



INSTITUT
POLYTECHNIQUE
DE PARIS

*J*ournal de l'École polytechnique *Mathématiques*

Alessandra CARACENI, Nicolas CURIEN, & Robin STEPHENSON

Where do (random) trees grow leaves?

Tome 13 (2026), p. 1151-1184.

<https://doi.org/10.5802/jep.345>

© Les auteurs, 2026.



Cet article est mis à disposition selon les termes de la licence
LICENCE INTERNATIONALE D'ATTRIBUTION CREATIVE COMMONS BY 4.0.
<https://creativecommons.org/licenses/by/4.0/>

Publié avec le soutien
du Centre National de la Recherche Scientifique



Publication membre du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 2270-518X

WHERE DO (RANDOM) TREES GROW LEAVES?

BY ALESSANDRA CARACENI, NICOLAS CURIEN
& ROBIN STEPHENSON

ABSTRACT. — We study a model of random binary trees grown “by the leaves” in the style of Luczak and Winkler [LW04]. If τ_n is a uniform plane binary tree of size n , Luczak and Winkler, and later explicitly Caraceni and Stauffer, constructed a measure ν_{τ_n} such that the tree obtained by adding a cherry on a leaf sampled according to ν_{τ_n} is still uniformly distributed on the set of all plane binary trees with size $n + 1$. It turns out that the measure ν_{τ_n} , which we call the leaf-growth measure, is noticeably different from the uniform measure on the leaves of the tree τ_n . In fact we prove that, as $n \rightarrow \infty$, with high probability it is almost entirely supported by a subset of only $n^{3(2-\sqrt{3})+o(1)} \approx n^{0.8038\dots}$ leaves. In the continuous setting, we construct the scaling limit of uniform binary trees equipped with this measure, which is the Brownian Continuum Random Tree equipped with a probability measure supported by a fractal set of dimension $6(2 - \sqrt{3})$. We also compute the full (discrete) multifractal spectrum. This work is a first step towards understanding the diffusion limit of the discrete leaf-growth procedure.

RÉSUMÉ. — Nous étudions un modèle d’arbres aléatoires croissant « par les feuilles », dans l’esprit de Luczak et Winkler [LW04]. Si τ_n désigne un arbre plan binaire uniforme de taille n , Luczak et Winkler, puis plus explicitement Caraceni et Stauffer, ont construit une mesure ν_{τ_n} telle que l’arbre obtenu en ajoutant une « cerise » à une feuille choisie selon la distribution ν_{τ_n} soit encore uniforme dans l’ensemble des arbres plans binaires de taille $n + 1$. La mesure ν_{τ_n} s’avère remarquablement différente de la mesure uniforme sur les feuilles de τ_n . En particulier, nous prouvons que, lorsque $n \rightarrow \infty$, avec grande probabilité, elle est presque entièrement supportée par $n^{3(2-\sqrt{3})+o(1)} \approx n^{0.8038\dots}$ feuilles. Nous construisons également sa limite d’échelle, une mesure de probabilité sur l’arbre continu brownien dont le support est fractal, de dimension $6(2 - \sqrt{3})$. Nous calculons de plus son spectre multifractal. Ce travail constitue une première étape vers la compréhension de la diffusion obtenue comme limite de la procédure de croissance par les feuilles.

CONTENTS

1. Introduction.....	1152
2. Building schemes for trees: leaf-growth measures.....	1157
3. The leaf-growth measure on the Brownian CRT.....	1163
4. Discrete fractal properties of leaf-growth measures.....	1171
5. Towards a diffusion limit.....	1178

MATHEMATICAL SUBJECT CLASSIFICATION (2020). — 05C05, 05C10, 05C80, 60D05, 60F17, 60G18.

KEYWORDS. — Random trees, scaling limits, self-similar fragmentations, Hausdorff dimension, multifractal spectrum.

The second author is supported by “SuperGrandMa”, the ERC Consolidator Grant No. 101087572.

Appendix A. Computing the integral (18).....	1180
Appendix B. Limits of trees and subtrees.....	1182
References.....	1183

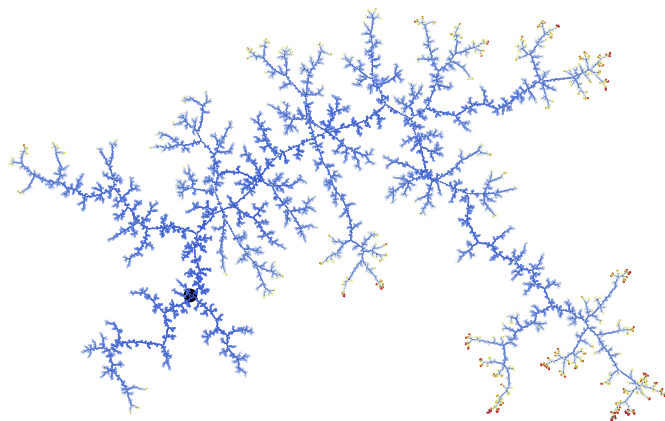


FIGURE 1. A uniform plane binary tree τ_n with 20 000 edges decorated with its leaf-growth measure ν_{τ_n} : the black disk represents the root vertex; and the colour of the other vertices (leaves) displays their ν_{τ_n} -mass (blue for small probability and red for high probability).

1. INTRODUCTION

Within the very broad and far-reaching topic of randomly generated trees, models involving some sort of growth procedure arise very naturally and lead to a myriad of interesting questions and results. Well-known examples include random recursive trees and preferential attachment trees (e.g., the Barabasi–Albert model), as well as many tree models arising as data structures (binary search trees, d -ary trees, quadrees, tries, etc.). We shall focus here on iterative growth procedures which, when performed up to size n , yield a tree that is uniformly distributed within the set of all size- n plane d -ary trees. One famous such procedure is given by Rémy’s algorithm and its variants [Rém85, Mar03, HS15], but others include [Mar23].

In this paper, we shall consider a model of growth “by the leaves” in the style of what Luczak and Winkler introduced in [LW04]. Their question was the following: letting $\tau_n^{(d)}$ be a uniformly random plane d -ary tree with n internal vertices (every vertex has either d or 0 children), is it possible to couple $\tau_n^{(d)}$ and $\tau_{n+1}^{(d)}$ in such a way that the latter is obtained from the former by adding d children to a leaf of $\tau_n^{(d)}$ (see Figure 2)? They answer the question in the affirmative, although their proof does not yield an explicit way of selecting a random leaf ℓ , conditionally on $\tau_n^{(d)}$, in such a way that giving it d new children will yield a uniform tree of size $n + 1$. On top of its

theoretical beauty, this result has had significant applications to stochastic domination of random trees, which was for example used to study planar maps [AB14, CS23], the parking model on trees [GP19], and the number of spanning trees in the Erdős–Rényi random graph [LPS08].

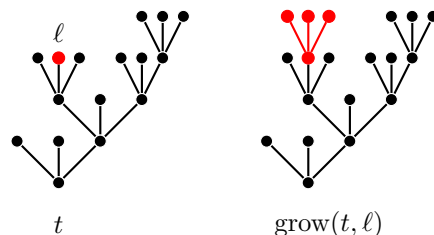


FIGURE 2. A plane ternary tree of size 5 (i.e., with 5 internal vertices and thus $3 + 2 \cdot 4 = 11$ leaves) is grown at the leaf ℓ by adding 3 children to ℓ , thus creating the tree $\text{grow}(t, \ell)$ of size 6.

In the case $d = 2$ of plane binary trees, however, such a coupling between τ_n and τ_{n+1} – where d is henceforth omitted – can be made very explicit (and in fact it is unique provided some natural symmetries are enforced). Given a uniform plane binary tree τ_n of size n (i.e., with n vertices that are *not* leaves, and hence with $n + 1$ leaves), one can produce an explicit measure ν_{τ_n} on the set of the $n + 1$ leaves of τ_n with the following property. If, conditionally on τ_n , one samples ℓ according to ν_{τ_n} , then the tree $\text{grow}(\tau_n, \ell)$ obtained by attaching a “cherry” (two sister leaves) to ℓ is uniformly distributed among all plane binary trees with $n + 1$ vertices. See Figure 2.

Construction of the leaf-growth measures. — The (random) measure ν_{τ_n} , which we call the *leaf-growth* measure, is explicit as a function of τ_n . Precisely, given $\tau_n = t$, and given any leaf ℓ of t , the mass given to ℓ is

$$(1) \quad \nu_t(\ell) = \prod_{i=1}^{|\ell|} C(a_i, b_i),$$

where

- $|\ell|$ is the height of ℓ in t ;
- if v_i is the ancestor of ℓ at height i , then a_i is the size of the subtree of descendants of v_i (including v_i itself);
- $b_i = a_{i-1} - a_i - 1$ (setting $a_0 = n$) is the size of the subtree of descendants of the *sibling* of v_i ;
- $C(a, b)$ is the explicit rational function

$$C(a, b) := \frac{(a+1)(2a+1)(a+3b+3)}{(a+b+1)(a+b+2)(2(a+b)+3)}.$$

See [CS20] for details. In fact, there is a nice combinatorial interpretation to the expression $C(a, b)$ based on a “best-of-three” match (as in plenty of competitive games

and sports) which begs for a bijective explanation for the function C , see Remark 2.5. The full details of the construction are given in Section 2.

Fractal properties of the leaf-growth measures. — In this paper, we intend to investigate the nature of the measure ν_{τ_n} : how much does it differ from the uniform measure on the set of all leaves of τ_n ? In other words, how “spread out” do we expect it to be among all possible leaves?

We provide a rather detailed and explicit answer to this question. We shall prove that, with high probability, the measure ν_{τ_n} is actually supported by a “small” set of leaves, whose size is of order $n^{3(2-\sqrt{3})+o(1)}$. Indeed, in Section 4.1 we will show the following:

THEOREM 1.1. — *Let τ_n be a uniformly distributed random binary tree of size n and let ν_{τ_n} be its leaf-growth measure; set*

$$\gamma = 3(2 - \sqrt{3}).$$

For all $\varepsilon > 0$, we have

$$\lim_{n \rightarrow \infty} \nu_{\tau_n}(\{\ell \in \tau_n \mid n^{-\gamma-\varepsilon} \leq \nu_{\tau_n}(\ell) \leq n^{-\gamma+\varepsilon}\}) = 1$$

in probability.

As a consequence, for all $\delta \in (0, 1)$ there exists with high probability a set of leaves $A_{n,\delta}$ of τ_n such that $\nu_{\tau_n}(A_{n,\delta}) \geq 1 - \delta$ and $|A_{n,\delta}| \leq n^{\gamma+\varepsilon}$. Moreover, the maximal measure according to ν_{τ_n} of a set of cardinality bounded above by $n^{\gamma-\varepsilon}$ tends to zero in probability as $n \rightarrow \infty$.

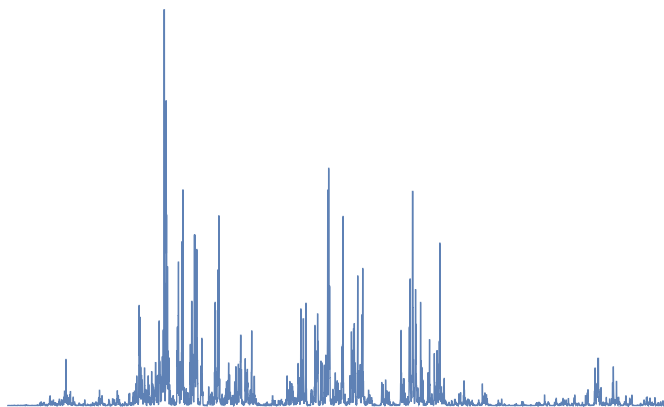
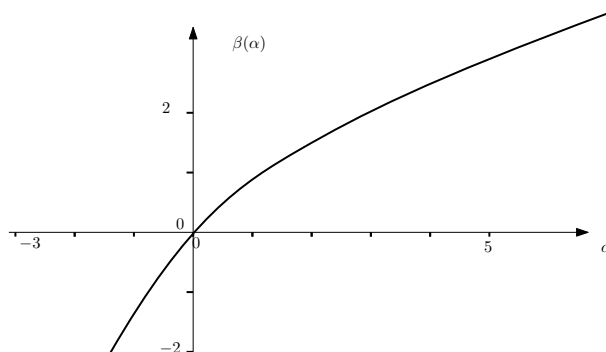


FIGURE 3. A plot of the density of the discrete leaf-growth measure with respect to the discrete uniform measure on the leaves of the tree depicted in Figure 1 (starting from the root and going counterclockwise around the tree). In the limit, those two measures are mutually singular.

FIGURE 4. Illustration of the function $\alpha \rightarrow \beta(\alpha)$.

Refer to [CLG17, Th. 1] for a similar phenomenon for the discrete harmonic measure in random trees and [Law93] for a discrete analogue of Makarov’s theorem. The above result is characteristic of a “fractal” behaviour of the measure ν_{τ_n} and indeed we shall see in Proposition 3.5 that the continuous analogue of the measure ν_{τ_n} is supported by a fractal subset of the Brownian CRT with dimension 2γ . We actually go further and compute the full (discrete) multifractal spectrum of the measure ν_{τ_n} . In plain terms, although Theorem 1.1 shows that the typical ν_{τ_n} -mass of a leaf sampled according to ν_{τ_n} is $n^{-\gamma+o(1)}$, its higher moments are not ruled by the typical behaviour (which is what we call multifractality):

THEOREM 1.2 (Discrete multifractal spectrum). — *Given $\alpha \in \mathbb{R}$, let $\beta(\alpha)$ be the unique value of β for which the integral*

$$(2) \quad I(\alpha, \beta) = \int_0^1 dx \frac{1}{\sqrt{x^3(1-x)^3}} \left(c(x)^{\alpha+1} x^{-\beta} + c(1-x)^{\alpha+1} (1-x)^{-\beta} - 1 \right)$$

is zero, where

$$c(x) = x^2(3-2x).$$

As $n \rightarrow \infty$, we have

$$\mathbb{E} \left[\sum_{\ell \in \tau_n} \nu_{\tau_n}(\ell)^{\alpha+1} \right] = n^{-\beta(\alpha)+o(1)}.$$

See Figure 4. In particular for $\alpha = k \in \{0, 1, 2, \dots\}$, the above display computes the (annealed) k th moment of the mass of a point sampled according to ν_{τ_n} which is asymptotic to $n^{-\beta(k)+o(1)}$. In the case of the first moment $\alpha = 1$ we have the explicit value $\beta(1) = \frac{1}{2}(5 - \sqrt{13})$. The proof of this result is presented in Section 4.2, along with some further comments on the minimal/maximal leaf-growth weight of a single leaf and the multi-fractal spectrum (see Remark 4.3).

Continuum limit. — The investigation of the natural discrete measure ν_{τ_n} inevitably leads to the construction of a corresponding limiting measure $\nu_{\mathcal{T}}$ on the Brownian

Continuum Random Tree (CRT) of David Aldous [Ald91a, Ald91b]. In Section 3, we show the following scaling limit result:

THEOREM 1.3. — *We have the following convergence in distribution for the Gromov–Hausdorff–Prokhorov topology (see Section 3.5 for a definition):*

$$(n^{-1/2} \cdot \tau_n, \nu_{\tau_n}) \xrightarrow[n \rightarrow \infty]{(d)} (2\sqrt{2} \mathcal{T}, \nu_{\mathcal{T}}),$$

where $\mathcal{T}$ is the Brownian CRT and $\nu_{\mathcal{T}}$ is a probability measure on the leaves of $\mathcal{T}$ with Hausdorff dimension $2\gamma = 6(2 - \sqrt{3}) \approx 1.6077$.

Recall that the dimension of a measure ν on a metric space E is the infimum of the Hausdorff dimensions of Borel subsets with full ν -measure. Our main tool is a description of the law of the spine towards a $\nu_{\mathcal{T}}$ -typical point which enables us to compute the Hausdorff dimension of the measure (Proposition 3.3), as well as distributional properties such as the (annealed) distribution of the height of a $\nu_{\mathcal{T}}$ -typical point.⁽¹⁾ Indeed, the continuous version of the measure sheds further light on many properties of its discrete counterparts, and techniques will be borrowed from the analysis of the continuous processes involved to produce the results already mentioned above.

Towards a diffusion on real trees. — The detailed analysis of the leaf-growth measure both in the discrete and continuous setting performed in this work can be seen as a first step in the ambitious program of understanding the scaling limit of the leaf-growth process itself. Indeed, it is natural to postulate that the rescaled leaf-growth dynamics would converge towards an appropriate continuous Markov process with values in real trees whose invariant measure is the Brownian CRT (such a dynamic should be non-trivial by Proposition 5.1). Several such dynamics have been the object of intense study in recent years: from the Rémy dynamic, which almost surely converges in the GHP sense [CH13], to the Aldous move on cladograms converging to the Aldous diffusion [LMW20, FPRW18, FPRW20, FPRW23]; see also the literature about the process of root growth with regrafting [EPW06]. The new continuous dynamic obtained by the leaf-growth process is constructed in the work [CFT25] dealing with the much more general self-similar Markov trees, while the discrete convergence will be addressed in a forthcoming work.

Acknowledgements. — The authors would like to thank Nic Freeman, Guillaume Conchon-Kerjan, Adrianus Twigt and Davide Lombardo for helpful discussions. They are grateful to the two anonymous referees for their careful readings and their useful remarks that helped improve the paper.

⁽¹⁾This height turns out to have density $8x^3e^{-2x^2}$ (see Proposition 3.3): as one might expect, this height stochastically dominates that of a random leaf chosen *uniformly*, since the leaf-growth measure must favour leaves belonging to larger subtrees. Unfortunately, we have no natural explanation for the surprisingly straightforward form of this density.

2. BUILDING SCHEMES FOR TREES: LEAF-GROWTH MEASURES

In this section, we set up some notation and give some context for what we are about to discuss, i.e., how to “uniformly grow” trees “from their leaves”. For completeness, we shall state some known results in their general form, which involves d -ary trees (see Luczak & Winkler [LW04]).

In the case $d = 2$, motivated by applications to the study of mixing times of flip chains, this procedure has been made completely explicit by Caraceni and Stauffer [CS20] and forms the basis of the current work.

2.1. BUILDING SCHEMES FOR d -ARY TREES

DEFINITION 2.1. — Let $d \geq 2$ and $n \geq 0$ be integers. We shall call a d -ary tree of size n a rooted plane tree whose vertices all have d children (we call such vertices *internal*) or no children (we call such vertices *leaves*).

We say that a d -ary tree τ has size $|\tau| = n$ if it has n internal vertices. For simplicity, we shall commit a slight abuse of notation and also write τ for the set of vertices of τ , so we can write $v \in \tau$ to mean that v is a vertex of τ .

We call $\mathsf{T}_n^{(d)}$ the set of all d -ary trees of size n and write $\mathsf{T}^{(d)}$ for $\bigcup_{n \geq 0} \mathsf{T}_n^{(d)}$. Note that the set $\mathsf{T}_0^{(d)}$ contains one element, the tree consisting of only its root.

A d -ary tree has a natural recursive structure: if $t \in \mathsf{T}_n^{(d)}$ with $n \geq 1$, it naturally induces a sequence of d -ary trees $(t^1, \dots, t^d)$: letting $\rho \in t$ be the root, one can erase all children of ρ other than the i th child v_i and consider the connected component of v_i , rooted at v_i , to obtain t^i . One then has $\sum_{i=1}^d |t^i| = n - 1$. From this decomposition one obtains the *profile* of the tree t :

DEFINITION 2.2. — Given $t \in \mathsf{T}_n^{(d)}$ with $n \geq 1$, its *profile* is the vector $\underline{x}(t) = (|t^1|, \dots, |t^d|) \in \mathbb{N}^d$.

Note that the term profile is often used with different meanings (for example, for the sequence of generation sizes) in similar contexts; however, we choose to employ it here to keep notation at least partially consistent with the original paper [LW04].

We now discuss a particular “growth” operation on d -ary trees. Given a pair (t, ℓ) , where t is a d -ary tree of size n and ℓ is a leaf of t , we construct a tree $\text{grow}(t, \ell)$ of size $n + 1$ by adding d leaves to t as children of the vertex ℓ , thus turning ℓ into an internal vertex.

Luczak and Winkler consider the following natural question in [LW04]: supposing τ_n is a uniform random element of $\mathsf{T}_n^{(d)}$ and τ_{n+1} is a uniform random element of $\mathsf{T}_{n+1}^{(d)}$, is it possible to couple τ_n and τ_{n+1} in such a way that τ_{n+1} is always of the form $\text{grow}(\tau_n, \ell_n)$, where ℓ_n is some leaf of the tree τ_n ?

They call a family of such couplings for all $n \geq 0$ a *building scheme* for d -ary trees and answer the question in the affirmative, thus proving

THEOREM 2.3 (Luczak, Winkler, 2004). — *For all $d \geq 2$ there exists a building scheme for d -ary trees.*

Note that one can present a building scheme as a family of probability measures ν_t , one for each d -ary tree t . The probability measure ν_t is defined on the set of leaves of t and has the following property. If (τ_n, ℓ) is a random variable such that τ_n is uniform in $\mathsf{T}_n^{(d)}$ and, conditionally on $\tau_n = t$, ℓ is a leaf of t distributed according to ν_t , then $\text{grow}(\tau_n, \ell)$ is uniform in $\mathsf{T}_{n+1}^{(d)}$.

Moreover, what Luczak and Winkler obtain is that there is a *recursive* building scheme for d -ary trees, which we can think of as a family $(\nu_t)_{t \in \mathsf{T}^{(d)}}$ such that for all (t, ℓ) with $t \in \mathsf{T}^{(d)}$ of size $|t| \geq 1$, ℓ leaf of t , we have

$$\nu_t(\ell) = \sum_{i=1}^d \mathbf{1}_{\ell \in t^i} f_\nu(i; \underline{x}(t)) \nu_{t^i}(\ell),$$

where $\underline{x}(t)$ is the profile of t and we have omitted curly brackets from the arguments of the measure $\nu_t(\cdot)$ (as we shall do consistently from now on). The function $f_\nu(\cdot; \cdot) : [d] \times \mathbb{N}^d \rightarrow [0, 1]$, when computed for a certain $i \in [d]$ and the profile $\underline{x}(t)$ of a tree t , yields the probability $\nu_t(\ell \in t^i)$, which only depends on $\underline{x}(t)$ and not on the actual shape of the trees $(t^1, \dots, t^d)$.

A recursive building scheme is completely determined by the functions $f_\nu(\cdot; \cdot)$, since

$$\nu_t(\ell) = \prod_{v \prec \ell} f_\nu(i_v; \underline{x}^v),$$

where the product is taken over all internal ancestors v of ℓ , v being the i_v th child of its parent and $\underline{x}^v$ being the profile of the tree of descendants of v , which serves as the root. Moreover, the functions f_ν must satisfy the obvious equation

$$\sum_{i=1}^d f_\nu(i; \underline{x}) = 1,$$

as well as a recurrence ensuring they are a building scheme. Letting τ_n and τ_{n+1} be uniform in $\mathsf{T}_n^{(d)}, \mathsf{T}_{n+1}^{(d)}$ and letting e_i be the i th canonical basis vector in $\mathbb{N}^d$, we must have

$$(3) \quad \Pr((|\tau_{n+1}^i|)_{i=1}^d = \underline{x}) = \sum_{i=1}^d \mathbf{1}_{x_i \geq 1} \Pr((|\tau_n^i|)_{i=1}^d = \underline{x} - e_i) f_\nu(i; \underline{x} - e_i),$$

where $\underline{x}_i$ is the i th entry of the vector $\underline{x}$, which simply amounts to requiring that the building scheme act correctly on the probabilities of the profiles.

The proof of Theorem 2.3 given in [LW04] leverages the fact that the probability that τ_n exhibits a certain profile can be computed explicitly in order to prove inequalities that are sufficient for the existence of a functions f_ν satisfying the above conditions and also taking values in $[0, 1]$. This is thanks to a max-flow-min-cut type of argument, which does not suggest explicit expressions for possible solutions f_ν , in spite of the fact that the coefficients of the recursions have fairly simple expressions.

It is at this point that we restrict ourselves to the simpler case $d = 2$: we will now introduce some more compact notation and describe an explicit building scheme as done in [CS20].

2.2. THE LEAF-GROWTH MEASURE FOR BINARY TREES. — We shall henceforth only deal with what we will simply call *binary trees*, that is, trees in $\mathbb{T}^{(2)}$. We will drop d from all notations, writing $\mathbb{T}$ for $\mathbb{T}^{(2)}$ and $\mathbb{T}_n$ for $\mathbb{T}_n^{(2)}$. Given $t \in \mathbb{T}$ such that $|t| \geq 1$, we shall write t^L and t^R for the trees t^1 and t^2 , identifying them as *the left subtree* and *the right subtree*. Moreover, we have the well-known result that binary trees are counted by Catalan numbers, i.e., that

$$|\mathbb{T}_n| = \text{Cat}(n) = \frac{1}{n+1} \binom{2n}{n}.$$

This implies that, given a profile $(a, b) \in \mathbb{N}^2$ and a uniform tree τ_{a+b+1} from $\mathbb{T}_{a+b+1}$, we have

$$\Pr(\underline{x}(\tau_{a+b+1}) = (a, b)) = \frac{\text{Cat}(a) \text{Cat}(b)}{\text{Cat}(a+b+1)}.$$

We shall need a compact notation for this expression, which from now on we will simply write as $P(a, b)$.

Now suppose we were to look for a recursive building scheme f_ν for binary trees, with the additional natural property of being *symmetric*, that is, that $f_\nu(1; (a, b)) = f_\nu(2; (b, a))$. In this case, f_ν can be computed explicitly and takes a very simple form. Indeed, as shown within the proof of [CS20, Prop. 5.1], we have

THEOREM 2.4. — *The function*

$$(4) \quad C(a, b) = \frac{(a+1)(2a+1)(a+3b+3)}{(a+b+1)(a+b+2)(2(a+b)+3)},$$

defined on pairs (a, b) of non-negative integers, is the only solution to the equations:

$$C(a, b) + C(b, a) = 1;$$

$$P(a+1, b+1) = P(a, b+1)C(a, b+1) + P(a+1, b)(1 - C(a+1, b))$$

for all integers $a, b \geq -1$, where we set $P(a, -1) = P(-1, b) = 0$ (and therefore $C(a, b)$ appears in the equation only if $a, b \geq 0$).

Therefore, the only symmetric recursive building scheme for binary trees is the family of measures $(\nu_t)_{t \in \mathbb{T}}$, where ν_t is the probability measure on the leaves of t satisfying the recursion

$$(5) \quad \nu_t(\ell) = 1_{\ell \in t^L} C(|t^L|, |t^R|) \nu_{t^L}(\ell) + 1_{\ell \in t^R} (1 - C(|t^L|, |t^R|)) \nu_{t^R}(\ell).$$

From now on, we shall write $(\nu_t)_{t \in \mathbb{T}}$ for the unique symmetric recursive building scheme for binary trees; we will call ν_t the *leaf-growth measure* on t . Notice that the equations in the statement above are precisely the requirement (3) in order for ν_t to be a building scheme, with the added requirement of symmetry.

Remark that the uniqueness of the solution is clear, since one can inductively compute $C(a, b)$, for $a > b$, by using equations of the third type to reduce the value

of b until it becomes 0 (while obviously $C(a, a) = 1/2$). The fact that the expression (4) is a solution can be easily checked by induction.

REMARK 2.5 (A sports question). — The expression (4) for $C(a, b)$ has an interesting combinatorial interpretation. Imagine a match played between L (“left”) and R (“right”), with L initially having $2a + 1$ tokens and R having $2b + 1$. Each player has probability of winning the match proportional to their number of tokens: for the first match, L wins with probability $(2a + 1)/(2a + 2b + 2)$ and R wins with probability $(2b + 1)/(2a + 2b + 2)$. After a match is played, the winner gains one token. Any subsequent matches have updated outcome probabilities, but are played independently.

In other words, the three matches are the first three draws from a classical two-colour Pólya urn with initial composition $(2a + 1, 2b + 1)$ (where the colour drawn receives one additional token).

The quantity $C(a, b)$ is precisely the probability that L wins a majority of three successive matches, that is,

$$C(a, b) = 3 \frac{(2a + 1)(2a + 2)(2b + 1)}{(2a + 2b + 2)(2a + 2b + 3)(2a + 2b + 4)} + \frac{(2a + 1)(2a + 2)(2a + 3)}{(2a + 2b + 2)(2a + 2b + 3)(2a + 2b + 4)}.$$

We do not have a direct combinatorial proof of the link between the tree growth measure and this presentation in terms of best-of-three winning probabilities, but this interpretation can be used to give alternative descriptions of the recursive symmetric building scheme, such as the one that follows.

COROLLARY 2.6. — *Given $n \geq 1$ and $t \in \mathsf{T}_n$, recall that t has $n + 1$ leaves; place exactly $n + 1$ tokens, numbered 1 to $n + 1$, one onto each leaf of t . Now play the following game:*

- *at each step, pick a vertex v with tokens such that its descendants have no tokens (initially, one can pick any leaf of t); move all the tokens of v to the parent of v ;*
- *if a vertex w contains $n_a > 0$ tokens numbered a and $n_b > 0$ tokens numbered $b \neq a$, play a best-of-three game on the vertex according to the rules described in Remark 2.5 with one player having n_a starting tokens and the other having n_b . After all three matches are played, three tokens have been added and the winner has been determined. Relabel all tokens at w with a or b , according to which player was the winner, and destroy two tokens (thus leaving $n_a + n_b + 1$ on w).*
- *unless all tokens have the same label, perform another step.*

The leaf corresponding to the final surviving label is distributed according to ν_t .

Proof. — We say that a pile of tokens on a vertex v is *movable* if they all have the same label and there are no tokens on any strict descendants of v . Letting t^v be the tree of descendants of v , if the tokens of v are movable, then they are $2|t^v| + 1$. Single tokens on leaves are movable, and they are indeed $2 \cdot 0 + 1$. Inductively, when a new movable pile is made on an internal vertex v , one can say that it must have been

created by merging a movable pile from its left child v_ℓ and one from its right child v_r , which must be of sizes $2|t^{v_\ell}| + 1$ and $2|t^{v_r}| + 1$, respectively. Thus a best-of-three game is played in v , yielding a movable pile of $2|t^{v_\ell}| + 2|t^{v_r}| + 2 + 3 - 2 = 2|t^v| + 1$ tokens. The probability that the label of the tokens is that of a leaf of t^{v_ℓ} is $C(|t^{v_\ell}|, |t^{v_r}|)$. The procedure as described plays a best-of-three game in each internal vertex of t . Letting t^L, t^R be the left and right subtrees, the probability that the final label is $\ell \in [n + 1]$ is indeed (inductively) given by (5). $\square$

2.3. SOME USEFUL ASYMPTOTICS. — Before we move on to investigate several aspects of the leaf-growth measures ν_t , it is convenient to state some of the main (easy or well known) approximation results we will need in what follows, both for the quantities $C(a, b)$ and the probabilities $P(a, b)$.

Note that, fixing $b = n - 1 - a$, we can write out the values of $P(a, b)$ as a quotient of factorials

$$P(a, b) = \frac{(2a)!(2b)!n!(n+1)!}{a!(a+1)!b!(b+1)!(2n)!}.$$

Using Stirling's approximation, it is clear that, for $x \in (0, 1)$,

$$(6) \quad 4\sqrt{\pi} \lim_{\substack{n \rightarrow \infty \\ a/n \rightarrow x}} n^{3/2} P(a, b) = \frac{1}{x^{3/2}(1-x)^{3/2}} =: \mathbf{p}(x);$$

moreover, for all $x \in [0, 1]$,

$$(7) \quad \lim_{\substack{n \rightarrow \infty \\ a/n \rightarrow x}} C(a, b) = x^2(3-2x) =: \mathbf{c}(x).$$

The more compact notation $\mathbf{p}(x)$ and $\mathbf{c}(x)$ for the functions $x^{-3/2}(1-x)^{-3/2}$ and $x^2(3-2x)$ (see Figure 5) that is being introduced here will be used frequently in future sections.

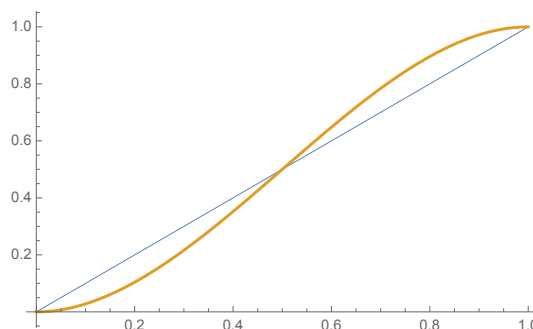


FIGURE 5. A plot of the function $x \mapsto \mathbf{c}(x)$. Notice that it is symmetric with respect to $(1/2, 1/2)$ and below the first bisector when $x \in [0, 1/2]$.

Over the course of the paper we will encounter several Riemann sums (of improper integrals) of the form $\sum_{a=1}^{n-1} f(a/n)P(a, n-1-a)C(a, n-1-a)^k$ for some integer

$k \in \{0, 1, 2, \dots\}$. Under some mild conditions on f and k , they exhibit the expected behaviour:

PROPOSITION 2.7. — *Let f be a continuous function on $(0, 1)$, and $k \in \{0, 1, 2, \dots\}$. Assume that $|f(x)| \leq K(x \wedge (1 - x))^{1-2k}$ for some $K > 0$, and if $k > 0$, assume further that $\sup\{f(a/n), a \in \{1, \dots, n-1\}\} = o(\sqrt{n})$. We then have*

$$(8) \quad \lim_{n \rightarrow \infty} \sqrt{n} \sum_{a=1}^{n-1} f(a/n) P(a, n-1-a) C(a, n-a-1)^k = \frac{1}{4\sqrt{\pi}} \int_0^1 f(x) c(x)^k p(x) dx.$$

Note that, while the $k = 0$ case is essentially a variant of [HM12, Prop. 39], we prefer giving a self-contained proof here.

Proof. — The following uniform bounds on C and P are easily checked: there exists a constant $K' > 0$ such that

$$(9) \quad n^{3/2} P(a, b) \leq \frac{K'}{(a/n)^{3/2} (b/n)^{3/2}},$$

and

$$(10) \quad |C(a, b) - c(a/n)| \leq \frac{1}{n},$$

where $n = a + b + 1$, and (9) is immediately obtained by Stirling's approximation.

We start with the $k = 0$ case. Specifically, we are going to show

$$\sqrt{n} \sum_{a=2}^{\lfloor n/2 \rfloor} f(a/n) P(a, n-1-a) \longrightarrow \frac{1}{4\sqrt{\pi}} \int_0^{1/2} f(x) p(x) dx.$$

This is enough to conclude by first exploiting the natural symmetry around $1/2$, then observing that $P(\lfloor n/2 \rfloor, \lfloor n/2 \rfloor) = O(n^{-3/2})$, so overcounting middle terms is not an issue, and finally noting that $f(1/n)P(1, n-2) = O(1/n)$, so removing the first term is also not an issue.

Write $n^{1/2} \sum_{a=2}^{\lfloor n/2 \rfloor} f(a/n) P(a, n-1-a)$ as $\int_0^{1/2} g_n(x) dx$, where g_n is the function defined by $g_n(x) = n^{3/2} f(a/n) P(a, n-1-a)$ for $x \in [\frac{a-1}{n}, \frac{a}{n}]$ and $a \in \{2, \dots, \lfloor n/2 \rfloor\}$, and $g_n(x) = 0$ otherwise. It is readily checked that $g_n(x)$ converges pointwise to $(1/4\sqrt{\pi})f(x)p(x)$, and we just need to check that we can apply the dominated convergence theorem. To do this, notice that, if $x \in [\frac{a-1}{n}, \frac{a}{n}]$ for $2 \leq a \leq \lfloor n/2 \rfloor$, then $x + 1/n \leq 2x$, while $n-1-a \geq n/3$. Hence, by (9), we have

$$\begin{aligned} g_n(x) &= n^{3/2} f(a/n) P(a, n-1-a) \\ &\leq K K' \frac{a}{n} \left(\frac{a}{n}\right)^{-3/2} \left(\frac{n-1-a}{n}\right)^{-3/2} \\ &\leq 3^{-3/2} K K' \left(x + \frac{1}{n}\right) x^{-3/2} \\ &\leq 2 \cdot 3^{-3/2} K K' x^{-1/2}, \end{aligned}$$

which ends the proof for $k = 0$.

Next we treat the $k > 0$ case. Notice first that, since $|C(a, n-1-a) - c(a/n)| \leq 1/n$ and both terms are less than one, we have $|C(a, n-1-a)^k - c(a/n)^k| \leq k/n$. Hence we can write

$$\begin{aligned} \sqrt{n} \sum_{a=1}^{n-1} f(a/n) P(a, n-1-a) C(a, n-a-1)^k \\ = \sqrt{n} \sum_{a=1}^{n-1} f(a/n) P(a, n-1-a) c(a/n)^k \\ + O\left(\frac{1}{\sqrt{n}} \sum_{a=1}^n f(a/n) P(a, n-1-a)\right) \\ = \sqrt{n} \sum_{a=1}^{n-1} f(a/n) P(a, n-1-a) c(a/n)^k \\ + O\left(\frac{1}{\sqrt{n}} \sup\{f(a/n), a \in \{1, \dots, n-1\}\} \sum_{a=1}^{n-1} P(n, n-1-a)\right). \end{aligned}$$

By assumption, the second term tends to 0. We treat the first term by applying the $k = 0$ case to the function $f(\cdot)c(\cdot)^k$, which does satisfy the desired condition that $|f(x)c(x)^k| \leq K(x \wedge (1-x))$ for some $K > 0$, because $|f(x)| \leq \frac{1}{3}K(x \wedge (1-x))^{1-2k}$ for well-chosen K and $c(x) \leq 3(x \wedge (1-x))^2$. $\square$

3. THE LEAF-GROWTH MEASURE ON THE BROWNIAN CRT

In this section we construct the leaf-growth measure $\nu_{\mathcal{T}}$ on the Brownian CRT and prove that (a scalar multiple of) the latter, equipped with the former, is the scaling limit of the discrete model for the Gromov–Hausdorff–Prokhorov metric. In order to establish the scaling limit, we provide an explicit description of the spine towards a typical $\nu_{\mathcal{T}}$ -point.

3.1. PRELIMINARY NOTATION. — The Brownian Continuum Random Tree (CRT) is the random real tree $(\mathcal{T}, d)$ coded by a standard Brownian excursion (length 1, variance 1). This differs by a factor 1/2 compared to Aldous’ original definition. We refer to [LG05] for its main properties and for standard notation about real trees, some of which is recalled below. It is given with a distinguished point ρ called the *root* of the tree. The degree of a point $x \in \mathcal{T}$ is the number of connected components of $\mathcal{T} \setminus \{x\}$. When the degree is equal to 1 we say x is a *leaf*; we call x a *point of the skeleton* when its degree is 2 and a *branch point* when its degree is 3 (there are almost surely no points with degree 4 or more in the Brownian CRT). For $x \in \mathcal{T}$, we call *height* of x its distance to ρ and denote it by $\text{ht}(x)$. We say that x is a *descendant* of another point $y \in \mathcal{T}$ (and equivalently that y is an *ancestor* of x) if y lies on the segment between ρ and x , and we call $\mathcal{T}_x$ the subset of $\mathcal{T}$ formed by x and all its descendants. Finally, for $x \in \mathcal{T}$ and $t \geq 0$ we define for the point $[x]_t$ as:

- the unique ancestor of x with height t , if $t < \text{ht}(x)$
- x itself, if $t \geq \text{ht}(x)$.

In particular, $\mathcal{T}_{[x]_t}$ denotes the subtree above height t which contains the point x . When $[x]_t$ is a branch point and if $t < \text{ht}(x)$, then $\mathcal{T}_{[x]_t} \setminus \{[x]_t\}$ has two components and we denote by $\mathcal{T}_{[x]_{t+}}$ the closure of component that contains x .

The Brownian CRT is also endowed with its uniform mass measure μ (also called the leaf measure), which can either be seen as the projection of the Lebesgue measure on $[0, 1]$ in the coding by the Brownian excursion, or more intrinsically as its natural Hausdorff measure, see [DLG05, DLG06].

In our work, another description of $\mathcal{T}$ will be of key importance. Namely, the Brownian CRT can (with this normalisation) be seen as a fragmentation tree in the sense of Haas and Miermont [HM04] with self-similarity index $\alpha = -1/2$ for the dislocation measure

$$(11) \quad \text{Dis} := \sqrt{\frac{2}{\pi}} \int_{1/2}^1 \frac{dx}{(x(1-x))^{3/2}} \delta_{(x, (1-x))}.$$

Heuristically, this means that it encodes the genealogy of a system of particles starting from a single particle of mass 1, where particles of mass m evolve independently and split into two particles of mass mx and $m(1-x)$ at rate $m^{-1/2} \cdot \sqrt{2/\pi} p(x)$, where $p(x)$ is as in (6). Following Bertoin [Ber02], those particles alive at time t correspond to the subtrees of $\mathcal{T}$ above height t and their masses correspond to the μ -masses of those subtrees.

3.2. DEFINITION OF ν . — Conditionally on the CRT $\mathcal{T}$, the leaf-growth measure $\nu_{\mathcal{T}}$, which throughout this section we will also denote by ν , rendering the dependence on $\mathcal{T}$ implicit, will be defined by specifying the value $\nu(\mathcal{T}_x)$ for all non-leaf $x \in \mathcal{T}$. Proposition 1 of [Ste13] will then guarantee its existence and uniqueness. The definition will be analogous to the discrete case: there, the relative masses given to the subtrees in an (a, b) split were $C(a, b)$ and $C(b, a)$, and here the relative masses given to the subtrees in a $(x, 1-x)$ split will be $c(x)$ and $1 - c(x) = c(1-x)$.

Note that the construction is purely deterministic: for the purposes of Proposition 3.1, one can take $\mathcal{T}$ to be any binary compact rooted real tree equipped with a measure μ supported by the leaves, and c to be any symmetric function satisfying $c(x) \leq x$ for $x \in [0, 1/2]$. Since our only application concerns the Brownian CRT with the specific function $c = x^2(3-2x)$, for the sake of simplicity we shall keep using the notations $\mathcal{T}$, μ and c throughout this section, though they should be interpreted here with this more general framework in mind.

Given $x \in \mathcal{T}$, recall from above that we denote by $\mathcal{T}_{[x]_t}$ and $\mathcal{T}_{[x]_{t+}}$, respectively, the subtree above height t which contains x , and in the case of a branch point, the closure of the connected component of $\mathcal{T} \setminus [x]_t$ which contains x . Finally we set

$$u_x(t) := \frac{\mu(\mathcal{T}_{[x]_{t+}})}{\mu(\mathcal{T}_{[x]_t})} \in [0, 1].$$

In particular, since the tree $\mathcal{T}$ is binary and compact, there are for any $x \in \mathcal{T}$ at most countably many branch points on the ancestral path of x , hence we have $u_x(t) < 1$ only

for a countable number of heights $t \leq \text{ht}(x)$. In particular, if $(t_i : i \in \mathbb{N})$ is an enumeration of the heights of those branch points then the family $(\mu(\mathcal{T}_{[x]_{t_i}}) - \mu(\mathcal{T}_{[x]_{t_i+}}) : i \in \mathbb{N})$ is summable, and since the measure μ does not charge the skeleton, its sum is equal to $1 - \mu(\mathcal{T}_x)$. We deduce then that for all $x \in \mathcal{T}$ we have

$$(12) \quad \mu(\mathcal{T}_x) = \prod_{0 \leq t < \text{ht}(x)} u_x(t).$$

PROPOSITION 3.1 (Definition of ν). — *Suppose that $\mathcal{T}$ is a compact binary rooted tree equipped with a measure μ supported by its leaves. If $\mathbf{c} : [0, 1] \rightarrow [0, 1]$ is symmetric with respect to $1/2$ and satisfies $\mathbf{c}(x) \leq x$ for $x \in [0, 1/2]$, then there exists a unique Borel measure ν supported by the leaves of $\mathcal{T}$ specified by*

$$(13) \quad \nu(\mathcal{T}_x) := \prod_{0 \leq t < \text{ht}(x)} \mathbf{c}(u_x(t)).$$

Proof. — Let, for $x \in \mathcal{T}$, $m(x) = \prod_{0 \leq t < \text{ht}(x)} \mathbf{c}(u_x(t))$. Since we have $\mathbf{c}(x) \leq x$ for $x \in [0, 1/2]$ the product is well-defined and is non-zero as soon as $\prod_{0 \leq t < \text{ht}(x)} u_x(t)$ is non-zero. One can straightforwardly check the following:

- m is a well-defined and nonnegative function, which is equal to 0 exactly on the leaves of $\mathcal{T}$.
- m is non-increasing, in the sense that, if x is an ancestor of y , then $m(x) \geq m(y)$.
- m is “left-continuous”, in the sense that $m(x) = \lim_{s \rightarrow (\text{ht}(x))^-} m([x]_s)$.
- m is “additively right-continuous”: if x is not a branch point, then m is continuous at x , while if x is a branch point, then, letting y and z be two descendants of x not in the same one of the two subtrees originating at x , then $m(x) = \lim_{s \rightarrow (\text{ht}(x))^+} m([y]_s) + m([z]_s)$.

These facts, combined, let us use [Ste13, Prop. 1], which then tells us there exists a unique measure ν such that $\nu(\mathcal{T}_a) = m(a)$ for a , and that it has no atoms. With the notation introduced just before the Proposition, since μ does not charge the skeleton, we can also check that $1 - \nu(\mathcal{T}_x) = \sum_{i \in \mathbb{N}} (\nu(\mathcal{T}_{[x]_{t_i}}) - \nu(\mathcal{T}_{[x]_{t_i+}}))$, in particular ν gives no mass to the ancestral path of x , and as a result is supported by the leaves of $\mathcal{T}$. $\square$

3.3. PATH TOWARDS A TYPICAL ν -LEAF AND SPINE DECOMPOSITION. — Recalling the definition of the Brownian dislocation measure in (11), we now build two measures π_μ and π_ν on $(0, \infty)$ defined by

$$\begin{aligned} \int_0^\infty d\pi_\mu(s) \phi(s) &:= \sqrt{\frac{2}{\pi}} \int_{1/2}^1 (\mathbf{c}(x)\phi(-\log(x)) + \mathbf{c}(1-x)\phi(-\log(1-x))) \mathbf{p}(x) dx \\ &= \sqrt{\frac{2}{\pi}} \int_0^1 \mathbf{c}(x)\phi(-\log(x)) \mathbf{p}(x) dx \end{aligned}$$

and

$$\begin{aligned} \int_0^\infty d\pi_\nu(s) \phi(s) &:= \sqrt{\frac{2}{\pi}} \int_{1/2}^1 (c(x)\phi(-\log(c(x))) + c(1-x)\phi(-\log(c(1-x)))) p(x) dx \\ &= \sqrt{\frac{2}{\pi}} \int_0^1 c(x)\phi(-\log(c(x))) p(x) dx. \end{aligned}$$

Consider then Π_μ a Poisson point process (P.p.p.) on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $d\pi_\mu \otimes dt$. By replacing each atom (p, t) of Π_μ by the atom $(-\log(c(e^{-p})), t)$ we obtain another Poisson point process with intensity $d\pi_\nu \otimes dt$. We can then define two coupled (non-decreasing) subordinators ξ_μ and ξ_ν started from 0, and with respective Lévy measures π_μ and π_ν , and obtained by summing the atoms of the respective P.p.p., i.e.,

$$(14) \quad \xi_\mu(t) = \sum_{\substack{(p,s) \in \Pi_\mu \\ s \leq t}} p \quad \text{and} \quad \xi_\nu(t) = \sum_{\substack{(p,s) \in \Pi_\nu \\ s \leq t}} p = \sum_{\substack{(p,s) \in \Pi_\mu \\ s \leq t}} -\log(c(e^{-p})).$$

Then we perform the Lamperti transformation with index $1/2$ using ξ_μ : we let

$$(15) \quad \theta(t) = \inf \left\{ u \geq 0 : \int_0^u e^{-\frac{1}{2}\xi_\mu(r)} dr > t \right\},$$

and build two processes X_μ (a so-called positive self-similar Markov process, see [Lam62]) and X_ν (which is not Markovian, since the time change uses information from X_μ) defined for $t \geq 0$ by

$$(16) \quad (X_\mu(t), X_\nu(t)) := (e^{-\xi_\mu(\theta(t))}, e^{-\xi_\nu(\theta(t))}).$$

Both of those processes are pure jump and non-increasing from 1 to 0, and are absorbed at 0 at time

$$(17) \quad I := \int_0^\infty e^{-\frac{1}{2}\xi_\mu(t)} dt$$

which is almost surely finite since ξ_μ drifts to $+\infty$ at linear speed.

THEOREM 3.2 (Spine decomposition). — *Conditionally on the Brownian CRT $\mathcal{T}$, let L be a ν -distributed random point of $\mathcal{T}$. Then the bivariate process*

$$(\mu(\mathcal{T}_{[L]_{t+}}), \nu(\mathcal{T}_{[L]_{t+}}))_{t \geq 0}$$

(under the annealed law) is distributed as (X_μ, X_ν) . As a consequence, L is almost surely a leaf, and its height has the law of I .

Moreover, let $(t_i, i \in \mathbb{N})$ be the times at which $u_L(\cdot) \in (0, 1)$, and for all $i \in \mathbb{N}$ let $\mathcal{T}_i$ be the subtrees which branch off $[[\rho, L]]$ at those times (ranked, say in decreasing order of their μ -masses). Then, conditionally on the whole process $(\mu(\mathcal{T}_{[L]_t}), t \geq 0)$, the trees

$$(u_L(t_i) \cdot \mu(\mathcal{T}_{L(t_i)}))^{-1/2} \cdot \mathcal{T}_i, \quad i \geq 1,$$

are independent Brownian CRT's.

Proof. — Informally, the idea is that the path to L encounters binary splits at the same rate as the path to a uniform leaf, but crosses those splits differently, using the probability c . This is formalised in [PW15] by the notion of *bifurcator*. If L' is a leaf with distribution μ , then the paths from ρ to L and to L' will encounter countable numbers of binary branch points. At any common such branch point, let x and $1-x$ be the relative μ -masses of both pending subtrees, choosing x to be the mass of the one containing L . Then the probability that L' is in the other subtree, known as the “switching probability” in [PW15], is equal to $c(1-x)$. While the process $(Y(t), t \geq 0)$ defined by $Y(t) = \mu(\mathcal{T}_{[L']_{t+}})$ is well-known [Ber02, §4] to be the Lamperti transform of a subordinator with Lévy measure π_{unif} , defined by

$$\begin{aligned} \int_0^\infty d\pi_{\text{unif}}(s) \phi(s) &= \sqrt{\frac{2}{\pi}} \int_{1/2}^1 (x\phi(-\log(x)) + (1-x)\phi(-\log(1-x))) p(x) dx \\ &= \sqrt{\frac{2}{\pi}} \int_0^1 x\phi(-\log(x)) p(x) dx, \end{aligned}$$

Proposition 1 from [PW15] states then that the process $\mu(\mathcal{T}_{[L]_{t+}}), t \geq 0$ will be the Lamperti transform of a subordinator where the initial x in the integrand of the above is replaced by $c(x)$, giving us π_μ .

That $(\nu(\mathcal{T}_{[L]_{t+}}), t \geq 0)$ has the appropriate paired distribution comes from the fact that its multiplicative jumps are obtained by applying c to those of $(\mu(\mathcal{T}_{[L]_{t+}}), t \geq 0)$, which is how X_ν is derived from X_μ .

Finally, the spinal decomposition itself is again a consequence of [PW15], this time Lemma 21. $\square$

3.4. APPLICATIONS: HEIGHT AND DIMENSION. — Let us see two direct applications of Theorem 3.2. In the first one we compute the (annealed) law of the height of a random ν -leaf, and in the second one we find the almost sure Hausdorff dimension of ν .

PROPOSITION 3.3. — *Let $\mathcal{T}$ be a Brownian CRT, and conditionally on $\mathcal{T}$, let $L \in \mathcal{T}$ be sampled according to ν . Then $\text{ht}(L)$ is a continuous random variable on $\mathbb{R}_+$, and its density is $8x^3e^{-2x^2}$.*

REMARK 3.4. — An integration by parts shows that this dominates the Rayleigh distribution $(4xe^{-2x^2}dx)$, which is the distribution of the height of a uniform leaf) stochastically. This is expected, as the path to L favours large subtrees more than the path to a uniform leaf. However, we have no natural explanation for the surprisingly simple form of this density. As one referee pointed out, this density is that of $\sqrt{G_2}/2$, where G_2 is a standard $\Gamma(2)$ random variable, which is also the density of the total length of the subtree spanned by three uniform independent leaves of $\mathcal{T}$.

Proof. — We know that $\text{ht}(L)$ has the same distribution as $I = \int_0^\infty e^{-\frac{1}{2}\xi_\mu(t)} dt$. The results of [CPY97] show that the integer moments of this distribution can be linked

to the Laplace exponent of ξ_μ , which we call Φ , by the following formula:

$$\mathbb{E}[(\text{ht}(L))^k] = \frac{k!}{\prod_{i=1}^k \Phi(i/2)}.$$

However, in our setting, we have for $\alpha \geq 0$

$$\begin{aligned} \Phi(\alpha) &= \sqrt{\frac{2}{\pi}} \int_{1/2}^1 (\mathbf{c}(x)(1-x^\alpha) + \mathbf{c}(1-x)(1-(1-x))^\alpha) \mathbf{p}(x) dx \\ &= \sqrt{\frac{2}{\pi}} \int_{1/2}^1 (1 - \mathbf{c}(x)x^\alpha - \mathbf{c}(1-x)(1-x)^\alpha) \mathbf{p}(x) dx \\ (18) \quad &= \frac{1}{\sqrt{2\pi}} \int_0^1 (1 - \mathbf{c}(x)x^\alpha - \mathbf{c}(1-x)(1-x)^\alpha) \mathbf{p}(x) dx. \end{aligned}$$

The latter integral can be computed (see Appendix A), and we obtain $\Phi(\alpha) = 2\sqrt{2}\alpha\Gamma(3/2 + \alpha)/\Gamma(2 + \alpha)$. Hence

$$\mathbb{E}[(\text{ht}(L))^k] = \frac{\Gamma((4+k)/2)}{2^{k/2}}.$$

A straightforward induction shows that $\Gamma((4+k)/2)/2^{k/2} = \int_0^\infty 8x^{3+k}e^{-2x^2} dx$. Given that the probability distribution with density $8x^3e^{-2x^2}$ is determined by its moments (since its moment generating function has positive radius of convergence), this ends the proof. $\square$

Hausdorff dimension. — The fact that the Hausdorff dimension of the measure ν is 2γ follows from the coming proposition together with standard results about Hausdorff dimension (see e.g. [LPP95, Lem. 4.1] and [Bil65, §14]):

PROPOSITION 3.5. — *Almost surely for $\mathcal{T}$ and a.e. for $\nu(d\ell)$ we have*

$$\lim_{r \rightarrow 0} \frac{\log \nu(B_r(\ell))}{\log r} = 2\gamma.$$

Proof. — Recall that $\gamma = 3(2 - \sqrt{3})$, and the definitions of $\mathbf{p}(x)$ in (6), $\mathbf{c}(x)$ in (7), ξ_μ and ξ_ν in (14), X_μ and X_ν in (16), the time change θ in (15) and extinction time I in (17). By Theorem 3.2, we can assume that the bivariate process $(\mu(\mathcal{T}_{[L]_{t+}}), \nu(\mathcal{T}_{[L]_{t+}}))_{t \geq 0}$ is given by (X_μ, X_ν) .

The local rate of growth of ν near a ν -typical point is related to the asymptotic behaviour of X_μ and X_ν near their extinction time. To understand this, let us compute the first moments of the Lévy processes ξ_μ and ξ_ν : we have

$$\mathbb{E}[\xi_\mu(t)] = t \int_0^\infty s \, d\pi_\mu(s) = \sqrt{2\pi} \quad \text{and} \quad \mathbb{E}[\xi_\nu(t)] = t \int_0^\infty s \, d\pi_\nu(s) = \gamma\sqrt{2\pi}.$$

Since the Lévy measure integrates the identity function, by the law of large numbers we have $t^{-1}\xi_\mu(t) \rightarrow \sqrt{2\pi}$ and $t^{-1}\xi_\nu(t) \rightarrow \gamma\sqrt{2\pi}$ as $t \rightarrow \infty$ almost surely. Through the

Lamperti transformation (16), this turns into estimate on the time change θ since for $\varepsilon > 0$,

$$\int_{\theta(I-\varepsilon)}^{\infty} e^{-\frac{1}{2}\xi_{\mu}(t)} dt = \varepsilon,$$

from which we straightforwardly deduce using the above law of large numbers that

$$\theta(I - \varepsilon) \underset{\varepsilon \rightarrow 0}{\sim} -\frac{2}{\sqrt{2\pi}} \log \varepsilon$$

almost surely. Substituting back into (16), we find using the same law of large numbers that

$$\lim_{\varepsilon \rightarrow 0} \frac{\log X_{\mu}(I - \varepsilon)}{\log \varepsilon} = 2 \quad \text{and} \quad \lim_{\varepsilon \rightarrow 0} \frac{\log X_{\nu}(I - \varepsilon)}{\log \varepsilon} = 2\gamma,$$

also almost surely.

Now, by the first item of the spine decomposition (Theorem 3.2), we clearly have $\nu(B_{\varepsilon}(L)) \leq X_{\nu}(I - \varepsilon)$ as a process in ε and together with the last display this entails that

$$\liminf_{\varepsilon \rightarrow 0} \frac{\log \nu(B_{\varepsilon}(L))}{\log \varepsilon} \geq 2\gamma, \quad \text{a.s..}$$

Let us now prove the upper bound for the dimension via a lower bound on $\nu(B_{\varepsilon}(L))$. For this, we concentrate on jumps of ξ_{μ} of size at least $\log 2$. Remember that thanks to the spine decomposition, conditionally on such a jump of size $s \geq \log 2$ appearing at time t , the subtree $\mathcal{T}_{[L]_u} \setminus \mathcal{T}_{[L]_{u+}} \cup \{[L]_u\}$ with $t = \theta(u)$ (which is the subtree which is “left on the side” at height u) is an independent Brownian CRT with μ -mass

$$(1 - e^{-s})e^{-\xi_{\mu}(t-)} \geq \frac{1}{2} e^{-\xi_{\mu}(t-)}.$$

Notice two things: first, the ν -mass of this subtree is at least $e^{-\xi_{\nu}(t-)} / 2$ (because $c(x) \leq x$ for $x \in [0, 1/2]$) and second, this tree has a fixed probability equal to $\mathbb{P}(\text{ht}(\mathcal{T}) \leq 1)$ of being of height less than the square root of its μ -mass. Under those conditions, this subtree is a subset of $B_{\varepsilon}(L)$ for $\varepsilon \geq e^{-\frac{1}{2}\xi_{\mu}(t-)} + \int_t^{\infty} e^{-\frac{1}{2}\xi_{\mu}(s)} ds$ and we have

$$(19) \quad \nu(B_{\varepsilon}(L)) \geq \frac{1}{2} e^{-\xi_{\nu}(t-)}.$$

By the law of large numbers on ξ_{μ} and ξ_{ν} , we can take $\varepsilon = c_1 e^{-t\sqrt{2\pi}/2}$ and we have $\nu(B_{\varepsilon}(L)) \geq c_2 e^{-t\sqrt{2\pi}\gamma/2} = \varepsilon^{2\gamma+o(1)}$ for some $c_1, c_2 > 0$. The preceding inequality is only valid for those ε associated to a jump time t satisfying the above conditions. However, if we can find a family $t_i \rightarrow \infty$ of such times satisfying

$$(20) \quad 0 < t_1 < t_2 < \dots < t_i \xrightarrow{i \rightarrow \infty} \infty \quad \text{and} \quad \lim_{t \rightarrow \infty} \frac{t_{i+1}}{t_i} = 1,$$

then the bound $\nu(B_{\varepsilon}(L)) \geq \varepsilon^{2\gamma+o(1)}$ still holds as $\varepsilon \rightarrow 0$ and proves the lower bound on the Hausdorff dimension. To see (20), notice that in the time parametrisation of ξ_{ν} , the intensity $C > 0$ of such times is positive: by a standard law of large numbers we can almost surely find a diverging sequence of jump times t_i satisfying all of the above assumptions and which satisfy $t_i \sim Ci$. This entails (20) and completes the proof. $\square$

3.5. SCALING LIMITS. — The aim of this section is to prove Theorem 1.3. For the reader's convenience, we quickly recall the definition of the (pointed) Gromov–Hausdorff–Prokhorov convergence and refer to [ADH13, Khe20] for more details, as well as [ALW16] for a presentation of the topology without the metric. A weighted pointed compact metric space (M, d, ρ, μ) is a compact metric space equipped with a finite Borel measure μ and a distinguished point $\rho \in M$. We let $\mathbb{M}$ be the set of all isometry classes of weighted pointed compact metric spaces. To lighten notation, we shall often identify a compact weighted metric space with its equivalence class. We equip $\mathbb{M}$ with the classical Gromov–Hausdorff–Prokhorov metric, defined for every $\mathbf{M} := (M, d, \rho, \mu)$ and $\mathbf{M}' := (M', d', \rho', \mu')$ in $\mathbb{M}$ by:

$$d_{\text{GHP}}(\mathbf{M}, \mathbf{M}') := \inf_{\phi, \phi'} \left(\delta_{\text{H}}(\phi(M), \phi'(M')) \vee \delta_{\text{P}}(\phi_*\mu, \phi'_*\mu') \vee \delta(\phi(\rho), \phi'(\rho')) \right),$$

where the infimum is taken over all isometries ϕ, ϕ' from M, M' into a metric space (Z, δ) and δ_{H} (resp. δ_{P}) stands for the classical Hausdorff distance (resp. the Prokhorov distance) in Z . The space $(\mathbb{M}, d_{\text{GHP}})$ is then a Polish space.

Our starting point is the well-known convergence in distribution from [Ald93]:

$$(21) \quad \left(\frac{1}{\sqrt{n}} \tau_n, \rho_n, \mu_n \right) \xrightarrow[n \rightarrow \infty]{(d)} (2\sqrt{2} \mathcal{T}, \rho, \mu),$$

for the pointed Gromov–Hausdorff–Prokhorov metric, where μ_n is the uniform measure on the vertices of τ_n . By analogy with this notation, we shall also use the notation ν_n for the measure ν_{τ_n} introduced in Section 2. We saw in Section 2 and in Proposition 3.1 that the measures ν_n and ν are obtained as measurable functions of $((1/\sqrt{n})\tau_n, \rho_n, \mu_n)$ and $(2\sqrt{2}\mathcal{T}, \rho, \mu)$ respectively. Furthermore, both measures are constructed similarly by modifying the splitting of the reference measures μ_n and μ using the functions $C(a, b)$ and c . In view of (21) and the convergence (7), it is natural to expect that those functions are almost “continuous”, letting us adjoin ν_n and ν to (21) using deterministic arguments.

Proof of Theorem 1.3. — Using the Skorokhod representation theorem, we can assume that (21) holds almost surely, and the proof will proceed deterministically. Next, by a general representation theorem for Gromov–Hausdorff-type topologies [Khe23, Lem. 2.5], we can embed all trees τ_n and $\mathcal{T}$ isometrically in the same compact metric space (E, d_E) —we keep the same notation for simplicity—, in such a way that

$$\left(\frac{1}{\sqrt{n}} \tau_n, \rho_n, \mu_n \right) \xrightarrow[n \rightarrow \infty]{} (2\sqrt{2} \mathcal{T}, \rho, \mu),$$

holds in pointed Hausdorff–Prokhorov sense inside E . Fix then any point x in the skeleton of $\mathcal{T}$ which is not a branch point. Lemma B.1 shows that there exists a sequence $(x_n, n \in \mathbb{N})$ with $x_n \in \tau_n$ such that $x_n \rightarrow x$ in E and $\mu_n((\tau_n)_{x_n}) \rightarrow \mu(\mathcal{T}_x)$. We will now show that we also have

$$(22) \quad \lim_{n \rightarrow \infty} \nu_n((\tau_n)_{x_n}) = \nu(\mathcal{T}_x),$$

which will guarantee via the Portmanteau theorem and [Ste13, Prop. 1], that any subsequential limit of (ν_n) is equal to ν , ending the proof of Theorem 1.3 since $\mathcal{T}$ is compact.

Let us denote by h_n the height of x_n and by $(a_i^{(n)}, b_i^{(n)})$ the sizes of the two subtrees above the ancestor of x_n at height $0 \leq i < h_n$, with the convention that $a_i^{(n)}$ is the size of the subtree containing x_n . It is then straightforward, but rather long, to prove that we also have the convergence

$$(23) \quad \left(\left(\frac{a_i^{(n)}}{a_i^{(n)} + b_i^{(n)}}, \frac{i}{2\sqrt{2n}} \right) : 0 \leq i < h_n \right) \xrightarrow{n \rightarrow \infty} \left(\left(\frac{\mu(\mathcal{T}_{[x]_{t+}})}{\mu(\mathcal{T}_{[x]_t})}, t \right) : 0 \leq t < \text{ht}(x) \right)$$

in terms of point processes on $\mathbb{R}_+^* \times \mathbb{R}_+$. In particular, recall that (12) holds almost surely. Recall first from the construction of ν_n and μ_n that we have

$$(24) \quad \nu_n((\tau_n)_{x_n}) = \prod_{i=0}^{h_n-1} C(a_i^{(n)}, b_i^{(n)}) \quad \text{and} \quad \mu_n((\tau_n)_{x_n}) = \prod_{i=0}^{h_n-1} \frac{a_i^{(n)}}{a_i^{(n)} + b_i^{(n)} + 1}.$$

On the continuous side, if we set $a_t = \mu(\mathcal{T}_{[x]_{t+}})$ and $b_t = \mu(\mathcal{T}_{[x]_t}) - \mu(\mathcal{T}_{[x]_{t+}})$ for $0 \leq t \leq \text{ht}(x)$ (note that $b_t = 0$ when $[x]_t$ is not a branch point), then by (12) and Proposition 3.1 we have

$$(25) \quad \nu(x) = \prod_{t < \text{ht}(x)} c(a_t/(a_t + b_t)) \quad \text{and} \quad \mu(x) = \prod_{t < \text{ht}(x)} \frac{a_t}{a_t + b_t}.$$

On the one hand, the convergences (23), (7) and Fatou's lemma directly imply that

$$\limsup_{n \rightarrow \infty} \nu_n((\tau_n)_{x_n}) = \limsup_{n \rightarrow \infty} \prod_{i=0}^{h_n-1} C(a_i^{(n)}, b_i^{(n)}) \leq \prod_{t < \text{ht}(x)} c(a_t/(a_t + b_t)) = \nu(x).$$

To get the matching lower bound $\nu(x) \leq \liminf_{n \rightarrow \infty} \nu_n((\tau_n)_{x_n})$, notice that the same argument yields

$$\limsup_{n \rightarrow \infty} \mu_n((\tau_n)_{x_n}) = \prod_{i=0}^{h_n-1} \frac{a_i^{(n)}}{a_i^{(n)} + b_i^{(n)} + 1} \leq \mu(x) = \prod_{t < \text{ht}(x)} \frac{a_t}{a_t + b_t};$$

here, however, we know that the matching lower bound holds since

$$\lim_{n \rightarrow \infty} \mu_n((\tau_n)_{x_n}) = \mu(\mathcal{T}_x)$$

is known. Using the fact that $C(a, b) \geq a/(a + b + 1)$ for $a \geq b$, we deduce that

$$1 \geq \lim_{\varepsilon \rightarrow 0} \prod_{i=0}^{h_n-1} C(a_i^{(n)}, b_i^{(n)}) \mathbf{1}_{b_i^{(n)} \leq \varepsilon a_i^{(n)}} \geq \lim_{\varepsilon \rightarrow 0} \prod_{i=0}^{h_n-1} \frac{a_i^{(n)}}{a_i^{(n)} + b_i^{(n)} + 1} \mathbf{1}_{b_i^{(n)} \leq \varepsilon a_i^{(n)}} = 1,$$

from which (22) follows. $\square$

4. DISCRETE FRACTAL PROPERTIES OF LEAF-GROWTH MEASURES

In this section, we study the discrete fractal properties of ν_{τ_n} and focus in particular on Theorem 1.2 and Theorem 1.1. Those results will be proved by investigating the asymptotic behaviour of the random variable $M_n = \nu_{\tau_n}(L_n)$, where τ_n is uniform in $\mathbb{T}_n$ and L_n is, conditionally on τ_n , a random leaf of τ_n sampled according to the leaf-growth measure ν_{τ_n} . This notation will be used throughout the section.

4.1. TYPICAL EXPONENT OF THE LEAF-GROWTH MASS OF A RANDOM LEAF. — Theorem 1.1 is a straightforward consequence of the following:

PROPOSITION 4.1. — *Setting $\gamma = 3(2 - \sqrt{3})$, we have the following convergence in probability:*

$$\frac{-\log M_n}{\log n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \gamma.$$

Proof. — For $k \in \mathbb{N}$, let L_n be a ν_{τ_n} -distributed random leaf of τ_n . Define two processes $X_\mu^{(n)}(\cdot)$ and $X_\nu^{(n)}(\cdot)$ as follows:

- For $k \leq \text{ht}(L_n)$, let x be the ancestor of V_n in τ_n with height k . Define $X_\mu^{(n)}(k)$ to be μ_n -mass of the subtree of τ_n rooted at x , and $X_\nu^{(n)}(k)$ to be the ν_{τ_n} -mass of that subtree.
- For $k > \text{ht}(L_n)$, let $X_\mu^{(n)}(k) = X_\nu^{(n)}(k) = 0$.

It is possible to show that the scaling limit of $X_\mu^{(n)}$ is X_μ , however here we are interested in the large-time behaviour of $X_\mu^{(n)}$ and $X_\nu^{(n)}$. Specifically, note then that $\text{ht}(L_n)$ is the time at which the process $X_\mu^{(n)}$ hits the value $1/(2n+1)$, and that $M_n = X_\nu^{(n)}(\text{ht}(L_n))$.

Setup. — The proof of the proposition will be mainly based on showing weak LLN-type behaviour for both processes $X_\mu^{(n)}$ and $X_\nu^{(n)}$ at large times. Taking inspiration from [BK16], we first switch to continuous time by using a “Poissonization” technique: let $(\mathcal{P}_n(t), t \geq 0)$ be a Poisson-like process which starts at 0 and increases by 1 periodically, the waiting time for the k -th jump being exponentially distributed with rate parameter $\sqrt{nX_\mu^{(n)}(k)}$. Now set, for all $t \geq 0$,

$$\mathcal{X}_\mu^{(n)}(t) = -\log X_\mu^{(n)}(\mathcal{P}_n(t)).$$

This forms a Markov process: when at position $\log(2n+1) - \log(2m+1)$ for $m \geq 1$, it waits an amount of time which is exponentially distributed with parameter $m^{1/2}$ and then jumps to $\log(2n+1) - \log(2a+1)$, for $a \in \{0, \dots, m-1\}$, with probability $2P(a, m-1-a)C(a, m-1-a)$. Notice that $\mathcal{X}_\mu^{(n)}$ grows from 0 to $\log(2n+1)$, since $X_\mu^{(n)}$ decreased from 1 to $1/(2n+1)$. Similarly, for all $t \geq 0$, set $\mathcal{X}_\nu^{(n)}(t) = -\log X_\nu^{(n)}(\mathcal{P}_n(t))$.

Letting $\theta_n = \inf\{t : \mathcal{X}_\mu^{(n)}(t) = \log(2n+1)\}$, our aim is to show that

$$(26) \quad \frac{\mathcal{X}_\nu^{(n)}(\theta_n)}{\log n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \gamma$$

To do this, we will first consider $\theta'_n = \inf\{t : \mathcal{X}_\mu^{(n)}(t) = \log(2n+1) - K\}$ for a large K that will be chosen later. Notice that, for any value of K , the distribution of $\theta_n - \theta'_n$ is tight as n tends to infinity, because it is a mixture of the distributions of (θ_i) for i in the finite set $\{1, \dots, \lfloor e^K \rfloor\}$. Similarly, the distribution of $\mathcal{X}_\nu^{(n)}(\theta_n) - \mathcal{X}_\nu^{(n)}(\theta'_n)$ is also tight, as it is a mixture of the distributions of $\mathcal{X}^{(i)}(\theta_i)$ for $i \in \{1, \dots, \lfloor e^K \rfloor\}$.

This implies that we only need to prove Equation (26) with θ'_n replacing θ_n :

$$(27) \quad \frac{\mathcal{X}_\nu^{(n)}(\theta'_n)}{\log n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \gamma.$$

We set out to prove (27).

Analysis of $\mathcal{X}_\mu^{(n)}$ and estimate of θ'_n . — Here we show that $\mathcal{X}_\mu^{(n)}(t)$ is mostly linear in t , and deduce that θ'_n is well approximated by $(2 \log n)/\sqrt{\pi}$.

Since $\mathcal{X}_\mu^{(n)}$ is a pure jump process, we can compute its predictable compensator. It is given by integrating the local drift $D_\mu^{(n)}(t) = \frac{1}{dt} \mathbb{E}[\mathcal{X}_\mu^{(n)}(t+dt) - \mathcal{X}_\mu^{(n)}(t) \mid \mathcal{F}_t]$ which is equal to the rate of jumps, times the expectation of the next jump, that is

$$D_\mu^{(n)}(t) = m^{1/2} \sum_{a=0}^{m-1} -\log\left(\frac{2a+1}{2m+1}\right) 2P(a, m-1-a)C(a, m-1-a),$$

with m such that $\mathcal{X}_\mu^{(n)}(t) = \log n - \log m$. We can apply Proposition 2.7 to this sum (since $\log(1/m) = o(\sqrt{m})$ and $x \log(x)$ tends to 0 as x tends to 0), obtaining that it converges as m tends to infinity to

$$\frac{1}{2\sqrt{\pi}} \int_0^1 -\log x \, c(x) p(x) dx = \frac{\sqrt{\pi}}{2}.$$

The fact that this is a limit as m tends to infinity is useful because the drift will be close to this value at any t , as long as m is big (i.e., as long as $\mathcal{X}_n(t) < \log(n) - K$ for large K). By definition of the predictable compensator, the process $(M_\mu^{(n)}(t), t \geq 0)$ defined by

$$M_\mu^{(n)}(t) = \mathcal{X}_\mu^{(n)}(t) - \int_0^t D_\mu(t) dt,$$

is a $(\mathcal{F}_t : t \geq 0)$ -martingale. Its quadratic variation is given by integrating the rate of jumps, times the expectation of the square of the next jump, which is

$$V_\mu^{(n)}(t) = m^{1/2} \sum_{a=0}^{m-1} \left(\log\left(\frac{2a+1}{2m+1}\right) \right)^2 2P(a, m-1-a)C(a, m-1-a),$$

where m is still defined by $\mathcal{X}_\mu^{(n)}(t) = \log n - \log m$. Again by Proposition 2.7, this converges as $m \rightarrow \infty$ to

$$\frac{1}{2\sqrt{\pi}} \int_0^1 (\log x)^2 c(x) p(x) dx.$$

In particular, this means that $B := \sup_{n,t} V_\mu^{(n)}(t) < \infty$.

Since $M_\mu^{(n)}(0) = 0$, we have

$$\mathbb{E}[M_\mu^{(n)}(t)^2] = \mathbb{E}\left[\int_0^t V_\mu(t) dt\right] \leq Bt,$$

implying via Doob's inequality that

$$\mathbb{P}\left[\sup_{u \leq t} |M_\mu^{(n)}(u)| > t^{0.6}\right] \leq \frac{Bt}{t^{1.2}} \xrightarrow[t \rightarrow \infty]{} 0,$$

uniformly in n .

Let now $\varepsilon > 0$, and choose K large enough to have $|D_\mu(t) - \sqrt{\pi}/2| \leq \varepsilon/2$ whenever $\mathcal{X}_\mu^{(n)}(t) \leq \log n - K$. Consider $t_n = (2(1 - \varepsilon) \log n)/\sqrt{\pi}$ and let E_n be the event that $\sup_{t \leq t_n} |M_\mu^{(n)}(t)| \leq t_n^{0.6}$. We claim that, for n large enough, on E_n , we have $\theta'_n \geq t_n$. To see this, notice that, on E_n , we have for all $t \leq t_n$

$$\mathcal{X}_\mu^{(n)}(t) = \int_0^t D_\mu^{(n)}(s) ds + O(\log(n)^{0.6})$$

However, if $\theta'_n \leq t_n$, then for n large enough we would have

$$\begin{aligned} \log(n) - K &\leq \mathcal{X}_\mu^{(n)}(\theta'_n) \\ &\leq \frac{\sqrt{\pi}}{2}(1 + \varepsilon/2)\theta'_n + O(\log(n)^{0.6}) \\ &\leq (1 + \varepsilon/2)(1 - \varepsilon) \log n + O(\log(n)^{0.6}), \end{aligned}$$

which eventually gives a contradiction. Put together, this means that

$$\mathbb{P}[\theta'_n \geq (1 - \varepsilon) \log n / A] \geq \mathbb{P}[E_n] \rightarrow 1.$$

Set now $u_n = (2(1 + \varepsilon) \log n)/\sqrt{\pi}$, and this time let F_n be the event that $\sup_{t \leq u_n} |M_\mu^{(n)}(t)| \leq u_n^{0.6}$, whose probability also tends to 1. Let us now check that, on F_n , $\theta'_n \leq u_n$. Similarly to before, observe that, on F_n , we have for all $t \leq u_n$ that

$$\mathcal{X}_\mu^{(n)}(t) = \int_0^t D_\mu^{(n)}(s) ds + O(\log(n)^{0.6}).$$

Thus, if we had $\mathcal{X}_\mu^{(n)}(u_n) < \log(n) - K$, then we would have $V_\mu^{(n)}(s) \geq A(1 - \varepsilon/2)$ for all $s \leq u_n$, and in particular

$$\log n \geq \mathcal{X}_\mu^{(n)}(u_n) \geq \frac{\sqrt{\pi}}{2}(1 - \varepsilon/2) \frac{2(1 + \varepsilon) \log n}{\sqrt{\pi}} + O((\log n)^{0.6}),$$

a contradiction for n large enough. Hence $\mathbb{P}[\theta'_n \leq (2(1 + \varepsilon) \log n)/\sqrt{\pi}] \geq \mathbb{P}[F_n] \rightarrow 1$.

Analysis of $\mathcal{X}_\nu^{(n)}$ and conclusion. — We can use the same method to establish the first order linear behaviour of $\mathcal{X}_\nu^{(n)}$. Keeping the notation $\mathcal{X}_\mu^{(n)}(t) = \log(2n+1) - \log(2m+1)$, we have that $\mathcal{X}_\nu^{(n)}$ jumps at rate $m^{1/2}$, and if $\mathcal{X}_\mu^{(n)}$ jumps to $\log(2n+1) - \log(2a+1)$, then the value of $\mathcal{X}_\nu^{(n)}$ increases by $\log(C(a, b))$. Hence

$$D_\nu(t) = m^{1/2} \sum_{a=0}^{m-1} \log(C(a, b)) 2P(a, b) C(a, b),$$

the limit of which is

$$\frac{1}{2\sqrt{\pi}} \int_0^1 c(x)^2 p(x) dx = \frac{\sqrt{\pi}}{2} \gamma.$$

Similarly, the local variance is then

$$V_\nu^{(n)}(t) = m^{1/2} \sum_{a=0}^{m-1} \log(C(a, b))^2 2P(a, b) C(a, b),$$

the limit of which is

$$\frac{1}{2\sqrt{\pi}} \int_0^1 \log(c(x))^2 c(x) p(x) dx,$$

hence we have $V_\nu(t) \leq B$ for all t , up to changing B .

We consider the martingale $(M_\nu^{(n)}(t), t \geq 0)$ defined by

$$M_\nu^{(n)}(t) = \mathcal{X}_\mu^{(n)}(t) - \int_0^t D_\nu(t) dt.$$

The same arguments as for $M_\nu^{(n)}$ show that, if G_n is the event that $\sup_{t \leq u_n} |M_\nu^{(n)}(t)| \leq u_n^{0.6}$, then $\mathbb{P}[G_n]$ tends to 1. Up to increasing K , we can moreover assume that $|D_\nu(t) - \frac{\sqrt{\pi}}{2}\gamma| \leq \varepsilon$ for $t \leq \theta'_n$. Then on the event $G_n \cap \{|\theta'_n - \frac{2}{\sqrt{\pi}} \log(n)| \leq \frac{2}{\sqrt{\pi}} \varepsilon \log(n)\}$, which has probability tending to 1, we have

$$|\mathcal{X}_\nu^{(n)}(\theta'_n) - \frac{\sqrt{\pi}}{2}\gamma\theta'_n| \leq \frac{\varepsilon}{2}\theta'_n + O((\log n)^{0.6}),$$

from which (27) readily follows. $\square$

4.2. MOMENTS OF THE LEAF-GROWTH MASS OF A RANDOM LEAF: THE MULTIFRACTAL SPECTRUM. — In contrast to the result of the previous section, we now highlight the multifractality phenomenon governing the behaviour of the random variable M_n and proceed to compute the full discrete multifractal spectrum of the measure ν_{τ_n} .

Our aim within this section will be to estimate all moments $\mathbb{E}[M_n^\alpha]$ for $\alpha \in \mathbb{R}$; by the end, we shall prove Theorem 1.2. Before we do, let us give a brief insight into the appearance of the integral $I(\alpha, \beta)$ from (2); we shall then show that, given $\alpha \in \mathbb{R}$, there is a unique $\beta(\alpha)$ such that $I(\alpha, \beta(\alpha)) = 0$. Finally, we dive into a formal proof of Theorem 1.2.

In order to compute $\mathbb{E}[M_n^\alpha]$, we can disintegrate according to the profile of τ_n ; indeed, letting τ_n^L and τ_n^R be the left and right subtrees, conditionally on $|\tau_n^L| = a$, the leaf L_n is distributed according to $\nu_{\tau_n^L}$ among the $a+1$ leaves of τ_n^L with probability $C(a, b)$ (where $b = n - 1 - a$), and distributed according to $\nu_{\tau_n^R}$ among the b leaves of τ_n^R with probability $1 - C(a, b)$. In the first case, the mass M_n has the distribution of $C(a, b)M_a$, and in the second case it is distributed as $C(b, a)M_b$. This yields

$$\mathbb{E}[M_n^\alpha] = \sum_{a=1}^{n-1} P(a, b) (C(a, b)^{\alpha+1} \mathbb{E}[M_a^\alpha] + C(b, a)^{\alpha+1} \mathbb{E}[M_b^\alpha]).$$

Since we have $\sum_{a=1}^{n-1} P(a, b) = 1$, we can rewrite this as

$$\sum_{a=0}^{n-1} P(a, b) (C(a, b)^{\alpha+1} \mathbb{E}[M_a^\alpha] + C(b, a)^{\alpha+1} \mathbb{E}[M_b^\alpha] - \mathbb{E}[M_n^\alpha]) = 0.$$

Now, if we informally assume an asymptotic of the form $\mathbb{E}[M_n^\alpha] \approx n^{-\beta}$, where β is some positive real number, and recall that $C(a, b)$ is close to $c(a/n)$, we obtain that

$$\sum_{a=1}^{n-2} P(a, b) \left(c(a/n)^{\alpha+1} (a/n)^{-\beta} + (1 - c(a/n))^{\alpha+1} (b/n)^{-\beta} - 1 \right)$$

should be close to 0. However, similarly to Proposition 2.7, we expect to this sum converge to $(1/4\sqrt{\pi})I(\alpha, \beta)$, hence if $\mathbb{E}[M_n^\alpha] \approx n^{-\beta}$ then we should have $I(\alpha, \beta) = 0$.

Let us now consider the problem of determining $\beta(\alpha)$ as a function of α :

LEMMA 4.2. — For $\alpha \in \mathbb{R}$, the quantity

$$I(\alpha, \beta) = \int_0^1 (\mathbf{c}(x)^{\alpha+1} x^{-\beta} + \mathbf{c}(1-x)^{\alpha+1} (1-x)^{-\beta} - 1) \mathbf{p}(x) dx$$

is finite for all $\beta \in (-\infty, 2\alpha + 3/2)$. It is continuous and strictly increasing as a function of β , with limits

$$\lim_{\beta \rightarrow -\infty} I(\alpha, \beta) = -\infty \quad \text{and} \quad \lim_{\beta \rightarrow 2\alpha + 3/2^-} I(\alpha, \beta) = +\infty.$$

Proof. — By symmetry, we only need to check what happens when x is close to 0. Since $\mathbf{c}(1) = 1$ and $\mathbf{c}$ is differentiable, we have $\mathbf{c}(1-x)^{\alpha+1} (1-x)^{-\beta} - 1 \sim Ax$ for some $A \in \mathbb{R}$, hence $(x(1-x))^{-3/2} (\mathbf{c}(1-x)^{\alpha+1} (1-x)^{-\beta} - 1)$ is integrable, while since $\mathbf{c}(x) \sim x^2$, $(x(1-x))^{-3/2} \mathbf{c}(x)^{\alpha+1} x^{-\beta} \sim x^{2\alpha-\beta+1/2}$ is integrable only if $\beta < 2\alpha + 3/2$.

Monotonicity, continuity and limit properties are direct consequences of the same properties for the integrand and the monotone convergence theorem. $\square$

It follows that, for each $\alpha \in \mathbb{R}$, there is a unique $\beta(\alpha)$ such that $I(\alpha, \beta(\alpha)) = 0$.

Proof of Theorem 1.2. — Fix $\alpha \in \mathbb{R}$, let $\beta = \beta(\alpha)$ as well as $e_n = \mathbb{E}[M_n^\alpha]$. We will prove that, for all $\varepsilon > 0$, $e_n = O(n^{-\beta+\varepsilon})$ and $e_n = \Omega(n^{-\beta-\varepsilon})$, from which the result follows straightforwardly.

We start with the upper bound. For $n \in \mathbb{N}$, let $C_n = \sup_{0 \leq k \leq n} e_k (k+1)^{\beta-\varepsilon}$; we aim to show by induction that C_n stays bounded as n tends to infinity. Recall the equation

$$e_n = \sum_{a=0}^{n-1} P(a, b) (C(a, b)^{\alpha+1} e_a + C(b, a)^{\alpha+1} e_b),$$

where b is implicitly always equal to $n-1-a$. We then have

$$e_n \leq C_{n-1} \sum_{a=0}^{n-1} P(a, b) (C(a, b)^{\alpha+1} (a+1)^{-\beta+\varepsilon} + C(b, a)^{\alpha+1} (b+1)^{-\beta+\varepsilon}).$$

Using the fact that $\sum_{a=0}^{n-1} P(a, b) = 1$, we can now write

$$\begin{aligned} & \frac{e_n}{C_{n-1} (n+1)^{-\beta+\varepsilon}} \\ & \leq 1 + \sum_{a=0}^{n-1} P(a, b) \left(C(a, b)^{\alpha+1} \left(\frac{a+1}{n+1} \right)^{-\beta+\varepsilon} + C(b, a)^{\alpha+1} \left(\frac{b+1}{n+1} \right)^{-\beta+\varepsilon} - 1 \right). \end{aligned}$$

While the sum on the right-hand side is not of the type featured in Proposition 2.7, it is reasonable to expect that

$$\begin{aligned} (28) \quad \lim_{n \rightarrow \infty} \sqrt{n} \sum_{a=0}^{n-1} P(a, b) & \left(C(a, b)^{\alpha+1} \left(\frac{a+1}{n+1} \right)^{-\beta+\varepsilon} + C(b, a)^{\alpha+1} \left(\frac{b+1}{n+1} \right)^{-\beta+\varepsilon} - 1 \right) \\ & = \frac{1}{4\sqrt{\pi}} I(\alpha, \beta - \varepsilon). \end{aligned}$$

We postpone checking (28) to the end of this proof.

Now, noting that $I(\alpha, \beta - \varepsilon) < 0$, we then get that, for n large enough,

$$\frac{e_n}{C_{n-1}n^{-\beta+\varepsilon}} \leq 1 - \frac{A}{\sqrt{n}}$$

for some $A > 0$, hence $e_n \leq C_{n-1}(1 - A/\sqrt{n})(n+1)^{-\beta+\varepsilon} \leq C_{n-1}(n+1)^{-\beta+\varepsilon}$. This shows that $C_{n+1} = C_n$, and so the sequence $(C_n, n \in \mathbb{N})$ is eventually constant.

The lower bound is proved the same way: letting $c_n = \inf_{k \leq n} e_k(k+1)^{\beta+\varepsilon}$, we obtain

$$\frac{e_n}{c_{n-1}(n+1)^{-\beta-\varepsilon}} \geq 1 + \frac{1}{4\sqrt{\pi n}} I(\alpha, \beta + \varepsilon) + o(1/\sqrt{n}),$$

and since $I(\alpha, \beta + \varepsilon) > 0$ this implies that the sequence $(c_n, n \in \mathbb{N})$ is eventually constant, completing our proof.

We now prove statement (28). We will use bounds (9) and (10) as well as

$$\frac{1}{3}c(a/n) \leq C(a, b) \leq 3c(a/n),$$

for $a + b = n - 1$. As with Proposition 2.7, it is sufficient to only consider

$$\sqrt{n} \sum_{a=2}^{\lfloor n/2 \rfloor} P(a, b) \left(C(a, b)^{\alpha+1} \left(\frac{a+1}{n+1} \right)^{-\beta+\varepsilon} + C(b, a)^{\alpha+1} \left(\frac{b+1}{n+1} \right)^{-\beta+\varepsilon} - 1 \right).$$

This sum is equal to $\int_0^{1/2} g_n(x) dx$, where

$$g_n(x) = n^{3/2} P(a, b) \left(C(a, b)^{\alpha+1} \left(\frac{a+1}{n+1} \right)^{-\beta+\varepsilon} + C(b, a)^{\alpha+1} \left(\frac{b+1}{n+1} \right)^{-\beta+\varepsilon} - 1 \right)$$

if $x \in [\frac{a-1}{n}, \frac{a}{n}]$ with $a \in \{2, \dots, \lfloor n/2 \rfloor\}$, and $g_n(x) = 0$ otherwise. It is readily checked that $g_n(x)$ converges pointwise towards

$$\frac{1}{4\sqrt{\pi}} p(x) (c(x)^{\alpha+1} x^{-\beta+\varepsilon} + c(1-x)^{\alpha+1} (1-x)^{-\beta+\varepsilon} - 1),$$

and we need to check that the convergence is dominated. Notice the following, where the constant $K > 0$ may vary from line to line:

- Since $n^{3/2} P(a, b) (a/n)^{3/2} (b/n)^{3/2}$ is uniformly bounded in a and n , and $x < 1/2$, we have $n^{3/2} P(a, b) \leq Kx^{-3/2}$.

- Since $(1/3)c(a/n) \leq C(a, b) \leq 3c(a/n)$ and $x \leq a/n \leq 2x$, we have $(1/3)c(x) \leq C(a, b) \leq 3c(2x)$. Since moreover $x \leq (a+1)/(n+1) \leq 3x$, we deduce that $C(a, b)^{\alpha+1} ((a+1)/(n+1))^{-\beta+\varepsilon} \leq Kx^{2(\alpha+1)} x^{-\beta+\varepsilon}$, independently of the signs of $\alpha + 1$ and β .

- Similarly, we have $|1 - C(b, a)| \leq Kx^2$, and $|1 - (b+1)/(n+1)| \leq Kx$ hence $|C(b, a)^{\alpha+1} ((b+1)/(n+1))^{-\beta+\varepsilon} - 1| \leq Kx$.

Combining these together yields a constant K such that

$$|g_n(x)| \leq \frac{K^{-3/2}}{x} (x^{2\alpha+2-\beta+\varepsilon} + x),$$

which is a uniform integrable bound on $(0, 1/2)$. $\square$

REMARK 4.3. — It is easy to prove by induction on $n \geq 1$ that the smallest possible leaf-growth weight of a leaf in a binary tree of size n is given by $C(0, n-1) \sim 3/2n^2$, which is the weight of a leaf attached to the root of such a tree. An interesting question

raised by a referee is whether one could show that there exists a constant $s_{\max}$ such that

$$\max_{\ell \in \tau_n} \nu_{\tau_n}(\ell) = n^{-s_{\max} + o_{\mathbb{P}}(1)},$$

where $o_{\mathbb{P}}(1)$ is a function which tends to 0 in probability. More generally, an alternative way to present the multi-fractal spectrum is to prove the existence of a function ζ such that

$$\nu_{\tau_n}(\ell \in \tau_n \mid \nu_{\tau_n}(\ell) = n^{-s+o(1)}) = n^{-\zeta(s)+o_{\mathbb{P}}(1)}.$$

Given the expression above, we would expect (but do not prove here) that ζ and β should be related by

$$\beta(\alpha) = \inf_{s \geq 0} (s \cdot (\alpha + 1) + \zeta(s)).$$

5. TOWARDS A DIFFUSION LIMIT

In this work we studied in details the leaf-growth measure on large binary trees and their continuous limit. This is the first step towards understanding the possible Markov process on real trees that one would obtain by passing the discrete Luczak–Winkler growth procedure to the limit. In what follows, we heuristically call “CRT dynamic” any Markov process whose values are compact real trees and whose invariant measure is given by the Brownian CRT.

Stationary dynamics. — A natural way to define a CRT dynamic is to take a scaling limit of a Markov chain $(T_k^{(n)} : k \geq 1)$ on (say) the space of binary trees with fixed size n with uniform invariant distribution. There are several natural discrete chains that one might use as a starting point for this kind of construction (e.g., the flip dynamic on trees [MT99], Aldous’ move on cladograms [LMW20, FPRW23], or subtree pruning and re-grafting procedures [EW06]), and various approaches have been taken to identifying potential continuous limits or proving convergence results. However, this has generally proved quite challenging. Notably, for Aldous’ chain on cladograms, convergence to a diffusion limit was proved by Löhr, Mytnik and Winter in the setting of algebraic measure trees; a closely related pathwise construction of the so-called Aldous diffusion as a stationary Brownian-CRT-valued process in the Gromov–Hausdorff–Prokhorov topology was developed by Forman, Pal, Rizzolo and Winkel.

One of the challenges in passing such dynamics to the limit as $n \rightarrow \infty$ is understanding the appropriate scaling in time, which is related to the mixing time of the chain: one would want to find $f(n)$ such that

$$(T_{[f(n)t]}^{(n)}, t \geq 0)$$

converges towards a CRT dynamic. When the Markov chain is itself reversible, the dynamic can even be extended to $\mathbb{R}$. Finding $f(n)$ is usually a hard problem and has been the subject of many investigations in recent years.

Growing dynamics. — Another route is to consider a growing chain $(T_n, n \geq 1)$ on (say) the set of binary trees so that T_n is uniformly distributed on plane binary trees of size n for each $n \geq 1$ fixed. Those chains do not require to understand the mixing time since the time and size are intimately tight and the limit CRT-dynamic should be the scaling limit as $n \rightarrow \infty$ of

$$(29) \quad (T_{[nt]}/\sqrt{nt}, t \geq 0).$$

The most well-known growth procedure is perhaps Rémy's algorithm [Rém85], where iteratively, a uniform edge of T_n is split in its middle to yield T_{n+1} . When the choices of edges are made independently for each n , the resulting Markov chain converges almost surely towards a Brownian CRT [CH13], so that the limiting CRT dynamic obtained in (29) is constant (which is not so interesting). However, still in the Rémy growth, there is a less well-known way of coupling the choices of edges, due to Bacher, Bodini and Jacquot [BBJ17], which yields a different Markov chain, whose growth is more “local”. The scaling limits of this chain given in (29) is under current investigation and is a CRT dynamic with “local and continuous growth”.

The Luczak–Winkler leaf-growth mechanism studied in these pages is totally different from Rémy's algorithm. We conjecture that if $(T_n^{\ell w}, n \geq 1)$ is a sequence of uniform binary trees so that conditionally on $(T_k^{\ell w}, 1 \leq k \leq n)$, the tree $T_{n+1}^{\ell w}$ is obtained by growing a cherry on point sample according to the leaf-growth measure on $T_n^{\ell w}$, then (29) yields a non-trivial CRT-dynamic $(\mathcal{LW}_t, t \geq 0)$. This dynamic is introduced in [CFT25] as a special case of a more general growing mechanism for self-similar Markov trees. The following proposition indicates that the limiting CRT dynamic is mixing:

PROPOSITION 5.1. — *Consider a sequence $(m_n, n \in \mathbb{N})$ which tends to ∞ such that $m_n = o(n)$. We abuse notation and skip the n index in the sequel. We have*

$$\left(\frac{1}{\sqrt{m}} T_m^{\ell w}, \frac{1}{\sqrt{n}} T_n^{\ell w} \right) \xrightarrow[n \rightarrow \infty]{(d)} (2\sqrt{2} \mathcal{T}, 2\sqrt{2} \mathcal{T}'),$$

where $\mathcal{T}'$ is an independent copy of $\mathcal{T}$.

Proof. — By construction, $T_n^{\ell w}$ can be seen as a tree obtained from $T_m^{\ell w}$ by grafting trees at each of the $m+1$ leaves of $T_m^{\ell w}$. More precisely, if we denote by $S_1(n, m), \dots, S_{m+1}(n, m)$ the sizes of the subtrees $(\theta_1(n, m), \dots, \theta_{m+1}(n, m))$ grafted on the leaves of $T_m^{\ell w}$ to get $T_n^{\ell w}$, it should be clear from the dynamics that conditionally on the vector $(S_1(n, m), \dots, S_{m+1}(n, m))$ and on $T_m^{\ell w}$, the random trees $(\theta_1(n, m), \dots, \theta_{m+1}(n, m))$ are independent and uniform binary trees of sizes $(S_1(n, m), \dots, S_{m+1}(n, m))$. Since the random tree $T_m^{\ell w}$ has diameter of order $\sqrt{m} = o_{\mathbb{P}}(\sqrt{n})$ the root of the trees $\theta_1(n, m), \dots, \theta_{m+1}(n, m)$ are almost confounded in $T_n^{\ell w}$. Since the root of the Brownian CRT is almost surely a point of degree 1, it follows that as $n, m \rightarrow \infty$ satisfying $n/m \rightarrow \infty$, then with high probability there

exists a unique index $1 \leq I(n, m) \leq m + 1$ so that

$$n^{-1} \cdot S_{I(n, m)}(n, m) \longrightarrow 1 \quad \text{and} \quad n^{-1/2} \cdot \max_{i \neq I(n, m)} \text{Height}(\theta_i(n, m)) \longrightarrow 0,$$

in probability. In particular, the geometry of $n^{-1/2} \cdot T_n^{\ell w}$ is close to that of $S_{I(n, m)}(n, m)^{-1/2} \cdot \theta_{I(n, m)}(n, m)$ which converges towards a Brownian CRT, while being independent of $T_m^{\ell w}$, and hence the respective scaling limits are also independent. $\square$

Let us conclude this paper with the following remark: as in [BBJ17] for the Rémy growth, it seems possible in our case to *couple* the sampling of the leaves in the leaf-growth mechanism to get even more “local” growth process $(\tilde{T}_n^{\ell w} : n \geq 0)$ whose dynamical scaling limits seem easier to establish than that $(T_n^{\ell w} : n \geq 0)$. We hope to address those questions in future work.

APPENDIX A. COMPUTING THE INTEGRAL (18)

We are going to prove that, for all real $\alpha \geq 0$,

$$(30) \quad \Phi(\alpha) = \frac{1}{\sqrt{2\pi}} \int_0^1 (1 - x^\alpha \mathbf{c}(x) - (1 - x)^\alpha \mathbf{c}(1 - x)) \mathbf{p}(x) dx = 2\sqrt{2}\alpha \frac{\Gamma(3/2 + \alpha)}{\Gamma(2 + \alpha)}.$$

Recall that for $x, y \in \mathbb{C}$ with $\Re x, \Re y > 0$ the Beta function is defined by

$$B(x, y) = \int_0^1 t^{x-1} (1 - t)^{y-1} dt.$$

and satisfies the fundamental equality

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)}.$$

Note that, since $\Gamma(1/2) = \sqrt{\pi}$, what we are claiming is that $\sqrt{\pi/2} \Phi(\alpha) = 2\alpha B(3/2 + \alpha, 1/2)$. In order to show this, one can use the following lemma:

LEMMA A.1. — For $\alpha \in \mathbb{C}$ such that $\Re \alpha > 1/2$, we have

$$\int_0^1 (1 - x^\alpha - (1 - x)^\alpha) (x(1 - x))^{-3/2} dx = 4\alpha (B(\alpha + 1/2, 1/2) - B(3/2, \alpha - 1/2)).$$

Proof. — Let $F(x) = 1 - x^\alpha - (1 - x)^\alpha$, $g(x) = (x(1 - x))^{-3/2}$ and

$$G(x) = \frac{2(2x - 1)}{\sqrt{x(1 - x)}}.$$

One can immediately check that $G(x)$ is an antiderivative of $g(x)$; thus, integrating by parts and noticing that the boundary term $F(x)G(x)$ has limit equal to 0 both as

$x \rightarrow 0$ and $x \rightarrow 1$, we find that the integral in the statement is equal to

$$\begin{aligned} - \int_0^1 F'(x)G(x)dx &= \int_0^1 (\alpha x^{\alpha-1} - \alpha(1-x)^{\alpha-1}) \frac{2(2x-1)}{\sqrt{x(1-x)}} dx \\ &= 2\alpha \int_0^1 (2x^\alpha - x^{\alpha-1} - 2x(1-x)^{\alpha-1} + (1-x)^{\alpha-1}) \frac{1}{\sqrt{x(1-x)}} dx \\ &= 2\alpha (2B(\alpha+1/2, 1/2) - B(\alpha-1/2, 1/2) - 2B(3/2, \alpha-1/2) \\ &\quad + B(1/2, \alpha-1/2)) \\ &= 2\alpha (2B(\alpha+1/2, 1/2) - 2B(3/2, \alpha-1/2)), \end{aligned}$$

where we have used the symmetry $B(1/2, \alpha-1/2) = B(\alpha-1/2, 1/2)$. $\square$

Now, in order to compute (30), we can just express $f_3(x)$ as $2x^2(1-x) + x^2$, thus decomposing $2\sqrt{\pi/2}\Phi(\alpha)$ as the sum of three integrals $I_1(\alpha), I_2(\alpha), I_3(\alpha)$:

$$\begin{aligned} I_1(\alpha) &:= \int_0^1 -x^\alpha 2x^2(1-x)x^{-3/2}(1-x)^{-3/2} dx = -2B(3/2 + \alpha, 1/2). \\ I_2(\alpha) &:= \int_0^1 -(1-x)^\alpha 2(1-x)^2 xx^{-3/2}(1-x)^{-3/2} dx = -2B(1/2, 3/2 + \alpha). \\ I_3(\alpha) &= \int_0^1 (1-x^{\alpha+2} - (1-x)^{\alpha+2})(x(1-x))^{-3/2} dx. \end{aligned}$$

Using Lemma A.1 with parameter $\alpha+2$ (thus merely requiring the condition that $\Re\alpha > -3/2$), the quantity $I_3(\alpha)$ can be rewritten as

$$I_3(\alpha) = 2(2\alpha+4)(B(5/2 + \alpha, 1/2) - B(3/2, 3/2 + \alpha)).$$

On the other hand, one has

$$\begin{aligned} B(5/2 + \alpha, 1/2) &= \Gamma(5/2 + \alpha)\Gamma(1/2)/\Gamma(3 + \alpha) \\ &= (3/2 + \alpha)\Gamma(3/2 + \alpha)\Gamma(1/2)/(2 + \alpha)\Gamma(2 + \alpha) \\ &= 2\alpha + 3/2\alpha + 4B(3/2 + \alpha, 1/2) \end{aligned}$$

and

$$\begin{aligned} B(3/2, 3/2 + \alpha) &= \Gamma(3/2)\Gamma(3/2 + \alpha)/\Gamma(3 + \alpha) \\ &= 1/2\Gamma(1/2)\Gamma(3/2 + \alpha)/(2 + \alpha)\Gamma(2 + \alpha) \\ &= 1/2\alpha + 4B(3/2 + \alpha, 1/2). \end{aligned}$$

Setting for simplicity $B := B(3/2 + \alpha, 1/2)$, we have obtained

$$\begin{aligned} 2\sqrt{\pi/2}\Phi(\alpha) &= I_1(\alpha) + I_2(\alpha) + I_3(\alpha) \\ &= -4B + 2(2\alpha+4)(2\alpha+3/2\alpha+4B-1/2\alpha+4B) \\ &= -4B + 2(2\alpha+2)B = 4\alpha B, \end{aligned}$$

as desired.

APPENDIX B. LIMITS OF TREES AND SUBTREES

LEMMA B.1. — Consider a sequence $((\mathcal{T}_n, \rho_n, \mu_n), n \in \mathbb{N})$ of deterministic rooted $\mathbb{R}$ -trees which are subsets of a fixed compact metric space (E, d) , such that $\mathcal{T}_n \rightarrow \mathcal{T}$, $\rho_n \rightarrow \rho$ and $\mu_n \rightarrow \mu$ for respectively the Hausdorff metric d_H on compact subsets of E , the metric d on E , and the Prokhorov metric d_P for Borel probability measures on E . Then, for any $z \in \mathcal{T}$, there exists a sequence $(z_n, n \in \mathbb{N})$ with $z_n \in \mathcal{T}_n$ such that $z_n \rightarrow z$ and $\mu_n((\mathcal{T}_n)_{z_n}) \rightarrow \mu(\mathcal{T}_z)$.

Proof. — Write

$$\varepsilon_n := 2 \max(d_H(\mathcal{T}_n, \mathcal{T}), d(\rho_n, \rho), d_P(\mu_n, \mu)),$$

this tends to 0 as n tends to infinity.

Let h be the height of z , for all n , x_n an element of $\mathcal{T}_n$ such that $d(z_n, x_n) \leq \varepsilon_n$, and y_n be the ancestor of x_n in $\mathcal{T}$ with height $\text{ht}(x_n) - \varepsilon_n$. Notice that $d(y_n, z) \leq 3\varepsilon_n$.

Let us show that $(\mathcal{T}_n)_{y_n}$ converges to $\mathcal{T}_z$ for the Hausdorff metric. We use the notation A^ε to denote the ε -enlargement of a subset A of E and will show that we have both $\mathcal{T}_z \subset (\mathcal{T}_n)_{y_n}^{\varepsilon_n}$ and $(\mathcal{T}_n)_{y_n} \subset \mathcal{T}_z^{\varepsilon'_n}$ with $\varepsilon'_n \rightarrow 0$.

Let $a \in \mathcal{T}_z$, we know that there is $b \in \mathcal{T}_n$ such that $d(a, b) \leq \varepsilon_n$, and we will show that b is in $(\mathcal{T}_n)_{y_n}$. By the triangle inequality we have $d(b, x_n) \leq d(b, a) + d(a, z) + d(z, x_n) \leq d(a, z) + 2\varepsilon_n$ and $d(a, z) \leq d(b, x_n) + 2\varepsilon_n$, hence $|d(b, x_n) - d(a, z)| \leq 2\varepsilon_n$. Similarly we obtain $|d(b, \rho_n) - d(a, \rho)| \leq 2\varepsilon_n$ and $|d(x_n, \rho_n) - d(z, \rho)| \leq 3\varepsilon_n$. Since $d(a, \rho) + d(a, z) + d(z, \rho)$, we deduce that $d(b, x_n) + d(x_n, \rho_n) - d(b, \rho_n) \leq 6\varepsilon_n$. Since $\mathcal{T}_n$ is an $\mathbb{R}$ -tree, we can see that this difference is equal to twice the distance from x_n to its most recent common ancestor with b is at most $3\varepsilon_n$, hence this most recent common ancestor is a descendant of y_n , hence $b \in (\mathcal{T}_n)_{y_n}$, hence $\mathcal{T}_z \subset (\mathcal{T}_n)_{y_n}^{\varepsilon_n}$.

Similarly, omitting the details, it can be seen that $(\mathcal{T}_n)_{y_n} \subset \mathcal{T}_{u_n}$, where u_n is an ancestor of z such that $d(u_n, z) \rightarrow 0$. However, since $\mathcal{T}$ is compact, we deduce that $\mathcal{T}_{u_n} \subset \mathcal{T}_z^{\varepsilon'_n}$, where ε'_n tends to 0, otherwise there would be a sequence with no subsequential limits.

We can now finish the proof. Letting $\eta_n = \max(\varepsilon_n, d_H((\mathcal{T}_n)_{y_n}, \mathcal{T}_z))$ for all n , we have $\mathcal{T}_z \subset (\mathcal{T}_n)_{z'_n}^{\varepsilon_n}$, we now finally define z_n as the ancestor of z'_n such that $\text{ht}(z_n) = \text{ht}(z'_n) - \eta_n$. The same argument as in the previous paragraph shows that $(\mathcal{T}_n)_{z_n}$ converges to $\mathcal{T}_z$ for the Hausdorff distance, hence $\eta'_n := \max(\eta_n, d_H((\mathcal{T}_n)_{z_n}, \mathcal{T}_z))$ tends to 0, and we have the following inclusions of sets:

$$\mathcal{T}_z \subset (\mathcal{T}_n)_{z'_n}^{\eta_n} \subset (\mathcal{T}_n)_{z_n} \subset \mathcal{T}_z^{\eta'_n}.$$

By the definition of the Prokhorov metric on measures, we then have

$$\mu(\mathcal{T}_z) \leq \mu_n((\mathcal{T}_n)_{z_n}) + \eta_n \leq \mu(\mathcal{T}_z^{\eta'_n}) + \eta_n + \eta'_n.$$

However, since $\mathcal{T}_z$ is a closed set, $\mu(\mathcal{T}_z^{\eta'_n})$ converges to $\mu(\mathcal{T}_z)$, and by the sandwich theorem, so does $\mu_n((\mathcal{T}_n)_{z_n})$. $\square$

REFERENCES

- [ADH13] R. ABRAHAM, J.-F. DELMAS & P. HOSCHEIT – “A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces”, *Electron. J. Probab.* **18** (2013), article no. 14 (21 pages).
- [AB14] L. ADDARIO-BERRY – “Growing random 3-connected maps”, *Electron. Comm. Probab.* **19** (2014), article no. 54 (12 pages).
- [Ald91a] D. ALDOUS – “The continuum random tree. I”, *Ann. Probab.* **19** (1991), no. 1, p. 1–28.
- [Ald91b] ———, “The continuum random tree. II. An overview”, in *Stochastic analysis (Durham, 1990)*, London Math. Soc. Lecture Note Ser., vol. 167, Cambridge Univ. Press, Cambridge, 1991, p. 23–70.
- [Ald93] ———, “The continuum random tree. III”, *Ann. Probab.* **21** (1993), no. 1, p. 248–289.
- [ALW16] S. ATHREYA, W. LÖHR & A. WINTER – “The gap between Gromov-vague and Gromov-Hausdorff-vague topology”, *Stochastic Process. Appl.* **126** (2016), no. 9, p. 2527–2553.
- [BBJ17] A. BACHER, O. BODINI & A. JACQUOT – “Efficient random sampling of binary and unary-binary trees via holonomic equations”, *Theoret. Comput. Sci.* **695** (2017), p. 42–53.
- [Ber02] J. BERTOIN – “Self-similar fragmentations”, *Ann. Inst. H. Poincaré Probab. Statist.* **38** (2002), no. 3, p. 319–340.
- [BK16] J. BERTOIN & I. KORTCHEMSKI – “Self-similar scaling limits of Markov chains on the positive integers”, *Ann. Appl. Probab.* **26** (2016), no. 4, p. 2556–2595.
- [Bil65] P. BILLINGSLEY – *Ergodic theory and information*, Wiley Series in Probab. and Math. Stat., Wiley, 1965.
- [CS20] A. CARACENI & A. STAUFFER – “Polynomial mixing time of edge flips on quadrangulations”, *Probab. Theory Relat. Fields* **176** (2020), no. 1-2, p. 35–76.
- [CS23] ———, “Growing uniform planar maps face by face”, *Random Structures Algorithms* **63** (2023), no. 4, p. 942–967.
- [CPY97] P. CARMONA, F. PETIT & M. YOR – “On the distribution and asymptotic results for exponential functionals of Lévy processes”, in *Exponential functionals and principal values related to Brownian motion*, Bibl. Rev. Mat. Iberoamericana, Revista Matemática Iberoamericana, Madrid, 1997, p. 73–130.
- [CFT25] N. CURIEN, W. FLEURAT & A. TWIGT – “Growing self-similar Markov trees”, 2025, work in progress.
- [CH13] N. CURIEN & B. HAAS – “The stable trees are nested”, *Probab. Theory Relat. Fields* **157** (2013), no. 3-4, p. 847–883.
- [CLG17] N. CURIEN & J.-F. LE GALL – “The harmonic measure of balls in random trees”, *Ann. Probab.* **45** (2017), no. 1, p. 147–209.
- [DLG05] T. DUQUESNE & J.-F. LE GALL – “Probabilistic and fractal aspects of Lévy trees”, *Probab. Theory Relat. Fields* **131** (2005), no. 4, p. 553–603.
- [DLG06] ———, “The Hausdorff measure of stable trees”, *Alea* **1** (2006), p. 393–415.
- [EPW06] S. N. EVANS, J. PITMAN & A. WINTER – “Rayleigh processes, real trees, and root growth with re-grafting”, *Probab. Theory Relat. Fields* **134** (2006), no. 1, p. 81–126.
- [EW06] S. N. EVANS & A. WINTER – “Subtree prune and regraft: a reversible tree-valued Markov process”, *Ann. Probab.* **34** (2006), no. 3, p. 81–126.
- [FPRW18] N. FORMAN, S. PAL, D. RIZZOLO & M. WINKEL – “Aldous diffusion I: a projective system of continuum k -tree evolutions”, 2018, [arXiv:1809.07756](https://arxiv.org/abs/1809.07756).
- [FPRW20] ———, “Projections of the Aldous chain on binary trees: intertwining and consistency”, *Random Structures Algorithms* **57** (2020), no. 3, p. 745–769.
- [FPRW23] ———, “The Aldous diffusion: a stationary evolution of the Brownian CRT”, 2023, [arXiv:2305.17269](https://arxiv.org/abs/2305.17269).
- [GP19] C. GOLDSCHMIDT & M. PRZYKUCKI – “Parking on a random tree”, *Combin. Probab. Comput.* **28** (2019), no. 1, p. 23–45.
- [HM04] B. HAAS & G. MIERMONT – “The genealogy of self-similar fragmentations with negative index as a continuum random tree”, *Electron. J. Probab.* **9**(4) (2004), p. 57–97.
- [HM12] ———, “Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees”, *Ann. Probab.* **40** (2012), no. 6, p. 2589–2666.

- [HS15] B. HAAS & R. STEPHENSON – “Scaling limits of k -ary growing trees”, *Ann. Inst. H. Poincaré Probab. Statist.* **51** (2015), no. 4, p. 1314–1341.
- [Khe20] A. KHEZELI – “Metriization of the Gromov–Hausdorff(-Prokhorov) topology for boundedly-compact metric spaces”, *Stochastic Process. Appl.* **130** (2020), no. 6, p. 3842–3864.
- [Khe23] ———, “A unified framework for generalizing the Gromov–Hausdorff metric”, *Probab. Surv.* **20** (2023), p. 837–896.
- [Lam62] J. LAMPERTI – “Semi-stable stochastic processes”, *Trans. Amer. Math. Soc.* **104** (1962), p. 62–78.
- [Law93] G. F. LAWLER – “A discrete analogue of a theorem of Makarov”, *Combin. Probab. Comput.* **2** (1993), no. 2, p. 181–199.
- [LG05] J.-F. LE GALL – “Random trees and applications”, *Probab. Surv.* **2** (2005), p. 245–311.
- [LMW20] W. LÖHR, L. MYTNIK & A. WINTER – “The Aldous chain on cladograms in the diffusion limit”, *Ann. Probab.* **48** (2020), no. 5, p. 2565–2590.
- [LW04] M. LUCZAK & P. WINKLER – “Building uniformly random subtrees”, *Random Structures Algorithms* **24** (2004), no. 4, p. 420–443.
- [LPS08] R. LYONS, R. PELED & O. SCHRAMM – “Growth of the number of spanning trees of the Erdős–Rényi giant component”, *Combin. Probab. Comput.* **17** (2008), no. 5, p. 711–726.
- [LPP95] R. LYONS, R. PEMANTLE & Y. PERES – “Ergodic theory on Galton–Watson trees: speed of random walk and dimension of harmonic measure”, *Ergodic Theory Dynam. Systems* **15** (1995), no. 3, p. 593–619.
- [Mar03] P. MARCHAL – “Constructing a sequence of random walks strongly converging to Brownian motion”, in *Discrete random walks (Paris, 2003)*, Discrete Math. Theor. Comput. Sci. Proc., AC, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2003, p. 181–190.
- [Mar23] J.-F. MARCKERT – “Growing random uniform d -ary trees”, *Ann. Comb.* **27** (2023), no. 1, p. 51–66.
- [MT99] L. MCSHINE & P. TETALI – “On the mixing time of the triangulation walk and other Catalan structures”, in *Randomization methods in algorithm design (Princeton, NJ, 1997)*, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 43, American Mathematical Society, Providence, RI, 1999, p. 147–160.
- [PW15] J. PITMAN & M. WINKEL – “Regenerative tree growth: Markovian embedding of fragmenters, bifurcators, and bead splitting processes”, *Ann. Probab.* **43** (2015), no. 5, p. 2611–2646.
- [Ré85] J.-L. RÉMY – “Un procédé itératif de dénombrement d’arbres binaires et son application à leur génération aléatoire”, *RAIRO Inform. Théor.* **19** (1985), no. 2, p. 179–195.
- [Ste13] R. STEPHENSON – “General fragmentation trees”, *Electron. J. Probab.* **18(101)** (2013), p. 1–45.

Manuscript received 15th January 2024

accepted 2nd July 2026

ALESSANDRA CARACENI, Scuola Normale Superiore di Pisa,
Piazza dei Cavalieri 7, 56121 Pisa, Italy

E-mail : alessandra.caraceni@sns.it

Url : <https://alessandracaraceni.altervista.org/>

NICOLAS CURIEN, Université Paris-Saclay, Laboratoire de mathématiques d’Orsay,
Bât. 307, 91400 Orsay, France

E-mail : nicolas.curien@gmail.com

Url : <https://www.imo.universite-paris-saclay.fr/~nicolas.curien/>

ROBIN STEPHENSON, School of Mathematical and Physical Sciences, University of Sheffield,
Hicks Building, Sheffield S3 7RH, UK

E-mail : robin.stephenson@normalesup.org

Url : <https://www.normalesup.org/~stephens/>