The Annals of Probability
2015, Vol. 43, No. 3, 1045–1081
DOI: 10.1214/13-AOP871

SCALING LIMITS OF RANDOM PLANAR MAPS WITH A UNIQUE LARGE FACE[1]

BY SVANTE JANSON AND SIGURDUR ÖRN STEFÁNSSON

Uppsala University, and Uppsala University and Nordita

We study random bipartite planar maps defined by assigning nonnegative weights to each face of a map. We prove that for certain choices of weights a unique large face, having degree proportional to the total number of edges in the maps, appears when the maps are large. It is furthermore shown that as the number of edges n of the planar maps goes to infinity, the profile of distances to a marked vertex rescaled by $n^{-1/2}$ is described by a Brownian excursion. The planar maps, with the graph metric rescaled by $n^{-1/2}$, are then shown to converge in distribution toward Aldous' Brownian tree in the Gromov–Hausdorff topology. In the proofs, we rely on the Bouttier–di Francesco–Guitter bijection between maps and labeled trees and recent results on simply generated trees where a unique vertex of a high degree appears when the trees are large.

1. Introduction. A planar map is an embedding of a finite connected graph into the two-sphere. Two planar maps are considered to be the same if one can be mapped to the other with an orientation-preserving homeomorphism of the sphere. The connected components of the complement of the edges of the graph are called faces. The degree of a vertex is the number of edges containing it and the degree of a face is the number of edges in its boundary where an edge is counted twice if both its sides are incident to the face.

In recent years, there has been great progress in understanding probabilistic aspects of large planar maps; we refer to [42] for a detailed overview. One approach has been to study the scaling limit of a sequence of random planar maps obtained by rescaling the graph distance on the maps appropriately with their size and taking the limit as the size goes to infinity. This notion of convergence involves viewing the maps as elements of the set of all compact metric spaces, up to isometries, equipped with the Gromov–Hausdorff topology. Le Gall showed that the scaling limit of uniform $2p$-angulations (all faces of degree $2p$) exists along a suitable subsequence and he furthermore showed that its topology is independent of the subsequence and proved that its Hausdorff dimension equals 4 [39]. Subsequently, Le Gall and Paulin proved that the limit has the topology of the sphere [43]. Recently, Miermont showed that in the case of uniform quadrangulations the choice

Received December 2012; revised June 2013.

[1]Supported in part by the Knut and Alice Wallenberg Foundation.

MSC2010 subject classifications. Primary 05C80; secondary 05C05, 60F17, 60J80.

Key words and phrases. Random maps, planar maps, mobiles, simply generated trees, continuum random tree, Brownian tree.

of subsequence is superfluous and the scaling limit in fact equals the so-called *Brownian map* up to a scale factor [46]. Le Gall proved independently that the same holds in the case of uniform $2p$-angulations and uniform triangulations [40].

The present work is motivated by a paper of Le Gall and Miermont [41] where the authors study random planar maps which roughly have the property that the distribution of the degree of a typical face is in the domain of attraction of a stable law with index $\alpha \in (1,2)$. The model belongs to a class of models in which Boltzmann weights are assigned to the faces of the map as we will now describe. Let $\mathcal{M}_n^*$ denote the set of rooted and pointed bipartite planar maps having n edges: the root is an oriented edge $e = (e_-, e_+)$ and pointed means that there is a marked vertex ρ in the planar map. The assumption of pointedness is for technical reasons. For a planar map $\mathbf{m} \in \mathcal{M}_n^*$, denote the set of faces in $\mathbf{m}$ by $F(\mathbf{m})$ and denote the degree of a face $f \in F(\mathbf{m})$ by $\deg(f)$. Note that the assumption that $\mathbf{m}$ is bipartite is equivalent to assuming that $\deg(f)$ is even for all f. Let $(q_i)_{i\geq 1}$ be a sequence of nonnegative numbers and assign a Boltzmann weight

$$W(\mathbf{m}) = \prod_{f \in F(\mathbf{m})} q_{\deg(f)/2} \tag{1.1}$$

to $\mathbf{m}$. The probability distribution μ_n is defined by normalizing $W(\mathbf{m})$

$$\mu_n(\mathbf{m}) = W(\mathbf{m})/Z_n, \tag{1.2}$$

where

$$Z_n = \sum_{\mathbf{m}' \in \mathcal{M}_n^*} W(\mathbf{m}') \tag{1.3}$$

is referred to as the finite volume partition function. We will always assume that $q_k > 0$ for some $k \geq 2$ to avoid the trivial case when all faces have degree 2. Note that for a given random element in $\mathcal{M}_n^*$ distributed by μ_n the marked vertex ρ is uniformly distributed. The motivation for studying these distributions is first of all related to questions of universality, namely, there is strong evidence that under certain integrability condition on the weights q_i the scaling limit of the maps distributed by μ_n is the Brownian map up to a scale factor [44]. Furthermore, the distributions are closely related to distributions arising in certain statistical mechanical models on random maps as is discussed in [41].

In [41], the authors show, among other things, that in the large planar maps under consideration there are many "macroscopic" faces present and that the scaling limit, if it exists, is different from the Brownian map. The presence of these large faces in the scaling limit can be understood by considering the labeled trees (mobiles) obtained from the planar maps using the Bouttier–di Francesco–Guitter (BDG) bijection [16]; see Section 2. For convenience, we rewrite the sequence $(q_i)_{i\geq 1}$ in terms of a new sequence $(w_i)_{i\geq 0}$ defined by $w_0 = 1$ and

$$w_i = \binom{2i-1}{i-1} q_i, \qquad i \geq 1. \tag{1.4}$$

Through yet another bijection between mobiles (with labels removed) and trees which we introduce in Section 3, the random trees corresponding to the maps distributed by μ_n can be viewed as so-called *simply generated trees* with weights w_i assigned to vertices of outdegree i. The choice of weights $(q_i)_{i\geq 1}$ in [41] corresponds to choosing the weights $(w_i)_{i\geq 0}$ as an offspring distribution of a critical Galton–Watson tree in the domain of attraction of a stable law of index $\alpha \in (1,2)$. In this case, the random trees converge, when scaled appropriately, to the so-called stable tree with index α. It follows from properties of the BDG bijection that the large faces in the planar maps correspond to individuals in the stable tree which have a macroscopic number of offspring, that is, vertices of large degree.

It was originally noted in [11] and recently developed further in [29, 30, 32, 36] that there exists a phase of simply generated trees where a unique vertex with a degree proportional to the size of the tree appears as the trees get large. This phenomenon has been referred to as condensation. The purpose of this paper is to study the scaling limit of planar maps corresponding to the condensation phase of the simply generated trees. The large vertex in the trees will produce a large face in the planar maps in analogy with the situation in [41]. The weights which we consider are chosen as explained below. Define the generating function

$$g(x)=\sum_{i=0}^{\infty} w_i x^i \tag{1.5}$$

and denote its radius of convergence by R. For $R>0$, define $\kappa=\lim_{t\nearrow R}\frac{tg'(t)}{g(t)}$ and for $R=0$ let $\kappa=0$. We will be interested in the following two cases, (C1) and (C2), which are known to be the only cases giving rise to condensation in the corresponding simply generated trees (see, e.g., [29]):

$$0<R<\infty \quad \text{and} \quad \kappa<1. \tag{C1}$$

$$R=0. \tag{C2}$$

In practice, we will consider the special case of (C1) when the weights furthermore obey

$$w_i=L(i)i^{-\beta} \tag{1.6}$$

for some $\beta>2$ and some slowly varying function L and the special case of (C2) when the weights furthermore obey

$$w_n=(n!)^{\alpha} \tag{1.7}$$

with $\alpha>0$. [By (1.4) and Stirling's formula, (1.6) is equivalent to $q_i = L'(i)4^{-i}i^{1/2-\beta}$ for another slowly varying function L'; the exponential factor 4^i does not matter when we fix the number of edges in the map, so we might as well take $q_i=L'(i)i^{1/2-\beta}$. However, we will in the sequel use w_i rather than q_i.]

We now introduce some formalism needed to state the results of the paper. Let $\mathbb{M}^*$ be the set of all pointed compact metric spaces viewed up to isometries, equipped with the pointed Gromov–Hausdorff metric d_{GH} [26]. Let $\mathbf{e}$ be a standard Brownian excursion on $[0, 1]$ and denote by $(\mathcal{T}_{\mathbf{e}}, \delta_{\mathbf{e}})$ Aldous' continuum random tree coded by $\mathbf{e}$. Recall that $\mathcal{T}_{\mathbf{e}} = [0, 1]/\{\delta_{\mathbf{e}} = 0\}$ where

$$\delta_{\mathbf{e}}(s, t) = \mathbf{e}(s) + \mathbf{e}(t) - 2 \inf_{s\wedge t<u<s\vee t} \mathbf{e}(u) \tag{1.8}$$

and by abuse of notation $\delta_{\mathbf{e}}$ is the induced distance on the quotient; see, for example, [3, 42]. From here on, we will denote a random element in $\mathcal{M}_n^*$ distributed by μ_n by (M_n, ρ); sometimes simplified to M_n. The graph distance in M_n will be denoted by d_n.

The main results of the paper are the following. In Theorem 4.2, we prove that for the weights (1.6) and (1.7), the limit as $n \to \infty$ of the profile of distances in M_n to the marked vertex ρ, rescaled by $(2(1-\kappa)n)^{-1/2}$, is described by a standard Brownian excursion; see Section 4 for definitions and a precise statement. Second, we prove the following theorem, which describes the limit of all distances (not just to the root).

THEOREM 1.1. *For the weights* (1.6) *and* (1.7), *the random planar maps* $((M_n, \rho), (2(1 - \kappa)n)^{-1/2}d_n)$ *distributed by* μ_n *and viewed as elements of* $\mathbb{M}^*$ *converge in distribution to* $((\mathcal{T}_{\mathbf{e}}, \rho^*), \delta_{\mathbf{e}})$, *where given* $\mathcal{T}_{\mathbf{e}}$, ρ^* *is a marked vertex chosen uniformly at random from* $\mathcal{T}_{\mathbf{e}}$.

Note that the root edge in $\mathcal{M}_n$ is forgotten when we regard the maps as elements of $\mathbb{M}^*$. We can reroot the random tree $\mathcal{T}_{\mathbf{e}}$ at the randomly chosen point ρ^*; this gives a new random rooted tree, which has the same distribution as $\mathcal{T}_{\mathbf{e}}$, as shown by [2], (20), but the point ρ^* is now the root. Hence, the result in Theorem 1.1 can also be formulated as follows.

THEOREM 1.2. *For the weights* (1.6) *and* (1.7), *the random planar maps in Theorem* 1.1 *converge in distribution in* $\mathbb{M}^*$ *to* $((\mathcal{T}_{\mathbf{e}}, 0), \delta_{\mathbf{e}})$, *where* 0 *denotes the root of* $\mathcal{T}_{\mathbf{e}}$.

Note that the limit $\mathcal{T}_{\mathbf{e}}$ is quite different from the Brownian map mentioned above; it is a (random) compact tree, and thus contractible, that is, of the same homotopy type as a point, and its Hausdorff dimension is 2 [24, 28]. Bettinelli [9] showed a similar convergence of uniform quadrangulations with a boundary toward Aldous' continuum random tree when the length of the boundary grows sufficiently fast and the distances in the quadrangulations are divided by the square root of the length of the boundary (see also the work of Bouttier and Guitter [17]). In this case, the boundary grows so fast that the faces disappear when rescaled and the boundary folds into a tree. This is analogous to our situation where the

FIG. 1. *The convergence in Theorem 1.1. A face of large degree (light gray) appears as the planar map gets larger and the boundary of that face collapses into a tree.*

boundary of the large face folds into a tree; see Figure 1. Other examples of planar maps converging to Aldous' continuum tree are stack triangulations [1] and random dissections of polygons [20].

The paper is organized as follows. We begin in Section 2 by recalling the BDG bijection between planar maps and planar mobiles. In Section 3, we introduce a bijection from the set of planar trees to itself which allows us to translate results on the condensed phase of simply generated trees to our setting. In Section 4, we state and prove Theorem 4.2 which was described informally above. Section 5 is devoted to the proof of Theorem 1.1. We end with some concluding remarks in Section 6 and Appendix containing further results on the random Galton–Watson trees used here and their relation to the two-type Galton–Watson trees used by Marckert and Miermont [44].

2. Planar mobiles and the BDG bijection. In this section, we define planar trees and mobiles and explain the BDG bijection between mobiles and planar maps. We consider rooted and pointed planar maps as is done in [43] which is different from the original case [16] where the maps were pointed but not rooted. (But see [16], Section 2.4.)

Planar trees are planar maps with a single face. It will be useful to keep this definition in mind later in the paper but we recall a more standard definition below and introduce some notation. The infinite Ulam–Harris tree T_∞ is the tree having a vertex set $\bigcup_{k=0}^{\infty} \mathbb{N}^k$, that is, the set of all finite sequences of natural numbers, and every vertex $v = v_1 \cdots v_k$ is connected to the corresponding vertex $v' = v_1 \cdots v_{k-1}$ with an edge. In this case, v is said to be a child of v' and v' is said to be the parent of v. The vertex belonging to $\mathbb{N}^0$ is called the root and denoted by r.

A rooted planar tree τ is defined as a rooted subtree of T_∞ having the properties that if $v = v_1 \cdots v_k$ is a vertex in τ then $v_1 \cdots v_{k-1} i$ is also a vertex in τ for every $i < v_k$. The vertices in a planar tree have a lexicographical ordering inherited from the lexicographical ordering of the vertices in T_∞. This order relation will be denoted by $\leq$. Let Γ_n be the set of rooted planar trees with n edges. We use the convention that the root vertex is connected to an extra half-edge (not counted as an edge) such that every vertex has degree $1+$ the number of its children ($1+$ its out degree). The number of edges in a planar tree τ will be denoted by $|\tau|$.

Consider a tree $\tau_n \in \Gamma_n$ and color its vertices with two colors, black and white, such that the root and vertices at even distance from the root are white and vertices at odd distance from the root are black. Denote the black vertex set of τ_n by $V^\bullet(\tau_n)$ and the white vertex set by $V^\circ(\tau_n)$. If u is a black vertex let u_0 be the (white) parent of u and denote by u_i the ith (white) neighbor of u going clockwise around u starting from u_0.

Assign integer labels $\ell_n : V^\circ(\tau) \to \mathbb{Z}$ to the white vertices of τ_n as follows: The root is labeled by 0. If u is black and has degree k then

$$\ell_n(u_{j+1}) \geq \ell_n(u_j) - 1 \tag{2.1}$$

for all $0 \leq j \leq k$, with the convention that $u_k = u_0$. The pair $\theta_n = (\tau_n, \ell_n)$ is called a mobile and we denote the set of mobiles having n edges by Θ_n.

The set $\Theta_n \times \{-1, 1\}$ is in a one to one correspondence with the set $\mathcal{M}_n^*$ according to (the rooted version of) the BDG bijection [16, 43]. We will denote the BDG bijection by $\mathcal{F}_n : \mathcal{M}_n^* \to \Theta_n \times \{-1, 1\}$ and we give an outline of its inverse direction below. Start with a planar mobile $\theta_n \in \Theta_n$ and an $\varepsilon \in \{-1, 1\}$. The contour sequence of θ_n is a list $a_0, a_1, \dots, a_{2n-1}$ of length $2n$ containing the vertices in the mobile (with repetitions allowed) constructed as follows. The first element is $a_0 = r$ and for each $i < 2n - 1$ the element following a_i is the first child (in the lexicographical order) of a_i which has still not appeared in the sequence or if all its children have appeared it is the parent of a_i. Extend this sequence to an infinite sequence by $2n$ periodicity. The white contour sequence is defined as $c_i = a_{2i}$, $i \geq 0$. The white contour sequence can be described as a list of the white vertices encountered in a clockwise walk around the contour of the tree, which starts at the root. For an index $i \in \mathbb{N}$, define its successor as

$$\sigma(i) = \inf\{j > i : \ell_n(c_j) = \ell_n(c_i) - 1\}, \tag{2.2}$$

where the infimum of the empty set is defined as ∞. Add an external vertex ρ to the mobile, disconnected from all other vertices, and write $\rho = c_\infty$. Also define the successor of a white vertex c_i as

$$\sigma(c_i) = c_{\sigma(i)}. \tag{2.3}$$

A planar map is constructed from θ_n by inserting an edge between c_i and $\sigma(c_i)$ for each $0 \leq i < n$ and deleting the edges and black vertices of the mobile. The vertex ρ corresponds to the marked vertex of the planar map. The root edge in the map is the edge between c_0 and $\sigma(c_0)$ and its direction is determined by the value of ε, if $\varepsilon = 1$ ($\varepsilon = -1$) the root edge points toward (away from) the root of the mobile.

Thus, the white vertices in the mobile along with an additional isolated white vertex ρ correspond to the vertices in the planar map and the black vertices in the mobile correspond to the faces in the planar map, a face having a degree two times the degree of its corresponding black vertex; see Figure 2 for an example. Moreover, the labels in a mobile give information on distances to the marked vertex ρ in the corresponding planar map $\mathbf{m}$. Define the label of ρ as $\ell_n(\rho) = \min_{u \in V^\circ(\mathbf{m})} \ell_n(u) - 1$. Then

$$d_n(v, \rho) = \ell_n(v) - \ell_n(\rho), \qquad v \in V(\mathbf{m}), \tag{2.4}$$

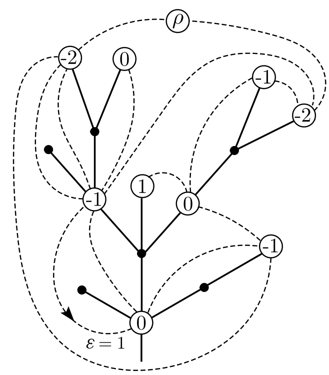

FIG. 2. *An illustration of the BDG bijection. The edges in the mobile are solid and the edges in the planar map are dashed.*

where by abuse of notation $\ell_n(v)$ stands for the label of the white vertex in the mobile corresponding to the vertex v in the planar map.

The probability distribution μ_n on $\mathcal{M}_n^*$ is carried to a probability distribution $\tilde{\mu}_n$ on $\Theta_n \times \{-1, 1\}$ through the BDG-bijection, that is, $\tilde{\mu}_n(A) = \mu_n(\mathcal{F}_n^{-1}(A))$ for any subset $A \subseteq \Theta_n \times \{-1, 1\}$ and $\tilde{\mu}_n$ can be described as follows: Let $\tau_n \in \Gamma_n$ and denote by $\lambda_n(\tau_n)$ the number of ways one can add labels to the white vertices of τ_n according to the above rules. One easily finds that

$$\lambda_n(\tau_n) = \prod_{v \in V^\bullet} \binom{2\deg(v) - 1}{\deg(v) - 1}. \tag{2.5}$$

This follows from counting the number of allowed label increments around each black vertex v. The number of label increments around v is $\deg(v)$, call them $x_1, x_2, \ldots, x_{\deg(v)}$ in say clockwise order. The number of different configurations is then given by

$$\sum_{\substack{x_1+\cdots+x_{\deg(v)}=0 \\ x_i \geq -1, \forall i}} 1 = \sum_{\substack{y_1+\cdots+y_{\deg(v)}=\deg(v) \\ y_i \geq 0, \forall i}} 1 = \binom{2\deg(v) - 1}{\deg(v) - 1}, \tag{2.6}$$

the number of compositions of $\deg(v)$ into $\deg(v)$ nonnegative parts.

A Boltzmann weight

$$\tilde{W}(\tau_n) = \prod_{v \in V^\bullet} \binom{2\deg(v) - 1}{\deg(v) - 1} q_{\deg(v)} = \prod_{v \in V^\bullet} w_{\deg(v)} \tag{2.7}$$

is assigned to the tree τ_n and

$$\tilde{\mu}_n\big(((\tau_n, \ell_n), \varepsilon)\big) = \tilde{W}(\tau_n) / \big(\lambda_n(\tau_n) Z_n\big), \tag{2.8}$$

where ℓ_n is any labeling of τ_n, $\varepsilon \in \{-1, 1\}$ and $Z_n = 2 \sum_{\tau_n \in \Gamma_n} \tilde{W}(\tau_n)$ is the finite volume partition function defined in (1.3). Note that given τ_n the labels ℓ_n are assigned uniformly at random from the set of all labelings and ε is chosen uniformly from $\{-1, 1\}$. We will also find it useful to study the distribution of τ_n after forgetting about the labeling and the value of ε. For that purpose, we define $\tilde{\nu}_n$ to be a probability distribution on Γ_n given by

$$\tilde{\nu}_n(\tau_n) = \sum_{\ell_n, \varepsilon} \tilde{\mu}_n\big(((\tau_n, \ell_n), \varepsilon)\big) = 2\tilde{W}(\tau_n)/Z_n. \tag{2.9}$$

This distribution was shown by Marckert and Miermont [44] to be the distribution of a certain two-type Galton–Watson tree; see Appendix.

2.1. *Distribution of labels in a fixed tree.* We provide a result which we will later need on the distribution of the maximum absolute value of the labels in a mobile.

LEMMA 2.1. *Let $\theta_n = (\tau_n, \ell_n) \in \Theta_n$ be a mobile with τ_n fixed (nonrandom) and the labels ℓ_n chosen uniformly from the allowed labelings of the white vertices of τ_n according to the rules (2.1). For every $p > 0$, there exists a constant $C(p) > 0$ independent of τ_n such that*

$$\mathbb{E}\Big(\sup_{v \in V^\circ(\tau_n)} \big|\ell_n(v)\big|^p\Big) \le C(p) n^{p/2}. \tag{2.10}$$

To prove this lemma, we relate the labels of τ_n to a random walk indexed by the white vertices in τ_n. We start by proving the result for $p > 2$ and the general case follows by Jensen's inequality. In the following, we will let $C_1, C_2, \dots$ be constants which do not depend on the tree τ_n but may depend on other quantities which will then be explicitly indicated. As before, denote the white contour sequence of a mobile (τ_n, ℓ_n) by $(c_i)_{0 \le i \le n}$ where by definition $c_n = c_0$. Let $\xi_1, \xi_2, \dots$ be a sequence of independent random variables identically distributed as

$$\mathbb{P}(\xi_1 = i) = 2^{-i-2}, \qquad i = -1, 0, 1, \dots. \tag{2.11}$$

(This is a shifted geometric distribution with mean 0.) The ξ_i will have the role of jumps of the random walk. For each black vertex $v \in \tau_n$, define the set $B_v \subseteq \mathbb{N}$ by

$$B_v = \{i \in \mathbb{N} | c_{i-1} \sim v \text{ and } c_i \sim v\}, \tag{2.12}$$

where $v \sim c_i$ means that v and c_i are nearest neighbors in τ_n. Define $S_m = \sum_{i=1}^m \xi_i$ and for any finite set $B \subset \mathbb{N}$ let $S_B = \sum_{i \in B} \xi_i$. Define the conditioned sequence of random variables

$$S_m^{\tau_n} = \big(S_m | S_{B_v} = 0 \text{ for all } v \in V^\bullet(\tau_n)\big), \qquad m = 0, \dots, n. \tag{2.13}$$

A simple calculation similar to the one in (2.6) shows that

$$\big(S_m^{\tau_n}\big)_{m=0}^n \overset{\mathrm{d}}{=} \big(\ell_n(c_m)\big)_{m=0}^n. \tag{2.14}$$

We have the following.

LEMMA 2.2. *Let τ_n be a fixed tree and let $\hat{S}_n^{\tau_n}(t)$ be the continuous function on* $[0,1]$ *defined by $\hat{S}_n^{\tau_n}(t) = n^{-1/2} S_{nt}^{\tau_n}$ when $t \in [0,1]$ and nt is an integer, and extended by linear interpolation to all $t \in [0,1]$. For every $p \geq 2$, there exists a constant $C_1(p)$ independent of n and τ_n such that*

$$\mathbb{E}\big|\hat{S}_n^{\tau_n}(t) - \hat{S}_n^{\tau_n}(s)\big|^p \leq C_1(p)|s-t|^{p/2} \tag{2.15}$$

for any $0 \leq s \leq t \leq 1$.

PROOF. First, consider the case when $s = k/n$ and $t = l/n$ for integers k and l. Suppose that $k < l$ and define $A = \{k+1, \dots, l\}$ and $A_v = A \cap B_v$, for every $v \in V^\bullet := V^\bullet(\tau_n)$. Then A is the disjoint union of the A_v, $v \in V^\bullet$, and thus

$$S_l - S_k = S_A = \sum_{v \in V^\bullet} S_{A_v}. \tag{2.16}$$

Conditioning on $S_{B_v} = 0$ for all $v \in V^\bullet$ now yields

$$S_l^{\tau_n} - S_k^{\tau_n} = \sum_{v \in V^\bullet} (S_{A_v} | S_{B_v} = 0). \tag{2.17}$$

Define $Y_v = (S_{A_v} | S_{B_v} = 0)$ for every $v \in V^\bullet$, and note that the random variables Y_v are independent. By [41], Lemma 1, there exists a constant $C_2(p) > 0$ such that for every v

$$\mathbb{E}|Y_v|^p \leq C_2(p)|A_v|^{p/2}. \tag{2.18}$$

Thus, by Rosenthal's inequality (see, e.g., [27], Theorem 3.9.1),

$$\begin{aligned}
&\mathbb{E}\big|S_l^{\tau_n} - S_k^{\tau_n}\big|^p \\
&\quad = \mathbb{E}\bigg|\sum_{v \in V^\bullet} Y_v\bigg|^p \leq C_3(p) \sum_{v \in V^\bullet} \mathbb{E}|Y_v|^p + C_4(p)\bigg(\sum_{v \in V^\bullet} \mathbb{E}|Y_v|^2\bigg)^{p/2} \\
&\quad \leq C_5(p) \sum_{v \in V^\bullet} |A_v|^{p/2} + C_6(p)\bigg(\sum_{v \in V^\bullet} |A_v|\bigg)^{p/2} \\
&\quad \leq C_7(p)\bigg(\sum_{v \in V^\bullet} |A_v|\bigg)^{p/2} = C_7(p)(l-k)^{p/2},
\end{aligned} \tag{2.19}$$

which is equivalent to (2.15) in this case. The case when $k/n \leq s \leq (k+1)/n$ follows directly since $\hat{S}_n^{\tau_n}(t)$ is linear on $[k/n, (k+1)/n]$ and the general case follows by splitting the interval $[s,t]$ into (at most) threes pieces and using Minkowski's inequality. $\square$

PROOF OF LEMMA 2.1. We will prove an equivalent statement for $S_m^{\tau_n}$. For any $t \in [0,1)$ define the dyadic approximations $t_j = 2^{-j}\lfloor 2^j t\rfloor$, $j = 0, 1, \dots$. Then $t_0 = 0$ and $t_j \to t$ as $j \to \infty$. Since $\hat{S}_n^{\tau_n}$ is continuous, it holds that $\hat{S}_n^{\tau_n}(t) =$

$\sum_{j=0}^{\infty}(\hat{S}_n^{\tau_n}(t_{j+1}) - \hat{S}_n^{\tau_n}(t_j))$. Fix $p > 2$. For any $\varepsilon > 0$, by Hölder's inequality, letting p' be the conjugate exponent

$$
\begin{aligned}
\left|\hat{S}_n^{\tau_n}(t)\right|^p &\le \left(\sum_{j=0}^{\infty} 2^{-p'\varepsilon j}\right)^{p/p'} \sum_{j=0}^{\infty} 2^{p\varepsilon j}\left|\hat{S}_n^{\tau_n}(t_{j+1}) - \hat{S}_n^{\tau_n}(t_j)\right|^p \\
&\le C_8(p,\varepsilon)\sum_{j=0}^{\infty} 2^{p\varepsilon j}\sum_{k=1}^{2^{j+1}}\left|\hat{S}_n^{\tau_n}(k/2^{j+1}) - \hat{S}_n^{\tau_n}((k-1)/2^{j+1})\right|^p.
\end{aligned}
\tag{2.20}
$$

The right-hand side is independent of t so taking the supremum over t and then taking the expectation and using (2.15) gives

$$
\begin{aligned}
\mathbb{E}\sup_{t\in[0,1]}\left|\hat{S}_n^{\tau_n}(t)\right|^p &\le C_8(p,\varepsilon)\sum_{j=0}^{\infty} 2^{p\varepsilon j}2^{j+1}C_1(p)2^{-jp/2} \\
&= C_9(p,\varepsilon)\sum_{j=0}^{\infty} 2^{(p\varepsilon+1-p/2)j}.
\end{aligned}
\tag{2.21}
$$

By choosing $\varepsilon < (p/2-1)/p$, the estimate (2.10) follows due to (2.14). $\square$

REMARK 2.3. By [13], Theorem 12.3 and (12.51), Lemma 2.2 implies also that the family of all random functions $\hat{S}_n^{\tau_n}(t)$, where $n \in \mathbb{N}$ and τ_n ranges over all rooted planar trees with n edges, is tight in $C([0,1])$; equivalently, we may consider $n^{-1/2}\ell_n(c_{nt})$, extended to $t \in [0,1]$ by linear interpolation. However, this family does not have a unique limit in distribution as $n \to \infty$. For example, if τ_n is a star, then $\hat{S}_n^{\tau_n}(t)$ converges to $\sqrt{2}\mathbf{b}(t)$, where $\mathbf{b}$ is a Brownian bridge, while if τ_n is a path, with the root at one endpoint, $\hat{S}_n^{\tau_n}(t)$ converges to $(2/3)^{1/2}\mathbf{B}(t \wedge (1-t))$ where $\mathbf{B}$ is a standard Brownian motion. And in many cases, $\hat{S}_n^{\tau_n}(t)$ converges to 0; if, for example, τ_n is a random binary tree, then $n^{-1/4}S_{nt}^{\tau_n}$ converges in distribution. See, for example, [31], and thus $\hat{S}_n^{\tau_n}(t)$ is typically of the order $n^{-1/4}$.

3. Another useful bijection and simply generated trees.

The coloring of the vertices in the mobiles is simply a bookkeeping device which groups together vertices in every second generation. We will continue referring to black and white vertices in trees even when no labels are assigned to white vertices. There exists a useful bijection from the set of trees Γ_n to itself which maps white vertices to vertices of degree 1 and black vertices of degree $k \ge 1$ to vertices of degree $k+1$. We will denote the bijection by $\mathcal{G}_n$. The bijection can be described informally in the following way: Start with a tree with vertices colored black and white as described above, the root being white. It will be mapped to a new tree which has the same vertex set as the old one but different edges. First consider the root, r, say of degree i and denote its black children by $r_1, \ldots, r_{i-1}$. Begin by attaching a

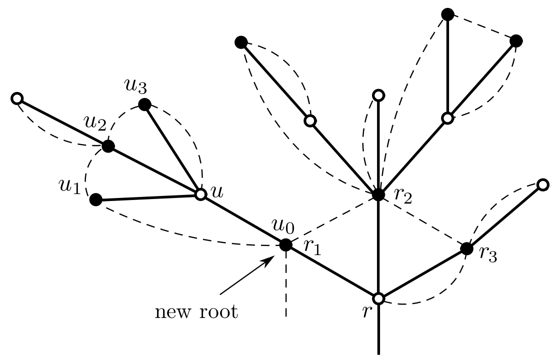

FIG. 3. *A diagram describing the bijection from Γ_n to itself which sends white vertices to vertices of degree 1 and black vertices of degree k to vertices of degree $k+1$.*

half-edge to r_1 which becomes the root of the new tree. Then connect r_j to r_{j+1} with an edge for $1 \le j \le i-1$ and finally connect r_{i-1} to the root r. Continue in the same way recursively for each of the subtrees attached to each of the r_j. More precisely, for a given white vertex $u \neq r$ of degree k denote its parent by u_0 and its children by $u_1, \dots, u_{k-1}$. Insert an edge between u_j and u_{j+1} for $0 \le j < k-1$ if possible (i.e., if $k > 0$), and finally connect u_{k-1} to u; see Figure 3.

To see that $\mathcal{G}_n$ is a bijection, we describe here its inverse. Start with a tree with all vertices black except the leaves which are white. Let $(a_i)_{i\ge 0}$ be the contour sequence of the tree. If a_j is a leaf let $\eta(j)$ denote the maximum number such that $a_j, a_{j+1}, \dots, a_{j+\eta(j)}$ all lie on the path from a_j to the root. Now, for each white a_j insert an edge between a_j and a_{j+k} for $1 \le k \le \eta(j)$ and remove the edges of the original tree. Let the last white vertex (within one period $[0, 2n) \cap \mathbb{Z}$) in the contour sequence be the root of the resulting tree. In the process, the degree of each black vertex is reduced by one and the degree of a white vertex a_j becomes $\eta(j)$ with the exception of the root in which case the degree becomes $\eta + 1$.

The usefulness of the bijection $\mathcal{G}_n$ is that it gives a simple description of the probability distribution $\tilde{\nu}_n$. Let ν_n be the push-forward of $\tilde{\nu}_n$ by $\mathcal{G}_n$. By (2.7) and the properties of $\mathcal{G}_n$,

$$\nu_n(\tau_n) = 2Z_n^{-1} \prod_{v \in V(\tau_n)} w_{\deg(v)-1}, \tag{3.1}$$

where we recall that w_i was defined in (1.4). The convenient thing is that now all vertices are treated equally. The probability measure ν_n describes simply generated trees, originally introduced by Meir and Moon [45] and has since been studied extensively; see, for example, [29] and references therein.

For the weights (1.6) in case (C1) in the Introduction, we define the probabilities

$$p_i = \frac{w_i}{g(1)}; \tag{3.2}$$

thus, for $i \geq 1$, with $\bar{L}(i) = g(1)^{-1} L(i)$,

$$p_i = \bar{L}(i) i^{-\beta}. \tag{3.3}$$

We let $\mathbb{P}_p$ be the law of a Galton–Watson tree with offspring distribution $(p_i)_{i \geq 0}$; see, for example, [6, 29]. Note that the expected number of offspring of an individual in the Galton–Watson process is equal to $g'(1)/g(1) = \kappa$. We will furthermore denote the variance of the number of offspring by

$$\sigma^2 = g''(1)/g(1) + \kappa(1 - \kappa), \tag{3.4}$$

which may be finite or infinite depending on the value of β. The measure ν_n viewed as a measure on the set of finite trees is in this case equal to the measure $\mathbb{P}_p(\cdot \| \tau| = n)$, where τ denotes a finite tree. In case (C2), ν_n has no such equivalent description in terms of a Galton–Watson process.

Using the bijection $\mathcal{G}_n$, one can translate known results on simply generated trees to the trees distributed by $\tilde{\nu}_n$. We will now introduce some notation and state a few technical results needed later on, some of which are interesting by themselves. In a random tree τ_n distributed by $\tilde{\nu}_n$ select a black vertex of maximum degree in some prescribed way (e.g., as the first such vertex encountered in the lexicographical order) and denote it by s. Denote the degree of s by Δ_n and the white vertices surrounding s by $s_0, s_1, \dots, s_{\Delta_n - 1}$ in a clockwise order, taking s_0 as the parent of s. For more compact notation, we do not explicitly write the dependency of s and s_i on n.

Denote by $\tau_{n,0}$ the tree which consists of all vertices in τ_n apart from s and its descendants. Let $\tau_{n,i}$ be the tree consisting of s_i and its descendants, $1 \leq i \leq \Delta_n - 1$. Furthermore, define $N^\circ_{n,i}$ as the number of white vertices in $\tau_{n,i}$. Write $\tau'_n = \mathcal{G}_n(\tau_n)$ and let s' be the vertex in τ'_n corresponding to the vertex s in τ_n. Then $\deg(s') = \Delta_n + 1$. Define the subtrees $\tau'_{n,i}$ around s' in τ'_n in an analogous way as above where $0 \leq i \leq \Delta_n$. It is then simple to check that

$$|\tau_{n,0}| = |\tau'_{n,0}| + |\tau'_{n,\Delta_n}| + 1 \quad \text{and} \quad |\tau_{n,i}| = |\tau'_{n,i}| \tag{3.5}$$

for $1 \leq i \leq \Delta_n - 1$. This is the key relation used to translate results from the simply generated trees to the mobiles.

Let $Y = (Y_t)_{t \geq 0}$ be the spectrally positive stable process with Laplace transform $\mathbb{E}(\exp(-\lambda Y_t)) = \exp(t \lambda^{2 \wedge (\beta - 1)})$. (This is a Lévy process with no negative jumps; the Lévy measure is $\Gamma(-\alpha)^{-1} x^{-\alpha - 1} \, dx$ on $x > 0$, where $\alpha = 2 \wedge (\beta - 1) \in (1, 2]$. See, for example, [8] and [51].) Denote by $\mathbb{D}([0, 1])$ the set of càdlàg functions $[0, 1] \to \mathbb{R}$ with the Skorohod topology; see [13], Section 14. We have the following proposition for the case (1.6), where $0 < \kappa < 1$.

PROPOSITION 3.1. *For the weights* (1.6), *the tree distributed by* $\tilde{\nu}_n$ *has the properties that*

(1)

$$\frac{\Delta_n}{n} \xrightarrow[n\to\infty]{\mathrm{p}} 1-\kappa. \tag{3.6}$$

(2)

$$\frac{N_n^\circ}{n} \xrightarrow[n\to\infty]{\mathrm{p}} p_0 \tag{3.7}$$

with $p_0 = 1/g(1)$ *defined in* (3.2).

(3) *For any fixed* $i \geq 0$, $|\tau_{n,i}|$ *converges in distribution as* $n \to \infty$ *to a finite random variable. For* $i \geq 1$, *the limit equals* $|\tau|$, *where* τ *is a Galton–Watson tree with offspring distribution* $(p_i)_{i\geq 0}$.

(4) *There exists a slowly varying function* $L_1(n)$ *such that for* $C_n = L_1(n) \times n^{1/(2\wedge(\beta-1))}$ *the following weak convergence holds in* $\mathbb{D}([0,1])$:

$$\left(\frac{\sum_{i=1}^{\lfloor(\Delta_n-1)t\rfloor} N_{n,i}^\circ - (p_0/1-\kappa)\Delta_n t}{C_n}\right)_{0\leq t\leq 1} \xrightarrow[n\to\infty]{\mathrm{d}} (Y_t)_{0\leq t\leq 1}. \tag{3.8}$$

(5) *It holds that*

$$\frac{1}{C_n} \sup_{1\leq i\leq \Delta_n-1} N_{n,i}^\circ \xrightarrow[n\to\infty]{\mathrm{d}} V \tag{3.9}$$

with C_n *from part* (4) *and the random variable* $V = \max_{0\leq t\leq 1} \Delta Y_t$.

PROOF. Part (1) follows from the corresponding result for simply generated trees which was originally proven in [32] in the case of an asymptotically constant slowly varying function L in (1.6) and then in [36] for a general slowly varying function L.

Part (2) follows from [29], Theorem 7.11(ii), since the number of white vertices N_n° in the tree τ_n equals the number of leaves in the simply generated tree τ_n', via the bijection $\mathcal{G}_n$.

For part (3), we note that the simply generated trees distributed by ν_n converge locally toward an infinite random tree; see [32], Theorem 5.3, in the case of an asymptotically constant slowly varying function L and [29], Theorem 7.1, for the most general case. Local convergence of the trees distributed by $\tilde{\nu}_n$ follows and the result in part (3) is then an immediate consequence; see the arguments in the proof of Theorem 3(iii) in [36].

Part (4) requires some explanation. We will prove a corresponding statement for the simply generated trees distributed by ν_n. Recall the notation τ_n for (colored) trees distributed by $\tilde{\nu}_n$ and τ_n' for (conditioned Galton–Watson) trees distributed by ν_n as explained in the paragraph above (3.5). First of all, note that the number of white vertices in $\tau_{n,i}$, which is denoted by $N_{n,i}^\circ$, corresponds to the number of leaves in the trees $\tau_{n,i}'$ for $1 \leq i \leq \Delta_n - 1$.

Recall that $\mathbb{P}_p$ is the law of a Galton–Watson process with the offspring distribution $(p_i)_{i\geq 0}$ defined in (3.2). Denote by N the total progeny (number of vertices) of the Galton–Watson process distributed by $\mathbb{P}_p$ and denote the random number of leaves by $N^{(0)}$. It is well known that $\mathbb{E}(N) = 1/(1-\kappa)$ (see, e.g., [6]), and furthermore,

$$\mathbb{E}N^{(0)} = \frac{p_0}{1-\kappa} = p_0\mathbb{E}N; \tag{3.10}$$

in fact, the expected number of vertices in generation $m \geq 0$ is κ^m, and the expected number of leaves among them is $p_0\kappa^m$, where summing over all $m \geq 0$ yields (3.10). This explains the linear term in (3.8).

Kortchemski [36], Theorem 4, proved a convergence result in $\mathbb{D}([0,1])$ which in our notation can be written as

$$\left(\frac{\sum_{i=1}^{\lfloor(\Delta_n-1)t\rfloor}(|\tau'_{n,i}|+1) - (1/(1-\kappa))\Delta_n t}{B'_n}\right)_{0\leq t\leq 1} \xrightarrow[n\to\infty]{\mathrm{d}} (Y_t)_{0\leq t\leq 1}, \tag{3.11}$$

where $B'_n = L_2(n)n^{1/(2\wedge(\beta-1))}$ for some slowly varying function L_2. The main idea of Kortchemski's proof is to use the fact that for n large, the subtrees $\tau'_{n,i}$ become asymptotically independent copies of a Galton–Watson process with law $\mathbb{P}_p$, and thus $|\tau'_{n,i}|+1$ appearing in the sum in (3.11) can be replaced by a sequence $(N_i)_{i\geq 1}$ of independent random variables distributed as N. (This is shown in [36] as a consequence of a corresponding result for random walks by Armendáriz and Loulakis [5].) Furthermore, it is well known (see, e.g., [34, 35, 48], [49], Section 6.1, [29], Theorem 15.5) that if ξ_i, $i = 1, 2, \dots$, is a sequence of independent random variables with the distribution $(p_i)_{i\geq 0}$, and we let $S_n = \sum_{i=1}^n \xi_i$, then

$$\mathbb{P}(N = n) = \frac{1}{n}\mathbb{P}(S_n = n-1). \tag{3.12}$$

Moreover, from the tail behavior (3.3) of $p_i = \mathbb{P}(\xi = i)$, it follows that, recalling that $\mathbb{E}\xi_i = \kappa$,

$$\begin{aligned}\mathbb{P}(S_n = n-1) &= \mathbb{P}\bigl(S_n - n\kappa = n(1-\kappa)-1\bigr)\\ &= n\bigl(1+o(1)\bigr)\mathbb{P}\bigl(\xi_1 = \lfloor n(1-\kappa)-1\rfloor\bigr)\end{aligned} \tag{3.13}$$

as $n\to\infty$, see [23] for more general statements. (In our case, (3.13) follows also directly by a modification of the proof of [29], Theorem 19.34.) Combining (3.12), (3.13) and (3.3), we obtain

$$\mathbb{P}(N = n) = \bigl(1+o(1)\bigr)(1-\kappa)^{-\beta}\bar{L}(n)n^{-\beta} = \bigl(1+o(1)\bigr)(1-\kappa)^{-\beta}p_n, \tag{3.14}$$

so the distribution of N also obeys (1.6) (with a different L), which by standard results (see, e.g., [25], Section XVII.5) implies that N is in the domain of attraction of a spectrally positive stable distribution of index $\alpha = 2\wedge(\beta-1)$, and thus

$$\left(\frac{\sum_{i=1}^{\lfloor nt\rfloor} N_i - (1/(1-\kappa))nt}{B'_n}\right)_{0\leq t\leq 1} \xrightarrow[n\to\infty]{\mathrm{d}} (Y_t)_{0\leq t\leq 1} \tag{3.15}$$

for a suitable $B_n' = L_2(n)n^{1/(2\wedge(\beta-1))}$. We refer to [36] for further details, and for the arguments using (3.15) to show (3.11).

Going through Kortchemski's proof, one sees that the latter arguments apply in our case also if we replace $\mathscr{X}^{(k)}$ in [36] by $(C_k^{-1}(\sum_{i=1}^{\lfloor kt\rfloor} N_i^{(0)} - \frac{p_0}{1-\kappa}kt))_{0\le t\le\eta}$ and the problem is reduced to showing that if $(N_i, N_i^{(0)})_{i\ge1}$ is a sequence of independent random vectors distributed as $(N, N^{(0)})$, then

$$\left(\frac{\sum_{i=1}^{\lfloor nt\rfloor} N_i^{(0)} - (p_0/(1-\kappa))nt}{C_n}\right)_{0\le t\le 1} \xrightarrow[n\to\infty]{\mathrm{d}} (\hat{Y}_t)_{0\le t\le 1}, \tag{3.16}$$

where $\hat{Y}$ has the same distribution as Y, and that this holds jointly with (3.15). (Joint convergence is used in the analogue of [36], (31), in the proof; however, the joint distribution of $(Y, \hat{Y})$ does not influence the result (3.8).) The proof of part (4) is thus completed by Lemma 3.4 below.

Finally, part (5) follows from part (4); see the proof of Corollary 2 in [36]. $\square$

REMARK 3.2. Actually, it would suffice to prove (3.16) separately; this and (3.15) show in particular that the left-hand sides are tight in $\mathbb{D}([0, 1])$, which implies that they are jointly tight in $\mathbb{D}([0, 1]) \times \mathbb{D}([0, 1])$, and we can obtain the desired joint convergence by considering suitable subsequences; this is enough to show (3.8) for the full sequence since the result does not depend on the joint distribution of $(Y_t)_{0\le t\le1}$ and $(\hat{Y}_t)_{0\le t\le1}$. We can show (3.16) by the same standard results as for (3.15) together with the estimate

$$\mathbb{P}(N^{(0)} = n) \sim c\bar{L}(n)n^{-\beta} \tag{3.17}$$

for some $c > 0$, see Lemma A.2, which shows that the distribution of $N^{(0)}$ has the same tail behavior as N and $(p_i)_{i\ge0}$. This thus yields an alternative proof of Proposition 3.1(4).

Before stating and proving Lemma 3.4 used above, we give another lemma.

LEMMA 3.3. *For the weights* (1.6), *with notation as above*, *as* $n \to \infty$,

$$\mathbb{P}(|N^{(0)} - p_0N| \ge n) = o(\bar{L}(n)n^{1-\beta}) = o(np_n) = o(\mathbb{P}(N \ge n)). \tag{3.18}$$

PROOF. Note first that (3.14) and (3.3) imply, by a standard calculation [14],

$$\begin{aligned}\mathbb{P}(N \ge n) &= (1+o(1))(1-\kappa)^{-\beta}(\beta-1)^{-1}\bar{L}(n)n^{1-\beta}\\ &= (1+o(1))(\beta-1)^{-1}(1-\kappa)^{-\beta}np_n.\end{aligned} \tag{3.19}$$

Let $a > 0$. Since $|N^{(0)} - p_0N| \le N$,

$$\begin{aligned}&\mathbb{P}(|N^{(0)} - p_0N| \ge n)\\ &\quad\le \mathbb{P}(N \ge an) + \mathbb{P}\left(|N^{(0)} - p_0N| \ge \frac{1}{a}N \text{ and } N \ge n\right).\end{aligned} \tag{3.20}$$

Let $\varepsilon > 0$. By [29], Theorem 7.11, $(N^{(0)}|N = n)/n \xrightarrow{\text{p}} p_0$ as $n \to \infty$. Thus, $\mathbb{P}(|N^{(0)} - p_0 N| \geq a^{-1}N|N = n) < \varepsilon$ if n is large enough, and for such n,

$$\begin{aligned}
&\mathbb{P}\Bigl(|N^{(0)} - p_0 N| \geq \frac{1}{a}N \text{ and } N \geq n\Bigr) \\
&\quad = \sum_{m=n}^{\infty} \mathbb{P}\Bigl(|N^{(0)} - p_0 N| \geq \frac{1}{a}N \Bigm| N = m\Bigr)\mathbb{P}(N = m) \qquad (3.21)\\
&\quad \leq \varepsilon \mathbb{P}(N \geq n).
\end{aligned}$$

Thus, (3.20) yields, for large n,

$$\mathbb{P}(|N^{(0)} - p_0 N| \geq n) \leq \mathbb{P}(N \geq an) + \varepsilon \mathbb{P}(N \geq n), \tag{3.22}$$

which by (3.19) yields, with $C = (\beta - 1)^{-1}(1 - \kappa)^{-\beta}$,

$$\mathbb{P}(|N^{(0)} - p_0 N| \geq n) \leq \bigl(1 + o(1)\bigr)C\bigl(a^{1-\beta} + \varepsilon\bigr)\bar{L}(n)n^{1-\beta}. \tag{3.23}$$

Since we may choose a arbitrarily large and ε arbitrarily small, (3.18) follows. □

LEMMA 3.4. *The limits* (3.15) *and* (3.16), *in distribution in* $\mathbb{D}([0, 1])$, *hold jointly.*

PROOF. Suppose first that the offspring distribution (3.3) has finite variance. [This implies $\beta \geq 3$ by (3.14).] It then follows from (3.14) that N and $N^{(0)} \leq N$ have finite variances and by a two-dimensional version of Donsker's theorem, the result follows with $2^{-1/2}Y_t$ and $2^{-1/2}\hat{Y}_t$ two different (dependent) standard Brownian motions, and $B'_n = \sqrt{\operatorname{Var}(N)n/2}$, $C_n = \sqrt{\operatorname{Var}(N^{(0)})n/2}$.

Suppose now instead that the variance of the offspring distribution is infinite; then $\mathbb{E}N^2 = \infty$. We follow [25], Section XVII.5, and let $\mu(x)$ be the truncated moment function

$$\mu(x) = \mathbb{E}\bigl(N^2 \mathbf{1}\{N \leq x\}\bigr). \tag{3.24}$$

Then $\mu(x) \to \infty$ as $x \to \infty$. Moreover, by [25], Theorem XVII.5.2 and XVII.(5.23), $\mu(x)$ is regularly varying with exponent $2 - \alpha = (3 - \beta) \vee 0$, and (3.15) holds with $n\mu(B'_n)/(B'_n)^2 \to C$ for some constant C.

If we similarly define the truncated moment function,

$$\mu_1(x) = \mathbb{E}\bigl((N^{(0)} - p_0 N)^2 \mathbf{1}\{|N^{(0)} - p_0 N| \leq x\}\bigr), \tag{3.25}$$

it follows easily by (3.18) [and $\mu(x) \to \infty$] that, as $x \to \infty$,

$$\mu_1(x) = o\bigl(\mu(x)\bigr) \tag{3.26}$$

and thus

$$\frac{n\mu_1(B'_n)}{(B'_n)^2} = o\biggl(\frac{n\mu(B'_n)}{(B'_n)^2}\biggr) \to 0, \qquad n \to \infty. \tag{3.27}$$

It follows by minor modifications of the arguments in [25], Section XVII.5, that

$$\frac{\sum_{i=1}^{n}(N_i^{(0)} - p_0 N_i)}{B'_n} \xrightarrow{\mathrm{p}} 0. \tag{3.28}$$

Moreover, by [33], Theorem 16.14, or by symmetrization and a stopping time argument, it follows that

$$\sup_{0\le t\le 1}\left|\sum_{i=1}^{\lfloor nt\rfloor}(N_i^{(0)} - p_0 N_i)\right|\Big/ B'_n \xrightarrow{\mathrm{p}} 0, \tag{3.29}$$

and thus (3.15) implies that (3.16) holds jointly with $\hat{Y}_t = Y_t$ and $C_n = p_0 B'_n$. (Note that $\hat{Y} = Y$ when the offspring variance is infinite, but not when it is finite.) □

For the case (1.7), where $\kappa = 0$, the proposition below follows immediately from [30], Theorems 2.4–2.5 and Remark 2.9.

PROPOSITION 3.5. *For $w_i = (i!)^\alpha$, $\alpha > 0$, the tree distributed by $\tilde{\nu}_n$ has the following properties*:

(1) *For $\alpha > 1$,*

$$n - \Delta_n \xrightarrow[n\to\infty]{\mathrm{p}} 0. \tag{3.30}$$

For $\alpha = 1$,

$$n - \Delta_n \xrightarrow[n\to\infty]{\mathrm{d}} \mathrm{Pois}(1). \tag{3.31}$$

For $\alpha < 1$

$$n - \Delta_n = O(n^{1-\alpha}) \tag{3.32}$$

with probability tending to 1 as $n \to \infty$.

(2)

$$\frac{N_n^\circ}{n} \xrightarrow[n\to\infty]{\mathrm{p}} 1. \tag{3.33}$$

(3) *The vertex s is the unique black child of the root r and*

$$\sup_{1\le i\le \Delta_n - 1} N_{n,i}^\circ \le \lfloor 1/\alpha\rfloor \vee 1 \tag{3.34}$$

with probability tending to 1 as $n \to \infty$.

The propositions above along with the correspondence between degrees of faces in the planar maps and degrees of black vertices in the mobiles show that a unique face of degree roughly equal to $(1-\kappa)n$ appears in the planar maps M_n with probability tending to 1 as $n \to \infty$.

4. Label process on mobiles. Let θ_n be a random mobile distributed by $\tilde{\mu}_n$, and denote by N_n° the random number of white vertices in θ_n. Order the white vertices in a lexicographical order $v_0, v_1, \dots, v_{N_n^\circ}$ (taking $v_{N_n^\circ} = v_0$). Again we do not write explicitly the dependency of v and v_i on n. Define the label process $L_n : \{0, 1, \dots, N_n^\circ\} \to \mathbb{Z}$ by $L_n(i) = \ell_n(v_i)$. Extend L_n to a function on $[0, N_n^\circ]$ by linear interpolation.

Denote the set of continuous functions from $[a, b]$ to $\mathbb{R}$ by $C([a, b])$ equipped with the topology of uniform convergence. Let $\mathbf{b}$ be the standard Brownian bridge on $[0, 1]$, starting and ending at 0. We will in this section prove the following result.

THEOREM 4.1. *For the weights* (1.6) *and* (1.7), *it holds that*

$$\left(\frac{1}{\sqrt{2(1-\kappa)n}} L_n\bigl(tN_n^\circ\bigr)\right)_{0\le t\le 1} \xrightarrow[n\to\infty]{\mathrm{d}} \bigl(\mathbf{b}(t)\bigr)_{0\le t\le 1} \tag{4.1}$$

with convergence in distribution in $C([0, 1])$.

Since the label function encodes information on distances, cf. (2.4), this result shows that the diameter of the maps grows like $n^{1/2}$. More precisely, we can translate Theorem 4.1 to a result on distances to the marked vertex ρ. Define the distance process $D_n : \{0, 1, \dots, N_n^\circ\} \to \mathbb{Z}$ by $D_n(i) = d(v_i, \rho)$. Extend D_n to a function on $[0, N_n^\circ]$ by linear interpolation, and then to a function on $\mathbb{R}$ with period N_n°. By (2.4),

$$D_n(t) = L_n(t) - \ell_n(\rho) = L_n(t) - \min_{0\le s\le N_n^\circ} L_n(s) + 1, \qquad 0 \le t \le N_n^\circ. \tag{4.2}$$

Further, let v_{i_*} be the first white vertex (in our ordering) that is a neighbor of ρ, that is, $i_* = \min\{i : \ell_n(v_i) = \min_j \ell_n(v_j)\}$.

THEOREM 4.2. *For the weights* (1.6) *and* (1.7), *it holds that*

$$\left(\frac{1}{\sqrt{2(1-\kappa)n}} D_n\bigl(tN_n^\circ + i_*\bigr)\right)_{0\le t\le 1} \xrightarrow[n\to\infty]{\mathrm{d}} \bigl(\mathbf{e}(t)\bigr)_{0\le t\le 1} \tag{4.3}$$

with convergence in distribution in $C([0, 1])$.

PROOF. The minimum of $\mathbf{b}$ is a.s. attained at a unique point, U say, and U is uniformly distributed on $[0, 1]$; moreover, by Vervaat's theorem [50], if this minimum is subtracted from $\mathbf{b}$ and the bridge is shifted (periodically) such that the minimum is located at 0 one obtains a standard Brownian excursion $\mathbf{e}$ on $[0, 1]$; see also [12].

By Skorohod's representation theorem, we may assume that the convergence in (4.1) holds a.s. Since the minimum point U is unique, it follows that the minimum point v_{i_*}/N_n° of the left-hand side converges to U a.s. (The minimum point v_{i_*} is typically not unique. We chose the first minimum point, but any other choice would also converge to U a.s.) The desired convergence (4.3) now follows from (4.2), (4.1) and Vervaat's theorem. □

We start by introducing some notation and proving a couple of lemmas before proceeding to the proof of Theorem 4.1. Begin by considering only the part of the label process which surrounds the vertex s, a black vertex of maximum degree. Let s_0 be the white parent of s and let s_i be its ith white child in clockwise order from s_0, where $i = 1, \dots, \Delta_n$ with the convention that $s_{\Delta_n} = s_0$. Define the function $L_n^\star : \{0, 1, \dots, \Delta_n\} \to \mathbb{Z}$ by $L_n^\star(i) = \ell_n(s_i)$. As before, extend $L_n^\star$ to a continuous function on $[0, \Delta_n]$ by linear interpolation.

LEMMA 4.3. *For the weights* (1.6) *and* (1.7), *it holds that*

$$\left(\frac{1}{\sqrt{2(1-\kappa)n}} L_n^\star(t\Delta_n)\right)_{0\le t\le 1} \xrightarrow[n\to\infty]{\mathrm{d}} (\mathbf{b}(t))_{0\le t\le 1} \tag{4.4}$$

with convergence in distribution in $C([0, 1])$.

PROOF. Let $\theta_n = (\tau_n, \ell_n)$ be a mobile distributed by $\tilde{\mu}_n$. By Propositions 3.1(1) and 3.5(1), $\Delta_n/n \longrightarrow^{\mathrm{p}} 1 - \kappa$ as $n \to \infty$. Using Skorohod's representation theorem, we may construct Δ_n and L_n on a common probability space such that this convergence holds almost surely, that is,

$$\Delta_n/n \xrightarrow[n\to\infty]{\mathrm{a.s.}} 1 - \kappa. \tag{4.5}$$

In the following, we will assume that this holds.

The label process $L_n^\star$, evaluated on the integers, is a random walk of length Δ_n having jump probabilities $\omega(k) = 2^{-k-2}$, $k = -1, 0, 1, \dots,$ starting at $L_n^\star(0) = \ell_n(s_0)$ and conditioned to end at $\ell_n(s_0)$; see [41], Section 3.3. It follows from Propositions 3.1(3) and 3.5(3) that $n^{-1/2}\ell_n(s_0) \longrightarrow^{\mathrm{p}} 0$ as $n \to \infty$. The jump distribution has mean 0 and variance $\sum_{k=-1}^{\infty} k^2\omega(k) = 2$. The result now follows by a conditional version of Donsker's invariance theorem; see, for example, [10], Lemma 10, for a detailed proof. □

LEMMA 4.4. *Let* $f_n, g_n : A_n \to [0, \Delta_n]$ *be random functions, for some* (*possibly random*) *set* A_n. *If*

$$\sup_{x\in A_n} n^{-1}\left|f_n(x) - g_n(x)\right| \xrightarrow{\mathrm{p}} 0 \tag{4.6}$$

then

$$n^{-1/2} \sup_{x\in A_n} \left|L_n^\star\big(f_n(x)\big) - L_n^\star\big(g_n(x)\big)\right| \xrightarrow[n\to\infty]{\mathrm{p}} 0. \tag{4.7}$$

PROOF. By the triangle inequality,

$$\begin{aligned}
&\big(2(1-\kappa)n\big)^{-1/2} \sup_{x\in A_n} \big|L_n^\star\big(f_n(x)\big) - L_n^\star\big(g_n(x)\big)\big| \\
&\quad\le \sup_{x\in A_n} \big|\mathbf{b}\big(f_n(x)/\Delta_n\big) - \mathbf{b}\big(g_n(x)/\Delta_n\big)\big| \\
&\qquad + \sup_{x\in A_n} \big|\big(2(1-\kappa)n\big)^{-1/2} L_n^\star\big(f_n(x)\big) - \mathbf{b}\big(f_n(x)/\Delta_n\big)\big| \\
&\qquad + \sup_{x\in A_n} \big|\big(2(1-\kappa)n\big)^{-1/2} L_n^\star\big(g_n(x)\big) - \mathbf{b}\big(g_n(x)/\Delta_n\big)\big|.
\end{aligned}$$

The first term converges to zero in probability by (4.6) and the fact that $\mathbf{b}$ is continuous on [0, 1], and hence uniformly continuous. The other terms converge to zero by Lemma 4.3, assuming as we may (by Skorohod's representation theorem) that (4.4) holds a.s. □

LEMMA 4.5. *As $n\to\infty$,*

$$n^{-1/2} \sup_{0\le i\le \Delta_n-1} \sup_{v\in\tau_{n,i}} \big|\ell_n(v) - \ell_n(s_i)\big| \xrightarrow{\mathrm{p}} 0. \tag{4.8}$$

PROOF. Write the left-hand side as $n^{-1/2}K$. Choose $\delta>0$ with $1-\delta > 1/(2\wedge(\beta-1))$, and choose $p>2/\delta$. We condition on τ_n and obtain, by using Lemma 2.1 for each subtree $\tau_{n,i}$ separately,

$$\begin{aligned}
\mathbb{E}\big(K^p|\tau_n\big) &\le \sum_{i=0}^{\Delta_n-1} \mathbb{E} \sup_{v\in\tau_{n,i}} \big|\ell_n(v)-\ell_n(s_i)\big|^p \le \sum_{i=0}^{\Delta_n-1} C(p)\big(N_{n,i}^\circ\big)^{p/2} \\
&\le C(p)n \sup_{0\le i<\Delta_n} \big(N_{n,i}^\circ\big)^{p/2}.
\end{aligned}\tag{4.9}$$

Then, by Propositions 3.1(3), (5) and 3.5(3),

$$\sup_{0\le i<\Delta_n} N_{n,i}^\circ/n^{1-\delta} \xrightarrow{\mathrm{p}} 0, \tag{4.10}$$

and thus, with probability tending to 1 as $n\to\infty$,

$$\sup_{0\le i<\Delta_n} N_{n,i}^\circ \le n^{1-\delta}. \tag{4.11}$$

If τ_n is such that (4.11) holds then (4.9), along with Markov's inequality, implies that for any $\varepsilon>0$,

$$\begin{aligned}
\mathbb{P}\big(K>\varepsilon n^{1/2}|\tau_n\big) &\le \varepsilon^{-p} n^{-p/2} C(p) n^{1+(1-\delta)p/2} \\
&= \varepsilon^{-p} C(p) n^{1-\delta p/2} \to 0.
\end{aligned}\tag{4.12}$$

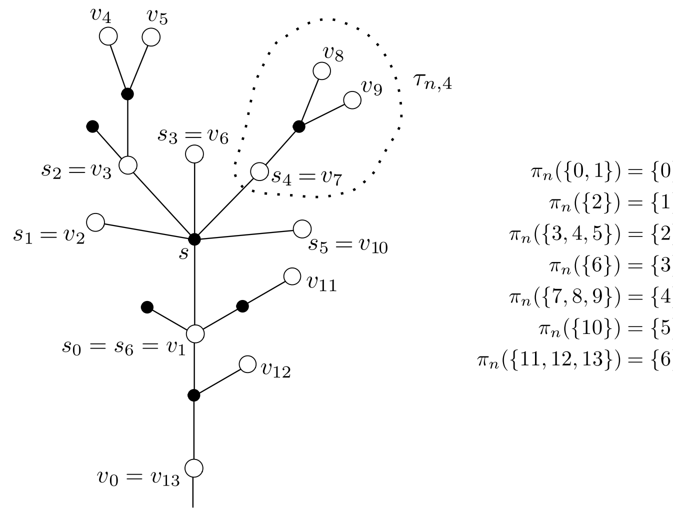

FIG. 4. *An example of the mapping π_n.*

Hence, $\mathbb{P}(K > \varepsilon n^{1/2}) \to 0$, as asserted. $\square$

PROOF OF THEOREM 4.1. To unify the treatment of the cases (1.6) and (1.7), we define $p_0 = 1$ for the weights in (1.7). By Lemma 4.3, it is sufficient to show that

$$n^{-1/2} \sup_{0 \le x \le N_n^\circ} \left| L_n^\star\left(x \frac{\Delta_n}{N_n^\circ}\right) - L_n(x) \right| \xrightarrow[n\to\infty]{\mathrm{p}} 0. \tag{4.13}$$

Note that L_n is a linear interpolation of its values on the integers. Using the triangle inequality (and Lemma 4.4) therefore allows us to restrict to integer values of x which we will write as k. Introduce the mapping $\pi_n : \{0, 1, \dots, N_n^\circ\} \to \{0, 1, \dots, \Delta_n\}$ defined as follows; see Figure 4: Let $\pi_n(0) = 0$ and $\pi_n(N_n^\circ) = \Delta_n$. If $v_i \in \tau_{n,j}$ for $j = 1, \dots, \Delta_n - 1$ then $\pi_n(i) = j$. If $v_0 < v_i \le s_0$ in the lexicographic ordering then $\pi_n(i) = 0$ and if $v_i \in \tau_{n,0}$ with $v_i > s_0$ then $\pi_n(i) = \Delta_n$. By the triangle inequality,

$$\begin{aligned} & n^{-1/2} \sup_{0 \le k \le N_n^\circ} \left| L_n^\star\left(k \frac{\Delta_n}{N_n^\circ}\right) - L_n(k) \right| \\ & \quad \le n^{-1/2} \sup_{0 \le k \le N_n^\circ} \left| L_n^\star\left(k \frac{\Delta_n}{N_n^\circ}\right) - L_n^\star(\pi_n(k)) \right| \\ & \qquad + n^{-1/2} \sup_{0 \le k \le N_n^\circ} \left| L_n(k) - L_n^\star(\pi_n(k)) \right|. \end{aligned} \tag{4.14}$$

We begin by showing that the first term on the right-hand side of (4.14) converges to 0 in probability. By Lemma 4.4, it suffices to show that

$$n^{-1} \sup_{0 \le k \le N_n^\circ} \left| k \frac{\Delta_n}{N_n^\circ} - \pi_n(k) \right| \xrightarrow{\mathrm{p}} 0. \tag{4.15}$$

We have the estimate

$$\sum_{i=1}^{\pi_n(k)-1} N_{n,i}^\circ \le k \le \sum_{i=0}^{\pi_n(k)} N_{n,i}^\circ. \tag{4.16}$$

Thus,

$$\left| k - \sum_{i=1}^{\pi_n(k)} N_{n,i}^\circ \right| \le N_{n,0}^\circ + N_{n,\pi_n(k)}^\circ \tag{4.17}$$

and hence, using Propositions 3.1(3), (5) and

$$n^{-1} \sup_{0 \le k \le N_n^\circ} \left| k - \sum_{i=1}^{\pi_n(k)} N_{n,i}^\circ \right| \xrightarrow{\mathrm{p}} 0. \tag{4.18}$$

Furthermore, in view of Propositions 3.1(1), (2) and 3.5(1), (2), $\Delta_n / N_n^\circ \xrightarrow{\mathrm{p}} (1-\kappa)/p_0$. It thus suffices to show that

$$\sup_{0 \le l \le \Delta_n} n^{-1} \left| \frac{1-\kappa}{p_0} \sum_{i=1}^{l} N_{n,i}^\circ - l \right| \xrightarrow[n\to\infty]{\mathrm{p}} 0, \tag{4.19}$$

which indeed follows from Propositions 3.1(4) and 3.5(1).

Next, consider the second term on the right-hand side of (4.14). This is exactly the left-hand side in Lemma 4.5, and thus it to tends to 0. □

5. Proof of Theorem 1.1. We start by recalling standard results on the Gromov–Hausdorff distance. A correspondence $\mathcal{R}$ between two metric spaces (E_1, d_1) and (E_2, d_2) is a subset of $E_1 \times E_2$ such that for every $x_1 \in E_1$ there exists an $x_2 \in E_2$ such that $(x_1, x_2) \in \mathcal{R}$ and vice versa. Denote the set of all correspondences between E_1 and E_2 by $\mathcal{C}(E_1, E_2)$. A distortion of a correspondence is defined as

$$\operatorname{dis}(\mathcal{R}) = \sup\{ |d_1(x_1, y_1) - d_2(x_2, y_2)| : (x_1, x_2), (y_1, y_2) \in \mathcal{R}\}. \tag{5.1}$$

The pointed Gromov–Hausdorff distance between (E_1, d_1) and (E_2, d_2) with marked points ρ_1 and ρ_2, respectively, can be conveniently expressed as, see [18], Theorem 7.3.25 (for the nonpointed version, the pointed version used here is similar)

$$d_{\mathrm{GH}}(E_1, E_2) = \frac{1}{2} \inf_{\mathcal{R} \in \mathcal{C}(E_1, E_2), (\rho_1, \rho_2) \in \mathcal{R}} \operatorname{dis}(\mathcal{R}). \tag{5.2}$$

In the proof of Theorem 1.1, we use similar ideas as in the previous section. Let M_n be a random planar map with a corresponding mobile $\theta_n = (\tau_n, \ell_n)$. As before, we denote the white vertices in θ_n by $v_0, \dots, v_{N_n^\circ}$ in lexicographical order and use the same notation for the corresponding white vertices in M_n. Also define the vertex s and its surrounding vertices $s_0, \dots, s_{\Delta_N}$ as before. Denote by $\theta_n^\star = (\tau_n^\star, \ell_n^\star)$ the mobile which is obtained by trimming θ_n such that it only consists of the black vertex s and its surrounding white vertices s_i, $0 \le i \le N_n^\circ$, and keeping the labels of these vertices the same as before. We add a superscript $\star$ to the notation when we consider these vertices as vertices in $\theta_n^\star$. Take $s_0^\star$ to be the root of $\theta_n^\star$. Note that if L_n is the label process corresponding to θ_n then $L_n^\star$, defined in Section 4, is the label process corresponding to $\theta_n^\star$. By definition, it holds that $\ell_n^\star(s_i^\star) = \ell_n(s_i)$ for all $0 \le i \le \Delta_n$. In general, the root $s_0^\star$ of $\theta_n^\star$ has a label different from zero, but note that the BDG bijection still works since it only depends on the increments of the labels in the white contour sequence. The planar map obtained from $\theta_n^\star$ is denoted by $M_n^\star$, the graph distance on $M_n^\star$ by $d_n^\star$ and the marked vertex by $\rho_n^\star$.

The planar map $M_n^\star$ has a single black vertex and has therefore a single face. Hence, it contains no cycles and is thus a planar tree with Δ_n edges. Given Δ_n, the map $M_n^\star$ is a uniformly distributed rooted planar tree and, given $M_n^\star$, the marked vertex $\rho_n^\star$ is chosen uniformly at random. (Note that the root edge of $M_n^\star$ yields both a root vertex and an ordering of the children of the root, and conversely; we may take the first child to be the other endpoint of the root edge.) Aldous [3] proved that the contour function of such a random rooted tree, after rescaling, converges in distribution to $\mathbf{e}$, which implies convergence of the tree to $\mathcal{T}_{\mathbf{e}}$ in Gromov–Hausdorff distance; see [38], Theorem 2.5. Hence we obtain, including also the marked vertex, the following.

THEOREM 5.1. *For the weights* (1.6) *and* (1.7), *the random planar map* $((M_n^\star, \rho_n^\star), (2(1-\kappa)n)^{-1/2} d_n^\star)$ *viewed as an element of* $\mathbb{M}^*$ *converges in distribution toward* $((\mathcal{T}_{\mathbf{e}}, \rho^*), \delta_{\mathbf{e}})$ *where given* $\mathcal{T}_{\mathbf{e}}$, ρ^* *is a marked vertex chosen uniformly at random from* $\mathcal{T}_{\mathbf{e}}$.

To complete the proof of Theorem 1.1, we construct the following correspondence between $((M_n, \rho_n), (2(1-\kappa)n)^{-1/2} d_n)$ and $((M_n^\star, \rho_n^\star), (2(1-\kappa)n)^{-1/2} d_n^\star)$:

$$\mathcal{R}_n = \{(\rho_n, \rho_n^\star)\} \cup \bigcup_{i=0}^{N_n^\circ - 1} \{(v_i, s_{\pi_n(i)}^\star)\} \tag{5.3}$$

with π_n the same as in the proof of Theorem 4.1. We then show that the distortion of this correspondence converges to zero in probability. Recall the definition of $\tau_{n,i}$ in Section 3. We have the following estimate.

LEMMA 5.2. *For any mobile* $\theta_n = (\tau_n, \ell_n)$ *it holds that*

$$\operatorname{dis}(\mathcal{R}_n) \le (2(1-\kappa)n)^{-1/2} \Big(14 \sup_{0 \le i \le \Delta_n - 1} \sup_{v \in \tau_{n,i}} |\ell_n(v) - \ell_n(s_i)| + 4 \Big). \tag{5.4}$$

By Lemma 4.5, the right-hand side tends to 0 in probability, which along with Theorem 5.1 completes the proof of Theorem 1.1. We conclude by proving Lemma 5.2.

PROOF OF LEMMA 5.2. Let $(x, x^\star), (y, y^\star) \in \mathcal{R}_n$. Write

$$K = \sup_{0 \le i \le \Delta_n - 1} \sup_{v \in \tau_{n,i}} \left|\ell_n(v) - \ell_n(s_i)\right|.$$

When we refer to ancestral relations in $M_n^\star$ we use $\rho_n^\star$ as the reference point, that is, we say that y is an ancestor of x in $M_n^\star$ if $x \neq y$ and the unique geodesic from x to $\rho_n^\star$ contains y. Consider separately the following three cases:

(1) $x^\star = y^\star$.
(2) $y^\star$ is an ancestor of $x^\star$ in $M_n^\star$, or conversely.
(3) $x^\star \neq y^\star$ and neither is an ancestor of the other.

Begin by studying case (1). The case $x^\star = y^\star = \rho_n^\star$ is trivial so we consider $x^\star \neq \rho_n^\star$. We can then write $x^\star = y^\star = s_i^\star$ for some i which will be fixed in this part of the proof. Let λ_0 denote the minimum label in $\tau_{n,i}$. For a vertex $v \in V(M_n)$ define the successor geodesic $\gamma(v)$ from v to ρ_n by $(v, \sigma(v), \sigma \circ \sigma(v), \dots, \rho_n)$ with σ defined in (2.3). Then there is a vertex w with label $\ell_n(w) = \lambda_0 - 1$ such that $\gamma(x)$ and $\gamma(y)$ contain w. Therefore, it follows from (2.4) and the definition of K that

$$\begin{aligned} d_n(x, w) &= \ell_n(x) - \lambda_0 + 1 \le 2K + 1 \quad \text{and} \\ d_n(y, w) &= \ell_n(y) - \lambda_0 + 1 \le 2K + 1. \end{aligned}$$

Thus, by the triangle inequality,

$$\left|d_n(x, y) - d_n^\star(x^\star, y^\star)\right| = d_n(x, y) \le d_n(x, w) + d_n(y, w) \le 4K + 2. \tag{5.5}$$

Next, consider case (2) and assume that $y^\star$ is an ancestor of $x^\star$. First, assume that $y^\star \neq \rho_n^\star$. Then there are unique i and j such that $x^\star = s_i^\star$, $y^\star = s_j^\star$ and without loss of generality we assume that $i < j$ (otherwise we shift the indices i and j modulo Δ_n). In this part, i and j are fixed. It holds that

$$\ell_n(s_m) > \ell_n(s_j) \tag{5.6}$$

for all m obeying $i \le m < j$. Let γ_i be a successor geodesic from s_i to ρ_n and let γ_j be a successor geodesic in M_n from s_j to ρ_n. We will show that the distance between γ_i and s_j is small (in terms of K). Define

$$\lambda_1 = \min\left\{\ell_n(v_m) : i \le \pi_n(m) < j\right\}. \tag{5.7}$$

It clearly holds that

$$\lambda_1 \le \ell_n(s_{j-1}) \le \ell_n(s_j) + 1. \tag{5.8}$$

Let l be an index for which the minimum in (5.7) is attained, that is, such that $\ell_n(v_l) = \lambda_1$ and $i \le \pi_n(l) < j$. Then, by (5.6),

$$\lambda_1 = \ell_n(v_l) \ge \ell_n(s_{\pi_n(l)}) - K \ge \ell_n(s_j) + 1 - K. \tag{5.9}$$

Now, γ_i and γ_j intersect for the first time at a vertex with label $\lambda_1 - 1$, and call this vertex z; see Figure 5. Then by (2.4) and (5.9)

$$d_n(z, s_j) = \ell_n(s_j) - \ell_n(z) \le K. \tag{5.10}$$

Furthermore, with the same argument leading to (5.5)

$$d_n(x, s_i) \le 3K + 2 \quad \text{and} \quad d_n(y, s_j) \le 3K + 2. \tag{5.11}$$

Finally, we get by repeatedly using the triangle inequality along with (2.4), (5.10) and (5.11)

$$\begin{aligned}
\bigl|d_n(x, y) - d_n^\star(x^\star, y^\star)\bigr| &\le \bigl|d_n(x, y) - d_n(s_i, s_j)\bigr| + \bigl|d_n(s_i, s_j) - d_n(z, s_i)\bigr| \\
&\quad + \bigl|d_n(z, s_i) - d_n^\star(x^\star, y^\star)\bigr| \\
&\le d_n(x, s_i) + d_n(y, s_j) + d_n(z, s_j) \\
&\quad + \bigl|\ell_n(s_i) - \ell_n(z) - \ell_n^\star(s_i^\star) + \ell_n^\star(s_j^\star)\bigr| \\
&\le 7K + 4 + \bigl|\ell_n(s_j) - \ell_n(z)\bigr| \le 8K + 4.
\end{aligned} \tag{5.12}$$

The case $y^\star = \rho_n^\star$ is treated in a simpler way leading to a similar upper bound as in (5.12); we omit the details.

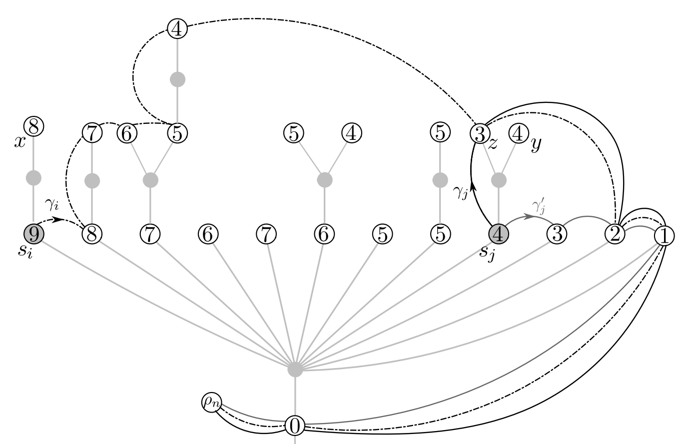

FIG. 5. *Illustration of the setup in part* (2) *of the proof with* $\ell_n(s_j) = 4$ *and* $\lambda_1 = 4$. *A planar mobile is shown with the edges and the black vertices colored light gray. The geodesic* γ_i *is black with dots and dashes and* γ_j *is black and solid. Since they are successor geodesics, they intersect at the vertex* z *with label* $\lambda_1 - 1 = 3$. *Another geodesic* γ'_j *from* s_j *to* ρ_n (*not a successor geodesic*) *is shown in dark gray and it does not intersect* γ_i *at a vertex with label* $\lambda_1 - 1$.

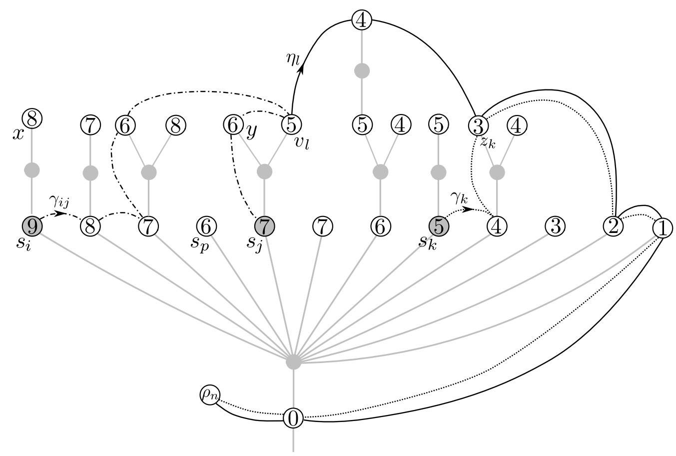

FIG. 6. *Illustration of the setup in part* (3) *of the proof with* $\ell_n(s_k) = 5$, $\lambda_2 = 5$ *and* $\lambda_3 = 4$. *A planar mobile is shown with the edges and the black vertices colored light gray. The geodesic* γ_{ij} *is dotted and dashed*, η_l *is solid and* γ_k *is dotted. Since* η_l *and* γ_k *are successor geodesics*, *they intersect at the vertex* z_k *with label* $\lambda_3 - 1 = 3$.

Finally, consider case (3) (see Figure 6 for an illustration). We keep writing $x^\star = s_i^\star$ and $y^\star = s_j^\star$. Denote the common ancestor of $x^\star$ and $y^\star$ having the largest label by $z^\star$. Assume that $z^\star \neq \rho_n^\star$ and write $z^\star = s_k^\star$; the case $z^\star = \rho_n^\star$ is treated in a similar but simpler way. We may assume without loss of generality that $i < j < k$ (by shifting the indices modulo Δ_n and possibly renaming x and y). In this part, i, j and k are fixed. Define the geodesics γ_i and γ_j as in case (2) and let γ_k be the successor geodesic from s_k to ρ_n. Furthermore, let γ_{ij} be a geodesic directed from s_i to s_j in M_n. Since $s_k^\star$ is an ancestor of both $x^\star$ and $y^\star$, it follows from (5.10) that there is a vertex z_i in γ_i and a vertex z_j in γ_j such that

$$d_n(z_m, s_k) = \ell_n(s_k) - \ell_n(z_m) \le K \qquad \text{for } m = i, j. \tag{5.13}$$

Moreover,

$$\ell_n(s_m) > \ell_n(s_k) \tag{5.14}$$

for all m obeying $i \le m < k$. We now show that γ_{ij} is also close to s_k. Define

$$\lambda_2 = \min\{\ell_n(v_m) : v_m \in \gamma_{ij}, i \le \pi_n(m) < k\}, \tag{5.15}$$

where by $v_m \in \gamma_{ij}$ we mean that v_m is visited by γ_{ij}. Condition (3) guarantees that there is an index, say p, such that $i \le p < j$ and $\ell_n(s_p) = \ell_n(s_k) + 1$. Let q be the first time at which $s_p < \gamma_{ij}(q) \le s_k$ in the lexicographic order on τ_n. Then q is well defined since γ_{ij} ends at s_j. If $\ell_n(\gamma_{ij}(q)) = \ell_n(\gamma_{ij}(q-1)) + 1$ then by the properties of the BDG bijection $\ell_n(\gamma_{ij}(q)) \le \ell_n(s_k)$. On the other hand, if $\ell_n(\gamma_{ij}(q)) = \ell_n(\gamma_{ij}(q-1)) - 1$ then by the same arguments $\ell_n(\gamma_{ij}(q-1)) \le \ell_n(s_p)$ which again yields $\ell_n(\gamma_{ij}(q)) \le \ell_n(s_k)$. We have thus established that

$$\lambda_2 \le \ell_n(s_k). \tag{5.16}$$

Let l be an index for which the minimum in (5.15) is attained, that is, such that $v_l \in \gamma_{ij}$, $i \le \pi_n(l) < k$ and $\ell_n(v_l) = \lambda_2$. By (5.14),

$$\lambda_2 = \ell_n(v_l) \ge \ell_n(s_{\pi_n(l)}) - K \ge \ell_n(s_k) + 1 - K. \tag{5.17}$$

Denote the successor geodesic from v_l to ρ_n by η_l. Next, define

$$\lambda_3 = \min\{\ell_n(v_m) : \pi_n(l) \le \pi_n(m) < k\}. \tag{5.18}$$

Now, η_l and γ_k intersect for the first time at a vertex having label $\lambda_3 - 1$, and call this vertex z_k. With same argument as in (5.9),

$$\lambda_3 \ge \ell_n(s_k) + 1 - K \tag{5.19}$$

and this yields, along with (2.4) and (5.16)

$$d_n(v_l, z_k) = \ell_n(v_l) - \ell_n(z_k) = \lambda_2 - \lambda_3 + 1 \le K. \tag{5.20}$$

Also, by (2.4) and (5.19)

$$d_n(s_k, z_k) = \ell_n(s_k) - \ell_n(z_k) \le K. \tag{5.21}$$

Using the triangle inequality along with (5.20) and (5.21), we get

$$d_n(v_\ell, s_k) \le d_n(v_l, z_k) + d_n(z_k, s_k) \le 2K. \tag{5.22}$$

Finally, we obtain by using the triangle inequality, (2.4), (5.11), (5.13) and (5.22)

$$\begin{aligned}
&\left|d_n(x, y) - d_n^\star(x^\star, y^\star)\right| \\
&\quad \le \left|d_n(x, y) - d_n(s_i, s_j)\right| \\
&\qquad + \left|d_n(s_i, v_l) - d_n(s_i, s_k)\right| + \left|d_n(s_i, s_k) - d_n(s_i, z_i)\right| \\
&\qquad + \left|d_n(s_j, v_l) - d_n(s_j, s_k)\right| + \left|d_n(s_j, s_k) - d_n(s_j, z_j)\right| \\
&\qquad + \left|d_n(s_i, z_i) + d_n(s_j, z_j) - d_n^\star(x^\star, y^\star)\right| \\
&\quad \le d_n(x, s_i) + d_n(y, s_j) + 2d_n(s_k, v_l) + d_n(z_i, s_k) + d_n(z_j, s_k) \\
&\qquad + \left|\ell_n(s_i) - \ell_n(z_i) + \ell_n(s_j) - \ell_n(z_j) - \ell_n^\star(s_i^\star) - \ell_n^\star(s_j^\star) + 2\ell_n^\star(s_k^\star)\right| \\
&\quad \le 12K + 4 + \left|\ell_n(s_k) - \ell_n(z_i)\right| + \left|\ell_n(s_k) - \ell_n(z_j)\right| \le 14K + 4. \qquad \square
\end{aligned}$$

6. Conclusions. We have shown that the random planar maps defined by the weights (1.6) and (1.7) converge to Aldous' Brownian tree. It is interesting to note that there does not seem to be a nontrivial scaling limit of the corresponding simply generated trees; see [36], Theorem 6, and thus the labels play a crucial role in obtaining a scaling limit for the random maps.

One can also study the so-called local limit of the planar maps M_n under consideration in this paper. The limit, when it exists, is an infinite graph M and convergence toward M roughly means that one considers all finite neighborhoods of faces around the root edge and shows that the probability that they appear in the maps M_n converges, as $n \to \infty$, to the probability that they appear in M. Angel and Schramm [4] studied local convergence in the case of uniformly distributed triangulations (all faces have degree 3) and later Durhuus and Chassaing [19] and Krikun [37] studied the case of uniformly distributed quadrangulations (all faces have degree 4). Recently, there have been several new results on the local limit of uniform quadrangulations concerning, for example, properties of infinite geodesics [21], random walks [7] and quadrangulations with a boundary [22]. In a forthcoming paper [15], it is shown that the local limit M of the maps M_n distributed by (1.2) exists for all choices of weights q_i. The proof involves using the bijection $\mathcal{G}_n$ introduced in the current paper along with theorems on local convergence of simply generated trees which we now briefly review.

The local limit of the simply generated trees corresponding to the weights (1.6) and (1.7) was established in [32] (with an asymptotically constant slowly varying function) and [30], respectively. Later it was established in full generality [covering cases (C1) and (C2)] in [29]. In case (C2), the local limit is deterministic and equals the infinite star, that is, the root has a single neighbor of infinite degree and all its neighbors are leaves. Therefore, the local limit M of the corresponding planar maps is simply the infinite uniform planar tree. In case (C1), the local limit of the trees is more complicated. It still has a unique vertex of infinite degree but the outgrowths from this vertex are now i.i.d. subcritical Galton–Watson trees. Therefore, the local limit M of the corresponding maps is not a tree. However, since subcritical Galton–Watson trees tend to be small, it is interesting to see how different M is from the uniform tree. It is, for example, interesting to study properties of random walks on M since random walks are sensitive to the presence of loops. In [15], it is shown (under some moment conditions on the weights w_i) that the spectral dimension of M, a number which characterizes the rate of decay of the return probability of the random walk, equals $4/3$ which is indeed the same value as for the uniform infinite planar tree.

A natural question to ask is how universal our results are, that is, is it enough to pose the conditions (C1) or (C2) in the Introduction or does one have to go to special cases? It is shown in [29], Examples 19.37–19.39, that by choosing irregular weights, still satisfying (C1) or (C2), the corresponding simply generated trees with n edges can have more than one vertex with a degree of the order of n; it is even possible that the large vertices have degrees $o(n)$ and that their number goes to infinity as $n \to \infty$ (at least along subsequences). In the case when there are two vertices with degrees of the order of n, it is plausible that the planar maps have a scaling limit which is roughly the Brownian tree with two points identified, forming a second macroscopic face. The more there is of large vertices in the simply generated trees the more faces should appear in the scaling limit of the maps. Thus, we conjecture that the Brownian tree only appears in special cases of (C1) and (C2). We consider, for simplicity, only one simple example (similar to [29], Example 19.38) illustrating this.

EXAMPLE 6.1. Let $(w_i)_{i\geq 0}$ be a weight sequence such that $w_i = 0$ unless $i \in \{0, 3^j : j \geq 0\}$. Further, let $w_0 = 1$ and let w_{3^j} increase so rapidly that (C2) holds, and moreover, with probability tending to 1 as $k \to \infty$, if $n = 3^k$, then the simply generated random tree τ_n with the distribution ν_n given by (3.1) is a star, while if $n = 2 \cdot 3^k$, then τ_n has two vertices of outdegree $n/2 = 3^k$ (and all other vertices are leaves).

For the subsequence $n = 3^k$, we then obtain the same results as above in the case (1.7).

For the subsequence $n = 2 \cdot 3^k$, the corresponding coloured tree distributed by $\tilde{\nu}$ has (with probability tending to 1) two black vertices of degrees $n/2$ connected by a single white vertex $\hat{v}$ of degree 2, and each of them joined to $n/2 - 1$ white leaves. For each choice of labels ℓ_n, the corresponding map M_n thus has two faces. The label processes around each black vertex converge to independent Brownian bridges, which together with the random choice of root implies that, in analogy to Theorem 4.1,

$$\left(\frac{1}{\sqrt{n}} L_n(tN_n^\circ)\right)_{0\leq t\leq 1} \xrightarrow[n\to\infty]{\mathrm{d}} \big(\mathbf{h}(t)\big)_{0\leq t\leq 1}, \tag{6.1}$$

where, for two Brownian bridges $\mathbf{b}_1, \mathbf{b}_2$ and U uniformly distributed on [0, 1], all independent,

$$\mathbf{h}(t) = \begin{cases} \mathbf{b}_1(2t + U) - \mathbf{b}_1(U), & 0 \leq t \leq (1 - U)/2, \\ \mathbf{b}_2(2t + U - 1) - \mathbf{b}_1(U), & (1 - U)/2 \leq t \leq 1 - U/2, \\ \mathbf{b}_1(2t + U - 2) - \mathbf{b}_1(U), & 1 - U/2 \leq t \leq 1. \end{cases} \tag{6.2}$$

Moreover, the label process visits $\hat{v}$, the unique white vertex of degree 2, twice. If we split this vertex into two, the corresponding map will be a tree, which after normalization converges in distribution in the Gromov–Hausdorff metric to a random real tree $\mathcal{T}_{\mathbf{h}'}$, where $\mathbf{h}'$ is the random function h above shifted to its minimum and with the minimum subtracted, so $\mathbf{h}' \geq 0$ and $\mathbf{h}'(0) = 0$. We may by the Skorohod representation theorem assume that the label processes converge a.s. Then the random maps with $\hat{v}$ split converge to $\mathcal{T}_{\mathbf{h}'}$ a.s. in the Gromov–Hausdorff metric, with the two halves of $\hat{v}$ corresponding to two different points in $\mathcal{T}_{\mathbf{h}'}$ [the points given by $t = (1 - U)/2$ and $t = 1 - U/2$], and it follows by combining the two parts of $\hat{v}$ again, that the random maps M_n converge to a limit that equals $\mathcal{T}_{\mathbf{h}'}$ with these two points identified. Note that this creates a cycle, so the limit is no longer a tree. (As a topological space, it is of the same homotopy type as a circle.)

APPENDIX: MORE ON GALTON–WATSON TREES

Marckert and Miermont [44] gave a description of the distribution $\tilde{\nu}$ in (2.9) as a conditioned two-type Galton–Watson tree, while we have used the bijection

$\mathcal{G}_n$ in Section 3 to obtain a simply generated tree (which in many cases is a conditioned Galton–Watson tree), with a single type only. In this appendix, we give some further comments on the relation between these two approaches.

Consider arbitrary weights $q_i \geq 0$, $i \geq 1$, assuming first only that $q_i > 0$ for some $i > 1$ (to avoid trivialities), and define w_i by (1.4) (and $w_0 = 1$) and their generating function $g(x)$ by (1.5). Marckert and Miermont [44] define another generating function $f(x)$ (denoted $f_{\mathbf{q}}(x)$ in [44]) by

$$f(x) = \sum_{k=0}^{\infty} w_{k+1} x^k; \tag{A.1}$$

thus

$$g(x) = 1 + x f(x). \tag{A.2}$$

We have seen in Sections 1 and 3 that a random planar map in $\mathcal{M}_n^*$ with Boltzmann weights (1.1) corresponds to a random mobile (τ_n, ℓ_n) (and a sign ε that we ignore here), and that τ_n corresponds by the bijection $\mathcal{G}_n$ to a random tree τ_n' that has the distribution of a simply generated tree with $|\tau_n'| = n$ edges, defined by the weights $(w_i)_{i\geq 0}$, cf. (3.1) [and note that $\deg(v) - 1$ is the outdegree, i.e., the number of children of v; see Section 2].

We consider first trees with unrestricted number of edges. We give a planar tree τ the weight

$$w(\tau) = \prod_{v\in V(\tau)} w_{\deg(v)-1}. \tag{A.3}$$

The generating function

$$G(x) = \sum_{\tau} x^{|\tau|+1} w(\tau) \tag{A.4}$$

summing over all planar trees, satisfies the well-known equation [48]

$$G(x) = x g\bigl(G(x)\bigr). \tag{A.5}$$

In particular, the total weight $Z = \sum_\tau w(\tau) = G(1)$ is finite if and only if the equation

$$z = g(z) \tag{A.6}$$

has a solution $z \in (0, \infty)$, and then Z is the smallest positive solution to (A.6). Using (A.2), we can write (A.6) as $z = 1 + z f(z)$, or

$$f(z) = 1 - 1/z, \tag{A.7}$$

the form of the equation used in [44].

If $Z = G(1) < \infty$ (such weights q_i are called *admissible* in [44]), define

$$p_i = w_i Z^{i-1}. \tag{A.8}$$

Then, by (A.6),

$$\sum_{i=0}^{\infty} p_i = Z^{-1} g(Z) = 1, \tag{A.9}$$

so $(p_i)_{i\geq 0}$ is a probability distribution on $\{0, 1, \ldots\}$. Let τ' be a random Galton–Watson tree with this offspring distribution. Then the probability of a particular realization τ' is

$$\prod_{v\in V(\tau')} p_{\deg(v)-1} = Z^{\sum_v(\deg(v)-2)} \prod_{v\in V(\tau')} w_{\deg(v)-1} = Z^{-1} w(\tau'), \tag{A.10}$$

recalling that the number of vertices in τ' is $|\tau'|+1$ and that $\sum_v \deg(v) = 2|\tau'|+1$ since we count an extra half-edge at the root. Hence, the distribution of the Galton–Watson tree τ' equals the distribution given by the weights $w(\tau)$ on the set of all planar trees.

REMARK A.1. The distribution $(p_i)_{i\geq 0}$ defined by (A.8) is not the same as the $(p_i)_{i\geq 0}$ used in Section 3, so they define different Galton–Watson trees τ'; however, they yield the same distribution ν_n when conditioned on a fixed size n of the tree.

Since $Z = \sum_\tau w(\tau)$, the sum of the probabilities (A.10) over all (finite) τ is 1; thus the Galton–Watson tree τ' is a.s. finite, which means that the offspring distribution has mean ≤ 1, that is, the Galton–Watson tree is subcritical or critical. Conversely, we can obtain any subcritical or critical probability distribution $(p_i)_{i\geq 0}$ by taking $w_i = p_0^{i-1} p_i$; then $w_0 = 1$ and $Z = p_0^{-1}$. (If we do not insist on $w_0 = 1$, we can simply take $w_i = p_i$.)

The offspring distribution (A.8) has probability generating function

$$g_{\mathbf{p}}(x) = \sum_{i=0}^{\infty} p_i x^i = Z^{-1} g(Zx) \tag{A.11}$$

and thus mean

$$g'_{\mathbf{p}}(1) = g'(Z), \tag{A.12}$$

which by (A.2) and (A.7) can be written as

$$g'_{\mathbf{p}}(1) = f(Z) + Zf'(Z) = 1 + \bigl(Z^2 f'(Z) - 1\bigr)/Z. \tag{A.13}$$

Hence, the Galton–Watson tree is critical if and only if $g'(Z) = 1$ or, equivalently, $Z^2 f'(Z) = 1$ (the form used in [44]). Moreover, the variance of the offspring distribution is

$$\sigma^2 = g''_{\mathbf{p}}(1) + g'_{\mathbf{p}}(1) - \bigl(g'_{\mathbf{p}}(1)\bigr)^2 = Zg''(Z) + g'(Z)\bigl(1 - g'(Z)\bigr), \tag{A.14}$$

which in the critical case $g'(Z) = 1$ can be written by (A.2) as

$$\sigma^2 = Zg''(Z) = Z\big(Zf''(Z) + 2f'(Z)\big) = \big(Z^3 f''(Z) + 2\big)/Z, \tag{A.15}$$

which in the notation of [44] is $\rho_{\mathbf{q}}/Z_{\mathbf{q}}$.

The two-type Galton–Watson tree defined by Marckert and Miermont [44], which we denote by τ, has a white root; a white vertex has only black children, and the number of them has the geometric distribution $\mathrm{Ge}(p_0) = ((1 - Z^{-1})^i Z^{-1})_{i \ge 0}$; a black vertex has only white children, and the number of them has the distribution $(p_{i+1}/(1 - p_0))_{i\ge 0} = (p_{i+1}/(1 - Z^{-1}))_{i \ge 0}$, that is, the conditional distribution of $(\xi - 1 | \xi > 0)$ if ξ has the distribution $(p_i)_{i\ge 0}$. Thus, the offspring distribution for the black vertices has the probability generating function

$$\sum_{i=0}^{\infty} \frac{p_{i+1}}{1 - Z^{-1}} x^i = \sum_{i=0}^{\infty} \frac{w_{i+1} Z^i x^i}{1 - Z^{-1}} = \frac{f(Zx)}{1 - Z^{-1}} = \frac{g(Zx) - 1}{(Z - 1)x}. \tag{A.16}$$

A simple calculation (which essentially is [44], Proposition 7) shows that the bijection in Section 3 maps this two-type Galton–Watson tree τ to the standard (single type) Galton–Watson tree τ' with offspring distribution (A.8). This can also be seen from the construction of the bijection; see Figure 3. In particular, note that the children of the root in τ are the vertices in the rightmost path from the root in τ', excluding its final leaf (and similarly for the children of other white vertices); this explains why the offspring distribution for a white vertex is geometric, since the length of the rightmost path in τ' obviously has a geometric distribution.

Restricting to trees with n edges (and thus $n + 1$ vertices) we see, by Remark A.1, that the tree τ_n' in Section 3 with distribution ν_n can be seen as τ' conditioned on $|\tau'| = n$, and thus the corresponding tree $\tau_n = \mathcal{G}_n^{-1}(\tau_n')$ has the same distribution as τ conditioned on $|\tau| = n$.

Although the Galton–Watson tree τ' is simpler than the two-type tree τ, the latter is more convenient for some purposes. For example, when considering the white vertices, as we do in parts of Section 3, it is immediate (by considering each second generation) that the number of white vertices in τ' is distributed as the total progeny (number of vertices) in a Galton–Watson tree with offspring distribution

$$\xi^{(0)} = \sum_{j=1}^{\zeta} \xi_j^*, \tag{A.17}$$

where $\zeta \sim \mathrm{Ge}(p_0) = \mathrm{Ge}(1 - Z^{-1})$ and $\xi_j^* = (\xi_j - 1 | \xi_j > 0)$ are independent of each other and of ζ, and each ξ_j has the distribution $(p_i)_{i \ge 0}$. We have, letting $\kappa = \mathbb{E}\xi = \sum_i i p_i \le 1$,

$$\mathbb{E}\xi_i^* = \mathbb{E}(\xi_i | \xi_i > 0) - 1 = \frac{\mathbb{E}\xi_i}{1 - p_0} - 1 = \frac{\kappa + p_0 - 1}{1 - p_0} \tag{A.18}$$

and

$$\mathbb{E}\xi^{(0)} = \mathbb{E}\zeta\mathbb{E}\xi_1^* = \frac{1-p_0}{p_0}\frac{\kappa+p_0-1}{1-p_0} = \frac{\kappa+p_0-1}{p_0} = 1-\frac{1-\kappa}{p_0}. \tag{A.19}$$

Furthermore, it is easy to see that $\xi^{(0)}$ has the probability generating function

$$\mathbb{E}x^{\xi^{(0)}} = \frac{p_0}{1-\sum_{k=1}^{\infty} p_k x^{k-1}}. \tag{A.20}$$

Note that $\mathbb{E}\xi^{(0)} < 1$ when $\mathbb{E}\xi < 1$, which says that the white tree consisting of each second generation in τ is subcritical if and only if τ' (or τ) is.

Translated to τ', this shows immediately that the number of leaves of the Galton–Watson tree τ' with offspring distribution ξ is distributed as the total progeny of a Galton–Watson process with offspring distribution $\xi^{(0)}$. In fact, this was shown by Minami [47]; one version of his argument is the following. Given a tree τ, we partition its vertex set into *twigs* as follows: Take the vertices in lexicographic order and stop each time we reach a leaf, that is, the first twig consists of the root and all vertices up to, and including, the first leaf; the second twig starts at the next vertex and ends at the next leaf, and so on. Thus, each twig ends with a leaf, and the number of twigs equals the number of leaves. If we start with a random Galton–Watson tree τ' with offspring distribution (p_i), the size of each twig has a geometric distribution $1+\zeta$ with $\zeta \sim \mathrm{Ge}(p_0)$ as above. Moreover, each nonleaf in the twig has further offspring distributed as ξ^*; hence, if we contract each twig to a single vertex, we obtain a new random Galton–Watson tree with offspring distributed as $\xi^{(0)}$; the number of vertices in this tree equals the number of twigs in τ', and thus the number of leaves in τ'.

In fact, these two arguments are essentially the same; if we use instead the reverse lexicographic order when defining the twigs, it is easy to see that each twig in τ' correspond to a white vertex and its (black) children in τ.

We use this representation to verify the tail estimate (3.17).

LEMMA A.2. *Let $N^{(0)}$ be the number of leaves in a Galton–Watson tree with offspring distribution $(p_i)_{i\ge 0}$ satisfying $\kappa < 1$ and (3.3) for some slowly varying function $\bar{L}(i)$. Then, as $n \to \infty$,*

$$\mathbb{P}(N^{(0)} = n) \sim c\bar{L}(n)n^{-\beta}, \tag{A.21}$$

with $c = {p_0}^{\beta-1}(1-\kappa)^{-\beta}$.

PROOF. We have seen that $N^{(0)}$ is distributed as the number of vertices in a Galton–Watson tree with offspring distribution (A.17). By (3.12) applied to a sequence $\xi_j^{(0)}$ of independent copies of $\xi^{(0)}$,

$$\mathbb{P}(N^{(0)} = n) = \frac{1}{n}\mathbb{P}(S_n^{(0)} = n-1), \tag{A.22}$$

where

$$S_n^{(0)} = \sum_{j=1}^{n} \xi_j^{(0)} \overset{\mathrm{d}}{=} \sum_{i=1}^{X_n} \xi_i^*, \tag{A.23}$$

where $X_n = \sum_{j=1}^n \zeta_j$ with $\zeta_j \sim \mathrm{Ge}(p_0)$ independent of each other and of $\{\xi_i^*\}$. [Thus, X_n has a negative binomial distribution $\mathrm{NegBin}(n, p_0)$.] Note that $\mathbb{E}X_n = n\mathbb{E}\zeta_1 = n(1-p_0)/p_0$, and that X_n is strongly concentrated about its mean; for example, moment convergence in the central limit theorem for X_n implies that

$$\mathbb{P}\left(\left|X_n - \frac{1-p_0}{p_0}n\right| > n^{2/3}\right) = O(n^{-b}) \tag{A.24}$$

for any fixed b. Furthermore,

$$\mathbb{P}(\xi_i^* = n) = (1-p_0)^{-1} p_{n+1} = (1+o(1))(1-p_0)^{-1}\bar{L}(n)n^{-\beta} \tag{A.25}$$

as $n \to \infty$, and thus, by a more general version of (3.13) applied to ξ_i^* and (A.18), uniformly for all k with $|k - n(1-p_0)/p_0| \le n^{2/3}$,

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{k} \xi_i^* = n-1\right) &= k(1+o(1))\mathbb{P}(\xi_1^* = \lfloor n - k\mathbb{E}\xi_1^* - 1\rfloor) \\ &= (1+o(1))\frac{n(1-p_0)}{p_0}\mathbb{P}(\xi_1^* = \lfloor n(1-\kappa)/p_0\rfloor + o(n)) \\ &= (1+o(1))\frac{n}{p_0}\bar{L}(n)(n(1-\kappa)/p_0)^{-\beta}. \end{aligned} \tag{A.26}$$

Choose $b = \beta + 1$. By (A.22)–(A.26),

$$\begin{aligned} \mathbb{P}(N^{(0)} = n) &= \frac{1}{n}\mathbb{P}\left(\sum_{i=1}^{X_n} \xi_i^* = n-1\right) \\ &= (1+o(1)){p_0}^{-1}\bar{L}(n)(n(1-\kappa)/p_0)^{-\beta}. \end{aligned}$$

□

Acknowledgements. S. Ö. Stefánsson would like to thank Jérémie Bouttier for telling him about [41] and for suggesting to relate the condensation phase in simply generated trees to planar maps. We would also like to thank Jakob Björnberg for discussions and comments.

REFERENCES

[1] ALBENQUE, M. and MARCKERT, J.-F. (2008). Some families of increasing planar maps. *Electron. J. Probab.* **13** 1624–1671. MR2438817

[2] ALDOUS, D. (1991). The continuum random tree. II. An overview. In *Stochastic Analysis* (*Durham*, 1990). *London Mathematical Society Lecture Note Series* **167** 23–70. Cambridge Univ. Press, Cambridge. MR1166406

[3] ALDOUS, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226

[4] ANGEL, O. and SCHRAMM, O. (2003). Uniform infinite planar triangulations. *Comm. Math. Phys.* **241** 191–213. MR2013797

[5] ARMENDÁRIZ, I. and LOULAKIS, M. (2011). Conditional distribution of heavy tailed random variables on large deviations of their sum. *Stochastic Process. Appl.* **121** 1138–1147. MR2775110

[6] ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, New York. MR0373040

[7] BENJAMINI, I. and CURIEN, N. (2013). Simple random walk on the uniform infinite planar quadrangulation: Subdiffusivity via pioneer points. *Geom. Funct. Anal.* **23** 501–531. MR3053754

[8] BERTOIN, J. (1996). *Lévy Processes. Cambridge Tracts in Mathematics* **121**. Cambridge Univ. Press, Cambridge. MR1406564

[9] BETTINELLI, J. (2015). Scaling limit of random planar quadrangulations with a boundary. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 432–477. MR3335010

[10] BETTINELLI, J. (2010). Scaling limits for random quadrangulations of positive genus. *Electron. J. Probab.* **15** 1594–1644. MR2735376

[11] BIALAS, P. and BURDA, Z. (1996). Phase transition in fluctuating branched geometry. *Phys. Lett. B* **384** 75–80. MR1410422

[12] BIANE, P. (1986). Relations entre pont et excursion du mouvement brownien réel. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** 1–7. MR0838369

[13] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York. MR0233396

[14] BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge Univ. Press, Cambridge. MR0898871

[15] BJÖRNBERG, J. E. and STEFÁNSSON, S. Ö. (2014). Recurrence of bipartite planar maps. *Electron. J. Probab.* **19** 1–40.

[16] BOUTTIER, J., DI FRANCESCO, P. and GUITTER, E. (2004). Planar maps as labeled mobiles. *Electron. J. Combin.* **11** Research Paper 69, 27. MR2097335

[17] BOUTTIER, J. and GUITTER, E. (2009). Distance statistics in quadrangulations with a boundary, or with a self-avoiding loop. *J. Phys. A* **42** 465208, 44. MR2552016

[18] BURAGO, D., BURAGO, Y. and IVANOV, S. (2001). *A Course in Metric Geometry. Graduate Studies in Mathematics* **33**. AMS, Boston, MA.

[19] CHASSAING, P. and DURHUUS, B. (2006). Local limit of labeled trees and expected volume growth in a random quadrangulation. *Ann. Probab.* **34** 879–917. MR2243873

[20] CURIEN, N., HAAS, B. and KORTCHEMSKI, I. (2014). The CRT is the scaling limit of random dissections. *Random Structures Algorithms*. DOI:10.1002/rsa.20554.

[21] CURIEN, N., MÉNARD, L. and MIERMONT, G. (2013). A view from infinity of the uniform infinite planar quadrangulation. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** 45–88. MR3083919

[22] CURIEN, N. and MIERMONT, G. (2014). Uniform infinite planar quadrangulations with a boundary. *Random Structures Algorithms*. DOI:10.1002/rsa.20531.

[23] DENISOV, D., DIEKER, A. B. and SHNEER, V. (2008). Large deviations for random walks under subexponentiality: The big-jump domain. *Ann. Probab.* **36** 1946–1991. MR2440928

[24] DUQUESNE, T. and LE GALL, J.-F. (2005). Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields* **131** 553–603. MR2147221

[25] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, *Vol. II*, 2nd ed. Wiley, New York. MR0270403

[26] GROMOV, M. (1999). *Metric Structures for Riemannian and Non-Riemannian Spaces*. *Progress in Mathematics* **152**. Birkhäuser, Boston, MA. MR1699320

[27] GUT, A. (2005). *Probability*: *A Graduate Course*. Springer, New York. MR2125120

[28] HAAS, B. and MIERMONT, G. (2004). The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electron*. *J*. *Probab*. **9** 57–97 (electronic). MR2041829

[29] JANSON, S. (2012). Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probab*. *Surv*. **9** 103–252. MR2908619

[30] JANSON, S., JONSSON, T. and STEFÁNSSON, S. Ö. (2011). Random trees with superexponential branching weights. *J*. *Phys*. *A* **44** 485002, 16. MR2860856

[31] JANSON, S. and MARCKERT, J.-F. (2005). Convergence of discrete snakes. *J*. *Theoret*. *Probab*. **18** 615–647. MR2167644

[32] JONSSON, T. and STEFÁNSSON, S. Ö. (2011). Condensation in nongeneric trees. *J*. *Stat*. *Phys*. **142** 277–313. MR2764126

[33] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. Springer, New York. MR1876169

[34] KEMPERMAN, J. H. B. (1950). *The General One-Dimensional Random Walk with Absorbing Barriers with Applications to Sequential Analysis*. Excelsiors Foto-Offset, 's-Gravenhage. The Hague. MR0038045

[35] KEMPERMAN, J. H. B. (1961). *The Passage Problem for a Stationary Markov Chain*. Univ. Chicago Press, Chicago, IL. MR0119226

[36] KORTCHEMSKI, I. (2012). Limit theorems for conditioned non-generic Galton–Watson trees. Available at arXiv:1205.3145v2.

[37] KRIKUN, M. (2005). Local structure of random quadrangulations. Available at arXiv:math/0512304.

[38] LE GALL, J.-F. (2005). Random trees and applications. *Probab*. *Surv*. **2** 245–311. MR2203728

[39] LE GALL, J.-F. (2007). The topological structure of scaling limits of large planar maps. *Invent*. *Math*. **169** 621–670. MR2336042

[40] LE GALL, J.-F. (2013). Uniqueness and universality of the Brownian map. *Ann*. *Probab*. **41** 2880–2960. MR3112934

[41] LE GALL, J.-F. and MIERMONT, G. (2011). Scaling limits of random planar maps with large faces. *Ann*. *Probab*. **39** 1–69. MR2778796

[42] LE GALL, J.-F. and MIERMONT, G. (2012). Scaling limits of random trees and planar maps. In *Probability and Statistical Physics in Two and More Dimensions*. *Clay Math*. *Proc*. **15** 155–211. Amer. Math. Soc., Providence, RI. MR3025391

[43] LE GALL, J.-F. and PAULIN, F. (2008). Scaling limits of bipartite planar maps are homeomorphic to the 2-sphere. *Geom*. *Funct*. *Anal*. **18** 893–918. MR2438999

[44] MARCKERT, J.-F. and MIERMONT, G. (2007). Invariance principles for random bipartite planar maps. *Ann*. *Probab*. **35** 1642–1705. MR2349571

[45] MEIR, A. and MOON, J. W. (1978). On the altitude of nodes in random trees. *Canad*. *J*. *Math*. **30** 997–1015. MR0506256

[46] MIERMONT, G. (2013). The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Math*. **210** 319–401. MR3070569

[47] MINAMI, N. (2005). On the number of vertices with a given degree in a Galton–Watson tree. *Adv*. *in Appl*. *Probab*. **37** 229–264. MR2135161

[48] OTTER, R. (1949). The multiplicative process. *Ann*. *Math*. *Statist*. **20** 206–224. MR0030716

[49] PITMAN, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math*. **1875**. Springer, Berlin. MR2245368

[50] VERVAAT, W. (1979). A relation between Brownian bridge and Brownian excursion. *Ann*. *Probab*. **7** 143–149. MR0515820

[51] ZOLOTAREV, V. M. (1986). *One-Dimensional Stable Distributions*. *Translations of Mathematical Monographs* **65**. Amer. Math. Soc., Providence, RI. MR0854867

DEPARTMENT OF MATHEMATICS
UPPSALA UNIVERSITY
PO BOX 480
SE-751 06 UPPSALA
SWEDEN
E-MAIL: svante.janson@math.uu.se

DEPARTMENT OF MATHEMATICS
UPPSALA UNIVERSITY
PO BOX 480
SE-751 06 UPPSALA
SWEDEN
AND
NORDITA, THE NORDIC INSTITUTE
FOR THEORETICAL PHYSICS
KTH ROYAL INSTITUTE OF TECHNOLOGY
AND STOCKHOLM UNIVERSITY
ROSLAGSTULLSBACKEN 23
SE 106 91 STOCKHOLM
SWEDEN
E-MAIL: rudrugis@gmail.com