DISCRETE SNAKES WITH GLOBALLY CENTERED DISPLACEMENTS

BY LOUIGI ADDARIO-BERRY[1], SERTE DONDERWINKEL[2], CHRISTINA GOLDSCHMIDT[3] AND RIVKA MITCHELL[4]

[1]*Department of Mathematics and Statistics, McGill University, louigi.addario@mcgill.ca*

[2]*Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, s.a.donderwinkel@rug.nl*

[3]*Department of Statistics and Lady Margaret Hall, University of Oxford, christina.goldschmidt@stats.ox.ac.uk*

[4]*Mathematical Institute, University of Oxford, rivka.mitchell@maths.ox.ac.uk*

We prove a scaling limit for globally centered discrete snakes on size-conditioned critical Bienaymé trees. More specifically, under a global finite variance condition, we prove convergence in the sense of random finite-dimensional distributions of the head of the discrete snake (suitably rescaled) to the head of the Brownian snake driven by a Brownian excursion. When the third moment of the offspring distribution is finite, we further prove uniform functional convergence under a necessary tail condition on the displacements. We also consider displacement distributions with heavier tails, for which we instead obtain convergence to a variant of the hairy snake introduced by Janson and Marckert. We further give two applications of our main result. Firstly, we obtain a scaling limit for the difference between the height process and the Łukasiewicz path of a size-conditioned critical Bienaymé tree. Secondly, we obtain a scaling limit for the difference between the height process of a size-conditioned critical Bienaymé tree and the height process of its associated looptree.

CONTENTS

MSC2020 subject classifications: 60J80, 60F17, 60C05.

Keywords and phrases: Discrete snakes, branching random walks, branching processes, random trees.

1. Introduction. We consider a branching random walk whose genealogy is given by the family tree of a Bienaymé branching process (which we refer to as a *Bienaymé tree*) conditioned to have n vertices. We assume that the offspring distribution $\mu = (\mu_k)_{k\geq 0}$ is critical and has finite, non-zero variance, so that the genealogical tree has the Brownian continuum random tree as its scaling limit.[1] Each vertex of the tree is endowed with a spatial location in $\mathbb{R}$: the root is fixed to be at 0; for every other vertex, its location is obtained via a random displacement away from the location of its parent. The random displacements of children of distinct vertices will always be independent but, in general, the displacements of siblings may be dependent and may, moreover, depend on the vertex degree. For a vertex v with k children, the distribution of the vector of displacements from v to its ordered children is denoted by ν_k. In the sequel, $Y_k = (Y_{k,1}, \ldots, Y_{k,k})$ always denotes a random vector with law ν_k. In this paper, we explore conditions on μ and $\nu = (\nu_k)_{k\geq 1}$ such that the whole object converges to a Brownian motion indexed by the Brownian tree.

A convenient formulation is via the notion of a *discrete snake*. We imagine exploring the vertices of the tree one by one in depth-first order (we shall give precise definitions in Section 2 below) and record a list of the spatial locations of the ancestors of the vertex we are currently visiting. In other words, the snake is a process taking values in the set of finite random walk paths (one should imagine it wiggling around as we explore the tree!). In fact, it turns out to be sufficient for many purposes to track the spatial location of the vertex that we are visiting only: this gives the so-called *head of the discrete snake*, which is our primary object of interest. We aim to prove convergence, after an appropriate rescaling, of the head of the discrete snake to the head of the Brownian snake driven by a normalised Brownian excursion (BSBE), first introduced by Le Gall [23, 24]. This is a stochastic process $(\mathbf{e}, \mathbf{r}) = (\mathbf{e}_t, \mathbf{r}_t)_{0\leq t\leq 1}$ taking values in $\mathbb{R}_+ \times \mathbb{R}$, such that $\mathbf{e}$ is a normalised Brownian excursion and, conditionally on $\mathbf{e}$, the second coordinate $\mathbf{r}$ is a centered Gaussian process taking values in $\mathbb{R}$ with covariance function

$$\operatorname{cov}(\mathbf{r}_s, \mathbf{r}_t) = \min_{u\in[s\wedge t, s\vee t]} \mathbf{e}_u. \tag{1.1}$$

Let us give some interpretation. For any pair of vertices in the Brownian tree, encoded by $s, t \in [0, 1]$, having heights $\mathbf{e}_s$ and $\mathbf{e}_t$, the spatial locations along their genealogical paths evolve as a common Brownian motion until their most recent common ancestor (which lies at

[1] To avoid technicalities, we shall also assume that the support of μ has greatest common divisor 1, so that the event that the tree has size n has strictly positive probability for all n large enough.

distance $\min_{u\in[s\wedge t,s\vee t]} \mathbf{e}_u$ from the root) is reached, and they evolve as independent Brownian motions thereafter.

The problem of proving convergence of rescaled discrete snakes to the BSBE has been studied by a number of authors, under a wide range of different conditions on μ and $(\nu_k)_{k\geq 1}$. We shall give a review of the literature after we state our main results.

In order to obtain a Brownian limit for the displacements along a lineage, we require appropriate centering and moment conditions, which we now explain. Let ξ be a random variable with distribution μ and let $\bar{\xi}$ be a size-biased version, that is, having distribution $\bar{\mu} := (\bar{\mu}_k)_{k\geq 1}$, where for all $k \geq 1$,

$$\bar{\mu}_k = \frac{k\mu_k}{\mathbf{E}[\xi]} = k\mu_k.$$

(Recall that the offspring distribution μ is assumed to be critical, so that $\mathbf{E}[\xi] = 1$.)

Conditionally on $\bar{\xi}$, let $Y_{\bar{\xi}} = (Y_{\bar{\xi},1}, \ldots, Y_{\bar{\xi},\bar{\xi}})$ be $\nu_{\bar{\xi}}$-distributed and, independently, let $U_{\bar{\xi}}$ be a Uniform($[\bar{\xi}]$) random variable (where $[m] := \{1, 2, \ldots, m\}$). Then we say that the discrete snake is *globally centered* if

$$\mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\right] = 0.$$

In other words, the expected displacement of a uniform child of a vertex with a size-biased number of offspring is 0. We define the *global variance* to be

$$\beta^2 := \mathbf{E}\left[Y^2_{\bar{\xi},U_{\bar{\xi}}}\right],$$

and will prove our results under the condition that $\beta^2 < \infty$. Since distances in the tree scale as $n^{1/2}$, the spatial displacements along a lineage will then scale as $n^{1/4}$.

1.1. *Main result.* Denote by T_n a Bienaymé tree with offspring distribution μ, conditioned to have n vertices. Write $v(\mathrm{T}_n)$ for the vertex set of T_n and $\partial\mathrm{T}_n$ for its set of leaves. Conditionally given T_n, let $Y = (Y^{(v)}, v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n)$ be independent random vectors, such that if $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ has k children then $Y^{(v)}$ has distribution ν_k. Endow the vertices of T_n with spatial locations using the displacement vectors $Y^{(v)}$ as described above. We call the pair $\mathbf{T}_n = (\mathrm{T}_n, Y)$ a (μ,ν)-*branching random walk conditioned to have n vertices*, or simply a (μ,ν)-branching random walk.

Let $H_n = (H_n(i))_{0\leq i\leq n}$ and $\widetilde{H}_n = (\widetilde{H}_n(i))_{0\leq i\leq 2n}$ be the height and contour processes of T_n, respectively. Let $R_n(i)$ be the spatial location of the i-th vertex visited in a depth-first exploration of $\mathbf{T}_n$. We call the process (H_n, R_n) the *head of the discrete snake* (see Section 2 for a careful description of this). We may alternatively encode the endpoints of the random walk trajectories using the process $(\widetilde{H}_n, \widetilde{R}_n)$, where $\widetilde{R}_n(i)$ is the spatial location of the i-th vertex visited in a contour exploration of $\mathbf{T}_n$. (Compared to (H_n, R_n), this simply revisits some vertices.) We interpolate all of these functions linearly between integer times, which turns H_n and R_n into elements of $\mathbf{C}([0,n],\mathbb{R})$ and turns $\widetilde{H}_n$ and $\widetilde{R}_n$ into elements of $\mathbf{C}([0,2n],\mathbb{R})$.

We use two different notions of convergence for a sequence of random elements $(f_n)_{n\geq 1}$ of $\mathbf{C}([0,1],\mathbb{R})$ such that $f_n(0) = f_n(1) = 0$ for all $n \geq 1$. Let $U_1, U_2, \ldots$ be IID Uniform($[0,1]$) random variables, independent of everything else. For $k \geq 1$, write $U^k_{(1)}, \ldots, U^k_{(k)}$ for the order statistics of $U_1, \ldots, U_k$. For another random element f of $\mathbf{C}([0,1],\mathbb{R})$ such that $f(0) = f(1) = 0$, we say that $f_n \xrightarrow{\mathrm{d}} f$ in the sense of *random finite-dimensional distributions* if, for every $k \geq 1$,

$$(f_n(U^k_{(1)}), \ldots, f_n(U^k_{(k)})) \xrightarrow{\mathrm{d}} (f(U^k_{(1)}), \ldots, f(U^k_{(k)}))$$

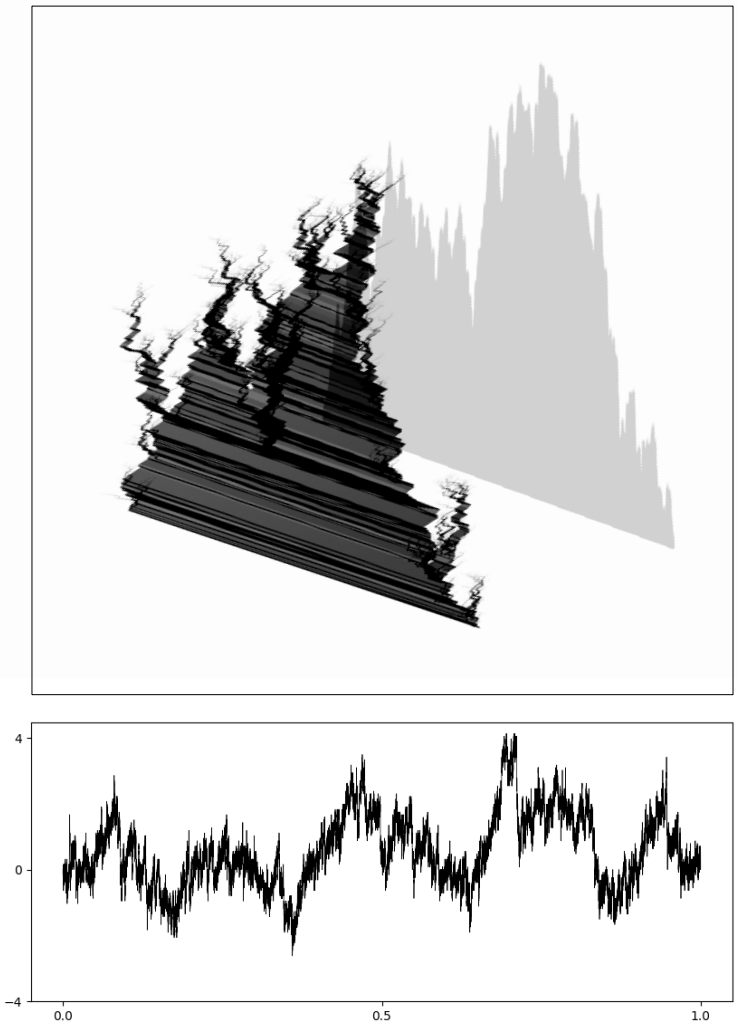

FIG 1. *Top: a (rescaled) discrete snake; bottom: its head. The underlying tree is a size-conditioned Poisson*(1) *Bienaymé tree with* $n = 25000$ *vertices and deterministic displacement distributions given by (1.5), below. The area under the contour process of the underlying tree is illustrated by the gray shaded region.*

as $n \to \infty$. (We will discuss our choice of this notion of convergence in more detail below.) We will also use the stronger notion of convergence with respect to the topology generated by the uniform norm.

THEOREM 1.1. *Let* $\mu = (\mu_k)_{k \geq 0}$ *be a critical offspring distribution with variance* $\sigma^2 \in (0, \infty)$. *If* $\nu = (\nu_k)_{k \geq 1}$ *is such that*

[A1] $$\mathbf{E}\left[Y_{\bar{\xi}, U_{\bar{\xi}}}\right] = 0 \quad \textit{and} \quad \beta^2 = \mathbf{E}\left[(Y_{\bar{\xi}, U_{\bar{\xi}}})^2\right] < \infty,$$

then as $n \to \infty$ the following joint convergence holds in the sense of random finite-dimensional distributions:

(1.2)
$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_n(nt)}{n^{1/4}}, \frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{\widetilde{R}_n(2nt)}{n^{1/4}}\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t, \frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\le t\le 1}.$$

The convergence (1.2) holds in distribution in $\mathbf{C}([0,1],\mathbb{R}^4)$ endowed with the topology of uniform convergence if, additionally,

[A2] $$\mathbf{P}\left\{\max_{1\le i\le \xi} |Y_{\xi,i}| > y\right\} = o(y^{-4}) \text{ as } y\to\infty \text{ and } \mathbf{E}\left[\xi^3\right] < \infty.$$

Theorem 1.1 follows immediately from Corollary 4.2 and Proposition 5.1 below. The analogue of Theorem 1.1 holds with $\mathbb{R}^d$-valued displacements for $d > 1$, and with essentially identical proofs to those in the current work; the only change in the conclusion is that the limit $\mathbf{r}$ of the rescaled spatial displacements takes values in $\mathbb{R}^d$ rather than $\mathbb{R}$, and that β^2 should be interpreted as the covariance matrix of $Y_{\bar{\xi},U_{\bar{\xi}}}$.

Let $\Phi_n(i)$ be the random walk trajectory associated with path from the root to the i-th vertex visited in the contour exploration of $\mathbf{T}_n$, for $0 \le i \le 2n$. Then $(\widetilde{H}_n, \Phi_n)$ is the *discrete snake driven by* $\widetilde{H}_n$. By the homeomorphism theorem of Marckert and Mokkadem (Theorem 2.1 of [31]), Theorem 1.1 entails also that $(\widetilde{H}_n, \Phi_n)$ has the BSBE as its scaling limit; see Figure 1 for an illustration.

By [12, Corollary 2.5.1] and [32] it turns out that the convergence of the parametrisations of the head of the snake via the height and contour processes are essentially equivalent. In particular, in order to prove Theorem 1.1, it suffices to show that, under assumption [**A1**], we have

(1.3)
$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_n(nt)}{n^{1/4}}\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\le t\le 1},$$

as $n\to\infty$ in the sense of random finite-dimensional distributions, and in $\mathbf{C}([0,1],\mathbb{R}^2)$ endowed with the topology of uniform convergence under the additional assumption [**A2**].

It is not clear to us whether the requirement that $\mathbf{E}\left[\xi^3\right] < \infty$ in [**A2**] is necessary or just an artefact of our approach to proving tightness. We shall see in the next subsection that the tail condition in [**A2**] *is* necessary.

1.2. *Necessity of the tail condition.* If we adjust assumption [**A2**] to allow for heavier tails, displacements start to appear near the leaves which are not negligible on the scale $n^{1/4}$. In this case, one can no longer expect a continuous limit process. This is a consequence of the following proposition on the largest displacement in the tree, which we prove in Section A.2 in the appendix.

PROPOSITION 1.2. *Let $\mu = (\mu_k)_{k\ge 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$. If $\limsup_{y\to\infty} y^4 \mathbf{P}\{\max_{1\le i\le\xi}|Y_{\xi,i}| > y\} > 0$, then there exists a $\delta > 0$ such that*

$$\limsup_{n\to\infty} \mathbf{P}\left\{\max_{v\in v(\mathrm{T}_n)\setminus\partial \mathrm{T}_n} \max_{j\ge 1} |Y_j^{(v)}| > \delta n^{1/4}\right\} > 0.$$

We may however still obtain a global convergence result on the appropriate scale, under an additional regularity assumption; see [**A3**] below. Since the large displacements from a

vertex with k children need not be independent, in this setting the limit depends on the joint distribution of the displacements from a vertex to its children. For $k \geq 1$ and $j \in [k]$ denote by

$$Y^+_{k,j} := Y_{k,j} \vee 0 \quad \text{and} \quad Y^-_{k,j} := (-Y_{k,j}) \vee 0$$

the positive and negative displacements of the j-th child of a vertex with k children, respectively. Further let $Y^+_k := (Y^+_{k,j})_{j\in[k]}$ and $Y^-_k = (Y^-_{k,j})_{j\in[k]}$.

[A3] Suppose that $\mathbf{E}\left[\xi^3\right] < \infty$. Furthermore, suppose that there exists a Borel measure π on $\mathbb{R}^2_+ \setminus \{(0,0)\}$ such that for any $\varepsilon > 0$, both $\pi(\mathbb{R}_+ \times (\varepsilon, \infty)) < \infty$ and $\pi((\varepsilon,\infty) \times \mathbb{R}_+) < \infty$, and there exists $\eta \in [0,2)$ such that for all Borel sets $A \subset \mathbb{R}^2_+ \setminus \{(0,0)\}$ for which $\pi(\partial A) = 0$, as $r \to \infty$

$$r^{4-\eta}\mathbf{P}\left\{\frac{1}{r}\left(\max_{1\leq i\leq \xi} Y^+_{\xi,i}, \max_{1\leq i\leq \xi} Y^-_{\xi,i}\right) \in A\right\} \to \pi(A).$$

We note the following lemma, whose proof may be found in Section A.2 in the appendix.

LEMMA 1.3. *[A3] implies that the projection of π onto either of its coordinates has no atom in $(0,\infty)$.*

Under assumption [A3] we prove convergence results for the head of the discrete snake in the space of non-empty compact subsets of $[0,1] \times \mathbb{R}$ equipped with the Hausdorff topology. In what follows, for a continuous function $f : [0,1] \to \mathbb{R}$ and a set $S \subset [0,1] \times \mathbb{R}^2_+ \setminus \{(0,0)\}$, write $U(f,S)$ for the union of the graph of f and the vertical line segments $[(t, f(t) - y), (t, f(t) + x)]$ for each $(t,x,y) \in S$. The next theorem relates to the case $\eta = 0$ in [A3].

THEOREM 1.4. *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$, and let $\nu = (\nu_k)_{k\geq 1}$ be such that [A1] holds and [A3] holds for a given measure π with $\eta = 0$. Then, taking Ξ to be a Poisson process on $[0,1] \times \mathbb{R}^2_+ \setminus \{(0,0)\}$ with intensity $dt \otimes \pi(dx,dy)$, we have*

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}\right)_{0\leq t\leq 1}, U\left(\frac{R_n(n\cdot)}{n^{1/4}}, \emptyset\right)\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\leq t\leq 1}, U\left(\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}, \Xi\right)\right), \tag{1.4}$$

as $n \to \infty$, where the convergence in the first coordinate is in $\mathbf{C}([0,1],\mathbb{R})$ endowed with the topology of uniform convergence, and the convergence in the second is in the space of non-empty, compact subsets of $[0,1] \times \mathbb{R}$ endowed with the Hausdorff topology.

We refer to the object on the right-hand side of (1.4) as the *hairy tour*, in keeping with the previous work of Janson and Marckert [18].

When $\eta \in (0,2)$, the large jumps dominate the smaller ones to such an extent that, in the limit, we obtain a pure jump process.

THEOREM 1.5. *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$, and let $\nu = (\nu_k)_{k\geq 1}$ be such that [A1] holds and [A3] holds for a given measure π with $\eta \in (0,2)$. Then, taking Ξ to be a Poisson process on $[0,1] \times \mathbb{R}^2_+ \setminus \{(0,0)\}$ with intensity $dt \otimes \pi(dx,dy)$, we have*

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}\right)_{0\leq t\leq 1}, U\left(\frac{R_n(n\cdot)}{n^{1/(4-\eta)}}, \emptyset\right)\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\leq t\leq 1}, U(0,\Xi)\right),$$

as $n \to \infty$, where the convergence in the first coordinate is in $\mathbf{C}([0,1],\mathbb{R})$ endowed with the topology of uniform convergence, and the convergence in the second is in the space of non-empty, compact subsets of $[0,1] \times \mathbb{R}$ endowed with the Hausdorff topology.

In contrast to Theorem 1.1, in Theorems 1.4 and 1.5 we need the condition $\mathbf{E}\left[\xi^3\right] < \infty$ not just for tightness but also for the convergence of the random finite-dimensional distributions. The reason for this is that we apply a quantitative local central limit theorem which requires a third moment on the offspring distribution. (See Theorem A.3 for the precise statement.)

The fact that we obtain a continuous function decorated by intervals in both Theorems 1.4 and 1.5 is really an artefact of the choice to interpolate R_n linearly between integer times. Indeed, the endpoints of the intervals capture the asymptotic behaviour of the two extremities of the displacements away from vertices, but tell us nothing about how the "point process" of displacements in between behaves. If we instead consider the graph of $\left(\frac{R_n(\lfloor nt \rfloor)}{n^{1/4}}\right)_{0 \le t \le 1}$ in the case where we do not have $\mathbf{P}\left\{\max_{1 \le i \le \xi} |Y_{\xi,i}| > y\right\} = o(y^{-4})$ there are, in fact, many possible behaviours. We will not undertake any sort of exhaustive classification here, but let us give a couple of illustrative examples.

Suppose first that the displacements are simply IID copies of a random variable Y such that, for some Borel measure π on $\mathbb{R}_+ \setminus \{0\}$ such that for any $\varepsilon > 0$, $\pi((\varepsilon,\infty)) < \infty$, we have $r^4\mathbf{P}\left\{Y \in rA\right\} \to \pi(A)$ as $r \to \infty$ for every Borel set $A \subset \mathbb{R} \setminus \{0\}$ such that $\pi(\partial A) = 0$. Then we will not, in the limit, observe two or more $\Theta(n^{1/4})$ displacements away from the same vertex of T_n (nor, indeed, from vertices at distance $o(n^{1/2})$ from one another), and so we just obtain the graph of $\mathbf{r}$ decorated by isolated points which occur as a Poisson process of intensity $dt \otimes \pi(dy)$ on $[0,1] \times \mathbb{R} \setminus \{0\}$.

On the other hand, suppose that we have the following deterministic displacements:

$$Y_{k,j} = \sigma - \frac{2}{\sigma}(k-j) \text{ for } \quad 1 \le j \le k. \tag{1.5}$$

These displacements have a particular significance, which we will discuss in the next subsection. For the moment, let us just observe that it is straightforward to check that they are globally centered and of finite global variance whenever the offspring distribution is critical and admits a finite third moment. Suppose that there exists a Borel measure π_1 on $(0,\infty)$ with $\pi_1((\varepsilon,\infty)) < \infty$ for all $\varepsilon > 0$, such that $r^4\mathbf{P}\left\{\xi \in rA\right\} \to \pi_1(A)$ as $r \to \infty$ for any Borel set $A \subset (0,\infty)$ with $\pi_1(\partial A) = 0$. Then all of the children of a vertex with $\Theta(n^{1/4})$ children will have $\Theta(n^{1/4})$ displacements which are regularly spaced with spacing $2/\sigma$. Again, with high probability, we will not see two vertices of degree $\Theta(n^{1/4})$ within distance $o(n^{1/2})$ in T_n. So in the limit for the graph of $\left(\frac{R_n(\lfloor nt \rfloor)}{n^{1/4}}\right)_{0 \le t \le 1}$ we will see decorations driven by a Poisson process on $[0,1] \times \mathbb{R}_+$ with intensity $dt \otimes \pi(dx)$ such that when we observe a point (t,x) of the Poisson process, we attach the whole interval $[-2x/\sigma, 0]$ to the graph of $\mathbf{r}$ at t.

1.3. *Related work.* As mentioned earlier, versions of the topic studied in this paper have received extensive attention in the literature. One reason for this is that discrete snakes play a crucial role in the study of random planar maps; see [1, 2, 8, 26, 30, 34].

The earliest discrete snake convergence results were proved in models with a fixed offspring distribution. Chassaing and Schaeffer [8] treated the setting of a Geometric$(1/2)$ offspring distribution (which results in uniformly random planar trees) with IID displacements uniform on $\{-1,0,1\}$. Marckert and Mokkadem [31] treated the same offspring distribution, but where the displacements away from a vertex all have the same centered marginal distribution (but may depend on one another) with a $6+\varepsilon$ moment. Gittenberger [13] later generalised these results to critical, finite variance offspring distributions with centered (but not necessarily IID) displacements having finite $8+\varepsilon$ moment.

Work on the IID displacement case culminated in a paper of Janson and Marckert [18] which established the following result.

THEOREM 1.6 ([18], Theorems 1 and 2). *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$ such that μ has a finite exponential moment. For each $k \geq 1$, let ν_k be the law of a vector of k IID copies of a random variable Y with $\mathbf{E}[Y] = 0$ and $\mathbf{E}\left[Y^2\right] = \beta^2 \in (0,\infty)$. Then*

$$\left(\frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{\widetilde{R}_n(2nt)}{n^{1/4}}\right)_{0\leq t\leq 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\leq t\leq 1} \tag{1.6}$$

as $n \to \infty$, in the sense of finite-dimensional distributions. The convergence also holds in distribution in $\mathbf{C}([0,1],\mathbb{R}^2)$ endowed with the topology of uniform convergence if and only if

$$\mathbf{P}\{|Y| > y\} = o(y^{-4}) \text{ as } y \to \infty. \tag{1.7}$$

The finite exponential moment condition on the offspring distribution has subsequently been shown to be unnecessary, and may be weakened to a finite second moment assumption; see, for example, Marzouk [33]. Our Theorem 1.1 recovers this theorem under the additional assumption of a finite third moment for μ (and replacing convergence in the sense of finite-dimensional distributions in the first statement with random finite-dimensional distributions).

Janson and Marckert [18] also considered what happens in some of the "heavy-tailed" cases for which the tail condition $\mathbf{P}\{|Y| > y\} = o(y^{-4})$ fails. In particular, they considered the setting in which

$$\mathbf{P}\{Y \geq y\} \sim a_+ y^{-q}, \quad \mathbf{P}\{Y \leq -y\} \sim a_- y^{-q} \quad \text{as } y \to \infty$$

for some constants $a_+, a_- \geq 0$ and $q \in (2,4]$, and prove analogues of Theorems 1.4 and 1.5 in such cases. They call the limiting object in this setting the *hairy tour*, and the associated snake the *jumping snake*. Their results were an important inspiration for Theorems 1.4 and 1.5.

Marzouk [33] later extended Janson and Marckert's results in [18] to the situation where the offspring distribution is in the domain of attraction of a stable law, and the displacements are IID.

Returning now to non-IID displacements, there are several notions of centering and finite variance which have been imposed in order to obtain convergence to the BSBE. Marckert and Miermont [30] worked under the "local centering" assumption that $\mathbf{E}[Y_{k,j}] = 0$ for all $1 \leq j \leq k$. For multi-type Bienaymé trees, [2] establishes convergence of discrete snakes under assumptions that impose in particular that the displacements away from vertices of each type are centered.

Most closely related to our results is a paper of Marckert [29], which proves the following theorem.

THEOREM 1.7 ([29], Theorem 1). *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with $\mu_0 + \mu_1 < 1$ and with bounded support. Suppose further that $\nu = (\nu_k)_{k\geq 1}$ is such that*

$$\mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\right] = 0 \quad \text{and} \quad \beta^2 = \mathbf{E}\left[(Y_{\bar{\xi},U_{\bar{\xi}}})^2\right] < \infty,$$

and that there exists $p > 4$ such that

$$\sup_{1\leq j\leq k\leq K} \mathbf{E}\left[|Y_{k,j} - \mathbf{E}[Y_{k,j}]|^p\right] < \infty.$$

where μ is supported by $\{0,\dots,K\}$. Then, as $n\to\infty$,

(1.8)
$$\left(\frac{H_n(nt)}{\sqrt{n}},\frac{R_n(nt)}{n^{1/4}},\frac{\widetilde{H}_n(2nt)}{\sqrt{n}},\frac{\widetilde{R}_n(2nt)}{n^{1/4}}\right)_{0\le t\le 1}\xrightarrow{\mathrm{d}}\left(\frac{2}{\sigma}\mathbf{e}_t,\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t,\frac{2}{\sigma}\mathbf{e}_t,\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\le t\le 1}$$

in $\mathbf{C}([0,1],\mathbb{R}^4)$ endowed with the topology of uniform convergence.

The boundedness condition is a necessary requirement of Marckert's proof technique, which is a *tour de force* involving tracking very detailed information about the number of vertices of each possible different degree along a lineage, which converge on appropriate rescaling to a Gaussian field. Our approach removes the boundedness requirement, but we do not obtain such fine information on the limit object.

Finally, we mention a forthcoming work of Duquesne and Rebei [10], which proves limit theorems for snakes whose jumps are centered and sibling-independent and such that the underlying family tree converges to a Lévy tree. Our understanding is that the results and technique of [10] are rather different from those of the current work.

1.4. *A first application.* One nice consequence of Theorem 1.1 is a strengthening of a result of Marckert and Mokkadem [32] concerning the difference between the height process, H_n, and the Łukasiewicz path, here denoted by W_n (and formally defined in Section 2) of T_n. It is proved in [32] that if ξ is critical with variance $\sigma^2\in(0,\infty)$ and has a finite exponential moment, then

$$\left(\frac{\widetilde{H}_n(2nt)}{\sqrt{n}},\frac{H_n(nt)}{\sqrt{n}},\frac{W_n(nt)}{\sqrt{n}}\right)_{0\le t\le 1}\xrightarrow{\mathrm{d}}\left(\frac{2}{\sigma}\mathbf{e}_t,\frac{2}{\sigma}\mathbf{e}_t,\sigma\mathbf{e}_t\right)_{0\le t\le 1}$$

as $n\to\infty$ in $\mathbf{C}([0,1],\mathbb{R}^3)$. (As mentioned after Theorem 1.6, the finite exponential moment condition is unnecessary and may be removed; see Duquesne [11] for this result in the context of trees rather than snakes.)

Moreover, under the same assumptions, [32] establishes that, for any $\varepsilon>0$, there exists $\gamma>0$ such that for $n>0$ sufficiently large

$$\mathbf{P}\left\{\sup_{0\le i\le n}\left|\sigma H_n(i)-2\sigma^{-1}W_n(i)\right|\ge n^{1/4+\varepsilon}\right\}\le\exp\left(-\gamma n^{\varepsilon}\right).$$

It is natural to conjecture that, under suitable conditions, the difference varies precisely on the order of $n^{1/4}$. We are able to prove this conjecture in a large degree of generality. It turns out that the difference $(\sigma H_n(i)-2\sigma^{-1}W_n(i),0\le i\le n)$ evolves precisely as the head of a discrete snake (see Lemma 2.1 for a proof of this fact). The relevant displacements are given by $Y_{k,j}=\sigma-(2/\sigma)(k-j)$; this formula already appeared at (1.5). We have

$$\sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\mathbf{E}\left[Y_{k,j}\right]=\sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\left(\sigma-\frac{2}{\sigma}(k-j)\right)=\sum_{k=1}^{\infty}\mu_k\left(\sigma k-\frac{k(k-1)}{\sigma}\right)=\sigma-\frac{\sigma^2}{\sigma}=0,$$

so that the associated discrete snake is globally centered. Moreover, the global variance is

$$\begin{aligned}\sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\mathbf{E}\left[Y_{k,j}^2\right]&=\sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\left(\sigma-\frac{2}{\sigma}(k-j)\right)^2\\&=\sum_{k=1}^{\infty}\mu_k\left(\sigma^2k-2k(k-1)+\frac{2}{3\sigma^2}k(k-1)(2k-1)\right)\\&=\frac{4}{3\sigma^2}(\mathbf{E}\left[\xi^3\right]-1)-(\sigma^2+2),\end{aligned}$$

which is finite provided that $\mathbf{E}\left[\xi^3\right] < \infty$. Also,

$$\mathbf{P}\left\{\max_{1\le i\le \xi} |Y_{\xi,i}| > y\right\} = \mathbf{P}\left\{\left|\sigma - \frac{2}{\sigma}(\xi-1)\right| \vee \sigma > y\right\} = o(y^{-4})$$

as $y \to \infty$ if and only if $\mathbf{P}\{\xi > y\} = o(y^{-4})$ as $y \to \infty$; moreover, the latter condition implies $\mathbf{E}\left[\xi^3\right] < \infty$. We obtain the following corollary of Theorem 1.1.

COROLLARY 1.8. *Let $\mu = (\mu_k)_{k\ge 1}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$. Let $\beta^2 = \frac{4}{3\sigma^2}(\mathbf{E}\left[\xi^3\right] - 1) - (\sigma^2 + 2)$. Then*

$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{\sigma H_n(nt) - 2\sigma^{-1}W_n(nt)}{n^{1/4}}\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\le t\le 1},$$

as $n \to \infty$ in $\mathbf{C}([0,1],\mathbb{R}^2)$ if and only if $\mathbf{P}\{\xi > y\} = o(y^{-4})$ as $y \to \infty$.

Let us observe that, while Corollary 1.8 concerns the difference between the height process and the Łukasiewicz path, the joint convergence in Theorem 1.1 can be used to prove an analogous result for the difference between the Łukasiewicz path and the contour process encoding of the head of the same discrete snake. (We leave the details of this statement to the reader.)

In the case where ξ is bounded, Marckert's result (Theorem 1.7) applies, so the corollary is new only in the case of unbounded offspring distributions. In an earlier paper [28], Marckert had already observed that the difference between the left and right pathlengths (also known as the *imbalance*) of a size-conditioned Bienaymé tree with offspring distribution $\mu_0 = \mu_2 = 1/2$ converges in distribution after rescaling to $2^{1/4}S$, where $S = \int_0^1 \mathbf{r}_t dt$. We note that such trees are binary, and recall that the left pathlength (resp. right pathlength) of a vertex v is the number of vertices in its ancestral lineage who precede (resp. succeed) their siblings in the lexicographical order. The left (resp. right) pathlength of binary trees is then the sum of the left (resp. right) path lengths over all vertices in the tree. Janson [15] later used the method of moments to give an alternate proof of this convergence in distribution.

It can also be the case that the sequence $(n^{-1/4}\max_{0\le i\le n}|\sigma H_n(i) - 2\sigma^{-1}W_n(i)|)_{n\ge 1}$ is tight *without* converging in distribution to the maximum modulus of the head of the BSBE; indeed, by Theorem 1.4, if $r^4\mathbf{P}\{\xi \in rA\} \to \pi_1(A)$ as $r \to \infty$ for all Borel sets A such that $\pi_1(\partial A) = 0$ and a Borel measure π_1 on $\mathbb{R}_+ \setminus \{0\}$ such that for any $\varepsilon > 0$, $\pi_1((\varepsilon,\infty)) < \infty$, then it is possible to prove that $n^{-1/4}\max_{0\le i\le n}|\sigma H_n(i) - 2\sigma^{-1}W_n(i)|$ converges in distribution to the maximum modulus of the appropriate hairy tour. If, on the other hand, we have $r^{4-\eta}\mathbf{P}\{\xi \in rA\} \to \pi_1(A)$ as $r \to \infty$ for some $\eta \in (0,2)$, then Theorem 1.5 yields the convergence

$$n^{-1/(4-\eta)}\max_{0\le i\le n}|\sigma H_n(i) - 2\sigma^{-1}W_n(i)| \xrightarrow{\mathrm{d}} L,$$

where $\mathbf{P}\{L \le \ell\} = \exp(-\int_\ell^\infty \pi_1(x)dx)$ is the probability that no point of a Poisson point process of intensity $dt\pi_1(dx)$ on $[0,1]\times\mathbb{R}_+$ has second co-ordinate greater than ℓ.

1.5. *A second application.* A second consequence of Theorem 1.1 concerns the difference between the height process of T_n and the height process of the corresponding *looptree*. The looptree corresponding to T_n, denoted by T_n°, is the connected multigraph obtained by replacing the edges from a vertex to its children by a cycle going through the parent and all of its children in order (whose length, therefore, equals its number of children plus one). See Figure 2 for an illustration. (It turns out that it is possible to make sense of a continuum analogue of this notion, as proved by Curien and Kortchemski [9].)

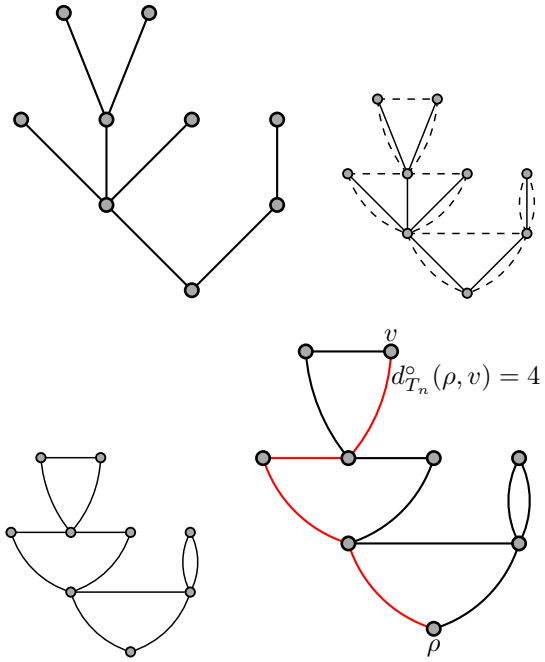

FIG 2. *In the top left figure, a tree, and in the bottom left figure its corresponding looptree. The top-right figure serves to aid in understanding the construction, and the bottom-right figure illustrates how distances are calculated in the loop-tree.*

Vertices in the original tree naturally correspond to vertices in the looptree. Let $v_1, \dots, v_n$ be the vertices of T_n listed in lexicographical order. We define the height function of the looptree, denoted $H^\circ_n : [0, n] \to \mathbb{R}$, to give the graph distance between the root and each of the vertices in the looptree, visited in the order $v_1, \dots, v_n$. This is the height process (in the usual sense) of the spanning tree of the looptree made up of the union of the geodesic paths from each of its vertices to the root. Formally, using the Ulam–Harris notation (see Section 2 for details) for $0 \le i \le n-1$ let

$$H^\circ_n(i) := \sum_{(u,uj) \in e(\mathrm{T}_n) \,:\, uj \preceq v_{i+1}} \min\{j, c(u, \mathrm{T}_n) + 1 - j\},$$

where for $u \in v(\mathrm{T}_n)$, $c(u, \mathrm{T}_n)$ denotes the number of children of u in T_n. Finally let $H^\circ_n(n) = 0$, and extend the domain to $[0, n]$ by linear interpolation. For $c \in \mathbb{R}$, it is readily seen that the difference $(cH_n(i) - H^\circ_n(i), 0 \le i \le n)$ evolves as the head of a discrete snake whose displacements are given by

$$Y_{k,j} = c - \min\{j, k + 1 - j\}.$$

Moreover, if we fix $c = \frac{1}{4}\mathbf{E}\left[\xi^3\right] + \frac{1}{2} + \frac{1}{4}\mathbf{P}\left\{\xi \in 2\mathbb{Z} + 1\right\}$, then

$$\begin{aligned}
\sum_{k=1}^{\infty} \mu_k \sum_{j=1}^{k} \mathbf{E}\left[Y_{k,j}\right] &= \sum_{k=1}^{\infty} \mu_k \sum_{j=1}^{k} \left(c - \min\{j, k+1-j\}\right) \\
&= \sum_{k=1}^{\infty} \mu_k \left(ck - 2 \sum_{i=1}^{\lfloor k/2 \rfloor} i - \left\lceil \frac{k}{2} \right\rceil \mathbf{1}_{[k \in 2\mathbb{Z}+1]} \right) \\
&= \sum_{k=1}^{\infty} \mu_k \left(ck - \frac{k}{2}\left(\frac{k}{2} + 1\right) - \frac{1}{4}\mathbf{1}_{[k \in 2\mathbb{Z}+1]} \right)
\end{aligned}$$

$$= c - \frac{1}{4}\mathbf{E}\left[\xi^3\right] - \frac{1}{2} - \frac{1}{4}\mathbf{P}\left\{\xi \in 2\mathbb{Z}+1\right\} = 0,$$

so that the associated discrete snake is globally centered. Moreover, the global variance is

(1.9)

$$\begin{aligned}
&\sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\mathbf{E}\left[Y_{k,j}\right] \\
&\quad= \sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\left(c-\min\{j,k+1-j\}\right)^2 \\
&\quad= \sum_{k=1}^{\infty}\mu_k\left(c^2k - ck\left(\frac{k}{2}+1\right) - \frac{c}{2}\mathbf{1}_{[k\in 2\mathbb{Z}+1]} + 2\sum_{i=1}^{\lfloor k/2\rfloor} i^2 + \left\lceil\frac{k}{2}\right\rceil^2\mathbf{1}_{[k\in 2\mathbb{Z}+1]}\right) \\
&\quad= \sum_{k=1}^{\infty}\mu_k\left(c^2k + k\left(\frac{k}{2}+1\right)\left(\frac{k}{6}+\frac{1}{6}-c\right) + \left(\frac{k^2}{12}+\frac{k}{3}+\frac{1}{4}-\frac{c}{2}\right)\mathbf{1}_{[k\in 2\mathbb{Z}+1]}\right) \\
&\quad= c^2 + \frac{\mathbf{E}\left[\xi^3\right]}{12} + \left(\frac{1}{4}-\frac{c}{2}\right)\mathbf{E}\left[\xi^2\right] + \frac{1}{6} - c + \mathbf{E}\left[\left(\frac{\xi^2}{12}+\frac{\xi}{3}+\frac{1}{4}-\frac{c}{2}\right)\mathbf{1}_{[\xi\in 2\mathbb{Z}+1]}\right] \\
&\quad=: \beta^2,
\end{aligned}$$

which is finite provided $\mathbf{E}\left[\xi^3\right] < \infty$.

Finally,

$$\mathbf{P}\left\{\max_{1\le i\le\xi}|Y_{\xi,i}| > y\right\} = \mathbf{P}\left\{|c-\lceil\xi/2\rceil| \vee |c-1| > y\right\} = o(y^{-4})$$

as $y\to\infty$ if and only if $\mathbf{P}\left\{\xi > y\right\} = o(y^{-4})$ as $y\to\infty$ and, moreover, the latter condition implies $\mathbf{E}\left[\xi^3\right] < \infty$. We obtain the following corollary of Theorem 1.1.

COROLLARY 1.9. *Let $\mu = (\mu_k)_{k\ge1}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$, and let ξ be a random variable with distribution μ. Let $c = \frac{1}{4}\mathbf{E}\left[\xi^2\right] + \frac{1}{2} + \frac{1}{4}\mathbf{P}\left\{\xi\in 2\mathbb{Z}+1\right\}$ and β^2 be as in (1.9). Then,*

$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{H_n^\circ(nt)}{\sqrt{n}}, \frac{cH_n(nt)-H_n^\circ(nt)}{n^{1/4}}\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \frac{2}{\sigma}\mathbf{e}_t, \sqrt{\frac{2}{\sigma}}\beta\mathbf{r}_t\right)_{0\le t\le 1}, \tag{1.10}$$

as $n\to\infty$ in $\mathbf{C}([0,1],\mathbb{R}^3)$ endowed with the topology of uniform convergence if and only if $\mathbf{P}\left\{\xi > y\right\} = o(y^{-4})$ as $y\to\infty$.

Analogues of this result also hold in the settings of Theorems 1.4 and 1.5. Even the functional convergence for the height process of looptrees of Bienaymé trees, expressed in the second coordinate of (1.10), is new, although pointwise convergence was proved by Kortchemski and Marzouk [21]. (The convergence of looptrees in the Gromov–Hausdorff topology was proved in [22] via spinal decomposition — see Theorem 1.2 and the generic case in Corollary 1.4 for the application to maps — and convergence in the Gromov–Hausdorff–Prokhorov topology was shown in [20, Theorem 15].)

1.6. *Overview of the proofs.* We will prove weak convergence of the head of the snake by making use of the following variant of the usual formulation of weak convergence for a sequence of random continuous functions. (This formulation is inspired by Theorem 20 of [4], and can be proved by essentially the same method as the second proof of Theorem 7.5 of [6].)

PROPOSITION 1.10. *Let $(f_n)_{n\geq 1}$ and f be random elements of $\mathbf{C}([0,1],\mathbb{R})$ such that $f_n(0) = f_n(1) = 0$ for every $n \geq 1$ and $f(0) = f(1) = 0$. Let $U_1, U_2, \ldots$ be* IID *$U[0,1]$ random variables, independent of $(f_n)_{n\geq 1}$ and f. For $k \geq 1$, write $U^k_{(1)}, U^k_{(2)}, \ldots, U^k_{(k)}$ for the values of $U_1, U_2, \ldots, U_k$ written in increasing order, and set $U^k_{(0)} = 0$ and $U^k_{(k+1)} = 1$.*

Suppose that for each $k \geq 1$ we have

$$\left(f_n(U^k_{(1)}), \ldots, f_n(U^k_{(k)})\right) \xrightarrow{\mathrm{d}} \left(f(U^k_{(1)}), \ldots, f(U^k_{(k)})\right), \tag{1.11}$$

as $n \to \infty$, and that for any $\varepsilon > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{\max_{0\leq i\leq k} \sup_{s,t\in[U^k_{(i)},U^k_{(i+1)}]} |f_n(s) - f_n(t)| > \varepsilon\right\} = 0. \tag{1.12}$$

Then $f_n \xrightarrow{\mathrm{d}} f$ as $n \to \infty$, for the topology generated by the uniform norm on $\mathbf{C}([0,1],\mathbb{R})$.

We will refer to assumption (1.11) as the *convergence of random finite-dimensional distributions* and to (1.12) as *tightness*. Observe that (1.11) is weaker than the usual convergence of finite-dimensional distributions. However, it is more natural in the context of random trees, and indeed plays a key role in Aldous' theory of continuum random trees as developed in [4]. (See the appendix of [5] for a discussion and for further references.)

Let $\mathcal{T}_{2\mathbf{e}}$ be the real tree encoded by $2\mathbf{e}$, where $\mathbf{e}$ is a normalised Brownian excursion. (We refer to the survey of Le Gall [25] for standard definitions concerning random real trees.) Fix $k \geq 1$ and let $U_1, \ldots, U_k$ be IID Uniform$([0,1])$ random variables. Furthermore, let $\mathcal{T}^k_{2\mathbf{e}}$ be the subtree of $\mathcal{T}_{2\mathbf{e}}$ spanned by the images of 0 and of $U_1, \ldots, U_k$ in $\mathcal{T}_{2\mathbf{e}}$. Formally, it is useful to think of this as an ordered rooted tree with leaves labeled by $1, 2, \ldots, k$ and edge-lengths, where we use the relative ordering of $U_1, U_2, \ldots, U_k$ to determine the planar ordering of the leaves. Using Aldous' line-breaking construction [4], we may construct a tree which is equal in distribution to $\mathcal{T}^k_{2\mathbf{e}}$ as follows.

Let $J_1, \ldots, J_k$ be the first k jump times of a Poisson point process on $[0,\infty)$ with intensity $t\,dt$ at time t. For $i = 1, \ldots, k-1$, sample an *attachment point* $A_i \sim \mathrm{Uniform}([0, J_i])$, independent of $(A_j)_{j\neq i}$. Take the completion of each of the line segments $[0, J_1], (J_1, J_2], \ldots, (J_{k-1}, J_k]$, and for each $i \in \{1, \ldots, k-1\}$ let J^*_i denote the limit point as $x \downarrow J_{i-1}$. Identify the points J^*_i and A_i, and think of the line-segment as being attached to the left side of the branch containing A_i with probability $1/2$ and to the right side with probability $1/2$. Denote the resulting rooted ordered tree with leaf-labels and edge-lengths by $\mathcal{T}^k$. Then, $\mathcal{T}^k_{2\mathbf{e}} \overset{\mathrm{d}}{=} \mathcal{T}^k$; see [4, p. 279].

The proof of Theorem 1.1 (and similarly Theorems 1.4 and 1.5) relies on proving that a certain discrete line-breaking construction of T_n, described formally in Section 2.2, converges to Aldous' line-breaking construction upon rescaling. The discrete construction builds a tree on $[n]$ by first constructing paths $P^{(1)}, \ldots, P^{(\ell^*)}$ and then attaching them to one another by identifying one endpoint of each path $P^{(i)}$ with a point in $(P^{(j)})_{j<i}$. The proof of convergence of the random finite-dimensional distributions relies on the observation that, along each path, the sequence of partial sums of the displacements is essentially a random walk

trajectory with IID steps with the same distribution as $Y_{\bar{\xi},U_{\bar{\xi}}}$ and, moreover, that random displacements appearing at branch points do not contribute to the displacements of the discrete snake on the "macroscopic" spatial scale of $\Theta(n^{1/4})$.

For the proofs of tightness, we adapt a method of Haas and Miermont [34] used to prove tightness for the height process of a Markov branching tree. (Note that size-conditioned Bienaymé trees are examples of Markov branching trees.) Let T_n^k be a subtree of T_n spanned by its root and k uniform vertices. The difference $\mathrm{T}_n \setminus \mathrm{T}_n^k$ is a forest F_n^k, and to prove tightness we bound the maximum modulus of the spatial locations in each tree in F_n^k. Following Haas and Miermont, we reduce this bound to an expression involving only a size-biased pick from among the trees in F_n^k. The proof of tightness then reduces to proving an explicit tail bound for the maximum modulus of the spatial location of a vertex in T_n when rescaled by $n^{-1/4}$. As a key part of our argument, we require a strong control on the total variation distance between the laws of $\bar{\xi}$ and of the number of children of the root of T_n, which we denote by $\widehat{D}_1^n$. For $k \in [n]$, by Kemperman's formula [37, Chapter 6],

$$\mathbf{P}\left\{\widehat{D}_1^n = k\right\} = \left(\frac{n}{n-1}\right) \frac{\mathbf{P}\left\{S_{n-1} = n-1-k\right\}}{\mathbf{P}\left\{S_n = n-1\right\}} \mathbf{P}\left\{\bar{\xi} = k\right\}, \tag{1.13}$$

where $(S_n)_{n\geq 1}$ is a random walk with IID μ-distributed increments. In order to control this total variation distance, we use a version of the local central limit theorem ([35, Theorem 13, Chapter VII] which, for completeness, we also state below in Theorem A.2) which holds whenever $\mathbf{E}\left[\xi^3\right] < \infty$; this is the origin of the third moment condition in our main theorem.

1.7. *Asymptotic notation.* We will use the following notation related to the asymptotics of random variables $(X_n)_{n\geq} \in \mathbb{R}$. (See Janson [16].) For $(y_n)_{n\geq 1} \in \mathbb{R}_{>0}$,

- $X_n = o_{\mathbf{P}}(y_n)$ means that $X_n/y_n \xrightarrow{\mathrm{p}} 0$ as $n \to \infty$;
- $X_n = \omega_{\mathbf{P}}(y_n)$ means that $X_n/y_n \xrightarrow{\mathrm{p}} \infty$ as $n \to \infty$;
- $X_n = O_{\mathbf{P}}(y_n)$ means that for all $\varepsilon > 0$, there exist constants n_ε, $C_\varepsilon > 0$ such that for all $n \geq n_\varepsilon$, $\mathbf{P}\left\{X_n \leq C_\varepsilon y_n\right\} \geq 1-\varepsilon$;
- $X_n = \Omega_{\mathbf{P}}(y_n)$ means that for all $\varepsilon > 0$ there exist constants n_ε, $C_\varepsilon > 0$ such that for all $n \geq n_\varepsilon$, $\mathbf{P}\left\{X_n \geq C_\varepsilon y_n\right\} \geq 1-\varepsilon$;
- $X_n = \Theta_{\mathbf{P}}(y_n)$ means that $X_n = O_{\mathbf{P}}(y_n)$ and $X_n = \Omega_{\mathbf{P}}(y_n)$;
- Lastly, "with high probability" always means "with probability tending to 1 as $n \to \infty$.

2. Trees, branching random walks, and their encodings. We require a number of different tree models, which we now define.

First, a *tree* is simply a connected acyclic graph $T = (v(T), e(T))$. A *rooted tree* consists of a tree together with a distinguished root vertex $\rho = \rho(T) \in v(T)$. Given a rooted tree T and a vertex v of T write $C(v,T)$ for the set of children of v in T and $c(v,T) = |C(v,T)|$; vertex v is a *leaf* of T if $c(v,T) = 0$. We write ∂T for the set of leaves of T. Also, for a non-root vertex v we write $p(v) = p(v,T)$ for the parent of v in T. For vertices $v, w \in v(T)$ we write $v \prec w$ if v is an ancestor of w, and for an edge e we also write $e \prec v$ if at least one endpoint of e is an ancestor of v. For $S \subset v(T)$, the *subtree of T spanned by S* is the minimal subtree of T containing all elements of S.

Letting $\mathbb{N}^0 := \{\emptyset\}$, the *Ulam–Harris tree* is the rooted tree with root $\emptyset$ and vertex set

$$\mathcal{U} := \bigcup_{n\geq 0} \mathbb{N}^n$$

in which, for each $v \in \mathcal{U}$, the set of children of v is $\{vi, i \in \mathbb{N}\}$. (Here, and in the sequel, for a string $v = (v_1, \ldots, v_k)$ we write $vi := (v_1, \ldots, v_k, i)$.) We say w is a younger sibling of u

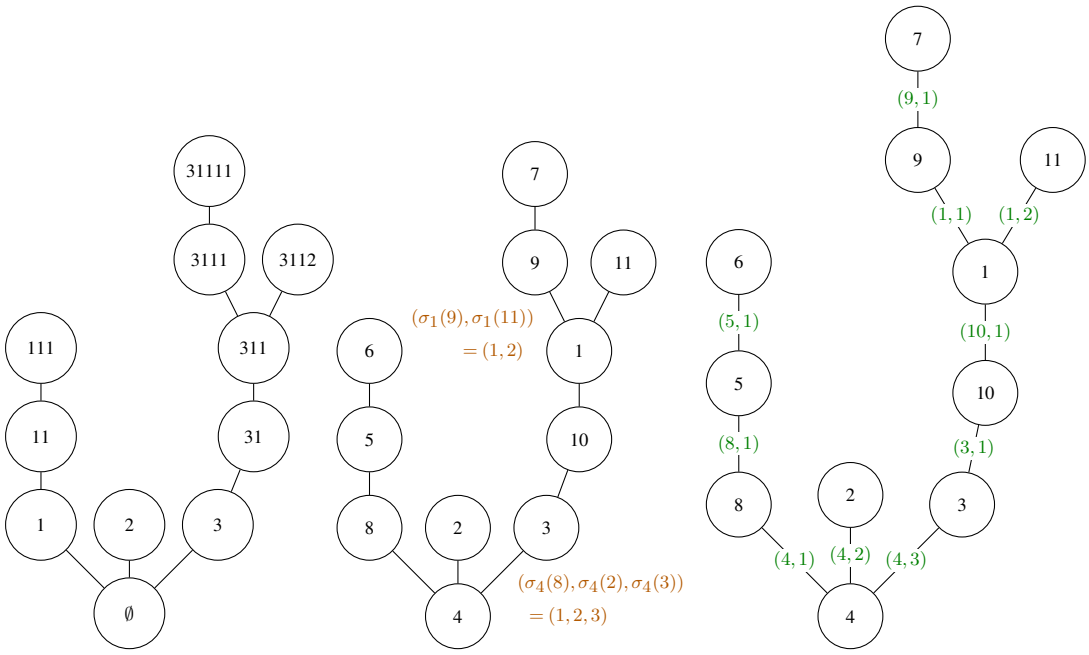

FIG 3. *Left: an ordered rooted tree. Center: a labeled ordered rooted tree, with the functions σ_v indicated for $v \in \{1,4\}$. Right: the edge labeling of T, introduced in Section 2.2.*

if $w = vj$, $u = vi$ and $j > i$. We will make use of the usual lexicographic order on $\mathcal{U}$, which is the total order in which each vertex precedes all of its descendants and all of its younger siblings. Also, for $v \in \mathbb{N}^n \subset \mathcal{U}$ we write $|v| = n$ for the depth of v in $\mathcal{U}$.

The definitions of the coming paragraph are illustrated in Figure 3. An *ordered rooted tree* is a tree T with $v(T) \subset \mathcal{U}$ and the following properties: (i) $\emptyset \in v(T)$; (ii) if $v \in v(T)$ then $p(v,\mathcal{U}) \in v(T)$; (iii) if $vi \in v(T)$ then $vj \in v(T)$ for all $1 \leq j \leq i$. Note that the edge set of an ordered rooted tree may be recovered from its vertex set, and we will often identify ordered rooted trees with their vertex sets. The lexicographic order on $v(T)$ is simply the restriction of the lexicographic order on $\mathcal{U}$ to $v(T)$.

A *labeled ordered rooted tree* is a finite rooted tree $T = (v(T), e(T))$ with $v(T) = [n]$ in which, for each non-leaf vertex of T, the set of children is endowed with a total order $\sigma_v = \sigma_{v,T} : C(v,T) \to [c(v,T)]$. We will sometimes abuse notation by writing $vi = \sigma_v^{-1}(i)$ for the i-th child of v under this total order. This abuse of notation is justified by the observation that the ordering of the children of each non-leaf induces an injection $\varphi : v(T) \to \mathcal{U}$ defined inductively by $\varphi(\rho(T)) = \emptyset$ and $\varphi(vi) = \varphi(\sigma_v^{-1}(i)) = \varphi(v)i$ for $i \in [c(v,T)]$; and $\varphi(v(T))$ is indeed (the vertex set of) an ordered rooted tree. As such, a labeled ordered rooted tree could equivalently be represented as a pair (T, f) where $T \subset \mathcal{U}$ is a finite ordered rooted tree and $f : T \to [n]$ is a bijection (so $n = |T|$). However, the first representation is more natural in the context of the methods we shall shortly use for constructing random labeled ordered rooted trees. Moreover, the second representation would be confusing, as it is very similar to our representations of *branching random walks* and of *spatial trees*, which we now describe.

2.1. *Branching random walks, Łukasiewicz path, contour and height processes.* A *branching random walk* is a pair $\mathrm{T} = (T, Y)$, where T is an ordered rooted tree (possibly labeled) and $Y = (Y^{(v)}, v \in v(T) \setminus \partial T)$, where $Y^{(v)} = (Y_j^{(v)}, j \in [c(v,T)]) \in \mathbb{R}^{c(v,T)}$. We

think of $Y^{(v)}$ as a set of spatial displacements from vertex v to its children, so $Y_j^{(v)}$ is the difference in the spatial locations of vertices v and vj. The spatial *location* of $u \in v(T)$ is then given by the sum of displacements along u's ancestral path:

$$\ell(u) = \ell(u, \mathrm{T}) := \sum_{\{(v,vj) \in e(T) : vj \preceq u\}} Y_j^{(v)}.$$

We refer to the pair (T, ℓ) as a *spatial tree*. The branching random walk (T, Y) can clearly be recovered from the spatial tree (T, ℓ), and vice versa.

Let $T = (v(T), e(T))$ be a finite ordered rooted tree and write $n = |T|$. The *Łukasiewicz path* of T is the function $W_T : [0, n] \to \mathbb{R}$ defined as follows. List the elements of $v(T)$ in lexicographic order as $v_1, \ldots, v_n$. Set $W_T(0) = 0$. For $1 \le i \le n$, set $W_T(i) = \sum_{j=1}^{i} (c(v_i, T) - 1)$, and then extend the domain of W_T to $[0, n]$ by linear interpolation.

The *height process* of T is the function $H_T : [0, n] \to \mathbb{R}_{\ge 0}$ defined as follows. For $0 \le i < n$ set $H_T(i) = |v_{i+1}|$ and set $H_T(n) = 0$; then extend the domain of H_T to $[0, n]$ by linear interpolation.

The *contour order* of $v(T)$ is the sequence $w_0, \ldots, w_{2(n-1)}$ of elements of $v(T)$ defined as follows. First, $w_0 = \emptyset$ is the root of T. Inductively, for each $0 \le i < 2(n-1)$, if w_i has at least one child in T which does not appear in the sequence $w_0, \ldots, w_{i-1}$, then let w_{i+1} be the lexicographically least such child. Otherwise, let $w_{i+1} = p(w_i, T)$. It is straightforward to verify that each vertex v of T appears in the resulting sequence exactly $1 + c(v, T)$ times. The *contour process* of T is the function $\widetilde{H}_T : [0, 2(n-1)] \to \mathbb{R}_{\ge 0}$ defined by setting $\widetilde{H}_T(i) = |w_i|$ for integers i with $0 \le i \le 2(n-1)$, letting $\widetilde{H}_T(2n) = 0$, and extending to $[0, 2n]$ by linear interpolation.

If $\mathrm{T} = (T, Y)$ is a branching random walk with underlying tree T then we encode the spatial locations by a function $R_{\mathrm{T}} : [0, n] \to \mathbb{R}$ given by setting $R_{\mathrm{T}}(i) = \ell(v_{i+1}, \mathrm{T})$, for $i \in \{0, \ldots, n-1\}$, $R_{\mathrm{T}}(n) = 0$, and extending to $[0, n]$ by linear interpolation. We also define a process $\widetilde{R}_{\mathrm{T}} : [0, 2n] \to \mathbb{R}$ by setting $\widetilde{R}_{\mathrm{T}}(i) = \ell(w_i, \mathrm{T})$ for integers i with $0 \le i \le 2(n-1)$, $\widetilde{R}_{\mathrm{T}}(2n) = 0$, and extending to $[0, 2n]$ by linear interpolation.

The following result appears somewhat implicitly in Section 3 of [7]. For completeness we give a proof.

LEMMA 2.1. *Fix $\alpha_1, \alpha_2 \ne 0$. Let $\mathrm{T} = (T, Y)$ be the branching random walk with $Y = (Y^{(v)}, v \in v(T) \setminus \partial T)$ such that $Y^{(v)} = (\alpha_1 - \frac{2}{\alpha_2}(c(v,T) - j), j \in [c(v,T)])$, $v \in v(T)$. Let R_{T} be the function encoding the spatial locations of T. Then for all $t \in [0, n]$,*

$$R_{\mathrm{T}}(t) = \alpha_1 H_T(t) - \frac{2}{\alpha_2} W_T(t).$$

PROOF. It is sufficient to prove that

$$R_{\mathrm{T}}(i) = \alpha_1 H_T(i) - \frac{2}{\alpha_2} W_T(i)$$

for $i \in \{0, 1, \ldots, n-1\}$. Let $i \in \{0, \ldots, n-1\}$. Then $H_T(i)$ is the number of ancestors of v_{i+1} in T. Further, $W_T(i)$ is the number of younger siblings of ancestors of v_{i+1}. It follows

that

$$\alpha_1 H_T(i) - \frac{2}{\alpha_2} W_T(i) = \alpha_1 \cdot \left(\sum_{(u,uj) \in e(T): uj \preceq v_{i+1}} 1 \right) - \frac{2}{\alpha_2} \sum_{(u,uj) \in e(T): uj \preceq v_{i+1}} (c(u,T) - j)$$

$$= \sum_{(u,uj) \in e(T): uj \preceq v_{i+1}} \left(\alpha_1 - \frac{2}{\alpha_2} (c(u,T) - j) \right)$$

$$= \sum_{(u,uj) \in e(T): uj \preceq v_{i+1}} Y_j^{(u)} = R_{\mathrm{T}}(i). \qquad \square$$

2.2. *Sequential encodings of labeled ordered rooted trees.* Given a labeled ordered rooted tree $T = ([n], e(T))$, we assign labels to the *edges* of T as follows. For $v \in [n]$ and $i \in [c(v,T)]$, assign label (v,i) to the edge $\{v, vi\} = \{v, \sigma_v^{-1}(i)\}$. The set of all edge labels is then $L(T) = \{(v,i) : v \in v(T), i \in [c(v,T)]\}$. Given any path $P = v_0 v_1 \dots v_k$ from a vertex v_0 of T to one of its descendants, let π_P be the sequence of edge labels along the path from v_0 to v_k: formally, $\pi_P = \pi_P(T) = ((v_0, c_0), \dots, (v_{k-1}, c_{k-1}))$, where $c_0, \dots, c_{k-1}$ are such that $v_j = v_{j-1} c_{j-1}$ for each $j \in [k]$.

We say a sequence $\mathrm{d} = (d_1, \dots, d_n)$ of non-negative integers is a *degree sequence* if $\sum_{v \in [n]} d_v = n - 1$. We say a labeled tree T with $v(T) = [n]$ has degree sequence d if $c(v,T) = d_v$ for all $v \in [n]$. Write $\mathcal{L}_{\mathrm{d}}$ for the set of labeled ordered rooted trees with degree sequence d. For any tree $T \in \mathcal{L}_{\mathrm{d}}$, it is the case that $L(T) = \{(v,c) : v \in [n], c \in [d_v]\}$. Write $\mathcal{P}_{\mathrm{d}}$ for the set of permutations of $\{(v,c) : v \in [n], c \in [d_v]\}$; this set has size $(n-1)!$. For a fixed degree sequence d, we will make extensive use of a bijection $B : \mathcal{P}_{\mathrm{d}} \to \mathcal{L}_{\mathrm{d}}$ for $\mathrm{d} = (d_1, \dots, d_n)$ a degree sequence of length $n \geq 2$, which we give below. We first describe B^{-1}, as it is slightly simpler.

The bijection $B^{-1} : \mathcal{L}_{\mathrm{d}} \to \mathcal{P}_{\mathrm{d}}$. **Input**: $T \in \mathcal{L}_{\mathrm{d}}$.

- Let $T^{(0)}$ be the subtree of T consisting of the root alone.
- For $\ell \geq 1$, if $T^{(\ell-1)} \neq T$ then let $y^{(\ell)}$ be the smallest label of a vertex in T which is not in $T^{(\ell-1)}$, let $P^{(\ell)}$ be the path in T from $T^{(\ell-1)}$ to $y^{(\ell)}$, and let $T^{(\ell)}$ be the subtree of T spanned by $\{P^{(\ell)}, y^{(1)}, \dots, y^{(\ell-1)}\}$.
- Let ℓ^* be the first value for which $T^{(\ell^*)} = T$.
- Let π_T be the concatenation of the sequences $\pi_{P^{(1)}}, \dots, \pi_{P^{(\ell^*)}}$, and set $B^{-1}(T) = \pi_T$.

In the example of Figure 3, $\ell^* = 6$ and the paths are $P^{(1)} = 4, 3, 10$, $P^{(2)} = 4$, $P^{(3)} = 4, 8$, $P^{(4)} = 5$, $P^{(5)} = 1, 9$ and $P^{(6)} = 1$, so

$$\pi_T = ((4,3), (3,1), (10,1), (4,2), (4,1), (8,1), (5,1), (1,1), (9,1), (1,2)). \tag{2.1}$$

We next describe B; for this we make use of the fact that to specify a labeled ordered rooted tree T with vertex set $[n]$ it suffices to specify the set $C(v,t)$ and the total orderings $\sigma_v : C(v,T) \to [c(v,T)]$ for each $v \in [n]$.

Informally, this construction can be thought of as a discrete analog of the continuous line-breaking construction from the second paragraph of Section 1.6. More specifically, given $\pi = ((v_1, c_1), \dots, (v_{n-1}, c_{n-1})) \in \mathcal{P}_{\mathrm{d}}$, certain substrings of π will correspond to paths in the tree $B(\pi)$. We will list these paths as $P^{(1)}, \dots, P^{(\ell^*)}$. As in the continuous line-breaking construction, for each $i \geq 2$ we will identify one endpoint of the path $P^{(i)}$, with a vertex in $(P^{(j)})_{j<i}$. In the following formal description we denote the i-th identified vertex by v_{j_i}.

When we identify the endpoint of path $P^{(i)}$ with vertex v_{j_i}, we use the second coordinate of the pair (v_{j_i}, c_{j_i}) to determine the position of the unique child of v_{j_i} belonging to $P^{(i)}$ among the children of v_{j_i}.

The bijection $B : \mathcal{P}_{\mathrm{d}} \to \mathcal{L}_{\mathrm{d}}$. **Input**: $\pi = ((v_1, c_1), \dots, (v_{n-1}, c_{n-1})) \in \mathcal{P}_{\mathrm{d}}$.

- Set $m_1 = \min\{m \in \mathbb{N} : m \neq v_1\}$ and let
$$j_1 = \inf\{j > 1 : v_j \in \{m_1, v_1, \dots, v_{j-1}\}\} \wedge n\,.$$
- For $i \geq 1$, if $j_i < n$ then:
  - set $m_{i+1} = \min\{m > m_i : m \notin \{v_1, \dots, v_{j_i}\}\}$;
  - let
$$j_{i+1} = \inf\{j > j_i : v_j \in \{m_1, \dots, m_{i+1}, v_1, \dots, v_{j-1}\}\} \wedge n.$$
- Let $\ell^* = \min\{i \geq 1 : j_i = n\}$.
- Define a labeled ordered rooted tree $T \in \mathcal{L}_{\mathrm{d}}$ as follows. For $1 \leq i \leq n-1$, if $i+1 \notin \{j_1, \dots, j_{\ell^*}\}$ then set $v_i c_i = \sigma_{v_i}^{-1}(c_i) := v_{i+1}$. If $i+1 = j_k$ for some $1 \leq k \leq \ell^*$ then set $\sigma_{v_i}^{-1}(c_i) = m_k$.
- Set $B(\pi) = T$.

The rightmost tree in Figure 3 is the tree $B(\pi)$ where π is equal to π_T from (2.1).

When needed, we will emphasise the dependence of the quantities m_i, j_i and ℓ^* on π by writing $m_i(\pi)$, $j_i(\pi)$ and $\ell^*(\pi)$. Setting $j_0 = 1$ for convenience, we may think of $T = B(\pi)$ as the union of the paths $P^{(1)}, \dots, P^{(\ell^*)}$, where $P_i = v_{j_{i-1}} \dots v_{j_i - 1} m_i$ is the path in T from $v_{j_{i-1}}$ to m_i. Note that since $m_i \geq i$ for all $i \in [\ell^*]$, vertices $1, \dots, k$ are contained within the union of paths $P^{(1)}, \dots, P^{(k \wedge \ell^*)}$ for all $k \in [n]$.

Recall from Section 1.1 that T_n denotes a Bienaymé tree with offspring distribution μ conditioned to have n vertices. Suppose now that $D^n = (D_1^n, \dots, D_n^n)$ is a sequence of IID μ-distributed random variables conditioned to have total sum $\sum_{i=1}^n D_i^n = n - 1$, and let $\Pi_{D^n} \in_{\mathcal{U}} \mathcal{P}_{D^n}$. Then the tree T_n has the same law as $B(\Pi_{D^n})$. Furthermore, $\mathbf{T}_n = (\mathrm{T}_n, Y)$, which we refer to as a (μ, ν)-branching random walk (conditioned to have size n), has the same law as $(B(\Pi_{D^n}), Y)$. (Here, conditionally on the underlying tree T, $Y = (Y^{(v)}, v \in v(T) \setminus \partial T)$ are independent random vectors such that if $c(v, T) = k$ then $Y^{(v)}$ has distribution ν_k). The associated spatial tree $(B(\Pi_{D^n}), \ell)$ is such that $\ell(v_1) = 0$, and for $0 \leq i \leq n-1$ if $i + 1 \notin \{j_1, \dots, j_{\ell^*}\}$,

$$\ell(v_{i+1}) = \ell(v_i) + Y_{c_i}^{(v_i)},$$

and if $i + 1 = j_k$ for some $1 \leq k \leq \ell^*$, then

$$\ell(m_k) = \ell(v_i) + Y_{c_i}^{(v_i)}.$$

In Section 3 we study the above bijective construction of uniform trees with a given deterministic degree sequence d; that is, for $T = B(\Pi_{\mathrm{d}})$ for $\Pi_{\mathrm{d}} \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$. We note however, that by conditioning on D^n, all results in Section 3 also apply to T_n and, consequently, to the underlying tree of $\mathbf{T}_n = (\mathrm{T}_n, Y)$.

3. Sampling from $\mathcal{L}_{\mathbf{d}}$. Fix a degree sequence $\mathrm{d} = (d_1, \dots, d_n)$. The bijection B applied to a uniform element $\Pi_{\mathrm{d}} \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$ yields a uniform element $T = B(\Pi_{\mathrm{d}})$ of $\mathcal{L}_{\mathrm{d}}$. We can think of the bijection as constructing T from Π_{d} by adding vertices one at a time in order of their first appearance in a pair (V, C) of Π_{d}. Below, we use this perspective to study properties of T, in particular the law of the sequence of vertices ordered by first appearance in a pair (V, C) of

Π_{d}, and the law of the number of vertices contained in the union of the paths $P^{(1)}, .., P^{(k)}$, for given $k \geq 1$.

3.1. *Size-biased random re-ordering.* For $n \geq 1$ let $\mathcal{S}_n$ denote the set of permutations of $[n]$. For $(k_1, \ldots, k_n) \in \mathbb{N}^n$, let $\Sigma = \Sigma_{(k_1,\ldots,k_n)}$ be the random permutation of $[n]$ with law given by

$$\mathbf{P}\{\Sigma = \sigma\} = \prod_{i=1}^{n} \frac{k_{\sigma(i)}}{\sum_{j=i}^{n} k_{\sigma(j)}}, \quad \text{for } \sigma \in \mathcal{S}_n.$$

We call $(k_{\Sigma(1)}, \ldots, k_{\Sigma(n)})$ the *size-biased random re-ordering* of $(k_1, \ldots, k_n)$.

For a degree sequence d, let $N_{\mathrm{d}} = |\{i \in [n] : d_i > 0\}|$. For $\pi = ((v_1, c_1), \ldots, (v_n, c_n)) \in \mathcal{P}_{\mathrm{d}}$ we let $\hat{v}_1(\pi), \ldots, \hat{v}_{N_{\mathrm{d}}}(\pi)$ denote the internal vertices in $T = B(\pi)$ ordered by their first appearance in a pair (v, c) in π. When $\pi = \Pi_{\mathrm{d}} = ((V_1, C_1), \ldots, (V_{n-1}, C_{n-1})) \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$ is random, we write $\widehat{V}_i(\Pi_{\mathrm{d}}) = \hat{v}_i(\Pi_{\mathrm{d}})$ to reinforce the fact that the order of the vertices is random. The next lemma states that $(\widehat{V}_1(\Pi_{\mathrm{d}}), \ldots, \widehat{V}_{N_{\mathrm{d}}}(\Pi_{\mathrm{d}}))$ are the vertices corresponding to a size-biased random reordering of $\{d_i \,:\, d_i > 0, i \in [n]\}$.

LEMMA 3.1. *Fix a degree sequence* $\mathrm{d} = (d_1, \ldots, d_n)$ *and let* $\Pi_{\mathrm{d}} \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$. *Then for any permutation* $(i_1, \ldots, i_{N_{\mathrm{d}}})$ *of* $\{i \in [n] : d_i > 0\}$,

$$\mathbf{P}\left\{(\widehat{V}_1(\Pi_{\mathrm{d}}), \ldots, \widehat{V}_{N_{\mathrm{d}}}(\Pi_{\mathrm{d}})) = (i_1, \ldots, i_{N_{\mathrm{d}}})\right\} = \frac{d_{i_1}}{n-1} \frac{d_{i_2}}{n-1-d_{i_1}} \cdots \frac{d_{i_{N_{\mathrm{d}}}}}{n-1-\sum_{j=1}^{N_{\mathrm{d}}-1} d_{i_j}}.$$

Consequently, the size-biased random reordering of the positive entries of d *is equal in distribution to* $(d_{\widehat{V}_1(\Pi_{\mathrm{d}})}, \ldots, d_{\widehat{V}_{N_{\mathrm{d}}}(\Pi_{\mathrm{d}})})$.

PROOF. We show the statement by induction on N_{d}. For $N_{\mathrm{d}} = 1$, the statement is immediate for all n and for all degree sequences of length n with $|\{i : d_i > 0\}| = 1$ since if $N_{\mathrm{d}} = 1$ there is a single vertex of positive degree and $\widehat{V}_1(\Pi_{\mathrm{d}}) = i_1$.

Next, fix $\ell \in \mathbb{N}$ and suppose the statement holds for all degree sequences d with $N_{\mathrm{d}} \leq \ell$. Then fix any degree sequence $\mathrm{d} = (d_1, \ldots, d_n)$ with $N_{\mathrm{d}} = \ell + 1$, and any permutation $(i_1, \ldots, i_{N_{\mathrm{d}}})$ of $\{i \in [n] : d_i > 0\}$. To specify an element of $\{\pi \in \mathcal{P}_{\mathrm{d}} : (\hat{v}_1(\pi), \ldots, \hat{v}_{\ell+1}(\pi)) = (i_1, \ldots, i_{\ell+1})\}$, it is necessary and sufficient to specify

1. $\pi_1 = (v_1, c_1) \in \{(i_1, c) : c \in [d_{i_1}]\}$;
2. The $d_{i_1} - 1$ values $j \in \{2, 3, \ldots, n-1\}$ for which $\pi_j = (i_1, c)$ for some $1 \leq c \leq d_{i_1}$;
3. The order of the $d_{i_1} - 1$ elements of $\{(i_1, c), 1 \leq c \leq d_{i_j}\} \backslash \{\pi_1\}$ in π;
4. The order of the elements of $\{(i_j, c), 2 \leq j \leq \ell + 1, 1 \leq c \leq d_{i_j}\}$ in π, which must ensure that $(\hat{v}_2(\pi), \ldots, \hat{v}_{\ell+1}(\pi)) = (i_2, \ldots, i_{\ell+1})$.

By the induction hypothesis applied to the degree sequence $(d_{i_2}, \ldots, d_{i_{\ell+1}}, 0, \ldots, 0) \in \mathbb{Z}_{\geq 0}^{n-d_{i_1}}$ this implies that

$$\begin{aligned} &|\{\pi \in \mathcal{P}_{\mathrm{d}} : (\hat{v}_1(\pi), \ldots, \hat{v}_{\ell+1}(\pi)) = (i_1, \ldots, i_{\ell+1})\}| \\ &\quad = d_{i_1} \binom{n-2}{d_{i_1}-1} (d_{i_1}-1)!(n-1-d_{i_1})! \frac{d_{i_2}}{n-1-d_{i_1}} \cdots \frac{d_{i_{\ell+1}}}{n-1-\sum_{j=1}^{\ell} d_{i_j}} \\ &\quad = (n-1)! \frac{d_{i_1}}{n-1} \frac{d_{i_2}}{n-1-d_{i_1}} \cdots \frac{d_{i_{\ell+1}}}{n-1-\sum_{j=1}^{\ell} d_{i_j}}; \end{aligned} \tag{3.1}$$

since $|\mathcal{P}_{\mathrm{d}}| = (n-1)!$, the claim follows. □

3.2. *Repeats in* Π_{d}. Let $\Pi_{\mathrm{d}} = ((V_1, C_1), \ldots, (V_{n-1}, C_{n-1})) \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$. Recall that $\widehat{V}_1(\Pi_{\mathrm{d}})$, $\ldots$, $\widehat{V}_{N_{\mathrm{d}}}(\Pi_{\mathrm{d}})$ are the internal vertices in $B(\Pi_{\mathrm{d}})$ ordered by their first appearance in a pair (V, C) in Π_{d}.

For $i \in [\ell^*(\Pi_{\mathrm{d}})]$ let $M_i^{\mathrm{d}} = m_i^{\mathrm{d}}(\Pi_{\mathrm{d}})$ and $J_i^{\mathrm{d}} = j_i^{\mathrm{d}}(\Pi_{\mathrm{d}})$. We introduce this notation to emphasise that $M_1^{\mathrm{d}}, \ldots, M_{\ell^*(\Pi_{\mathrm{d}})}^{\mathrm{d}}$ and $J_1^{\mathrm{d}}, \ldots, J_{\ell^*(\Pi_{\mathrm{d}})}^{\mathrm{d}}$ are random variables. We will see later that for the random degree sequences $D^n = (D_1^n, \ldots, D_n^n)$ arising in this paper, for $k \geq 1$ fixed and for n large, $\{V_1, \ldots, V_{J_k^{D^n}}\} \cap [k] = \emptyset$ with high probability. In this case, for each $i \in [k]$ the first coordinate of the pair $(V_{J_i^{D^n}}, C_{J_i^{D^n}}) \in \Pi_{D^n}$, corresponds to a *repeated* first coordinate of Π_{D^n}. It is therefore convenient to define a second set of indices which correspond to the indices of Π_{d} for which the first coordinate is a repeat. Specifically, let $\widetilde{J}_1^{\mathrm{d}} = \inf\{j > 1 : V_j \in \{V_1, \ldots, V_{j-1}\}\} \wedge n$, and for $i \geq 1$, let

$$\widetilde{J}_{i+1}^{\mathrm{d}} = \inf\{j > \widetilde{J}_i^{\mathrm{d}} : V_j \in \{V_1, \ldots, V_{j-1}\}\} \wedge n.$$

The next two lemmas describe the laws of $\widetilde{J}_1^{\mathrm{d}}$ and $(\widetilde{J}_i^{\mathrm{d}}, i \geq 2)$, respectively.

LEMMA 3.2. *Fix an integer* $n \geq 2$ *and a degree sequence* $\mathrm{d} = (d_1, \ldots, d_n)$ *and let* $\Pi_{\mathrm{d}} \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$. *Then for* $1 \leq k \leq N_{\mathrm{d}}$,

$$\mathbf{P}\left\{\widetilde{J}_1^{\mathrm{d}} > k \;\middle|\; \widehat{V}_1(\Pi_{\mathrm{d}}), \ldots, \widehat{V}_{N_d}(\Pi_{\mathrm{d}})\right\} = \prod_{j=1}^{k}\left(1 - \frac{\sum_{i=1}^{j}(d_{\widehat{V}_i(\Pi_{\mathrm{d}})} + 1)}{n - 1 - j}\right),$$

and, for $k > N_{\mathrm{d}}$,

$$\mathbf{P}\left\{\widetilde{J}_1^{\mathrm{d}} > k \;\middle|\; \widehat{V}_1(\Pi_{\mathrm{d}}), \ldots, \widehat{V}_{N_d}(\Pi_{\mathrm{d}})\right\} = 0.$$

PROOF. Observe that $\widetilde{J}_1^{\mathrm{d}} \leq N_{\mathrm{d}} + 1$ deterministically, so the statement for $k > N_{\mathrm{d}}$ is immediate. To prove the statement for $1 \leq k \leq N_{\mathrm{d}}$, fix any ordering $i_1, \ldots, i_{N_{\mathrm{d}}}$ of $\{i \in [n] : d_i > 0\}$. Then using Bayes' formula and the fact that $|\mathcal{P}_{\mathrm{d}}| = (n-1)!$, the probability

$$\mathbf{P}\left\{\widetilde{J}_1^{\mathrm{d}} > k \;\middle|\; (\widehat{V}_1(\Pi_{\mathrm{d}}), \ldots, \widehat{V}_{N_d}(\Pi_{\mathrm{d}})) = (i_1, \ldots, i_{N_{\mathrm{d}}})\right\}$$

may be expressed as a ratio with denominator

$$|\{\pi = ((v_i, c_i), i \in [n]) \in \mathcal{P}_{\mathrm{d}} : (\hat{v}_1(\pi), \ldots, \hat{v}_{N_{\mathrm{d}}}(\pi)) = (i_1, \ldots, i_{N_d})\}|$$

and numerator

$$|\{\pi = ((v_i, c_i), i \in [n]) \in \mathcal{P}_{\mathrm{d}} : (\hat{v}_1(\pi), ..., \hat{v}_{N_{\mathrm{d}}}(\pi)) = (i_1, ..., i_{N_d}), (v_1, ..., v_k) = (i_1, ..., i_k)\}|.$$

Equation (3.1) directly yields a formula for the denominator. Also, letting d' be the degree sequence $(d_{i_k+1}, \ldots, d_{N_d}, 0, \ldots, 0) \in \mathbb{Z}_{\geq 0}^{n-d_{i_1}-\ldots-d_{i_k}}$, then the numerator is

$$\prod_{j=1}^{k} d_{i_j} \cdot (n-1-k)_{d_{i_1}+\ldots+d_{i_k}-k} \cdot |\{\pi' \in \mathcal{P}_{\mathrm{d}'} : (\hat{v}_1(\pi'), \ldots, \hat{v}_{N_{\mathrm{d}}-k}(\pi')) = (i_{k+1}, \ldots, i_{N_d})\}|.$$

The first term selects $c_j \in [d_{i_j}]$ for each $j \in [k]$; the second, falling factorial term selects the locations of the remaining entries of π whose first coordinate belongs to $\{i_1, \ldots, i_k\}$; and the third term specifies the order of the remaining entries of $\pi \in \mathcal{P}_{\mathrm{d}}$. Equation (3.1) also gives a formula for this final term, and the lemma then follows by routine algebra. □

LEMMA 3.3. *Fix a degree sequence* $\mathrm{d} = (d_1, \dots, d_n)$ *and let* $\Pi_\mathrm{d} \in_{\mathcal{U}} \mathcal{P}_\mathrm{d}$. *Let* $i \geq 1$. *Then for* $n \geq 2$ *and* k *such that* $\widetilde{J}_i^\mathrm{d} + k \in [N_\mathrm{d}]$,

$$\mathbf{P}\left\{\widetilde{J}_{i+1}^\mathrm{d} > \widetilde{J}_i^\mathrm{d} + k \;\middle|\; \widetilde{J}_1^\mathrm{d}, \dots, \widetilde{J}_i^\mathrm{d}, \widehat{V}_1(\Pi_\mathrm{d}), \dots, \widehat{V}_{N_\mathrm{d}}(\Pi_\mathrm{d})\right\}$$
$$= \prod_{j=\widetilde{J}_i^\mathrm{d}}^{\widetilde{J}_i^\mathrm{d}+k} \left(1 - \frac{\sum_{\ell=1}^{j}(d_{\widehat{V}_\ell(\Pi_\mathrm{d})} - 1) - i}{n-j}\right)$$

and, for $k > N_\mathrm{d}$,

$$\mathbf{P}\left\{\widetilde{J}_{i+1}^\mathrm{d} > k \;\middle|\; \widehat{V}_1(\Pi_\mathrm{d}), \dots, \widehat{V}_{N_\mathrm{d}}(\Pi_\mathrm{d})\right\} = 0.$$

The proof of Lemma 3.3 is analogous to that of Lemma 3.2 and is therefore omitted. Finally, a bound we will need in Section 5, whose proof relies on the bijective construction of T_n, is the following; its proof is postponed to the appendix.

LEMMA 3.4. *Let* $\mathrm{d} = (d_1, \dots, d_n)$ *be a degree sequence and let* $\mathcal{B} \subset [n]$ *be a set of vertices. Suppose that* $|\mathcal{B}| \leq K$ *and suppose that* $\max_{1 \leq i \leq n} d_i \leq \Delta$. *Let* B_d *be the smallest distance between two vertices in* $\mathcal{B}$ *that are ancestrally related in* $T_\mathrm{d} = B(\Pi_\mathrm{d})$ *(with* $B_\mathrm{d} = \infty$ *if no vertices in* $\mathcal{B}$ *are ancestrally related). Then, for any* $b \geq 0$

$$\mathbf{P}\{B_\mathrm{d} \leq b\} \leq K\left(1 - \left(1 - \frac{K\Delta}{n-1-b\Delta}\right)^b\right).$$

4. Random finite-dimensional distributions. In this section we use the bijection B to prove the convergence of the random finite-dimensional distributions of the head of the discrete snake (H_n, R_n). We assume throughout this section that μ is critical and has variance $\sigma^2 \in (0, \infty)$, and that assumption [**A1**] holds.

Recall that T_n is a Bienaymé tree with offspring distribution μ conditioned to have n vertices, and that $\mathbf{T}_n = (\mathrm{T}_n, Y)$ denotes the conditioned (μ, ν)-branching random walk. By Section 2.2, $\mathbf{T}_n$ has the same distribution as $(B(\Pi_{D^n}), Y)$, where $D^n = (D_1^n, \dots, D_n^n)$ is a sequence of IID μ-distributed random variables conditioned to have total sum $\sum_{i=1}^n D_i^n = n-1$ and, conditionally on D^n, $\Pi_{D^n} = ((V_1, C_1), \dots, (V_{n-1}, C_{n-1})) \in_{\mathcal{U}} \mathcal{P}_{D^n}$.

Fix $k \geq 1$. Let $U_1^n, \dots, U_k^n$ be a uniformly random k-set of indices chosen from $[n]$. Let $\mathrm{T}_n(U_1^n, \dots, U_k^n)$ be the subtree of T_n spanned by the root of T_n and the vertices $v_{U_1^n}, \dots, v_{U_k^n}$, where for $i \in [n]$, we recall that v_i is the i-th vertex in the lexicographical order of T_n. (For fixed k, as $n \to \infty$, a collection of k IID Uniform$([n])$ random variables will be distinct with probability tending to 1, so we can treat $U_1^n, \dots, U_k^n$ as indistinguishable from independent uniform picks from the vertices.) We immediately observe that $\mathrm{T}_n(U_1^n, \dots, U_k^n)$ has the same distribution as T_n^k, the subtree of $B(\Pi_{D^n})$ spanned by the root and the vertices $1, \dots, k$. Since T_n^k is more convenient for our analysis, we will work with it instead. Note that T_n^k is a labeled ordered rooted tree whose leaves are labeled by $1, 2, \dots, k$. Write ℓ_n^k for the map from T_n^k into $\mathbb{R}$ which gives the spatial locations of the vertices, so that $(\mathrm{T}_n^k, \ell_n^k)$ is the spatial tree (T_n, ℓ) restricted to the subtree spanned by the root and the vertices $1, \dots, k$.

Let $\mathcal{T}_{2\mathbf{e}}$ denote the Brownian tree encoded by the excursion $2\mathbf{e}$, and let $U_1, \dots, U_k$ be IID Uniform$([0,1])$ random variables, independent of $\mathbf{e}$. Recall that $\mathcal{T}_{2\mathbf{e}}^k = \mathcal{T}_{2\mathbf{e}}(U_1, \dots, U_k)$ denotes the subtree of $\mathcal{T}_{2\mathbf{e}}$ spanned by the images of 0 and $U_1, \dots, U_k$ in $\mathcal{T}_{2\mathbf{e}}$, thought of as an ordered rooted tree with leaves labeled by $1, 2, \dots, k$ and with real-valued edge lengths. Recall that $\mathcal{T}_{2\mathbf{e}}^k$ has the same distribution as the tree $\mathcal{T}^k$ built by Aldous' line-breaking construction. We now introduce a version of the line-breaking construction which incorporates spatial locations.

Line-breaking construction of the Brownian tree with spatial locations

We construct a sequence $(\mathcal{T}^k)_{k\geq 1}$ of trees along with two functions $\mathbf{h} : [0,\infty) \to [0,\infty)$ and $\mathbf{l} : [0,\infty) \to \mathbb{R}$ recursively. Let $J_1, J_2, \ldots$ be the jump times of a Poisson point process on $[0,\infty)$ with intensity tdt at time t, listed in increasing order. Independently, let $(B_t)_{t\geq 0}$ be a standard Brownian motion. Start from the tree $\mathcal{T}^1$ which consists of the line-segment $[0, J_1]$. Define $\mathbf{h}(t) = t$ and $\mathbf{l}(t) = B_t$ for $0 \leq t \leq J_1$. Recursively, for $k \geq 2$, conditionally on J_{k-1}, sample an attachment point $A_{k-1} \sim \mathrm{Uniform}([0, J_{k-1}])$, independent of $(A_j)_{j<k-1}$. Take the completion of the line segment $(J_{k-1}, J_k]$, and let J^*_{k-1} denote the limit point as $x \downarrow J_{k-1}$. Identify the points J^*_{k-1} and A_{k-1}. This has the effect of gluing the line-segment $(J_{k-1}, J_k]$ onto $\mathcal{T}^{k-1}$. We do this with probability $1/2$ to the left side and with probability $1/2$ to the right side. This yields $\mathcal{T}^k$. Define $\mathbf{h}(t) = \mathbf{h}(A_{k-1}) + t - J_{k-1}$ and $\mathbf{l}(t) = \mathbf{l}(A_{k-1}) + B_t - B_{J_{k-1}}$ for $t \in (J_{k-1}, J_k]$ to determine the height and location processes on the new line-segment.

The planar embedding of $\mathcal{T}^k$ is captured by a permutation $\tau^k : [k] \to [k]$ which is such that $\tau^k(1), \ldots, \tau^k(k)$ is the order in which we observe the leaves when exploring the tree from left to right. Using the notation $U^k_{(1)}, \ldots, U^k_{(k)}$ for the increasing ordering of $U_1, \ldots, U_k$ as in Proposition 1.10, we then have

$$\begin{aligned} &\left(\mathbf{h}(J_{\tau^k(1)}), \ldots, \mathbf{h}(J_{\tau^k(k)}), \mathbf{l}(J_{\tau^k(1)}), \ldots, \mathbf{l}(J_{\tau^k(k)})\right) \\ &\stackrel{\mathrm{d}}{=} \left(2\mathbf{e}_{U^k_{(1)}}, \ldots, 2\mathbf{e}_{U^k_{(k)}}, \sqrt{2}\mathbf{r}_{U^k_{(1)}}, \ldots, \sqrt{2}\mathbf{r}_{U^k_{(k)}}\right), \end{aligned} \tag{4.1}$$

where the equality in distribution of the first k co-ordinates on the two sides is a consequence of Corollary 22 of Aldous [4], and that of the final k co-ordinates is a consequence of the definition of the Brownian snake given at (1.1). So the line-breaking construction indeed realises the random finite-dimensional distributions of the head of the Brownian snake.

We show that the scaling limit of $(\mathrm{T}^k_n, \ell^k_n)$ is $(\mathcal{T}^k, \mathbf{l}|_{[0,J_k]})$ in an appropriate sense, which will allow us to prove the convergence of the random finite-dimensional distributions, along with a certain amount of extra information which will be useful to us in Section 5 where we prove tightness.

Recall that the tree T^k_n necessarily sits within the first k paths, $P^{(1)}, \ldots, P^{(k)}$, in the discrete line-breaking construction. We need to understand the lengths of these paths, and the positions at which the paths are glued onto one another. It is convenient to use the *indices* of the vertices in Π_{D^n} for this purpose rather than the vertex labels themselves. Recall that $J^{D^n}_1, J^{D^n}_2, \ldots, J^{D^n}_k$ are the first k indices at which we see either a repeat or an element of $\{1, 2, \ldots, k\}$. Let us henceforth write $J^n_i = J^{D^n}_i$ (and also $\widetilde{J}^n_i = \widetilde{J}^{D^n}_i$) for $i \geq 1$. Then the lengths of the paths $P^{(1)}, \ldots, P^{(k)}$ are given by $J^n_1, J^n_2 - J^n_1, \ldots, J^n_k - J^n_{k-1}$. For $1 \leq m \leq k-1$, the index at which the path $P^{(m+1)}$ attaches onto the subtree constructed from the first m paths is given by the value i such that $V_{J^n_m} = \widehat{V}_i(\Pi_{D^n})$ (i.e. we find the index of the vertex $V_{J^n_m}$ within the vector $(\widehat{V}_1(\Pi_{D^n}), \ldots, \widehat{V}_{N_n}(\Pi_{D^n}))$). We write A^n_m for this value i and call this the m-th *attachment point*. See Figure 4.

Since T^k_n is an ordered tree, we will need to understand where the paths $P^{(2)}, \ldots, P^{(k)}$ attach relative to the pre-existing children of their attachment points. If we are looking to attach to a vertex which has only one pre-existing child (i.e. for which there has been no previous repeat) then that vertex must have degree $d \geq 2$, and then whether we attach to the left or to the right of the pre-existing child is simply determined by the relative ordering of the corresponding second coordinates in the sequence Π_{D^n}. If there has been no previous

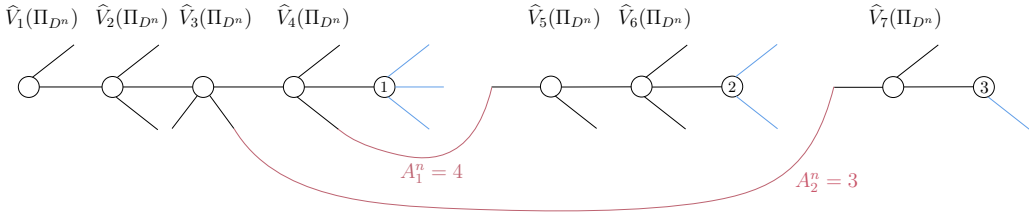

FIG 4. *Illustration of the first and second attachment points, A_1^n and A_2^n.*

repeat at this vertex then this pair of second coordinates is chosen uniformly at random without replacement from $[d]$ and, in particular, we attach to the left and right sides each with probability $1/2$. This ceases to be true after the first repeat (not least because then there are three or more children whose relative ordering we need to understand), but as we shall show below, we observe a second repeat of any vertex in T_n^k with vanishing probability as $n \to \infty$. Let $F_1^n, \ldots, F_k^n$ be random variables taking values in $\{0, 1, 2\}$ such that $F_i^n = 1$ if $P^{(i+1)}$ attaches at a first repeat and to the left-hand side, $F_i^n = 2$ if $P^{(i+1)}$ attaches at a first repeat and to the right-hand side and $F_{i+1}^n = 0$ otherwise, for $1 \le i \le k$.

Finally, recall that vertex $\widehat{V}_i(\Pi_{D^n})$ has degree $D^n_{\widehat{V}_i(\Pi_{D^n})}$ for $i \le N_{D^n} = |\{i \in [n] : D_i^n > 0\}|$. Let $L^n(0) = 0$ and let $L^n(i)$ be the spatial location of the C_i-th child of vertex V_i in line-breaking construction $B(\Pi_{D^n})$, for $1 \le i \le n-1$.

The following proposition shows that, on rescaling, these quantities converge in distribution to their analogues in the line-breaking construction of the Brownian tree with spatial locations.

PROPOSITION 4.1. *Fix $k \ge 1$. Then*

$$\frac{\sigma}{\sqrt{n}}(J_1^n, J_2^n, \ldots, J_k^n, A_1^n, \ldots, A_k^n) \xrightarrow{\mathrm{d}} (J_1, J_2, \ldots, J_k, A_1, \ldots, A_k) \tag{4.2}$$

as $n \to \infty$. Jointly with this convergence, we have that

$$(F_1^n, F_2^n, \ldots, F_k^n) \xrightarrow{\mathrm{d}} (F_1, F_2, \ldots, F_k), \tag{4.3}$$

where $F_1, F_2, \ldots, F_k$ are IID *random variables, independent of everything else, such that $\mathbf{P}\{F_i = 1\} = \mathbf{P}\{F_i = 2\} = 1/2$ and*

$$\begin{aligned} &\left(\frac{L^n(\lfloor tn^{1/2} \rfloor \wedge (J_1^n - 1))}{n^{1/4}}\right)_{t \ge 0} \xrightarrow{\mathrm{d}} \beta(B_{t \wedge (J_1/\sigma)})_{t \ge 0}, \\ &\left(\frac{L^n((J_i^n + \lfloor tn^{1/2} \rfloor) \wedge (J_{i+1}^n - 1))}{n^{1/4}}\right)_{t \ge 0} \xrightarrow{\mathrm{d}} \beta(B_{A_i/\sigma} + B_{((J_i/\sigma)+t) \wedge (J_{i+1}/\sigma)} - B_{(J_i/\sigma)})_{t \ge 0} \end{aligned} \tag{4.4}$$

for $1 \le i \le k-1$, in each case for the uniform norm.

As a corollary, we obtain the convergence of the random finite-dimensional distributions in (1.3).

COROLLARY 4.2. *For any $k \ge 1$, as $n \to \infty$*

$$\left(\frac{H_n(nU_{(1)}^k)}{\sqrt{n}}, \ldots, \frac{H_n(nU_{(k)}^k)}{\sqrt{n}}, \frac{R_n(nU_{(1)}^k)}{n^{1/4}}, \ldots, \frac{R_n(nU_{(k)}^k)}{n^{1/4}}\right)$$

$$\xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma} \mathbf{e}_{U^k_{(1)}}, \dots, \frac{2}{\sigma} \mathbf{e}_{U^k_{(k)}}, \beta \sqrt{\frac{2}{\sigma}} \mathbf{r}_{U^k_{(1)}}, \dots, \beta \sqrt{\frac{2}{\sigma}} \mathbf{r}_{U^k_{(k)}} \right),$$

where $U^k_{(1)}, \dots, U^k_{(k)}$ *are the order statistics of* k IID *Uniform*$([0,1])$ *random variables.*

PROOF. Let $(U^n_1, \dots, U^n_k)$ be a uniformly random k-set chosen from $[n]$, and $(U^{n,k}_{(1)}, \dots, U^{n,k}_{(k)})$ be the order statistics of $(U^n_1, \dots, U^n_k)$. As argued above, we may straightforwardly replace $nU^k_{(1)}, \dots, nU^k_{(k)}$ by $(U^{n,k}_{(1)} - 1, \dots, U^{n,k}_{(k)} - 1)$ at no asymptotic cost. Recall that $H_n(i)$ gives the distance from the root of the $(i+1)$-th vertex visited in a depth-first exploration of the tree. The random variables

$$(H_n(U^n_1 - 1), \dots, H_n(U^n_k - 1))$$

have the joint law of the distances from the root to the leaves labeled $1, 2, \dots, k$ in T^k_n (these may be expressed in terms of sums and differences of elements of $(J^n_1, \dots, J^n_k, A^n_1, \dots, A^n_k)$ analogously to the definition of $\mathbf{h}$ in the line-breaking construction of the Brownian tree with spatial locations), and

$$(R_n(U^n_1 - 1), \dots, R_n(U^n_k - 1)) = (L^n(J^n_1 - 1), \dots, L^n(J^n_k - 1)).$$

The effect of ordering the uniforms is simply to apply the same permutation of the entries to each of $(H_n(U^n_1 - 1), \dots, H_n(U^n_k - 1))$ and $(R_n(U^n_1 - 1), \dots, R_n(U^n_k - 1))$. This permutation is straightforwardly induced by the choices $(F^n_1, \dots, F^n_{k-1})$. By (4.3), this permutation then converges in distribution to τ^k. But then the claimed convergence follows from Proposition 4.1 using the scaling property of Brownian motion and (4.1). □

We begin by studying the vertex degrees at the start of the bijective construction, and show that, on the timescale of $\sqrt{n}$, the degrees that we observe are asymptotically indistinguishable from IID copies of $\bar{\xi}$. We show further that the subtree T^k_n is constructed on a timescale of order $\sqrt{n}$. This allows us to prove (4.2) in Proposition 4.7. To get the convergence of the spatial locations, we observe that, with the exception of branch points, the displacements along the ancestral lineages in T^k_n are asymptotically indistinguishable from IID copies of $Y_{\bar{\xi}, U_{\bar{\xi}}}$. Combining this with the convergence of the tree allows us to obtain the convergence of the spatial locations along the branches of the subtree.

4.1. *A discrete change of measure.* In this subsection, we show that the size-biased random re-ordering of the positive entries of D^n may be viewed as a vector of IID copies of the size-biased offspring random variable $\bar{\xi}$ up to a change of measure. We study the behaviour of the Radon–Nikodym derivative and show that its effect is trivial on the first $O(\sqrt{n})$ entries of the vector. Recall that $N_{D^n} = |\{i \in [n] : D^n_i > 0\}|$. To ease the notation, we write $N_n = N_{D^n}$. Let

$$\widehat{D}^n = \left(\widehat{D}^n_1, \dots, \widehat{D}^n_{N_n} \right)$$

be the size-biased random re-ordering of the positive entries of D^n. We note that

$$\widehat{D}^n \stackrel{\mathrm{d}}{=} \left(D^n_{\widehat{V}_1(\Pi_{D^n})}, \dots, D^n_{\widehat{V}_{N_n}(\Pi_{D^n})} \right).$$

Later we will often somewhat abuse notation and write $(\widehat{D}^n_1, \dots, \widehat{D}^n_{N_n})$ in place of

$$(D^n_{\widehat{V}_1(\Pi_{D^n})}, \dots, D_{\widehat{V}_{N_n}(\Pi_{D^n})}),$$

for example in the proof of Proposition 4.7.

PROPOSITION 4.3. *Let $\xi_1, \xi_2, \ldots, \xi_n$ be* IID *random variables with distribution μ. Further, let $\bar{\xi}_1, \bar{\xi}_2, \ldots$ be* IID *samples from the size-biased distribution of ξ_1. Then for $1 \le m < n$, and any non-negative measurable function $f : \mathbb{Z}^m \to \mathbb{R}_+$,*

$$\mathbf{E}\left[f(\widehat{D}_1^n, \ldots, \widehat{D}_m^n)\mathbf{1}_{[N_n \ge m]}\right] = \mathbf{E}\left[f(\bar{\xi}_1, \ldots, \bar{\xi}_m)\Theta^n(\bar{\xi}_1, \ldots, \bar{\xi}_m)\right],$$

where for $k_1, \ldots, k_m \in \mathbb{N}$,

$$\Theta^n(k_1, \ldots, k_m) = \frac{\mathbf{P}\left\{\sum_{i=m+1}^n \xi_i = n - 1 - \sum_{i=1}^m k_i\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \prod_{i=1}^m \left(\frac{n-i+1}{n-1-\sum_{j=1}^{i-1} k_j}\right), \tag{4.5}$$

if $k_1 + \ldots + k_m \le n-1$, and $\Theta^n(k_1, \ldots, k_m) = 0$ otherwise.

Proposition 4.3 is a special case of Proposition A.4 that we state and prove in the appendix, and use in full generality to prove Theorems 1.4 and 1.5 in Section 7. We state only the special case here as the more general formulation is much more technical and requires definitions that are only relevant in settings where assumption [**A3**] holds.

The next lemma shows that the change of measure Θ^n appearing in Proposition 4.3 is asymptotically unimportant provided that $m = \Theta(\sqrt{n})$.

LEMMA 4.4. *Let μ be a critical offspring distribution with variance $\sigma^2 \in (0, \infty)$, and let $(\bar{\xi}_i)_{i \ge 1}$ be* IID *samples from the size-biased distribution of μ. Suppose that $m = m(n) = \Theta(\sqrt{n})$. Then as $n \to \infty$*

$$\Theta^n(\bar{\xi}_1, \ldots, \bar{\xi}_m) \xrightarrow{\mathrm{p}} 1,$$

and $(\Theta^n(\bar{\xi}_1, \ldots, \bar{\xi}_m))_{n \ge 1}$ is a uniformly integrable sequence of random variables.

PROOF. By a subsubsequence argument we may assume that $m/\sqrt{n} \to t$ as $n \to \infty$ for some $t > 0$. Let $\xi_1, \ldots, \xi_n$ be IID random variables with distribution μ. We deal with the ratio of probabilities in the definition of Θ^n using the local central limit theorem. Specifically, since $\mathbf{E}[\xi_1] = 1$ and $\mathbf{Var}\{\xi_1\} = \sigma^2$, we have that

$$\sup_{k \in \mathbb{Z}} \left| \sqrt{n-m} \cdot \mathbf{P}\left\{\sum_{i=m+1}^n \xi_i = n - 1 - m + k\right\} - \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{k^2}{2\sigma^2(n-m)}\right)\right| \to 0$$

as $n \to \infty$, so for $k_1, \ldots, k_m \in \mathbb{N}$,

$$\begin{aligned}
&\mathbf{P}\left\{\sum_{i=m+1}^n \xi_i = n - 1 - \sum_{i=1}^m k_i\right\} \\
&= \mathbf{P}\left\{\sum_{i=m+1}^n \xi_i - (n-m) = -1 - m\sigma^2 - \sum_{i=1}^m (k_i - 1 - \sigma^2)\right\} \\
&= \frac{\exp\left(-\frac{1}{2\sigma^2(n-m)}\left(1 + m\sigma^2 + \sum_{i=1}^m (k_i - 1 - \sigma^2)\right)^2\right)}{\sqrt{2\pi\sigma^2(n-m)}} + o(n^{-1/2}).
\end{aligned}$$

Similarly, we have that

$$\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n - 1\right\} = \frac{1}{\sqrt{2\pi\sigma^2 n}} + o(n^{-1/2}).$$

Therefore,

$$\frac{\mathbf{P}\left\{\sum_{i=m+1}^{n}\xi_i = n-1-\sum_{i=1}^{m}k_i\right\}}{\mathbf{P}\left\{\sum_{i=1}^{n}\xi_i = n-1\right\}}$$

$$= \exp\left(-\left(\frac{1+m\sigma^2+\sum_{i=1}^{m}(k_i-1-\sigma^2)}{\sqrt{2\sigma^2(n-m)}}\right)^2\right)+o(1). \tag{4.6}$$

Since the random variables $\bar{\xi}_1,\dots,\bar{\xi}_n$ are IID with mean σ^2+1, by the functional strong law of large numbers (as stated in Lemma A.1), as $n\to\infty$,

$$\frac{1}{\sqrt{n}}\max_{1\le i\le\lfloor t\sqrt{n}\rfloor}\left|\sum_{j=1}^{i}(\bar{\xi}_j-1-\sigma^2)\right| \xrightarrow{\text{a.s.}} 0. \tag{4.7}$$

Since $m=(1+o(1))t\sqrt{n}$ this, in particular, yields that

$$\exp\left(-\left(\frac{1+m\sigma^2+\sum_{i=1}^{m}(\bar{\xi}_i-(1+\sigma^2))}{\sqrt{2\sigma^2(n-m)}}\right)^2\right) \xrightarrow{\text{p}} \exp\left(-\frac{t^2\sigma^2}{2}\right). \tag{4.8}$$

We claim that, as $n\to\infty$,

$$\prod_{i=1}^{m}\left(\frac{n-i+1}{n-1-\sum_{j=1}^{i-1}\bar{\xi}_j}\right) \xrightarrow{\text{p}} \exp\left(\frac{t^2\sigma^2}{2}\right). \tag{4.9}$$

Indeed,

$$\prod_{i=1}^{m}\left(\frac{n-i+1}{n-1-\sum_{j=1}^{i-1}\bar{\xi}_j}\right) = \exp\left(-\sum_{i=1}^{m}\log\left(1-\frac{\sum_{j=1}^{i-1}(\bar{\xi}_j-1-\sigma^2)+\sigma^2(i-1)}{n-i+1}\right)\right).$$

It follows by Taylor's theorem and (4.7) that the last expression is equal to

$$\exp\left(\sum_{i=1}^{m}\frac{\sum_{j=1}^{i-1}(\bar{\xi}_j-1-\sigma^2)+\sigma^2(i-1)}{n-i+1}+o_{\mathbf{P}}(1)\right)$$

$$= \exp\left(\frac{\sigma^2\lfloor t\sqrt{n}\rfloor(\lfloor t\sqrt{n}\rfloor-1)}{2n}+o_{\mathbf{P}}(1)\right) \xrightarrow{\text{p}} \exp\left(\frac{t^2\sigma^2}{2}\right), \tag{4.10}$$

establishing (4.9). Combining this with (4.6) and (4.8) yields that

$$\Theta^n(\bar{\xi}_1,\dots,\bar{\xi}_m) \xrightarrow{\text{p}} 1.$$

To prove uniform integrability, notice that, by applying Proposition 4.3 with $f\equiv 1$,

$$\mathbf{E}\left[\Theta^n(\bar{\xi}_1,\dots,\bar{\xi}_m)\right] = \mathbf{P}\left\{N_n\ge m\right\}.$$

We claim that this tends to 1 as $n\to\infty$. To this end, note that

$$\#\{i\in[n]\,:\,\xi_i>0\} \overset{\text{d}}{=} \text{Binomial}(n,1-\mu_0).$$

So by conditioning on the event $\{\sum_{i=1}^{n}\xi_i=n-1\}$, which occurs with probability $\Theta(n^{-1/2})$, there are $(1+o_{\mathbf{P}}(1))n(1-\mu_0)$ non-zero entries of $(\xi_1,\dots,\xi_n)$. Since $m=(1+o(1))t\sqrt{n}$ it follows that as $n\to\infty$,

$$\mathbf{P}\left\{N_n\ge m\right\} = \mathbf{P}\left\{\#\{i\in[n]\,:\,\xi_i>0\}\ge m \;\middle|\; \sum_{i=1}^{n}\xi_i=n-1\right\}\to 1.$$

Uniform integrability then follows by the generalised Scheffé lemma, see [19, Theorem 5.12]. □

Lemma 4.4 implies that if the offspring distribution has finite variance then, on a timescale of $\sqrt{n}$ in the bijective construction $B(\Pi_{D^n})$ of T_n, the degrees we observe are asymptotically indistiguishable from IID copies of $\bar{\xi}$. To prove Proposition 4.1, we use this fact in the form of Proposition 4.5 stated below.

PROPOSITION 4.5. *Given $\widehat{D}^n$, let $U_1, \dots, U_n$ be independent random variables such that, for each $i \in [n]$, U_i is uniformly distributed on $[\widehat{D}_i^n]$. Further, let $Y_{\widehat{D}_1^n,U_1}, \dots Y_{\widehat{D}_n^n,U_n}$ be independent random variables such that, for each $i \in [n]$, $Y_{\widehat{D}_i^n,U_i}$ is a uniform entry of a $\nu_{\widehat{D}_i^n}$-distributed displacement vector. If [**A1**] holds then as $n \to \infty$,*

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor t\sqrt{n} \rfloor} (\widehat{D}_i^n - 1), \frac{1}{n^{1/4}} \sum_{i=1}^{\lfloor t\sqrt{n} \rfloor} Y_{\widehat{D}_i^n,U_i} \right)_{t \geq 0} \xrightarrow{\mathrm{d}} (\sigma^2 t, \beta B_t)_{t \geq 0},$$

for the topology of uniform convergence on compact time-intervals, where $(B_t)_{t\geq 0}$ is a standard Brownian motion.

PROOF. Fix $T > 0$ and let $F : \mathbf{D}([0,T], \mathbb{R})^2 \to \mathbb{R}$ be a bounded continuous function, where $\mathbf{D}([0,T], \mathbb{R})$ is the space of real-valued functions on $[0,T]$ that are right-continuous with left limits equipped with the Skorokhod topology. Let $\bar{\xi}_1, \bar{\xi}_2, \dots$ be IID samples from the size biased distribution of ξ. Further, independently for $i \geq 1$, let $\overline{U}_i$ be a Uniform$([\bar{\xi}_i])$ random variable.

By Proposition 4.3,

$$\mathbf{E}\left[F\left(\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor t\sqrt{n} \rfloor} (\widehat{D}_i^n - 1), \frac{1}{n^{1/4}} \sum_{i=1}^{\lfloor t\sqrt{n} \rfloor} Y_{\widehat{D}_i^n,U_i} \right)_{0 \leq t \leq T} \right) \mathbf{1}_{[N_n \geq \lfloor T\sqrt{n} \rfloor]} \right]$$

(4.11)

$$= \mathbf{E}\left[F\left(\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor t\sqrt{n} \rfloor} (\bar{\xi}_i - 1), \frac{1}{n^{1/4}} \sum_{i=1}^{\lfloor t\sqrt{n} \rfloor} Y_{\bar{\xi},\overline{U}_i} \right)_{0 \leq t \leq T} \right) \Theta^n\left(\bar{\xi}_1, \dots, \bar{\xi}_{\lfloor T\sqrt{n} \rfloor} \right) \right],$$

where the random variables $(Y_{\bar{\xi}_i,\overline{U}_i})_{i\geq 1}$ are independent and, given $\bar{\xi}_i$, we have that $Y_{\bar{\xi}_i,\overline{U}_i}$ is a uniform entry of a $\nu_{\bar{\xi}_i}$ distributed displacement vector. Since $\mathbf{E}\left[\bar{\xi}_1\right] = \sigma^2 + 1$, by the functional strong law of large numbers (Lemma A.1), as $n \to \infty$,

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor t\sqrt{n} \rfloor} (\bar{\xi}_i - 1) \right)_{t \geq 0} \xrightarrow{\mathrm{p}} (\sigma^2 t)_{t \geq 0}$$

in $\mathbf{D}((0,T), \mathbb{R})$.

Furthermore, the random variables $(Y_{\bar{\xi},\overline{U}_i})_{i\geq 1}$ are IID with mean and variance given by

$$\mathbf{E}\left[Y_{\bar{\xi}_1,\overline{U}_1} \right] = \sum_{k=1}^{\infty} \mu_k \sum_{j=1}^{k} \mathbf{E}\left[Y_{k,j} \right] = 0, \quad \mathbf{Var}\left\{ Y_{\bar{\xi}_1,\overline{U}_1} \right\} = \sum_{k=1}^{\infty} \mu_k \sum_{j=1}^{k} \mathbf{E}\left[Y_{k,j}^2 \right] = \beta^2.$$

It then follows from Donsker's theorem that as $n \to \infty$

$$\left(\frac{1}{n^{1/4}} \sum_{i=1}^{\lfloor t\sqrt{n} \rfloor} Y_{\bar{\xi}_i,\overline{U}_i} \right)_{t \geq 0} \xrightarrow{\mathrm{d}} (\beta B_t)_{t \geq 0}$$

in $\mathbf{D}([0,T],\mathbb{R})$. Therefore, by the continuity of F, as $n\to\infty$,

$$\mathbf{E}\left[F\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor t\sqrt{n}\rfloor}(\bar{\xi}_i-1),\frac{1}{n^{1/4}}\sum_{i=1}^{\lfloor t\sqrt{n}\rfloor}Y_{\bar{\xi}_i,\overline{U}_i}\right)_{0\le t\le T}\right)\right]\to\mathbf{E}\left[F\left((\sigma^2 t,\beta B_t)_{0\le t\le T}\right)\right].$$

Combining this with Lemma 4.4 and the boundedness of F yields that (4.11) converges to

$$\mathbf{E}\left[F\left((\sigma^2 t,\beta B_t)_{0\le t\le T}\right)\right],$$

as $n\to\infty$, and the result follows. □

4.2. *Bijective construction on the timescale $\sqrt{n}$.* In this subsection we show that the subtree T_n^k is constructed on a timescale of order $\sqrt{n}$ with high probability. We then prove that the lengths of the paths which are glued together to form T_n^k converge on rescaling, as do the positions at which they attach to one another.

We begin by showing that, with high probability, the vertices $1,\dots,k$ do not appear in the first $\Theta(\sqrt{n})$ entries of Π_{D^n}.

LEMMA 4.6. *Fix $T>0$ and $k\ge 1$, and let*

$$\mathcal{G}_{n,k}(T)=\left\{\left\{\widehat{V}_1(\Pi_{D^n}),\dots,\widehat{V}_{\lfloor T\sqrt{n}\rfloor}(\Pi_{D^n})\right\}\cap\{1,\dots,k\}=\emptyset,N_n\ge\lfloor T\sqrt{n}\rfloor\right\}.$$

Then $\mathbf{P}\{\mathcal{G}_{n,k}(T)\}\to 1$ as $n\to\infty$.

Notice that on the *good event* $\mathcal{G}_{n,k}(T)$, if $J_k^n\le\lfloor T\sqrt{n}\rfloor$ then $J_i^{D^n}=\widetilde{J}_i^{D^n}$ for all $i\in[k]$, and T_n^k is precisely the tree spanned by the root and the paths $P^{(1)},\dots,P^{(k)}$ in the bijective construction $B(\Pi_{D^n})$ of T_n.

PROOF. We have

$$\begin{aligned}\mathbf{P}\{\mathcal{G}_{n,k}(T)\}&=\mathbf{P}\left\{\{\widehat{V}_1(\Pi_{D^n}),\dots,\widehat{V}_{\lfloor T\sqrt{n}\rfloor}(\Pi_{D^n}))\}\cap\{1,\dots,k\}=\emptyset,N_n\ge\lfloor T\sqrt{n}\rfloor\right\}\\&\ge\mathbf{E}\left[\left(1-\frac{D_1^n+\cdots+D_k^n}{n-1-T\sqrt{n}\max_{1\le i\le n}D_i^n}\right)^{\lfloor T\sqrt{n}\rfloor}\right]-\mathbf{P}\left\{N_n<\lfloor T\sqrt{n}\rfloor\right\}.\end{aligned}$$

Let $\varepsilon>0$ and $(\xi_i)_{i\ge1}$ be a sequence of IID random variables with distribution μ. Then,

$$\begin{aligned}\mathbf{P}\left\{\max_{1\le i\le n}D_i^n>\varepsilon\sqrt{n}\right\}&=\frac{\mathbf{P}\{\max_{1\le i\le n}\xi_i>\varepsilon\sqrt{n},\sum_{i=1}^n\xi_i=n-1\}}{\mathbf{P}\{\sum_{i=1}^n\xi_i=n-1\}}\\&\le\frac{n\mathbf{P}\{\xi_1>\varepsilon\sqrt{n},\sum_{i=1}^n\xi_i=n-1\}}{\mathbf{P}\{\sum_{i=1}^n\xi_i=n-1\}}.\end{aligned}\tag{4.12}$$

Since $\mathbf{E}\left[\xi^2\right]<\infty$ we have $n\mathbf{P}\{\xi_1>\varepsilon\sqrt{n}\}\to0$ as $n\to\infty$. Hence,

$$\begin{aligned}&\frac{n\mathbf{P}\{\xi_1>\varepsilon\sqrt{n},\sum_{i=1}^n\xi_i=n-1\}}{\mathbf{P}\{\sum_{i=1}^n\xi_i=n-1\}}\\&\le\frac{n\mathbf{P}\{\xi_1>\varepsilon\sqrt{n}\}\max_{\varepsilon\sqrt{n}<m\le n-1}\mathbf{P}\{\sum_{i=2}^n\xi_i=n-1-m\}}{\mathbf{P}\{\sum_{i=1}^n\xi_i=n-1\}}\to0\end{aligned}$$

as $n\to\infty$. Combining this with (4.12) gives that

$$\frac{1}{\sqrt{n}}\max_{1\le i\le n}D_i^n\xrightarrow{\mathrm{p}}0$$

and so

$$\frac{D_1^n + \cdots + D_k^n}{\sqrt{n}} \xrightarrow{\text{p}} 0$$

as $n \to \infty$. Therefore, by the bounded convergence theorem,

$$\mathbf{E}\left[\left(1 - \frac{D_1^n + \cdots + D_k^n}{n - 1 - T\sqrt{n}\max_{1\le i\le n} D_i}\right)^{\lfloor T\sqrt{n}\rfloor}\right] \to 1$$

as $n \to \infty$. The result then follows by noting (as at the end of the proof of Lemma 4.4) that as $n \to \infty$,

$$\mathbf{P}\left\{N_n < \lfloor T\sqrt{n}\rfloor\right\} \le \frac{\mathbf{P}\left\{\text{Binomial}(n, 1-\mu_0) < \lfloor T\sqrt{n}\rfloor\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \to 0. \qquad \square$$

PROPOSITION 4.7. *Fix $k \ge 1$. Then*

$$\frac{\sigma}{\sqrt{n}}(J_1^n, J_2^n, \ldots, J_k^n, A_1^n, \ldots, A_k^n) \xrightarrow{\text{d}} (J_1, J_2, \ldots, J_k, A_1, \ldots, A_k) \tag{4.13}$$

as $n \to \infty$, where $J_1, J_2, \ldots, J_k$ are the first k jump-times of an inhomogeneous Poisson process of intensity t with respect to the Lebesgue measure at $t \in \mathbb{R}_+$ and, for $i \in [k]$, conditionally on $J_1, \ldots, J_i$, A_i is uniform on $[0, J_i]$, independently of $A_1, \ldots, A_{i-1}$.

PROOF. Fix $T > 0$. Let $0 \le t_1 \le \cdots \le t_k \le T$ and $s_1 < t_1, \ldots, s_k < t_k$. We will prove that

$$\begin{aligned}
&\mathbf{P}\left\{J_1^n \le t_1\sqrt{n}, \ldots, J_k^n \le t_k\sqrt{n}, A_1^n \le s_1\sqrt{n}, \ldots, A_k^n \le s_k\sqrt{n}\right\} \\
&\quad \to \sigma^{2k}\left(\prod_{j=1}^k s_j\right)\int_0^{t_1}\cdots\int_{r_{k-1}}^{t_k} \exp(-\sigma^2 t_k^2/2)dr_k\ldots dr_1 \\
&\quad = \mathbf{P}\left\{J_1 \le \sigma t_1, \ldots J_k \le \sigma t_k, A_1 \le \sigma s_1, \ldots, A_k \le \sigma s_k\right\}. \qquad (4.14)
\end{aligned}$$

We will often work conditionally on the random variables $\widehat{D}^n = (\widehat{D}_1^n, \ldots, \widehat{D}_{N_n}^n)$. To make the equations easier to read, we write $\mathbf{P}_{\widehat{D}^n}$ for the conditional probability given $\widehat{D}^n$ and $\mathbf{E}_{\widehat{D}^n}$ for the corresponding expectation.

Fix $T' > T$. By Skorokhod's representation theorem, there exists a probability space on which the uniform convergence

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor t\sqrt{n}\rfloor}(\widehat{D}_i^n - 1)\right)_{0\le t\le T'} \xrightarrow{\text{d}} (\sigma^2 t)_{0\le t\le T'} \tag{4.15}$$

from Proposition 4.5 occurs in the almost sure sense. We work on this probability space for the rest of the proof. Note, in particular, that if the above convergence occurs almost surely then it is also the case that

$$\mathbf{1}_{[N_n \ge \lfloor T\sqrt{n}\rfloor]} \xrightarrow{\text{a.s.}} 1.$$

We first show that, as $n \to \infty$,

$$\begin{aligned}
&n^{k/2}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor, \widetilde{J}_2^n = \lfloor t_2\sqrt{n}\rfloor, \ldots, \widetilde{J}_k^n = \lfloor t_k\sqrt{n}\rfloor\right\}\mathbf{1}_{[N_n\ge\lfloor T\sqrt{n}\rfloor]} \\
&\quad \xrightarrow{\text{a.s.}} \sigma^{2k}t_1t_2\ldots t_k\exp\left(-\sigma^2 t_k^2/2\right). \qquad (4.16)
\end{aligned}$$

By Lemma 3.1, whenever the bijective construction $B(\Pi_{D^n})$ of T_n encounters a new vertex, its degree is distributionally equivalent to the next one on the list $(\widehat{D}_1^n, \ldots, \widehat{D}_{N_n}^n)$. So by Lemma 3.2, on the event $\{N_n \geq \lfloor T\sqrt{n} \rfloor\}$, we have

$$\begin{aligned}
&\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor\right\} \\
&= \frac{\sum_{\ell=1}^{\lfloor t_1\sqrt{n}\rfloor - 1}(\widehat{D}_\ell^n - 1)}{n - \lfloor t_1\sqrt{n}\rfloor} \prod_{j=1}^{\lfloor t_1\sqrt{n}\rfloor - 2}\left(1 - \frac{\sum_{\ell=1}^{j}(\widehat{D}_\ell^n - 1)}{n-1-j}\right) \\
&= \frac{\sum_{\ell=1}^{\lfloor t_1\sqrt{n}\rfloor - 1}(\widehat{D}_\ell^n - 1)}{n - \lfloor t_1\sqrt{n}\rfloor} \exp\left(\sum_{j=1}^{\lfloor t_1\sqrt{n}\rfloor - 2}\log\left(1 - \frac{\sum_{\ell=1}^{j}(\widehat{D}_\ell^n - 1)}{n-1-j}\right)\right) \\
&= \frac{\sum_{\ell=1}^{\lfloor t_1\sqrt{n}\rfloor - 1}(\widehat{D}_\ell^n - 1)}{n - \lfloor t_1\sqrt{n}\rfloor} \exp\left(\sum_{j=1}^{\lfloor t_1\sqrt{n}\rfloor - 2}\log\left(1 - \frac{\sum_{\ell=1}^{j}(\widehat{D}_\ell^n - 1 - \sigma^2) + \sigma^2 j}{n-1-j}\right)\right).
\end{aligned}$$

By (4.15) and a similar argument to that used in the proof of (4.10), we get that as $n \to \infty$,

$$\sqrt{n}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor\right\}\mathbf{1}_{[N_n \geq \lfloor T\sqrt{n}\rfloor]} \xrightarrow{\text{a.s.}} \sigma^2 t_1 \exp\left(-\frac{t_1^2\sigma^2}{2}\right).$$

We now proceed to prove the joint convergence of the first k coordinates in (4.16) by induction. Suppose that the claimed convergence holds for $\widetilde{J}_1^n, \ldots, \widetilde{J}_{m-1}^n$. By Lemma 3.3, on the event $\{N_n \geq \lfloor T\sqrt{n}\rfloor\}$,

$$\begin{aligned}
&\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_m^n - \widetilde{J}_{m-1}^n = \lfloor t_m\sqrt{n}\rfloor - \lfloor t_{m-1}\sqrt{n}\rfloor \;\middle|\; \widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor, \ldots, \widetilde{J}_{m-1}^n = \lfloor t_{m-1}\sqrt{n}\rfloor\right\} \\
&= \frac{\sum_{\ell=1}^{\lfloor t_m\sqrt{n}\rfloor - m}(\widehat{D}_\ell^n - 1) - m + 1}{n - \lfloor t_m\sqrt{n}\rfloor} \prod_{j=\lfloor t_{m-1}\sqrt{n}\rfloor}^{\lfloor t_m\sqrt{n}\rfloor - 2}\left(1 - \frac{\sum_{\ell=1}^{j-m+1}(\widehat{D}_\ell^n - 1) - m + 1}{n-1-j}\right).
\end{aligned}$$

Arguing as above, we obtain

$$\begin{aligned}
&\sqrt{n}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_m^n = \lfloor t_m\sqrt{n}\rfloor \;\middle|\; \widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor, \ldots, \widetilde{J}_{m-1}^n = \lfloor t_{m-1}\sqrt{n}\rfloor\right\}\mathbf{1}_{[N_n \geq \lfloor t_m\sqrt{n}\rfloor]} \\
&\quad\xrightarrow{\text{a.s.}} \sigma^2 t_m \exp\left(-\int_{t_{m-1}}^{t_m}\sigma^2 r dr\right).
\end{aligned}$$

By induction on m, we get this for all $1 \leq m \leq k$. Taking the product of the conditional probabilities, we obtain (4.16).

We now wish to add in the random variables $(A_i^n)_{i\in[k]}$. We work conditionally on the event $\mathcal{G}_{n,k}(T)$. Given also $J_1^n = \lfloor t_1\sqrt{n}\rfloor, \ldots, J_m^n = \lfloor t_m\sqrt{n}\rfloor$, $\widehat{D}_1^n, \ldots, \widehat{D}_{N_n}^n$ and $A_1^n, \ldots, A_{m-1}^n$, since $N_n \geq \lfloor T\sqrt{n}\rfloor$, at time J_m^n there are $\widehat{D}_i^n - 1 - \sum_{\ell=1}^{m-1}\mathbf{1}_{[A_\ell^n = i]}$ remaining instances of the vertex $\widehat{V}_i(\Pi_{D^n})$ to appear in the bijective construction. So, the repeated vertex that we see is $\widehat{V}_i(\Pi_{D^n})$, i.e. $A_m^n = i$, with probability

$$\frac{\widehat{D}_i^n - 1 - \sum_{\ell=1}^{m-1}\mathbf{1}_{[A_\ell^n = i]}}{\sum_{j=1}^{\lfloor t_m\sqrt{n}\rfloor - m}(\widehat{D}_i^n - 1) - m},$$

for $1 \leq i \leq \lfloor t_m\sqrt{n}\rfloor - m$. Hence,

$$\mathbf{P}_{\widehat{D}^n}\left\{A_m^n \leq s_m\sqrt{n} \mid \mathcal{G}_{n,k}(T), J_1^n = \lfloor t_1\sqrt{n}\rfloor, \ldots, J_m^n = \lfloor t_m\sqrt{n}\rfloor, A_1^n, \ldots, A_{m-1}^n\right\}$$

$$= \frac{\sum_{i=1}^{\lfloor s_m\sqrt{n}\rfloor}(\widehat{D}_i^n - 1) - \sum_{\ell=1}^{m-1}\mathbf{1}_{[A_\ell^n \le s_m\sqrt{n}]}}{\sum_{j=1}^{\lfloor t_m\sqrt{n}\rfloor - m}(\widehat{D}_j^n - 1) - m}.$$

This quantity lies in the interval

$$\left[\frac{\sum_{i=1}^{\lfloor s_m\sqrt{n}\rfloor}(\widehat{D}_i^n - 1) - m + 1}{\sum_{j=1}^{\lfloor t_m\sqrt{n}\rfloor - m}(\widehat{D}_j^n - 1) - m}, \frac{\sum_{i=1}^{\lfloor s_m\sqrt{n}\rfloor}(\widehat{D}_i^n - 1)}{\sum_{j=1}^{\lfloor t_m\sqrt{n}\rfloor - m}(\widehat{D}_j^n - 1) - m}\right]$$

whose end-points do not depend on $A_1^n, A_2^n, \dots, A_{m-1}^n$. Iterating, we thus obtain that

$$\mathbf{P}_{\widehat{D}^n}\left\{A_1^n \le s_1\sqrt{n}, \dots, A_k^n \le s_k\sqrt{m} \;\middle|\; \mathcal{G}_{n,k}(T), J_1^n = \lfloor t_1\sqrt{n}\rfloor, \dots, J_m^n = \lfloor t_m\sqrt{n}\rfloor\right\}$$

lies in a random interval depending only on $\widehat{D}_1^n, \dots, \widehat{D}_{\lfloor t_k\sqrt{n}\rfloor}$, both of whose end-points converge almost surely to $\prod_{m=1}^k (s_m/t_m)$ by (4.15). So the same is true by sandwiching for our conditional probability which lies in that interval.

Putting everything together, we then have

$$\begin{aligned}
&\mathbf{P}\left\{J_1^n \le t_1\sqrt{n}, \dots, J_k^n \le t_k\sqrt{n}, A_1^n \le s_1\sqrt{n}, \dots, A_k \le s_k\sqrt{n}\right\}\\
&\quad = \mathbf{P}\left\{J_1^n \le t_1\sqrt{n}, \dots, J_k^n \le t_k\sqrt{n}, A_1^n \le s_1\sqrt{n}, \dots, A_k^n \le s_k\sqrt{n}, \mathcal{G}_{n,k}(T)^c\right\}\\
&\qquad + \mathbf{E}\left[\mathbf{P}_{\widehat{D}^n}\left\{J_1^n \le t_1\sqrt{n}, \dots, J_k^n \le t_k\sqrt{n}, A_1^n \le s_1\sqrt{n}, \dots, A_k^n \le s_k\sqrt{n}, \mathcal{G}_{n,k}(T)\right\}\right].
\end{aligned}$$

The first term on the right-hand side of this equation clearly tends to 0 by Lemma 4.6. Since the second is the expectation of a conditional probability, it is sufficient to show that the conditional probability itself tends to $\exp(-\sigma^2 t_k/2)\prod_{m=1}^k s_k$ in distribution. For $1 \le m \le k$ and $n \ge 1$, let us write

$$t_m^n = \frac{\lfloor t_m\sqrt{n}\rfloor + 1}{\sqrt{n}}.$$

Then we have

$$\begin{aligned}
&\mathbf{P}_{\widehat{D}^n}\left\{J_1^n \le t_1\sqrt{n}, \dots, J_k^n \le t_k\sqrt{n}, A_1^n \le s_1\sqrt{n}, \dots, A_k^n \le s_k\sqrt{n}, \mathcal{G}_{n,k}(T)\right\}\\
&= \int_0^{t_1^n}\!\!\cdots\int_{r_{k-1}}^{t_k^n} \mathbf{P}_{\widehat{D}^n}\left\{A_1^n \le s_1\sqrt{n}, \dots, A_k \le s_k\sqrt{n} \,\middle|\, \mathcal{G}_{n,k}(T), J_1^n = \lfloor r_1\sqrt{n}\rfloor, \dots, J_k^n = \lfloor r_k\sqrt{n}\rfloor\right\}\\
&\qquad \times n^{k/2}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor r_1\sqrt{n}\rfloor, \dots, \widetilde{J}_k^n = \lfloor r_k\sqrt{n}\rfloor, \mathcal{G}_{n,k}(T)\right\} dr_k \dots dr_1\\
&= \int_0^{t_1^n}\!\!\cdots\int_{r_{k-1}}^{t_k^n} \mathbf{P}_{\widehat{D}^n}\left\{A_1^n \le s_1\sqrt{n}, \dots, A_k \le s_k\sqrt{n} \,\middle|\, \mathcal{G}_{n,k}(T), J_1^n = \lfloor r_1\sqrt{n}\rfloor, \dots, J_k^n = \lfloor r_k\sqrt{n}\rfloor\right\}\\
&\qquad \times n^{k/2}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor r_1\sqrt{n}\rfloor, \dots, \widetilde{J}_k^n = \lfloor r_k\sqrt{n}\rfloor\right\}\mathbf{1}_{[N_n \ge \lfloor T\sqrt{n}\rfloor]} dr_k \dots dr_1 - E_n,
\end{aligned}$$

where E_n is an error term with the property that $0 \le E_n \le \mathbf{P}_{\widehat{D}^n}\{\mathcal{G}_{n,k}(T)^c\}$, and so tends to 0 in distribution as $n \to \infty$. The first term in the product which forms the integrand tends to $\prod_{m=1}^k (s_m/r_m)$ as $n \to \infty$ and the second term tends to $\sigma^{2k} r_1 \dots r_k \exp(-\sigma^2 r_k^2/2)$, both almost surely. Write $g_n(r_1, \dots, r_k)$ for the integrand above, considered as a function of $r_1, \dots, r_k$. Then we have just shown that

$$g_n(r_1, \dots, r_k) \xrightarrow{\text{a.s.}} \sigma^{2k}\prod_{m=1}^k s_m \exp(-\sigma^2 r_k^2/2),$$

It is straightforward to see that this convergence is, in fact, uniform on compacts. Hence,

$$\int_0^{t_1^n}\cdots\int_{r_{k-1}}^{t_k^n} g_n(r_1,\dots,r_k)dr_k\dots dr_1$$
$$\xrightarrow{\text{a.s.}} \sigma^{2k}\left(\prod_{m=1}^{k} s_m\right)\int_0^{t_1}\cdots\int_{r_{k-1}}^{t_k}\exp(-\sigma^2 r_k^2/2)dr_k\dots dr_1,$$

which yields (4.14). The result follows, since $T>0$ was arbitrary. □

This completes the proof of (4.2) in Proposition 4.1.

4.3. *Displacements at repeats.* As shown above, for fixed k and large n, T_n^k is with high probability the subtree of $B(\Pi_{D^n})$ composed of the union of the paths $P^{(1)},\dots,P^{(k)}$. Moreover, for $i\in[k]$, under the bijective construction, by Proposition 4.5, with the exception of the first vertex in each path $P^{(i)}$, the displacements of the vertices in $P^{(i)}$ away from their parents are asymptotically indistinguishable from IID copies of uniform entries of a $\nu_{\bar\xi}$ distributed displacement vector. On the other hand, the displacement away from of the first vertex in $P^{(i)}$ cannot be compared to a random variable with the same distribution as a uniform entry of a $\nu_{\bar\xi}$ distributed displacement vector. However, in the following lemma we will prove that such displacements are $O_{\mathbf{P}}(1)$ and so negligible on the scale of $n^{1/4}$.

We first introduce some notation. Recall that for $i\in[\ell^*(\Pi_{D^n})]$, vertex $V_{J_i^n}$ is the i-th repeated vertex encountered in the bijective construction $(B(\Pi_{D^n}),Y)$ of $\mathbf{T}_n=(\mathrm{T}_n,Y)$ (and hence a branchpoint). For $i\in[\ell^*(\Pi_{D^n})]$, let Δ_i^n be the displacement of $V_{J_i^n+1}$ away from its parent $V_{J_i^n}$ in $\mathbf{T}_n$.

LEMMA 4.8. *For any $\ell\geq 0$, $\max\{|\Delta_1^n|,\dots,|\Delta_\ell^n|\}$ is a tight sequence of random variables for $n\geq 1$.*

PROOF. We will prove that for all $\varepsilon>0$ there exists $N>0$ such that for all $n\geq N$, $\mathbf{P}\{|\Delta_1^n|>N\}<\varepsilon$.

To prove the result for $|\Delta_2^n|,\dots,|\Delta_\ell^n|$, note that by Proposition 4.7, since $(A_i)_{i\in[k]}$ are almost surely distinct, we have

$$\mathbf{P}\left\{(A_i^n)_{i\in[k]}\text{ are distinct}\right\}\to 1$$

as $n\to\infty$. On the event $\{(A_i^n)_{i\in[k]}\text{ are distinct}\}$ the proof for $|\Delta_2^n|,\dots,|\Delta_\ell^n|$ is analogous to that for $|\Delta_1^n|$ and so we omit it.

Recall from Proposition 4.7 that $\sigma n^{-1/2}J_1^n\xrightarrow{\mathrm{d}} J_1$. Recalling also that A_1^n is such that $V_{J_1^n}=\widehat{V}_{A_1^n}(\Pi_{D^n})$, it follows that conditionally on $\widehat{D}_{A_1^n}^n=k$, $\Delta_1^n\overset{\mathrm{d}}{=}Y_{k,U_k}$, where $U_k\overset{\mathrm{d}}{=}\mathrm{Uniform}([k])$ and Y_{k,U_k} is distributed as a uniform entry of a displacement vector with law ν_k, independent of D^n. Fix $T>0$ large. We work on the event $\{J_1^n\leq T\sqrt{n}\}$. For $N>0$ and $K\geq 1$,

$$\begin{aligned}
&\mathbf{P}\left\{|\Delta_1^n|>N,\ J_1^n\leq T\sqrt{n}\right\}\\
&\leq \mathbf{P}\left\{\widehat{D}_{A_1^n}^n>K,\ J_1^n\leq T\sqrt{n}\right\}+\mathbf{P}\left\{|\Delta_1^n|>N,\ \widehat{D}_{A_1^n}^n\leq K,\ J_1^n\leq T\sqrt{n}\right\}\\
&\leq \mathbf{P}\left\{\widehat{D}_{A_1^n}^n>K,\ J_1^n\leq T\sqrt{n}\right\}+\sum_{k=2}^{K}\frac{k(k-1)\mu_k}{\sigma^2}\mathbf{P}\left\{|Y_{k,U_k}|>N\right\}\mathbf{P}\left\{J_1\leq\sigma T\right\}\\
&+\sum_{k=2}^{K}\left|\mathbf{P}\left\{|\Delta_1^n|>N,\widehat{D}_{A_1^n}^n=k,J_1^n\leq T\sqrt{n}\right\}-\frac{k(k-1)\mu_k}{\sigma^2}\mathbf{P}\left\{|Y_{k,U_k}|>N\right\}\mathbf{P}\left\{J_1\leq\sigma T\right\}\right|.
\end{aligned}$$

We have

$$\mathbf{P}\left\{|\Delta_1^n|>N, \widehat{D}^n_{A_1^n}=k, J_1^n\le T\sqrt{n}\right\}=\mathbf{P}\left\{|Y_{k,U_k}|>N\right\}\mathbf{P}\left\{\widehat{D}^n_{A_1^n}=k, J_1^n\le T\sqrt{n}\right\}$$

and so

$$\begin{aligned}&\mathbf{P}\left\{|\Delta_1^n|>N,\ J_1^n\le T\sqrt{n}\right\}\\&\le \mathbf{P}\left\{\widehat{D}^n_{A_1^n}>K,\ J_1^n\le T\sqrt{n}\right\}+\sum_{k=2}^{K}\frac{k(k-1)\mu_k}{\sigma^2}\mathbf{P}\left\{|Y_{k,U_k}|>N\right\}\\&\quad+\sum_{k=2}^{K}\left|\mathbf{P}\left\{\widehat{D}^n_{A_1^n}=k,\ J_1^n\le T\sqrt{n}\right\}-\frac{k(k-1)\mu_k}{\sigma^2}\mathbf{P}\left\{J_1\le\sigma T\right\}\right|.\end{aligned}\tag{4.17}$$

Since $k\mu_k\le 1$ for all k, it follows that (4.17) is at most

$$\begin{aligned}&\mathbf{P}\left\{\widehat{D}^n_{A_1^n}>K,\ J_1^n\le T\sqrt{n}\right\}+\frac{K-1}{\sigma^2}\mathbf{P}\left\{|Y_{\bar\xi,U_{\bar\xi}}|>N\right\}\\&\quad+K\max_{2\le k\le K}\left|\mathbf{P}\left\{\widehat{D}^n_{A_1^n}=k,\ J_1^n\le T\sqrt{n}\right\}-\frac{k(k-1)\mu_k}{\sigma^2}\mathbf{P}\left\{J_1\le\sigma T\right\}\right|.\end{aligned}\tag{4.18}$$

Fix $\varepsilon>0$. Since $|Y_{\bar\xi,U_{\bar\xi}}|$ is a random variable with support in $[0,\infty)$, we may take $M=M(K)>0$ large enough so that

$$\frac{K-1}{\sigma^2}\mathbf{P}\left\{|Y_{\bar\xi,U_{\bar\xi}}|>N\right\}<\frac{\varepsilon}{4}.$$

It remains to prove that for sufficiently large $n\ge 1$ and $K\ge 1$ the sum of the first and third terms in (4.18) is at most $3\varepsilon/4$. To this end, observe that for $i\ge 1$,

$$\mathbf{P}_{\widehat{D}^n}\left\{A_1^n=i,\ J_1^n\le T\sqrt{n}\ \middle|\ J_1^n\right\}=\frac{\widehat{D}_i^n-1}{\sum_{j=1}^{J_1^n-1}(\widehat{D}_j^n-1)}\mathbf{1}_{[1\le i\le J_1^n]}\mathbf{1}_{[J_1^n\le T\sqrt{n}]}.$$

Therefore, for any $k\ge 2$,

$$\mathbf{P}_{\widehat{D}^n}\left\{\widehat{D}^n_{A_1^n}=k,\ J_1^n\le T\sqrt{n}\ \middle|\ J_1^n\right\}=(k-1)\frac{\left|\left\{1\le i\le J_1^n\ :\ \widehat{D}_i^n=k\right\}\right|}{\sum_{j=1}^{J_1^n-1}(\widehat{D}_j^n-1)}\mathbf{1}_{[J_1^n\le T\sqrt{n}]}.$$

It follows that

$$\begin{aligned}&\mathbf{P}\left\{\widehat{D}^n_{A_1^n}=k, J_1^n\le T\sqrt{n}\right\}\\&=(k-1)\mathbf{E}\left[\frac{\left|\left\{1\le i<J_1^n\ :\ \widehat{D}_i^n=k\right\}\right|}{\sum_{j=1}^{J_1^n-1}(\widehat{D}_j^n-1)}\mathbf{1}_{[J_1^n\le T\sqrt{n}]}\right]\\&=(k-1)\mathbf{E}\left[\frac{\left|\left\{1\le i<J_1^n\ :\ \bar\xi_i=k\right\}\right|}{\sum_{j=1}^{J_1^n-1}(\bar\xi_j-1)}\mathbf{1}_{[J_1^n\le T\sqrt{n}]}\Theta^n\left(\bar\xi_1,\dots,\bar\xi_{\lfloor T\sqrt{n}\rfloor}\right)\right],\end{aligned}$$

where the final equality holds by Proposition 4.3.

By Proposition 4.7, $J_1^n=\Theta_{\mathbf{P}}(\sqrt{n})$, and so by a functional law of large numbers (see Lemma A.1 in the appendix),

$$\frac{\left|\left\{1\le i<J_1^n\ :\ \bar\xi_i=k\right\}\right|}{\sum_{j=1}^{J_1^n-1}(\bar\xi_j-1)}\xrightarrow{\mathrm{p}}\frac{k\mu_k}{\sigma^2}.$$

Combining this with Lemma 4.4 we obtain that as $n \to \infty$

$$\mathbf{P}\left\{\widehat{D}^n_{A^n_1} = k,\ J^n_1 \le T\sqrt{n}\right\} \to \frac{k(k-1)\mu_k}{\sigma^2}\mathbf{P}\left\{J_1 \le \sigma T\right\}. \tag{4.19}$$

Since $\sum_{k=1}^{\infty} k(k-1)\mu_k/\sigma^2 = 1$, we can take $K \ge 1$ and $T > 0$ large enough so that

$$\mathbf{P}\left\{J_1 \le \sigma T\right\} \sum_{k=2}^{K} \frac{k(k-1)\mu_k}{\sigma^2} > 1 - \frac{\varepsilon}{4}.$$

Further, by (4.19) we can take $n \ge 1$ large enough such that

$$\max_{2\le k\le K}\left|\mathbf{P}\left\{\widehat{D}^n_{A^n_1} = k,\ J^n_1 \le T\sqrt{n}\right\} - \frac{k(k-1)\mu_k}{\sigma^2}\mathbf{P}\left\{J_1 \le \sigma T\right\}\right| < \frac{\varepsilon}{4K}.$$

For such n and T, we have $\mathbf{P}\left\{\widehat{D}^n_{A^n_1} > K,\ J^n_1 \le T\sqrt{n}\right\} < \varepsilon/2$. The result follows. □

4.4. *Convergence to the continuous line-breaking construction.* We are now ready to complete the proof of Proposition 4.1.

PROOF OF PROPOSITION 4.1. In view of Proposition 4.7, it remains to prove (4.3) and (4.4).

For (4.3), we recall from the discussion at the start of Section 4 (where $(F^n_1, \dots, F^n_k)$ were defined) that at attachment points which are first repeats, the attachment is to the left with probability $1/2$ and to the right with probability $1/2$. By Proposition 4.7, the first k attachment points are distinct and are, therefore, all first repeats with probability tending to 1 as $n \to \infty$. The statement (4.3) follows.

For (4.4), we must consider the spatial locations of the vertices along the first k paths in the bijective construction. We work on the event that the paths $P^{(1)}, \dots, P^{(k)}$ terminate in vertices $1, 2, \dots, k$ respectively, which we have already shown holds with high probability as $n \to \infty$. For the first path, we have

$$L^n(\lfloor tn^{1/2}\rfloor \wedge (J^n_1 - 1)) = \sum_{j=1}^{\lfloor tn^{1/2}\rfloor \wedge (J^n_1 - 1)} Y_{\widehat{D}^n_j, U_j}$$

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

$$L^n((J^n_i + \lfloor tn^{1/2}\rfloor) \wedge (J^n_{i+1} - 1)) = L^n(A^n_i + i - 2) + \Delta^n_i + \sum_{j=J^n_i+1}^{(J^n_i + \lfloor tn^{1/2}\rfloor)\wedge(J^n_{i+1}-1)} Y_{\widehat{D}^n_j, U_j}.$$

The desired convergence then follows from Propositions 4.5 and 4.7 and Lemma 4.8. □

5. Tightness. We assume throughout the section that μ is critical and has finite variance $\sigma^2 \in (0, \infty)$, and that [**A1**] and [**A2**] hold.

Let $k \ge 1$. Recall that T^k_n is the subtree of T_n spanned by the root and the vertices $v_{U^n_1}, \dots v_{U^n_k} \in \mathrm{T}_n$, where $(U^n_1, \dots, U^n_k)$ is a uniformly random k-set sampled from $[n]$ and, for $i \in [n]$, v_i is the i-th vertex in the lexicographical order of T_n. In what follows we write $(U^{n,k}_{(1)}, \dots, U^{n,k}_{(k)})$ for the increasing rearrangement of $(U^n_1, \dots, U^n_k)$. Further, recall from Section 2 that $\mathbf{T}_n = (\mathrm{T}_n, Y)$ is the (μ, ν)-branching random walk conditioned to have size n.

PROPOSITION 5.1. *Suppose that [**A1**] holds. Then for all $\gamma > 0$,*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U_{(i)}^{n,k}-1, U_{(i+1)}^{n,k}-1]} |H_n(s) - H_n(t)| > \gamma n^{1/2} \right\} = 0 \tag{5.1}$$

*and, additionally, if [**A2**] holds, then*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U_{(i)}^{n,k}-1, U_{(i+1)}^{n,k}-1]} |R_n(s) - R_n(t)| > \gamma n^{1/4} \right\} = 0. \tag{5.2}$$

Under [**A1**], we have that

$$\left(\frac{H_n(nt)}{\sqrt{n}}\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\le t\le 1}$$

in $\mathbf{C}([0,1],\mathbb{R})$ and so (5.1) holds. It follows that we only need to prove (5.2).

Let us immediately observe that the vertices of the tree T_n either belong to T_n^k or belong to a subtree hanging off T_n^k. In Proposition 4.1, we showed the convergence of the spatial locations along the subtree T_n^k to those given by a Brownian motion indexed by $\mathcal{T}^k$. This has the consequence that for values $s,t \in [U_{(i)}^{n,k}-1, U_{(i+1)}^{n,k}-1]$ such that both corresponding vertices lie in T_n^k, we have that $|R_n(s) - R_n(t)|$ is bounded above by the maximum modulus $\Upsilon_i^{n,k}$ of an increment of the location process along the path from $U_{(i)}^{n,k}$ to $U_{(i+1)}^{n,k}$ in T_n^k. Moreover, this upper bound converges in distribution on rescaling to the analogous quantity in the limit tree, which has the same distribution as the maximum modulus Υ_i^k of an increment of β times a Brownian motion run for time D_i^k, where D_i^k is the distance between the ith and $(i+1)$st leaves of $(2/\sigma)\mathcal{T}^k$ in planar order. We thus have that

$$\max_{0\le i\le k} \Upsilon_i^k \overset{\mathrm{d}}{=} \beta\sqrt{\frac{2}{\sigma}} \max_{0\le i\le k} \sup_{s,t\in[U_{(i)}^k, U_{(i+1)}^k]} |\mathbf{r}_s - \mathbf{r}_t|.$$

But

$$\max_{0\le i\le k} \left(U_{(i+1)}^k - U_{(i)}^k\right) \xrightarrow{\text{a.s.}} 0$$

as $k\to\infty$ and so, since $\mathbf{r}$ is uniformly continuous, we may deduce that for any $\gamma > 0$,

$$\lim_{k\to\infty} \lim_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \Upsilon_i^{n,k} > \gamma n^{1/4} \right\} = \lim_{k\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \Upsilon_i^k > \gamma \right\} = 0. \tag{5.3}$$

For values $s,t \in [U_{(i)}^{n,k}, U_{(i+1)}^{n,k}]$ for some $0 \le i \le k$ such that at least one of the corresponding vertices does not lie in T_n^k, we may bound $|R_n(s) - R_n(t)|$ by $\Upsilon_i^{n,k}$ plus twice the maximum modulus of the difference in spatial location between the parent in T_n^k of the root of a pendant subtree and some other vertex inside the tree. We have already dealt with $\Upsilon_i^{n,k}$, and so it remains to deal with the pendant subtrees. Before we can do so, we need to do some truncation of the displacements.

Fix $\gamma > 0$ and $\delta \in (0, 1/4)$. We will consider three "restrictions" of the branching random walk $\mathbf{T}_n = (\mathrm{T}_n, Y)$, which we denote by $\mathbf{T}_{n,\delta} = (\mathrm{T}_n, Y_{n,\delta})$, $\mathbf{T}_{n,\delta}^\gamma = (\mathrm{T}_n, Y_{n,\delta}^\gamma)$, and $\mathbf{T}_n^\gamma = (\mathrm{T}_n, Y_n^\gamma)$. These branching random walks capture the "typical", "mid-range", and "large" spatial displacements in $\mathbf{T}_n$.

1. **(typical displacements):** $Y_{n,\delta} = (Y_{n,\delta}^{(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ is such that for $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$,
$$Y_{n,\delta}^{(v)} = Y^{(v)} \mathbf{1}_{[\|Y^{(v)}\|_\infty \le n^{1/4-\delta}]}.$$
2. **(mid-range displacements):** $Y_{n,\delta}^{\gamma} = (Y_{n,\delta}^{\gamma,(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ is such that for all $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$,
$$Y_{n,\delta}^{\gamma,(v)} = Y^{(v)} \mathbf{1}_{[n^{1/4-\delta} < \|Y^{(v)}\|_\infty \le \gamma n^{1/4}]}.$$
3. **(large displacements):** $Y_{n}^{\gamma} = (Y_{n}^{\gamma,(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ is such that for $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$,
$$Y_{n}^{\gamma,(v)} = Y^{(v)} \mathbf{1}_{[\|Y^{(v)}\|_\infty > \gamma n^{1/4}]}.$$

For $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$, the vectors $Y_{n,\delta}^{(v)}, Y_{n,\delta}^{\gamma,(v)}, Y_{n}^{\gamma,(v)}$ are all of length $c(v, \mathrm{T}_n)$; however, in what follows we will not refer to their individual entries.

Let $R_{n,\delta}$, $R_{n,\delta}^{\gamma}$, and R_n^{γ} denote the functions encoding the spatial locations of the branching random walks $\mathbf{T}_{n,\delta}$, $\mathbf{T}_{n,\delta}^{\gamma}$, and $\mathbf{T}_n^{\gamma}$, respectively. Then, for all n large enough so that $n^{1/4-\delta} \le \gamma n^{1/4}$,
$$R_n = R_{n,\delta} + R_{n,\delta}^{\gamma} + R_n^{\gamma}.$$
By the triangle inequality, for all $\gamma > 0$, we then have
$$\begin{aligned} &\max_{0 \le i \le k} \sup_{s,t \in [U_{(i)}^{n,k}-1, U_{(i+1)}^{n,k}-1]} |R_n(s) - R_n(t)| \\ &\le \max_{0 \le i \le k} \sup_{s,t \in [U_{(i)}^{n,k}-1, U_{(i+1)}^{n,k}-1]} |R_{n,\delta}(s) - R_{n,\delta}(t)| + 2\|R_{n,\delta}^{\gamma}\|_\infty + 2\|R_n^{\gamma}\|_\infty. \end{aligned} \tag{5.4}$$
We deal with each of these three terms separately.

5.1. *Large and mid-range displacements.* Under assumption [**A2**], we show that the probability that there is a displacement in $\mathbf{T}_n$ with modulus exceeding $\gamma n^{1/4}$ goes to zero, so that the contribution of the large displacements is negligible.

PROPOSITION 5.2. *For all $\gamma > 0$, as $n \to \infty$,*
$$\mathbf{P}\left\{\|R_n^{\gamma}\|_\infty > \gamma n^{1/4}\right\} = o(1).$$

PROOF. Let
$$M_n^{\gamma} := \left|\left\{v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n \ : \ \|Y^{(v)}\|_\infty > \gamma n^{1/4}\right\}\right|.$$
It suffices to prove that $\mathbf{P}\{M_n^{\gamma} > 0\} \to 0$ as $n \to \infty$. To this end, let $\xi_1, \dots, \xi_n$ be IID random variables with distribution μ. By assumption [**A2**],
$$\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > \gamma n^{1/4}\right\} = o(n^{-1}).$$
Fixing $\varepsilon > 0$, this implies that for n large enough,
$$\widetilde{M}_n^{\gamma} := \left|\left\{i \in [n] \ : \ \|Y_{\xi_i}\|_\infty > \gamma n^{1/4}\right\}\right| \preceq_{st} \mathrm{Bin}\left(n, \frac{\varepsilon}{n}\right), \tag{5.5}$$
where $\preceq_{st}$ denotes stochastic domination. It follows from a Chernoff bound that there exists $c > 0$ such that for n sufficiently large,
$$\mathbf{P}\left\{\widetilde{M}_n^{\gamma} \ge n^{\varepsilon}\right\} \le \exp\left(-cn^{\varepsilon}\right).$$

Since $\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\} = \Theta(n^{-1/2})$, we obtain

$$\begin{aligned}\mathbf{P}\{M_n^\gamma \geq n^\varepsilon\} &= \mathbf{P}\left\{\widetilde{M}_n^\gamma \geq n^\varepsilon \mid \sum_{i=1}^n \xi_i = n-1\right\} \\ &\leq \frac{\mathbf{P}\left\{\widetilde{M}_n^\gamma \geq n^\varepsilon\right\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}} = O\left(n^{1/2}\exp(-cn^\varepsilon)\right).\end{aligned}$$

Let $\widetilde{S}_n^\gamma := \sum_{i=1}^n \xi_i \mathbf{1}_{[\|Y_{\xi_i}\|_\infty > \gamma n^{1/4}]}$ and let $S_n^\gamma := \sum_{v\in v(\mathrm{T}_n)} c(v,\mathrm{T}_n)\mathbf{1}_{[\|Y^{(v)}\|_\infty > \gamma n^{1/4}]}$. Since $\mathbf{E}\left[\xi^3\right] < \infty$, by [17, Corollary 19.11], both $\max_{1\leq i\leq n}\xi_i$ and $\max_{v\in v(\mathrm{T}_n)} c(v,\mathrm{T}_n)$ are $O_{\mathbf{P}}(n^{1/3})$, and so

$$\begin{aligned}\mathbf{P}\left\{M_n^\gamma \geq n^\varepsilon \text{ or } S_n^\gamma \geq n^{1/3+\varepsilon}\right\} &\leq o(1) + \mathbf{P}\left\{S_n^\gamma \geq n^{1/3+\varepsilon} \cap \max_{v\in v(\mathrm{T}_n)} c(v,\mathrm{T}_n) \leq n^{1/3}\right\} \\ &\leq o(1) + \mathbf{P}\left\{\sum_{v\in v(\mathrm{T}_n)} \mathbf{1}_{[\|Y^{(v)}\|_\infty \geq \gamma n^{1/4}]} > n^\varepsilon\right\} \\ &\leq o(1) + \frac{\mathbf{P}\left\{\mathrm{Bin}\left(n, \frac{\varepsilon}{n}\right) > n^\varepsilon\right\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}} = o(1), \qquad (5.6)\end{aligned}$$

where the final inequality holds by (5.5). Further, for $\xi_1^n, \xi_2^n, \ldots$ independent random variables such that for each $i \geq 1$, ξ_i^n is distributed as ξ_i conditional on $\|Y_{\xi_i}\|_\infty < \gamma n^{1/4}$, we have that

$$\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1 \,\middle|\, \widetilde{S}_n^\gamma, \widetilde{M}_n^\gamma\right\} = \mathbf{P}\left\{\widetilde{S}_n^\gamma + \sum_{i=1}^{n-\widetilde{M}_n^\gamma} \xi_i^n = n-1 \,\middle|\, \widetilde{S}_n^\gamma, \widetilde{M}_n^\gamma\right\}.$$

Therefore,

$$\begin{aligned}\mathbf{P}\{M_n^\gamma > 0\} &= \mathbf{P}\left\{0 < M_n^\gamma < n^\varepsilon,\ S_n^\gamma < n^{1/3+\varepsilon}\right\} + o(1) \\ &= \mathbf{P}\left\{0 < \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon} \,\middle|\, \sum_{i=1}^n \xi_i = n-1\right\} + o(1) \\ &= \frac{\mathbf{P}\left\{0 < \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}, \sum_{i=1}^n \xi_i = n-1\right\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}} + o(1) \\ &= \mathbf{E}\left[\frac{\mathbf{P}\left\{\widetilde{S}_n^\gamma + \sum_{i=1}^{n-\widetilde{M}_n^\gamma} \xi_i^n = n-1 \,\middle|\, \widetilde{S}_n^\gamma, \widetilde{M}_n^\gamma\right\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}} \mathbf{1}_{[0<\widetilde{M}_n^\gamma<n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}]}\right] + o(1).\end{aligned}$$

By a quantitative local limit theorem (see Lemma A.3 in the appendix), we obtain that as $n \to \infty$

$$\frac{\mathbf{P}\{\sum_i^{n-m} \xi_i^n = n-1-s\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}} \to 1,$$

uniformly over all $0 < m < n^\varepsilon$ and $0 < s < n^{1/3+\varepsilon}$. It follows that

$$\mathbf{P}\{M_n^\gamma > 0\} = \mathbf{P}\left\{0 < \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}\right\} + o(1) \leq \mathbf{P}\left\{\widetilde{M}_n^\gamma > 0\right\} + o(1).$$

The result follows since for n sufficiently large $\widetilde{M}_n^\gamma \preceq_{st} \mathrm{Bin}(n, \varepsilon/n)$, and $\varepsilon > 0$ is arbitrary. □

Similarly to the large displacements, the mid-range displacements are also negligible on the order of $n^{-1/4}$. However, the argument required to prove this is more refined.

PROPOSITION 5.3. *Fix $\gamma > 0$. For $\delta > 0$ sufficiently small, as $n \to \infty$,*

$$\mathbf{P}\left\{\|R_{n,\delta}^\gamma\|_\infty > \gamma n^{1/4}\right\} = o(1).$$

To prove this proposition, we will require some further results pertaining to the positions of non-typical displacements in the branching random walk $\mathbf{T}_n$. More specifically, we will need to study the law of the *number* and *positions* of the vertices $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ such that $\|Y^{(v)}\|_\infty > n^{1/4-\delta}$, for fixed, small $\delta > 0$. The next lemma pertains to the number of such vertices.

LEMMA 5.4. *For $\delta > 0$ sufficiently small,*

$$\left|\left\{v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n \text{ such that } \|Y^{(v)}\|_\infty > n^{1/4-\delta}\right\}\right| = o_{\mathbf{P}}(n^{1/12}).$$

PROOF. Let $\xi_1, \dots, \xi_n$ be IID with distribution μ. By [**A2**] there exists $C > 0$ such that $\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > n^{1/4-\delta}\right\} \leq Cn^{-1+4\delta}$. It follows that

$$A_n := \left|\left\{i \in [n] \,:\, \|Y_{\xi_i}\|_\infty > n^{1/4-\delta}\right\}\right| \preceq_{st} \mathrm{Bin}\left(n, Cn^{-1+4\delta}\right).$$

By a Chernoff bound, this implies that for $\delta \in (0, 1/48)$, and $n \geq 1$ sufficiently large, for any $\varepsilon > 0$,

$$\begin{aligned}
\mathbf{P}\left\{A_n > \varepsilon n^{1/12}\right\} &\leq \mathbf{P}\left\{\mathrm{Bin}\left(n, Cn^{-1+4\delta}\right) > \varepsilon n^{1/12}\right\} \\
&= \mathbf{P}\left\{\mathrm{Bin}\left(n, Cn^{-1+4\delta}\right) > Cn^{4\delta}\left(1 + \left(\frac{\varepsilon}{C} n^{1/12-4\delta} - 1\right)\right)\right\} \\
&= O\left(\exp(-n^{4\delta})\right),
\end{aligned}$$

and so

$$\mathbf{P}\left\{A_n > \varepsilon n^{1/12} \;\middle|\; \sum_{i=1}^n \xi_i = n - 1\right\} = O\left(n^{1/2}\exp(-n^{4\delta})\right) = o(1). \qquad \square$$

We say that two vertices $u, v \in \mathcal{U}$ are *ancestrally related* if either $u \prec v$ or $v \prec u$. The following lemma establishes that with high probability there are no ancestrally related vertices $u, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ such that $\|Y^{(u)}\|_\infty \wedge \|Y^{(v)}\|_\infty > n^{1/4-\delta}$.

PROPOSITION 5.5. *For $\delta > 0$ sufficiently small, as $n \to \infty$,*

$$\mathbf{P}\left\{\exists u, v \in \mathbf{T}_n, u \prec v, \text{ such that } \|Y^{(u)}\|_\infty \wedge \|Y^{(v)}\|_\infty > n^{1/4-\delta}\right\} = o(1).$$

The proof of this proposition relies on an application of the technical lemma, Lemma 3.4, which we prove in the appendix.

PROOF. We generate T_n using the bijective construction $B(\Pi_{D^n})$ described in Section 2.2. Sample the displacement vectors $(Y_{D_i^n})_{1\le i\le n}$ with $Y_{D_i^n} = (Y_{D_i^n,1}, \dots, Y_{D_i^n,D_i^n})$, and let

$$\mathcal{B} = \left\{ i \in [n] \,:\, \|Y_{D_i^n}\|_\infty > n^{1/4-\delta} \right\}.$$

Then

$$\begin{aligned}
&\mathbf{P}\left\{ \exists u, v \in \mathbf{T}_n, u \prec v, \text{ such that } \|Y^{(u)}\|_\infty \wedge \|Y^{(v)}\|_\infty > n^{1/4-\delta} \right\} \\
&\quad \le \mathbf{P}\left\{ \max_{0\le i\le n} H_n(i) > t\sqrt{n} \right\} + \mathbf{P}\left\{ |\mathcal{B}| > sn^{1/12} \right\} + \mathbf{P}\left\{ \max_{1\le i\le n} D_i^n > Tn^{1/3} \right\} \\
&\quad + \mathbf{P}\left\{ \left\{ \max_{1\le i\le n} D_i^n \le Tn^{1/3}, |\mathcal{B}| \le sn^{1/12} \right\} \cap \left\{ \exists i,j \in \mathcal{B} : i \prec j, d_n(i,j) \le t\sqrt{n} \right\} \right\} \qquad (5.7)
\end{aligned}$$

where, for vertices $i, j \in v(\mathrm{T}_n)$, $d_n(i,j)$ denotes the length of the shortest path between i and j in $B(\Pi_{D^n}) \overset{\mathrm{d}}{=} \mathrm{T}_n$. Take t and T large enough so that $\mathbf{P}\left\{\max_{0\le i\le n} H_n(i) > t\sqrt{n}\right\} < \varepsilon/4$, and $\mathbf{P}\left\{\max_{1\le i\le n} D_i^n > Tn^{1/3}\right\} < \varepsilon/4$. (The latter inequality is possible by [17, Corollary 19.11] since $\mathbf{E}\left[\xi^3\right] < \infty$.) By Lemma 5.4, we may take n large enough so that $\mathbf{P}\left\{|\mathcal{B}| > sn^{1/12}\right\} < \varepsilon/4$. Therefore, for t, T and n sufficiently large, (5.7) is at most

$$\frac{3\varepsilon}{4} + \mathbf{P}\left\{ \left\{ \max_{1\le i\le n} D_i^n \le Tn^{1/3}, |\mathcal{B}| \le sn^{1/12} \right\} \cap \left\{ \exists i,j \in \mathcal{B} : i \prec j, d_n(i,j) \le t\sqrt{n} \right\} \right\}.$$

Then by Lemma 3.4 with $\mathrm{d} = D^n$, $K \le sn^{1/12}$, $\Delta \le Tn^{1/3}$, and $b = t\sqrt{n}$, for n sufficiently large, this is at most

$$\frac{3\varepsilon}{4} + sn^{1/12}\left(1 - \left(1 - \frac{sTn^{-7/12}}{1 - n^{-1} - tTn^{-1/6}}\right)^{t\sqrt{n}}\right).$$

The result follows by taking $s > 0$ small enough and n large enough so that

$$sn^{1/12}\left(1 - \left(1 - \frac{sTn^{-7/12}}{1 - n^{-1} - tTn^{-1/6}}\right)^{t\sqrt{n}}\right) < \frac{\varepsilon}{4},$$

which is possible since

$$sn^{1/12}\left(1 - \left(1 - \frac{sTn^{-7/12}}{1 - n^{-1} - tTn^{-1/6}}\right)^{t\sqrt{n}}\right) < s^2Tt\frac{1}{1 - n^{-1} - tTn^{-1/6}},$$

for n large enough because $(1-x)^r > 1 - rx$ for $x < 1$ and $r > 1$. □

LEMMA 5.6. *Let $v^*(\mathrm{T}_n) \subseteq v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ be the set of vertices $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ such that $\|Y^{(v)}\|_\infty \le n^{1/4-\delta}$ and there exists an ancestor $u \prec v$ with $\|Y^{(u)}\|_\infty > n^{1/4-\delta}$. For $\delta > 0$ sufficiently small, $v^*(\mathrm{T}_n) = o_{\mathbf{P}}(n)$.*

PROOF. The result holds if and only if the probability that a uniformly random vertex in $v \in v(\mathrm{T}_n)$ is ancestrally related to a vertex $u \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ with $\|Y^{(u)}\|_\infty > n^{1/4-\delta}$ is $o_{\mathbf{P}}(1)$. By exchangeability, this holds if and only if the probability that vertex 1 is ancestrally related to a vertex $u \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ with $\|Y^{(u)}\|_\infty > n^{1/4-\delta}$ is $o_{\mathbf{P}}(1)$. To prove this we may adapt the proof of Proposition 5.5 by including vertex 1 in the set $\mathcal{B}$. Then by Lemma 5.4, $|\mathcal{B}| = o_{\mathbf{P}}(n^{1/12})$ still holds, and so the proof carries over verbatim. □

As an immediate consequence of Lemma 5.6, with probability $1 - o(1)$ none of the increments of the branching random walk $\mathbf{T}_{n,\delta}^{\gamma}$ are ancestrally related with high probability, and Proposition 5.3 follows.

5.2. *Typical displacements.* In this subsection we will prove the following proposition.

PROPOSITION 5.7. *For all $\gamma > 0$,*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |R_{n,\delta}(s) - R_{n,\delta}(t)| > \gamma n^{1/4} \right\} = 0.$$

Notice that $R_{n,\delta}$ is equal in distribution to the function encoding the spatial locations of the branching random walk with underlying tree T_n and displacements $Y^{n,\delta} = (Y^{n,\delta,(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ such that if $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ has k children, then $Y^{n,\delta,(v)}$ has the same distribution as

$$Y^{n,\delta}_k = (Y^{n,\delta}_{k,1}, \dots Y^{n,\delta}_{k,k}) := \begin{cases} (Y_{k,1}, \dots, Y_{k,k}) & \text{if } \max_{1\le j\le k} |Y_{k,j}| \le n^{1/4-\delta}, \\ (0, \dots, 0) & \text{else.} \end{cases}$$

This branching random walk is *not* globally centered, and in particular has "global" drift $\mathbf{E}\left[Y^{n,\delta}_{\bar\xi, U_{\bar\xi}}\right]$. Thus for all $t \in [0, n]$ we have that

$$R_{n,\delta}(t) \stackrel{\mathrm{d}}{=} \breve{R}_{n,\delta}(t) + \mathbf{E}\left[Y^{n,\delta}_{\bar\xi, U_{\bar\xi}}\right] \cdot H_n(t),$$

where $\breve{R}_{n,\delta} : [0, n] \to \mathbb{R}$ is the function encoding the spatial locations of the globally centered branching random walk $(\mathrm{T}_n, \breve{Y}^{n,\delta})$ for which, conditionally on T_n, $\breve{Y}^{n,\delta} = (\breve{Y}^{n,\delta,(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ is a vector of independent random variables, such that if $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ has k children then $\breve{Y}^{n,\delta,(v)}$ has the same distribution as

$$\breve{Y}^{n,\delta}_k := Y^{n,\delta}_k - \mathbf{E}\left[Y^{n,\delta}_{\bar\xi, U_{\bar\xi}}\right] = Y^{n,\delta}_k - \mathbf{E}\left[Y_{\bar\xi, U_{\bar\xi}} \mathbf{1}_{[\|Y_{\bar\xi}\|_\infty \le n^{1/4-\delta}]}\right].$$

Moreover, by the triangle inequality, for all $\gamma > 0$,

$$\begin{aligned} &\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |R_{n,\delta}(s) - R_{n,\delta}(t)| > \gamma n^{1/4} \right\} \\ &\le \lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |\breve{R}_{n,\delta}(s) - \breve{R}_{n,\delta}(t)| > \frac{\gamma}{2} n^{1/4} \right\} \\ &\quad + \limsup_{n\to\infty} \mathbf{P}\left\{ \left|\mathbf{E}\left[Y^{n,\delta}_{\bar\xi, U_{\bar\xi}}\right]\right| \cdot \|H_n\|_\infty > \frac{\gamma}{4} n^{1/4} \right\}. \end{aligned}$$

LEMMA 5.8. *It holds that, as $n \to \infty$,*

$$\left|\mathbf{E}\left[Y^{n,\delta}_{\bar\xi, U_{\bar\xi}}\right]\right| = O(n^{-5/12+5\delta/3}) \quad \textit{and} \quad \mathbf{Var}\left\{Y^{n,\delta}_{\bar\xi, U_{\bar\xi}}\right\} \to \beta^2.$$

This result is a special case of Lemma A.11, which is stated and proved in the appendix. Since $\|H_n\|_\infty = O_{\mathbf{P}}(\sqrt{n})$, Lemma 5.8 implies that for δ sufficiently small,

$$\limsup_{n\to\infty} \mathbf{P}\left\{ \left|\mathbf{E}\left[Y^{n,\delta}_{\bar\xi, U_{\bar\xi}}\right]\right| \cdot \|H_n\|_\infty > \gamma n^{1/4} \right\} = 0.$$

It follows that to prove Proposition 5.7, it suffices to prove that for all $\gamma > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |\breve{R}_{n,\delta}(s) - \breve{R}_{n,\delta}(t)| > \gamma n^{1/4} \right\} = 0.$$

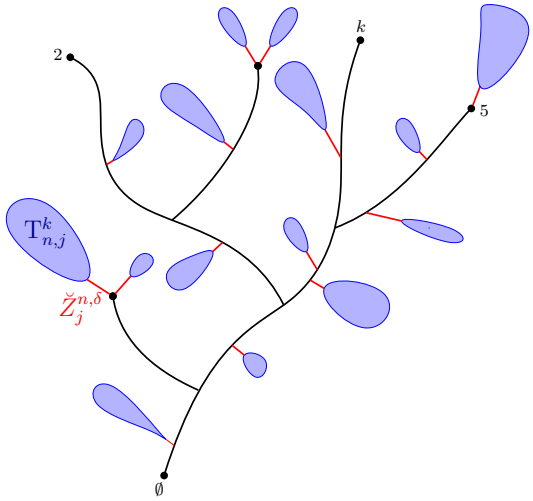

FIG 5. *In black, the tree* T^k_n. *In blue, the forest* $\mathrm{F}^k_n = (\mathrm{T}^k_{n,j})_{j\geq 1}$. *The root of tree* $\mathrm{T}^k_{n,j}$ *is displaced* $\breve{Z}^{n,\delta}_j$ *away from its parent in* T^k_n.

As discussed above, we need to deal with the maximum modulus of the difference in spatial location (for the branching random walk $\mathrm{T}_{n,\delta}$) between the parent of the root of a pendant subtree and a vertex of that subtree. There are

$$c(\mathrm{T}^k_n) := \sum_{v\in V(\mathrm{T}^k_n)} (c(v, T_n) - 1) + 1$$

edges in T_n with one endpoint in T^k_n and another in $\mathrm{T}_n \setminus \mathrm{T}^k_n$. Conditionally on T^k_n, if we remove all such edges we obtain a Bienaymé(μ) forest conditioned to have $n - |V(\mathrm{T}^{(k)}_n)|$ vertices and $c(\mathrm{T}^k_n)$ trees. We denote this forest by $\mathrm{F}^k_n = (\mathrm{T}^k_{n,j})_{j\geq 1}$, where the trees are listed in decreasing order of size, and $|\mathrm{T}^k_{n,j}| = 0$ for $j > c(\mathrm{T}^k_n)$. Write $\|\breve{R}_{n,\delta}(\mathrm{T}^k_{n,j})\|_\infty$ for maximum modulus of the difference in spatial location between the root and any other vertex of $\mathrm{T}^k_{n,j}$.

The trees $(\mathrm{T}^k_{n,j})_{j\geq 1}$ are independent Bienaymé trees, conditioned on their sizes. Therefore, conditionally on F^k_n, we have $\|\breve{R}_{n,\delta}(\mathrm{T}^k_{n,j})\|_\infty \overset{\mathrm{d}}{=} \|\breve{R}_{|\mathrm{T}^k_{n,j}|,\delta}\|_\infty$. Moreover, displacements on the tree $\mathrm{T}^k_{n,j}$ (from the branching random walk $\mathbf{T}_{n,\delta}$) depend on those in other parts of T_n only through the displacement $\breve{Z}^{n,\delta}_j$ of the root of $\mathrm{T}^k_{n,j}$ away from its parent in T^k_n; see Figure 5.

It follows that

$$\max_{0\leq i\leq k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |\breve{R}_{n,\delta}(s) - \breve{R}_{n,\delta}(t)|$$
$$\leq \max_{0\leq i\leq k} \Upsilon^{n,k}_i + 2 \max_{1\leq j\leq c(\mathrm{F}^k_n)} \left(\left\| \breve{R}_{n,\delta}(\mathrm{T}^k_{n,j}) \right\|_\infty + |\breve{Z}^{n,\delta}_j| \right).$$

Consequently, using (5.3), in order to prove Proposition 5.7, it is sufficient to prove that for $\gamma > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{1\leq j\leq c(\mathrm{F}^k_n)} \left(\left\| \breve{R}_{n,\delta}(\mathrm{T}^k_{n,j}) \right\|_\infty + |\breve{Z}^{n,\delta}_j| \right) \geq \gamma n^{1/4} \right\} = 0.$$

The proof requires two key ingredients: (1) a scaling limit for the sizes of the trees in F^k_n; (2) quantitative control on the tail of $\|\breve{R}_{n,\delta}\|_\infty$. We begin by establishing (1).

PROPOSITION 5.9. *As $n \to \infty$,*

$$\frac{c(\mathrm{T}_n^k)}{\sigma\sqrt{n}} \xrightarrow{\mathrm{d}} J_k, \tag{5.8}$$

where J_k is $\mathrm{Gamma}(k, 1/2)$ distributed. Jointly with this convergence, we have

$$\frac{\sigma}{n - |V(\mathrm{T}_n^k)|}(|\mathrm{T}_{n,j}^k|, j \geq 1) \xrightarrow{\mathrm{d}} (|\gamma_j^k|, j \geq 1), \tag{5.9}$$

where, conditionally on J_k, $(|\gamma_j^k|, j \geq 1)$ lists the sizes of the excursions above the past minimum of a Brownian motion stopped on first hitting $-J_k$, listed in decreasing order.

PROOF. By Skorokhod's representation theorem we may work in a probability space where the convergence in Proposition 4.7 holds almost surely so that in particular as $n \to \infty$, $\sigma n^{-1/2} J_k^n \xrightarrow{\text{a.s.}} J_k$.

Let $T > 0$ and recall the event

$$\mathcal{G}_{n,k}(T) = \left\{ \left\{ \widehat{V}_1(\Pi_{D^n}), \ldots, \widehat{V}_{\lfloor T\sqrt{n} \rfloor}(\Pi_{D^n}) \right\} \cap \{1, \ldots, k\} = \emptyset, N_n \geq \lfloor T\sqrt{n} \rfloor \right\}$$

from Lemma 4.6. On $\mathcal{G}_{n,k}(T) \cap \{J_k^n \leq \lfloor T\sqrt{n} \rfloor\}$, the tree T_n^k is precisely the subtree of T_n spanned by the root and the vertices $1, \ldots, k$. Therefore, on $\mathcal{G}_{n,k}(T) \cap \{J_k^n \leq \lfloor T\sqrt{n} \rfloor\}$,

$$\frac{V(\mathrm{T}_n^k)}{\sqrt{n}} = \frac{J_k^n}{\sqrt{n}}.$$

Since $T > 0$ is arbitrary and $\sigma n^{-1/2} J_k^n \xrightarrow{\text{a.s.}} J_k$ we obtain that $n - |V(\mathrm{T}_n^k)| = n - o_{\mathbf{P}}(n)$. Hence, we are essentially considering a forest of Bienaymé trees conditioned to have n vertices. We now need to show that the number of trees in such a forest is $\sim \sigma\sqrt{n} J_k$. We note that on the event $\mathcal{G}_{n,k}(T) \cap \{J_k^n \leq T\sqrt{n}\}$, there are $\sum_{i=1}^{J_k^n - k}(\widehat{D}_i^n - 1) + \sum_{i=1}^k D_{J_i^n}^n$ subtrees of T_n whose roots have a parent in T_n^k, and $(k-1)$ branch points in T_n^k. Therefore, for $s \geq 0$,

$$\begin{aligned} &\mathbf{P}\left\{ \frac{c(\mathrm{T}_n^k)}{\sigma\sqrt{n}} \geq s,\ \mathcal{G}_{n,k}(T),\ J_k^n \leq T\sqrt{n} \right\} \\ &= \mathbf{P}\left\{ \frac{1}{\sigma\sqrt{n}} \left(\sum_{i=1}^{J_k^n - k} (\widehat{D}_i^n - 1) + \sum_{i=1}^k D_{J_i^n}^n - (k-1) \right) \geq s,\ \mathcal{G}_{n,k}(T),\ J_k^n \leq T\sqrt{n} \right\}. \end{aligned} \tag{5.10}$$

Since $\sigma n^{-1/2} J_k^n \xrightarrow{\text{a.s.}} J_k$, by Proposition 4.5,

$$\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{J_k^n - k} (\widehat{D}_i - 1) \xrightarrow{\mathrm{d}} J_k.$$

Combining this with Lemma 4.6 and Proposition 4.7, we obtain that (5.10) converges to

$$\mathbf{P}\left\{ J_k > s, J_k \leq \sigma T \right\}.$$

Then (5.8) follows as $T > 0$ is arbitrary. The scaling limit in (5.9) now follows from [27, Proposition 1.4] and [7, Lemma 11]. □

The control on $\|\breve{R}_{n,\delta}\|_\infty$ needed to prove Proposition 5.7 is given by the next proposition.

PROPOSITION 5.10. *There exists $A > 0$ such that for all $\gamma > 0$, $\delta \in (0, 1/4)$, and $n \geq 1$,*

$$\mathbf{P}\left\{\|\breve{R}_{n,\delta}\|_\infty > \gamma n^{1/4}\right\} \leq \frac{A}{\gamma^8}.$$

The proof of this proposition is long and somewhat technical, so we postpone it until Section 6.

PROOF OF PROPOSITION 5.7 ASSUMING PROPOSITION 5.10. By Skorokhod's representation theorem, we may assume that we are working on a probability space where the convergence in Proposition 4.7 is almost sure. In particular, $\sigma n^{-1/2} J_k^n \xrightarrow{\text{a.s.}} J_k$ as $n \to \infty$.

As argued above, it remains to show that, for $\gamma > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{\max_{1\leq j\leq c(\mathrm{F}_n^k)} \left(\left\|\breve{R}_{n,\delta}(\mathrm{T}_{n,j}^k)\right\|_\infty + |\breve{Z}_j^{n,\delta}|\right) \geq \gamma n^{1/4}\right\} = 0.$$

Since $(\mathrm{T}_{n,j}^k)_{1\leq j\leq c(\mathrm{F}_n^k)}$ are independent Bienaymé(μ) trees conditionally on their sizes, we obtain

$$\begin{aligned}
&\mathbf{P}\left\{\max_{1\leq j\leq c(\mathrm{F}_n^k)} \left(\left\|\breve{R}_{n,\delta}(\mathrm{T}_{n,j}^k)\right\|_\infty + |\breve{Z}_j^{n,\delta}|\right) \geq \gamma n^{1/4}\right\} \\
&= \mathbf{E}\left[\mathbf{P}\left\{\max_{1\leq j\leq c(\mathrm{F}_n^k)} \left(\left\|\breve{R}_{|\mathrm{T}_{n,j}^k|,\delta}\right\|_\infty + |\breve{Z}_j^{n,\delta}|\right) \geq \gamma n^{1/4} \mid \mathrm{F}_n^k, (\breve{Z}_j^{n,\delta})_{j\geq 1}\right\}\right] \\
&\leq \mathbf{E}\left[\sum_{j=1}^{c(\mathrm{F}_n^k)} \mathbf{P}\left\{\left\|\breve{R}_{|\mathrm{T}_{n,j}^k|,\delta}\right\|_\infty \geq n^{1/4}(\gamma - |\breve{Z}_j^{n,\delta}|/n^{1/4}) \mid \mathrm{F}_n^k, (\breve{Z}_j^{n,\delta})_{j\geq 1}\right\}\right] \\
&\leq \mathbf{E}\left[\sum_{j=1}^{\infty} \mathbf{P}\left\{\|\breve{R}_{|\mathrm{T}_{n,j}^k|}\|_\infty \geq \frac{\gamma n^{1/4}}{2} \,\middle|\, \mathrm{F}_n^k\right\}\right],
\end{aligned}$$

for all n sufficiently large, since $|\breve{Z}_j^{n,\delta}| \leq n^{1/4-\delta}$ for all $1 \leq j \leq c(\mathrm{T}_k^n)$ and all $n \geq 1$. Applying Proposition 5.10 to each of the conditional probabilities in the above sum, we obtain that the right-hand side is at most

$$\frac{2^8 A}{\gamma^8} \mathbf{E}\left[\sum_{j=1}^{\infty} \left(\frac{|\mathrm{T}_{n,j}^k|}{n}\right)^2\right] = \frac{2^8 A}{\gamma^8} \mathbf{E}\left[\left(\frac{n - |V(\mathrm{T}_n^k)|}{n}\right)^2 \cdot \frac{\widehat{|\mathrm{T}^k|}}{n - |V(\mathrm{T}_n^k)|}\right], \tag{5.11}$$

where $\widehat{|\mathrm{T}^k|}$ is a size-biased pick from $(|\mathrm{T}_{n,j}^k|)_{j\geq 1}$. Clearly,

$$\frac{n - |V(\mathrm{T}_n^k)|}{n} \leq 1.$$

By Proposition 5.9, as $n \to \infty$, $\widehat{|\mathrm{T}^k|}/(n - |V(\mathrm{T}_n^k)|) \xrightarrow{\text{d}} \sigma^{-1}\widehat{|\gamma^k|}$ where $\widehat{|\gamma^k|}$ is a size-biased pick from $(|\gamma_j^k|, j \geq 1)$. By [36, Section 8.1], conditionally on J_k,

$$\widehat{|\gamma^k|} \stackrel{\text{d}}{=} \frac{B^2}{J_k^2 + B^2},$$

where B is a $N(0,1)$ random variable independent of J_k. Combining this with (5.11), we obtain that

$$\limsup_{n\to\infty} \mathbf{P}\left\{\max_{1\leq j\leq c(\mathrm{F}_n^k)} \left(\left\|\breve{R}_{n,\delta}(\mathrm{T}_{n,j}^k)\right\|_\infty + |\breve{Z}_j^{n,\delta}|\right) \geq \gamma n^{1/4}\right\} \leq \frac{2^8 A}{\sigma\gamma^8} \mathbf{E}\left[\frac{B^2}{J_k^2 + B^2}\right].$$

As $k \to \infty$, $J_k \xrightarrow{\mathrm{p}} \infty$. Therefore, by bounded convergence,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{1\le j\le c(\mathrm{F}_n^k)} \left(\left\| \breve{R}_{n,\delta}(\mathrm{T}_{n,j}^k) \right\|_\infty + |\breve{Z}_j^{n,\delta}| \right) \ge \gamma n^{1/4} \right\} = 0. \qquad \square$$

Assuming Proposition 5.10, Proposition 5.1 now follows from (5.4) by taking $\delta \in (0, 1/4)$ sufficiently small so that Proposition 5.3 holds, and combining that with Propositions 5.2 and 5.7.

6. The maximum spatial location: proof of Proposition 5.10. We assume throughout this section that μ is critical and has finite variance $\sigma^2 \in (0, \infty)$, and that [**A1**] and [**A2**] hold.

For $n \ge 1$, let $\Lambda^{(n)} := (\Lambda_1^{(n)}, \Lambda_2^{(n)}, \dots, \Lambda_{\widehat{D}_1^n}^{(n)})$ be the sizes of the subtrees of the root of T_n, so that $\Lambda_i^{(n)}$ is the size of the subtree rooted at the i-th child of the root. We will make extensive use of the fact that, conditionally on $\widehat{D}_1^n$, these are exchangeable random variables (i.e. their distribution is invariant under permutations of the labels). To prove Proposition 5.10 we will make extensive use of the following consequence of Lemma 25 of Haas and Miermont [14] which, roughly speaking, tells us that typically only one subtree of a child of the root is macroscopic and, moreover, the probability of a non-trivial macroscopic split at the root is on the order of $n^{-1/2}$.

LEMMA 6.1 (Lemma 25 of [14]). *It holds that*

$$\mathbf{E}\left[1 - \sum_{i=1}^{\widehat{D}_1^n} \left(\frac{\Lambda_i^{(n)}}{n} \right)^2 \right] = \Theta(n^{-1/2}). \tag{6.1}$$

In the proof of Proposition 5.10, we encounter terms directly related to the global centering and global finite variance conditions, respectively. The latter is more challenging to control, and is the reason for the third moment condition on the offspring distribution. These terms, and the control we will require on them, are stated in the following technical lemma. Recall the definition of $\widehat{D}^m$, the size-biased ordering of $D^m = (D_1^m, \dots, D_m^m)$, IID samples random distribution μ conditioned to sum to $m - 1$.

The proof of Proposition 5.10 is inductive, and requires that we control the maximum of $\breve{R}_{n,\delta}^k$ when restricted to subtrees of T_n^k. We henceforth use $m \ge 1$ to denote the number of vertices in the underlying tree, T_m, and $n \ge 1$ to denote the truncation threshold $n^{1/4-\delta}$ on the displacements. More specifically, in this section, we will consider branching random walks on T_m with displacements $\breve{Y}_k^{n,\delta}$, $k \ge 1$.

LEMMA 6.2. *Let $n \ge 1$ and $m \le n$. There exists $B > 0$ such that*

$$\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m} \right)^2 (\breve{Y}_{\widehat{D}_1^m, i}^{n,\delta})^2 \right] \le B. \tag{6.2}$$

*If in addition (μ, ν) satisfies [**A1**] and [**A2**], then there exists $B' > 0$ such that*

$$\left| \mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m} \right)^2 \breve{Y}_{\widehat{D}_1^m, i}^{n,\delta} \right] \right| \le \frac{B' n^{1/4-\delta}}{\sqrt{m}} + \frac{B'}{m^{1/4}}. \tag{6.3}$$

Condition [**A1**] pertains to the mean and variance of the displacement of a uniform child of a vertex with a size-biased number of offspring, $Y_{\bar{\xi},U_{\bar{\xi}}}$. The displacement from the root of T_m to a uniform child is distributed as $Y_{\widehat{D}_1^m,U_{\widehat{D}_1^m}}$ and we have $\widehat{D}_1^m \xrightarrow{\mathrm{d}} \bar{\xi}$ as $m \to \infty$. However, in order to use the global centering and global finite variance conditions in the proof of Lemma 6.2, we need something stronger, namely an explicit rate of decay for the total variation distance between the laws of $\bar{\xi}$ and $\widehat{D}_1^m$. This is provided by the next lemma.

LEMMA 6.3. *As $m \to \infty$,*

$$d_{\mathrm{TV}}(\widehat{D}_1^m, \bar{\xi}) = \frac{1}{2}\sum_{k=1}^{\infty}\left|\mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\}\right| = o(m^{-1/2}).$$

PROOF. Let $k \geq 1$, and let $(S_m)_{m\geq 1}$ be a random walk with IID μ-distributed increments. Recall from (1.13) that

$$\mathbf{P}\left\{\widehat{D}_1^m = k\right\} = \left(\frac{m}{m-1}\right)\frac{\mathbf{P}\left\{S_{m-1} = m-1-k\right\}}{\mathbf{P}\left\{S_m = m-1\right\}}\mathbf{P}\left\{\bar{\xi} = k\right\}.$$

Since $\mathbf{E}\left[\xi\right] = 1$ and $\mathbf{E}\left[\xi^3\right] < \infty$, by Theorem A.2,

$$\begin{aligned}&\sqrt{2\pi(m-1)}\sigma\mathbf{P}\left\{S_{m-1} = m-1-k\right\}\\ &\quad = e^{-k^2/(2\sigma^2(m-1))}\left(1 + \frac{1}{\sqrt{m-1}}\frac{\gamma_3}{6\sigma^3}\left(\frac{k^3}{\sigma^3(m-1)^{3/2}} - \frac{3k}{\sigma\sqrt{m-1}}\right)\right) + o(m^{-1/2}).\end{aligned}$$

If $k = O(m^{1/4})$,

$$\left|\frac{k^3}{\sigma^3(m-1)^{3/2}} - \frac{3k}{\sigma\sqrt{m-1}}\right| = O(m^{-1/4}),$$

and

$$e^{-k^2/(2\sigma^2(m-1))} = 1 - \frac{k^2}{2\sigma^2(m-1)} + O(m^{-1}).$$

Hence for $k = O(m^{1/4})$,

$$\sqrt{2\pi(m-1)}\sigma\mathbf{P}\left\{S_{m-1} = m-1-k\right\} = 1 - \frac{k^2}{2\sigma^2(m-1)} + o(m^{-1/2}).$$

It follows that for $k = O(m^{1/4})$,

$$\left(\frac{m}{m-1}\right)\frac{\mathbf{P}\left\{S_{m-1}=m-1-k\right\}}{\mathbf{P}\left\{S_m=m-1\right\}} = \frac{1-\frac{k^2}{2\sigma^2(m-1)}+o(m^{-1/2})}{1+o(m^{-1/2})} = 1-\frac{k^2}{2\sigma^2(m-1)}+o(m^{-1/2}),$$

and, consequently,

$$\mathbf{P}\left\{\widehat{D}_1^m = k\right\} = \left(1 - \frac{k^2}{2\sigma^2(m-1)} + o(m^{-1/2})\right)\mathbf{P}\left\{\bar{\xi} = k\right\}.$$

Therefore,

$$\sum_{k=1}^{\infty}\left|\mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\}\right|$$

$$= \sum_{k=1}^{\lfloor m^{1/4} \rfloor} \left| \mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\} \right| + \sum_{k=\lfloor m^{1/4} \rfloor + 1}^{\infty} \left| \mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\} \right|$$

$$= \sum_{k=1}^{\lfloor m^{1/4} \rfloor} \left(\frac{k^2}{2\sigma^2(m-1)} + o(m^{-1/2}) \right) \mathbf{P}\left\{\bar{\xi} = k\right\} + \sum_{k=\lfloor m^{1/4} \rfloor + 1}^{\infty} \left| \mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\} \right|$$

$$\leq \frac{\mathbf{E}\left[\xi^3\right]}{2\sigma^2(m-1)} + o(m^{-1/2}) + \sum_{k=\lfloor m^{1/4} \rfloor + 1}^{\infty} \left(\mathbf{P}\left\{\widehat{D}_1^m = k\right\} + \mathbf{P}\left\{\bar{\xi} = k\right\} \right)$$

$$\leq o(m^{-1/2}) + (c+1) \cdot \mathbf{P}\left\{\bar{\xi} > m^{1/4}\right\},$$

where the final inequality follows since $\mathbf{P}\left\{\widehat{D}_1^m = k\right\} \leq c\mathbf{P}\left\{\bar{\xi} = k\right\}$ for all $k \in [m]$. Since $\mathbf{E}\left[\xi^3\right] < \infty$, $\bar{\xi}$ has a finite second moment. Therefore, $\mathbf{P}\left\{\bar{\xi} > k\right\} = o(k^{-2})$ as $k \to \infty$ and so $\mathbf{P}\left\{\bar{\xi} > m^{1/4}\right\} = o(m^{-1/2})$. The result follows. □

The terms (6.2) and (6.3) relate to the variance and mean (respectively) of the displacement of a uniform child of the root in branching random walk $(\mathrm{T}_m, \breve{Y}^{n,\delta})$. Since this branching random walk is globally centered, it is reasonable to expect that the mean will be small and that the second moment will be bounded. A key technical lemma follows.

LEMMA 6.4. *There exists a constant $C > 0$ such that for $m \leq n$,*

$$\left| \mathbf{E}\left[\frac{1}{\widehat{D}_1^m} \sum_{i=1}^{\widehat{D}_1^m} \breve{Y}^{n,\delta}_{\widehat{D}_1^m, i} \right] \right| \leq \frac{C n^{1/4-\delta}}{\sqrt{m}}.$$

PROOF. Let $(\widehat{D}_1^m, \bar{\xi})$ be a coupling of the degree of the root of T_m and the size-biased distribution of μ. We consider the events $\{\bar{\xi} = \widehat{D}_1^m\}$ and $\{\bar{\xi} \neq \widehat{D}_1^m\}$ separately:

$$\left| \mathbf{E}\left[\frac{1}{\widehat{D}_1^m} \sum_{i=1}^{\widehat{D}_1^m} \breve{Y}^{n,\delta}_{\widehat{D}_1^m, i} \right] \right| \leq \left| \mathbf{E}\left[\frac{1}{\bar{\xi}} \sum_{i=1}^{\bar{\xi}} \breve{Y}^{n,\delta}_{\bar{\xi}, i} \mathbf{1}_{[\bar{\xi} = \widehat{D}_1^m]} \right] \right| + \mathbf{E}\left[\frac{1}{\widehat{D}_1^m} \sum_{i=1}^{\widehat{D}_1^m} |\breve{Y}^{n,\delta}_{\widehat{D}_1^m, i}| \mathbf{1}_{[\bar{\xi} \neq \widehat{D}_1^m]} \right]$$

$$= \left| \mathbf{E}\left[\frac{1}{\bar{\xi}} \sum_{i=1}^{\bar{\xi}} \breve{Y}^{n,\delta}_{\bar{\xi}, i} \mathbf{1}_{[\bar{\xi} \neq \widehat{D}_1^m]} \right] \right| + \mathbf{E}\left[\frac{1}{\widehat{D}_1^m} \sum_{i=1}^{\widehat{D}_1^m} |\breve{Y}^{n,\delta}_{\widehat{D}_1^m, i}| \mathbf{1}_{[\bar{\xi} \neq \widehat{D}_1^m]} \right], \tag{6.4}$$

where the equality holds since

$$\mathbf{E}\left[\frac{1}{\bar{\xi}} \sum_{i=1}^{\bar{\xi}} \breve{Y}^{n,\delta}_{\bar{\xi}, i} \right] = \mathbf{E}\left[\breve{Y}^{n,\delta}_{\bar{\xi}, U_{\bar{\xi}}} \right] = 0.$$

Since $|\breve{Y}^{n,\delta}_{k,j}| \leq 2n^{1/4-\delta}$ for all $k \geq 1$ and $j \in [k]$, it follows that (6.4) is at most

$$4n^{1/4-\delta} \mathbf{P}\left\{\bar{\xi} \neq \widehat{D}_1^m\right\}.$$

The result follows from Lemma 6.3 by taking an optimal coupling of $(\widehat{D}_1^m, \bar{\xi})$. □

We now proceed to prove Lemma 6.2.

PROOF OF LEMMA 6.2. We first prove (6.2). Note that by exchangeability of $(\Delta_1^{(m)}, \dots, \Delta_{\widehat{D}_m^1}^{(m)})$ and linearity of conditional expectation

$$\begin{aligned}
\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2 (\breve{Y}^{n,\delta}_{\widehat{D}_m^1,i})^2\right] &= \mathbf{E}\left[\left(\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}(\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i})^2\right)\left(\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)\right] \\
&\le \mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}(\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i})^2\right] \\
&\le c\mathbf{E}\left[\frac{1}{\bar{\xi}}\sum_{i=1}^{\bar{\xi}}(\breve{Y}^{n,\delta}_{\bar{\xi},i})^2\right] \\
&= c\mathbf{E}\left[(\breve{Y}^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}})^2\right], \qquad (6.5)
\end{aligned}$$

where the second inequality follows since $\mathbf{P}\left\{\widehat{D}_1^m = k\right\} \le c\mathbf{P}\left\{\bar{\xi} = k\right\}$. By Lemma 5.8, (6.5) tends to β^2 as $n \to \infty$ and hence (6.2) holds.

We now proceed to proving (6.3). By linearity and the triangle inequality we have

$$\begin{aligned}
&\left|\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2 \breve{Y}^{n,\delta}_{\widehat{D}_m^1,i}\right]\right| \\
&\le \left|\mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right]\right| + \left|\mathbf{E}\left[\left(\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right)\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)\right]\right|. \qquad (6.6)
\end{aligned}$$

By Lemma 6.4, (6.6) is at most

$$\frac{Cn^{1/4-\delta}}{\sqrt{m}} + \left|\mathbf{E}\left[\left(\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right)\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)\right]\right|.$$

Applying the Cauchy–Schwarz inequality to the second term yields an upper bound of

$$\begin{aligned}
&\frac{Cn^{1/4-\delta}}{\sqrt{m}} + \mathbf{E}\left[\left(\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}|\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}|\right)^2\right]^{1/2}\mathbf{E}\left[\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)^2\right]^{1/2} \\
&\le \frac{Cn^{1/4-\delta}}{\sqrt{m}} + \mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}(\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i})^2\right]^{1/2}\mathbf{E}\left[\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)^2\right]^{1/2}, \qquad (6.7)
\end{aligned}$$

where we have again used the Cauchy–Schwarz inequality on the sum inside the expectation to get the second inequality. Since $\mathbf{P}\left\{\widehat{D}_1^m = k\right\} \le c\mathbf{P}\left\{\bar{\xi} = k\right\}$, by the same methods used in (6.2), there exists $c' > 0$ such that (6.7) is at most

$$\frac{Cn^{1/4-\delta}}{\sqrt{m}} + c'\mathbf{E}\left[\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)^2\right]^{1/2}.$$

Lastly, since $x^2 \leq x$ for all $x \in [0,1]$, we obtain the bound

$$\left| \mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m} \right)^2 \breve{Y}^{n,\delta}_{\widehat{D}_1^m,i} \right] \right| \leq \frac{Cn^{1/4-\delta}}{\sqrt{m}} + c'\mathbf{E}\left[1 - \sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m} \right)^2 \right]^{1/2},$$

and the result follows by Lemma 6.1. □

We now present the proof of Proposition 5.10.

PROOF OF PROPOSITION 5.10. For $n \geq 1$ and $m \leq n$, let $\breve{R}_{m,n,\delta}$ be the spatial process of a branching random walk $\breve{\mathbf{T}}_{m,n} = (\mathrm{T}_m, \breve{Y}^{n,\delta})$ where the displacement vector of a vertex $v \in v(\mathrm{T}_m) \setminus \partial \mathrm{T}_m$ with k children is distributed as

$$\breve{Y}^{n,\delta}_k = Y^{n,\delta}_k - \mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}} \right].$$

Furthermore, let

$$\breve{R}^+_{m,n,\delta} := \max\left\{ 0, \max_{0\leq i\leq m} \breve{R}_{m,n,\delta} \right\},$$

and

$$\breve{R}^-_{m,n,\delta} := -\min\left\{ 0, \min_{0\leq i\leq m} \breve{R}_{m,n,\delta} \right\}.$$

It suffices to prove that there exists $A > 0$ such that for all $m \geq 0$, all $n \geq m$ and all $\gamma > 0$,

$$\mathbf{P}\left\{ \breve{R}^+_{m,n,\delta} > \gamma n^{1/4} \right\} \leq \frac{A}{\gamma^8} \quad \text{and} \quad \mathbf{P}\left\{ \breve{R}^-_{m,n,\delta} > \gamma n^{1/4} \right\} \leq \frac{A}{\gamma^8},$$

since Proposition 5.10 then follows by taking $n = m$. We only prove the tail bound for $\breve{R}^+_{m,n,\delta}$, as the bound for $\breve{R}^-_{m,n,\delta}$ then follows by symmetry.

Notice that $\breve{R}^+_{1,n,\delta} = 0$ for all $n \geq 0$, and so the claim holds trivially if $m = 1$. Moreover, at the cost of taking $A > 0$ larger, it is sufficient to prove the result for $\gamma > 0$ sufficiently large. We will proceed by induction on $m \geq 2$, and hence assume that for $1 \leq k \leq m-1$ and $\gamma > 0$,

$$\mathbf{P}\left\{ \breve{R}^+_{k,n,\delta} > \gamma n^{1/4} \right\} \leq \frac{A}{\gamma^8},$$

for all $n \geq k$.

Observe that conditionally on $\widehat{D}_1^m$ and $\Lambda^{(m)}$,

$$\breve{R}^+_{m,n,\delta} \stackrel{\mathrm{d}}{=} \max\left\{ 0, \max_{1\leq i\leq \widehat{D}_1^m} \left\{ \breve{R}^+_{\Lambda_i^{(m)},n,\delta} + \breve{Y}^{n,\delta}_{\widehat{D}_1^m,i} \right\} \right\}.$$

For the rest of the proof, we write $\breve{Y}^{n,\delta}_i$ in place of $\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}$ to ease the notation.

Take $u_0 \in (0,1)$ such that for all $0 < u < u_0$, $(1-u)^{-8} \leq 1 + 8u + 72u^2$. Then, taking $\gamma > 2/u_0$ (recall this is possible at the cost of taking $A > 0$ larger), it follows that

$$\begin{aligned}
&\mathbf{P}\left\{ \breve{R}^+_{m,n,\delta} \leq \gamma n^{1/4} \right\} \\
&= \mathbf{E}\left[\mathbf{P}\left\{ \max_{1\leq i\leq \widehat{D}_1^m} \left\{ \breve{R}^+_{\Lambda^{(m)},n,\delta} + \breve{Y}^{n,\delta}_i \right\} \leq \gamma n^{1/4} \middle| \widehat{D}_1^m, \Lambda^{(m)} \right\} \right] \\
&= \mathbf{E}\left[\prod_{i=1}^{\widehat{D}_1^m} \mathbf{P}\left\{ \breve{R}^+_{\Lambda_i^{(m)},n,\delta} \leq \gamma n^{1/4} - \breve{Y}^{n,\delta}_i \middle| \widehat{D}_1^m, \Lambda^{(m)}, \breve{Y}^{n,\delta}_i \right\} \right],
\end{aligned}$$

where in the second equality we have used the tower law and the branching property. We will bound the right-hand side of the above equality by applying induction to each term in the product. More specifically, taking $n = k = \Lambda_i^{(m)}$ and for the i-th term of the product, by the induction hypothesis, we obtain

$$\begin{aligned}
&\mathbf{E}\left[\prod_{i=1}^{\widehat{D}_1^m} \mathbf{P}\left\{\breve{R}_{\Lambda_i^{(m)},n,\delta} \le \gamma n^{1/4} - \breve{Y}_i^{n,\delta} \middle| \widehat{D}_1^m, \Lambda^{(m)}, \breve{Y}_i^{n,\delta}\right\}\right] \\
&\quad = \mathbf{E}\left[\prod_{i=1}^{\widehat{D}_1^m} \mathbf{P}\left\{\breve{R}_{\Lambda_i^{(m)},n,\delta} \le \left(\frac{\gamma n^{1/4} - \breve{Y}_i^{n,\delta}}{(\Lambda_i^{(m)})^{1/4}}\right)(\Lambda_i^{(m)})^{1/4} \middle| \widehat{D}_1^m, \Lambda^{(m)}, \breve{Y}_i^{n,\delta}\right\}\right] \\
&\quad \ge \mathbf{E}\left[\prod_{i=1}^{\widehat{D}_1^m} \left(1 - \frac{A(\Lambda_i^{(m)})^2}{(\gamma n^{1/4} - \breve{Y}_i^{n,\delta})^8}\right)_+\right].
\end{aligned}$$

Furthermore, since $\prod_{i=1}^k (1-x_i)_+ \ge 1 - \sum_{i=1}^k x_i$ for any non-negative sequence $(x_i)_{i\ge 1}$, we may lower bound the above as

$$\begin{aligned}
\mathbf{E}\left[\prod_{i=1}^{\widehat{D}_1^m} \left(1 - \frac{A(\Lambda_i^{(m)})^2}{(\gamma n^{1/4} - \breve{Y}_i^{n,\delta})^8}\right)_+\right] &\ge 1 - \frac{A}{\gamma^8}\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{n}\right)^2 \left(1 - \frac{\breve{Y}_i^{n,\delta}}{\gamma n^{1/4}}\right)^{-8}\right] \\
&\ge 1 - \frac{A}{\gamma^8}\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m}\right)^2 \left(1 - \frac{\breve{Y}_i^{n,\delta}}{\gamma n^{1/4}}\right)^{-8}\right],
\end{aligned}$$

where the final inequality holds since $m \le n$. Moreover, since $\gamma > 2/u_0$, we have that $|\breve{Y}_i^{n,\delta}|/(\gamma n^{1/4}) < u_0$ for any n, and so $(1 - \breve{Y}_i^{n,\delta}/(\gamma n^{1/4}))^{-8} \le 1 + 8\breve{Y}_i^{n,\delta}/(\gamma n^{1/4}) + 72(\breve{Y}_i^{n,\delta})^2/(\gamma^2\sqrt{n})$. Hence,

$$\mathbf{P}\left\{\breve{R}^+_{m,n,\delta} \le \gamma n^{1/4}\right\} \ge 1 - \frac{A}{\gamma^8} + \frac{A}{\gamma^8}\mathbf{E}\left[1 - \sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right] \tag{6.8}$$

$$- \frac{8A}{\gamma^9 n^{1/4}}\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m}\right)^2 \breve{Y}_i^{n,\delta}\right] \tag{6.9}$$

$$- \frac{72A}{\gamma^{10}\sqrt{m}}\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m}\right)^2 (\breve{Y}_i^{n,\delta})^2\right], \tag{6.10}$$

where we may take the denominator of the final term of the above expression to be $\gamma^{10}\sqrt{m}$ rather than $\gamma^{10}\sqrt{n}$ as $m \le n$ and the expectation in this term is non-negative. Applying (6.1), (6.2), and (6.3) to bound the expectations in (6.8), (6.10), and (6.9), respectively, we obtain that there exist constants $B, B', B'' > 0$ such that

$$\begin{aligned}
\mathbf{P}\left\{\breve{R}^+_{m,n,\delta} \le \gamma n^{1/4}\right\} &\ge 1 - \frac{A}{\gamma^8} + \frac{AB''}{\gamma^8\sqrt{m}} - \frac{8A}{\gamma^9 n^{1/4}}\left(\frac{B' n^{1/4-\delta}}{\sqrt{m}} + \frac{B'}{m^{1/4}}\right) - \frac{72AB}{\gamma^{10}\sqrt{m}} \\
&\ge 1 - \frac{A}{\gamma^8} + \frac{A}{\gamma^8\sqrt{m}}\left(B'' - \frac{8B'}{\gamma n^\delta} - \frac{8B'}{\gamma} - \frac{72B}{\gamma^2}\right).
\end{aligned}$$

For $\gamma > 0$ large enough, the final term in parentheses is positive so the whole expression is at least $1 - A/\gamma^8$. The result follows by induction on m. □

7. The hairy tour. In this section we prove Theorems 1.4 and 1.5. In particular, we show that under assumptions [**A1**] and [**A3**] for a given measure π with $\eta \in [0, 2)$, we have that $(n^{-1/2}H_n, n^{-1/(4-\eta)}R_n)$ converges in distribution to a generalisation of the hairy tour introduced by Janson and Marckert [18] if $\eta = 0$, and to a process whose second coordinate is a pure jump process if $\eta \in (0, 2)$. Recall that by [**A3**], π is a Borel measure on $\mathbb{R}^2 \setminus \{(0,0)\}$ such that for for any $\varepsilon > 0$, both $\pi(\mathbb{R}_+ \times (\varepsilon, \infty)) < \infty$ and $\pi((\varepsilon, \infty) \times \mathbb{R}_+) < \infty$, and that for all Borel sets $A \subset \mathbb{R}^2_+ \setminus \{(0,0)\}$ for which $\pi(\partial A) = 0$,

$$r^{4-\eta}\mathbf{P}\left\{\frac{1}{r}\left(\max_{1\le i\le \xi} Y^+_{\xi,i}, \max_{1\le i\le \xi} Y^-_{\xi,i}\right) \in A)\right\} \to \pi(A)$$

as $r \to \infty$, where $Y^+_{k,j} = Y_{k,j} \vee 0$ and $Y^-_{k,j} = (-Y_{k,j}) \vee 0$. The measure π will be the intensity measure for a Poisson point process which drives the second coordinate of the limit.

Recall that $\mathbf{T}_n = (\mathrm{T}_n, Y)$ is such that given T_n, $Y = (Y^{(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ is a collection of independent random vectors, where if $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ has k children then $Y^{(v)}$ has distribution ν_k. Observe that, for fixed $\eta \in [0, 2)$, by assumption [**A3**], if the measure π has non-zero mass then

$$\max_{v \in v(\mathrm{T}_n)} \|Y^{(v)}\|_\infty = \Theta_{\mathbf{P}}(n^{1/(4-\eta)}).$$

Fix $\gamma > 0$, $\delta \in (0, 1/(4-\eta))$, and suppose that $n \ge 1$ is sufficiently large so that $n^{1/(4-\eta)-\delta} \le \gamma n^{1/(4-\eta)}$. As in the proof of tightness for Theorem 1.1, and more specifically as in Section 5, in order to prove Theorems 1.4 and 1.5, we will need to consider three "restrictions" of the branching random walk $\mathbf{T}_n$. These restrictions are a generalisation of those used in Section 5 from the case $\eta = 0$ to that of general $\eta \in [0, 2)$; the modified definitions are given below.

We denote the restrictions of $\mathbf{T}_n$ by $\mathbf{T}_{n,\delta} = (\mathrm{T}_n, Y_{n,\delta})$, $\mathbf{T}^\gamma_{n,\delta} = (\mathrm{T}_n, Y^\gamma_{n,\delta})$, and $\mathbf{T}^\gamma_n = (\mathrm{T}_n, Y^\gamma_n)$. Again, these branching random walks will respectively capture the "typical", "mid-range", and "large" displacements in $\mathbf{T}_n$, as follows:

1. **(typical displacements):** For all $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$,
$$Y^{(v)}_{n,\delta} = Y^{(v)}\mathbf{1}_{[\|Y^{(v)}\|_\infty \le n^{1/(4-\eta)-\delta}]};$$
2. **(mid-range displacements):** For all $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$,
$$Y^{\gamma,(v)}_{n,\delta} = Y^{(v)}\mathbf{1}_{[n^{1/(4-\eta)-\delta} < \|Y^{(v)}\|_\infty \le \gamma n^{1/(4-\eta)}]};$$
3. **(large displacements):** For all $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$,
$$Y^{\gamma,(v)}_{n} = Y^{(v)}\mathbf{1}_{[\|Y^{(v)}\|_\infty > \gamma n^{1/(4-\eta)}]}.$$

We note that, informally, taking $\gamma \downarrow 0$ in $\mathbf{T}^\gamma_n$ captures all displacements of the largest order. We define $R_{n,\delta}$, $R^\gamma_{n,\delta}$, and R^γ_n to be the functions encoding the spatial locations of the vertices of $\mathbf{T}_{n,\delta}$, $\mathbf{T}^\gamma_{n,\delta}$, and $\mathbf{T}^\gamma_n$, respectively.

Before studying the convergence of the functions $R_{n,\delta}$, $R^\gamma_{n,\delta}$, and R^γ_n, we will prove convergence upon rescaling of the *values* of the large displacements. For $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$, let

$$Y^{(v,+)} := 0 \vee \max_{j \in [c(v,\mathrm{T}_n)]} Y^{(v)}_j \quad \text{and} \quad Y^{(v,-)} := 0 \vee \max_{j \in [c(v,\mathrm{T}_n)]} (-Y^{(v)}_j)$$

be the largest positive and negative terms (respectively) in the displacement vector $Y^{(v)}$ from v to its children and, for $v \in \partial \mathrm{T}_n$, set $Y^{(v,+)} = Y^{(v,-)} = 0$. For a finite multiset $S \subset \mathbb{R}^2$, by

"the decreasing ordering of S" we mean the vector $(s_1, \dots, s_m)$ which lists the elements of S in decreasing order of their largest coordinate, breaking ties in decreasing order of their smallest coordinate. Let $L_n^{\eta,\gamma}$ be the decreasing ordering of the multiset

$$\left\{(Y^{(v,+)}, Y^{(v,-)})\mathbf{1}_{[\|Y^{(v)}\|_\infty > \gamma n^{1/(4-\eta)}]},\ v \in v(\mathrm{T}_n)\right\}, \tag{7.1}$$

concatenated with an infinite sequence with all entries $(0,0)$.

LEMMA 7.1. *Fix $\gamma > 0$ and suppose that [A1] holds and [A3] holds for a given measure π with $\eta \in [0,2)$. Then as $n \to \infty$,*

$$\frac{L_n^{\eta,\gamma}}{n^{1/(4-\eta)}} \xrightarrow{\mathrm{d}} L^{\eta,\gamma}$$

in ℓ_∞, where $L^{\eta,\gamma}$ is the decreasing ordering of the points of a Poisson process on $\mathbb{R}^2_{\geq 0}$ with intensity $\pi(dx,dy)\mathbf{1}_{[(x\vee y)>\gamma]}$ concatenated with an infinite sequence with all entries $(0,0)$.

PROOF. Let $(\xi_i, i \geq 1)$ be IID samples from the offspring distribution μ. Further, for $i \geq 1$, sample Y_{ξ_i} independently and let

$$Y^+_{\xi_i} := 0 \vee \max_{j \in [\xi_i]} Y_{\xi_i, j} \quad \text{and} \quad Y^-_{\xi_i} := 0 \vee \max_{j \in [\xi_i]} (-Y_{\xi_i, j}).$$

By definition, the multiset $\{(Y^{(v,+)}, Y^{(v,-)}), v \in v(\mathrm{T}_n)\}$ is distributed as $\{(Y^+_{\xi_i}, Y^-_{\xi_i}), i \in [n]\}$ conditioned on the event that $\sum_{i=1}^n \xi_i = n-1$.

For $n \geq 1$, let $\widetilde{L}_n^{\eta,\gamma}$ be the decreasing ordering of

$$\left\{(Y^+_{\xi_i}, Y^-_{\xi_i})\mathbf{1}_{[\|Y_{\xi_i}\|_\infty > \gamma n^{1/(4-\eta)}]},\ i \in [n]\right\},$$

concatenated with an infinite sequence with all entries $(0,0)$. We will first show that

$$n^{-1/(4-\eta)} \widetilde{L}_n^{\eta,\gamma} \xrightarrow{\mathrm{d}} L^{\eta,\gamma} \tag{7.2}$$

in ℓ_∞ as $n \to \infty$. To this end, note that by [A3], for any $x, y \geq 0$ such that $x \vee y > \gamma$ and such that $\pi((\{x\} \times [y,\infty)) \cup ([x,\infty) \times \{y\})) = 0$,

$$n\mathbf{P}\left\{Y^+_{\xi_i} > x n^{1/(4-\eta)},\ Y^-_{\xi_i} > y n^{1/(4-\eta)}\right\} \to \pi((x,\infty) \times (y,\infty)),$$

as $n \to \infty$ and, moreover, $\pi((x,\infty) \times (y,\infty)) < \infty$. Therefore,

$$\begin{aligned}
&\left|\left\{i \in [n]\ :\ Y^+_{\xi_i} > x n^{1/(4-\eta)},\ Y^-_{\xi_i} > y n^{1/(4-\eta)}\right\}\right| \\
&\qquad \stackrel{\mathrm{d}}{=} \mathrm{Binomial}\left(n\mathbf{P}\left\{Y^+_{\xi_i} > x n^{1/(4-\eta)},\ Y^-_{\xi_i} > y n^{1/(4-\eta)}\right\}\right) \\
&\qquad \xrightarrow{\mathrm{d}} \mathrm{Poisson}(\pi((x,\infty) \times (y,\infty))),
\end{aligned} \tag{7.3}$$

and (7.2) follows from the fact that a Poisson process on $\mathbb{R}^2$ is determined by its distribution on half-infinite rectangles and the continuity of the function $x, y \mapsto x \vee y, x \wedge y$ that we use to order the multisets.

We now show that the convergence in (7.2) still holds when we condition on $\sum_{i=1}^n \xi_i = n-1$. We note that the remainder of this proof is similar to the end of the proof of Proposition 5.2.

Let $\widetilde{M}_n^\gamma$ be the number of elements in $\widetilde{L}_n^{\eta,\gamma}$ which are not equal to $(0,0)$. Note that by (7.3), the sequence $(\widetilde{M}_n^\gamma)_{n\geq 1}$ is tight. Further, let $\widetilde{S}_n^\gamma := \sum_{i \in [n]} \xi_i \mathbf{1}_{[\|Y_{\xi_i}\|_\infty > \gamma n^{1/(4-\eta)}]}$. Since

$\xi_1, \dots \xi_n$ are IID, the law of $\sum_{i=1}^n \xi_i$ depends on $\widetilde{L}_n^{\eta,\gamma}$ solely through $\widetilde{M}_n^\gamma$ and $\widetilde{S}_n^\gamma$. To be precise, let $\xi_1^n, \xi_2^n, \dots$ be independent random variables such that for each $i \geq 1$, ξ_i^n is distributed as ξ_i conditional on $\|Y_{\xi_i}\|_\infty < \gamma n^{1/(4-\eta)}$. Then,

$$\mathbf{P}\left\{\sum_{i=1}^n \xi_i = k \;\middle|\; \widetilde{L}_n^{\eta,\gamma}\right\} = \mathbf{P}\left\{\widetilde{S}_n^\gamma + \sum_{i=1}^{n-\widetilde{M}_n^\gamma} \xi_i^n = k \;\middle|\; \widetilde{S}_n^\gamma,\ \widetilde{M}_n^\gamma\right\}. \tag{7.4}$$

Let $F : \ell_\infty \to \mathbb{R}$ be a bounded measurable function. Then, by analogous arguments to those used to prove (5.6),

$$\begin{aligned}
\mathbf{E}\left[F(L_n^{\eta,\gamma})\right] &= \mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\mathbf{1}_{[\widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}]} \;\middle|\; \sum_{i=1}^n \xi_i = n-1\right] + o(1) \\
&= \frac{\mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\mathbf{1}_{[\sum_{i=1}^n \xi_i = n-1,\ \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}]}\right]}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} + o(1) \\
&= \frac{\mathbf{E}\left[\mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\mathbf{1}_{[\sum_{i=1}^n \xi_i = n-1,\ \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}]} \;\middle|\; \widetilde{L}_n^{\eta,\gamma}\right]\right]}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} + o(1) \\
&= \mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\right] \frac{\mathbf{P}\left\{\sum_i^{n-\widetilde{M}_n^\gamma} \xi_i^n = n-1-\widetilde{S}_n^\gamma \;\middle|\; \widetilde{M}_n^\gamma < n^\varepsilon, \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} + o(1),
\end{aligned}$$

where the last equality holds by (7.4). By a quantitative local limit theorem (see Lemma A.3 in the appendix), we obtain that as $n \to \infty$

$$\frac{\mathbf{P}\left\{\sum_i^{n-m} \xi_i^n = n-1-s\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \to 1,$$

uniformly over all $m < n^\varepsilon$ and $s < n^{1/3+\varepsilon}$. It follows that

$$\mathbf{E}\left[F(L_n^{\eta,\gamma})\right] = \mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\right] + o(1).$$

The result then follows by (7.2). □

In the remainder of the section, we continue to use $L^{n,\gamma}$ to refer to a random vector with the distribution given in Lemma 7.1.

To prove Theorems 1.4 and 1.5, we use similar methods to those used to prove Theorem 1.1. First, we will prove convergence of the branching random walk restricted to the subtree spanned by k uniform vertices, by showing that the convergence from Proposition 4.1 holds jointly with that in Lemma 7.1, and that the limits are independent. This, in particular, implies the convergence of the random finite-dimensional distributions in Theorems 1.4 and 1.5. The independence is the key issue here, and in order to obtain it, we require adaptations of Proposition 4.3 and Lemma 4.4 to the setting of n-dependent offspring distributions. The required technical results may be found in the appendix.

Following this, using similar techniques to those used in Sections 5 and 6 to prove tightness for the discrete snake in Theorem 1.1, and applying the aforementioned joint convergence, we will show that a discrete snake comprised solely of the "typical" displacements converges to the head of the BSBE on rescaling by $n^{-1/4}$ if $\eta = 0$, and to 0 on rescaling by $n^{-1/(4-\eta)}$ if $\eta \in (0,2)$. Furthermore, this discrete snake is asymptotically independent of the large displacements. In Section 7.2 we show that for $\eta \in [0,2)$ the mid-range displacements make

only a vanishing contribution to the head of the discrete snake on the scale of $n^{1/(4-\eta)}$. Next, by a small variant of Lemma 5.6, we deduce that the large displacements appear near the leaves. We apply this result to prove Lemma 7.11, which states that the discrete snake associated with the branching random walk $\mathbf{T}'_n$ obtained by pruning sub-branching random walks rooted at vertices with large displacements in $\mathbf{T}_n$ converges upon rescaling by $n^{-1/(4-\eta)}$ to the same limit as that of the "typical displacement" discrete snake (with the limit depending on whether $\eta = 0$ or $\eta \in (0,2)$). Theorems 1.4 and 1.5 then follow by showing that the branching random walk obtained by regrafting these pruned sub-branching random walks to uniform leaves of $\mathbf{T}'_n$ has the same law as $\mathbf{T}_n$.

The following proposition establishes the convergence of the branching random walk restricted to the subtree spanned by k uniform vertices, as well as the its asymptotic independence from the large displacements.

PROPOSITION 7.2. *Fix $\gamma > 0$ and suppose that [**A1**] holds and [**A3**] holds for a given measure π with $\eta \in [0,2)$. Fix $k \geq 1$. Then*

$$\frac{\sigma}{\sqrt{n}}(J_1^n, J_2^n, \dots, J_k^n, A_1^n, \dots, A_k^n) \xrightarrow{\mathrm{d}} (J_1, J_2, \dots, J_k, A_1, \dots, A_k)$$

as $n \to \infty$. Jointly with this convergence, we have that

$$(F_1^n, F_2^n, \dots, F_k^n) \xrightarrow{\mathrm{d}} (F_1, F_2, \dots, F_k),$$

where $F_1, F_2, \dots, F_k$ are IID *random variables, independent of everything else, such that $\mathbf{P}\{F_i = 1\} = \mathbf{P}\{F_i = 2\} = 1/2$ and*

$$\left(\frac{L^n(\lfloor tn^{1/2}\rfloor \wedge (J_1^n - 1))}{n^{1/4}}\right)_{t\geq 0} \xrightarrow{\mathrm{d}} \beta(B_{t\wedge(J_1/\sigma)})_{t\geq 0},$$

$$\left(\frac{L^n((J_i^n + \lfloor tn^{1/2}\rfloor) \wedge (J_{i+1}^n - 1))}{n^{1/4}}\right)_{t\geq 0} \xrightarrow{\mathrm{d}} \beta(B_{A_i/\sigma} + B_{((J_i/\sigma)+t)\wedge(J_{i+1}/\sigma)} - B_{(J_i/\sigma)})_{t\geq 0}$$

for $1 \leq i \leq k-1$, in each case for the uniform norm. Moreover, jointly with this convergence,

$$\frac{L_n^{\eta,\gamma}}{n^{1/(4-\eta)}} \xrightarrow{\mathrm{d}} L^{\eta,\gamma},$$

in ℓ_∞, where $L^{\eta,\gamma}$ is independent of all the other limiting random variables.

PROOF. Fix $k \geq 1$ and $\gamma > 0$ and write

$$V_n = (J_1^n, J_2^n, \dots, J_k^n, A_1^n, \dots, A_k^n, F_1^n, F_2^n, \dots, F_k^n,$$
$$(L^n(\lfloor tn^{1/2}\rfloor \wedge (J_1^n - 1)))_{t\geq 0}), (L^n((J_1^n + \lfloor tn^{1/2}\rfloor) \wedge (J_2^n - 1)))_{t\geq 0}), \dots,$$
$$(L^n((J_{k-1}^n + \lfloor tn^{1/2}\rfloor) \wedge (J_k^n - 1)))_{t\geq 0})$$

for the vector containing all variables that, in Proposition 4.1, have already been shown to converge jointly under rescaling when we equip the first $3k$ entries with the Euclidean topology on $\mathbb{R}$, the last k entries with the topology of uniform convergence, and the whole vector with the product topology. Then, let g be an $\mathbb{R}$-valued bounded continuous function (for this topology), and $h : \ell_\infty \to \mathbb{R}$ be another bounded continuous function. By Proposition 4.1 and Lemma 7.1, it suffices to prove that

$$\left|\mathbf{E}[g(V_n)h(L_n^{\eta,\gamma})] - \mathbf{E}[g(V_n)]\,\mathbf{E}[h(L_n^{\eta,\gamma})]\right| \to 0 \tag{7.5}$$

as $n \to \infty$.

Let $(M_n^{\eta,\gamma}, S_n^{\eta,\gamma})$ have the joint distribution of the number of vertices with a large displacement, $\sum_{v \in \mathrm{T}_n} \mathbf{1}_{[\|Y^{(v)}\| > \gamma n^{1/(4-\eta)}]}$, and the total number of children of such vertices, $\sum_{v \in \mathrm{T}_n} c(v, \mathrm{T}_n) \mathbf{1}_{[\|Y^{(v)}\| > \gamma n^{1/(4-\eta)}]}$. Fix $\varepsilon \in (0, 1/6)$ and define the good event

$$\mathcal{G}_1 = \{M_n^{\eta,\gamma} \le n^{\varepsilon}, S_n^{\eta,\gamma} \le n^{1/3+\varepsilon}\}.$$

By analogous arguments to those used to prove (5.6), $\mathcal{G}_1$ occurs with high probability. Now recall that $\sigma n^{-1/2} J_k^n \xrightarrow{\mathrm{d}} J_k$ as $n \to \infty$. Fix $T > 0$ and let $\mathcal{G}_2$ be the (good) event that $J_k^n \le T\sqrt{n}$. (We observe that by choosing T large we may make $\mathbf{P}\{\mathcal{G}_2\}$ as close to 1 as we like, uniformly in n sufficiently large.) Then,

$$\mathbf{E}\left[g(V_n)h(L_n^{\eta,\gamma})\right] = \mathbf{E}\left[g(V_n)h(L_n^{\eta,\gamma})\mathbf{1}_{[\mathcal{G}_1 \cap \mathcal{G}_2]}\right] + o(1),$$

where $o(1)$ is to be understood as an error that tends to 0 as $n \to \infty$ and then $T \to \infty$.

Let $\mathcal{F}_n^{\eta,\gamma}$ denote the σ-algebra generated by the degrees and displacement vectors of the vertices v with $\|Y^{(v)}\| > \gamma n^{1/(4-\eta)}$. We see that $\mathcal{G}_1$ and $L_n^{\eta,\gamma}$ are measurable with respect to $\mathcal{F}_n^{\eta,\gamma}$, and so

$$\mathbf{E}\left[g(V_n)h(L_n^{\eta,\gamma})\mathbf{1}_{[\mathcal{G}_1 \cap \mathcal{G}_2]}\right] = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \mid \mathcal{F}_n^{\eta,\gamma}\right] h(L_n^{\eta,\gamma})\mathbf{1}_{[\mathcal{G}_1]}\right].$$

Therefore, since g and h are bounded, to prove (7.5) it suffices to show that as $n \to \infty$ and $T \to \infty$,

$$\left|\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \mid \mathcal{F}_n^{\eta,\gamma}\right]\mathbf{1}_{[\mathcal{G}_1]} - \mathbf{E}\left[g(V_n)\right]\right| \xrightarrow{\mathrm{p}} 0. \tag{7.6}$$

To prove (7.6), we will use the measure change between a size-biased random array and a vector of IID size-biased random variables which may be found in Proposition A.4 below. To this end, let ξ^n denote a random variable with distribution μ, conditioned not to yield a large displacement vector (i.e. conditioned on $\max_{1 \le i \le \xi^n} |Y_{\xi^n,i}| \le \gamma n^{1/(4-\eta)}$), and let μ^n denote the distribution of ξ^n. Using similar notation to that in Proposition A.4, write r_n for the value of $M_n^{\eta,\gamma}$, s_n for the value of $S_n^{\eta,\gamma}$ and $d_1, \ldots, d_{r_n}$ for the degrees of the vertices v with $\|Y^{(v)}\| > \gamma n^{1/(4-\eta)}$. Then, let $\xi_{r_n+1}^n, \ldots, \xi_n^n$ be IID samples from μ^n and write $\vec{Z} = (Z_1, \ldots, Z_n) = (d_1, \ldots, d_{r_n}, \xi_{r_n+1}^n, \ldots, \xi_n^n)$. Further, conditionally given $\vec{Z}$, let $\Sigma = \Sigma_{\vec{Z}}$ be the random permutation in (A.1), so that $(Z_{\Sigma(1)}, \ldots, Z_{\Sigma(n)})$ is a size-biased random re-ordering of $\vec{Z}$. Also define $\tau_{r_n}(\Sigma) = \min\{j \in [n] : \Sigma(j) \in [r_n]\}$. Finally, write $N = N_{n,r_n} = |\{i \in \{r_n + 1, \ldots, n\} : \xi_i^n > 0\}|$.

Note that conditionally on $\mathcal{F}_n^{\eta,\gamma}$, the remaining vertex degrees are distributed as $\xi_{r_n+1}^n, \ldots, \xi_n^n$ conditioned on $\xi_{r_n+1}^n + \cdots + \xi_n^n = n - 1 - s_n$. Therefore,

$$\begin{aligned}
&\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \mid \mathcal{F}_n^{\eta,\gamma}\right] \\
&\quad = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \,\middle|\, \xi_{r_n+1}^n, \ldots, \xi_n^n, \sum_{i=r_n+1}^{n} \xi_i^n = n - 1 - s_n, \mathcal{F}_n^{\eta,\gamma}\right] \,\middle|\, \mathcal{F}_n^{\eta,\gamma}\right].
\end{aligned}$$

By (A.18), $\tau_{r_n}(\Sigma) > T\sqrt{n}$ with high probability. Furthermore, by a Chernoff bound, $N \ge T\sqrt{n}$ with high probability. It follows that

$$\begin{aligned}
&\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \,\middle|\, \xi_{r_n+1}^n, \ldots, \xi_n^n, \sum_{i=r_n+1}^{n} \xi_i^n = n - 1 - s_n, \mathcal{F}_n^{\eta,\gamma}\right] \,\middle|\, \mathcal{F}_n^{\eta,\gamma}\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\mathbf{1}_{[N \ge T\sqrt{n}, \tau_{r_n}(\Sigma) > T\sqrt{n}]} \,\middle|\, \xi_{r_n+1}^n, \ldots, \xi_n^n, \sum_{i=r_n+1}^{n} \xi_i^n = n - 1 - s_n, \mathcal{F}_n^{\eta,\gamma}\right] \,\middle|\, \mathcal{F}_n^{\eta,\gamma}\right] \\
&\quad + o_{\mathbf{P}}(1).
\end{aligned}$$

Now observe that, on the event $\mathcal{G}_2$, T_n^k contains at most $T\sqrt{n}$ vertices, and further on the event $\tau_{r_n}(\Sigma) > T\sqrt{n}$, none of these vertices have a displacement exceeding $\gamma n^{1/(4-\eta)}$. This implies that $g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]}$ only depends on $\xi^n_{r_n+1},\dots,\xi^n_n$ and $\mathcal{F}_n^{\eta,\gamma}$ through $Z_{\Sigma(1)},\dots,Z_{\Sigma(\lfloor T\sqrt{n}\rfloor)}$ and $\Sigma(1),\dots,\Sigma(\lfloor T\sqrt{n}\rfloor)$. Therefore,

$$\begin{aligned}
&\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\,\middle|\,\mathcal{F}_n^{\eta,\gamma}\right]\\
&=\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]}\,\middle|\,\left(Z_{\Sigma(i)}\right)_{i\in[\lfloor T\sqrt{N}\rfloor]},(\Sigma(i))_{i\in[\lfloor T\sqrt{N}\rfloor]}\right]\middle|\mathcal{F}_n^{\eta,\gamma}\right]\\
&\quad+o_{\mathbf{P}}(1).\\
&=\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\,\middle|\,\left(Z_{\Sigma(i)}\right)_{i\in[\lfloor T\sqrt{N}\rfloor]},(\Sigma(i))_{i\in[\lfloor T\sqrt{N}\rfloor]}\right]\mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]}\middle|\mathcal{F}_n^{\eta,\gamma}\right]\\
&\quad+o_{\mathbf{P}}(1),
\end{aligned}$$

where the last equality is implied by the fact that the events $N\geq T\sqrt{n}$ and $\tau_{r_n}(\Sigma)>T\sqrt{n}$ are measurable with respect to $Z_{\Sigma(1)},\dots,Z_{\Sigma(\lfloor T\sqrt{n}\rfloor)},\Sigma(1),\dots,\Sigma(\lfloor T\sqrt{n}\rfloor)$. However, observe that $g(V_n)\mathbf{1}_{[\mathcal{G}_2]}$ is independent of $\Sigma(1),\dots,\Sigma(\lfloor T\sqrt{n}\rfloor)$ given $Z_{\Sigma(1)},\dots,Z_{\Sigma(\lfloor T\sqrt{n}\rfloor)}$, and so

$$\begin{aligned}
&\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\,\middle|\,\left(Z_{\Sigma(i)}\right)_{i\in[\lfloor T\sqrt{N}\rfloor]},(\Sigma(i))_{i\in[\lfloor T\sqrt{N}\rfloor]}\right]\mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]}\,\middle|\,\mathcal{F}_n^{\eta,\gamma}\right]\\
&=\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\,\middle|\,\left(Z_{\Sigma(i)}\right)_{i\in[\lfloor T\sqrt{N}\rfloor]}\right]\mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]}\,\middle|\,\mathcal{F}_n^{\eta,\gamma}\right].
\end{aligned}\tag{7.7}$$

We now apply the measure change from Proposition A.4 to obtain that, for $\bar{\xi}^n_1,\bar{\xi}^n_2,\dots$ IID samples from the size-biased law of μ^n, (7.7) is equal to

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor}\right]\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor})\mid\mathcal{F}_n^{\eta,\gamma}\right],\tag{7.8}$$

where the inner conditional expectation of $g(V_n)\mathbf{1}_{[\mathcal{G}_2]}$ is now thought of as a measurable functional of the IID random variables $\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor}$ in place of $Z_{\Sigma(1)},\dots,Z_{\Sigma(\lfloor T\sqrt{n}\rfloor)}$. This implies that

$$\begin{aligned}
&\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\mid\mathcal{F}_n^{\eta,\gamma}\right]\mathbf{1}_{[\mathcal{G}_1]}\\
&\quad=\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor}\right]\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor})\middle|\mathcal{F}_n^{\eta,\gamma}\right]\mathbf{1}_{[\mathcal{G}_1]}+o_{\mathbf{P}}(1).
\end{aligned}$$

By applying Lemma A.5 on $\mathcal{G}_1$ (which occurs with high probability),

$$\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor})\xrightarrow{\mathrm{P}}1$$

as $n\to\infty$ and $(\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor}))_{n\geq 0}$ is uniformly integrable, so (7.8) is equal to

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor}\right]\middle|\mathcal{F}_n^{\eta,\gamma}\right]+o_{\mathbf{P}}(1).$$

Since $\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor}\right]$ does not depend on $\mathcal{F}_n^{\eta,\gamma}$, it follows that

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor}\right]\middle|\mathcal{F}_n^{\eta,\gamma}\right]=\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}^n_1,\dots,\bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor}\right]\right].$$

By Corollary A.10, the total variation distance between $(\bar{\xi}^n_1, \dots, \bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor})$ and IID size-biased samples from μ, henceforth denoted by $(\bar{\xi}_1, \dots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor})$, tends to 0 as $n \to \infty$. Therefore, since g is bounded,

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}^n_1, \dots, \bar{\xi}^n_{\lfloor T\sqrt{n}\rfloor}\right]\right] = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}_1, \dots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}\right]\right] + o(1).$$

Finally, by Lemma 4.4 and Proposition 4.3, this is in turn equal to

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\widehat{D}^n_1, \dots, \widehat{D}^n_{\lfloor T\sqrt{n}\rfloor}\right]\mathbf{1}_{[N_n \geq T\sqrt{n}\rfloor]}\right] + o(1),$$

where we recall that $N_n = |\{i \in [n] \ : \ D^n_i > 0\}|$. Again, since the probability of $\mathcal{G}_2$ and $N_n \geq \lfloor T\sqrt{n}\rfloor$ occurring tends to 1 as $n \to \infty$ and subsequently $T \to \infty$, we see that

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\widehat{D}^n_1, \dots, \widehat{D}^n_{\lfloor T\sqrt{n}\rfloor}\right]\mathbf{1}_{[N_n \geq T\sqrt{n}\rfloor]}\right] = \mathbf{E}\left[g(V_n)\right] + o(1),$$

which proves (7.6). The result follows. □

7.1. *Typical displacements.* Fix $\eta \in [0, 2)$ and $\delta \in (0, 1/(10 - 4\eta)) \subset (0, 1/(4 - \eta))$. In this section we will study the function encoding the spatial locations of the branching random walk $\mathbf{T}_{n,\delta} = (\mathrm{T}_n, Y_{n,\delta})$, namely $R_{n,\delta} : [0, n] \to \mathbb{R}$.

PROPOSITION 7.3. *Fix $\gamma > 0$ and suppose that [**A1**] holds and [**A3**] holds for a given measure π with $\eta \in [0, 2)$. Let $\delta \in (0, 1/(10 - 4\eta))$. If $\eta = 0$, then*

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_{n,\delta}(nt)}{n^{1/4}}\right)_{0\leq t\leq 1}, \frac{L^{0,\gamma}_n}{n^{1/4}}\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\leq t\leq 1}, L^{0,\gamma}\right),$$

as $n \to \infty$, in $\mathbf{C}([0, 1], \mathbb{R}^2) \times \ell_\infty$. Furthermore, $L^{0,\gamma}$ is independent of $((\mathbf{e}_t, \mathbf{r}_t))_{0\leq t\leq 1}$.

If $\eta \in (0, 2)$, then

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_{n,\delta}(nt)}{n^{1/(4-\eta)}}\right)_{0\leq t\leq 1}, \frac{L^{\eta,\gamma}_n}{n^{1/(4-\eta)}}\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t, 0\right)_{0\leq t\leq 1}, L^{\eta,\gamma}\right),$$

in $\mathbf{C}([0, 1], \mathbb{R}^2) \times \ell_\infty$, where $L^{\eta,\gamma}$ is independent of $(\mathbf{e}_t)_{0\leq t\leq 1}$.

PROOF. The convergence of the random finite-dimensional distributions follows from Proposition 7.2 exactly as Corollary 4.2 follows from Proposition 4.1, but now with the additional independence from $L^{\eta,\gamma}$.

We will obtain tightness (now on the scale of $n^{1/(4-\eta)}$) via arguments very similar to those in Section 5, where we replace the truncations with those defined in Section 7. In particular, the key point is that we must show the analogue of Proposition 5.7, which states that

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{\max_{0\leq i\leq k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |R_{n,\delta}(s) - R_{n,\delta}(t)| > \gamma n^{1/(4-\eta)}\right\} = 0.$$

Fix $\delta \in (0, 1/(10 - 4\eta))$. For all $n \geq 1$ and $k \geq 1$ let $Y^{n,\delta}_k \in \mathbb{R}^k$ be such that

$$Y^{n,\delta}_k = (Y^{n,\delta}_{k,1}, \dots, Y^{n,\delta}_{k,k}) := \begin{cases} (Y_{k,1}, \dots, Y_{k,k}) & \text{if } \max_{1\leq j\leq k} |Y_{k,j}| \leq n^{1/(4-\eta)-\delta}, \\ (0, \dots, 0) & \text{else.} \end{cases}$$

As discussed in Section 6, the displacements of the branching random walk $\mathbf{T}_{n,\delta}$ are not necessarily globally centered and so may not satisfy [**A1**]. Thus to prove the result, we will need to instead consider the re-centered branching random walk $(\mathrm{T}_n, Y^*_{n,\delta})$ where conditionally on T_n, the entries of $Y^*_{n,\delta} = (Y^{*,(v)}_{n,\delta},\ v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n)$ are independent random vectors, such that if $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ has k children then $Y^{*,(v)}_{n,\delta}$ has the same distribution as

$$Y^{n,\delta}_k - \mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right].$$

The function $R^*_{n,\delta} : [0,n] \to \mathbb{R}$ encoding the spatial locations of $(\mathrm{T}_n, Y^*_{n,\delta})$ is such that for all $t \in [0,n]$,

$$R^*_{n,\delta}(t) \overset{\mathrm{d}}{=} R_{n,\delta}(t) - \mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right] \cdot H_n(t). \tag{7.9}$$

By Lemma A.11,

$$\left|\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right]\right| = O\left((n^{1/(4-\eta)-\delta})^{1-2(4-\eta)/3}\right).$$

Since $(n^{-1/2}H_n(nt))_{0\le t\le 1} \xrightarrow{\mathrm{d}} \frac{2}{\sigma}(\mathbf{e}_t)_{0\le t\le 1}$ as $n\to\infty$ in $\mathbf{C}([0,1],\mathbb{R})$, it then follows that

$$\frac{\|H_n\|_\infty}{n^{1/(4-\eta)}}\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right] \xrightarrow{\mathrm{p}} 0 \tag{7.10}$$

as long as $\delta > 0$ satisfies

$$\left(\frac{1}{4-\eta} - \delta\right)\left(1 - \frac{2(4-\eta)}{3}\right) < \frac{1}{4-\eta} - \frac{1}{2}.$$

Rearranging, this is equivalent to requiring that $\delta < (10-4\eta)^{-1}$. For these values of δ, we then have

$$\sup_{t\in[0,1]} |R^*_{n,\delta}(t) - R_{n,\delta}(t)| \xrightarrow{\mathrm{d}} 0,$$

and so there is no asymptotic cost in doing this re-centering. Arguing again exactly as in Section 5, it is sufficient to prove the analogue of Lemma 5.10, which states that there exists $A > 0$ such that for any $\gamma > 0$, $\delta \in (0, 1/(4-\eta))$ and $n \ge 1$ we have

$$\mathbf{P}\left\{\|R^*_{n,\delta}\|_\infty > \gamma n^{1/(4-\eta)}\right\} \le \frac{A}{\gamma^8}.$$

It is straightforward to verify that the proof of Lemma 5.10 given in Section 6 generalises immediately to this setting, on replacing $n^{1/4}$ by $n^{1/(4-\eta)}$. □

7.2. *Mid-range and large displacements.* We will adapt the proof of Proposition 5.3 to the case where [**A3**] holds instead of [**A2**]. The proof of Proposition 5.3 uses Lemma A.12 to show that, with high probability, there are no vertices with a mid-range or large displacement that are ancestrally related. To apply that lemma, it is sufficient to bound both the maximum degree in the tree and the number of vertices with a mid-range or large displacement, with high probability. The required bound on the maximal degree follows from the assumption that $\mathbf{E}\left[\xi^3\right] < \infty$. Therefore, for the adaptation, we need to obtain the same control on the number of mid-range displacements under [**A3**] as we obtained under [**A2**] in Lemma 5.4.

LEMMA 7.4. *Suppose that [**A3**] holds for a given measure π and $\eta \in [0,2)$. For $\delta > 0$ sufficiently small,*

$$\left|\left\{v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n \text{ such that } \|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta}\right\}\right| = o_{\mathbf{P}}(n^{1/12}).$$

PROOF. Let $\xi_1, \dots, \xi_n$ be IID with distribution μ. Let $x \in (0,1)$ be such that $\pi(\{x\} \times \mathbb{R}_+) = \pi(\mathbb{R}_+ \times \{x\}) = 0$. Then, by [A3],

$$n^{1-(4-\eta)\delta}\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > n^{1/(4-\eta)-\delta}\right\} \le n^{1-(4-\eta)\delta}\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > xn^{1/(4-\eta)-\delta}\right\}$$
$$\to \pi\big(((x,\infty)\times\mathbb{R}_+) \cup (\mathbb{R}_+ \times (x,\infty))\big) < \infty.$$

This in particular implies that there exists $C > 0$ such that $\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > n^{1/(4-\eta)-\delta}\right\} \le Cn^{-1+(4-\eta)\delta}$ for all $n \ge 1$. It follows that

$$A_n := \left|\left\{i \in [n] \,:\, \|Y_{\xi_i}\|_\infty > n^{1/(4-\eta)-\delta}\right\}\right| \preceq_{st} \mathrm{Bin}\left(n, Cn^{-1+(4-\eta)\delta}\right).$$

By a Chernoff bound, this implies that for $\delta \in (0, (12(4-\eta))^{-1})$, and $n \ge 1$ sufficiently large, for any $\varepsilon > 0$,

$$\begin{aligned}\mathbf{P}\left\{A_n > \varepsilon n^{1/12}\right\} &\le \mathbf{P}\left\{\mathrm{Bin}\left(n, Cn^{-1+(4-\eta)\delta}\right) > \varepsilon n^{1/12}\right\} \\ &= \mathbf{P}\left\{\mathrm{Bin}\left(n, Cn^{-1+(4-\eta)\delta}\right) > Cn^{(4-\eta)\delta}\left(1 + \left(\frac{\varepsilon}{C}n^{1/12-(4-\eta)\delta} - 1\right)\right)\right\} \\ &= O\left(\exp(-n^{(4-\eta)\delta})\right),\end{aligned}$$

so

$$\mathbf{P}\left\{A_n > \varepsilon n^{1/12} \;\middle|\; \sum_{i=1}^n \xi_i = n-1\right\} = O\left(n^{1/2}\exp(-n^{(4-\eta)\delta})\right) = o(1). \qquad \square$$

We now obtain the following result on the mid-range displacements under [A3] with a proof that is analogous to that of Proposition 5.3; we omit the details.

PROPOSITION 7.5. *Fix $\gamma > 0$ and suppose that [A3] holds for a given measure π and $\eta \in [0,2)$. For $\delta > 0$ sufficiently small, as $n \to \infty$,*

$$\mathbf{P}\left\{\|R^\gamma_{n,\delta}\|_\infty > \gamma n^{1/(4-\eta)}\right\} = o(1).$$

In the remainder of this section we will study the function encoding the spatial locations of the "large-displacement" branching random walk $\mathbf{T}^\gamma_n = (\mathrm{T}_n, Y^\gamma_n)$, namely $R^\gamma_n : [0,n] \to \mathbb{R}$.

Let Ξ be a Poisson process on $[0,1] \times \mathbb{R}^2_+ \setminus \{(0,0)\}$ with intensity $dt \otimes \pi(dx,dy)$, and let Ξ^γ be the restriction of Ξ to $[0,1] \times (\mathbb{R}^2_+ \setminus ([0,\gamma]\times[0,\gamma]))$. Also, recall the definition of the function U from just before Theorem 1.4.

PROPOSITION 7.6. *Fix $\gamma > 0$ and suppose that [A1] holds and that [A3] holds for a given measure π and $\eta \in [0,2)$. Let $\delta \in \left(0, \frac{1}{64} \wedge \frac{1}{10-4\eta}\right)$.*

If $\eta = 0$ then as $n \to \infty$,

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_{n,\delta}(nt)}{n^{1/4}}\right)_{0\le t\le 1}, U\left(\frac{R^\gamma_n}{n^{1/4}}, \emptyset\right)\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\le t\le 1}, U(0,\Xi^\gamma)\right)$$

with convergence in the first coordinate in $\mathbf{C}([0,1],\mathbb{R}^2)$, and convergence in the second coordinate with respect to the Hausdorff topology on non-empty compact subsets. Furthermore, $U(0,\Xi^\gamma)$ is independent of $(\mathbf{e}_t, \mathbf{r}_t,\ 0 \le t \le 1)$.

If $\eta \in (0,2)$ then as $n \to \infty$,

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_{n,\delta}(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1}, U\left(\frac{R_n^\gamma}{n^{1/(4-\eta)}}, \emptyset\right)\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t, 0\right)_{0\le t\le 1}, U(0, \Xi^\gamma)\right)$$

with convergence in the first coordinate in in $\mathbf{C}([0,1], \mathbb{R}^2)$, and convergence in the second coordinate with respect to the Hausdorff topology on non-empty compact subsets. Furthermore, $U(0,\Xi^\gamma)$ is independent of $(\mathbf{e}_t, 0 \le t \le 1)$.

We first prove Theorems 1.4 and 1.5 assuming Proposition 7.6.

PROOF OF THEOREMS 1.4 AND 1.5 ASSUMING PROPOSITION 7.6. For γ and δ as in Proposition 7.6,

$$\left(\frac{R_n(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1} = \left(\frac{R_{n,\delta}(nt)}{n^{1/(4-\eta)}} + \frac{R_n^\gamma(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1} + \left(\frac{R_{n,\delta}^\gamma(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1}.$$

By Proposition 7.6, as $n \to \infty$, $U(n^{-1/(4-\eta)} R_n^\gamma(n\,\cdot), \emptyset) \xrightarrow{\mathrm{d}} U(0,\Xi^\gamma)$ with respect to the Hausdorff topology on non-empty compact subsets, jointly with convergence

$$\left(\frac{R_{n,\delta}(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} \begin{cases} \beta\sqrt{\frac{2}{\sigma}}\mathbf{r} & \text{if } \eta = 0, \\ 0 & \text{if } \eta \in (0,2) \end{cases}$$

in $\mathbf{C}([0,1],\mathbb{R}^2)$ where, for $\eta = 0$, $U(0,\Xi^\gamma)$ and $(\mathbf{r}_t)_{0\le t\le 1}$ are independent. Therefore,

$$U\left(\frac{R_{n,\delta}(n\cdot)}{n^{1/(4-\eta)}} + \frac{R_n^\gamma(n\cdot)}{n^{1/(4-\eta)}}, \emptyset\right) \xrightarrow{\mathrm{d}} \begin{cases} U\left(\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}, \Xi^\gamma\right) & \text{if } \eta = 0, \\ U(0,\Xi^\gamma) & \text{if } \eta \in (0,2). \end{cases}$$

Note that $U(0,\Xi)$ is a compact set by our assumptions on π, and that $U(0,\Xi^\gamma) \xrightarrow{\text{a.s.}} U(0,\Xi)$ in the Hausdorff sense as $\gamma \downarrow 0$. We have

$$d_{\mathrm{H}}\left(U\left(\frac{R_n(n\cdot)}{n^{1/(4-\eta)}}, \emptyset\right), U\left(\frac{R_{n,\delta}(n\cdot)}{n^{1/(4-\eta)}} + \frac{R_n^\gamma(n\cdot)}{n^{1/(4-\eta)}}, \emptyset\right)\right) \le n^{-1/(4-\eta)} \|R_{n,\delta}^\gamma\|_\infty$$

and, by Proposition 7.5,

$$\lim_{\gamma\to 0} \limsup_{n\to\infty} \mathbf{P}\left\{\|R_{n,\delta}^\gamma\|_\infty > \gamma n^{1/(4-\eta)}\right\} = 0.$$

We may now apply the principle of accompanying laws [6, Theorem 3.2] in order to obtain that

$$U\left(\frac{R_n(n\cdot)}{n^{1/(4-\eta)}}, \emptyset\right) \xrightarrow{\mathrm{d}} \begin{cases} U\left(\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}, \Xi\right) & \text{if } \eta = 0, \\ U(0,\Xi) & \text{if } \eta \in (0,2), \end{cases}$$

which yields Theorems 1.4 and 1.5. □

The remainder of this section is devoted to the proof of Proposition 7.6. We will need a notion of pruning and grafting of branching random walks. We refer to Figure 6 as a visual aid in understanding the following three definitions.

DEFINITION 7.7 (Pruning branching random walks). Let $\mathrm{T} = (T, Y)$ be a branching random walk with displacements $Y = (Y^{(v)}, v \in v(T) \setminus \partial T)$. Let $v \in v(T)$ and $T^{(v)}$ be the subtree of T rooted at v. The sub-branching random walk of T rooted at v is the branching random walk $\mathrm{T}^{(v)} = (T^{(v)}, Y')$ with displacements $Y' = (Y'^{(u)}, u \in v(T^{(v)}) \setminus \partial T^{(v)})$. Also, $\mathrm{T}^{\uparrow v}$ is the branching random walk obtained by removing all descendants of v from T. More generally, for $\mathrm{v} = (v_1, \ldots, v_k)$ a sequence of distinct vertices in $v(T)$ such that no two vertices in v are ancestrally related, we set $\mathrm{T}^{\mathrm{v}} = (\mathrm{T}^{(v)}, v \in \mathrm{v})$, and define $\mathrm{T}^{\uparrow \mathrm{v}}$ inductively as $\mathrm{T}^{\uparrow \mathrm{v}} = (\mathrm{T}^{\uparrow (v_1, \ldots, v_{k-1})})^{\uparrow v_k}$.

DEFINITION 7.8 (Grafting branching random walks). For branching random walks $\mathrm{T} = (T, Y)$ and $\mathrm{T}' = (T', Y')$, and for a leaf $l \in \partial T$, let $\mathrm{T} \oplus_l \mathrm{T}' = (T \oplus_l T, Y \oplus_l Y')$ be the branching random walk defined by setting $T \oplus_l T' = T \cup lT'$ and, for $v \in v(T \oplus_l T') \setminus \partial(T \oplus_l T')$, setting

$$(Y \oplus_l Y')^{(v)} = \begin{cases} Y^{(v)} & \text{if } v \in v(T) \setminus \partial T, \\ Y'^{(u)} & \text{if } v = lu \text{ for some } u \in v(T') \setminus \partial T'. \end{cases}$$

More generally, for branching random walks $\mathrm{T}, \mathrm{T}^1, \ldots, \mathrm{T}^k$ and distinct leaves $l_1, \ldots, l_k \in \partial T$, define $\mathrm{T} \oplus_{l_1, \ldots, l_k} (\mathrm{T}^1, \ldots, \mathrm{T}^k)$ recursively over k as

$$\mathrm{T} \oplus_{l_1, \ldots, l_k} (\mathrm{T}^1, \ldots, \mathrm{T}^k) = (\mathrm{T} \oplus_{l_1, \ldots, l_{k-1}} (\mathrm{T}^1, \ldots, \mathrm{T}^{k-1})) \oplus_{l_k} \mathrm{T}^k.$$

The previous definitions imply that for a branching random walk $\mathrm{T} = (T, Y)$ and $v \in v(T)$,

$$\mathrm{T}^{\uparrow v} \oplus_v \mathrm{T}^{(v)} = \mathrm{T},$$

and, more generally, for a sequence of distinct vertices $\mathrm{v} = (v_1, \ldots, v_k)$ of T such that no two vertices in v are ancestrally related in T, that $\mathrm{T}^{\uparrow \mathrm{v}} \oplus_{\mathrm{v}} \mathrm{T}^{\mathrm{v}} = \mathrm{T}$.

We next use the above definitions to define a map that prunes the sub-branching random walks of branching random walks that are rooted at ancestrally minimal vertices v with $\|Y^{(v)}\|_\infty \geq \tau$. See Figure 6 for an illustration of the coming definition.

DEFINITION 7.9. For a branching random walk $\mathrm{T} = (T, Y)$ and for $\tau > 0$, let $\mathrm{v}_\tau = (v_1, \ldots, v_m)$ be the set of vertices $v \in v(T)$ such that $\|Y^{(v)}\|_\infty > \tau$ and for all ancestors $u \preceq v$, $\|Y^{(u)}\| \leq \tau$, listed in depth-first order. Define a map f_τ by

$$\mathrm{T} \overset{f_\tau}{\longmapsto} (\mathrm{T}^{\uparrow \mathrm{v}_\tau}, \{\mathrm{T}^{(v_1)}, \ldots, \mathrm{T}^{(v_m)}\}),$$

where the second coordinate is a multiset with elements $\mathrm{T}^{(v_1)}, \ldots, \mathrm{T}^{(v_m)}$ which are the branching random walks rooted at the vertices $v_1, \ldots, v_m$.

For $\tau \geq 0$, let

$$v^\tau(\mathbf{T}_n) := \left\{ v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n \,:\, \|Y^{(v)}\|_\infty > \tau, \text{ and } \|Y^{(u)}\|_\infty \leq \tau \; \forall u \prec v \right\}.$$

We will apply f_τ to $\mathbf{T}_n$, and then study the law of $\mathbf{T}_n$ conditional on $f_\tau(\mathbf{T}_n)$. Observe that, given $f_\tau(\mathbf{T}_n)$, $\mathbf{T}_n$ is determined by $\mathrm{v}^\tau(\mathbf{T}_n)$. We will show that conditional on $f_\tau(\mathbf{T}_n)$, $\mathrm{v}^\tau(\mathbf{T}_n)$ is distributed as a uniformly random subset of leaves in $(f_\tau(\mathbf{T}_n))_1$. We make this formal in the next lemma.

LEMMA 7.10. *Let $\tau > 0$ and write $f_\tau(\mathbf{T}_n) = (\mathbf{T}'_n, \{\mathbf{T}^1_n, \ldots, \mathbf{T}^m_n\})$. Fix $m \geq 1$ and let $\Sigma \in_{\mathcal{U}} \mathcal{S}_m$, where $\mathcal{S}_m$ is the symmetric group of order m. Further, let $(\mathcal{L}_1, \ldots, \mathcal{L}_m)$ be a uniformly random vector of leaves in $\mathbf{T}'_n$ listed in depth-first order. Then, given $f_\tau(\mathbf{T}_n)$, $\mathbf{T}_n$ is equal in distribution to*

$$\mathbf{T}'_n \oplus_{\mathcal{L}_1, \ldots, \mathcal{L}_m} (\mathbf{T}^{\Sigma(1)}_n, \ldots, \mathbf{T}^{\Sigma(m)}_n).$$

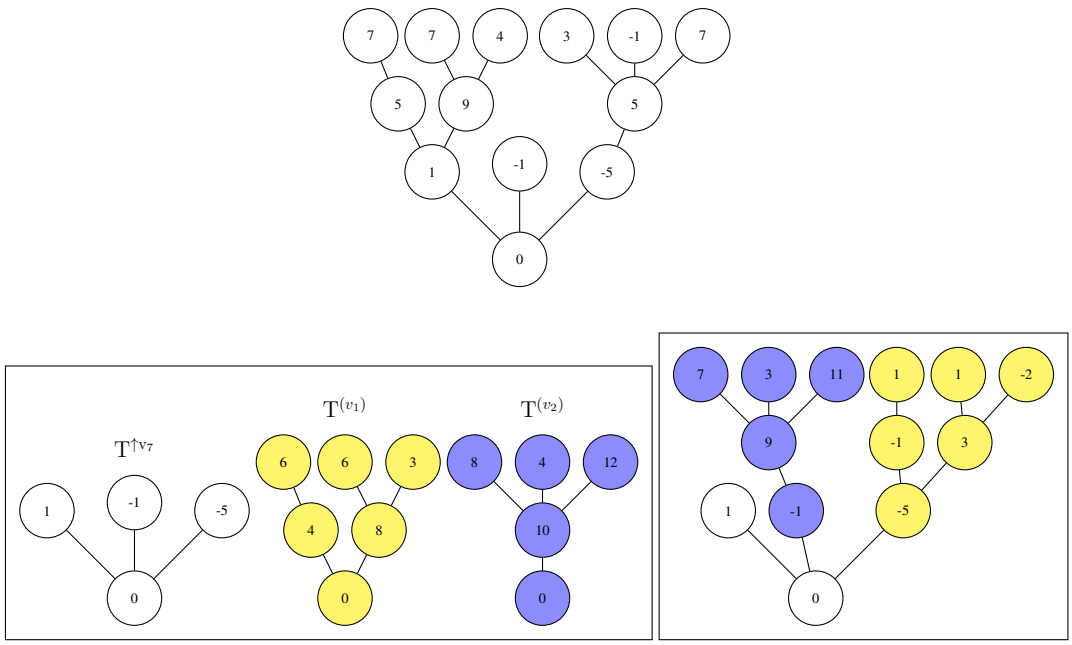

FIG 6. *On top, a spatial tree. We denote the associated branching walk by* T. *On the bottom left, we depict* $f_7(\mathrm{T}) = (\mathrm{T}^{\uparrow v_7}, \{\mathrm{T}^{(v_1)}, \mathrm{T}^{(v_2)}\})$, *which is obtained from* T *by pruning the sub-branching walks of* T *that have a displacement with absolute value exceeding* 7 *in their first generation. On the bottom right is a spatial tree obtained by grafting the branching walks* $(f_7(\mathrm{T}))_2$ *to leaves of* $(f_7(\mathrm{T}))_1$. $\mathrm{T}^{(v_1)}$ *and* $\mathrm{T}^{(v_2)}$ *to leaves of* $\mathrm{T}^{\uparrow v_7}$.

PROOF. Let $(\mathrm{t}', \{\mathrm{t}^1, \dots, \mathrm{t}^m\})$ be in the support of $f_\tau(\mathbf{T}_n)$. We will first show that

(7.11)
$$f_\tau^{-1}(\mathrm{t}', \{\mathrm{t}^1, \dots, \mathrm{t}^m\}) = \left\{ \mathrm{t}' \oplus_{l_1,\dots,l_m} (\mathrm{t}^{\pi(1)}, \dots, \mathrm{t}^{\pi(m)}) : (l_1, \dots, l_m) \text{ leaves in } \mathrm{t}'; \pi \in \mathcal{S}_m \right\},$$

where on the right-hand side, $(l_1, \dots, l_m)$ are listed in depth-first order. Following this, we will show that the law of $\mathbf{T}_n$ conditional on its degrees and displacement vectors assigns equal mass to all elements of the right-hand set in (7.11), and that each element of the right-hand set corresponds to the same number of sets of leaves $(l_1, \dots, l_m)$ listed in depth-first order and permutations π.

For the inclusion of the left-hand set in the right-hand set, observe that if for some spatial tree t it holds that $f_\tau(\mathrm{t}) = (\mathrm{t}', \{\mathrm{t}^1, \dots, \mathrm{t}^m\})$ then $(l_1, \dots, l_m)$, the minimal vertices in t that have a displacement vector with sup-norm lower bounded by τ listed in depth-first order, are leaves in t'. Thus there is some $\pi \in \mathcal{S}_m$ such that for all $i = 1, \dots, m$, $\mathrm{t}^{\pi(i)} = \mathrm{t}^{(l_i)}$. This implies that $\mathrm{t} = \mathrm{t}' \oplus_{l_1,\dots,l_m} (\mathrm{t}^{\pi(1)}, \dots, \mathrm{t}^{\pi(m)})$.

For the other inclusion, it is straightforward to see that for leaves $(l_1, \dots, l_m)$ in t', listed in depth-first order, and $\pi \in \mathcal{S}_m$, it holds that $f_\tau(\mathrm{t}' \oplus_{l_1,\dots,l_m} (\mathrm{t}^{\pi(1)}, \dots, \mathrm{t}^{\pi(m)}) = (\mathrm{t}', \{\mathrm{t}^1, \dots, \mathrm{t}^m\})$.

We now show that the law of $\mathbf{T}_n$ conditional on its degrees and displacement vectors assigns equal mass to all elements of the right-hand set. This follows from the observation that, conditional on its degrees and displacement vectors, $\mathbf{T}_n$ is uniform on all branching random walks with those degrees and displacement vectors. Each element in

(7.12)
$$\left\{ \mathrm{t}' \oplus_{l_1,\dots,l_m} (\mathrm{t}^{\pi(1)}, \dots, \mathrm{t}^{\pi(m)}) : (l_1, \dots, l_m) \text{ leaves in } \mathrm{t}'; \pi \in \mathcal{S}_m \right\}$$

with $l_1, \dots, l_m$ listed in depth-first order has the same degrees and displacement vectors.

Finally, we show that each element of (7.12) corresponds to the same number of sets of leaves $(l_1, \dots, l_m)$ and permutations $\pi \in \mathcal{S}_m$. To this end, note that every vertex in a spatial

tree t with a displacement vector whose sup-norm is at least τ has a non-zero number of children, so for each t in the set (7.12), we can recognise $(l_1,\dots,l_m)$ as the vertices v that are leaves in t' and not leaves in t; thus, the choice of $(l_1,\dots,l_m)$ is unique. Moreover, if the multiset $\{\mathrm{t}^1,\dots,\mathrm{t}^m\}$ contains j different spatial trees with multiplicities $m_1,\dots,m_j$ respectively, then t corresponds to $m!/(m_1!\dots m_j!)$ different permutations π. This number does not depend on t, and the statement follows. □

For $n \geq 1$, let $\tau_n = n^{1/(4-\eta)-\delta}$. Further, let $\mathbf{T}'_n = (\mathrm{T}'_n, Y')$ denote the first coordinate of $f_{\tau_n}(\mathbf{T}_n)$, and $\mathbf{F}_n^{\mathrm{pr}} = (\mathbf{T}_n^{(v)})_{v\in v^{\tau_n}(\mathrm{T}_n)}$ denote the second coordinate of $f_{\tau_n}(\mathbf{T}_n)$, where we assume that the trees in $\mathbf{F}_n^{\mathrm{pr}}$ are ordered according to the depth-first order of their roots in T_n. We require one further lemma to prove Proposition 7.6.

LEMMA 7.11. *Fix $\gamma > 0$. Suppose that [**A1**] holds and that [**A3**] holds for a given measure π and $\eta \in [0,2)$. For $n \geq 1$, let H'_n be the height function of T'_n and R'_n be the function encoding the spatial locations of $\mathbf{T}'_n$. Extend their domains to $[0,n]$ by setting $H'_n(t) = R'_n(t) = 0$ for all $t > |\mathrm{T}'_n|$. If $\eta = 0$, then as $n \to \infty$,*

$$\left(\left(\frac{H'_n(nt)}{\sqrt{n}}, \frac{R'_n(nt)}{n^{1/4}}\right)_{0\leq t\leq 1}, \frac{L_n^{0,\gamma}}{n^{1/4}},\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\leq t\leq 1}, L^{0,\gamma}\right), \tag{7.13}$$

and if $\eta \in (0,2)$, then

$$\left(\left(\frac{H'_n(nt)}{\sqrt{n}}, \frac{R'_n(nt)}{n^{1/(4-\eta)}}\right)_{0\leq t\leq 1}, \frac{L_n^{\eta,\gamma}}{n^{1/(4-\eta)}},\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t, 0\right)_{0\leq t\leq 1}, L^{\eta,\gamma}\right), \tag{7.14}$$

with convergence in the first coordinate in $\mathbf{C}([0,1],\mathbb{R}^2)$ endowed with the topology of uniform convergence, and the convergence in the second coordinate in ℓ_∞.

PROOF. We prove (7.13). The proof of (7.14) then follows by identical arguments. By Proposition 7.3 it suffices to prove that as $n \to \infty$,

$$\sup_{1\leq j\leq n}\left\{n^{-1/2}|H_n(j) - H'_n(j)| \vee n^{-1/4}|R_{n,\delta}(j) - R'_n(j)|\right\} \xrightarrow{\mathrm{p}} 0. \tag{7.15}$$

We also prove (7.15) using Proposition 7.3. Fix $\varepsilon > 0$. The sample paths of both $\mathbf{e}$ and $\mathbf{r}$ are almost surely continuous so, since $[0,1]$ is compact, they are in fact almost surely uniformly continuous. This implies that there exists $\rho > 0$ so that

$$\mathbf{P}\left\{\sup_{0\leq s<t\leq 1,|s-t|<\rho}\left|\frac{2}{\sigma}\mathbf{e}_s - \frac{2}{\sigma}\mathbf{e}_t\right| \vee \left|\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_s - \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right| > \varepsilon/2\right\} < \varepsilon/2.$$

Then, the convergence in Proposition 7.3 implies that for n sufficiently large, the probability that the event

$$B_n := \left\{\sup_{0\leq k<\ell\leq n,|k-\ell|<\rho n}\left\{\frac{|H_n(k) - H_n(\ell)|}{n^{1/2}} \vee \frac{|R_{n,\delta}(k) - R_{n,\delta}(\ell)|}{n^{1/4}}\right\} \geq \varepsilon\right\}$$

occurs is less than ε.

Next, let

$$v^*(\mathrm{T}_n) := \left\{v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n : \|Y^{(v)}\|_\infty \leq n^{1/4-\delta}, \exists\, u \prec v \text{ with } \|Y^{(u)}\|_\infty > n^{1/4-\delta}\right\}.$$

By identical methods as those used to prove Lemma 5.6, it can be seen that $v^*(\mathrm{T}_n) = o_{\mathbf{P}}(n)$ and so for n sufficiently large, $\mathbf{P}\{|\mathrm{T}'_n| \leq n - \rho n\} \leq \varepsilon$.

Now suppose that neither of the (unlikely, bad) events $\{|\mathrm{T}'_n| \le n - \rho n\}$ or B_n hold. Observe that H'_n and R'_n can respectively be obtained from H_n and $R_{n,\delta}$ by "skipping" all the vertices in $v^*(\mathrm{T}_n)$. To be precise, for $1 \le k \le |\mathrm{T}'_n|$, let $P_n(k)$ be the position of the k-th vertex that is not in $v^*(\mathrm{T}_n)$ in the depth-first order of T_n. Then,

$$(H'_n(k), R'_n(k)) = \begin{cases} (H_n(P_n(k)), R_{n,\delta}(P_n(k))) & \text{for } k = 1, \dots, |\mathrm{T}'_n|, \\ (0,0) & \text{for } k > |\mathrm{T}'_n|. \end{cases}$$

By our assumption that $n - |\mathrm{T}'_n| < \rho n$, we have $|P_n(k) - k| < \rho n$ for all k; by our assumption that

$$\sup_{0 \le k < \ell \le n, |k-\ell| < \rho n} \left\{ \frac{|H_n(k) - H_n(\ell)|}{n^{1/2}} \vee \frac{|R_{n,\delta}(k) - R_{n,\delta}(\ell)|}{n^{1/4}} \right\} < \varepsilon,$$

we then also have

$$\sup_{0 \le k \le n} \left\{ \frac{|H_n(k) - H'_n(k)|}{\sqrt{n}} \vee \frac{|R_{n,\delta}(k) - R'_n(k)|}{n^{1/4}} \right\} < \varepsilon. \tag{7.16}$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

With Lemma 7.11 in hand, we proceed to proving Proposition 7.6. In the proof, the pair $(\mathbf{T}'_n, \mathbf{F}^{\mathrm{pr}}_n)$ is as in Lemma 7.11. Observe that by Lemma 7.10, given $f_{\tau_n}(\mathbf{T}_n)$, we can obtain an object with the same law as $\mathbf{T}_n$ by grafting the branching random walks in $\mathbf{F}^{\mathrm{pr}}_n$ at uniformly random leaves of the first coordinate of $f_{\tau_n}(\mathbf{T}_n)$.

PROOF OF PROPOSITION 7.6. Let $n \ge 1$ be large enough so that $n^{1/(4-\eta)-\delta} < \gamma n^{1/(4-\eta)}$. Then if $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ is such that $\|Y^{(v)}\|_\infty > \gamma n^{1/(4-\eta)}$, it also holds that $\|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta}$. The proof of Proposition 5.5 can be adapted so that under [**A3**] for a given measure π and $\eta \in [0,2)$, for $\delta > 0$ sufficiently small, as $n \to \infty$

$$\mathbf{P}\left\{ \exists u, v \in \mathbf{T}_n, u \prec v, \text{ such that } \|Y^{(u)}\|_\infty \wedge \|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta} \right\} = o(1).$$

If follows that at the cost of throwing away an event of asymptotically vanishing probability, we may work on the event that there are no ancestrally related vertices $u, v \in v(\mathrm{T}_n)$ such that both $\|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta}$ and $\|Y^{(u)}\|_\infty > n^{1/(4-\eta)-\delta}$.

By Skorokhod's representation theorem, we may work on a probability space where the convergence in Lemma 7.11 holds almost surely.

We now use Lemma 7.10 to study the asymptotic law of R^γ_n conditional on $(\mathbf{T}'_n, \mathbf{F}^{\mathrm{pr}}_n)$. Lemma 7.10 implies that given $(\mathbf{T}'_n, \mathbf{F}^{\mathrm{pr}})$, we can obtain an object with the law of $\mathbf{T}_n$ by grafting each of the branching random walks in $\mathbf{F}^{\mathrm{pr}}_n$ onto uniformly random leaves in $\mathbf{T}'_n$. In fact, in order to obtain the (conditional) law of R^γ_n we only need to sample the positions of the vertices in $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ whose displacement vectors $Y^{(v)}$ satisfy that $\|Y^{(v)}\|_\infty \ge \gamma n^{1/(4-\eta)}$, since the trees of $\mathbf{F}^{\mathrm{pr}}_n$ attach to these vertices in exchangeable random order. We denote the branching random walks in $\mathbf{F}^{\mathrm{pr}}_n$ by $\mathbf{T}^{(v_1)}, \dots, \mathbf{T}^{(v_{M_n})}$ (ordered according to the depth-first order of their roots $v_1, \dots, v_{M_n} \in \mathrm{T}_n$). By symmetry we may assume that for $1 \le j \le M_n$, the largest and smallest displacement at the root of $\mathbf{T}^{(v_j)}$ (i.e., $Y^{(v_j,+)}$ and $Y^{(v_j,-)}$) are described by the j-th entry of $L^{\eta,\gamma}_n$.

We claim that as $n \to \infty$, $M_n \xrightarrow{\mathrm{d}} M$ for some finite, random variable M. Indeed, as $n \to \infty$, $n^{-1/(4-\eta)} L^{\eta,\gamma}_n \xrightarrow{\mathrm{a.s.}} L^{\eta,\gamma}$. Furthermore, since $\gamma > 0$, almost surely $L^{\eta,\gamma}$ has finitely many non-zero terms and each non-zero entry of $n^{-1/(4-\eta)} L^{\eta,\gamma}_n$ is at ℓ^∞ distance at least γ from $(0,0)$, there are finite random variables M and N such that $L^{\eta,\gamma}_n$ has M non-zero terms

for all $n > N$ large enough; i.e., the number of vertices $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ such that $\|Y^{(v)}\|_\infty \geq \gamma n^{1/(4-\eta)}$ is equal to M.

Now, for $k \geq 1$, let $\mathcal{L}'_n(k)$ denote the number of leaves in T'_n which are among the first k vertices in the depth-first order of the vertices of T_n. Then $\mathcal{L}'_n(k)$ is bounded from above by the number of down-steps of the Łukasiewicz path $W_n(k)$ of T_n by time k. It is bounded from below by this same number minus $|T_n \setminus T'_n|$ which is $o(n)$ by (7.13).

Therefore, by Lemma A.1, as $n \to \infty$,

$$\left(\frac{\mathcal{L}'_n(\lfloor nt \rfloor)}{n} \right)_{0 \leq t \leq 1} \xrightarrow{\mathrm{p}} (\mu_0 t)_{0 \leq t \leq 1} ,$$

so that the positions of M_n uniform leaves in T'_n in depth-first order converge upon rescaling by n^{-1} to M independent uniform samples from $[0,1]$, which we denote by $U_1, \ldots, U_M$ respectively. For all $1 \leq j \leq M_n$, we graft $\mathbf{T}^{(v_j)}$ (which has size $o(n)$ since T'_n has size $n - o(n)$ by (7.13)) onto the j-th such leaf of $\mathbf{T}'_n$, using the operation in Definition 7.8.

The branching random walk $\mathbf{T}^{(v_j)}$ contains exactly one vertex (namely the root) with displacement vector $\|Y^{(v)}\|_\infty > \gamma n^{1/(4-\eta)}$ (since we assumed that such vertices are not ancestrally related) and the largest and smallest displacements of this vertex are given by $L_n^{\eta,\gamma}(j)$. Therefore, asymptotically, $n^{-1/(4-\eta)} R_n^\gamma$ will contain a line segment from $(U_j, -Y_j^-)$ to $(U_j, -Y_j^+)$. This implies that if $\eta = 0$

$$\left(\left(\frac{H'_n(nt)}{\sqrt{n}}, \frac{R'_n(nt)}{n^{1/4}} \right)_{0 \leq t \leq 1}, U \left(\frac{R_n^\gamma}{n^{1/4}}, \emptyset \right) \right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma} \mathbf{e}_t, \beta \sqrt{\frac{2}{\sigma}} \mathbf{r}_t \right)_{0 \leq t \leq 1}, U(0, \Xi^\gamma) \right),$$

and if $\eta \in (0,2)$

$$\left(\left(\frac{H'_n(nt)}{\sqrt{n}}, \frac{R'_n(nt)}{n^{1/(4-\eta)}} \right)_{0 \leq t \leq 1}, U \left(\frac{R_n^\gamma}{n^{1/(4-\eta)}}, \emptyset \right) \right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma} \mathbf{e}_t, 0 \right)_{0 \leq t \leq 1}, U(0, \Xi^\gamma) \right).$$

The result then follows from (7.15). □

APPENDIX: STANDARD RESULTS AND REMAINING PROOFS

A.1. Standard results. In this section we provide standard results which we use throughout this work without proof. We start by stating a functional strong law of large numbers for sums of IID non-negative random variables that we use at multiple points in proofs of convergence of finite-dimensional distributions.

LEMMA A.1. *Let $X_1, X_2, \ldots$ be* IID *random variables with $X_1 \geq 0$ almost surely and $\mathbf{E}[X_1] = \mu < \infty$. Then, for any $a_n \uparrow \infty$,*

$$\left(\frac{1}{a_n} \sum_{i=1}^{\lfloor a_n t \rfloor} X_i, t \geq 0 \right) \xrightarrow{\text{a.s.}} (\mu t, t \geq 0),$$

uniformly on compact sets as $n \to \infty$.

The next result is a generalised local central limit theorem, from Theorem 13, Chapter VII of Petrov [35], which we use to prove tightness in Theorem 1.1.

THEOREM A.2 (Theorem 13, Chapter VII of Petrov [35]). *Let $(X_n)_{n \geq 1}$ be a sequence of* IID *integer-valued random variables. Suppose that $\mathbf{E}[X_1] = 0$, $\mathbf{Var}\{X_1\} = \sigma^2 > 0$,*

$\mathbf{E}\left[|X_1|^3\right] < \infty$, *and the maximal span of the distribution of* X_1 *is equal to* 1. *Let* $S_n = \sum_{i=1}^n X_i$. *Then,*

$$\sqrt{2\pi n\sigma}\mathbf{P}\left\{S_n = k\right\} = e^{-k^2/(2\sigma^2 n)}\left(1 + \frac{1}{\sqrt{n}}\frac{\gamma_3}{6\sigma^3}\left(\frac{k^3}{\sigma^3 n^{3/2}} - \frac{3k}{\sigma\sqrt{n}}\right)\right) + o(n^{-1/2}),$$

uniformly in $k \in \mathbb{Z}$, *where* γ_3 *is the third central moment of* X_1.

The last result is a quantitative local central limit theorem proved in [3, Lemma 5.5] for $k = 1$, which we use in the proof of Theorems 1.4 and 1.5. The generalisation to $k \geq 1$ is standard.

LEMMA A.3. *Fix* $\eta, \beta > 0$, $0 < \gamma < 1/2$ *and* $k \in \mathbb{N}$. *Then, there exist constants* $C = C(\eta, \beta, \gamma, k)$ *and* $M = M(\eta, \beta, \gamma, k)$ *so that for all random variables* X *on* $\mathbb{Z}_{\geq 0}$ *that satisfy the following conditions:*

1. *the greatest common divisor of the support of* X *is* 1*;*
2. $\mathbf{P}\left\{X = 0\right\} > \gamma$ *and* $\mathbf{P}\left\{X = k\right\} > \gamma$*;*
3. $\mathbf{E}\left[X^2\right] < \eta$ *and*
4. $\mathbf{E}\left[X^3\right] < \beta$,

it holds that for all $m > M$

$$\sup_{\ell \in \mathbb{Z}}\left|\sqrt{m}\mathbf{P}\left\{\sum_{i=1}^m X_i = \ell\right\} - \phi\left(\frac{\ell - m\mathbf{E}\left[X\right]}{\sqrt{\mathbf{Var}\left\{X\right\}m}}\right)\right| \leq \frac{C}{\sqrt{m}},$$

where $X_1, X_2, \ldots,$ *are* IID *copies of* X, *and* $\phi(t) = e^{-t^2/2}$ *is the standard normal density.*

A.2. Supporting results from the introduction.

PROOF OF LEMMA 1.3. We argue by contradiction. Fix $T > 0$. Without loss of generality, assume that $\pi(\{x\} \times \mathbb{R}_+) = \delta > 0$ for some $x > 0$. We show that this implies that $\pi((x/2, \infty) \times \mathbb{R}_+) > T$, which contradicts the requirement that $\pi((x/2, \infty) \times \mathbb{R}_+) < \infty$ because $T > 0$ was chosen arbitrarily. Fix $0 < \varepsilon < x/4$ small enough that $\delta\lfloor\frac{x}{8\varepsilon}\rfloor > T$ and $\pi(\{x - \varepsilon, x + \varepsilon\} \times \mathbb{R}_+) = 0$. Define $A_0 = (x - \varepsilon, x + \varepsilon)$, so that by [**A3**],

$$r^{4-\eta}\mathbf{P}\left\{\frac{1}{r}\max_{1 \leq i \leq \xi} Y^+_{\xi,i} \in A_0\right\} \to \pi(A_0 \times \mathbb{R}_+) \geq \delta \quad \text{as } r \to \infty.$$

Then, letting $J = \lfloor\frac{x}{8\varepsilon}\rfloor - 1$, for $j \in \{1, \ldots, J\}$, we can find $\theta_j \in (0, 1]$ such that $A_j := \theta_j(x - \varepsilon, x + \varepsilon) \subset (x - (2j + 1)\varepsilon, x - (2j - 1)\varepsilon)$ and $\pi(\{\theta_j(x - \varepsilon), \theta_j(x + \varepsilon)\} \times \mathbb{R}_+) = 0$. By definition, $A_0, \ldots, A_J$ are pairwise disjoint, and by our choice for J, $\cup_{0 \leq j \leq J} A_j \subset (x/2, x + \varepsilon)$, so $\pi((x/2, \infty) \times \mathbb{R}_+) \geq \sum_{0 \leq j \leq J} \pi(A_j \times \mathbb{R}_+)$. Moreover, setting $r = \theta_j s$ in the above limit shows that

$$s^{4-\eta}\mathbf{P}\left\{\frac{1}{s}\max_{1 \leq i \leq \xi} Y^+_{\xi,i} \in \theta_j(x - \epsilon, x + \epsilon)\right\} \to \theta_j^{\eta-4}\pi(A_0 \times \mathbb{R}_+) \geq \delta \text{ as } s \to \infty.$$

But [**A3**] implies that

$$s^{4-\eta}\mathbf{P}\left\{\frac{1}{s}\max_{1 \leq i \leq \xi} Y^+_{\xi,i} \in \theta_j(x - \epsilon, x + \epsilon)\right\} \to \pi(A_j \times \mathbb{R}_+),$$

so $\pi((x/2, \infty) \times \mathbb{R}_+) \geq (J + 1)\delta > T$, which implies the claim. □

PROOF OF PROPOSITION 1.2. To ease notation, we write $n/2$ instead of $\lfloor n/2 \rfloor$ throughout the proof.

First observe that, by assumption, there exist $\varepsilon, \delta > 0$ such that, for $\xi_1, \dots, \xi_n$ IID samples from μ,

$$\limsup_{n\to\infty} \mathbf{P}\left\{ \max_{1\le i\le n} \max_{1\le j\le \xi_i} |Y_{\xi_i,j}| > \delta n^{1/4} \right\} > \varepsilon.$$

By the central limit theorem, we may pick K large enough that

$$\liminf_{n\to\infty} \mathbf{P}\left\{ n/2 - Kn^{1/2} \le \sum_{i=1}^{n/2} \xi_i \le n/2 + Kn^{1/2} \right\} > 1 - \varepsilon/2,$$

so that by a union bound

$$\limsup_{n\to\infty} \mathbf{P}\left\{ \max_{1\le i\le n} \max_{1\le j\le \xi_i} |Y_{\xi_i,j}| > \delta n^{1/4},\ n/2 - Kn^{1/2} \le \sum_{i=1}^{n/2} \xi_i \le n/2 + Kn^{1/2} \right\} > \varepsilon/2.$$

Denote the event inside the probability by $\mathcal{E}_n$. We see that

$$\begin{aligned}
&\mathbf{P}\left\{ \max_{v\in v(\mathrm{T}_n)\setminus \partial \mathrm{T}_n} \max_{j\ge 1} |Y_j^{(v)}| > \delta n^{1/4} \right\} = \mathbf{P}\left\{ \max_{1\le i\le n} \max_{1\le j\le D_i^n} |Y_{D_i^n,j}| > \delta n^{1/4} \right\} \\
&\quad \ge \mathbf{P}\left\{ \max_{1\le i\le n/2} \max_{1\le j\le D_i^n} |Y_{D_i^n,j}| > \delta n^{1/4},\ n/2 - Kn^{1/2} \le \sum_{i=1}^{n/2} D_i^n \le n/2 + Kn^{1/2} \right\} \\
&\quad = \frac{\mathbf{P}\left\{ \mathcal{E}_n \cap \{\sum_{i=1}^n \xi_i = n-1\} \right\}}{\mathbf{P}\left\{ \sum_{i=1}^n \xi_i = n-1 \right\}} \\
&\quad = \frac{\mathbf{E}\left[\mathbf{1}_{[\mathcal{E}_n]} \mathbf{P}\left\{ \sum_{i=n/2+1}^n \xi_i = n - 1 - \sum_{i=1}^{n/2} \xi_i \mid \xi_1, \dots, \xi_{n/2}, Y_{\xi_1}, \dots, Y_{\xi_n/2} \right\} \right]}{\mathbf{P}\left\{ \sum_{i=1}^n \xi_i = n-1 \right\}} \\
&\quad \ge \mathbf{P}\left\{ \mathcal{E}_n \right\} \frac{\min_{n/2-1-Kn^{1/2} \le m \le n/2-1+Kn^{1/2}} \mathbf{P}\left\{ \sum_{i=n/2+1}^n \xi_i = m \right\}}{\mathbf{P}\left\{ \sum_{i=1}^n \xi_i = n-1 \right\}}.
\end{aligned}$$

By the local central limit theorem, there exist constants $c, C > 0$ such that

$$\liminf_{n\to\infty} n^{1/2} \min_{n/2-1-Kn^{1/2} \le m \le n/2-1+Kn^{1/2}} \mathbf{P}\left\{ \sum_{i=n/2+1}^n \xi_i = m \right\} > c$$

$$\text{and} \quad \limsup_{n\to\infty} n^{1/2} \mathbf{P}\left\{ \sum_{i=1}^n \xi_i = n-1 \right\} < C.$$

Thus, as claimed

$$\limsup_{n\to\infty} \mathbf{P}\left\{ \max_{1\le i\le n} \max_{1\le j\le D_i^n} |Y_{D_i^n,j}| > \delta n^{1/4} \right\} \ge \frac{\varepsilon c}{2C} > 0. \qquad \square$$

A.3. Measure change. For $n \ge 1$ let $\mathcal{S}_n$ denote the set of permutations of $[n]$. For $(k_1, \dots, k_n) \in \mathbb{N}^n$, let $\Sigma = \Sigma_{(k_1,\dots,k_n)}$ be the random permutation of $[n]$ with law given by

$$\mathbf{P}\left\{ \Sigma = \sigma \right\} = \prod_{i=1}^n \frac{k_{\sigma(i)}}{\sum_{j=i}^n k_{\sigma(j)}}, \qquad \text{for } \sigma \in \mathcal{S}_n.$$

We call $(k_{\Sigma(1)}, \dots k_{\Sigma(n)})$ the *size-biased random re-ordering* of $(k_1, \dots, k_n)$. It will be convenient to extend this definition to vectors $(k_1, \dots, k_n)$ that contain 0-valued entries. We start with a size-biased random re-ordering of the non-zero entries of $(k_1, \dots, k_n)$ and then append to this the correct number of zeroes. Formally, if $(k_1, \dots, k_n) \in \mathbb{Z}_{\geq 0}^n$ has $N \geq 0$ non-zero entries, let $\Sigma_{(k_1,\dots,k_n)}$ be the random permutation of $[n]$ with

$$\mathbf{P}\left\{\Sigma_{(k_1,\dots,k_n)} = \sigma\right\} = \frac{1}{(n-N)!} \prod_{i=1}^{N} \frac{k_{\sigma(i)}}{\sum_{j=1}^{N} k_{\sigma(j)}}, \tag{A.1}$$

for $\sigma \in \mathcal{S}_n$, and still refer to $(k_{\Sigma(1)}, \dots k_{\Sigma(n)})$ as the size-biased random re-ordering of $(k_1, \dots, k_n)$.

For a permutation $\sigma \in \mathcal{S}_n$ and $r \in \{0, 1, \dots, n\}$ define

$$\tau_r(\sigma) = \begin{cases} \min\{j \in [n] : \sigma(j) \in [r]\} & \text{if } r \in [n], \\ n+1 & \text{if } r = 0\,. \end{cases}$$

As discussed in Section 4, the proof of Theorem 1.1 relies on establishing a change of measure, (4.5), which relates the size-biased random re-ordering of the positive entries of the degree sequence of T_n, and IID samples from the offspring distribution. The proofs of Theorems 1.4 and 1.5 rely on establishing a similar change of measure, which is a generalisation of (4.5) to the situation where instead of an IID sequence, the first r elements are non-zero and are fixed in advance; the whole sequence is conditioned to have sum $n-1$; and we consider the first m elements of the size-biased random reordering of the sequence. Specifically, let $m, n, r, s \in \mathbb{Z}_{\geq 0}$ with $m, r, s < n$, and μ be a distribution on $\mathbb{Z}_{\geq 0}$. For $k_1, \dots, k_m \in \mathbb{N}$, we define

$$\begin{aligned} \Theta_\mu(k_1, \dots, k_m) &= \Theta_\mu^{n,r,s}(k_1, \dots, k_m) \\ &= \frac{\mathbf{P}\left\{X_{m+1} + \dots + X_{n-r} = n - 1 - s - \sum_{i=1}^{m} k_i\right\}}{\mathbf{P}\left\{X_1 + \dots + X_{n-r} = n - 1 - s\right\}} \cdot \left(\mathbf{E}\left[X_1\right]\right)^m \cdot \prod_{i=1}^{m} \frac{n - r - i + 1}{n - 1 - \sum_{j=1}^{i-1} k_j} \end{aligned} \tag{A.2}$$

if $k_1 + \dots + k_m \leq n - 1 - s$, and otherwise $\Theta_\mu(k_1, \dots, k_m) = 0$, where $(X_i, i \geq 1)$ are IID random variables with distribution μ. We note that when $r = s = 0$, and μ is a critical offspring distribution, we recover (4.5).

PROPOSITION A.4. *Fix $n, r, s \in \mathbb{Z}_{\geq 0}$ with $r, s < n$, and $d_1, \dots, d_r \in \mathbb{N}$ with $\sum_{i=1}^{r} d_i = s$. Let μ be a distribution on $\mathbb{Z}_{\geq 0}$ and $(X_i, i \geq 1)$ be* IID *random variables with distribution μ. Further, let*

$$N = N_{n,r} = |\{i \in \{r+1, \dots, n\} : X_i > 0\}|.$$

Let $\vec{Z} = (Z_1, \dots, Z_n) = (d_1, \dots, d_r, X_{r+1}, \dots, X_n)$, and conditionally given $\vec{Z}$, let $\Sigma = \Sigma_{\vec{Z}}$ be given by (A.1). Finally, let $(\overline{X}_i, i \in [n])$ be IID *samples from the size-biased distribution of X_1. Suppose that $\mathbf{E}\left[X_1\right] < \infty$. Then for any $m \in [n-r]$ and any function $f : \mathbb{N}^m \to \mathbb{R}$, if $\mathbf{P}\left\{X_{r+1} + \ldots + X_n = n - 1 - s)\right\} > 0$, then*

$$\begin{aligned} &\mathbf{E}\left[f\left(Z_{\Sigma(1)}, \dots, Z_{\Sigma(m)}\right) \mathbf{1}_{[N \geq m]} \mathbf{1}_{[\tau_r(\Sigma) > m]} \,\middle|\, X_{r+1} + \ldots + X_n = n - 1 - s \right] \\ &\qquad = \mathbf{E}\left[f(\overline{X}_1, \dots, \overline{X}_m) \Theta_\mu(\overline{X}_1, \dots, \overline{X}_m) \right], \end{aligned} \tag{A.3}$$

where $\Theta_\mu(\overline{X}_1, \dots \overline{X}_m) = \Theta_\mu^{n,r,s}(\overline{X}_1, \dots \overline{X}_m)$ is as in (A.2).

We observe that when $r \neq 0$, $\tau_r(\Sigma) > m$ implies that $N \geq m$ because all positive entries occur before zero-valued entries in the size-biased random reordering. However, when $r = 0$, the former event is vacuously true for all $m \in [n]$, but we still enforce that $N \geq m$ in (A.3). It follows that Proposition 4.3 is the special case when $r = 0$, $s = 0$ and $X_1, X_2, \ldots$ are IID samples from the offspring distribution μ.

PROOF. In this proof, for $n \geq 1$, and $r \geq 1$, we let

$$[n]_r = \{(n_1, \ldots, n_r) \in \{1, \ldots, n\}^r \ : \ n_i \neq n_j \text{ for all } i \neq j\}.$$

Furthermore, for a set A we write A_r for the set of ordered sequences $(s_1, \ldots, s_r)$ of r distinct elements of A. We also let $\mu_i = \mathbf{P}\{X_1 = i\}$ for $i \in \mathbb{Z}_{\geq 0}$.

We first prove the proposition assuming that $\mu_0 = 0$; we will later generalise this by conditioning on the number of non-zero entries of $\vec{Z}$ and sampling a size-biased re-ordering of only these entries. When $\mu_0 = 0$, we have $\mathbf{P}\{N = n\} = 1$, so the indicator $\mathbf{1}_{[N \geq m]}$ in (A.3) equals 1 and may be ignored.

For $\sigma \in S_n$, we write $\vec{Z}_\sigma = (Z_{\sigma(1)}, \ldots, Z_{\sigma(n)})$ and $\sigma^{-1}[r] = (\sigma^{-1}(1), \ldots, \sigma^{-1}(r))$. Observe that for $m \in [n - r]$, we have the equality of events

$$\{\tau_r(\Sigma) > m\} = \left\{\Sigma^{-1}[r] \in ([n] \backslash [m])_r\right\} .$$

It is thus useful to determine the law of $(\vec{Z}_\Sigma, \Sigma^{-1}[r])$. Note that for any $\vec{k} = (k_1, \ldots, k_n) \in \mathbb{N}^n$ and $\vec{j} = (j_1, \ldots, j_r) \in [n]_r$, if $(\vec{k}, \vec{j})$ is in the support of $(\vec{Z}_\Sigma, \Sigma^{-1}[r])$ then $k_{j_i} = d_i$ for each $i \in [r]$. For such $(\vec{k}, \vec{j})$,

$$\begin{aligned} \mathbf{P}\left\{\vec{Z}_\Sigma = \vec{k}, \Sigma^{-1}[r] = \vec{j}\right\} &= \sum_{\sigma \in \mathcal{S}_n : \sigma^{-1}[r] = \vec{j}} \mathbf{P}\left\{\vec{Z}_\sigma = \vec{k}, \Sigma = \sigma\right\} \\ &= \prod_{i=1}^{n} \frac{k_i}{\sum_{j=i}^{n} k_j} \cdot \sum_{\sigma \in \mathcal{S}_n : \sigma^{-1}[r] = \vec{j}} \mathbf{P}\left\{\vec{Z}_\sigma = \vec{k}\right\}. \end{aligned} \tag{A.4}$$

Since we fixed $\sigma^{-1}[r]$, the sum (A.4) ranges over exactly $(n - r)!$ elements of $\mathcal{S}_n$ and each term of the sum is equal to

$$\prod_{j \in [n] \backslash \{j_1, \ldots, j_r\}} \mu_{k_j}.$$

Hence, for any $\vec{k} \in \mathbb{N}^n$ and $\vec{j} \in [n]_r$,

$$\mathbf{P}\left\{\vec{Z}_\Sigma = \vec{k}, \Sigma^{-1}[r] = \vec{j}\right\} = (n-r)! \left(\prod_{i=1}^{r} \mathbf{1}_{[k_{j_i} = d_i]}\right) \left(\prod_{j \in [n] \backslash \{j_1, \ldots, j_r\}} \mu_{k_j}\right) \left(\prod_{i=1}^{n} \frac{k_i}{\sum_{j=i}^{n} k_j}\right). \tag{A.5}$$

Now, fix $m \in [n - r]$ and $k_1, \ldots, k_m \in \mathbb{N}$. Note that it suffices to prove (A.3) when $f : \mathbb{N}^m \to \mathbb{R}$ has the form

$$f(z_1, \ldots, z_m) = \prod_{i=1}^{m} \mathbf{1}_{[z_i = k_i]}, \tag{A.6}$$

so we now restrict our attention to this case. Since

$$\sum_{i \in [n]} Z_{\Sigma(i)} = \sum_{i=1}^{r} d_i + \sum_{i=r+1}^{n} X_i = s + \sum_{i=r+1}^{n} X_i \,,$$

for any $k_1, \dots, k_m \in \mathbb{N}$, by summing over the possible values of $Z_{\Sigma(m+1)}, \dots, Z_{\Sigma(n)}$, we can use (A.5) to find that

(A.7)

$$\mathbf{P}\left\{(Z_{\Sigma(1)}, \dots, Z_{\Sigma(m)}) = (k_1, \dots, k_m), \tau_r(\Sigma) > m, \sum_{i=r+1}^{n} X_i = n-1-s\right\}$$

$$= \sum_{(k_{m+1},\dots,k_n)\in\mathbb{N}^{n-m}} \sum_{\vec{j}\in([n]\setminus[m])_r} \mathbf{1}_{[\sum_{i=1}^n k_i = n-1]} \mathbf{P}\left\{\vec{Z}_\Sigma = (k_1, \dots, k_n), \Sigma^{-1}[r] = \vec{j}\right\}$$

$$= (n-r)! \left(\prod_{i=1}^{m} \mu_{k_i}\right) \left(\prod_{i=1}^{m} \frac{k_i}{n-1-\sum_{j=1}^{i-1} k_i}\right)$$

$$\cdot \sum_{\substack{(k_{m+1},\dots,k_n)\in\mathbb{N}^{n-m} \\ \vec{j}\in([n]\setminus[m])_r}} \mathbf{1}_{[\sum_{i=1}^n k_i = n-1]} \left(\prod_{i=1}^{r} \mathbf{1}_{[k_{j_i}=d_i]}\right) \left(\prod_{i\in([n]\setminus[m])\setminus\{j_1,\dots,j_r\}} \mu_{k_i}\right) \left(\prod_{i=m+1}^{n} \frac{k_i}{\sum_{j=i}^n k_j}\right).$$

Using that $k\mu_k = \mathbf{P}\left\{\overline{X}_1 = k\right\} \mathbf{E}[X_1]$ for all $k \in \mathbb{N}$, writing $n' = n-m$, and re-indexing the above sum, this yields that (A.7) is equal to

$$\mathbf{P}\left\{(\overline{X}_1, \dots, \overline{X}_m) = (k_1, \dots, k_m)\right\} \mathbf{E}[X_1]^m \left(\prod_{i=1}^{m} \frac{n-r-i+1}{n-1-\sum_{j=1}^{i-1} k_j}\right) (n'-r)!$$

$$\cdot \sum_{\substack{(k'_1,\dots,k'_{n'})\in\mathbb{N}^{n'} \\ \vec{j}\in[n']_r}} \mathbf{1}_{[\sum_{i=1}^{n'} k'_i = n-1-\sum_{i=1}^m k_i]} \left(\prod_{i=1}^{r} \mathbf{1}_{[k'_{j_i}=d_i]}\right) \left(\prod_{i\in[n']\setminus\{j_1,\dots,j_r\}} \mu_{k'_i}\right) \cdot \prod_{i\in[n']} \frac{k'_i}{\sum_{j=i}^{n'} k'_j}.$$

Now, define $\vec{Z}' = (d_1, \dots, d_r, X_{r+1}, \dots, X_{n-m})$ and, conditionally given $\vec{Z}'$, let $\Sigma' = \Sigma_{\vec{Z}'}$ be given by (A.1). Applying (A.5) to $\vec{Z}'$ and Σ', we thus find that (A.7) equals

$$\mathbf{P}\left\{(\overline{X}_1, \dots, \overline{X}_m) = (k_1, \dots, k_m)\right\} \prod_{i=1}^{m} \left(\frac{n-1-i+1}{n-1-\sum_{j=1}^{i-1} k_j} \mathbf{E}[X_1]\right)$$

$$\cdot \sum_{\substack{(k'_1,\dots,k'_{n'})\in\mathbb{N}^{n'} \\ \vec{j}\in[n']_r}} \mathbf{1}_{[\sum_{i=1}^{n'} k'_i = n-1-\sum_{i=1}^m k_i]} \mathbf{P}\left\{\vec{Z}'_{\Sigma'} = \vec{k}, (\Sigma')^{-1}[r] = \vec{j}\right\}$$

$$= \mathbf{P}\left\{(\overline{X}_1, \dots, \overline{X}_m) = (k_1, \dots, k_m)\right\} \prod_{i=1}^{m} \left(\frac{n-1-i+1}{n-1-\sum_{j=1}^{i-1} k_j} \mathbf{E}[X_1]\right)$$

$$\cdot \mathbf{P}\left\{\sum_{i=1}^{n'} Z'_{\Sigma'(i)} = n-1-\sum_{i=1}^{m} k_i\right\}.$$

Finally, since the sum of the entries of $\vec{Z}'_{\Sigma'}$ is unaffected by the random reordering and is the same as $s + \sum_{i=1}^{n'-r+m} X_i = s + \sum_{i=1}^{n-(r+m)} X_i$, we deduce that (A.7) equals

$$\mathbf{P}\left\{(\overline{X}_1, \dots, \overline{X}_m) = (k_1, \dots, k_m)\right\} \prod_{i=1}^{m} \left(\frac{n-1-i+1}{n-1-\sum_{j=1}^{i-1} k_j} \mathbf{E}[X_1]\right)$$

$$\cdot \mathbf{P}\left\{\sum_{i=1}^{n-(r+m)} X_i = n-1-s-\sum_{i=1}^{m} k_i\right\}.$$

Dividing the above expression by $\mathbf{P}\left\{\sum_{i=1}^{n-r} X_i = n-1-s\right\}$ yields the statement when $\mu_0 = 0$, in the special case that f has the form given in (A.6), and thus for general f.

For the general case with $\mu_0 > 0$, we let $p = 1-\mu_0$. Further, we let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be IID copies of X_1 conditioned to be positive. Notice that $\mathbf{E}[\mathbf{X}_1] = p^{-1}\mathbf{E}[X_1]$ and that the size-biased distributions of $\mathbf{X}_1$ and of X_1 are identical. We let $\widehat{\mathbf{X}}_1, \widehat{\mathbf{X}}_2, \ldots$ denote IID samples from the size-biased distribution of $\mathbf{X}_1$. Finally, fix $m' \geq 0$ and define

$$\vec{\mathbf{Z}}' = (\mathbf{Z}'_1, \ldots, \mathbf{Z}'_{m'+r}) = (d_1, \ldots, d_r, \mathbf{X}_{r+1}, \ldots, \mathbf{X}_{r+m'}),$$

and conditionally given $\vec{\mathbf{Z}}'$, let $\Sigma' = \Sigma_{\vec{\mathbf{Z}}'}$ be given by (A.1).

Now fix $m \in [n-r]$ with $m \leq m'$. For $k_1, \ldots, k_m \in \mathbb{N}$, we have that

$$\mathbf{P}\left\{(Z_{\Sigma(1)}, \ldots, Z_{\Sigma(m)}) = (k_1, \ldots, k_m), \tau_r(\Sigma) > m, \sum_{i=r+1}^{n} X_i = n-1-s \,\middle|\, N = m'\right\}$$

(A.8)

$$= \mathbf{P}\left\{(\mathbf{Z}'_{\Sigma'(1)}, \ldots, \mathbf{Z}'_{\Sigma'(m)}) = (k_1, \ldots, k_m),\ \tau_r(\Sigma') > m, \sum_{i=r+1}^{m'} \mathbf{X}_i = n-1-s\right\}.$$

By the proof of the case where $\mu_0 = 0$, if $k_1 + \cdots + k_m \leq n-1-s$ this is equal to

$$\mathbf{P}\left\{(\overline{\mathbf{X}}_1, \ldots, \overline{\mathbf{X}}_m) = (k_1, \ldots, k_m)\right\}$$
$$\cdot \mathbf{P}\left\{\sum_{i=m+1}^{m'} \mathbf{X}_i = n-1-s-\sum_{i=1}^{m} k_i\right\} \prod_{i=1}^{m} \left(\frac{m'-i+1}{n-1-\sum_{j=1}^{i-1} k_j} \mathbf{E}[\mathbf{X}_1]\right),$$

and otherwise is equal to 0. For the remainder of the proof we may thus assume that $k_1 + \cdots + k_m \leq n-1-s$. Since $\overline{\mathbf{X}}_1 \stackrel{\mathrm{d}}{=} \overline{X}_1$ and $\mathbf{E}[\mathbf{X}_1] = p^{-1}\mathbf{E}[X_1]$ this is in turn equal to

$$\mathbf{P}\left\{(\overline{X}_1, \ldots, \overline{X}_m) = (k_1, \ldots, k_m)\right\}$$

(A.9)
$$\cdot \frac{1}{p^m} \mathbf{P}\left\{\sum_{i=m+1}^{m'} \mathbf{X}_i = n-1-s-\sum_{i=1}^{m} k_i\right\} \prod_{i=1}^{m} \left(\frac{m'-i+1}{n-1-\sum_{j=1}^{i-1} k_j} \mathbf{E}[X_1]\right).$$

It then follows from (A.8) and (A.9) that

$$\mathbf{P}\left\{(Z_{\Sigma(1)}, \ldots, Z_{\Sigma(m)}) = (k_1, \ldots, k_m), N \geq m, \tau_r(\Sigma) > m, \sum_{i=r+1}^{n} X_i = n-1-s\right\}$$
$$= \mathbf{P}\left\{(\overline{X}_1, \ldots, \overline{X}_m) = (k_1, \ldots, k_m)\right\}$$

(A.10)
$$\cdot \sum_{m'=m}^{n-r} \frac{\mathbf{P}\{N = m'\}}{p^m} \mathbf{P}\left\{\sum_{i=m+1}^{m'} \mathbf{X}_i = n-1-s-\sum_{i=1}^{m} k_i\right\} \prod_{i=1}^{m} \left(\frac{m'-i+1}{n-1-\sum_{j=1}^{i-1} k_j} \mathbf{E}[X_1]\right).$$

Notice now that $N \stackrel{\mathrm{d}}{=} \mathrm{Binomial}(n-r, p)$. So using the change of variable $\ell = m'-m$ and letting M be a $\mathrm{Binomial}(n-(r+m), p)$, by routine algebra we obtain that (A.10) equals

$$\sum_{\ell=0}^{n-(r+m)} \mathbf{P}\{M = \ell\} \mathbf{P}\left\{\sum_{i=m'-\ell+1}^{m+\ell} \mathbf{X}_i = n-1-s-\sum_{i=1}^{m} k_i\right\} \cdot \prod_{i=1}^{m} \left(\frac{n-r-i+1}{n-1-\sum_{j=1}^{i-1} k_i} \mathbf{E}[X_1]\right)$$

$$= \sum_{\ell=0}^{n-(r+m)} \mathbf{P}\{M=\ell\}\,\mathbf{P}\left\{\sum_{i=1}^{\ell}\mathbf{X}_i = n-1-s-\sum_{i=1}^{m}k_i\right\}\cdot\prod_{i=1}^{m}\left(\frac{n-r-i+1}{n-1-\sum_{j=1}^{i-1}k_i}\mathbf{E}[X_1]\right)$$

$$= \mathbf{P}\left\{\sum_{i=1}^{M}\mathbf{X}_i = n-1-s-\sum_{i=1}^{m}k_i\right\}\prod_{i=1}^{m}\left(\frac{n-r-i+1}{n-1-\sum_{j=1}^{i-1}k_i}\mathbf{E}[X_1]\right).$$

Since $\sum_{i=1}^{M}\mathbf{X}_i \overset{\mathrm{d}}{=} \sum_{i=1}^{n-(r+m)}X_i$, dividing the above expression (which is equal to (A.10)) by $\mathbf{P}\left\{\sum_{i=1}^{n-r}X_i = n-1-s\right\}$ yields the result for the special case that f has the form given in (A.6), and thus for general f. □

The next proposition gives conditions under which the change of measure $\Theta_{\mu}^{n,r,s}$ appearing in (A.2) is asymptotically unimportant in the specific case when $m=\Theta(\sqrt{n})$ and $(X_i, i\geq 1)$ are IID samples from the offspring distribution μ conditioned to yield a displacement vector such that $\max_{1\leq j\leq X_i}|Y_{X_i,j}|\leq \gamma n^{1/(4-\eta)}$. This then allows us to use the measure change in the proofs of Theorems 1.4 and 1.5.

LEMMA A.5. *Let μ be a critical offspring distribution with variance $\sigma^2\in(0,\infty)$, and let $\nu=(\nu_k)_{k\geq 1}$ be such that [**A1**] holds and [**A3**] holds for a given measure π with $\eta\in[0,2)$. Fix $\gamma>0$. Let ξ denote a random variable with distribution μ, and for $n\geq 1$ let ξ^n be distributed as ξ, conditioned to not yield a displacement vector with $\max_{1\leq i\leq\xi^n}|Y_{\xi^n,i}|>\gamma n^{1/(4-\eta)}$. Further, let μ^n denote the distribution of ξ^n, and let $\bar{\xi}_1^n,\bar{\xi}_2^n,\ldots$ be* IID *samples from the size-biased law of ξ^n.*

Finally, fix $\varepsilon\in(0,1/6)$ and let $(r_n)_{n\geq 1}$ and $(s_n)_{n\geq 1}$ be sequences such that for all $n\geq 1$, $r_n<n^{\varepsilon}$, $s_n<n^{1/3+\varepsilon}$ and $n-1-s_n$ is in the support of $\sum_{i=r_n+1}^{n}\xi_i^n$.

Suppose that $m=\Theta(\sqrt{n})$. Then as $n\to\infty$,

$$\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n,\ldots,\bar{\xi}_m^n)\xrightarrow{\mathrm{p}}1, \tag{A.11}$$

and $(\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n,\ldots,\bar{\xi}_m^n))_{n\geq 1}$ is a uniformly integrable sequence of random variables.

The proof of Lemma A.5 is very similar to that of Lemma 4.4. However, in this case instead of the standard local central limit theorem, we will require a quantitative local central limit theorem in order to get uniform estimates on local probabilities for the family of random variables $\{\xi^n, n\geq 1\}$.

LEMMA A.6. *Let μ be a critical offspring distribution with variance $\sigma^2\in(0,\infty)$, and let $\nu=(\nu_k)_{k\geq 1}$ be such that [**A1**] holds and [**A3**] holds for a given measure π with $\eta\in[0,2)$. Let $\gamma>0$. Further, let ξ denote a random variable with distribution μ and for $n\geq 1$ let $\xi^n=(\xi_i^n, i\geq 1)$ be* IID *copies of ξ each conditioned to satisfy $\{\max_{1\leq i\leq\xi_j^n}|Y_{\xi_j^n,i}|\leq\gamma n^{1/(4-\eta)}\}$. Then there exist $C,N>0$ and M such that for all $m,n>N$,*

$$\sup_{k\in\mathbb{Z}}\left|\sqrt{m}\mathbf{P}\left\{\sum_{i=1}^{m}\xi_i^n=k\right\}-\phi\left(\frac{k-m\mathbf{E}[\xi_1^n]}{\sqrt{\mathbf{Var}\{\xi_1^n\}\,m}}\right)\right|\leq\frac{C}{\sqrt{m}},$$

where $\phi(t)=e^{-t^2/2}$ is the standard normal density.

This lemma is immediate from Lemma A.3 as soon as we show that the family $\{\xi^n,\ n\geq 1\}$ satisfies the conditions of that lemma. This is verified in Lemmas A.7 and A.8.

LEMMA A.7. *For all n sufficiently large, the support of ξ^n has greatest common divisor* 1.

PROOF. By assumption, the support of ξ has greatest common divisor 1, so we can find an M such that the greatest common divisor of the support of ξ restricted to $\{0,\dots,M\}$ is 1. Since $\gamma n^{1/(4-\eta)} > M$ for n sufficiently large, the result follows. □

LEMMA A.8. *As $n \to \infty$,*

$$\mathbf{E}\left[(\xi^n)^j\right] \to \mathbf{E}\left[\xi^j\right] \quad \text{for } j = 1,2,3 \tag{A.12}$$

and

$$|\mathbf{E}\left[\xi^n\right] - 1| = O(n^{-2/3}). \tag{A.13}$$

PROOF. For $j \in \{1,2,3\}$ we have

$$\begin{aligned}\mathbf{E}\left[(\xi^n)^j\right] &= \sum_{k=1}^{\infty} k^j \mathbf{P}\left\{\xi^n = k\right\} \\ &\le \sum_{k=1}^{\infty} k^j \frac{\mathbf{P}\left\{\xi = k\right\}}{\mathbf{P}\left\{\max_{1\le i\le \xi} |Y_{\xi,i}| \le \gamma n^{1/(4-\eta)}\right\}} = \left(1 + O\left(\frac{1}{n}\right)\right)\mathbf{E}\left[\xi^j\right],\end{aligned}$$

where the final equality follows by assumption [**A3**]. By the bounded convergence theorem, as $n \to \infty$,

$$\mathbf{E}\left[(\xi^n)^j\right] \ge \mathbf{E}\left[\xi^j\right] - \mathbf{E}\left[\xi^j \mathbf{1}_{[\max_{1\le i\le \xi} |Y_{\xi,i}| > \gamma n^{1/(4-\eta)}]}\right] \to \mathbf{E}\left[\xi^j\right],$$

where we have used assumption [**A3**] again. (A.12) follows.

To get the more precise lower bound for $j = 1$ in (A.13), observe that

$$\mathbf{E}\left[\xi \mathbf{1}_{[\max_{1\le i\le \xi} |Y_{\xi,i}| > \gamma n^{1/(4-\eta)}]}\right] \le n^{1/3}\mathbf{P}\left\{\max_{1\le i\le \xi} |Y_{\xi,i}| > \gamma n^{1/(4-\eta)}\right\} + \mathbf{E}\left[\xi \mathbf{1}_{[\xi > n^{1/3}]}\right].$$

The first term on the right-hand side of this inequality is $O(n^{-2/3})$ by [**A3**]. Also, $\mathbf{E}\left[\xi^3\right] < \infty$ and so the second term is also $O(n^{-2/3})$, thus establishing (A.13). □

The last tool that we need to prove Lemma A.5 is an upper bound on the total variation distance between $\bar{\xi}_1^n$ and $\bar{\xi}$ where $\bar{\xi}$, a sample from the size-biased law of ξ.

LEMMA A.9. *Let X be a random variable taking values in $\mathbb{N}$ such that $\mathbf{E}[X] \ge 0$ and $\mathbf{E}\left[X^3\right] < \infty$. Let $(\mathcal{E}_n)_{n\ge 1}$ be a sequence of events with $\mathbf{P}\{\mathcal{E}_n\} = 1 - O(1/n)$. Let X_n be distributed as X conditional on $\mathcal{E}_n$. Let $\overline{X}_n$ have the size-biased law of X_n and let $\overline{X}$ have the size-biased law of X. Then,*

$$d_{\mathrm{TV}}(\overline{X}_n, X) = \frac{1}{2}\sum_{k=1}^{\infty} \left|\mathbf{P}\left\{\overline{X}_n = k\right\} - \mathbf{P}\left\{\overline{X} = k\right\}\right| = O(n^{-2/3}).$$

PROOF. By definition,

$$\mathbf{P}\left\{\overline{X}_n = k\right\} = \frac{k\mathbf{P}\left\{X = k, \mathcal{E}_n\right\}}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]}, \text{ and } \mathbf{P}\left\{\overline{X} = k\right\} = \frac{k\mathbf{P}\left\{X = k\right\}}{\mathbf{E}\left[X\right]}. \tag{A.14}$$

Since $\mathbf{E}\left[X^3\right] < \infty$ we have that $\mathbf{P}\left\{X > n^{1/3}\right\} = o(n^{-1})$ as $n \to \infty$ and so Hölder's inequality yields that

$$\mathbf{E}\left[X\mathbf{1}_{[X>n^{1/3}]}\right] \le \mathbf{E}\left[X^3\right]^{1/3}\mathbf{P}\left\{X > n^{1/3}\right\}^{2/3} = o(n^{-2/3}).$$

Next,

$$\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n^c]}\right] \le n^{1/3}\mathbf{P}\left\{\mathcal{E}_n^c\right\} + \mathbf{E}\left[X\mathbf{1}_{[X>n^{1/3}]}\right] = O(n^{-2/3}),$$

so that $\mathbf{E}\left[X\mathbf{1}_{\mathcal{E}_n}\right] = \mathbf{E}\left[X\right] + O(n^{-2/3})$ and the difference between the denominators in (A.14) is $O(n^{-2/3})$. It follows that

$$\begin{aligned}
&\sum_{k=1}^{\infty}\left|\mathbf{P}\left\{\overline{X}_n = k\right\} - \mathbf{P}\left\{\overline{X} = k\right\}\right| \\
&\quad= \sum_{k=1}^{\infty}\left|\frac{k\mathbf{P}\left\{X=k,\ \mathcal{E}_n\right\}}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n^c]}\right]} - \frac{k\mathbf{P}\left\{X=k\right\}}{\mathbf{E}\left[X\right]}\right| \\
&\quad\le \frac{1}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]}\left(\sum_{k=1}^{\infty} k\mathbf{P}\left\{X=k, \mathcal{E}_n^c\right\}\right) + O(n^{-2/3}) \\
&\quad\le \frac{1}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]}\left(\sum_{k=1}^{n^{1/3}} k\mathbf{P}\left\{X=k, \mathcal{E}_n^c\right\} + \mathbf{E}\left[X\mathbf{1}_{[X>n^{1/3}]}\right]\right) + O(n^{-2/3}) \\
&\quad\le \frac{n^{1/3}\mathbf{P}\left\{\mathcal{E}_n^c\right\}}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]} + \frac{\mathbf{E}\left[X\mathbf{1}_{[X>n^{1/3}]}\right]}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]} + O(n^{-2/3}).
\end{aligned}$$

The first term in the last line is also is $O(n^{-2/3})$ since $\mathbf{P}\left\{\mathcal{E}_n^c\right\} = O(1/n)$. □

This lemma has the following corollary.

COROLLARY A.10. *Let μ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$, and let $\nu = (\nu_k)_{k\ge1}$ be such that [**A1**] holds and [**A3**] holds for a given measure π with $\eta \in [0,2)$. Fix $\gamma > 0$. Let ξ denote a random variable with distribution μ and let $\bar{\xi}_1, \bar{\xi}_2, \dots$ be* IID *samples from the size-biased law of ξ. For $n \ge 1$ let ξ^n be distributed as ξ, conditioned to not yield a displacement vector with $\max_{1\le i\le \xi^n}|Y_{\xi^n,i}| > \gamma n^{1/(4-\eta)}$. Further, let μ^n denote the distribution of ξ^n, and let $\bar{\xi}_1^n, \bar{\xi}_2^n, \dots$ be* IID *samples from the size-biased law of ξ^n. Then for $m = \Theta(\sqrt{n})$,*

$$d_{\mathrm{TV}}((\bar{\xi}_1^n, \dots, \bar{\xi}_m^n), (\bar{\xi}_1, \dots, \bar{\xi}_m)) = O(n^{-1/6}).$$

PROOF. By [**A3**], $\bar{\xi}_1^n$ is obtained from $\bar{\xi}_1$ by conditioning on an event which occurs with probability $1 - O(1/n)$. Therefore, by Lemma A.9, the total variation distance between $\bar{\xi}_1^n$ and $\bar{\xi}_1$ is $O(n^{-2/3})$. Since $m = \Theta(\sqrt{n})$, the conclusion follows. □

We now prove Lemma A.5. Since the proof is very similar to that of Lemma 4.4 we will be brief.

PROOF OF LEMMA A.5. As in the proof of Lemma 4.4, we may assume that there exists $t > 0$ such that $m/\sqrt{n} \to t$ as $n \to \infty$.

Suppose that $k_1, \dots, k_m \in \mathbb{Z}_{\ge0}$. Then by almost identical techniques to those used to prove (4.6) (replacing the local central limit theorem by Lemma A.6), we obtain that

$$\frac{\mathbf{P}\left\{\sum_{i=(r_n+m)+1}^{n}\xi_i^n = n - 1 - s_n - \sum_{i=1}^{m} k_i\right\}}{\mathbf{P}\left\{\sum_{i=r_n+1}^{n}\xi_i^n = n - 1 - s_n\right\}}$$

$$= \exp\left(-\left(\left(\frac{1+s_n-r_n+m\sigma^2+\sum_{i=1}^m(k_i-(1+\sigma^2))}{\sqrt{2\sigma^2(n-(r_n+m))}}\right)^2+o(1)\right)\right)+o(1). \tag{A.15}$$

Recall that, for $i\in[m]$, $\bar{\xi}_i$ is sample from the size-biased distribution of ξ. We claim that instead of substituting $\bar{\xi}^n_1,\dots,\bar{\xi}^n_m$ in the place of $k_1,\dots,k_m$ we can substitute $\bar{\xi}_1,\dots,\bar{\xi}_m$. Indeed, by Corollary A.10, the total variation distance between $\bar{\xi}^n_1,\dots,\bar{\xi}^n_m$ and $\bar{\xi}_1,\dots,\bar{\xi}_m$ tends to 0 as $n\to\infty$. Therefore, by (4.8), we obtain that (A.15) tends to $\exp(-(t^2\sigma^2)/2)$ in probability as $n\to\infty$. This convergence is analogous to (4.8) in the proof of Lemma 4.4.

It remains to establish an analogue of (4.9), i.e.,

$$\prod_{i=1}^m\left(\frac{n-r_n-i+1}{n-1-\sum_{j=1}^{i-1}\bar{\xi}^n_j}\mathbf{E}\left[\xi^n\right]\right)=\mathbf{E}\left[\xi^n\right]^m\prod_{i=1}^m\left(\frac{n-r_n-i+1}{n-1-\sum_{j=1}^{i-1}\bar{\xi}_j}\right)\xrightarrow{\mathrm{p}}\exp\left(\frac{t^2\sigma^2}{2}\right), \tag{A.16}$$

as $n\to\infty$. By Lemma A.8,

$$\mathbf{E}\left[\xi^n\right]=\mathbf{E}\left[\xi\right]+O(n^{-2/3})=1+O(n^{-2/3}),$$

and so, since $m=(1+o(1))t\sqrt{n}$, we obtain that $\mathbf{E}\left[\xi^n\right]^m=1+o(1)$. Therefore (A.16) follows from (4.9).

We now prove uniform integrability of the family $(\Theta^{n,r_n,s_n}_{\mu^n}(\bar{\xi}^n,\dots,\bar{\xi}^n_m))_{n\geq1}$. Again, by the generalised Scheffé lemma [19, Theorem 5.12], since $\Theta^{n,r_n,s_n}_{\mu^n}(\bar{\xi}^n_1,\dots,\bar{\xi}^n_m)\xrightarrow{\mathrm{p}}1$ it suffices to show that $\mathbf{E}\left[\Theta^{n,r_n,s_n}_{\mu^n}(\bar{\xi}^n_1,\dots,\bar{\xi}^n_m)\right]\to1$ as $n\to\infty$. By Proposition A.4 with $f\equiv1$,

$$\mathbf{E}\left[\Theta^{n,r_n,s_n}_{\mu^n}(\bar{\xi}^n_1,\dots,\bar{\xi}^n_m)\right]=\mathbf{P}\left\{N\geq m\,\tau_{r_n}(\Sigma)>m\;\middle|\;\sum_{i=r_n+1}^n\xi^n_i=n-1-s_n\right\}, \tag{A.17}$$

where $\Sigma=\Sigma_{\vec{Z}}$ with $\vec{Z}=(Z_1,\dots,Z_n)=(d_1,\dots,d_{r_n},\xi^n_{r_n+1},\dots,\xi^n_n)$ such that $d_1=s_n$, and $d_2,\dots d_{r_n}=0$. (Indeed, any fixed choice of $d_1,\dots,d_{r_n}$ with $\sum_{i=1}^{r_n}d_i=s_n$ would suffice.) To see that the probability on the right-hand side of (A.17) tends to 1 as $n\to\infty$, first note that $N\overset{\mathrm{d}}{=}\text{Binomial}(n-r_n,1-\mu_0)$ where $r_n<n^\varepsilon$. So even after conditioning on the event $\{\sum_{i=r_n+1}^n\xi^n_i=n-1-s_n\}$, which occurs with probability $O(n^{-1/2})$, there are $(1+o_{\mathbf{P}}(1))(n-r_n)(1-\mu_0)$ non-zero entries of $(\xi^n_{r_n+1},\dots,\xi^n_n)$. Therefore, to prove uniform integrability it remains to show that $\tau_{r_n}(\Sigma)=\omega_{\mathbf{P}}(\sqrt{n})$.

To see this, observe that for any $k\in[n]$,

$$\mathbf{P}\left\{\tau_{r_n}(\Sigma)=k+1\mid(Z_{\Sigma(1)},\dots,Z_{\Sigma(k)}),\ \tau_{r_n}(\Sigma)\geq k\right\}=\frac{s_n}{\sum_{i=k+1}^nZ_{\Sigma(i)}}.$$

Since $\vec{Z}$ contains $(1+o_{\mathbf{P}}(1))(n-r_n)(1-\mu_0)+1$ positive entries, this denominator is $(1+o_{\mathbf{P}}(1))(n-r_n)(1-\mu_0)+1$ uniformly over all $k\leq m=O(\sqrt{n})$, and all labeled random reorderings of $\vec{Z}$. Moreover, since $s_n=o(\sqrt{n})$ by assumption,

$$\mathbf{P}\left\{\tau_{r_n}(\Sigma)=k+1\mid\tau_{r_n}(\Sigma)\geq k\right\}=o(n^{-1/2}),$$

uniformly across all $k\leq m$. The claim follows by summing these probabilities over $k\leq m$, since by the above $\mathbf{P}\left\{\tau_{r_n}(\Sigma)>k\right\}=(1-o(n^{-1/2}))^k$, and in particular for $T>0$,

$$\mathbf{P}\left\{\tau_{r_n}(\Sigma)>T\sqrt{n}\right\}=(1-o(n^{-1/2}))^{T\sqrt{n}}, \tag{A.18}$$

which tends to 1 as $n\to\infty$. □

A.4. Backstage at the hairy tour. To control the restrictions of the discrete snake introduced in the proofs of Theorems 1.1, 1.4 and 1.5 we require a couple of technical lemmas. The first of these results shows that if we truncate the displacements of the discrete snake by $n^{1/(4-\eta)-\delta}$, then the global moments agree with assumption [**A1**] in the limit.

Fix $\eta \in (0,2]$, and $\delta \in (0, 1/(4-\eta))$. For $n \geq 1$, and $k \geq 1$ let

$$Y_k^{n,\delta} = (Y_{k,1}^{n,\delta}, \dots, Y_{k,k}^{n,\delta}) = \begin{cases} (Y_{k,1}, \dots, Y_{k,k}) & \text{if } \max_{1\leq j\leq k} |Y_{k,j}| \leq n^{1/(4-\eta)-\delta}, \\ 0 & \text{else.} \end{cases}$$

LEMMA A.11. *As $n \to \infty$, it holds that*

$$\left|\mathbf{E}\left[Y^{n,\delta}_{\bar\xi,U_{\bar\xi}}\right]\right| = O\left((n^{1/(4-\eta)-\delta})^{1-2(4-\eta)/3}\right) \quad \textit{and} \quad \mathbf{Var}\left(Y^{n,\delta}_{\bar\xi,U_{\bar\xi}}\right) \to \beta^2.$$

PROOF. First, observe that by Hölder's inequality, there exists a constant $c > 0$ such that

$$\mathbf{P}\left\{\max_{1\leq i\leq \bar\xi} |Y_{\bar\xi,i}| > y\right\} \leq \mathbf{E}\left[\xi^3\right]^{1/3} \mathbf{P}\left\{\max_{1\leq i\leq \xi} |Y_{\xi,i}| > y\right\}^{2/3} \leq cy^{-2(4-\eta)/3}.$$

Then, by global centering,

$$\begin{aligned}
\left|\mathbf{E}\left[Y^{n,\delta}_{\bar\xi,U_{\bar\xi}}\right]\right| &= \left|\mathbf{E}\left[Y_{\bar\xi,U_{\bar\xi}} \mathbf{1}_{[\max_{1\leq i\leq\bar\xi}|Y_{\bar\xi,i}|>n^{1/(4-\eta)-\delta}]}\right]\right| \\
&\leq \int_0^\infty \mathbf{P}\left\{\left|Y_{\bar\xi,U_{\bar\xi}} \mathbf{1}_{[\max_{1\leq i\leq\bar\xi}|Y_{\bar\xi,i}|>n^{1/(4-\eta)-\delta}]}\right| > y\right\} dy \\
&\leq n^{1/(4-\eta)-\delta} \mathbf{P}\left\{\max_{1\leq i\leq\bar\xi} |Y_{\bar\xi,i}| > n^{1/(4-\eta)-\delta}\right\} \\
&\quad + \int_{n^{1/(4-\eta)-\delta}}^\infty \mathbf{P}\left\{\max_{1\leq i\leq\bar\xi} |Y_{\bar\xi,i}| > y\right\} dy \\
&\leq c(n^{1/(4-\eta)-\delta})^{1-2(4-\eta)/3} - \frac{c}{2(4-\eta)/3-1}\left[y^{-2(4-\eta)/3+1}\right]_{n^{1/(4-\eta)-\delta}}^\infty \\
&= O\left((n^{1/(4-\eta)-\delta})^{1-2(4-\eta)/3}\right),
\end{aligned}$$

as claimed.

As for the variance,

$$\begin{aligned}
\mathbf{Var}\left(Y^{n,\delta}_{\bar\xi,U_{\bar\xi}}\right) &= \mathbf{E}\left[\left(Y^{n,\delta}_{\bar\xi,U_{\bar\xi}}\right)^2\right] - \left(\mathbf{E}\left[Y^{n,\delta}_{\bar\xi,U_{\bar\xi}}\right]\right)^2 \\
&= \mathbf{E}\left[Y^2_{\bar\xi,U_{\bar\xi}} \mathbf{1}_{[\max_{1\leq i\leq\bar\xi}|Y_{\bar\xi,i}|\leq n^{1/(4-\eta)-\delta}]}\right] - \mathbf{E}\left[Y_{\bar\xi,U_{\bar\xi}} \mathbf{1}_{[\max_{1\leq i\leq\bar\xi}|Y_{\bar\xi,i}|\leq n^{1/(4-\eta)-\delta}]}\right]^2 \\
&\to \mathbf{E}\left[Y^2_{\bar\xi,U_{\bar\xi}}\right] = \beta^2,
\end{aligned}$$

as $n \to \infty$, by dominated convergence and the result for the mean. □

The above lemma pertains to snakes where the displacements which are above $n^{1/(4-\eta)-\delta}$ are all set to 0. The next lemma in this section will help us to understand the asymptotics of the head of the discrete snake where displacements which are *below* $n^{1/(4-\eta)-\delta}$ are set to 0. More specifically, we present a tail bound for the size of a set of marked vertices in random trees, which we apply in Proposition 5.5 where the marked vertices pertain to vertices $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ for which $\|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta}$.

LEMMA A.12. *Let* $\mathrm{d} = (d_1, \ldots, d_n)$ *be a degree sequence, fix* $\mathcal{B} \subset [n]$ *and write* $K = |\mathcal{B}|$, *and* $\Delta = \max_{1 \le i \le n} d_i$. *Let* B_{d} *be the smallest distance between two vertices in* $\mathcal{B}$ *that are ancestrally related in* $T_{\mathrm{d}} = B(\Pi_{\mathrm{d}})$ *(with* $B_{\mathrm{d}} = \infty$ *if no vertices in* $\mathcal{B}$ *are ancestrally related). Then, for any* $b \ge 0$

$$\mathbf{P}\{B_{\mathrm{d}} \le b\} \le K\left(1 - \left(1 - \frac{K\Delta}{n-1-b\Delta}\right)^b\right).$$

PROOF. It suffices to show the statement for integer b since for general b, $\mathbf{P}\{B_{\mathrm{d}} \le b\} = \mathbf{P}\{B_{\mathrm{d}} \le \lfloor b \rfloor\}$ and the upper bound is increasing in b.

Fix a degree sequence and a set $\mathcal{B}$. Without loss of generality, assume that $\mathcal{B} = \{n-K+1, \ldots, n\}$.

For $v \in [n]$, let $p(v)$ be the parent of v in $\mathrm{T}_{\mathbf{d}}$ (with $p(v) = v$ if v is the root of $\mathrm{T}_{\mathbf{d}}$). Also set $p^0(v) = v$ and recursively for $k \ge 1$ define the k-th ancestor of v as $p^k(v) = p(p^{k-1}(v))$.

We will show that

$$\mathbf{P}\left\{\{p^1(n), \ldots, p^b(n)\} \cap \{n-K+1, \ldots, n-1\} = \emptyset\right\} \ge \left(1 - \frac{K\Delta}{n-1-b\Delta}\right)^b, \tag{A.19}$$

after which the statement follows by symmetry and the union bound.

We will prove (A.19) by induction. To ease notation, write $p^k = p^k(n)$ for $k \ge 0$. For (i,c) such that $i \in [n], c \in [d_i]$ write $\Pi_{\mathrm{d}}^{-1}(i,c)$ for the position of (i,c) in Π_{d}.

We will define a sequence of σ-algebras $(\mathcal{F}_k)_{k \ge 0}$ such that for each $k \ge 1$, $\mathcal{F}_k$ is the σ-algebra generated by the first k ancestors of n and the positions of their corresponding entries in Π_{d}. Let, $\mathcal{F}_0 = \sigma(\Pi_{\mathrm{d}}^{-1}(n,c) : c \in [d_n])$ contain the information on the position of vertex n in Π_{d}.

If $d_n = 0$ and $\{\Pi_{\mathrm{d}}^{-1}(n,c) : c \in [d_n]\} = \emptyset$ then n is the final vertex in the final path of the line-breaking construction, and the last entry of Π_{d} gives its parent. Thus, in this case, we reveal $\Pi_{\mathrm{d}}(n-1)$ and we have $\Pi_{\mathrm{d}}(n-1) = (p^1, c')$ for some $c' \in [d_{p^1}]$. Then, we reveal all other entries of the form $(p^1, c), c \in [d_{p^1}]$ in Π_{d} and this yields $\mathcal{F}_1$.

If $d_n > 0$, then set $m_0 = \min\{\Pi_{\mathrm{d}}^{-1}(n,c) : c \in [d_n]\}$. If $m_0 = 1$ then n is the root of $\mathbf{T}_{\mathrm{d}}$ so $p^\ell = n$ for all $\ell \ge 1$, and so we let $\mathcal{F}_\ell = \mathcal{F}_0$ for all $\ell \ge 1$. Otherwise, the entry before the first occurrence of an entry of the form $(n,c), c \in [d_n]$ in Π_{d} gives the parent of n so then we obtain $\mathcal{F}_1$ as follows. We reveal $\Pi_{\mathrm{d}}(m_0 - 1)$. In that case $\Pi_{\mathrm{d}}(m_0 - 1) = (p^1, c')$ for some $c' \in [d_{p^1}]$. Secondly, we reveal all other entries of the form $(p^1, c), c \in [d_{p^1}]$ in Π_{d} and this yields $\mathcal{F}_1$.

For $k \ge 1$, given $\mathcal{F}_k$, let

$$m_k = \min\{\Pi_{\mathrm{d}}^{-1}(p^k, c) : c \in [d_{p^k}]\}.$$

If $m_k = 1$ then p^k is the root of $\mathbf{T}_{\mathrm{d}}$ so $p^\ell = p^k$ and we take $\mathcal{F}_\ell = \mathcal{F}_k$ for all $\ell \ge k$. If $m_k > 1$, we obtain $\mathcal{F}_{k+1}$ as follows. First, we reveal $\Pi_{\mathrm{d}}(m_k - 1)$. In that case $\Pi_{\mathrm{d}}(m_k - 1) = (p^{k+1}, c')$ for some $c' \in [d_{p^{k+1}}]$. Secondly, we reveal all other entries of the form $(p^{k+1}, c), c \in [d_{p^{k+1}}]$ in Π_{d} and this yields $\mathcal{F}_{k+1}$.

Now, observe that, for $k \ge 0$, given $\mathcal{F}_k$, the unrevealed entries of Π_{d} occur in an order given by a uniformly random permutation. So given $\mathcal{F}_k$ if $m_k > 1$ the k-th ancestor of n is the first coordinate of a uniformly random sample from

$$\{(i,c) : i \in [n] \backslash \{p^0, \ldots, p^k\}, c \in [d_i]\}$$

and

$$\begin{aligned}\mathbf{P}&\left\{p^{k+1} \in \{n-K+1, \ldots, n-1\} \;\middle|\; \mathcal{F}_k, \{p^1, \ldots, p^k\} \cap \{n-K+1, \ldots, n-1\} = \emptyset\right\} \\ &= \frac{d_{n-K+1} + \cdots + d_{n-1}}{n - 1 - \sum_{i=0}^k d_{p^i}} \le \frac{K\Delta}{n-1-(k+1)\Delta}.\end{aligned}$$

If $m_k = 1$ then the conditional probability above is 0 so the inequality also holds.

Therefore, we see inductively that

$$\mathbf{P}\left\{\{p^1,\dots,p^b\}\cap\{n-K+1,\dots,n-1\}=\emptyset\right\} \geq \prod_{k=1}^{b}\left(1-\frac{K\Delta}{n-1-k\Delta}\right) \geq \left(1-\frac{K\Delta}{n-1-b\Delta}\right)^b.$$

□

Funding. This work was partially supported by the Marie Skłodowska–Curie RISE grant RandNET (Randomness and Learning in Networks), grant agreement ID 101007705. LAB acknowledges the financial support of NSERC and of the Canada Research Chairs program. SD acknowledges the financial support of the CogniGron research center and the Ubbo Emmius Funds (University of Groningen). Her research was also partially supported by the Marie Skłodowska-Curie grant GraPhTra (Universality in phase transitions in random graphs), grant agreement ID 101211705. RM's work is supported by the EPSRC Centre for Doctoral Training in the Mathematics of Random Systems: Analysis, Modelling, and Simulation (EP/S023925/1). The final part of the work was carried out while LAB, SD and CG were in residence at the Simons Laufer Mathematical Sciences Institute in Berkeley, California, during the Spring 2025 semester. They gratefully acknowledge the financial support of the National Science Foundation under Grant No. DMS-1928930.

REFERENCES

[1] ADDARIO-BERRY, L. and ALBENQUE, M. (2017). The scaling limit of random simple triangulations and random simple quadrangulations. *Ann. Probab.* **45** 2767–2825. https://doi.org/10.1214/16-AOP1124 MR3706731

[2] ADDARIO-BERRY, L. and ALBENQUE, M. (2021). Convergence of non-bipartite maps via symmetrization of labeled trees. *Ann. H. Lebesgue* **4** 653–683. https://doi.org/10.5802/ahl.84 MR4315765

[3] ADDARIO-BERRY, L., DONDERWINKEL, S. and KORTCHEMSKI, I. (2025). Critical trees are neither too short nor too fat. *Ann. H. Lebesgue* **8** 113–149. https://doi.org/10.5802/ahl.231 MR4912670

[4] ALDOUS, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226

[5] ALDOUS, D. and PITMAN, J. (2002). Invariance principles for non-uniform random mappings and trees. In *Asymptotic combinatorics with application to mathematical physics (St. Petersburg, 2001). NATO Sci. Ser. II Math. Phys. Chem.* **77** 113–147. Kluwer Acad. Publ., Dordrecht. MR1999358

[6] BILLINGSLEY, P. (1999). *Convergence of probability measures*, second ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. John Wiley & Sons, Inc., New York. https://doi.org/10.1002/9780470316962 MR1700749

[7] BROUTIN, N. and MARCKERT, J.-F. (2014). Asymptotics of trees with a prescribed degree sequence and applications. *Random Structures Algorithms* **44** 290–316. https://doi.org/10.1002/rsa.20463 MR3188597

[8] CHASSAING, P. and SCHAEFFER, G. (2004). Random planar lattices and integrated superBrownian excursion. *Probab. Theory Related Fields* **128** 161–212. https://doi.org/10.1007/s00440-003-0297-8 MR2031225

[9] CURIEN, N. and KORTCHEMSKI, I. (2014). Random stable looptrees. *Electron. J. Probab.* **19** no. 108, 35. https://doi.org/10.1214/EJP.v19-2732 MR3286462

[10] DUQUENSE, T. and REBEI, F. (2025). Scaling limits of critical branching random walks via their snakes. In preparation.

[11] DUQUESNE, T. (2003). A limit theorem for the contour process of conditioned Galton–Watson trees. *Ann. Probab.* **31** 996–1027. https://doi.org/10.1214/aop/1048516543 MR1964956

[12] DUQUESNE, T. and LE GALL, J.-F. (2002). Random trees, Lévy processes and spatial branching processes. *Astérisque* **281** vi+147. MR1954248

[13] GITTENBERGER, B. (2003). A note on: "State spaces of the snake and its tour—convergence of the discrete snake" [31] by J.-F. Marckert and A. Mokkadem. *J. Theoret. Probab.* **16** 1063–1067. https://doi.org/10.1023/B:JOTP.0000012006.91251.e8 MR2033198

[14] HAAS, B. and MIERMONT, G. (2012). Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees. *Ann. Probab.* **40** 2589–2666. https://doi.org/10.1214/11-AOP686 MR3050512

[15] JANSON, S. (2006). Left and right pathlengths in random binary trees. *Algorithmica* **46** 419–429. https://doi.org/10.1007/s00453-006-0099-3 MR2291963

[16] JANSON, S. (2011). Probability asymptotics: notes on notation. Preprint https://arxiv.org/pdf/1108.3924.

[17] JANSON, S. (2012). Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probab. Surv.* **9** 103–252. https://doi.org/10.1214/11-PS188 MR2908619

[18] JANSON, S. and MARCKERT, J.-F. (2005). Convergence of discrete snakes. *J. Theoret. Probab.* **18** 615–647. https://doi.org/10.1007/s10959-005-7252-9 MR2167644

[19] KALLENBERG, O. (2021). *Foundations of modern probability*, third ed. *Probability Theory and Stochastic Modelling* **99**. Springer, Cham. https://doi.org/10.1007/978-3-030-61871-1 MR4226142

[20] KHANFIR, R. (2022). Convergences of looptrees coded by excursions. *Ann. Inst. Henri Poincaré Probab. Stat.* to appear. https://doi.org/10.48550/arXiv.2208.11528

[21] KORTCHEMSKI, I. and MARZOUK, C. (2024). Random Lévy looptrees and Lévy maps. Preprint. https://doi.org/10.48550/arXiv.2402.04098

[22] KORTCHEMSKI, I. and RICHIER, L. (2020). The boundary of random planar maps via looptrees. *Ann. Fac. Sci. Toulouse Math. (6)* **29** 391–430. https://doi.org/10.5802/afst.1636 MR4150547

[23] LE GALL, J.-F. (1993). A class of path-valued Markov processes and its applications to superprocesses. *Probab. Theory Related Fields* **95** 25–46. https://doi.org/10.1007/BF01197336 MR1207305

[24] LE GALL, J.-F. (1995). The Brownian snake and solutions of $\Delta u = u^2$ in a domain. *Probab. Theory Related Fields* **102** 393–432. https://doi.org/10.1007/BF01192468 MR1339740

[25] LE GALL, J.-F. (2005). Random trees and applications. *Probab. Surv.* **2** 245–311. https://doi.org/10.1214/154957805100000140 MR2203728

[26] LE GALL, J.-F. (2013). Uniqueness and universality of the Brownian map. *Ann. Probab.* **41** 2880–2960. https://doi.org/10.1214/12-AOP792 MR3112934

[27] LEI, T. (2019). Scaling limit of random forests with prescribed degree sequences. *Bernoulli* **25** 2409–2438. https://doi.org/10.3150/18-BEJ1058 MR4003553

[28] MARCKERT, J.-F. (2004). The rotation correspondence is asymptotically a dilatation. *Random Structures Algorithms* **24** 118–132. https://doi.org/10.1002/rsa.10111 MR2035871

[29] MARCKERT, J.-F. (2008). The lineage process in Galton–Watson trees and globally centered discrete snakes. *Ann. Appl. Probab.* **18** 209–244. https://doi.org/10.1214/07-AAP450 MR2380897

[30] MARCKERT, J.-F. and MIERMONT, G. (2007). Invariance principles for random bipartite planar maps. *Ann. Probab.* **35** 1642–1705. https://doi.org/10.1214/009117906000000908 MR2349571

[31] MARCKERT, J.-F. and MOKKADEM, A. (2003). States spaces of the snake and its tour—convergence of the discrete snake. *J. Theoret. Probab.* **16** 1015–1046. https://doi.org/10.1023/B:JOTP.0000012004.73051.f3 MR2033196

[32] MARCKERT, J.-F. and MOKKADEM, A. (2003). The depth first processes of Galton–Watson trees converge to the same Brownian excursion. *Ann. Probab.* **31** 1655–1678. https://doi.org/10.1214/aop/1055425793 MR1989446

[33] MARZOUK, C. (2020). Scaling limits of discrete snakes with stable branching. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 502–523. https://doi.org/10.1214/19-AIHP970 MR4058997

[34] MIERMONT, G. (2013). The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Math.* **210** 319–401. https://doi.org/10.1007/s11511-013-0096-8 MR3070569

[35] PETROV, V. V. (1975). *Sums of independent random variables*. *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]* **82**. Springer-Verlag, New York-Heidelberg Translated from the Russian by A. A. Brown. MR388499

[36] PITMAN, J. (2003). Poisson–Kingman partitions. In *Statistics and Science: a Festschrift for Terry Speed*. *IMS Lecture Notes Monogr. Ser.* **40** 1–34. Inst. Math. Statist., Beachwood, OH. https://doi.org/10.1214/lnms/1215091133 MR2004330

[37] PITMAN, J. (2006). *Combinatorial stochastic processes*. *Lecture Notes in Mathematics* **1875**. Springer-Verlag, Berlin Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, with a foreword by Jean Picard. MR2245368