Check for updates

Received: 23 March 2023 | Revised: 23 October 2023 | Accepted: 20 December 2023

DOI: 10.1002/rsa.21206

RESEARCH ARTICLE

WILEY

Coupling Bertoin's and Aldous–Pitman's representations of the additive coalescent

Igor Kortchemski[1] | **Paul Thévenin**[2]

[1]CMAP, CNRS, École Polytechnique, Institut Polytechnique de Paris, Palaiseau, France
[2]Institut für Mathematik, University of Vienna, Vienna, Austria

Correspondence
Paul Thévenin, Institut für Mathematik, University of Vienna, Vienna, Austria.
Email: paul.thevenin@univie.ac.at

Abstract

We construct a coupling between two seemingly very different constructions of the standard additive coalescent, which describes the evolution of masses merging pairwise at rates proportional to their sums. The first construction, due to Aldous and Pitman, involves the components obtained by logging the Brownian continuum random tree (CRT) by a Poissonian rain on its skeleton as time increases. The second one, due to Bertoin, involves the excursions above its running infimum of a linear-drifted standard Brownian excursion as its drift decreases. Our main tool is the use of an exploration algorithm of the so-called cut-tree of the Brownian CRT, which is a tree that encodes the genealogy of the fragmentation of the CRT.

KEYWORDS

Brownian excursion, fragmentation, probability, random trees

1 | INTRODUCTION

1.1 | Cutting down trees

Starting with a rooted tree, a natural logging operation consists in choosing and removing one of its edges uniformly at random, thus splitting the tree into two connected components. Iterating and removing edges one after another, one obtains a fragmentation process of this tree. This model was introduced by Meir and Moon [31, 32] for random Cayley and recursive trees. They focused on the connected component containing the root, and investigated the number of cuts needed to isolate it. Since then,

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2024 The Authors. *Random Structures & Algorithms* published by Wiley Periodicals LLC.

this subject has brought considerable interest for a number of classical models of deterministic and random trees, including random binary search trees [25, 26], random recursive trees [8, 12, 21, 27] and Bienaymé–Galton–Watson trees conditioned on the total progeny [1, 11, 13, 24, 28, 34].

It is remarkable that the so-called standard additive coalescent, which describes the evolution of masses merging pairwise at rates proportional to their sums, can be defined (after time-reversal) using a continuous analogue of the cutting procedure on discrete trees.

1.2 | The Aldous–Pitman construction

The Aldous–Pitman fragmentation, introduced in [6], describes the evolution of the masses of the connected components of a Brownian CRT $\mathcal{T}$ cut according to a Poissonian rain $\mathcal{P}$ of intensity $\mathrm{d}\lambda \otimes \mathrm{d}t$ on $\mathrm{Sk}(\mathcal{T}) \times \mathbb{R}_+$, where $\mathrm{d}t$ is the Lebesgue measure on $\mathbb{R}_+$ and λ is the length measure on the skeleton $\mathrm{Sk}(\mathcal{T})$ on $\mathcal{T}$ (see Section 2.1 for precise definitions). We set, for every $t \geq 0$,

$$\mathcal{P}_t := \{c \in \mathrm{Sk}(\mathcal{T}), \exists\, s \in [0,t], (c,s) \in \mathcal{P}\}.$$

Then, for every $t \geq 0$, we define $X_{\mathsf{AP}}(t)$ to be the sequence of μ-masses of the connected components of $\mathcal{T} \setminus \mathcal{P}_t$, sorted in nonincreasing order, where μ is the so-called mass measure on $\mathcal{T}$. Then $(X_{\mathsf{AP}}(t))_{t\geq 0}$ is a fragmentation process with explicit characteristics (see [6] and more generally the book [10] for a general theory of stochastic coalescence and fragmentation processes, as well as examples and motivation). Up to time-reversal, X_{AP} is closely related to the well-known standard additive coalescent [6, 22]. It is also interesting to mention that very recently X_{AP} has naturally appeared in the study of random uniform factorizations of large permutations into products of transpositions [23, 36].

1.3 | The Bertoin construction

Bertoin [9] gave another construction of this fragmentation process from a drifted standard Brownian excursion the following way. Let $\mathbb{e}$ be a standard Brownian excursion on $[0, 1]$. For every $t \geq 0$, consider the function $f_t : [0,1] \to \mathbb{R}$ defined by $f_t(s) = \mathbb{e}_s - ts$ for $s \in [0,1]$ and denote by $X_{\mathsf{B}}(t)$ the sequence of lengths of the excursions of f_t above its running infimum, sorted in nonincreasing order. Bertoin [9] proves that this process has the same distribution as the Aldous–Pitman fragmentation of a Brownian CRT (normalized so that it is coded by $2\mathbb{e}$), that is, that X_{B} and X_{AP} have the the same distribution.

It may be puzzling that these two constructions define the same object, since the Aldous–Pitman representation involves two independent levels of randomness (the CRT and the Poissonian rain), while the Bertoin representation only involves a Brownian excursion. For the analog representations in the discrete framework of finite trees, several connections have been recently discovered. We present them just after the statement of Theorem 1.1, which is our main contribution: in the continuous framework, we unify these two constructions and explain why they are actually intimately related.

1.4 | Coupling the two constructions

Our main result consists in coupling X_{B} and X_{AP}.

Theorem 1.1. *The following assertions hold.*

(i) Let $\mathcal{T}$ be a Brownian CRT equipped with the Poissonian rain $\mathcal{P}$. On the same probability space, there is a function $F_{(\mathcal{T},\mathcal{P})}$, measurable with respect to $(\mathcal{T}, \mathcal{P})$, having

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the law of the Brownian excursion and such that almost surely, for every $t \geq 0$, the nonincreasing rearrangement of the masses of the connected components of $\mathcal{T} \setminus \mathcal{P}_t$ is the same as the nonincreasing rearrangement of the lengths of the excursions of $(F_{(\mathcal{T},\mathcal{P})}(s) - ts)_{0 \leq s \leq 1}$ above its running infimum.

(ii) Conversely, given a Brownian excursion $\mathbb{e}$ and an independent sequence of i.i.d. uniform random variables on $[0, 1]$, on the same probability space there is a Brownian CRT $\mathcal{T}$ equipped with a Poissonian rain $\mathcal{P}$ (which are measurable with respect to $\mathbb{e}$ and the latter sequence) such that almost surely $F_{(\mathcal{T},\mathcal{P})} = \mathbb{e}$.

Some comments are in order. In (i), the construction of $F_{(\mathcal{T},\mathcal{P})}$ from $(\mathcal{T}, \mathcal{P})$ is explicit using the so-called “Pac-man algorithm,” which we briefly describe below. In (ii), unlike (i), $\mathbb{e}$ alone is not enough to build $(\mathcal{T}, \mathcal{P})$: an additional independent source of randomness is needed (see Section 4 for details).

Let us first give some underlying intuition. In the discrete world of finite trees, it turns out that there is a beautiful explicit exact relation, due to Broutin and Marckert [15], between masses of components obtained after removing edges one after the other, and lengths of excursions above its running infimum of a certain function. The idea is to use the so-called Prim order to explore an edge-labeled tree. Let us explain this in more detail.

Consider a rooted tree T_n with n vertices, whose edges are labeled from 1 to $n-1$. Its vertices $u_1, \dots, u_n$ listed in Prim order are defined as follows: u_1 is the root of the tree and, for every $i \in [\![1, n-1]\!]$, u_{i+1} is the vertex, among all children of a vertex of $\{u_j, j \leq i\}$, whose edge to its parent has minimum label.

Let T_n be a Cayley tree with n vertices (that is, a tree with n vertices labeled from 1 to n, rooted at 1), whose edges are labeled from 1 to $n-1$ uniformly at random conditionally on T_n. Let $(u_i)_{1 \leq i \leq n}$ be the vertices of T_n sorted according to the Prim order. For every $1 \leq i \leq n$ and $1 \leq k \leq n-1$, let $X_i(k)$ be the number of children of u_i in the forest $F_n(k)$ obtained by deleting all edges with labels belonging to $[\![n-k, n-1]\!]$. Finally, set $S_x(k) := \sum_{i=1}^{\lfloor xn \rfloor}(X_i(k) - 1)$ for every $0 \leq x \leq 1$ (which is called the Prim path of the forest explored in Prim order). Then Broutin and Marckert [15] establish that:

- the lengths of excursions of $(S_x(k))_{0 \leq x \leq 1}$ above its running infimum are equal to the sizes of the connected components of $F_n(k)$ (see Figure 1);
- the following convergence holds in $\mathbb{D}(\mathbb{R}_+, \mathbb{D}([0,1], \mathbb{R}))$:

$$\left(\frac{(S_x(\lfloor tn \rfloor))_{x \in [0,1]}}{\sqrt{n}} \right)_{t \geq 0} \xrightarrow[n \to \infty]{(d)} \left((\mathbb{e}_x - tx)_{x \in [0,1]} \right)_{t \geq 0},$$

  where, for $I \subset \mathbb{R} \cup \{\pm\infty\}$ an interval and E a metric space, $\mathbb{D}(I, E)$ denotes the space of càdlàg functions from I to E endowed with the J_1 Skorokhod topology (we refer to Annex $A2$ in [29] for further definitions and details).

Using the fact that the Aldous–Pitman fragmentation is the continuum analog of this discrete logging procedure, this gives another proof of the fact that X_{B} and X_{AP} have the same distribution. This also indicates that if one couples X_{B} and X_{AP}, then the Brownian excursion appearing in the definition of X_{B} should represent, in a certain sense, the encoding of the exploration of a Brownian CRT equipped with a Poissonian rain using an associated Prim order.

Also, still in the discrete world of finite trees, Marckert and Wang [30] couple a uniform Cayley tree with n vertices and edges labeled from 1 to $n-1$ uniformly at random with a uniform Cayley tree equipped with an independent uniform decreasing edge-labeling (i.e., labels decrease along paths

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

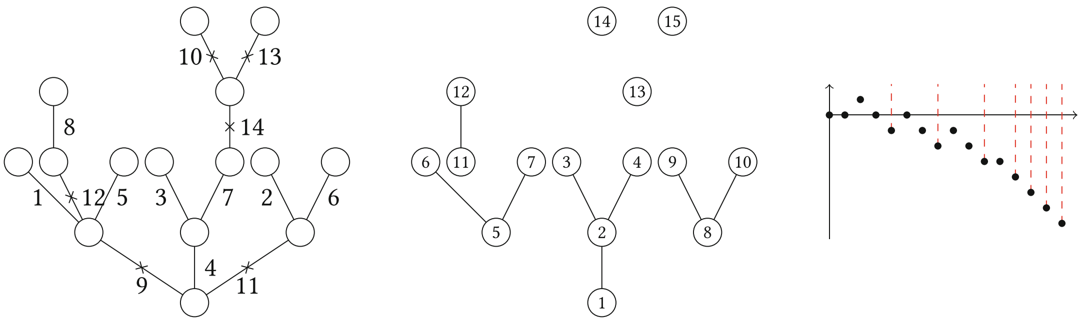

FIGURE 1 Left: an edge-labeled plane tree. Middle: The forest obtained by labeling the vertices in Prim order and removing the six edges with largest labels. Right: The Prim path of the forest explored using the Prim order.

directed from the root towards the leaves) in such a way that edge removals give the same sizes of connected components. A connection with discrete cut-trees is also given (see [30, Section 3]). Indeed, it turns out that this second tree is closely related to the cut-tree of the first one, which allows Marckert and Wang to give a nice simple proof of the well-known fact that the number of cuts needed to isolate a uniform vertex in a uniform Cayley tree, scaled by $\sqrt{n}$, converges in distribution to a Rayleigh random variable (see also [30] for other connections between the standard additive coalescent and other combinatorial and probabilistic models such as size biased percolation and parking schemes in a tree).

However, how to make such statements precise in the realm of continuous trees remains unclear. For this reason, we use a different route to define a coupling between X_{B} and X_{AP}. The main tool in the proof of Theorem 1.1 is the use of the so-called cut-tree $\mathcal{C}$, defined by Bertoin and Miermont in [13], which roughly speaking encodes the genealogy of the fragmentation of a Brownian CRT by a Poissonian rain. Indeed, one of our contributions is to use the cut-tree to define the "Bertoin" excursion $F_{(\mathcal{T},\mathcal{P})}$ from the Aldous–Pitman fragmentation by using an algorithm which we call the "Pac-Man" algorithm, roughly described as follows (see Section 3.1 for a precise definition). With every value $h \in [0, 1]$, using a local exploration procedure, we associate a final target point of $\mathcal{C}$. Then the value $F_{(\mathcal{T},\mathcal{P})}(h)$ is defined using the genealogy of this exploration. Strictly speaking, the cut-tree $\mathcal{C}$ is defined from $(\mathcal{T}, \mathcal{P})$ by using additional randomness (namely a collection of independent i.i.d. points sampled according to the mass measure). However, the quantity $F_{(\mathcal{T},\mathcal{P})}$ can be directly defined as a measurable function of $(\mathcal{T}, \mathcal{P})$ without any reference to the cut-tree, which still serves as a useful tool to check that $F_{(\mathcal{T},\mathcal{P})}$ satisfies the desired properties (see Section 3.5 for details).

Conversely, given the "Bertoin" excursion, we will also see that coupling X_{B} and X_{AP} is closely related to the question of reconstruction of the original CRT from its cut-tree (see Section 4 for more details).

Quite interestingly it seems that, while there is no simple analog of the coupling between drifted excursions and sizes of connected components using Prim's order in the continuous framework, there is no simple analog of the coupling between drifted excursions and sizes of connected components using the cut-tree in the discrete framework. Indeed, in Marckert and Wang's coupling, given the Cayley tree, the decreasing edge labeling is random, while in the continuous framework, given the cut-tree, its labeling is deterministic; see Remark 3.3. In particular, new ideas and techniques are required to analyze the continuous framework.

Finally, let us mention that the question of reconstructing the Brownian CRT from the "Bertoin" excursion appears in an independent work by Nicolas Broutin and Jean-François Marckert [16] in the different context of the study of scaling limits of minimal spanning trees on complete graphs.

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1.5 | Perspectives

There has been some recent developments in studying analogs of the Aldous–Pitman fragmentation for different classes of trees and their associated cut-trees, see [2, 14, 17, 20, 33]. In particular, it has been shown that by an appropriate tuning (fragmentation along the skeleton and/or at nodes), the law of the cut-tree is equal to the law of the original tree.

We expect that our coupling can be extended to more general classes of trees, with fragmentation on the skeleton only, such as stable trees (the study of this fragmentation is mentioned in [2, Section 5, (6)]). Indeed, for stable trees, it is known [33] that the analog of Aldous–Pitman fragmentation can be obtained using the normalized excursion of a stable process. One of the main issues is that in this case the associated cut-tree is not compact anymore; hence, our main argument, which consists in comparing distances in this cut-tree, does not apply directly. Furthermore, new results (see [37]) suggest strong connections between the so-called ICRT (inhomogeneous continuum random trees) and stable trees. Thus similar techniques could work in both cases, and it is plausible that analog couplings exist. We plan to investigate this in future work.

1.6 | Overview of the article

Section 2 is devoted to the definition of our main tool, the cut-tree associated to the fragmentation of the Brownian CRT; we also prove there some preliminary structural results on this object. In Section 3, we prove the first part of our main result, Theorem 1.1, with the help of our so-called Pac-Man algorithm. Finally, we prove Theorem 1.1 (ii) in Section 4, essentially making use of results from [2].

2 | THE CUT-TREE OF THE BROWNIAN CRT

An important object in our study is the cut-tree of a Brownian CRT, which roughly speaking encodes the genealogy of its fragmentation by a Poissonian rain. We recall here its construction and main properties, and refer to [2, 13] for details and proofs.

2.1 | Definitions

Let us first introduce some definitions and notation for trees.

2.1.1 | Real trees

We say that a complete metric space (T, d) is a real tree if, for every $u, v \in T$:

- there exists a unique isometry $f_{u,v} : [0, d(u, v)] \to T$ such that $f_{u,v}(0) = u$ and $f_{u,v}(d(u, v)) = v$;
- for any continuous injective map $f : [0, 1] \to T$ such that $f(0) = u$ and $f(1) = v$, we have $f([0, 1]) = f_{u,v}([0, d(u, v)]) =: [\![u, v]\!]$.

A rooted real tree is a real tree with a distinguished vertex, called the root.

2.1.2 | Tree structure

Let $\mathcal{T}$ be a real tree. We say that a point $x \in \mathcal{T}$ is a leaf if $\mathcal{T} \setminus \{x\}$ is connected, and a branchpoint if $\mathcal{T} \setminus \{x\}$ has at least three connected components. We denote by $\mathcal{B}(\mathcal{T})$ the set of all branchpoints of the tree $\mathcal{T}$ and by $\varnothing$ the root of $\mathcal{T}$.

We let $\mathrm{Sk}(\mathcal{T})$ be the skeleton of $\mathcal{T}$, that is, the set of all points of $\mathcal{T}$ that are not leaves nor branchpoints. We emphasize that this definition differs from the usual one, where the skeleton is defined as the complement of the set of leaves. The reason is that in the sequel of the article, we usually need to treat differently leaves, branchpoints, and non-branchpoints which are not leaves.

We denote by $\mathcal{T}_x$ the tree which is the set of all (weak) descendents of x in $\mathcal{T}$, rooted at x. If the tree $\mathcal{T}$ is binary (that is, for all $x \in \mathcal{T}$, $\mathcal{T} \setminus \{x\}$ has at most three connected components, and $\mathcal{T} \setminus \{\emptyset\}$ has at most two), when x is an ancestor of y, we denote by $\mathcal{T}_x^y$ the subtree above x containing y, rooted at x and by $\overline{\mathcal{T}}_x^y$ the subtree above x *not* containing y, rooted at x (which is unique if it exists).

Furthermore, for any two vertices $x, y \in \mathcal{T}$, we write $x \preceq y$ when x is an ancestor of y, and $x \prec y$ when $x \preceq y$ and $x \neq y$. In particular $\emptyset \preceq x$ for every $x \in \mathcal{T}$ and $\preceq$ is a partial order on $\mathcal{T}$ called the genealogical order.

Finally, for $x, y \in \mathcal{T}$, we denote by $x \wedge y$ their closest common ancestor, that is, the unique $z \in \mathcal{T}$ satisfying $[\![\emptyset, x]\!] \cap [\![\emptyset, y]\!] = [\![\emptyset, z]\!]$.

2.1.3 | Brownian excursion and Brownian tree

The Brownian CRT, introduced by Aldous [3–5], is a random real tree constructed from twice a standard Brownian excursion $\mathbb{e} : [0,1] \to \mathbb{R}_+$ the following way. The function $2\mathbb{e}$ induces an equivalence relation $\sim_{2\mathbb{e}}$ on $[0,1]$: define a pseudo-distance d on $[0,1]$ by setting $d(u,v) = 2\mathbb{e}_u + 2\mathbb{e}_v - 2\min_{[u,v]} 2\mathbb{e}$, and say that, for all $0 \leq u, v \leq 1$, $u \sim_{2\mathbb{e}} v$ if and only if $d(u,v) = 0$. Define now $\mathcal{T} := [0,1]/\sim_{2\mathbb{e}}$, endowed with the distance which is the projection of d on the quotient space (which we also denote by d by convenience). It is standard that $(\mathcal{T}, d)$ is a real tree, called the Brownian CRT.

2.1.4 | Length and mass measures

For any real tree (T, d), observe that the distance d on T induces a length measure λ on $\mathrm{Sk}(T)$, defined as the only σ-finite measure such that $\lambda([\![u, v]\!]) = d(u,v)$ for all $u, v \in T$. In the case of the Brownian CRT $(\mathcal{T}, d)$, we can furthermore endow it with a mass measure μ, which is the pushforward of the Lebesgue measure on $[0,1]$ by the quotient map $[0,1] \to [0,1]/\sim_{2\mathbb{e}}$. Roughly speaking, μ accounts for the proportion of leaves in a given component of $\mathcal{T}$.

2.2 | Construction of the Brownian cut-tree

We first need some notation. Let $\mathcal{T}$ be a Brownian CRT, μ its mass measure and let $\mathcal{P}$ be a Poissonian rain of intensity $\mathrm{d}\lambda \otimes \mathrm{d}t$, where λ denotes the length measure on $\mathrm{Sk}(\mathcal{T})$ and $\mathrm{d}t$ is the Lebesgue measure on $\mathbb{R}_+$. For every $t \geq 0$, we set

$$\mathcal{P}_t := \{c \in \mathrm{Sk}(\mathcal{T}), \exists\, s \in [0,t], (c,s) \in \mathcal{P}\}.$$

The elements of $\mathcal{P}_\infty := \cup_{t \geq 0} \mathcal{P}_t$ are called *cutpoints*. For every $t \geq 0$ and $x \in \mathcal{T}$, we denote by $\mathcal{T}_t(x)$ the connected component of $\mathcal{T} \setminus \mathcal{P}_t$ containing x and $\mu_t(x) = \mu(\mathcal{T}_t(x))$ its μ-mass. If $x \in \mathcal{P}_t$, we set $\mathcal{T}_t(x) = \emptyset$ and $\mu_t(x) = 0$.

Let $U_0 = \emptyset$ be the root of $\mathcal{T}$, and let $(U_i)_{i \geq 1}$ be a sequence of i.i.d. leaves of $\mathcal{T}$ sampled according to the mass measure μ, independently of $\mathcal{P}$. For every $i, j \in \mathbb{Z}_+$, we let $t_{i,j} := \inf\{t \geq 0, \mathcal{T}_t(U_i) \neq \mathcal{T}_t(U_j)\}$ be the first time a cutpoint appears on $[\![U_i, U_j]\!]$. Then by [2, 13] there exists almost surely a metric space $C^\circ := (C^\circ, d^\circ, \rho)$ containing the set $\{\rho\} \cup \mathbb{Z}_+$ such that $C^\circ = \bigcup_{i \in \mathbb{Z}_+} [\![\rho, i]\!]$ and for every $i, j \in \mathbb{Z}_+$,

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$d^\circ(\rho, i) = \int_0^\infty \mu_s(U_i)\,\mathrm{d}s \qquad \text{and} \qquad d^\circ(i,j) = \int_{t_{i,j}}^\infty \big(\mu_s(U_i) + \mu_s(U_j)\big)\mathrm{d}s.$$

We denote by $\mathcal{C} := (\mathcal{C}, d_{\mathcal{C}}, \rho)$ the completion of this metric space, which is a real tree called the cut-tree. The sequence $(i)_{i\geq 1}$ is dense in $\mathcal{C}$, and in particular every branchpoint of $\mathcal{C}$ can be written (non uniquely) as $i \wedge j$ with $i, j \geq 1$. We also endow the set of leaves of $\mathcal{C}$ with the measure ν defined as:

$$\nu = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \delta_i. \tag{1}$$

For a measurable subset $\mathcal{A} \subset \mathcal{C}$ with $\nu(\mathcal{A}) > 0$ we use the notation $\nu_{\mathcal{A}}$ for the probability measure on $\mathcal{C}$ defined by $\nu_{\mathcal{A}}(B) = \nu(\mathcal{A} \cap B)/\nu(\mathcal{A})$.

The main result of [13] is the following:

Theorem 2.1 (Bertoin and Miermont [13]). *The cut-tree of the Brownian CRT has the law of a Brownian CRT*:

$$\mathcal{C} \overset{(d)}{=} \mathcal{T}.$$

The cut-tree $\mathcal{C}$ encodes the genealogical structure, as time increases, of the cuts of the subtrees which contain the points $(U_i)_{i\geq 0}$, in such a way that branchpoints of $\mathcal{C}$ are in correspondence with $\mathcal{P}_\infty$. Let us make this more explicit.

Informally speaking, for every $i, j \in \mathbb{Z}_+$, the branchpoint $i \wedge j$ of $\mathcal{C}$ encodes the (a.s. unique) cutpoint appearing on $[\![U_i, U_j]\!]$ at time $t_{i,j}$. The subtree of $\mathcal{C}$ above $i \wedge j$ containing i (resp. j) is then the cut-tree of the subtree of $\mathcal{T} \setminus \mathcal{P}_{t_{i,j}}$ containing U_i (resp. U_j).

The measures μ on $\mathcal{T}$ and ν on $\mathcal{C}$ are related in the following way (see [2, Proposition 7]):

$$\mu(\mathcal{T}_{t_{i,j}}(U_i)) = \nu(\mathcal{C}^i_{i\wedge j}), \qquad \mu(\mathcal{T}_{t_{i,j}}(U_j)) = \nu(\mathcal{C}^j_{i\wedge j}). \tag{2}$$

In particular, using the fact that $\mathcal{C}$ is binary, since $\mathcal{T}_{t_{i,j}-}(U_i) = \mathcal{T}_{t_{i,j}-}(U_j) = \mathcal{T}_{t_{i,j}}(U_i) \cup \mathcal{T}_{t_{i,j}}(U_j) \cup \{i \wedge j\}$, we have $\mu(\mathcal{T}_{t_{i,j}-}(U_i)) = \nu(\mathcal{C}_{i\wedge j})$.

The leaf $0 \in \mathcal{C}$ will play a distinguished role in the sequel. In the terminology of [2, Section 3.3], it can be seen as the "image" in $\mathcal{C}$ of the root $\varnothing$ of $\mathcal{T}$ in the following sense: if $(U_{i_n})_{n\geq 1}$ is a sequence converging to $\varnothing$, then i_n converges to 0 (this is shown in [2, Section 3.3]). Similarly, for every branchpoint x of $\mathcal{C}$, each of the two subtrees grafted on x comes with a distinguished leaf. Indeed, consider $x \prec y$ two points of $\mathcal{C}$ with x a branchpoint. Recall that $\mathcal{C}_x^y$ denotes the subtree above x containing y, rooted at x and $\overline{\mathcal{C}}_x^y$ the subtree above x *not* containing y, rooted at x. Intuitively speaking, $\Lambda(\mathcal{C}_x^y)$ is the image of x obtained by considering the Poissonian rain in the the subtree $\mathcal{C}_x^y$ rooted at x, and $\Lambda(\overline{\mathcal{C}}_x^y)$ is the image of x obtained by considering the Poissonian rain in the the subtree $\overline{\mathcal{C}}_x^y$ rooted at x.

More formally, we may find $i, j \geq 1$ such that $x = i \wedge j$, $i \in \mathcal{C}_x^y$ and $j \in \overline{\mathcal{C}}_x^y$. Let $c \in [\![U_i, U_j]\!]$ be the cutpoint appearing at time $t_{i,j}$. Consider a sequence $(U_{i_n})_{n\geq 1}$ of elements of $\mathcal{T}_{t_{i,j}}(U_i)$ converging to c and a sequence $(U_{j_n})_{n\geq 1}$ of elements of $\mathcal{T}_{t_{i,j}}(U_j)$ converging to c. Then by [2, Section 3.3] the sequence $(i_n)_{n\geq 1}$ converges in $\mathcal{C}_x^y$ to a leaf denoted by $\Lambda(\mathcal{C}_x^y)$ (which does not depend on the sequence $(U_{i_n})_{n\geq 1}$), and the sequence $(j_n)_{n\geq 1}$ converges in $\overline{\mathcal{C}}_x^y$ to a leaf denoted by $\Lambda(\overline{\mathcal{C}}_x^y)$ (which does not depend on the sequence $(U_{j_n})_{n\geq 1}$). To see that the limiting points have to be leaves, simply observe that $t_{i_n, i_{n+1}} \underset{n\to\infty}{\to} +\infty$, so that

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\nu\left(C^{i_n}_{i_n\wedge i_{n+1}}\right) = \mu_{t_{i_n,i_{n+1}}}\left(U_{i_n}\right) \underset{n\to\infty}{\to} 0.$$

Finally, setting for every $x \in \mathcal{C}$:

$$\tau_x = \int_{[\![\rho,x]\!]} \frac{1}{\nu(\mathcal{C}_z)} \lambda(\mathrm{d}z), \tag{3}$$

we have $t_{i,j} = \tau_{i\wedge j}$ (see e.g., the end of the proof of Theorem 16 in [2]). In other words, the times at which cutpoints appear can be recovered from the cut-tree. Observe that τ is increasing along any branch of $\mathcal{C}$.

We end this section with a result which tells how to find the connected components of $\mathcal{T} \setminus \mathcal{P}_t$ using the cut-tree.

Lemma 2.2. *Fix $t > 0$. The connected components of $\mathcal{T} \setminus \mathcal{P}_t$ are in bijection with the subtrees of $\mathcal{C}$ of the form $\mathcal{C}_x^y$, with $x, y \in \mathcal{C}$ satisfying $\tau_x = t$ and $x \prec y$. Furthermore, this bijection conserves the masses.*

Proof. Let C be a connected component of $\mathcal{T} \setminus \mathcal{P}_t$, and let $x_C \in \mathcal{C}$ be the most recent common ancestor of the leaves $\mathcal{L}_C := \{i \in \mathbb{Z}_+ : U_i \in C\}$. Notice that x_C is not necessarily a branchpoint (indeed x_C belongs to the skeleton when the component C is the same at times t and $t-$).

First let us show that $\tau_{x_C} = t$. Since $\mathcal{C}$ is a CRT, almost surely there exists a sequence of branchpoints of the form $i_n \wedge j_n$ converging to x with $i_n, j_n \in \mathcal{L}_C$. For all n, since U_{i_n} and U_{j_n} are in the same connected component of $\mathcal{T} \setminus \mathcal{P}_t$ it follows that $\tau_{i_n\wedge j_n} \geq t$ for every $n \geq 1$. Since $z \mapsto \tau_z$ is continuous, we obtain that $\tau_{x_C} \geq t$. Now assume by contradiction that $\tau_{x_C} > t$. Then, again by continuity of $z \mapsto \tau_z$ and since $\mathcal{C}$ is a CRT, there exists a branchpoint b such that $b \prec x_C$ and $\tau_b > t$. Now take two leaves $j \in \overline{\mathcal{C}}_b^{x_C}$ and $i \in \mathcal{L}_C$. We have $i \wedge j = b$ and thus $t_{i,j} = \tau_b > t$, so that $j \in \mathcal{L}_C$. This contradicts the definition of x_C. Hence, $\tau_{x_C} = t$.

Now take $i \in \mathcal{L}_C$ and set

$$\Phi(C) = \mathcal{C}^i_{x_C}.$$

Let us check that Φ is well defined by showing that for $i, j \in \mathcal{L}_C$ we have $\mathcal{C}^i_{x_C} = \mathcal{C}^j_{x_C}$. To this end, observe that by definition $x_C \prec i \wedge j$, and argue by contradiction assuming that $i \wedge j = x_C$. Then $t_{i,j} = \tau_{x_C} = t$, so that U_i and U_j do not belong to the same connected component of $\mathcal{T} \setminus \mathcal{P}_t$. Hence, i and j cannot be both in $\mathcal{L}_C$, which leads to a contradiction.

Finally, let us establish that Φ is bijective. To this end, we exhibit the reverse bijection. Consider a subtree of $\mathcal{C}$ of the form $\mathcal{C}_x^y$, with $x \in \mathcal{C}$ satisfying $\tau_x = t$ and $x \prec y$. Consider $i \in \mathbb{Z}_+$ such that $i \in \mathcal{C}_x^y$ and denote by C the connected component of $\mathcal{T} \setminus \mathcal{P}_t$ containing U_i. We set

$$\Psi(\mathcal{C}_x^y) = C.$$

The map Ψ is well defined since, if $i, j \in \mathcal{C}_x^y$ then $t_{i,j} > \tau_x = t$ so that the connected component of $\mathcal{T} \setminus \mathcal{P}_t$ containing U_i also contains U_j.

Now, if C is a connected component of $\mathcal{T} \setminus \mathcal{P}_t$, then $\Psi \circ \Phi(C) = C$. Indeed, let $i \in \mathbb{Z}_+$ be such that $U_i \in C$. By definition of Φ, $i \in \Phi(C)$. In turn by definition of Ψ, $\Psi \circ \Phi(C)$ is the connected component of $\mathcal{T} \setminus \mathcal{P}_t$ containing U_i, which is precisely C.

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Conversely, consider a subtree of $\mathcal{C}$ of the form $\mathcal{C}_x^y$, with $x \in \mathcal{C}$ satisfying $\tau_x = t$ and $x \prec y$. We check that $\Phi \circ \Psi(\mathcal{C}_x^y) = \mathcal{C}_x^y$. Let $i \in \mathbb{Z}_+$ be such that $i \in \mathcal{C}_x^y$. Then $C = \Psi(\mathcal{C}_x^y)$ is the connected component of $\mathcal{T} \setminus \mathcal{P}_t$ containing U_i. It follows that $\Phi(C) = \mathcal{C}_{x_C}^i$. In particular, $x, x_C \in [\![\rho, i]\!]$ and $\tau_x = \tau_{x_C} = t$. Since τ is increasing on $[\![\varnothing, i]\!]$, it follows that $x = x_C$, and thus $\mathcal{C}_x^y = \mathcal{C}_{x_C}^i$. This completes the proof.

The fact that this bijection conserves the masses is a consequence of the fact that $\Phi(C) = \overline{\cup_{i:U_i \in C} [\![x_C, i]\!]}$ for every connected component C of $\mathcal{T} \setminus \mathcal{P}_t$ and that $\nu = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \delta_i$ and $\mu = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \delta_{U_i}$. ∎

Recall that our main result, Theorem 1.1, consists in coupling both processes X_{AP} and X_{B}. First in Section 3 we start from the Aldous–Pitman fragmentation on a CRT and we construct an excursion-type function, which we show to be continuous and to be equal in law to a Brownian excursion, thus proving Theorem 1.1 (i). Then in Section 4 we explain how to recover the Brownian CRT with its Poissonian rain from the "Bertoin excursion," proving Theorem 1.1 (ii).

3 | DEFINING THE "BERTOIN" EXCURSION FROM THE ALDOUS–PITMAN FRAGMENTATION

Here we start from a Brownian CRT $\mathcal{T}$ equipped with a Poissonian rain $\mathcal{P}$, and using the the cut-tree $\mathcal{C}$ of $\mathcal{T}$ we construct a function F having the law of a Brownian excursion, in such a way that, for all $t \geq 0$, the nonincreasing rearrangement of the masses of the connected components of $\mathcal{T} \setminus \mathcal{P}_t$ are the same as the nonincreasing rearrangement of the lengths of the excursions of $(F(s) - ts)_{0\leq s\leq 1}$ above its running infimum.

The algorithm used to construct the function F, which we call the Pac-Man algorithm and which we define in Section 3.1, consists in exploring the cut-tree $\mathcal{C}$ from its root ρ, associating with each element $h \in [0,1]$ a final target point π_h in $\mathcal{C}$ in a surjective way, and a value $F(h)$. We then investigate the properties of this function F, showing that it is continuous (Section 3.2) and that it has the law of a Brownian excursion (Section 3.4).

3.1 | Defining an excursion-type function from the Aldous–Pitman representation

We keep the notation of Section 2.2. Here we shall construct an excursion-type function F from a Brownian CRT $\mathcal{T}$ equipped with a Poissonian rain $\mathcal{P}$ which will turn out to meet the requirements of Theorem 1.1 (i). To this end, we shall use the Brownian cut-tree $\mathcal{C}$ associated with $\mathcal{T}$ as defined in Section 2.

With every value $h \in [0,1]$ we shall associate one point π_h of $\mathcal{C}$ using a recursive procedure. It can be informally presented as follows. Imagine Pac-Man starting at the root of $\mathcal{C}$ and wanting to eat exactly an amount h of mass. It has an initial target leaf ℓ, and starts going towards this target. As soon as it encounters a point x such that the subtree $\mathcal{C}_x^\ell$ containing the target leaf has mass at most h, Pac-Man eats this subtree; if this mass was strictly less than h, then it turns out that this point was necessarily a branchpoint, and Pac-Man continues his journey in the remaining subtree equipped with a new target. Pac-Man stops when it has eaten an amount h of mass, and we denote by π_h the ending point of the process.

Given a tree $T = (T, r, \nu)$ with root r and mass measure ν, a distinguished leaf ℓ and a value $0 \leq h \leq \nu(T)$, we set

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$B(T,\ell,h)=\inf\left\{x\in [\![r,\ell]\!]\ :\ \nu\left(T_x^{\ell}\right)\le h\right\}.$$

Observe that $B(T,\ell,0)=\ell$, $B(T,\ell,\nu(T))=r$ and that for $0<h<\nu(T)$, if $B(T,\ell,h)$ is a point of the skeleton of T, then necessarily $\nu(T^{\ell}_{B(T,\ell,h)})=h$.

3.1.1 | Pac-Man algorithm

Given $h\in[0,1]$, we define a sequence $(B_i,L_i,H_i)_{0\le i<N+1}$ with $N\in\mathbb{Z}_+\cup\{+\infty\}$ as follows (we drop the dependence in h to simplify notation). First, set $B_0=\rho$, $L_0=0$, $H_0=h$. Then, by induction, if $(B_i,L_i,H_i)_{0\le i\le k}$ have been constructed, we set:

$$B_{k+1}=B\left(\mathcal{C}^{L_k}_{B_k},L_k,H_k\right),\qquad H_{k+1}=h-\nu\left(\mathcal{C}^{L_k}_{B_{k+1}}\right).$$

If $H_{k+1}=0$, we set $N=k+1$ and stop, otherwise we set $L_{k+1}=\Lambda\left(\overline{\mathcal{C}}^{L_k}_{B_{k+1}}\right)$ and continue (see Figure 2 for an illustration). In particular, observe that $h=\sum_{1\le k<N+1}\nu\left(\mathcal{C}^{L_{k-1}}_{B_k}\right)$ by construction. When $N<\infty$, we set $\pi_h=B_N$. When $N=\infty$, we define π_h as the point which is the limit $(B_i)_{i\ge0}$ (we will later see that when $N=\infty$ the point π_h is necessarily a leaf). Observe that the limit exists by compactness, since the sequence $(B_i)_{i\ge0}$ is increasing for the genealogical order.

Finally, we set

$$F(h)=\sum_{1\le k<N+1}\tau_{B_k}\cdot\nu\left(\mathcal{C}^{L_{k-1}}_{B_k}\right),\tag{4}$$

where we recall from (3) the notation τ_x for $x\in\mathcal{C}$. In order to unify the treatment, here and in the sequel, when $N=\infty$, the notation $\sum_{1\le i<N+1}$ simply means $\sum_{i=1}^{\infty}$; similarly $(\cdot)_{1\le i<N+1}$ means $(\cdot)_{i\ge1}$.

Remark 3.1. Observe that the Pac-Man algorithm is initialized with the target leaf 0, which is the "image" of $\varnothing$ in the cut-tree $\mathcal{C}$. Changing the initial target leaf amounts to rerooting the tree $\mathcal{T}$, and will produce another function F (for which the assertion of Theorem 1.1 (i) still holds).

It is rather straightforward to check that (4) defines a bounded function on [0, 1].

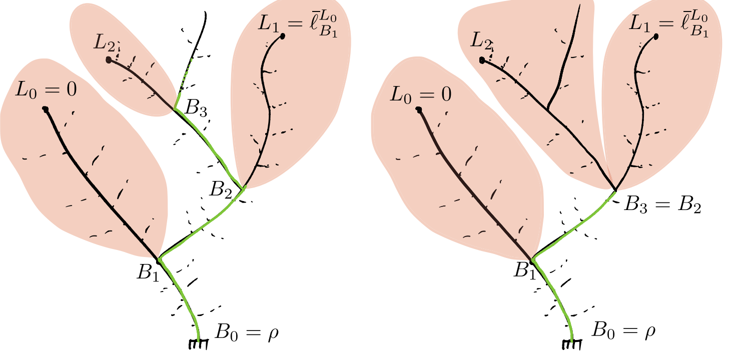

FIGURE 2 Two illustrations of the construction: Pac-Man's trajectory in the cut-tree is represented in green and the eaten subtrees in red. For example, when reaching B_1, Pac-Man eats the subtree containing L_0 (it is the first time the mass of the subtree above is less than h), and continues in direction of L_1. The value $F(h)$ is obtained by summing weighted masses of the eaten subtrees.

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Lemma 3.2. *We have* $\sup_{h\in[0,1]} F(h) \leq \mathrm{Height}(C)$.

Proof. Take $h \in [0,1]$ and keep the previous notation. To simplify notation, set $m_k = \nu(C_{B_k}^{L_{k-1}})$ for $1 \leq k < N+1$. Since B_i is an ancestor of B_j for $i < j$, for every $0 \leq i < N+1$ we have

$$\int_{[\![B_i,B_{i+1}]\!]} \frac{1}{\nu(C_z)}\lambda(\mathrm{d}z) \leq \frac{d(B_i,B_{i+1})}{\sum_{i+1\leq j<N+1} m_j}.$$

It readily follows that for every $1 \leq k < N+1$:

$$\tau_{B_k} = \int_{[\![\rho,B_k]\!]} \frac{1}{\nu(C_z)}\lambda(\mathrm{d}z) \leq \sum_{i=0}^{k-1} \frac{d(B_i,B_{i+1})}{\sum_{i+1\leq j<N+1} m_j}.$$

Thus

$$\begin{aligned} F(h) &\leq \sum_{1\leq k<N+1} m_k \sum_{i=0}^{k-1} \frac{d(B_i,B_{i+1})}{\sum_{i+1\leq j<N+1} m_j} = \sum_{0\leq i<N+1} \sum_{i+1\leq k<N+1} \frac{m_k}{\sum_{i+1\leq j<N+1} m_j} d(B_i,B_{i+1}) \\ &= \sum_{0\leq i<N+1} d(B_i,B_{i+1}), \end{aligned}$$

and the desired conclusion follows. ∎

Remark 3.3. Given the rescaled convergence of discrete cut-trees to continuous cut-trees [13], it is natural to expect that when one equips the discrete cut-trees with labels corresponding to cutting times, a joint convergence holds towards C equipped with the labeling $(\tau_x)_{x\in C}$. It is interesting to notice that in the discrete case, given the cut-tree, the labeling is random (see [30, Section 3]), while in the continuous case, given C, the labeling $(\tau_x)_{x\in C}$ is deterministic.

3.2 | Continuity of the function F

We now prove that the function F constructed this way is a.s. continuous. To this end, we rely on the fact that C is distributed as a Brownian CRT, comparing as in Lemma 3.2 values of F with distances in C.

Proposition 3.4. *Almost surely, F is continuous on* $[0,1]$.

In order to establish the continuity of the function F, it is useful to define a "backward" construction. Fix $x \in C$. We define a sequence $(B_i, L_i)_{0\leq i<N+1}$ with $N \in \mathbb{Z}_+ \cup \{+\infty\}$ as follows (we drop the dependence in x to simplify notation). Set $B_0 = \rho, L_0 = 0$. Then, by induction, if $(B_i, L_i)_{0\leq i\leq k}$ have been constructed, we define B_{k+1} by:

$$[\![B_k, x]\!] \cap [\![B_k, L_k]\!] = [\![B_k, B_{k+1}]\!], \qquad L_{k+1} = \begin{cases} \Lambda\left(\overline{C}_{B_{k+1}}^{L_k}\right) & \text{if } B_{k+1} \text{ is a branchpoint} \\ B_{k+1} & \text{otherwise} \end{cases}$$

If $x = B_{k+1}$ we set $N = k+1$ and stop, otherwise we continue. We say that $(B_i, L_i)_{0\leq i<N+1}$ is the *record sequence* associated with x (see Figure 3 for an illustration).

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

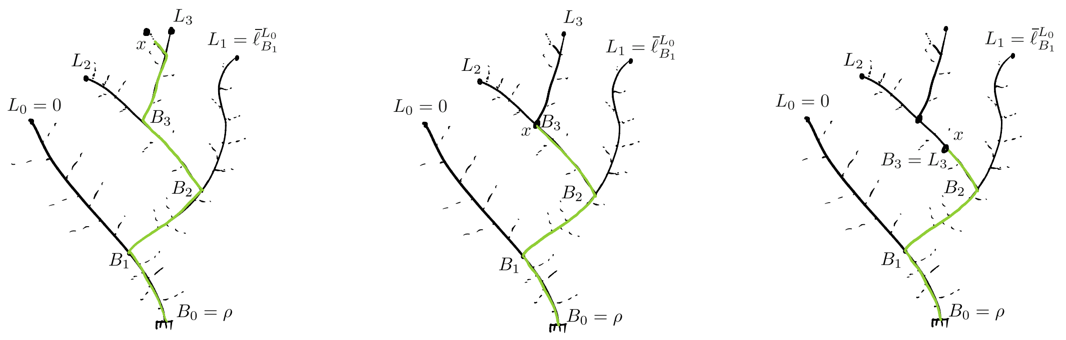

FIGURE 3 Illustration of the backward construction: on the left, x is a leaf and the record sequence associated with x is infinite; in the middle, x is a branchpoint and $(B_i, L_i)_{0\leq i\leq 3}$ is the associated record sequence; on the right, x belongs to the skeleton and $(B_i, L_i)_{0\leq i\leq 3}$ is the associated record sequence.

The next result characterizes points of $\mathcal{C}$ whose record sequence is infinite.

Lemma 3.5. *When $N = \infty$, x is a leaf, and $(B_i)_{i\geq 0}$ converges to x.*

Proof. First, let us show that if x is not a leaf, then $N < \infty$. It suffices to show that $N < \infty$ when x is a branchpoint (this will entail that $N < \infty$ when $x \in \mathrm{Sk}(\mathcal{C})$ by considering a descendent of x which is a branchpoint). To this end, recall from Section 2.2 that x can be written as $i \wedge j$ for some $i, j \geq 1$. Let $c \in [\![U_i, U_j]\!]$ be the cutpoint appearing at time $t_{i,j}$ (observe that c is the first cutpoint falling on $[\![U_i, U_j]\!]$). Let $(c_p)_{1\leq p\leq M}$ be the cutpoints falling on $[\![\emptyset, U_i]\!] \cup [\![\emptyset, U_j]\!]$ before time $t_{i,j}$, ordered by their time of appearance (observe that almost surely $M < \infty$, and that all these points fall on $[\![\emptyset, U_i \wedge U_j]\!]$ except for c). We construct a subsequence $(c_{p_k})_{1\leq k\leq N}$ of cutpoints precisely corresponding to $(B_k)_{1\leq i\leq N}$ in $\mathcal{C}$ as follows. Set $p_1 = 1$. If $c = c_1$, we set $N = 1$. Otherwise, assuming that $(p_k)_{1\leq k\leq n}$ has been defined, set $p_{n+1} = \min\{p > p_n : c_{p_n} \prec c_p\}$; if $c_{p_{n+1}} = c$ we set $N = p_{n+1}$ and stop, otherwise we continue. It is clear that, for all $1 \leq k \leq N$, $c_{p_k} = B_k$. Furthermore, since $M < \infty$, this procedure eventually stops, thus showing that $N < \infty$.

Next, argue by contradiction and assume that $N = \infty$, x is a leaf and $(B_k)_{k\geq 0}$ converges to a point $u \in \mathcal{C}$ with $u \neq x$ (observe that $(B_k)_{k\geq 0}$ always converges as it is increasing). Since $B_k \prec x$ for every $k \geq 0$, we have $u \prec x$. Choose a branchpoint $b \in \mathcal{C}$ such that $u \prec b \prec x$. Then the record sequence associated with b has infinitely many terms, in contradiction with the previous paragraph. ∎

When x is a branchpoint (then $N < \infty$), we set:

$$\ell(x) = L_{N-1}, \qquad \overline{\ell}(x) = L_N. \tag{5}$$

In terms of the Pac-Man construction, these leaves can be interpreted as follows: when passing at x during its journey, if possible, the Pac-Man eats the subtree above x containing $\ell(x)$ and continues towards $\overline{\ell}(x)$; otherwise it continues towards $\ell(x)$.

We also define for every $x \in \mathcal{C}$:

$$h_1(x) = \sum_{1\leq k<N+1} \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right). \tag{6}$$

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Then, by Lemma 3.5, for every $x \in \mathcal{C}$, if one takes $h = h_1(x)$ in the Pac-Man construction, one precisely gets the sequence $(B_k, L_k)_{0 \le k < N+1}$ with $\pi_{h_1} = x$.

The following result is an immediate consequence of the definition:

Lemma 3.6. *For every branchpoint B, h_1 is decreasing on $[\![B, \ell(B)]\!]$.*

When x is a branchpoint, we also set

$$h_2(x) = \sum_{1 \le k \le N+1} \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right); \tag{7}$$

observe that if one takes $h = h_2(x)$ in the Pac-Man construction, we also get $\pi_{h_2} = x$ (see Figure 4 for an illustration). Also note that $h_2(x) = h_1(x) + \nu(\mathcal{C}_x^{\overline{\ell}(x)})$.

Finally, we define

$$\mathbb{h}_2 = \{h_2(x) : x \in \mathcal{B}(\mathcal{C})\}, \tag{8}$$

where we recall that $\mathcal{B}(\mathcal{C})$ denotes the set of all branchpoints of $\mathcal{C}$, and, to simplify notation, when $x \in \mathcal{C}$ is not a leaf, we set

$$h_0(x) = \begin{cases} h_1(x) - \nu(\mathcal{C}_x) & \text{if } x \in \mathrm{Sk}(\mathcal{C}) \\ h_1(x) - \nu\left(\mathcal{C}_x^{\ell(x)}\right) & \text{if } x \in \mathcal{B}(\mathcal{C}). \end{cases} \tag{9}$$

The following result is also an immediate consequence of the definition of the Pac-Man construction, which we record for future use.

Lemma 3.7. *Fix $x \in \mathcal{C}$.*

- *If $x \in \mathrm{Sk}(\mathcal{C})$, then for every $h \in (h_0(x), h_1(x)]$, the final target point π_h belongs to $\mathcal{C}_x$. Conversely, each element of $\mathcal{C}_x$ is the final target point of an element $h \in (h_0(x), h_1(x)]$.*
- *If $x \in \mathcal{B}(\mathcal{C})$, then for every $h \in (h_0(x), h_1(x)]$, the final target point π_h belongs to $\mathcal{C}_x^{\ell(x)}$ and for every $h \in [h_1(x), h_2(x)]$, π_h belongs to $\mathcal{C}_x^{\overline{\ell}(x)}$. Conversely, every point of $\mathcal{C}_x^{\ell(x)}$ is the final target point of some $h \in (h_0(x), h_1(x)]$, and every point of $\mathcal{C}_x^{\overline{\ell}(x)}$ is the final target point of some $h \in [h_1(x), h_2(x)]$.*

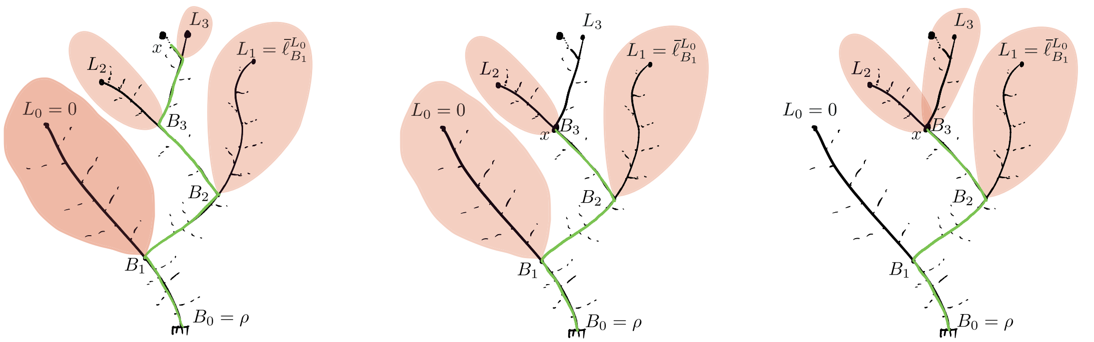

FIGURE 4 Illustration of the definitions of $h_1(x)$ and $h_2(x)$: on the left, x is a leaf and $h_1(x)$ is the sum of the masses of the red subtrees (there are infinitely many of them); in the middle, x is a branchpoint and $h_1(x)$ is the sum of the masses of the three red subtrees; on the right, x is the same branchpoint as in the middle and $h_2(x)$ is the sum of the masses of the four red subtrees.

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Before proving the continuity of F, we gather some preparatory lemmas.

Lemma 3.8. *Let $\ell \in \mathcal{C}$ be a leaf. Then*

$$\nu(\mathcal{C}_x) \cdot \tau_x \quad \to_{x \to \ell} \quad 0.$$

Proof. Write

$$\nu(\mathcal{C}_x) \cdot \tau_x = \nu(\mathcal{C}_x) \cdot \int_{[\![\rho,x]\!]} \frac{1}{\nu(\mathcal{C}_z)} \lambda(\mathrm{d}z) = \int_{[\![\rho,\ell]\!]} \frac{\nu(\mathcal{C}_x)}{\nu(\mathcal{C}_z)} \mathbb{1}_{z \in [\![\rho,x]\!]} \lambda(\mathrm{d}z).$$

Then observe that the quantity $\nu(\mathcal{C}_x)/\nu(\mathcal{C}_z)\mathbb{1}_{z \in [\![\rho,x]\!]}$ is bounded by 1 and tends to 0 since $\nu(\mathcal{C}_x) \to 0$ as $x \to \ell$. The conclusion follows by dominated convergence. ∎

The next lemma compares the difference between two values of h with masses of subtrees in the cut-tree $\mathcal{C}$.

Lemma 3.9. *Take $x \prec y$ in $\mathcal{C}$ with $x \in \mathrm{Sk}(\mathcal{C})$ and assume that $\pi_{h'} = y$. Then:*

(i) when $y \in \mathrm{Sk}(\mathcal{C})$ we have $|h_1(x) - h'| \leq \nu(\mathcal{C}_x \setminus \mathcal{C}_y)$.
(ii) $|h_1(x) - h'| \leq \nu(\mathcal{C}_x)$;

Proof. Let $(B_i^1, L_i^1)_{0 \leq i < N_1+1}$ and respectively $(B_i^2, L_i^2)_{0 \leq i < N_2+1}$ be the record sequences associated with x and y. Since $x \prec y$, we have $N_1 \leq N_2$. Also, x is not a leaf, so that $N_1 < \infty$ and $x = B_{N_1}^1$. Observe that we may have $B_{N_1}^1 \neq B_{N_1}^2$ (e.g., if $x \in \mathrm{Sk}(\mathcal{C})$). Then $(B_i^1, L_i^1)_{0 \leq i < N_1} = (B_i^2, L_i^2)_{0 \leq i < N_1}$ and $B_{N_1}^2 \in [\![x, L_{N_1-1}^1]\!]$,

Recall the notation $\mathbb{h}_2$ from (8). Then

$$h_1(x) = \sum_{1 \leq k < N_1+1} \nu\left(\mathcal{C}_{B_k^1}^{L_{k-1}^1}\right) \quad \text{and} \quad h' = \sum_{1 \leq k < N_2+1} \nu\left(\mathcal{C}_{B_k^2}^{L_{k-1}^2}\right) + \mathbb{1}_{h' \in \mathbb{h}_2} \nu\left(\mathcal{C}_{B_{N_2}^2}^{L_{N_2}^2}\right),$$

so that

$$h_1(x) - h' = \sum_{N_1 \leq k < N_2+1} \nu\left(\mathcal{C}_{B_k^2}^{L_{k-1}^2}\right) + \mathbb{1}_{h' \in \mathbb{h}_2} \nu\left(\mathcal{C}_{B_{N_2}^2}^{L_{N_2}^2}\right) - \nu\left(\mathcal{C}_{B_{N_1}^1}^{L_{N_1-1}^1}\right).$$

Now define

$$\mathcal{A} = \mathcal{C}_{B_{N_1}^1}^{L_{N_1-1}^1}, \qquad \mathcal{B} = \bigcup_{N_1 \leq k < N_2+1} \mathcal{C}_{B_k^2}^{L_{k-1}^2}.$$

Observe that the union defining $\mathcal{B}$ is disjoint. When $y \in \mathrm{Sk}(\mathcal{C})$ we have $h' \notin \mathbb{h}_2$, $\mathcal{A} \setminus \mathcal{B} \subset \mathcal{C}_x \setminus \mathcal{C}_y$. When y is a branchpoint, setting $\mathcal{B}' = \mathcal{B} \cup \mathcal{C}_{B_N^2}^{L_N^2}$, observe that this union is disjoint and $\mathcal{A} \setminus \mathcal{B} \subset \mathcal{C}_x$. The conclusion follows. ∎

The next lemma bounds from above the difference between two values of F.

Lemma 3.10. *Take $x \prec y$ in $\mathcal{C}$ and $h, h' \in [0,1]$ such that $\pi_h = x$ and $\pi_{h'} = y$. Then*

$$|F(h) - F(h')| \leq \tau_x |h - h'| + d_{\mathcal{C}}(x, y).$$

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Proof. We keep the notation introduced in the beginning of Lemma 3.9 (in particular it may be helpful to also refer to Figure 5): we denote by $(B_i^1, L_i^1)_{0\le i<N_1+1}$ and respectively $(B_i^2, L_i^2)_{0\le i<N_2+1}$ the record sequences associated with x and y, so that $N_1 \le N_2$, x is not a leaf, $N_1 < \infty$, $x = B^1_{N_1}$, $(B_i^1, L_i^1)_{0\le i<N_1} = (B_i^2, L_i^2)_{0\le i<N_1}$ and $B^2_{N_1} \in [\![x, L^1_{N_1-1}]\!]$. Recall the definitions of F in (4) and of h_1, h_2 in (6), (7).

We have

$$h = \sum_{1\le k<N_1+1} \nu\left(C_{B_k^1}^{L_{k-1}^1}\right) + \mathbb{1}_{h\in \mathbb{h}_2}\nu\left(C_{B_{N_1}^1}^{L_{N_1}^1}\right), \quad F(h) = \sum_{1\le k<N_1+1} \tau_{B_k^1}\nu\left(C_{B_k^1}^{L_{k-1}^1}\right) + \mathbb{1}_{h\in \mathbb{h}_2}\tau_{B_{N_1}^1}\nu\left(C_{B_{N_1}^1}^{L_{N_1}^1}\right),$$

and

$$h' = \sum_{1\le k<N_2+1} \nu\left(C_{B_k^2}^{L_{k-1}^2}\right) + \mathbb{1}_{h'\in \mathbb{h}_2}\nu\left(C_{B_{N_2}^2}^{L_{N_2}^2}\right), \quad F(h') = \sum_{1\le k<N_2+1} \tau_{B_k^2}\nu\left(C_{B_k^2}^{L_{k-1}^2}\right) + \mathbb{1}_{h'\in \mathbb{h}_2}\tau_{B_{N_2}^2}\nu\left(C_{B_{N_2}^2}^{L_{N_2}^2}\right).$$

Thus, setting

$$m_k = \nu\left(C_{B_k^2}^{L_{k-1}^2}\right) \text{ for } k < N_2, \qquad m_{N_2} = \nu\left(C_{B_{N_2}^2}^{L_{N_2-1}^2}\right) + \mathbb{1}_{h'\in \mathbb{h}_2}\nu\left(C_{B_{N_2}^2}^{L_{N_2}^2}\right)$$

with the convention that $m_{N_2} = 0$ if $N_2 = \infty$ and remembering that $x = B^1_{N_1}$, we have

$$F(h) - F(h') = \tau_x \nu\left(C_{B_{N_1}^1}^{L_{N_1-1}^1}\right) + \mathbb{1}_{h\in \mathbb{h}_2}\tau_x\nu\left(C_{B_{N_1}^1}^{L_{N_1}^1}\right) - \sum_{N_1\le k<N_2+1} \tau_{B_k^2} m_k,$$

and

$$h - h' = \nu\left(C_{B_{N_1}^1}^{L_{N_1-1}^1}\right) + \mathbb{1}_{h\in \mathbb{h}_2}\nu\left(C_{B_{N_1}^1}^{L_{N_1}^1}\right) - \sum_{N_1\le k<N_2+1} m_k.$$

It follows that

$$F(h) - F(h') = \tau_x(h - h') + \sum_{N_1\le k<N_2+1} (\tau_x - \tau_{B_k^2}) m_k. \tag{10}$$

In particular,

$$|F(h) - F(h')| \le \tau_x |h - h'| + \sum_{N_1\le k<N_2+1} (\tau_{B_k^2} - \tau_x) m_k.$$

To control the sum, we perform an Abel transformation by setting $\sigma_k = \tau_{B_k^2} - \tau_{B_{k-1}^2}$ for $N_1 < k < N_2 + 1$ and $\sigma_{N_1} = \tau_{B^2_{N_1}} - \tau_x$. Then

$$\sum_{N_1\le k<N_2+1} (\tau_{B_k^2} - \tau_x)\cdot m_k = \sum_{N_1\le k<N_2+1} \sum_{i\in [\![N_1,k]\!]} \sigma_i \cdot m_k = \sum_{N_1\le i<N_2+1} \sigma_i \sum_{k\in [\![i,N_2+1[\![} m_k.$$

Then observe that, as in the proof of Lemma 3.9, for every $N_1 \le i < N_2 + 1$ we have $\sum_{k\in [\![i,N_2+1[\![} m_k \le \nu\left(C_{B_i^2}\right)$. Also, $\sigma_{N_1} \le d_C\left(x, B^2_{N_1}\right)/\nu\left(C_{B^2_{N_1}}\right)$ and $\sigma_i \le d_C(B^2_{i-1}, B^2_i)/\nu\left(C_{B_i^2}\right)$ for $N_1 < i < N_2 + 1$. The conclusion follows. ■

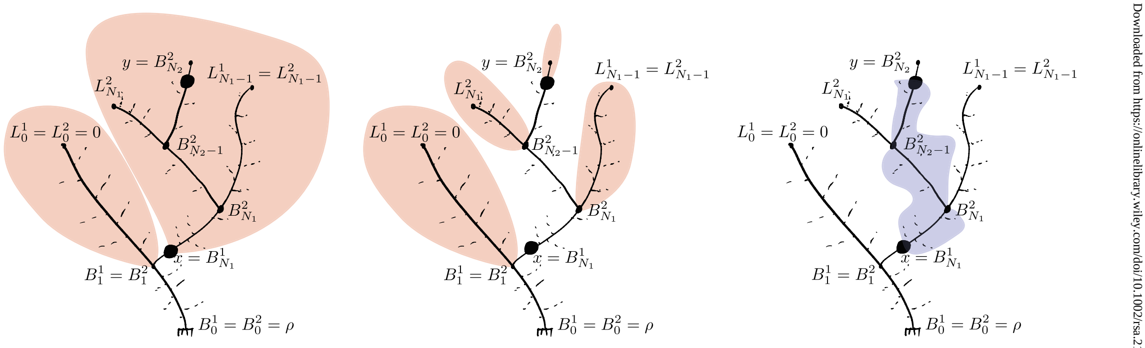

FIGURE 5 Illustration of the proof of Lemma 3.9 when $x, y \in \mathrm{Sk}(C)$. Here $N_1 = 2$ and $N_2 = 4$. Left: The sum of the masses of the two red subtrees is $h_1(x)$. Middle: The sum of the masses of the four red subtrees is $h_1(y)$. Right: The difference $|h_1(x) - h'|$ is at most the mass of the blue subtree $\mathcal{A} \setminus \mathcal{B}$.

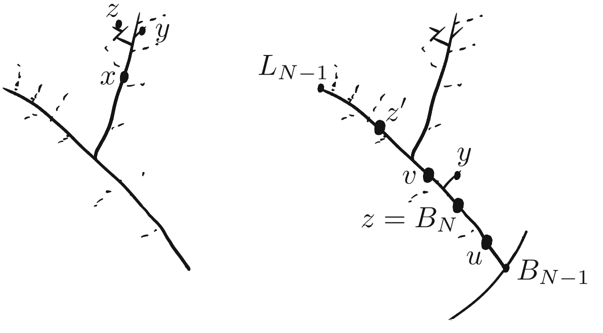

FIGURE 6 Illustration of the first case when z is a leaf (left) and the second case where $z \in \mathrm{Sk}(C)$ (right).

We can now prove the continuity of the function F.

Proof of Proposition 3.4. Fix $h \in [0, 1]$. We want to prove that F is continuous at h. We distinguish several cases according to the nature of the final target point $z := \pi_h$.

- ⋆ *First case*: z is a leaf (see Figure 6, left). Fix $\varepsilon > 0$. Using in particular Lemma 3.8 and the fact that C is a Brownian CRT, we may choose $x \in \mathrm{Sk}(C)$ such that $x \prec z$ and $\tau_x\ \nu(C_x) \le \varepsilon$, $\nu(C_x) \le \varepsilon$ and $\mathrm{Diam}(C_x) < \varepsilon$. Observe that $h_1(x) - \nu(C_x) < h < h_1(x)$ by Lemma 3.7. Then, by Lemmas 3.9 (ii) and 3.10, we have

$$|F(h_1(x)) - F(h)| \le \tau_x |h_1(x) - h| + d_C(x, z) \le \tau_x \nu(C_x) + \varepsilon \le 2\varepsilon.$$

Next, take $h' \in (h_1(x) - \nu(C_x), h_1(x))$. By Lemma 3.7, there exists $y \in C_x$ such that $\pi_{h'} = y$. Then, as before, again by Lemmas 3.9 (ii) and 3.10, we have

$$|F(h_1(x)) - F(h')| \le \tau_x |h_1(x) - h'| + d_C(x, z) \le \tau_x \nu(C_x) + \varepsilon \le 2\varepsilon.$$

We conclude that $|F(h) - F(h')| \le 4\varepsilon$.

- ⋆ *Second case*: $z \in \mathrm{Sk}(C)$ (see Figure 6, right). Fix $\epsilon > 0$. Let $(B_i, L_i)_{0 \le i < N+1}$ be the record sequence associated with z, so that $z = B_N$ and $z \in]\!] B_{N-1}, L_{N-1} [\![$. Fix $z' \in C$ such that $z \prec z' \prec L_{N-1}$. Take $\varepsilon > 0$. Then choose $u, v \in \mathrm{Sk}(C)$ such that $B_{N-1} \prec u \prec$

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$z \prec v \prec z'$ such that $\nu(\mathcal{C}_u \setminus \mathcal{C}_v) \le \varepsilon/\tau_{z'}$ and $\mathsf{Diam}(\mathcal{C}_u \setminus \mathcal{C}_v) \le \varepsilon$. Then, by Lemmas 3.9 (i) and 3.10, we have

$$|F(h_1(u)) - F(h)| \le \tau_u|h_1(u) - h| + d_{\mathcal{C}}(u, z) \le \tau_{z'}\nu(\mathcal{C}_u \setminus \mathcal{C}_z) + \varepsilon \\ \le \tau_{z'}\nu(\mathcal{C}_u \setminus \mathcal{C}_v) + \varepsilon \le 2\varepsilon.$$

Next, observe that by Lemma 3.6 we have $h_1(v) < h < h_1(u)$. Take $h' \in (h_1(v), h_1(u))$. By Lemma 3.7, since $v \in \mathcal{C}_u$, we have $h_1(u)-\nu(\mathcal{C}_u) < h_1(v)$, so we have $h_1(u)-\nu(\mathcal{C}_u) < h' < h_1(u)$.

By Lemma 3.7, there exists $y \in \mathcal{C}_u$ such that $\pi_{h'} = y$. Then, again by Lemma 3.10, we have

$$|F(h_1(u)) - F(h')| \le \tau_u|h_1(u) - h'| + d_{\mathcal{C}}(u, y) \le \tau_{z'}(h_1(u) - h_1(v)) + \varepsilon \\ \le \tau_{z'}\nu(\mathcal{C}_u \setminus \mathcal{C}_v) + \varepsilon \le 2\varepsilon,$$

where we have also used Lemma 3.9 (i) to write $h_1(u)-h_1(v) \le \nu(\mathcal{C}_u\setminus\mathcal{C}_v)$. We conclude that $|F(h) - F(h')| \le 4\varepsilon$.

★ *Third case*: z is a branchpoint. Let $(B_i, L_i)_{0\le i<N+1}$ be the record sequence associated with z with $N < \infty$. We consider two subcases.

★★ *First subcase*: h is of the form $h = h_1(z)$ (see Figure 7). We first show that F is right-continuous at h. Fix $\epsilon > 0$. Choose a point $u \in \mathrm{Sk}(\mathcal{C})$ such that $z \prec u \prec L_N$ and $\nu(\mathcal{C}_u) \le \varepsilon/\tau_u$ and $\mathsf{Diam}(\mathcal{C}_u) < \varepsilon$. By definition of the Pac-Man construction, $h_1(u) = h + \nu(\mathcal{C}_u)$ and $F(h_1(u)) = F(h) + \tau_u\nu(\mathcal{C}_u)$. In particular, $|F(h_1(u)) - F(h)| \le \varepsilon$.

Now take $h' \in (h, h + \nu(\mathcal{C}_u))$. Since $h_1(u) - \nu(\mathcal{C}_u) < h' < h_1(u)$, by Lemma 3.7, there exists $y \in \mathcal{C}_u$ such that $\pi_{h'} = y$. Then, as before, by Lemma 3.10, we have

$$|F(h_1(u)) - F(h')| \le \tau_y|h_1(u) - h'| + d_{\mathcal{C}}(u, y) \le \tau_u\nu(\mathcal{C}_u) + \varepsilon \le 2\varepsilon.$$

We conclude that $|F(h) - F(h')| \le 3\varepsilon$.

Let us next show that F is left-continuous at h. Fix a point $v \in \mathrm{Sk}(\mathcal{C})$ such that $z \prec v \prec L_{N-1}$. Take $\varepsilon > 0$ and choose a point $u \in \mathrm{Sk}(\mathcal{C})$ such that $z \prec u \prec v$ and $\nu(\mathcal{C}_z^{L_{N-1}} \setminus \mathcal{C}_u) \le \varepsilon/\tau_v$ and $\mathsf{Diam}(\mathcal{C}_z^{L_{N-1}} \setminus \mathcal{C}_u) < \varepsilon$. Observe that by definition of the Pac-Man construction, $h_1(u) + \nu(\mathcal{C}_z^{L_{N-1}} \setminus \mathcal{C}_u) = h$. Take

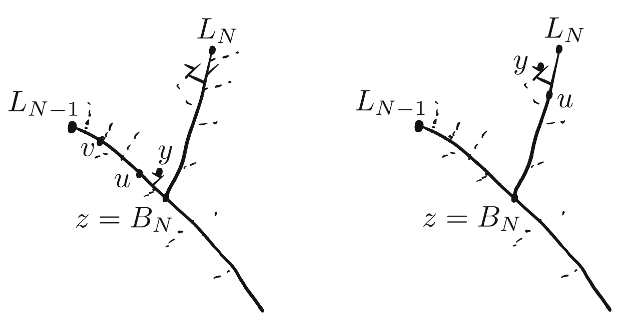

FIGURE 7 Illustration of the case $h = h_1(z)$. Left: Proof of the left-continuity; Right: Proof of the right-continuity.

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

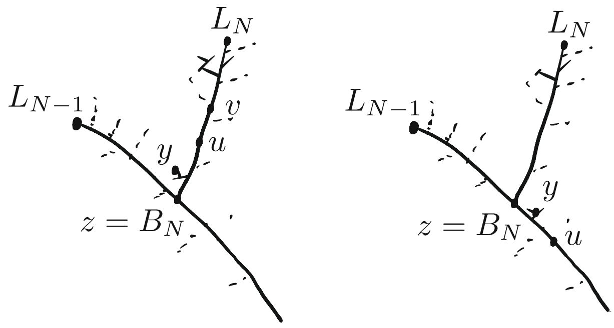

FIGURE 8 Illustration of the case $h = h_2(z)$. Left: Proof of the left-continuity; Right: Proof of the right-continuity.

$h' \in (h_1(u), h)$. By Lemma 3.7, there exists $y \in \mathcal{C}_z^{L_{N-1}} \setminus \mathcal{C}_u$ such that $\pi_{h'} = y$. Then, as before, by Lemma 3.10, we have

$$|F(h) - F(h')| \leq \tau_z |h - h'| + d_{\mathcal{C}}(z, y) \leq \tau_v \nu(\mathcal{C}_z^{L_{N-1}} \setminus \mathcal{C}_u) + \varepsilon \leq 2\varepsilon.$$

★★ *Second subcase*: h is of the form $h = h_2(z)$ (see Figure 8).

Let us first show that F is left-continuous at h. Fix a point $v \in \mathrm{Sk}(\mathcal{C})$ such that $z \prec v \prec L_N$. Take $\varepsilon > 0$ and choose a point $u \in \mathrm{Sk}(\mathcal{C})$ such that $z \prec u \prec v$, $\nu(\mathcal{C}_z^{L_N} \setminus \mathcal{C}_u) \leq \varepsilon/\tau_v$ and $\mathsf{Diam}(\mathcal{C}_z^{L_N} \setminus \mathcal{C}_u) < \varepsilon$. Observe that by definition of the Pac-Man construction, $h_1(u) + \nu(\mathcal{C}_z^{L_{N-1}} \setminus \mathcal{C}_u) = h$. Take $h' \in (h_1(u), h)$. By Lemma 3.7, there exists $y \in \mathcal{C}_z^{L_{N-1}} \setminus \mathcal{C}_u$ such that $\pi_{h'} = y$. Then, as before, by Lemma 3.10, we have

$$|F(h) - F(h')| \leq \tau_z |h_1(z) - h'| + d_{\mathcal{C}}(z, y) \leq \tau_v \nu(\mathcal{C}_z^{L_{N-1}} \setminus \mathcal{C}_u) + \varepsilon \leq 2\varepsilon.$$

Let us next show that F is right-continuous at h. Take $\varepsilon > 0$ and fix a point $u \in \mathrm{Sk}(\mathcal{C})$ such that $B_{N-1} \prec u \prec z$, $\nu(\mathcal{C}_u \setminus \mathcal{C}_z) \leq \varepsilon/\tau_z$ and $\mathsf{Diam}(\mathcal{C}_u \setminus \mathcal{C}_z) < \varepsilon$. Observe that by definition of the Pac-Man construction, $h_1(u) = h + \nu(\mathcal{C}_u \setminus \mathcal{C}_z)$. Take $h' \in (h, h_1(u))$. By Lemma 3.7, there exists $y \in \mathcal{C}_u \setminus \mathcal{C}_z$ such that $\pi_{h'} = y$. Then, as before, by Lemma 3.10, we have

$$|F(h) - F(h')| \leq \tau_u |h - h'| + d_{\mathcal{C}}(u, y) \leq \tau_z \nu(\mathcal{C}_u \setminus \mathcal{C}_z) + \varepsilon \leq 2\varepsilon.$$

This completes the proof. ■

3.3 | The function F codes the Aldous–Pitman fragmentation of the CRT

We have constructed a continuous excursion-type function F from the Aldous–Pitman fragmentation of the Brownian CRT $\mathcal{T}$. In order to prove Theorem 1.1 (i), before showing that F is in law the Brownian excursion, we first show the following result.

Proposition 3.11. *A.s. for every $t \geq 0$, the nonincreasing rearrangement of the masses of the connected components of $\mathcal{T} \setminus \mathcal{P}_t$ is the same as the nonincreasing rearrangement of the lengths of the excursions of $(F(h) - th)_{0 \leq h \leq 1}$ above its running infimum.*

To simplify notation, set $F_t(h) = F(h) - th$ for $0 \leq h \leq 1$.

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Lemma 3.12. *Take $x \prec y$ in $\mathcal{C}$ and $h, h' \in [0,1]$ such that $\pi_h = x$ and $\pi_{h'} = y$. Then $F_{\tau_x}(h') > F_{\tau_x}(h)$.*

Proof. This readily follows from the identity (10) appearing in the proof of Lemma 3.10. Indeed, keeping the same notation, if $x = B^1_{N_1} = B^2_{N_1}$ then $N_2 > N_1$ and $\tau_x < \tau_{B^2_k}$ for every $N_1 < k < N_2 + 1$ since x is an ancestor of B^2_k. If $x = B^1_{N_1} \neq B^2_{N_1}$ then similarly $\tau_x < \tau_{B^2_{N_1}}$. ■

Proof of Proposition 3.11. By Lemma 2.2, for every $t \geq 0$, the connected components of $\mathcal{T} \setminus \mathcal{P}_t$ are in bijection with subtrees of $\mathcal{C}$ of the form $\mathcal{C}_x^{\ell(x)}$ or $\mathcal{C}_x^{\overline{\ell}(x)}$ for $x \in \mathcal{C}$ with $\tau_x = t$; and this bijection conserves the masses. If C is a connected component, recall that we denote by $\Phi(C)$ the corresponding subtree of $\mathcal{C}$ (in particular $\nu(\Phi(C)) = \mu(C)$).

Let C be a connected component of $\mathcal{T} \setminus \mathcal{P}_t$ and $x \in \mathcal{C}$ with $\tau_x = t$. Observe that then x is not a leaf; we denote by $(B_i, L_i)_{0 \leq i \leq N}$ the record sequence of x. Recall from (9) the definition of $h_0(x)$.

We claim that $F_t(h_0(x)) = F_t(h_1(x))$ and that

$$\forall h \in (h_0(x), h_1(x)), \quad F_t(h) > F_t(h_1(x)); \qquad \forall h \in (0, h_0(x)), \quad F_t(h) > F_t(h_1(x)). \tag{11}$$

This implies that when $x \in \mathrm{Sk}(\mathcal{C})$, we have an excursion of length $\nu(\mathcal{C}_x)$ of F_t above its running infimum, and when x is a branchpoint we have an excursion of length $\nu(\mathcal{C}_x^{\ell(x)})$ of F_t above its running infimum.

To check that $F_t(h_0(x)) = F_t(h_1(x))$, observe that by definition

$$x = B_N, \quad t = \tau_{B_N}, \quad h_1(x) = \sum_{1 \leq k \leq N} \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right), \quad F(h_1(x)) = \sum_{1 \leq k \leq N} \tau_{B_k} \cdot \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right),$$

so

$$F_t(h_1(x)) = \sum_{1 \leq k \leq N} \tau_{B_k} \cdot \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right) - \tau_{B_N} \sum_{1 \leq k \leq N} \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right).$$

In addition, observe that

$$h_0(x) = \sum_{1 \leq k \leq N-1} \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right) = h_1(B_{N-1}).$$

It follows that $\pi_{h_0(x)} = B_{N-1}$, and

$$F(h_1(B_{N-1})) = \sum_{1 \leq k \leq N-1} \tau_{B_k} \cdot \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right).$$

Hence

$$F_t(h_0(x)) = F_t(h_1(B_{N-1})) = \sum_{1 \leq k \leq N-1} \tau_{B_k} \cdot \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right) - \tau_{B_N} \sum_{1 \leq k \leq N-1} \nu\left(\mathcal{C}_{B_k}^{L_{k-1}}\right) = F_t(h_1(x)).$$

The first inequality in (11) readily follows from Lemma 3.12, since the final target point π_h of any element $h \in (h_0(x), h_1(x))$ belongs to $\mathcal{C}_x \setminus \{x\}$ by Lemma 3.7.

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To establish the second inequality in (11), take $h \in (0, h_0(x))$. Let $1 \leq m \leq N-1$ be such that

$$\sum_{1\leq k\leq m-1} \nu\left(C_{B_k}^{L_{k-1}}\right) \leq h < \sum_{1\leq k\leq m} \nu\left(C_{B_k}^{L_{k-1}}\right) = h_1(B_m).$$

Then $\pi_h \in C_{B_m}^{L_{m-1}}$, so that by Lemma 3.12 we have $F_{\tau_{B_m}}(h) \geq F_{\tau_{B_m}}(h_1(B_m))$ (with the inequality being strict if $\pi_h \neq B_m$). Since $t \geq \tau_{B_m}$, this entails $F_t(h) \geq F_t(h_1(B_m))$. It remains to note that

$$F_t(h_1(B_m)) \geq F_t(h_1(x))$$

with the inequality being strict if $m < N-1$. Indeed,

$$F_t(h_1(B_m)) = \sum_{1\leq k\leq m} \tau_{B_k} \cdot \nu\left(C_{B_k}^{L_{k-1}}\right) - \tau_{B_N} \sum_{1\leq k\leq m} \nu\left(C_{B_k}^{L_{k-1}}\right),$$

so

$$F_t(h_1(B_m)) - F_t(h_1(x)) = \tau_{B_N} \sum_{m+1\leq k\leq N-1} \nu\left(C_{B_k}^{L_{k-1}}\right) - \sum_{m+1\leq k\leq N-1} \tau_{B_k} \cdot \nu\left(C_{B_k}^{L_{k-1}}\right),$$

which entails the result since $\tau_{B_N} > \tau_{B_k}$ for $m+1 \leq k \leq N-1$. Since one of the two inequalities $F_t(h) \geq F_t(h_1(B_m))$ or $F_t(h_1(B_m)) \geq F_t(h_1(x))$ is strict (because $h < h_0(x)$ so $\pi_h \neq B_{N-1}$), the second inequality in (11) follows.

To finish the proof, assume that x is a branchpoint. We claim that $F_t(h_1(x)) = F_t(h_2(x))$ and that

$$\forall h \in (h_1(x), h_2(x)), \qquad F_t(h) > F_t(h_1(x)). \tag{12}$$

This implies that we have an excursion of length $\nu\left(C_x^{\overline{\ell}(x)}\right)$ of F_t above its running infimum.

The fact that $F_t(h_1(x)) = F_t(h_2(x))$ is proved exactly in the same way as the identity $F_t(h_0(x)) = F_t(h_1(x))$, by using the definition of F_t. Finally, (12) follows from Lemma 3.12 by observing that when $h \in (h_1(x), h_2(x))$, π_h belongs to $C_x \setminus \{x\}$. ∎

We establish the following result for future use.

Lemma 3.13. *The probability measure ν on C is the push-forward of the Lebesgue measure on $[0,1]$ by $h \mapsto \pi_h$.*

Proof. Recall that for $x \in C$ the tree C_x is the set of all (weak) descendents of x in C. First, observe that $\{C_x : x \in \mathrm{Sk}(C)\}$ is a generating π-system of C. Hence, if two probability measures ν and $\tilde{\nu}$ supported on the set of leaves of C satisfy

$$\forall x \in \mathrm{Sk}(C), \qquad \nu(C_x) = \tilde{\nu}(C_x),$$

then $\nu = \tilde{\nu}$.

Now, let $\tilde{\nu}$ be the push-forward of the Lebesgue measure *Leb* on $[0,1]$ by π. For $x \in \mathrm{Sk}(C)$, keeping the notation of (9), by Lemma 3.7 we have

$$\tilde{\nu}(C_x) = Leb\{x \in [0,1] : \pi_x \in C_x\} = Leb([h_0(x), h_1(x)]) = \nu(C_x),$$

and the desired result follows. ∎

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Remark 3.14. By [2, Proposition 12], $\{\rho\} \cup \mathbb{N}$ are i.i.d. with law ν and conditionally given $(\mathcal{C}, \nu, \mathbb{N})$, for $b \in \mathcal{B}(\mathcal{C})$, $\overline{\ell}(b)$ has law $\nu_{C_b^{\overline{\ell}(b)}}$ and these random variables are independent. Indeed, in the notation of the latter reference, we have $Z_b^y = \overline{\ell}(b)$ for $y \in C_b^{\overline{\ell}(b)}$ since by definition $\overline{\ell}(b)$ is the image of b in $\nu_{C_b^{\overline{\ell}(b)}}$.

3.4 | The function F is a Brownian excursion

In order to prove that F is distributed as a Brownian excursion $\mathbb{e}$, we show that both F and $\mathbb{e}$ satisfy a same recursive equation, which has a unique solution in distribution.

For every continuous function $f : [0,1] \to \mathbb{R}_+$, define $(P_t(f), t \geq 0)$ as follows:

$$\forall t > 0, \qquad P_t(f) = \inf\{u > 0, f(u) - tu = 0\}.$$

Proposition 3.15. *Let $f : [0,1] \to \mathbb{R}_+$ be a (random) continuous function satisfying the following*:

(i) $f(0) = 0$ *a.s.*
(ii) The two processes $(P_t(f))_{t\geq 0}$ and $\left(\frac{1}{1+S_t}\right)_{t\geq 0}$ have the same law, where S is a $1/2$-stable subordinator with Laplace exponent $\Phi(\lambda) = (2\lambda)^{1/2}$.
(iii) Define $(t_i, a_i, b_i)_{i\geq 1}$ as follows: $\{t_i, i \geq 1\}$ is the set of jump times of $(P_t(f), t \geq 0)$ and, for all $i \geq 1$, $(a_i, b_i) = \left(P_{t_i}(f), P_{t_i-}(f)\right)$. Furthermore, the t_i's are sorted in nonincreasing values of $b_i - a_i$. Then, conditionally given $(P_t(f))_{t\geq 0}$, $\left\{\left((b_i - a_i)^{-1/2}(f(a_i + (b_i - a_i)u) - t_i(a_i + (b_i - a_i)u))\right)_{0\leq u\leq 1}\right\}_{i\geq 1}$ are i.i.d. random variables distributed as $(f(u), 0 \leq u \leq 1)$.

Then f and $\mathbb{e}$ have the same law.

In particular, observe that if f satisfies these assumptions, then $f(1) = 0$ a.s.

Lemma 3.16. *The standard Brownian excursion $\mathbb{e}$ on $[0,1]$ satisfies (i)–(iii).*

In order to prove that $\mathbb{e}$ satisfies these properties, we need the following result from Chassaing and Janson [19].

Theorem 3.17 ([19, Theorem 2.6])**.** *Consider b and $\mathbb{e}$, respectively a Brownian bridge and a Brownian excursion on $[0,1]$, extended on $\mathbb{R}$ so that they are 1-periodic. For all $a \geq 0$, define the following two processes*:

- *X_a the reflected Brownian bridge $|b|$ conditioned to have local time at time 1 at 0 equal to a*;
- *$Z_a = \Psi_a \mathbb{e}$, where $\Psi_a f(t) = f(t) - at - \inf_{-\infty<s\leq t}\{f(s) - as\}$.*

Denote by $L_t(X_a)$ the local time of X_a up to time t, and $V \in [0,1]$ the unique point at which $t \mapsto L_t(X_a) - at$ is maximum. Then,

$$Z_a \overset{(d)}{=} X_a(V + \cdot).$$

Let us show Lemma 3.16.

Proof of Lemma 3.16. First observe that it is clear by the Markov property and [9, Proposition 11] that $\mathbb{e}$ satisfies (i) and (ii). Let us prove that it satisfies (iii). To this end, observe that almost surely $X_a(V) = 0$ by definition of V. Hence, the excursions of Z_a, ordered by non-increasing order of length, are distributed as the excursions, ordered by non-increasing order of length, of X_a. As a consequence, conditionally given their endpoints, the excursions of X_a ordered by non-increasing order of length, are independent and distributed as appropriately rescaled standard Brownian excursions (see e.g., the proof of [35, Lemma 12]). This proves that $\mathbb{e}$ satisfies (iii). ■

Let us now prove Proposition 3.15.

Proof of Proposition 3.15. It thus remains to check that these assumptions characterize the distribution of f. Let f be a function satisfying (i)–(iii). For every $\varepsilon > 0$ we shall construct a coupling $(f, \mathbb{e})$ such that $\mathbb{P}(\|f - \mathbb{e}\| > \varepsilon) < \varepsilon$, which will imply that f and $\mathbb{e}$ have the same law (indeed, this implies e.g., that the Lévy-Prokhorov distance between the laws of f and $\mathbb{e}$ is at most ε). To simplify notation, set

$$\mathcal{U} = \bigcup_{k \in \mathbb{Z}_+} \mathbb{N}^k.$$

If $\mathbf{s} \in \mathbb{N}^k$ with $k \geq 0$ we set $|\mathbf{s}| = k$. Consider a family of i.i.d. $1/2$-stable subordinators $(S^{\mathbf{s}})_{\mathbf{s} \in \mathcal{U}}$ with the law of S, where by convention $\mathbb{N}^0 = \{\emptyset\}$. For each $\mathbf{s} \in \mathcal{U}$, set $Z^{\mathbf{s}}_t = (1 + S^{\mathbf{s}}_t)^{-1}$ for $t \geq 0$. Denote by $\mathbb{T}^{\mathbf{s}} := (t^{\mathbf{s}}_i)_{i \geq 1}$ the set of jump times of $S^{\mathbf{s}}$, ordered by decreasing values of $d^{\mathbf{s}}_i - g^{\mathbf{s}}_i$, where $d^{\mathbf{s}}_i = Z^{\mathbf{s}}_{t_i -}$ and $g^{\mathbf{s}}_i = Z^{\mathbf{s}}_{t_i}$ (in case of equality, sort them by increasing value of $g^{\mathbf{s}}_i$).

Now we define by induction sets of points $(\mathcal{R}^{\mathbf{s}})_{\mathbf{s} \in \mathcal{U}}$ as follows. Define intervals $[a^{\mathbf{s}}, b^{\mathbf{s}}]_{\mathbf{s} \in \mathcal{U}}$ as:

- for $\mathbf{s} = \emptyset$, $a^{\emptyset} = 0, b^{\emptyset} = 1$;
- for $\mathbf{s} \neq \emptyset$, let $\bar{\mathbf{s}} \in \mathcal{U}, i \in \mathbb{N}$ be such that $\mathbf{s} = \bar{\mathbf{s}} \cdot i$. Then, we define

$$a^{\mathbf{s}} := a^{\bar{\mathbf{s}}} + \left(b^{\bar{\mathbf{s}}} - a^{\bar{\mathbf{s}}}\right) g^{\mathbf{s}}_i, \qquad b^{\mathbf{s}} := a^{\bar{\mathbf{s}}} + \left(b^{\bar{\mathbf{s}}} - a^{\bar{\mathbf{s}}}\right) d^{\mathbf{s}}_i.$$

For every $\varepsilon > 0$, we shall now use the subordinators $(S^{\mathbf{s}})_{\mathbf{s} \in \mathcal{U}}$ to construct a coupling between f and $\mathbb{e}$ such that $\mathbb{P}(\|f - \mathbb{e}\| > \varepsilon) < \varepsilon$. For every fixed $k \geq 1$, set $\mathcal{R}_k := \{a^{\mathbf{s}}, |\mathbf{s}| \leq k\} \cup \{b^{\mathbf{s}}, |\mathbf{s}| \leq k\}$. Since f and $\mathbb{e}$ both satisfy (ii) and (iii), we can couple them using the subordinators $(S^{\mathbf{s}})_{|\mathbf{s}| \leq k}$, so that a.s.

$$\forall u \in \mathcal{R}_k, \qquad f(u) = \mathbb{e}(u).$$

Next, for every $\epsilon > 0$, one can find $K_\epsilon \in \mathbb{Z}_+$ such that

$$\mathbb{P}\left(\sup\{b - a : a, b \in \mathcal{R}_{K_\epsilon}, (a, b) \cap \mathcal{R}_{K_\epsilon} = \emptyset\} > \epsilon\right) < \epsilon.$$

Furthermore, since $\mathbb{e}$ and f are continuous on $[0, 1]$, they are uniformly continuous. In particular, there exists $C_\eta > 0$ such that $\mathbb{P}\left(\omega(\mathbb{e}) > C_\eta\right) < \eta$ and $\mathbb{P}\left(\omega(f) > C_\eta\right) < \eta$. Now,

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

on the event that $\{\sup\{b-a : a,b \in \mathcal{R}_{K_\epsilon}, (a,b)\cap\mathcal{R}_{K_\epsilon} = \emptyset < \epsilon, \omega(\mathbb{e}) < C_\eta, \omega(f) < C_\eta\}$, we have clearly $\|f-\mathbb{e}\| < 2C_\eta\epsilon$. Thus $\mathbb{P}(\|f-\mathbb{e}\| > 2C_\eta\epsilon) \leq \varepsilon + 2\eta$ (observe that the choice of C_η is independent of ϵ). This completes the proof. ∎

Proposition 3.18. *We have*

$$(F(t), 0 \leq t \leq 1) \overset{(d)}{=} (\mathbb{e}_t, 0 \leq t \leq 1).$$

Proof. We need to prove that F is continuous and satisfies (i)–(iii) in Proposition 3.15. We immediately obtain (i) from Lemma 3.8 applied to the leaf 0, continuity from Proposition 3.4 and (ii) from [9, Theorem 1 and Proposition 11] combined with invariance by uniform rerooting of the Brownian CRT. Finally, (iii) comes from [18, Corollary 2.3]. ∎

Remark 3.19. The previous considerations entail that $\mathcal{C}$ can be decomposed as follows, where we set $P_t = P_t(F)$ and $y_t = P_{t-} - P_t$ to simplify notation.

Step 1. The branch $[\![\rho, \overline{\ell}(\rho)]\!]$ isometric to a line segment with length $\int_0^\infty P_s \mathrm{d}s$.

Step 2. For every $t > 0$ such that $P_t < P_{t-}$ there is a branchpoint b_t on $[\![\rho, \overline{\ell}(\rho)]\!]$ at distance $\int_0^t P_s \mathrm{d}s$ from the root, and on b_t is grafted the tree obtained by iteration with the function $F^{(t)}$ defined by $F^{(t)} = y_t^{-1/2}(F(P_t + sy_t) - t(P_t + sy_t))_{0\leq s\leq 1}$, with distances renormalized by $y_t^{1/2}$. The leaf $\overline{\ell}(b_t)$ is then the leaf associated with b_t in the first step of this iteration. This construction provides us with an increasing sequence of trees (for the inclusion), and $\mathcal{C}$ is the completion of the union of these trees.

In order to see this, observe that each t such that $P_t \neq P_{t-}$, we have $P_t = \mu_t(\emptyset)$. By monotonicity and density of such instants, this holds for all $t \geq 0$. In particular, for each branchpoint $b \in [\![\rho, \overline{\ell}(\rho)]\!]$ corresponding to a cutpoint appearing at time t, we have

$$d(\rho, b_t) = \int_0^t \mu_s(\emptyset) ds = \int_0^t P_s \mathrm{d}s.$$

Letting $t \to \infty$, we have $d(\rho, \overline{\ell}(\rho)) = \int_0^\infty P_s \mathrm{d}s$. In order to show Step 2, consider a branchpoint $b_t \in [\![\rho, \overline{\ell}(\rho)]\!]$ and a branchpoint $b'_u \in [\![b_t, \overline{\ell}(b_t)]\!]$ corresponding to a cutpoint c appearing at time $u > t$. Again, it holds that

$$d(b_t, b'_u) = \int_t^u \mu_s(c) \mathrm{d}s = y_t^{1/2} \int_0^{y_t^{1/2}(u-t)} P_s(F^{(t)}) \mathrm{d}s,$$

as $\mu_s(c) = \mu_t(c) \cdot P_{y_t^{1/2}(s-t)}(F^{(t)}) = y_t \cdot P_{y_t^{1/2}(s-t)}(F^{(t)})$ for all $s \geq t$ by the same argument.

3.5 | Proof of Theorem 1.1 (i)

Roughly speaking, to establish Theorem 1.1 (i), we shall consider the function F obtained by the Pac-Man algorithm from the Brownian CRT $\mathcal{T}$ and the Poissonian rain $\mathcal{P}$. However, one has to be slightly careful since the Pac-Man algorithm has been defined using the cut-tree $\mathcal{C}$, which is itself defined by using an additional source of randomness, namely the points $(U_i)_{i\geq 1}$, so it is not clear whether the function F defined this way is a measurable function of $(\mathcal{T}, \mathcal{P})$.

To overcome this issue, we explain how the "Bertoin function" F can be directly defined from $(\mathcal{T}, \mathcal{P})$ in a measurable way. Recall that for every $t \geq 0$ and $x \in \mathcal{T}$, we denote by $\mathcal{T}_t(x)$ the connected component of $\mathcal{T} \setminus \mathcal{P}_t$ containing x and $\mu_t(x) = \mu(\mathcal{T}_t(x))$ its μ-mass. We first set $F_{(\mathcal{T},\mathcal{P})}(0) = 0$. Now, take $h \in (0, 1]$. There are three cases:

(i) If there exists $t_1 \geq 0$ such that $\mu_{t_1}(\varnothing) = h$, then we set $F_{(\mathcal{T},\mathcal{P})}(h) = t_1 h$.
(ii) If there exists $t_1 \geq 0$ such that $\mu_{t_1-}(\varnothing) = h$, then we set $F_{(\mathcal{T},\mathcal{P})}(h) = t_1 h$.
(iii) Otherwise, there exists $t_1 \geq 0$ such that $\mu_{t_1}(\varnothing) < h < \mu_{t_1-}(\varnothing)$.

In case (iii), notice that there exists a unique cutpoint $c_1 \in \mathcal{P}_\infty$ that has appeared at time t_1. Furthermore, $c_1 \in \mathcal{T}_{t_1-}(\varnothing)$. We now reason inductively. There are again three cases:

(i) If there exists $t_2 \geq t_1$ such that $\mu_{t_1}(\varnothing)+\mu_{t_2}(c_1) = h$, then we set $F_{(\mathcal{T},\mathcal{P})}(h) = t_1\mu_{t_1}(\varnothing)+t_2\mu_{t_2}(c_1)$.
(ii) If there exists $t_2 \geq t_1$ such that $\mu_{t_1}(\varnothing) + \mu_{t_2-}(c_1) = h$, then we set $F_{(\mathcal{T},\mathcal{P})}(h) = t_1\mu_{t_1}(\varnothing) + t_2\mu_{t_2-}(c_1)$.
(iii) Otherwise, there exists $t_2 \geq t_1$ such that $\mu_{t_2}(c_1) < h - \mu_{t_1}(\varnothing) < \mu_{t_2-}(c_1)$.

Thus, there are finally three cases, depending on h:

(a) either there exists a finite sequence $c_1, \dots, c_k$ of cutpoints appeared at respective times $t_1, \dots, t_k$ such that $h = \sum_{i=1}^{k} \mu_{t_i}(c_{i-1})$, $c_i \in \mathcal{T}_{t_i-}(c_{i-1})$ for all $i \leq k$; in this case, $F_{(\mathcal{T},\mathcal{P})}(h) = \sum_{i=1}^{k} t_i\mu_{t_i}(c_{i-1})$;
(b) there exists a finite sequence $c_1, \dots, c_k$ of cutpoints appeared at respective times $t_1, \dots, t_k$ such that $h = \sum_{i=1}^{k-1} \mu_{t_i}(c_{i-1}) + \mu_{t_k-}(c_{k-1})$, in which case $F_{(\mathcal{T},\mathcal{P})}(h) = \sum_{i=1}^{k-1} t_i\mu_{t_i}(c_{i-1}) + t_k\mu_{t_k-}(c_{k-1})$;
(c) there exists an infinite sequence $c_1, \dots$ of cutpoints appeared at respective times $t_1, \dots, t_k$ such that $h = \sum_{i=1}^{\infty} \mu_{t_i}(c_{i-1})$, in which case $F_{(\mathcal{T},\mathcal{P})}(h) = \sum_{i=1}^{\infty} t_i\mu_{t_i}(c_{i-1})$.

We shall now prove that $F_{(\mathcal{T},\mathcal{P})}$ meets the requirements of Theorem 1.1 (i).

Proof of Theorem 1.1(i). Consider a Brownian CRT $\mathcal{T}$, the Poissonian rain $\mathcal{P}$ on $\mathcal{T}$ and and let $U = (U_i)_{i\geq 1}$ be a sequence of i.i.d. leaves of $\mathcal{T}$ sampled according to the mass measure, independently of $\mathcal{P}$. Denote by $F_{(\mathcal{T},U),\mathcal{P}}$ the function defined by (4) using the Pac-Man algorithm. By the correspondence between cutpoints in $\mathcal{T}$ and branchpoints in $\mathcal{C}$ and the associated masses (Lemma 2.2), almost surely $F_{(\mathcal{T},\mathcal{U},\mathcal{P})} = F_{(\mathcal{T},\mathcal{P})}$.

By Propositions 3.11 and 3.18, almost surely, $F_{(\mathcal{T},\mathcal{P})}$ has the law of the Brownian excursion and almost surely, for every $t \geq 0$, the nonincreasing rearrangement of the masses of the connected components of $\mathcal{T} \setminus \mathcal{P}_t$ is the same as the nonincreasing rearrangement of the lengths of the excursions of $(F_{(\mathcal{T},\mathcal{P})}(s) - ts)_{0\leq s\leq 1}$ above its running infimum. This completes the proof. ∎

4 | RECOVERING THE ORIGINAL TREE TOGETHER WITH ITS POISSONIAN RAIN

Let $\mathcal{T}$ be a Brownian CRT with mass measure μ, $U := (U_i)_{i\geq 1}$ a sequence of i.i.d. leaves of $\mathcal{T}$ with common distribution μ, and $\mathcal{P}$ a Poissonian rain on $\mathrm{Sk}(\mathcal{T})$ independent of U. An important question in the literature (see [2, 14, 18]) concerns the problem of reconstruction of the original tree: is it possible to reconstruct $(\mathcal{T}, U, \mathcal{P})$ being given the cut-tree?

It turns out that there is a loss of information when one goes from a triple $(\mathcal{T}, U, \mathcal{P})$ to the cut-tree $(\mathcal{C}, \mathbb{N})$, where $\mathbb{N}$ denotes the subset $\{i : i \geq 1\}$ of points of $\mathcal{C}$ built in Section 2.2. More precisely, the following holds:

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Theorem 4.1 ([18, Theorem 3.2 (c)]). *Let $\mathcal{T}$ be a Brownian CRT. Then there exists a (random) tree* $\mathsf{shuff}(\mathcal{T})$ *such that in distribution*:

$$(\mathcal{T}, \mathsf{shuff}(\mathcal{T})) \overset{(d)}{=} (\mathrm{Cut}(\mathcal{T}), \mathcal{T}).$$

Here one recovers $\mathcal{T}$ from $\mathrm{Cut}(\mathcal{T})$ only in distribution. Later, Addario-Berry et al. [2] have shown that, if one considers an enrichment of the cut-tree transform with information called *routings*, it is possible to almost surely recover the initial tree (along with U and $\mathcal{P}$) from this enriched cut-tree (see [14] for an extension to ICRT). In short, routing variables are a collection of uniform points in every subtree dangling on the branchpoints.

More precisely, recall that for every $b \in \mathcal{B}(C)$ the subtree C_b of all descendents of b is $C_b = C_b^{\ell(b)} \cup C_b^{\overline{\ell}(b)}$. Denote by S the set of all subtrees of the form $C_b^{\ell(b)}$ or $C_b^{\overline{\ell}(b)}$ for $b \in \mathcal{B}(C)$. Consider the set of so-called routing variables $Z = (Z_{\mathcal{A}}, \mathcal{A} \in S)$ where $Z_{C_b^{\ell(b)}}$ is the image $\Lambda(C_b^{\ell(b)})$ of b in $C_b^{\ell(b)}$ and $Z_{C_b^{\overline{\ell}(b)}}$ is the image $\Lambda(C_b^{\overline{\ell}(b)})$ of b in $C_b^{\overline{\ell}(b)}$. By [2, Proposition 12], for every $\mathcal{A} \in S$ the random variable $Z_{\mathcal{A}}$ has law $\nu_{\mathcal{A}}$, and these random variables are conditionally independent given $(C, \nu, \mathbb{N})$. We also have:

Theorem 4.2 ([2, Proposition 12 and Corollary 17]). *There exists a (deterministic) measurable map Φ such that, almost surely*:

$$(\mathcal{T}, U, \mathcal{P}) = \Phi((C, \mathbb{N}, Z)).$$

The question answered by Theorem 1.1 (ii) is quite similar: being given the "Bertoin" Brownian excursion $\mathbb{e}$, is it possible to construct a map Ψ such that $\Psi(\mathbb{e})$ has the law of $(\mathcal{T}, \mathcal{P})$ and $F_{\Psi(\mathbb{e})} = \mathbb{e}$? The answer is positive, when adding an independent source of randomness.

The strategy of the proof is divided in two steps: first, in Section 4.1 we show that having the "Bertoin" excursion $F_{(\mathcal{T},\mathcal{P})}$ obtained from $(\mathcal{T}, \mathcal{P})$ is equivalent to having the cut-tree C, along with "half" of the routing variables. Second, in Section 4.2 we add the additional information (the remaining "half" routings) that allows us to reconstruct a tree with a Poissonian rain.

4.1 | From Bertoin's excursion to the semi-enriched cut-tree

We first prove that having the "Bertoin" excursion $F_{(\mathcal{T},\mathcal{P})}$ obtained from $(\mathcal{T}, \mathcal{P})$ is equivalent to having the cut-tree C, along with a collection of points that we call half-routings, a notion which we shall now define.

Let T be a compact binary real tree with root ρ, and recall that $\mathcal{B}(T)$ denotes the set of branchpoints of the tree T. We call half-routings on T a collection of leaves $H := \{H_b, b \in \mathcal{B}(T)\}$ such that $H_b \in T_b$ for every $b \in \mathcal{B}(T)$, where we recall that T_b is set of all (weak) descendents of b in T. For every $b \in \mathcal{B}(T)$, we define its associated record sequence $(b_i)_{i \geq 0} \in (\mathcal{B}(T) \cup \{\rho\})^{\mathbb{Z}_+}$ by induction as follows. Set $b_0 = \rho$. Then for every $k \geq 0$, assuming that $(b_i)_{0 \leq i \leq k}$ has been defined, we set $b_{k+1} = \ell_{b_k} \wedge b$. Since $(b_k)_{k \geq 0}$ is increasing for the genealogical order, it converges in T.

We say that the collection H of half-routings is *consistent* if the following holds:

$$\forall b \in \mathcal{B}(T) \cup \{\rho\}, \exists N_b \geq 0, b_{N_b} = b.$$

In particular, when H is consistent, then for all $b \in \mathcal{B}(T)$ the sequence $(b_k)_{k \geq 0}$ is stationary after time N_b and, for all $k \geq 1$ such that $b_k \neq b$, we have $\ell_{b_k} \in \overline{T}_{b_k}^{\ell_{b_{k-1}}}$. Indeed, by definition, $b_k = \ell_{b_{k-1}} \wedge b$, so that $b \in \overline{T}_{b_k}^{\ell_{b_{k-1}}}$. Thus, if we had $\ell_{b_k} \in T_{b_k}^{\ell_{b_{k-1}}}$, then it would follow that $b_{k+1} = \ell_{b_k} \wedge b = b_k$ and hence $b_j = b_k$ for $j \geq k$, so that $b = b_k$ since H is consistent. Furthermore, every branchpoint has a finite record sequence. Finally, denote by $\mathbb{T}_{HR}$ the set of compact rooted binary real trees enriched with the a consistent collection of half-routings and a mass measure supported on its set of leaves.

Observe that the Pac-Man algorithm can be applied mutatis mutandis to any enriched tree (T, H, M), where H is a consistent collection of half-routings on T and M a probability measure on T. We denote by $\mathsf{X}(T, H, M)$ the function obtained from this algorithm.

We keep the notation $(\mathcal{T}, U, \mathcal{P})$ for a Brownian CRT along with a sequence of i.i.d. leaves and a Poissonian rain on its skeleton, $\mathcal{C}$ its associated cut-tree, and set $\mathcal{H} := \left\{\overline{\ell}(b), b \in \mathcal{B}(\mathcal{C})\right\}$, with $\overline{\ell}(b)$ defined by (5). Observe that $\mathcal{H}$ is a.s. a consistent collection of half-routings on $\mathcal{C}$. Indeed, it is clear by (5) that $\overline{\ell}(b) \in \mathcal{C}_b$ for all $b \in \mathcal{B}(\mathcal{C})$. In addition, by Lemma 3.5, the collection $\mathcal{H}$ on $\mathcal{C}$ is consistent. Denoting by $C([0, 1], [0, \infty])$ the set of continuous maps from $[0, 1]$ to $[0, \infty]$, we have the following result:

Proposition 4.3. *There exists a (deterministic) map* $\Xi : C([0, 1], [0, \infty]) \to \mathbb{T}_{HR}$ *such that the following properties hold almost surely*:

(i) *We have* $\Xi \circ \mathsf{X}(\mathcal{C}, \mathcal{H}, \nu) = (\mathcal{C}, \mathcal{H}, \nu)$.
(ii) *Let* $\mathbb{e}$ *be a standard Brownian excursion. Then* $\mathsf{X} \circ \Xi(\mathbb{e}) = \mathbb{e}$.

Proof of Proposition 4.3. Let us start by defining the map Ξ. By construction, we have $\mathsf{X}(\mathcal{C}, \mathcal{H}, \nu) = F_{(\mathcal{T},\mathcal{P})}$. Since $F_{(\mathcal{T},\mathcal{P})}$ follows the law of a Brownian excursion by Theorem 1.1 (i), it is therefore enough to define $\Xi(\mathbb{e})$. We use a stick-breaking construction in the spirit of [7] to define $\Xi(\mathbb{e}) = (\tilde{\mathcal{C}}, \tilde{\mathcal{H}}, \tilde{\nu})$. We start with $\{\tilde{\rho}\}$, which will be the root of the tree $\Xi(\mathbb{e})$.

Step 1. Construct a branch $[\![\tilde{\rho}, \tilde{\ell}(\tilde{\rho})]\!]$ isometric to a line segment with length $\int_0^\infty P_s \mathrm{d}s$, where $P_s := \inf\{u > 0, \mathbb{e}_u = su\}$.

Step 2. For every $t > 0$ such that $P_t < P_{t-}$ we put a branchpoint $\tilde{b}_t$ on $[\![\tilde{\rho}, \tilde{\ell}(\tilde{\rho})]\!]$ at distance $\int_0^t P_s \mathrm{d}s$ from the root, and on $\tilde{b}_t$ we graft the tree obtained by iteration with the function $\mathbb{e}^{(t)}$ defined by $\mathbb{e}^{(t)} = (P_{t-} - P_t)^{-1/2}(\mathbb{e}_{P_t + s(P_{t-} - P_t)} - t(P_t + s(P_{t-} - P_t))_{0 \leq s \leq 1}$, which is a Brownian excursion by Proposition 3.15. In the first step of this iteration, a leaf, denoted by $\tilde{\ell}(\tilde{b}_t)$, is associated with $\tilde{b}_t$.

This construction provides us with an increasing sequence of trees (for the inclusion). We denote by $\tilde{\mathcal{C}}$ the completion of the union of these trees with a collection of half-routings $\tilde{\mathcal{H}} = (\tilde{\ell}(\tilde{b}) : \tilde{b} \in \mathcal{B}(\tilde{\mathcal{C}}) \cup \{\tilde{\rho}\})$.

We check that $\tilde{\mathcal{C}}$ is compact and establish (i) at the same time. Since the desired properties involve only the law of $\mathbb{e}$, without loss of generality, we may assume that $\mathbb{e} = \mathsf{X}(\mathcal{C}, \mathcal{H}, \nu) = F_{(\mathcal{T},\mathcal{P})}$. Observe that by Remark 3.19, the tree $\mathcal{C}$ satisfies the same recursive construction as $\tilde{\mathcal{C}}$. As a consequence, setting $\overline{\ell}(\rho) = 0$ (which we recall to be the "image" in $\mathcal{C}$ of the root $\varnothing$ of $\mathcal{T}$), the trees $\mathcal{C}^0 := \bigcup_{b \in \mathcal{B}(\mathcal{C}) \cup \{\rho\}} [\![b, \overline{\ell}(b)]\!]$ and $\tilde{\mathcal{C}}^0 := \bigcup_{\tilde{b} \in \mathcal{B}(\tilde{\mathcal{C}}) \cup \{\tilde{\rho}\}} [\![\tilde{b}, \tilde{\ell}(\tilde{b})]\!]$ are isometric, and so are their completions, which implies that $\mathcal{C} = \tilde{\mathcal{C}}$.

Now let us explain how to endow $\tilde{\mathcal{C}}$ with a mass measure. Roughly speaking, given $h \in [0, 1]$, we explain how to define the final target point of the Pacman algorithm just from $\mathsf{X}(\mathcal{C}, \mathcal{H}, \nu)$. We construct a sequence of branchpoints associated with h as follows. If there exists t such that $P_t \neq P_{t-}$ and $h = P_t$ or $h = P_{t-}$, let $t_1(h) = \dagger$. Otherwise, let $t_1(h) := \inf\{t \in [0, 1], P_t \neq P_{t-}, P_t < h\}$. From Step 2 above, with $t_1(h)$ is associated a branchpoint in $\tilde{\mathcal{C}}$ denoted by $b_1(h)$. Then, as for the stick breaking construction of $\mathcal{C}$,

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

we iterate this in the excursion $\mathbb{e}^{(t_1(h))}$. In the end, we obtain an increasing sequence (for the genealogical order) of branchpoints associated with h. Also observe that this sequence stops at † only for countably many values of $h \in [0,1]$. For every $h \in [0,1]$ for which this does not happen, we define $\ell(h)$ as the limit of the sequence $(b_i(h))_{i\geq 1}$. Then given the description of the Pacman algorithm in the beginning of Section 3.5 we have $\ell(h) = \pi_h$, in the sense that $\ell(h)$ is the final target point of h in $\mathcal{C}$ by the Pacman algorithm. We then define the mass measure $\tilde{\nu}$ on $\tilde{\mathcal{C}}$ as the pushforward of the Lebesgue measure on $[0,1]$ by ℓ, which by Lemma 3.13 coincides with ν.

We finally prove (ii). By step (i), we have $\mathsf{X}(\Xi(\mathbb{e})) \overset{(d)}{=} \mathbb{e}$, so it suffices to check that they coincide a.s. on a dense subset of $[0,1]$. Now, for any $t \geq 0$ such that $P_{t-} \neq P_t$, by the discussion in the beginning of Section 3.5 we have $\mathsf{X}(\Xi(\mathbb{e}))(P_t) = \mathbb{e}_{P_t}$ and $\mathsf{X}(\Xi(\mathbb{e}))(P_{t-}) = \mathbb{e}_{P_{t-}}$. Iterating this with the functions $\mathbb{e}^{(t)}$, we obtain that a.s. $\mathsf{X}(\Xi(\mathbb{e}))$ and $\mathbb{e}$ coincide on a dense set, and this completes the proof. ■

Remark 4.4. By Remark 3.14 the previous proof also shows that if $(\mathcal{C}, \mathcal{H}, \nu) = \Xi(\mathbb{e})$ with $\mathcal{H} = (\overline{\ell}(b), b \in \mathcal{B}(\mathcal{C}))$, then conditionally given $(\mathcal{C}, \nu, \mathbb{N})$, $\overline{\ell}(b)$ has law $\nu_{\mathcal{C}_b^{\overline{\ell}(b)}}$ for $b \in \mathcal{B}(\mathcal{C})$, and these random variables are independent.

4.2 | From the semi-enriched cut-tree to the initial tree

To establish Theorem 1.1 (ii) we apply the map Ξ introduced in Proposition 4.3 to a Brownian excursion, which allows to reconstruct a cut-tree. We then want to apply Theorem 4.2 to reconstruct a Brownian CRT with its Poissonian rain, and to this end we need a set of routing random variables. Additional randomness is required because Ξ gives only "half" of the rootings.

Proof of Theorem 1.1(ii). Consider a Brownian excursion $\mathbb{e}$ and an independent sequence of i.i.d. uniform random variables on $[0,1]$. Recall from Proposition 4.3 the map Ξ. Set $(\mathcal{C}, \mathcal{H}, \nu) = \Xi(\mathbb{e})$ with $\mathcal{H} = (\overline{\ell}(b), b \in \mathcal{B}(\mathcal{C}))$. Recall also from the proof of Proposition 4.3 the map $\pi : [0,1] \to \mathcal{C}$.

For convenience, we split the i.i.d. uniform random variables on $[0,1]$ into two independent collections of i.i.d. uniform random variables on $[0,1]$: the first one $(V_i)_{i\geq 1}$ indexed by $\mathbb{N}$ and the second one $(W_b)_{b\in\mathcal{B}(\mathcal{C})}$ indexed by the branchpoints of $\mathcal{C}$ (this can be done in a deterministic measurable way).

We set $i = \pi(V_i) \in \mathcal{C}$ for every $i \geq 1$, so that $\mathbb{N} \cup \{\rho\}$ are i.i.d. with law ν, where ρ is the root of $\mathcal{C}$.

We shall now define routing variables $Z = (Z_{\mathcal{A}}, \mathcal{A} \in \mathcal{S})$ such that for every $\mathcal{A} \in \mathcal{S}$ the random variable $Z_{\mathcal{A}}$ has law $\nu_{\mathcal{A}}$, and these random variables are conditionally independent given $(\mathcal{C}, \nu)$, where we recall that $\mathcal{S}$ denotes the set of all subtrees of the form $\mathcal{C}_b^{\overline{\ell}(b)}$ or $\mathcal{C}_b \setminus \mathcal{C}_b^{\overline{\ell}(b)} \cup \{b\}$ for $b \in \mathcal{B}(\mathcal{C})$.

First, for every $b \in \mathcal{B}(\mathcal{C})$ we set $Z_{\mathcal{C}_b^{\overline{\ell}(b)}} = \overline{\ell}(b)$. Then, by Remark 4.4 the random variables $(Z_{\mathcal{C}_b^{\overline{\ell}(b)}} : b \in \mathcal{B}(\mathcal{C}) \cup \{\rho\})$ have respective laws $\nu_{\mathcal{C}_b^{\overline{\ell}(b)}}$, and these random variables are conditionally independent given $(\mathcal{C}, \mathbb{N})$.

Second take $b \in \mathcal{B}(\mathcal{C})$. To define $Z_{\mathcal{C}_b \setminus \mathcal{C}_b^{\ell(b)} \cup \{b\}}$ we proceed as follows. Assume for convenience that $b \in [\![\rho, 0]\!]$. Keeping the notation introduced in the proof of Proposition 4.3, with b is associated a time $t \geq 0$ such that $P_t \neq P_{t-}$. Then set $Z_{\mathcal{C}_b \setminus \mathcal{C}_b^{\ell(b)} \cup \{b\}} = \pi_{W_b P_t}$.

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Define it in the same way for every $b \in \mathcal{B}(C)$, in the spirit of the previous stick-breaking construction.

Finally define $Z = (Z_{\mathcal{A}}, \mathcal{A} \in \mathcal{S})$. By construction, for every $\mathcal{A} \in \mathcal{S}$ the random variable $Z_{\mathcal{A}}$ has law $\nu_{\mathcal{A}}$, and these random variables are conditionally independent given $(C, \mathbb{N}, \nu)$.

Applying to $(C, \mathbb{N}, Z)$ the map Φ of [2, Corollary 17] provides a triple $\Phi(C, \mathbb{N}, Z)$ having the law of $(\mathcal{T}, U, \mathcal{P})$, such that almost surely, by Proposition 4.3, $F_{\Phi(C,\mathbb{N},Z)} = \mathbb{e}$. This completes the proof. ■

ACKNOWLEDGMENTS

I.K. thanks Jean Bertoin for asking him if there is a connection between the two seemingly different constructions X_{B} and X_{AP}, and we thank Jean-François Marckert for mentioning the use of cut-trees in [30]. We also thank the anonymous referees for their careful review of the article and for their constructive comments which have greatly contributed to improve the article.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in arXiv at https://arxiv.org/abs/2301.01153.

ORCID

Paul Thévenin https://orcid.org/0009-0006-6441-196X

REFERENCES

1. L. Addario-Berry, N. Broutin, and C. Holmgren, *Cutting down trees with a Markov chainsaw*, Ann. Appl. Probab. **24** (2014), no. 6, 2297–2339.
2. L. Addario-Berry, D. Dieuleveut, and C. Goldschmidt, *Inverting the cut-tree transform*, Ann. Inst. Henri Poincaré Probab. Stat. **55** (2019), no. 3, 1349–1376.
3. D. Aldous, *The continuum random tree. I*, Ann. Probab. **19** (1991), no. 1, 1–28.
4. D. Aldous, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289.
5. D. Aldous, *The continuum random tree. II*, Stoch. Anal. **19** (1991), 23–70.
6. D. Aldous and J. Pitman, *The standard additive coalescent*, Ann. Probab. **26** (1998), 1703–1726.
7. D. Aldous and J. Pitman, *Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent*, Probab. Theory Relat. Fields **118** (2000), no. 4, 455–482.
8. E. Baur and J. Bertoin, "*Cutting edges at random in large recursive trees*," Stochastic analysis and applications 2014, D. Crisan, B. Hambly, and T. Zariphopoulou (eds.), Springer, Cham, 2014, pp. 51–76.
9. J. Bertoin, *A fragmentation process connected to Brownian motion*, Probab. Theory Relat. Fields **117** (2000), no. 2, 289–301.
10. J. Bertoin, Random fragmentation and coagulation processes, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, UK, 2006.
11. J. Bertoin, *Fires on trees*, Ann. I H P Probab. Stat. **48** (2012), 909–921.
12. J. Bertoin, *The cut-tree of large recursive trees*, Ann. I H P Probab. Stat. **51** (2015), 478–488.
13. J. Bertoin and G. Miermont, *The cut-tree of large Galton–Watson trees and the Brownian CRT*, Ann. Appl. Probab. **23** (2013), no. 4, 1469–1493.
14. N. Broutin, H. He, and M. Wang. Pruning, cut trees, and the reconstruction problem. arXiv preprint arXiv: 2210.13948, 2022.
15. N. Broutin and J.-F. Marckert, *A new encoding of coalescent processes: Applications to the additive and multiplicative cases*, Probab. Theory Relat. Fields **166** (2016), no. 1-2, 515–552.
16. N. Broutin and J.-F. Marckert. Convex minorant trees associated with Brownian paths and the continuum limit of the minimum spanning tree. arXiv preprint arXiv:2307.12260, 2023.

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

17. N. Broutin and M. Wang, *Cutting down p-trees and inhomogeneous continuum random trees*, Bernoulli **23** (2017), no. 4A, 2380–2433.
18. N. Broutin and M. Wang, *Reversing the cut tree of the Brownian continuum random tree*, Electron. J. Probab. **22** (2017), no. 80, 1–23.
19. P. Chassaing and S. Janson, *A Vervaat-like path transformation for the reflected Brownian bridge conditioned on its local time at 0*, Ann. Probab. **29** (2001), no. 4, 1755–1779.
20. D. Dieuleveut, *The vertex-cut-tree of Galton–Watson trees converging to a stable tree*, Ann. Appl. Probab. **25** (2015), no. 4, 2215–2262.
21. M. Drmota, A. Iksanov, M. Moehle, and U. Roesler, *A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree*, Random Struct. Algoritm. **34** (2009), no. 3, 319–336.
22. S. N. Evans and J. Pitman, *Stationary Markov processes related to stable Ornstein–Uhlenbeck processes and the additive coalescent*, Stoch. Process. Appl. **77** (1998), no. 2, 175–185.
23. V. Féray and I. Kortchemski, *The geometry of random minimal factorizations of a long cycle via biconditioned bitype random trees*, Ann. Henri Lebesgue **1** (2018), 149–226.
24. J. A. Fill, N. Kapur, and A. Panholzer, *Destruction of very simple trees*, Algorithmica **46** (2006), no. 3, 345–366.
25. C. Holmgren, *Random records and cuttings in binary search trees*, Comb. Probab. Comput. **19** (2010), no. 3, 391–424.
26. C. Holmgren, *A weakly 1-stable distribution for the number of random records and cuttings in split trees*, Adv. Appl. Probab. **43** (2011), no. 1, 151–177.
27. A. Iksanov and M. Möhle, *A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree*, Electron. Commun. Probab. **12** (2007), 28–35.
28. S. Janson, *Random cutting and records in deterministic and random trees*, Random Struct. Algoritm. **29** (2006), no. 2, 139–179.
29. O. Kallenberg, Foundations of modern probability, Probability and its Applications, 2nd ed., Springer-Verlag, New York, 2002.
30. J.-F. Marckert and M. Wang, *A new combinatorial representation of the additive coalescent*, Random Struct. Algoritm. **54** (2019), no. 2, 340–370.
31. A. Meir and J. Moon, *Cutting down recursive trees*, Math. Biosci. **21** (1974), no. 3-4, 173–181.
32. A. Meir and J. W. Moon, *Cutting down random trees*, J. Austral. Math. Soc. **11** (1970), 313–324.
33. G. B. Ojeda and C. Holmgren. Invariance principle for fragmentation processes derived from conditioned stable Galton-Watson trees. arXiv preprint arXiv:2010.07880, 2020.
34. A. Panholzer, *Cutting down very simple trees*, Quaest. Math. **29** (2006), no. 2, 211–227.
35. J. Pitman, *The SDE solved by local times of a Brownian excursion or bridge derived from the height profile of a random tree or forest*, Ann. Probab. **27** (1999), no. 1, 261–283.
36. P. Thévenin, *A geometric representation of fragmentation processes on stable trees*, Ann. Probab. **49** (2021), no. 5, 2416–2476.
37. M. Wang. Stable trees as mixings of inhomogeneous continuum random trees. arXiv preprint arXiv:2211.07253, 2022.

How to cite this article: I. Kortchemski and P. Thévenin, *Coupling Bertoin's and Aldous–Pitman's representations of the additive coalescent*, Random Struct. Alg. **65** (2024), 17–45. https://doi.org/10.1002/rsa.21206

10982418, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/rsa.21206 by Cochrane France, Wiley Online Library on [16/06/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License