Reversing the cut tree of the Brownian continuum random tree

Article (Published Version)

Electron. J. Probab. **22** (2017), no. 80, 1–23.
ISSN: 1083-6489 DOI: 10.1214/17-EJP105

Reversing the cut tree of the Brownian continuum random tree

Nicolas Broutin* Minmin Wang†

Abstract

Consider the Aldous–Pitman fragmentation process [7] of a Brownian continuum random tree $\mathcal{T}^{\mathrm{br}}$. The associated cut tree $\mathrm{cut}(\mathcal{T}^{\mathrm{br}})$, introduced by Bertoin and Miermont [13], is defined in a measurable way from the fragmentation process, describing the genealogy of the fragmentation, and is itself distributed as a Brownian CRT. In this work, we introduce a *shuffle* transform, which can be considered as the reverse of the map taking $\mathcal{T}^{\mathrm{br}}$ to $\mathrm{cut}(\mathcal{T}^{\mathrm{br}})$.

Keywords: Brownian continuum random tree; Aldous–Pitman fragmentation; cut tree; random cutting of random trees.
AMS MSC 2010: Primary 60J80; 60C05, Secondary 60G18; 60F15.
Submitted to EJP on November 4, 2016, final version accepted on September 8, 2017.
Supersedes arXiv:1408.2924v2.

1 Introduction

1.1 Motivation and literature

Let $\mathcal{T}^{\mathrm{br}}$ be Aldous' Brownian continuum random tree (CRT). We consider the fragmentation process introduced by Aldous & Pitman [7]: informally, the process describes the time evolution of the masses of the connected components of a forest $\mathcal{F}_t$, $t \geq 0$, where $\mathcal{F}_t$ results from a logging of $\mathcal{T}^{\mathrm{br}}$ with cuts falling uniformly per unit of time and length in $\mathcal{T}^{\mathrm{br}}$. The Aldous–Pitman fragmentation is an instance of a self-similar fragmentation such as studied in Bertoin's book [9]. There is a natural genealogical structure associated with the fragmentation process, and it is as a representation of this genealogy that Bertoin and Miermont [13] constructed the so-called cut tree of $\mathcal{T}^{\mathrm{br}}$, hereafter denoted by $\mathrm{cut}(\mathcal{T}^{\mathrm{br}})$. A rather remarkable fact is that $\mathrm{cut}(\mathcal{T}^{\mathrm{br}})$ is itself also distributed as the Brownian CRT. In this work, we are interested in defining the reverse of the map $\mathcal{T}^{\mathrm{br}} \mapsto \mathrm{cut}(\mathcal{T}^{\mathrm{br}})$. This has been motivated by a seemingly natural question: given the cut tree, can one reconstruct the initial tree? We will see that the cut tree does

*INRIA de Paris, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12 - France. E-mail: nicolas.broutin@m4x.org
†Université Pierre et Marie Curie, 4 place Jussieu, 75005 Paris - France. E-mail: wangminmin03@gmail.com

not contain all the information necessary for such a reconstruction; this observation leads us then to introduce a reverse transform. However, giving a proper meaning to the reverse transform requires some explanation, which we postpone to Sections 1.2 and 1.3. For the time being, we provide some background on cut trees, which can be traced back to some work in combinatorics dating from the seventies.

Random cutting of trees. The idea of cut trees is closely related to random cutting of trees, a subject initiated by Meir and Moon [28], which has since then been largely studied. The initial question concerns discrete trees, and we present here a version of the random cutting problem where cuts happen at nodes (one can also define a version where cuts happen at edges): take a rooted tree (random or not) on a finite vertex set; at each step, sample a node uniformly at random and remove it along with all the edges adjacent to it (the removed node is then referred to as a *cut*); this disconnects the tree into connected components (maybe one, if we picked a leaf, for instance); discard the components that are now disconnected from the root; keep going until the root is finally picked. The main questions addressed by Meir and Moon and many subsequent researchers concern mostly the number of cuts that are needed for the process to terminate. This problem has been considered for a number of classical models of deterministic and random trees, including random binary search trees [22, 23], random recursive trees [24, 17, 11, 8] and Galton–Watson trees conditioned on the total progeny [25, 4, 20, 30, 10, 13].

The CRT being the scaling limits of Galton–Watson trees with finite-variance offspring distribution [6], the case of Galton–Watson trees is the most related to our matters, and we now focus on that case: let T_n be a Galton–Watson tree conditioned to have n nodes, and whose offspring distribution has variance $\sigma^2 < \infty$; denote by N_n the number of cuts until the root V is picked in the above process. Then Janson [25] showed by moment calculations that, as $n \to \infty$, $N_n/(\sigma\sqrt{n})$ converges in distribution to the Rayleigh distribution. Incidentally, the Rayleigh distribution is also the limit law of $(\sigma H_n/\sqrt{n})_{n\geq 1}$, where H_n denotes the height (distance to the root) of a randomly picked node in T_n. Thus, Janson's result can be rephrased as follows: the limit distribution of $(N_n/(\sigma\sqrt{n}))_{n\geq 1}$ coincides with that of $(\sigma H_n/\sqrt{n})_{n\geq 1}$. It turns out that an even stronger statement holds true in the case that T_n is a uniform labelled tree of n nodes (this is equivalent to take the offspring distribution to be Poisson(1), and is sometimes referred to as the Cayley tree): we have that N_n and $H_n + 1$ actually have exactly the same distribution. This result is due to Addario-Berry, Broutin and Holmgren [4], and relies on the following bijective method: one can construct another tree $\mathrm{cut}(T_n; V)$ (on the same vertex set as T_n) which encodes the isolation of V by the successive cuts (see Section 1.2 for the details) such that (1) the node V lies at distance $N_n - 1$ from the root, and (2) $\mathrm{cut}(T_n; V)$ has the same distribution as T_n, while (3) V is a uniform node in $\mathrm{cut}(T_n; V)$. The above distributional identity then follows. We call $\mathrm{cut}(T_n; V)$ the 1-partial cut tree, since it keeps track of the way one node (here V) was isolated. More generally, Addario-Berry, Broutin and Holmgren [4] have considered cutting procedures resulting in the isolation of k nodes and introduced the corresponding k-partial cut trees. For these cutting procedures, one only discards the portions of the tree that do not contain any of the k marked nodes to be isolated. Moreover, by first taking a uniform permutation of the vertex set, we can define simultaneously all the k-isolation processes, so that letting $k \to \infty$ we obtain the (complete) cut tree $\mathrm{cut}(T_n)$ of T_n, whose graph distance encodes the number of cuts required to isolate every single one of the n nodes. In this case, since all the nodes are marked, no portion of the tree is ever discarded, and the tree $\mathrm{cut}(T_n)$ actually encodes the genealogy of a discrete fragmentation of the tree (we refer to Section 1.2 for details). A similar notion appears in Bertoin [10] and Bertoin & Miermont [13], where

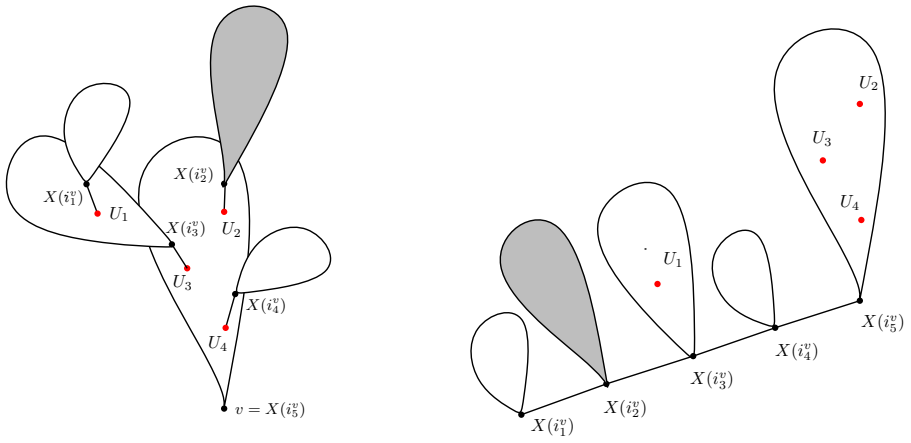

Figure 1: An example of one-node isolation and the associated 1-partial cut tree. On the left, the initial tree T and the sequence of cuts $X(i_1^v), X(i_2^v), \dots, X(i_5^v)$, which are responsible for the isolation of $v = X(i_5^v)$. On the right, the corresponding cut tree. Note that due to the cut at $X(i_m^v)$, a small subtree of T is discarded; it then appears in $\mathrm{cut}(T; v)$ as the subtree above the path $[\![X(i_1^v), X(i_5^v)]\!]$.

they define a (different) cut tree $\widehat{\mathrm{cut}}(T_n)$ for T_n directly as the genealogy tree of the discrete fragmentation process induced by the cutting of T_n.

Random cutting of continuum trees and fragmentation processes. More recently, such cutting processes have been considered for the Brownian CRT [4, 2, 13, 10]. The cutting on the Brownian CRT is of course closely related to the Aldous–Pitman fragmentation mentioned in the first paragraph. Moreover, Bertoin and Miermont [13] proved that if T_n is a Galton–Watson tree with a finite variance offspring distribution and conditioned to have n nodes, then the pair $(T_n, \widehat{\mathrm{cut}}(T_n))$, after suitable scaling in the graph distance, converges in distribution in the sense of Gromov–Prokhorov, to a pair $(\mathcal{T}^{\mathrm{br}}, \mathrm{cut}(\mathcal{T}^{\mathrm{br}}))$ of continuum random trees; furthermore the tree $\mathrm{cut}(\mathcal{T}^{\mathrm{br}})$ can be defined directly from the fragmentation process of $\mathcal{T}$ and indeed encodes its genealogy. A similar result holds for $(T_n, \mathrm{cut}(T_n))$ in the case where T_n is a uniform Cayley tree; see [15].

Let us also mention that cut trees have been introduced for other models of continuum random trees, including Lévy trees under excursion measures [1], stable trees conditioned on the total masses [16], and inhomogeneous continuum random trees [15].

1.2 Reversing the cut trees of Cayley trees

Although our main concern is the case of the continuum tree, we think it will be helpful to explain here the question we address and our approach to its solution in the setting of discrete trees. The case of Cayley trees, for which the question has been studied in [4] for partial reversals and then in [15] for the complete reversal, is especially adapted to our presentation since many of the correspondences are then exact. We refer to these two papers for proofs and further details.

Throughout this part, let $\mathbb{T}_n$ denote the set of rooted labelled trees on the vertex set $[n] := \{1, 2, \dots, n\}$ and let $T \in \mathbb{T}_n$. For $u, v \in [n]$, we write $[\![u, v]\!]$ for the set of vertices that lie on the shortest path joining u to v in T. Let $X(1), X(2), \dots, X(n)$ be a uniform permutation of the vertex set. We will use the sequence $(X(i))_{1 \le i \le n}$ to define various isolation processes on T.

One-node isolation and the 1-partial cut tree. Let v be any node of T and consider the following isolation process of v. Let $F_0^v = T$ and for $1 \leq m \leq n$, let F_m^v be the connected component of $T \setminus \{X(1), \ldots, X(m)\}$ containing v, with the convention that $F_{m'}^v = \varnothing$ for all $m' \geq m$ if $X(m) = v$. We say that $X(m)$ is a *cut* in this process if $X(m) \in F_{m-1}^v$. Namely, the cuts are those elements of $[n]$ whose removal have reduced the size of the current connected component of v. Let $\mathbf{X}(v) := \{X(i_m^v) : 1 \leq m \leq N(v), i_1^v < i_2^v < \cdots < i_{N(v)}^v\}$ be the sequence formed by the successive cuts. Observe in particular that $X(i_{N(v)}^v) = v$. For $1 \leq m \leq N(v) - 1$, we let U_m be the unique neighbor of $X(i_m^v)$ which belongs to $[\![X(i_m^v), v]\!]$. The following algorithm returns a tree denoted by $G(T, v, \mathbf{X}(v))$ as a function of T, v and $\mathbf{X}(v)$.

Algorithm 1. **Construction of the 1-partial cut tree.** Apply the following transformations to the tree T:
- for $1 \leq m \leq N(v) - 1$, remove the edge $\{X(i_m^v), U_m\}$;
- for $1 \leq m \leq N(v) - 1$, add the edge $\{X(i_m^v), X(i_{m+1}^v)\}$;
- declare $X(i_1^v)$ as the root.

Denote by $\mathrm{cut}(T; v) = G(T, v, \mathbf{X}(v))$ the graph thereby obtained (see also Figure 1). Then we have $G(T, v, \mathbf{X}(v)) \in \mathbb{T}_n$ and it contains a path consisting of the sequence $\mathbf{X}(v)$. Moreover, if we remove all the edges in that path, then each $X(i_k^v)$ remains connected to the subgraph $F_{i_k^v - 1} \setminus F_{i_k^v}$, namely the part discarded at step k due to the cut at $X(i_k^v)$.

k-node isolation and the k-partial cut tree. The above isolation process can be generalized to the case of multiple nodes. Let $\mathbf{v} = \{v_1, v_2, \ldots, v_k\}$ be $k \geq 1$ vertices of T (not necessarily distinct). For each v_j and $m \geq 0$, let $F_m^{v_j}$ denote the connected component of $T \setminus \{X(i) : i \leq m\}$ containing v_j; then the sequence of cuts responsible for the isolation of v_j, namely, those X_m satisfying $X_m \in F_{m-1}^{v_j}$, is denoted by $\{X(i_m^{v_j}) : 1 \leq m \leq N(v_j), i_1^{v_j} < i_2^{v_j} < \cdots < i_{N(v_j)}^{v_j}\}$.

To define the associated partial cut tree, we adopt a recursive approach. For $2 \leq j \leq k$, let $m_j = \max\{m : v_j \in \cup_{1 \leq j' \leq j-1} F_m^{v_{j'}}\}$. Then set $\mathbf{X}(v_1) = \{X(i_m^{v_1}) : 1 \leq m \leq N(v_1)\}$ and for $2 \leq j \leq k$, set $\mathbf{X}(v_j) = \{X(i_m^{v_j}) : i_m^{v_j} > m_j, m \leq N(v_j)\}$.

Algorithm 2. **Construction of the k-partial cut tree.** Set $G_1 = G(T, v_1, \mathbf{X}(v_1))$, and for $2 \leq j \leq k$, do the following: if $\mathbf{X}(v_j) = \varnothing$, set $G_j = G_{j-1}$; otherwise,
- locate in G_{j-1} the connected component of $G_{j-1} \setminus (\mathbf{X}(v_1) \cup \cdots \cup \mathbf{X}(v_{j-1}))$ which contains v_j, denote it by T_j; let w_j be the node of $\mathbf{X}(v_1) \cup \cdots \cup \mathbf{X}(v_{j-1})$ that is closest to v_j in G_{j-1};
- replace in G_{j-1} the subgraph T_j by $G(T_j, v_j, \mathbf{X}(v_j))$: to do so, remove the edge between w_j and T_j and add the edge $\{w_j, X(m_j + 1)\}$; let G_j be the graph obtained, still rooted at $X(i_1^{v_1})$.

Denote by $\mathrm{cut}(T; \mathbf{v}) = G_k$ the graph thereby obtained. Then $\mathrm{cut}(T; \mathbf{v}) \in \mathbb{T}_n$ and the subset of nodes $\mathbf{X}(v_1) \cup \cdots \cup \mathbf{X}(v_k)$ forms a subtree that contains the root $X(i_1^{v_1})$ (See Figure 2).

The complete cut tree. Let $(V_i)_{i \geq 1}$ be a sequence of independent and uniform nodes of T and write $\mathbf{V}_k = \{V_1, \ldots, V_k\}$, $k \geq 1$. Observe that almost surely the sequence $\{\mathrm{cut}(T; \mathbf{V}_k) : k \geq 1\}$ becomes stationary after some k. Denote by $\mathrm{cut}(T)$ the limit of this sequence; it has the following remarkable properties. First, for each vertex v, the path in $\mathrm{cut}(T)$ from the root to v consists of precisely the cuts responsible for the isolation of v. Note there is at most one tree in $\mathbb{T}_n$ satisfying this property. We conclude that $\mathrm{cut}(T)$ does not depend on $(V_i)_{i \geq 1}$. Second, $\mathrm{cut}(T)$ is uniformly distributed in $\mathbb{T}_n$.

Note that we cannot recover T from $\mathrm{cut}(T)$. To explain this, let us first introduce the following notation. For $T \in \mathbb{T}_n$ and two vertices $u \neq v$, let $\mathrm{S}(T, u \,|\, v)$ denote the

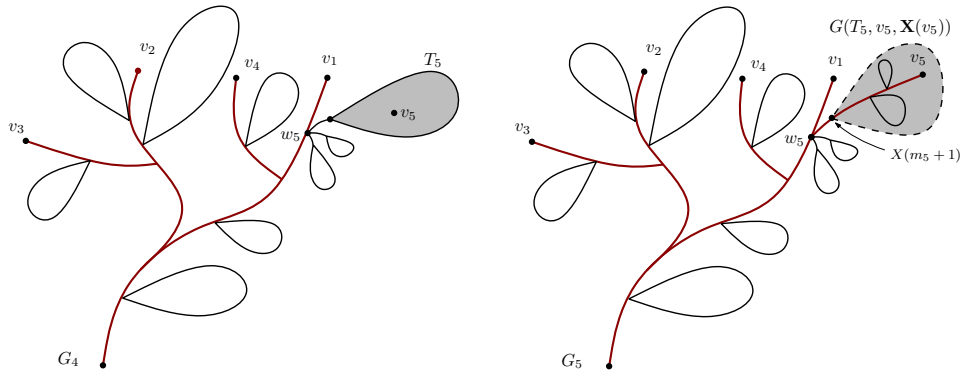

Figure 2: The recursive construction of partial cut trees. To obtain G_5 from G_4, we simply replace the subtree T_5 (the grey part on the left) by the 1-partial cut tree $G(T_5, v_5, \mathbf{X}(v_5))$ (the grey one on the right). The red paths on the left consist of $\mathbf{X}(v_1) \cup \cdots \cup \mathbf{X}(v_4)$; the ones on the right consist of $\mathbf{X}(v_1) \cup \cdots \cup \mathbf{X}(v_5)$.

connected component of $T \setminus \{u\}$ which contains v. If $X(i_m^{V_j})$ is a cut in the isolation of V_j as defined above, denote by U_m^j the neighbor of $X(i_m^{V_j})$ that belongs to $[\![V_j, X(i_m^{V_j})]\!]$. Then one can show that U_m^j is uniformly random in $\mathrm{S}(\mathrm{cut}(T), X(i_m^{V_j}) \,|\, V_j)$ (see also Figure 1). In particular, this means that the information concerning the whereabouts of U_m^j is partially lost in $\mathrm{cut}(T)$; therefore we cannot know the initial tree just from its cut tree. On the other hand, we know the distribution of (U_m^j) conditional on $\mathrm{cut}(T)$. Relying on this, we can "resample" the lost information, namely, take a random collection $(\hat{U}_m^j)$ according to the distribution of (U_m^j) conditional on $\mathrm{cut}(T)$; we then "reconstruct" T from $\mathrm{cut}(T)$ by assuming $(\hat{U}_m^j)$ are the actual (U_m^j). Of course, we will not obtain from this procedure the actual T with probability 1, but instead a random tree which has the distribution of T given $\mathrm{cut}(T)$. This is the basic idea of our reverse transform for the mapping $T \mapsto \mathrm{cut}(T)$. The following paragraphs explain how to proceed in the case of Cayley trees.

Reversing the 1-partial cut tree transform. Let v be a vertex of T. Suppose that $X_1^v, \ldots, X_L^v = v$ is the sequence of vertices along the path of T from the root to v. For $1 \le l \le L-1$, sample a random vertex U_i^v uniformly in $\mathrm{S}(T, X_l^v \,|\, v)$. Write $\mathbf{U}(v) = \{U_i^v : 1 \le l \le L-1\}$. Sample a uniform vertex ρ_1 of T. The following algorithm returns a tree $H(T, v, \mathbf{U}(v), \rho_1)$ as a function of T, v, $\mathbf{U}(v)$ and ρ_1.

Algorithm 3. **1-partial shuffle transform.** Take T and do the following:
– for $1 \le l \le L-1$, remove the edge $\{X_l^v, X_{l+1}^v\}$;
– for $1 \le l \le L-1$, add the edge $\{X_l^v, U_l^v\}$;
– declare ρ_1 as the root.

Denote by $H = H(T, v, \mathbf{U}(v), \rho_1)$ the graph thereby obtained, which turns out to be an element of $\mathbb{T}_n$. Observe that the above algorithm is the exact reverse of Algorithm 1. Let us also make the following observation, which is useful for the generalization to continuum trees. Let w, w' be two distinct vertices of T and let x (resp. x') be the vertex among $\{X_l^v : 1 \le l \le L\}$ which is closest to w (resp. w'). To simplify the discussion, suppose that x, x', v are all distinct; the path in H joining w to w' contains a sub-collection of $(X_l^v)_{1 \le l \le L} \cup \mathbf{U}(v) \cup \{w, w'\}$; we then match the elements of this sub-collection into pairs: $\{(a_i, p_i) : 1 \le i \le m\}$, so that for each i, a_i and p_i are in the same connected component after removing the edges $\{X_l^v, X_{l+1}^v\}, 1 \le l \le L-1$. See Fig. 3 for an example. If we write respectively d for the graph distance in T and d_H for that in H, then we have

(see also Fig. 3)

$$d_H(w,w') = \sum_{1\le i\le m} d(a_i,p_i) + m - 1. \tag{1.1}$$

Reversing the cut tree transform. The above partial shuffle transform can be extended to several vertices. To explain this, let $(V_i)_{i\ge 1}$ be a sequence of independent uniform vertices of T and write $\mathbf{V}_k = \{V_1,\dots,V_k\}$, $k \ge 1$. Let r denote the root of T. Let $\mathrm{Span}(T,\mathbf{V}_k)$ be the smallest connected subgraph of T containing $\mathbf{V}_k \cup \{r\}$. Let $T_1 = T$ and for $k \ge 2$, if $V_k \in \mathrm{Span}(T,\mathbf{V}_{k-1})$, set $T_k = \mathbf{U}_k = \rho_k = \varnothing$; otherwise let T_k be the connected component of $T \setminus \mathrm{Span}(T,\mathbf{V}_{k-1})$ which contains V_k. For each $k \ge 1$ and $T_k \ne \varnothing$, sample a uniform vertex ρ_k in T_k; let $\{X_1^k,\dots,X_{L_k}^k\} = [\![r,V_k]\!] \cap T_k$, with the X_i^k ordered by decreasing distances to V_k; then, for $1 \le l \le L_k - 1$, sample a uniform vertex U_l^k among $\mathrm{S}(T,X_l^k \,|\, V_k)$; let $\mathbf{U}_k = \{U_l^k : 1 \le l \le L_k - 1\}$. Then, the following algorithm returns a tree $H(T,\mathbf{V}_k,(\mathbf{U}_j)_{1\le j\le k},(\rho_j)_{1\le j\le k})$ as a function of $T,\mathbf{V}_k,(\mathbf{U}_j)_{1\le j\le k}$ and $(\rho_j)_{1\le j\le k}$.

Algorithm 4. **k-partial shuffle transform.** Take T and do the following:
- Remove the edges of $\mathrm{Span}(T,\mathbf{V}_k)$.
- For $1 \le j \le k$ and $T_k \ne \varnothing$, do the following: [1]
  - For $1 \le l \le L_j - 1$, add an edge $\{X_l^j, U_l^j\}$ for $1 \le l \le L_j - 1$;
  - add an edge $\{\rho_k, w\}$, where w is the vertex of $\mathrm{Span}(T,\mathbf{V}_{k-1})$ closest to T_k;[2]
- Declare ρ_1 as the root.

Denote by $H_k = H(T,\mathbf{V}_k,(\mathbf{U}_j)_{1\le j\le k},(\rho_j)_{1\le j\le k})$ the graph produced, which is an element of $\mathbb{T}_n$. Alternatively, H_k can be obtained in the following recursive way, which can be seen as the dual of Algorithm 2. (See also Figure 2, from right to left.) Let $k \ge 2$.

- If $T_k = \varnothing$, then $H_k = H_{k-1}$.
- Otherwise, replace in T the subgraph T_k by $H(T_k,\mathbf{U}_k,\rho_k)$ and denote by R_k the resulting graph. Then we have $H_k = H(R_k,\mathbf{V}_{k-1},(\mathbf{U}_j)_{1\le j\le k-1},(\rho_1)_{1\le j\le k-1})$.

Note that we have $\mathrm{Span}(R_k,\mathbf{V}_{k-1}) = \mathrm{Span}(T,\mathbf{V}_{k-1})$ (with obvious notation); therefore Algorithm 4 can still be applied to define $H(R_k,\mathbf{V}_{k-1},(\mathbf{U}_j)_{1\le j\le k-1},(\rho_1)_{1\le j\le k-1})$. It is not difficult to see that the sequence $(H_k)_{k\ge 1}$ is eventually stationary, and denote by $\mathrm{shuff}(T)$ its limit. If T is distributed uniformly in $\mathbb{T}_n$, then we have

$$\big(T,\mathrm{shuff}(T)\big) \overset{d}{=} \big(\mathrm{cut}(T),T\big).$$

Thanks to this identity in distribution, we can consider the shuffle transform as the reverse of the transform $T \mapsto \mathrm{cut}(T)$. Let us also keep in mind that in the definition of $\mathrm{shuff}(T)$, we have sampled the following random variables for each $k \ge 1$: i) a uniform vertex ρ_k in T_k which "serves" as the initial root of the subtree containing V_k; ii) for each vertex X_l^k on the path of T_k from its root to V_k, a uniform vertex U_l^k in $\mathrm{S}(T,X_l^k \,|\, V_k)$ which "serves" as a neighbor of X_l^k in the initial tree.

1.3 An overview of the paper

The aim of the paper is to introduce the *shuffle transform* for real trees, and more specifically for the Brownian continuum random tree: for a rooted real tree

[1] The formulation here is slightly different from the one in [15], which is stated as follows: replace each edge (x,w) of $\mathrm{Span}(T,\mathbf{V}_k)$ by (x,U_w), where x is the one closer to the root than w and U_w is a uniform node sampled in the subtree of T above (x,w), and then root the obtained tree at ρ_1. It is not difficult to see that this gives the same transform as Algorithm 4.

[2] Observe that this step amounts to rooting the subgraph replacing T_k at ρ_k.

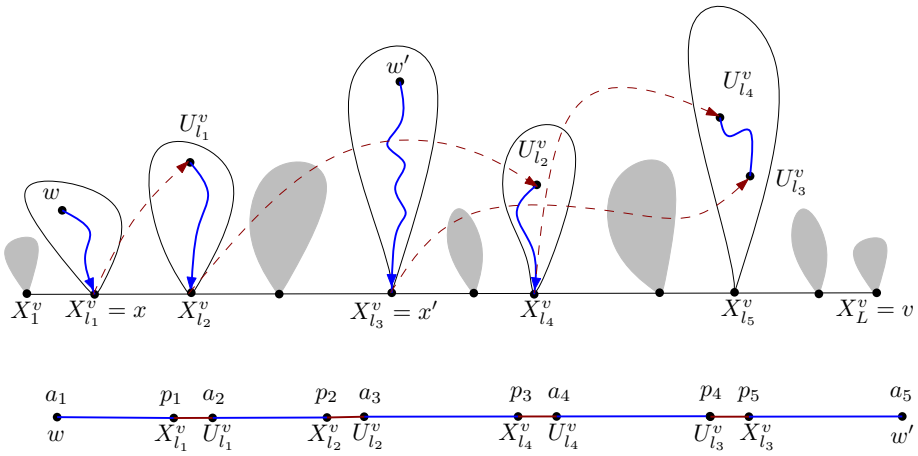

Figure 3: The 1-partial shuffle transform as viewed from two points w, w'. In the upper line is illustrated the tree T; the straight line represents the path $\{X^v_l : 1 \leq l \leq L\}$ joining the root to v. Removing the edges $\{X^v_l, X^v_{l+1}\}, 1 \leq l \leq L-1$, T falls into L subtrees (blocks in white and in gray in above). For $l < L$, the point U^v_l is located in one of these subtrees right to X^v_l. By adding an edge between X^v_l and U^v_l (represented by the red arrows in above), we shuffle these subtrees and make them into a new graph H. Observe that the path of H joining w to w' only crosses a sub-collection of these subtrees. In the example above, they correspond to the white subtrees. In the line below is depicted this path. Observe that it consists of m blue segments and $m-1$ red segments, where m is the number of subtrees it crosses ($m = 5$ in the example above). Each of the red segments has length one; they are the edges added between X^v_l and U^v_l. Each of the blue segments was contained in a white subtree; so that its length is the same in H as in T. This explains Equation (1.1). The pair (a_i, p_i) consists of the endpoints of the i-th blue segment. Their relative positions (which one is on the left) are in fact unimportant. Here, we have followed a choice convenient for generalization. In this example, their respective values in $(X^v_l)_{1\leq l\leq L} \cup \mathbf{U}(v) \cup \{w, w'\}$ are given by the labels below.

$(\mathcal{T}, d, \rho)$ equipped with a finite measure μ, we define a random symmetric matrix $\Gamma_\infty = (\gamma_\infty(i,j))_{1\leq i,j<\infty}$ whose entries take values in $\mathbb{R}_+ \cup \{\infty\}$, such that when $(\mathcal{T}, d, \rho)$ is distributed according to the distribution of the Brownian CRT, which we denote by $\mathbf{P}$, almost surely Γ_∞ is well-defined and characterizes a random measured and rooted real tree $\mathrm{shuff}(\mathcal{T}) = (\mathcal{H}, d_\mathcal{H}, \mu_\mathcal{H}, \rho_\mathcal{H})$; moreover, the law under $\mathbf{P}$ of the pair $(\mathcal{H}, \mathcal{T})$ seen as measured real trees is the same as that of $(\mathcal{T}, \mathrm{cut}(\mathcal{T}))$ under $\mathbf{P}$.

This shuffle transform can be viewed as an extension to the Brownian CRT of the construction in Section 1.2 for Cayley trees. As we have already mentioned, for a discrete tree T, the associated cut tree does not contain all the information necessary to reconstruct T. This is also true for continuum trees. Therefore, in defining the shuffle transform, we begin by sampling a collection of random points of $\mathcal{T}$ referred to as the *marks* which replaces the lost information in the cut tree. We then construct a real tree which corresponds to the initial tree, if the "resampled" information corresponds to the actual one. Recall that for discrete trees, we have used an approximation procedure: we first consider the reversals of the partial cut trees and then define the shuffle transform as the limit of the partial reversals. The advantage of this approach lies in that the partial cut trees retain some unmodified portions of the initial tree, therefore their reversals are easier to handle. This is even more important in the continuum tree setting. In that case, the reconstruction consists in roughly three steps:

– Define the 1-partial shuffle transform for the CRT. This has been done in [15]. Let us briefly explain the idea there. Algorithm 3 does not generalize directly to the CRT, but Equation (1.1) does. Indeed, if η, η' are two independent uniform points of the Brownian CRT, then during the one-node isolation process, there is only a finite number of cuts falling on the geodesic of the CRT which joins η to η' (see Lemma 2.5 below for a precise statement). This suggests that the number of summands in (1.1) remains bounded for the CRTs and we can "recover" the distance between η and η' by a generalization of (1.1).

– Define the k-partial shuffle transform for the CRT, for $k \geq 2$. We use a recursive procedure, similar to the one for Cayley trees. In particular, the recursive construction provides a natural coupling between the different partial reversals, which is convenient for the proof of convergence in the next step.

– The convergence of k-partial shuffle transforms as $k \to \infty$. Contrary to the case of discrete trees, for CRTs, this convergence is non trivial, and a significant part of the paper is devoted to its proof. Note that our proof relies crucially on some specific properties of the Brownian CRT, especially the scaling property.

The rest of the paper is organized as follows. In Section 2, we introduce the necessary notation and recall from [13] the definition of the cut tree of the Brownian CRT. We also collect some results from [15] that are useful later on. In Section 3, we give the formal definition of the shuffle transform, which is defined as the limit of partial reversal transforms and state our main result (Theorem 3.2). The proof for the convergence of the partial reversals is found in Section 4.

2 Notation and preliminaries

2.1 Notation and background on continuum random trees

We only give here a short overview, the interested reader may consult [6], [27] or [19] for more details.

Pointed measured metric spaces. A *pointed measured metric space* is a quadruple (X, d, μ, ρ) where (X, d) is a compact metric space equipped with a *finite* Borel measure μ, and ρ is a distinguished point that is usually referred to as being the root. Two pointed measured metric spaces (X, d, μ, ρ) and (X', d', μ', ρ') are *equivalent* if there exists an isometry $f : (X, d) \to (X', d')$ satisfying $\mu' = f_*\mu$ and $f(\rho) = \rho'$. Let $\mathbb{M}$ denote the set of equivalence classes of pointed measured metric spaces. Then $\mathbb{M}$ is a Polish space when endowed with the pointed Gromov–Hausdorff–Prokhorov topology ([19, 29]).

The following functional defined on $\mathbb{M}$ is useful in our treatment. Let (X, d, μ, ρ) be a pointed measured metric space where μ is a probability measure and has X as its support. Let $(\eta_i, i \geq 1)$ be a sequence of i.i.d. points of X with common distribution μ and set $\eta_0 = \rho$. We define a random symmetric and semi-infinite matrix

$$\mathrm{M}(X, d, \mu, \rho) = (d(\eta_i, \eta_j) : 0 \leq i, j < \infty).$$

Note that the distribution of $\mathrm{M}(X, d, \mu, \rho)$ only depends on the equivalence class of (X, d, μ, ρ). Thanks to the Gromov reconstruction theorem ([21, Section $3\frac{1}{2}.5$]), the matrix $\mathrm{M}(X, d, \mu, \rho)$ characterizes this equivalence class. In the rest of the paper and when no confusion arises, we often use the short-hand notation $\mathrm{X} = (X, d, \mu, \rho)$ to indicate that X stands for the whole equivalence class of (X, d, μ, ρ).

Real trees. The metric spaces of interest here are real trees. A compact metric space (X, d) is a *real tree* if for any two points $u, v \in X$, the following two properties hold. First, there exists a unique isometry $\varphi : [0, d(u, v)] \to X$ such that $\varphi(0) = u$ and $\varphi(d(u, v)) = v$; in this case, we denote by $[\![u, v]\!]_X := \varphi([0, d(u, v)])$ or sometimes simply by $[\![u, v]\!]$ if the underlined metric space (X, d) is clear from the context. Second, if $f : [0, 1] \to X$ is an injective continuous map satisfying $f(0) = u$ and $f(1) = v$, then necessarily $f([0, 1]) = [\![u, v]\!]_X$. A *rooted real tree* (T, d, ρ) is a real tree (T, d) with a distinguished point called the *root*.

Let (T, d, ρ) be a rooted real tree. The *degree* of a point $u \in T$, which we denote by $\deg(u, T)$, is the number of connected components of $T\backslash\{u\}$. We let

$$\mathrm{Lf}(T) = \{u \in T : \deg(u, T) = 1\},\ \mathrm{Br}(T) = \{u \in T : \deg(u, T) \geq 3\} \text{ and } \mathrm{Sk}(T) = T\backslash\mathrm{Lf}(T)$$

denote the set of the *leaves*, the set of *branch points* and the *skeleton* of T, respectively. Note that the distance d induces a sigma-finite measure ℓ on T satisfying $\ell([\![u, v]\!]) = d(u, v)$, for any $u, v \in T$. We refer to ℓ as the *length measure* of T. A *subtree* S of T is a closed and connected nonempty subset of T. Observe that (S, d) is itself a real tree. We often root S at the point $\mathrm{root}(S)$, which is defined to be the unique point of S minimising the distance to ρ; in that case, we say that S is a *rooted subtree*.

Let (T, d, ρ) be a rooted real tree and let $u, v \in T$ be two distinct points. We denote by $\mathrm{S}(T, u \,|\, v)$ the connected component of $T\backslash\{u\}$ which contains v. Namely,

$$\mathrm{S}(T, u \,|\, v) = \{w \in T : u \notin [\![w, v]\!]\}. \tag{2.1}$$

Now let $\mathbf{v} = \{v_1, \ldots, v_k\}$ be a set of k points of T. We write

$$\mathrm{Span}(T, \mathbf{v}) = \cup_{1 \leq i \leq k} [\![\rho, v_i]\!]$$

for the subtree of T spanning $v_1, \ldots, v_k$. Next, observe that there is at most countably infinite collection of the connected components of $T\backslash\mathrm{Span}(T, \mathbf{v})$, since (T, d) is compact. Let $\{C_i^\circ : i \geq 1\}$ be this collection. For each $i \geq 1$, let C_i be the closure of C_i° in T. Then one can check that C_i is a subtree and there exists a unique $b_i \in \mathrm{Br}(T) \cap \mathrm{Span}(T, \mathbf{v})$ such that $C_i = C_i^\circ \cup \{b_i\}$ and $b_i = \mathrm{root}(C_i)$. Set $h_i = d(\rho, b_i)$. If T is further equipped with a finite (Borel) measure μ, then each C_i is also equipped with a finite (Borel) measure which is the restriction of μ to C_i; by a slight abuse of notation we still denote this measure by μ. In that case, we denote by $\mathrm{C}_i = (C_i, d, \mu, b_i)$ for the (equivalence class of) pointed measured metric space, $i \geq 1$; then the *k-spine decomposition* of T with respect to $\mathbf{v}$ is the point measure on $\mathbb{R}_+ \times \mathbb{M}$ defined as

$$\mathrm{Decomp}(T, \mathbf{v}) = \sum_{i \geq 1} \delta_{(h_i,\, \mathrm{C}_i)} \tag{2.2}$$

A *measured real tree* is a pointed measured metric space (T, d, μ, ρ) where (T, d) is a real tree. For instance, each C_i in (2.2) is a measured real tree.

The Brownian continuum random tree. One way to obtain a measured real tree starts from an *excursion*: a continuous nonnegative function $e : \mathbb{R}_+ \to \mathbb{R}_+$ is said to be an excursion if e has compact support and satisfies that $e_0 = 0$, $\zeta_e := \sup\{s > 0 : e_s > 0\} \in (0, \infty)$ and $e_s > 0$, $\forall s \in (0, \zeta_e)$. Let e be an excursion. For $s, t \in [0, \zeta_e]$, let

$$\tfrac{1}{2}\, d_e(s, t) := e_s + e_t - 2 \inf_{u \in [s \wedge t,\, s \vee t]} e_u. \tag{2.3}$$

The factor $1/2$ in the above definition is unconventional but suits our purpose here. Define $s \sim_e t$ if $d_e(s, t) = 0$. Then d_e induces a metric on the quotient space $[0, \zeta_e]/\!\sim_e$,

which we still denote by d_e. Moreover, the metric space $(\mathcal{T}_e := [0,\zeta_e]/\!\sim, d_e)$ is a real tree (see e.g. Theorem 2.1 of [18]). Write $p_e : [0,\zeta_e] \to \mathcal{T}_e$ for the canonical projection. We denote by μ_e the push-forward of the Lebesgue measure on $[0,\zeta_e]$ by p_e and set $\rho_e := p_e(0)$. Then $(\mathcal{T}_e, d_e, \mu_e, \rho_e)$ is a measured real tree as defined previously. Moreover, it follows from the above construction that the support of μ_e is $\mathcal{T}_e$ and that $\mu_e(\mathcal{T}_e) = \zeta_e$.

Let e denote the canonical process of $\mathbf{C}(\mathbb{R}_+,\mathbb{R})$. For $a \in (0,\infty)$, let $\mathbf{P}^{(a)}$ be the probability distribution on $\mathbf{C}(\mathbb{R}_+,\mathbb{R})$ of the normalized Brownian excursion of length a, namely, e under $\mathbf{P}^{(a)}$ is distributed as a Brownian excursion conditioned on $\zeta_e = a$. The following *scaling property* of Brownian excursions plays a crucial role in our treatment: for each $a > 0$,

$$\Big(\tfrac{1}{\sqrt{a}} e_{at}, t \geq 0\Big) \text{ under } \mathbf{P}^{(a)} \overset{d}{=} e \text{ under } \mathbf{P}^{(1)}. \tag{2.4}$$

Recall the measured real tree $(\mathcal{T}_e, d_e, \mu_e, \rho_e)$ from the paragraph above. We view (the equivalence class of) $(\mathcal{T}_e, d_e, \mu_e, \rho_e)$ under $\mathbf{P}^{(a)}$ as a random variable taking values in $\mathbb{M}$, whose distribution we still denote as $\mathbf{P}^{(a)}$. In particular, $\mathbf{P} := \mathbf{P}^{(1)}$ is the *law of the Brownian continuum random tree*. A real-valued random variable R is a *Rayleigh random variable* if R has density $\mathbf{1}_{\{x>0\}} x e^{-x^2/2}$. The following well-known fact will be used implicitly at various places: let $\eta \in \mathcal{T}_e$ be a random point of distribution μ_e and let η' be either another independent point of distribution μ_e or the root ρ; then, under $\mathbf{P}$, in both cases $d_e(\eta,\eta')$ is a Rayleigh random variable.

In this work, we study stochastic processes defined on random measured real trees. In a general way, we construct these processes first for the canonical process e, or equivalently the real tree $(\mathcal{T}_e, d_e, \mu_e, \rho_e)$; we then consider the ensemble under the law $\mathbf{P}^{(a)}$, for $a > 0$. See e.g. [3] for a construction of the Aldous–Pitman fragmentation in this manner. In the rest of the paper, $(\mathcal{T}, d, \mu, \rho)$ stands for $(\mathcal{T}_e, d_e, \mu_e, \rho_e)$ and $\boldsymbol{\mathcal{T}}$ the corresponding measured real tree.

2.2 Cut tree of the Brownian continuum random tree

Let $\boldsymbol{\mathcal{T}} = (\mathcal{T}, d, \mu, \rho)$ be as defined above, where $\mathcal{T}$ is further equipped with the length measure ℓ. Recall that $\mathbf{P}$ is the law of the Brownian CRT. We define the cut tree for $\boldsymbol{\mathcal{T}}$, following Bertoin and Miermont [13]. To that end, let $\mathcal{P}$ be a Poisson point process on $\mathbb{R}_+ \times \mathcal{T}$ of intensity measure $dt \otimes \ell(dx)$. Every point $(t,x) \in \mathcal{P}$ is seen as a *cut* on $\mathcal{T}$ at location x and arriving at time t. Given $(\mathcal{T}, d, \mu, \rho)$, let $(V_i)_{i\geq 1}$ be a sequence of i.i.d. points of $\mathcal{T}$ with common distribution $\mu(\cdot)/\mu(\mathcal{T})$. Then for each $i \geq 1$ and $t \geq 0$, let $\mathcal{T}_i(t)$ be the set of those points in $\mathcal{T}$ which are still connected to V_i at time t, that is

$$\mathcal{T}_i(t) := \{u \in \mathcal{T} : \mathcal{P} \cap ([0,t] \times [\![V_i, u]\!]) = \varnothing\}. \tag{2.5}$$

For $i \neq j$, set $t_{ij} = \inf\{t \geq 0 : \mathcal{P} \cap ([0,t] \times [\![V_i, V_j]\!]) \neq \varnothing\} \in [0,\infty]$. Define a symmetric function $\delta : \mathbb{Z}_+^2 \to \mathbb{R}_+$ by setting $\delta(i,i) = 0$ for $i \geq 0$ and

$$\delta(0,i) = \int_0^\infty \mu\big(\mathcal{T}_i(s)\big) ds, \quad \delta(i,j) = \int_{t_{ij}}^\infty \Big\{\mu\big(\mathcal{T}_i(s)\big) + \mu\big(\mathcal{T}_j(s)\big)\Big\} ds, \quad i,j \geq 1, i \neq j. \tag{2.6}$$

Proposition 2.1 ([13])**.** *Under* $\mathbf{P}$*, the following statements (I-II) hold almost surely.*

I. For all $i,j \geq 1$ *and* $i \neq j$*,* $\delta(0,i) \in (0,\infty)$*,* $t_{ij} \in (0,\infty)$*, and* $\delta(i,j) \in (0,\infty)$*.*

II. There exists a measured real tree $\mathrm{cut}(\mathcal{T}) = (\mathcal{G}, d_\mathcal{G}, \mu_\mathcal{G}, \rho_\mathcal{G})$ *and a sequence of its points* $(V_i')_{i\geq 0}$ *with* $V_0' = \rho_\mathcal{G}$ *such that*

$$d_\mathcal{G}(V_i', V_j') = \delta(i,j), \quad \forall\, 0 \leq i,j < \infty.$$

Also, conditional on $\mu_\mathcal{G}$*,* $(V_i')_{i\geq 1}$ *has the distribution of a sequence of i.i.d. points with common probability distribution* $\mu_\mathcal{G}$*.*

Moreover, $\mathrm{cut}(\mathcal{T})$ *under* $\mathbf{P}$ *has the same distribution as* $\mathcal{T}$ *under* $\mathbf{P}$*.*

Part of the above Proposition says that $(\mathcal{G}, d_{\mathcal{G}}, \mu_{\mathcal{G}}, \rho_{\mathcal{G}})$ is uniquely determined by $(\delta(i,j) : 0 \le i, j < \infty)$ (up to measure-preserving isometry), by Gromov's reconstruction theorem. Therefore, the measured real tree $\mathrm{cut}(\mathcal{T})$ is well-defined, $\mathbf{P}$-a.s. It also follows from the above construction that the mapping $(\mathcal{T}, \mathcal{P}) \mapsto \mathrm{cut}(\mathcal{T})$ is measurable; see a related discussion in [13].

If, for $t \ge 0$, we write $i \sim_t j$ if and only if $\mathcal{T}_i(t) = \mathcal{T}_j(t)$, then $\sim_t$ defines an exchangeable random partition of $\mathbb{N}$. Moreover, the family of partitions $\{\sim_t\colon t \ge 0\}$ has a natural genealogical structure, which is described by $\mathrm{cut}(\mathcal{T})$; we refer to [13, 9] for more details. Note that the root of $\mathcal{T}$ is irrelevant in the above cutting process, whereas the root of $\mathrm{cut}(\mathcal{T})$ is meaningful for the genealogy it describes.

2.3 Partial cut trees as an approximation of $\mathrm{cut}(\mathcal{T})$

We recall here the definition of the k-partial cut tree of the measured real tree $\boldsymbol{\mathcal{T}} = (\mathcal{T}, d, \mu, \rho)$, as well as some of its properties, which will be useful for the proof later. These properties are mostly proven in [4] and [15] for $k = 1$, but the case $k > 1$ in general also follows from the arguments there.

For $i, j \ge 1$ and $t \ge 0$, recall $\mathcal{T}_i(t)$ from (2.5) and $t_{i,j}$ just below. Observe that $\mathcal{T}_i(t) \subseteq \mathcal{T}_i(t')$ if $t \ge t'$. We then denote $\mathcal{T}_i(t-) = \cap_{s<t}\mathcal{T}_i(s)$, for $t > 0$. Set $t_i^* = \max_{1\le j<i} t_{i,j}$ for $i \ge 2$ and $t_1^* = 0$. For each $i \ge 1$, let $\{t_m^{(i)} : m \in \mathbb{N}\}$ be the set of discontinuity points of the mapping $t \in (t_i^*, \infty) \mapsto \mu(\mathcal{T}_i(t))$. Then for each $m \in \mathbb{N}$, almost surely there exists a unique point $x_m^{(i)} \in \mathcal{T}_i(t_m^{(i)}-)$ such that $(t_m^{(i)}, x_m^{(i)}) \in \mathcal{P}$. It follows that $\Delta_m^{\circ,(i)} := \mathcal{T}_i(t_m^{(i)}-) \setminus \mathcal{T}_i(t_m^{(i)})$ is non empty and $x_m^{(i)} \in \Delta_m^{\circ,(i)}$. Let $\Delta_m^{(i)}$ be the closure of $\Delta_m^{\circ,(i)}$ and let $\boldsymbol{\Delta}_m^{(i)} = (\Delta_m^{(i)}, d, \mu, x_m^{(i)})$ be the (equivalence class of) measured real tree induced. Note in particular that $\mu(\Delta_m^{(i)}) = \mu(\Delta_m^{\circ,(i)})$. We also define $h_m^{(i)} := \int_{[0,\, t_m^{(i)}]} \mu(\mathcal{T}_i(s))ds$ for each $m \in \mathbb{N}, i \ge 1$. Let us recall from (2.2) the k-spine decomposition of a real tree.

Proposition 2.2 ([15])**.** *Under* $\mathbf{P}$*, the following holds almost surely: for each* $k \ge 1$*, there exists a measured real tree* $\mathrm{cut}(\mathcal{T}; V_1, \dots, V_k) = (\mathcal{G}_k, d_{\mathcal{G}_k}, \mu_{\mathcal{G}_k}, \rho_{\mathcal{G}_k})$ *and* k *points* $V_1', \dots, V_k' \in \mathcal{G}_k$ *such that*

$$\mathrm{Span}(\mathcal{G}_k, \{V_1', \dots, V_k'\}) \text{ is isometric to } \mathrm{Span}(\mathcal{G}, \{V_1', \dots, V_k'\}),$$
$$\mathrm{Decomp}(\mathcal{G}_k, \{V_1', \dots, V_k'\}) = \sum_{1\le i\le k} \sum_{m\in\mathbb{N}} \delta_{(h_m^{(i)},\, \boldsymbol{\Delta}_m^{(i)})},$$

where the real tree $(\mathcal{G}, d_{\mathcal{G}})$ *is defined in Proposition 2.1; conditional on* $\mu_{\mathcal{G}_k}$*,* $(V_1', \dots, V_k')$ *are distributed as* k *independent points of common probability distribution* $\mu_{\mathcal{G}_k}$*. Moreover for each* k*,* $\mathrm{cut}(\mathcal{T}; V_1, \dots, V_k)$ *under* $\mathbf{P}$ *has the same distribution as* $\mathcal{T}$ *under* $\mathbf{P}$*.*

The case $k = 1$ of Proposition 2.2 corresponds to Theorem 1.7 of [4] (see also Theorem 3.2 of [15]). The arguments there can be straightforwardly adapted to yield a proof of the general case $k \ge 1$. Now recall from page 9 the probability measure $\mathbf{P}^{(a)}$ for the measured real tree encoded by a Brownian excursion of length a. Recall also (2.4), the scaling property of Brownian excursions. The following is a direct consequence of Proposition 2.2 and a multi-point version of the Bismut decomposition for the Brownian CRT ([26, Theorem 3]).

Corollary 2.3 (Scaling property)**.** *Let* $k \ge 1$*. For* $1 \le i \le k$ *and* $m \in \mathbb{N}$*, denote by* $\mu_{i,m} = \mu(\Delta_m^{(i)})$*. Then under* $\mathbf{P}$*, conditional on the collection* $\{\mu_{i,m} : 1 \le i \le k, m \in \mathbb{N}\}$*, the measured real trees* $\{\boldsymbol{\Delta}_m^{(i)} : 1 \le i \le k, m \in \mathbb{N}\}$ *are independent and* $\boldsymbol{\Delta}_m^{(i)}$ *has distribution* $\mathbf{P}^{(\mu_{i,m})}$*.*

The following is the analog of Algorithm 2 for the continuum trees.

Lemma 2.4 (Recurrence relation for $(\mathcal{G}_k)$)**.** *Let $k \geq 2$. Let $i_k = \min\{i \in \{1,\dots,k-1\} : \exists\, m \in \mathbb{N}$ such that $V_k \in \Delta_m^{(i)}\}$ and let $m_k \in \mathbb{N}$ be the index such that $V_k \in \Delta_{m_k}^{(i_k)}$. Then under $\mathbf{P}$, we have a.s.*

$$\mathrm{Decomp}(\mathcal{G}_k, \{V_1', \dots, V_{k-1}'\}) = \sum_{\substack{1 \leq i \leq k: \\ i \neq i_k}} \sum_{m \in \mathbb{N}} \delta_{\left(h_m^{(i)},\, \boldsymbol{\Delta}_m^{(i)}\right)} + \sum_{\substack{m \in \mathbb{N}: \\ m \neq m_k}} \delta_{\left(h_m^{(i_k)},\, \boldsymbol{\Delta}_m^{(i_k)}\right)} + \delta_{\left(h_{m_k}^{(i_k)},\, \mathrm{cut}(\boldsymbol{\Delta}_{m_k}^{(i_k)}; V_k)\right)}.$$

Note that in the above formula, $\mathrm{cut}(\Delta_{m_k}^{(i_k)}; V_k)$ is well-defined under $\mathbf{P}$ thanks to Corollary 2.3. Lemma 2.4 is easily seen to hold true by comparing the definitions of $\mathcal{G}_k$ and $\mathcal{G}_{k-1}$.

The following observation constitutes the foundations of our partial reconstructions.

Lemma 2.5. *Let $k \geq 1$ and let η, η' be two independent points of $\mathcal{T}$ sampled according to the distribution $\mu(\cdot)/\mu(\mathcal{T})$. We have the following.*

a) Set $\mathcal{I}_k(\eta, \eta') = \{(i,m) : 1 \leq i \leq k, m \in \mathbb{N}, \Delta_m^{\circ,(i)} \cap [\![\eta, \eta']\!]_{\mathcal{T}} \neq \varnothing\}$. Then $\mathbf{E}(|\mathcal{I}_k(\eta, \eta')|) < \infty$.

b) For $(i,m) \in \mathcal{I}_k(\eta, \eta')$, set $\mathrm{R}'_{i,m} = \mu(\Delta_m^{(i)})^{-\frac{1}{2}} \cdot \ell(\Delta_m^{\circ,(i)} \cap [\![\eta, \eta']\!]_{\mathcal{T}})$. Under $\mathbf{P}$, conditional on the set $\mathcal{I}_k(\eta, \eta')$, $\{\mathrm{R}'_{i,m} : (i,m) \in \mathcal{I}_k(\eta, \eta')\}$ are independent Rayleigh random variables which are independent of the collection $\{\mu(\Delta_m^{(i)}) : 1 \leq i \leq k, m \in \mathbb{N}\}$.

Proof. Proof of a). For $i \geq 1$, let b_i be the unique point of $\mathcal{T}$ satisfying $[\![b_i, V_i]\!] = [\![\eta, V_i]\!] \cap [\![\eta', V_i]\!]$, and let $\tau_i = \inf\{t > 0 : \mathcal{P} \cap ([0,t] \times [\![b, V_i]\!]) \neq \varnothing\}$, the time of the first cut on $[\![b_i, V_i]\!]$. Observe that τ_i is an exponential random variable of parameter $d(V_i, b_i)$. Let $\tau = \max_{1 \leq i \leq k} \tau_i$, which has finite expectation under $\mathbf{P}$. Denote by N the cardinality of $\mathcal{P} \cap ([0,\tau] \times [\![\eta, \eta']\!])$; then N is distributed as a Poisson random variable of rate $\tau \cdot d(\eta, \eta')$. Note that $|\mathcal{I}_k(\eta, \eta')|$ is stochastically dominated by N, since for all $t > \tau$, $[\![\eta, \eta']\!] \cap \mathcal{T}_i(t) = \varnothing$ for all $i = 1, \dots, k$. This yields $\mathbf{E}(|\mathcal{I}_k(\eta, \eta')|) < \infty$.

Proof of b). The case $k = 1$ is a consequence of Theorem 5.1 in [4]. The general case follows by adapting the arguments there and we omit the formal proof. □

Lemma 2.5 says that $[\![\eta, \eta']\!]$ only intersects a finite sub-collection of $\{\Delta_m^{(i)} : 1 \leq i \leq k, m \in \mathbb{N}\}$. This suggests that to reverse the mapping of the k-partial cut tree $\boldsymbol{\mathcal{T}} \mapsto \mathrm{cut}(\mathcal{T}; V_1, \dots, V_k)$, which boils down to "reconstructing" the distance $d(\eta, \eta')$ from the partial cut tree, we should first sample the collection $\{\Delta_m^{(i)} : (i,m) \in \mathcal{I}_k(\eta, \eta')\}$ from the k-spine decomposition of $\mathcal{G}_k$. In the next section, we develop this idea into a definition of the partial shuffle transforms.

3 The shuffle transform

In this section, we give the definition of the shuffle transform by generalizing the construction in Section 1.2. Recall that $\boldsymbol{\mathcal{T}} = (\mathcal{T}, d, \mu, \rho)$ is a measured real tree and $\mathbf{P}$ is the law of the Brownian CRT. We aim at defining for each $k \geq 1$, a (random) semi-infinite matrix $\Gamma_k = (\gamma_k(i,j) : 1 \leq i, j < \infty)$ which will play the role of the k-partial reversal for $\boldsymbol{\mathcal{T}}$. Indeed, $\gamma_k(i,j)$ will represent the distance between two independent uniform points η_i, η_j obtained from the k-partial reversal. The main theorem (Theorem 3.2) then states that under $\mathbf{P}$, the sequence $(\gamma_k(i,j))_{k \geq 1}$ converges almost surely to a limit $\gamma_\infty(i,j)$, for all (i,j). Moreover, the limit $(\gamma_\infty(i,j) : 1 \leq i, j < \infty)$ characterizes a measured real tree which will be the image of $\mathcal{T}$ by the shuffle transform.

To define Γ_k, we do the following. First, for each k, we sample a collection of random points or marks, in an analogous way as we have done for Cayley trees. We then explain

how to build a path between two independent points in the 1-partial reversal using these marks, based on an intuitive idea illustrated in Fig. 3. Relying on a recursive procedure, this construction is then extended to more general partial reversals, which gives us the definition of Γ_k.

Sampling the marks. Let $(V_i)_{i\geq 1}$ be a sequence of i.i.d. points of $\mathcal{T}$ whose common distribution is $\mu(\cdot)/\mu(\mathcal{T})$. For $k \geq 1$, write $\mathbf{V}_k = \{V_1, \ldots, V_k\}$ and then $\mathcal{S}_k = \mathrm{Span}(\mathcal{T}, \mathbf{V}_k)$. Recall from (2.1) the notation $\mathrm{S}(\mathcal{T}, u \,|\, v)$. Set $\mathcal{T}_1 = \mathcal{T}$. For $k \geq 2$, with probability 1, there is a unique connected component of $\mathcal{T} \setminus \mathcal{S}_{k-1}$ containing V_k; set $\mathcal{T}_k$ to be the smallest rooted subtree of $\mathcal{T}$ containing that connected component. We define the sequences $(\rho_k)_{k\geq 1}$ and $(\mathcal{M}_k)_{k\geq 1}$ as follows:

- For each $k \geq 1$, let $\rho_k \in \mathcal{T}_k$ be a random point having the following distribution

$$\forall \text{ Borel set } B \subset \mathcal{T}: \ \mathbf{P}(\rho_k \in B) = \mu|_{\mathcal{T}_k}(B), \quad \text{where } \mu|_A := \frac{\mu(\cdot \cap A)}{\mu(A)}, \ A \subset \mathcal{T}. \tag{3.1}$$

- For each $k \geq 1$, let $\{C^\circ_{k,i} : i \in \mathbb{N}\}$ be the collection of the connected components of $\mathcal{T}_k \setminus \mathrm{Span}(\mathcal{T}_k, \{V_k\})$ and let $C_{k,i}$ be the smallest rooted subtree of $\mathcal{T}$ containing $C^\circ_{k,i}$. Note that $b_{k,i} := \mathrm{root}(C_{k,i})$ is the only element of $\mathcal{S}_k \cap C_{k,i}$. For each $i \geq 1$, let $U_{k,i} \in \mathrm{S}(\mathcal{T}_k, b_{k,i} \,|\, V_k)$ be a random point with distribution $\mu|_{\mathrm{S}(\mathcal{T}_k,\, b_{k,i} \,|\, V_k)}$. Observe that $\{C_{k,i} : i \geq 1\}$ is a collection of disjoint rooted subtrees of $\mathcal{T}_k$ and almost surely $U_i \in C_{k,i'}$ for some $i' \neq i$. We define $\mathcal{M}_k$ to be the collection

$$\mathcal{M}_k = \{(C_{k,i}, U_{k,i}) : i \in \mathbb{N}\}. \tag{3.2}$$

Building paths from the marks. Suppose that we have a collection

$$\mathcal{N} = \{(C_i,\, u_i) : i \geq 1\},$$

where $\{C_i : i \geq 1\}$ is a collection of disjoint rooted subtrees of $\mathcal{T}$ and $u_i \in C_{i'}$ for some $i' \neq i$, for all $i \geq 1$. Then for each $u \in \cup_{i\geq 1} C_i$, we introduce the following sequence

$$\chi^u(\mathcal{T}, \mathcal{N}) = \big\{(C^u_j, a^u_j, p^u_j) : j \geq 1\big\},$$

which is defined in the following inductive way: $a^u_1 = u$ and for each $j \geq 1$,

$$\text{let } i_j \text{ be the index s.t. } a^u_j \in C_{i_j}, \text{ then } \ C^u_j = C_{i_j}, \ p^u_j = \mathrm{root}(C_{i_j}) \text{ and } \ a^u_{j+1} = u_{i_j}. \tag{3.3}$$

Note that such an i_j exists since u and all u_i belong to $\cup_i C_i$ and i_j is unique as the C_i's are disjoint. Next, let $u, u' \in \cup_i C_i$ be two distinct points; set

$$\mathcal{I}(u, u'; \mathcal{T}, \mathcal{N}) = \inf\big\{j \geq 1 : C^u_j \in \{C^{u'}_{j'} : j' \geq 1\}\big\}, \tag{3.4}$$

with the convention that $\inf \varnothing = \infty$. Observe that $\mathcal{I}(u, u'; \mathcal{T}, \mathcal{N}) < \infty$ if and only if $\mathcal{I}(u', u; \mathcal{T}, \mathcal{N}) < \infty$; in that case, $C^u_{\mathcal{I}(u,u';\mathcal{T},\mathcal{N})} = C^{u'}_{\mathcal{I}(u',u;\mathcal{T},\mathcal{N})}$. Let $\tilde{\chi}(u, u'; \mathcal{T}, \mathcal{N})$ be a (possibly infinite) collection

$$\tilde{\chi}(u, u'; \mathcal{T}, \mathcal{N}) = \big\{\big(C^{u,u'}_m, a^{u,u'}_m, p^{u,u'}_m\big) : m \in \mathbb{N}, m \leq N^{u,u'}\big\}, \tag{3.5}$$

where $N^{u,u'} = |\tilde{\chi}(u, u'; \mathcal{T}, \mathcal{N})| \in \mathbb{N} \cup \{\infty\}$ and $\{C^{u,u'}_m : m \in \mathbb{N}, m \leq N^{u,u'}\}$ consists of disjoint rooted subtrees. We further requires $\tilde{\chi}(u, u'; \mathcal{T}, \mathcal{N})$ to satisfy the following conditions:

$$\begin{aligned} &\big\{C^{u,u'}_m : m \in \mathbb{N}, m \leq N^{u,u'}\big\} \\ &\quad = \big\{C^u_j : 1 \leq j < \mathcal{I}(u, u'; \mathcal{T}, \mathcal{N})\big\} \cup \big\{C^{u'}_j : 1 \leq j < \mathcal{I}(u', u; \mathcal{T}, \mathcal{N})\big\} \cup \big\{C^u_{\mathcal{I}(u,u';\mathcal{T},\mathcal{N})}\big\}, \end{aligned} \tag{3.6}$$

where the last term is taken to be $\varnothing$ if $\mathcal{I}(u,u';\mathcal{T},\mathcal{N})=\infty$; and

case 1: if $C_m^{u,u'}=C_j^u$ with $1\le j<\mathcal{I}(u,u';\mathcal{T},\mathcal{N})$, then $a_m^{u,u'}=a_j^u$ and $p_m^{u,u'}=p_j^u$;
case 2: if $C_m^{u,u'}=C_j^{u'}$ with $1\le j<\mathcal{I}(u',u;\mathcal{T},\mathcal{N})$, then $a_m^{u,u'}=a_j^{u'}$ and $p_m^{u,u'}=p_j^{u'}$;
case 3: if $\mathcal{I}(u,u';\mathcal{T},\mathcal{N})<\infty$ and $C_m^{u,u'}=C^u_{\mathcal{I}(u,u';\mathcal{T},\mathcal{N})}$, then $a_m^{u,u'}=a^u_{\mathcal{I}(u,u';\mathcal{T},\mathcal{N})}$ and $p_m^{u,u'}=a^{u'}_{\mathcal{I}(u',u;\mathcal{T},\mathcal{N})}$.

For the sake of definiteness, we can require the elements of $\tilde\chi(u,u';\mathcal{T},\mathcal{N})$ to be ranked in decreasing order according to $\mu(C_m^{u,u'})$, say, but the order of $\tilde\chi(u,u';\mathcal{T},\mathcal{N})$ is irrelevant for the rest of the construction. The above definition can be seen as an analog of the 1-partial shuffle transform for Cayley trees, which has been illustrated in Figure 3. Indeed, each C_i in the collection $\mathcal{N}$ can be understood as a subtree isolated from the rest of the tree due to the cutting at its root and the point u_i represents the neighbor of this cut in the original tree. Next, we choose a subcollection of subtrees $(C_m^{u,u'})_{m\ge1}$ which contain a portion of the path between u to u' in the original tree (i.e. the white trees in Figure 3). The subcollection is chosen by following this path: we start from $u=a_1^u$ and look for the nearest cut to u (i.e. p_1^u) on the path; the subtree rooted at p_1^u is then C_1^u and the neighbor of p_1^u is a_2^u, etc. Proceeding in the same way from u' yields another sequence $\chi^{u'}(\mathcal{T},\mathcal{N})$. Merging the two sequences up to the point where they coincide gives $\tilde\chi(u,u';\mathcal{T},\mathcal{N})$.

Defining partial reversals. Let $(\rho_k)_{k\ge1}$ and $(\mathcal{M}_k)_{k\ge1}$ be as defined in (3.1) and (3.2). Let $(\eta_i)_{i\ge2}$ be an independent sequence of i.i.d. points of common distribution $\mu|_{\mathcal{T}}$ and set $\eta_1=\rho_1$. Recall the definition of $\tilde\chi(u,u';\mathcal{T},\mathcal{N})$ from (3.5). Recall $\mathcal{T}_k$ is a rooted subtree containing V_k. For all $i,i'\ge1$ such that $i\ne i'$, we define

$$\chi_k^{i,i'}=\big\{(C_{k,m}^{i,i'},a_{k,m}^{i,i'},p_{k,m}^{i,i'}):m\in\mathbb{N},m\le N_k^{i,i'}\big\},\quad k\ge1, \tag{3.7}$$

in the following inductive way. Let $\chi_1^{i,i'}=\tilde\chi(\eta_i,\eta_{i'};\mathcal{T},\mathcal{M}_1)$, which is well-defined since almost surely we have $\eta_i,\eta_{i'}\in\mathcal{T}\backslash\mathcal{S}_1\subset\cup_iC_{k,i}$. Suppose that $\chi_k^{i,i'}$ has been defined. If $\mathcal{T}_{k+1}\notin\{C_{k,m}^{i,i'}:m\in\mathbb{N},m\le N_k^{i,i'}\}$, then we set $\chi_{k+1}^{i,i'}=\chi_k^{i,i'}$. Otherwise, let m_k be the index such that $\mathcal{T}_{k+1}=C_{k,m_k}^{i,i'}$; then we define $\chi_{k+1}^{i,i'}$ to be the collection

$$\chi_{k+1}^{i,i'}=\big\{(C_{k,m}^{i,i'},a_{k,m}^{i,i'},p_{k,m}^{i,i'}):m\in\mathbb{N}\backslash\{m_k\},m\le N_k^{i,i'}\big\}\cup\tilde\chi(a_{k,m_k}^{i,i'},\rho_{k+1};\mathcal{T}_{k+1},\mathcal{M}_{k+1}). \tag{3.8}$$

Note that $\tilde\chi(a_{k,m_k}^{i,i'},\rho_{k+1};\mathcal{T}_{k+1},\mathcal{M}_{k+1})$ is well-defined, since $\mathcal{T}_{k+1}$ is a rescaled version of $\mathcal{T}$ and almost surely $a_{k,m_k}^{i,i'},\rho_{k+1}\in\mathcal{T}_k\setminus\mathcal{S}_k$. Moreover, the role of ρ_{k+1} could be understood as follows: analogously to the discrete construction (Algorithm 4), the "replacement" of $\mathcal{T}_{k+1}$ will be rooted at ρ_{k+1}.

For each $k\ge1$, let us define a symmetric matrix $\Gamma_k=\big(\gamma_k(i,i'):1\le i,i'<\infty\big)$ where

$$\gamma_k(i,i)=0\quad\text{and}\quad\gamma_k(i,i')=\sum_{m\in\mathbb{N}:m\le N_k^{i,i'}}d(a_{k,m}^{i,i'},p_{k,m}^{i,i'}),\quad i\ne i'. \tag{3.9}$$

Lemma 3.1. *Let $k\ge1$, $i,i'\ge1$ and $\gamma_k(i,i')$ be defined as in* (3.9)*. Then under $\mathbf{P}$, we have $\gamma_k(i,i')<\infty$ almost surely.*

Theorem 3.2. *Under $\mathbf{P}$, the following statements hold almost surely.*

a) The sequence of matrices $\{\Gamma_k=(\gamma_k(i,j))_{1\le i,j<\infty}:k\ge1\}$ converges almost surely in the product topology of $\mathbb{R}^{\mathbb{Z}_+\times\mathbb{Z}_+}$. Denote by $\Gamma_\infty=(\gamma_\infty(i,j))_{1\le i,j<\infty}$ the almost sure limit.

b) There exists a measured real tree $\mathrm{shuff}(\mathcal{T}) = (\mathcal{H}, d_{\mathcal{H}}, \mu_{\mathcal{H}}, \rho_{\mathcal{H}})$ *and a sequence of its points* $(\varsigma_i)_{i\geq 1}$ *with* $\varsigma_1 = \rho_{\mathcal{H}}$ *such that*

$$d_{\mathcal{H}}(\varsigma_i, \varsigma_j) \,=\, \gamma_\infty(i,j), \quad \forall\, 1 \leq i, j < \infty.$$

Moreover, conditional on $\mathcal{H}$, $(\varsigma_i)_{i\geq 2}$ *is distributed as a sequence of i.i.d. points with common probability distribution* $\mu_{\mathcal{H}}$.

c) Set $\eta_0 = \rho$. *Recall from* (2.6) *the matrix* $(\delta(i,j))_{0\leq i,j<\infty}$. *We have*

$$\big(d(\eta_i,\eta_j), \gamma_\infty(i+1,j+1)\big)_{0\leq i,j<\infty} \text{ under } \mathbf{P} \;\overset{d}{=}\; \big(\delta(i,j), d(V_{i+1},V_{j+1})\big)_{0\leq i,j<\infty} \text{ under } \mathbf{P}. \tag{3.10}$$

In particular, this implies that the pair $(\mathcal{T}, \mathrm{shuff}(\mathcal{T}))$ *under* $\mathbf{P}$ *has the same distribution as* $(\mathrm{cut}(\mathcal{T}), \mathcal{T})$ *under* $\mathbf{P}$.

The mapping $\mathcal{T} \mapsto \mathrm{shuff}(\mathcal{T})$ is measurable. Indeed, in the above construction, we have performed a sequence of measurable operations on $\mathcal{T}$ with respect to the Gromov–Prokhorov topology. These operations can be seen as compositions of the following basic ones:

– sample two independent random points V and V' according to the mass measure μ and output the distances $d(\rho, V), d(V, V')$;
– denote by b the branch point of V and V', that is, the element of $[\![V, V']\!]$ minimising the distance to the root ρ; for an independent point ξ of law $\mu|_{\mathcal{T}}$, determine if $\xi \in \mathrm{S}(\mathcal{T}, b \,|\, V')$, which is the same as seeing if $d(\xi, V') < d(\xi, b) + d(b, V')$;
– output $\mu(\mathrm{S}(\mathcal{T}, b \,|\, V'))$, which a.s. equals $\lim_{k\to\infty} \frac{1}{k} \sum_{1\leq i\leq k} \mathbf{1}_{\{\xi_i \in \mathrm{S}(\mathcal{T}, b \,|\, V')\}}$, where $(\xi_i)_{i\geq 1}$ are independent points of common law $\mu|_{\mathcal{T}}$;
– determine if two rooted subtrees C and C' are identical, which reduces to comparing the two matrices $\mathrm{M}(C, d, \mu|_C, \mathrm{root}(C))$ and $\mathrm{M}(C', d, \mu|_{C'}, \mathrm{root}(C'))$.

Remark. As suggested by the discrete construction, the random points $(U_{k,i})_{k,i\geq 1}$ and $(\rho_k)_{k\geq 2}$ are the "traces" of the cuts left in $\mathrm{cut}(\mathcal{T})$. After completing an earlier version of this work, we have learned of the approach of Addario-Berry, Dieuleveut & Goldschmidt [5], who have made a rigorous statement out of this intuition. In their work, they enrich $\mathcal{T}$ with a collection of points formed by the cuts and a sequence of i.i.d. leaves $(\xi_i)_{i\geq 1}$; similarly, $\mathrm{cut}(\mathcal{T})$ is enriched with the images of the cuts and $(\xi_i)_{i\geq 1}$ by the cut tree transform; then they give a reconstruction procedure (different from ours) which allows them to reconstruct almost surely the enriched $\mathcal{T}$ from the enriched $\mathrm{cut}(\mathcal{T})$. Let us also remark that we can sample the randomness of the marks prior to the choice of $(V_i)_{i\geq 1}$. For this, simply sample for each branch point b of $\mathcal{T}$ a pair of independent points (r_0^b, r_1^b), each uniformly distributed in one of the two subtrees above b. After taking $(V_i)_{i\geq 1}$, we can define the sequence (ρ_k) and $(U_{k,i})_{k,i\geq 1}$ from $\{r_i^b : b \in \mathrm{Br}(\mathcal{T}), i = 0, 1\}$ as a function of (V_i). For example, $\rho_k = r_i^b$ where b is the root of $\mathcal{T}_k$ and r_i^b is the random point associated to b which belongs to $\mathcal{T}_k$ (since $\mathcal{T}_k$ is a subtree above b, there must exist one). The choice of $(U_{k,i})_{k,i\geq 1}$ is more tedious to put down; we omit it. This construction bears some resemblance to the one given in [5].

4 Partial reversals and their convergence

In this section, we give the proofs of Lemma 3.1 and Theorem 3.2.

4.1 Preliminary and proof of Lemma 3.1

Recall from (3.4) the notation $\mathcal{I}(u, u'; \mathcal{T}, \mathcal{N})$. We first recall the following result, obtained in [15].

Lemma 4.1 ([15], Theorem 6.1)**.** *If η, η' are two independent points of $\mathcal{T}$ with common distribution $\mu|_{\mathcal{T}}$, and if $\mathcal{M}_1$ is the collection in* (3.2)*, then* $\mathbf{E}[\mathcal{I}(\eta, \eta'; \mathcal{T}, \mathcal{M}_1) + \mathcal{I}(\eta', \eta\,; \mathcal{T}, \mathcal{M}_1)] < \infty$.

Lemma 4.2. *Let $k \geq 1$ and $i, i' \geq 1$, $i \neq i'$. Let $\chi_k^{i,i'}$ be as defined in* (3.7)*. For each $m \in \mathbb{N}, m \leq N_k^{i,i'}$, denote $\nu_{k,m} = \mu(C_{k,m}^{i,i'})$. Then $\mathbf{E}[N_k^{i,i'}] < \infty$. Moreover under $\mathbf{P}$, conditional on the collection $\{\nu_{k,m} : 1 \leq m \leq N_k^{i,i'}\}$, the collection $\chi_k^{i,i'}$ consists of independent elements which are distributed as follows:*
(a) $C_{k,m}^{i,i'}$ has the law $\mathbf{P}^{(\nu_{k,m})}$; and given $C_{k,m}^{i,i'}$:
(b) $a_{k,m}^{i,i'}$ is an independent point of law $\mu|_{C_{k,m}^{i,i'}}$,
(c) $p_{k,m}^{i,i'}$ is either the root of $C_{k,m}^{i,i'}$ or another independent point of law $\mu|_{C_{k,m}^{i,i'}}$.

Proof. We proceed by induction on k. For $k = 1$, by definition, $\chi_1^{i,i'} = \tilde{\chi}(\eta_i, \eta_{i'}; \mathcal{T}, \mathcal{M}_1)$; then by the definition of the latter, $N_1^{i,i'} = \mathcal{I}(\eta_i, \eta_{i'}; \mathcal{T}, \mathcal{M}_1) + \mathcal{I}(\eta_{i'}, \eta_i\,; \mathcal{T}, \mathcal{M}_1) - 1$. It follows from Lemma 4.1 that $\mathbf{E}(N_1^{i,i'}) < \infty$. Recall from (2.2) the 1-spine decomposition of $\mathcal{T}$ with respect to V_1. By construction, $\{C_{k,m}^{i,i'} : 1 \leq m \leq N_1^{i,i'}\}$ is a sub-collection of $\{C_{1,i} : i \geq 1\}$, the closures of the connected components of $\mathcal{T} \setminus [\![\rho, V_1]\!]$. Moreover, from the definition of $(U_{1,i})$ and the definition in (3.3) we see that the event that $C_{1,i}$ belongs to this sub-collection only depends on its μ-mass. We then deduce from Corollary 2.3 that conditional on their masses, $C_{k,m}^{i,i'}, 1 \leq m \leq N_1^{i,i'}$, are independent, each one being a rescaled Brownian CRT. We next check that each $a_{k,m}^{i,i'}$ is a μ-random point restricted to $C_{k,m}^{i,i'}$. But this follows from the definition of $(U_{1,i})$, (3.3) and the definition of $a_{k,m}^{u,u'}$ on page 14. On the other hand, $p_{k,m}^{i,i'}$ is either the root of $C_{k,m}^{i,i'}$ (**case 1** & **2** on page 14) or another point independent of $a_{k,m}^{i,i'}$ with distribution $\mu|_{C_{k,m}^{i,i'}}$ (**case 3** on page 14). In this way, we verify the statements of the lemma for $k = 1$.

Now we assume that the lemma holds up to $k-1$, for some $k \geq 2$. Let us show that it also holds for k. Recall that $\mathcal{T}_k$ is the closure of the connected component of $\mathcal{T} \setminus \mathcal{S}_{k-1}$ which contains V_k. If $\mathcal{T}_k \notin \{C_{k-1,m}^{i,i'} : 1 \leq m \leq N_{k-1}^{i,i'}\}$, then the statement for k follows trivially from the induction hypothesis. If instead, $\mathcal{T}_k = C_{k-1,m_{k-1}}^{i,i'}$ with $1 \leq m_{k-1} \leq N_{k-1}^{i,i'}$, then by the inductive hypothesis, $\mathcal{T}_k$ is a rescaled Brownian CRT with total mass $\mu(\mathcal{T}_k)$ and $a_k := a_{k-1,m_{k-1}}^{i,i'}$ is a point of $\mathcal{T}_k$ with distribution $\mu|_{\mathcal{T}_k}$; moreover, ρ_k is another independent point of the same distribution, by (3.1). We can then apply the statements for the case $k = 1$ to $\tilde{\chi}(a_k, \rho_k; \mathcal{T}_k, \mathcal{M}_k)$ and find that (i) $\mathbf{E}[N^{a_k,\rho_k}] < \infty$; (ii) conditional on their masses, $C_m^{a_k,\rho_k}, 1 \leq m \leq N^{a_k,\rho_k}$, are independent and distributed as rescaled Brownian CRT; (iii) for each $1 \leq m \leq N^{a_k,\rho_k}$, $a_m^{a_k,\rho_k}$ is a μ-random point restricted to $C_m^{a_k,\rho_k}$ and $p_m^{a_k,\rho_k}$ is either its root or another independent point. Combined with (3.8) and the induction hypothesis, this leads to the statements for k. □

Proof of Lemma 3.1. It follows from Lemma 4.2 that the number of summands in (3.9) is finite. Then $\gamma_k(i,j) < \infty$, a.s. □

4.2 Convergence of $(\Gamma_k)_{k\geq 1}$

This subsection is devoted to proving the almost sure convergence of the sequence $\{\Gamma_k = (\gamma_k(i,j))_{1\leq i,j<\infty} : k \geq 1\}$ in the product topology. Note that since $(\eta_i)_{i\geq 1}$ is an i.i.d. sequence, it suffices to prove the convergence for the sequence $\{\gamma_k(1,2) : k \geq 1\}$.

4.2.1 A Markov chain representation of $(\gamma_k(1,2))_{k\geq 1}$

Let $\mathcal{S}^{\downarrow} = \{\mathbf{x} = (x_1, x_2, \dots) : x_1 \geq x_2 \geq \cdots \geq 0 \text{ and } \sum_{i\geq 1} x_i \leq 1\}$. For $\mathbf{x} \in \mathcal{S}^{\downarrow}$, its ℓ_1-norm is $\|\mathbf{x}\|_1 = \sum_i x_i$. Let $\mathcal{S}_f^{\downarrow}$ be the subset of $\mathcal{S}^{\downarrow}$ which consists of those $\mathbf{x} \in \mathcal{S}^{\downarrow}$ for which

there exists some $n \in \mathbb{N}$ such that $x_i = 0$ for all $i \geq n$.

For each $k \geq 1$, recall the collection $\chi_k^{1,2}$ from (3.7) and recall that $N_k := N_k^{1,2} = |\chi_k^{1,2}|$ is finite $\mathbf{P}$-a.s. according to Lemma 4.2. Then let $\mathbf{m}_k = (\mathrm{m}_{k,n} : n \geq 1)$ be the sequence obtained from $\{\mu(C_{k,m}^{1,2}) : 1 \leq m \leq N_k^{i,i'}\}$ by a re-ordering in decreasing order and completed with infinitely many 0. Observe that $\mathbf{m}_k \in \mathcal{S}_f^{\downarrow}$ since $m_{k,n} = 0$ for all $n > N_k$.

Proposition 4.3. *For $k \geq 1$, let $\mathbf{m}_k = (\mathrm{m}_{k,n})_{n\geq 1}$ be defined as above. Under $\mathbf{P}$, the sequence $(\mathbf{m}_k)_{k\geq 1}$ is a Markov chain taking values in $\mathcal{S}_f^{\downarrow}$ which evolves in the following way: for each $k \geq 1$,*

- *with probability $1 - \|\mathbf{m}_k\|_1$, $\mathbf{m}_{k+1} = \mathbf{m}_k$, and*
- *for $1 \leq n \leq N_k$, with probability $\mathrm{m}_{k,n}$, $\mathbf{m}_{k+1}$ is obtained by replacing in $\mathbf{m}_k$ the element $m_{k,n}$ by $m_{k,n} \cdot \widetilde{\mathbf{m}}$, where $\widetilde{\mathbf{m}}$ is an independent copy of $\mathbf{m}_1$, and then sorting the sequence thereby obtained in decreasing order.*

Moreover for each $k \geq 1$, there exists a sequence of positive real numbers $(\mathrm{R}_{k,n})_{1\leq n\leq N_k}$ such that

$$\gamma_k(1,2) = \sum_{1\leq n\leq N_k} \sqrt{m_{k,n}}\, \mathrm{R}_{k,n}\,. \tag{4.1}$$

Under $\mathbf{P}$ and given that $N_k = p \in \mathbb{N}$, $(\mathrm{R}_n^k)_{1\leq n\leq N_k}$ consists of p independent Rayleigh random variables which are independent of $\mathbf{m}_k$.

Proof. By the definition (3.8) of $\chi_k^{1,2}$, $\mathbf{m}_{k+1} = \mathbf{m}_k$ iff $V_{k+1} \in \{C_{k,m}^{1,2} : 1 \leq m \leq N_k\}$. We can readily check by an induction on k that $\{C_{k,m}^{1,2} : 1 \leq m \leq N_k\}$ is a sub-collection of $\{C_{1,i} : i \in \mathbb{N}\}$, the closures of the connected components of $\mathcal{T} \setminus \mathrm{Span}(\mathcal{T}, \mathbf{V}_k)$. On the other hand, V_{k+1} is a point independent of $\mathbf{V}_k$ with the law $\mu|_{\mathcal{T}}$. Thus, under $\mathbf{P}$, the event $V_{k+1} \in \{C_{k,m}^{1,2} : 1 \leq m \leq N_k\}$ takes place with probability $1 - \sum_{1\leq m\leq N_k} \mu(C_{k,m}^{1,2}) = 1 - \|\mathbf{m}_k\|_1$. Next, suppose that $V_{k+1} \in C_{k,m_k}^{1,2}$ for some index $1 \leq m_k \leq N_k$, which takes place with probability $\mu(C_{k,m_k}^{1,2})$. In that case, we have $\mathcal{T}_{k+1} = C_{k,m_k}^{1,2}$. We have seen in Lemma 4.2 that $\mathcal{T}_{k+1} = C_{k,m_k}^{1,2}$ is a rescaled Brownian CRT and that the points $a_{k,m_k}^{1,2}, \rho_{k+1}$ are independent and distributed according to $\mu|_{\mathcal{T}_{k+1}}$. We then deduce from (3.8) the distribution of $\mathbf{m}_{k+1}$ in this case. In this way, we have checked the transition probabilities of $(\mathbf{m}_k)_{k\geq 1}$. The expression (4.1) is a direct consequence of (3.9) and the statements (a-c) in Lemma 4.2. □

4.2.2 Polynomial decay of a self-similar fragmentation chain

The dynamic of $(\mathbf{m}_k)_{k\geq 1}$ as described in Proposition 4.3 is that of a discrete-time self-similar fragmentation chain with index of self-similarity 1. Self-similar fragmentation chains are studied in Bertoin [9, Chapter 1] and a series of papers including Bertoin and Gnedin [12]. Here, we apply their results on the asymptotic behavior of fragmentation chains in order to obtain the following.

Lemma 4.4. *Let $(\mathbf{m}_k)_{k\geq 1}$ be as in Proposition 4.3. There exists some $\alpha \in (0,1)$ such that*

$$\lim_{k\to\infty} k^{\alpha} \|\mathbf{m}_k\|_1 = 0, \quad \mathbf{P}\text{-almost surely}\,.$$

The proof of Lemma 4.4 will occupy the rest of this part. Note that Chapter 1 of [9] studies the continuous-time version of self-similar fragmentation chains, which can be related to the discrete-time version by a time-change. Let us first recall some terminology from there.

Continuous-time fragmentation chain. Let ϖ denote the law of $\mathbf{m}_1$ under $\mathbf{P}$. We consider a self-similar fragmentation chain $\{\tilde{\mathbf{Z}}(t) = (\tilde{Z}_i(t)_{i\geq 1} : t \geq 0\}$ with index of self-similarity 1 and dislocation measure ϖ starting from the initial state $(1, 0, 0, \cdots)$ as defined in [9, Definition 1.1], which is a continuous-time Markov chain taking values in $\mathcal{S}^{\downarrow}$ whose total jump rate at time t is $\|\tilde{\mathbf{Z}}(t)\|_1 \in [0, 1]$. Using standard facts about Poisson processes and the construction of fragmentation chains in [9], we can construct on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ the following processes:

- a Poisson process of rate 1 which jumps at times $\tau_1 < \tau_2 < \cdots < \tau_i < \cdots$;
- a process $(\mathbf{Z}(t))_{t\geq 0}$ having the same distribution as $(\tilde{\mathbf{Z}}(t))_{t\geq 0}$ such that the set of discontinuities of $t \mapsto \tilde{\mathbf{Z}}(t)$ is a subset of $(\tau_i : i \geq 1)$.

Then by Proposition 4.3, we have

$$(\mathbf{Z}(\tau_k))_{k\geq 1} \text{ under } \mathbb{P} \overset{d}{=} (\mathbf{m}_k)_{k\geq 1} \text{ under } \mathbf{P}\,. \tag{4.2}$$

Note that, in particular, we have $\mathbf{Z}(\tau_1) \overset{d}{=} \mathbf{m}_1 = (m_{1,n})_{n\geq 1}$.

Let p^* be the *Malthusian exponent* associated with ϖ, namely, $p^* \in [0, 1]$ is such that

$$\mathbf{E}\left[\sum_{1\leq n\leq N_1} \mathrm{m}_{1,n}^{p^*}\right] = 1$$

(see [9, Section 1.2.2]). The following is a consequence of Theorem 1 of [12].

Lemma 4.5. *Let $\{\mathbf{Z}(t) = (Z_i(t)_{i\geq 1} : t \geq 0\}$ be a self-similar fragmentation chain with index of self-similarity 1 and dislocation measure ϖ which is defined on $(\Omega, \mathcal{F}, \mathbb{P})$ as above. Then $p^* \in (0, 1)$ and for any $\delta \in (0, 1)$,*

$$\limsup_{t\to\infty} t^{\,\delta(1-p^*)} \sum_{i\geq 1} Z_i(t) = 0 \quad \mathbb{P}\text{-almost surely}\,. \tag{4.3}$$

Proof. First, let us show that $p^* \in (0, 1)$. Recall N_1 is the number of non zero elements of $\mathbf{m}_1$, which is also equal to $\mathcal{I}(\eta_1, \eta_2; \mathcal{T}, \mathcal{M}_1) + \mathcal{I}(\eta_2, \eta_1; \mathcal{T}, \mathcal{M}_1) - 1$ by our previous definitions. Then Lemma 4.1 tells that $N_1 < \infty$, $\mathbf{P}$-a.s. On the other hand, $\mathbf{m}_1$ is a sub-collection of the μ-masses of the connected components of $\mathcal{T}\setminus \mathrm{Span}(\mathcal{T}, \{V_1\})$. Therefore, we must have $\|\mathbf{m}_1\|_1 < 1$, $\mathbf{P}$-a.s. This shows $p^* < 1$. To see why $p^* > 0$, note that $N_1 = 1$ if and only if η_1 and η_2 are found in the same component of $\mathcal{T}\setminus \mathrm{Span}(\mathcal{T}, \{V_1\})$, which occurs with probability strictly smaller than 1. Then on the event $N_1 \geq 2$, we have $\lim_{p\to 0+} \sum_{1\leq n\leq N_1} \mathrm{m}_{1,n}^{p} \geq 2$. This shows $p^* > 0$.

We introduce

$$Y(t) := \sum_{i\geq 1} Z_i^{p^*}(t), \tag{4.4}$$

which is a strictly positive martingale by the choice of p^*. Denote by $Y(\infty) \in [0, \infty)$ its almost sure limit. Next, we check that the hypothesis of Theorem 1 in [12] are fulfilled: with the notation there, we see that $\beta^* = p^*$, $\beta_a \leq 0$ (since $\mathbf{E}[N_1] < \infty$ by Lemma 4.1), $\sigma = \varpi$ is diffuse and the conditions (1) on page 577 all hold for ϖ. Then, by the above mentioned theorem, for every $k \geq 1$, there exists a constant $C_k \in (0, \infty)$ such that

$$\sup_{t\geq 0} t^k \cdot \mathbb{E}\left[\sum_{i\geq 1} Z_i^{p^*+k}(t)\right] \leq C_k,$$

from which it follows immediately that $\sup_{t\geq 0} t^k \cdot \mathbb{E}Z_1^{p^*+k}(t) \leq C_k$. Now, for any $\delta \in (0, 1)$ and $\epsilon > 0$, by Markov's inequality, we obtain that at time $m > 0$,

$$\mathbb{P}\big(Z_1(m) \geq \epsilon m^{-\delta}\big) \leq \frac{m^k \cdot \mathbb{E}[Z_1^{p^*+k}(m)]}{\epsilon^{p^*+k} m^{(1-\delta)k-\delta p^*}} \leq \frac{C_k}{\epsilon^{p^*+k} m^{(1-\delta)k-\delta p^*}}. \tag{4.5}$$

Choosing k large enough so that $k > (1+\delta p^*)/(1-\delta)$, one sees that (4.5) implies that $\sum_{m\geq 1}\mathbb{P}(Z_1(m)\geq \epsilon m^{-\delta})<\infty$ and $\limsup_{m\to\infty} m^{\delta}Z_1(m)\leq \epsilon$ almost surely, by the Borel–Cantelli lemma. As ϵ was chosen arbitrary, we then obtain that $m^{\delta}Z_1(m)\to 0$ a.s. as $m\to\infty$, and $t^{\delta}Z_1(t)\to 0$ a.s. as $t\to\infty$ as well by monotonicity. Now note that for any $t\geq 0$,

$$\sum_{i\geq 1} Z_i(t) = \sum_{i\geq 1} Z_i^{1-p^*}(t)\cdot Z_i^{p^*}(t) \leq Z_1^{1-p^*}(t)\cdot Y(t).$$

Then, for any $\delta\in(0,1)$,

$$\limsup_{t\to\infty} t^{\delta(1-p^*)}\sum_{i\geq 1} Z_i(t) \leq Y(\infty)\cdot \limsup_{t\to\infty}(t^{\delta}Z_1(t))^{1-p^*}=0,$$

almost surely, since $p^*\in(0,1)$. This completes the proof of the lemma. □

Proof of Lemma 4.4. We work on a probability space where the equality in (4.2) holds almost surely. By the strong law of large numbers, we have $\tau_k/k\to 1$ almost surely as $k\to\infty$. Therefore, we obtain that for any $\delta\in(0,1)$,

$$\limsup_{k\to\infty} k^{\,\delta(1-p^*)}\|\mathbf{m}_k\|_1 = \limsup_{k\to\infty} k^{\,\delta(1-p^*)}\sum_{i\geq 1} Z_i(\tau_k)=0, \quad \text{almost surely,}$$

by (4.3). This proves Lemma 4.4 by taking $\alpha=\delta(1-p^*)$. □

4.2.3 Concentration around the conditional expectations

In this part, we rely on Lemma 4.4 and the exponential tail of a Rayleigh random variable to show the following result.

Lemma 4.6. $\mathbf{P}$*-almost surely,* $\gamma_k(1,2)-\mathbf{E}[\gamma_k(1,2)\,|\,\mathbf{m}_k]\to 0$ *as* $k\to\infty$.

Let R be a Rayleigh random variable defined on some probability space $(\Omega,\mathcal{F},\mathbb{P})$, namely, R has density $xe^{-x^2/2}\mathbf{1}_{\{x>0\}}$. Then, one readily verifies that $\mathrm{R}-\mathbb{E}[\mathrm{R}]$ is *sub-Gaussian* in the sense that there exists a constant v such that for every $\lambda\in\mathbb{R}$, one has

$$\log\mathbb{E}[e^{\lambda(\mathrm{R}-\mathbb{E}[\mathrm{R}])}]\leq \tfrac{1}{2}\,\lambda^2 v.$$

(See [14, Theorem 2.1, p. 25].) We may thus apply concentration results for sub-Gaussian random variables such as the ones presented in Section 2.3 of [14]. To that end, we set for each $k\geq 1$,

$$\sigma_k := \gamma_k(1,2)-\mathbf{E}[\gamma_k(1,2)\,|\,\mathbf{m}_k] = \sum_{1\leq n\leq N_k}\sqrt{\mathrm{m}_{k,n}}\,\big(\mathrm{R}_{k,n}-\mathbf{E}[\mathrm{R}_{k,n}]\big),$$

by (4.1), where according to Proposition 4.3, conditional on $\mathbf{m}_k$, $(\mathrm{R}_{k,n}, 1\leq n\leq N_k)$ are i.i.d. copies of R. Therefore, by the above mentioned concentration results, we find

$$\mathbf{P}(|\sigma_k|\geq\epsilon\,|\,\mathbf{m}_k)\leq 2\exp\Big(-\frac{\epsilon^2}{2v\|\mathbf{m}_k\|_1}\Big),\quad \forall\,\epsilon>0. \tag{4.6}$$

If $(A_k,B_k,C_k)_{k\geq 1}$ are sequence of events satisfying that $A_k\subset B_k\cup C_k$ for each $k\geq 1$, then it is elementary that $\mathbf{P}(\limsup_k A_k)\leq \mathbf{P}(\limsup_k B_k)+\mathbf{P}(\limsup_k C_k)$. Here, we take

$$A_k=\{|\sigma_k|\geq\epsilon\},\quad B_k=A_k\cap\{\|\mathbf{m}_k\|_1\leq k^{-\alpha}\},\quad C_k=\{\|\mathbf{m}_k\|_1>k^{-\alpha}\}$$

with the same α as in Lemma 4.4. Then $\mathbf{P}(\limsup_k C_k)=0$ by Lemma 4.4. On the other hand, we deduce from (4.6) that

$$\sum_{k\geq 1}\mathbf{P}(B_k)=\sum_{k\geq 1}\mathbf{E}\Big[\mathbf{P}\big(|\sigma_k|\geq\epsilon\,|\,\mathbf{m}_k\big)\cdot\mathbf{1}_{\{\|\mathbf{m}_k\|_1\leq k^{-\alpha}\}}\Big]\leq\sum_{k\geq 1}2e^{-\epsilon^2 k^{\alpha}/(2v)}<\infty,$$

which entails that $\mathbf{P}(\limsup_k B_k) = 0$ by the Borel–Cantelli lemma. It follows that $\mathbf{P}(\limsup_k A_k) = 0$, which means $\limsup_k |\sigma_k| < \epsilon$ almost surely. Since $\epsilon > 0$ was arbitrary, the proof of Lemma 4.6 is now complete.

4.2.4 A coupling via partial cut trees

Thanks to Lemma 4.6, the last step to show the convergence of $(\gamma_k(1,2))_{k\geq 1}$ consists in proving that under $\mathbf{P}$, $\mathbf{E}[\gamma_k(1,2)\,|\,\mathbf{m}_k]$ converges almost surely as $k \to \infty$. For this, we rely on a coupling of $(\mathbf{m}_k)_{k\geq 1}$ with a sequence of masses defined on the partial cut trees.

Let $k \geq 1$. We recall the following notation in Lemma 2.5: η, η' are two independent points of $\mathcal{T}$ with distribution $\mu|_{\mathcal{T}}$; the collection $\{\Delta_m^{(i)} : (i,m) \in \mathcal{I}_k(\eta,\eta')\}$ consists of the subsets of $\mathcal{G}_k$ (the k-partial cut tree of $\boldsymbol{\mathcal{T}}$) which intersect the geodesic $[\![\eta,\eta']\!]_{\mathcal{T}}$. Denote $N'_k = |\mathcal{I}_k(\eta,\eta')|$. Let $\{(\mathrm{m}'_{k,n}, \mathrm{R}'_{k,n}) : 1 \leq n \leq N'_k\}$ be the sequence obtained from $\{(\mu(\Delta_m^{(i)}), \mathrm{R}_{i,m}) : (i,m) \in \mathcal{I}_k(\eta,\eta')\}$ by arranging the first coordinates in decreasing order. It follows from Lemma 2.5 that

$$D := d(\eta,\eta') = \sum_{1\leq n\leq N'_k} \sqrt{\mathrm{m}'_{k,n}}\,\mathrm{R}'_{k,n}\,, \tag{4.7}$$

where, given $N'_k = p \in \mathbb{N}$, $(\mathrm{R}'_{k,n} : 1 \leq n \leq N'_k)$ are p independent Rayleigh random variables which are independent of $\mathbf{m}'_k := (\mathrm{m}'_{k,n})_{n\geq 1}$ with $m'_{k,n} = 0$ for $n > N'_k$.

Lemma 4.7. *Under* $\mathbf{P}$, $(\mathbf{m}'_k)_{k\geq 1}$ *has the same distribution as the Markov chain* $(\mathbf{m}_k)_{k\geq 1}$.

Proof. We first show that $\mathbf{m}_1 \overset{d}{=} \mathbf{m}'_1$. Recall from (3.2) the collection $\mathcal{M}_1$. For $i = 1, 2$, recall from (3.3) the definition of $\chi^{\eta_i}(\mathcal{T}, \mathcal{M}_1)$. Note that it tells that the sequence $(\mu(C_j^{\eta_i}) : j \geq 1)$ has the following distribution. Conditional on $(\boldsymbol{\mathcal{T}}, V_1)$, $\mu(C_1^{\eta_i})$ is a random element of $\{\mu(C_{1,m}) : m \in \mathbb{N}\}$ chosen by size-biased sampling, that is, for $m \in \mathbb{N}$,

$$\mathbf{P}\big(\mu(C_1^{\eta_i}) = \mu(C_{1,m})\,|\,\boldsymbol{\mathcal{T}}, V_1\big) = \mathbf{P}(\eta_i \in C_{1,m}) = \mu(C_{1,m}),$$

since $\mathbf{P}$-a.s, the μ-masses of $C_{1,m}, m \in \mathbb{N}$, are all distinct. More generally for $j \geq 1$, given $d(\rho, p_{j-1}^{\eta_i})$, $\mu(C_j^{\eta_i})$ is chosen from $\{\mu(C_{1,m}) : m \in \mathbb{N}, d(\rho, \mathrm{root}(C_{1,m})) > d(\rho, p_{j-1}^{\eta_i})\}$ by size-biased sampling. Moreover, conditional on $\boldsymbol{\mathcal{T}}$ and V_1, the two sequences $(\mu(C_j^{\eta_1}) : j \in \mathbb{Z}_+)$ and $(\mu(C_j^{\eta_2}) : j \in \mathbb{Z}_+)$ are independent, since η_1 and η_2 are independent. Note that we also have $\mathcal{I}(\eta_1,\eta_2;\mathcal{T},\mathcal{M}_1) = \inf\{j \in \mathbb{N} : \exists\, j' \in \mathbb{N} \text{ s.t. } \mu(C_j^{\eta_1}) = \mu(C_{j'}^{\eta_2})\}$, by the fact that $\mu(C_{1,m}), m \geq 1$, are a.s. distinct. It follows that we can write $\mathbf{m}_1 = F(\boldsymbol{\mathcal{T}}, V_1; \eta_1, \eta_2)$ for some measurable function F, outside a $\mathbf{P}$-null set. Next, we show that a.s. $\mathbf{m}'_1 = F(\mathcal{G}_1, V'_1; \eta, \eta')$. But this is a consequence of Theorem 5.1 in [4]. Indeed, recall from Proposition 2.2 the spinal decomposition $\mathrm{Decomp}(\mathcal{G}_1, \{V'_1\})$. Denote by $\{\tilde{\Delta}_{m_j} : j \in \mathbb{Z}_+\}$ the sequence of those $\Delta_m^{(1)}, m \geq 1$, which satisfy $\Delta_m^{(1)} \cap [\![\eta, V_1]\!] \neq \varnothing$ such that $h_{m_1}^{(1)} < h_{m_2}^{(1)} < \cdots$. Then the above mentioned theorem identifies the distribution of $\{\tilde{\Delta}_{m_j} : j \in \mathbb{Z}_+\}$ as that of $\{\mu(C_j^{\eta_1}) : j \in \mathbb{N}\}$. This then entails that $\mathbf{m}'_1$ can be written as $F(\mathcal{G}_1, V'_1; \eta, \eta')$ with the same F as before. Then we have $\mathbf{m}_1 \overset{d}{=} \mathbf{m}'_1$, since $(\mathcal{G}_1, V'_1) \overset{d}{=} (\mathcal{T}, V_1)$ by Proposition 2.2.

The rest of the proof is very similar to that of Proposition 4.3. The main differences lie in that we use Corollary 2.3 and Lemma 2.4 instead of Lemma 4.2 and (3.8). We omit the details. □

As $(\mathbf{m}'_k)_{k\geq 1}$ has the same distribution as $(\mathbf{m}_k)_{k\geq 1}$, Lemma 4.4 also holds for $(\mathbf{m}'_k)_{k\geq 1}$. Furthermore, combined with Lemma 4.7, the concentration arguments already used in the course of the proof of Lemma 4.6 imply that, a.s.,

$$D - \mathbf{E}[D\,|\,\mathbf{m}'_k] = \sum_{1\leq n\leq N'_k} \sqrt{\mathrm{m}'_{k,n}} \cdot \big(\mathrm{R}'_{k,n} - \mathbf{E}[\mathrm{R}'_{k,n}]\big) \xrightarrow{k\to\infty} 0, \quad \mathbf{P}\text{-a.s.}\,.$$

This entails that

$$\mathbf{E}[D \,|\, \mathbf{m}'_k] = \mathbf{E}[\mathrm{R}'_{1,1}] \cdot \sum_{1 \le n \le N'_k} \sqrt{\mathrm{m}'_{k,n}}\,, \quad k \ge 1,$$

converges almost surely to D. Since the sequence $(\mathbf{E}[\gamma_k(1,2) \,|\, \mathbf{m}_k])_{k\ge 1}$ has the same distribution by Lemma 4.7, it also converges almost surely to some random variable that has the same distribution as D. Combined with Lemma 4.6, we have shown the following.

Proposition 4.8. *Under* $\mathbf{P}$*, there exists a Rayleigh random variable* $\gamma_\infty(1,2)$ *such that* $\gamma_k(1,2) \to \gamma_\infty(1,2)$ *almost surely.*

4.3 Proof of Theorem 3.2

By exchangeability and by Proposition 4.8, for all $i, i' \ge 1$, $i \ne i'$, there exists a Rayleigh random variable $\gamma_\infty(i,j)$ such that $\lim_{k\to\infty} \gamma_k(i,j) = \gamma_\infty(i,j)$ $\mathbf{P}$-almost surely. This proves the point a) of Theorem 3.2. For the rest of the statement, let us begin with the proof of (3.10).

We start with a distributional identity for Γ_k. Let $(\xi_i)_{i\ge 1}$ be an independent sequence of i.i.d. points of $\mathcal{T}$ with common distribution $\mu|_{\mathcal{T}}$ and set $\xi_0 = \rho$. Then Equation (6.7) and Equation (6.1) in [15] entail that

$$\big(d(\eta_i,\eta_j), \gamma_k(i+1,j+1)\big)_{0\le i,j<\infty} \text{ under } \mathbf{P} \overset{d}{=} \big(d_{\mathcal{G}_k}(\xi_i,\xi_j), d(\xi_{i+1},\xi_{j+1})\big)_{0\le i,j<\infty} \text{ under } \mathbf{P}, \tag{4.8}$$

in the case where $k = 1$. However, the arguments in [15, Section 6] can be readily adapted to a general proof for $k \ge 1$. Therefore, (4.8) holds for any $k \ge 1$. On the other hand, for $k \ge 1$, recall from Proposition 2.2 that conditional on $\mu_{\mathcal{G}_k}$, the points $V'_i, 1 \le i \le k$, of $\mathcal{G}_k$ are distributed as k independent points with common distribution $\mu_{\mathcal{G}_k}$. Then,

$$\Big(\big(d_{\mathcal{G}_k}(\xi_i,\xi_j)\big)_{0\le i,j\le k}, \big(d(\xi_i,\xi_j)\big)_{i,j\ge 0}\Big) \overset{d}{=} \Big(d_{\mathcal{G}_k}(V'_i,V'_j)_{0\le i,j\le k}, \big(d(V_i,V_j)\big)_{i,j\ge 0}\Big). \tag{4.9}$$

Now let $n \ge 1$ and let $G, H : \mathbb{R}^{\mathbb{Z}_+\times\mathbb{Z}_+} \to \mathbb{R}$ be two continuous bounded functions with respect to the product topology which are supported on $\mathbb{R}^{n\times n}$. Set V'_0 to be the root of $\mathcal{G}_k$. We have used the same sequence $(V'_i)_{i\ge 0}$ for different $\mathcal{G}_k$; this will not cause confusion, since by Proposition 2.2, $\mathrm{Span}(\mathcal{G}_k; \{V'_1,\dots,V'_k\})$ is isometric to $\mathrm{Span}(\mathcal{G}; \{V'_1,\dots,V'_k\})$. In particular, we have the fact that $d_{\mathcal{G}_k}(V'_i,V'_j) = \delta(i,j)$, $0 \le i,j \le k$, for all k. We then obtain from the convergence of $(\Gamma_k)_{k\ge 1}$, Equations (4.8), (4.9) and this fact that

$$\begin{aligned}
&\mathbf{E}\Big[G\big((d(\eta_i,\eta_j))_{0\le i,j<\infty}\big) H\big((\gamma_\infty(i,j))_{1\le i,j<\infty}\big)\Big] \\
&\qquad = \lim_{k\to\infty} \mathbf{E}\Big[G\big((d(\eta_i,\eta_j))_{0\le i,j<\infty}\big) H\big((\gamma_k(i,j))_{1\le i,j<\infty}\big)\Big] \\
&\qquad = \mathbf{E}\Big[G\big((\delta(i,j))_{0\le i,j<\infty}\big) H\big((d(V_i,V_j))_{1\le i,j<\infty}\big)\Big],
\end{aligned}$$

which proves (3.10), as n is arbitrary.

Next, we follow Aldous [6] to construct the measured real tree $\mathrm{shuff}(\mathcal{T})$; see also [13, Section 1.4]. First, we observe that we can construct a family of rooted real trees $(\mathcal{R}_k, d_{\mathcal{H}}), k \ge 1$, such that 1) $\mathcal{R}_1 \subseteq \mathcal{R}_2 \subseteq \cdots$ as metric spaces; 2) each $\mathcal{R}_k$ has exactly $k+1$ leaves which we denote as $\varsigma_0, \varsigma_1, \dots, \varsigma_k$ and a common root $\rho_{\mathcal{H}} = \varsigma_0$; 3) for each $k \ge 1$, the distance between ς_i and ς_j is given by $\gamma_\infty(i+1, j+1)$, $0 \le i,j \le k$. On the other hand, note that (3.10) entail that $(\gamma_\infty(i,j))_{1\le i,j<\infty}$ under $\mathbf{P}$ has the same distribution as $(d(V_i,V_j))_{1\le i,j<\infty}$. Since $\mathcal{T}$ satisfied the so-called *leaf-tight property*, we deduce this also holds for $(\mathcal{R}_k)_{k\ge 1}$, namely, $\inf_{j\ge 1} d_{\mathcal{H}}(\varsigma_0,\varsigma_j) = 0$, $\mathbf{P}$-a.s. Note that all

this still holds if we have first conditioned on $\boldsymbol{\mathcal{T}}$. Then by Theorem 3 in [6], given $\boldsymbol{\mathcal{T}}$, $(\mathcal{R}_k)_{k\geq 1}$ under $\mathbf{P}$ allows for a representation as a measured real tree, which we denote as $\mathrm{shuff}(\mathcal{T}) = (\mathcal{H}, d_{\mathcal{H}}, \mu_{\mathcal{H}}, \rho_{\mathcal{H}})$. Moreover, if $(\zeta_i)_{i\geq 1}$ is a sequence of i.i.d. points of $\mathcal{H}$ with common distribution $\mu_{\mathcal{H}}$ and $\zeta_0 = \rho_{\mathcal{H}}$, then $(d_{\mathcal{H}}(\zeta_i, \zeta_j))_{0\leq i,j<\infty}$ has the same distribution as $(d_{\mathcal{H}}(\varsigma_i, \varsigma_j))_{0\leq i,j<\infty}$. For this reason, we can then view $(\varsigma_i)_{i\geq 1}$ as an i.i.d. sequence of common law $\mu_{\mathcal{H}}$. As Γ_∞ characterizes $\mathrm{shuff}(\mathcal{T})$, (3.10) entails that $(\boldsymbol{\mathcal{T}}, \mathrm{shuff}(\mathcal{T})) \overset{d}{=} (\mathrm{cut}(\mathcal{T}), \boldsymbol{\mathcal{T}})$. The proof of Theorem 3.2 is now complete.

References

[1] Romain Abraham and Jean-François Delmas, *The forest associated with the record process on a Lévy tree*, Stochastic Process. Appl. **123** (2013), no. 9, 3497–3517. MR-3071387

[2] Romain Abraham and Jean-François Delmas, *Record process on the continuum random tree*, ALEA Lat. Am. J. Probab. Math. Stat. **10** (2013), no. 1, 225–251. MR-3083925

[3] Romain Abraham and Laurent Serlet, *Poisson snake and fragmentation*, Electron. J. Probab. **7** (2002), no. 17, 15. MR-1943890

[4] Louigi Addario-Berry, Nicolas Broutin, and Cecilia Holmgren, *Cutting down trees with a Markov chainsaw*, Ann. Appl. Probab. **24** (2014), no. 6, 2297–2339. MR-3262504

[5] Louigi Addario-Berry, Daphné Dieuleveut, and Christina Goldschmidt, *Inverting the cut-tree transform*, arXiv:1606.04825, 2016.

[6] David Aldous, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR-1207226

[7] David Aldous and Jim Pitman, *The standard additive coalescent*, Ann. Probab. **26** (1998), no. 4, 1703–1726. MR-1675063

[8] Erich Baur and Jean Bertoin, *Cutting edges at random in large recursive trees*, Stochastic analysis and applications 2014, Springer Proc. Math. Stat., vol. 100, Springer, Cham, 2014, pp. 51–76. MR-3332709

[9] Jean Bertoin, *Random fragmentation and coagulation processes*, Cambridge Studies in Advanced Mathematics, vol. 102, Cambridge University Press, Cambridge, 2006. MR-2253162

[10] Jean Bertoin, *Fires on trees*, Ann. Inst. Henri Poincaré Probab. Stat. **48** (2012), no. 4, 909–921. MR-3052398

[11] Jean Bertoin, *The cut-tree of large recursive trees*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 478–488. MR-3335011

[12] Jean Bertoin and Alexander V. Gnedin, *Asymptotic laws for nonconservative self-similar fragmentations*, Electron. J. Probab. **9** (2004), no. 19, 575–593. MR-2080610

[13] Jean Bertoin and Grégory Miermont, *The cut-tree of large Galton-Watson trees and the Brownian CRT*, Ann. Appl. Probab. **23** (2013), no. 4, 1469–1493. MR-3098439

[14] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013, A nonasymptotic theory of independence, With a foreword by Michel Ledoux. MR-3185193

[15] Nicolas Broutin and Minmin Wang, *Cutting down* $\mathbf{p}$*-trees and inhomogeneous continuum random trees*, Bernoulli **23** (2017), no. 4A, 2380–2433. MR-3648034

[16] Daphné Dieuleveut, *The vertex-cut-tree of Galton-Watson trees converging to a stable tree*, Ann. Appl. Probab. **25** (2015), no. 4, 2215–2262. MR-3349006

[17] Michael Drmota, Alex Iksanov, Martin Moehle, and Uwe Roesler, *A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree*, Random Structures Algorithms **34** (2009), no. 3, 319–336. MR-2504401

[18] Thomas Duquesne and Jean-François Le Gall, *Probabilistic and fractal aspects of Lévy trees*, Probab. Theory Related Fields **131** (2005), no. 4, 553–603. MR-2147221

[19] Steven N. Evans, *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, Berlin, 2008, Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR-2351587

[20] James Allen Fill, Nevin Kapur, and Alois Panholzer, *Destruction of very simple trees*, Algorithmica **46** (2006), no. 3-4, 345–366. MR-2291960

[21] Misha Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, english ed., Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA, 2007, Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates. MR-2307192

[22] Cecilia Holmgren, *Random records and cuttings in binary search trees*, Combin. Probab. Comput. **19** (2010), no. 3, 391–424. MR-2607374

[23] Cecilia Holmgren, *A weakly 1-stable distribution for the number of random records and cuttings in split trees*, Adv. in Appl. Probab. **43** (2011), no. 1, 151–177. MR-2761152

[24] Alex Iksanov and Martin 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. Comm. Probab. **12** (2007), 28–35. MR-2407414

[25] Svante Janson, *Random cutting and records in deterministic and random trees*, Random Structures Algorithms **29** (2006), no. 2, 139–179. MR-2245498

[26] Jean-François Le Gall, *The uniform random tree in a Brownian excursion*, Probab. Theory Related Fields **96** (1993), no. 3, 369–383. MR-1231930

[27] Jean-François Le Gall, *Random trees and applications*, Probab. Surv. **2** (2005), 245–311. MR-2203728

[28] A. Meir and J. W. Moon, *Cutting down random trees*, J. Austral. Math. Soc. **11** (1970), 313–324. MR-0284370

[29] Grégory Miermont, *Tessellations of random maps of arbitrary genus*, Ann. Sci. Éc. Norm. Supér. (4) **42** (2009), no. 5, 725–781. MR-2571957

[30] Alois Panholzer, *Cutting down very simple trees*, Quaest. Math. **29** (2006), no. 2, 211–227. MR-2233368

Acknowledgments. This work has undergone significant changes since its first version. We thank the referees and editors for their critics and suggestions, which have helped us greatly improve the presentation. This work is partially supported by ANR-14-CE25-0014 (ANR GRAAL).

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe `http://www.lockss.org/`
[2]EJMS: Electronic Journal Management System `http://www.vtex.lt/en/ejms.html`
[3]IMS: Institute of Mathematical Statistics `http://www.imstat.org/`
[4]BS: Bernoulli Society `http://www.bernoulli-society.org/`
[5]Project Euclid: `https://projecteuclid.org/`
[6]IMS Open Access Fund: `http://www.imstat.org/publications/open.htm`