Journal of Theoretical Probability
https://doi.org/10.1007/s10959-019-00975-0

Self-similar Growth Fragmentations as Scaling Limits of Markov Branching Processes

Benjamin Dadoun[1]

Received: 22 November 2017 / Revised: 26 June 2019

Abstract
We provide explicit conditions, in terms of the transition kernel of its driving particle, for a Markov branching process to admit a scaling limit toward a self-similar growth fragmentation with negative index. We also derive a scaling limit for the genealogical embedding considered as a compact real tree.

Keywords Growth fragmentation · Scaling limit · Markov branching tree

Mathematics Subject Classification (2010) 60F17 · 60J80

1 Introduction

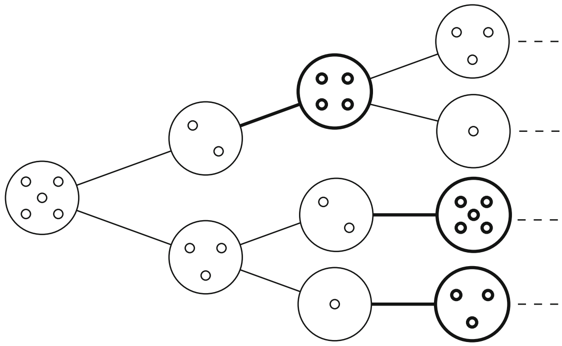
Imagine a bin containing n balls which is repeatedly subject to random (binary) divisions at discrete times, until every ball has been isolated. There is a natural random (binary) tree with n leaves associated with this partitioning process, where the subtrees above a given height $k \geq 0$ represent the different subcollections of all n balls at time k, and the number of leaves of each subtree matches the number of balls in the corresponding subcollection. The habitual *Markov branching property* stipulates that these subtrees must be independent conditionally on their respective size. In the literature on random trees, a central question is the approximation of so-called continuum random trees (CRT) as the size of the discrete trees tends to infinity. We mention especially the works of Aldous [1–3] and Haas, Miermont, et al. [14–17,23]. Concerning the above example, Haas and Miermont [16] obtained, under some natural assumption on the splitting laws, distributional scaling limits regarded in the Gromov–Hausdorff–Prokhorov topology. In the Gromov–Hausdorff sense where trees are considered as compact metric spaces, they especially identified the so-called *self-similar fragmentation trees* as the scaling limits. The latter describe the genealogy of *self-similar*

✉ Benjamin Dadoun
benjamin.dadoun@math.uzh.ch

1 Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

Published online: 17 December 2019

Fig. 1 An example of dynamics with growth transitions (in bold)



fragmentation processes, which, reciprocally, are known to record the size of the components of a (continuous) fragmentation tree above a given height [14], and thus correspond to scaling limits for partition sequences of balls as their number n tends to infinity. One key tool in the work of Haas and Miermont [16] is provided by some non-increasing integer-valued Markov chain which, roughly speaking, depicts the size of the subcollection containing a randomly tagged ball. This Markov chain essentially captures the dynamics of the whole fragmentation and, by their previous work [15], itself possesses a scaling limit.

The purpose of the present work is to study more general dynamics which incorporate *growth*, that is the addition of new balls in the system (see Fig. 1). One example of recent interest lies in the exploration of random planar maps [7,8], which exhibits "holes" (the yet unexplored areas) that split or grow depending on whether the new edges being discovered belong to an already known face or not. We thus consider a Markov branching system in discrete time and space where at each step every particle is replaced by either one particle with a bigger size (growth) or by two smaller particles in a conservative way (fragmentation). We condition the system to start from a single particle with size n (we use the superscript $\cdot^{(n)}$ in this respect) and we are interested in its behavior as $n \to \infty$. Namely, we are looking for:

1. A functional scaling limit for the process in time $(\mathbb{X}(k) \colon k \geq 0)$ of all particle sizes:
$$\left(\frac{\mathbb{X}^{(n)}(\lfloor a_n t \rfloor)}{n} : t \geq 0 \right) \xrightarrow[n\to\infty]{(d)} \big(\mathbf{Y}(t) \colon t \geq 0 \big),$$
in some sequence space, where the a_n are positive (deterministic) numbers;
2. A scaling limit for the system's genealogical tree, seen as a random metric space $(\chi^{(n)}, d_n)$:
$$\left(\chi^{(n)}, \frac{d_n}{a_n} \right) \xrightarrow[n\to\infty]{(d)} \mathcal{Y},$$
in the Gromov–Hausdorff topology.

Like in the pure fragmentation setting, we may single out some specific integer-valued Markov chain, but which of course is no longer non-increasing. To derive a scaling limit for this chain, a first idea is to apply, as a substitution to [15], the more general criterion of Bertoin and Kortchemski [9] in terms of the asymptotic behavior of its transition kernel at large states. However, this criterion is clearly insufficient for the convergences stated above as it provides no control on the "microscopic" particles. To circumvent this issue, we choose to "prune" the system by freezing the particles below a (large but fixed) threshold. That is to say, we let the system evolve from a large size n but stop every individual as soon as it is no longer bigger than some threshold $M > 0$ which will be independent of n, and we rather study the modifications $\mathbf{X}^{(n)}$ and $\mathcal{X}^{(n)}$ of the process and the genealogical tree that are induced by this procedure.

The limits $\mathbf{Y}$ and $\mathcal{Y}$ that we aim at are, respectively, a *self-similar growth fragmentation process* and its associated genealogical structure. Indeed, the scaling limits of integer-valued Markov chains investigated in [9], which we build our work upon, belong to the class of so-called *positive self-similar Markov processes* (pssMp), and these processes constitute the cornerstone of Bertoin's self-similar growth fragmentations [6,7]. Besides, in the context of random planar maps [7,8], they have already been identified as scaling limits for the sequences of perimeters of the separating cycles that arise in the exploration of large Boltzmann triangulations. Informally, a self-similar growth fragmentation $\mathbf{Y}$ depicts a system of particles which all evolve according to a given pssMp and whose each negative jump $-y < 0$ begets a new independent particle with initial size y. In our setting, the self-similarity property has a negative index and makes the small particles split at higher rates, in such a way that the system becomes eventually extinct [6, Corollary 3]. The genealogical embedding $\mathcal{Y}$ is thus a compact real tree; its formal construction is presented in [24].

Because of growth, one main difference with the conservative case is, of course, that the mass of a particle at a given time no longer equals the size (number of leaves) of the corresponding genealogical subtree. In a similar vein, choosing the uniform distribution to mark a ball at random will appear less relevant than a size-biased pick from an appropriate (non-degenerate) supermartingale. This will highlight a Markov chain admitting a self-similar scaling limit (thanks to the criterion [9]), and which we can plug into a many-to-one formula. Under an assumption preventing an explosive production of relatively small particles, we will then be able to establish our first desired convergence. Concerning the convergence of the (rescaled) trees $\mathcal{X}^{(n)}$, we shall employ a Foster–Lyapunov argument to obtain an uniform control on their heights, which are nothing else than the extinction times of the processes $\mathbf{X}^{(n)}$. Contrary to what one would first expect, it turns out that a good enough Lyapunov function is not simply a power of the size, but merely depends on the scaling sequence (a_n). This brings a tightness property that, together with the convergence of "finite-dimensional marginals", will allow us to conclude.

In the next section, we set up the notation and the assumptions more precisely and state our main two results.

2 Assumptions and Results

Our basic data are probability transitions $p_{n,m}$, $m \geq n/2$ and $n \in \mathbb{N}$ "sufficiently large", with which we associate a Markov chain, generically denoted X, that governs the law of the particle system $\mathbb{X}$: at each time $k \in \mathbb{N}$ and with probability $p_{n,m}$, every particle with size n either grows up to a size $m > n$, or fragmentates into two independent particles with sizes $m \in \{\lceil n/2 \rceil, \ldots, n-1\}$ and $n-m$. That is to say, $X^{(n)}(0) = n$ is the size of the initial particle in $\mathbb{X}^{(n)}$, and given $X(k)$ for some $k \geq 0$, $X(k+1)$ is the largest among the (one or two) particles replacing $X(k)$. We must emphasize that the transitions $p_{n,m}$ from n "small" are irrelevant since our assumptions shall only rest upon the asymptotic behavior of $p_{n,m}$ as n tends to infinity. Indeed, for the reason alluded in Introduction that we explain further below, we rather study the pruned version $\mathbf{X}$ where particles are frozen (possibly at birth) when they become not bigger than a threshold parameter $M > 0$, which we will fix later on. Keeping the same notation, this means that X is a Markov chain stopped upon hitting $\{1, 2, \ldots, M\}$. For convenience, *we omit to write the dependency in* M, and set $p_{n,n} := 1$ for $n \leq M$.

In turn, the law of the genealogical tree $\mathcal{X}$ can be defined inductively as follows. (We give a more rigorous construction in Sect. 3.) Let $1 \leq k_1 \leq \cdots \leq k_p$ enumerate the instants during the lifetime $\zeta^{(n)}$ of $X^{(n)}$ when $n_i := X^{(n)}(k_i - 1) - X^{(n)}(k_i) > 0$. Then $\mathcal{X}^{(n)}$ consists in a branch with length $\zeta^{(n)}$ to which are, respectively, attached, at positions k_i from the root, independent trees distributed like $\mathcal{X}^{(n_i)}$ (agreeing that $\mathcal{X}^{(n)}$ degenerates into a single vertex for $n \leq M$). We view $\mathcal{X}^{(n)}$ as a metric space with metric denoted by d_n.

Suppose $(a_n)_{n\in\mathbb{N}}$ is a sequence of positive real numbers which is regularly varying with index $\gamma > 0$, in the sense that for every $x > 0$,

$$\lim_{n\to\infty} \frac{a_{\lfloor nx \rfloor}}{a_n} = x^{\gamma}. \tag{1}$$

Our starting requirement will be the convergence in distribution

$$\left(\frac{X^{(n)}(\lfloor a_n t \rfloor)}{n} : t \geq 0 \right) \xrightarrow[n\to\infty]{(d)} \big(Y(t) : t \geq 0 \big), \tag{2}$$

in the space $\mathbb{D}([0,\infty), \mathbb{R})$ of càdlàg functions on $[0, \infty)$ (endowed with Skorokhod's J_1 topology), toward a positive strong Markov process $(Y(t) : t \geq 0)$, continuously absorbed at 0 in an almost surely finite time ζ, and with the following self-similarity property:

$$\text{The law of } Y \text{ started from } x > 0 \text{ is that of } (xY(x^{-\gamma} t) : t \geq 0) \text{ when } Y \text{ starts from } 1. \tag{3}$$

Since the seminal work of Lamperti [20], this simply means that

$$\log Y(t) = \xi \left(\int_0^t Y(s)^{-\gamma} \, \mathrm{d}s \right), \qquad t \geq 0, \tag{4}$$

with ξ a Lévy process which drifts to $-\infty$ as $t \to \infty$. We denote by Ψ the characteristic exponent of ξ (so there is the Lévy–Khintchine formula $E[\exp(q\xi(t))] = \exp(t\Psi(q))$ for every $t \geq 0$ and every $q \in \mathbb{C}$, wherever this makes sense) and by Λ the Lévy measure of its jumps (that is a measure on $\mathbb{R} \setminus \{0\}$ with $\int (1 \wedge y^2)\, \Lambda(\mathrm{d}y) < \infty$).

In order to state precisely our assumptions, we need to introduce some more notation. First, we define the exponent

$$\kappa(q) := \Psi(q) + \int_{(-\infty,0)} \left(1 - e^y\right)^q \Lambda(\mathrm{d}y),$$

whose meaning will be discussed shortly (in the paragraph "Discussion on the assumptions"). Next, we also define, for every $n \in \mathbb{N}$, the discrete versions

$$\Psi_n(q) := a_n \sum_{m=1}^{\infty} p_{n,m} \left[\left(\frac{m}{n}\right)^q - 1\right], \quad \text{and} \quad \kappa_n(q) := \Psi_n(q) + a_n \sum_{m=1}^{n-1} p_{n,m} \left(1 - \frac{m}{n}\right)^q.$$

Finally, we fix some parameter $q^* > 0$. After [9, Theorem 2], convergence (2) holds under the following two assumptions:

(**H1**) For every $t \in \mathbb{R}$,

$$\lim_{n\to\infty} \Psi_n(\mathrm{i}t) = \Psi(\mathrm{i}t).$$

(**H2**) We have

$$\limsup_{n\to\infty} a_n \sum_{m=2n}^{\infty} p_{n,m} \left(\frac{m}{n}\right)^{q^*} < \infty.$$

Indeed, by [19, Theorems 15.14 and 15.17], Assumption (H1) is essentially equivalent to (A1) and (A2) of [9], while (H2) rephrases Assumption (A3) there. We now introduce the new assumption:

(**H3**) We have either $\kappa(q^*) < 0$, or $\kappa(q^*) = 0$ and $\kappa'(q^*) > 0$. Moreover, for some $\varepsilon > 0$,

$$\lim_{n\to\infty} a_n \sum_{m=1}^{n-1} p_{n,m} \left(1 - \frac{m}{n}\right)^{q^*-\varepsilon} = \int_{(-\infty,0)} \left(1 - e^y\right)^{q^*-\varepsilon} \Lambda(\mathrm{d}y). \tag{5}$$

Postponing the description of the limits, we can already state our two convergence results formally:

Theorem 1 *Suppose* (H1)–(H3). *Then we can fix a freezing threshold M sufficiently large so that, for every $q \geq 1 \vee q^*$, the convergence in distribution*

$$\left(\frac{\mathbf{X}^{(n)}(\lfloor a_n t \rfloor)}{n} : t \geq 0\right) \xrightarrow[n\to\infty]{(d)} \left(\mathbf{Y}(t) : t \geq 0\right),$$

holds in the space $\mathbb{D}([0,\infty),\ell^{q\downarrow})$*, where* $\mathbf{Y}$ *is the self-similar growth fragmentation driven by* Y*, and*

$$\ell^{q\downarrow} := \left\{\mathbf{x} := (x_1 \geq x_2 \geq \cdots \geq 0) \colon \sum_{i=1}^{\infty} (x_i)^q < \infty\right\}$$

(that is, the family of particles at a given time is always ranked in the non-increasing order).

Theorem 2 *Suppose* (H1)–(H3)*, and* $q^* > \gamma$*. Then we can fix a freezing threshold* M *sufficiently large so that there is the convergence in distribution*

$$\left(\mathcal{X}^{(n)}, \frac{d_n}{a_n}\right) \xrightarrow[n\to\infty]{(d)} \mathcal{Y},$$

in the Gromov–Hausdorff topology, where $\mathcal{Y}$ *is the random compact real tree that represents the genealogy of* $\mathbf{Y}$*.*

Description of the limits As explained in Introduction, the process Y portrays the size of particles in the self-similar growth fragmentation process $\mathbf{Y} := (\mathbf{Y}(t) \colon t \geq 0)$, whose construction we briefly recall (referring to [6,7] for more details): The Eve particle $Y_\varnothing$ is distributed like Y. We rank the negative jumps of a particle Y_u in the decreasing order of their absolute sizes (and chronologically in case of *ex aequo*). When this particle makes its jth negative jump, say with size $-y_j < 0$, then a daughter particle Y_{uj} is born at this time and evolves, independently of its siblings, according to the law of Y started from y_j. (Recall that Y is eventually absorbed at 0, so we can indeed rank the negative jumps in this way; for definiteness, we set $b_{uj} := \infty$ and $Y_{uj} :\equiv 0$ if Y_u makes less than j negative jumps during its lifetime.) Particles are here labeled on the Ulam–Harris tree $\mathbb{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$, the set of finite words over $\mathbb{N}$, where $\mathbb{N}^0 = \{\varnothing\}$ is reduced to the root of the tree, and a vertex $u := (u_1, u_2, \ldots, u_k) \in \mathbb{U}$, at generation $|u| := k$, has $uj := (u_1, u_2, \ldots, u_k, j)$ as jth descendent. Write b_u for the birth time of Y_u. Then

$$\mathbf{Y}(t) = \Big(Y_u(t - b_u) \colon u \in \mathbb{U},\ b_u \leq t\Big), \qquad t \geq 0.$$

After [6,7], this process is self-similar with index $-\gamma$. Roughly speaking, this means that a particle with size $x > 0$ evolves $x^{-\gamma}$ times "faster" than a particle with size 1. Since here $-\gamma < 0$, there is the snowball effect that particles get rapidly absorbed toward 0 as time passes, and it has been shown [6, Corollary 3] that such a growth fragmentation becomes eventually extinct, namely that $\epsilon := \inf\{t \geq 0 \colon \mathbf{Y}(t) = \emptyset\}$ is almost surely finite.

The extinction time ϵ is also the height of the genealogical structure $\mathcal{Y}$ seen as a compact real tree. Referring to [24] for details, we shall just sketch the construction. Let $\mathcal{Y}_{u,0}$ consists in a segment with length $\zeta_u := \inf\{t \geq 0 \colon Y_u(t) = 0\}$ rooted at a vertex u. Recursively, define $\mathcal{Y}_{u,h+1}$ by attaching to the segment $\mathcal{Y}_{u,0}$ the trees $\mathcal{Y}_{uj,h}$ at respective distances $b_{uj} - b_u$, for each born particle uj, $j \leq h+1$. The limiting tree

$\mathcal{Y} := \lim \uparrow_{h\to\infty} \mathcal{Y}_h$, where $\mathcal{Y}_h := \mathcal{Y}_{\varnothing,h}$ fulfills a so-called *recursive distributional equation*. Namely, by [24, Corollary 4.2], given the sequence of negative jump times and sizes (b_j, y_j) of Y and an independent sequence $\mathcal{Y}^1, \mathcal{Y}^2, \ldots$ of copies of $\mathcal{Y}$, the action of grafting, on a branch with length $\zeta := \inf\{t \geq 0 \colon Y(t) = 0\}$ and at distances b_j from the root, the trees $\mathcal{Y}^j$ rescaled by the multiplicative factor y_j^γ, yields a tree with the same law as $\mathcal{Y}$. With this connection, Rembart and Winkel [24, Corollary 4.4] proved that ϵ admits moments up to the order $\sup\{q > 0 \colon \kappa(q) < 0\}/\gamma$. When particles do not undergo sudden positive growth (i.e., $\Lambda((0,\infty)) = 0$), Bertoin et al. [7, Corollary 4.5] more precisely exhibited a polynomial tail behavior of this order for the law of ϵ.

Discussion on the assumptions Observe that (H1) entails (H2) when the Lévy measure Λ of ξ is bounded from above (in particular, when ξ has no positive jumps). By analyticity, Assumptions (H1) and (H2) imply that $\Psi_n(z) \to \Psi(z)$ as $n \to \infty$, for $0 \leq \Re z \leq q^*$. Adding the condition (5) in (H3) then yields the convergence $\kappa_n(z) \to \kappa(z)$ for $\Re z$ in a left-neighborhood of q^*. Lastly, the first condition in (H3) itself implies $\Psi(q^*) < 0$ (since $\Psi < \kappa$) and, together with the other assumptions, that there must exist $q_* \in (0, q^*)$ and some threshold M such that

$$\kappa_n(q) < 0 \text{ and } \kappa(q) < 0, \qquad \text{for all } q \in [q_*, q^*) \ \text{ and } \ n > M,$$

which is all but a superfluous requirement. Indeed, the condition $\kappa(q) \leq 0$ for some $q > 0$ is necessary (and sufficient) [10] to prevent local explosion of the growth fragmentation $\mathbf{Y}$ (a phenomenon which would not allow us to view it in some ℓ^q-space). Informally, the *cumulant function* $\kappa(q)$ captures the expected value of the sum of the particle sizes raised to the power q immediately after the first birth event. This function constitutes a key feature of branching processes and, in particular, of self-similar growth fragmentations [25]. Of course, the meanings of the quantity $\kappa_n(q)$ and of the condition $\kappa_n(q) < 0$ should be regarded the same but at the discrete level (that is, w.r.t. $\mathbf{X}^{(n)}$).

We stress that our assumptions do not provide any control on the "small particles" ($n \leq M$). This explains why we need to "freeze" them (meaning that they no longer grow or beget children); otherwise, their number could become quickly very high and make the system explode, as we illustrate in the example below. We will basically choose M as above, so that $\kappa_n(q) \leq 0$ for some q and all n, once we take the freezing into account (which is tantamount to resetting[1] $\kappa_n :\equiv 0$ for $n \leq M$).

Example 2.1 Suppose that a particle with size n increases to size $n+1$ with probability $p < 1/2$ and, at least when n is small, splits into two particles with sizes 1 and $n-1$ with probability $1-p$. Thus, at small sizes, the unstopped Markov chain essentially behaves like a simple random walk. On the one hand, we know from Cramér's theorem (see, e.g., [12, Theorem 2.2.3]) that for every $\varepsilon > 0$ sufficiently small,

$$\mathbb{P}\Big(X^{(1)}(k) > (1-2p+\varepsilon)k\Big), \qquad k \geq 0,$$

[1] We make here a slight abuse on the notation. Again, even though the dependency is not explicitly written, the discrete objects such as $\kappa_n, \mathbf{X}^{(n)}, \ldots$ all ultimately depend on the freezing threshold M.

decreases exponentially at a rate $c_p(\varepsilon) > 0$. On the other hand, keeping only track of particles with size 1 or 2, the number of particles with size 1 is bounded from below by $Z^{[1]}$, where $\mathbf{Z} := (Z^{[1]}, Z^{[2]})$ is a 2-type Galton–Watson process whose mean-matrix

$$\begin{pmatrix} 0 & 1 \\ 2(1-p) & 0 \end{pmatrix}$$

has spectral radius $r_p := \sqrt{2(1-p)} > 1$, so that by the Kesten–Stigum theorem [5, Theorem V.6.1] the number of particles with size 1 at time $k \to \infty$ is of order at least r_p^k, almost surely. Consequently, the expected number of particles which are above $(1-2p+\varepsilon)k$ at time $2k$ is of exponential order at least $m_p(\varepsilon) := \log r_p - c_p(\varepsilon)$. It is easily checked that this quantity may be positive (e.g., $m_{1/4}(1/4) \geq 0.16$). Thus, without any "local" assumption on the reproduction law at small sizes, the number of small particles may grow exponentially and we cannot in general expect $\mathbb{X}^{(n)}(\lfloor a_n \cdot \rfloor)/n$ to be tight in $\ell^{q\downarrow}$, for some $q > 0$. However, this happens to be the case for the perimeters of the cycles in the branching peeling process of random Boltzmann triangulations [8], where versions of Theorems 1 and 2 hold for $\gamma = 1/2$, $q^* = 3$, and $M = 0$, although $\kappa_n(3) \leq 0$ seems fulfilled only for $M \geq 3$ (which should mean that the holes with perimeter 1 or 2 do not contribute to a substantial part of the triangulation).

We start with the relatively easy convergence of finite-dimensional marginals (Sect. 3). Then, we develop a few key results (Sect. 4) that will be helpful to complete the proofs of Theorem 1 (Sect. 5) and Theorem 2 (Sect. 6).

3 Convergence of Finite-Dimensional Marginals

In this section, we prove finite-dimensional convergences for both the particle process $\mathbf{X}$ and its genealogical structure $\mathcal{X}$. (We mention that the freezing procedure is of no relevance here as it will be only useful in the next section to establish tightness results; in particular the freezing threshold M will be fixed later.)

We start by adopting a representation of the particle system $\mathbf{X}$ that better matches that of $\mathbf{Y}$ given above. We define, for every word $u := u_1 \dots u_i \in \mathbb{N}^i$, the *u-locally largest particle*[2] $(X_u(k) \colon k \geq 0)$ by induction on $i = 0, 1, \dots$ Initially, for $i = 0$, there is a single particle $X_\varnothing$ labeled by $u = \varnothing$, born at time $\beta_\varnothing := 0$ and distributed like X. Then, we enumerate the sequence $(\beta_{u1}, n_1), (\beta_{u2}, n_2), \dots$ of the negative jump times and sizes of X_u so that $n_1 \geq n_2 \geq \cdots$ and $\beta_{uj} < \beta_{u(j+1)}$ whenever $n_j = n_{j+1}$. Conditionally on $(n_j)_{n\geq 1}$, the processes X_{uj}, $j = 1, 2, \dots$ are independent and distributed like $X^{(n_j)}$, respectively (for definiteness, we set $\beta_{uj} := \infty$ and $X_{uj} :\equiv 0$ if X_u makes less than j negative jumps during its lifetime), and we have

$$\big(\mathbf{X}(k)\big)_{k\geq 0} \overset{d}{=} \Big(X_u(k-\beta_u) \colon u \in \mathbb{U},\ \beta_u \leq k\Big)_{k\geq 0}.$$

[2] In the peeling of random Boltzmann maps [8], the locally largest cycles are called *left-twigs*.

Recall the notation $\cdot^{(n)}$ to stress that the system is started from a particle with size $X_\varnothing(0) = n$.

Lemma 3.1 *Suppose* (H1) *and* (H2)*. Then for every finite subset $U \subset \mathbb{U}$, there is the convergence in $\mathbb{D}([0,\infty), \mathbb{R}^U)$:*

$$\left(\frac{X_u^{(n)}(\lfloor a_n t\rfloor - \beta_u^{(n)})}{n} : u \in U\right)_{t\geq 0} \xrightarrow[n\to\infty]{(d)} \left(Y_u(t-b_u) : u \in U\right)_{t\geq 0}. \tag{6}$$

Proof We follow the argument used to prove the second part of [8, Lemma 17]. For $h \geq 0$, let $\mathbb{U}_h := \{u \in \mathbb{U} : |u| \leq h\}$ be the set of vertices with height at most h in the tree $\mathbb{U}$. It suffices to show

$(\mathscr{I}_h)$: Convergence (6) holds in $\mathbb{D}([0,\infty), \mathbb{R}^U)$ for every finite subset $U \subset \mathbb{U}_h$,

by induction on h. The statement $(\mathscr{I}_0)$ is given by (2). Now, if U is a finite subset of $\mathbb{U}_{h+1}$ and F_u, $u \in U$, are continuous bounded functions from $\mathbb{D}([0,\infty), \mathbb{R})$ to $\mathbb{R}$, then the branching property entails that, for $\hat{X}_u^{(n)} := X_u^{(n)}(\lfloor a_n \cdot\rfloor - \beta_u)/n$,

$$\mathbb{E}\left[\prod_{u\in U} F_u\big(\hat{X}_u^{(n)}\big) \,\middle|\, \big(X_u^{(n)} : u \in \mathbb{U}_h\big)\right] = \prod_{u\in U\cap \mathbb{U}_h} F_u\big(\hat{X}_u^{(n)}\big) \cdot \prod_{\substack{u\in U\\ |u|=h+1}} E^{(n)}_{\hat{X}_u^{(n)}(0)}[F_u],$$

where $E_x^{(n)}$ stands for expectation under the law $P_x^{(n)}$ of $\hat{X}_\varnothing$ started from x, which by (1), (2) and (3), converges weakly as $n \to \infty$ to the law P_x of Y started from x. The values $\hat{X}_u^{(n)}(0)$ for $|u| = h+1$ correspond to (rescaled) negative jump sizes of particles at height h. With [18, Corollary VI.2.8] and our convention of ranking the jump sizes in the non-increasing order, the convergence in distribution $(\hat{X}_u^{(n)}(0) : u \in U, |u| = h+1) \to (Y_u(0) : u \in U, |u| = h+1)$ as $n \to \infty$ thus holds jointly with $(\mathscr{I}_h)$. Further, thanks to the Feller property [20, Lemma 2.1] of Y, its distribution is weakly continuous in its starting point. By the continuous mapping theorem we therefore obtain, applying back the branching property, that

$$\mathbb{E}\left[\prod_{u\in U} F_u\big(\hat{X}_u^{(n)}\big)\right] \xrightarrow[n\to\infty]{} \mathbb{E}\left[\prod_{u\in U} F_u\big(Y_u(\cdot - b_u)\big)\right].$$

A priori, this establishes the convergence in distribution $(\hat{X}_u^{(n)} : u \in U) \to (Y_u(\cdot - b_u) : u \in U)$ only in the product space $\mathbb{D}([0,\infty), \mathbb{R})^U$. By [18, Proposition 2.2] it will also hold in $\mathbb{D}([0,\infty), \mathbb{R}^U)$ provided that the processes Y_u, $u \in U$, almost surely never jump simultaneously. But this is plain since particles evolve independently and the jumps of Y are totally inaccessible. Thus, $(\mathscr{I}_h) \implies (\mathscr{I}_{h+1})$. □

Next, we proceed to the convergence of the finite-dimensional marginals of $\mathcal{X}$, which we shall first formally construct. For each $u \in \mathbb{U}$ with $\beta_u < \infty$, let ζ_u denote the lifetime of the stopped Markov chain X_u. Recall the definition in Sect. 2 of the trees $\mathcal{Y}, \mathcal{Y}_h,\ h \geq 0$, related to $\mathbf{Y}$'s genealogy, that echoes Rembart and Winkel's construction [24]. Similarly, let $\mathcal{X}_{u,0}$ simply consist of an edge with length ζ_u, rooted at a vertex u. Recursively, define $\mathcal{X}_{u,h+1}$ by attaching to the edge $\mathcal{X}_{u,0}$ the trees $\mathcal{X}_{uj,h}$ at a distance $\beta_{uj} - \beta_u$ from the root u, respectively, for each born particle $uj,\ j \leq h+1$, descending from u. The tree $\mathcal{X}_h := \mathcal{X}_{\varnothing,h}$ is a finite tree whose vertices are labeled by the set $\mathbb{U}^{(h)}$ of words over $\{1, \ldots, h\}$ with length at most h. Plainly, the sequence $\mathcal{X}_h,\ h \geq 0$, is consistent, in that $\mathcal{X}_h$ is the subtree of $\mathcal{X}_{h+1}$ with vertex set $\mathbb{U}^{(h)}$, and we may consider the inductive limit $\mathcal{X} := \lim \uparrow_{h\to\infty} \mathcal{X}_h$. We write $d_n(v, v')$ for the length of the unique path between v and v' in $\mathcal{X}^{(n)}$. All these trees belong to the space $\mathscr{T}$ of (equivalence classes of) compact, rooted, real trees and can be embedded as subspaces of a large metric space (such as, for instance, the space $\ell^1(\mathbb{N})$ of summable sequences [3, Sect. 2.2]). Irrespective of the embedding, they can be compared one with each other through the so-called Gromov–Hausdorff metric $\mathrm{d_{GH}}$ on $\mathscr{T}$. We forward the reader to [13,21] and references therein.

Lemma 3.2 *Suppose* (H1)–(H3). *Then for all* $h \in \mathbb{N}$, *there is the convergence in* $(\mathscr{T}, \mathrm{d_{GH}})$*:*

$$\left(\mathcal{X}_h^{(n)}, \frac{d_n}{a_n}\right) \xrightarrow[n\to\infty]{(d)} \mathcal{Y}_h.$$

Proof It suffices to show the joint convergence of all branches. The branch going from the root $\varnothing$ through the vertex $u \in \mathbb{U}^{(h)}$ has total length $\mathcal{E}_u^{(n)} := \beta_u^{(n)} + \zeta_u^{(n)}$ in $(\mathcal{X}_h^{(n)}, d_n)$, and length $\epsilon_u := b_u + \zeta_u$ in $\mathcal{Y}_h$. Recall that conditionally on $\{X_u(0) = n\}$, the random variable ζ_u has the same distribution as $\zeta^{(n)} := \zeta_\varnothing^{(n)}$. By [9, Theorem 3.(i)], the convergence

$$\frac{\zeta^{(n)}}{a_n} \xrightarrow[n\to\infty]{(d)} \zeta := \inf\{t \geq 0 \colon Y(t) = 0\}$$

holds jointly with (2). Adapting the proof of Lemma 3.1, we can more generally check that for every finite subset $U \subset \mathbb{U}$, we have, jointly with (6),

$$\left(\frac{\mathcal{E}_u^{(n)}}{a_n} : u \in U\right) \xrightarrow[n\to\infty]{(d)} \big(\epsilon_u : u \in U\big).$$

In particular, this is true for $U := \mathbb{U}^{(h)}$. □

To conclude this section, we restate an observation of Bertoin, Curien, and Kortchemski [8, Lemma 21] which results from the convergence of finite-dimensional marginals (Lemma 3.1): with high probability as $h \to \infty$, "non-negligible" particles have their labels in $\mathbb{U}^{(h)}$. Specifically, say that an individual $u \in \mathbb{U}$ is (n, ε)-good, and write $u \in \mathcal{G}(n, \varepsilon)$, if the particles X_v labeled by each ancestor v of u (including u itself) have size at birth at least $n\varepsilon$. Then:

Lemma 3.3 *We have*

$$\lim_{h\to\infty}\limsup_{n\to\infty}\ \mathbb{P}^{(n)}\left(\mathcal{G}(n,\varepsilon)\not\subseteq\mathbb{U}^{(h)}\right)=0.$$

4 A Size-Biased Particle and a Many-to-One Formula

We now introduce a "size-biased particle" and relate it to a many-to-one formula. This will help us derive tightness estimates in Sects. 5 and 6 and thus complement the finite-dimensional convergence results of the preceding section. Recall from Assumptions (H1)–(H3) that we can find $q_* \in (0, q^*)$ such that, as $n \to \infty$, $\kappa_n(q) \to \kappa(q) < 0$ for every $q \in [q_*, q^*)$. Consequently, *we may and will suppose for the remainder of this section that the freezing threshold M is taken sufficiently large so that $\kappa_n(q_*) \le 0$ for every $n > M$*. (Note that $\kappa_n(q_*) = 0$ for $n \le M$, by our convention $p_{n,n} := 1$.)

Lemma 4.1 *For every $n \in \mathbb{N}$,*

$$\mathbb{E}^{(n)}\left[\left(X(1)\right)^{q_*}+\left(n-X(1)\right)_+^{q_*}\right]\le n^{q_*}. \tag{7}$$

Therefore, the process[3]

$$\sum_{u\in\mathbb{U}}\left(X_u(k-\beta_u)\right)^{q_*},\qquad k\ge 0,$$

is a supermartingale under $\mathbb{P}^{(n)}$.

Proof The left-hand side of (7) is

$$n^{q_*}\sum_{m=0}^{\infty}p_{n,m}\left[\left(\frac{m}{n}\right)^{q_*}+\left(1-\frac{m}{n}\right)_+^{q_*}\right]=n^{q_*}\left(1+\frac{\kappa_n(q_*)}{a_n}\right),$$

where $\kappa_n(q_*) \le 0$. Hence, the first part of the statement follows. The second part follows by applying the branching property at any given time $k \ge 0$:

$$\begin{aligned}\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}}\left(X_u(k+1-\beta_u)\right)^{q_*}\,\middle|\,\mathbf{X}(k)=(x_i:i\in I)\right]&=\sum_{i\in I}\mathbb{E}^{(x_i)}\left[\left(X(1)\right)^{q_*}+\left(x_i-X(1)\right)_+^{q_*}\right]\\&\le\sum_{i\in I}(x_i)^{q_*}\\&=\sum_{u\in\mathbb{U}}\left(X_u(k-\beta_u)\right)^{q_*}.\end{aligned}$$

□

[3] We set here $X_u(i) := 0$ for $i < 0$ in order to not burden the notation with the indicator $\mathbb{1}_{\{\beta_u\le k\}}$.

Remark 4.2 Put differently, the condition "$\kappa_n(q_*) \le 0$" entails that $n \mapsto n^{q_*}$ is superharmonic with respect to the "fragmentation operator". This map plays the same role as the function f in [8], where it takes the form of a cubic polynomial ($q_* = 3$) and

$$\sum_{u \in \mathbb{U}} f\big(X_u(k - \beta_u)\big), \qquad k \ge 0,$$

is actually a martingale. More generally, the map $n \mapsto n^{q_*}$ could be replaced by any regularly varying sequence with index q_*, but probably at the cost of heavier notation.

As we see in the proof of Lemma 4.1, the fact that $\kappa_n(q_*) \le 0$ allows us to introduce a (defective) Markov chain $(\bar{X}(k) \colon k \ge 0)$ on $\mathbb{N}$, to which we add 0 as cemetery state, with transition

$$\mathbb{E}^{(n)}\big[f\big(\bar{X}(1)\big);\ \bar{X}(1) \neq 0\big] = \sum_{m=1}^{\infty} p_{n,m}\left[\left(\frac{m}{n}\right)^{q_*} f(m) + \left(1 - \frac{m}{n}\right)_+^{q_*} f(n-m)\right]. \tag{8}$$

We let $\bar{\zeta} := \inf\{k \ge 0 \colon \bar{X}(k) = 0\}$ denote its lifetime. Up to a change of probability measure, $\bar{X}$ follows the trajectory of a randomly selected particle in $\mathbf{X}$, until it is eventually absorbed to the cemetery state 0. It admits the following scaling limit (which could also be seen as a randomly selected particle in $\mathbf{Y}$; see [7, Sect. 4]):

Proposition 4.3 *There is the convergence in distribution*

$$\left(\frac{\bar{X}^{(n)}(\lfloor a_n t \rfloor)}{n} \colon t \ge 0\right) \xrightarrow[n\to\infty]{(d)} \big(\bar{Y}(t) \colon t \ge 0\big) \tag{9}$$

in $\mathbb{D}([0,\infty), \mathbb{R})$*, where the limit* $\bar{Y}$ *fulfills the same identity* (4) *as* Y*, but for a (killed) Lévy process* $\bar{\xi}$ *with characteristic exponent* $\bar{\kappa}(q) := \kappa(q_* + q)$*. Further, if* $\bar{\zeta}$ *denotes the lifetime of* $\bar{Y}$*, then the convergence*

$$\frac{\bar{\zeta}^{(n)}}{a_n} \xrightarrow[n\to\infty]{(d)} \bar{\zeta}$$

holds jointly with (9).

Proof Write $\bar{\Lambda}_n$ for the law of $\log(\bar{X}^{(n)}(1)/n)$, with the convention $\log 0 := -\infty$. We see from (8) that $a_n\,\mathbb{P}(\bar{X}(1) = 0) = -\kappa_n(q_*)$, and, for every $0 \le q \le q^* - q_*$,

$$\begin{aligned}\int_{\mathbb{R}} (e^{qy} - 1)\, a_n \bar{\Lambda}_n(\mathrm{d}y) &= a_n \sum_{m=0}^{\infty} p_{n,m}\left[\left(\frac{m}{n}\right)^{q_*}\left(\left(\frac{m}{n}\right)^{q} - 1\right) + \left(1 - \frac{m}{n}\right)_+^{q_*}\left(\left(1 - \frac{m}{n}\right)^{q} - 1\right)\right] \\ &= \kappa_n(q_* + q) - \kappa_n(q_*).\end{aligned}$$

Hence,

$$-a_n \bar{\Lambda}_n(\{-\infty\}) + \int_{\mathbb{R}} (e^{qy} - 1)\, a_n \bar{\Lambda}_n(\mathrm{d}y) = \kappa_n(q_* + q) \xrightarrow[n\to\infty]{} \bar{\kappa}(q).$$

Furthermore, by (H2),

$$\limsup_{n\to\infty} a_n \int_1^\infty e^{(q^*-q_*)y}\,\bar{\Lambda}_n(\mathrm{d}y) \le \limsup_{n\to\infty} a_n \sum_{m=2n}^{\infty} p_{n,m}\left(\frac{m}{n}\right)^{q^*} < \infty.$$

In other words, assumptions (A1), (A2) and (A3) of [9] are satisfied (w.r.t the Markov chain $\bar{X}$ and the limiting process $\bar{Y}$). Our statement thus follows from Theorems 1 and 2 there.[4] □

Heading now toward pathwise and optional many-to-one formulae, we first set up some notation. Let $\mathtt{A} \subseteq \mathbb{N}$ be a fixed subset of states, and let $\ell \in \partial\mathbb{U}$ refer to an infinite word over $\mathbb{N}$, which we see as a branch of $\mathbb{U}$. For every $u \in \mathbb{U}\cup\partial\mathbb{U}$ and every $k \ge 0$, set

$$\widetilde{X}_u(k) := X_{u[k]}(k-\beta_{u[k]}),$$

where $u[k]$ is the youngest ancestor v of u with $\beta_v \le k$, and write $\tau_u^{\mathtt{A}} := \inf\{k \ge 0\colon \widetilde{X}_u(k) \in \mathtt{A}\}$ for the first hitting time of $\mathtt{A}$ by $\widetilde{X}_u$. Let also $\bar{\tau}^{\mathtt{A}} := \inf\{k \ge 0\colon \bar{X}(k) \in \mathtt{A}\}$. Now, imagine that once a particle hits $\mathtt{A}$, it is stopped and thus has no further progeny. The state when all particles have hit $\mathtt{A}$ in finite time is $x_u^{\mathtt{A}} := \widetilde{X}_u(\tau_u^{\mathtt{A}})$, $u \in \mathbb{U}_{\mathtt{A}}$, where $\mathbb{U}_{\mathtt{A}} := \{u \in \mathbb{U}\colon \ell[\tau_\ell^{\mathtt{A}}] = u \text{ for some } \ell \in \partial\mathbb{U} \text{ with } \tau_\ell^{\mathtt{A}} < \infty\}$.

Lemma 4.4 (Many-to-one formula)

(i) *For every $n \in \mathbb{N}$, every $k \ge 0$, and every $f\colon \mathbb{N}^{k+1} \to \mathbb{R}_+$,*

$$\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}} \big(X_u(k-\beta_u)\big)^{q_*} f\left(\widetilde{X}_u(i)\colon i \le k\right)\right] = n^{q_*}\,\mathbb{E}^{(n)}\left[f\left(\bar{X}(i)\colon i \le k\right);\ \bar{\zeta} > k\right].$$

(ii) *For every $n \in \mathbb{N}$, every $\mathtt{A} \subseteq \mathbb{N}$, and every $f\colon \mathbb{Z}_+ \times \mathbb{N} \to \mathbb{R}_+$,*

$$\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}_{\mathtt{A}}} \big(x_u^{\mathtt{A}}\big)^{q_*} f\big(\tau_u^{\mathtt{A}}, x_u^{\mathtt{A}}\big)\right] = n^{q_*}\,\mathbb{E}^{(n)}\left[f\big(\bar{\tau}^{\mathtt{A}}, \bar{X}(\bar{\tau}^{\mathtt{A}})\big);\ \bar{\zeta} > \bar{\tau}^{\mathtt{A}}\right].$$

Proof (i) The proof is classical (see, e.g., [26, Theorem 1.1]) and proceeds by induction on k. The identity clearly holds for $k = 0$. Using (8) together with the branching property at time k,

$$\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}} \big(X_u(k+1-\beta_u)\big)^{q_*} f\left(\widetilde{X}_u(i)\colon i \le k+1\right) \,\middle|\, \widetilde{X}_u(i) = x_{u,i},\ i \le k\right]$$

$$= \sum_{u\in\mathbb{U}} \sum_{m=0}^{\infty} p_{x_{u,k},m}\left(m^{q_*} f\big(x_{u,0},\ldots,x_{u,k},m\big)\right.$$

[4] Strictly speaking, the results are only stated when there is no killing, that is $\kappa(q_*) = 0$, but as mentioned by the authors [9, p. 2562, §2], they can be extended using the same techniques to the case where some killing is involved.

$$+\left(x_{u,k}-m\right)_+^{q_*} f\left(x_{u,0},\ldots,x_{u,k},x_{u,k}-m\right)\Big)$$
$$=\sum_{u\in\mathbb{U}}\left(x_{u,k}\right)^{q_*}\mathbb{E}^{(x_{u,k})}\left[f\left(x_{u,0},\ldots,x_{u,k},\bar{X}(1)\right);\ \bar{X}(1)\neq 0\right].$$

By taking expectations on both sides and applying the induction hypothesis with the function $\widetilde{f}(x_0,\ldots,x_k):=\mathbb{E}^{(x_k)}[f(x_0,\ldots,x_k,\bar{X}(1));\ \bar{X}(1)\neq 0]$ on the one hand, and by applying the Markov property of $\bar{X}$ at time k on the other hand, we derive the identity at time $k+1$:

$$\begin{aligned}&\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}}\left(X_u(k+1-\beta_u)\right)^{q_*} f\left(\widetilde{X}_u(i)\colon i\le k+1\right)\right]\\&\quad=n^{q_*}\,\mathbb{E}^{(n)}\left[\widetilde{f}\left(\bar{X}(i)\colon i\le k\right);\ \bar{\zeta}>k\right]\\&\quad=n^{q_*}\,\mathbb{E}^{(n)}\left[f\left(\bar{X}(i)\colon i\le k+1\right);\ \bar{\zeta}>k+1\right].\end{aligned}$$

(ii) For every $k\ge 0$ and every $x_0,\ldots,x_k\in\mathbb{N}$, we set $f_k^{\mathtt{A}}(x_0,\ldots,x_k):=\mathbb{1}_{\{x_0\notin\mathtt{A},\ldots,x_{k-1}\notin\mathtt{A},x_k\in\mathtt{A}\}}f(k,x_k)$. Then

$$\begin{aligned}\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}_{\mathtt{A}}}\left(x_u^{\mathtt{A}}\right)^{q_*} f\left(\tau_u^{\mathtt{A}},x_u^{\mathtt{A}}\right)\right]&=\sum_{k=0}^{\infty}\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}}\left(X_u(k-b_u)\right)^{q_*} f_k^{\mathtt{A}}\left(\widetilde{X}_u(i)\colon i\le k\right)\right]\\&=n^{q_*}\sum_{k=0}^{\infty}\mathbb{E}^{(n)}\left[f\left(k,\bar{X}(k)\right);\ \bar{\tau}^{\mathtt{A}}=k;\ \bar{\zeta}>k\right]\\&=n^{q_*}\,\mathbb{E}^{(n)}\left[f\left(\bar{\tau}^{\mathtt{A}},\bar{X}\left(\bar{\tau}^{\mathtt{A}}\right)\right);\ \bar{\zeta}>\bar{\tau}^{\mathtt{A}}\right],\end{aligned}$$

by (i) and the monotone convergence theorem. □

We now combine Proposition 4.3 and Lemma 4.4 to derive the following counterpart of [8, Lemma 14] that we will apply in the next two sections. Consider the hitting set $\mathtt{A}:=\{1,\ldots,\lfloor n\varepsilon\rfloor\}$ and denote by $x_u^{\le n\varepsilon}:=x_u^{\mathtt{A}}$, $u\in\mathbb{U}^{\le n\varepsilon}:=\mathbb{U}_{\mathtt{A}}$, the population of particles stopped below $n\varepsilon$.

Corollary 4.5 *We have*

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\ n^{-q_*}\,\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}^{\le n\varepsilon}}\left(x_u^{\le n\varepsilon}\right)^{q_*}\right]=0.$$

Proof By Lemma 4.4,

$$n^{-q_*}\,\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}^{\le n\varepsilon}}\left(x_u^{\le n\varepsilon}\right)^{q_*}\right]=\mathbb{P}^{(n)}\left(\bar{\zeta}>\bar{\tau}^{\le n\varepsilon}\right),$$

where $\bar{\tau}^{\le n\varepsilon} := \inf\{k \ge 0 \colon \bar{X}(k) \le n\varepsilon\}$. Thus, if $\bar{\zeta}$ is the lifetime of $\bar{Y}$ and $\bar{\tau}^{\le\varepsilon} := \inf\{t \ge 0 \colon \bar{Y}(t) \le \varepsilon\}$, then by Proposition 4.3 and the continuous mapping theorem,

$$\limsup_{n\to\infty} n^{-q_*}\, \mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}^{\le n\varepsilon}} \left(x_u^{\le n\varepsilon}\right)^{q_*}\right] \le \mathbb{P}\left(\bar{\zeta} > \bar{\tau}^{\le\varepsilon}\right),$$

which tends to 0 as $\varepsilon \to 0$. □

5 Proof of Theorem 1

We prove Theorem 1 by combining Lemma 3.1 with the next two "tightness" properties. We suppose that Assumptions (H1)–(H3) hold and recall that $\mathbb{U}^{(h)} \subset \mathbb{U}$ refers to the set of words over $\{1, \ldots, h\}$ with length at most h.

Lemma 5.1 *For every $\delta > 0$,*

$$\lim_{h\to\infty} \mathbb{P}\left(\sup_{t\ge0} \sum_{u\in\mathbb{U}\setminus\mathbb{U}^{(h)}} \left(Y_u(t-b_u)\right)^{q^*} > \delta\right) = 0.$$

Proof This was already derived in [8, Lemma 20] and results from the following fact [6, Corollary 4]:

$$\mathbb{E}\left[\sum_{u\in\mathbb{U}} \sup_{t\ge0} \left(Y_u(t-b_u)\right)^{q}\right] < \infty \quad \text{for } \kappa(q) < 0.$$

□

Lemma 5.2 *If M is sufficiently large, then for every $\delta > 0$,*

$$\lim_{h\to\infty} \limsup_{n\to\infty} \mathbb{P}^{(n)}\left(\sup_{k\ge0} \sum_{u\in\mathbb{U}\setminus\mathbb{U}^{(h)}} \left(X_u(k-\beta_u)\right)^{q^*} > \delta n^{q^*}\right) = 0.$$

Proof Let us first take $q_* < q^*$ and M as in Sect. 4. As in the proof of [8, Lemma 22] and by definition of $\mathcal{G}(n, \varepsilon)$ in Sect. 3, we claim that each particle in $\{X_u(k-\beta_u) \colon u \in \mathbb{U}\setminus\mathcal{G}(n,\varepsilon)\}$ has an ancestor with size at birth smaller than $n\varepsilon$. Thanks to the branching property, we may therefore consider that these particles derive from a system that has first been "frozen" below the level $n\varepsilon$, that is, with the notations of Sect. 4, from a particle system having $x_u^{\le n\varepsilon}$, $u \in \mathbb{U}^{\le n\varepsilon}$, as initial population. Hence, by Lemma 4.1 and Doob's maximal inequality,

$$\mathbb{P}^{(n)}\left(\sup_{k\ge0} \sum_{u\in\mathbb{U}\setminus\mathcal{G}(n,\varepsilon)} \left(X_u(k-\beta_u)\right)^{q^*} > \delta n^{q^*}\right) \le \frac{1}{\delta^{q_*/q^*}\, n^{q_*}}\, \mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}^{\le n\varepsilon}} \left(x_u^{\le n\varepsilon}\right)^{q_*}\right]$$

(bounding from above the ℓ^{q^*}-norm by the ℓ^{q_*}-norm). We conclude by Corollary 4.5 and Lemma 3.3. □

Proof of Theorem 1 From Lemmas 3.1, 5.1 and 5.2, we deduce the convergence in distribution

$$\left(\frac{X_u^{(n)}(\lfloor a_n t\rfloor - \beta_u^{(n)})}{n} : u \in \mathbb{U}\right)_{t\geq 0} \xrightarrow[n\to\infty]{(d)} \left(Y_u(t-b_u) \colon u \in \mathbb{U}\right)_{t\geq 0},$$

in the space $\mathbb{D}([0,\infty), \ell^{q^*}(\mathbb{U}))$ of $\ell^{q^*}(\mathbb{U})$-valued càdlàg functions on $[0,\infty)$, where

$$\ell^{q^*}(\mathbb{U}) := \left\{\mathbf{x} := (x_u : u \in \mathbb{U}) \colon \sum_{u\in\mathbb{U}} (x_u)^{q^*} < \infty\right\}.$$

Since for $q \geq 1$, rearranging sequences in the non-increasing order does not increase their q-distance [22, Theorem 3.5], the convergence in $\ell^{q^*}(\mathbb{U})$ implies that in $\ell^{q\downarrow}$, $q \geq 1 \vee q^*$. □

6 Proof of Theorem 2

Similarly to the previous section, by Lemma 3.2 the proof of Theorem 2 is complete once we have established that

$$\lim_{h\to\infty} \mathbb{P}\Big(\mathrm{d}_{\mathrm{GH}}(\mathcal{Y}, \mathcal{Y}_h) > \delta\Big) = 0,$$

and

$$\lim_{h\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\mathrm{d}_{\mathrm{GH}}\left(\mathcal{X}^{(n)}, \mathcal{X}_h^{(n)}\right) > \delta a_n\right) = 0, \tag{10}$$

for all $\delta > 0$. The first display is clear since the tree $\mathcal{Y}$ is compact. The second will be a consequence of the following counterpart of [8, Conjecture 1]:

Lemma 6.1 *Suppose* (H1)–(H3)*, and $q^* > \gamma$. Then for every $q < q^*$, and for M sufficiently large,*

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[\left(\frac{\mathrm{ht}\big(\mathcal{X}^{(n)}\big)}{a_n}\right)^{q/\gamma}\right] < \infty,$$

where $\mathrm{ht}\big(\mathcal{X}^{(n)}\big) := \sup_{x\in\mathcal{X}^{(n)}} d_n(\varnothing, x)$ is the height of the tree $\mathcal{X}^{(n)}$.

The proof of Lemma 6.1 involves martingale arguments. Prior to writing it, we need a preparatory lemma. Let us define

$$\widetilde{\kappa}_n(q) := a_n \sum_{m=1}^{\infty} p_{n,m} \left[\left(\frac{a_m}{a_n} \right)^{q/\gamma} - 1 + \left(\frac{a_{n-m}}{a_n} \right)^{q/\gamma} \right],$$

which slightly differs from $\kappa_n(q)$ to the extent that we have replaced the map $m \mapsto m^q$ by the q-regularly varying sequence $A_q(m) := a_m^{q/\gamma}$, $m \in \mathbb{N}$ (for convenience, we have set $a_m := 0$, $m \leq 0$). Of course, $\widetilde{\kappa}_n = \kappa_n$ if $a_m = m^\gamma$ for every $m \in \mathbb{N}$.

Lemma 6.2 *Suppose $q^* > \gamma$. Then we can find $q_* \in (0, q^*)$ such that, for every $q \in [q_*, q^*)$,*

$$\lim_{n\to\infty} \widetilde{\kappa}_n(q) = \kappa(q) < 0.$$

Proof We will more generally show that for every q-regularly varying sequence (r_n),

$$\left| a_n \sum_{m=1}^{\infty} p_{n,m} \left[\frac{r_m}{r_n} - \left(\frac{m}{n} \right)^q \right] \right| + \left| a_n \sum_{m=1}^{n-1} p_{n,m} \left[\frac{r_{n-m}}{r_n} - \left(1 - \frac{m}{n} \right)^q \right] \right| \xrightarrow[n\to\infty]{} 0,$$

provided $q < q^*$ is close enough to q^*. Denoting by Λ_n the law of $\log(X^{(n)}(1)/n)$, we observe that

$$a_n \sum_{m=1}^{\infty} p_{n,m} \left[\frac{r_m}{r_n} - \left(\frac{m}{n} \right)^q \right] = a_n \int_{-\infty}^{\infty} \left[\left(\frac{r_{ne^x}}{r_n} \right) - e^{qx} \right] \Lambda_n(\mathrm{d}x),$$

which, by repeating the arguments in [9, Proof of Lemma 4.9], tends to 0 as $n \to \infty$. Next, an appeal to Potter's bounds [11, Theorem 1.5.6] shows that for every $c > 1$ and $\delta > 0$ arbitrary small,

$$\frac{1}{c} \left(\frac{m}{n} \right)^{q+\delta} \leq \frac{r_m}{r_n} \leq c \left(\frac{m}{n} \right)^{q-\delta}$$

whenever $m < n$ are sufficiently large. Thus, recalling that $\Psi_n(q) \to \Psi(q)$ and $\kappa_n(q) \to \kappa(q)$ for every q in some left-neighborhood of q^*, we have

$$\liminf_{n\to\infty} a_n \sum_{m=1}^{n-1} p_{n,m} \left[\frac{r_{n-m}}{r_n} - \left(1 - \frac{m}{n} \right)^q \right] \geq \frac{1}{c} \Big(\kappa(q+\delta) - \Psi(q+\delta) \Big) - \Big(\kappa(q) - \Psi(q) \Big),$$

and

$$\limsup_{n\to\infty} a_n \sum_{m=1}^{n-1} p_{n,m} \left[\frac{r_{n-m}}{r_n} - \left(1 - \frac{m}{n} \right)^q \right] \leq \frac{1}{c} \Big(\kappa(q-\delta) - \Psi(q-\delta) \Big) - \Big(\kappa(q) - \Psi(q) \Big).$$

We conclude by letting $c \to 1$ and $\delta \to 0$. □

We can now prove Lemma 6.1.

Proof of Lemma 6.1 We shall rely on a Foster-type technique close to the machinery developed in [4]; see in particular the proof of Theorem 2' there. First, observe that $\mathrm{ht}(\mathcal{X})$ is distributed like the extinction time $\mathcal{E}$ of $\mathbf{X}$:

$$\mathrm{ht}\big(\mathcal{X}^{(n)}\big) \stackrel{d}{=} \sup_{u\in\mathbb{U}} \mathcal{E}_u^{(n)} =: \mathcal{E}^{(n)}.$$

Fix $q \in (\gamma, q^*)$ arbitrary close to q^* and set $r := q/\gamma$. By Lemma 6.2, suppose M large enough so that $\widetilde{\kappa}_m(q) < 0$ for every $m > M$. It is easy to see as in the proof of Lemma 4.1 that the process

$$\Gamma(k) := \sum_{u\in\mathbb{U}} A_q\big(X_u(k-\beta_u)\big), \qquad k \geq 0,$$

is a supermartingale under $\mathbb{P}^{(n)}$ (with respect to the natural filtration $(\mathcal{F}_k)_{k\geq 0}$ of $\mathbf{X}$): indeed, for $\mathbf{X}(k) = (x_i : i \in I)$,

$$\mathbb{E}^{(n)}\Big[\Gamma(k+1) - \Gamma(k) \,\Big|\, \mathcal{F}_k\Big] = \sum_{i\in I} \widetilde{\kappa}_{x_i}(q)\, A_{q-\gamma}(x_i),$$

where the right-hand side is (strictly) negative on the event $\{\mathcal{E} > k\} = \{\exists i \in I : x_i > M\}$. We will more precisely show the existence of $\eta > 0$ sufficiently small such that the process

$$G(k) := \Big(\Gamma(k)^{1/r} + \eta\left(\mathcal{E}\wedge k\right)\Big)^r, \qquad k \geq 0,$$

is a $(\mathcal{F}_k)_{k\geq 0}$-supermartingale under $\mathbb{P}^{(n)}$, for any $n \in \mathbb{N}$. Then, the result will be readily obtained from $\eta^r\, \mathbb{E}^{(n)}[(\mathcal{E}\wedge k)^r] \leq \mathbb{E}^{(n)}[G(k)] \leq \mathbb{E}^{(n)}[G(0)] = A_q(n) = a_n^r$ and an appeal to Fatou's lemma.

On the one hand, we have

$$\sigma := \sum_{i\in I} A_{q-\gamma}(x_i) \geq \left(\sum_{i\in I} A_q(x_i)\right)^{1-\gamma/q}$$

because

$$\frac{A_q(x_i)}{\sigma^{q/(q-\gamma)}} = \left(\frac{A_{q-\gamma}(x_i)}{\sigma}\right)^{q/(q-\gamma)} \leq \frac{A_{q-\gamma}(x_i)}{\sigma},$$

where $q/(q-\gamma) > 1$ and the right-hand side sums to 1 as i ranges over I. Then, if we let $\eta > 0$ sufficiently small such that $\widetilde{\kappa}_m(q) \leq -r\eta$ for every $m > M$, we deduce that

$$\mathbb{E}^{(n)}\Big[\Gamma(k+1) \,\Big|\, \mathcal{F}_k\Big] \leq \Gamma(k)\left(1 - r\eta\,\Gamma(k)^{-\gamma/q}\, \mathbb{1}_{\{\mathcal{E}>k\}}\right).$$

Raising this to the power $1/r = \gamma/q$ yields

$$\mathbb{E}^{(n)}\left[\Gamma(k+1) \;\middle|\; \mathcal{F}_k\right]^{1/r} \le \Gamma(k)^{1/r}\left(1-\eta\,\Gamma(k)^{-\gamma/q}\,\mathbb{1}_{\{\mathcal{E}>k\}}\right) = \Gamma(k)^{1/r} - \eta\,\mathbb{1}_{\{\mathcal{E}>k\}}, \tag{11}$$

by concavity of $x \mapsto x^{1/r}$. On the other hand, the supermartingale property also implies that $(\Gamma(k+1)^{1/r}+a)^r$ is integrable for every constant $a > 0$; we may thus apply the generalized triangle inequality [4, Lemma 1] with the positive, convex increasing function $x \mapsto x^r$, the positive random variable $\Gamma(k+1)^{1/r}$, and the probability $\mathbb{P}^{(n)}(\,\cdot \mid \mathcal{F}_k)$ (under which $\mathcal{E} \wedge (k+1)$ can be seen as a positive constant):

$$\begin{aligned}&\mathbb{E}^{(n)}\left[\left(\Gamma(k+1)^{1/r} + \eta\left(\mathcal{E}\wedge(k+1)\right)\right)^r \;\middle|\; \mathcal{F}_k\right]^{1/r}\\ &\quad\le \mathbb{E}^{(n)}\left[\Gamma(k+1) \mid \mathcal{F}_k\right]^{1/r} + \eta\left(\mathcal{E}\wedge(k+1)\right).\end{aligned}$$

Reporting (11) shows as desired that $(G(k)\colon k \ge 0)$ is a supermartingale. □

We are finally ready to derive (10) and complete the proof of Theorem 2.

Proof of (10) We start as in the proof of Theorem 1: thanks to Lemma 3.3 and the branching property, with high probability as $h \to \infty$, the connected components of $\mathcal{X}^{(n)} \setminus \mathcal{X}_h^{(n)}$ are included in independent copies of $\mathcal{X}$ stemming from the population $x_u^{\le n\varepsilon}$, $u \in \mathbb{U}^{\le n\varepsilon}$, of particles frozen below $n\varepsilon$. Specifically,

$$\begin{aligned}&\mathbb{P}^{(n)}\left(\mathrm{d_{GH}}\left(\mathcal{X}^{(n)}, \mathcal{X}_h^{(n)}\right) > \delta a_n\right)\\ &\quad\le \mathbb{P}^{(n)}\left(\mathcal{G}(n,\varepsilon) \not\subseteq \mathbb{U}^{(h)}\right) + \mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}^{\le n\varepsilon}} \mathbb{P}^{(x_u^{\le n\varepsilon})}\left(\mathrm{ht}(\mathcal{X}) > \delta a_n\right)\right].\end{aligned}$$

Now, take $q_* < q < q^*$ and M large enough so that both Lemma 6.1 and the results of Sect. 4 hold. So, there exists a constant $C > 0$ such that

$$\begin{aligned}\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}^{\le n\varepsilon}} \mathbb{P}^{(x_u^{\le n\varepsilon})}\left(\mathrm{ht}(\mathcal{X}) > \delta a_n\right)\right] &= \mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}^{\le n\varepsilon}} \mathbb{P}^{(x_u^{\le n\varepsilon})}\left(\mathrm{ht}(\mathcal{X}) > \delta a_{x_u^{\le n\varepsilon}} \frac{a_n}{a_{x_u^{\le n\varepsilon}}}\right)\right]\\ &\le C\,\mathbb{E}^{(n)}\left[\sum_{u\in\mathbb{U}^{\le n\varepsilon}} \left(\frac{a_{x_u^{\le n\varepsilon}}}{a_n}\right)^{q/\gamma}\right].\end{aligned}$$

But we know, thanks to another application of Potter's bounds, that we may find $c > 0$ such that

$$\left(\frac{a_m}{a_n}\right)^{q/\gamma} \le c\left(\frac{m}{n}\right)^{q_*},$$

whenever n is sufficiently large and $m \le n$. Since $x_u^{\le n\varepsilon} \le n$ (for $0 < \varepsilon < 1$), we can again conclude by Corollary 4.5 and Lemma 3.3. □

Acknowledgements The author would like to thank Jean Bertoin for his constant suggestions and guidance and also Bastien Mallein for helpful discussions and comments. Thanks are also directed to an anonymous referee for their careful reading.

References

1. Aldous, D.: The continuum random tree. I. Ann. Probab. **19**(1), 1–28 (1991)
2. Aldous, D.: The Continuum Random Tree. II. An Overview, Stochastic Analysis (Durham, 1990), London Mathematical Society Lecture Note series, vol. 167, pp. 23–70. Cambridge University Press, Cambridge (1991)
3. Aldous, D.: The continuum random tree. III. Ann. Probab. **21**(1), 248–289 (1993)
4. Aspandiiarov, S., Iasnogorodski, R.: General criteria of integrability of functions of passage-times for non-negative stochastic processes and their applications. Teor. Veroyatnost. i Primenen. **43**(3), 509–539 (1998)
5. Athreya, K.B., Ney, P.E.: Branching Processes. Springer, New York (1972). (Die Grundlehren der mathematischen Wissenschaften, Band 196)
6. Bertoin, J.: Markovian growth-fragmentation processes. Bernoulli **23**(2), 1082–1101 (2017)
7. Bertoin, J., Budd, T., Curien, N., Kortchemski, I.: Martingales in self-similar growth-fragmentations and their connections with random planar maps. Probab. Theory Relat. Fields **172**(3–4), 663–724 (2018)
8. Bertoin, J., Curien, N., Kortchemski, I.: Random planar maps and growth-fragmentations. Ann. Probab. **46**(1), 207–260 (2018)
9. Bertoin, J., Kortchemski, I.: Self-similar scaling limits of Markov chains on the positive integers. Ann. Appl. Probab. **26**(4), 2556–2595 (2016)
10. Bertoin, J., Stephenson, R.: Local explosion in self-similar growth-fragmentation processes. Electron. Commun. Probab. **21**(66), 12 (2016)
11. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation, Encyclopedia of Mathematics and Its Applications, vol. 27. Cambridge University Press, Cambridge (1987)
12. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications, Stochastic Modelling and Applied Probability, vol. 38. Springer, Berlin (2010). (corrected reprint of the second (1998) edition)
13. Evans, S.N.: Probability and real trees, lecture notes in mathematics. In: 2008, Lectures from the 35th Summer School on Probability Theory Held in Saint-Flour, July 6–23, vol. 1920, Springer, Berlin (2005)
14. Haas, B., Miermont, G.: The genealogy of self-similar fragmentations with negative index as a continuum random tree. Electron. J. Probab. **9**(4), 57–97 (2004)
15. Haas, B., Miermont, G.: Self-similar scaling limits of non-increasing Markov chains. Bernoulli **17**(4), 1217–1247 (2011)
16. Haas, B., Miermont, G.: Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees. Ann. Probab. **40**(6), 2589–2666 (2012)
17. Haas, B., Miermont, G., Pitman, J., Winkel, M.: Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. Ann. Probab. **36**(5), 1790–1837 (2008)
18. Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288. Springer, Berlin (1987)
19. Kallenberg, O.: Foundations of Modern Probability. Probability and Its Applications (New York), 2nd edn. Springer, New York (2002)
20. Lamperti, J.: Semi-stable Markov processes. I. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **22**, 205–225 (1972)
21. Le Gall, J.-F.: Random real trees. Ann. Fac. Sci. Toulouse Math. (6) **15**(1), 35–62 (2006)
22. Lieb, E.H., Loss, M.: Analysis, Graduate Studies in Mathematics, vol. 14, 2nd edn. American Mathematical Society, Providence (2001)
23. Miermont, G.: Self-similar fragmentations derived from the stable tree. II. Splitting at nodes. Probab. Theory Relat. Fields **131**(3), 341–375 (2005)

24. Rembart, F., Winkel, M.: Recursive construction of continuum random trees. Ann. Probab. **46**(5), 2715–2748 (2018)
25. Shi, Q.: Growth-fragmentation processes and bifurcators. Electron. J. Probab. **22**, 25 (2017)
26. Shi, Z.: Branching random walks, lecture notes in mathematics. In: Lecture Notes from the 42nd Probability Summer School Held in Saint Flour, 2012, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School], vol. 2151, Springer, Cham (2015)