

THE HARMONIC MEASURE OF BALLS IN RANDOM TREES

BY NICOLAS CURIEN AND JEAN-FRANÇOIS LE GALL

Université Paris-Sud

We study properties of the harmonic measure of balls in typical large discrete trees. For a ball of radius n centered at the root, we prove that, although the size of the boundary is of order n , most of the harmonic measure is supported on a boundary set of size approximately equal to n^β , where $\beta \approx 0.78$ is a universal constant. To derive such results, we interpret harmonic measure as the exit distribution of the ball by simple random walk on the tree, and we first deal with the case of critical Galton–Watson trees conditioned to have height greater than n . An important ingredient of our approach is the analogous continuous model (related to Aldous’ continuum random tree), where the dimension of harmonic measure of a level set of the tree is equal to β , whereas the dimension of the level set itself is equal to 1. The constant β is expressed in terms of the asymptotic distribution of the conductance of large critical Galton–Watson trees.

1. Introduction. The main goal of this work is to study properties of the harmonic measure of balls in large discrete trees. From a probabilistic point of view, the harmonic measure of a set is the exit distribution of that set by random walk, in the discrete setting, or by Brownian motion, in the continuous setting. Harmonic measure has been studied in depth both in harmonic analysis and in probability theory, and it would be hopeless to try to survey the literature on this subject. It has been observed in different contexts that the harmonic measure of a set with a fractal-like boundary is often supported on a subset of the boundary of strictly smaller dimension. For example, the famous Makarov theorem [30] states that harmonic measure on the boundary of a simply connected planar domain is always supported on a subset of Hausdorff dimension equal to 1, regardless of the dimension of the boundary (see [21] for similar results in a discrete setting and [8] for higher dimensional analogs). This “dimension drop” phenomenon also appears in the context of (infinite) discrete random trees. In [27], Lyons, Pemantle and Peres studied the harmonic measure at infinity for simple random walk on

AMS 2000 subject classifications: Primary 05C81,31C05,60J45; secondary 05C80,60J80

Keywords and phrases: harmonic measure, Brownian motion, random walk, random tree, Galton-Watson tree, Hausdorff dimension, conductance

an infinite supercritical Galton–Watson tree and proved that the harmonic measure is supported on a boundary set of dimension strictly less than the dimension of the whole boundary. The same authors then extended this result to biased random walk on a supercritical Galton–Watson tree [28].

In the present work, we study a similar phenomenon in the context of finite discrete trees. Our results apply to several combinatorial classes of discrete trees, such as plane trees, binary trees or Cayley trees in particular. For a typical tree with a (fixed) large size chosen in any of these classes, we obtain that the harmonic measure of a ball of radius n is supported, up to a small mass, on a subset of about n^β vertices, despite the fact that the boundary of the ball has of order n vertices. Here $\beta \approx 0.78$ is a universal constant that does not depend on the combinatorial class.

In order to obtain these results for “combinatorial trees”, we interpret them as conditioned Galton–Watson trees. Recall that a Galton–Watson tree describes the genealogy of a population starting with an ancestor or root, where each individual has, independently of the others, a number of children distributed according to a given offspring distribution (see Section 4.1 for a precise definition). We first study harmonic measure on generation n of a *critical* Galton–Watson tree, whose offspring distribution has mean 1 and finite variance, and which is conditioned to have height greater than n . In this setting, we obtain that most of the harmonic measure on generation n is concentrated on a set of approximately n^β vertices, with high probability. Again, this should be contrasted with the fact that the generation n of the tree has about n vertices. The constant β has an explicit expression in terms of the law of a random variable \mathcal{C} , which is the limit in distribution of the (scaled) conductance of the tree between the root and generation n – again this limiting distribution does not depend on the offspring distribution. In the related continuous model, we show that the Hausdorff dimension of the harmonic measure is almost surely equal to β , whereas the dimension of the boundary is known to be equal to 1. Let us describe our results in a more precise way.

Discrete setting. Let θ be a probability measure on \mathbb{Z}_+ , and assume that θ has mean one and finite variance $\sigma^2 > 0$. Under the probability \mathbb{P} , for every integer $n \geq 0$, we let $\mathbb{T}^{(n)}$ be a Galton–Watson tree with offspring distribution θ , conditioned on non-extinction at generation n . Conditionally on the tree $\mathbb{T}^{(n)}$, we then consider simple random walk on $\mathbb{T}^{(n)}$, starting from the root, and we let Σ_n be the first hitting point of generation n by random walk. The harmonic measure μ_n is the law of Σ_n . Notice that μ_n is a random probability measure supported on the set $\mathbb{T}_n^{(n)}$ of all vertices

of $\mathsf{T}^{(n)}$ at generation n . By a classical theorem of the theory of branching processes, $n^{-1} \#\mathsf{T}_n^{(n)}$ converges in distribution to an exponential distribution with parameter $2/\sigma^2$.

THEOREM 1. *There exists a constant $\beta \in (0, 1)$, which does not depend on the offspring distribution θ , such that, for every $\delta > 0$, we have the convergence in \mathbb{P} -probability*

$$\mu_n(\{v \in \mathsf{T}_n^{(n)} : n^{-\beta-\delta} \leq \mu_n(v) \leq n^{-\beta+\delta}\}) \xrightarrow[n \rightarrow \infty]{(\mathbb{P})} 1.$$

Consequently, for every $\varepsilon \in (0, 1)$, there exists, with \mathbb{P} -probability tending to 1 as $n \rightarrow \infty$, a subset $A_{n,\varepsilon}$ of $\mathsf{T}_n^{(n)}$ such that $\#A_{n,\varepsilon} \leq n^{\beta+\delta}$ and $\mu_n(A_{n,\varepsilon}) \geq 1 - \varepsilon$. Conversely, the maximal μ_n -measure of a set of cardinality bounded by $n^{\beta-\delta}$ tends to 0 as $n \rightarrow \infty$, in \mathbb{P} -probability.

Although we have no exact numerical expression for β , calculations using the formulas in Proposition 4 below indicate that $\beta \approx 0.78$. See the discussion at the end of Section 3.4. This approximate numerical value confirms simulations made in physics [18].

The last two assertions of the theorem are easy consequences of the first one. Indeed, $A_{n,\varepsilon} := \{v \in \mathsf{T}_n^{(n)} : \mu_n(v) \geq n^{-\beta-\delta}\}$ has cardinality smaller than $n^{\beta+\delta}$, and the first assertion of the theorem shows that the μ_n -measure of the latter set is greater than $1 - \varepsilon$ with \mathbb{P} -probability tending to 1 as $n \rightarrow \infty$. On the other hand, if A is any subset of $\mathsf{T}_n^{(n)}$ with cardinality smaller than $n^{\beta-\delta}$, we have

$$\mu_n(A) \leq \mu_n(\{v \in \mathsf{T}_n^{(n)} : \mu_n(v) > n^{-\beta+\delta/2}\}) + n^{\beta-\delta} n^{-\beta+\delta/2}$$

and the first term in the right-hand side tends to 0 in \mathbb{P} -probability by the first assertion of the theorem.

Theorem 1 implies a similar result for Galton-Watson trees conditioned to have a fixed size. For every integer $N \geq 0$ such that this makes sense, let $\mathbf{T}(N)$ be distributed under the probability measure \mathbb{P} as a Galton-Watson tree with offspring distribution θ conditioned to have N edges. For every integer $n \geq 1$, let $\mathbf{T}_n(N)$ be the set of all vertices of $\mathbf{T}(N)$ at generation n . The harmonic measure μ_n^N is defined on the event $\{\mathbf{T}_n(N) \neq \emptyset\}$ as the hitting distribution of $\mathbf{T}_n(N)$ by simple random walk on $\mathbf{T}(N)$ started from the root.

COROLLARY 2. *Let $\delta > 0$ and $\varepsilon > 0$. Then,*

$$\mathbb{P}\left(\left\{\mu_n^N(\{v \in \mathbf{T}_n(N) : \mu_n^N(v) \notin [n^{-\beta-\delta}, n^{-\beta+\delta}]\}) > \varepsilon\right\} \cap \left\{\mathbf{T}_n(N) \neq \emptyset\right\}\right) \xrightarrow{n, N \rightarrow \infty} 0.$$

As in Theorem 1, this implies that, with high probability on the event $\{\mathbf{T}_n(N) \neq \emptyset\}$, the harmonic measure μ_n^N is supported, up to a mass less than ε , on a set of $n^{\beta+\delta}$ vertices, and conversely the maximal μ_n^N -measure of a set of cardinality bounded above by $n^{\beta-\delta}$ is small.

If $h(\mathbf{T}(N))$ denotes the height (maximal distance from the root) of the tree $\mathbf{T}(N)$, it is well known that $N^{-1/2}h(\mathbf{T}(N))$ converges in distribution to a positive random variable, see (47) below. Therefore, if we let n and N tend to infinity in such a way that $n = o(\sqrt{N})$, the probability $\mathbb{P}(\mathbf{T}_n(N) \neq \emptyset)$ tends to 1. It is worth pointing that Corollary 2 applies to balls of radius n which is large but small in comparison with the diameter of the tree – a similar extension would in fact hold also for Theorem 1.

For specific choices of θ , the tree $\mathbf{T}(N)$ is uniformly distributed over certain classes of combinatorial trees, and Corollary 2 yields the results that were mentioned earlier in this introduction. In particular, if θ is the geometric distribution $\theta(k) = 2^{-k-1}$, $\mathbf{T}(N)$ is uniformly distributed over plane trees with N edges. If θ is the Poisson distribution with mean 1, and if we assign labels $1, \dots, N+1$ to vertices in a random manner and then “forget” the ordering of $\mathbf{T}(N)$, we get a random tree uniformly distributed over Cayley trees on $N+1$ vertices. In a similar manner, for every integer $p \geq 2$, we can handle p -ary trees (where the number of children of every vertex belongs to $\{0, 1, \dots, p\}$) or strictly p -ary trees (where each vertex has 0 or p children).

Continuous setting. A key ingredient of the proof of Theorem 1 is a similar result in the continuous setting. A critical Galton–Watson tree conditioned on having height greater than n , viewed as a metric space for the graph distance normalized by the factor n^{-1} , is close in the Gromov–Hausdorff sense to a variant of Aldous’ Brownian continuum random tree [2], also called the CRT. So a continuous analog of the harmonic measure μ_n would be the hitting distribution of height 1 by Brownian motion on the CRT starting from the root. Although the construction of Brownian motion on the CRT has been carried out in [20] (see also [10] for a simpler approach, and [4] for a general construction of Brownian motion on \mathbb{R} -trees), we will not follow this approach, because there is a simpler way of looking at the continuous setting.

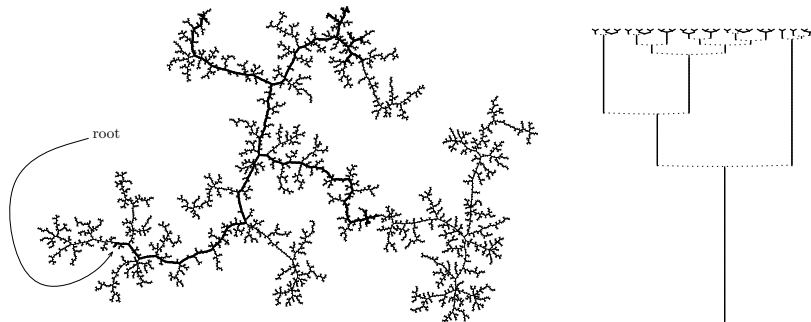


FIG 1. A large (binary) Galton-Watson tree and the reduced tree at a given level.

The point is that properties of the harmonic measure μ_n on $\mathbb{T}_n^{(n)}$ can be read from the reduced tree \mathbb{T}^{*n} that consists only of vertices of $\mathbb{T}^{(n)}$ that have descendants at generation n . In other words, we can chop off the branches of the discrete tree that do not reach the level n . Indeed, a simple argument shows that the hitting distribution of generation n is the same for simple random walk on $\mathbb{T}^{(n)}$ and on the reduced tree \mathbb{T}^{*n} . See Figure 1 for a simulation of a large Galton-Watson tree and the associated reduced tree.

The scaling limit of the discrete reduced trees \mathbb{T}^{*n} (when distances are scaled by the factor n^{-1}) is particularly simple. We define a random compact \mathbb{R} -tree by the following device. We start from an (oriented) line segment whose length U_\emptyset is uniformly distributed over $[0, 1]$ and whose origin will serve as the root of our tree. At the other end of this initial line segment, we attach the initial point of two other line segments with respective lengths U_1 and U_2 such that, conditionally given U_\emptyset , U_1 and U_2 are independent and uniformly distributed over $[0, 1 - U_\emptyset]$. At the other end of the first of these segments, respectively of the second one, we attach two line segments whose lengths are again independent and uniformly distributed over $[0, 1 - U_\emptyset - U_1]$, resp. over $[0, 1 - U_\emptyset - U_2]$, conditionally on the triplet (U_\emptyset, U_1, U_2) . We continue the construction by induction and after an infinite number of steps we get a random (non-compact) rooted \mathbb{R} -tree, whose completion is denoted by Δ . This is the scaling limit of the discrete reduced trees \mathbb{T}^{*n} . See Section 2.1 for a more precise construction.

The metric on Δ is denoted by \mathbf{d} . By definition, the boundary $\partial\Delta$ consists of all points of Δ at height 1, that is, at distance 1 from the root: These are exactly the points that are added when taking the completion in the preceding construction.

It is then easy to define Brownian motion on Δ starting from the root and up to the first hitting time of $\partial\Delta$ (it would be possible to extend the

definition of Brownian motion beyond the first hitting time of $\partial\Delta$, but this is not relevant to our purposes). Roughly speaking, this process behaves like linear Brownian motion as long as it stays on an “open interval” of the tree. It is reflected at the root of the tree and when it arrives at a branching point, it chooses each of the three possible line segments incident to this point with equal probabilities. The harmonic measure μ is then the (quenched) distribution of the first hitting point of $\partial\Delta$ by Brownian motion (see Section 2.1 for details).

THEOREM 3. *With the same constant β as in Theorem 1, we have \mathbb{P} a.s. $\mu(dx)$ a.e.,*

$$\lim_{r \downarrow 0} \frac{\log \mu(\mathcal{B}_{\mathbf{d}}(x, r))}{\log r} = \beta,$$

where $\mathcal{B}_{\mathbf{d}}(x, r)$ stands for the closed ball of radius r centered at x in the metric space (Δ, \mathbf{d}) . Consequently, the Hausdorff dimension of μ is \mathbb{P} a.s. equal to β .

The fact that the second assertion of the theorem follows from the first one is standard. See e.g. Lemma 4.1 in [27].

Remark. It is not hard to prove that the Hausdorff dimension of $\partial\Delta$ (with respect to \mathbf{d}) is a.s. equal to 1. An exact Hausdorff measure function is given by Theorem 1.3 in Duquesne and Le Gall [12].

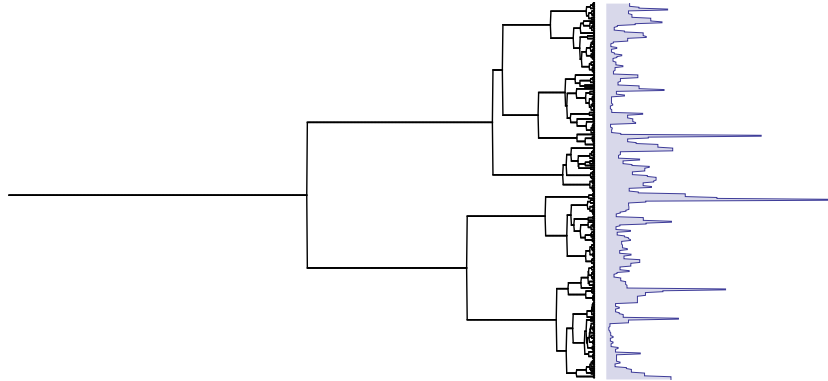


FIG 2. A simulation of the reduced tree Δ and the harmonic measure on its boundary. Clearly the measure is not uniformly spread and exhibits a fractal behavior.

Let us give some ideas of the proof of Theorem 3. It is well known that one can turn the tree Δ , or rather the subtree $\Delta \setminus \partial\Delta$, into a “stationary” object

via a logarithmic transformation. Roughly speaking, we introduce a new tree which has the same binary branching structure as Δ , such that each point of Δ at height $s \in [0, 1)$ corresponds to a point of the new tree at height $-\log(1-s) \in [0, \infty)$. The resulting non-compact tree is called the Yule tree because it describes the genealogy of the classical Yule process, where individuals have (independent) exponential lifetimes with parameter 1 and each individual has exactly two offspring. We define the boundary of the Yule tree as the collection of all its geodesic rays, where a geodesic ray is just a semi-infinite geodesic path starting from the root. This boundary is easily identified with $\partial\Delta$. An application of Itô's formula shows that the logarithmic transformation turns Brownian motion on Δ into a time-changed Brownian motion *with drift* $1/2$ toward infinity on the Yule tree. Consequently, the probability measure μ corresponds via the preceding transformation to the distribution ν of the geodesic ray that is “selected” by Brownian motion with drift $1/2$ (that is, the unique ray of the Yule tree that is visited by Brownian motion at arbitrarily large times). The first assertion of Theorem 3 is then equivalent to proving that, \mathbb{P} a.s., $\nu(dy)$ a.e.,

$$(1) \quad \lim_{r \rightarrow \infty} \frac{1}{r} \log \nu(\mathcal{B}(y, r)) = -\beta$$

where $\mathcal{B}(y, r)$ denotes the set of all geodesic rays of the Yule tree that coincide with y up to height r .

The next step is then to identify a kind of “stationary environment seen from the particle” for Brownian motion on the Yule tree. More precisely, we show in Section 3.1 that the law of the subtree above level $r \geq 0$ that is selected by Brownian motion (with drift $1/2$) converges as $r \rightarrow \infty$ to a limiting probability measure that we explicitly describe. This allows us to construct an ergodic invariant measure for the natural shifts on the space of all pairs consisting of a (deterministic) Yule-type tree and a distinguished geodesic ray on this tree, and moreover this measure is absolutely continuous with respect to the law of the random pair formed by the Yule tree and the ray selected by Brownian motion. The limiting result (1) then follows from an application of Birkhoff's ergodic theorem to a suitable functional. In this part of our work, we use several ideas that have been developed by Lyons, Pemantle and Peres [27] in a slightly different setting.

The random conductance. The constant β in Theorems 1 and 3 can be expressed in terms of the (continuous) conductance of Δ . Roughly speaking, if one considers Δ as a network of resistors with unit resistance per unit length, then the effective resistance between height 0 and height 1 is a random variable, which we denote by \mathcal{C} . With this interpretation, it is clear that $\mathcal{C} > 1$

a.s. Alternatively, \mathcal{C} is the mass under the Brownian excursion measure from the root of those excursion paths that hit height 1. Note that \mathcal{C} is also the limit in distribution of the (scaled) conductance between generations 0 and n in $\mathbb{T}^{(n)}$. The distribution of \mathcal{C} satisfies the following recursive equation

$$(2) \quad \mathcal{C} \stackrel{(d)}{=} \left(U + \frac{1-U}{\mathcal{C}_1 + \mathcal{C}_2} \right)^{-1},$$

where \mathcal{C}_1 and \mathcal{C}_2 are independent copies of \mathcal{C} , and U is uniformly distributed over $[0, 1]$ and independent of the pair $(\mathcal{C}_1, \mathcal{C}_2)$. Despite this rather simple recursive equation, the law $\gamma(ds)$ of \mathcal{C} is not completely understood (in particular, its mean is unknown). We prove that, although γ has a continuous density f over $[1, \infty)$, the function f is not twice continuously differentiable at the point 2 (and we expect a similar singular behavior at all integer values). See Figure 3.

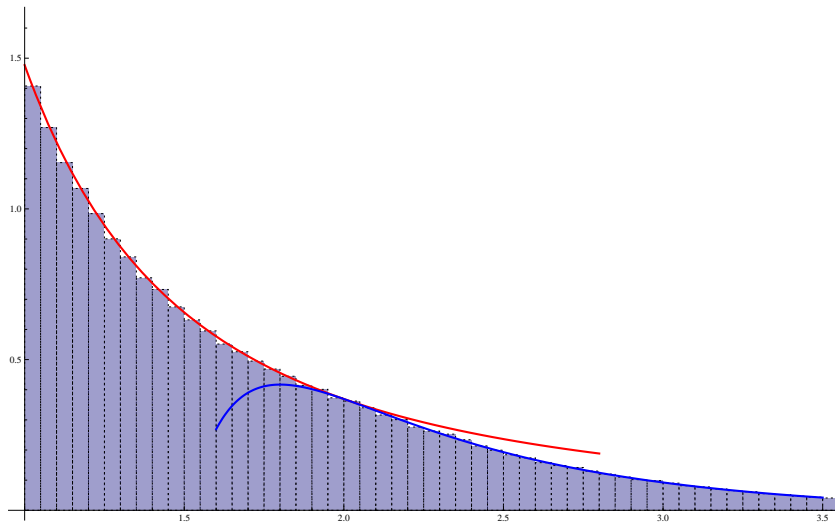


FIG 3. A histogram of the distribution of γ over $(1, \infty)$ from simulations based on the recursive equation (2). There are explicit formulas for the density of γ over $[1, 2]$ and over $[2, 3]$, which however depend on the (unknown) density at 1. The red and the blue curves correspond to these explicit formulas, with a numerical approximation of the density at 1.

In many respects, the distribution γ governs the behavior of harmonic measure. In particular the constant β has an explicit expression in terms of γ .

PROPOSITION 4. *The distribution γ is characterized in the class of all probability measures on $[1, \infty)$ by the distributional equation (2). The con-*

stant β appearing in Theorems 1 and 3 is given by

$$(3) \quad \beta = 2 \frac{\iiint \gamma(dr)\gamma(ds)\gamma(dt) \frac{rs}{r+s+t-1} \log\left(\frac{s+t}{s}\right)}{\iint \gamma(ds)\gamma(dt) \frac{st}{s+t-1}} = \frac{1}{2} \left(\frac{(\int \gamma(ds)s)^2}{\iint \gamma(ds)\gamma(dt) \frac{st}{s+t-1}} - 1 \right).$$

We finally mention that some extensions of the results of the present work have been obtained by Lin [24, 25]. A version of Theorem 1 for Galton–Watson trees whose offspring distribution belongs to the domain of attraction of a stable law with index $\alpha \in (1, 2)$ is derived in [24]. In the setting of the present work, the article [25] gives an analog of Theorem 1 for the harmonic measure of a vertex chosen according to the uniform probability measure on generation n of the tree $T^{(n)}$. This is another step toward a full multifractal analysis of the harmonic measure μ_n .

The paper is organized as follows. We start by studying the continuous model. In Section 2 we introduce the basic set-up and we relate the random tree Δ to the Yule tree. The law of the random conductance \mathcal{C} is studied in Section 2.3. Section 3 gathers the ingredients of the proof of Theorem 3. In particular, Section 3.2 identifies the limiting distribution of the subtree above level r selected by Brownian motion, and Section 3.3 explains the application of the ergodic theorem needed to derive (1). Section 4 is devoted to the proof of Theorem 1 and Corollary 2. Let us emphasize that Theorem 1 is not a straightforward consequence of Theorem 3, and that the proof of our results in the discrete setting requires a number of additional estimates, even though a key role is played by Theorem 3. The last section is devoted to a few complements. In particular, we comment on the connection between the present paper and the recent work of Aïdékon [1].

2. The continuous setting. In this section we give a formal definition of the (continuous) reduced tree Δ . We then explain the connection between the reduced tree and the Yule tree. We finally introduce and study the conductance of these trees, which plays a key role in the next sections.

2.1. *The reduced tree Δ .* We set

$$\mathcal{V} = \bigcup_{n=0}^{\infty} \{1, 2\}^n$$

where $\{1, 2\}^0 = \{\emptyset\}$. If $v = (v_1, \dots, v_n) \in \mathcal{V}$, we set $|v| = n$ (in particular, $|\emptyset| = 0$), and if $n \geq 1$, we define the parent of v as $\widehat{v} = (v_1, \dots, v_{n-1})$ (we then say that v is a child of \widehat{v}). If $v = (v_1, \dots, v_n)$ and $v' = (v'_1, \dots, v'_m)$ belong to \mathcal{V} , the concatenation of v and v' is $vv' := (v_1, \dots, v_n, v'_1, \dots, v'_m)$.

The notions of a descendant and an ancestor of an element of \mathcal{V} are defined in the obvious way, with the convention that a vertex $v \in \mathcal{V}$ is both an ancestor and a descendant of itself. If $v, w \in \mathcal{V}$, $v \wedge w$ is the unique element of \mathcal{V} that is an ancestor of both v and w and such that $|v \wedge w|$ is maximal.

We then consider a collection

$$(U_v)_{v \in \mathcal{V}}$$

of independent real random variables uniformly distributed over $[0, 1]$ under the probability measure \mathbb{P} . We set

$$Y_\emptyset = U_\emptyset$$

and then, by induction, for every $v \in \{1, 2\}^n$, with $n \geq 1$,

$$Y_v = Y_{\hat{v}} + U_v(1 - Y_{\hat{v}}).$$

Note that $0 \leq Y_v < 1$ for every $v \in \mathcal{V}$, a.s. Consider then the set

$$\Delta_0 := (\{\emptyset\} \times [0, Y_\emptyset]) \cup \left(\bigcup_{v \in \mathcal{V} \setminus \{\emptyset\}} \{v\} \times (Y_{\hat{v}}, Y_v) \right).$$

There is a straightforward way to define a metric \mathbf{d} on Δ_0 , so that (Δ_0, \mathbf{d}) is a (noncompact) \mathbb{R} -tree and, for every $x = (v, r) \in \Delta_0$, we have $\mathbf{d}((\emptyset, 0), x) = r$. To be specific, let $x = (v, r) \in \Delta_0$ and $y = (w, r') \in \Delta_0$:

- If v is a descendant of w or w is a descendant of v , we set $\mathbf{d}(x, y) = |r - r'|$.
- Otherwise, $\mathbf{d}(x, y) = \mathbf{d}((v \wedge w, Y_{v \wedge w}), x) + \mathbf{d}((v \wedge w, Y_{v \wedge w}), y) = (r - Y_{v \wedge w}) + (r' - Y_{v \wedge w})$.

See Figure 4 for an illustration of the tree Δ_0 .

We let Δ be the completion of Δ_0 with respect to the metric \mathbf{d} . Then

$$\Delta = \Delta_0 \cup \partial\Delta$$

where by definition $\partial\Delta = \{x \in \Delta : \mathbf{d}((\emptyset, 0), x) = 1\}$, which is canonically identified with $\{1, 2\}^{\mathbb{N}}$ (here and below, $\mathbb{N} = \{1, 2, \dots\}$ is the set of all positive integers). Note that (Δ, \mathbf{d}) is a compact \mathbb{R} -tree.

The point $(\emptyset, 0)$ is called the root of Δ . For every $x \in \Delta$, we set $H(x) = \mathbf{d}((\emptyset, 0), x)$ and call $H(x)$ the height of x . We can define a genealogical order on Δ by setting $x \prec y$ if and only if x belongs to the geodesic path from the root to y .

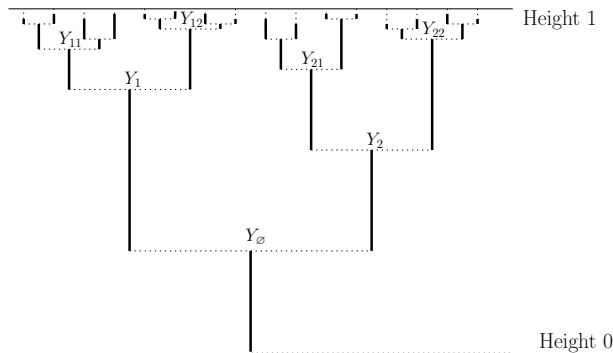


FIG 4. The random tree Δ_0

For every $\varepsilon \in (0, 1)$, we set

$$\Delta_\varepsilon = \{x \in \Delta : H(x) \leq 1 - \varepsilon\},$$

which is also a compact \mathbb{R} -tree for the metric \mathbf{d} . The leaves of Δ_ε are the points of the form $(v, 1 - \varepsilon)$ for all $v \in \mathcal{V}$ such that $Y_{\hat{v}} < 1 - \varepsilon \leq Y_v$. The branching points of Δ_ε are the points of the form (v, Y_v) for all $v \in \mathcal{V}$ such that $Y_v < 1 - \varepsilon$. We can then define Brownian motion on Δ_ε as a special case of a diffusion on a graph (see in particular [15], [13] and the references therein, and note that the definition of Brownian motion on Δ_ε can also be viewed as a very special case of the construction of Brownian motion on \mathbb{R} -trees given in [4]). Informally, this process behaves like linear Brownian motion as long as it stays on an “open interval” of the form $\{v\} \times (Y_{\hat{v}}, Y_v \wedge (1 - \varepsilon))$. It is reflected at the root $(\emptyset, 0)$ and at the leaves of Δ_ε , and when it arrives at a branching point of the tree, it chooses each of the three possible line segments ending at this point with equal probabilities.

Write $B^\varepsilon = (B_t^\varepsilon)_{t \geq 0}$ for Brownian motion on Δ_ε starting from the root, which is defined under the probability measure P (for our purposes, it will be important to carefully distinguish the probability measure \mathbb{P} governing the random trees and the one governing Brownian motions on these trees). We let

$$T_\varepsilon := \inf\{t \geq 0 : H(B_t^\varepsilon) = 1 - \varepsilon\},$$

be the hitting time of the set of all leaves of Δ_ε .

If we now set $\varepsilon_n = 2^{-n}$ for every $n \geq 1$, we may define all processes B^{ε_n} on the same probability space, in such a way that $B_{t \wedge T_{\varepsilon_m}}^{\varepsilon_n} = B_{t \wedge T_{\varepsilon_m}}^{\varepsilon_m}$ for every $t \geq 0$ and every choice of $m \leq n$, P a.s. Assuming that the latter property holds, we set

$$T = \lim_{n \uparrow \infty} \uparrow T_{\varepsilon_n}$$

and we define the process $(B_t)_{t \geq 0}$ by requiring that $B_t = \dagger$ if $t \geq T$ (where \dagger is a cemetery point) and, for every $n \geq 1$ and $t \geq 0$, $B_{t \wedge T_{\varepsilon_n}} = B_{t \wedge T_{\varepsilon_n}}^{\varepsilon_n}$. It is easy to verify that the left limit

$$B_{T-} = \lim_{t \uparrow T, t < T} B_t$$

exists in Δ and belongs to $\partial\Delta$, P a.s. The harmonic measure μ is the distribution of B_{T-} under P , which is a (random) probability measure on $\partial\Delta = \{1, 2\}^{\mathbb{N}}$.

2.2. *The Yule tree.* For the proof of Theorem 3, it will be more convenient to reformulate the problem in terms of Brownian motion on the Yule tree. To define the Yule tree, consider now a collection

$$(V_v)_{v \in \mathcal{V}}$$

of independent real random variables exponentially distributed with mean 1 under the probability measure \mathbb{P} . We set

$$\mathcal{Y}_{\emptyset} = V_{\emptyset}$$

and then by induction, for every $v \in \{1, 2\}^n$, with $n \geq 1$,

$$\mathcal{Y}_v = \mathcal{Y}_{\hat{v}} + V_v.$$

The Yule tree is the set

$$\Gamma := (\{\emptyset\} \times [0, \mathcal{Y}_{\emptyset}]) \cup \left(\bigcup_{v \in \mathcal{V} \setminus \{\emptyset\}} \{v\} \times (\mathcal{Y}_{\hat{v}}, \mathcal{Y}_v] \right),$$

which is equipped with the metric d defined in the same way as \mathbf{d} in the preceding section. For this metric, Γ is again a non-compact \mathbb{R} -tree. For every $x = (v, r) \in \Gamma$, we keep the notation $H(x) = r = d((\emptyset, 0), x)$ for the height of the point x .

Now observe that if U is uniformly distributed over $[0, 1]$, the random variable $-\log(1 - U)$ is exponentially distributed with mean 1. Hence we may and will suppose that the collection $(V_v)_{v \in \mathcal{V}}$ is constructed from the collection $(U_v)_{v \in \mathcal{V}}$ in the previous section via the formula $V_v = -\log(1 - U_v)$, for every $v \in \mathcal{V}$. Then, the mapping Ψ defined on Δ_0 by $\Psi(v, r) = (v, -\log(1 - r))$, for every $(v, r) \in \Delta_0$, is a homeomorphism from Δ_0 onto Γ .

Stochastic calculus shows that we can write, for every $t \in [0, T)$,

$$(4) \quad \Psi(B_t) = W\left(\int_0^t (1 - H(B_s))^{-2} ds\right)$$

where $(W(t))_{t \geq 0}$ is Brownian motion with constant drift $1/2$ toward infinity on the Yule tree (this process is defined in a similar way as Brownian motion on Δ_ε , except that it behaves like Brownian motion with drift $1/2$ on every “open interval” of the tree). Note that W is again defined under the probability measure P . From now on, when we speak about Brownian motion on the Yule tree or on other similar trees, we will always mean Brownian motion with drift $1/2$ toward infinity.

By definition, the boundary of Γ is the set of all infinite geodesics in Γ starting from the root $(\emptyset, 0)$ (these are called geodesic rays). The boundary of Γ is canonically identified with $\{1, 2\}^{\mathbb{N}}$. From the transience of Brownian motion on Γ , there is an a.s. unique geodesic ray denoted by W_∞ that is visited by $(W(t), t \geq 0)$ at arbitrarily large times. We sometimes say that W_∞ is the exit ray of Brownian motion on Γ . The distribution of W_∞ under P yields a probability measure ν on $\{1, 2\}^{\mathbb{N}}$. Thanks to (4), we have in fact $\nu = \mu$, provided we view both μ and ν as (random) probability measures on $\{1, 2\}^{\mathbb{N}}$. The statement of Theorem 3 is then reduced to checking that (1) holds $\nu(dy)$ a.e., \mathbb{P} a.s.

Yule-type trees. Our proof of (1) makes a heavy use of tools of ergodic theory applied to certain transformations on a space of trees that we now describe. We let \mathbb{T} be the set of all collections $(z_v)_{v \in \mathcal{V}}$ of nonnegative real numbers such that the following properties hold:

- (i) $z_{\hat{v}} < z_v$ for every $v \in \mathcal{V} \setminus \{\emptyset\}$;
- (ii) for every $\mathbf{v} = (v_1, v_2, \dots) \in \{1, 2\}^{\mathbb{N}}$,

$$\lim_{n \rightarrow \infty} z_{(v_1, \dots, v_n)} = +\infty.$$

Notice that we allow the possibility that $z_\emptyset = 0$. We equip \mathbb{T} with the σ -field generated by the coordinate mappings. If $(z_v)_{v \in \mathcal{V}} \in \mathbb{T}$, we can consider the associated “tree”

$$\mathcal{T} := (\{\emptyset\} \times [0, z_\emptyset]) \cup \left(\bigcup_{v \in \mathcal{V} \setminus \{\emptyset\}} \{v\} \times (z_{\hat{v}}, z_v] \right),$$

equipped with the distance defined as above. We will keep the notation $H(x) = r$ if $x = (v, r)$ for the height of a point $x \in \mathcal{T}$. The genealogical

order on \mathcal{T} is defined as previously and will again be denoted by \prec . If $\mathbf{u} = (u_1, u_2, \dots, u_n, \dots) \in \{1, 2\}^{\mathbb{N}}$, and $x = (v, r) \in \mathcal{T}$, we write $x \prec \mathbf{u}$ if $v = (u_1, u_2, \dots, u_k)$ for some integer $k \geq 0$.

We will often abuse notation and say that we consider a tree $\mathcal{T} \in \mathbb{T}$: This really means that we are given a collection $(z_v)_{v \in \mathcal{V}}$ satisfying the above properties, and we consider the associated tree \mathcal{T} . In particular \mathcal{T} has an order structure (in addition to the genealogical partial order) given by the lexicographical order on \mathcal{V} . Elements of \mathbb{T} will be called Yule-type trees.

Clearly, the Yule tree can be viewed as a random variable with values in \mathbb{T} , and we write $\Theta(d\mathcal{T})$ for its distribution.

Let us fix $\mathcal{T} \in \mathbb{T}$. If $r > 0$, the level set at height r is

$$\mathcal{T}_r = \{x \in \mathcal{T} : H(x) = r\}.$$

If $x \in \mathcal{T}_r$, we can consider the subtree $\mathcal{T}[x]$ of descendants of x in \mathcal{T} . Formally, we view $\mathcal{T}[x]$ as an element of \mathbb{T} : We write v_x for the unique element of \mathcal{V} such that $x = (v_x, r)$, and define $\mathcal{T}[x]$ as the Yule-type tree corresponding to the collection $(z_{v_x v} - r)_{v \in \mathcal{V}}$. Similarly, if $[[0, x]]$ denotes the geodesic segment between the root and x , we can define the subtrees of \mathcal{T} branching off $[[0, x]]$. To this end, let $n_x = |v_x|$ and let $v_{x,0} = \emptyset, v_{x,1}, \dots, v_{x,n_x} = v_x$ be the successive ancestors of v_x from generation 0 to generation n_x . Set $r_{x,i} = z_{v_{x,i-1}}$ for every $1 \leq i \leq n_x$. Then, for every $1 \leq i \leq n_x$, the i -th subtree branching off $[[0, x]]$, which is denoted by $\mathcal{T}_{x,i}$, corresponds to the collection

$$(z_{\tilde{v}_{x,i} v} - r_{x,i})_{v \in \mathcal{V}}$$

where $\tilde{v}_{x,i}$ is the child of $v_{x,i-1}$ that is not $v_{x,i}$. To simplify notation, we introduce the point measure

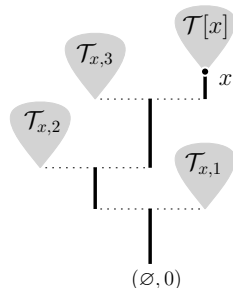
$$\xi_{r,x}(\mathcal{T}) = \sum_{i=1}^{n_x} \delta_{(r_{x,i}, \mathcal{T}_{x,i})},$$

which belongs to the set $\mathcal{M}_p(\mathbb{R}_+ \times \mathbb{T})$ of all finite point measures on $\mathbb{R}_+ \times \mathbb{T}$.

We now state a ‘‘spine’’ decomposition of the Yule tree, which plays an important role in our approach.

PROPOSITION 5 (Spine decomposition). *Let F be a nonnegative measurable function on \mathbb{T} , and let G be a nonnegative measurable function on $\mathcal{M}_p(\mathbb{R}_+ \times \mathbb{T})$. Let $r > 0$. Then,*

$$\mathbb{E} \left[\sum_{x \in \Gamma_r} F(\Gamma[x]) G(\xi_{r,x}(\Gamma)) \right] = e^r \mathbb{E}[F(\Gamma)] \times \mathbb{E}[G(\mathcal{N})],$$

FIG 5. *The spine decomposition*

where $\mathcal{N}(ds d\mathcal{T})$ is, under the probability measure \mathbb{P} , a Poisson point measure on $\mathcal{M}_p(\mathbb{R}_+ \times \mathbb{T})$ with intensity $2 \mathbf{1}_{[0,r]}(s) ds \Theta(d\mathcal{T})$.

This result is part of the folklore of the subject (see Theorem 2 in [9] for essentially the same Palm decomposition in the more general setting where branching is combined with spatial motion), and is closely related to the spine decomposition of size-biased Galton-Watson trees in the discrete setting (see e.g. [29, Section 12.1]). For the reader's convenience, we sketch a proof of Proposition 5 in the appendix below. This proof is based on a relation between the continuous reduced tree Δ and the Brownian excursion conditioned to hit level 1, which is recalled in Section 4.2 below (see in particular Figure 7).

2.3. The continuous conductance. Before we proceed to the proof of Theorem 3, we will define and study the continuous conductance \mathcal{C} of the tree Δ , which plays a major role in this proof. Informally, the random variable \mathcal{C} is defined by viewing the random tree Δ as a network of ideal resistors with unit resistance per unit of length and letting \mathcal{C} be the conductance between the root and the set $\partial\Delta$ in this network. We will give a more formal definition using excursion measures of Brownian motion. To this end, and in view of further applications in the next section, we first define the excursion measure on a (deterministic) Yule-type tree.

So let $\mathcal{T} \in \mathbb{T}$, and consider the associated collection $(z_v)_{v \in \mathcal{V}}$ as explained in the preceding section. We suppose that $z_\emptyset > 0$. We write $C(\mathbb{R}_+, \mathcal{T})$ for the set of all continuous functions from \mathbb{R}_+ into \mathcal{T} . We also let $\mathcal{E}_{\mathcal{T}}$ be the subset of $C(\mathbb{R}_+, \mathcal{T})$ consisting of all “excursions” in \mathcal{T} : an element ω of $C(\mathbb{R}_+, \mathcal{T})$ belongs to $\mathcal{E}_{\mathcal{T}}$ if and only if $\omega(0) = (\emptyset, 0)$ and there exists a number $\zeta(\omega) \in (0, \infty]$ such that $\omega(t) \neq (\emptyset, 0)$ if and only if $0 < t < \zeta(\omega)$. For every $r \geq 0$

and $\omega \in C(\mathbb{R}_+, \mathcal{T})$, set

$$T_r(\omega) := \inf\{t \geq 0 : H(\omega(t)) = r\},$$

where we recall that $H(\omega(t))$ is the height (or distance from the root) of $\omega(t)$, and we make the usual convention $\inf \emptyset = \infty$. For every $\varepsilon \in (0, z_\emptyset)$, there is a unique $y_\varepsilon \in \mathcal{T}$ whose height is equal to ε . Let $n_{\mathcal{T}, \varepsilon}$ be the law of Brownian motion on \mathcal{T} with drift 1/2 started from y_ε and stopped when it hits the root $(\emptyset, 0)$ (if this event occurs). Then $n_{\mathcal{T}, \varepsilon}$ is a probability measure on the space $C(\mathbb{R}_+, \mathcal{T})$. If $0 < \varepsilon' < \varepsilon$ an application of the strong Markov property shows that the distribution of $(\omega(T_\varepsilon(\omega) + t))_{t \geq 0}$ under $n_{\mathcal{T}, \varepsilon'}(\cdot | T_\varepsilon < \infty)$ is $n_{\mathcal{T}, \varepsilon}$. Furthermore, $n_{\mathcal{T}, \varepsilon'}(T_\varepsilon < \infty) = (1 - e^{-\varepsilon'})/(1 - e^{-\varepsilon})$, by the formula for the scale function of linear Brownian motion with drift. From these properties, it is an easy exercise to verify that the measures $\varepsilon^{-1}n_{\mathcal{T}, \varepsilon}$ converge when $\varepsilon \rightarrow 0$ toward a σ -finite measure $n_{\mathcal{T}}$ on the set $\mathcal{E}_{\mathcal{T}}$ of all excursions in \mathcal{T} . The convergence holds in the sense that

$$\varepsilon^{-1} n_{\mathcal{T}, \varepsilon}(g_1(\omega(t_1)) \cdots g_p(\omega(t_p))) \xrightarrow{\varepsilon \rightarrow 0} n_{\mathcal{T}}(g_1(\omega(t_1)) \cdots g_p(\omega(t_p)))$$

for every choice of $0 < t_1 < \cdots < t_p$ and of the bounded continuous functions g_1, \dots, g_p on \mathcal{T} that vanish on a neighborhood of $(\emptyset, 0)$ in \mathcal{T} . Alternatively, the measure $n_{\mathcal{T}}$ is the unique σ -finite measure on $\mathcal{E}_{\mathcal{T}}$ such that, for every $\varepsilon > 0$, one has $n_{\mathcal{T}}(T_\varepsilon < \infty) = (1 - e^{-\varepsilon})^{-1}$ and the law of $(\omega(T_\varepsilon(\omega) + t))_{t \geq 0}$ under $n_{\mathcal{T}}(\cdot | T_\varepsilon < \infty)$ is $n_{\mathcal{T}, \varepsilon}$.

The measure $n_{\mathcal{T}}$ is called the excursion measure of Brownian motion (with drift 1/2) in the tree \mathcal{T} . The preceding construction is an analog of a classical construction of the Itô excursion measure of linear Brownian motion, see e.g. [32, Chapter XII]. Of course, it is also a special case of the definition of the excursion measure of a general Markov process from a regular point (see Blumenthal [6]).

The conductance $\mathcal{C}(\mathcal{T})$ is then defined by

$$\mathcal{C}(\mathcal{T}) = n_{\mathcal{T}}(\zeta = \infty) = \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} n_{\mathcal{T}, \varepsilon}(T_0 = \infty).$$

Note that we have $1 \leq \mathcal{C}(\mathcal{T}) \leq (1 - e^{-z_\emptyset})^{-1} < \infty$. The bound $\mathcal{C}(\mathcal{T}) \geq 1$ is obtained by saying that $\mathcal{C}(\mathcal{T})$ is greater than the conductance of the trivial tree that consists only of a half-line. The other bound follows from the form of the scale function of Brownian motion with drift, which yields an explicit expression for the probability under $n_{\mathcal{T}, \varepsilon}$ that the process comes back to 0 before hitting the first branching point.

To simplify notation, we set $\mathcal{C} = \mathcal{C}(\Gamma)$, which is a random variable with values in $[1, \infty)$. Because of the relations between the Yule tree Γ and the

reduced tree Δ , the random conductance \mathcal{C} may also be defined as the mass assigned by the excursion measure of Brownian motion on Δ (away from the root), to the set of trajectories that reach height 1 before coming back to the root.

The distributional identity (2) is obvious from the electric network interpretation: just view Δ as a series of two conductors, the first one being a segment of length U and the second one consisting of two independent copies of Δ (scaled by the factor $1 - U$) in parallel. Alternatively, it is also easy to derive (2) from the probabilistic definition in terms of excursion measures, by applying the strong Markov property at the hitting time of the first node of the tree. We leave the details to the reader.

Let us now prove that (2) characterizes the law of \mathcal{C} and discuss some of the properties of this law. For $u \in (0, 1)$ and $x, y \geq 1$, we define

$$(5) \quad G(u, x, y) := \left(u + \frac{1 - u}{x + y} \right)^{-1},$$

so that (2) can be rewritten as

$$(6) \quad \mathcal{C} \stackrel{(d)}{=} G(U, \mathcal{C}_1, \mathcal{C}_2)$$

where $U, \mathcal{C}_1, \mathcal{C}_2$ are as in (2). Let \mathcal{M} be the set of all probability measures on $[1, \infty]$ and let $\Phi : \mathcal{M} \rightarrow \mathcal{M}$ map a distribution λ to

$$\Phi(\lambda) = \text{Law}(G(U, X_1, X_2))$$

where X_1 and X_2 are independent and distributed according to λ and U is uniformly distributed over $[0, 1]$ and independent of the pair (X_1, X_2) .

PROPOSITION 6. *The law γ of \mathcal{C} is the unique fixed point of the mapping Φ on \mathcal{M} , and we have $\Phi^k(\lambda) \rightarrow \gamma$ weakly as $k \rightarrow \infty$, for every $\lambda \in \mathcal{M}$. Furthermore all moments of γ are finite, and γ has a continuous density over $[1, \infty)$. Finally, the Laplace transform*

$$\varphi(\ell) = \mathbb{E}[\exp(-\ell \mathcal{C}/2)] = \int_1^\infty e^{-\ell r/2} \gamma(dr), \quad \ell \geq 0$$

solves the differential equation

$$(7) \quad 2\ell \varphi''(\ell) + \ell \varphi'(\ell) + \varphi^2(\ell) - \varphi(\ell) = 0.$$

Remark. In [27] the authors discuss the conductance of an infinite supercritical Galton–Watson tree with offspring distribution θ . This conductance also satisfies a recursive distributional equation, which depends on θ . In that setting, it is conjectured that the distribution of the conductance is absolutely continuous with respect to Lebesgue measure at least if $\theta(k) = 0$ for all sufficiently large k , see [26, 29].

PROOF. We start with a few preliminary observations. If $\lambda, \lambda' \in \mathcal{M}$ we say that a random pair (X, Y) is a coupling of λ and λ' if X is distributed according to λ and Y is distributed according to λ' . The stochastic partial order \preceq on \mathcal{M} is defined by saying that $\lambda \preceq \lambda'$ if and only if there exists a coupling (X, Y) of λ and λ' such that $X \leq Y$ a.s. It is then clear that the mapping Φ is increasing for the stochastic partial order.

We endow the set \mathcal{M}_1 of all probability measures on $[1, \infty]$ that have a finite first moment with the 1-Wasserstein metric

$$d_1(\lambda, \lambda') := \inf \{E[|X - Y|] : (X, Y) \text{ coupling of } (\lambda, \lambda')\}.$$

The metric space (\mathcal{M}_1, d_1) is Polish and its topology is finer than the weak topology on \mathcal{M}_1 . From the easy bound $G(u, x, y) \leq x + y$, we immediately see that Φ maps \mathcal{M}_1 into \mathcal{M}_1 . We then observe that the mapping Φ is strictly contractant on \mathcal{M}_1 . To see this, let (X_1, Y_1) and (X_2, Y_2) be two independent copies of a coupling between $\lambda, \lambda' \in \mathcal{M}_1$ and let U be uniformly distributed over $[0, 1]$ and independent of (X_1, Y_1, X_2, Y_2) . Then the two variables $G(U, X_1, X_2)$ and $G(U, Y_1, Y_2)$ give a coupling of $\Phi(\lambda)$ and $\Phi(\lambda')$. Using the fact that $X_1, Y_1, X_2, Y_2 \geq 1$ we have

$$\begin{aligned} & |G(U, X_1, X_2) - G(U, Y_1, Y_2)| \\ &= \left| \left(U + \frac{1-U}{X_1 + X_2} \right)^{-1} - \left(U + \frac{1-U}{Y_1 + Y_2} \right)^{-1} \right| \\ &= \left| \frac{(X_1 + X_2 - Y_1 - Y_2)(1-U)}{(U(X_1 + X_2) + 1 - U)(U(Y_1 + Y_2) + 1 - U)} \right| \\ &\leq (|X_1 - Y_1| + |X_2 - Y_2|) \frac{1-U}{(1+U)^2}. \end{aligned}$$

Taking expected values and minimizing over the choice of the coupling between λ and λ' , we get $d_1(\Phi(\lambda), \Phi(\lambda')) \leq 2(1 - \log(2))d_1(\lambda, \lambda')$. Since $2(1 - \log(2)) < 1$, the mapping Φ is contractant on \mathcal{M}_1 and by completeness it has a unique fixed point γ_0 in \mathcal{M}_1 . Furthermore, for every $\lambda \in \mathcal{M}_1$, we have $\Phi^k(\lambda) \rightarrow \gamma_0$ for the metric d_1 , hence also weakly, as $k \rightarrow \infty$.

Since we know from (6) that γ is also a fixed point of Φ , the equality $\gamma = \gamma_0$ will follow if we can verify that γ_0 is the unique fixed point of Φ in \mathcal{M} . To

this end, it will be enough to verify that we have $\Phi^k(\lambda) \rightarrow \gamma_0$ as $k \rightarrow \infty$, for every $\lambda \in \mathcal{M}$. Let $\lambda \in \mathcal{M}$ and for every $t \in \mathbb{R}$ set $F_\lambda(t) = \lambda([t, \infty])$. Also set $F_\lambda^{(2)}(t) = P(X_1 + X_2 \geq t)$ where X_1 and X_2 are independent and distributed according to λ . Then we have, for every $t > 1$,

$$\begin{aligned}
 F_{\Phi(\lambda)}(t) &= P\left(U + \frac{1-U}{X_1 + X_2} \leq t^{-1}\right) \\
 &= P\left(U < t^{-1} \text{ and } \frac{t-Ut}{1-Ut} \leq X_1 + X_2\right) \\
 &= \int_0^{1/t} du F_\lambda^{(2)}\left(\frac{t-ut}{1-ut}\right) \\
 (8) \qquad &= \frac{t-1}{t} \int_t^\infty \frac{dx}{(x-1)^2} F_\lambda^{(2)}(x).
 \end{aligned}$$

It follows that, for every $t \geq 1$,

$$(9) \qquad F_{\Phi(\lambda)}(t) \leq \frac{F_\lambda^{(2)}(t)}{t} \leq \frac{2F_\lambda(t/2)}{t}.$$

We apply this to $\lambda = \Phi(\delta_\infty)$, where δ_∞ is the Dirac measure at ∞ . We have $F_{\Phi(\delta_\infty)}(t) = t^{-1}$, and it follows that, for every $t \geq 1$,

$$F_{\Phi^2(\delta_\infty)}(t) \leq \frac{4}{t^2}.$$

This implies that $\Phi^2(\delta_\infty) \in \mathcal{M}_1$. By monotonicity, we have also $\Phi^2(\lambda) \in \mathcal{M}_1$ for every $\lambda \in \mathcal{M}$, and from the preceding results we get $\Phi^k(\lambda) \rightarrow \gamma_0$ for every $\lambda \in \mathcal{M}$. As explained above this implies that $\gamma = \gamma_0$ is the unique fixed point of Φ in \mathcal{M} .

Let us now check that all moments of γ are finite. To simplify notation, we write $F = F_\gamma$ and $F^{(2)} = F_\gamma^{(2)}$. By (8) we have for every $t > 1$,

$$(10) \qquad F(t) = \frac{t-1}{t} \int_t^\infty \frac{dx}{(x-1)^2} F^{(2)}(x)$$

which implies that $F(t) \leq 2F(t/2)/t$ for every $t \geq 1$, by the same argument as above. Iterating this inequality, we get that $F(t) \leq c_1 \exp(-c_2(\log t)^2)$, with certain constants $c_1, c_2 > 0$. It follows that all moments of γ are finite.

By construction, we have $F^{(2)}(t) = 1$ for every $t \in [1, 2]$. It then immediately follows from (10) that we have

$$(11) \qquad F(t) = \frac{K_0}{t} + 1 - K_0, \qquad \forall t \in [1, 2],$$

where

$$K_0 = 2 - \int_2^\infty \frac{dx}{(x-1)^2} F^{(2)}(x) \in [1, 2].$$

Then we observe that the right-hand side of (10) is a continuous function of $t \in (1, \infty)$, so that F is continuous on $[1, \infty)$ (the right-continuity at 1 is obvious from (11)). Thus γ has no atoms and it follows that the function $F^{(2)}$ is also continuous on $[1, \infty)$. Using (10) again we obtain that F is continuously differentiable on $[1, \infty)$ and consequently γ has a continuous density $f = -F'$ with respect to Lebesgue measure on $[1, \infty)$. By (11), $f(t) = K_0 t^{-2}$ for $t \in [1, 2]$ and in particular $f(1) = K_0$.

Let us finally derive the differential equation (7). To this end, we first differentiate (10) with respect to t to get that the linear differential equation

$$(12) \quad t(t-1)F'(t) - F(t) = -F^{(2)}(t).$$

holds for $t \in [1, \infty)$. Then let $g : [1, \infty) \rightarrow \mathbb{R}_+$ be a continuously differentiable function such that $g(x)$ and $g'(x)$ are both $o(x^\alpha)$ when $x \rightarrow \infty$, for some $\alpha \in (0, \infty)$. From the definition of F and Fubini's theorem, we have

$$\int_1^\infty dt g'(t)F(t) = \mathbb{E}[g(\mathcal{C})] - g(1)$$

and similarly

$$\int_1^\infty dt g'(t)F^{(2)}(t) = \mathbb{E}[g(\mathcal{C}_1 + \mathcal{C}_2)] - g(1)$$

where \mathcal{C}_1 and \mathcal{C}_2 are independent copies of \mathcal{C} under the probability \mathbb{P} . We then multiply both sides of (12) by $g'(t)$ and integrate for t running from 1 to ∞ to get

$$(13) \quad \mathbb{E}[\mathcal{C}_1(\mathcal{C}_1 - 1)g'(\mathcal{C}_1)] + \mathbb{E}[g(\mathcal{C}_1)] = \mathbb{E}[g(\mathcal{C}_1 + \mathcal{C}_2)].$$

When $g(x) = \exp(-x\ell/2)$ for $\ell > 0$, this readily gives (7). \square

Remark. We may also take $g(x) = x^m$ for $m \in \{1, 2, 3, 4, \dots\}$ in (13). This leads to recursive formulas for the moments of \mathcal{C} in terms of the first moment $\mathbb{E}[\mathcal{C}]$ (simulations give $\mathbb{E}[\mathcal{C}] \approx 1.72$).

Singular behavior of the density of γ . By (11), the values of F and $f = -F'$ on the interval $[1, 2]$ are determined by the constant $K_0 = f(1)$. We have not been able to obtain an exact numerical value for K_0 , but simulations

indicate that $K_0 \approx 1.47$ (see Figure 3). We may now observe that the values of F over $[1, 2]$ determine the values of $F^{(2)}$ over $[2, 3]$, via the formula

$$1 - F^{(2)}(t) = \int_1^{t-1} ds f(s) (1 - F(t - s)), \quad \forall t \in [2, 3].$$

We can then use either (12) or (10) to get a complicated explicit expression for F over $[2, 3]$, again in terms of K_0 . By iterating the argument, we can in principle determine F by solving linear differential equations on the successive intervals $[n, n + 1]$, $n = 1, 2, \dots$. Unfortunately, the calculations become tedious and we have not been able to find a closed expression for $F(t)$. However, from the expressions found for the first two intervals $[1, 2]$ and $[2, 3]$, one can verify that, although the function f is continuously differentiable on $(1, 3)$, one has

$$f''(2-) = \frac{3K_0}{8}, \quad \text{whereas} \quad f''(2+) = \frac{3K_0 - 4K_0^2}{8},$$

so that f is not twice differentiable at the point 2. See the inflection point at 2 on Figure 3.

2.4. The flow property of harmonic measure. In this section we establish a property of harmonic measure that plays an important role in the proof of Theorem 3. This property is well known in the discrete setting, but perhaps less standard in the continuous setting, and we sketch a short proof.

We fix a Yule-type tree $\mathcal{T} \in \mathbb{T}$. In this section only, we slightly abuse notation by writing $W = (W_t)_{t \geq 0}$ for Brownian motion with drift 1/2 on \mathcal{T} started from the root. As previously, W_∞ is the exit ray of W , and the distribution of W_∞ is the harmonic measure of \mathcal{T} . For every $r > 0$, if x is the unique point of \mathcal{T}_r such that $x \prec W_\infty$, we write $W_\infty^{(r)}$ for the ray of $\mathcal{T}[x]$ that is obtained by shifting W_∞ at time r .

LEMMA 7. *Let $r > 0$ and $x \in \mathcal{T}_r$. Conditionally on $\{x \prec W_\infty\}$, the law of $W_\infty^{(r)}$ is the harmonic measure of $\mathcal{T}[x]$.*

PROOF. For simplicity, we suppose that x is not a branching point of \mathcal{T} , and then we can choose $\varepsilon > 0$ sufficiently small so that there is a unique descendant x_ε of x in \mathcal{T} at distance ε from x . Clearly the harmonic measure of $\mathcal{T}[x]$ can be obtained by considering the distribution of the exit ray of Brownian motion started from x_ε and conditioned never to hit x . On the other hand, by considering the successive passage times at x_ε , we can also verify that the conditional law of $W_\infty^{(r)}$ knowing that $x \prec W_\infty$ corresponds to the same distribution. We leave the details to the reader. \square

3. Proof of Theorem 3. Let us outline the main steps of the proof of Theorem 3. Proposition 8 below uses the spine decomposition (Proposition 5) and the Ray-Knight theorem for local times of Brownian motion with drift to determine the exact distribution of the subtree of the Yule tree above level r that is selected by harmonic measure. In Section 3.2, we use stochastic calculus to prove that this law converges as $r \rightarrow \infty$ to an explicit distribution, which is absolutely continuous with respect to Θ (Corollary 11). In the last two subsections, we rely on arguments of ergodic theory, mainly inspired by [27], to complete the proof of Theorem 3.

We recall that \mathbb{P} stands for the probability measure under which the Yule tree is defined, whereas Brownian motion (with drift $1/2$) on the Yule tree is defined under the probability measure P .

3.1. The subtree above level r selected by harmonic measure. In this subsection, we fix $r > 0$. We will implicitly use the fact that Γ has a.s. no branching point at height r .

There is a unique point $x \in \Gamma_r$ such that $x \prec W_\infty$, and we set $\Gamma^{(r)} = \Gamma[x]$, which is the subtree above level r selected by harmonic measure. We are interested in the distribution of $\Gamma^{(r)}$. Let F be a nonnegative measurable function on \mathbb{T} , and consider the quantity

$$(14) \quad I_r := \mathbb{E} \otimes E[F(\Gamma^{(r)})] = \mathbb{E} \otimes E \left[\sum_{x \in \Gamma_r} F(\Gamma[x]) \mathbf{1}_{\{x \prec W_\infty\}} \right],$$

where the notation $\mathbb{E} \otimes E$ means that we consider the expectation first under the probability measure P (under which the Brownian motion W is defined) and then under \mathbb{P} . We will use Proposition 5 to evaluate I_r . In the first part of the argument, until the derivation of formula (17) below, we argue under the probability measure P , that is, conditionally given the tree Γ .

Let us fix $x \in \Gamma_r$ and $R > r$. We will use the notation $\tilde{\Gamma}[x] := \{y \in \Gamma : x \prec y\}$. This is just the set of all descendants of x in Γ , now viewed as a subset of Γ and not as a Yule-type tree as in the definition of $\Gamma[x]$. Define

$$\Gamma^{x,R} = \{y \in \Gamma \setminus \tilde{\Gamma}[x] : H(y) \leq R\} \cup \tilde{\Gamma}[x].$$

Let $W^{x,R}$ be Brownian motion (with drift $1/2$) on $\Gamma^{x,R}$ (we assume that $W^{x,R}$ is reflected both at the root and at the leaves of $\Gamma^{x,R}$, which are the points y of $\Gamma \setminus \tilde{\Gamma}[x]$ such that $H(y) = R$). We look for an expression of the probability that $W^{x,R}$ never hits the leaves of $\Gamma^{x,R}$, or equivalently that $W^{x,R}$ escapes to infinity in $\tilde{\Gamma}[x]$ before hitting any leaf of $\Gamma^{x,R}$. Write $(\ell_t^{x,R})_{t \geq 0}$ for the local time process of $W^{x,R}$ at x . Note that we use here the

standard normalization of local time as an occupation time density. With this normalization, $\ell_t^{x,R}$ is the a.s. limit as $\varepsilon \rightarrow 0$ of the quantities $2\varepsilon N_t^{x,\varepsilon}$, where $N_t^{x,\varepsilon}$ is the number of “upcrossings” of $W^{x,R}$ from x to the point $x_\varepsilon \in \Gamma$ such that $x \prec x_\varepsilon$ and $d(x, x_\varepsilon) = \varepsilon$ (this point is unique for ε small) before time t . We claim that $\ell_\infty^{x,R}$ has an exponential distribution with parameter $\mathcal{C}(\Gamma[x])/2$. This is easy from excursion theory, but an elementary argument can be given as follows. Each time $W^{x,R}$ does an upcrossing from x to x_ε , there is a probability of order $\varepsilon \mathcal{C}(\Gamma[x])$ that it escapes to infinity before coming back to x (by the very definition of $\mathcal{C}(\Gamma[x])$). Hence the total number of upcrossings from x to x_ε before escaping to infinity is geometric with parameter of order $\varepsilon \mathcal{C}(\Gamma[x])$, and our claim follows from the approximation of local time by upcrossing numbers.

We then consider, for every $a \in [0, r]$, the local time process $(L_t^{a,R})_{t \geq 0}$ of $W^{x,R}$ at the unique point of $\llbracket 0, x \rrbracket$ at distance a from the root. Note in particular that $L_t^{r,R} = \ell_t^{x,R}$. The distribution of the process $(L_\infty^{a,R})_{0 \leq a \leq r}$ can be derived via a time change argument, which consists in looking at $W^{x,R}$ only when it visits $\llbracket 0, x \rrbracket$. More precisely, we set, for every $s \geq 0$,

$$\tau_s := \inf \left\{ t \geq 0 : \int_0^t \mathbf{1}_{\llbracket 0, x \rrbracket} (W_r^{x,R}) \, dr > s \right\}$$

with $\inf \emptyset = \infty$ as usual. Setting $Z_s = W_{\tau_s}^{x,r}$ if $\tau_s < \infty$ and $Z_s = x$ otherwise, we obtain that the process $(Z_s)_{s \geq 0}$ is under P a Brownian motion (with drift $1/2$) on $\llbracket 0, x \rrbracket$ started from the root, reflected at both ends of the segment $\llbracket 0, x \rrbracket$, and stopped when its local time at x hits an independent exponential variable with parameter $\mathcal{C}(\Gamma[x])/2$. The latter exponential random variable is of course the local time $\ell_\infty^{x,R} = L_\infty^{r,R}$, and the independence property in the last sentence corresponds to the independence of excursions of $W^{x,R}$ “below” and “above” x . The preceding assertions can be obtained either by arguments of excursion theory, or, via scaling limits, from the (easy) corresponding properties for random walk on discrete trees.

Observe that, for every $y \in \llbracket 0, x \rrbracket$, the total local time of $W^{x,R}$ at y coincides with the total local time of Z at y . Using the Ray-Knight theorem for Brownian motion with drift (see e.g. [7], p.93) we get that, conditionally on $\ell_\infty^{x,R} = \ell$, the process $(L_\infty^{r-a,R})_{0 \leq a \leq r}$ is distributed as the process $(X_a)_{0 \leq a \leq r}$ which solves the stochastic differential equation

$$(15) \quad \begin{cases} dX_a = 2\sqrt{X_a} \, d\eta_a + (2 - X_a) \, da \\ X_0 = \ell \end{cases}$$

where $(\eta_a)_{a \geq 0}$ is a standard linear Brownian motion. In what follows, we will write P_ℓ for the probability measure under which the process X starts from

ℓ , and $P_{(c)}$ for the probability measure under which the process X starts with an exponential distribution with parameter $c/2$.

Now write x_j , $1 \leq j \leq k$ for the branching points of $\Gamma^{x,R}$ (or equivalently of Γ) that belong to $\llbracket 0, x \rrbracket$, and set $a_j = H(x_j)$ for $1 \leq j \leq k$. Also let $\Gamma_{x,j,R}$ be the subtree of $\Gamma^{x,R}$ that branches off $\llbracket 0, x \rrbracket$ at x_j . We consider the event $A_{x,R}$ where $W^{x,R}$ never hits the leaves of $\Gamma^{x,R}$. We can compute the conditional probability of $A_{x,R}$ knowing the local times $(L_\infty^{a,R})_{0 \leq a \leq r}$, using arguments of excursion theory. For $1 \leq j \leq k$, write $n_{(j)}$ for the excursion measure of Brownian motion with drift $1/2$ in the tree $\Gamma_{x,j,R}$ (defined as in the beginning of Section 2.3) and let $\mathcal{C}(\Gamma_{x,j,R})$ be the conductance of $\Gamma_{x,j,R}$ between its root x_j and the set of its leaves. This conductance may be defined as the measure under $n_{(j)}$ of the event E_j where the excursion hits the leaves before returning to the root. Conditionally given $(L_\infty^{a,R})_{0 \leq a \leq r}$, the excursions of $W^{x,R}$ inside the tree $\Gamma_{x,j,R}$ form a Poisson point process with intensity $\frac{1}{2} L_\infty^{a_j,R} n_{(j)}(\cdot)$ and these point processes are independent when j varies (once again, the reader who is unfamiliar with excursion theory may find it easier to deduce these statements from their discrete versions, which are elementary). Consequently, the conditional probability for a fixed j that no excursion in $\Gamma_{x,j,R}$ hits the leaves is $\exp(-\frac{1}{2} L_\infty^{a_j,R} n_{(j)}(E_j)) = \exp(-\frac{1}{2} \mathcal{C}(\Gamma_{x,j,R}) L_\infty^{a_j,R})$. Finally, thanks to the conditional independence of the point processes of excursions in the different trees $\Gamma_{x,j,R}$, we get that the conditional probability of $A_{x,R}$ knowing $(L_\infty^{a,R})_{0 \leq a \leq r}$ is

$$\exp\left(-\frac{1}{2} \sum_{j=1}^k \mathcal{C}(\Gamma_{x,j,R}) L_\infty^{a_j,R}\right).$$

Using the distribution of the process $(L_\infty^{r-a,R})_{0 \leq a \leq r}$, we have thus

$$\begin{aligned} P(A_{x,R}) &= E\left[\exp\left(-\frac{1}{2} \sum_{j=1}^k \mathcal{C}(\Gamma_{x,j,R}) L_\infty^{a_j,R}\right)\right] \\ (16) \quad &= E_{(\mathcal{C}(\Gamma[x]))}\left[\exp\left(-\frac{1}{2} \sum_{j=1}^k \mathcal{C}(\Gamma_{x,j,R}) X_{r-a_j}\right)\right]. \end{aligned}$$

At this point, we let R tend to infinity. It is easy to verify that $P(A_{x,R})$ increases to $P(A_x)$, where $A_x = \{x \prec W_\infty\}$. Furthermore, for every $j \in \{1, \dots, k\}$, $\mathcal{C}(\Gamma_{x,j,R})$ decreases to $\mathcal{C}(\Gamma_{x,j})$, where $\Gamma_{x,j}$ is the subtree of Γ branching off $\llbracket 0, x \rrbracket$ at x_j . Consequently, we obtain that

$$(17) \quad P(x \prec W_\infty) = E_{(\mathcal{C}(\Gamma[x]))}\left[\exp\left(-\frac{1}{2} \sum_{j=1}^k \mathcal{C}(\Gamma_{x,j}) X_{r-a_j}\right)\right].$$

We can now return to the computation of the quantity I_r defined in (14). By integrating (17) with respect to \mathbb{P} , we get

$$\begin{aligned} I_r &= \mathbb{E} \left[\sum_{x \in \Gamma_r} F(\Gamma[x]) P(x \prec W_\infty) \right] \\ &= \mathbb{E} \left[\sum_{x \in \Gamma_r} F(\Gamma[x]) E_{(\mathcal{C}(\Gamma[x]))} \left[\exp \left(-\frac{1}{2} \sum_{j=1}^k \mathcal{C}(\Gamma_{x,j}) X_{r-a_j} \right) \right] \right]. \end{aligned}$$

Note that the quantity inside the sum over $x \in \Gamma_r$ is a function of $\Gamma[x]$ and of the subtrees of Γ branching off the segment $[[0, x]]$. We can thus apply Proposition 5 and we get

$$I_r = e^r \int \Theta(d\mathcal{T}) F(\mathcal{T}) \mathbf{E} \left[E_{(\mathcal{C}(\mathcal{T}))} \left[\exp \left(-\frac{1}{2} \int \mathcal{N}_r(da d\mathcal{T}') \mathcal{C}(\mathcal{T}') X_{r-a} \right) \right] \right],$$

where under the probability measure \mathbf{P} , $\mathcal{N}_r(da d\mathcal{T}')$ is a Poisson point measure on $[0, r] \times \mathbb{T}$ with intensity $2 da \Theta(d\mathcal{T}')$. We can interchange the expectation under \mathbf{P} and the one under $P_{(\mathcal{C}(\mathcal{T}))}$, and using the exponential formula for Poisson measures, we arrive at

$$(18) \quad I_r = e^r \int \Theta(d\mathcal{T}) F(\mathcal{T}) E_{(\mathcal{C}(\mathcal{T}))} \left[\exp -2 \int_0^r da (1 - \varphi(X_a)) \right],$$

where we recall that for every $s \geq 0$,

$$\varphi(s) = \mathbb{E}[\exp(-s\mathcal{C}/2)] = \Theta \left(\exp(-s\mathcal{C}(\mathcal{T})/2) \right)$$

is the Laplace transform (evaluated at $s/2$) of the distribution of the conductance of the Yule tree. We have thus proved the following proposition.

PROPOSITION 8. *The distribution under $\mathbb{P} \otimes P$ of the subtree $\Gamma^{(r)}$ has a density with respect to the law $\Theta(d\mathcal{T})$ of the Yule tree, which is given by $\Phi_r(\mathcal{C}(\mathcal{T}))$, where, for every $c > 0$,*

$$\Phi_r(c) = E_{(c)} \left[\exp - \int_0^r da (1 - 2\varphi(X_a)) \right].$$

3.2. Asymptotics. In this section, we study the asymptotic behavior of $\Phi_r(c)$ when r tends to ∞ . We first observe that, in terms of the law $\gamma(ds)$ of $\mathcal{C}(\Gamma)$, we have

$$\varphi(\ell) = \int_{[1, \infty)} e^{-\ell s/2} \gamma(ds), \quad \varphi'(\ell) = -\frac{1}{2} \int_{[1, \infty)} s e^{-\ell s/2} \gamma(ds).$$

It follows that $\varphi(\ell) \leq e^{-\ell/2}$ and $|\varphi'(\ell)| \leq \frac{1}{2}(\int s\gamma(ds))e^{-\ell/2}$. By differentiating (7), we have also

$$(19) \quad 2\ell \varphi'''(\ell) + (2 + \ell)\varphi''(\ell) + 2\varphi(\ell)\varphi'(\ell) = 0.$$

Our main tool is the next proposition.

PROPOSITION 9. *For every $\ell \geq 0$,*

$$\lim_{r \rightarrow \infty} E_\ell \left[\exp - \int_0^r da (1 - 2\varphi(X_a)) \right] = - \frac{\varphi'(\ell)e^{\ell/2}}{\int_0^\infty ds \varphi'(s)^2 e^{s/2}}.$$

Additionally, there exists a constant $K < \infty$ such that, for every $\ell \geq 0$ and $r > 0$,

$$E_\ell \left[\exp - \int_0^r da (1 - 2\varphi(X_a)) \right] \leq K.$$

PROOF. Under the probability measure $P_{(\mathcal{C}(\mathcal{T}))}$ the process X starts with an initial distribution which is exponential with parameter $\mathcal{C}(\mathcal{T})/2$. Consequently, under $\int \Theta(d\mathcal{T}) P_{(\mathcal{C}(\mathcal{T}))}$, the initial density of X is

$$(20) \quad q(\ell) = \int_{[1, \infty)} \gamma(ds) \frac{s}{2} e^{-s\ell/2} = -\varphi'(\ell).$$

However, from (18) with $F = 1$, we have

$$\begin{aligned} 1 &= \int \Theta(d\mathcal{T}) E_{(\mathcal{C}(\mathcal{T}))} \left[\exp - \int_0^r da (1 - 2\varphi(X_a)) \right] \\ &= - \int d\ell \varphi'(\ell) E_\ell \left[\exp - \int_0^r da (1 - 2\varphi(X_a)) \right]. \end{aligned}$$

We can generalize the last identity via a minor extension of the calculations of the preceding section. We let L_∞^0 be the total local time accumulated by the process W at the root of Γ . Let $r > 0$ and let h be a bounded non-negative continuous function on $(0, \infty)$, and instead of the quantity I_r of the preceding section, set

$$I_r^h := \mathbb{E} \otimes E \left[h(L_\infty^0) \sum_{x \in \Gamma_r} F(\Gamma[x]) \mathbf{1}_{\{x \prec W_\infty\}} \right],$$

where F is a given nonnegative measurable function on \mathbb{T} . The same calculations that led to (16) give, for every $x \in \Gamma_r$ and $R > r$,

$$\begin{aligned} E[h(L_\infty^{0,R}) \mathbf{1}_{A_{x,R}}] &= E\left[h(L_\infty^{0,R}) \exp\left(-\frac{1}{2} \sum_{j=1}^k \mathcal{C}(\Gamma_{x,j,R}) L_\infty^{a_j,R}\right)\right] \\ (21) \qquad &= E_{(\mathcal{C}(\Gamma[x]))}\left[h(X_r) \exp\left(-\frac{1}{2} \sum_{j=1}^k \mathcal{C}(\Gamma_{x,j,R}) X_{r-a_j}\right)\right]. \end{aligned}$$

When $R \rightarrow \infty$, $L_\infty^{0,R}$ converges to L_∞^0 , and so we get

$$E[h(L_\infty^0) \mathbf{1}_{\{x < W_\infty\}}] = E_{(\mathcal{C}(\Gamma[x]))}\left[h(X_r) \exp\left(-\frac{1}{2} \sum_{j=1}^k \mathcal{C}(\Gamma_{x,j}) X_{r-a_j}\right)\right].$$

We then sum over $x \in \Gamma_r$ and integrate with respect to \mathbb{P} . By the same manipulations as in the preceding section, we arrive at

$$(22) \quad I_r^h = e^r \int \Theta(d\mathcal{T}) F(\mathcal{T}) E_{(\mathcal{C}(\mathcal{T}))}\left[h(X_r) \exp -2 \int_0^r da (1 - \varphi(X_a))\right].$$

Note that if $F = 1$,

$$I_r^h = \mathbb{E} \otimes E[h(L_\infty^0)] = - \int_0^\infty \varphi'(\ell) h(\ell) d\ell$$

since given $\Gamma = \mathcal{T}$ the local time L_∞^0 follows an exponential distribution with parameter $\mathcal{C}(\mathcal{T})/2$, and we use the same calculation as in (20). Hence the case $F = 1$ of (22) gives

$$(23) \quad \int_0^\infty d\ell \varphi'(\ell) E_\ell\left[h(X_r) \exp - \int_0^r da (1 - 2\varphi(X_a))\right] = \int_0^\infty d\ell \varphi'(\ell) h(\ell).$$

By an obvious truncation argument, this identity also holds if h is unbounded.

At this point, we need a lemma.

LEMMA 10. *The process*

$$M_a := -\varphi'(X_a) \exp\left(\frac{X_a}{2} - \int_0^a ds (1 - 2\varphi(X_s))\right)$$

is a martingale under P_ℓ , for every $\ell \geq 0$.

Proof of Lemma 10. From the stochastic differential equation (15), an application of Itô's formula shows that the finite variation part of the semimartingale $-M_a$ is

$$\int_0^a (2X_s\varphi'''(X_s) + (2 + X_s)\varphi''(X_s) + 2\varphi(X_s)\varphi'(X_s)) \times \exp\left(\frac{X_s}{2} - \int_0^s du(1 - 2\varphi(X_u))\right) ds$$

and this vanishes thanks to (19). Hence M is a local martingale. Furthermore, we already noticed that, for every $\ell \geq 0$, $|\varphi'(\ell)| \leq Ce^{-\ell/2}$, where $C := \frac{1}{2} \int s \gamma(ds)$. It follows that $|M|$ is bounded by Ce^a over the time interval $[0, a]$, and thus M is a (true) martingale. \square

We return to the proof of Proposition 9. Let $\ell \geq 0$ and $t > 0$. On the probability space where X is defined, we introduce a new probability measure Q_ℓ^t by setting

$$Q_\ell^t = \frac{M_t}{M_0} \cdot P_\ell.$$

Note that the fact that Q_ℓ^t is a probability measure follows from the martingale property derived in Lemma 10. Furthermore, we have P_ℓ a.s.

$$\frac{M_t}{M_0} = \frac{\varphi'(X_t)}{\varphi'(\ell)} \exp\left(\frac{X_t - \ell}{2} - \int_0^t ds(1 - 2\varphi(X_s))\right),$$

so that the martingale part of $\log \frac{M_t}{M_0}$ is

$$\int_0^t \sqrt{X_s} d\eta_s + 2 \int_0^t \frac{\varphi''(X_s)}{\varphi'(X_s)} \sqrt{X_s} d\eta_s,$$

where η is the linear Brownian motion in (15). An application of Girsanov's theorem shows that the process

$$\tilde{\eta}_s := \eta_s - \int_0^s \sqrt{X_u} \left(1 + \frac{2\varphi''(X_u)}{\varphi'(X_u)}\right) du, \quad 0 \leq s \leq t,$$

is a linear Brownian motion over the time interval $[0, t]$, under Q_ℓ^t . Furthermore, still on the time interval $[0, t]$, the process X satisfies the stochastic differential equation

$$dX_s = 2\sqrt{X_s} d\tilde{\eta}_s + 2X_s \left(1 + \frac{2\varphi''(X_s)}{\varphi'(X_s)}\right) ds + (2 - X_s) ds,$$

or equivalently, using (7),

$$(24) \quad dX_s = 2\sqrt{X_s} d\tilde{\eta}_s + \left(2 - X_s + 2\frac{\varphi - \varphi^2}{\varphi'}(X_s)\right) ds.$$

Notice that the function

$$\ell \mapsto \frac{\varphi - \varphi^2}{\varphi'}(\ell)$$

is continuously differentiable over $[0, \infty)$, takes negative values on $(0, \infty)$ and vanishes at 0. Pathwise uniqueness, and therefore also weak uniqueness, holds for (24) by an application of the classical Yamada-Watanabe criterion (see e.g. [32, Theorem IX.3.5]). The preceding considerations show that, under the probability measure Q_ℓ^t and on the time interval $[0, t]$, the process X is distributed as the diffusion process on $[0, \infty)$ with generator

$$\mathcal{L} = 2r \frac{d^2}{dr^2} + \left(2 - r + 2\frac{\varphi - \varphi^2}{\varphi'}(r)\right) \frac{d}{dr}$$

started from ℓ . Write \tilde{X} for this diffusion process, and assume that \tilde{X} starts from ℓ under the probability measure P_ℓ . Note that 0 is an entrance point for \tilde{X} , but, independently of its starting point, \tilde{X} does not visit 0 at a positive time (indeed this follows from the fact that X does not visit 0 at a positive time).

Standard comparison theorems for stochastic differential equations (see e.g. [32, Theorem IX.3.7]) can be used to compare the solutions of (15) and (24), and it follows that \tilde{X} is recurrent on $(0, \infty)$.

We next observe that the finite measure ρ on $(0, \infty)$ defined by

$$\rho(d\ell) := \varphi'(\ell)^2 e^{\ell/2} d\ell$$

is invariant for \tilde{X} . Indeed, we have, for any bounded continuous function h on $(0, \infty)$,

$$\begin{aligned} & \int_{(0, \infty)} d\ell \varphi'(\ell)^2 e^{\ell/2} E_\ell[h(\tilde{X}_t)] \\ &= \int_{(0, \infty)} d\ell \varphi'(\ell)^2 e^{\ell/2} Q_\ell^t[h(X_t)] \\ &= \int_{(0, \infty)} d\ell \varphi'(\ell) E_\ell \left[h(X_t) \varphi'(X_t) \exp \left(\frac{X_t}{2} - \int_0^t ds (1 - 2\varphi(X_s)) \right) \right] \\ &= \int_{(0, \infty)} d\ell \varphi'(\ell)^2 e^{\ell/2} h(\ell), \end{aligned}$$

where the last equality follows from (23). We normalize ρ by setting

$$\widehat{\rho} = \frac{\rho}{\rho((0, \infty))}.$$

From the known results about the convergence of positive recurrent diffusion processes toward their stationary distribution (see Chapter 23 in Kallenberg [19]), the distribution of \widetilde{X}_t under P_ℓ converges to $\widehat{\rho}$ in variation norm as $t \rightarrow \infty$, for any $\ell \geq 0$. Consequently, for any bounded Borel function g on $[0, \infty)$, and every $\ell \geq 0$,

$$(25) \quad E_\ell[g(\widetilde{X}_t)] \xrightarrow[t \rightarrow \infty]{} \int g \, d\widehat{\rho}.$$

We claim that (25) still holds if g may be unbounded but is assumed to be nonnegative, monotone increasing and such that $\int g \, d\widehat{\rho} < \infty$. To see this, fix $\ell \geq 0$ and write $\Pi_t(\ell, d\ell')$ for the distribution of \widetilde{X}_t under P_ℓ . Using comparison theorems for stochastic differential equations (see e.g. [32, Theorem IX.3.7]), we can, for every choice of $\ell' \geq \ell$, couple a solution \widetilde{X}_t^1 of (24) starting from ℓ and a solution \widetilde{X}_t^2 starting from ℓ' so that $\widetilde{X}_t^2 \geq \widetilde{X}_t^1$ for all $t \geq 0$. It follows that

$$\begin{aligned} \int \Pi_t(\ell, d\ell') g(\ell') &\leq \frac{1}{\widehat{\rho}(\ell, \infty)} \int_\ell^\infty \widehat{\rho}(du) \int \Pi_t(u, d\ell') g(\ell') \\ &\leq \frac{1}{\widehat{\rho}(\ell, \infty)} \int_0^\infty \widehat{\rho}(du) \int \Pi_t(u, d\ell') g(\ell') \\ &= \frac{1}{\widehat{\rho}(\ell, \infty)} \int \widehat{\rho}(d\ell') g(\ell'). \end{aligned}$$

Applying the above display to the function $g \mathbf{1}_{(A, \infty)}$, where $A > 0$, we get that

$$0 \leq \int \Pi_t(\ell, d\ell') g(\ell') - \int_{[0, A]} \Pi_t(\ell, d\ell') g(\ell') \leq \frac{1}{\widehat{\rho}(\ell, \infty)} \cdot \int_{(A, \infty)} \widehat{\rho}(d\ell') g(\ell').$$

Since g is integrable with respect to $\widehat{\rho}$, the right-hand side can be made arbitrarily small, by choosing A large enough. Our claim now follows by letting $t \rightarrow \infty$, using the fact that $g \mathbf{1}_{[0, A]}$ is a bounded Borel function.

We can thus apply (25) to the nonnegative increasing function

$$g(\ell) = -\frac{1}{\varphi'(\ell)} e^{-\ell/2}$$

which is such that $\int g d\rho = -\int \varphi'(\ell) d\ell = 1$. Note that for this particular function g ,

$$E_\ell[g(\tilde{X}_t)] = Q_\ell^t[g(X_t)] = -\frac{e^{-\ell/2}}{\varphi'(\ell)} E_\ell\left[\exp\left(-\int_0^t ds (1 - 2\varphi(X_s))\right)\right].$$

It follows from (25) that, for every $\ell \geq 0$,

$$\begin{aligned} \lim_{t \rightarrow \infty} -\frac{e^{-\ell/2}}{\varphi'(\ell)} E_\ell\left[\exp\left(-\int_0^t ds (1 - 2\varphi(X_s))\right)\right] &= \int g d\hat{\rho} = \frac{1}{\rho((0, \infty))} \\ &= \frac{1}{\int_{(0, \infty)} ds \varphi'(s)^2 e^{s/2}}. \end{aligned}$$

This gives the first assertion of the proposition.

The second assertion is now easy. By the first assertion, there exists a constant K such that, for every $r \geq 0$,

$$E_0\left[\exp\left(-\int_0^r ds (1 - 2\varphi(X_s))\right)\right] \leq K.$$

Since the function φ is monotone decreasing, a comparison argument gives for every $\ell \geq 0$ and $r \geq 0$,

$$E_\ell\left[\exp\left(-\int_0^r ds (1 - 2\varphi(X_s))\right)\right] \leq E_0\left[\exp\left(-\int_0^r ds (1 - 2\varphi(X_s))\right)\right] \leq K.$$

This completes the proof of the proposition. \square

To simplify notation, we set

$$C_0 := \int_0^\infty ds \varphi'(s)^2 e^{s/2} = \iint \gamma(d\ell)\gamma(d\ell') \frac{\ell\ell'}{2(\ell + \ell' - 1)}.$$

COROLLARY 11. *For every $c > 0$,*

$$\lim_{r \rightarrow \infty} \Phi_r(c) = \Phi_\infty(c)$$

where

$$\Phi_\infty(c) = \frac{1}{C_0} \int \gamma(ds) \frac{cs}{2(c + s - 1)}.$$

PROOF. By definition, we have

$$\Phi_r(c) = \frac{c}{2} \int_0^\infty d\ell e^{-c\ell/2} E_\ell\left[\exp\left(-\int_0^r ds (1 - 2\varphi(X_s))\right)\right].$$

From Proposition 9 and an application of the dominated convergence theorem, we get

$$\lim_{r \rightarrow \infty} \Phi_r(c) = \frac{c}{2} \int_0^\infty d\ell e^{-c\ell/2} \times \left(-\frac{\varphi'(\ell)e^{\ell/2}}{C_0} \right).$$

The limit is identified with $\Phi_\infty(c)$ by a straightforward calculation. \square

3.3. *The invariant measure.* For the purposes of this section, it will be useful to introduce the set of all pairs consisting of a tree $\mathcal{T} \in \mathbb{T}$ and a distinguished geodesic ray \mathbf{v} , which we can represent by an element of $\{1, 2\}^{\mathbb{N}}$. We formally set

$$\mathbb{T}^* = \mathbb{T} \times \{1, 2\}^{\mathbb{N}}.$$

We can define shifts $(\sigma_r)_{r \geq 0}$ on \mathbb{T}^* in the following way. For $r = 0$, σ_r is just the identity mapping of \mathbb{T}^* . Then let $r > 0$ and $(\mathcal{T}, \mathbf{v}) \in \mathbb{T}^*$. Write $\mathbf{v} = (v_1, v_2, \dots)$ and $\mathbf{v}_n = (v_1, \dots, v_n)$ for every $n \geq 0$. Also let $x_{r, \mathbf{v}}$ be the unique element of \mathcal{T}_r such that $x_{r, \mathbf{v}} \prec \mathbf{v}$. Then, if $k = \min\{n \geq 0 : z_{\mathbf{v}_n} \geq r\}$, we set

$$\sigma_r(\mathcal{T}, \mathbf{v}) = \left(\mathcal{T}[x_{r, \mathbf{v}}], (v_{k+1}, v_{k+2}, \dots) \right).$$

Informally, $\sigma_r(\mathcal{T}, \mathbf{v})$ is obtained by taking the subtree of \mathcal{T} consisting of descendants of the vertex at height r on the distinguished geodesic ray, and keeping in this subtree the “same” geodesic ray. It is straightforward to verify that $\sigma_r \circ \sigma_s = \sigma_{r+s}$ for every $r, s \geq 0$.

Under the probability measure $\mathbb{P} \otimes P$, we can view (Γ, W_∞) as a random variable with values in \mathbb{T}^* . Write Θ^* for the distribution of (Γ, W_∞) . Then Θ^* is not invariant under the shifts σ_r , but Corollary 11 will give an invariant measure absolutely continuous with respect to Θ^* .

PROPOSITION 12. *The probability measure*

$$\Lambda^*(d\mathcal{T} d\mathbf{v}) := \Phi_\infty(\mathcal{C}(\mathcal{T})) \Theta^*(d\mathcal{T} d\mathbf{v})$$

is invariant under the shifts σ_r , $r \geq 0$.

PROOF. Let $r > 0$. We have

$$\sigma_r(\Gamma, W_\infty) = (\Gamma^{(r)}, W_\infty^{(r)}),$$

where $\Gamma^{(r)}$ and $W_\infty^{(r)}$ are as in the previous sections.

By Proposition 8, we have, for any bounded measurable function F on \mathbb{T} ,

$$\mathbb{E} \otimes E[F(\Gamma^{(r)})] = \int \Theta(d\mathcal{T}) \Phi_r(\mathcal{C}(\mathcal{T})) F(\mathcal{T}).$$

Write $\nu_{\mathcal{T}}$ for the harmonic measure of a Yule-type tree \mathcal{T} . At this point we use the flow property of harmonic measure. By Lemma 7 and the preceding identity, we have also, for any bounded measurable function F on \mathbb{T}^* ,

$$\begin{aligned}
\mathbb{E} \otimes E[F(\Gamma^{(r)}, W_{\infty}^{(r)})] &= \mathbb{E} \otimes E \left[\int \nu_{\Gamma^{(r)}}(d\mathbf{v}) F(\Gamma^{(r)}, \mathbf{v}) \right] \\
&= \int \Theta(d\mathcal{T}) \Phi_r(\mathcal{C}(\mathcal{T})) \int \nu_{\mathcal{T}}(d\mathbf{v}) F(\mathcal{T}, \mathbf{v}) \\
(26) \qquad \qquad \qquad &= \int \Theta^*(d\mathcal{T} d\mathbf{v}) \Phi_r(\mathcal{C}(\mathcal{T})) F(\mathcal{T}, \mathbf{v}),
\end{aligned}$$

since $\Theta^*(d\mathcal{T} d\mathbf{v}) = \Theta(d\mathcal{T})\nu_{\mathcal{T}}(d\mathbf{v})$ by construction.

If we now let $r \rightarrow \infty$, Corollary 11 gives

$$\lim_{r \rightarrow \infty} \mathbb{E} \otimes E[F(\Gamma^{(r)}, W_{\infty}^{(r)})] = \int \Theta^*(d\mathcal{T} d\mathbf{v}) \Phi_{\infty}(\mathcal{C}(\mathcal{T})) F(\mathcal{T}, \mathbf{v})$$

noting that the functions Φ_r are uniformly bounded thanks to the last assertion of Proposition 9. Let $s > 0$. If we replace F by $F \circ \sigma_s$ in the last convergence, observing that $F \circ \sigma_s(\Gamma^{(r)}, W_{\infty}^{(r)}) = F \circ \sigma_s \circ \sigma_r(\Gamma, W_{\infty}) = F(\Gamma^{(s+r)}, W_{\infty}^{(s+r)})$, we get

$$\int \Theta^*(d\mathcal{T} d\mathbf{v}) \Phi_{\infty}(\mathcal{C}(\mathcal{T})) F(\mathcal{T}, \mathbf{v}) = \int \Theta^*(d\mathcal{T} d\mathbf{v}) \Phi_{\infty}(\mathcal{C}(\mathcal{T})) F \circ \sigma_s(\mathcal{T}, \mathbf{v}),$$

which was the desired result. \square

PROPOSITION 13. *For every $r > 0$, the shift σ_r acting on the probability space $(\mathbb{T}^*, \Lambda^*)$ is ergodic.*

PROOF. We take $r = 1$ in this proof, and we write $\sigma = \sigma_1$ for simplicity. We essentially rely on ideas of [27] (see also [29, Chapter 16]). However our setting is different, because our trees are not discrete, and also because we consider ordered trees rather than unordered trees in [27]. For this reason, we will provide some details. We write π_1 for the canonical projection from \mathbb{T}^* onto \mathbb{T} , and let Λ be the image of Λ^* under this projection, so that

$$\Lambda(d\mathcal{T}) = \Phi_{\infty}(\mathcal{C}(\mathcal{T})) \Theta(d\mathcal{T}).$$

We define a transition kernel $\mathbf{p}(\mathcal{T}, d\mathcal{T}')$ on \mathbb{T} by setting

$$\mathbf{p}(\mathcal{T}, d\mathcal{T}') = \sum_{x \in \mathcal{T}_1} \nu_{\mathcal{T}}(\{\mathbf{v} \in \{1, 2\}^{\mathbb{N}} : x \prec \mathbf{v}\}) \delta_{\mathcal{T}[x]}(d\mathcal{T}').$$

Informally, under the probability measure $\mathbf{p}(\mathcal{T}, d\mathcal{T}')$, we choose one of the subtrees of \mathcal{T} above level 1 with probability equal to its harmonic measure. Then it follows from Proposition 12 that Λ is a stationary probability measure for the Markov chain with transition kernel \mathbf{p} . Indeed, Lemma 7 shows that we may obtain this Markov chain under its stationary measure Λ by considering the process

$$\mathcal{Z}_n(\mathcal{T}, \mathbf{v}) := \pi_1(\sigma_n(\mathcal{T}, \mathbf{v})), \quad n = 0, 1, 2, \dots$$

on the probability space $(\mathbb{T}^*, \Lambda^*)$. Note that $\mathcal{Z}_0(\mathcal{T}, \mathbf{v}) = \mathcal{T}$.

Write \mathbb{T}^∞ for the set of all sequences $(\mathcal{T}^0, \mathcal{T}^1, \dots)$ of elements of \mathbb{T} . By [29, Proposition 16.2], if a measurable subset F of \mathbb{T}^∞ is shift-invariant for the Markov chain \mathcal{Z} , in the sense that $\mathbf{1}_F(\mathcal{Z}_0, \mathcal{Z}_1, \dots) = \mathbf{1}_F(\mathcal{Z}_1, \mathcal{Z}_2, \dots)$ a.s., then there exists a measurable subset A of \mathbb{T} such that

$$\mathbf{1}_F(\mathcal{Z}_0, \mathcal{Z}_1, \dots) = \mathbf{1}_A(\mathcal{Z}_0), \quad \text{a.s.}$$

and moreover

$$\mathbf{p}(\mathcal{T}, A) = \mathbf{1}_A(\mathcal{T}), \quad \Lambda(d\mathcal{T}) \text{ a.s.}$$

We let $\widehat{\mathbb{T}}^\infty$ be the set of all sequences $(\mathcal{T}^0, \mathcal{T}^1, \dots)$ in \mathbb{T}^∞ , such that, for every integers $0 \leq i < j$, \mathcal{T}^j is a subtree of \mathcal{T}^i above generation $j - i$ (i.e., there exists a point $x \in \mathcal{T}_{j-i}^i$ such that $\mathcal{T}^j = \mathcal{T}^i[x]$). Note that $\widehat{\mathbb{T}}^\infty$ is a measurable subset of \mathbb{T}^∞ and that $(\mathcal{Z}_0(\mathcal{T}, \mathbf{v}), \mathcal{Z}_1(\mathcal{T}, \mathbf{v}), \dots) \in \widehat{\mathbb{T}}^\infty$ for every $(\mathcal{T}, \mathbf{v}) \in \mathbb{T}^*$. If $(\mathcal{T}^0, \mathcal{T}^1, \dots) \in \widehat{\mathbb{T}}^\infty$, there exists $\mathbf{v} \in \{1, 2\}^{\mathbb{N}}$ such that $\mathcal{T}^j = \mathcal{Z}_j(\mathcal{T}^0, \mathbf{v})$ for every $j \geq 0$, and we set $\Psi(\mathcal{T}^0, \mathcal{T}^1, \dots) = (\mathcal{T}^0, \mathbf{v})$. Note that \mathbf{v} is a priori not unique, but for the previous definition to make sense we take the smallest possible \mathbf{v} in lexicographical ordering (of course for the random trees that we consider later this uniqueness problem does not arise). In this way, we define a measurable mapping Ψ from $\widehat{\mathbb{T}}^\infty$ into \mathbb{T}^* , and we have $\Psi(\mathcal{Z}_0(\mathcal{T}, \mathbf{v}), \mathcal{Z}_1(\mathcal{T}, \mathbf{v}), \dots) = (\mathcal{T}, \mathbf{v})$, Λ^* a.s.

Let us now prove the statement of the proposition. We let B be a measurable subset of \mathbb{T}^* such that $\sigma^{-1}(B) = B$, and we aim at proving that $\Lambda^*(B) = 0$ or 1. To this end, we set $F = \Psi^{-1}(B)$, which is a measurable subset of $\widehat{\mathbb{T}}^\infty \subset \mathbb{T}^\infty$. Furthermore, we claim that F is shift-invariant. To see this, we have to verify that

$$\{(\mathcal{Z}_0, \mathcal{Z}_1, \dots) \in F\} = \{(\mathcal{Z}_1, \mathcal{Z}_2, \dots) \in F\}, \quad \text{a.s.}$$

or equivalently

$$\{\Psi(\mathcal{Z}_0, \mathcal{Z}_1, \dots) \in B\} = \{\Psi(\mathcal{Z}_1, \mathcal{Z}_2, \dots) \in B\}, \quad \text{a.s.}$$

But this is immediate since by construction $\Psi(\mathcal{Z}_1, \mathcal{Z}_2, \dots) = \sigma \circ \Psi(\mathcal{Z}_0, \mathcal{Z}_1, \dots)$ a.s. and $\sigma^{-1}(B) = B$ by assumption.

From preceding considerations, we then obtain that there exists a measurable subset A of \mathcal{T} , such that $(\mathcal{Z}_0, \mathcal{Z}_1, \dots) \in F$ if and only if $\mathcal{Z}_0 \in A$, a.s., and moreover $\mathbf{p}(\mathcal{T}, A) = \mathbf{1}_A(\mathcal{T})$, $\Lambda(\mathrm{d}\mathcal{T})$ a.s. Since $\Psi(\mathcal{Z}_0(\mathcal{T}, \mathbf{v}), \mathcal{Z}_1(\mathcal{T}, \mathbf{v}), \dots) = (\mathcal{T}, \mathbf{v})$, Λ^* a.s., it also follows that we have $(\mathcal{T}, \mathbf{v}) \in B$ if and only if $\mathcal{T} \in A$, Λ^* a.s.

However, from the property $\mathbf{p}(\mathcal{T}, A) = \mathbf{1}_A(\mathcal{T})$, $\Lambda(\mathrm{d}\mathcal{T})$ a.s., one can verify that $\Lambda(A) = 0$ or 1 . First note that this property also implies that $\mathbf{p}(\mathcal{T}, A) = \mathbf{1}_A(\mathcal{T})$, $\Theta(\mathrm{d}\mathcal{T})$ a.s. Hence, $\Theta(\mathrm{d}\mathcal{T})$ a.s., the tree \mathcal{T} belongs to A if and only if each of its subtrees above level 1 belong to A (it is clear that that the measure $\mathbf{p}(\mathcal{T}, \cdot)$ assigns a positive mass to each of these subtrees). Then, if $p_k = \mathbb{P}(\#\Gamma_1 = k)$, for every $k \geq 1$, the branching property of the Yule tree shows that

$$\Theta(A) = \sum_{k=1}^{\infty} p_k \Theta(A)^k$$

which is only possible if $\Theta(A) = 0$ or 1 , or equivalently $\Lambda(A) = 0$ or 1 . Finally, we also get that $\Lambda^*(B) = 0$ or 1 , which completes the proof. \square

3.4. End of the proof. Recall that $\nu_{\mathcal{T}}$ stands for the harmonic measure of a tree $\mathcal{T} \in \mathbb{T}$. With this notation, we have $\nu = \nu_{\Gamma}$. For every $r > 0$, we then consider the nonnegative measurable function F_r defined on \mathbb{T}^* by the formula

$$F_r(\mathcal{T}, \mathbf{v}) = \nu_{\mathcal{T}}(\mathcal{B}_{\mathcal{T}}(\mathbf{v}, r)),$$

where $\mathcal{B}_{\mathcal{T}}(\mathbf{v}, r)$ denotes the set of all geodesic rays of \mathcal{T} that coincide with the ray \mathbf{v} over the interval $[0, r]$. We claim that, for every $r, s > 0$, we have

$$F_{r+s} = F_r \times F_s \circ \sigma_r.$$

Indeed, if we write $\sigma_r(\mathcal{T}, \mathbf{v}) = (\mathcal{T}^{(r)}, \mathbf{v}^{(r)})$, this is equivalent to saying that

$$\frac{\nu_{\mathcal{T}}(\mathcal{B}_{\mathcal{T}}(\mathbf{v}, s+r))}{\nu_{\mathcal{T}}(\mathcal{B}_{\mathcal{T}}(\mathbf{v}, r))} = \nu_{\mathcal{T}^{(r)}}(\mathcal{B}_{\mathcal{T}^{(r)}}(\mathbf{v}^{(r)}, s)),$$

and the latter equality is an immediate consequence of Lemma 7.

If we set $G_r = -\log F_r \geq 0$, we have for every $r, s > 0$,

$$G_{s+r} = G_r + G_s \circ \sigma_r$$

and the ergodic theorem (with Proposition 13) implies that

$$\frac{1}{s} G_s \xrightarrow[s \rightarrow \infty]{\Lambda^* \text{ a.s.}} \Lambda^*(G_1).$$

Since Λ^* has a strictly positive density with respect to Θ^* , the latter convergence also holds Θ^* a.s. Recalling that Θ^* is the distribution of (Γ, W_∞) , this exactly gives the convergence (1), with $\beta = \Lambda^*(G_1)$. This completes the proof of Theorem 3, except that we have not checked that $\beta < 1$. We will do this in the next proposition, and then we will complete the proof of Proposition 4 by deriving the explicit formulas (3) for β in terms of the law γ of the conductance $\mathcal{C}(\Gamma)$.

PROPOSITION 14. *We have $\beta < 1$.*

PROOF. Here again, we strongly rely on ideas from [27] (see also [29, Chapter 16]). We start with some notation. If $\mathcal{T} \in \mathbb{T}$ and $x \in \mathcal{T}_1$, we set

$$\nu_{\mathcal{T}}^*(x) = \nu_{\mathcal{T}}(\{\mathbf{v} \in \{1, 2\}^{\mathbb{N}} : x \prec \mathbf{v}\}).$$

Clearly $(\nu_{\mathcal{T}}^*(x))_{x \in \mathcal{T}_1}$ is a probability distribution on \mathcal{T}_1 . We also set, for every $\mathcal{T} \in \mathbb{T}$,

$$U(\mathcal{T}) = \liminf_{r \rightarrow \infty} e^{-r} \#\mathcal{T}_r \in [0, \infty].$$

It is well known that the preceding liminf is a limit, $\Theta(d\mathcal{T})$ a.s., and that the distribution of $U(\mathcal{T})$ under $\Theta(d\mathcal{T})$ is exponential. It follows that, for Θ -almost every $\mathcal{T} \in \mathbb{T}$, we can also define, for every $x \in \mathcal{T}_1$,

$$U_{\mathcal{T}}(x) = \lim_{r \rightarrow \infty} e^{-r} \#\mathcal{T}_{r-1}[x] = \frac{1}{e} U(\mathcal{T}[x]),$$

and, if we set

$$u_{\mathcal{T}}(x) = \frac{U_{\mathcal{T}}(x)}{U(\mathcal{T})}$$

the collection $(u_{\mathcal{T}}(x))_{x \in \mathcal{T}_1}$ is a probability distribution on \mathcal{T}_1 .

By a concavity argument, we have

$$(27) \quad \sum_{x \in \mathcal{T}_1} \nu_{\mathcal{T}}^*(x) \log \left(\frac{u_{\mathcal{T}}(x)}{\nu_{\mathcal{T}}^*(x)} \right) \leq 0$$

and the inequality is even strict if $(\nu_{\mathcal{T}}^*(x))_{x \in \mathcal{T}_1} \neq (u_{\mathcal{T}}(x))_{x \in \mathcal{T}_1}$. It is easy to verify that the latter property holds with positive probability under Θ . To give a precise argument, recall the notation used in Section 2.2 to define the Yule tree Γ , and consider the event

$$E := \left\{ \mathcal{Y}_\emptyset < \frac{1}{4}, \mathcal{Y}_1 > 1, \mathcal{Y}_2 > \frac{3}{4} \right\} \cap \{\#\Gamma_1 = 9\},$$

which clearly has positive probability. On this event, write x_1, \dots, x_9 for the elements of Γ_1 listed in the lexicographical order of \mathcal{V} . On the event E , we have $\nu_\Gamma^*(x_1) > 1/8$, because Brownian motion (with drift $1/2$) has probability $1/2$ to hit the ancestor of x_1 at height $3/4$ before the other point of $\Gamma_{3/4}$, and then probability at least $1 - e^{-1/2} > 1/4$ to escape to infinity before returning to the first branching point of Γ . On the other hand, the branching property of the Yule tree shows that, conditionally on the event E (which only involves the part of the tree below height 1), the random variables $u_\Gamma(x_1), \dots, u_\Gamma(x_9)$ have the same distribution. It follows that we have $u_\Gamma(x_1) < \nu_\Gamma^*(x_1)$ with positive probability on E .

Next we have

$$\begin{aligned} \beta &= \Lambda^*(G_1) = \int \log \left(\frac{1}{\nu_{\mathcal{T}}(\mathcal{B}_{\mathcal{T}}(\mathbf{v}, 1))} \right) \Lambda^*(d\mathcal{T} d\mathbf{v}) \\ &= \int \sum_{x \in \mathcal{T}_1} \nu_{\mathcal{T}}^*(x) \log \left(\frac{1}{\nu_{\mathcal{T}}^*(x)} \right) \Lambda(d\mathcal{T}) \\ &< \int \sum_{x \in \mathcal{T}_1} \nu_{\mathcal{T}}^*(x) \log \left(\frac{1}{u_{\mathcal{T}}(x)} \right) \Lambda(d\mathcal{T}), \end{aligned}$$

where the strict inequality follows from (27) and the fact that $(\nu_{\mathcal{T}}^*(x))_{x \in \mathcal{T}_1} \neq (u_{\mathcal{T}}(x))_{x \in \mathcal{T}_1}$ with positive probability under Θ , hence also under Λ . Next, recalling the Markov chain (\mathcal{Z}_n) introduced in the proof of Proposition 13, we have

$$\begin{aligned} \int \sum_{x \in \mathcal{T}_1} \nu_{\mathcal{T}}^*(x) \log \left(\frac{1}{u_{\mathcal{T}}(x)} \right) \Lambda(d\mathcal{T}) &= \int \sum_{x \in \mathcal{T}_1} \nu_{\mathcal{T}}^*(x) \log \left(\frac{e^{U(\mathcal{T})}}{U(\mathcal{T}[x])} \right) \Lambda(d\mathcal{T}) \\ &= 1 + \int \log \left(\frac{U(\mathcal{Z}_0)}{U(\mathcal{Z}_1)} \right) \Lambda^*(d\mathcal{T} d\mathbf{v}) \\ &= 1 \end{aligned}$$

because \mathcal{Z}_0 and \mathcal{Z}_1 have the same distribution under Λ^* , and we also use the fact that $\log U(\mathcal{T})$ is integrable under $\Theta(d\mathcal{T})$ hence under $\Lambda(d\mathcal{T})$. Together with the preceding display, this completes the proof. \square

Proof of Proposition 4. The first assertion of Proposition 4 follows from Proposition 6. To complete the proof of Proposition 4, we start by establishing the first half of formula (3), that is,

$$(28) \quad \beta = \frac{2 \iiint \gamma(dr) \gamma(ds) \gamma(dt) \frac{rs}{r+s+t-1} \log \frac{r+t}{r}}{\iint \gamma(dr) \gamma(ds) \frac{rs}{r+s-1}}.$$

We use the notation of the beginning of this section, and we first fix $\varepsilon > 0$ and define a function H_ε on \mathbb{T}^* by setting

$$H_\varepsilon(\mathcal{T}, \mathbf{v}) = \begin{cases} 0 & \text{if } z_\emptyset \geq \varepsilon, \\ -\log \nu_{\mathcal{T}}(\{\mathbf{v}' \in \{1, 2\}^{\mathbb{N}} : \mathbf{v}_1 \prec \mathbf{v}'\}) & \text{if } z_\emptyset < \varepsilon, \end{cases}$$

where we write $\mathcal{T} = (z_v)_{v \in \mathcal{V}}$ as previously, and we recall the notation \mathbf{v}_n from the beginning of Section 3.3. Clearly, $H_\varepsilon(\mathcal{T}, \mathbf{v}) \leq G_\varepsilon(\mathcal{T}, \mathbf{v})$, and $H_\varepsilon(\mathcal{T}, \mathbf{v}) = G_\varepsilon(\mathcal{T}, \mathbf{v})$ if $z_{\mathbf{v}_1} \geq \varepsilon$. More generally, $H_\varepsilon \circ \sigma_r(\mathcal{T}, \mathbf{v}) = G_\varepsilon \circ \sigma_r(\mathcal{T}, \mathbf{v})$ if there is at most one index $i \geq 0$ such that $r \leq z_{\mathbf{v}_i} < r + \varepsilon$. It follows from these remarks that, for every integer $n \geq 1$,

$$(29) \quad G_1 \geq \sum_{k=0}^{n-1} H_{1/n} \circ \sigma_{k/n}$$

and, for every $(\mathcal{T}, \mathbf{v}) \in \mathbb{T}^*$,

$$(30) \quad G_1(\mathcal{T}, \mathbf{v}) = \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} H_{1/n} \circ \sigma_{k/n}(\mathcal{T}, \mathbf{v}).$$

Let us then investigate the behavior of $\Lambda^*(H_\varepsilon)$ when $\varepsilon \rightarrow 0$. It will be convenient to write $\mathcal{T}_{(1)}$ and $\mathcal{T}_{(2)}$ for the two ‘‘subtrees’’ of \mathcal{T} obtained at the first branching point (formally $\mathcal{T}_{(i)}$ corresponds to the collection $(z_{iv} - z_\emptyset)_{v \in \mathcal{V}}$, for $i = 1$ or 2). We observe that, if $i = 1$ or $i = 2$, the exit ray of Brownian motion on \mathcal{T} will belong to $\{(i, v_2, v_3, \dots) : (v_2, v_3, \dots) \in \{1, 2\}^{\mathbb{N}}\}$ with probability

$$\frac{\mathcal{C}(\mathcal{T}_{(i)})}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})}.$$

Thanks to this observation, we can write

$$\begin{aligned} & \Lambda^*(H_\varepsilon) \\ &= - \int \Theta(d\mathcal{T}) \Phi_\infty(\mathcal{C}(\mathcal{T})) \mathbf{1}_{\{z_\emptyset < \varepsilon\}} \left(\frac{\mathcal{C}(\mathcal{T}_{(1)})}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})} \log \frac{\mathcal{C}(\mathcal{T}_{(1)})}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})} \right. \\ & \quad \left. + \frac{\mathcal{C}(\mathcal{T}_{(2)})}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})} \log \frac{\mathcal{C}(\mathcal{T}_{(2)})}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})} \right) \\ &= -2 \int \Theta(d\mathcal{T}) \Phi_\infty(\mathcal{C}(\mathcal{T})) \mathbf{1}_{\{z_\emptyset < \varepsilon\}} \frac{\mathcal{C}(\mathcal{T}_{(1)})}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})} \log \frac{\mathcal{C}(\mathcal{T}_{(1)})}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})}, \end{aligned}$$

by a symmetry argument. An easy calculation gives

$$\mathcal{C}(\mathcal{T}) = \frac{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})}{e^{-z_\emptyset} + (1 - e^{-z_\emptyset})(\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)}))}.$$

Since, under $\Theta(d\mathcal{T})$, $\mathcal{T}_{(1)}$ and $\mathcal{T}_{(2)}$ are independent and distributed according to Θ , and are also independent of z_\emptyset , we get

$$\begin{aligned} \Lambda^*(H_\varepsilon) &= -2 \iint \Theta(d\mathcal{T})\Theta(d\mathcal{T}') \frac{\mathcal{C}(\mathcal{T})}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')} \log \frac{\mathcal{C}(\mathcal{T})}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')} \\ &\quad \times \int_0^\varepsilon dz e^{-z} \Phi_\infty \left(\frac{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')}{e^{-z} + (1 - e^{-z})(\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}'))} \right). \end{aligned}$$

Note that the function $(\mathcal{T}, \mathcal{T}') \mapsto \frac{\mathcal{C}(\mathcal{T})}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')} \log \frac{\mathcal{C}(\mathcal{T})}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')}$ is integrable with respect to the measure $\Theta(d\mathcal{T})\Theta(d\mathcal{T}')$, and that Φ_∞ is bounded and continuous. We can thus let $\varepsilon \rightarrow 0$ in the preceding expression and get

$$(31) \quad \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \Lambda^*(H_\varepsilon) = -2 \iint \Theta(d\mathcal{T})\Theta(d\mathcal{T}') \Phi_\infty(\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')) \\ \times \frac{\mathcal{C}(\mathcal{T})}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')} \log \frac{\mathcal{C}(\mathcal{T})}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')}.$$

Since the limit in the preceding display is finite, we can use (30) and Fatou's lemma to get that $\Lambda^*(G_1) < \infty$, and then (29) (to justify dominated convergence) and (30) again to obtain that

$$\Lambda^*(G_1) = \lim_{n \rightarrow \infty} n \Lambda^*(H_{1/n})$$

coincides with the right-hand side of (31). Finally, we use the expression of Φ_∞ to obtain formula (28).

We will now establish the second half of formula (3), which will complete the proof of Proposition 4. We let $\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2$ be independent and distributed according to γ under the probability measure \mathbb{P} . Then, the denominator of the right-hand side of (28) can be written as

$$\mathbb{E} \left[\frac{\mathcal{C}_0 \mathcal{C}_1}{\mathcal{C}_0 + \mathcal{C}_1 - 1} \right].$$

On the other hand, the numerator is equal to

$$\begin{aligned} &2\mathbb{E} \left[\frac{\mathcal{C}_0 \mathcal{C}_1}{\mathcal{C}_0 + \mathcal{C}_1 + \mathcal{C}_2 - 1} \log \left(\frac{\mathcal{C}_1 + \mathcal{C}_2}{\mathcal{C}_1} \right) \right] \\ &= \mathbb{E} \left[\frac{\mathcal{C}_0(\mathcal{C}_1 + \mathcal{C}_2) \log(\mathcal{C}_1 + \mathcal{C}_2)}{\mathcal{C}_0 + \mathcal{C}_1 + \mathcal{C}_2 - 1} \right] - \mathbb{E} \left[\frac{(\mathcal{C}_0 + \mathcal{C}_2)\mathcal{C}_1 \log(\mathcal{C}_1)}{\mathcal{C}_0 + \mathcal{C}_1 + \mathcal{C}_2 - 1} \right] \\ &= \mathbb{E}[f(\mathcal{C}_1 + \mathcal{C}_2)] - \mathbb{E}[g(\mathcal{C}_1 + \mathcal{C}_2)] \end{aligned}$$

where we have set, for every $x \geq 1$,

$$f(x) = \mathbb{E} \left[\frac{\mathcal{C}_0 x}{\mathcal{C}_0 + x - 1} \log x \right] \quad \text{and} \quad g(x) = \mathbb{E} \left[\frac{\mathcal{C}_0 x}{\mathcal{C}_0 + x - 1} \log \mathcal{C}_0 \right].$$

Using (13), we replace $\mathbb{E}[f(\mathcal{C}_1 + \mathcal{C}_2)]$ by $\mathbb{E}[f(\mathcal{C}_1)] + \mathbb{E}[\mathcal{C}_1(\mathcal{C}_1 - 1)f'(\mathcal{C}_1)]$, and similarly for g , to obtain

$$\begin{aligned} \mathbb{E}[f(\mathcal{C}_1 + \mathcal{C}_2)] - \mathbb{E}[g(\mathcal{C}_1 + \mathcal{C}_2)] &= \mathbb{E} \left[\frac{\mathcal{C}_0 \mathcal{C}_1}{\mathcal{C}_0 + \mathcal{C}_1 - 1} \log \mathcal{C}_1 \right] \\ &\quad + \mathbb{E} \left[\frac{\mathcal{C}_0(\mathcal{C}_0 - 1)\mathcal{C}_1(\mathcal{C}_1 - 1)}{(\mathcal{C}_0 + \mathcal{C}_1 - 1)^2} \log \mathcal{C}_1 \right] + \mathbb{E} \left[\frac{\mathcal{C}_0 \mathcal{C}_1 (\mathcal{C}_1 - 1)}{\mathcal{C}_0 + \mathcal{C}_1 - 1} \right] \\ &\quad - \mathbb{E} \left[\frac{\mathcal{C}_0 \mathcal{C}_1}{\mathcal{C}_0 + \mathcal{C}_1 - 1} \log \mathcal{C}_0 \right] - \mathbb{E} \left[\frac{\mathcal{C}_0(\mathcal{C}_0 - 1)\mathcal{C}_1(\mathcal{C}_1 - 1)}{(\mathcal{C}_0 + \mathcal{C}_1 - 1)^2} \log \mathcal{C}_0 \right] \\ &= \mathbb{E} \left[\frac{\mathcal{C}_0 \mathcal{C}_1 (\mathcal{C}_0 - 1)}{\mathcal{C}_0 + \mathcal{C}_1 - 1} \right] \\ &= \frac{1}{2} \mathbb{E} \left[\frac{\mathcal{C}_0 \mathcal{C}_1 (\mathcal{C}_0 + \mathcal{C}_1 - 1) - \mathcal{C}_0 \mathcal{C}_1}{\mathcal{C}_0 + \mathcal{C}_1 - 1} \right] \\ &= \frac{1}{2} \left(\mathbb{E}[\mathcal{C}_0]^2 - \mathbb{E} \left[\frac{\mathcal{C}_0 \mathcal{C}_1}{\mathcal{C}_0 + \mathcal{C}_1 - 1} \right] \right). \end{aligned}$$

If we substitute this in (28), we arrive at

$$2\beta = \frac{\mathbb{E}[\mathcal{C}_0]^2}{\mathbb{E} \left[\frac{\mathcal{C}_0 \mathcal{C}_1}{\mathcal{C}_0 + \mathcal{C}_1 - 1} \right]} - 1,$$

which gives the second half of (3) and completes the proof of Proposition 4.

Remark. Despite all that is known about the distribution γ (see Section 2.3), it requires some work to derive the fact that $\beta < 1$ (Proposition 14) from the explicit formulas of Proposition 4. The approximate numerical value $\beta = 0.78\dots$ is obtained by first estimating γ using Proposition 6 (or more precisely the convergence of $\Phi^k(\lambda)$ to γ , for any probability measure λ on $[1, \infty)$), and then applying a Monte-Carlo method to evaluate the integrals in the right-hand side of (3).

4. Discrete random trees. In this section we prove Theorem 1 and Corollary 2. We first explain why discrete reduced trees converge modulo a suitable rescaling toward the continuous reduced tree Δ . This leads to a first connection between the discrete harmonic measures and the continuous one (Proposition 18). Combining this result with Theorem 3, one gets a first estimate in the direction of Theorem 1 (Corollary 19). The recursive properties of Galton–Watson trees are then used to complete the proof of Theorem 1. Corollary 2 is proved at the end of the section.

4.1. *Notation for trees.* We consider discrete rooted ordered trees, which are also called plane trees in combinatorics. A plane tree τ is a finite subset of

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n,$$

where $\mathbb{N}^0 = \{\emptyset\}$, such that the following holds:

- (i) $\emptyset \in \tau$.
- (ii) If $u = (u_1, \dots, u_n) \in \tau \setminus \{\emptyset\}$ then $\hat{u} := (u_1, \dots, u_{n-1}) \in \tau$.
- (iii) For every $u = (u_1, \dots, u_n) \in \tau$, there exists an integer $k_u(\tau) \geq 0$ such that, for every $j \in \mathbb{N}$, $(u_1, \dots, u_n, j) \in \tau$ if and only if $1 \leq j \leq k_u(\tau)$.

In this section, we say tree instead of plane tree. We often view a tree τ as a graph whose vertices are the elements of τ and whose edges are the pairs $\{\hat{u}, u\}$ for all $u \in \tau \setminus \{\emptyset\}$.

We will use the notation and terminology introduced at the beginning of Section 2.1 in a slightly different setting. In particular, $|u|$ is the generation of u , uv denotes the concatenation of u and v , \prec stands for the genealogical order and $u \wedge v$ is the maximal element of $\{w \in \mathcal{U} : w \prec u \text{ and } w \prec v\}$.

The height of a tree τ is

$$h(\tau) = \max\{|v| : v \in \tau\}.$$

We write \mathcal{T} for the set of all trees, and \mathcal{T}_n for the set of all trees with height n .

Let τ be a tree. The set τ is equipped with the distance

$$d(v, w) = \frac{1}{2}(|v| + |w| - 2|v \wedge w|).$$

Notice that this is half the usual graph distance. We will write $B_\tau(v, r)$, or simply $B(v, r)$ if there is no ambiguity, for the closed ball of radius r centered at v , with respect to the distance d , in the tree τ .

The set of all vertices of τ at generation n is denoted by

$$\tau_n := \{v \in \tau : |v| = n\}.$$

If $v \in \tau$, the subtree of descendants of v is

$$\tilde{\tau}[v] := \{v' \in \tau : v \prec v'\}.$$

Note that $\tilde{\tau}[v]$ is not a tree with our definitions, but we turn it into a tree by relabelling its vertices, setting

$$\tau[v] := \{w \in \mathcal{U} : vw \in \tau\}.$$

If $v \in \tau$, then for every $i \in \{0, 1, \dots, |v|\}$ we write $\langle v \rangle_i$ for the ancestor of v at generation i . Suppose that $|v| = n$. Then $B_\tau(v, i) \cap \tau_n = \tilde{\tau}[\langle v \rangle_{n-i}] \cap \tau_n$, for every $i \in \{0, 1, \dots, n\}$. This simple observation will be used several times below.

Galton–Watson trees. Let θ be a probability measure on \mathbb{Z}_+ , and assume that θ has mean one and finite variance $\sigma^2 > 0$. There exists a unique probability measure $\text{GW}_\theta(d\tau)$ on \mathcal{T} such that the following two properties hold:

- (i) The law of $k_\emptyset(\tau)$ under $\text{GW}_\theta(d\tau)$ is θ .
- (ii) Let $k \geq 1$ such that $\theta(k) > 0$. Then under $\text{GW}_\theta(d\tau \mid k_\emptyset(\tau) = k)$, the subtrees $\tau[1], \dots, \tau[k]$ are independent and distributed according to GW_θ .

A random tree distributed according to GW_θ will be called a Galton–Watson tree with offspring distribution θ (see e.g. [23] for a discussion of Galton–Watson trees).

For every integer $n \geq 0$, we let $\mathbb{T}^{(n)}$ be a Galton–Watson tree with offspring distribution θ , conditioned on non-extinction at generation n . In particular, $\mathbb{T}^{(0)}$ is just a Galton–Watson tree with offspring distribution θ . We suppose that the random trees $\mathbb{T}^{(n)}$ are defined under the probability measure \mathbb{P} .

We let \mathbb{T}^{*n} be the reduced tree associated with $\mathbb{T}^{(n)}$, which consists of all vertices of $\mathbb{T}^{(n)}$ that have (at least) one descendant at generation n . A priori \mathbb{T}^{*n} is not a tree in the sense of the preceding definition. However we can relabel the vertices of \mathbb{T}^{*n} , preserving both the lexicographical order and the genealogical order, so that \mathbb{T}^{*n} becomes a tree in the sense of our definitions. We will always assume that this relabelling has been done.

Note that $|u| \leq n$ for every $u \in \mathbb{T}^{*n}$. It will be convenient to introduce truncations of \mathbb{T}^{*n} . For every $s \in [0, n]$, we set

$$R_s(\mathbb{T}^{*n}) = \{v \in \mathbb{T}^{*n} : |v| \leq n - \lfloor s \rfloor\}.$$

We then consider simple random walk on \mathbb{T}^{*n} , starting from the root \emptyset , which we denote by $Z^n = (Z_k^n)_{k \geq 0}$. This random walk is defined under the probability measure P (as previously, it is important to distinguish the probability measures governing the trees on one hand, the random walks on the other hand).

We let

$$H_n = \inf\{k \geq 0 : |Z_k^n| = n\}$$

be the first hitting time of generation n by Z^n , and we set

$$\Sigma_n = Z_{H_n}^n.$$

The discrete harmonic measure μ_n , is the law of Σ_n under P . Notice that μ_n is a probability measure on the set T_n^{*n} of all vertices of T^{*n} at generation n .

We start with a lemma that gives bounds on the size of level sets in T^{*n} .

LEMMA 15. *There exists a constant C depending only on θ such that, for every integer $n \geq 2$ and every integer p such that $1 \leq p \leq n/2$,*

$$\mathbb{E}[(\log \#\mathsf{T}_{n-p}^{*n})^4]^{1/4} \leq C \log \frac{n}{p} \quad \text{and} \quad \mathbb{E}[(\log \#\mathsf{T}_n^{*n})^4]^{1/4} \leq C \log n.$$

PROOF. Set $q_n = \mathbb{P}(h(\mathsf{T}^{(0)}) \geq n)$. By a standard result (Theorem 9.1 of [5, Chapter 1]), we have

$$(32) \quad q_n \sim \frac{2}{n\sigma^2}, \quad \text{as } n \rightarrow \infty.$$

Then, for every $p \in \{0, 1, \dots, n\}$,

$$\begin{aligned} \mathbb{E}[\#\mathsf{T}_{n-p}^{*n}] &= \mathbb{E}[\#\{v \in \mathsf{T}_{n-p}^{(n)} : h(\mathsf{T}^{(n)}[v]) \geq p\}] \\ &= (q_n)^{-1} \mathbb{E}[\#\{v \in \mathsf{T}_{n-p}^{(0)} : h(\mathsf{T}^{(0)}[v]) \geq p\}]. \end{aligned}$$

By the branching property of Galton–Watson trees, the conditional distribution of $\#\{v \in \mathsf{T}_{n-p}^{(0)} : h(\mathsf{T}^{(0)}[v]) \geq p\}$ knowing that $\#\mathsf{T}_{n-p}^{(0)} = k$ is the binomial distribution $\mathcal{B}(k, q_p)$. Hence,

$$\mathbb{E}[\#\mathsf{T}_{n-p}^{*n}] = \frac{q_p \mathbb{E}[\#\mathsf{T}_{n-p}^{(0)}]}{q_n} = \frac{q_p}{q_n}.$$

We can find $a > 0$ such that the function $x \rightarrow (\log(a+x))^4$ is concave over $[1, \infty)$. Then,

$$\begin{aligned} \mathbb{E}[(\log \#\mathsf{T}_{n-p}^{*n})^4]^{1/4} &\leq \mathbb{E}[(\log(a + \#\mathsf{T}_{n-p}^{*n}))^4]^{1/4} \leq \log(a + \mathbb{E}[\#\mathsf{T}_{n-p}^{*n}]) \\ &= \log(a + \frac{q_p}{q_n}), \end{aligned}$$

and the bounds of the lemma easily follow from (32). □

4.2. Discrete and continuous reduced trees.

4.2.1. *Convergence of discrete reduced trees.* Recall from Section 2.1 the definition of the continuous reduced tree Δ . For every $\varepsilon \in (0, 1)$, we have set $\Delta_\varepsilon = \{x \in \Delta : H(x) \leq 1 - \varepsilon\}$. We will implicitly use the fact that, for every fixed ε , there is a.s. no branching point of Δ at height $1 - \varepsilon$. The skeleton of Δ_ε is defined as

$$\begin{aligned} \text{Sk}(\Delta_\varepsilon) &= \{\emptyset\} \cup \{v \in \mathcal{V} \setminus \{\emptyset\} : Y_{\hat{v}} \leq 1 - \varepsilon\} \\ &= \{\emptyset\} \cup \{v \in \mathcal{V} \setminus \{\emptyset\} : (\hat{v}, Y_{\hat{v}}) \in \Delta_\varepsilon\}. \end{aligned}$$

Consider then a tree $\tau \in \mathcal{T}$ such that every vertex of τ has either 0, 1 or 2 children. It will be convenient to write \mathcal{T}_{bin} for the collection of all such trees. With τ we associate another tree denoted by $[\tau]$, which is obtained by “removing” all vertices that have exactly one child. More precisely, write $\mathcal{S}(\tau)$ for the set of all vertices v of τ having 0 or 2 children. Then we can find a unique tree $[\tau]$ such that there exists a bijection $u \rightarrow w_u$ from $[\tau]$ onto $\mathcal{S}(\tau)$ that preserves both the genealogical order and the lexicographical order of vertices. We call this bijection the canonical bijection from $[\tau]$ onto $\mathcal{S}(\tau)$.

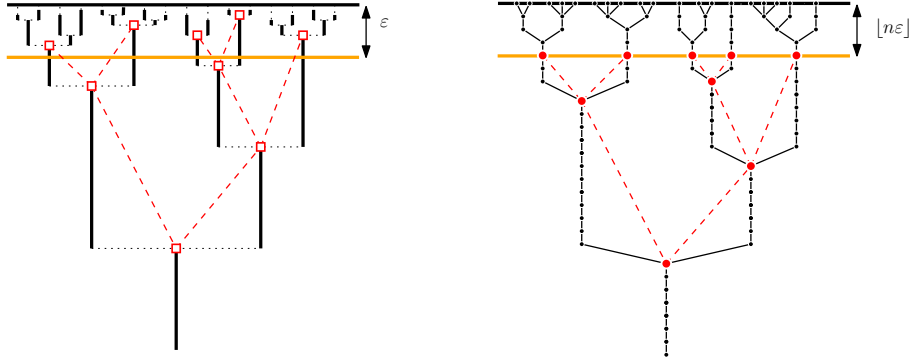


FIG 6. *Setting of Proposition 16. On the left, the tree Δ , its truncation Δ_ε and the skeleton $\text{Sk}(\Delta_\varepsilon)$. On the right, a large reduced tree \mathbb{T}^{*n} of height n , its truncation $R_{\varepsilon n}(\mathbb{T}^{*n})$ and the associated binary tree $[\mathbb{T}^{*n}]$.*

PROPOSITION 16. *We can construct the reduced trees \mathbb{T}^{*n} and the (continuous) tree Δ on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ so that the following properties hold for every fixed $\varepsilon \in (0, 1)$ with \mathbb{P} -probability one.*

- (i) *For every sufficiently large integer n , we have $R_{\varepsilon n}(\mathbb{T}^{*n}) \in \mathcal{T}_{\text{bin}}$ and $[\mathbb{T}^{*n}] = \text{Sk}(\Delta_\varepsilon)$.*

- (ii) For every sufficiently large n , such that the properties stated in (i) hold, and for every $u \in \text{Sk}(\Delta_\varepsilon)$, let $w_u^{n,\varepsilon}$ denote the vertex of $\mathcal{S}(R_{\varepsilon n}(\mathbb{T}^{*n}))$ corresponding to u via the canonical bijection from $[R_{\varepsilon n}(\mathbb{T}^{*n})]$ onto $\mathcal{S}(R_{\varepsilon n}(\mathbb{T}^{*n}))$. Then we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} |w_u^{n,\varepsilon}| = Y_u \wedge (1 - \varepsilon).$$

See Figure 6 for an illustration of Proposition 16. This proposition is essentially a consequence of classical results on the convergence in distribution of reduced critical Galton–Watson trees, see in particular [33] and [14]. A simple way of proving Proposition 16 is to use the convergence in distribution of the rescaled contour functions associated with the trees $\mathbb{T}^{(n)}$ toward a Brownian excursion with height greater than 1 (see [23, Corollary 1.13]). By using the Skorokhod representation theorem, one may assume that the trees $\mathbb{T}^{(n)}$ and the Brownian excursion are constructed so that the latter convergence holds almost surely. We then use the relation between the Brownian excursion with height greater than 1 and the continuous reduced tree Δ , which can be found in [22, Section 5] (this is a particular case of a more general result connecting reduced Lévy trees with the so-called height process, see [11, Section 2.7]). Let us briefly explain this relation.

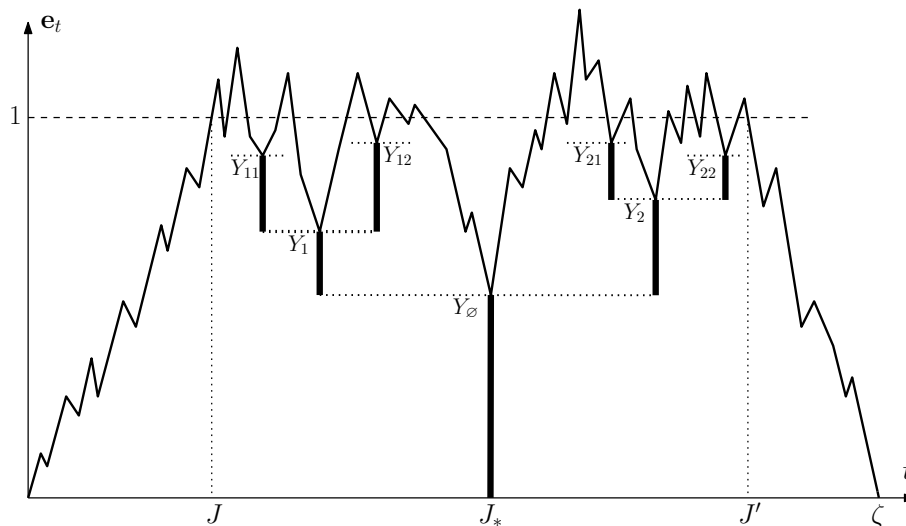


FIG 7. The relation between the Brownian excursion with height greater than 1 and the continuous reduced tree Δ . For each $0 < a < 1$, the number of vertices of Δ at height a corresponds to the number of “subexcursions” above height a that hit level 1.

We write $(\mathbf{e}_t)_{0 \leq t \leq \zeta}$ for a Brownian excursion conditioned to hit level 1. We associate with this process a collection $(Y_v)_{v \in \mathcal{V}}$ of nonnegative random variables defined as follows. If $J = \inf\{t \in [0, \zeta] : \mathbf{e}_t = 1\}$, $J' = \sup\{t \in [0, \zeta] : \mathbf{e}_t = 1\}$, we set $Y_\emptyset := \min\{\mathbf{e}_t : J \leq t \leq J'\}$ and we also let J_* be the a.s. unique time in $[J, J']$ such that $\mathbf{e}_{J_*} = Y_\emptyset$. Then, we let Y_1 be the minimum of \mathbf{e} between J and $\sup\{t \leq J_* : \mathbf{e}_t = 1\}$, and Y_2 be the minimum of \mathbf{e} between $\inf\{t \geq J_* : \mathbf{e}_t = 1\}$ and J' . The construction is continued by induction (compare Figure 7 and Figure 4). According to [22, Section 5], the collection $(Y_v)_{v \in \mathcal{V}}$ has the distribution described in Section 2.1: This shows that the tree Δ can be embedded in the graph of \mathbf{e} in the way suggested by Figure 7. Moreover, for every $a \in [0, 1)$, the number of vertices of the tree Δ at height a corresponds to the number of excursions of \mathbf{e} above level a that hit height 1 (to be precise, a vertex which is a branching point of the tree should be counted twice).

Once we know that the rescaled contour functions associated with the trees $\mathbb{T}^{(n)}$ converge a.s. to a Brownian excursion with height greater than 1, the various assertions of Proposition 16 follow, with the continuous reduced tree Δ constructed as explained above from the limiting Brownian excursion. We leave the details to the reader.

Let us comment on the properties stated in Proposition 16. In property (ii), we have $Y_u > 1 - \varepsilon$ if and only if u is a leaf (i.e. a vertex with no child) of $\text{Sk}(\Delta_\varepsilon)$. Furthermore, if u is a vertex of $\text{Sk}(\Delta_\varepsilon)$ which is not a leaf, the vertex $w_u^{n, \varepsilon}$, which is well defined for n large enough, does not depend on ε . More precisely, suppose that $0 < \delta < \varepsilon$, and suppose that n is sufficiently large so that the properties stated in (i) hold as well as the same properties with ε replaced by δ . Then, if $u \in \text{Sk}(\Delta_\varepsilon)$ is not a leaf of $\text{Sk}(\Delta_\varepsilon)$, we must have $w_u^{n, \varepsilon} = w_u^{n, \delta}$. On the other hand, if u is a leaf of $\text{Sk}(\Delta_\varepsilon)$, then we must have $|w_u^{n, \varepsilon}| = n - \lfloor \varepsilon n \rfloor$, and $w_u^{n, \varepsilon}$ is an ancestor of $w_u^{n, \delta}$. We leave the verification of these properties to the reader.

4.2.2. Convergence of conductances. Let i be a positive integer and let $\tau \in \mathcal{T}$ be a tree such that $h(\tau) \geq i$. Consider the new graph τ' obtained by adding to the graph τ an edge between the root \emptyset and an extra vertex ∂ . We let $\mathcal{C}_i(\tau)$ be the probability that simple random walk on τ' starting from \emptyset hits generation i of τ before hitting the vertex ∂ . The notation is justified by the fact that $\mathcal{C}_i(\tau)$ can be interpreted as the effective conductance between ∂ and generation i of τ in the graph τ' , see [29, Chapter 2].

PROPOSITION 17. *Suppose that the reduced trees \mathbb{T}^{*n} and the (continuous) tree Δ are constructed so that the properties stated in Proposition 16 hold, and that the Yule tree Γ is obtained from Δ as explained in Section 2.2.*

Then

$$n \mathcal{C}_n(\mathbb{T}^{*n}) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \mathcal{C}(\Gamma).$$

We omit the easy proof, as this result is not needed for the proof of Theorem 1.

4.2.3. *Convergence of harmonic measures.* Our goal is now to verify that the discrete harmonic measures μ_n converge in some sense to the continuous harmonic measure μ defined in Section 2.1.

For every $x \in \partial\Delta_\varepsilon = \{z \in \Delta : H(z) = 1 - \varepsilon\}$, we set

$$\mu^\varepsilon(x) = \mu(\{y \in \partial\Delta : x \prec y\}).$$

Similarly, we define a probability measure μ_n^ε on $\mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}$ by setting

$$\mu_n^\varepsilon(u) = \mu_n(\{v \in \mathbb{T}_n : u \prec v\}),$$

for every $u \in \mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}$. Clearly, μ_n^ε is also the distribution of $\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}$.

PROPOSITION 18. *Suppose that the reduced trees \mathbb{T}^{*n} and the (continuous) tree Δ have been constructed so that the properties of Proposition 16 hold, and recall the notation $(w_u^{n,\varepsilon})_{u \in \text{Sk}(\Delta_\varepsilon)}$ introduced in this proposition. Then \mathbb{P} a.s. for every $x = (v, 1 - \varepsilon) \in \partial\Delta_\varepsilon$,*

$$\lim_{n \rightarrow \infty} \mu_n^\varepsilon(w_v^{n,\varepsilon}) = \mu^\varepsilon(x).$$

PROOF. Let $\delta \in (0, \varepsilon)$ and set $T_\delta = \inf\{t \geq 0 : H(B_t) = 1 - \delta\} < T$. Define a probability measure $\mu^{\varepsilon,(\delta)}$ on $\partial\Delta_\varepsilon$ by setting for every $x \in \partial\Delta_\varepsilon$,

$$\mu^{\varepsilon,(\delta)}(x) = P(x \prec B_{T_\delta}).$$

Similarly, we write $\mu_n^{(\delta)}$ for the distribution of the hitting point of generation $n - \lfloor \delta n \rfloor$ by random walk on \mathbb{T}^{*n} started from \emptyset , and we define a probability measure $\mu_n^{\varepsilon,(\delta)}$ on $\mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}$ by setting

$$\mu_n^{\varepsilon,(\delta)}(v) = \mu_n^{(\delta)}(\{w \in \mathbb{T}_{n-\lfloor \delta n \rfloor}^{*n} : v \prec w\}),$$

for every $v \in \mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}$.

It is easy to verify that

$$\lim_{\delta \rightarrow 0} \mu^{\varepsilon,(\delta)}(x) = \mu^\varepsilon(x)$$

for every $x \in \partial\Delta_\varepsilon$, \mathbb{P} -a.s. Indeed we have the bound $|\mu^{\varepsilon,(\delta)}(x) - \mu^\varepsilon(x)| \leq \delta/\varepsilon$, which follows from the fact that there is probability at least $1 - \delta/\varepsilon$ that after time T_δ Brownian motion will hit the boundary $\partial\Delta$ before returning to height $1 - \varepsilon$ (and if this event occurs then for $x \in \partial\Delta_\varepsilon$, we have $x \prec B_T$ if and only if $x \prec B_{T_\delta}$). By similar arguments, one has \mathbb{P} -a.s.

$$\lim_{\delta \rightarrow 0} \left(\limsup_{n \rightarrow \infty} \left(\sup_{v \in \mathbb{T}_{n - \lfloor \varepsilon n \rfloor}^{*n}} |\mu_n^{\varepsilon,(\delta)}(v) - \mu_n^\varepsilon(v)| \right) \right) = 0.$$

In view of the preceding remarks, the convergence of the proposition will follow if we can verify that for every fixed $\delta \in (0, \varepsilon)$, we have a.s. for every $x = (u, 1 - \varepsilon) \in \partial\Delta_\varepsilon$,

$$(33) \quad \lim_{n \rightarrow \infty} \mu_n^{\varepsilon,(\delta)}(w_u^{n,\varepsilon}) = \mu^{\varepsilon,(\delta)}(x).$$

By considering the successive passage times of Brownian motion stopped at time T_δ in the set $\{(v, Y_v \wedge (1 - \delta)) : v \in \text{Sk}(\Delta_\delta)\}$, we get a Markov chain $X^{(\delta)}$, which is absorbed in the set $\{(v, 1 - \delta) : v \text{ is a leaf of } \text{Sk}(\Delta_\delta)\}$, and whose transition kernels are explicitly described in terms of the quantities Y_v , $v \in \text{Sk}(\Delta_\delta)$.

Let n be sufficiently large so that assertions (i) and (ii) of Proposition 16 hold with ε replaced by δ , and consider random walk on \mathbb{T}^{*n} started from \emptyset and stopped at the first hitting time of generation $n - \lfloor \delta n \rfloor$. By considering the successive passage times of this random walk in the set $\{w_v^{n,\delta} : v \in \text{Sk}(\Delta_\delta)\}$, we again get a Markov chain $X^{(\delta),n}$, which is absorbed in the set $\{w_v^{n,\delta} : v \text{ is a leaf of } \text{Sk}(\Delta_\delta)\}$ and whose transition kernels are explicit in terms of the quantities $|w_v^n|$, $v \in \text{Sk}(\Delta_\delta)$.

Identifying both sets $\{(v, Y_v \wedge (1 - \delta)) : v \in \text{Sk}(\Delta_\delta)\}$ and $\{w_v^{n,\delta} : v \in \text{Sk}(\Delta_\delta)\}$ with $\text{Sk}(\Delta_\delta)$, we can view $X^{(\delta)}$ and $X^{(\delta),n}$ as Markov chains with values in the set $\text{Sk}(\Delta_\delta)$, and then assertion (ii) of Proposition 16 implies that the transition kernels of $X^{(\delta),n}$ converge to those of $X^{(\delta)}$. Write $X_\infty^{(\delta)}$ for the absorption point of $X^{(\delta)}$, and similarly write $X_\infty^{(\delta),n}$ for the absorption point of $X^{(\delta),n}$. We thus obtain that the distribution of $X_\infty^{(\delta),n}$ converges to that of $X_\infty^{(\delta)}$. Consequently, for every $u \in \mathcal{V}$ such that $x = (u, 1 - \varepsilon) \in \partial\Delta_\varepsilon$, we have

$$\lim_{n \rightarrow \infty} P(u \prec X_\infty^{(\delta),n}) = P(u \prec X_\infty^{(\delta)}).$$

However, from our definitions, we have

$$P(u \prec X_\infty^{(\delta)}) = \mu^{\varepsilon,(\delta)}(x),$$

and, for n sufficiently large, noting that $w_u^{n,\varepsilon}$ coincides with the ancestor of $w_u^{n,\delta}$ at generation $n - \lfloor \varepsilon n \rfloor$ (see the remarks after Proposition 16),

$$P(u \prec X_\infty^{(\delta),n}) = \mu_n^{\varepsilon,(\delta)}(w_u^{n,\varepsilon}).$$

This completes the proof of (33) and of the proposition. \square

Recall that, if $v \in \mathcal{U}$, $\langle v \rangle_i$ is the ancestor of v at generation $i \leq |v|$.

COROLLARY 19. *Let $\xi \in (0, 1)$. We can find $\varepsilon_0 \in (0, 1/2)$ such that the following holds. For every $\varepsilon \in (0, \varepsilon_0)$, there exists $n_0 \geq 0$ such that for every $n \geq n_0$ we have*

$$\mathbb{E} \otimes E \left[\left| \log \mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}) - \beta \log \varepsilon \right|^2 \right] \leq \xi |\log \varepsilon|^2.$$

PROOF. Recall our notation $\mathcal{B}_d(x, r)$ for the closed ball of radius r centered at $x \in \Delta$. Fix $\eta \in (0, 1)$. Since B_T is distributed according to μ , it follows from Theorem 3 that there exists $\varepsilon_0 \in (0, 1/2)$ such that for every $\varepsilon \in (0, \varepsilon_0)$ we have

$$(34) \quad \mathbb{P} \otimes P \left(\left| \log \mu(\mathcal{B}_d(B_T, 2\varepsilon)) - \beta \log \varepsilon \right| > (\eta/2) |\log \varepsilon| \right) < \eta/2.$$

Let us fix $\varepsilon \in (0, \varepsilon_0)$. We now claim that, under $\mathbb{P} \otimes P$,

$$(35) \quad \mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}) \xrightarrow[n \rightarrow \infty]{(d)} \mu(\mathcal{B}_d(B_T, 2\varepsilon)).$$

To see this, let f be a continuous function on $[0, 1]$. Since the distribution of $\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}$ under P is μ_n^ε , we have

$$\mathbb{E} \otimes E \left[f(\mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor})) \right] = \mathbb{E} \left[\sum_{u \in \mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}} \mu_n^\varepsilon(u) f(\mu_n^\varepsilon(u)) \right].$$

By Proposition 16, we know that \mathbb{P} a.s. for n sufficiently large,

$$\sum_{u \in \mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}} \mu_n^\varepsilon(u) f(\mu_n^\varepsilon(u)) = \sum_{x=(v, 1-\varepsilon) \in \partial \Delta_\varepsilon} \mu_n^\varepsilon(w_v^{n,\varepsilon}) f(\mu_n^\varepsilon(w_v^{n,\varepsilon}))$$

and, by Proposition 18, the latter quantities converge as $n \rightarrow \infty$ toward

$$\sum_{x \in \partial \Delta_\varepsilon} \mu^\varepsilon(x) f(\mu^\varepsilon(x)) = E[f(\mu(\mathcal{B}_d(B_T, 2\varepsilon)))].$$

Our claim (35) now follows.

By (34) and (35), we can find $n_0 = n_0(\varepsilon) \geq \varepsilon^{-1}$ such that for $n \geq n_0$ we have

$$\mathbb{P} \otimes P \left(\left| \log \mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}) - \beta \log \varepsilon \right| > \eta |\log \varepsilon| \right) < \eta.$$

It follows that

$$\begin{aligned} & \mathbb{E} \otimes E \left[\left| \log \mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}) - \beta \log \varepsilon \right|^2 \right] \\ & \leq \eta^2 |\log \varepsilon|^2 + \eta^{1/2} \mathbb{E} \otimes E \left[\left| \log \mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}) - \beta \log \varepsilon \right|^4 \right]^{1/2} \\ (36) \quad & \leq (\eta^2 + 2\eta^{1/2}\beta^2) |\log \varepsilon|^2 + 2\eta^{1/2} \mathbb{E} \otimes E \left[\left| \log \mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}) \right|^4 \right]^{1/2}. \end{aligned}$$

Let us bound the last term in the right-hand side. It is elementary to verify that the function $g(r) = (r \wedge e^{-4}) |\log(r \wedge e^{-4})|^4$ is nondecreasing and concave over $[0, 1]$. It follows that

$$\begin{aligned} E \left[\left| \log \mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}) \right|^4 \right] &= \sum_{u \in \mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}} \mu_n^\varepsilon(u) |\log \mu_n^\varepsilon(u)|^4 \\ &\leq \sum_{u \in \mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}} (\mu_n^\varepsilon(u) \wedge e^{-4}) |\log(\mu_n^\varepsilon(u) \wedge e^{-4})|^4 + 4^4 \\ &= \sum_{u \in \mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}} g(\mu_n^\varepsilon(u)) + 4^4 \\ &\leq \#\mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n} \times g\left(\left(\#\mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n}\right)^{-1}\right) + 4^4 \\ &\leq \left| \log \#\mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n} \right|^4 + 2 \times 4^4. \end{aligned}$$

We now use Lemma 15 to get

$$\begin{aligned} \mathbb{E} \otimes E \left[\left| \log \mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}) \right|^4 \right] &\leq 2 \times 4^4 + \mathbb{E} \left[\left| \log \#\mathbb{T}_{n-\lfloor \varepsilon n \rfloor}^{*n} \right|^4 \right] \\ &\leq 2 \times 4^4 + C^4 \left(\log \frac{n}{\lfloor \varepsilon n \rfloor} \right)^4. \end{aligned}$$

By combining the last estimate with (36), we get that, for every $n \geq n_0(\varepsilon)$,

$$\begin{aligned} & \mathbb{E} \otimes E \left[\left| \log \mu_n^\varepsilon(\langle \Sigma_n \rangle_{n-\lfloor \varepsilon n \rfloor}) - \beta \log \varepsilon \right|^2 \right] \\ & \leq (\eta^2 + 2\eta^{1/2}\beta^2) |\log \varepsilon|^2 + 2\eta^{1/2}(2^{9/2} + C^2 |\log \varepsilon|^2). \end{aligned}$$

The statement of the corollary follows since η was arbitrary. \square

4.3. *Proof of the main result.* We need a few preliminary lemmas before we can proceed to the proof of Theorem 1.

4.3.1. *Preliminary lemmas.* Our first lemma is a discrete version of Lemma 7. This result is well known and corresponds to the “flow rule” for harmonic measure in [27]. We provide a detailed statement and a brief proof because this result plays a key role in what follows.

We consider a plane tree $\tau \in \mathcal{T}_n$, and we write $Z^{(\tau)} = (Z_k^{(\tau)})_{k \geq 0}$ for simple random walk on τ starting from \emptyset (we may assume that this process is defined under the probability measure P). We set

$$H_n^{(\tau)} = \inf\{k \geq 0 : |Z_k^{(\tau)}| = n\},$$

and $\Sigma_n^{(\tau)} = Z_{H_n^{(\tau)}}^{(\tau)}$. We let $\mu_n^{(\tau)}$ be the distribution of $\Sigma_n^{(\tau)}$. We view $\mu_n^{(\tau)}$ as a measure on τ , which is supported on τ_n .

For $0 \leq p \leq n$, we set

$$L_p^{(\tau)} = \sup\{k \leq H_n^{(\tau)} : |Z_k^{(\tau)}| = p\}.$$

Clearly $\Sigma_n^{(\tau)} \in \tilde{\tau}[Z_{L_p^{(\tau)}}^{(\tau)}]$, and therefore $Z_{L_p^{(\tau)}}^{(\tau)} = \langle \Sigma_n^{(\tau)} \rangle_p$.

LEMMA 20. *Let $p \in \{0, 1, \dots, n-1\}$ and $z \in \tau_p$. Then, conditionally on $\langle \Sigma_n^{(\tau)} \rangle_p = z$, the process*

$$\left(Z_{(L_p^{(\tau)} + k) \wedge H_n^{(\tau)}}^{(\tau)} \right)_{k \geq 0}$$

is distributed as simple random walk on $\tilde{\tau}[z]$ starting from z and conditioned to hit $\tilde{\tau}[z] \cap \tau_n$ before returning to z , and stopped at this hitting time. Consequently, for every integer $q \in \{0, 1, \dots, n-p\}$, the conditional distribution of

$$\frac{\mu_n^{(\tau)}(B_\tau(\Sigma_n^{(\tau)}, q))}{\mu_n^{(\tau)}(B_\tau(\Sigma_n^{(\tau)}, n-p))}$$

knowing that $\langle \Sigma_n^{(\tau)} \rangle_p = z$ is equal to the distribution of

$$\mu_{n-p}^{(\tau[z])}(B_{\tau[z]}(\Sigma_{n-p}^{(\tau[z])}, q)).$$

PROOF. The first assertion is easy from the fact that the successive (non-trivial) excursions of $Z^{(\tau)}$ in the subtree $\tilde{\tau}[z]$ are independent (and independent of the behavior of $Z^{(\tau)}$ outside $\tilde{\tau}[z]$) and have the same distribution as

the excursion of random walk in $\tilde{\tau}[z]$ away from z . We leave the details to the reader.

Let us explain why the second assertion of the lemma follows from the first one. Clearly, the distribution of the hitting point of $\tilde{\tau}[z] \cap \tau_n$ by simple random walk on $\tilde{\tau}[z]$ starting from z and conditioned to hit $\tilde{\tau}[z] \cap \tau_n$ before returning to z is the same as the distribution of the hitting point of $\tilde{\tau}[z] \cap \tau_n$ by simple random walk on $\tilde{\tau}[z]$ starting from z . Let $\mu_n^{(\tau),z}$ be the conditional distribution of $\Sigma_n^{(\tau)}$ knowing that $\langle \Sigma_n^{(\tau)} \rangle_p = z$. We get from the first assertion of the lemma that $\mu_n^{(\tau),z}$ is equal to the hitting distribution of $\tilde{\tau}[z] \cap \tau_n$ for simple random walk on $\tilde{\tau}[z]$ started from z (note that we are here interested in the subgraph $\tilde{\tau}[z]$ of τ and not in the “relabelled” tree $\tau[z]$). It also follows that, for every integer $q \in \{0, 1, \dots, n - p\}$, the conditional distribution of

$$\mu_n^{(\tau),z}(B_\tau(\Sigma_n^{(\tau)}, q))$$

knowing that $\langle \Sigma_n^{(\tau)} \rangle_p = z$ coincides with the distribution of

$$\mu_{n-p}^{(\tau[z])}(B_{\tau[z]}(\Sigma_{n-p}^{(\tau[z])}, q)).$$

Now notice that, on the event $\{\langle \Sigma_n^{(\tau)} \rangle_p = z\}$, $\mu_n^{(\tau),z}(B_\tau(\Sigma_n^{(\tau)}, q))$ is equal to

$$\frac{\mu_n^{(\tau)}(B_\tau(\Sigma_n^{(\tau)}, q))}{\mu_n^{(\tau)}(B_\tau(\Sigma_n^{(\tau)}, n - p))}.$$

This gives the second assertion of the lemma. \square

Let us come back to the (random) reduced tree \mathbb{T}^{*n} . If $1 \leq i \leq n$, $\tilde{\mathbb{T}}^{*n}[\langle \Sigma_n \rangle_{n-i}]$ is the subtree of \mathbb{T}^{*n} above generation $n - i$ that is “selected” by harmonic measure, and $\mathbb{T}^{*n}[\langle \Sigma_n \rangle_{n-i}]$ is the tree obtained by relabelling the vertices of $\tilde{\mathbb{T}}^{*n}[\langle \Sigma_n \rangle_{n-i}]$ as explained above. It is not true that the distribution of $\mathbb{T}^{*n}[\langle \Sigma_n \rangle_{n-i}]$ under $\mathbb{P} \otimes P$ coincides with the distribution of \mathbb{T}^{*i} under \mathbb{P} , because harmonic measure induces a distributional bias. Still the next lemma gives a useful bound for the distribution of $\mathbb{T}^{*n}[\langle \Sigma_n \rangle_{n-i}]$ in terms of that of \mathbb{T}^{*i} . We recall the notation $\mathcal{C}_i(\tau)$ from Section 4.2.2.

LEMMA 21. *For every $i \in \{1, \dots, n - 1\}$ and every nonnegative function F on \mathcal{T} ,*

$$\mathbb{E} \otimes E [F(\mathbb{T}^{*n}[\langle \Sigma_n \rangle_{n-i})] \leq (i + 1) \mathbb{E}[\mathcal{C}_i(\mathbb{T}^{*i}) F(\mathbb{T}^{*i})].$$

PROOF. Fix $i \in \{1, \dots, n-1\}$ in this proof. Recall our notation $R_i(\mathbb{T}^{*n})$ for the tree \mathbb{T}^{*n} truncated at level $n-i$. From the branching property of Galton–Watson trees, one easily verifies the following fact: Under \mathbb{P} , conditionally on $R_i(\mathbb{T}^{*n})$, the (relabelled) subtrees $\mathbb{T}^{*n}[v]$, $v \in \mathbb{T}_{n-i}^{*n}$ are independent and distributed as \mathbb{T}^{*i} (to make this statement precise we can order the subtrees according to the lexicographical order on \mathbb{T}_{n-i}^{*n}).

Consider the stopping times of the random walk Z^n which are defined inductively as follows,

$$\begin{aligned} U_0^n &= \inf\{k \geq 0 : |Z_k^n| = n-i\}, \\ V_0^n &= \inf\{k \geq U_0^n : |Z_k^n| = n-i-1\}, \end{aligned}$$

and, for every $j \geq 0$,

$$\begin{aligned} U_{j+1}^n &= \inf\{k \geq V_j^n : |Z_k^n| = n-i\}, \\ V_{j+1}^n &= \inf\{k \geq U_{j+1}^n : |Z_k^n| = n-i-1\}. \end{aligned}$$

Set $W_j^n = Z_{U_j^n}^n$ for every $j \geq 0$. Then, under the probability measure P , $(W_j^n)_{j \geq 0}$ is a Markov chain on \mathbb{T}_{n-i}^{*n} , whose initial distribution and transition kernel only depend on $R_i(\mathbb{T}^{*n})$.

Now observe that

$$\langle \Sigma_n \rangle_{n-i} = W_{j_0}^n$$

where j_0 is the first index j such that

$$(37) \quad \sup_{U_j^n \leq k \leq V_j^n} |Z_k^n| = n.$$

If $j \geq 0$ is fixed, then, conditionally on the Markov chain W^n , the probability that (37) holds is $\mathcal{C}_i(\mathbb{T}^{*n}[W_j^n])$.

Thanks to these observations, we have

$$\begin{aligned} &E[F(\mathbb{T}^{*n}[\langle \Sigma_n \rangle_{n-i}])] \\ &= \sum_{j=0}^{\infty} E\left[F(\mathbb{T}^{*n}[W_j^n]) \mathcal{C}_i(\mathbb{T}^{*n}[W_j^n]) \prod_{\ell=0}^{j-1} \left(1 - \mathcal{C}_i(\mathbb{T}^{*n}[W_\ell^n])\right)\right]. \end{aligned}$$

We then use the simple bound $\mathcal{C}_i(\mathbb{T}) \geq \frac{1}{i+1}$, which holds for any tree \mathbb{T} with height greater than or equal to i . It follows that

$$E[F(\mathbb{T}^{*n}[\langle \Sigma_n \rangle_{n-i}])] \leq \sum_{j=0}^{\infty} \left(1 - \frac{1}{i+1}\right)^j E\left[F(\mathbb{T}^{*n}[W_j^n]) \mathcal{C}_i(\mathbb{T}^{*n}[W_j^n])\right].$$

For every $u \in \mathcal{U}$ with $|u| = n - i$, let $\pi_j^n(u) = P(W_j^n = u)$, and recall that $\pi_j^n(u)$ only depends on the truncated tree $R_i(\mathbb{T}^{*n})$. Then, for every $j \geq 0$,

$$\begin{aligned} \mathbb{E} \otimes E \left[F(\mathbb{T}^{*n}[W_j^n]) \mathcal{C}_i(\mathbb{T}^{*n}[W_j^n]) \right] &= \mathbb{E} \left[\sum_{u \in \mathbb{T}_{n-i}^{*n}} \pi_j^n(u) F(\mathbb{T}^{*n}[u]) \mathcal{C}_i(\mathbb{T}^{*n}[u]) \right] \\ &= \mathbb{E}[F(\mathbb{T}^{*i}) \mathcal{C}_i(\mathbb{T}^{*i})], \end{aligned}$$

by the observation of the beginning of the proof. We conclude that

$$\begin{aligned} E[F(\mathbb{T}^{*n}[\langle \Sigma_n \rangle_{n-i})] &\leq \sum_{j=0}^{\infty} \left(1 - \frac{1}{i+1}\right)^j \mathbb{E}[F(\mathbb{T}^{*i}) \mathcal{C}_i(\mathbb{T}^{*i})] \\ &= (i+1) \mathbb{E}[F(\mathbb{T}^{*i}) \mathcal{C}_i(\mathbb{T}^{*i})], \end{aligned}$$

as desired. \square

Our last lemma gives an estimate for the conductance $\mathcal{C}_i(\mathbb{T}^{*i})$.

LEMMA 22. *There exists a constant $K \geq 1$ such that, for every integer $n \geq 1$,*

$$\mathbb{E}[\mathcal{C}_n(\mathbb{T}^{*n})^2] \leq \frac{K}{(n+1)^2}.$$

PROOF. Obviously we can assume that $n \geq 2$, and we set $j = \lfloor n/2 \rfloor \geq 1$. An immediate application of the Nash–Williams inequality [29, Chapter 2] gives

$$\mathcal{C}_n(\mathbb{T}^{*n}) \leq \frac{\#\mathbb{T}_j^{*n}}{j}$$

(just consider the cutsets obtained by looking for every integer $\ell \in \{1, \dots, j\}$ at the collection of edges of \mathbb{T}^{*n} between generation $\ell - 1$ and generation ℓ). Then,

$$\begin{aligned} \mathbb{E}[(\#\mathbb{T}_j^{*n})^2] &= \mathbb{E}[(\#\{v \in \mathbb{T}_j^{(0)} : h(\mathbb{T}^{(0)}[v]) \geq n - j\})^2 \mid h(\mathbb{T}^{(0)}) \geq n] \\ &= q_n^{-1} \mathbb{E}[(\#\{v \in \mathbb{T}_j^{(0)} : h(\mathbb{T}^{(0)}[v]) \geq n - j\})^2]. \end{aligned}$$

As we already observed in the proof of Lemma 15, the conditional distribution of $\#\{v \in \mathbb{T}_j^{(0)} : h(\mathbb{T}^{(0)}[v]) \geq n - j\}$ knowing that $\#\mathbb{T}_j^{(0)} = k$ is the binomial distribution $\mathcal{B}(k, q_{n-j})$. It follows that

$$\begin{aligned} &\mathbb{E}[(\#\{v \in \mathbb{T}_j^{(0)} : h(\mathbb{T}^{(0)}[v]) \geq n - j\})^2] \\ &= q_{n-j}^2 \mathbb{E}[(\#\mathbb{T}_j^{(0)})^2] + (q_{n-j} - q_{n-j}^2) \mathbb{E}[\#\mathbb{T}_j^{(0)}] \\ &= q_{n-j}^2 \sigma^2 j + q_{n-j}. \end{aligned}$$

We conclude that

$$\mathbb{E}[\mathcal{C}_n(\mathbb{T}^{*n})^2] \leq (j^2 q_n)^{-1} \left(q_{n-j}^2 \sigma^2 j + q_{n-j} \right),$$

and the statement of the lemma follows from (32). \square

4.3.2. *Proof of Theorem 1.* We will prove that

$$(38) \quad \mathbb{E} \otimes E [|\log \mu_n(\Sigma_n) + \beta \log n|] = o(\log n), \quad \text{as } n \rightarrow \infty.$$

Theorem 1 follows, since (38) and the Markov inequality give, for any $\delta > 0$,

$$\mathbb{P} \otimes P (|\log \mu_n(\Sigma_n) + \beta \log n| \geq \delta \log n) \xrightarrow[n \rightarrow \infty]{} 0,$$

and therefore

$$\mathbb{E} \left[P(\mu_n(\Sigma_n) \leq n^{-\beta-\delta} \text{ or } \mu_n(\Sigma_n) \geq n^{-\beta+\delta}) \right] \xrightarrow[n \rightarrow \infty]{} 0.$$

Since by definition μ_n is the distribution of Σ_n under P , the last convergence is equivalent to the first assertion of Theorem 1.

Fix $\xi > 0$ and let $\varepsilon > 0$ and $n_0 \geq 0$ be such that the conclusion of Corollary 19 holds for every $n \geq n_0$. Without loss of generality, we may and will assume that $\varepsilon = 1/N$, for some integer $N \geq 4$, which is fixed throughout the proof. We also fix a constant $\alpha > 0$, such that $\alpha \log N < 1/2$.

Let $n > N$ be sufficiently large so that $N^{\lfloor \alpha \log n \rfloor} \geq n_0$. We then let $\ell \geq 1$ be the unique integer such that

$$N^\ell < n \leq N^{\ell+1}.$$

Notice that

$$(39) \quad \frac{\log n}{\log N} - 1 \leq \ell \leq \frac{\log n}{\log N}.$$

Our starting point is the equality

$$(40) \quad \log \mu_n(\Sigma_n) = \log \frac{\mu_n(\Sigma_n)}{\mu_n(B(\Sigma_n, N))} + \sum_{j=2}^{\ell} \log \frac{\mu_n(B(\Sigma_n, N^{j-1}))}{\mu_n(B(\Sigma_n, N^j))} + \log \mu_n(B(\Sigma_n, N^\ell)).$$

To simplify notation, we set

$$\begin{aligned} A_1^n &:= \log \frac{\mu_n(\Sigma_n)}{\mu_n(B(\Sigma_n, N))} + \beta \log N, \\ A_j^n &:= \log \frac{\mu_n(B(\Sigma_n, N^{j-1}))}{\mu_n(B(\Sigma_n, N^j))} + \beta \log N \quad \text{for every } j \in \{2, \dots, \ell\}, \\ A_{\ell+1}^n &:= \log \mu_n(B(\Sigma_n, N^\ell)) + \beta \log(n/N^\ell). \end{aligned}$$

From (40), we see that

$$(41) \quad \mathbb{E} \otimes E \left[\left| \log \mu_n(\Sigma_n) + \beta \log n \right| \right] = \mathbb{E} \otimes E \left[\left| \sum_{j=1}^{\ell+1} A_j^n \right| \right] \leq \sum_{i=1}^{\ell+1} \mathbb{E} \otimes E \left[|A_j^n| \right].$$

We will now bound the different terms in the sum of the right-hand side.

FIRST STEP: A PRIORI BOUNDS. We verify that, for $j \in \{1, \dots, \ell + 1\}$, we have

$$(42) \quad \mathbb{E} \otimes E[|A_j^n|] \leq (C\sqrt{K} + \beta) \log N,$$

where C is the constant in Lemma 15, and K is the constant in Lemma 22. Suppose first that $2 \leq j \leq \ell$. Applying the second assertion of Lemma 20 (with $p = n - N^j$ and $q = N^{j-1}$) to the tree \mathbb{T}^{*n} , we obtain that, for every $z \in \mathbb{T}_{n-N^j}^{*n}$, the conditional distribution of A_j^n under P , knowing that $\langle \Sigma_n \rangle_{n-N^j} = z$, is the same as the distribution of

$$\log \mu_{N^j}^{(\mathbb{T}^{*n}[z])}(B(\Sigma_{N^j}^{(\mathbb{T}^{*n}[z])}, N^{j-1})) + \beta \log N.$$

Recalling that $\mu_{N^j}^{(\mathbb{T}^{*n}[z])}$ is the distribution of $\Sigma_{N^j}^{(\mathbb{T}^{*n}[z])}$ under P , we get

$$\begin{aligned} E[|A_j^n| \mid \langle \Sigma_n \rangle_{n-N^j} = z] &\leq E \left[\left| \log \mu_{N^j}^{(\mathbb{T}^{*n}[z])}(B(\Sigma_{N^j}^{(\mathbb{T}^{*n}[z])}, N^{j-1})) \right| \right] + \beta \log N \\ (43) \quad &= G_j(\mathbb{T}^{*n}[z]) + \beta \log N, \end{aligned}$$

where for any tree $\tau \in \mathcal{T}_{N^j}$,

$$\begin{aligned} G_j(\tau) &= \int \mu_{N^j}^{(\tau)}(dy) \left| \log \mu_{N^j}^{(\tau)}(B_\tau(y, N^{j-1})) \right| \\ &= \sum_{z \in \tau_{N^j-N^{j-1}}} \mu_{N^j}^{(\tau)}(\tilde{\tau}[z]) \left| \log \mu_{N^j}^{(\tau)}(\tilde{\tau}[z]) \right|. \end{aligned}$$

In the latter form, $G_j(\tau)$ is just the entropy of the probability measure that assigns mass $\mu_{N^j}^{(\tau)}(\tilde{\tau}[z])$ to every point $z \in \tau_{N^j-N^{j-1}}$. By a standard bound

for the entropy of probability measures on finite sets, we have $G_j(\tau) \leq \log \#\tau_{N^j-N^{j-1}}$ for any tree $\tau \in \mathcal{T}_{N^j}$. Recalling (43), we get

$$\begin{aligned} \mathbb{E} \otimes E[|A_j^n|] &\leq \mathbb{E} \otimes E[\log \#\mathbf{T}_{N^j-N^{j-1}}^{*n}[\langle \Sigma_n \rangle_{n-N^j}]] + \beta \log N \\ &\leq (N^j + 1) \mathbb{E} \left[\mathcal{C}_{N^j}(\mathbf{T}^{*N^j}) \log \#\mathbf{T}_{N^j-N^{j-1}}^{*N^j} \right] + \beta \log N \\ &\leq (N^j + 1) \mathbb{E} \left[(\mathcal{C}_{N^j}(\mathbf{T}^{*N^j}))^2 \right]^{1/2} \mathbb{E} \left[(\log \#\mathbf{T}_{N^j-N^{j-1}}^{*N^j})^2 \right]^{1/2} + \beta \log N \\ &\leq \sqrt{K} \mathbb{E} \left[(\log \#\mathbf{T}_{N^j-N^{j-1}}^{*N^j})^2 \right]^{1/2} + \beta \log N, \end{aligned}$$

using successively Lemma 21, the Cauchy–Schwarz inequality and Lemma 22. Finally, Lemma 15 gives

$$\mathbb{E} \left[(\log \#\mathbf{T}_{N^j-N^{j-1}}^{*N^j})^2 \right]^{1/2} \leq C \log N,$$

and this completes the proof of (42) when $2 \leq j \leq \ell$.

The cases $j = 1$ and $j = \ell + 1$ are treated on a similar manner. For $j = \ell + 1$, we observe that the same entropy bound gives

$$E[|\log \mu_n(B(\Sigma_n, N^\ell))|] = \sum_{y \in \mathbf{T}_{n-N^\ell}^{*n}} \mu_n(\tilde{\mathbf{T}}^{*n}[y]) |\log \mu_n(\tilde{\mathbf{T}}^{*n}[y])| \leq \log \#\mathbf{T}_{n-N^\ell}^{*n}.$$

It follows that

$$\mathbb{E} \otimes E[|\log \mu_n(B(\Sigma_n, N^\ell))|] \leq \mathbb{E}[\log \#\mathbf{T}_{n-N^\ell}^{*n}] \leq C \log N,$$

by Lemma 15 and using the fact that $N^\ell < n \leq N^{\ell+1}$.

Finally, for the case $j = 1$, we use exactly the same argument as in the case $2 \leq j \leq \ell$, to get

$$E \left[\left| \log \frac{\mu_n(\Sigma_n)}{\mu_n(B(\Sigma_n, N))} \right| \right] \leq E[\log \#\mathbf{T}_n^{*n}[\langle \Sigma_n \rangle_{n-N}]],$$

and we obtain similarly, using Lemma 21, Lemma 22 and Lemma 15,

$$\begin{aligned} \mathbb{E} \otimes E[\log \#\mathbf{T}_n^{*n}[\langle \Sigma_n \rangle_{n-N}]] &\leq (N + 1) \mathbb{E}[\mathcal{C}_N(\mathbf{T}^{*N}) \log \#\mathbf{T}_N^{*N}] \\ &\leq \sqrt{K} \mathbb{E}[(\log \#\mathbf{T}_N^{*N})^2]^{1/2} \\ &\leq C\sqrt{K} \log N. \end{aligned}$$

This completes the proof of (42)

SECOND STEP: REFINED BOUNDS. We will get a better bound than (42) for certain values of j . Precisely we prove that, if $\lfloor \alpha \log n \rfloor \leq j \leq \ell$, we have

$$(44) \quad \mathbb{E} \otimes E[|A_j^n|] \leq \sqrt{\xi K} \log N.$$

Let us fix $j \in \{\lfloor \alpha \log n \rfloor, \dots, \ell\}$. Recall that we have then $N^j \geq n_0$. From (43), we have

$$(45) \quad E[|A_j^n|] = E[F_j(\mathbb{T}^{*n}[\langle \Sigma_n \rangle_{n-N^j}])],$$

where, if $\tau \in \mathcal{T}_{N^j}$,

$$F_j(\tau) = |\beta \log N - G_j(\tau)| = \left| \int \mu_{N^j}^{(\tau)}(dy) \left(\log \mu_{N^j}^{(\tau)}(B_\tau(y, N^{j-1})) + \beta \log N \right) \right|.$$

Using Lemma 21 as in the first step, we have

$$\mathbb{E} \otimes E[|A_j^n|] = \mathbb{E} \otimes E[F_j(\mathbb{T}^{*n}[\langle \Sigma_n \rangle_{n-N^j}])] \leq (N^j + 1) \mathbb{E} \left[\mathcal{C}_{N^j}(\mathbb{T}^{*N^j}) F_j(\mathbb{T}^{*N^j}) \right].$$

We then apply the Cauchy-Schwarz inequality together with the bound of Lemma 22 to get

$$\begin{aligned} \mathbb{E} \otimes E[|A_j^n|] &\leq \sqrt{K} \mathbb{E}[F_j(\mathbb{T}^{*N^j})^2]^{1/2} \\ &= \sqrt{K} \mathbb{E} \left[\left(\int \mu_{N^j}(dy) \left| \log \mu_{N^j}(B(y, N^{j-1})) + \beta \log N \right| \right)^2 \right]^{1/2} \\ &\leq \sqrt{K} \mathbb{E} \left[\int \mu_{N^j}(dy) \left| \log \mu_{N^j}(B(y, N^{j-1})) + \beta \log N \right|^2 \right]^{1/2} \\ &= \sqrt{K} \mathbb{E} \otimes E \left[\left| \log \mu_{N^j}(B(\Sigma_{N^j}, N^{j-1})) + \beta \log N \right|^2 \right]^{1/2} \\ &= \sqrt{K} \cdot \mathbb{E} \otimes E \left[\left| \log \mu_{N^j}^{1/N} \left(\langle \Sigma_{N^j} \rangle_{N^j - N^{j-1}} \right) + \beta \log N \right|^2 \right]^{1/2}, \end{aligned}$$

where the last equality follows from the definition of the measures μ_n^ε at the beginning of Section 4.2.3. Now recall that $1/N = \varepsilon$ and note that $N^j - N^{j-1} = N^j - \varepsilon N^j$. Since we have $N^j \geq n_0$, we can apply the bound of Corollary 19 and we get that the right-hand side of the preceding display is bounded above by $\sqrt{\xi K} \log N$, which completes the proof of (44).

By combining (42) and (44), and using (41), we arrive at the bound

$$\begin{aligned} \mathbb{E} \otimes E \left[\left| \log \mu_n(\Sigma_n) + \beta \log n \right| \right] &\leq \lfloor \alpha \log n \rfloor (C\sqrt{K} + \beta) \log N + \ell \sqrt{\xi K} \log N \\ &\leq \left(\alpha (C\sqrt{K} + \beta) \log N + \sqrt{\xi K} \right) \log n, \end{aligned}$$

which holds for every sufficiently large n . Now note that $\xi > 0$ can be chosen arbitrarily small. The choice of ξ determines the choice of N , but afterwards we can also choose α arbitrarily small given this choice. We thus see that our claim (38) follows from the last bound, and this completes the proof of Theorem 1.

4.4. *Proof of Corollary 2.* In what follows we always **implicitly restrict our attention** to integers $N \geq 1$ such that $\mathbb{P}(\#\mathbf{T}^{(0)} = N + 1) > 0$. For such values of N , $\mathbf{T}(N)$ is distributed as $\mathbf{T}^{(0)}$ conditioned on the event $\{\#\mathbf{T}^{(0)} = N + 1\}$. We write $(C_t^{(N)})_{0 \leq t \leq 2N}$ for the contour function of the tree $\mathbf{T}(N)$ (see e.g. [23] or [31, Figure 6.2], where the contour function is called the Harris walk of the tree). By a famous theorem of Aldous [3], we have the convergence in distribution

$$(46) \quad \left(\frac{\sigma}{2\sqrt{N}} C_{2Nt}^{(N)}, 0 \leq t \leq 1 \right) \xrightarrow[N \rightarrow \infty]{(d)} (\mathbf{e}_t, 0 \leq t \leq 1)$$

where $(\mathbf{e}_t, 0 \leq t \leq 1)$ stands for a Brownian excursion with duration 1. Since $h(\mathbf{T}(N))$ is just the maximum of the contour function, it follows that

$$(47) \quad \frac{1}{\sqrt{N}} h(\mathbf{T}(N)) \xrightarrow[N \rightarrow \infty]{(d)} \frac{2}{\sigma} \max_{0 \leq t \leq 1} \mathbf{e}_t.$$

Consequently, for every $\eta > 0$, we can choose a constant $A > 0$ such that for all sufficiently large N , the probability $P(h(\mathbf{T}(N)) > A\sqrt{N})$ is bounded above by η . Thanks to this remark, it is enough to prove that the convergence of Corollary 2 holds when n and N tend to infinity in such a way that $n \leq B\sqrt{N}$, for some fixed constant B . For future reference, we note that (47) implies that, for every sufficiently large N and every nonnegative integer n such that $n \leq B\sqrt{N}$,

$$(48) \quad \mathbb{P}(h(\mathbf{T}(N)) \geq n) \geq c,$$

for some constant $c > 0$.

If $\tau \in \mathcal{T}$ is a tree, we write $\tau_{\leq n}$ for the tree that consists of all vertices of τ at generation less than or equal to n .

LEMMA 23. *Let $\varepsilon > 0$. We can find $\delta \in (0, \frac{1}{2})$ such that, for every sufficiently large N , and every nonnegative integer n with $n \leq B\sqrt{N}$, we have*

$$P\left(\#\mathbf{T}(N)_{\leq n} \leq (1 - \delta)N \mid h(\mathbf{T}(N)) \geq n\right) \geq 1 - \varepsilon.$$

PROOF. As a simple consequence of (46), we can find $\eta > 0$ sufficiently small so that, for every sufficiently large N and for every integer n with $0 \leq n \leq \eta\sqrt{N}$,

$$\mathbb{P}(\{h(\mathbf{T}(N)) \geq n\} \cap \{\#\mathbf{T}(N)_{\leq n} < \frac{N}{2}\}) > 1 - \varepsilon.$$

So we may concentrate on values of n such that $\eta\sqrt{N} \leq n \leq B\sqrt{N}$.

We then observe that there exists $\delta \in (0, \frac{1}{2})$ such that, for every $a \in [\frac{1}{2}\sigma\eta, \frac{1}{2}\sigma B]$,

$$(49) \quad \mathbb{P}\left(\int_0^1 dt \mathbf{1}_{\{\mathbf{e}_t \geq a\}} \leq \delta \mid \sup_{0 \leq t \leq 1} \mathbf{e}_t \geq a\right) < \varepsilon.$$

This bound follows from standard properties of linear Brownian motion. We omit the details.

We now claim that the result of the lemma holds with the preceding value of δ . To verify the claim, observe that from the properties of the contour function,

$$N + 1 - \#\mathbf{T}(N)_{\leq n} = \frac{1}{2} \int_0^{2N} dt \mathbf{1}_{\{C_t^{(N)} > n\}}.$$

It readily follows that

$$\begin{aligned} & \mathbb{P}\left(\#\mathbf{T}(N)_{\leq n} > (1 - \delta)N \mid h(\mathbf{T}(N)) \geq n\right) \\ &= \mathbb{P}\left(\frac{1}{2} \int_0^{2N} dt \mathbf{1}_{\{C_t^{(N)} > n\}} < \delta N + 1 \mid \sup_{0 \leq t \leq 2N} C_t^{(N)} \geq n\right) \\ &= \mathbb{P}\left(\int_0^1 dt \mathbf{1}_{\{(\sigma/2\sqrt{N})C_{2Nt}^{(N)} > (\sigma/2\sqrt{N})n\}} < \delta + \frac{1}{N} \mid \sup_{0 \leq t \leq 1} \frac{\sigma}{2\sqrt{N}}C_{2Nt}^{(N)} \geq \frac{n}{\sqrt{2N}}\right). \end{aligned}$$

If the conclusion of the lemma does not hold, we can find a sequence N_k converging to $+\infty$, and, for every k , an integer n_k with $\eta\sqrt{N_k} \leq n_k \leq B\sqrt{N_k}$, such that the probability in the last display, evaluated with $N = N_k$ and $n = n_k$ is bounded below by ε . But then, by extracting a convergent subsequence from the sequence $(n_k/\sqrt{N_k})$ and using the convergence (46), we get a contradiction with (49). This contradiction completes the proof. \square

As previously, we let $\mathbb{T}^{(n)}$ stand for a Galton–Watson tree with offspring distribution θ , conditioned on non-extinction at generation n . Corollary 2 is a simple consequence of Theorem 1 and the following comparison lemma applied, for every fixed $\delta > 0$ and $\varepsilon > 0$, with

$$A_n = \left\{ \tau \in \mathcal{T}_n : \mu_n^{(\tau)}(\{v \in \tau_n : n^{-\beta-\delta} \leq \mu_n^{(\tau)}(v) \leq n^{-\beta+\delta}\}) \leq 1 - \varepsilon \right\}.$$

LEMMA 24. *For every $n \geq 0$, let A_n be a subset of \mathcal{T}_n . Assume that $\mathbb{P}(\mathbf{T}_{\leq n}^{(n)} \in A_n) \rightarrow 0$ as $n \rightarrow \infty$. Then we have*

$$\mathbb{P}\left(\mathbf{T}(N)_{\leq n} \in A_n \mid h(\mathbf{T}(N)) \geq n\right) \xrightarrow[\substack{n, N \rightarrow \infty \\ n \leq B\sqrt{N}}]{} 0.$$

PROOF. Throughout the proof, we consider positive integers n and N such that $n \leq B\sqrt{N}$. Let $\tau \in \mathcal{T}_n$ and set $m = \#\tau - 1$ (m is the number of edges of τ) and $p = \#\tau_n$. From (32), we see that there exists a constant $c_0 > 0$ such that, for every n ,

$$(50) \quad \mathbb{P}(\mathbf{T}_{\leq n}^{(n)} = \tau) \geq c_0 n \mathbb{P}(\mathbf{T}_{\leq n}^{(0)} = \tau).$$

We then evaluate

$$\mathbb{P}(\mathbf{T}(N)_{\leq n} = \tau) = \frac{\mathbb{P}(\{\mathbf{T}_{\leq n}^{(0)} = \tau\} \cap \{\#\mathbf{T}^{(0)} = N + 1\})}{\mathbb{P}(\#\mathbf{T}^{(0)} = N + 1)}.$$

Let $\tilde{\theta}$ be the probability measure on \mathbb{Z} defined by $\tilde{\theta}(k) = \theta(k + 1)$ for every $k \geq -1$, and let Z be a random walk on \mathbb{Z} with jump distribution $\tilde{\theta}$ started from 0. A standard result (see e.g. [23, Section 1]) states that $\#\mathbf{T}^{(0)}$ has the same distribution as the hitting time of -1 by Z , and by Kemperman's formula (see e.g. Pitman [31, p.122]), we get that $\mathbb{P}(\#\mathbf{T}^{(0)} = N + 1) = (N + 1)^{-1} \mathbb{P}(Z_{N+1} = -1)$. From a classical local limit theorem, we obtain the existence of a constant $c_1 > 0$ such that

$$(51) \quad \mathbb{P}(\#\mathbf{T}^{(0)} = N + 1) \geq c_1 N^{-3/2}$$

(recall that we consider only values of N such that $\mathbb{P}(\#\mathbf{T}^{(0)} = N + 1) > 0$). Then, using the branching property of Galton-Watson trees, we have, if $m \leq N$,

$$(52) \quad \mathbb{P}(\{\mathbf{T}_{\leq n}^{(0)} = \tau\} \cap \{\#\mathbf{T}^{(0)} = N + 1\}) = \mathbb{P}(\mathbf{T}_{\leq n}^{(0)} = \tau) \times F(p, N - m + p),$$

where, for every integer $\ell \geq p$, $F(p, \ell)$ is the probability that a forest of p independent Galton-Watson trees with offspring distribution θ has exactly ℓ vertices. By the same arguments as in the derivation of (51),

$$(53) \quad F(p, \ell) = \frac{p}{\ell} \mathbb{P}(Z_\ell = -p) \leq c_2 p \ell^{-3/2},$$

with some constant c_2 .

Next let $\delta \in (0, 1)$ and suppose that $m \leq (1 - \delta)N$ and $p \leq \delta^{-1}n$, so that in particular $N - m + p \geq \delta N$. Under these conditions, (51), (52) and (53) give

$$(54) \quad \mathbb{P}(\mathbf{T}(N)_{\leq n} = \tau) \leq c_1^{-1} c_2 \delta^{-5/2} n \mathbb{P}(\mathbf{T}_{\leq n}^{(0)} = \tau).$$

Let $G'_{N,n,\delta}$ be the set of all trees $\tau \in \mathcal{T}_n$ such that $\#\tau \leq (1 - \delta)N$, let $G''_{n,\delta}$ be the set of all trees $\tau \in \mathcal{T}_n$ such that $\#\tau_n \leq \delta^{-1}n$, and set $G_{N,n,\delta} = G'_{N,n,\delta} \cap G''_{n,\delta}$. Comparing (50) and (54), we obtain that the density of the law of $\mathbf{T}(N)_{\leq n}$ with respect to that of $\mathbf{T}_{\leq n}^{(n)}$ is bounded above, on the set $G_{N,n,\delta}$, by a positive constant C_δ independent of n and N (but depending on δ). If $\varepsilon > 0$ is given, we can use Lemma 23 to find $\delta > 0$ such that for every sufficiently large N and every integer n with $1 \leq n \leq B\sqrt{N}$ we have

$$\mathbb{P}(\mathbf{T}(N)_{\leq n} \in G'_{N,n,\delta} \mid h(\mathbf{T}(N)) \geq n) \geq 1 - \frac{\varepsilon}{2}.$$

On the other hand, Theorem 1.13 in Janson [17] gives the existence of a constant K independent of N such that, for every integer $n \geq 1$, $\mathbb{E}[\#\mathbf{T}(N)_n] \leq Kn$. Choosing δ smaller if necessary, and using (48), we see that we have also, for every integer n with $1 \leq n \leq B\sqrt{N}$,

$$\mathbb{P}(\mathbf{T}(N)_{\leq n} \in G''_{n,\delta} \mid h(\mathbf{T}(N)) \geq n) \geq 1 - \frac{\varepsilon}{2}.$$

Finally, if A_n is a subset of \mathcal{T}_n , with $1 \leq n \leq B\sqrt{N}$, we have

$$\begin{aligned} & \mathbb{P}(\mathbf{T}(N)_{\leq n} \in A_n \mid h(\mathbf{T}(N)) \geq n) \\ & \leq \mathbb{P}(\mathbf{T}(N)_{\leq n} \notin G_{N,n,\delta} \mid h(\mathbf{T}(N)) \geq n) \\ & \quad + \mathbb{P}(\mathbf{T}(N)_{\leq n} \in A_n \cap G_{N,n,\delta} \mid h(\mathbf{T}(N)) \geq n) \\ & \leq \varepsilon + \frac{C_\delta}{\mathbb{P}(h(\mathbf{T}(N)) \geq n)} \mathbb{P}(\mathbf{T}_{\leq n}^{(n)} \in A_n). \end{aligned}$$

Letting $n, N \rightarrow \infty$ with the constraint $n \leq B\sqrt{N}$, and using the assumption of the lemma together with (48), we see that the last display eventually becomes less than 2ε . This proves the lemma. \square

5. Complements.

5.1. *A different approach to the continuous results.* In this section, we briefly outline another approach to Theorem 3, which is based on a different shift transformation on the space \mathbb{T}^* . Informally, if $(\mathcal{T}, \mathbf{v}) \in \mathbb{T}^*$, we let

$S(\mathcal{T}, \mathbf{v})$ be obtained by shifting $(\mathcal{T}, \mathbf{v})$ at the first node of \mathcal{T} . More precisely, if \mathcal{T} corresponds to the collection $(z_v)_{v \in \mathcal{V}}$, and $\mathbf{v} = (v_1, v_2, \dots)$, we set

$$S(\mathcal{T}, \mathbf{v}) = (\mathcal{T}_{(v_1)}, \tilde{\mathbf{v}})$$

where $\tilde{\mathbf{v}} = (v_2, v_3, \dots)$ and, for $i = 1$ or $i = 2$, $\mathcal{T}_{(i)}$ is the tree corresponding to the collection $(z_{iv} - z_\emptyset)_{v \in \mathcal{V}}$, in agreement with the notation of Section 3.4.

PROPOSITION 25. *For every $r \geq 1$, set*

$$\kappa(r) = \iint \gamma(ds)\gamma(dt) \frac{rs}{r+s+t-1}.$$

The finite measure $\kappa(\mathcal{C}(\mathcal{T})) \cdot \Theta^(d\mathcal{T} d\mathbf{v})$ is invariant under S .*

PROOF. Let F be a bounded measurable function on \mathbb{T}^* . We have to prove that

$$(55) \quad \int F \circ S(\mathcal{T}, \mathbf{v}) \kappa(\mathcal{C}(\mathcal{T})) \Theta^*(d\mathcal{T} d\mathbf{v}) = \int F(\mathcal{T}, \mathbf{v}) \kappa(\mathcal{C}(\mathcal{T})) \Theta^*(d\mathcal{T} d\mathbf{v}).$$

Recall that $\Theta^*(d\mathcal{T} d\mathbf{v}) = \Theta(d\mathcal{T})\nu_{\mathcal{T}}(d\mathbf{v})$ by construction. If we fix $\mathcal{T} \in \mathbb{T}$, the distribution of the pair $(v_1, \tilde{\mathbf{v}})$ under $\nu_{\mathcal{T}}$ is given by

$$\int \nu_{\mathcal{T}}(d\mathbf{v}) \mathbf{1}_{\{v_1=i\}} g(\tilde{\mathbf{v}}) = \frac{\mathcal{C}(\mathcal{T}_{(i)})}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})} \int \nu_{\mathcal{T}_{(i)}}(d\mathbf{u}) g(\mathbf{u})$$

where $i \in \{1, 2\}$ and g is any bounded measurable function on $\{1, 2\}^{\mathbb{N}}$. It follows that the left-hand side of (55) may be written as

$$(56) \quad \sum_{i=1}^2 \int F(\mathcal{T}_{(i)}, \mathbf{u}) \kappa(\mathcal{C}(\mathcal{T})) \frac{\mathcal{C}(\mathcal{T}_{(i)})}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})} \Theta(d\mathcal{T}) \nu_{\mathcal{T}_{(i)}}(d\mathbf{u}).$$

We then observe that under $\Theta(d\mathcal{T})$ the subtrees $\mathcal{T}_{(1)}$ and $\mathcal{T}_{(2)}$ are independent and distributed according to Θ , and moreover we have

$$\mathcal{C}(\mathcal{T}) = \left(U + \frac{1-U}{\mathcal{C}(\mathcal{T}_{(1)}) + \mathcal{C}(\mathcal{T}_{(2)})} \right)^{-1}$$

where U is uniformly distributed over $[0, 1]$ and independent of $(\mathcal{T}_{(1)}, \mathcal{T}_{(2)})$. Using these observations, and a simple symmetry argument, we get that the

quantity (56) is also equal to

$$\begin{aligned}
& 2 \int_0^1 dx \int \Theta(d\mathcal{T}) \Theta(d\mathcal{T}') \nu_{\mathcal{T}}(d\mathbf{u}) F(\mathcal{T}, \mathbf{u}) \\
& \quad \times \frac{\mathcal{C}(\mathcal{T})}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')} \kappa\left(\left(x + \frac{1-x}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')}\right)^{-1}\right) \\
& = \int \Theta^*(d\mathcal{T} d\mathbf{u}) F(\mathcal{T}, \mathbf{u}) \\
& \quad \times \left(2 \int_0^1 dx \int \Theta(d\mathcal{T}') \frac{\mathcal{C}(\mathcal{T})}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')} \kappa\left(\left(x + \frac{1-x}{\mathcal{C}(\mathcal{T}) + \mathcal{C}(\mathcal{T}')}\right)^{-1}\right)\right).
\end{aligned}$$

Hence, the proof of (55) reduces to checking that, for every $r \geq 1$,

$$(57) \quad \kappa(r) = 2 \int_0^1 dx \int \Theta(d\mathcal{T}') \frac{r}{r + \mathcal{C}(\mathcal{T}')} \kappa\left(\left(x + \frac{1-x}{r + \mathcal{C}(\mathcal{T}')}\right)^{-1}\right).$$

To verify (57), let $\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2$ be independent and distributed according to γ , and let U be uniformly distributed over $[0, 1]$ and independent of $(\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2)$ under the probability measure \mathbb{P} . Note that by definition, for every $x \geq 1$,

$$\kappa(x) = \mathbb{E} \left[\frac{x\mathcal{C}_1}{x + \mathcal{C}_1 + \mathcal{C}_2 - 1} \right].$$

It follows that the right-hand side of (57) can be written as

$$\begin{aligned}
& 2 \mathbb{E} \left[\frac{r}{r + \mathcal{C}_0} \frac{\mathcal{C}_1 \left(U + \frac{1-U}{r+\mathcal{C}_0}\right)^{-1}}{\mathcal{C}_1 + \mathcal{C}_2 + \left(U + \frac{1-U}{r+\mathcal{C}_0}\right)^{-1} - 1} \right] \\
& = 2r \mathbb{E} \left[\frac{\mathcal{C}_1}{(\mathcal{C}_1 + \mathcal{C}_2 - 1)(U(\mathcal{C}_0 + r) + 1 - U) + \mathcal{C}_0 + r} \right] \\
& = r \mathbb{E} \left[\frac{\mathcal{C}_1 + \mathcal{C}_2}{(\mathcal{C}_1 + \mathcal{C}_2 - 1)(U(\mathcal{C}_0 + r) + 1 - U) + \mathcal{C}_0 + r} \right] \\
& = r \mathbb{E} \left[\frac{\mathcal{C}_1 + \mathcal{C}_2}{(\mathcal{C}_1 + \mathcal{C}_2)(U(\mathcal{C}_0 + r - 1) + 1) + (\mathcal{C}_0 + r - 1)(1 - U)} \right] \\
& = r \mathbb{E} \left[\frac{1}{(\mathcal{C}_0 + r - 1)\left(U + \frac{1-U}{\mathcal{C}_1 + \mathcal{C}_2}\right) + 1} \right] \\
& = r \mathbb{E} \left[\frac{\tilde{\mathcal{C}}}{r + \mathcal{C}_0 + \tilde{\mathcal{C}} - 1} \right]
\end{aligned}$$

where $\tilde{\mathcal{C}} = (U + \frac{1-U}{\mathcal{C}_1 + \mathcal{C}_2})^{-1}$. By (2), $\tilde{\mathcal{C}}$ is distributed according to γ . Since $\tilde{\mathcal{C}}$ is also independent of \mathcal{C}_0 , we immediately see that the right-hand side of the last display is equal to $\kappa(r)$, which completes the proof of (57) and of the proposition. \square

One can verify that the shift S is ergodic with respect to the invariant probability measure obtained by normalizing $\kappa(\mathcal{C}(\mathcal{T})) \cdot \Theta^*(d\mathcal{T} d\mathbf{v})$ (we omit the proof). One then applies the ergodic theorem to the two functionals defined as follows. First we let $Z_n(\mathcal{T}, \mathbf{v})$ denote the height of the n -th branching point on the geodesic ray \mathbf{v} . One immediately verifies that, for every $n \geq 1$,

$$Z_n = \sum_{i=0}^{n-1} Z_1 \circ S^i.$$

If $A = \int \kappa(\mathcal{C}(\mathcal{T})) \Theta^*(d\mathcal{T} d\mathbf{v})$, it follows that

$$(58) \quad \frac{1}{n} Z_n \xrightarrow[n \rightarrow \infty]{\Theta^* \text{ a.s.}} A^{-1} \int Z_1(\mathcal{T}, \mathbf{v}) \kappa(\mathcal{C}(\mathcal{T})) \Theta^*(d\mathcal{T} d\mathbf{v}).$$

Note that the limit can also be written as

$$A^{-1} \mathbb{E} \left[|\log(1-U)| \kappa \left(\left(U + \frac{1-U}{\mathcal{C}_1 + \mathcal{C}_2} \right)^{-1} \right) \right]$$

with the notation of the preceding proof. Secondly, if $\mathbf{x}_{n,\mathbf{v}}$ stands for the $n+1$ -st branching point on the geodesic ray \mathbf{v} (with the notation of Section 2.2, $\mathbf{x}_{n,\mathbf{v}} = ((v_1, \dots, v_n), Z_{n+1}(\mathcal{T}, \mathbf{v}))$ if $\mathbf{v} = (v_1, v_2, \dots)$), we set for every $n \geq 1$,

$$H_n(\mathcal{T}, \mathbf{v}) = \log \nu_{\mathcal{T}}(\{\mathbf{u} \in \{1, 2\}^{\mathbb{N}} : \mathbf{x}_{n,\mathbf{v}} \prec \mathbf{u}\}).$$

It is then also easy to verify that

$$H_n = \sum_{i=0}^{n-1} H_1 \circ S^i$$

and we have thus

$$(59) \quad \frac{1}{n} H_n \xrightarrow[n \rightarrow \infty]{\Theta^* \text{ a.s.}} A^{-1} \int H_1(\mathcal{T}, \mathbf{v}) \kappa(\mathcal{C}(\mathcal{T})) \Theta^*(d\mathcal{T} d\mathbf{v}).$$

The limit can be written as

$$2 A^{-1} \mathbb{E} \left[\frac{\mathcal{C}_1}{\mathcal{C}_1 + \mathcal{C}_2} \log \left(\frac{\mathcal{C}_1}{\mathcal{C}_1 + \mathcal{C}_2} \right) \kappa \left(\left(U + \frac{1-U}{\mathcal{C}_1 + \mathcal{C}_2} \right)^{-1} \right) \right].$$

By combining (58) and (59), we now obtain that the convergence (1) holds with limit

$$-\beta = \frac{2\mathbb{E}\left[\frac{c_1}{c_1+c_2} \log\left(\frac{c_1}{c_1+c_2}\right) \kappa\left(\left(U + \frac{1-U}{c_1+c_2}\right)^{-1}\right)\right]}{\mathbb{E}\left[|\log(1-U)| \kappa\left(\left(U + \frac{1-U}{c_1+c_2}\right)^{-1}\right)\right]}.$$

We leave it as an exercise for the reader to check that this is consistent with the other formulas for β in Proposition 4.

5.2. Supercritical Galton–Watson trees. One may compare our results about Brownian motion on the Yule tree to the recent paper of Aïdékon [1], which deals with biased random walk on supercritical Galton–Watson trees. To this end, consider the supercritical offspring distribution $\theta^{(n)}$ given by $\theta^{(n)}(1) = 1 - \frac{1}{n}$ and $\theta^{(n)}(2) = \frac{1}{n}$. If $\mathcal{T}^{(n)}$ is the (infinite) Galton–Watson tree with offspring distribution $\theta^{(n)}$, then $\mathcal{T}^{(n)}$, viewed as a metric space for the graph distance rescaled by the fact n^{-1} , converges in distribution in an appropriate sense (e.g. for the local Gromov–Hausdorff topology) to the Yule tree Γ .

Consider then the biased random walk $(Z_k^{(n)})_{k \geq 0}$ on $\mathcal{T}^{(n)}$ with bias parameter $\lambda^{(n)} = 1 - \frac{1}{n}$ (see e.g. [28] or [1] for a definition of this process). Since the “mean drift” of $Z^{(n)}$ away from the root is $\frac{1}{2n} + o(n^{-1})$, it should be clear that the rescaled process $(Z_{[n^2 t]}^{(n)})_{t \geq 0}$ is asymptotically close to Brownian motion with drift 1/2 on the Yule tree, in a sense that can easily be made precise.

An explicit form of an invariant measure for the “environment seen from the particle” has been derived by Aïdékon [1, Theorem 4.1] for biased random walk on a supercritical Galton–Watson tree (see also [16] for a related result in a different setting). In the unbiased case such an explicit formula already appeared in the work of Lyons, Pemantle and Peres [27], but in the subsequent work of the same authors [28] dealing with the biased case, only the existence of the invariance measure was derived by general arguments. It is tempting to use Aïdékon’s formula and the connection between the $\lambda^{(n)}$ -biased random walk on $\mathcal{T}^{(n)}$ and Brownian motion with drift $\frac{1}{2}$ on the Yule tree to recover our formulas for invariant measures in Propositions 12 and 25. Note however that the continuous analog of Aïdékon’s formula would be an invariant measure for the environment seen from Brownian motion on the Yule tree at a fixed time, whereas we have obtained invariant measures for the environment at the last visit of a fixed height (Proposition 12) or the last visit of a node of the n -th generation (Proposition 25). Still the reader should note the similarity between the limiting distribution in [1, Theorem

4.1] and the formula for the invariant measure in Proposition 25. Indeed, we were able to guess the formula for κ in Proposition 25 from a (non-rigorous) passage to the limit from the corresponding formula in [1].

Appendix. In this appendix, we sketch a proof of Proposition 5, which is based on the relation between the continuous reduced tree Δ of Section 2.1 and the Brownian excursion conditioned to hit level 1. This relation was described after Proposition 16 (see Figure 7) and we retain the notation introduced after this proposition. In particular, $(\mathbf{e}_t)_{0 \leq t \leq \zeta}$ is a Brownian excursion conditioned to hit level 1 defined under the probability measure \mathbb{P} , and Δ is the associated continuous reduced tree.

We fix $\varepsilon \in (0, 1)$ and let $N_\varepsilon \geq 1$ be the number of excursions of \mathbf{e} from $1 - \varepsilon$ to 1. We let $(R_1^\varepsilon, S_1^\varepsilon), (R_2^\varepsilon, S_2^\varepsilon), \dots, (R_{N_\varepsilon}^\varepsilon, S_{N_\varepsilon}^\varepsilon)$ be the time intervals corresponding to these excursions listed in chronological order. For convenience, we also set $R_i^\varepsilon = S_i^\varepsilon = \infty$ if $i > N_\varepsilon$. The key ingredient of our proof is the following lemma. We write $(B_t)_{t \geq 0}$ for a linear Brownian motion that starts from x under the probability measure P_x , and $T_0 = \inf\{t \geq 0 : B_t = 0\}$. We also let n_ε be the law of a Brownian excursion above level $1 - \varepsilon$ conditioned to hit level 1. Agreeing that the excursion stays constant after returning to $1 - \varepsilon$, we can view n_ε as a probability measure on the space $C(\mathbb{R}_+, \mathbb{R}_+)$ of all continuous functions from \mathbb{R}_+ into \mathbb{R}_+ .

LEMMA 26. *Let F, G, H be three nonnegative measurable functions on $C(\mathbb{R}_+, \mathbb{R}_+)$. Then,*

$$\begin{aligned} & \mathbb{E} \left[\sum_{i=1}^{N_\varepsilon} F((\mathbf{e}_{(R_i^\varepsilon - t)^+})_{t \geq 0}) G((\mathbf{e}_{(R_i^\varepsilon + t) \wedge S_i^\varepsilon})_{t \geq 0}) H((\mathbf{e}_{(S_i^\varepsilon + t) \wedge \zeta})_{t \geq 0}) \right] \\ &= \frac{1}{\varepsilon} E_{1-\varepsilon} [F((B_{t \wedge T_0})_{t \geq 0})] n_\varepsilon(G) E_{1-\varepsilon} [H((B_{t \wedge T_0})_{t \geq 0})]. \end{aligned}$$

The proof of this lemma is straightforward. First note that, for every $i \geq 1$, the law of $(\mathbf{e}_{(R_i^\varepsilon + t) \wedge S_i^\varepsilon})_{t \geq 0}$ under $\mathbb{P}(\cdot \mid N_\varepsilon \geq i)$ is n_ε . Then, since S_i^ε is a stopping time for every integer $i \geq 1$, we get by applying the strong Markov property at time S_i^ε ,

$$\begin{aligned} & \mathbb{E} \left[\sum_{i=1}^{N_\varepsilon} F((\mathbf{e}_{(R_i^\varepsilon - t)^+})_{t \geq 0}) G((\mathbf{e}_{(R_i^\varepsilon + t) \wedge S_i^\varepsilon})_{t \geq 0}) H((\mathbf{e}_{(S_i^\varepsilon + t) \wedge \zeta})_{t \geq 0}) \right] \\ &= \sum_{i=1}^{\infty} \mathbb{E} \left[\mathbf{1}_{\{S_i^\varepsilon < \infty\}} F((\mathbf{e}_{(R_i^\varepsilon - t)^+})_{t \geq 0}) G((\mathbf{e}_{(R_i^\varepsilon + t) \wedge S_i^\varepsilon})_{t \geq 0}) \right] \\ & \quad \times E_{1-\varepsilon} [H((B_{t \wedge T_0})_{t \geq 0})]. \end{aligned}$$

On the other hand, using the fact that the law of $(\mathbf{e}_t)_{0 \leq t \leq \zeta}$ is invariant under time reversal ($(\mathbf{e}_t)_{0 \leq t \leq \zeta}$ and $(\mathbf{e}_{\zeta-t})_{0 \leq t \leq \zeta}$ have the same law), we also obtain that the sum in the second line of the last display is equal to

$$\begin{aligned} & \mathbb{E} \left[\sum_{i=1}^{N_\varepsilon} G((\mathbf{e}_{(S_i^\varepsilon-t) \vee R_i^\varepsilon})_{t \geq 0}) F((\mathbf{e}_{(S_i^\varepsilon+t) \wedge \zeta})_{t \geq 0}) \right] \\ &= \mathbb{E} \left[\sum_{i=1}^{N_\varepsilon} G((\mathbf{e}_{(S_i^\varepsilon-t) \vee R_i^\varepsilon})_{t \geq 0}) \right] \times E_{1-\varepsilon} [F((B_{t \wedge T_0})_{t \geq 0})] \\ &= \mathbb{E}[N_\varepsilon] n_\varepsilon(G) E_{1-\varepsilon} [F((B_{t \wedge T_0})_{t \geq 0})] \end{aligned}$$

giving the desired result since $\mathbb{E}[N_\varepsilon] = \frac{1}{\varepsilon}$.

Let us informally explain why Proposition 5 (or the equivalent statement in terms of the tree Δ) follows from the lemma. To make the connection with Proposition 5 we take $\varepsilon = e^{-r}$, so that the factor $\frac{1}{\varepsilon}$ becomes the multiplicative factor e^r in the formula of Proposition 5. We first recall that every vertex v of Δ at height $1 - \varepsilon$ corresponds to one excursion of \mathbf{e} above height $1 - \varepsilon$ that hits level 1, and we observe that the tree of descendants of v will be coded by this excursion in the same way as Δ is coded by \mathbf{e} . Hence this tree of descendants is distributed as a scaled copy of Δ (and the scaling factor will disappear when we do the logarithmic scale transformation to return to the Yule tree). Then, we need to consider the subtrees branching off the ancestral line of v , and we can first look at those subtrees branching on the right of the ancestral line. Supposing that v corresponds to the excursion during the time interval $(R_i^\varepsilon, S_i^\varepsilon)$, the latter subtrees exactly correspond to all excursions of the process $(\mathbf{e}_{(S_i^\varepsilon+t) \wedge \zeta})_{t \geq 0}$ above its past minimum process that hit level 1, and the level at which a subtree branches is the starting level of the corresponding excursion. The formula of Lemma 26 then leads us to consider the excursions of $(B_{t \wedge T_0})_{t \geq 0}$ above its past minimum process, under the measure $P_{1-\varepsilon}$. If f_1, \dots, f_N stand for these excursions, and if h_i denotes the starting level of the excursion f_i , Itô's excursion theory shows that the point measure $\sum_{i=1}^N \delta_{(h_i, f_i)}$ is Poisson with intensity

$$\mathbf{1}_{[0, 1-\varepsilon]}(h) \frac{dh}{1-h} n_{1-h}(df).$$

Recalling that $\varepsilon = e^{-r}$, the image of the measure $\mathbf{1}_{[0, 1-\varepsilon]}(h) \frac{dh}{1-h}$ under the logarithmic scale transformation $h = 1 - e^{-s}$ is the measure $\mathbf{1}_{[0, r]}(s) ds$. This explains the form of the intensity of the Poisson measure \mathcal{N} in the statement of Proposition 5, noting that the factor 2 comes from the fact that we also need to consider the subtrees that branch on the left of the ancestral line

(these are treated in a similar manner, considering now the excursions of $(\mathbf{e}_{(R_i^\varepsilon - t)^+})_{t \geq 0}$ above its past minimum process that hit 1).

Although we avoided introducing the notation that would be needed to make the previous arguments precise, the reader will easily turn these arguments into a rigorous proof of Proposition 5 based on Lemma 26.

Acknowledgment. We would like to thank Thordur Jonsson for suggesting the study of the harmonic measure on critical Galton–Watson trees during Spring 2012.

REFERENCES

- [1] E. AÏDÉKON (2014). Speed of the biased random walk on a Galton–Watson tree. *Probab. Theory Related Fields* **159** 597–617.
- [2] D. ALDOUS (1991) The continuum random tree I. *Ann. Probab.* **19** 1–28.
- [3] D. ALDOUS (1993). The continuum random tree III. *Ann. Probab.* **21** 248–289.
- [4] S. ATHREYA, M. ECKHOFF, AND A. WINTER (2013). Brownian motion on \mathbb{R} -trees. *Trans. Amer. Math. Soc.* **365** 3115–3150.
- [5] K.B. ATHREYA, P. E. NEY (1972). *Branching Processes*. Springer, New York-Heidelberg.
- [6] R.M. BLUMENTHAL (1992). *Excursions of Markov Processes*. Birkhäuser, Boston.
- [7] A.N. BORODIN, P. SALMINEN (2002). *Handbook of Brownian Motion - Facts and Formulae*. Birkhäuser, Basel.
- [8] J. BOURGAIN (1987). On the Hausdorff dimension of harmonic measure in higher dimension. *Invent. Math.* **87** 477–483.
- [9] B. CHAUVIN, A. ROUAULT, A. WAKOLBINGER (1991). Growing conditioned trees. *Stoch. Process. Appl.* **39** 117–130.
- [10] D. CROYDON (2008). Volume growth and heat kernel estimates for the continuum random tree. *Probab. Theory Relat. Fields* **140** 207–238.
- [11] T. DUQUESNE, J.-F. LE GALL (2002). Random Trees, Lévy Processes and Spatial Branching Processes, *Astérisque* **281**.
- [12] T. DUQUESNE, J.-F. LE GALL (2006). The Hausdorff measure of stable trees. *Alea* **1** 393–415.
- [13] N. ENRIQUEZ, Y. KIFER (2001). Markov chains on graphs and Brownian motion, *J. Theoret. Probab.* **14** 495–510.
- [14] K. FLEISCHMANN, R. SIEGMUND-SCHULTZE (1977). The structure of reduced critical Galton–Watson processes, *Math. Nachr.* **79** 233–241.
- [15] M. FREIDLIN, S.-J. SHEU (2000). Diffusion processes on graphs: stochastic differential equations, large deviation principle. *Probab. Theory Related Fields* **116** 181–220.
- [16] N. GANTERT, S. MÜLLER, S. POPOV, AND M. VACHKOVSKAIA (2012). Random walks on Galton–Watson trees with random conductances. *Stoch. Process. Appl.* **122** 1652–1671.
- [17] S. JANSON (2006). Random cutting and records in deterministic and random trees. *Random Structures Algorithms* **29** 139–179.
- [18] T. JONSSON (2012). Private communication.

- [19] O. KALLENBERG (2010). *Foundations of Modern Probability*. Springer-Verlag, Berlin.
- [20] W. B. KREBS (1995). Brownian motion on the continuum tree. *Probab. Theory Related Fields* **101** 421–433.
- [21] G. F. LAWLER (1993). A discrete analogue of a theorem of Makarov. *Combin. Probab. Comput.* **2** 181–199.
- [22] J.-F. LE GALL (1989). Marches aléatoires, mouvement brownien et processus de branchement. *Séminaire de Probabilités XXIII. Lecture Notes Math.* **1372** 258–274. Springer, Berlin.
- [23] J.-F. LE GALL (2005). Random trees and applications. *Probab. Surveys* **2** 245–311.
- [24] S. LIN (2014). The harmonic measure of balls in critical Galton–Watson trees with infinite variance offspring distribution. *Electron. J. Probab.* **19** 1–35.
- [25] S. LIN (2015). Typical behavior of the harmonic measure in critical Galton-Watson trees. Preprint available at arXiv:1502.05584
- [26] R. LYONS (2000). Singularity of some random continued fractions. *J. Theoret. Probab.* **13** 535–545.
- [27] R. LYONS, R. PEMANTLE, AND Y. PERES (1995). Ergodic theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynam. Systems* **15** 593–619.
- [28] R. LYONS, R. PEMANTLE, AND Y. PERES (1996). Biased random walks on Galton-Watson trees. *Probab. Theory Related Fields* **106** 249–264.
- [29] R. LYONS AND Y. PERES (2015) *Probability on Trees and Networks*. Book in preparation, current version available at <http://mypage.iu.edu/~rdlyons/>
- [30] N. G. MAKAROV (1985). On the distortion of boundary sets under conformal mappings. *Proc. London Math. Soc.* (3) **51** 369–384.
- [31] J. PITMAN (2006). *Combinatorial Stochastic Processes. Lecture Notes Math.* **1875**. Springer-Verlag, Berlin.
- [32] D. REVUZ, M. YOR (1991). *Continuous Martingales and Brownian Motion*, Springer-Verlag, Berlin.
- [33] A.M. ZUBKOV (1975). Limit distributions of the distance to the nearest common ancestor (Russian). *Teor. Veroyatnost. i Primenen.* **20** 614–623.

MATHÉMATIQUES, BÂT. 425
 UNIVERSITÉ PARIS-SUD
 91405 ORSAY CEDEX
 FRANCE
 E-MAIL: nicolas.curien@gmail.com

MATHÉMATIQUES, BÂT. 425
 UNIVERSITÉ PARIS-SUD
 91405 ORSAY CEDEX
 FRANCE
 E-MAIL: jean-francois.legall@math.u-psud.fr