Y^{(N)}_m > \varepsilon$ and (A.3.9). Multiplying by κ^m and taking logarithms on both sides yields (A.4.2).

We now turn to (A.4.1(a)). Since

$$M^{(N)}_{k-1} = \sum_{l=1}^{k-1} Z^{(N)}_l \leq k \max_{1\leq l\leq k-1} Z^{(N)}_l,$$

the inequality (A.4.1(a)) follows from (A.4.2), when we show that for any $l \leq k-1$,

$$Z_l^{(N)} \leq N^{\frac{1-\varepsilon^4}{\kappa(\tau-1)}}. \tag{A.4.4}$$

Observe that for $l < m$ we have that, due to (A.3.9),(A.3.12(c)) and (A.3.12(d)), for any $\varepsilon > 0$ and m fixed and by taking N sufficiently large,

$$Z_l^{(N)} \leq M_m^{(N)} \leq 2Z_m^{(N)} \leq 2e^{\kappa^m \varepsilon^{-1}} < N^{\frac{1-\varepsilon^4}{\kappa(\tau-1)}}. \tag{A.4.5}$$

Consider $m \leq l \leq k-1$. Due to (A.3.9), inequality (A.4.4) is equivalent to

$$\kappa^{l+1} Y_l^{(N)} \leq \frac{1-\varepsilon^4}{\tau-1} \log N. \tag{A.4.6}$$

To obtain (A.4.6) we will need two inequalities. Firstly, (A.3.12(a)) and $l+1 \leq k$ imply that

$$\kappa^{l+1} Y_m^{(N)} \leq \frac{1-\varepsilon^2}{\tau-1} \log N. \tag{A.4.7}$$

Given (A.4.7) and (A.3.12(b)), we obtain, when $Y_m^{(N)} \geq \varepsilon$, and for $m \leq l \leq k-1$,

$$\begin{aligned} \kappa^{l+1} Y_l^{(N)} &\leq \kappa^{l+1}(Y_m^{(N)} + \varepsilon^3) \leq \kappa^{l+1} Y_m^{(N)}(1+\varepsilon^2) \\ &\leq \tfrac{(1-\varepsilon^2)(1+\varepsilon^2)}{\tau-1} \log N = \tfrac{1-\varepsilon^4}{\tau-1} \log N. \end{aligned} \tag{A.4.8}$$

Hence we have (A.4.6) or equivalently (A.4.4) for $m \leq l \leq k-1$.

The bound in (A.4.1(c)) is an immediate consequence of (A.3.9) and (A.3.12(b,c)) that imply for $k-1 > m$,

$$\varepsilon - \varepsilon^3 \leq Y_{k-1}^{(N)} \leq \varepsilon^{-1} + \varepsilon^3.$$

We complete the proof by establishing (A.4.1(d)). We use induction to prove that for all $l \geq m$, the bound $M_l^{(N)} \leq 2Z_l^{(N)}$ holds. The initialization of the induction hypothesis for $l = m$ follows from (A.3.12(d)). So assume that for some $m \leq l < k-1$ the inequality $M_l^{(N)} \leq 2Z_l^{(N)}$ holds, then

$$M_{l+1}^{(N)} = Z_{l+1}^{(N)} + M_l^{(N)} \leq Z_{l+1}^{(N)} + 2Z_l^{(N)}, \tag{A.4.9}$$

so that it suffices to bound $2Z_l^{(N)}$ by $Z_{l+1}^{(N)}$. We note that $\hat{F}_{m,k}(\varepsilon)$ implies that

$$|Y_{l+1}^{(N)} - Y_l^{(N)}| \leq |Y_{l+1}^{(N)} - Y_m^{(N)}| + |Y_l^{(N)} - Y_m^{(N)}| \leq 2\varepsilon^3 \leq 3\varepsilon^2 Y_{l+1}^{(N)}, \tag{A.4.10}$$

where in the last inequality we used that $Y_{l+1}^{(N)} \geq Y_m^{(N)} - \varepsilon^3 \geq \varepsilon - \varepsilon^3 > \frac{2}{3}\varepsilon$, as $\varepsilon \downarrow 0$. Therefore,

$$2Z_l^{(N)} = 2e^{\kappa^l Y_l^{(N)}} \leq 2e^{(1+3\varepsilon^2)\kappa^l Y_{l+1}^{(N)}} = 2\big(Z_{l+1}^{(N)}\big)^{(1+3\varepsilon^2)\kappa^{-1}} \leq Z_{l+1}^{(N)}, \tag{A.4.11}$$

when $\varepsilon > 0$ is so small that $\omega = (1+3\varepsilon^2)\kappa^{-1} < 1$ and where we take m large enough to ensure that for $l \geq m$, the lower bound $Z_{l+1}^{(N)} = \exp\{\kappa^{l+1} Y_{l+1}^{(N)}\} > \exp\{\kappa^{l+1}\varepsilon\} > 2^{\frac{1}{1-\omega}}$ is satisfied. $\square$

Proof of Proposition A.3.2(a). Recall that $\alpha = \frac{1-\varepsilon^5}{\tau-1}$. We write

$$\begin{aligned}\mathbb{P}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \geq \big(l^3 x\big)^{\kappa + c_\gamma \gamma(x)}\big\}\Big) &\leq \mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \geq \big(l^3 x\big)^{\kappa + c_\gamma \gamma(x)}\big\}\Big) + \mathbb{P}(F_{\varepsilon,N^\alpha}^c) \\ &\leq \mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \geq \big(l^3 x\big)^{\kappa + c_\gamma \gamma(x)}\big\}\Big) + l^{-3},\end{aligned} \tag{A.4.12}$$

where $\mathbb{P}_M$ is the conditional probability given that $F_{\varepsilon,M}$ holds, and where we have used Lemma A.3.3 with $N^{-h} < l^{-3}$. It remains to bound the first term on the right-hand side of (A.4.12). For this bound we aim to use Lemma A.1.4. Clearly because loops and cycles can occur,

$$Z_{x,l}^{(N)} \leq \sum_{i=1}^{x} X_{i,l-1}^{(N)}, \tag{A.4.13}$$

where for $1 \leq i \leq x$, $X_{i,l-1}^{(N)}$ denotes the number of brother stubs of the i^{th}-attached node. Since the free stubs of SPG_{l-1} are exchangeable, each free stub will choose any stub with label unequal to 3 with the same probability. Therefore, all $X_{i,l-1}^{(N)}$ have the same law which we denote by $H^{(N)}$. Then we observe that due to (A.3.8), $X_{i,l-1}^{(N)}$ can be coupled with $\overline{X}_{i,l-1}^{(N)}$ having law $\overline{G}_M^{(N)}$, where M is equal to the number of stubs with label 3 at the moment we generate $X_{i,l-1}^{(N)}$, which is at most the number of stubs with label 3 in SPG_l plus 1. The last number is due to Lemma A.3.4 at most $2M_{l-1}^{(N)} + 1$. By Lemma A.4.1(a), we have that

$$2M_{l-1}^{(N)} + 1 \leq 2N^{\frac{1-3\varepsilon^4/4}{\kappa(\tau-1)}} + 1 \leq N^{\frac{1-\varepsilon^5}{\tau-1}} = N^\alpha, \tag{A.4.14}$$

and hence, due to (A.3.8), we can take as the largest possible number $M = N^\alpha$. We now verify whether we can apply Lemma A.1.4(i). Observe that $x \leq N^{\frac{1-\varepsilon/2}{\kappa(\tau-1)}}$ so that for N large and each c_γ, we have

$$y = (l^3 x)^{\kappa + c_\gamma \gamma(x)} < N^\alpha, \tag{A.4.15}$$

since by (A.4.2), we can bound l by $O(\log\log N)$. Hence (A.1.28) holds, because we condition on F_{ε,N^α}. We therefore can apply Lemma A.1.4(i), with $\hat{S}_x^{(N)} = \sum_{i=1}^{x} \overline{X}_{i,l-1}^{(N)}$, $\hat{H}^{(N)} = \overline{G}_{N^\alpha}^{(N)}$, and, also using the upper bound in (A.1.1), we obtain,

$$\begin{aligned}\mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \geq \big(l^3 x\big)^{\kappa + c_\gamma \gamma(x)}\big\}\Big) &\leq b' x [1 + 2N^{-h}]\big[1 - G\big(y\big)\big] \\ \leq 2b' x y^{-\kappa^{-1} + K_\tau \gamma(y)} = 2b' x (l^3 x)^{(-\kappa^{-1} + K_\tau \gamma(y))(\kappa + c_\gamma \gamma(x))} &\leq b l^{-3},\end{aligned} \tag{A.4.16}$$

if we show that

$$c_\gamma \gamma(x)\left(-\kappa^{-1} + K_\tau \gamma(y)\right) + \kappa K_\tau \gamma(y) < 0. \tag{A.4.17}$$

Inequality (A.4.17) holds, because $\gamma(y) = (\log y)^{\gamma-1}$, $\gamma \in [0,1)$, can be made arbitrarily small by taking y large. The fact that y is large follows from (A.4.1(c)) and (A.4.15), and since $l^3 x \geq l^3 \exp\{\kappa^m \varepsilon/2\}$, which can be made large by taking m large. $\square$

Proof of Proposition A.3.2(b). Similarly to (A.4.12), we have

$$\mathbb{P}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \leq \big(\tfrac{x}{l^3}\big)^{\kappa - c_\gamma \gamma(x)}\big\}\Big) \leq \mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \leq \big(\tfrac{x}{l^3}\big)^{\kappa - c_\gamma \gamma(x)}\big\}\Big) + l^{-3}, \tag{A.4.18}$$

and it remains to bound the first term on the right-hand side of (A.4.18). Recall that

$$\underline{Z}_{x,l}^{(N)} = \max_{1\le i\le x} X_{i,l-1}^{(N)},$$

where, for $1 \le i \le x$, $X_{i,l-1}^{(N)}$ is the number of brother stubs of a stub attached to the i^{th} free stub of SPG_{l-1}. Suppose we can bound the first term on the right-hand side of (A.4.18) by bl^{-3}, when $Z_{x,l}^{(N)}$ is replaced by $\underline{Z}_{x,l}^{(N)}$ after adding an extra factor 2, e.g., suppose that

$$\mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{\underline{Z}_{x,l}^{(N)} \le 2\left(\tfrac{x}{l^3}\right)^{\kappa - c_\gamma\gamma(x)}\big\}\Big) \le bl^{-3}. \tag{A.4.19}$$

Then we bound

$$\begin{aligned}
&\mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \le \left(\tfrac{x}{l^3}\right)^{\kappa - c_\gamma\gamma(x)}\big\}\Big) \qquad\qquad (\text{A.4.20})\\
&\quad\le \mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{\underline{Z}_{x,l}^{(N)} \le 2\left(\tfrac{x}{l^3}\right)^{\kappa - c_\gamma\gamma(x)}\big\}\Big)\\
&\qquad+ \mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \le \left(\tfrac{x}{l^3}\right)^{\kappa - c_\gamma\gamma(x)}\big\} \cap \big\{\underline{Z}_{x,l}^{(N)} > 2\left(\tfrac{x}{l^3}\right)^{\kappa - c_\gamma\gamma(x)}\big\}\Big).
\end{aligned}$$

By assumption, the first term is bounded by bl^{-3}, and we must bound the second term. We will prove that the second term in (A.4.20) is equal to 0.

For x sufficiently large we obtain from $l \le C\log x$, $\kappa > 1$, and $\gamma(x) \to 0$,

$$2\left(\frac{x}{l^3}\right)^{\kappa - c_\gamma\gamma(x)} > 6x. \tag{A.4.21}$$

Hence for $x = Z_{l-1}^{(N)} > (\varepsilon - \varepsilon^3)\kappa^{l-1}$, it follows from Lemma A.4.1(d), that $\underline{Z}_{x,l}^{(N)} > 2\left(\tfrac{x}{l^3}\right)^{\kappa - c_\gamma\gamma(x)}$ induces

$$\underline{Z}_{x,l}^{(N)} > 6Z_{l-1}^{(N)} \ge 2M_{l-1}^{(N)} + 2Z_{l-1}^{(N)}. \tag{A.4.22}$$

On the other hand, when $x = N^{\frac{(1-\varepsilon/2)}{\kappa(\tau-1)}} < Z_{l-1}^{(N)}$, then, by Lemma A.4.1(a), and where we use again $l \le C\log x$, $\kappa > 1$, and $\gamma(x) \to 0$,

$$\begin{aligned}
\underline{Z}_{x,l}^{(N)} &\ge 2\left(\tfrac{x}{l^3}\right)^{\kappa - c_\gamma\gamma(x)} = 2\left(\frac{N^{\frac{1-\varepsilon/2}{\kappa(\tau-1)}}}{l^3}\right)^{\kappa - c_\gamma\gamma(x)}\\
&> 2N^{\frac{1-3\varepsilon^4/4}{\kappa(\tau-1)}} + 2N^{\frac{1-\varepsilon/2}{\kappa(\tau-1)}} > 2M_{l-1}^{(N)} + 2x.
\end{aligned} \tag{A.4.23}$$

We conclude that in both cases we have that $\underline{Z}_{x,l}^{(N)} \ge 2M_{l-1}^{(N)} + 2x \ge 2M_{l-2}^{(N)} + 2x$. We claim that the event $\underline{Z}_{x,l}^{(N)} > 2M_{l-2}^{(N)} + 2x$ implies that

$$Z_{x,l}^{(N)} \ge \underline{Z}_{x,l}^{(N)} - x. \tag{A.4.24}$$

Indeed, let $i_0 \in \{1, \ldots, N\}$ be the node such that

$$D_{i_0} = \underline{Z}_{x,l}^{(N)} + 1,$$

and suppose that $i_0 \in \mathrm{SPG}_{l-1}$. Then D_{i_0} is at most the total number of stubs with labels 2 and 3, i.e., at most $2M_{l-2}^{(N)} + 2x$. Hence $\underline{Z}_{x,l}^{(N)} < D_{i_0} \le 2M_{l-2}^{(N)} + 2x$, and this is a contradiction with the assumption that $\underline{Z}_{x,l}^{(N)} > 2M_{l-2}^{(N)} + 2x$. Since by definition $i_0 \in \mathrm{SPG}_l$, we conclude that $i_0 \in \mathrm{SPG}_l \setminus \mathrm{SPG}_{l-1}$, which is equivalent to saying that the chosen stub with $\underline{Z}_{x,l}^{(N)}$ brother stubs had label 1. Then, on $\underline{Z}_{x,l}^{(N)} > 2M_{l-2}^{(N)} + 2x$, we have (A.4.24). Indeed, the one stub from level $l-1$ connected to i_0 gives us $\underline{Z}_{x,l}^{(N)}$ free stubs at level l and the other $x-1$ stubs from level $l-1$ can 'eat up' at most x stubs.

We conclude from the above that

$$\begin{aligned} &\mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \le \big(\tfrac{x}{l^3}\big)^{\kappa - c_\gamma \gamma(x)}\big\} \cap \big\{\underline{Z}_{x,l}^{(N)} > 2\big(\tfrac{x}{l^3}\big)^{\kappa - c_\gamma \gamma(x)}\big\}\Big) \qquad \text{(A.4.25)}\\ &\quad \le \mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{Z_{x,l}^{(N)} \le \big(\tfrac{x}{l^3}\big)^{\kappa - c_\gamma \gamma(x)}\big\} \cap \big\{Z_{x,l}^{(N)} > 2\big(\tfrac{x}{l^3}\big)^{\kappa - c_\gamma \gamma(x)} - x\big\}\Big) = 0, \end{aligned}$$

since (A.4.21) implies that

$$2\left(\frac{x}{l^3}\right)^{\kappa - c_\gamma \gamma(x)} - x \ge \left(\frac{x}{l^3}\right)^{\kappa - c_\gamma \gamma(x)}.$$

This completes the proof that the second term on the right-hand side of (A.4.20) is 0.

We are left to prove that there exists a value of b such that (A.4.19) holds, which we do in two steps. First we couple $\{X_{i,l-1}^{(N)}\}_{i=1}^x$ with a sequence of i.i.d. random variables $\{\underline{X}_{i,l-1}^{(N)}\}_{i=1}^x$ with law $\underline{G}_{N^\alpha}^{(N)}$, such that almost surely,

$$X_{i,l-1}^{(N)} \ge \underline{X}_{i,l-1}^{(N)}, \qquad i = 1, 2, \ldots, x, \tag{A.4.26}$$

and hence

$$\underline{Z}_{x,l}^{(N)} \ge V_x^{(N)} \stackrel{def}{=} \max_{1 \le i \le x} \underline{X}_{i,l-1}^{(N)}. \tag{A.4.27}$$

Then we apply Lemma A.1.4(ii) with $\hat{X}_i^{(N)} = \underline{X}_{i,l-1}^{(N)}$ and $y = 2(x/l^3)^{\kappa - c_\gamma \gamma(x)}$.

We use Lemma A.3.1 to couple $\{X_{i,l-1}^{(N)}\}_{i=1}^x$ with a sequence of i.i.d. random variables $\{\underline{X}_{i,l-1}^{(N)}\}_{i=1}^x$, with law $\underline{G}_{N^\alpha}^{(N)}$. Indeed, Lemma A.3.1 can be applied, when at all times $i = 1, 2, \ldots, x$ sampling is performed, excluding at most N^α stubs with label 3. Since the number of stubs increases with i, we hence have to verify that $M \le N^\alpha$, when M is the maximal possible number of stubs with label 3 at the moment we generate $X_{x,l-1}^{(N)}$. The number M is bounded from above by

$$2M_{l-2}^{(N)} + 2x \le N^{\frac{1-3\varepsilon^4/4}{\kappa(\tau-1)}} + 2x \le N^\alpha,$$

using (A.4.1)(a) for the first inequality and $x \le N^{\frac{(1-\varepsilon/2)}{\kappa(\tau-1)}} \le \frac{1}{4}N^\alpha$ for the second one.

We finally restrict to $x = Z_{l-1}^{(N)} \wedge N^{\frac{1-\varepsilon/2}{\kappa(\tau-1)}}$, as required in Proposition A.3.2(b) . Note that $y = 2(x/l^3)^{\kappa - c_\gamma \gamma(x)} \le N^\alpha$, so that F_{α, N^α} holds, which in turn implies condition (A.1.30). We can therefore apply Lemma A.1.4(ii) with $\hat{X}_i^{(N)} = \underline{X}_{i,l-1}^{(N)}$, $i = 1, 2, \ldots, x$, $\hat{H}^{(N)} = \underline{G}_{N^\alpha}^{(N)}$, and $y = 2(x/l^3)^{\kappa - c_\gamma \gamma(x)}$, to obtain from (A.4.27),

$$\begin{aligned} &\mathbb{P}_{N^\alpha}\Big(\hat{F}_{m,l}(\varepsilon) \cap \big\{\underline{Z}_{x,l}^{(N)} \le 2\big(\tfrac{x}{l^3}\big)^{\kappa - c_\gamma \gamma(x)}\big\}\Big) \\ &\quad \le \mathbb{P}\left(V_x^{(N)} \le y\right) \le \Big(1 - [1 - 2N^{-h}][1 - G(y)]\Big)^x. \end{aligned} \tag{A.4.28}$$

From the lower bound of (A.1.1),

$$[1-G(y)] \geq y^{-\kappa^{-1}-K_\tau\gamma(y)} = 2^{-\kappa^{-1}-K_\tau\gamma(y)}(x/l^3)^{(-\kappa^{-1}-K_\tau\gamma(y))(\kappa-c_\gamma\gamma(x))} \geq \frac{l^3}{2x}, \qquad \text{(A.4.29)}$$

because $x/l^3 > 1$ and

$$\kappa^{-1}c_\gamma\gamma(x) - \kappa K_\tau\gamma(y) + c_\gamma K_\tau\gamma(x)\gamma(y) \geq c_\gamma\kappa^{-1}\gamma(x) - \kappa K_\tau\gamma(y) \geq 0,$$

by choosing c_γ large and using $\gamma(x) \geq \gamma(y)$. Combining (A.4.28) and (A.4.29) and taking $1-2N^{-h} > \frac{1}{2}$, we conclude that

$$\Big(1-[1-2N^{-h}][1-G\left(y\right)]\Big)^x \leq \left(1-\frac{l^3}{4x}\right)^x \leq e^{-l^3/4} \leq l^{-3}, \qquad \text{(A.4.30)}$$

because $l > m$ and m can be chosen large. This yields (**??**) with $b = 1$. □

In the proof of Proposition 3.2, in Section A.5, we often use a corollary of Proposition A.3.2 that we formulate and prove below.

Corollary A.4.2. *Let F satisfy Assumption 1.1(ii). For any $\varepsilon > 0$ sufficiently small, there exists an integer m such that such that for any $k > m$,*

$$\mathbb{P}\left(\hat{F}_{m,k}(\varepsilon) \cap \{|Y_k^{(N)} - Y_{k-1}^{(N)}| > (\tau-2+\varepsilon)^{k(1-\gamma)}\}\right) \leq k^{-2}, \qquad \text{(A.4.31)}$$

for sufficiently large N.

Proof. We use that part (a) and part (b) of Proposition A.3.2 together imply:

$$\mathbb{P}\Big(\hat{F}_{m,k}(\varepsilon) \cap \Big\{\, \big|\log\big(Z_k^{(N)}\big) - \kappa\log\big(Z_{k-1}^{(N)}\big)\big| \geq \kappa\log(k^3) + c_\gamma\gamma\left(Z_{k-1}^{(N)}\right)\log\left(k^3 Z_{k-1}^{(N)}\right)\,\Big\}\Big) \leq 2bk^{-3}. \qquad \text{(A.4.32)}$$

Indeed applying Proposition A.3.2, with $l = k$ and $x = Z_{k-1}^{(N)}$, and hence $Z_{x,k}^{(N)} = Z_k^{(N)}$, yields:

$$\mathbb{P}\Big(\hat{F}_{m,k}(\varepsilon) \cap \big\{Z_k^{(N)} \geq (k^3x)^{\kappa+c_\gamma\gamma(x)}\big\}\Big) \leq bk^{-3}, \qquad \text{(A.4.33)}$$

$$\mathbb{P}\Big(\hat{F}_{m,k}(\varepsilon) \cap \big\{Z_k^{(N)} \leq (x/k^3)^{\kappa-c_\gamma\gamma(x)}\big\}\Big) \leq bk^{-3}, \qquad \text{(A.4.34)}$$

and from the identities

$$\{Z_k^{(N)} \geq (k^3x)^{\kappa+c_\gamma\gamma(x)}\} = \{\log(Z_k^{(N)}) - \kappa\log(Z_{k-1}^{(N)}) \geq \log((k^3x)^{\kappa+c_\gamma\gamma(x)}) - \kappa\log x\},$$
$$\{Z_k^{(N)} \leq (x/k^3)^{\kappa-c_\gamma\gamma(x)}\} = \{\log(Z_k^{(N)}) - \kappa\log(Z_{k-1}^{(N)}) \leq \log((x/k^3)^{\kappa+c_\gamma\gamma(x)}) - \kappa\log x\},$$

we obtain (A.4.32).

Applying (A.4.32) and (A.3.9), we arrive at

$$\begin{aligned}&\mathbb{P}\left(\hat{F}_{m,k}(\varepsilon) \cap \{|Y_k^{(N)} - Y_{k-1}^{(N)}| > (\tau-2+\varepsilon)^{k(1-\gamma)}\}\right) &\text{(A.4.35)}\\ &\quad\leq \mathbb{P}\left(\hat{F}_{m,k}(\varepsilon) \cap \left\{\kappa^{-k}\left[\kappa\log(k^3) + c_\gamma\gamma\left(Z_{k-1}^{(N)}\right)\log\left(k^3 Z_{k-1}^{(N)}\right)\right] > (\tau-2+\varepsilon)^{k(1-\gamma)}\right\}\right) + 2bk^{-3}.\end{aligned}$$

Observe that, due to Lemma A.4.1(c), and since $\gamma(x) = (\log x)^{\gamma-1}$, where $0 \leq \gamma < 1$, we have on $\hat{F}_{m,k}(\varepsilon)$,

$$
\begin{aligned}
&\kappa^{-k}\left[\kappa \log(k^3) + c_\gamma \gamma\left(Z_{k-1}^{(N)}\right) \log\left(k^3 Z_{k-1}^{(N)}\right)\right] \\
&\quad = \kappa^{-k}\left[\kappa \log(k^3) + c_\gamma \left(\log\left(Z_{k-1}^{(N)}\right)\right)^{\gamma-1} \left(\log(k^3) + \log\left(Z_{k-1}^{(N)}\right)\right)\right] \\
&\quad \leq \kappa^{-k}\left[\kappa \log(k^3) + c_\gamma \log(k^3) + c_\gamma \left(\log\left(Z_{k-1}^{(N)}\right)\right)^{\gamma}\right] \\
&\quad \leq \kappa^{-k}\left[(c_\gamma + \kappa)\log(k^3) + c_\gamma \left(\kappa^{k-1}(\varepsilon^{-1} + \varepsilon^3)\right)^{\gamma}\right] \\
&\quad \leq \kappa^{-k(1-\gamma)}\left[\kappa^{-k\gamma}(c_\gamma + \kappa)\log(k^3) + c_\gamma \left(\kappa^{-1}(\varepsilon^{-1} + \varepsilon^3)\right)^{\gamma}\right] \leq (\tau - 2 + \varepsilon)^{k(1-\gamma)},
\end{aligned}
$$

because, for k large, and since $\kappa^{-1} = \tau - 2$,

$$
\left(\tfrac{\tau-2}{\tau-2+\varepsilon}\right)^{k(1-\gamma)} \left[\kappa^{-k\gamma}(c_\gamma + \kappa)\log(k^3)\right] \leq \tfrac{1}{2}, \qquad \left(\tfrac{\tau-2}{\tau-2+\varepsilon}\right)^{k(1-\gamma)} c_\gamma \left(\kappa^{-1}(\varepsilon^{-1} + \varepsilon^3)\right)^{\gamma} \leq \tfrac{1}{2}.
$$

We conclude that the first term on the right-hand side of (A.4.35) is 0, for sufficiently large k, and the second term is bounded by $2bk^{-3} \leq k^{-2}$, and hence the statement of the corollary follows. □

A.5 Proof of Proposition 3.2 and Proposition 3.4

Proof of Proposition 3.2(a). We have to show that

$$
\mathbb{P}\Big(\varepsilon \leq Y_m^{(i,N)} \leq \varepsilon^{-1}, \max_{k \in \mathcal{T}_m^{(i,N)}(\varepsilon)} |Y_k^{(i,N)} - Y_m^{(i,N)}| > \varepsilon^3\Big) = o_{N,m,\varepsilon}(1).
$$

Fix $\varepsilon > 0$, such that $\tau - 2 + \varepsilon < 1$. Then, take $m = m_\varepsilon$, such that (A.3.10) holds, and increase m, if necessary, until (A.4.31) holds.

We use the inclusion

$$
\Big\{\max_{k \in \mathcal{T}_m^{(N)}(\varepsilon)} |Y_k^{(N)} - Y_m^{(N)}| > \varepsilon^3\Big\} \subseteq \Big\{\sum_{k \in \mathcal{T}_m^{(N)}(\varepsilon)} |Y_k^{(N)} - Y_{k-1}^{(N)}| > \sum_{k \geq m} (\tau - 2 + \varepsilon)^{k(1-\gamma)}\Big\}. \tag{A.5.1}
$$

If the event on the right-hand side of (A.5.1) holds, then there must be a $k \in \mathcal{T}_m^{(N)}(\varepsilon)$ such that $|Y_k^{(N)} - Y_{k-1}^{(N)}| > (\tau - 2 + \varepsilon)^{k(1-\gamma)}$, and therefore

$$
\{\max_{k \in \mathcal{T}_m^{(N)}(\varepsilon)} |Y_k^{(N)} - Y_m^{(N)}| > \varepsilon^3\} \subseteq \bigcup_{k \in \mathcal{T}_m^{(N)}(\varepsilon)} G_{m,k-1} \cap G_{m,k}^c, \tag{A.5.2}
$$

where we denote

$$
G_{m,k} = G_{m,k}(\varepsilon) = \bigcap_{j=m+1}^{k} \Big\{|Y_j^{(N)} - Y_{j-1}^{(N)}| \leq (\tau - 2 + \varepsilon)^{j(1-\gamma)}\Big\}. \tag{A.5.3}
$$

Since (A.3.10) implies that on $G_{m,k-1}$ we have $|Y_j^{(N)} - Y_m^{(N)}| \leq \varepsilon^3$, $m < j \leq k-1$, we find that,

$$G_{m,k-1} \cap G_{m,k}^c \subseteq \Big\{|Y_l^{(N)} - Y_m^{(N)}| \leq \varepsilon^3, \forall l\, : m < l \leq k-1\Big\} \cap \Big\{|Y_k^{(N)} - Y_{k-1}^{(N)}| > (\tau-2+\varepsilon)^{k(1-\gamma)}\Big\}. \tag{A.5.4}$$

Take N sufficiently large such that, by Proposition 3.1,

$$\begin{aligned}\mathbb{P}\left(M_m^{(N)} > 2Z_m^{(N)}, \varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}\right) &\leq \mathbb{P}\left(\exists l \leq m : Y_l^{(N)} \neq Y_l\right) + \mathbb{P}\left(M_m > 2\mathcal{Z}_m, \varepsilon \leq Y_m \leq \varepsilon^{-1}\right)\\ &\leq \mathbb{P}\left(M_m > 2\mathcal{Z}_m, \varepsilon \leq Y_m \leq \varepsilon^{-1}\right) + \varepsilon/4,\end{aligned} \tag{A.5.5}$$

where we recall the definition of $M_m = \sum_{j=1}^{m} \mathcal{Z}_j$ in (A.3.11). Next, we use that

$$\lim_{m\to\infty} \mathbb{P}\left(M_m > 2\mathcal{Z}_m, \varepsilon \leq Y_m \leq \varepsilon^{-1}\right) = 0, \tag{A.5.6}$$

since $Y_l = (\tau-2)^l \log \mathcal{Z}_l$ converges a.s., so that when $Y_m \geq \varepsilon$ and m is large, M_{m-1} is much smaller than Z_m, so that $M_m = M_{m-1} + \mathcal{Z}_m > 2\mathcal{Z}_m$ has small probability, as m is large.

Then we use (A.5.1)–(A.5.6), together with (A.3.12), to derive that

$$\begin{aligned}&\mathbb{P}\left(\varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}, \max_{k\in\mathcal{T}_m^{(N)}(\varepsilon)} |Y_k^{(N)} - Y_m^{(N)}| > \varepsilon^3\right) \\ &\quad\leq \mathbb{P}\left(\varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}, \max_{k\in\mathcal{T}_m^{(N)}(\varepsilon)} |Y_k^{(N)} - Y_m^{(N)}| > \varepsilon^3,\, Y_l^{(N)} = Y_l,\, \forall l \leq m\right) + \mathbb{P}\left(\exists l \leq m : Y_l^{(N)} \neq Y_l\right)\\ &\quad\leq \sum_{k>m} \mathbb{P}\left(G_{m,k-1} \cap G_{m,k}^c \cap \{k \in \mathcal{T}_m^{(N)}(\varepsilon)\} \cap \{\varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}\} \cap \{Y_l^{(N)} = Y_l,\, \forall l \leq m\}\right) + \frac{\varepsilon}{2}\\ &\quad\leq \sum_{k>m} \mathbb{P}\left(\hat{F}_{m,k}(\varepsilon) \cap \Big\{|Y_k^{(N)} - Y_{k-1}^{(N)}| > (\tau-2+\varepsilon)^{k(1-\gamma)}\Big\}\right) + \varepsilon < 3\varepsilon/2,\end{aligned} \tag{A.5.7}$$

by Corollary A.4.2 and (A.3.10). □

Proof of Proposition 3.2(b). We first show (3.6), then (3.5). Due to Proposition 3.2(a), and using that $\{Y_m^{(N)} \geq \varepsilon\}$, we find

$$Y_k^{(N)} \leq Y_m^{(N)} + \varepsilon^3 \leq Y_m^{(N)}(1+\varepsilon^2),$$

apart from an event with probability $o_{N,m,\varepsilon}(1)$, for all $k \in \mathcal{T}_m^{(N)}$. By (A.3.9) and because $k \in \mathcal{T}_m^{(N)}$, this is equivalent to

$$Z_k^{(N)} \leq \left(Z_m^{(N)}\right)^{\kappa^{k-m}(1+\varepsilon^2)} \leq N^{\frac{1-\varepsilon^2}{\tau-1}(1+\varepsilon^2)} = N^{\frac{1-\varepsilon^4}{\tau-1}},$$

which implies (3.6).

We next show (3.5). Observe that $k \in \mathcal{T}_m^{(N)}$ implies that either $k-1 \in \mathcal{T}_m^{(N)}$, or $k-1 = m$. Hence, from $k \in \mathcal{T}_m^{(N)}$ and Proposition 3.2(a), we obtain, apart from an event with probability $o_{N,m,\varepsilon}(1)$,

$$Y_{k-1}^{(N)} \geq Y_m^{(N)} - \varepsilon^3 \geq \varepsilon - \varepsilon^3 \geq \frac{\varepsilon}{2}, \tag{A.5.8}$$

for $\varepsilon > 0$ sufficiently small, and

$$Y_k^{(N)} = Y_k^{(N)} - Y_m^{(N)} + Y_m^{(N)} - Y_{k-1}^{(N)} + Y_{k-1}^{(N)} \geq Y_{k-1}^{(N)} - 2\varepsilon^3 \geq Y_{k-1}^{(N)}(1 - 4\varepsilon^2). \tag{A.5.9}$$

By (A.3.9) this is equivalent to

$$Z_k^{(N)} \geq \big(Z_{k-1}^{(N)}\big)^{\kappa(1-4\varepsilon^2)} \geq Z_{k-1}^{(N)},$$

when $\varepsilon > 0$ is so small that $\kappa(1 - 4\varepsilon^2) \geq 1$, since $\tau \in (2, 3)$, and $\kappa = (\tau - 2)^{-1}$. □

Proof of Proposition 3.4. We must show that

$$\mathbb{P}\Big(k \in \partial\mathcal{T}_m^{(N)}(\varepsilon), \varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}, Z_{k+1}^{(N)} \leq N^{\frac{1-\varepsilon}{\tau-1}}\Big) = o_{N,m,\varepsilon}(1), \tag{A.5.10}$$

where we recall that

$$\{k \in \partial\mathcal{T}_m^{(N)}\} = \{k \in \mathcal{T}_m^{(N)}\} \cap \{k + 1 \notin \mathcal{T}_m^{(N)}\}.$$

In the proof, we will make repeated use of Propositions 3.1 and 3.2, whose proofs are now complete.

According to the definition of $\hat{F}_{m,k}(\varepsilon)$ in (A.3.12),

$$\begin{aligned} &\mathbb{P}\Big(\{k \in \partial\mathcal{T}_m^{(N)}(\varepsilon)\} \cap \{\varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}\} \cap \hat{F}_{m,k}(\varepsilon)^c\Big) \qquad \text{(A.5.11)}\\ &\quad\leq \mathbb{P}\Big(\varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}, \max_{l \in \mathcal{T}_m^{(N)}(\varepsilon)} |Y_l^{(N)} - Y_m^{(N)}| > \varepsilon^3\Big) + \mathbb{P}\Big(\varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}, M_m^{(N)} > 2Z_m^{(N)}\Big). \end{aligned}$$

In turn Propositions 3.2(a), as well as (A.5.5–A.5.6) imply that both probabilities on the right-hand side of (A.5.11) are $o_{N,m,\varepsilon}(1)$. Therefore, it suffices to show that

$$\begin{aligned} &\mathbb{P}\Big(\{k \in \partial\mathcal{T}_m^{(N)}(\varepsilon), \varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}, Z_{k+1}^{(N)} \leq N^{\frac{1-\varepsilon}{\tau-1}}\} \cap \hat{F}_{m,k}(\varepsilon)\Big)\\ &\quad = \mathbb{P}\Big(\{k + 1 \notin \mathcal{T}_m^{(N)}(\varepsilon), Z_{k+1}^{(N)} \leq N^{\frac{1-\varepsilon}{\tau-1}}\} \cap \hat{F}_{m,k}(\varepsilon)\Big) = o_{N,m,\varepsilon}(1). \qquad \text{(A.5.12)} \end{aligned}$$

Let $x = N^{\frac{1-\varepsilon/2}{\kappa(\tau-1)}}$, and define the event $I_{N,k} = I_{N,k}(a) \cap I_{N,k}(b) \cap I_{N,k}(c) \cap I_{N,k}(d)$, where

$$\begin{aligned} I_{N,k}(a) &= \{M_{k-1}^{(N)} < N^{\frac{1-3\varepsilon^4/4}{\kappa(\tau-1)}}\}, & \text{(A.5.13)}\\ I_{N,k}(b) &= \{x \leq Z_k^{(N)}\}, & \text{(A.5.14)}\\ I_{N,k}(c) &= \{Z_k^{(N)} \leq N^{\frac{1-\varepsilon^4}{\tau-1}}\}, & \text{(A.5.15)}\\ I_{N,k}(d) &= \{Z_{k+1}^{(N)} \geq Z_{x,k+1}^{(N)} - Z_k^{(N)}\}. & \text{(A.5.16)} \end{aligned}$$

We split

$$\begin{aligned} &\mathbb{P}\big(\{k + 1 \notin \mathcal{T}_m^{(N)}(\varepsilon), Z_{k+1}^{(N)} \leq N^{\frac{1-\varepsilon}{\tau-1}}\} \cap \hat{F}_{m,k}(\varepsilon)\big) \qquad \text{(A.5.17)}\\ &\quad = \mathbb{P}\big(\{k + 1 \notin \mathcal{T}_m^{(N)}(\varepsilon), Z_{k+1}^{(N)} \leq N^{\frac{1-\varepsilon}{\tau-1}}\} \cap \hat{F}_{m,k}(\varepsilon) \cap I_{N,k}\big)\\ &\qquad + \mathbb{P}\big(\{k + 1 \notin \mathcal{T}_m^{(N)}(\varepsilon), Z_{k+1}^{(N)} \leq N^{\frac{1-\varepsilon}{\tau-1}}\} \cap \hat{F}_{m,k}(\varepsilon) \cap I_{N,k}^c\big). \end{aligned}$$

We claim that both probabilities are $o_{N,m,\varepsilon}(1)$, which would complete the proof. We start to show that

$$\mathbb{P}\Big(\{k+1 \notin \mathcal{T}_m^{(N)}(\varepsilon), Z_{k+1}^{(N)} \leq N^{\frac{1-\varepsilon}{\tau-1}}\} \cap \hat{F}_{m,k}(\varepsilon) \cap I_{N,k}\Big) = o_{N,m,\varepsilon}(1). \tag{A.5.18}$$

Indeed, by (A.3.12), (3.6), and Lemma 5.1, with $u = (\tau-1)^{-1}$, for the second inequality,

$$\begin{aligned}
&\mathbb{P}\Big(\{k+1 \notin \mathcal{T}_m^{(N)}(\varepsilon)\} \cap \{Z_k^{(N)} \geq N^{\frac{1-\varepsilon}{\tau-1}}\} \cap \hat{F}_{m,k}(\varepsilon) \cap I_{N,k}\Big) \qquad \text{(A.5.19)}\\
&\quad \leq \mathbb{P}\Big(\{k \in \mathcal{T}_m^{(N)}(\varepsilon)\} \cap \{\varepsilon \leq Y_m^{(N)} \leq \varepsilon^{-1}\} \cap \{Z_k^{(N)} \in [N^{\frac{1-\varepsilon}{\tau-1}}, N^{\frac{1-\varepsilon^4}{\tau-1}}]\}\Big) + o_{N,m,\varepsilon}(1) = o_{N,m,\varepsilon}(1).
\end{aligned}$$

Therefore, we are left to deal with the case where $Z_k^{(N)} \leq N^{\frac{1-\varepsilon}{\tau-1}}$. For this, and assuming $I_{N,k}$, we can use Proposition A.3.2(b) with $x = N^{\frac{1-\varepsilon/2}{\kappa(\tau-1)}} \leq Z_k^{(N)}$ by $I_{N,k}(b)$, and $l = k+1$ to obtain that, **whp**,

$$Z_{k+1}^{(N)} \geq Z_{x,k+1}^{(N)} - Z_k^{(N)} \geq x^{\kappa(1-\varepsilon/2)} - N^{\frac{1-\varepsilon}{\tau-1}} = N^{\frac{(1-\varepsilon/2)^2}{\tau-1}} - N^{\frac{1-\varepsilon}{\tau-1}} > N^{\frac{1-\varepsilon}{\tau-1}}, \tag{A.5.20}$$

where we have used that when $k \in \mathcal{T}_m^{(N)}(\varepsilon)$ and $Y_m^{(N)} > \varepsilon$, then we have $k \leq c \log\log N$, for some $c = c(\tau, \varepsilon)$, and hence, for N large enough,

$$(k+1)^{3(\kappa - c_\gamma \gamma(x))} x^{c_\gamma \gamma(x)} \leq (k+1)^{3\kappa} x^{c_\gamma \gamma(x)} \leq x^{\varepsilon\kappa/2}.$$

This proves (A.5.18).

For the second probability on the right-hand side of (A.5.17) it suffices to prove that

$$\mathbb{P}\big(\{k+1 \notin \mathcal{T}_m^{(N)}(\varepsilon)\} \cap \hat{F}_{m,k}(\varepsilon) \cap I_{N,k}^c\big) = o_{N,m,\varepsilon}(1). \tag{A.5.21}$$

In order to prove (A.5.21), we prove that (A.5.21) holds with $I_{N,k}^c$ replaced by each one of the four events $I_{N,k}^c(a), \ldots, I_{N,k}^c(d)$. For the intersection with the event $I_{N,k}^c(a)$, we apply Lemma A.4.1(a), which states that $\hat{F}_{m,k}(\varepsilon) \cap I_{N,k}^c(a)$ is the empty set.

It follows from (3.3) that if $k+1 \notin \mathcal{T}_m^{(N)}(\varepsilon)$, then

$$\kappa^{k+1} Y_m^{(N)} > \frac{1-\varepsilon^2}{\tau-1} \log N. \tag{A.5.22}$$

If $\hat{F}_{m,k}(\varepsilon)$ holds then by definition (A.3.12), and Corollary A.4.2, **whp**,

$$Y_k^{(N)} \geq Y_{k-1}^{(N)} \geq Y_m^{(N)} - \varepsilon^3 \geq Y_m^{(N)}(1-\varepsilon^2). \tag{A.5.23}$$

Hence, if $\hat{F}_{m,k}(\varepsilon)$ holds and $k+1 \notin \mathcal{T}_m^{(N)}(\varepsilon)$, then, by (A.5.22)–(A.5.23), **whp**,

$$\kappa \log(Z_k^{(N)}) = \kappa^{k+1} Y_k^{(N)} \geq (1-\varepsilon^2)\kappa^{k+1} Y_m^{(N)} \geq \tfrac{(1-\varepsilon^2)^2}{\tau-1} \log N, \tag{A.5.24}$$

so that, **whp**,

$$Z_k^{(N)} \geq x = N^{\frac{1-\varepsilon/2}{\kappa(\tau-1)}}, \tag{A.5.25}$$

for small enough $\varepsilon > 0$ and sufficiently large N, i.e., we have

$$\mathbb{P}(\{k+1 \notin \mathcal{T}_m^{(N)}(\varepsilon)\} \cap \hat{F}_{m,k}(\varepsilon) \cap I_{N,k}^c(b)) = o_{N,m,\varepsilon}(1).$$

From Proposition 3.2(b) it is immediate that

$$\mathbb{P}(\{k+1 \notin \mathcal{T}_m^{(N)}(\varepsilon)\} \cap \hat{F}_{m,k}(\varepsilon) \cap I_{N,k}^c(c)) = o_{N,m,\varepsilon}(1).$$

Finally, recall that $Z_{x,k+1}^{(N)}$ is the number of constructed free stubs at level $k+1$ after pairing of the first x stubs at level k. The pairing of the remaining $Z_k^{(N)} - x$ stubs at level k can 'eat up' at most $Z_k^{(N)} - x \leq Z_k^{(N)}$ stubs, so that $I_{N,k}(d)$ holds with probability 1.

This completes the proof of (A.5.21) and hence of Proposition 3.4. □

Acknowledgement

We thank Eric Cator for useful discussions on uniform continuity used in the proof of Lemma 5.1. The work of all authors was performed in part at the Mittag-Leffler Institute. The work of RvdH was performed in part also at Microsoft Research.

References

[1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Modern Phys.,* **74**, 47-97, 2002. MR1895096

[2] N. Alon and J.H. Spencer. *The Probabilistic Method.* 2$^{\text{nd}}$ *Edition.* Wiley, New York, 2000. MR1885388

[3] B. Bollobás. *Random Graphs,* 2nd edition. Cambridge University Press, Cambridge, 2001. MR1864966

[4] B. Bollobás, S. Janson and O. Riordan. The phase transition in inhomogeneous random graphs. To appear in Random Structures Algorithms. Available on http://arxiv.org/abs/math.PR/0504589, 2006.

[5] T. Britton, M. Deijfen, and A. Martin-Löf. Generating simple random graphs with arbitrary degree sequences, Journal of Statistical Physics **124**(6), 1377-1397, 2006. MR2266448

[6] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci.,* **25**, 15879-15882, 2002. MR1944974

[7] R. Cohen and S. Havlin. Scale free networks are ultrasmall. *Phys. Rev. Lett.*, **90**, 058701, 1-4, 2003.

[8] P.L. Davies. The simple branching process: a note on convergence when the mean is infinite. J. Appl. Probab. **15**, 466-480, 1978. MR0483050

[9] R. Durrett. *Random Graph Dynamics.* Cambridge University Press, Cambridge (2006). MR2271734

[10] H. van den Esker, R. van der Hofstad, G. Hooghiemstra and D. Znamenski. Distances in random graphs with infinite mean degrees. *Extremes,* **8**, 111-141, 2006. MR2275914

[11] A. Gut. *Probability: A graduate course.* Springer, 2005. MR2125120

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

[13] W. Feller. *An Introduction to Probability Theory and Its Applications.* Volume II, 2nd edition, Wiley, New York, 1971. MR0270403

[14] R. van der Hofstad, G. Hooghiemstra and P. Van Mieghem. Random graphs with finite variance degrees. *Random Structures Algorithms*, **27**(1), 76-123, 2005. MR2150017

[15] R. van der Hofstad, G. Hooghiemstra and D. Znamenski. Random graphs with arbitrary i.i.d. degrees, Available on `http://arxiv.org/abs/math.PR/0502580`, 2006.

[16] S. Janson. On concentration of probability. Contemporary Combinatorics, ed. B. Bollobás, Bolyai Soc. Math. Stud. **10**, 289-301, János Bolyai Mathematical Society, Budapest, 2002. MR1919574

[17] S. Janson, T. Łuczak and A. Rucinski. *Random Graphs.* Wiley, New York, 2000. MR1782847

[18] M.E.J. Newman. The structure and function of complex networks. *SIAM Rev.* **45**(2), 167-256, 2003. MR2010377

[19] M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Random graphs with arbitrary degree distribution and their application. *Phys. Rev. E*, **64**, 026118, 2001.

[20] I. Norros and H. Reittu. On a conditionally Poissonean graph process. *Adv. Appl. Prob.* **38**, 59-75, 2006. MR2213964

[21] H. Reittu and I. Norros. On the power law random graph model of massive datanetworks. *Performance Evalution*, **55** (1-2), 3-23, 2004.

[22] H-J. Schuh and A. D. Barbour. On the asymptotic behaviour of branching processes with infinite mean. *Adv. Appl. Probab.*, **9**, 681-723, 1977. MR0478378

[23] E. Seneta. The simple branching processes with infinite mean.I. *J. Appl. Probab.*, **10**, 206-212, 1973. MR0413293

[24] H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Network topology generators: Degree-based vs. structural. *Proceedings of ACM Sigcomm, Pittsburgh, Pennsylvania, USA*, 147-159, 2002.

[25] D. J. Watts. *Small Worlds, The Dynamics of Networks between Order and Randomness.* Princeton University Press, Princeton, New Jersey, 1999. MR1716136