Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 104, 1–32.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1160

Quenched critical percolation on Galton–Watson trees

Eleanor Archer* Quirin Vogel†

Abstract

We consider critical percolation on a supercritical Galton–Watson tree. We show that, when the offspring distribution is in the domain of attraction of an α-stable law for some $\alpha \in (1,2)$, or has finite variance, several annealed properties also hold in a quenched setting. In particular, the following properties hold for the critical root cluster on almost every realisation of the tree: (1) the rescaled survival probabilities converge; (2) the Yaglom limit or its stable analogue hold – in particular, conditioned on survival, the number of vertices at generation n that are connected to the root cluster rescale to certain (explicit) random variable; (3) conditioned on initial survival, the sequence of generation sizes in the root cluster rescales to a continuous–state branching process. This strengthens some earlier results of Michelen (2019) who proved (1) and (2) in the case where the initial tree has an offspring distribution with all moments finite.

Keywords: critical percolation; incipient infinite cluster; branching processes.
MSC2020 subject classifications: Primary 60K35; 60J80, Secondary 60G52.
Submitted to EJP on December 20, 2023, final version accepted on June 11, 2024.

1 Introduction

Let $\mathbf{T}$ be a supercritical Galton–Watson tree with law $\mathbf{P}$ such that the law of its offspring distribution is in the domain of attraction of a stable law with parameter $\alpha \in (1,2)$, or has finite variance, in which case we set $\alpha = 2$. Suppose moreover that the offspring distribution is supported on $\{1,2,\dots\}$, and that $\mu > 1$ is its mean. The restriction to distributions with at least one descendant is for technical reasons and is discussed later in the introduction, after Assumption 1.1. The aim of this paper is to obtain quenched results for critical (Bernoulli) percolation on $\mathbf{T}$.

Let $\mathbf{T}_n$ be the vertices of generation n in $\mathbf{T}$ and denote its cardinality by $|\mathbf{T}_n|$. It is well-known that there is a random variable $\mathbf{W}$ such that

$$\mathbf{W}_n := \frac{|\mathbf{T}_n|}{\mu^n} \to \mathbf{W}\,, \tag{1.1}$$

*Modal'X, UMR CNRS 9023, UPL, Université Paris Nanterre, 92000 Nanterre, France.
E-mail: earcher@parisnanterre.fr

†Department of Mathematics, CIT, Technische Universität München, Boltzmannstr. 3, D-85748, Garching bei München, Germany. E-mail: quirinvogel@outlook.com

as $n \to \infty$ almost surely and in L^p if $\mathbf{E}\left[|\mathbf{T}_1|^p\right] < \infty$, see [5, Theorems 0 and 5] (in particular this holds whenever $p < \alpha$). Moreover, it was shown by Lyons [13, Theorem 6.2 and Proposition 6.4] that $\mathbf{P}$-almost surely, the critical percolation probability on $\mathbf{T}$ is equal to $\frac{1}{\mu}$. Given $\mathbf{T}$, we run critical percolation (i.e. with retention probability $\frac{1}{\mu}$) on the edges of $\mathbf{T}$. Let $\mathbb{P}_{\mathbf{T}}$ denote the (random) law of percolation on $\mathbf{T}$ and consider the (critical) root cluster on $\mathbf{T}$. Under the *annealed* law (i.e. sampling first the tree then running the percolation process and then averaging over the tree), denoted $\mathbb{P} = \mathbf{P} \circ \mathbb{P}_{\mathbf{T}}$, the root cluster has the law of a critical Galton–Watson tree with offspring distribution in the domain of attraction an α-stable law (with the same α, or with $\alpha = 2$ in the finite variance case). As a result, precise asymptotics for the annealed critical cluster are well-known. The purpose of this paper is to show that the same results hold in the *quenched* setting, i.e. for $\mathbf{P}$-almost every realisation of $\mathbf{T}$. We note that some of our results were previously obtained by Michelen [15] under the assumption that the offspring distribution has all moments finite.

We start with a brief recap of known results in the annealed setting.

Firstly, for each $n \geq 0$ let $Y_n = |\{v \in \mathbf{T}_n \text{ connected to the root}\}|$, and set $\beta = \frac{1}{\alpha-1}$ if $\alpha < 2$ and $\beta = 1$ if the variance is finite. Then there exists an explicit constant $C_\alpha \in (0, \infty)$, depending on the offspring distribution, such that

$$n^\beta \mathbb{P}\left(Y_n > 0\right) \to C_\alpha \, . \tag{1.2}$$

This result was first obtained by Kolmogorov [10] in the case where the offspring distribution has has finite third moment, then by Kesten, Ney and Spitzer [9] under a finite variance assumption, and was finally extended by Slack [18, Lemma 2] to the stable window.

One can also say a lot more. In the finite variance case, Kesten, Ney and Spitzer [9] also showed that the law of $n^{-1}Y_n$ conditioned on the event $Y_n > 0$ converges in distribution to an exponential random variable. This is in fact known as *Yaglom's limit* as it was first proved by Yaglom [19] under a third moment assumption. In the stable case, this was again extended by Slack [18, Theorem 1] who showed that the rescaled random variables $\left(n^{-\beta}Y_n\right)_{n\geq 1}$ under the conditioning $Y_n > 0$ converge to an α-stable random variable Y via the following convergence of the Laplace transforms: for all $\theta > 0$,

$$\mathbb{E}\left[\mathrm{e}^{-\theta n^{-\beta} Y_n} \mid Y_n > 0\right] \to \phi(\theta) := 1 - C_\alpha^{-1}\theta(1 + (C_\alpha^{-1}\theta)^{\alpha-1})^{-\beta}. \tag{1.3}$$

Moreover, in the finite variance case, the limiting random variable Y has the exponential distribution with parameter C_α and

$$\mathbb{E}\left[Y\right] = \lim_{\theta \to 0} \frac{1 - \phi(\theta)}{\theta} = C_\alpha^{-1}. \tag{1.4}$$

One can in fact strengthen this last result quite dramatically to ask about convergence of the whole sequence of generation sizes under the conditioning that $Y_{n\varepsilon} > 0$ for some fixed $\varepsilon > 0$. Under this conditioning, the process $(n^{-\beta}Y_{n(t+\varepsilon)})_{t\geq 0}$ converges to a *continuous–state branching process* (CSBP) with initial condition determined by (1.3). More precisely, the limiting process $(Y_t)_{t\geq 0}$ is a Markov process such that Y_0 is a random variable with Laplace transform given by $\phi(\theta\varepsilon^\beta)$, and for all $t, x, y \geq 0$ its transition kernels $P_t(x, y)$ satisfy the *branching property*

$$P_t(x + y, \cdot) = P_t(x, \cdot) * P_t(y, \cdot). \tag{1.5}$$

The form of the transition kernel P_t above is determined by a function known as the *branching mechanism* which takes the form $\psi(\lambda) = \tilde{c}\lambda^\alpha$ in the stable case, and $\psi(\lambda) = \tilde{c}\lambda^2$ in the finite variance case, where $\tilde{c}$ is a constant depending on the precise form of the offspring distribution; we compute it in Lemma A.1. We refer to [12] for further background on CSBPs.

Finally, rather than conditioning on the event $\{Y_n > 0\}$, one can also look at the effect of conditioning on survival to infinity, and construct the corresponding law. In the annealed model, the law of the root cluster conditioned to survive forever has the law of a critical Galton–Watson process conditioned to survive forever and was constructed by Kesten [8, Lemma 1.14]. The critical cluster conditioned to survive is known as the *incipient infinite cluster (IIC)*.

Our goal in this paper is to establish quenched versions of the above results. First steps in this direction were achieved by Michelen [15, Theorem 1.3] who proved that a quenched version of (1.2) holds under a fourth moment assumption on the offspring distribution, and that a quenched version of (1.3) holds under the assumption that all moments are finite. In this paper we extend this to the general case of finite variance or stable tails. In particular we will work under the following assumption for all of our results.

Assumption 1.1. *Assume that the offspring distribution of* $\boldsymbol{T}$ *is supported on* $\{1, 2, \ldots\}$*, its mean is given by* $\mu > 1$*, and that one of the following conditions holds.*

(a) The offspring distribution of $\boldsymbol{T}$ *has finite variance. In this case set* $\alpha = 2$*,* $C_\alpha = \frac{2\boldsymbol{E}[|\boldsymbol{T}_1|]^2}{\boldsymbol{E}[|\boldsymbol{T}_1|(|\boldsymbol{T}_1|-1)]}$*,* $\beta = 1$ *and* $\phi(\theta) = C_\alpha/(\theta + C_\alpha)$*.*

(b) The offspring distribution of $\boldsymbol{T}$ *has infinite variance with stable (power-law) tails, meaning that there exist* $c_1 \in (0, \infty)$ *and* $\alpha \in (1, 2)$ *such that* $\boldsymbol{P}(|\boldsymbol{T}_1| \geq x) \sim c_1 x^{-\alpha}$ *as* $x \to \infty$*. Here, we write* $a_n \sim b_n$ *for two sequences* $(a_n)_n$ *and* $(b_n)_n$ *if* $a_n = b_n(1+o(1))$*. In this case, we set* $\beta = (\alpha - 1)^{-1}$*,*

$$C_\alpha = c_1^{-\beta} \mu^{\alpha\beta} \Gamma(1-\alpha)^{-\beta} \beta^\beta \qquad \text{and} \qquad \phi(\theta) = 1 - C_\alpha^{-1}\theta(1 + (C_\alpha^{-1}\theta)^{\alpha-1})^{-\beta}. \tag{1.6}$$

The derivation of the constant C_α and the Laplace transform are in the Appendix; see Lemma A.1. We would like to point out that we could also deal with the case where $\alpha = 2$ and/or where there are slowly-varying corrections to the tail decay in case (b) above (using exactly the same proofs), but have omitted this in order to lighten the notation.

We furthermore note that the assumption that $\mathbf{T}$ has no leaves is not a serious restriction since any supercritical Galton–Watson tree with leaves can be decomposed into a core consisting of a supercritical Galton–Watson tree with no leaves, to which several finite subcritical Galton–Watson trees are attached (this is the Harris decomposition; see [14, Proposition 5.28]). The p.g.f. of the law of the core as well as the distribution of the additional subcritical trees can be obtained explicitly from the original probability distribution. We anticipate that this can be used to extend our theorems to the case where leaves are permitted (though it requires some work and would impact the constants).

The first result is the following.

Theorem 1.2. *We have that for* $\boldsymbol{P}$*-almost every* $\boldsymbol{T}$*, as* $n \to \infty$

$$n^\beta \mathbb{P}_{\mathbf{T}}(Y_n > 0) \to C_\alpha \boldsymbol{W}, \tag{1.7}$$

where $\boldsymbol{W}$ *is as defined in* (1.1)*.*

In the finite variance case this result was also recently recovered as a special case of [1, Theorem 2.6].

In addition, we obtain convergence of the rescaled generation sizes.

Theorem 1.3. *For* $\boldsymbol{P}$*-almost every* $\boldsymbol{T}$*, we have that the law of* Y_n *conditioned on survival converges, i.e.*

$$(Y_n | Y_n > 0) \xrightarrow[n\to\infty]{(d)} Y\,, \tag{1.8}$$

where Y *has Laplace transform* ϕ *as in* (1.3)*.*

We also add to this with convergence of the whole branching process. This was not considered in [15].

Theorem 1.4. *For $\boldsymbol{P}$-almost every $\boldsymbol{T}$, under the conditioning $Y_n > 0$ we have that the process $(n^{-\beta} Y_{n(t+1)})_{t\geq 0}$ converges in distribution to a continuous state branching process with the same branching mechanism ψ appearing below* (1.5)*, and with initial condition determined by Theorem 1.3. This convergence holds with respect to the Skorokhod-J_1 topology on the space $D([0,\infty),[0,\infty))$.*

We anticipate that it should also be possible to prove that the process $(n^{-\beta} Y_{nt})_{t\geq 0}$ converges to its annealed analogue, but the proof is more technically complicated (due to conditioning on a future event) and does not seem to add further insight.

Finally, we turn to the incipient infinite cluster (IIC), which we define as the distribution of the tree conditioned to survive "to infinity":

$$\mu_{\mathbf{T}}(\,\cdot\,) := \lim_{M\to\infty} \mathbb{P}_{\mathbf{T}}\left(\,\cdot\,|Y_M > 0\right). \tag{1.9}$$

Michelen shows in [15, Lemma 3.9] that one can construct the IIC measure in the quenched setting essentially provided that Theorem 1.2 holds (in his case this requires a fourth moment assumption; his proof works under our assumptions as well); in particular, for every tree $t \subset \mathbf{T}$ of height n,

$$\mu_{\mathbf{T}}(\mathcal{C}_\infty[n] = t) = \frac{\sum_{v\in t_n} \mathbf{W}_v}{\mathbf{W}} \mathbb{P}_{\mathbf{T}}(\mathcal{C}[n] = t), \tag{1.10}$$

where $\mathcal{C}_\infty[n]$ denotes the IIC restricted to the first n generations, t_n denotes vertices in the n^{th} generation of t, $\mathbf{W}_v$ denotes the value of $\mathbf{W}$ for the subtree rooted at v and $\mathcal{C}[n]$ denotes the first n generations of an unconditioned critical cluster.

Michelen shows that under the assumption that all offspring moments are finite, the law of the size of the n^{th} generation in $\mathcal{C}_\infty$ is a size-biased version of the law appearing in Theorem 1.3. We establish a similar result under our assumption.

To state it, we first let Y denote the random variable with Laplace transform appearing in (1.3), and let Y^* denote its size-biased version, meaning that (also using (1.4)):

$$\mathbb{P}\left(Y^* \in [a,b]\right) = \frac{\mathbb{E}\left[Y\mathbb{1}\{Y\in[a,b]\}\right]}{\mathbb{E}\left[Y\right]} = C_\alpha \mathbb{E}\left[Y\mathbb{1}\{Y\in[a,b]\}\right]. \tag{1.11}$$

Note that, under Assumption 1.1(a), Y^* has finite mean. Under Assumption 1.1(b), since Y is α-stable with index $\alpha > 1$, Y^* is well-defined.

We let $\mathbf{Z}_n$ denote $n^{-\beta} Y_n$ under the law $\mu_{\mathbf{T}}$.

Theorem 1.5. *For $\boldsymbol{P}$-a.e. $\boldsymbol{T}$, we have that (under $\mu_{\boldsymbol{T}}$)*

$$\boldsymbol{Z}_n \xrightarrow[n\to\infty]{(d)} Y^*. \tag{1.12}$$

We anticipate that an analogue of Theorem 1.4 should also hold under $\mu_{\mathbf{T}}$, with convergence to a CSBP conditioned to survive, but we have not proved this here.

The main observation that allows us to strengthen the results of Michelen is that it in fact suffices to prove all of the almost sure convergence results along an appropriate subsequence $(n_k)_{k\geq 0}$, and then extend to all n using continuity properties of the various probabilities and processes. For Theorem 1.2 this continuity is an immediate consequence of the monotonicity of the connection probabilities, and the subsequential convergence simply follows from a refinement of the arguments of Michelen. The proof of Theorem 1.3 is quite different to that of Michelen, however, who used a martingale originally studied in [16] to apply the method of moments. Instead, we use a byproduct of the proof of Theorem 1.2 which tells us that for each $n \geq 0$, we can choose $m = m(n)$

satisfying $0 \ll m \ll n$ and such that, with high probability on the event $\{Y_n > 0\}$, there will only be a single vertex at level m that connects directly upwards to level n. Therefore, by averaging over the choice of this vertex and the subtree emanating from it, we can establish good enough concentration of the Laplace transform to show that Theorem 1.3 holds along a subsequence. The bulk of the proof is devoted to proving continuity estimates for the process $(Y_n)_{n\geq 0}$ which allow us to extend this to convergence for all n. The strategy used to prove Theorem 1.4 is very similar. The proof of Theorem 1.5 is also quite different to that of Michelen, who used Chebyshev's inequality to directly analyse the expression appearing in (1.10). This is not possible for us since under our assumptions $\mathbf{W}$ does not in general have a finite second moment; however by some careful analysis we are still able to evaluate the limiting expression in (6.1) to obtain almost sure subsequential convergence, and once again lift this to full convergence using similar ideas to the previous theorems. The main idea to evaluate (6.1) is that once we have shown that the limit in (6.1) exists (Lemma 6.2), it suffices to evaluate the limit of $\mu_{\mathbf{T}}(\mathbf{Z}_n \in (a,b) | Y_M > 0)$ along *any* subsequence $M_j \to \infty$ (Lemma 6.4) to obtain an expression for $\mu_{\mathbf{T}}(\mathbf{Z}_n \in (a,b))$.

It may be surprising at first that the limits appearing in Theorems 1.2, 1.3, 1.4 and 1.5 do not depend on $\mathbf{T}$. However, this is easily explained by the role of the vertex at level $m(n)$ mentioned in the previous paragraph (also see Remark 3.4). In particular, since there will only be one subtree from level $m(n)$ that will survive to level n under the conditioning $\{Y_n > 0\}$, the behaviour of the cluster at level n is essentially independent of its behaviour below level $m(n)$, so as n goes to infinity we successively lose the dependence on more and more of the earlier generations in the tree.

This paper also leaves some remaining questions open. In particular we have not considered the *genealogy* of a critical cluster. In the annealed model it is known that the Gromov–Hausdorff–Prokhorov scaling limit of critical cluster (i.e. when viewing the cluster as a metric-measure space) conditioned to be large is either the Brownian continuum random tree (in the case of finite variance offspring distribution) or a stable Lévy tree (in the stable case); see [7, Theorem 3.1] and recall that the cluster has the law of a critical Galton–Watson tree under the annealed law. In light of our other theorems we anticipate that this should also be true in the quenched setting, but this would require somewhat different techniques to establish.

Finally, we note that our proofs depend on the annealed results in many places, so our results do not provide an independent proofs of the annealed results.

Organisation of the paper. In Section 3 we prove Theorem 1.2 and introduce some notation used throughout the rest of the paper. In Section 4, we prove the convergence of the rescaled generation sizes given in Theorem 1.3. This includes a strategy to lift the convergence from a subsequence to the full sequence that will also be re-used in following sections. In Section 5, we prove Theorem 1.4, the convergence of the full branching process. The measure of the IIC is then constructed in Section 6 and Theorem 1.5 is proved. Finally, in the Appendix we calculate the constant C_α and give a large deviation estimate for the sum of independent random variables in the domain of attraction of a stable law.

2 Notation and set up

Before we begin, we fix some notation that we will use throughout the paper. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space such that under $\mathbf{P}$, $\mathbf{T}$ has the law of a supercritical Galton–Watson tree with no leaves and mean number of descendants $\mu > 1$. Let $\mathcal{F}_n$ be the sigma-algebra generated by the Galton–Watson tree up to generation n. For a tree $\mathbf{T}$, we define the probability space $\left(\Omega^{\mathbf{T}}, \mathcal{G}^{\mathbf{T}}, \mathbb{P}_{\mathbf{T}}\right)$ in which $\mathbb{P}_{\mathbf{T}}$ is the law of an independent

Bernoulli percolation on the edges of $\mathbf{T}$ with retention probability μ^{-1}, $\Omega^{\mathbf{T}}$ is the collection of all possible subtrees of $\mathbf{T}$ and $\mathcal{G}^{\mathbf{T}}$ is the canonical sigma-algebra (generated by cylinder sets). We write $\mathcal{G}_n^{\mathbf{T}}$ for the sigma-algebra generated by the percolation process up to level n in the tree. For events $A \in \mathcal{G}^{\mathbf{T}}$, $B \in \mathcal{G}_n^{\mathbf{T}}$ and $p \in (0,1)$, we abbreviate

$$\left\{\mathbb{P}_{\mathbf{T}}\left(A|\mathcal{G}_n^{\mathbf{T}}, B\right) < p\right\} := \left\{\sup_{\omega \in B} \mathbb{P}_{\mathbf{T}}\left(A|\mathcal{G}_n^{\mathbf{T}}\right)(\omega) < p\right\}. \tag{2.1}$$

For two subsets A, B of the tree $\mathbf{T}$, $A \leftrightarrow B$ means A is connected to B by open edges. $A \overset{*}{\longleftrightarrow} B$ means A is connected to B by a path along which the distance to the root is monotone. We also let $\mathbf{0}$ denote the root of $\mathbf{T}$. For $u, v \in \mathbf{T}$, we let $|u|$ denote the distance of u from $\mathbf{0}$, we let $u \wedge v$ denote the most recent common ancestor of u and v, and we write $u \preceq v$ if u is an ancestor of v.

We choose the following subsequence $n_k \in \mathbb{N}$ such that

$$n_k \sim k^{\frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1}} = k^A \qquad \text{where} \qquad A = \frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1}. \tag{2.2}$$

Note that $A\beta > 2$ (this fact will be used several times). We remark that for some proofs, we could have chosen subsequences which grow slower (though still polynomially). However, we work with the subsquence n_k, in order to unify the different proofs.

Furthermore, unless stated otherwise, constants denoted by c, C only depend on $\mathbf{P}$ and can change from line to line.

3 Convergence of the survival probabilities: proof of Theorem 1.2

In this section we prove Theorem 1.2.

3.1 Outline of strategy and notation

The proof of Theorem 1.2 is divided into a series of smaller lemmas.

Lemma 3.1 (Convergence along a subsequence is sufficient)**.** *Recall that $n_k \sim k^A$, with A defined in Eq.* (2.2)*. Then convergence of* (1.7) *along the subsequence $(n_k)_{k\ge 1}$ implies convergence for all n.*

Proof. This follows from the fact that $n_{k+1}/n_k = 1 + o(1)$. Therefore if n and k are such that $n \in [n_k, n_{k+1}]$, and the subsequential convergence holds, we have that

$$n^\beta \mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right) \le \left(\frac{n_{k+1}}{n_k}\right)^\beta n_k^\beta \mathbb{P}_{\mathbf{T}}\left(Y_{n_k} > 0\right) = \left(1 + o(1)\right) C_\alpha \mathbf{W}, \tag{3.1}$$

and similarly for the lower bound. □

In light of Lemma 3.1, we will prove almost sure convergence only along the subsequence $(n_k)_{k\ge1}$. The proof will depend on the following inequality, which is a consequence of the inclusion-exclusion principle. For every $k \ge 1$ and $1 \le m \le n_k$, we have that

$$\left| n_k^\beta \mathbb{P}_{\mathbf{T}}\left(Y_{n_k} > 0\right) - n_k^\beta \sum_{v \in \mathbf{T}_m} \mathbb{P}_{\mathbf{T}}\left(\mathbf{0} \leftrightarrow v \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k}\right)\right| \le n_k^\beta \sum_{\substack{u,v \in \mathbf{T}_m \\ u \ne v}} \mathbb{P}_{\mathbf{T}}\left(\mathbf{0} \leftrightarrow (u,v) \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k}\right), \tag{3.2}$$

where by $\mathbf{0} \leftrightarrow (u,v) \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k}$ we mean that both u and v are connected to the root as well as directly to $\mathbf{T}_{n_k}$ (by paths along which the distance to the root is monotone).

Note that the sum on the left hand side counts the events in which there is at least a single vertex (connected to the root) at level m that connects directly to level n_k, and the sum on the right-hand side counts the events in which there are at least two vertices (both connected to the root) at level m that connect directly to level n_k.

In what follows we will make an appropriate choice of m so that the right-hand side of (3.2) will be small, and the sum on the left hand side will be well-approximated by its mean, which is close to the desired limit.

In particular, for each $k \in \mathbb{N}$ we set for $\varepsilon = \frac{\alpha-1}{2}$ (fixed for the entire section)

$$m_k = \left\lfloor \frac{(1+\varepsilon)}{(\alpha-1)\log\mu} \log n_k \right\rfloor . \tag{3.3}$$

The proof of Theorem 1.2 then rests on the following two lemmas. In the first lemma we prove that, with this choice of m_k, the right-hand side of (3.2) goes to 0 almost surely. In the second lemma we prove that the sum appearing on the left hand side is well-approximated by its mean.

Lemma 3.2 (Connection through a single node at lower level)**.** ***P****-almost surely,*

$$\lim_{k\to\infty} n_k^\beta \sum_{\substack{u,v\in \boldsymbol{T}_{m_k} \\ u\neq v}} \mathbb{P}_{\mathbf{T}}\left(\mathbf{0} \leftrightarrow (u,v) \overset{*}{\longleftrightarrow} \boldsymbol{T}_{n_k}\right) = 0\,. \tag{3.4}$$

Lemma 3.3 (A concentration estimate)**.** ***P****-almost surely,*

$$\lim_{k\to\infty} n_k^\beta \sum_{v\in \boldsymbol{T}_{m_k}} \left[\mathbb{P}_{\mathbf{T}}\left(\mathbf{0} \leftrightarrow v \overset{*}{\longleftrightarrow} \boldsymbol{T}_{n_k}\right) - \mu^{-m_k}\mathbb{P}\left(\mathbf{0} \leftrightarrow \boldsymbol{T}_{n_k-m_k}\right)\right] = 0. \tag{3.5}$$

Theorem 1.2 for a subsequence follows immediately from the above three lemmas, (1.1) and (1.2), since they imply that

$$\begin{aligned}\lim_{k\to\infty} n_k^\beta \mathbb{P}_{\mathbf{T}}\left(Y_{n_k}>0\right) &= \lim_{k\to\infty} n_k^\beta \sum_{v\in\mathbf{T}_{m_k}} \mu^{-m_k}\mathbb{P}\left(\mathbf{0}\leftrightarrow \mathbf{T}_{n_k-m_k}\right) \\ &= \lim_{k\to\infty} \frac{|\mathbf{T}_{m_k}|}{\mu^{m_k}} n_k^\beta \mathbb{P}\left(\mathbf{0}\leftrightarrow \mathbf{T}_{n_k-m_k}\right) = C_\alpha \mathbf{W}. \end{aligned} \tag{3.6}$$

Remark 3.4. Define ℓ_n so that $\ell_{n_k} = m_k$ for all $k \geq 1$ and extend to all n by setting $\ell_n = \ell_{n_k}$ when $n \in [n_k, n_{k+1})$. Let A_n be the event that there exist at least two vertices at level m_n that connect the root to level n, and let B_n be the event that there exist at least two vertices at level ℓ_n that connect the root to level n. By Theorem 1.2 and Lemma 3.2, $\mathbb{P}_{\mathbf{T}}\left(A_{n_k} \text{ i.o.} | Y_{n_k}>0\right) \to 0$, $\mathbf{P}$-almost surely. This also implies that $\mathbb{P}_{\mathbf{T}}\left(B_n \text{ i.o.} | Y_n > 0\right) \to 0$, $\mathbf{P}$-almost surely. This follows since if $n \in [n_k, n_{k+1})$ and there are distinct u, v in generation ℓ_n connecting the root to level n using a path of length n, then since $\ell_n = m_k \leq n_k \leq n$ we obtain two distinct vertices in level m_k that connect to level n_k by following these two paths up to level m_k and onwards to level n_k. Therefore, if B_n occurred infinitely often, then so would A_{n_k}. Moreover, $\frac{\mathbb{P}_{\mathbf{T}}\left(Y_{n_k}>0\right)}{\mathbb{P}_{\mathbf{T}}(Y_n>0)} \to 1$, so we can exchange the two conditioning events.

3.2 Proof of lemmas

In this subsection we prove Lemma 3.2 and Lemma 3.3. We first fix constants conveniently chosen such that the Borel–Cantelli arguments applied later work

Definition 3.5. *Fix* $\varepsilon = \frac{\alpha-1}{2}$ *and* $p = \frac{\alpha-\varepsilon}{2} = \frac{\alpha+1}{4} \in (0,\alpha/2)$. *Recall* $n_k \sim k^A$ *from Equation* (2.2) *and* m_k *from Equation* (3.3)*. For this choice, we have that* $A > 2p^{-1}(\alpha-1)\varepsilon^{-1}$*, which we use repeatedly later.*

Proof of Lemma 3.2. (Connection through a single node at lower level.)

We want to show that, $\mathbf{P}$-almost surely,

$$\lim_{k\to\infty} n_k^\beta \sum_{\substack{u,v\in\mathbf{T}_{m_k}\\ u\neq v}} \mathbb{P}_{\mathbf{T}}\left(\mathbf{0}\leftrightarrow(u,v)\overset{*}{\longleftrightarrow}\mathbf{T}_{n_k}\right)=0\,. \tag{3.7}$$

We abbreviate $m=m_k$ and $n=n_k$ for the rest of the proof.

Recall that $\mathcal{F}_i=\sigma(\mathbf{T}_r\colon 0\le r\le i)$ is the sigma algebra generated by the first i levels of the tree. For $u,v\in\mathbf{T}_m$, we set $\boldsymbol{p}_{u,v}=\mathbb{P}_{\mathbf{T}}\left(u\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)\mathbb{P}_{\mathbf{T}}\left(v\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)$. By Markov's inequality, we have (taking p as defined in Definition 3.5) that

$$\begin{aligned}
&\mathbf{P}\left(\sum_{\substack{u,v\in\mathbf{T}_m\\ u\neq v}} \mathbb{P}_{\mathbf{T}}\left(\mathbf{0}\leftrightarrow(u,v)\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)>x\right)\le \mathbf{E}\left[\left(\sum_{\substack{u,v\in\mathbf{T}_m\\ u\neq v}} \mathbb{P}_{\mathbf{T}}\left(\mathbf{0}\leftrightarrow(u,v)\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)\right)^p\right]x^{-p}\\
&=\mathbf{E}\left[\mathbf{E}\left[\left(\sum_{\substack{u,v\in\mathbf{T}_m\\ u\neq v}} \mathbb{P}_{\mathbf{T}}\left(\mathbf{0}\leftrightarrow(u,v)\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)\right)^p\middle|\mathcal{F}_m\right]\right]x^{-p}\\
&=\mathbf{E}\left[\mathbf{E}\left[\left(\sum_{\substack{u,v\in\mathbf{T}_m\\ u\neq v}} \frac{\boldsymbol{p}_{u,v}}{\mu^{2m-|u\wedge v|}}\right)^p\middle|\mathcal{F}_m\right]\right]x^{-p}\,,
\end{aligned} \tag{3.8}$$

where $u\wedge v$ is the most recent common ancestor (i.e. furthest from the root) of u and v, and $|u\wedge v|$ is the distance of $u\wedge v$ from the root of $\mathbf{T}$.

Given $u\in\mathbf{T}_m$, let $\mathbf{0}=u_0,\dots,u_m=u$ denote the ancestors of u ordered by distance from the root. Let $\preceq$ denote the partial ordering induced by generations and let

$$\mathbf{T}_m^{(u_i)}=\{v\in\mathbf{T}_m\colon u_i\preceq v\}\ , \tag{3.9}$$

or in other words the set of vertices in $\mathbf{T}_m$ that descend from u_i. Reordering the sum and using Jensen's inequality (recall that $p<1$), we deduce that

$$\begin{aligned}
\mathbf{E}\left[\left(\sum_{\substack{u,v\in\mathbf{T}_m\\ u\neq v}} \frac{\boldsymbol{p}_{u,v}}{\mu^{2m-|u\wedge v|}}\right)^p\middle|\mathcal{F}_m\right]&\le \mathbf{E}\left[\left(\sum_{i=1}^m\sum_{u\in\mathbf{T}_i}\sum_{\substack{a,b\in\mathbf{T}_m^{(u)}\\ a\neq b}} \frac{\boldsymbol{p}_{a,b}}{\mu^{2m-i}}\right)^p\middle|\mathcal{F}_m\right]\\
&\le \sum_{i=1}^m\sum_{u\in\mathbf{T}_i}\mathbf{E}\left[\left(\sum_{\substack{a,b\in\mathbf{T}_m^{(u)}\\ a\neq b}} \frac{\boldsymbol{p}_{a,b}}{\mu^{2m-i}}\right)\middle|\mathcal{F}_m\right]^p.
\end{aligned} \tag{3.10}$$

(Note that $p<1$ for the final line.) Since each term of the form $\boldsymbol{p}_{a,b}$ is independent of $\mathcal{F}_m$ and moreover its conditional law does not depend on the choice of a and b, this term can be factorised outside the internal expectation.

Write $\overline{\mathbf{W}}=\sup_n\mathbf{W}_n$. Applying the tower property once more we deduce that there exists $C<\infty$ such that the expectation of this expression is upper bounded by (now letting a, b denote arbitrary distinct vertices in $\mathbf{T}_m$)

$$\begin{aligned}\mathbf{E}\left[\sum_{i=1}^{m}\sum_{u\in\mathbf{T}_i}\left(\frac{\left(\mathbf{T}_m^{(u)}\right)^2}{\mu^{2m-i}}\right)^p\right]\mathbf{E}\left[\boldsymbol{p}_{a,b}\right]^p &= \mathbf{E}\left[\sum_{i=1}^{m}\mathbf{E}\left[\sum_{u\in\mathbf{T}_i}\mu^{-ip}\left(\frac{\left(\mathbf{T}_m^{(u)}\right)^2}{\mu^{2m-2i}}\right)^p\middle|\mathcal{F}_i\right]\right]\mathbf{E}\left[\boldsymbol{p}_{a,b}\right]^p \\ &\le \mathbf{E}\left[\sum_{i=1}^{m}|\mathbf{T}_i|\,\mu^{-ip}\right]\mathbf{E}\left[\overline{\mathbf{W}}^{2p}\right]\mathbf{E}\left[\boldsymbol{p}_{a,b}\right]^p \\ &\le C\mu^{m(1-p)}\mathbf{E}\left[\overline{\mathbf{W}}^{2p}\right]\mathbf{E}\left[\boldsymbol{p}_{a,b}\right]^p .\end{aligned} \tag{3.11}$$

Since $2p < \alpha$ (recall Definition 3.5), we have that $\overline{\mathbf{W}}^{2p}$ is bounded in $\mathrm{L}^1(\mathbf{P})$ and hence, applying (1.2) to bound $\mathbf{E}\left[\boldsymbol{p}_{a,b}\right] = \mathbb{P}\left(Y_{n-m} > 0\right)^2$ for $a \neq b$ and substituting back into (3.8), we obtain that there exists $C < \infty$ such that

$$\mathbf{P}\left(\sum_{\substack{u,v\in\mathbf{T}_m \\ u\neq v}}\mathbb{P}_{\mathbf{T}}\left(\mathbf{0}\leftrightarrow(u,v)\overset{*}{\longleftrightarrow}\mathbf{T}_n\right) > x\right) \le C\mu^{m(1-p)}\left(n-m\right)^{-2p/(\alpha-1)}x^{-p}. \tag{3.12}$$

Note that for p as in Definition 3.5, we have that $\frac{p}{1-p} = \frac{1+\varepsilon}{1-\varepsilon}$. Recall also that $m_k = \left\lfloor\frac{(1+\varepsilon)}{(\alpha-1)\log\mu}\log n_k\right\rfloor$. To prove the proposition, we choose $x = n^{-\beta}(\log n)^{-1}$. Noting that $\mu^{m(1-p)} = n^{\frac{p(1-\varepsilon)}{\alpha-1}}$, it follows that the right-hand side of (3.12) is upper bounded by $n^{-\frac{\varepsilon p}{\alpha-1}}(\log n)^p$. By our choice of A we have $A > 2p^{-1}(\alpha-1)\varepsilon^{-1}$. Hence, this bound is summable along the sequence $(n_k)_{k\ge1}$ and hence the claim follows by Borel–Cantelli. □

We now turn to the proof of Lemma 3.3, which consists of a simple second moment estimate.

Proof of Lemma 3.3. (A concentration estimate.) Again write $n = n_k$ and $m = m_k$, and recall the constant from Definition 3.5. Note that given $\mathcal{F}_m$, we have that the collection of random variables

$$\left(\mathbb{P}_{\mathbf{T}}\left(v\overset{*}{\longleftrightarrow}\mathbf{T}_n\right) - \mathbb{P}\left(\mathbf{0}\leftrightarrow\mathbf{T}_{n-m}\right)\right)_{v\in\mathbf{T}_m}, \tag{3.13}$$

are i.i.d. with mean zero under $\mathbf{P}$. It follows that

$$\begin{aligned}&\mathbf{Var}\left(\sum_{v\in\mathbf{T}_m}\left[\mathbb{P}_{\mathbf{T}}\left(0\leftrightarrow v\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)-\mu^{-m}\mathbb{P}\left(\mathbf{0}\leftrightarrow\mathbf{T}_{n-m}\right)\right]\middle|\mathcal{F}_m\right)\\ &\quad= \mu^{-2m}\sum_{v\in\mathbf{T}_m}\mathbf{Var}\left(\mathbb{P}_{\mathbf{T}}\left(v\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)\right) \le \mu^{-2m}\sum_{v\in\mathbf{T}_m}\mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(v\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)\right].\end{aligned} \tag{3.14}$$

Here the final inequality follows since $\mathbb{P}_{\mathbf{T}}\left(v\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)^2 \le \mathbb{P}_{\mathbf{T}}\left(v\overset{*}{\longleftrightarrow}\mathbf{T}_n\right)$. Note that, by (1.2), the final line in (3.14) is bounded by

$$O\left(|\mathbf{T}_m|\,(n-m)^{-\beta}\mu^{-2m}\right). \tag{3.15}$$

It therefore follows from Chebyshev's inequality that there exists $C < \infty$ such that for any $x > 0$,

$$\mathbf{P}\left(\left|\sum_{v\in\mathbf{T}_m}\left[\mathbb{P}_{\mathbf{T}}\left(0\leftrightarrow v\leftrightarrow\mathbf{T}_n\right)-\mu^{-m}\mathbb{P}\left(\mathbf{0}\leftrightarrow\mathbf{T}_{n-m}\right)\right]\right| > x\middle|\mathcal{F}_m\right) \le C\frac{|\mathbf{T}_m|\,n^{-\beta}}{x^2\mu^{2m}}. \tag{3.16}$$

Setting $x = n^{-\beta}(\log n)^{-1}$, it follows from the tower property and the fact that $\mathbf{E}[\overline{\mathbf{W}}] < \infty$ that there exists $C' < \infty$ such that

$$\begin{aligned}\mathbf{P}&\left(n^{\beta}\left|\sum_{v\in\mathbf{T}_m}\left[\mathbb{P}_{\mathbf{T}}\left(0\leftrightarrow v\leftrightarrow\mathbf{T}_n\right)-\mu^{-m}\mathbb{P}\left(\mathbf{0}\leftrightarrow\mathbf{T}_{n-m}\right)\right]\right|>(\log n)^{-1}\right)\\&\leq\mathbf{E}\left[C\frac{|\mathbf{T}_m|\,n^{-\beta}}{x^2\mu^{2m}}\right]\leq\frac{C'n^{-\beta}}{x^2\mu^m}=C'n^{-\frac{\varepsilon}{\alpha-1}}(\log n)^2\,,\end{aligned}\tag{3.17}$$

which is summable along the subsequence $(n_k)_{k\geq1}$ by our choice of A (recall Definition 3.5). Hence

$$\mathbf{P}\left(n_k^{\beta}\left|\sum_{v\in\mathbf{T}_{m_k}}\left[\mathbb{P}_{\mathbf{T}}\left(0\leftrightarrow v\leftrightarrow\mathbf{T}_{n_k}\right)-\mu^{-m_k}\mathbb{P}\left(\mathbf{0}\leftrightarrow\mathbf{T}_{n_k-m_k}\right)\right]\right|>(\log n)^{-1}\text{ infinitely often}\right)=0,\tag{3.18}$$

thus proving the lemma. □

4 Yaglom's limit: proof of Theorem 1.3

In this section we examine the limit of the distribution of $n^{-\beta}Y_n$ conditioned on the event $\{Y_n > 0\}$. We do this through a series of intermediate results, all concerning the Laplace transform. An advantage of our method is that we do not need to make any moment assumption on our offspring distribution. This is in contrast to the proof of Michelen who proved a similar result using the method of moments, and therefore needed to assume that all moments were finite. We first state a preparatory lemma.

For $\theta \geq 0, n \geq 1$ let $\phi_n(\theta)$ be the critical (annealed) Laplace transform

$$\phi_n(\theta) = \mathbb{E}\left[\mathrm{e}^{-\theta n^{-\beta}Y_n}|Y_n > 0\right]\,.\tag{4.1}$$

Recall from (1.3) that $\phi_n(\theta)$ converges to $\phi(\theta)$ as $n \to \infty$ for all $\theta \geq 0$, where

$$\phi(\theta) = 1 - C_\alpha^{-1}\theta(1 + \left(C_\alpha^{-1}\theta\right)^{\alpha-1})^{-\beta}.$$

As in the proof of Theorem 1.2, our strategy to prove Theorem 1.3 is to start by proving almost sure convergence along a subsequence $(n_k)_{k\geq1}$, and then extend this to full convergence by showing that Y_n is very likely to be close to Y_{n_k} if $n \in [n_k, n_{k+1}]$. We start with the subsequential convergence.

Proposition 4.1. *Recall the definition of the subsequences $(n_k)_{k\geq1}$ and $(m_k)_{k\geq1}$ from Lemma 3.1 and* (3.3)*. For every $\theta \geq 0$, we have that, **P**-almost surely*

$$\phi_{\mathbf{T}}^{(n_k)}(\theta) := \mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n_k^{-\beta}Y_{n_k}}|Y_{n_k} > 0\right] \to \phi(\theta)\tag{4.2}$$

as $k \to \infty$.

Proof. By the monotonicity of the Laplace transform in $\theta \geq 0$, we get from Theorem 1.3 and (3.12) that

$$\mathbf{P}\left(\exists\theta\geq0\colon\sum_{u\neq v\in\mathbf{T}_{m_k}}\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n_k^{-\beta}Y_{n_k}}\mathbb{1}\{0\leftrightarrow(u,v)\overset{*}{\longleftrightarrow}\mathbf{T}_{n_k}\}|Y_{n_k}>0\right]>\frac{1}{\log(n_k)}\text{ i.o.}\right)=0\,.\tag{4.3}$$

Indeed, $\theta = 0$ is the extremal case here, in which case we are reduced to (3.12). By the inclusion-exclusion principle, this means that **P**-almost surely, we can write (henceforth writing m and n in place of m_k and n_k respectively)

$$\phi_{\mathbf{T}}^{(n)}(\theta) = \sum_{v\in\mathbf{T}_m}\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n^{-\beta}Y_n^{(v)}}\mathbb{1}\{0\leftrightarrow v\overset{*}{\longleftrightarrow}\mathbf{T}_n\}\middle|Y_n>0\right]+o(1)\,,\tag{4.4}$$

where $Y_n^{(v)}$ denotes elements in the root cluster at level n that are also descendants of v.

We now examine the unconditional Laplace transform:

$$\sum_{v\in\mathbf{T}_m} \mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n^{-\beta} Y_n^{(v)}} \mathbb{1}\{0 \leftrightarrow v \overset{*}{\longleftrightarrow} \mathbf{T}_n\}\right] . \tag{4.5}$$

Note that given $\mathcal{F}_m$, we have for any $v \in \mathbf{T}_m$ that the collection of random variables

$$\left(\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n^{-\beta} Y_n^{(v)}} \mathbb{1}\{v \overset{*}{\longleftrightarrow} \mathbf{T}_n\}\right] - \mathbb{E}\left[\mathrm{e}^{-\theta n^{-\beta} Y_{n-m}} \mathbb{1}\{Y_{n-m} > 0\}\right]\right)_{v\in\mathbf{T}_m} \tag{4.6}$$

are i.i.d. with mean zero under $\mathbf{P}$. Hence, as in (3.14) and (3.15), we have that

$$\mathbf{Var}\left(\sum_{v\in\mathbf{T}_m}\left[\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n^{-\beta} Y_n^{(v)}} \mathbb{1}\{0 \leftrightarrow v \overset{*}{\longleftrightarrow} \mathbf{T}_n\}\right] - \mu^{-m}\mathbb{E}\left[\mathrm{e}^{-\theta n^{-\beta} Y_{n-m}} \mathbb{1}\{Y_{n-m} > 0\}\right]\right] \middle| \mathcal{F}_m\right) = O\left(|\mathbf{T}_m|\, n^{-\beta} \mu^{-2m}\right) . \tag{4.7}$$

Thus we have by Chebyshev's inequality and the first Borel–Cantelli lemma (exactly as in (3.16) and (3.17)) that eventually $\mathbf{P}$-almost surely,

$$n_k^{\beta}\left|\sum_{v\in\mathbf{T}_{m_k}}\left[\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n_k^{-\beta} Y_{n_k}^{(v)}} \mathbb{1}\{0 \leftrightarrow v \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k}\}\right] - \mu^{-m}\mathbb{E}\left[\mathrm{e}^{-\theta n_k^{-\beta} Y_{n_k-m_k}} \mathbb{1}\{Y_{n_k-m_k} > 0\}\right]\right]\right| < (\log n_k)^{-1} . \tag{4.8}$$

To conclude, note that it follows by definition, continuity of ϕ and (1.2) that

$$\begin{aligned} n_k^{\beta} \sum_{v\in\mathbf{T}_m} \mu^{-m}\mathbb{E}\left[\mathrm{e}^{-\theta n_k^{-\beta} Y_{n_k-m_k}} \mathbb{1}\{Y_{n_k-m_k} > 0\}\right] &= n_k^{\beta}\mathbf{W}_m\mathbb{E}\left[\mathrm{e}^{-\theta n_k^{-\beta} Y_{n_k-m_k}} \mathbb{1}\{Y_{n_k-m_k} > 0\}\right] \\ &= \mathbf{W}_m C_\alpha \phi(\theta)\,(1+o(1)) \ , \end{aligned} \tag{4.9}$$

and hence combining with (4.4), (4.8) and Theorem 1.2 gives

$$\phi_{\mathbf{T}}^{(n_k)}(\theta) = \frac{\mathbf{W}_{m_k} C_\alpha}{n_k^{\beta}\mathbb{P}_{\mathbf{T}}(Y_{n_k} > 0)}\phi(\theta)\,(1+o(1)) \to \phi(\theta). \qquad \square$$

Next, we verify that Proposition 4.1 implies the almost sure convergence in law of $n_k^{-\beta} Y_{n_k}$ conditioned on survival.

Corollary 4.2. *For $\mathbf{P}$-almost every tree $\mathbf{T}$, we have that $n_k^{-\beta} Y_{n_k}$ conditioned to survive up to time n_k converges in distribution as $k \to \infty$ to a law with Laplace transform ϕ defined in* (1.3).

Proof. By Proposition 4.1, we have that

$$\mathbf{P}\left(\forall\theta \in [0,\infty)\cap\mathbb{Q}\colon\ \lim_{k\to\infty}\phi_{\mathbf{T}}^{(n_k)}(\theta) = \phi(\theta)\right) = 1\,. \tag{4.10}$$

It just remains to remove the restriction $\theta \in \mathbb{Q}$. To this end, note that for every $\mathbf{T}$, the function $\theta \mapsto \phi_{\mathbf{T}}^{(n_k)}(\theta)$ is monotone and that furthermore the limiting function ϕ is continuous. Hence, this immediately implies that

$$\mathbf{P}\left(\forall\theta \in [0,\infty)\colon\ \lim_{k\to\infty}\phi_{\mathbf{T}}^{(n_k)}(\theta) = \phi(\theta)\right) = 1\,, \tag{4.11}$$

and the result follows. $\square$

It remains now to prove that the subsequential convergence of Corollary 4.2 can be lifted to the entire sequence $\left(n^{-\beta}Y_n\right)_{n\geq 1}$. Our strategy is to work via a proof by contradiction: if $\left|\phi_{\mathbf{T}}^{(n)}(\theta)-\phi(\theta)\right|$ was large but $\left|\phi_{\mathbf{T}}^{(n)}(\theta)-\phi_{\mathbf{T}}^{(n_k)}(\theta)\right|$ was small, then $\left|\phi_{\mathbf{T}}^{(n_k)}(\theta)-\phi(\theta)\right|$ would also be large, which would contradict Corollary 4.2. Hence if we can show that $\left|\phi_{\mathbf{T}}^{(n)}(\theta)-\phi_{\mathbf{T}}^{(n_k)}(\theta)\right|$ is consistently small, the result will follow.

We start in Lemma 4.3 by defining an event that indeed implies that $\left|\phi_{\mathbf{T}}^{(n)}(\theta)-\phi_{\mathbf{T}}^{(n_k)}(\theta)\right|$ is small. Afterwards, we will explain how lower bounding the probability of this event directly implies Theorem 1.3 and in particular in Lemma 4.4 we will give a formal criterion of this ilk that does indeed imply the theorem. Then we just need to verify that the criterion is satisfied: this is the subject of Proposition 4.5.

Lemma 4.3 (Upper bounding $|\phi_{\mathbf{T}}^{(n)}(\theta)-\phi_{\mathbf{T}}^{(n_k)}(\theta)|$)**.** *Fix some $\delta\in(0,\frac{\beta}{2})$. $\boldsymbol{P}$-almost surely, we have for all $\theta>0$ and $\varepsilon>0$ that there exists $K=K_{\delta,\varepsilon,\theta,\boldsymbol{T}}<\infty$ such that for all $k\geq K$ and all $n_{k-1}<n\leq n_k$,*

$$\left\{\left|\phi_{\mathbf{T}}^{(n)}(\theta)-\phi_{\mathbf{T}}^{(n_k)}(\theta)\right|<\varepsilon\right\}\supset\left\{\mathbb{P}_{\mathbf{T}}\left(|Y_n-Y_{n_k}|>Y_n^{1-\delta}\vee n^{\beta-\delta}\middle|Y_n>0\right)<\varepsilon/4\right\}. \tag{4.12}$$

Proof. Set $\delta'=\frac{1}{2A}$ (recall from Lemma 3.1 that $n_k\sim k^A$). For each $n\geq 1$ take k such that $n\in[n_{k-1},n_k]$ and write

$$D_n=\left\{|Y_n-Y_{n_k}|>Y_n^{1-\delta}\vee n^{\beta-\delta}\right\}\cup\left\{Y_{n_k}<n^{\beta+\delta'}\right\}. \tag{4.13}$$

Now set $\gamma_n=\frac{\mathbb{P}_{\mathbf{T}}(Y_n>0)}{\mathbb{P}_{\mathbf{T}}\left(Y_{n_k}>0\right)}$ and note that, $\mathbf{P}$-almost surely,

$$1-\frac{\varepsilon}{8}<\gamma_n^{-1}\leq 1\leq\gamma_n<1+\frac{\varepsilon}{8} \tag{4.14}$$

for all sufficiently large n by Theorem 1.2. Now note that if the event on the right-hand side of (4.12) occurs, then provided n is sufficiently large (depending on ε), it follows by Corollary 4.2 and (4.14) that $\mathbb{P}_{\mathbf{T}}\left(D_n|Y_n>0\right)+\mathbb{P}_{\mathbf{T}}\left(D_n|Y_{n_k}>0\right)<\varepsilon/2$ (on a $\mathbf{P}$-almost sure event). Since the Laplace transform is bounded by 1, it is therefore sufficient to show that

$$\left|\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n_k^{-\beta}Y_{n_k}}\mathbb{1}\left\{D_n^c\right\}|Y_{n_k}>0\right]-\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n^{-\beta}Y_n}\mathbb{1}\left\{D_n^c\right\}|Y_n>0\right]\right|\leq\varepsilon/2 \tag{4.15}$$

in order to prove the lemma. To this end, note that on the event D_n^c we can write

$$\left|{n_k}^{-\beta}Y_{n_k}-n^{-\beta}Y_n\right|\leq\left|{n_k}^{-\beta}-n^{-\beta}\right|Y_{n_k}+n^{-\beta}\left|Y_{n_k}-Y_n\right|\leq n^{-\beta}Y_n^{1-\delta}+O\left(n^{-(\frac{1}{2A}\wedge\delta)}\right). \tag{4.16}$$

Here the last line follows by our choice of δ' since $n_k^{-\beta}-n^{-\beta}=O\left(n^{-\beta-\frac{1}{A}}\right)$. Note also that $\mathbb{P}_{\mathbf{T}}\left(Y_{n_k}=0|Y_n>0\right)=1-\gamma_n^{-1}$. Combining the previous observations, we see that there exists $C<\infty$ such that we can bound the left-hand side of Equation (4.15) from above by

$$\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n^{-\beta}Y_n}\left|1-\gamma_n\mathrm{e}^{C\theta(n^{-\beta}Y_n^{1-\delta}+n^{-(\frac{1}{2A}\wedge\delta)})}\right|\mathbb{1}\left\{D_n^c\right\}\middle|Y_n>0\right]+\frac{\varepsilon}{8}. \tag{4.17}$$

In the above expectation, we have that $\mathrm{e}^{-\theta n^{-\beta}(Y_n-CY_n^{1-\delta})}=o(1)$ when $Y_n>n^{\frac{\beta}{1-\delta/2}}$ and also that $n^{-\beta}Y_n^{1-\delta}+n^{-(\frac{1}{2A}\wedge\delta)}=o(1)$ when $Y_n\leq n^{\frac{\beta}{1-\delta/2}}$. We combine this with (4.14) to bound (4.17) by

$$\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta n^{-\beta}Y_n}\left|1-\gamma_n+\gamma_n(1-\mathrm{e}^{Cn^{-\beta}Y_n^{1-\delta}+Cn^{-(\frac{1}{2A}\wedge\delta)}})\right|\mathbb{1}\left\{D_n^c\right\}|Y_n>0\right]+\frac{\varepsilon}{8}<\frac{\varepsilon}{4}+o(1), \tag{4.18}$$

thus completing the proof of (4.15). □

To ease notation, we fix $\delta \in (0,1)$ which will be made small in Proposition 4.5, in a way that depends only on α and other constants, and set, for $n_{k-1} < n \leq n_k$,

$$E_n = \left\{ |Y_n - Y_{n_k}| > Y_n^{1-\delta} \vee n^{\beta-\delta} \right\}. \tag{4.19}$$

We now set $\varepsilon > 0$ and abbreviate the event F_n which depends on the tree up to generation n (recall also the notation of (2.1)):

$$F_n = F_n(\mathbf{T}, \varepsilon) = \left\{ \mathbb{P}_{\mathbf{T}} \left(E_n | \mathcal{G}_n^{\mathbf{T}}, Y_n > 0 \right) < \varepsilon/4 \right\}. \tag{4.20}$$

The aforementioned criterion will be as follows: there exists $c > 0$ such that for all n large enough

$$\mathbf{P}\left(F_n | \mathcal{F}_n\right) \geq c > 0 \qquad \text{everywhere}. \tag{4.21}$$

We now prove that (4.21) is a sufficient condition for the final result.

Lemma 4.4 (Criterion for full convergence)**.** *If there exists a choice of $\delta > 0$ such that for every $\varepsilon > 0$ there exists $c > 0$ such that Equation (4.21) holds for all sufficiently large n, we have that $n^{-\beta} Y_n$, conditioned on survival up to time n, converges in distribution to the random variable with distribution function ϕ.*

Proof. Assume that $\delta \in (0, \frac{\beta}{2})$ is fixed. Now fix some $\varepsilon > 0$ and for all $n, k \geq 0$ write

$$\begin{aligned} A_n &= \left\{ \left| \phi_{\mathbf{T}}^{(n)}(\theta) - \phi(\theta) \right| > 2\varepsilon \right\}, \\ B_k &= \left\{ \left| \phi_{\mathbf{T}}^{(n_k)}(\theta) - \phi(\theta) \right| > \varepsilon \right\}, \end{aligned} \tag{4.22}$$

Note that A_n is $\mathcal{F}_n$ measurable and B_k is $\mathcal{F}_{n_k}$ measurable.

By Lemma 4.3, we have that there $\mathbf{P}$-almost surely exists $K_o < \infty$ such that for all $k \geq K_o$ and all $n \in [n_{k-1}, n_k]$, A_n implies B_k on the event F_n. In other words we have for all $n_{k-1} \leq n \leq n_k$ that

$$A_n \cap F_n \subset B_k. \tag{4.23}$$

It follows that for any $K > K_o$,

$$\mathbf{P}\left(\exists k \geq K \colon B_k \text{ occurs}\right) \geq \mathbf{P}\left(\exists n \geq n_K \colon A_n \cap F_n \text{ occurs}\right). \tag{4.24}$$

Set $\tau_K = \inf\{n \geq n_K : A_n \text{ occurs}\}$. Since τ_K is a stopping time, on the event $\{\tau_K < \infty\}$ we have that $\mathbf{P}\left(F_{\tau_K} | \tau_K < \infty, \mathcal{F}_{\tau_K}\right) \geq c$ by (4.21). Hence we can lower bound the right-hand side of (4.24) by

$$\mathbf{E}\left[\mathbf{P}\left(\tau_K < \infty \text{ and } F_{\tau_K} | \mathcal{F}_{\tau_K}\right)\right] \geq c\mathbf{P}\left(\tau_K < \infty\right). \tag{4.25}$$

Combining (4.24) and (4.25) gives

$$\mathbf{P}\left(\exists k \geq K \colon B_k \text{ occurs}\right) \geq c\mathbf{P}\left(\exists n \geq n_K \colon A_n \text{ occurs}\right). \tag{4.26}$$

On account of Corollary 4.2, we have that B_k occurs only finitely many times $\mathbf{P}$-almost surely, and hence the left hand side above tends to 0 as $K \to \infty$. We deduce that the same holds for the right hand side, and therefore that $\mathbf{P}\left(A_n \text{ i.o.}\right) = 0$. Since this holds for any $\varepsilon > 0$, this proves the claim. □

For the rest of the proof we therefore focus on establishing (4.21). Let $y_n = \{v \in \mathbf{T}_n \colon 0 \leftrightarrow \mathbf{T}_n\}$. To this end we fix $\varepsilon > 0$ and define a collection of "good" sets of vertices at level n in $\mathbf{T}$ by setting

$$A_{\mathbf{T}}^{(n)} = \{x_n \subset \mathbf{T}_n \colon \mathbb{P}_{\mathbf{T}}\left(E_n | y_n = x_n\right) < \varepsilon/4\}. \tag{4.27}$$

Note that $\{y_n = x_n\}$ is measurable with respect to the σ-algebra generated by the set of possible percolation configurations up to level n on $\mathbf{T}$, which we recall is denoted $\mathcal{G}_n^{\mathbf{T}}$.

The next proposition contains the key estimate.

Proposition 4.5 (All subsets are good)**.** *There is a choice of $\delta > 0$ in* (4.19) *such that, for all $x_n \subset \boldsymbol{T}_n$ we have that (uniformly over all choices of x_n and all possibilities for $\boldsymbol{T}_n$)*

$$\boldsymbol{P}\left(x_n \in A^{(n)}_{\boldsymbol{T}} | \mathcal{F}_n\right) = 1 - o(1)\,. \tag{4.28}$$

Proof. First some notation: in this proof, for $a, b, c, d \in \mathbb{R}$ with $c < d$ we write $a = b[c, d]$ if there exists $C \in [c, d]$ such that $a = bC$.

Now to the proof: fix $n \geq 0$ and assume that $n_{k-1} \leq n \leq n_k$. Take $x_n \subset \mathbf{T}_n$, set $N = |x_n|$ and write $\{v_1, \dots, v_N\}$ for the (left to right) enumeration of the vertices contained in x_n. With respect to $\mathbf{P}$, conditionally on $\mathcal{F}_n$, we write $\mathbf{T}^{(i)}_n$ for the subtree emanating from v_i for $i = 1, \dots, N$. Note that $\left(\mathbf{T}^{(i)}_n\right)_{i=1}^N$ is a family of i.i.d. Galton–Watson trees. For each $1 \leq i \leq N$ and $k \geq 0$ also let $Y^{(i)}_k$ denote the number of vertices in the k^{th} generation of $\mathbf{T}^{(i)}_n$ that are connected to v_i by an open path. We can then write, conditionally on the event $\{y_n = x_n\}$,

$$|Y_n - Y_{n_k}| = \left|\sum_{i=1}^N (Y^{(i)}_l - 1)\right|\,, \tag{4.29}$$

where $l = n_k - n = O\left(k^{A-1}\right) = O\left(n^{\frac{A-1}{A}}\right)$.

Recall from the introduction that $\mathbb{P}$ is the law of a critical Galton–Watson tree corresponding to the annealed percolation model. We therefore have (where E_n is as in (4.19)) that

$$\mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(E_n | y_n = x_n\right) | \mathcal{F}_n\right] = \mathbb{P}^{\otimes N}\left(\left|\sum_{i=1}^N (Y^{(i)}_l - 1)\right| > N^{1-\delta} \vee n^{\beta-\delta}\right)\,, \tag{4.30}$$

where we abuse notation and use $Y^{(i)}_l$ as the number of surviving vertices in the annealed critical tree for the rest of this proof.

Let $\kappa = \left|\{i \leq N : Y^{(i)}_l \neq 0\}\right|$, so that by (1.2) κ is binomially distributed with parameters $p_l := \mathbb{P}\left(Y_l > 0\right) \sim C_\alpha l^{-\beta}$ and sample size N under $\mathbb{P}$. We then have that

$$\begin{aligned}&\mathbb{P}^{\otimes N}\left(\left|\sum_{i=1}^N (Y^{(i)}_l - 1)\right| > N^{1-\delta} \vee n^{\beta-\delta}\right)\\ &\quad= \sum_{\kappa=0}^N \binom{N}{\kappa} p_l^\kappa (1-p_l)^{N-\kappa} \mathbb{P}^{\otimes\kappa}\left(\left|-(N-\kappa) + \sum_{i=1}^\kappa (\tilde{Y}^{(i)}_l - 1)\right| > N^{1-\delta} \vee n^{\beta-\delta}\right),\end{aligned} \tag{4.31}$$

where $\tilde{Y}^{(i)}_l$ are now the i.i.d. sizes of the populations conditioned to survive to level l. Note that $\mathbb{E}\left[\tilde{Y}^{(i)}_l\right] = p_l^{-1}$ by Bayes' formula and since the original process was critical. Now pick some $\varepsilon > 0$ and for each $l \geq 1$ and $\kappa \geq 1$ define the event

$$A_l(\kappa) = \left\{\sum_{i=1}^\kappa \tilde{Y}^{(i)}_l = \kappa p_l^{-1} + \kappa^{1/\alpha+\varepsilon} p_l^{-1}[-1, 1]\right\}. \tag{4.32}$$

On the event $A_l(\kappa)$, we get that our sum $\sum_{i=1}^N (Y^{(i)}_l - 1)$ is equal to $-N + \kappa p_l^{-1} + \kappa^{1/\alpha+\varepsilon} p_l^{-1}[-1, 1]$. Now for an additional parameter $M > 0$ also consider the event

$$B_{N,M} = \left\{\kappa = Np_l + \sqrt{Np_l}[-M, M]\right\}\,. \tag{4.33}$$

We henceforth fix some small $\eta > 0$ and take $M = n^{\varepsilon}$. By Lemma A.2 (in the Appendix), we get that $A_l(\kappa)$ occurs with probability at least $1-\eta$, given $\kappa > \kappa_0$ for some κ_0 depending on $\alpha > 1$ and η. We will treat the case $\kappa \le \kappa_0$ later.

Recall the Chebychev bound $\mathbb{P}\left(|X - np| \ge t\sqrt{np(1-p)}\right) \le t^{-2}$, for X a binomial random variable with parameters n and p. Since we are only interested in upper bounds, we bound $1-p_l$ by 1. This, together with the previous paragraph shows that we can conclude that $B_{N,M}$ and $A_l(\kappa)$ occur with probability at least $1-\eta$, and we henceforth assume they occur.

Now note that there exists some $c = c(M) = OM(Np_l)^{-1/2} < \infty$ such that on $B_{N,M}$, we have that

$$\kappa^{1/\alpha+\varepsilon} \le c\,[Np_l \vee 1]^{1/\alpha+\varepsilon}\,, \tag{4.34}$$

and hence there exists another constant $c < \infty$ such that the sum in the equation below is bounded by (also using the fact that $p_l \sim C_\alpha l^{-\beta}$)

$$\begin{aligned}\sum_{i=1}^{\kappa} \tilde{Y}_l^{(i)} - N &\le \kappa p_l^{-1} + \kappa^{1/\alpha+\varepsilon} p_l^{-1}[-1,1] - N \\ &\le \sqrt{N/p_l}[-M,M] + c\,[Np_l \vee 1]^{1/\alpha+\varepsilon}\, p_l^{-1}[-1,1] \\ &\le cN^{1/2} l^{\beta/2}[-M,M] + c(l^{\beta} \vee N^{1/\alpha+\varepsilon} l^{\beta(1-1/\alpha-\varepsilon)}[-1,1]) \\ &\le cN^{1/2} l^{\beta/2}[-M,M] + c(l^{\beta} \vee N^{1/\alpha+\varepsilon} l^{1/\alpha}[-1,1])\,.\end{aligned} \tag{4.35}$$

Now take some $\xi > 0$ and wlog assume that we originally picked $\varepsilon > 0$ small enough that $\varepsilon < \xi/2\alpha$. Then note that if $l \le N^{-\xi} \min\left\{N^{1-2\delta}, N^{\frac{1-\delta-1/\alpha}{\beta(1-1/\alpha)}}\right\} = N^{-\xi} N^{\alpha-\alpha\delta-1}$ (since $\delta < 1$ and $\alpha = (\beta+1)/\beta$), we have that the equation above is at most $cn^{\varepsilon} N^{1-\delta-\frac{\xi}{2\alpha}}$ and hence smaller than $N^{1-\delta} \vee n^{\beta-\delta}$ (for all sufficiently large n).

On the other hand, assume now that $l > N^{-\xi} N^{\alpha-\alpha\delta-1}$. Write $B = (A-1)/A$ and set $\zeta = \frac{\alpha\delta+\xi}{\alpha-1}$ and $\zeta' = \frac{2\zeta}{\alpha} + \varepsilon$. Recall that there exists $C < \infty$, not depending on n, such that $l \le Cn^{B}$ (see the remark under (4.29)) and recall that $l \ge N^{-\xi+\alpha(1-\delta)-1} = N^{(\alpha-1)(1-\zeta)}$. Provided that we chose δ and ξ sufficiently small in the first place, we have that $N \le l^{\beta+2\zeta}$. This implies that the latter expression in (4.35) is upper bounded by

$$cl^{\beta+\zeta}[-M,M] + cl^{(\beta+1)/\alpha+2\zeta/\alpha+\varepsilon}[-1,1] = cl^{\beta+\zeta}[-M,M] + cl^{\beta+\zeta'}[-1,1] \le cn^{B(\beta+\zeta')+\varepsilon}\,. \tag{4.36}$$

Here the final inequality holds on account of the condition imposed on l, and on substituting $M = n^{\varepsilon}$.

Now note that $\zeta' > 0$ can be made as small as we want, by decreasing δ, ε and ξ towards 0. Hence, since $B < 1$, we can decrease these all sufficiently that the right-hand side of (4.36) is at most $n^{\beta-\zeta}$.

It remains to show treat the case $\kappa \le \kappa_0$. However, in that case, we can assume that there is some $L = L(\eta, \kappa_0)$ such that $p_l N \le L$; otherwise $B_{N,M}$ doesn't hold, which can only happen with probability $1-\eta$, provided we chose L large enough. This implies that for some L', we have $N \le L' n^{\beta\left(1-A^{-1}\right)}$. From there on, we can proceed as before, as $\sum_{i=1}^{\kappa_0} \tilde{Y}_l^{(i)} < n^{\beta-\delta}$ outside a set of vanishing probability.

It follows that, for all $\eta > 0$ and n sufficiently large

$$\mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(E_n | y_n = x_n\right) | \mathcal{F}_n\right] \le \eta\,. \tag{4.37}$$

Hence using Markov's inequality we deduce that

$$\mathbf{P}\left(x_n \notin A_{\mathbf{T}}^{(n)} | \mathcal{F}_n\right) \le \eta \tag{4.38}$$

for all sufficiently large n. As $\eta > 0$ was arbitrary, the claim follows. □

Proof of Theorem 1.3. Theorem 1.3 now follows as a direct consequence of Lemma 4.4 and Proposition 4.5. □

5 Convergence of the branching process: proof of Theorem 1.4

In this section we will consider the convergence of the process $(n^{-\beta}Y_{n(1+t)})_{t\geq 0}$ under the conditioning $Y_n > 0$, in the Skorokhod-J_1 topology. Note that the Skorokhod-$J1$ topology is the the strongest amongst the different Skorokhod topologies, see [17]. For further background and definitions on the Skorokhod space of càdlàg functions (denoted $D[0,\infty)$) and the Skorokhod-J_1 topology we refer to [4, Section 12]. As outlined in the introduction, it is known that under the annealed law the process $(n^{-\beta}Y_{n(1+t)})_{t\geq 0}$ under the conditioning $Y_n > 0$ converges to a *continuous-state branching process* with entrance measure determined by (1.3); see for example [12, Theorem 2.1.9]. More precisely, the limiting process $(\tilde{Y}_t)_{t\geq 0}$ is a Markov process on $[0,\infty)$ such that $\tilde{Y}_0$ is a random variable with Laplace transform given by $\phi(\theta)$, and for all $t,x,y \geq 0$ its transition kernels $P_t(x,y)$ satisfy the *branching property*

$$P_t(x+y,\cdot) = P_t(x,\cdot) * P_t(y,\cdot). \tag{5.1}$$

We will continue to denote these transition kernels by $P_t(\cdot,\cdot)$ throughout this section. We also remark that for $a > 0$ and $t > s > 0$, the law of $\tilde{Y}_t$ conditionally on the event $\{\tilde{Y}_s = a\}$ is that of a sum of N independent copies of X where X has Laplace transform $\phi(\theta(t-s)^{\beta})$ and N has the distribution $\mathsf{Poisson}(aC_\alpha(t-s)^{-\beta})$. In particular its Laplace transform is therefore given by

$$Q_{a,s,t}(\theta) := \mathbb{E}\left[e^{-\theta\tilde{Y}_t}|\tilde{Y}_s = a\right] = \mathbb{E}\left[\mathbb{E}\left[e^{-\theta X}\right]^N\right] = \exp\{-aC_\alpha(t-s)^{-\beta}(1-\phi(\theta(t-s)^{\beta}))\}. \tag{5.2}$$

In fact we can and will use (5.2) as a characterisation of this limiting branching process.

By [4, Theorem 13.1], we can prove Theorem 1.4 by first showing that the sequence of processes is tight, and then showing that the finite dimensional marginals converge for all $t \in [0,\infty)$.

5.1 Tightness

We use a standard criteria for tightness in the Skorokhod-J_1 topology given in [4, Theorem 13.2]. If $X = (X_t)_{t\geq 0}$ is a càdlàg function, we define

$$||X||_\infty = \sup_{t\geq 0}|X_t|, \qquad \omega'_\delta(X) := \inf_{\{t_i\}_{i=0}^m}\sup_{1\leq i\leq m}\sup_{s,t\in[t_{i-1}t_i)}|X_t - X_s|,$$

where the infimum is taken over all partitions $0 = t_0 < t_1 < \ldots < t_m$ with $t_i - t_{i-1} > \delta$ for all i. For $K \geq 0$, we also let X^K denote the function defined by

$$X^K_t = X_t\mathbb{1}\{t \leq K\}.$$

We will use the following result.

Proposition 5.1 ([4, Theorem 13.2])**.** *Let $(X^{(n)})_{n\geq 0}$ be a sequence of càdlàg functions $[0,\infty) \to \mathbb{R}$ defined on a probability space $(\Omega,\mathcal{F},P)$. For $K > 0$ set $X^{(n,K)} = (X^{(n)})^K$. Then the sequence is tight if and only if for every $K > 0$, the following conditions are satisfied.*

(a) $\lim_{C\to\infty}\limsup_{n\to\infty} P\left(||X^{(n,K)}||_\infty \geq C\right) = 0.$

(b) For all $\varepsilon > 0$, $\lim_{\delta\downarrow 0}\limsup_{n\to\infty} P\left(\omega'_\delta(X^{(n,K)}) > \varepsilon\right) = 0.$

This leads us rather straightforwardly to the following proposition.

Proposition 5.2. *For each $n \geq 1$ let $(\tilde{Y}_t^{(n)})_{t\geq 0}$ denote the process $(n^{-\beta} Y_{n(1+t)})_{t \geq 0}$ under the conditioning $Y_n > 0$ (i.e., set $\tilde{Y}_t^{(n)} = n^{-\beta} Y_{n(1+t)}$ under this conditioning). Then, **P**-almost surely, the sequence of processes $(\tilde{Y}^{(n)})_{n \geq 1}$ is tight with respect to the Skorokhod-J_1 topology.*

Proof. If A_K is the subset of Ω where either (a) or (b) of Proposition 5.1 fail for K, then we have that A_K is an increasing function of K. As a result, in order to prove this proposition it is sufficient to show that, for every $K > 0$, the following hold $\mathbf{P}$-almost surely.

(a) $\lim_{C\to\infty} \limsup_{n\to\infty} \mathbb{P}_{\mathbf{T}} \left(||\tilde{Y}^{(n,K)}||_\infty \geq C \right) = 0.$

(b) For all $\varepsilon > 0$, $\lim_{\delta \downarrow 0} \limsup_{n\to\infty} \mathbb{P}_{\mathbf{T}} \left(\omega'_\delta(\tilde{Y}^{(n,K)}) > \varepsilon \right) = 0.$

We will show that, thanks to the convergence of the annealed process, the following two statements hold.

(a ′) $\mathbf{E}\left[\lim_{C\to\infty} \limsup_{n\to\infty} \mathbb{P}_{\mathbf{T}} \left(||\tilde{Y}^{(n,K)}||_\infty \geq C \right) \right] = 0.$

(b ′) For all $\varepsilon > 0$, $\mathbf{E}\left[\lim_{\delta \downarrow 0} \limsup_{n\to\infty} \mathbb{P}_{\mathbf{T}} \left(\omega'_\delta(\tilde{Y}^{(n,K)}) > \varepsilon \right) \right] = 0.$

Note that the expectation of a sequence of non-negative random variables can only be zero if the random variables in question are zero almost surely. Hence (a ′) and (b ′) directly imply that (a) and (b) are satisfied, $\mathbf{P}$-almost surely, and it just remains to establish (a ′) and (b ′).

To this end, set $x_C = \limsup_{n\to\infty} \mathbb{P}_{\mathbf{T}} \left(||\tilde{Y}^{(n,K)}||_\infty \geq C \right)$. Suppose for a contradiction that the expectation in (a ′) is not 0, and is in fact lower bounded by $2c > 0$. Then $\mathbf{P}\left(\lim_{C\to\infty} x_C \geq c \right) \geq c$.

Moreover, on the event $\{x_C > c\}$, we have by Fatou's lemma that

$\mathbb{P}_{\mathbf{T}} \left(||\tilde{Y}^{(n,K)}||_\infty \geq C \text{ i.o.} \right) \geq c$. Since this holds for all C on the event in question (and the event is monotone in C), it follows from monotone convergence that

$\mathbb{P}_{\mathbf{T}} \left(\forall C > 0 : ||\tilde{Y}^{(n,K)}||_\infty \geq C \text{ i.o.} \right) \geq c$. However, on the event $\{\forall C > 0 : ||\tilde{Y}^{(n,K)}||_\infty \geq C \text{ i.o.}\}$ it follows from [4, Theorem 12.3] that $\tilde{Y}^{(n,K)}$ cannot belong to a compact set. Since this event occurs with $\mathbb{P}$-probability at least c^2, this would imply that the annealed law is not tight, which is a contradiction.

Similarly, suppose for a contradiction that there exists $\varepsilon > 0$ such that the expectation in (b ′) is lower bounded by $2c$. Set $y_\delta := \limsup_{n\to\infty} \mathbb{P}_{\mathbf{T}} \left(\omega'_\delta(\tilde{Y}^{(n,K)}) > \varepsilon \right)$. Then necessarily $\mathbf{P}\left(\lim_{\delta\to 0} y_\delta > c \right) > c$, and moreover on the event $\{y_\delta > c\}$, we have by Fatou's lemma that $\mathbb{P}_{\mathbf{T}} \left(\omega'_\delta(\tilde{Y}^{(n,K)}) > \varepsilon \text{ i.o.} \right) > c$.

By monotone convergence, it follows that $\mathbb{P}_{\mathbf{T}} \left(\forall \delta > 0 : \omega'_\delta(\tilde{Y}^{(n,K)}) > \varepsilon \text{ i.o.} \right) > c$, and moreover on the event $\{\forall \delta > 0 : \omega'_\delta(\tilde{Y}^{(n,K)}) > \varepsilon \text{ i.o.}\}$, it follows from [4, Theorem 12.3] that $\tilde{Y}^{(n,K)}$ cannot belong to a compact set. Since this event occurs with $\mathbb{P}$-probability at least c^2, this would similarly imply that the annealed law is not tight, which is a contradiction. □

5.2 Convergence of the finite dimensional distributions along a subsequence

Throughout this section, we abbreviate $Y_t^{(n)} = n^{-\beta} Y_{\lfloor tn \rfloor}$. As in the previous subsections, the strategy will be to show that the finite dimensional marginals converge almost surely along a subsequence, and then lift this to convergence along the full sequence using the continuity event F_n that we used in Lemma 4.4.

We start with a variance bound that will be useful for some technical estimates.

Lemma 5.3 (A second moment estimate)**.** *For any* $\varepsilon > 0$*, as* $n \to \infty$

$$\mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right)^2\right] = O\left(n^{-\beta\alpha+\varepsilon}\right) . \tag{5.3}$$

Proof. Fix $\delta > 0$ and write $p = \alpha - \delta$. We bound the second moment by the p^{th} moment:

$$\mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right)^2\right] \le \mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right)^p\right] . \tag{5.4}$$

We now set $m = \lceil \frac{\beta}{\log\mu} \log n \rceil$. Note that, by a union bound, we have that

$$\begin{aligned} &\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right) \le \sum_{v\in\mathbf{T}_m} \mu^{-m}\mathbb{P}_{\mathbf{T}}\left(v \overset{*}{\longleftrightarrow} \mathbf{T}_n\right) \\ &= \sum_{v\in\mathbf{T}_m} \mu^{-m}\mathbb{P}\left(\mathbf{0} \overset{*}{\longleftrightarrow} \mathbf{T}_{n-m}\right) + \sum_{v\in\mathbf{T}_m} \mu^{-m}\left[\mathbb{P}_{\mathbf{T}}\left(v \overset{*}{\longleftrightarrow} \mathbf{T}_n\right) - \mathbb{P}\left(\mathbf{0} \overset{*}{\longleftrightarrow} \mathbf{T}_{n-m}\right)\right] := X_1 + X_2 \,, \end{aligned} \tag{5.5}$$

where X_1 and X_2 denote the first and second sums respectively. Now note that

$$\mathbf{E}\left[X_1^p\right] \le O\left((n-m)^{-\beta p}\right)\mathbf{E}\left[\mathbf{W}_m^p\right] = O\left(n^{-\beta p}\right) , \tag{5.6}$$

by the tower property, (1.2) and since $(\mathbf{W}_m)_m$ converges in $\mathrm{L}^p(\mathbf{P})$ when $p < \alpha$.

Furthermore, by Equation (3.16) and the tower property, we have that there exists $C < \infty$ such that for any $x > 0$ (note that although we had set $m = \lfloor \frac{(1+\varepsilon)}{(\alpha-1)\log\mu} \log n \rfloor$ in the proof of Lemma 3.3, the proof up until (3.16) did not use this, so it also holds for the present value of m)

$$\mathbf{P}\left(|X_2| > x\right) \le C\frac{n^{-\beta}}{x^2\mu^m} . \tag{5.7}$$

In particular, we have for any $x > 0$ that

$$\mathbf{P}\left(|X_2| > xn^{-\beta}\right) \le C\frac{n^{\beta}}{x^2\mu^m} \le \frac{C}{x^2}. \tag{5.8}$$

Here the final inequality follows by our choice of m. Hence, the p-moment of $|X_2|$ exists and is bounded by $O\left(n^{-\beta p}\right)$. To tie up, note that

$$\mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right)^p\right] = \mathbf{E}\left[(X_1+X_2)^p\right] \le 2^p(\mathbf{E}\left[X_1^p\right] + \mathbf{E}\left[|X_2|^p\right]) = O\left(n^{-\beta p}\right) = O\left(n^{-\beta\alpha+\varepsilon}\right) , \tag{5.9}$$

for some $\varepsilon > 0$ depending on δ, which can be as small as we want to. Substituting back into (5.4) concludes the proof. □

Remark 5.4. We note that the above lemma allows for a subsequent improvement of the estimate made in Equation (3.14).

We are now ready to prove convergence of the marginals along an appropriate subsequence. Recall the definition of the subsequence $(n_k)_{k\ge1}$ that we defined in Section 2 (in fact it will only be important in this section that n_k grow slower than k^2).

The convergence follows from Lemma 5.3 and a simple application of Chebyshev's inequality. Before giving the proof, we remind the reader of the following formula for the variance of a product of independent variables $X_1, \dots, X_M$:

$$\mathbf{Var}\left(\prod_{i=1}^M X_i\right) = \prod_{i=1}^M \left(\mathbf{Var}\left(X_i\right) + \mathbf{E}\left[X_i\right]^2\right) - \prod_{i=1}^M \mathbf{E}\left[X_i\right]^2 . \tag{5.10}$$

Lemma 5.5 (Pointwise convergence of the Laplace transform)**.** *Take any $a > 0, t > 0, s > 0$, and for each $n \geq 1$ set $a_n = n^{-\beta}\lfloor an^{\beta}\rfloor$. For each $n \geq 1$ set*

$$Q^{n,\boldsymbol{T}}_{a,s,t}(\theta) = \mathbb{E}_{\mathbf{T}}\left[\exp\{-\theta Y_t^{(n)}\}\middle| Y_s^{(n)} = a_n\right]$$

and let $Q_{a,s,t}(\theta)$ denote the analogous quantity in the limit of the annealed model, i.e. as in (5.2)*. Then for $\boldsymbol{P}$-almost every tree $\boldsymbol{T}$ and for all $\theta > 0$,*

$$Q^{n_k,\boldsymbol{T}}_{a_n,s,t}(\theta) \to Q_{a,s,t}(\theta) \tag{5.11}$$

as $k \to \infty$.

Proof. In the following proof we will abuse notation slightly by writing $t-s$ in place of $\frac{1}{n}\left(\lfloor tn\rfloor - \lfloor sn\rfloor\right)$, and $(t-s)n$ in place of $\lfloor tn\rfloor - \lfloor sn\rfloor$ (any resulting errors can be absorbed into the $(1+o(1))$ factors that are already present in the proof).

Fix some $s, t, a, \theta > 0$. We condition on $\mathbf{T}_{\lfloor sn\rfloor}$ and write

$$S = \sum_{v \in \mathbf{T}_{\lfloor sn\rfloor}:\, 0 \overset{*}{\longleftrightarrow} v} Y^{(v)}_{(t-s)n},$$

where for $v \in \mathbf{T}_{\lfloor sn\rfloor}$ and $\ell \in \mathbb{N}$ we write $Y_\ell^{(v)}$ to denote vertices in generation ℓ in the subtree rooted at v.

For each $n \geq 1$ recall that $\mathcal{F}_n$ is the σ-algebra generated by possible realisations of the first n levels of $\mathbf{T}$, so that for the set $y_{\lfloor sn\rfloor} := \{v \in \mathbf{T}_{\lfloor sn\rfloor} : 0 \overset{*}{\longleftrightarrow} v\}$, the probability $\mathbb{P}_{\mathbf{T}}\left(y_{\lfloor sn\rfloor} = A | Y_s^{(n)} = a_n\right)$ is $\mathcal{F}_{\lfloor sn\rfloor}$-measurable, for any $A \subset \mathbf{T}_{\lfloor sn\rfloor}$. We now compute the moment-generating function $\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta S} | Y_s^{(n)} = a_n\right]$ conditionally on $\mathcal{F}_{\lfloor sn\rfloor}$.

Write $X_n = \mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta S} | Y_s^{(n)} = a_n\right]$. Note that $\mathbf{E}\left[X_n | \mathcal{F}_{\lfloor sn\rfloor}\right]$ is given by

$$\left[1 - \mathbb{P}\left(Y^{(n)}_{t-s} > 0\right)\left(1 - \mathbb{E}\left[\mathrm{e}^{-\theta Y^{(n)}_{t-s}} | Y^{(n)}_{t-s} > 0\right]\right)\right]^{a_n n^{\beta}} = \mathrm{e}^{a(t-s)^{-\beta} C_\alpha\left(\phi(\theta(t-s)^{\beta}) - 1\right)}\left(1 + o(1)\right), \tag{5.12}$$

where the final equality holds by (1.3). On the other hand, we have from (5.10) that

$$\mathbf{Var}\left(X_n | \mathcal{F}_{\lfloor sn\rfloor}\right) = \left[\mathbf{Var}\left(\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta Y^{(n)}_{t-s}}\right]\right) + \mathbb{E}\left[\mathrm{e}^{-\theta Y^{(n)}_{t-s}}\right]^2\right]^{a_n n^{\beta}} - \mathbb{E}\left[\mathrm{e}^{-\theta Y^{(n)}_{t-s}}\right]^{2a_n n^{\beta}}. \tag{5.13}$$

Now take some small $\varepsilon > 0$. We note that, by Lemma 5.3, there exists $C < \infty$ such that for all $n \geq 1$,

$$\begin{aligned} \mathbf{Var}\left(\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta Y^{(n)}_{t-s}}\right]\right) &= \mathbf{Var}\left(1 - \mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta Y^{(n)}_{t-s}}\right]\right) \\ &= \mathbf{Var}\left(\mathbb{P}_{\mathbf{T}}\left(Y^{(n)}_{t-s} > 0\right)\left(1 - \mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta Y^{(n)}_{t-s}} | Y^{(n)}_{t-s} > 0\right]\right)\right) \\ &\leq \mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(Y^{(n)}_{t-s} > 0\right)^2\right] \leq C n^{-\beta\alpha + \varepsilon}. \end{aligned} \tag{5.14}$$

Moreover, by (1.2) and (1.3), we have that

$$\mathbb{E}\left[\mathrm{e}^{-\theta Y^{(n)}_{t-s}}\right] = 1 - C_\alpha (t-s)^{-\beta} n^{-\beta}(1 - \phi(\theta(t-s)^{\beta}))(1 + o(1)). \tag{5.15}$$

Now recall that for $x, y > 0$ and $m \geq 1$ we have that $(x+y)^m - y^m \leq mx(x+y)^{m-1}$. Applying this to (5.13) and substituting the estimates from (5.14) and (5.15), we deduce that there exist $c' > 0$ and other $c, C, C' < \infty$, possibly depending on α, θ, s and t such that

$$\begin{aligned}\mathbf{Var}\left(X_n|\mathcal{F}_{\lfloor sn\rfloor}\right) &\leq a_n n^\beta \mathbf{Var}\left(\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta Y_{t-s}^{(n)}}\right]\right)\left[\mathbf{Var}\left(\mathbb{E}_{\mathbf{T}}\left[\mathrm{e}^{-\theta Y_{t-s}^{(n)}}\right]\right)+\mathbb{E}\left[\mathrm{e}^{-\theta Y_{t-s}^{(n)}}\right]^2\right]^{a_n n^\beta-1}\\ &\leq Cn^{-\beta(\alpha-1)+\varepsilon}(1+cn^{-\alpha\beta+\varepsilon}-c'n^{-\beta})^{a_n n^\beta-1}\\ &\leq C'n^{-\beta(\alpha-1)+\varepsilon}=C'n^{-1+\varepsilon}.\end{aligned} \tag{5.16}$$

Hence Chebyshev's inequality shows that $X_n \to \mathrm{e}^{a(t-s)^{-\beta}C_\alpha\left(\phi(\theta(t-s)^\beta)-1\right)} = Q_{a,s,t}(\theta)$ almost surely along the subsequence $(n_k)_{k\geq 1}$

This currently proves the result for any fixed θ, $\mathbf{P}$-almost surely. Note that we can extend this to hold simultaneously for all $\theta > 0$ using the same argument as in Corollary 4.2. □

In fact we need to strengthen Lemma 5.5 to the following.

Lemma 5.6 (Subsequential convergence of the Laplace transform)**.** *For* $\mathbf{P}$*-almost every tree* $\mathbf{T}$ *the following holds for almost every* $a > 0$ *and every* $t > s > 0, \theta > 0$*. For each* $n \geq 1$ *again set* $a_n = n^{-\beta}\lfloor an^\beta\rfloor$ *and*

$$Q_{a,s,t}^{n,\mathbf{T}}(\theta) = \mathbb{E}_{\mathbf{T}}\left[\exp\{-\theta Y_t^{(n)}\}\middle| Y_s^{(n)} = a_n\right].$$

Let $Q_{a,s,t}(\theta)$ *denote the analogous quantity in the annealed model, i.e. as in* (5.2)*. Then*

$$Q_{a,s,t}^{n_k,\mathbf{T}}(\theta) \to Q_{a,s,t}(\theta) \tag{5.17}$$

as $k \to \infty$*.*

Proof. In the previous lemma, we showed that for any fixed s, t and a, the convergence holds for all $\theta > 0$ for $\mathbf{P}$-almost every tree. This immediately implies the following: for $\mathbf{P}$-almost every tree, the statement is true for Lebesgue almost every $a, s, t \geq 0$, including all rationals, using Fubini's theorem.

Now note that, $\mathbf{P}$-almost surely, for any fixed θ, any $s > 0$ and almost every $a > 0$, the function $Q_{a_n,s,t}^{n,\mathbf{T}}(\theta)$ is equal to the limit of $Q_{a_n,s,t'}^{n,\mathbf{T}}(\theta)$ as $t' \downarrow t$, and this continues to hold as $n \to \infty$. Otherwise, it would contradict point (ii) in Proposition 5.2, which says that for all $\varepsilon > 0$, $\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathbb{P}_{\mathbf{T}}\left(\omega'_\delta(\tilde{Y}^{(n,K)}) > \varepsilon\right) = 0$ (here, we take $K > t$ w.l.o.g.). Since the limiting function $Q_{a,s,t}(\theta)$ is continuous in all variables, we deduce the following: for almost every s, it holds for almost every choice of a that the function $Q_{a,s,t}(\theta)$ converges for all θ and all t. Denote the set of good s where this holds by S^{good}.

Then if $s \notin S^{\text{good}}$, choose some other $s' \in S^{\text{good}}$ and set $c = \frac{s}{s'}$. Given $t > s$, also set $t' = \frac{t}{c}$. Then $Y_s^{(n)} = c^\beta Y_{s'}^{(cn)}$ and $Y_t^{(n)} = c^\beta Y_{t'}^{(cn)}$. Suppose moreover that a is such that $Q_{ac^{-\beta},s',t'}^{cn,\mathbf{T}}(\theta)$ converges for all θ (this will be the case for almost every a by virtue of the fact that $s' \in S^{\text{good}}$ and c is fixed). It then follows that (also using (5.1))

$$\begin{aligned}Q_{a,s,t}^{n,\mathbf{T}}(\theta) = \mathbb{E}_{\mathbf{T}}\left[\exp\{-\theta Y_t^{(n)}\}\middle| Y_s^{(n)} = a_n\right] &= \mathbb{E}_{\mathbf{T}}\left[\exp\{-c^\beta\theta Y_{t'}^{(cn)}\}\middle| Y_{s'}^{(cn)} = a_n c^{-\beta}\right]\\ &= Q_{ac^{-\beta},s',t'}^{cn,\mathbf{T}}(c^\beta\theta)\\ &\to \exp\{-c^{-\beta}aC_\alpha(t'-s')^{-\beta}(1-\phi(c^\beta\theta(t'-s')^\beta)\}\\ &= \exp\{-aC_\alpha(t-s)^{-\beta}(1-\phi((t-s)^\beta\theta))\}\\ &= Q_{a,s,t}(\theta).\end{aligned} \tag{5.18}$$

We therefore now have the following statement for **P**-almost every tree: for all s, we have for almost every a that the convergence holds for all t, θ. ☐

Remark 5.7. Similar results that could be applied to lift the convergence of Lemma 5.5 to that of Lemma 5.6 appear in [2, Theorem 12.12].

5.3 Lifting to the full sequence

Similarly to Section 4, we fix a small $\delta > 0$ and define for every $\varepsilon > 0$ the events

$$\begin{aligned} E_n &= \left\{ |Y_n - Y_{n_k}| > Y_n^{1-\delta} \vee n^{\beta-\delta} \right\}, \\ F_n &= F_n(\mathbf{T}, \varepsilon) = \left\{ \mathbb{P}_{\mathbf{T}} \left(E_n | \mathcal{G}_n^{\mathbf{T}}, Y_n > 0 \right) < \varepsilon/4 \right\}, \end{aligned} \tag{5.19}$$

where k is chosen so that $n_{k-1} \leq n \leq n_k$. We then moreover showed in Proposition 4.5 that for all $\varepsilon > 0$,

$$\mathbf{P}\left(F_n | \mathcal{F}_n\right) \geq 1 - o(1) \qquad \text{everywhere}. \tag{5.20}$$

For this proof we slightly modify the events in question: assume $sn, tn \in \mathbb{N}$ and set

$$F_n(a, s, t) = \left\{ \mathbb{P}_{\mathbf{T}} \left(E_{tn} | \mathcal{G}_{tn}^{\mathbf{T}}, Y_{tn} > 0, Y_{sn} = \lfloor an^{\beta} \rfloor \right) < \varepsilon/4 \right\}. \tag{5.21}$$

Lemma 5.8. *For all $a > 0, t > s > 0$, as $n \to \infty$,*

$$\boldsymbol{P}\left(F_n(a, s, t) | \mathcal{F}_{tn}\right) \geq 1 - o(1). \tag{5.22}$$

Proof. Since the event $\{Y_{sn} = \lfloor an^{\beta} \rfloor\}$ is $\mathcal{G}_{tn}^{\mathbf{T}}$-measurable, this is a direct consequence of Proposition 4.5. ☐

We now prove that Lemma 5.8 is a sufficient condition to lift Lemma 5.6 to the whole sequence.

Lemma 5.9 (Criterion for full convergence)**.** *Suppose that for every $\varepsilon > 0$ Equation* (5.8) *holds. Then for $\boldsymbol{P}$-almost every tree $\boldsymbol{T}$ the following holds for every $t > 0, s > 0, \theta > 0$, for almost every $a > 0$. For all $n \geq 1$ set $a_n = n^{-\beta} \lfloor an^{\beta} \rfloor$ and*

$$Q_{a,s,t}^{n,\boldsymbol{T}}(\theta) = \mathbb{E}_{\mathbf{T}} \left[\exp\{-\theta Y_t^{(n)}\} \middle| Y_s^{(n)} = a_n \right],$$

and let $Q_{a,s,t}(\theta)$ denote the analogous quantity in the annealed model, i.e. as in (5.2)*. Then*

$$Q_{a,s,t}^{n,\boldsymbol{T}}(\theta) \to Q_{a,s,t}(\theta) \tag{5.23}$$

as $n \to \infty$.

Proof. The proof is similar to that of Lemma 4.4 and we just write the details for completeness. In particular we already showed that the result holds almost surely along the subsequence $(n_k)_{k \geq 1}$ in Lemma 5.6.

Now fix some $\varepsilon > 0$ and for $n, k \geq 0$ write

$$\begin{aligned} A_n &= \left\{ \left| Q_{a,s,t}^{n,\mathbf{T}}(\theta) - Q_{a,s,t}(\theta) \right| > 2\varepsilon \right\}, \\ B_k &= \left\{ \left| Q_{a,s,t}^{n_k,\mathbf{T}}(\theta) - Q_{a,s,t}(\theta) \right| > \varepsilon \right\}. \end{aligned} \tag{5.24}$$

Note that A_n is $\mathcal{F}_{tn}$ measurable and B_k is $\mathcal{F}_{tn_k}$ measurable. For the rest of the proof we treat a, s and t as fixed and write F_n in place of $F_n(a, s, t)$.

By the same arguments as in Lemma 4.3, we have that there exists $K_o < \infty$ such that for all $k \geq K_o$ and all $n \in [n_{k-1}, n_k]$, A_n implies B_k on the event F_n. In other words we have for all $n_{k-1} \leq n \leq n_k$ that

$$A_n \cap F_n \subset B_k. \tag{5.25}$$

It follows that for any $K > K_o$,

$$\mathbf{P}\left(\exists k \geq K \colon B_k \text{ occurs}\right) \geq \mathbf{P}\left(\exists n \geq n_K \colon A_n \cap F_n \text{ occurs}\right). \tag{5.26}$$

Set $\tau_K = \inf\{n \geq n_K : A_n \text{ occurs}\}$. Since τ_K is a stopping time, on the event $\{\tau_K < \infty\}$ we have that $\mathbf{P}\left(F_{\tau_K} | \tau_K < \infty, \mathcal{F}_{t\tau_K}\right) \geq \frac{1}{2}$ (provided K was large enough) by Lemma 5.8. Hence we can lower bound the right-hand side of (5.26) by

$$\mathbf{E}\left[\mathbf{P}\left(\tau_K < \infty \text{ and } F_{t\tau_K} | \mathcal{F}_{\tau_K}\right)\right] \geq \frac{1}{2}\mathbf{P}\left(\tau_K < \infty\right). \tag{5.27}$$

Combining (5.26) and (5.27) gives

$$\mathbf{P}\left(\exists k \geq K \colon B_k \text{ occurs}\right) \geq \frac{1}{2}\mathbf{P}\left(\exists n \geq n_K \colon A_n \text{ occurs}\right). \tag{5.28}$$

On account of Lemma 5.6, we have that B_k occurs only finitely many times $\mathbf{P}$-almost surely, and hence the left hand side above tends to 0 as $K \to \infty$. We deduce that the same holds for the right hand side, and therefore that $\mathbf{P}\left(A_n \text{ i.o.}\right) = 0$. Since this holds for any $\varepsilon > 0$, this proves the claim. □

Corollary 5.10. *We have that* $\mathbf{P}$*-almost surely, for all* $t > s > 0$ *it holds for almost every* $a > 0$ *and any Borel set* B *that*

$$\mathbb{P}_{\mathbf{T}}(Y_t^{(n)} \in B | Y_s^{(n)} = a) = P_{t-s}(a, B) + o(1). \tag{5.29}$$

Proof. Lemma 5.9 implies that for all $t > s > 0$ that for almost every $a > 0$, we have almost sure convergence in distribution of the law $\mathbb{P}_{\mathbf{T}}\left(Y_t^{(n)} \in \cdot | Y_s^{(n)} = a\right)$ to $\mathbb{P}(Y_t \in \cdot | Y_s = a)$. □

By iterating this result and combining with Theorem 1.3, we deduce the following.

Lemma 5.11. *For* $\mathbf{P}$*-almost every tree* $\mathbf{T}$ *the following holds. Take any* $d \geq 1$*, any* $t_1 < t_2 < \ldots < t_d \in (1, \infty)$ *and let* f *be the density function of the random variable with Laplace transform* $\phi(\theta)$ *as in* (1.3)*. Take any Borel sets* $B_1, \ldots, B_d \subset [0, \infty)$*. Also set* $t_0 = 1$ *and let* P_t *be as in* (5.1)*. Then*

$$\begin{aligned} &\lim_{n \to \infty} \mathbb{P}_{\mathbf{T}}\left(n^{-\beta} Y_{\lfloor t_i n \rfloor} \in B_i \text{ for all } i \in \{1, \ldots, d\} \,\middle|\, Y_n > 0\right) \\ &\qquad = \int_{B_0} \mathrm{d}x_0 \int_{B_1} \mathrm{d}x_1 \ldots \int_{B_d} \mathrm{d}x_d f(x_0) \prod_{i=1}^{d} P_{t_i - t_{i-1}}(x_{i-1}, x_i). \end{aligned} \tag{5.30}$$

In particular, Proposition 5.2 and Lemma 5.11 together imply the result of Theorem 1.4.

6 The IIC: proof of Theorem 1.5

Let us first define the measure on the IIC. For this, for a tree $T \subset \mathbf{T}$ we let $T[n]$ denote the subtree of T obtained by keeping only the vertices at level n and below, and the edges between them. Recall also that for a finite tree t, t_n denotes the collection of vertices at exactly level n in t. If $\mathbf{0} \subset t \subset \mathbf{T}$, the height of t is equal to the maximal generation reached by t (i.e. $\sup_{v \in t} |v|$).

In this section we let $\mathcal{H}_n(\mathbf{T})$ denote the set of subtrees of $\mathbf{T}$ that contain $\mathbf{0}$ and have height n.

Definition 6.1 (Law of the IIC)**.** *For any tree $t \in \mathcal{H}_n(\boldsymbol{T})$, set*

$$\mu_{\mathbf{T}}\Big|_n(t) = \frac{\sum_{v \in t_n} \boldsymbol{W}(v)}{\boldsymbol{W}} \mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n] = t\right), \tag{6.1}$$

where $\mathcal{C}_{p_c}$ is the root cluster, and hence $\mathcal{C}_{p_c}[n]$ is the root cluster restricted to levels n and below.

We emphasise that for this definition to make sense we must view t as a subset of $\mathbf{T}$, rather than as an arbitrary plane tree. We see that this is well-defined in the following lemma.

Lemma 6.2. *For every $t \in \mathcal{H}_n(\boldsymbol{T})$,*

$$\mu_{\mathbf{T}}\Big|_n(t) = \lim_{M \to \infty} \mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n] = t \middle| 0 \leftrightarrow \boldsymbol{T}_M\right). \tag{6.2}$$

Moreover, the collection of measures $\left(\mu_{\mathbf{T}}\Big|_n\right)_{n\geq 1}$ are consistent and therefore extend to a unique measure on infinite trees, which we will henceforth denote by $\mu_{\mathbf{T}}$.

Proof. The first part of the proof is a carbon copy of that of [16, Lemma 3.9]. In particular, we can write (applying Theorem 1.2 in the second and third steps, and where $O(\cdot)$ refers to the limit as $M \to \infty$ with n fixed)

$$\begin{aligned}
\mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n] = t \middle| 0 \leftrightarrow \mathbf{T}_M\right) &= \frac{\mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n] = t \text{ and } 0 \leftrightarrow \mathbf{T}_M\right)}{\mathbb{P}_{\mathbf{T}}\left(0 \leftrightarrow \mathbf{T}_M\right)} \\
&= \mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n] = t\right) \frac{\sum_{v \in \mathbf{T}_n} \mathbb{P}_{\mathbf{T}}\left(v \leftrightarrow \mathbf{T}_M\right) + O(|t_n|^2 M^{-2\beta})}{\mathbb{P}_{\mathbf{T}}\left(0 \leftrightarrow \mathbf{T}_M\right)} \\
&\to \frac{\sum_{v \in \mathbf{T}_n} \mathbf{W}(v)}{\mathbf{W}} \mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n] = t\right).
\end{aligned} \tag{6.3}$$

To prove that the measures are consistent, we let $\mathsf{Ch}(t_n)$ denote the set of children of t_n, and let $\mathsf{Ch}_k(t_n)$ denote the collection of subsets of $\mathsf{Ch}(t_n)$ of size k. Note that, if v is a child of a vertex in t_n, then

$$|\{S \subset \mathsf{Ch}_k(t_n) : v \in S\}| = \binom{|\mathsf{Ch}(t_n)| - 1}{k - 1}.$$

Hence, if $t \in \mathcal{H}_n(\mathbf{T})$, we can write (note that going from the third to the fourth line we can eliminate the option $k = 0$, since the corresponding tree has probability 0 under $\mu_{\mathbf{T}}\Big|_{n+1}$ by Equation 6.1)

$$\begin{aligned}
\mu_{\mathbf{T}}\Big|_{n+1}(t) &= \mu_{\mathbf{T}}\Big|_{n+1}\left(\{t' \in \mathcal{H}_{n+1}(\mathbf{T}) : t'[n] = t\}\right) \\
&= \sum_{S \subset \mathsf{Ch}(t_n)} \frac{\sum_{v \in S} \mathbf{W}(v)}{\mathbf{W}} \mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n+1] = t \cup \{S\}\right) \\
&= \sum_{S \subset \mathsf{Ch}(t_n)} \frac{\sum_{v \in S} \mathbf{W}(v)}{\mathbf{W}} p_c^{|S|} (1 - p_c)^{|\mathsf{Ch}(t_n)| - |S|} \mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n] = t\right) \\
&= \sum_{k=1}^{|\mathsf{Ch}(t_n)|} \sum_{S \subset \mathsf{Ch}_k(t_n)} \frac{\sum_{v \in S} \mathbf{W}(v)}{\mathbf{W}} p_c^k (1 - p_c)^{|\mathsf{Ch}(t_n)| - k} \mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n] = t\right) \\
&= \sum_{k=1}^{|\mathsf{Ch}(t_n)|} \sum_{v \in \mathsf{Ch}(t_n)} \binom{|\mathsf{Ch}(t_n)| - 1}{k - 1} \frac{\mathbf{W}(v)}{\mathbf{W}} p_c^k (1 - p_c)^{|\mathsf{Ch}(t_n)| - k} \mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n] = t\right).
\end{aligned} \tag{6.4}$$

Now note that for any $v \in \mathbf{T}$, we have that $W(v) = p_c \sum W(w)$, where the sum is over all children of v. Hence the last line above is equal to

$$\sum_{k=1}^{|\mathsf{Ch}(t_n)|} \binom{|\mathsf{Ch}(t_n)|-1}{k-1} \frac{\sum_{v\in t_n} \mathbf{W}(v)}{\mathbf{W}} p_c^{k-1}(1-p_c)^{|\mathsf{Ch}(t_n)|-k} \mathbb{P}_{\mathbf{T}}\left(\mathcal{C}_{p_c}[n]=t\right)$$
$$= \sum_{k=1}^{|\mathsf{Ch}(t_n)|} \binom{|\mathsf{Ch}(t_n)|-1}{k-1} p_c^{k-1}(1-p_c)^{|\mathsf{Ch}(t_n)|-k} \mu_{\mathbf{T}}\Big|_n(t) = \mu_{\mathbf{T}}\Big|_n(t)\,, \tag{6.5}$$

as required. The measures therefore extend by Kolmogorov's consistency theorem. □

Now let $\mathbf{Z}_n$ be the number of vertices in the IIC at level n, rescaled by (i.e. divided by) a factor of n^{β}. Recall that we let Y denote the random variable with Laplace transform appearing in (1.3), and let Y^* denote its size-biased version, meaning that (also using (1.4)):

$$\mathbb{P}\left(Y^* \in [a,b]\right) = \frac{\mathbb{E}\left[Y\mathbb{1}\{Y\in[a,b]\}\right]}{\mathbb{E}\left[Y\right]} = C_\alpha \mathbb{E}\left[Y\mathbb{1}\{Y\in[a,b]\}\right]. \tag{6.6}$$

The main aim of this section is to prove Theorem 1.5, which says that for $\mathbf{P}$-almost every tree, $\mathbf{Z}_n$ converges in distribution to Y^*.

Before launching into the proof, we give a short lemma, for which we use the following notation. For $n \geq 0, M \geq 0, K > 0$ and $v \in \mathbf{T}_n$, we write

$$p_v^{n,M} = \mathbb{P}_{\mathbf{T}}\left(v \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M}\right), \quad p_v^{n,M}(K) = p_v^{n,M}\mathbb{1}\{p_v^{n,M} \leq KM^{-\beta}\}\,.$$

Lemma 6.3. *There exists $K < \infty$ such that for all $n \geq 1$,*

$$\mathbf{Var}\left(\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right)\mathbb{1}\{\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right) \leq Kn^{-\beta}\}\right) \leq C_\alpha K n^{-2\beta}\,, \tag{6.7}$$

Proof.

$$\begin{aligned}&\mathbf{Var}\left(\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right)\mathbb{1}\{\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right) \leq Kn^{-\beta}\}\right)\\ &\quad\leq \mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right)^2\mathbb{1}\{\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right) \leq Kn^{-\beta}\}\right]\\ &\quad\leq Kn^{-\beta}\mathbf{E}\left[\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right)\right] = C_\alpha K n^{-2\beta}.\end{aligned}$$

□

This enables us to prove the following lemma, which is the key technical estimate in the proof of Theorem 1.5. In the following proof and for the rest of this section, we recall $n_k = k^{\frac{\sqrt{\alpha}+1}{\sqrt{\alpha}-1}}$ for all $k \geq 1$.

Lemma 6.4 (Subsequential control of connection probabilities)**.** *Fix $n \geq 0$ and an interval $(a,b) \subset (0,\infty)$ and suppose that A_n is a sequence of (potentially random) sets such that $A_n \subset \boldsymbol{T}_n$, such that $|A_n| \in [an^\beta, bn^\beta]$, and such that the law of A_n, denoted $\mathbb{P}_{\boldsymbol{T}}^{A_n}$, is measurable with respect to $\mathcal{F}_n$. Fix any $\delta > 0$ and let E_n^δ denote the event that there is a subsequence $(M_j)_{j=1}^\infty$ such that*

$$\mathbb{P}_{\boldsymbol{T}}^{A_n}\left(\left|\frac{M_j^\beta}{n^\beta}\sum_{v\in A_n} p_v^{n,M_j} - |A_n|n^{-\beta}C_\alpha\right| < \delta \text{ eventually as } j\to\infty\right) > 1-\delta.$$

Then E_n^δ occurs eventually $\boldsymbol{P}$-almost surely along the subsequence $(n_k)_{k=1}^\infty$.

Proof. First set $\delta' = \frac{\delta^2}{4bC_\alpha}$ and choose $\tilde{K} < \infty$ such that $\mathbf{P}\left(\mathbb{P}_{\mathbf{T}}\left(Y_n > 0\right) \geq \tilde{K}n^{-\beta}\right) \leq \delta'$ for all $n \geq 1$ (note that this is possible by Theorem 1.2). We will abbreviate A_n by A, let

$\mathbb{P}_{\mathbf{T}}^A$ denote the law of A, and write estimates conditionally on both A and $\mathbb{P}_{\mathbf{T}}^A$ (which is $\mathcal{F}_n$-measurable by assumption) in what follows.

First note that we can write for any $K \geq \tilde{K}$ and all sufficiently large M that for any fixed $A' \subset \mathbf{T}_n$ satisfying $|A'| \leq bn^\beta$,

$$\mathbf{E}\left[\left|\frac{M^\beta}{n^\beta}\sum_{v\in A'} p_v^{n,M}(K) - |A'|n^{-\beta}C_\alpha\right|\middle|\mathcal{F}_n\right] \leq 2\delta'|A'|n^{-\beta}C_\alpha \leq 2\delta' bC_\alpha, \tag{6.8}$$

by (1.2) and by our definition of $\tilde{K}$. Conditionally on $\mathcal{F}_n$, the same estimate holds when A is random with law $\mathbb{P}_{\mathbf{T}}^A$ by averaging its possible realisations according to $\mathbb{P}_{\mathbf{T}}^A$. Hence, since $\mathbb{P}_{\mathbf{T}}^A$ is $\mathcal{F}_n$-measurable, we can apply the tower property to deduce that

$$\mathbf{E}\left[\mathbb{E}_{\mathbf{T}}^A\left[\left|\frac{M^\beta}{n^\beta}\sum_{v\in A} p_v^{n,M}(K) - |A|n^{-\beta}C_\alpha\right|\right]\right] \leq 2\delta' bC_\alpha. \tag{6.9}$$

Similarly, by Lemma 6.3, we have for any fixed $A' \subset \mathbf{T}_n$ that

$$\mathbf{Var}\left(\frac{M^\beta}{n^\beta}\sum_{v\in A'} p_v^{n,M}(K)\middle|\mathcal{F}_n\right) \leq \frac{M^{2\beta}}{n^{2\beta}}|A'|C_\alpha K M^{-2\beta} \leq \frac{C_\alpha bK}{n^\beta}. \tag{6.10}$$

Again averaging over realisations of A according to $\mathbb{P}_{\mathbf{T}}^A$, and then applying the tower property, we deduce that

$$\mathbf{Var}\left(\mathbb{E}_{\mathbf{T}}^A\left[\frac{M^\beta}{n^\beta}\sum_{v\in A} p_v^{n,M}(K)\right]\right) \leq \frac{C_\alpha bK}{n^\beta}. \tag{6.11}$$

Hence it follows from Chebyshev's inequality that

$$\mathbf{P}\left(\mathbb{E}_{\mathbf{T}}^A\left[\left|\frac{M^\beta}{n^\beta}\sum_{v\in A} p_v^{n,M}(K) - |A|n^{-\beta}C_\alpha\right|\right] \geq 4\delta' bC_\alpha\right) \leq \frac{K}{4(\delta')^2 bC_\alpha n^\beta} = O(n^{-\beta\varepsilon}), \tag{6.12}$$

where we have set $K = n^{\beta(1-\varepsilon)}$ for some $\varepsilon > 0$ small enough that $\alpha(1-\varepsilon)^2 > 1+\varepsilon$ (e.g. $\varepsilon = \frac{\sqrt{\alpha}-1}{\sqrt{\alpha}+1}$ will do – note that we then have that $K > \tilde{K}$ for all large enough n). By Theorem 1.2, we know that for each $v \in \mathbf{T}_n$, $M^\beta p_v^{n,M}$ converges almost surely to $C_\alpha \mathbf{W}_v$ as $M \to \infty$ (keeping n fixed). Since $\mathbf{W}$ is bounded in $\mathrm{L}^{\alpha(1-\varepsilon)}(\mathbf{P})$, we have for all sufficiently large M that there exists $C < \infty$, not depending on n, and an event $F_{M,n}$ that occurs eventually $\mathbf{P}$-almost surely as $M \to \infty$, such that, for all $v \in \mathbf{T}_n$,

$$\mathbf{P}\left(F_{M,n} \text{ and } M^\beta p_v^{n,M} > K\middle|\mathcal{F}_n\right) \leq CK^{-\alpha(1-\varepsilon)} = Cn^{-\beta\alpha(1-\varepsilon)^2} \leq Cn^{\beta(1+\varepsilon)}. \tag{6.13}$$

In particular, for all such M, we can write for any fixed $A' \subset \mathbf{T}_n$ with $|A'| \leq bn^\beta$:

$$\mathbf{P}\left(F_{M,n} \text{ and } \exists v \in A' : M^\beta p_v^{n,M} \geq K\middle|\mathcal{F}_n\right) \leq C|A'|K^{-\alpha(1-\varepsilon)} \leq C|A'|n^{-\beta(1+\varepsilon)} \leq Cbn^{-\beta\varepsilon}. \tag{6.14}$$

Again averaging over A and applying the tower property as before, we deduce that

$$\mathbf{E}\left[\mathbb{1}\{F_{M,n}\}\mathbb{P}_{\mathbf{T}}^A\left(\exists v \in A : M^\beta p_v^{n,M} \geq K\right)\right] \leq Cbn^{-\beta\varepsilon}. \tag{6.15}$$

In particular, (6.12), (6.15), Markov's inequality, a union bound Fatou's lemma and Borel–Cantelli imply that, eventually $\mathbf{P}$-almost surely along the subsequence $(n_k)_{k=1}^\infty$, the following events occur simultaneously infinitely often as $M \to \infty$:

- $\mathbb{E}_{\mathbf{T}}^{A_n}\left[\left|\frac{M^\beta}{n^\beta}\sum_{v\in A_n} p_v^{n,M}(K) - |A|n^{-\beta}C_\alpha\right|\right] < 4\delta' b C_\alpha$.
- $\mathbb{1}\{F_{M,n}\}\mathbb{P}_{\mathbf{T}}^{A_n}\left(\exists v \in A_n : M^\beta p_v^{n,M} \geq K\right) < \delta$.

Recall that $\delta^2 = 4\delta' b C_\alpha$ and fix an n and an M such that the above two points occur. On the event $F_{M,n}$, we have that

$$\begin{aligned}&\mathbb{P}_{\mathbf{T}}^{A}\left(\left|\frac{M^\beta}{n^\beta}\sum_{v\in A} p_v^{n,M} - |A|n^{-\beta}C_\alpha\right| \geq \delta\right)\\ &\leq \mathbb{P}_{\mathbf{T}}^{A}\left(\left|\frac{M^\beta}{n^\beta}\sum_{v\in A} p_v^{n,M}(K) - |A|n^{-\beta}C_\alpha\right| > \delta\right) + \mathbb{P}_{\mathbf{T}}^{A}\left(\exists v \in A : M^\beta p_v^{n,M} \geq K\right) \leq 2\delta.\end{aligned} \tag{6.16}$$

In particular, we know that $\mathbf{P}$-almost surely along the subsequence, we have for all sufficiently large k that there exists a sequence $(M_j^{(k)})_{j\geq 1}$ where the above estimate holds for $n = n_k$ and $M = M_j^{(k)}$. It therefore follows from Fatou's lemma that for all sufficiently large k,

$$\mathbb{P}_{\mathbf{T}}^{A}\left(\left|\frac{M^\beta}{{n_k}^\beta}\sum_{v\in A} p_v^{n_k,M} - |A|{n_k}^{-\beta}C_\alpha\right| < \delta \text{ i.o. as } M\to\infty\right) \geq (1-2\delta)\liminf_{M\to\infty}\mathbb{1}\{F_{M,n_k}\}. \tag{6.17}$$

Since F_{M,n_k} holds eventually $\mathbf{P}$-almost surely as $M \to \infty$ for every k, this implies the claim. □

We now prove the main theorem in two steps as follows.

Lemma 6.5 (Subsequential convergence of generation sizes in the IIC)**.** ***P****-almost surely, the convergence of Theorem 1.5 holds along the subsequence* $(n_k)_{k\geq 1}$.

Proof. Note that our aim is equivalent to calculate for every $a < b$, the a.s. limit

$$\lim_{n\to\infty}\mu_{\mathbf{T}}\left(\mathbf{Z}_n \in (a,b)\right) \tag{6.18}$$

along the subsequence $(n_k)_{k\geq 1}$. By Lemma 6.2, we have for each $k \geq 1$ that

$$\mu_{\mathbf{T}}\left(\mathbf{Z}_{n_k} \in (a,b)\right) = \lim_{M\to\infty}\mathbb{P}_{\mathbf{T}}\left(Y^{(n_k)} \in (a,b) | 0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k+M}\right). \tag{6.19}$$

Since we already know from Lemma 6.2 that the limit as $M \to \infty$ exists $\mathbf{P}$-almost surely, it will suffice to evaluate it along a subsequence.

We first write using Bayes' formula that $\mathbb{P}_{\mathbf{T}}\left(Y^{(n_k)} \in (a,b) | 0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k+M}\right)$ is equal to

$$\frac{\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k+M} | Y^{(n_k)} \in (a,b)\right)\mathbb{P}_{\mathbf{T}}\left(Y^{(n_k)} \in (a,b) | 0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k}\right)\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k}\right)}{\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k+M}\right)}. \tag{6.20}$$

We start with the first term and write n in place of n_k for the time being. Note that, by Theorem 1.2, the expression $\frac{\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M} | Y^{(n)} \in (a,b)\right)\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n}\right)}{\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M}\right)}$ is equal to (the $O(\cdot)$ refers to the limit as $M \to \infty$ with n fixed):

$$\frac{(1+o(1))(n+M)^\beta}{n^\beta}\sum_{\substack{A\subset\mathbf{T}_n:\\ n^{-\beta}|A|\in(a,b)}}\mathbb{P}_{\mathbf{T}}\left(y^{(n)} = A | Y^{(n)} \in (a,b)\right)\left(\sum_{v\in A} p_v^{n,M} + O(M^{-2\beta})\right). \tag{6.21}$$

(The latter statement holds for all sufficiently large n appearing in the subsequence, $\mathbf{P}$-almost surely.) Recalling that in fact $n = n_k$ and applying Lemma 6.4, we deduce that, $\mathbf{P}$-almost surely, the following holds for all sufficiently large n: with $\mathbb{P}_{\mathbf{T}}$-probability at least $1-\delta$, there is a subsequence $(M_j^{(n)})_{j\geq 1}$ along which this is equal to

$$\begin{aligned}&(1+o(1)) \sum_{\substack{A\subset \mathbf{T}_n:\\ n^{-\beta}|A|\in(a,b)}} \mathbb{P}_{\mathbf{T}}\left(y^{(n)} = A | Y^{(n)} \in (a,b)\right) \left(|A|n^{-\beta}C_\alpha + O(\delta) + O((M_j^{(n)})^{-\beta})\right)\\ &\to (1+O(\delta))\mathbb{E}_{\mathbf{T}}\left[Y^{(n)}|Y^{(n)}\in(a,b)\right] C_\alpha.\end{aligned} \tag{6.22}$$

All that remains is therefore to multiply this by $\mathbb{P}_{\mathbf{T}}\left(Y^{(n)} \in (a,b) | 0 \overset{*}{\longleftrightarrow} \mathbf{T}_n\right)$, and then take the limit as $n\to\infty$. In particular, we see that $\mathbf{P}$-almost surely, the limit of (6.20) (for all sufficiently large n in the subsequence) along the subsequence $(M_j^{(n)})_{j\geq 1}$ is equal to

$$(1+O(\delta))C_\alpha \mathbb{E}_{\mathbf{T}}\left[Y^{(n)}|Y^{(n)}\in(a,b)\right] \mathbb{P}_{\mathbf{T}}\left(Y^{(n)} \in (a,b) | 0 \overset{*}{\longleftrightarrow} \mathbf{T}_n\right). \tag{6.23}$$

In particular, since we already know that the limit exists (by Lemma 6.2), we see that the limit of (6.20) (for fixed n) is

$$(1+O(\delta))C_\alpha \mathbb{E}_{\mathbf{T}}\left[Y^{(n)}\mathbb{1}\{Y^{(n)}\in(a,b)\}|Y_n>0\right]. \tag{6.24}$$

So to establish the subsequential convergence, it just remains to take the limit of this latter expression. However, we already know from Theorem 1.3 that $Y^{(n)}$ converges to Y, and hence this converges to the size-biased version Y^* appearing in Theorem 1.5. Since $\delta>0$ was arbitrary, this implies the result. ☐

It remains to lift this from the subsequence $(n_k)_{k\geq 1}$ to the whole sequence $n \geq 1$.

Proof of Theorem 1.5. It is sufficient to evaluate the limit (as $n\to\infty$) of the limit (as $M\to\infty$) of $\frac{M^\beta}{n^\beta}\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M} | Y^{(n)}\in(a,b)\right)$, since we already know that the other terms appearing in (6.20) converge, $\mathbf{P}$-almost surely. To this end, we take E_n as in (4.19), i.e.

$$E_n = \left\{|Y_n - Y_{n_k}| > Y_n^{1-\delta} \vee n^{\beta-\delta}\right\}, \tag{6.25}$$

but this time take $\varepsilon>0$ and

$$F_n' = F_n'(\mathbf{T},\varepsilon) = \left\{\mathbb{P}_{\mathbf{T}}\left(E_n|\mathcal{G}_n^{\mathbf{T}}, Y^{(n)}\in(a,b)\right) < \varepsilon/4\right\}. \tag{6.26}$$

Since the event $\{Y^{(n)}\in(a,b)\}$ is $\mathcal{G}_n^{\mathbf{T}}$-measurable, it follows from Proposition 4.5 that

$$\mathbf{P}\left(F_n'|\mathcal{F}_n\right) \geq 1-o(1) \tag{6.27}$$

as well. Our strategy is therefore to follow that of Lemma 4.4, and first show an analogous statement to (4.23) and Lemma 4.3, that is that on the event F_n' the n^{th} quantity should be close to the n_k^{th} quantity when $n\in(n_{k-1},n_k]$, and then apply the same logic as Lemma 4.4 to lift the result of Lemma 6.5 to the whole sequence.

To this end, first note that on the event F_n' we have by Theorem 1.3 that for all $\varepsilon>0$ that there exists $\delta>0$ such that

$$\begin{aligned}\mathbb{P}_{\mathbf{T}}\left(Y^{(n_k)}\in(a,b)|Y^{(n)}\in(a,b)\right) &\geq \mathbb{P}_{\mathbf{T}}\left(Y^{(n_k)}\in(a,b)|Y^{(n)}\in(a+\delta,b-\delta)\right)(1-\varepsilon/4)\\ &\geq 1-\varepsilon/2.\end{aligned} \tag{6.28}$$

Hence we deduce that on the event F_n', for any $n \in (n_{k-1}, n_k]$,

$$\begin{aligned}\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M} | Y^{(n)} \in (a,b)\right) &\geq \mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k+M} | Y^{(n)} \in (a,b)\right) \\ &\geq \mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k+M} | Y^{(n_k)} \in (a,b)\right)(1-\varepsilon/2).\end{aligned} \tag{6.29}$$

In the other direction, first note that similarly to (6.28), on the event F_n' we have that

$$\mathbb{P}_{\mathbf{T}}\left(Y^{(n_k)} \notin (a,b) | Y^{(n)} \notin (a,b)\right) \geq 1-\varepsilon/2,$$

which rearranges to

$$\mathbb{P}_{\mathbf{T}}\left(Y^{(n)} \in (a,b) | Y^{(n_k)} \in (a,b)\right) \geq 1-\varepsilon/2. \tag{6.30}$$

Hence we deduce that on the event F_n', for any $n \in (n_{k-1}, n_k]$,

$$\begin{aligned}\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M} | Y^{(n)} \in (a,b)\right) &\leq \mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M} | Y^{(n_k)} \in (a,b)\right)(1+\varepsilon/2) \\ &= \mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k+(M+n-n_k)} | Y^{(n_k)} \in (a,b)\right)(1+\varepsilon/2).\end{aligned} \tag{6.31}$$

Taking limits as $M \to \infty$ gives

$$\begin{aligned}&\lim_{M\to\infty} M^{\beta}\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M} | Y^{(n)} \in (a,b)\right) \\ &\qquad \leq \lim_{M\to\infty} M^{\beta}\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k+M} | Y^{(n_k)} \in (a,b)\right)(1+\varepsilon/2).\end{aligned} \tag{6.32}$$

To summarise, and using the fact that $\frac{n}{n_k} \to 1$ as $n \to \infty$, we showed in (6.29) and (6.32) that on the event F_n' (and provided n is sufficiently large),

$$\left|\lim_{M\to\infty}\frac{M^{\beta}}{n^{\beta}}\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M} | Y^{(n)} \in (a,b)\right) - \lim_{M\to\infty}\frac{M^{\beta}}{n_k^{\beta}}\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n_k+M} | Y^{(n_k)} \in (a,b)\right)\right| \leq \varepsilon. \tag{6.33}$$

We can therefore use this latter statement in combination with (6.27) to imitate the proof of Lemma 4.4 and deduce that

$$\lim_{M\to\infty}\left|\frac{M^{\beta}}{n^{\beta}}\mathbb{P}_{\mathbf{T}}\left(0 \overset{*}{\longleftrightarrow} \mathbf{T}_{n+M} | Y^{(n)} \in (a,b)\right) - C_{\alpha}\mathbb{E}_{\mathbf{T}}\left[Y^{(n)}\mathbb{1}\{Y^{(n)} \in (a,b)\} | Y_n > 0\right]\right| \leq \varepsilon, \tag{6.34}$$

for all sufficiently large n, $\mathbf{P}$-almost surely, and then combine with (6.20) exactly as in the previous proof to deduce the result. □

A Appendix

Lemma A.1. *Under Assumption 1.1 (i.e. $\boldsymbol{P}(|\boldsymbol{T}_1| > x) \sim c_1 x^{-\alpha}$ for $\alpha \in (1,2)$ or $\mathbf{Var}(|\boldsymbol{T}_1|) < \infty$), we have that $\mathbb{P}(Y_1 > x) \sim c_1 \mu^{-\alpha} x^{-\alpha}$, as well as*

$$\phi(\theta) = 1 - C_{\alpha}^{-1}\theta(1+(C_{\alpha}^{-1}\theta)^{\alpha-1})^{-\beta} \quad \textit{with} \quad C_{\alpha} = c_1^{-\beta}\mu^{\alpha\beta}\Gamma(1-\alpha)^{-\beta}\beta^{\beta}. \tag{A.1}$$

Moreover, the branching mechanism (see Section 1) satisfies

$$\psi(\lambda) = \begin{cases} \beta C_{\alpha}^{1-\alpha}\lambda^{\alpha} & \textit{if } \alpha < 2, \\ C_{\alpha}^{-1}\lambda^2 & \textit{if the variance is finite.}\end{cases} \tag{A.2}$$

Proof. According to [18, Lemma 2], if the generating function $f(s) = \mathbb{E}\left[s^{Y_1}\right]$ of the critical annealed process satisfies

$$f(s) = s + (1-s)^{\alpha} L(1-s) \quad \text{with} \quad \lim_{x \to 0^+} L(x) = c_o\,, \tag{A.3}$$

then

$$\mathbb{P}\left(Y_n > 0\right)^{\alpha-1} c_o \sim \left(n(\alpha-1)\right)^{-1} \quad \Longleftrightarrow \quad \mathbb{P}\left(Y_n > 0\right) \sim \left(c_o n(\alpha-1)\right)^{-\beta}\,. \tag{A.4}$$

Write $C_\alpha = C(\alpha, c_o) = c_o^{-\beta}\beta^{\beta}$ so that $\mathbb{P}\left(Y_n > 0\right) \sim C_\alpha n^{-\beta}$. It is then shown that (see [18, Theorem 1])

$$\mathbb{E}\left[\mathrm{e}^{-un^{-\beta}Y_n} | Y_n > 0\right] \sim 1 - (u/C_\alpha)\left(1 + (u/C_\alpha)^{\alpha-1}\right)^{-\beta}\,. \tag{A.5}$$

By Equation (A.3), we get that, as $u \to 0$,

$$\mathbb{E}\left[\mathrm{e}^{-uY_1}\right] = f\left(\mathrm{e}^{-u}\right) = \mathrm{e}^{-u} + \left(1 - \mathrm{e}^{-u}\right)^{\alpha} L\left(1 - \mathrm{e}^{-u}\right) = 1 - u + c_o u^{\alpha} + o(u^{\alpha})\,. \tag{A.6}$$

This gives by [6, Theorem 8.1.6], that

$$\mathbb{P}\left(Y_1 > x\right) \sim \frac{c_o}{\Gamma(1-\alpha)} x^{-\alpha}\,. \tag{A.7}$$

Now, suppose that we have in the quenched case that the mean $\mu > 1$ and that

$$\mathbf{P}\left(|\mathbf{T}_1| > x\right) \sim c_1 x^{-\alpha}\,. \tag{A.8}$$

By standard concentration estimates for binomial random variables, it follows that $\mathbb{P}\left(Y_1 > x\right) \sim c_1 \mu^{-\alpha} x^{-\alpha}$. This implies that, in the notation above,

$$c_o = c_1 \mu^{-\alpha}\Gamma(1-\alpha) \quad \text{and hence} \quad C_\alpha = c_1^{-\beta}\mu^{\alpha\beta}\Gamma(1-\alpha)^{-\beta}\beta^{\beta}\,. \tag{A.9}$$

Next, we compute the branching mechanism ψ. Note that by (5.2), if

$$\mathbb{E}\left[e^{-\theta\tilde{Y}_t} | \tilde{Y}_s = a\right] = \mathrm{e}^{-au_{t-s}(\theta)} \qquad \text{then} \qquad u_t(\lambda) = \lambda\left(1 + \left(C_\alpha^{-1}\lambda\right)^{\alpha-1} t\right)^{-\beta}$$
$$\text{and} \qquad u_t'(\lambda) = -\beta C_\alpha^{1-\alpha}\lambda^{\alpha}\left(1 + \left(C_\alpha^{-1}\lambda\right)^{\alpha-1} t\right)^{-\beta-1}\,. \tag{A.10}$$

By [11, Section 1.3], we have that $\psi(\lambda) = \tilde{c}\lambda^{\alpha}$ (in the case $\alpha < 2$) and $\psi(\lambda) = \tilde{c}\lambda^2$ in the finite variance case. Furthermore, we have that $u_t(\lambda)$ satisfies the following integral equation:

$$u_t(\lambda) + \int_0^t \psi\left(u_s(\lambda)\right) \mathrm{d}s = \lambda\,. \tag{A.11}$$

Then, if $\alpha < 2$

$$u_t(\lambda) + \tilde{c}\int_0^t \left(u_s(\lambda)\right)^{\alpha} \mathrm{d}s = \lambda \quad \Longrightarrow \quad u_0'(\lambda) + \tilde{c}u_0(\lambda)^{\alpha} = 0\,, \tag{A.12}$$

From this, (A.10) and the fact that $\frac{\beta+1}{\beta} = \alpha$ we infer that $\tilde{c} = \beta C_\alpha^{1-\alpha}$. The finite variance case is analogous. ☐

Lemma A.2. *Let $\mathbb{P}$ be the law of the critical annealed tree with stable law $\alpha \in (1, 2]$. Let Y_l be the number of individuals at generation l conditioned to survive and let $(Y_l^{(i)})_{i\geq 1}$ be i.i.d. copies of Y_l. Then for all $\varepsilon > 0$, there exists $C > 0$ such that for all N, l, x with $x \geq N^{1/(\alpha-\varepsilon)}$*

$$\mathbb{P}^{\otimes N}\left(\sum_{i=1}^{N}(p_l Y_l^{(i)} - 1) > x\right) < CNx^{-(\alpha-\varepsilon)}\,, \tag{A.13}$$

for $p_l := \mathbb{P}\left(Y_l > 0\right) \sim C_\alpha l^{-\beta}$.

Proof. Note that the expectation of $p_l Y_l^{(i)}$ is equal to 1, and hence the summands are centred. We begin by developing a bound for $p_l Y_l^{(i)}$ away from its mean, i.e.,

$$\mathbb{P}\left(l^{-\beta} Y_l > z\right) , \tag{A.14}$$

for z large. Let f_n be the probability generating function of the critical process after n generations, i.e. $f_n(s) = \mathbb{E}\left[s^{Y_n}\right]$. Note that, by (1.2) (see also [18, Lemma 2]) there exists $e_n = o(1)$ such that

$$1 - f_n(0) = C_\alpha n^{-\beta}\left(1 + e_n\right) . \tag{A.15}$$

Abbreviate $c = C_\alpha^\beta$ from now on. We now assume that $0 < u < 1$. Define

$$y_n(u) = y_n = \mathrm{e}^{-u(1-f_n(0))} . \tag{A.16}$$

Hence, there exist sequences $(h_n)_n = o(1)$ (independent of u, as u remains bounded from above) and $(k_n)_n \in \mathbb{N}$ such that

$$k_n = k_n(u) = u^{-1/\beta} n\left(1 + h_n\right) \qquad \text{and} \qquad f_{k_n}(0) \le y_n < f_{k_n+1}(0) , \tag{A.17}$$

By [18, Eq. (2.6)] that there exists $(g_n)_n = o(1)$ such that

$$\frac{1 - f_{r+1}(0)}{1 - f_r(0)} = 1 - \beta r^{-1}\left(1 + g_r\right) , \tag{A.18}$$

for all $r \ge 1$, where we have used the relation $C_\alpha^{\alpha-1} c_o = \beta$ from the previous lemma. Hence

$$\frac{1 - f_{n+M}(0)}{1 - f_n(0)} = \exp\left(\sum_{l=0}^{M-1} \log\left(1 - \beta(n+l)^{-1}\left(1 + g_{n+l}\right)\right)\right) . \tag{A.19}$$

Therefore, there exists a $(j_n)_n = o(1)$ such that for all $n, M \ge 1$:

$$\frac{1 - f_{n+M}(0)}{1 - f_n(0)} = \mathrm{e}^{-\beta(\log(M+n-1) - \log(n)))(1+j_n)} . \tag{A.20}$$

Now let $\phi_n(t)$ be the Laplace transform of (the rescaled) size of generation n conditioned on survival, so that by definition,

$$\phi_n(u) = \mathbb{E}\left[\mathrm{e}^{-u(1-f_n(0))Y_n} | Y_n > 0\right] = 1 - \frac{1 - f_n(y_n(u))}{1 - f_n(0)} . \tag{A.21}$$

By definition of k_n, we moreover have that

$$\frac{1 - f_{n+k_n(u)}(0)}{1 - f_n(0)} \ge \frac{1 - f_n(y_n)}{1 - f_n(0)} \ge \frac{1 - f_{n+k_n(u)+1}(0)}{1 - f_n(0)} , \tag{A.22}$$

Hence, by (A.20), we have for all $0 < u < 1$

$$\begin{aligned} \phi_n(u) = 1 - \left(\frac{n}{n + k_n(u)(1+o(1))}\right)^{\beta(1+o(1))} &= 1 - \left(\frac{1}{1 + u^{-1/\beta}(1+o(1))}\right)^{\beta(1+o(1))} \\ &= 1 - u\left(1 + u^{\alpha-1}\right)^{-\beta(1+o(1))} , \end{aligned} \tag{A.23}$$

where the $o(1)$ are independent of u. Note that if X is a positive random variable with mean 1, then

$$\{X \ge 2r\} \subset \left\{\mathrm{e}^{-X/r} - 1 + X/r \ge 1\right\} , \tag{A.24}$$

and moreover the latter random variable is non-negative. Hence, using Markov's inequality, we obtain that

$$\mathbb{P}(X \geq 2r) \leq \mathbb{E}\left[\mathrm{e}^{-X/r}\right] - 1 + \mathbb{E}\left[X/r\right] = \phi_X(1/r) - 1 + 1/r\,. \tag{A.25}$$

Hence, plugging Equation (A.25) into (A.23), we get that there exists a $C > 0$ such that for all $z > 1$, as $l \to \infty$

$$\mathbb{P}\left(p_l Y_l > z\right) \leq C z^{-\alpha(1+o(1))}\,. \tag{A.26}$$

Using [3, Theorem 5.1 (ii)] (with $y = x$ in Berger's notation) Equation (A.26) implies that for every $\tilde{\alpha} < \alpha$ that there exists an $l_o > 0$ and $C > 0$ such that for all $l > l_o$ and $n \geq 1$ and for all $x \geq n^{1/\tilde{\alpha}}$, we have that

$$\mathbb{P}\left(\sum_{i=1}^{n}\left(p_l Y_l^{(i)} - 1\right) > x\right) \leq C n x^{-\tilde{\alpha}}\,. \tag{A.27}$$

However, note that (by adjusting the constant C) Equation (A.27) continues to hold for $l \in \{1, \ldots, l_o\}$. Indeed, as l_o is fixed, the probability generating function of Y_k is given $f^{\circ k}(x) = f(f(\ldots f(x)\ldots))$, where we apply f exactly k times. Since this is a finite convolution, we get the same tail bounds as in the case $l = 1$. This concludes the proof. □

References

[1] I. Ayuso Ventura and Q. Berger. Non-linear conductances of galton-watson trees and application to the (near) critical random cluster model. *arXiv:2404.11564*, 2024.

[2] R. Balan and B. Saidani. Weak convergence and tightness of probability measures in an abstract Skorohod space. *arXiv:1907.10522*, 2019. MR4113449

[3] Q. Berger. Notes on random walks in the Cauchy domain of attraction. *Probability Theory and Related Fields*, 175(1):1–44, 2019. MR4009704

[4] P. Billingsley. *Convergence of Probability Measures, second edition*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1999. MR1700749

[5] N. Bingham and R. Doney. Asymptotic properties of supercritical branching processes I: The Galton–Watson process. *Advances in Applied Probability*, 6(4):711–731, 1974. MR0362525

[6] N. Bingham, C. Goldie, and J. Teugels. *Regular Variation*. Cambridge University Press, 1989. MR1015093

[7] T. Duquesne. A limit theorem for the contour process of conditioned Galton–Watson trees *The Annals of Probability*, 31(2):996–1027, 2003. MR1964956

[8] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Annales de l'IHP Probabilités et Statistiques*, 22:425–487, 1986. MR0871905

[9] H. Kesten, P. Ney, and F. Spitzer. The Galton-Watson process with mean one and finite variance. *Theory of Probability & Its Applications*, 11(4):513–540, 1966. MR0207052

[10] A. Kolmogorov. On the solution of a problem in Biology. *Izv. NII Matem. Mekh. Tomskogo Univ*, 2:7–12, 1938.

[11] J. Le Gall. *Spatial Branching Processes, Random Snakes and Partial Differential Equations*. Springer Science & Business Media, 1999. MR1714707

[12] Z. Li. Continuous-state branching processes. *arXiv:1202.3223*, 2012.

[13] R. Lyons. Random walks and percolation on trees. *The Annals of Probability*, 18(3):931–958, 1990. MR1062053

[14] R. Lyons and Y. Peres. *Probability on Trees and Networks*, volume 42. Cambridge University Press, 2017. MR3616205

[15] M. Michelen. Critical percolation and the incipient infinite cluster on Galton–Watson trees. *Electronic Communications in Probability*, 24:1–13, 2019. MR3916342

[16] M. Michelen, R. Pemantle, and J. Rosenberg. Quenched survival of Bernoulli percolation on Galton–Watson trees. *Journal of Statistical Physics*, 181(4):1323–1364, 2020. MR4163504

[17] A. Skorokhod. Limit theorems for stochastic processes. *Theory of Probability & Its Applications*, 1(3):261–290, 1956. MR0084897

[18] R. Slack. A branching process with mean one and possibly infinite variance. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 9(2):139–145, 1968. MR0228077

[19] A. Yaglom. Certain limit theorems of the theory of branching random processes. *Doklady Akad. Nauk SSSR (NS)*, 56:3, 1947. MR0022045

Acknowledgments. The authors would like to thank the two anonymous referees for their comments which helped to improve that paper. The authors would like to express their gratitude to the program *Global Challenges for Women in Math Science* at the Technical University Munich which enabled a research visit of EA to Munich during which some of this work was written, and to thank Silke Rolles for her comments on an earlier draft of the paper. EA would also like to thank Matan Shalev and Pengfei Tang for helpful conversations on this topic. EA was supported by the ANR project ProGraM, reference ANR-19-CE40-0025.