The Annals of Probability
2017, Vol. 45, No. 1, 225–277
DOI: 10.1214/16-AOP1112

DOOB–MARTIN BOUNDARY OF RÉMY'S TREE GROWTH CHAIN

BY STEVEN N. EVANS[1], RUDOLF GRÜBEL AND ANTON WAKOLBINGER[2]

University of California, Leibniz Universität Hannover and Goethe-Universität

Rémy's algorithm is a Markov chain that iteratively generates a sequence of random trees in such a way that the nth tree is uniformly distributed over the set of rooted, planar, binary trees with $2n + 1$ vertices. We obtain a concrete characterization of the Doob–Martin boundary of this transient Markov chain and thereby delineate all the ways in which, loosely speaking, this process can be conditioned to "go to infinity" at large times. A (deterministic) sequence of finite rooted, planar, binary trees converges to a point in the boundary if for each m the random rooted, planar, binary tree spanned by $m + 1$ leaves chosen uniformly at random from the nth tree in the sequence converges in distribution as n tends to infinity—a notion of convergence that is analogous to one that appears in the recently developed theory of graph limits.

We show that a point in the Doob–Martin boundary may be identified with the following ensemble of objects: a complete separable $\mathbb{R}$-tree that is rooted and binary in a suitable sense, a diffuse probability measure on the $\mathbb{R}$-tree that allows us to make sense of sampling points from it, and a kernel on the $\mathbb{R}$-tree that describes the probability that the first of a given pair of points is below and to the left of their most recent common ancestor while the second is below and to the right. Two such ensembles represent the same point in the boundary if for each m the random, rooted, planar, binary trees spanned by $m + 1$ independent points chosen according to the respective probability measures have the same distribution. Also, the Doob–Martin boundary corresponds bijectively to the set of extreme point of the closed convex set of nonnegative harmonic functions that take the value 1 at the binary tree with 3 vertices; in other words, the minimal and full Doob–Martin boundaries coincide. These results are in the spirit of the identification of graphons as limit objects in the theory of graph limits.

CONTENTS

Received November 2014; revised December 2015.

[1]Supported in part by NSF Grants DMS-09-07630 and DMS-15-12933, and NIH Grant 1R01GM109454-01.

[2]Supported in part by DFG priority program 1590.

MSC2010 subject classifications. Primary 60J50; secondary 60J10, 68W40.

Key words and phrases. Binary tree, tail σ-field, Doob–Martin compactification, Poisson boundary, bridge, real tree, exchangeability, continuum random tree, Catalan number, graph limit, graphon, partial order.

1. Introduction.

Rémy's algorithm [28] iteratively generates a sequence of random binary trees $T_1, T_2, \dots$ in a Markovian manner in such a way that T_n is uniformly distributed on the set of binary trees with $2n+1$ vertices (see [4] for a textbook discussion of this procedure). Here (and throughout this paper), a *binary tree* is a finite rooted tree in which every vertex has zero or two children and the tree is *planar*, so it is possible to distinguish between the *left* and *right* children of a vertex with two children.

A binary tree has $2n+1$ vertices for some $n \in \mathbb{N}$: $n+1$ leaves and n interior vertices. The number of binary trees with $2n+1$ vertices is the *Catalan number* $C_n := \frac{1}{n+1}\binom{2n}{n}$ [32].

Writing $\{0,1\}^\star := \bigsqcup_{k=0}^\infty \{0,1\}^k$ for the set of finite words drawn from the alphabet $\{0,1\}$ (with the empty word $\emptyset$ allowed), any binary tree can be identified with a unique finite subset $\mathbf{t} \subset \{0,1\}^\star$ that has the properties:

- $v_1 \cdots v_k \in \mathbf{t} \Longrightarrow v_1 \cdots v_{k-1} \in \mathbf{t}$,
- $v_1 \cdots v_k 0 \in \mathbf{t} \Longleftrightarrow v_1 \cdots v_k 1 \in \mathbf{t}$.

The empty word $\emptyset \in \{0,1\}^\star$ is the root of the tree. See Figure 1 for an example of this representation. Rémy's algorithm begins by setting T_1 to be the unique binary tree $\aleph$ with 3 vertices (a root and two leaves). Supposing that $T_1, \dots, T_n$ have been generated, the algorithm generates T_{n+1} as follows (see Figures 2, 3 and 4 for a depiction of the steps that make up a single iteration of the algorithm).

- Pick a vertex v of T_n uniformly at random.
- Cut off the subtree of T_n rooted at v and set it aside.
- Attach a copy of the tree $\aleph$ with 3 vertices to the end of the edge in T_n that previously led to v.
- Reattach the subtree that was rooted at v in T_n uniformly at random to one of the two leaves in the copy of $\aleph$.

We call the two new vertices that are produced in the above iteration *clones* of v.

It is not too difficult to see that the algorithm does produce uniformly distributed binary trees. Indeed, suppose that the algorithm is modified so that it starts with the leaves of $\aleph$ labeled with 1 and 2, with each of the two labelings being equally likely, a random leaf-labeled tree that we denote by $\tilde{T}_1$. Suppose further that we

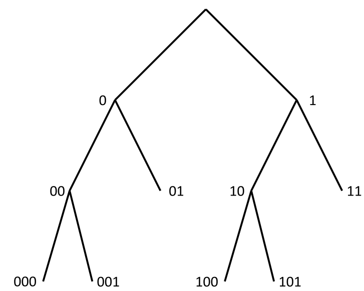

FIG. 1. *An example of a binary tree as a subset of* $\{0, 1\}^{\star}$.

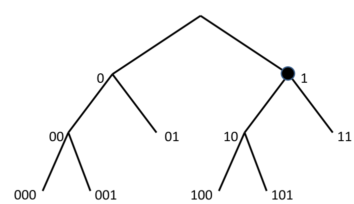

FIG. 2. *First step in an iteration of Rémy's algorithm*: *pick a vertex* v *uniformly at random*.

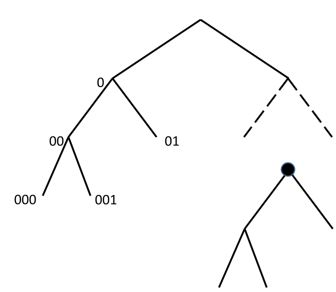

FIG. 3. *Second step in an iteration of Rémy's algorithm*: *cut off the subtree rooted at* v *and attach a copy of* $\aleph$ *to the end of the edge that previously led to* v.

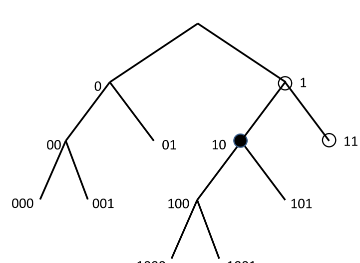

FIG. 4. *Third step in an iteration of Rémy's algorithm*: *reattach the subtree rooted at* v *to one of the two leaves of the copy of* $\aleph$, *and relabel the vertices appropriately. The solid circle is the new location of* v *and the open circles are the clones of* v.

begin the $(n+1)$st step with a leaf-labeled binary tree $\tilde{T}_n$ that has $n+1$ leaves labeled with $[n+1] := \{1, \dots, n+1\}$ in some order and that this step produces a random leaf-labeled binary tree $\tilde{T}_{n+1}$ labeled with $[n+2]$ as follows.

- Use the same randomization as in the algorithm described above to produce a tree with a single new leaf.
- Leave the labels of the old leaves unchanged.
- Label the new leaf with $n+2$.

It will certainly suffice to show that this enhanced algorithm produces a sequence $\tilde{T}_1, \tilde{T}_2, \dots$ such that for all $n \in \mathbb{N}$ the random leaf-labeled binary tree $\tilde{T}_n$ is uniformly distributed on the set of binary trees with $2n+1$ vertices that have their $n+1$ leaves labeled by $[n+1]$. This, however, is almost immediate from an inductive argument and the observation that in order for the value of $\tilde{T}_{n+1}$ to be a particular labeled binary tree, there is a unique possibility for the value of $\tilde{T}_n$, the choice of vertex v to clone, and the left-right choice for reattaching the subtree below v; see [4, 28] for more details.

Following [4, 28], we note that this argument also shows that if we let p_n be the common value of $\mathbb{P}\{\tilde{T}_n = \tilde{\mathbf{t}}\}$ as $\tilde{\mathbf{t}}$ ranges over the binary trees with $2n+1$ vertices and their $n+1$ leaves labeled by $[n+1]$, then $p_{n+1} = \frac{1}{2n+1}\frac{1}{2}p_n$, so that $p_n = \frac{1}{1\times 3\times\cdots\times(2n-1)}\frac{1}{2^n}$. It follows that the number of binary trees with $2n+1$ vertices and their $n+1$ leaves labeled by $[n+1]$ is

$$\big(1 \times 3 \times \cdots \times (2n-1)\big)2^n = \frac{(2n)!}{n!},$$

and so the number of binary trees with $2n+1$ vertices and $n+1$ leaves is

$$\frac{(2n)!}{(n+1)!n!} = C_n,$$

as expected.

As well as counting the number of binary trees with $2n+1$ vertices for $n \in \mathbb{N}$, the Catalan number C_n counts the number of functions $f : \{0, 1, \dots, 2n\} \to \mathbb{N}_0$ such that $f(0) = f(2n) = 0$ and $f(k+1) = f(k) \pm 1$ for $0 \le k < 2n$. It is shown in [25] that there are particular bijections ϕ_n (sometimes credited to Łukasiewicz or Dwass) between the former and latter sets such that if $f_n := \phi_n(T_n)$, then $(n^{-\frac{1}{2}} f_n(\lfloor 2nt \rfloor))_{t\in[0,1]}$ converges almost surely in the supremum norm to a standard Brownian excursion. A similar result is given in [27], Exercise 7.4.11, where T_n is represented as a function from $\{0, 1, \dots, 4n\}$ to $\mathbb{N}_0$ using a coding where one "walks around the outside" of the tree visiting left children before right children (so that each edge is traversed twice, leaves are visited once and other vertices are visited three times), and recording the distance from the root to the vertex visited at each step—this coding, or a minor modification of it, is sometimes called the Harris path of the tree.

The standard Brownian excursion induces a random metric space, that is, up to a scaling factor, Aldous' Brownian continuum random tree (CRT) [1–3]. More precisely, if $(E_t)_{t\in[0,1]}$ is the standard Brownian excursion, then $d(s,t) := E_s + E_t - 2\min_{u\in[s,t]} E_u$, $s,t \in [0,1]$ defines a pseudo-metric on $[0,1]$ that becomes a metric on the collection of equivalence classes for the equivalence relation $s \equiv t \Leftrightarrow d(s,t) = 0$, and the latter random metric space is a random $\mathbb{R}$-tree, that is, by definition, a scaled version of the Brownian CRT (see [13] for a general treatment of $\mathbb{R}$-trees directed at probabilists). This definition carries with it a natural rooting and hence a natural genealogical structure: the most recent common ancestor of the equivalence classes containing s and t is the equivalence class of the almost surely unique $v \in [s,t]$ such that $E_v = \min_{u\in[s,t]} E_u$. The Brownian CRT with this rooting is almost surely binary in the sense that almost surely for all $r,s,t \in [0,1]$ coming from distinct equivalence classes the most recent common ancestors of the pairs (r,s), (r,t), (s,t) are not all equal. Moreover, this construction also endows the Brownian CRT with a natural planar structure: for $s,t \in [0,1]$ coming from distinct equivalence classes, the equivalence class containing s may be consistently declared to be below and to the left of the most recent common ancestor of the two equivalence classes (and the equivalence class containing t is below and to the right) if $\min\{q : \min_{u\in[q,s]} E_u = E_s\} < \min\{r : \min_{v\in[r,t]} E_v = E_t\}$ (in other words, the time parameter in the Brownian excursion induces a traversal of the points of the Brownian CRT that starts and ends at the root, and we say that one equivalence class is below and to the left of the most recent common ancestor it shares with another equivalence class if this traversal encounters the former equivalence class before the latter).

Conversely, it is observed in [21] that if one samples i.i.d. uniformly distributed points $U_0, U_1, \dots$ from $[0,1]$ and lets $\hat{T}_n$ be the binary tree spanned by the equivalence classes of $U_0, \dots, U_n$ for $n \in \mathbb{N}$ (more fully, one takes the subtree of the rescaled Brownian CRT thought of as a random $\mathbb{R}$-tree but equipped with the additional rooting and left–right ordering described above, forgets the metric structure on the subtree, but keeps the rooting and left–right ordering), then $(\hat{T}_n)_{n\in\mathbb{N}}$ has the same distribution as $(T_n)_{n\in\mathbb{N}}$; that is, $(\hat{T}_n)_{n\in\mathbb{N}}$ is an instance of Rémy's chain.

As we shall explain soon, these last two results and several more are parallels of classical results about the simplest Pólya urn scheme in which one starts with an urn containing one black and one white ball and at each step one picks a ball uniformly at random and replaces it along with another of the same color.

If we write N_n for the number of new black balls that have been added to the urn up to and including the nth step of the Pólya urn chain, then $((N_n, n - N_n))_{n\in\mathbb{N}}$ is a Markov chain with the following properties. For each $n \in \mathbb{N}$, the random variable N_n is uniformly distributed on $\{0,1,\dots,n\}$ and N_n/n converges almost surely as $n \to \infty$ to a random variable U that is uniformly distributed on the interval $[0,1]$. If $(X_n)_{n\in\mathbb{N}}$ is a sequence of $\{0,1\}$-valued random variables that are conditionally independent given U with $\mathbb{P}\{X_n = 1|U\} = U$, then [6] $(N_n)_{n\in\mathbb{N}}$ has the same distribution as $(X_1 + \cdots + X_n)_{n\in\mathbb{N}}$. It follows from this observation and the

Hewitt–Savage zero–one law that the tail σ-field of the Pólya urn chain is generated up to null sets by the random variable U. By the martingale convergence theorem, the vector space of bounded harmonic functions for the Pólya urn chain (i.e., the *Poisson boundary* of the chain) can thus be identified with L^∞ of the unit interval equipped with Lebesgue measure. Another consequence is the well-known fact that the colors of the successive balls form an exchangeable sequence and so the backward dynamics of the Pólya urn chain from step n to step $n-1$ can be thought of as removing one of the n added balls present at step n uniformly at random and discarding it.

We will show that the backward transitions of the Rémy chain are as follows:

- Pick a leaf uniformly at random.
- Delete the chosen leaf and its sibling (the sibling may or may not be a leaf).
- Close up the gap if there is one (there will be a gap if the sibling is not a leaf).

Note how these dynamics are reminiscent of the backward transitions of the Pólya urn chain. It is a consequence of the exchangeability inherent in these dynamics and the Hewitt–Savage zero–one law that the tail σ-field of the Rémy chain is generated up to null sets by the limiting Brownian CRT augmented by the additional rooting and left–right ordering described above. More precisely, we may assume that the entire Rémy chain has been built from a Brownian excursion (equivalently, the augmented Brownian CRT) and an independent, identically distributed sequence $(U_k)_{k\in\mathbb{N}_0}$ of random variables that are each uniformly distributed on $[0, 1]$. If the first $n+1$ of these random variables are permuted in any way, then the values of the Rémy chain from time n onward are unchanged, and so the Hewitt–Savage zero–one law gives that the tail σ-field of the Rémy chain is, up to null sets, contained in the σ-field generated by the augmented Brownian CRT. Conversely, since one can build the Brownian CRT as an almost sure limit (as $n \to \infty$) of the rescaled Rémy chain, the tail σ-field is equal to the σ-field generated by the augmented Brownian CRT up to null sets. Hence, the Poisson boundary of the Rémy chain can be identified with L^∞ of a space of suitably defined "rooted, planar, binary" $\mathbb{R}$-trees equipped with the distribution of the augmented Brownian CRT or, equivalently, with L^∞ of the space of continuous excursion paths indexed by $[0, 1]$ equipped with the standard Brownian excursion measure.

The Rémy chain is not the only Markov chain which at step n produces uniformly distributed binary trees with $2n + 1$ vertices. Another example is the Markov chain proposed in [24] which, unlike the Rémy chain, has the property that the state at time n is a subtree of the state at time $n + 1$ for all $n \in \mathbb{N}$. The Poisson boundary of this chain, which was described in [14], turns out to be quite different from that of the Rémy chain.

The object of the present paper is to go further and investigate the *Doob–Martin compactification* of the Rémy chain. Before giving a formal definition of the Doob–Martin compactification in Section 2, let us illustrate the concept with the archetypal example of the Pólya urn chain. Given $(\mathbf{b}, \mathbf{w}) \in (\mathbb{N}_0 \times \mathbb{N}_0) \setminus \{0, 0\}$, let $N_1^{(\mathbf{b},\mathbf{w})}, \dots, N_{\mathbf{b}+\mathbf{w}}^{(\mathbf{b},\mathbf{w})}$ be the bridge process obtained by conditioning the initial segment $N_1, \dots, N_{\mathbf{b}+\mathbf{w}}$ of the Pólya urn chain on the event $\{N_{\mathbf{b}+\mathbf{w}} = \mathbf{b}\}$. The backward transitions of such a bridge are the same as those of the Pólya urn chain itself and it is not hard to show that if $((\mathbf{b}_k, \mathbf{w}_k))_{k\in\mathbb{N}}$ is a sequence such that $\mathbf{b}_k + \mathbf{w}_k \to \infty$ as $k \to \infty$, then the finite-dimensional distributions of the corresponding bridges converge if and only if $\lim_{k\to\infty} \frac{\mathbf{b}_k}{\mathbf{b}_k+\mathbf{w}_k} \in [0, 1]$ exists. It is a classical result ([20], Chapter 10) that a sequence $((\mathbf{b}_k, \mathbf{w}_k))_{k\in\mathbb{N}}$ such that $\mathbf{b}_k + \mathbf{w}_k \to \infty$ as $k \to \infty$ converges in the Doob–Martin compactification of the Pólya urn chain if and only if $\lim_{k\to\infty} \frac{\mathbf{b}_k}{\mathbf{b}_k+\mathbf{w}_k} \in [0, 1]$ exists and, as we recall in Section 2, a general result from [15] establishes the equivalence between convergence of bridges and convergence in the Doob–Martin compactification under suitable conditions. It follows that the Doob–Martin boundary of the Pólya urn chain is (homeomorphic to) the unit interval $[0, 1]$. There is thus a nonnegative harmonic function associated with each point $u \in [0, 1]$ and the corresponding Doob h-transform process can be interpreted as $(N_n)_{n\in\mathbb{N}}$ conditioned on the event $\{U = u\}$. As one would expect, the distribution of the Doob h-transform process is nothing other than that of the process of partial sums of independent, identically distributed Bernoulli random variables with success probability u.

We will investigate the bridges of the Rémy chain and thereby identify its Doob–Martin boundary. This boundary of the space of (finite) binary trees determines the compact convex set of nonnegative harmonic functions normalized to take the value 1 at the binary tree with two leaves. We show that the set of extreme points of the latter compact convex set corresponds bijectively to the Doob–Martin boundary, and hence the boundary delineates all the ways that the Rémy chain can be conditioned to "behave at infinity" in such a way that any randomness disappears asymptotically in the sense that the tail σ-field of the conditioned chain is trivial.

We will show that a sequence of finite binary trees with the number of vertices going to infinity converges in the Doob–Martin topology if and only if for all m the sequence of random binary trees spanned by $m + 1$ leaves sampled uniformly at random from those of the nth tree in the original sequence converges in distribution as $n \to \infty$. Moreover, two convergent sequences converge to the same limit if and only if the corresponding limit distributions of these "sampled subtrees" are the same for all m. (The analogous fact is also true for the Pólya urn: a sequence of states converges in the Doob–Martin topology if and only if for any m when we sample m balls uniformly at random from the urn composition specified by the nth state, the distribution of the number of black balls in the sample converges as $n \to \infty$.) This type of convergence of a sequence of large combinatorial objects in terms of the convergence in distribution of randomly sampled sub-objects of a given but arbitrary size is similar to a notion of convergence of finite graphs investigated in the theory of graph limits where a sequence of graphs with increasing numbers of vertices converges if for each m the distributions of the random finite graphs induced by m vertices sampled uniformly at random converge (see [22] for a recent monograph and [5, 7–9, 11, 23, 33] for some examples of papers in this area). A binary tree encodes two partial orders on its set of vertices [one vertex can be below and to the left (resp., right) of another vertex], and so the work in [18] on limits of large partially ordered sets is particularly close in spirit to our work. A further connection between our work and graph limits is the result of [16] that the above notion of graph convergence is nothing other than convergence in the Doob–Martin topology for the graph-valued *Erdős–Rényi chain* in which at each step an additional vertex is added with the possible edges connecting it to each of the existing vertices independently present with probability p and absent with probability $1 - p$ for some fixed $0 < p < 1$ (the Doob–Martin compactification does not depend on the value of p).

One of the major achievements of the theory of graph limits has been to obtain concrete representations of the limit objects corresponding to a convergent sequence of graphs as so-called *graphons*. A graphon is a symmetric Borel function $K : [0, 1]^2 \to [0, 1]$ and a random graph with the distribution of the limit of the randomly sampled subgraphs of size m corresponding to a convergent sequence of graphs is obtained by choosing m points $U_1, \dots, U_m$ uniformly at random from $[0, 1]$ and connecting vertex i and j with conditional probability $K(U_i, U_j)$.

In our main result, Theorem 8.2, we obtain a similar concrete representation of a point in the Doob–Martin boundary of the Rémy chain as a rooted $\mathbb{R}$-tree $\mathbf{S}$ equipped with a probability measure μ and a function $V : \mathbf{S}^2 \to [0, 1]$. The limit in distribution of the subtrees spanned by $m + 1$ uniformly chosen leaves is obtained by, loosely speaking, looking at the subtree of $\mathbf{S}$ spanned by independent random points $\xi_1, \dots, \xi_{m+1}$ with distribution μ and declaring that with probability $V(\xi_i, \xi_j)$ leaf i is below and to the left while j is below and to the right of the most recent common ancestor of leaves i and j. Like all transient Markov chains, the Rémy chain has the property that T_n converges almost surely as $n \to \infty$ in the Doob–Martin topology to a random element of the Doob–Martin boundary. The distribution of the limit may be identified with that of the augmented Brownian CRT described above: the underlying $\mathbb{R}$-tree and its root come from the Brownian excursion, and the probability measure on the $\mathbb{R}$-tree is the one lifted by the Brownian excursion from Lebesgue measure on $[0, 1]$. In this case the function V takes values in the set $\{0, 1\}$ and is determined by the left–right ordering coming from the Brownian excursion. We will see that it is not always possible to have the left–right ordering be induced from one on the underlying $\mathbb{R}$-tree $\mathbf{S}$ and that cases do arise where it is necessary to work with functions V that take values strictly between 0 and 1.

Briefly, the strategy of the proof of Theorem 8.2 will be as follows.

(i) First, we determine the backward transition dynamics of the Rémy chain in Section 4. Understanding the Doob–Martin compactification is equivalent to understanding all Markov chains with initial state $\aleph$ that have these backward transition dynamics. We call any such chain $(T_n^\infty)_{n\in\mathbb{N}}$ an *infinite Rémy bridge*:

the class of infinite Rémy bridges corresponds bijectively with the class of Doob h-transforms of the original Rémy chain as h ranges over the nonnegative harmonic functions for the original chain normalized so that $h(\aleph) = 1$. This class of nonnegative harmonic functions is a compact convex set, and an arbitrary such function has a unique representation as an integral over extremal elements. For a general Markov chain, an extremal nonnegative harmonic function corresponds to a point in the Doob–Martin boundary, but there may be points in the Doob–Martin boundary that correspond to harmonic functions which are not extremal. We show that this is not the case for the Rémy chain, and it follows that the elements of the Doob–Martin boundary of the Rémy chain correspond bijectively to the infinite Rémy bridges that are *extremal* in the sense that they are not nondegenerate mixtures of infinite Rémy bridges (equivalently, to the infinite Rémy bridges that have trivial tail σ-fields).

(ii) A key tool for obtaining a concrete characterization of the extremal infinite Rémy bridges will be the introduction of an auxiliary labeling of the $n+1$ leaves of the tree T_n^∞ by $[n+1] := \{1, \dots, n+1\}$ that has the properties that the labeling is uniformly distributed over the $(n+1)!$ possible labelings for each n and the new leaf added at step $n+1$ is labeled with $n+2$ while the other leaves keep the labels they had at step n. Such a labeling scheme is "projective" in the sense that the leaf-labeled subtree of T_{m+n}^∞ spanned by the leaves with labels in $[m+1]$ coincides with the leaf-labeled version of T_m^∞; more precisely, T_m^∞ embeds into T_{m+n}^∞ in the manner defined in Section 3 via an injective map from the vertices of T_m^∞ into the vertices of T_{m+n}^∞ that maps leaves to leaves in such a way that the image of the leaf labeled k in T_m^∞ is mapped to the leaf labeled k in T_{m+n}^∞ for $k \in [m+1]$. As we observe in Section 5, for any $i, j, k \in [m+1]$ there are twelve possibilities for how the leaves labeled i, j, k in the tree T_m^∞ sit in relation to each other; for example, one possibility is that the most recent common ancestor of i and j is a descendant of the most recent common ancestor of i and k which is also the most recent common ancestor of j and k, that i is below and to the left and j is below and to the right of their most recent common ancestor, that i is below and to the left and k is below and to the right of their most recent common ancestor, and that j is below and to the left and k is below and to the right of their most recent common ancestor. Moreover, knowing which of these possibilities holds for each triple i, j, k uniquely determines the tree T_m^∞ and its leaf labels. A key feature of this labeling is that the relative positions of the leaves labeled i, j, k in the tree T_m^∞ is the same as the relative positions of the leaves labeled i, j, k in the tree T_{m+n}^∞. Because of this consistency there is a well-defined random array indexed by $\{(i, j, k) \in \mathbb{N}^3 : i, j, k \text{distinct}\}$ that for any indices (i, j, k) records for all m such that $\{i, j, k\} \subseteq [m+1]$ which of the twelve possibilities holds for the relative positions of the leaves labeled i, j, k in the tree T_m^∞. This random array is jointly exchangeable. It is possible to reconstruct the entire leaf-labeled version of the infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$ from this array, and hence the infinite Rémy bridge itself by then simply discarding the leaf labels. The infinite Rémy bridge is extremal if and only if this jointly exchangeable random array is ergodic in the usual sense for jointly exchangeable random arrays.

(iii) In Sections 6, 7 and 8, we use ideas related to those in [10, 26] and the Aldous–Hoover–Kallenberg theory of jointly exchangeable random arrays to obtain a concrete description of the jointly exchangeable random arrays that can arise from extremal infinite Rémy bridges, and it is the ingredients in this description that appear in our above sketch of the statement of Theorem 8.2. The $\{0, 1\}$-valued random variables $W(\xi_i, U_i, \xi_j, U_j)$ figuring in the actual statement of Theorem 8.2 indicate whether leaf i is below and to the left while j is below and to the right of the most recent common ancestor of leaves i and j, with the above-mentioned $V(\xi_i, \xi_j)$ as the corresponding probabilities.

2. Background on Doob–Martin compactifications. We restrict the following sketch of Doob–Martin compactification theory for discrete time Markov chains to the situation of interest in the present paper. The primary reference is [12], but useful reviews may be found in [20], Chapter 10, [29], Chapter 7, [31], [35], Chapter 7, [30], Chapter III.

Suppose that $(X_n)_{n\in\mathbb{N}_0}$ is a discrete time Markov chain with countable state space E and transition matrix P. Suppose in addition that E can be partitioned as $E = \bigsqcup_{n\in\mathbb{N}_0} E_n$, where $E_0 = \{e\}$ for some distinguished state e, each set E_n is finite, and the transition matrix P is such that $P(k, \ell) = 0$ unless $k \in E_n$ and $\ell \in E_{n+1}$ for some $n \in \mathbb{N}_0$. Define the *Green kernel* or *potential kernel* G of P by

$$G(i, j) := \sum_{n=0}^{\infty} P^n(i, j) = \mathbb{P}^i\{X_n = j \text{ for some } n \in \mathbb{N}_0\} =: \mathbb{P}^i\{X \text{ hits } j\},$$

$i, j \in E$, and assume that $G(e, j) > 0$ for all $j \in E$, so that any state can be reached with positive probability starting from e. The Rémy chain belongs to this class. The state space E of the Rémy chain is the set of all binary trees, the distinguished state e is the binary tree $\aleph$ with 3 vertices and E_n is the set of binary trees with $2n + 3$ vertices.

If Z is a $\mathbb{P}^e$-a.s. bounded random variable that is measurable with respect to the tail σ-field of $(X_n)_{n\in\mathbb{N}_0}$, then $\mathbb{E}^e[Z|X_0, \ldots, X_n] = h(X_n)$ for some bounded harmonic function h; that is, $\sum_{j\in E} P(i, j)h(j) = h(i)$ for $i \in E$. By the martingale convergence theorem, $\lim_{n\to\infty} h(X_n) = Z$ $\mathbb{P}^e$-a.s. Conversely, if h is a bounded harmonic function, then $\lim_{n\to\infty} h(X_n)$ exists $\mathbb{P}^e$-a.s. and the limit random variable is $\mathbb{P}^e$-a.s. equal to a random variable that is measurable with respect to the tail σ-field of $(X_n)_{n\in\mathbb{N}_0}$.

In order to characterize the bounded harmonic functions (and hence the tail σ-field), it certainly suffices to determine what the nonnegative harmonic functions are. The key to doing so is the introduction of the *Doob–Martin kernel with reference state e* given by

$$K(i,j) := \frac{G(i,j)}{G(e,j)} = \frac{\mathbb{P}^i\{X \text{ hits } j\}}{\mathbb{P}^e\{X \text{ hits } j\}}.$$

Observe that

$$\sum_{j \in E} P(i,j)K(j,k) = K(i,k), \qquad i \neq k,$$

and so the function $K(\cdot,k)$ is, in some sense, "almost harmonic" and becomes closer to being harmonic as k "goes to infinity". With this intuition in mind, it is natural to investigate sequences $(j_n)_{n\in\mathbb{N}}$ in E such that the sequence of real numbers $(K(i,j_n))_{n\in\mathbb{N}}$ converges for all $i \in E$.

These considerations lead to the following construction. If $j,k \in E$ with $j \neq k$, then $K(\cdot,j) \neq K(\cdot,k)$, and so E can be identified with the collection of functions $K(\cdot,j)$, $j \in E$. Note that

$$0 \leq K(i,j) \leq \frac{\mathbb{P}^i\{X \text{ hits } j\}}{\mathbb{P}^e\{X \text{ hits } i\}\mathbb{P}^i\{X \text{ hits } j\}} = \frac{1}{\mathbb{P}^e\{X \text{ hits } i\}},$$

and so the set of functions $\{K(\cdot,j) : j \in E\}$ is a pre-compact subset of $\mathbb{R}_+^E$ equipped with the usual product topology. Its closure $\bar{E}$ is the *Doob–Martin compactification* of E. The set $\partial E := \bar{E} \setminus E$ is the *Doob–Martin boundary* of E. By construction, a sequence $(j_n)_{n\in\mathbb{N}}$ in E converges to a point in $\bar{E}$ if and only if the sequence of real numbers $(K(i,j_n))_{n\in\mathbb{N}}$ converges for all $i \in E$, and each function $K(i,\cdot)$ extends continuously to $\bar{E}$. The resulting function $K : E \times \bar{E} \to \mathbb{R}$ is the *extended Doob–Martin kernel*.

A specific subset $\partial_{\min} E$, the *minimal* boundary, of the full boundary ∂E is of particular importance from a geometric as well as probabilistic point of view. Let $\mathbb{H}_{1,+}$ be the set of harmonic functions $h : E \to \mathbb{R}_+$ with $h(e) = 1$. This is a compact convex set, and its extreme points are those harmonic functions $h \in \mathbb{H}_{1,+}$ with the property that $ag < h$ implies $g = h$ whenever $0 < a < 1$ and $g \in \mathbb{H}_{1,+}$. We have $K(\cdot,y) \in \mathbb{H}_{1,+}$ for all $y \in \partial E$, and we write $\partial_{\min} E$ for the set of those boundary points that correspond to extremal harmonic functions. The set $\partial_{\min} E$ is a G_δ. With this notation in place, any $h \in \mathbb{H}_{1,+}$ has a unique representation

$$h(x) = \int K(x,y)\mu(dy),$$

where μ is a probability measure that assigns all of its mass to $\partial_{\min} E$.

A first major probabilistic consequence of the Doob–Martin compactification is that the limit $X_\infty := \lim_{n\to\infty} X_n$ exists $\mathbb{P}^e$-almost surely in the topology of $\bar{E}$ and that the distribution of this limit is given by the measure μ representing the trivial element $h \equiv 1$ of $\mathbb{H}_{1,+}$.

In terms of analysis, the vector space $\mathbb{H}_b$ of bounded harmonic functions endowed with the supremum norm is a Banach space (the *Poisson boundary* of the Markov chain) and this Banach space is isomorphic to the L^∞ space associated with the measure space consisting of ∂E equipped with its Borel σ-field and the probability measure given by the distribution of X_∞ under $\mathbb{P}^e$. The tail σ-field of $(X_n)_{n\in\mathbb{N}_0}$ coincides $\mathbb{P}^e$-almost surely with the σ-field generated by X_∞ and so the Poisson boundary captures how the process can "go to infinity" and what probabilities are associated with the various alternatives.

The second consequence of the Doob–Martin compactification is that not only does it contain information about how the Markov chain behaves at large times when "left to its own devices", but also, somewhat loosely speaking, how it can be conditioned to behave at large times. Each $j \in E = \bigsqcup_{n\in\mathbb{N}_0} E_n$ belongs to a unique E_n whose index n we denote by $N(j)$. If the chain starts in state e, then $N(j)$ is the only time that there is positive probability the chain will be in state j. Write $(X_0^j, \dots, X_{N(j)}^j)$ for the *bridge* obtained by starting the chain in state e and conditioning it to be in state j at time $N(j)$. This process is a Markov chain with forward transition probabilities

$$\begin{aligned}
\mathbb{P}\{X_{n+1}^j = i'' | X_n^j = i'\} &= \frac{\mathbb{P}^e\{X_n = i', X_{n+1} = i'', X_{N(j)} = j\}}{\mathbb{P}^e\{X_n = i', X_{N(j)} = j\}} \\
&= \frac{\mathbb{P}^e\{X \text{ hits } i'\} P(i', i'') \mathbb{P}^{i''}\{X \text{ hits } j\}}{\mathbb{P}^e\{X \text{ hits } i'\} \mathbb{P}^{i'}\{X \text{ hits } j\}} \\
&= \frac{P(i', i'') \mathbb{P}^{i''}\{X \text{ hits } j\} / \mathbb{P}^e\{X \text{ hits } j\}}{\mathbb{P}^{i'}\{X \text{ hits } j\} / \mathbb{P}^e\{X \text{ hits } j\}} \\
&= K(i', j)^{-1} P(i', i'') K(i'', j).
\end{aligned}$$

Moreover,

$$\begin{aligned}
\mathbb{P}\{X_n^j = i' | X_{n+1}^j = i''\} &= \frac{\mathbb{P}^e\{X_n = i', X_{n+1} = i'', X_{N(j)} = j\}}{\mathbb{P}^e\{X_{n+1} = i'', X_{N(j)} = j\}} \\
&= \frac{\mathbb{P}^e\{X_n = i', X_{n+1} = i''\} \mathbb{P}^{i''}\{X_{N(j)} = j\}}{\mathbb{P}^e\{X_{n+1} = i''\} \mathbb{P}^{i''}\{X_{N(j)} = j\}} \\
&= \mathbb{P}^e\{X_n = i' | X_{n+1} = i''\},
\end{aligned}$$

and so $(X_0^j, \dots, X_{N(j)}^j)$ has the same backward transition probabilities as $(X_n)_{n\in\mathbb{N}_0}$.

Suppose now that $(j_k)_{k\in\mathbb{N}}$ is a sequence of elements of the state space E that converges to infinity in the one-point compactification of E or, equivalently, $N(j_k) \to \infty$ as $k \to \infty$. As observed in [15], such a sequence $(j_k)_{k\in\mathbb{N}}$ converges in the Doob–Martin topology if and only if finite initial segments of the corresponding bridges converge in distribution. Moreover, two sequences of states converge to the same limit if and only if the limiting distributions of finite initial segments are the same. For a sequence $(j_k)_{k\in\mathbb{N}}$ that converges to the point y in the Doob–Martin boundary, the limiting distributions of the initial segments define the distribution of an E-valued process $(X_n^{(h)})_{n\in\mathbb{N}_0}$ that is Markovian with forward transition probabilities $P^{(h)}$ given by

$$P^{(h)}(i,j) := h(i)^{-1}P(i,j)h(j), \qquad i,j \in E^{(h)},$$

where $h(i) = \lim_{k\to\infty} K(i,j_k) = K(i,y)$ and

$$E^{(h)} := \{i \in E : h(i) > 0\} = \Big\{i \in E : \lim_{k\to\infty} \mathbb{P}\{X_{N(i)} = i \mid X_{N(j_k)} = j_k\} > 0\Big\},$$

and the same backward transition probabilities as $(X_n)_{n\in\mathbb{N}_0}$. This Markov chain $(X_n^{(h)})_{n\in\mathbb{N}_0}$ is an h-transform using the harmonic function h. Moreover, if $(y_k)_{k\in\mathbb{N}}$ is a sequence of points in the Doob–Martin boundary, then $\lim_{k\to\infty} y_k = y$ for some point y in the Doob–Martin boundary if and only if the initial segments of $(X_n^{(K(\cdot,y_k))})_{n\in\mathbb{N}_0}$ converge in distribution to the corresponding initial segments of $(X_n^{(K(\cdot,y))})_{n\in\mathbb{N}_0}$.

We call any Markov process $(Y_n)_{n\in\mathbb{N}_0}$ with $Y_0 = e$ and the same backward transition probabilities as $(X_n)_{n\in\mathbb{N}_0}$ an *infinite bridge* for $(X_n)_{n\in\mathbb{N}_0}$. The distribution of an infinite bridge is a mixture of distributions of infinite bridges that have trivial tail σ-fields, and we call the latter *extremal* infinite bridges. If $(j_k)_{k\in\mathbb{N}}$ converges to a point y in the Doob–Martin boundary, then the corresponding harmonic function $h = K(\cdot, y)$ is extremal if and only if the limit infinite bridge $(X_n^{(h)})_{n\in\mathbb{N}_0}$ is extremal.

3. The Doob–Martin kernel of the Rémy chain. We return from the general setting of the previous section to consideration of the Rémy chain. Given two binary trees $\mathbf{s}$ and $\mathbf{t}$ with $2m+1$ and $2(m+n)+1$ vertices, we wish to derive a formula for the multi-step transition probability

$$p(\mathbf{s},\mathbf{t}) := \mathbb{P}\{T_{m+n} = \mathbf{t} \mid T_m = \mathbf{s}\}$$

and hence obtain a formula for the Doob–Martin kernel with reference state $\aleph$, since

$$\begin{aligned}
K(\mathbf{s},\mathbf{t}) &:= \frac{p(\mathbf{s},\mathbf{t})}{p(\aleph,\mathbf{t})} \\
&= \frac{\mathbb{P}\{T_{m+n} = \mathbf{t} \mid T_m = \mathbf{s}\}}{\mathbb{P}\{T_{m+n} = \mathbf{t}\}} \\
&= \frac{\mathbb{P}\{T_{m+n} = \mathbf{t}, T_m = \mathbf{s}\}}{\mathbb{P}\{T_m = \mathbf{s}\}\mathbb{P}\{T_{m+n} = \mathbf{t}\}} \\
&= \frac{1}{\mathbb{P}\{T_m = \mathbf{s}\}}\mathbb{P}\{T_m = \mathbf{s} \mid T_{m+n} = \mathbf{t}\} \\
&= C_m \mathbb{P}\{T_m = \mathbf{s} \mid T_{m+n} = \mathbf{t}\},
\end{aligned}$$

where we recall that the mth Catalan number C_m is the number of binary trees with $2m + 1$ vertices. For this, we need the notion of one binary tree being embedded in another, and this requires us to introduce some preliminary definitions.

To begin, we define a partial order $<$ on the vertices of a binary tree by declaring that $u < v$ for two vertices u and v if $u \neq v$ and u is on the (unique) path leading from the root to v. We say that v is *below* u. Given two vertices x and y, there is a unique vertex z such that $z \le x$, $z \le y$, and $w < z$ for any other vertex w such that $w \le x$ and $w \le y$. We say that z is the *most recent common ancestor* of x and y and write $z = x \wedge y$.

If $u < v$ and the unique path from u to v passes through the left (resp., right) child of u, then we write $u <_L v$ (resp., $u <_R v$) and say that v is *below and to the left* (resp., *below and to the right*) of u. Note that $<_L$ and $<_R$ are partial orders with the property that if two vertices of the tree are comparable in one order, then they are not comparable in the other. Note also that $u < v$ if and only if $u <_L v$ or $u <_R v$.

If we think of a binary tree as a subset of $\{0, 1\}^*$, then for two vertices $u = u_1 \cdots u_m$ and $v = v_1 \cdots v_n$ we have:

- $u < v$ if and only if $m < n$ and $u_k = v_k$ for $1 \le k \le m$,
- the most recent common ancestor $u \wedge v$ of u and v is the vertex $w = w_1 \cdots w_p$, where $p = \max\{k : u_k = v_k\}$ (where the maximum of the empty set is 0) and $w_k = u_k = v_k$ for $1 \le k \le p$,
- $u <_L v$ if and only if $m < n$, $u_k = v_k$ for $1 \le k \le m$, and $v_{m+1} = 0$,
- $u <_R v$ if and only if $m < n$, $u_k = v_k$ for $1 \le k \le m$, and $v_{m+1} = 1$.

DEFINITION 3.1. An *embedding* of a binary tree $\mathbf{s}$ into a binary tree $\mathbf{t}$ is a map from the vertex set of $\mathbf{s}$ into the vertex set of $\mathbf{t}$ such that the following hold:

- The image of a leaf of $\mathbf{s}$ is a leaf of $\mathbf{t}$.
- If u, v are vertices of $\mathbf{s}$ such that v is below and to the left (resp., right) of u, then the image of v in $\mathbf{t}$ is below and to the left (resp., right) of the image of u in $\mathbf{t}$.

Figure 5 illustrates this definition.

REMARK 3.2. Note that an embedding of $\mathbf{s}$ into $\mathbf{t}$ is uniquely determined by the images of the leaves of $\mathbf{s}$, because if x and y are vertices of $\mathbf{s}$, then the image of the most recent common ancestor of x and y in $\mathbf{s}$ must be the most recent common ancestor in $\mathbf{t}$ of the images of x and y.

NOTATION 3.3. Write $N(\mathbf{s}, \mathbf{t})$ for the number of embeddings of $\mathbf{s}$ into $\mathbf{t}$. For $k = 1, 2, \ldots$, write $\mathbf{t}_k^c$ for the complete binary tree with 2^k leaves, that is the binary tree with 2^k leaves such that every leaf is graph distance k from the root. (In the representation of binary trees as subsets of $\{0, 1\}^\star$, $\mathbf{t}_k^c$ is the subset consisting of words with length at most k and the leaves are the words with length k.)

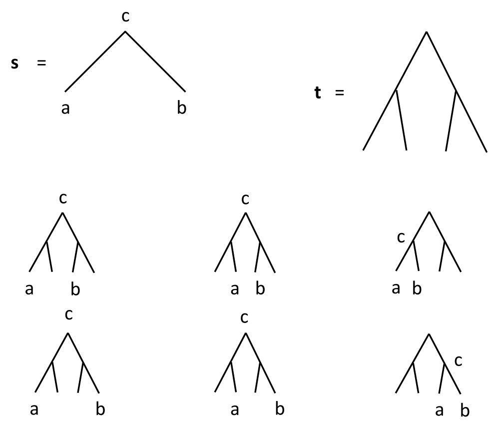

FIG. 5. *All the embeddings of the unique binary tree* $\mathbf{s} = \aleph$ *with* 3 *vertices into a particular tree* $\mathbf{t}$ *with* 7 *vertices.*

EXAMPLE 3.4. We want to identify the number $N(\mathbf{s}, \mathbf{t}_k^c)$ of embeddings of a given tree $\mathbf{s}$ into $\mathbf{t}_k^c$, the complete binary tree with 2^k leaves.

It will be useful to introduce the infinite complete binary tree. This is the set $\{0,1\}^* \sqcup \{0,1\}^\infty$. For distinct points x and y in $\{0,1\}^\infty$ with $x = u_1u_2\cdots$ and $y = v_1v_2\cdots$, set $x \wedge y = u_1\cdots u_h = v_1\cdots v_h \in \{0,1\}^\star$, where $h = \max\{g : u_1\cdots u_g = v_1\cdots v_g\}$. We say that x is below and to the left of $x \wedge y$ and y is below and to the right of $x \wedge y$ if $u_{h+1} = 0$ and $v_{h+1} = 1$.

Using the same notation, put $r(x,y) = 2^{-h}$ and set $r(z,z) = 0$ for $z \in \{0,1\}^\infty$. Then r is a metric on $\{0,1\}^\infty$ that induces the (compact) product topology on $\{0,1\}^\infty$. We can equip $\{0,1\}^\infty$ with the probability measure κ that is the product of the uniform probability measures on each of the factors (i.e., κ is fair coin-tossing measure). The κ-measure of any ball with diameter $2^{-\ell}$ is $2^{-\ell}$.

If $x_1, \dots, x_{m+1}$ are distinct points in $\{0,1\}^\infty$, then these points induce a (finite) binary tree with $m+1$ leaves in the obvious way: we may identify the most recent common ancestor of the leaves corresponding to x_i and x_j with $x_i \wedge x_j$ and say that the point corresponding to x_i is below and to the left of the most recent common ancestor of the points corresponding to x_i and x_j in the reduced tree if x_i is below and to the left of $x_i \wedge x_j$, etc. Call this tree $T(x_1, \dots, x_{m+1})$. Observe that $T(x_1, \dots, x_{m+1}) = T(x_{\pi(1)}, \dots, x_{\pi(m+1)})$ for any permutation π of $\{1, 2, \dots, m+1\}$.

Suppose that the tree $\mathbf{s}$ has $m+1$ leaves. If the leaves of an embedding of $\mathbf{s}$ into $\mathbf{t}_k^c$ are $y_i = u_{i1}\cdots u_{ik}$ for $1 \le i \le m+1$ and we set $x_i = u_{i1}\cdots u_{ik}u_{i,k+1}u_{i,k+2}\cdots$ for any choice of $u_{i,k+1}, u_{i,k+2}, \dots, 1 \le i \le m+1$, then $T(x_1, \dots, x_{m+1}) = \mathbf{s}$. Conversely, if $x_i = u_{i1}u_{i2}\cdots \in \{0,1\}^\infty$, $1 \le i \le m+1$, are such that $T(x_1, \dots, x_{m+1}) = \mathbf{s}$ and $r(x_i, x_j) > 2^{-k}$ for $1 \le i \ne j \le m+1$, then putting $y_i = u_{i1}\cdots u_{ik}$ for $1 \le i \le m+1$ gives the leaves of an embedding of $\mathbf{s}$ into $\mathbf{t}_k^c$.

With the notation $x = (x_1, \dots, x_{m+1})$, it follows that

$$\frac{1}{(m+1)!} \frac{\kappa^{\otimes(m+1)}\{x : T(x) = \mathbf{s} \text{ and } r(x_i, x_j) > 2^{-k} \text{ for } 1 \le i \ne j \le m+1\}}{\kappa^{\otimes(m+1)}\{x : r(x_i, x_j) > 2^{-k} \text{ for } 1 \le i \ne j \le m+1\}}$$
$$= 2^{-(m+1)k} N(\mathbf{s}, \mathbf{t}_k^c).$$

Indeed, the left-hand side counts the fraction of all those of the (in total $2^{k(m+1)}$) mappings from $[m+1]$ to $[2^k]$ which correspond to an embedding of $\mathbf{s}$ into $\mathbf{t}_k^c$. In particular,

$$\lim_{k\to\infty} 2^{-(m+1)k} N(\mathbf{s}, \mathbf{t}_k^c)$$
$$= \frac{1}{(m+1)!} \kappa^{\otimes(m+1)}\{(x_1, \dots, x_{m+1}) : T(x_1, \dots, x_{m+1}) = \mathbf{s}\}.$$

THEOREM 3.5. *Suppose that* $\mathbf{s}$ *and* $\mathbf{t}$ *are two binary trees with, respectively,* $2m+1$ *and* $2(m+n)+1$ *vertices. Then the probability that the Rémy chain transitions from* $\mathbf{s}$ *to* $\mathbf{t}$ *in* n *steps is*

$$p(\mathbf{s}, \mathbf{t}) = n! \frac{1}{(2m+1) \times (2m+3) \times \cdots \times (2(m+n)-1)} \frac{1}{2^n} N(\mathbf{s}, \mathbf{t}),$$

where $N(\mathbf{s}, \mathbf{t})$ *is the number of ways of embedding* $\mathbf{s}$ *into* $\mathbf{t}$.

PROOF. We condition on the event $\{T_m = \mathbf{s}\}$ and say that a vertex of T_{m+n} is a *clonal descendant* of a vertex $v \in \mathbf{s}$ if it is v itself, a clone of v, a clone-of-a-clone of v, etc. We can then decompose T_{m+n} into connected pieces according to their clonal descent from the vertices of $\mathbf{s}$; see Figure 6 for a schematic representation of such a decomposition.

It follows from the definition of the Rémy chain that the numbers of clonal descendants of the $2m+1$ vertices of $\mathbf{s}$ are the result of n steps in a *Pólya urn*

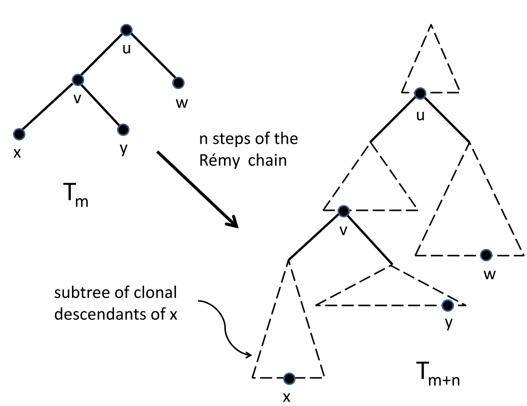

FIG. 6. *Decomposition of the tree* T_{m+n} *via clonal descent from the vertices of the tree* $T_m = \mathbf{s}$.

that starts with $2m + 1$ balls of different colors and at each stage a ball is chosen uniformly at random and replaced along with **two** balls of the same color.

Because the Rémy chain generates uniformly distributed binary trees, it further follows that conditional on the various numbers of clonal descendants, the respective binary trees of clonal descendants are independent and uniformly distributed.

Moreover, a straightforward induction shows that, conditional on the trees of clonal descendants, the ancestral vertices from $\mathbf{s}$ are located at independently and uniformly chosen leaves of their respective trees of clonal descendants.

Therefore, if we label the vertices of $\mathbf{s}$ with $1, \ldots, 2m+1$, then the conditional probability given $\{T_m = \mathbf{s}\}$ that the operation of a further n steps of Rémy's algorithm results in a binary tree $\mathbf{t}$ enhanced with a particular clonal descent decomposition in which $2n_j + 1$ vertices are clonal descendants of vertex j of $\mathbf{s}$ for $1 \le j \le 2m+1$ is

$$\begin{aligned}
&\frac{n!}{n_1! \cdots n_{2m+1}!} \frac{\prod_{j=1}^{2m+1} [1 \times 3 \times \cdots \times (2n_j - 1)]}{(2m+1) \times (2m+3) \times \cdots \times (2(m+n)-1)} \\
&\quad \times \prod_{j=1}^{2m+1} \frac{1}{C_{n_j}} \prod_{j=1}^{2m+1} \frac{1}{n_j + 1} \\
&\quad = n! \frac{1}{(2m+1) \times (2m+3) \times \cdots \times (2(m+n)-1)} \frac{1}{2^n},
\end{aligned}$$

and the result is immediate. $\square$

REMARK 3.6. An alternative method for proving Theorem 3.5 is to use arguments similar to those used in the Introduction to show that Rémy's algorithm does indeed generate uniform random binary trees. More precisely, let $\tilde{\mathbf{s}}$ be a tree with $m+1$ leaves labeled by $[m+1]$ and let $\tilde{\mathbf{t}}$ be a tree with $(m+n)+1$ leaves labeled by $[(m+n)+1]$. Recalling the construction of the enhanced chain $\tilde{T}_1, \tilde{T}_2, \ldots$, the conditional probability of the event $\{\tilde{T}_{m+n} = \tilde{\mathbf{t}}\}$ given the event $\{\tilde{T}_m = \tilde{\mathbf{s}}\}$ is either zero if the leaf-labeled binary tree induced by the leaves of $\tilde{\mathbf{t}}$ labeled with $[m+1]$ is not $\tilde{\mathbf{s}}$ or

$$\frac{1}{(2m+1) \times (2m+3) \times \cdots \times (2(m+n)-1)} \frac{1}{2^n}$$

if it is, because, as in the Introduction, the leaf-labeling dictates the order in which vertices must be cloned, as well as the associated choices of left-right reattachments. If $\mathbf{s}$ and $\mathbf{t}$ are unlabeled binary trees with $m+1$ and $(m+n)+1$ leaves, respectively, then for any labeling of the leaves of $\mathbf{s}$ to give a leaf–labeled binary tree $\tilde{\mathbf{s}}$, the number of ways of labeling $\mathbf{t}$ to give a leaf-labeled binary tree $\tilde{\mathbf{t}}$ such that the leaf-labeled binary tree induced by the leaves labeled with $[m+1]$ is just $n!N(\mathbf{s}, \mathbf{t})$, because any admissible labeling of $\mathbf{t}$ corresponds to an embedding of $\mathbf{s}$ into $\mathbf{t}$ composed with a labeling [with $\{m+1, \ldots, (m+n)+1\}$] of those leaves of $\mathbf{t}$ that are not in the image of $\mathbf{s}$.

COROLLARY 3.7. *Suppose that* $\mathbf{s}$ *and* $\mathbf{t}$ *are two binary trees with, respectively,* $2m+1$ *and* $2(m+n)+1$ *vertices. Then the corresponding Doob–Martin kernel is*

$$K(\mathbf{s},\mathbf{t}) = C_{m+n}\, p(\mathbf{s},\mathbf{t}) = 2^m \frac{1 \times 3 \times \cdots \times (2m-1)}{(n+1) \times (n+2) \times \cdots \times (m+n+1)} N(\mathbf{s},\mathbf{t}).$$

PROOF. This is immediate from the definition of the Doob–Martin kernel and Theorem 3.5. □

NOTATION 3.8. Given $m \in \mathbb{N}$ and a binary tree $\mathbf{t}$ with $2(m+n)+1$ vertices for some $n \in \mathbb{N}$, define $S_m^{\mathbf{t}}$ to be the random binary tree embedded in $\mathbf{t}$ that is obtained by picking $m+1$ leaves of $\mathbf{t}$ uniformly at random without replacement; see Figure 7.

COROLLARY 3.9. *Suppose that* $\mathbf{s}$ *and* $\mathbf{t}$ *are two binary trees with, respectively,* $2m+1$ *and* $2(m+n)+1$ *vertices. Then*

$$\mathbb{P}\{S_m^{\mathbf{t}} = \mathbf{s}\} = \frac{1}{C_m} K(\mathbf{s},\mathbf{t}).$$

PROOF. It suffices to observe that

$$\begin{aligned}\mathbb{P}\{S_m^{\mathbf{t}} = \mathbf{s}\} &= \frac{N(\mathbf{s},\mathbf{t})}{\binom{m+n+1}{m+1}} \\ &= \frac{(m+1)!}{(n+1) \times (n+2) \times \cdots \times (m+n+1)} N(\mathbf{s},\mathbf{t})\end{aligned}$$

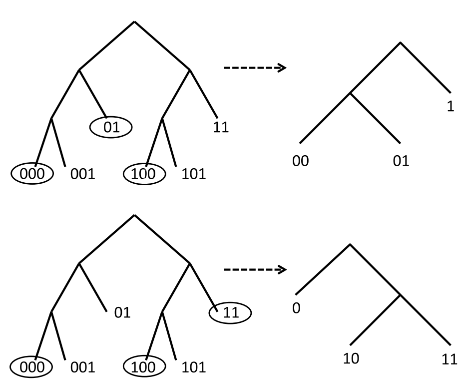

FIG. 7. *Two possible realizations of the random tree* $S_m^{\mathbf{t}}$ *when* $m=2$, $n=3$ *and* $\mathbf{t}$ *is the binary tree with* $11 = 2(2+3)+1$ *vertices depicted twice on the left-hand side along with an indication of its representation as a subset of* $\{0,1\}^*$. *Picking the* $m+1=3$ *leaves* 000, 01 *and* 100 *out of the* $(m+n)+1=6$ *leaves of* $\mathbf{t}$ *as shown in the top row results in a realization of* $S_m^{\mathbf{t}}$ *that has leaves* 00, 01 *and* 1 *in its representation as a subset of* $\{0,1\}^*$, *while picking the leaves* 001, 100 *and* 11 *of* $\mathbf{t}$ *as shown in the bottom row results in a realization of* $S_m^{\mathbf{t}}$ *that has leaves* 0, 10 *and* 11 *in its representation as a subset of* $\{0,1\}^*$.

$$= \frac{(m+1)!}{2^m \times 1 \times 3 \times \cdots \times (2m-1)} K(\mathbf{s}, \mathbf{t})$$

$$= \frac{1}{C_m} K(\mathbf{s}, \mathbf{t}). \qquad \square$$

The following result is immediate from Corollary 3.9. It shows that convergence of a sequence of binary trees to a point in the Doob–Martin boundary is equivalent to the convergence in distribution of the random embedded subtrees resulting from sampling a finite number of leaves uniformly at random. Thus, convergence in our setting is, as remarked in the Introduction, analogous to the notion of convergence of dense graph sequences as explored in the theory of graph limits, where a sequence of larger and larger graphs converges to a limit if the random subgraphs defined by restriction to a finite number of vertices sampled uniformly at random converge in distribution (see, e.g., [22], Chapter 13). The latter notion of convergence is metrized by a very natural metric called the *cut metric* that is, a priori, unrelated to sampling from a graph and it would be interesting to know if there is an analogous object that metrizes the notion of convergence of binary trees in our setting.

COROLLARY 3.10. *A sequence $(\mathbf{t}_k)_{k\in\mathbb{N}}$ of binary trees with the number of leaves of $\mathbf{t}_k$ going to infinity as $k \to \infty$ converges in the Doob–Martin compactification if and only if for each $m \in \mathbb{N}$ the sequence of random binary trees $(S_m^{\mathbf{t}_k})_{k\in\mathbb{N}}$ converges in distribution. Moreover, two such convergent sequences of binary trees $(\mathbf{t}'_k)_{k\in\mathbb{N}}$ and $(\mathbf{t}''_k)_{k\in\mathbb{N}}$ converge to the same point in the Doob–Martin boundary if and only if for all $m \in \mathbb{N}$ the limiting distribution of $S_m^{\mathbf{t}'_k}$ as $k \to \infty$ coincides with the limiting distribution of $S_m^{\mathbf{t}''_k}$ as $k \to \infty$.*

EXAMPLE 3.11. Recall that $\mathbf{t}_k^c$ is the complete binary tree with 2^k leaves. It follows from Corollary 3.7, the last equality in Example 3.4 and Corollary 3.10 that the sequence $(\mathbf{t}_k^c)_{k\in\mathbb{N}}$ converges in the Doob–Martin topology with

$$\begin{aligned}
\lim_{k\to\infty} K\bigl(\mathbf{s}, \mathbf{t}_k^c\bigr) &= 2^m\bigl(1 \times 3 \times \cdots \times (2m-1)\bigr)\frac{1}{(m+1)!} \\
&\quad \times \kappa^{\otimes(m+1)}\bigl\{(x_1,\dots,x_{m+1}) : T(x_1,\dots,x_{m+1}) = \mathbf{s}\bigr\} \\
&= C_m \kappa^{\otimes(m+1)}\bigl\{(x_1,\dots,x_{m+1}) : T(x_1,\dots,x_{m+1}) = \mathbf{s}\bigr\}
\end{aligned}$$

for a binary tree $\mathbf{s}$ with $m+1$ leaves. Equivalently,

$$\lim_{k\to\infty} \mathbb{P}\bigl\{T_m = \mathbf{s} \mid T_{2^k-1} = \mathbf{t}_k^c\bigr\} = \kappa^{\otimes(m+1)}\bigl\{(x_1,\dots,x_{m+1}) : T(x_1,\dots,x_{m+1}) = \mathbf{s}\bigr\}.$$

The latter probability can be evaluated quite explicitly. Let $X_1, \dots, X_{m+1}$ be independent, identically distributed $\{0,1\}^\infty$-valued random variables with common distribution κ. We label the balls of $\{0,1\}^\infty$ that have diameter 2^{-k} with the elements of $\{0,1\}^k$ by declaring that $B_{u_1\cdots u_k}$ is the unique ball containing all points of the form $u_1\cdots u_k u_{k+1}u_{k+2}\cdots$ for arbitrary $u_{k+1},u_{k+2},\ldots\in\{0,1\}$. There is a random integer R such that $\{X_1,\ldots,X_{m+1}\}\subset B_{u_1\cdots u_R}$ for some $u_1,\ldots,u_R\in\{0,1\}$, but $\{X_1,\ldots,X_{m+1}\}\not\subset B_{u_1\cdots u_R0}$ and $\{X_1,\ldots,X_{m+1}\}\not\subset B_{u_1\cdots u_R1}$. Observe that

$$\mathbb{P}\bigl\{\#\{i:X_i\in B_{u_1\cdots u_R0}\}=h,\#\{j:X_j\in B_{u_1\cdots u_R1}\}=m+1-h|R,u_1,\ldots,u_R\bigr\}$$
$$=\frac{\binom{m+1}{h}(\frac{1}{2})^{m+1}}{1-2(\frac{1}{2})^{m+1}}$$

for $1\le h\le m$. Moreover, given that $\#\{i:X_i\in B_{u_1\cdots u_R0}\}=h$ and $\#\{j:X_j\in B_{u_1\cdots u_R1}\}=m+1-h$, the set of locations of the X_i that fall in $B_{u_1\cdots u_R0}$ and the set of locations of the X_i that fall in $B_{u_1\cdots u_R1}$ are independent, with the former random set being conditionally distributed as h i.i.d. draws from the probability measure κ restricted to $B_{u_1\cdots u_R0}$ and renormalized to be a probability measure, and with the latter random set being conditionally distributed as $m+1-h$ i.i.d. draws from the probability measure κ restricted to $B_{u_1\cdots u_R1}$ and renormalized to be a probability measure. Label the internal vertices of $\mathbf{s}$ with $1,\ldots,m$. Let α_ℓ (resp., β_ℓ) be the number of leaves of $\mathbf{s}$ that are below and to the left (resp., below and to the right) of vertex ℓ and write $\gamma_\ell:=\alpha_\ell+\beta_\ell$ for the total number of leaves below the vertex labeled ℓ. It follows that

$$\begin{aligned}
&\kappa^{\otimes(m+1)}\bigl\{(x_1,\ldots,x_{m+1}):T(x_1,\ldots,x_{m+1})=\mathbf{s}\bigr\}\\
&\quad=\prod_{\ell=1}^m\frac{(\alpha_\ell+\beta_\ell)!}{\alpha_\ell!\beta_\ell!}\left(\frac{1}{2}\right)^{\alpha_\ell+\beta_\ell}\left(1-\left(\frac{1}{2}\right)^{\alpha_\ell+\beta_\ell-1}\right)^{-1}\\
&\quad=(m+1)!\frac{1}{2^m}\prod_{\ell=1}^m\bigl(2^{\alpha_\ell+\beta_\ell-1}-1\bigr)^{-1}\\
&\quad=(m+1)!\frac{1}{2^m}\prod_{\ell=1}^m\bigl(2^{\gamma_\ell-1}-1\bigr)^{-1},
\end{aligned}$$

where the second equality results from a telescope product along the binary tree $\mathbf{s}$. In particular, the function that maps $\mathbf{s}$ to

$$C_m(m+1)!\frac{1}{2^m}\prod_{\ell=1}^m\bigl(2^{\gamma_\ell-1}-1\bigr)^{-1}=1\times3\times\cdots\times(2m-1)\times\prod_{\ell=1}^m\bigl(2^{\gamma_\ell-1}-1\bigr)^{-1}$$

is harmonic for the Rémy chain. We can write this function more compactly as

$$\mathfrak{h}:\mathbf{s}\mapsto1\times3\times\cdots\times(2m-1)\times\prod_v\bigl(2^{\#\mathbf{s}(v)-1}-1\bigr)^{-1},$$

where the product is over the interior vertices of $\mathbf{s}$, and $\#\mathbf{s}(v)$ is the number of leaves below the interior vertex v.

It is instructive to check directly that this function is indeed harmonic. Suppose that in one step of the chain starting from the tree $\mathbf{s}$ with $2m+1$ vertices the vertex v of $\mathbf{s}$ is chosen to be cloned. This produces a tree $\mathbf{t}$ with $2m+1$ old vertices that we can identify with the vertices of $\mathbf{s}$ and two new vertices that we will call x and y, with x an interior vertex and y a leaf. If $u \neq v$ is an interior vertex of $\mathbf{s}$ that is on the path from the root to v (i.e., u is an ancestor of v), then $\#\mathbf{t}(u) = \#\mathbf{s}(u) + 1$. For any other interior vertex u of $\mathbf{s}$, we have $\#\mathbf{t}(u) = \#\mathbf{s}(u)$. Lastly, $\#\mathbf{t}(x) = \#\mathbf{s}(v) + 1$, where we put $\#\mathbf{s}(v) = 1$ if v is a leaf of $\mathbf{s}$. Therefore, if v is an interior vertex of $\mathbf{s}$, then

$$\begin{aligned}
&1 \times 3 \times \cdots \times (2m+1) \times \prod_w \left(2^{\#\mathbf{t}(w)-1} - 1\right)^{-1} \\
&\quad = 1 \times 3 \times \cdots \times (2m+1) \times \left[\prod_{u<v} \left(2^{\#\mathbf{s}(u)} - 1\right)^{-1}\right] \\
&\quad \times \left(2^{\#\mathbf{s}(v)-1} - 1\right)^{-1} \times \left(2^{\#\mathbf{s}(v)} - 1\right)^{-1} \\
&\quad \times \left[\prod_{u \not\leq v} \left(2^{\#\mathbf{s}(u)-1} - 1\right)^{-1}\right],
\end{aligned}$$

whereas if v is a leaf, then

$$\begin{aligned}
&1 \times 3 \times \cdots \times (2m+1) \times \prod_w \left(2^{\#\mathbf{t}(w)-1} - 1\right)^{-1} \\
&\quad = 1 \times 3 \times \cdots \times (2m+1) \times \left[\prod_{u<v} \left(2^{\#\mathbf{s}(u)} - 1\right)^{-1}\right] \times \left[\prod_{u \not\leq v} \left(2^{\#\mathbf{s}(u)-1} - 1\right)^{-1}\right].
\end{aligned}$$

Writing I for the set of internal vertices of $\mathbf{s}$ and L for the leaves, we therefore see that harmonicity of $\mathfrak{h}$ is equivalent to

$$\begin{aligned}
&\sum_{v \in I} \prod_{u<v} \frac{2^{\#\mathbf{s}(u)-1} - 1}{2^{\#\mathbf{s}(u)} - 1} \frac{1}{2^{\#\mathbf{s}(v)} - 1} + \sum_{v \in L} \prod_{u<v} \frac{2^{\#\mathbf{s}(u)-1} - 1}{2^{\#\mathbf{s}(u)} - 1} \\
&\quad = \sum_{v \in \mathbf{s}} \prod_{u<v} \frac{2^{\#\mathbf{s}(u)-1} - 1}{2^{\#\mathbf{s}(u)} - 1} \frac{1}{2^{\#\mathbf{s}(v)} - 1} \\
&\quad = 1.
\end{aligned}$$

This, however, is clear by induction. It is certainly true if $\mathbf{s}$ is the trivial binary tree with a single vertex or the binary tree $\aleph$ with three vertices. Assuming for some binary tree $\mathbf{s}$ with $m+1$ leaves that it is true for all binary trees with fewer leaves, we see from a consideration of the left and right subtrees below the root of $\mathbf{s}$ that the sum in question is

$$\frac{1}{2^{m+1} - 1} + \frac{2^m - 1}{2^{m+1} - 1}[1 + 1] = 1,$$

as required.

The one-step transition probability for the corresponding Doob h-transformed chain is, for binary trees $\mathbf{s}$ and $\mathbf{t}$ with $2m+1$ and $2m+3$ vertices,

$$\begin{aligned}&\left[1 \times 3 \times \cdots \times (2m-1) \times \prod_u (2^{\#\mathbf{s}(u)-1}-1)^{-1}\right]^{-1}\\ &\quad\times \frac{1}{2(2m+1)} N(\mathbf{s},\mathbf{t})\\ &\quad\times 1 \times 3 \times \cdots \times (2m+1) \times \prod_v (2^{\#\mathbf{t}(v)-1}-1)^{-1}\\ &\quad= \frac{1}{2}\frac{\prod_u (2^{\#\mathbf{s}(u)-1}-1)}{\prod_v (2^{\#\mathbf{t}(v)-1}-1)} N(\mathbf{s},\mathbf{t}),\end{aligned}$$

where the products in u run over the interior vertices of $\mathbf{s}$ and the products in v run over the interior vertices of $\mathbf{t}$.

It is apparent from the above that one step of the h-transformed chain starting from the state $\mathbf{s}$ can be described as follows:

- Pick a vertex v of $\mathbf{s}$ with probability
$$\prod_{u<v} \frac{2^{\#\mathbf{s}(u)-1}-1}{2^{\#\mathbf{s}(u)}-1} \frac{1}{2^{\#\mathbf{s}(v)}-1}.$$
- Cut off the subtree rooted at v and set it aside.
- Attach a copy of the tree $\aleph$ with 3 vertices to the end of the edge that previously led to v.
- Reattach the subtree rooted at v uniformly at random to one of the two leaves in the copy of $\aleph$.

4. Infinite Rémy bridges. Given a binary tree $\mathbf{t}$ with $2m(\mathbf{t})+1$ vertices, write $T_1^{\mathbf{t}}(=\aleph), T_2^{\mathbf{t}}, \ldots, T_{m(\mathbf{t})}^{\mathbf{t}}$ for the bridge process obtained by conditioning $T_1, \ldots, T_{m(\mathbf{t})}$ on the event $\{T_{m(\mathbf{t})} = \mathbf{t}\}$.

Recall from Section 2 that a sequence $(\mathbf{t}_k)_{k\in\mathbb{N}}$ with $m(\mathbf{t}_k) \to \infty$ converges in the Doob–Martin topology if and only if for each $\ell \in \mathbb{N}$ the random ℓ-tuple $(T_1^{\mathbf{t}_k}, \ldots, T_\ell^{\mathbf{t}_k})$ converges in distribution. Moreover, the various limits define a set of consistent distributions, and hence the distribution of a Markov chain $(T_n^\infty)_{n\in\mathbb{N}}$ with $T_1^\infty = \aleph$.

Note that if $\mathbf{s}, \mathbf{t}$ are binary trees with $2m+1$ and $2m+3$ vertices, respectively, then, using Theorem 3.5,

$$\begin{aligned}\mathbb{P}\{T_m^{\mathbf{t}_k} = \mathbf{s} | T_{m+1}^{\mathbf{t}_k} = \mathbf{t}\} &= \frac{\mathbb{P}\{T_m^{\mathbf{t}_k} = \mathbf{s}, T_{m+1}^{\mathbf{t}_k} = \mathbf{t}\}}{\mathbb{P}\{T_{m+1}^{\mathbf{t}_k} = \mathbf{t}\}}\\ &= \frac{\mathbb{P}\{T_{m+1}^{\mathbf{t}_k} = \mathbf{t} | T_m^{\mathbf{t}_k} = \mathbf{s}\}\mathbb{P}\{T_m^{\mathbf{t}_k} = \mathbf{s}\}}{\mathbb{P}\{T_{m+1}^{\mathbf{t}_k} = \mathbf{t}\}}\end{aligned}$$

$$
\begin{aligned}
&= \frac{p(\aleph, \mathbf{s})p(\mathbf{s}, \mathbf{t})p(\mathbf{t}, \mathbf{t}_k)}{p(\aleph, \mathbf{t})p(\mathbf{t}, \mathbf{t}_k)} \\
&= C_m^{-1} \frac{1}{2m+1} \frac{1}{2} N(\mathbf{s}, \mathbf{t}) / C_{m+1}^{-1} \\
&= \frac{(m+1)!m!}{(2m)!} \frac{1}{2m+1} \frac{1}{2} N(\mathbf{s}, \mathbf{t}) \frac{(2(m+1))!}{(m+2)!(m+1)!} \\
&= \frac{1}{m+2} N(\mathbf{s}, \mathbf{t}).
\end{aligned}
$$

Therefore, any finite bridge $(T_n^{\mathbf{t}})_{n=1}^{m(\mathbf{t})}$ and any limit of finite bridges $(T_n^\infty)_{n\in\mathbb{N}}$ evolves one step backward in time as follows:

- Pick a leaf uniformly at random.
- Delete the chosen leaf and its sibling (which may or may not be a leaf).
- If the sibling is not a leaf, then close up the resulting gap by attaching the subtree below the sibling to the parent of the chosen leaf and the sibling.

As we have already explained in the Introduction, understanding the Doob–Martin compactification is equivalent to understanding all Markov chains with initial state $\aleph$ that have these backward transition dynamics. We call any such a process an *infinite Rémy bridge*.

EXAMPLE 4.1. Suppose that $(\mathbf{t}_k)_{k\in\mathbb{N}}$ is the binary tree depicted in Figure 8. It is not hard to see that the sequence $(\mathbf{t}_k)_{k\in\mathbb{N}}$ converges in the Doob–Martin topology and that the value at time n of the corresponding limit bridge $(T_n^\infty)_{n\in\mathbb{N}}$ can be represented as the subset of $\{0,1\}^\star$ that consists of the vertices $\emptyset, \varepsilon_1, \varepsilon_1\varepsilon_2, \dots, \varepsilon_1\varepsilon_2\cdots\varepsilon_n$, where $\varepsilon_1, \dots, \varepsilon_n$ are independent $\{0,1\}$-valued random variables with $\mathbb{P}\{\varepsilon_i = 0\} = \mathbb{P}\{\varepsilon_i = 1\} = \frac{1}{2}$ for $1 \le i \le n$, plus the vertices $\bar{\varepsilon}_1, \varepsilon_1\bar{\varepsilon}_2, \dots, \varepsilon_1\varepsilon_2\cdots\bar{\varepsilon}_n$, where $\bar{\varepsilon}_i = 1 - \varepsilon_i$ for $1 \le i \le n$. In other words, T_n^∞ consists of a "spine" that moves to the left or right depending on the suc-

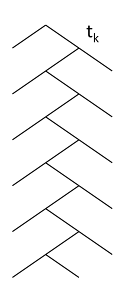

FIG. 8. *The binary tree* $\mathbf{t}_k$ *of Example* 4.1. *This tree has* $2k+1$ *vertices and consists of a single spine with leaves hanging off to the left and right alternately.*

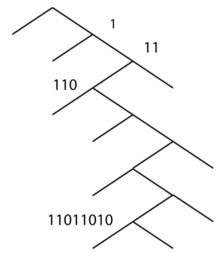

FIG. 9. *A realization at time $n = 9$ of the infinite Rémy bridge arising from the sequence of trees depicted in Figure 8. The random tree consists of leaves hanging off a single spine that moves to the left or right according to successive tosses of a fair coin.*

cessive tosses of a fair coin plus the minimal set of extra leaves that are required to yield a valid binary tree; see Figure 9. We stress that T_{n+1}^∞ is not obtained by simply appending an extra independent fair coin toss ε_{n+1} to the end of the sequence $\varepsilon_1, \dots, \varepsilon_n$. Rather, if we write $\varepsilon_1^n, \dots, \varepsilon_n^n$ for the coin tosses that correspond to T_n^∞ and $\varepsilon_1^{n+1}, \dots, \varepsilon_{n+1}^{n+1}$ for the coin tosses that correspond to T_{n+1}^∞, then $\varepsilon_1^{n+1}, \dots, \varepsilon_{n+1}^{n+1}$ is obtained from $\varepsilon_1^n, \dots, \varepsilon_n^n$ by inserting an additional independent toss uniformly at random into one of the $n + 1$ "slots" associated with the latter sequence—before the first toss, between two successive tosses, or after the last toss.

EXAMPLE 4.2. We know from Example 3.11 that if $\mathbf{t}_k^c$ is the complete binary tree with 2^k leaves, then the sequence $(\mathbf{t}_k^c)_{k\in\mathbb{N}}$ converges in the Doob–Martin topology. Moreover, it is clear that the value at time n of the corresponding infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$ is obtained by picking $n + 1$ points from $\{0, 1\}^\infty$ independently according to the probability measure κ and taking the finite binary tree they induce; see Figure 10.

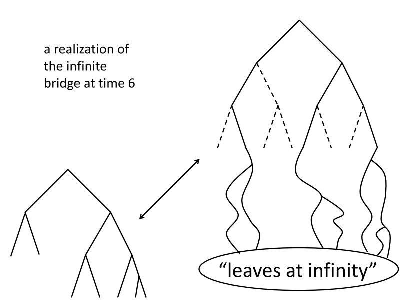

FIG. 10. *If $\mathbf{t}_k^c$ is the complete binary tree with 2^k leaves, then $\lim_k \mathbf{t}_k^c$ exists in the Doob–Martin topology. The random value at time n of the resulting infinite Rémy bridge can be built by choosing $n+1$ points independently and uniformly at random from the leaves at infinity of the infinite complete binary tree and constructing the tree they induce.*

5. Labeled infinite Rémy bridges and didendritic systems. Consider a binary tree T'' with $n+2$ leaves. Label the leaves of T'' with $[n+2]$ uniformly at random [i.e., all $(n+2)!$ labelings are equally likely]. Now apply the following deterministic procedure to produce a binary tree T' with $n+1$ leaves and a labeling of those leaves by $[n+1]$:

- Delete the leaf labeled $n+2$, along with its sibling (which may or may not be a leaf).
- If the sibling of the leaf labeled $n+2$ is also a leaf, then assign the sibling's label to the common parent (which is now a leaf).
- If the sibling of the leaf labeled $n+2$ is not a leaf, then attach the subtree below the sibling to the common parent with its leaf labels unchanged and leave all other leaf labels unchanged.

Clearly, the distribution of T' is that arising from one step starting from T'' of the backward Rémy dynamics (i.e., the common backward dynamics of all infinite Rémy bridges). Moreover, the labeling of T' by $[n+1]$ is uniformly distributed over the $(n+1)!$ possible labelings.

Now suppose that $(T_n^\infty)_{n\in\mathbb{N}}$ is an infinite Rémy bridge. For some N, let S_N be a random binary tree with the same distribution as T_N^∞. Label S_N uniformly at random with $[N+1]$ to produce a labeled binary tree $\tilde{S}_N$. Apply the above deterministic procedure successively for $n = N-1, \dots, 1$ to produce labeled binary trees $\tilde{S}_{N-1}, \dots, \tilde{S}_1$, where $\tilde{S}_n$ has $n+1$ leaves labeled by $[n+1]$ for $1 \le n \le N-1$. Write S_n for the underlying binary tree obtained by removing the labels of $\tilde{S}_n$. It follows from the observation above that the sequence $(S_1, \dots, S_N)$ has the same joint distribution as $(T_1^\infty, \dots, T_N^\infty)$. Note that the distribution of the sequence $(\tilde{S}_1, \dots, \tilde{S}_N)$ is uniquely determined by the distribution of T_N^∞, and hence, a fortiori, by the joint distribution of $(T_n^\infty)_{n\in\mathbb{N}}$. Note also that if we perform this construction for two different values of N, say $N' < N''$, to produce, with the obvious notation, sequences $(\tilde{S}'_1, \dots, \tilde{S}'_{N'})$ and $(\tilde{S}''_1, \dots, \tilde{S}''_{N''})$, then $(\tilde{S}'_1, \dots, \tilde{S}'_{N'})$ has the same distribution as $(\tilde{S}''_1, \dots, \tilde{S}''_{N'})$.

By Kolmogorov's extension theorem, we may therefore suppose that there is a Markov process $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ such that for each $n \in \mathbb{N}$ the random element $\tilde{T}_n^\infty$ is a leaf-labeled binary tree with $n+1$ leaves labeled by $[n+1]$ and the following hold:

- The binary tree obtained by removing the labels of $\tilde{T}_n^\infty$ is T_n^∞.
- For every $n \in \mathbb{N}$, the conditional distribution of $\tilde{T}_n^\infty$ given T_n^∞ is uniform over the $(n+1)!$ possible labelings of T_n^∞.
- In going backward from time $n+1$ to time n, $\tilde{T}_{n+1}^\infty$ is transformed into $\tilde{T}_n^\infty$ according to the deterministic procedure described above.

The distribution of $(\tilde{T}_n)_{n\in\mathbb{N}}$ is uniquely specified by the distribution of $(T_n^\infty)_{n\in\mathbb{N}}$ and the above requirements. Because of this distributional uniqueness, we refer to $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ as **the** *labeled version* of $(T_n^\infty)_{n\in\mathbb{N}}$ and $(T_n^\infty)_{n\in\mathbb{N}}$ as the *unlabeled version* of $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$. In a similar vein, we will talk about objects such as the "leaf of T_n^∞ labeled with $i \in [n+1]$".

We have just described the construction of a labeled infinite Rémy bridge from an unlabeled one. Using the Doob–Martin boundary, we can view this construction from a slightly different point of view as follows.

REMARK 5.1. Recall that for binary trees $\mathbf{s}$ and $\mathbf{t}$ with $n+1$ and $n+2$ leaves, respectively, the backward transition probability $q(\mathbf{s},\mathbf{t}) := \mathbb{P}\{T_n^\infty = \mathbf{s} | T_{n+1}^\infty = \mathbf{t}\}$ is the same for all infinite Rémy bridges $(T_n^\infty)_{n\in\mathbb{N}}$. For y in the Doob–Martin boundary of the Rémy chain, write $\mathbb{Q}^y$ for the distribution of the infinite Rémy bridge associated with y; that is, $\mathbb{Q}^y$ is the distribution of the Doob h-transform of the Rémy chain for the harmonic function $K(\cdot, y)$. Thus, $\mathbb{Q}^y$ assigns mass $q(\mathbf{t}_1,\mathbf{t}_2)q(\mathbf{t}_2,\mathbf{t}_3)\cdots q(\mathbf{t}_{n-1},\mathbf{t}_n)K(\mathbf{t}_n, y)$ to the set of paths that begin with the sequence of states $\mathbf{t}_1,\mathbf{t}_2,\dots,\mathbf{t}_{n-1},\mathbf{t}_n$. The distribution of an arbitrary infinite Rémy bridge is of the form $\int \mathbb{Q}^y \mu(dy)$ for some probability measure μ concentrated on the Doob–Martin boundary of the Rémy chain, and this representation is unique if μ is required to be concentrated on the minimal boundary. For labeled binary trees $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{t}}$ with $n+1$ and $n+2$ leaves, respectively, write $\tilde{q}(\tilde{\mathbf{s}},\tilde{\mathbf{t}}) := \mathbb{P}\{\tilde{T}_n^\infty = \tilde{\mathbf{s}} | \tilde{T}_{n+1}^\infty = \tilde{\mathbf{t}}\}$ for the backward transition probability common to all labeled infinite Rémy bridges $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$. The construction of a labeled infinite Rémy bridge $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ corresponding to an infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$ can be described as follows: if $(T_n^\infty)_{n\in\mathbb{N}}$ has distribution $\int \mathbb{Q}^y \mu(dy)$, then $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ has distribution $\int \tilde{\mathbb{Q}}^y \mu(dy)$, where $\tilde{\mathbb{Q}}^y$ is the probability measure that assigns mass $\tilde{q}(\tilde{\mathbf{t}}_1,\tilde{\mathbf{t}}_2)\tilde{q}(\tilde{\mathbf{t}}_2,\tilde{\mathbf{t}}_3)\cdots\tilde{q}(\tilde{\mathbf{t}}_{n-1},\tilde{\mathbf{t}}_n)\frac{1}{(n+1)!}K(\mathbf{t}_n, y)$ to the set of paths that begin with the sequence of states $\tilde{\mathbf{t}}_1,\tilde{\mathbf{t}}_2,\dots,\tilde{\mathbf{t}}_{n-1},\tilde{\mathbf{t}}_n$ and $\mathbf{t}_n$ is the binary tree obtained by removing the labels from $\tilde{\mathbf{t}}_n$.

It will be convenient for later use to be more concrete about the structure of the extra randomness introduced by labeling.

DEFINITION 5.2. Define a sequence of random variables $(L_n)_{n\in\mathbb{N}}$ by setting $L_n := k \in [n+1]$ if the leaf labeled $n+1$ in $\tilde{T}_n^\infty$ (and hence the one removed to form T_{n-1}^∞ from T_n^∞ and $\tilde{T}_{n-1}^\infty$ from $\tilde{T}_n^\infty$) is the kth smallest of the leaves of T_n^∞ in the lexicographic order on $\{0,1\}^*$ (recall that $v_1\cdots v_s$ is smaller than $w_1\cdots w_t$ in the lexicographic order if there is some $r < s \wedge t$ such that $v_q = w_q$ for $q \le r$, $v_{r+1} = 0$, and $w_{r+1} = 1$). For any $n \in \mathbb{N}$, it is clear that $L_1, L_2, \dots, L_n, (T_n^\infty, T_{n+1}^\infty, \dots)$ are independent and that L_n is uniformly distributed on $[n+1]$.

By construction, $(\tilde{T}_1^\infty, \dots, \tilde{T}_n^\infty)$ is a measurable function of $(L_1, \dots, L_n)$ and T_n^∞. It might be expected from this observation that the entire labeled infinite

Rémy bridge $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ [and hence, a fortiori, the infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$] may be constructed from $(L_n)_{n\in\mathbb{N}}$ and "boundary conditions" in the tail σ-field $\bigcap_{m\in\mathbb{N}}\sigma\{T_n^\infty : n \geq m\}$. The next result shows that this is indeed the case.

LEMMA 5.3. *For an infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$, its labeled version $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$, and the selection sequence $(L_n)_{n\in\mathbb{N}}$,*

$$\begin{aligned}\sigma\{\tilde{T}_n^\infty : n\in\mathbb{N}\} &= \bigcap_{m\in\mathbb{N}}\sigma\{\tilde{T}_n^\infty : n\geq m\}\\ &= \sigma\{L_p : p\in\mathbb{N}\}\vee\bigcap_{m\in\mathbb{N}}\sigma\{T_n^\infty : n\geq m\},\qquad \mathbb{P}\text{-}a.s.\end{aligned}$$

PROOF. Because $\tilde{T}_m^\infty$ is a measurable function of $\tilde{T}_n^\infty$ for $m < n$, the first two σ-fields are clearly equal, and since $(L_n)_{n\in\mathbb{N}}$ and $(T_n^\infty)_{n\in\mathbb{N}}$ are measurable functions of $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$, it is also clear that these two σ-fields both contain the third σ-field. Now $\tilde{T}_m^\infty$ is a measurable function of T_n^∞ and $L_1, \dots, L_n$ for $m < n$, and so to complete the proof it suffices to show that

$$\begin{aligned}&\sigma\{L_p : p\in\mathbb{N}\}\vee\bigcap_{m\in\mathbb{N}}\sigma\{T_n^\infty : n\geq m\}\\ &\quad\supseteq\bigcap_{m\in\mathbb{N}}\bigl(\sigma\{L_p : p\in\mathbb{N}\}\vee\sigma\{T_n^\infty : n\geq m\}\bigr),\qquad \mathbb{P}\text{-a.s.}\end{aligned}$$

That is, setting $\mathcal{F} := \sigma\{L_p : p\in\mathbb{N}\}$, $\mathcal{G}_m := \sigma\{T_n^\infty : n\geq m\}$, and $\mathcal{G}_\infty := \bigcap_{m\in\mathbb{N}}\mathcal{G}_m$, it is enough to establish that

$$\bigcap_{m\in\mathbb{N}}(\mathcal{F}\vee\mathcal{G}_m) = \mathcal{F}\vee\mathcal{G}_\infty.$$

Let $(\omega, A)\mapsto\mathbb{P}^{\mathcal{F}}(\omega, A)$, $\omega\in\Omega$, $A\in\mathcal{G}_1$, be the conditional probability kernel on $\mathcal{G}_1$ given $\mathcal{F}$. Because each σ-field $\mathcal{G}_m$ is countably generated, the desired equality will follow from the implication (d) $\Longrightarrow$ (a) of the main theorem of [34] if we can show that there is a countably generated σ-field $\mathcal{H}$ such that $\mathcal{G}_\infty = \mathcal{H} \mod \mathbb{P}^{\mathcal{F}}(\omega,\cdot)$ for $\mathbb{P}$-a.e. $\omega\in\Omega$. Because $\mathcal{F}$ and $\mathcal{G}_\infty$ are independent, $\mathbb{P}^{\mathcal{F}}(\omega,\cdot)$ restricted to $\mathcal{G}_\infty$ coincides with $\mathbb{P}$ restricted to $\mathcal{G}_\infty$ for $\mathbb{P}$-a.e. $\omega\in\Omega$. Now $(\omega, A)\mapsto\mathbb{P}(A)$, $\omega\in\Omega$, $A\in\mathcal{G}_1$, is certainly the conditional probability kernel on $\mathcal{G}_1$ given the trivial σ-field $\{\emptyset,\Omega\}$. Moreover, since

$$\bigcap_{m\in\mathbb{N}}\bigl(\{\emptyset,\Omega\}\vee\mathcal{G}_m\bigr) = \{\emptyset,\Omega\}\vee\mathcal{G}_\infty$$

obviously holds, it follows from the implication (a) $\Longrightarrow$ (d) of the main theorem of [34] that such a countably generated $\mathcal{H}$ does indeed exist. Alternatively, because $\mathcal{G}_1$ is countably generated, $L^1(\Omega,\mathcal{G}_1,\mathbb{P})$ contains a countable dense subset C. Let D be a collection of random variables that contains a version of $\mathbb{E}[\xi\,|\,\mathcal{G}_\infty]$ for each

$\xi \in C$. It is clear that D is dense in $L^1(\Omega, \mathcal{G}_\infty, \mathbb{P})$ and it suffices to take $\mathcal{H}$ to be the σ-field generated by D. $\square$

We now want to use the labeled infinite Rémy bridge to build an infinite binary-tree-like structure for which the set $\mathbb{N}$ plays the role of the leaves. Interior vertices in this infinite binary-tree-like structure have a left child and a right child, but we show that if we forget about this ordering, then there is an $\mathbb{R}$-tree such that, loosely speaking, the tree-like structure is that of the tree spanned by countably many points picked independently according to a certain probability measure on the $\mathbb{R}$-tree. The $\mathbb{R}$-tree is nonrandom when the infinite Rémy bridge is extremal, but even in that case further randomization may be necessary to reconstitute the left versus right ordering of children.

As will become clear in Section 8, Example 4.1 gives rise to a situation in which additional randomization is required to "distinguish left from right" after the countable collection of points has been sampled in order to fully reconstitute the binary tree-like structure; that is, it is not possible to impose a planar structure on the $\mathbb{R}$-tree so that the left versus right ordering of children in the subtree spanned by the sampled points is inherited from the planar structure on the $\mathbb{R}$-tree. The essential point here is that there is no Borel subset A of the unit interval with Lebesgue measure $\frac{1}{2}$ such that if U is a uniform random variable on the unit interval the random variables U and $\mathbb{1}_A(U)$ are independent.

However, no such additional randomization is necessary in Example 4.2 and the associated $\mathbb{R}$-tree can be augmented with a planar structure that induces the desired one on the subtree spanned by the sampled points.

DEFINITION 5.4. If $i, j \in [n+1]$ are the labels of two leaves of T_n^∞ that are represented by the words $u_1 \cdots u_k$ and $v_1 \cdots v_\ell$ in $\{0,1\}^\star$, then set $[i,j]_n := u_1 \cdots u_m = v_1 \cdots v_m$, where $m := \max\{h : u_h = v_h\}$. That is, $[i,j]_n$ is the *most recent common ancestor* in T_n^∞ of the leaves labeled i and j. Note that $[i,i]_n$ is just the leaf labeled i and every internal vertex of T_n^∞ is of the form $[i,j]_n$ for at least one pair (i,j) with $i \neq j$.

DEFINITION 5.5. Define an equivalence relation $\equiv$ on the Cartesian product $\mathbb{N} \times \mathbb{N}$ by declaring that $(i', j') \equiv (i'', j'')$ if and only if $[i', j']_n = [i'', j'']_n$ for some (and hence all) n such that $i', j', i'', j'' \in [n+1]$. We write $\langle i, j \rangle$ for the equivalence class of the pair (i,j). We will see that we can think of the equivalence classes as being the vertices of a binary-tree-like object. For $i \in \mathbb{N}$ the equivalence class of the pair (i,i) has only one element and it will sometimes be convenient to denote this equivalence class simply by i. With this convention, we regard $\langle i, j \rangle$ as being the most recent common ancestor of the leaves i and j.

DEFINITION 5.6. Define a partial order $<_L$ on the set of equivalence classes by declaring for $(i', j'), (i'', j'') \in \mathbb{N} \times \mathbb{N}$ that $\langle i', j' \rangle <_L \langle i'', j'' \rangle$ if and only if for some (and hence all) n such that $i', j', i'', j'' \in [n+1]$ we have $[i', j']_n = u_1 \cdots u_k$ and $[i'', j'']_n = u_1 \cdots u_k 0 v_1 \cdots v_\ell$ for some $u_1, \ldots, u_k, v_1, \ldots, v_\ell \in \{0, 1\}$. We interpret the ordering $\langle i', j' \rangle <_L \langle i'', j'' \rangle$ as the "vertex" $\langle i'', j'' \rangle$ being below and to the left of the "vertex" $\langle i', j' \rangle$. Similarly, we define another partial order $<_R$ by declaring that $\langle i', j' \rangle <_R \langle i'', j'' \rangle$ if and only if for some (and hence all) n such that $i', j', i'', j'' \in [n + 1]$ we have $[i', j']_n = u_1 \cdots u_k$ and $[i'', j'']_n = u_1 \cdots u_k 1 v_1 \cdots v_\ell$ for some $u_1, \ldots, u_k, v_1, \ldots, v_\ell \in \{0, 1\}$. We interpret the ordering $\langle i', j' \rangle <_R \langle i'', j'' \rangle$ as the "vertex" $\langle i'', j'' \rangle$ being below and to the right of the "vertex" $\langle i', j' \rangle$.

REMARK 5.7. The equivalence relation $\equiv$ and the partial orders $<_L$ and $<_R$ have a number of simple properties that it is useful to record:

- For $i, j \in \mathbb{N}$, $(i, j) \equiv (j, i)$.
- For $i, j, k \in \mathbb{N}$, $(i, j) \not\equiv (k, k)$ unless $i = j = k$.
- For $i, j \in \mathbb{N}$ with $i \neq j$, either $\langle i, j \rangle <_L \langle i, i \rangle$ and $\langle i, j \rangle <_R \langle j, j \rangle$, or $\langle i, j \rangle <_R \langle i, i \rangle$ and $\langle i, j \rangle <_L \langle j, j \rangle$.
- For $h, i, j, k \in \mathbb{N}$, if $\langle h, i \rangle <_L \langle j, k \rangle$, then $\langle h, i \rangle \not<_R \langle j, k \rangle$.
- For $h, i, j, k \in \mathbb{N}$, if $\langle h, i \rangle <_R \langle j, k \rangle$, then $\langle h, i \rangle \not<_L \langle j, k \rangle$.
- For $f, g, h, i, j, k \in \mathbb{N}$, if $\langle f, g \rangle <_L \langle h, i \rangle$ and $\langle h, i \rangle <_R \langle j, k \rangle$, then $\langle f, g \rangle <_L \langle j, k \rangle$.
- For $f, g, h, i, j, k \in \mathbb{N}$, if $\langle f, g \rangle <_R \langle h, i \rangle$ and $\langle h, i \rangle <_L \langle j, k \rangle$, then $\langle f, g \rangle <_R \langle j, k \rangle$.

DEFINITION 5.8. An equivalence relation on $\mathbb{N} \times \mathbb{N}$ and two partial orders on the associated equivalence classes form a *didendritic system* if they satisfy the conditions listed in Remark 5.7. (We have coined the word "didendritic" from the Greek roots "$\delta\iota\varsigma$" = "two, twice or double" and "$\delta\varepsilon\nu\delta\rho\iota\tau\eta\varsigma$" = "of or pertaining to a tree, tree-like" as an adjective meaning "binary tree-like".)

NOTATION 5.9. From now on, we will use the notation $\equiv$, $\langle \cdot, \cdot \rangle$, $<_L$ and $<_R$ to denote the equivalence relation, equivalence classes and the two partial orders of an arbitrary didendritic system.

REMARK 5.10. Given a didendritic system $(\equiv, \langle \cdot, \cdot \rangle, <_L, <_R)$ and $n \in \mathbb{N}$, there is a unique binary tree with $n+1$ leaves labeled by $i = \langle i, i \rangle$, $i \in [n+1]$, and internal vertices labeled by $\langle i, j \rangle$, $i, j \in [n + 1]$, $i \neq j$. Using the representation of binary trees as subsets of $\{0, 1\}^*$, the root $\varnothing$ is labeled by the unique equivalence class $\langle p, q \rangle$, $p, q \in [n + 1]$, such that there is no equivalence class $\langle r, s \rangle$, $r, s \in [n+1]$, for which $\langle r, s \rangle <_L \langle p, q \rangle$ or $\langle r, s \rangle <_R \langle p, q \rangle$. If the equivalence class $\langle h, i \rangle$, $h, i \in [n + 1]$, is the label of the vertex $v_1 \cdots v_r$ of the tree and the equivalence class $\langle j, k \rangle$, $j, k \in [n+1]$ is such that $\langle h, i \rangle <_L \langle j, k \rangle$ (resp., $\langle h, i \rangle <_R \langle j, k \rangle$) and there is no equivalence class $\langle \ell, m \rangle$ with $\langle h, i \rangle <_L \langle \ell, m \rangle$ (resp., and $\langle h, i \rangle <_R$

$\langle \ell, m\rangle$) and $\langle \ell, m\rangle <_L \langle j, k\rangle$ or $\langle \ell, m\rangle <_R \langle j, k\rangle$, then $v_1 \cdots v_r 0$ (resp., $v_1 \cdots v_r 1$) is a vertex of the tree with label $\langle j, k\rangle$.

If a labeled binary tree is constructed in this way and another one is constructed from the same didendritic system with n replaced by $n + 1$, then the first labeled binary tree can be produced from the second as follows:

- The leaf labeled $n + 2 = \langle n + 2, n + 2\rangle$ is deleted, along with its sibling (which may or may not be a leaf).
- If the sibling of the leaf labeled $n + 2$ is also a leaf, then the common parent (which is now a leaf) is assigned the sibling's label.
- If the sibling of the leaf labeled $n + 2$ is not a leaf, then the subtree below the sibling is attached to the common parent. The labelings of the vertices in the subtree are unchanged and the common parent is assigned the sibling's label.

DEFINITION 5.11. Given a didendritic system $\mathbf{D} = (\equiv, \langle\cdot,\cdot\rangle, <_L, <_R)$ and a permutation σ of $\mathbb{N}$ such that $\sigma(i) = i$ for all but finitely many $i \in \mathbb{N}$, the didendritic system $\mathbf{D}^\sigma = (\equiv^\sigma, \langle\cdot,\cdot\rangle^\sigma, <_L^\sigma, <_R^\sigma)$ is defined by:

- $(i', j') \equiv^\sigma (i'', j'')$ if and only if $(\sigma(i'), \sigma(j')) \equiv (\sigma(i''), \sigma(j''))$,
- $\langle i, j\rangle^\sigma$ is the equivalence class of the pair (i, j) for the equivalence relation $\equiv^\sigma$,
- $\langle h, i\rangle^\sigma <_L^\sigma \langle j, k\rangle^\sigma$ if and only if $\langle\sigma(h), \sigma(i)\rangle <_L \langle\sigma(j), \sigma(k)\rangle$,
- $\langle h, i\rangle^\sigma <_R^\sigma \langle j, k\rangle^\sigma$ if and only if $\langle\sigma(h), \sigma(i)\rangle <_R \langle\sigma(j), \sigma(k)\rangle$.

A random didendritic system $\mathbf{D} = (\equiv, \langle\cdot,\cdot\rangle, <_L, <_R)$ is *exchangeable* if for each permutation σ of $\mathbb{N}$ such that $\sigma(i) = i$ for all but finitely many $i \in \mathbb{N}$ the random didendritic system $\mathbf{D}^\sigma$ has the same distribution as $\mathbf{D}$.

In view of the similarity of the procedure described before Definition 5.11 and the procedure described at the beginning of this section, the following result is obvious and shows that characterizing the family of infinite Rémy bridges is equivalent to characterizing the family of exchangeable random didendritic systems.

LEMMA 5.12. *The random didendritic system corresponding to the labeled version of an infinite Rémy bridge is exchangeable. Conversely, the sequence of random labeled binary trees produced from an exchangeable random didendritic system by the procedure described in Remark 5.10 is an infinite Rémy bridge.*

With this result in mind, we now explore what sort of information is required to uniquely specify a didendritic system. From Remark 5.10, the subtree spanned by three distinct labeled leaves $i, j, k \in \mathbb{N}$ is one of twelve isomorphism types that we depict in Figure 11 along with notation for each one.

LEMMA 5.13. *Any didendritic system $(\equiv, \langle\cdot,\cdot\rangle, <_L, <_R)$ is uniquely determined by the isomorphism types of the subtrees spanned by all triples of distinct labeled leaves.*

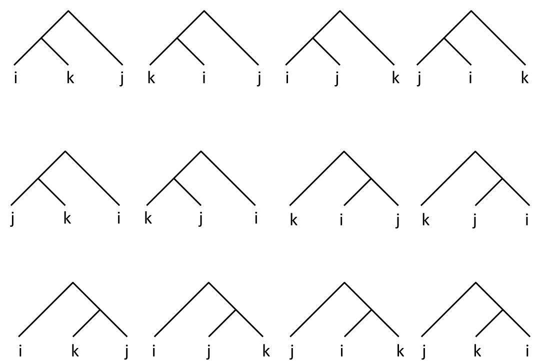

FIG. 11. *The isomorphism types for the subtree spanned by* 3 *leaves of a leaf-labeled binary tree. Going left to right and from top to bottom, we denote these types by* $((i,k),j)$, $((k,i),j)$, $((i,j),k),\dots,(i,(j,k)),(j,(i,k)),(j,(k,i))$.

PROOF. Observe that, $\langle h,i\rangle <_L \langle j,k\rangle$ for $h,i,j,k\in\mathbb{N}$ if and only if either one of the following six conditions holds or one of the three similar sets of six conditions with the roles of h and i interchanged or the roles of k and j interchanged holds:

- $((i,j),h)$ and $((i,k),h)$ and $((j,k),h)$ and $(i,(j,k))$,
- $((j,i),h)$ and $((k,i),h)$ and $((j,k),h)$ and $((j,k),i)$,
- $((i,j),h)$ and $((i,k),h)$ and $((j,k),h)$ and $((i,j),k)$,
- $((j,i),h)$ and $((i,k),h)$ and $((j,k),h)$ and $((j,i),k)$,
- $((j,i),h)$ and $((i,k),h)$ and $((j,k),h)$ and $(j,(i,k))$,
- $((j,i),h)$ and $((k,i),h)$ and $((j,k),h)$ and $(j,(k,i))$.

Moreover, $(h,i)\equiv(j,k)$ (that is, $\langle h,i\rangle=\langle j,k\rangle$ if and only if for all $\ell,m\in\mathbb{N}$, $\langle h,i\rangle <_L \langle \ell,m\rangle \iff \langle j,k\rangle <_L \langle \ell,m\rangle$ and $\langle h,i\rangle <_R \langle \ell,m\rangle \iff \langle j,k\rangle <_R \langle \ell,m\rangle$). $\square$

REMARK 5.14. It follows from Lemma 5.13 that any didendritic system has a unique coding as an array indexed by $\{(i,j,k)\in\mathbb{N}^3 : i,j,k\text{distinct}\}$, where the (i,j,k) entry records the isomorphism type of the subtree spanned by the leaves labeled i,j,k. The triply indexed random array corresponding to an exchangeable random didendritic system is jointly exchangeable in the usual sense for random arrays (see, e.g., [19], Section 7.1).

DEFINITION 5.15. Define a third partial order $<$ on the set of equivalence classes of $\mathbb{N}\times\mathbb{N}$ by declaring that $\langle h,i\rangle < \langle j,k\rangle$ if either $\langle h,i\rangle <_L \langle j,k\rangle$ or $\langle h,i\rangle <_R \langle j,k\rangle$. We interpret the ordering $\langle h,i\rangle < \langle j,k\rangle$ as the "vertex" $\langle j,k\rangle$ being below the "vertex" $\langle h,i\rangle$.

REMARK 5.16. It is easy to see that if $\langle h, i\rangle$ and $\langle j, k\rangle$ are two equivalence classes, then there is a unique "most recent common ancestor" $\langle \ell, m\rangle$ such that $\langle \ell, m\rangle \le \langle h, i\rangle$, $\langle \ell, m\rangle \le \langle j, k\rangle$, and if $\langle p, q\rangle$ also has these two properties, then $\langle p, q\rangle \le \langle \ell, m\rangle$. Moreover, we can choose ℓ, m so that $\ell \in \{h, i\}$ and $m \in \{j, k\}$. Indeed, for any $n \in \mathbb{N}$ we can, by Remark 5.10, think of the equivalence classes $\{\langle i, j\rangle : i, j \in [n+1]\}$ as the vertices of a binary tree with its leaves labeled by $[n+1]$. When the didendritic system was constructed from the labeled version $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ of an infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$, this leaf-labeled binary tree is just $\tilde{T}_n^\infty$.

LEMMA 5.17. *Any didendritic system $(\equiv, \langle\cdot,\cdot\rangle, <_L, <_R)$ is uniquely determined by the equivalence relation $\equiv$, the partial order $<$, and a determination for each pair of distinct labeled leaves $i, j \in \mathbb{N}$ whether*

$$\langle i, j\rangle <_L i \quad \textit{and} \quad \langle i, j\rangle <_R j$$

or

$$\langle i, j\rangle <_L j \quad \textit{and} \quad \langle i, j\rangle <_R i.$$

PROOF. Because of Lemma 5.13, it suffices to show that it is possible to reconstruct from the given data the isomorphism types of the subtrees spanned by all triples of distinct labeled leaves. For distinct $i, j, k \in \mathbb{N}$, the isomorphism type assignment $((i, k), j)$ is equivalent to

$$\langle i, j\rangle = \langle k, j\rangle < \langle i, k\rangle$$

and

$$\begin{aligned} \langle i, k\rangle <_L i \quad &\text{and} \quad \langle i, k\rangle <_R k, \\ \langle i, j\rangle <_L i \quad &\text{and} \quad \langle i, j\rangle <_R j, \\ \langle k, j\rangle <_L k \quad &\text{and} \quad \langle k, j\rangle <_R j. \end{aligned}$$

Similar observations for the other eleven isomorphism types establish the result. $\square$

REMARK 5.18. We have seen that any infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$ has a uniquely defined labeled version $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ [in the sense that the distribution of the sequence $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ is uniquely specified by the distribution of the sequence $(T_n^\infty)_{n\in\mathbb{N}}$] and also that a labeled infinite Rémy bridge corresponds, via a bijection between infinite bridge paths and didendritic systems, to a unique exchangeable random didendritic system.

Our aim is to find concrete representations of the extremal infinite Rémy bridges (recall that an infinite Rémy bridge is extremal if it has a trivial tail σ-field). To this end, it will be useful to relate the extremality of an infinite Rémy bridge to properties of the associated exchangeable random didendritic system. We say that an exchangeable random didendritic system $\mathbf{D}$ is *ergodic* if

$$\mathbb{P}(\{\mathbf{D} \in A\} \triangle \{\mathbf{D}^\sigma \in A\}) = 0$$

for all permutations σ of $\mathbb{N}$ such that $\sigma(i) = i$ for all but finitely many $i \in \mathbb{N}$ implies that

$$\mathbb{P}\{\mathbf{D} \in A\} \in \{0, 1\}.$$

By classical results on ergodic decompositions (see, e.g., [19], Theorem A 1.4), an exchangeable random didendritic system with distribution ε is ergodic if and only if there is no decomposition $\varepsilon = p'\varepsilon' + p''\varepsilon''$, where $\varepsilon', \varepsilon''$ are distinct distributions of exchangeable random didendritic systems, $p', p'' > 0$, and $p' + p'' = 1$. Also, it follows from Remark 5.14 and a result of Aldous (see, e.g., [19], Lemma 7.35) that ergodicity is further equivalent to the independence of the exchangeable random didendritic systems induced by disjoint subsets of $\mathbb{N}$, where here we extend the definition of a didendritic system in the obvious manner to allow an equivalence relation and partial orders that are defined on an underlying countable (possibly finite) set other than $\mathbb{N}$.

PROPOSITION 5.19. *An infinite Rémy bridge is extremal if and only if the associated exchangeable random didendritic system is ergodic.*

PROOF. Let $\mathcal{T}$ be the set of sequences of binary trees that can arise as a sample path of an infinite Rémy bridge and let $\tilde{\mathcal{T}}$ be the set of sequences of leaf-labeled binary trees that can arise as a sample path of a labeled infinite Rémy bridge.

The distribution α of an infinite Rémy bridge has a unique representation of the form $\alpha = \int \mathbb{Q}^y \mu(dy)$ for a probability measure μ concentrated on the minimal Doob–Martin boundary of the Rémy chain, where $\mathbb{Q}^y$ is the distribution of the infinite Rémy bridge corresponding to the boundary point y. The infinite Rémy bridge is extremal if and only if μ is a point mass, which is in turn equivalent to the condition that it is not possible to write $\alpha = p'\alpha' + p''\alpha''$, where α', α'' are distributions of infinite Rémy bridges, $p', p'' > 0$, and $p' + p'' = 1$.

Recall from Remark 5.1 that if $\alpha = \int \mathbb{Q}^y \mu(dy)$ is the distribution of an infinite Rémy bridge, where $\mathbb{Q}^y$ is the distribution of the infinite Rémy bridge corresponding to the boundary point y, then $\Lambda(\alpha) := \int \tilde{\mathbb{Q}}^y \mu(dy)$ is the distribution of the associated labeled infinite Rémy bridge, where $\tilde{\mathbb{Q}}^y$ is the distribution of the labeled infinite Rémy bridge corresponding to the boundary point y. Writing $\phi : \tilde{\mathcal{T}} \to \mathcal{T}$ for the map that removes the labels from each tree in a path, we see that the map Λ is bijective with inverse Υ given by $\Upsilon(\tilde{\alpha}) = \tilde{\alpha} \circ \phi^{-1}$ when $\tilde{\alpha}$ is the distribution of a labeled infinite Rémy bridge.

It is clear that if $\alpha, \alpha', \alpha''$ are distributions of infinite Rémy bridges, $p', p'' > 0$, $p' + p'' = 1$ and $\alpha = p'\alpha' + p''\alpha''$, then $\Lambda(\alpha) = p'\Lambda(\alpha') + p''\Lambda(\alpha'')$. Similarly, if $\tilde{\alpha}, \tilde{\alpha}', \tilde{\alpha}''$ are distributions of labeled infinite Rémy bridges, $p', p'' > 0$, $p' + p'' = 1$, and $\tilde{\alpha} = p'\tilde{\alpha}' + p''\tilde{\alpha}''$, then $\Upsilon(\tilde{\alpha}) = p'\Upsilon(\tilde{\alpha}') + p''\Upsilon(\tilde{\alpha}'')$. In short, an infinite Rémy bridge has a nontrivial tail σ-field if and only if it is distributed as a nontrivial mixture of infinite Rémy bridges, and this in turn is equivalent to the associated labeled infinite Rémy bridge being distributed as a nontrivial mixture of labeled infinite Rémy bridges.

Let $\mathcal{D}$ be the set of didendritic systems. Write $\psi : \tilde{\mathcal{T}} \to \mathcal{D}$ for the map that takes a sequence that can arise as a sample path of a labeled infinite Rémy bridge and turns it into a didendritic system.

Because ψ is a bijection, a probability measure γ on $\tilde{\mathcal{T}}$ that is the distribution of a labeled infinite Rémy bridge is a nontrivial mixture $\gamma = p'\gamma' + p''\gamma''$, where $p', p'' > 0$, $p' + p'' = 1$ and γ', γ'' are distinct distributions of labeled infinite Rémy bridges, if and only if $\gamma \circ \psi^{-1} = p'\varepsilon' + p''\varepsilon''$, where $\varepsilon', \varepsilon''$ are distinct probability measures on $\mathcal{D}$ (in which case $\varepsilon' = \gamma' \circ \psi^{-1}$ and $\varepsilon'' = \gamma'' \circ \psi^{-1}$).

Combining all of the above equivalent conditions establishes the result. $\square$

REMARK 5.20. The equivalence of Proposition 5.19 is central to the subsequent development and so we sketch the following alternative "bare hands" proof that is also interesting in its own right.

Consider an infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$, its labeled version $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$, the corresponding sequence $(L_n)_{n\in\mathbb{N}}$ defined in Definition 5.2, and the associated exchangeable random didendritic system $\mathbf{D} = (\equiv, \langle\cdot,\cdot\rangle, <_L, <_R)$.

Fix $m \in \mathbb{N}$. For $n \geq m$, let $\tilde{T}_{m,n}^\infty$ be the random partially leaf-labeled tree that is obtained from $\tilde{T}_n^\infty$ by removing those labels that belong to $[m+1]$. Thus, $m+1$ leaves of $\tilde{T}_{m,n}^\infty$ have no labels and the remaining $(n+1) - (m+1)$ leaves are labeled by elements of $[n+1] \setminus [m+1]$. The σ-field consisting of events of the form $\{\mathbf{D} \in A\}$ where A is such that $\mathbb{P}(\{\mathbf{D} \in A\} \triangle \{\mathbf{D}^\sigma \in A\}) = 0$ for all permutations σ of $\mathbb{N}$ that fix $\mathbb{N} \setminus [m+1]$ is $\mathbb{P}$-a.s. equal to $\sigma\{\tilde{T}_{m,n}^\infty : n \geq m\}$.

To establish Proposition 5.19, it will therefore suffice to show that the σ-field $\bigcap_{m\in\mathbb{N}} \sigma\{\tilde{T}_{m,n}^\infty : n \geq m\}$ is $\mathbb{P}$-trivial if and only if the σ-field $\bigcap_{m\in\mathbb{N}} \sigma\{T_n^\infty : n \geq m\}$ is $\mathbb{P}$-trivial. The former σ-field contains the latter, and hence it further suffices to show that if the latter σ-field is $\mathbb{P}$-trivial, then so is the former. We therefore suppose from now on that $\bigcap_{m\in\mathbb{N}} \sigma\{T_n^\infty : n \geq m\}$ is $\mathbb{P}$-trivial.

For any $m \leq n \leq p$, the random partially leaf-labeled tree $\tilde{T}_{m,n}^\infty$ is a measurable function of $L_{m+1}, \dots, L_p$ and T_p^∞, so that

$$\sigma\{\tilde{T}_{m,n}^\infty : n \geq m\} \subseteq \sigma\{L_k : k > m\} \vee \sigma\{T_q^\infty : q \geq p\},$$

for any $p \geq m$.

An argument similar to that in the proof of Lemma 5.3 combined with the $\mathbb{P}$-triviality of $\bigcap_{p\ge m}\sigma\{T_q^\infty : q\ge p\}$ gives

$$\begin{aligned}\sigma\{\tilde{T}_{m,n}^\infty : n\ge m\} &\subseteq \bigcap_{p\ge m}\big(\sigma\{L_k : k>m\}\vee\sigma\{T_q^\infty : q\ge p\}\big)\\ &= \sigma\{L_k : k>m\}\vee\bigcap_{p\ge m}\sigma\{T_q^\infty : q\ge p\}\\ &= \sigma\{L_k : k>m\} \qquad \mathbb{P}\text{-a.s.}\end{aligned}$$

Since $\bigcap_{m\in\mathbb{N}}\sigma\{L_k : k>m\}$ is $\mathbb{P}$-trivial by Kolmogorov's zero-one law, it follows that $\bigcap_{m\in\mathbb{N}}\sigma\{\tilde{T}_{m,n}^\infty : n\ge m\}$ is also $\mathbb{P}$-trivial, as required.

The next result shows that identifying the Doob–Martin boundary of the Rémy chain is equivalent to characterizing the extremal infinite Rémy bridges.

COROLLARY 5.21. *If y is an element of the Doob–Martin boundary of the Rémy chain, then the corresponding nonnegative harmonic function $K(\cdot, y)$ is extremal; equivalently, the corresponding infinite Rémy bridge is extremal. There is thus a bijective correspondence between the Doob–Martin boundary of the Rémy chain and the set of extremal infinite Rémy bridges.*

PROOF. Suppose that $(\mathbf{t}_p)_{p\in\mathbb{N}}$ is a sequence of binary trees, where $\mathbf{t}_p$ has $m(\mathbf{t}_p)+1$ leaves and $m(\mathbf{t}_p)\to\infty$ as $p\to\infty$. Suppose, moreover, that $\lim_{p\to\infty}\mathbf{t}_p = y$ for some y in the Doob–Martin boundary of the Rémy chain. We have to show that the harmonic function $K(\cdot, y)$ is extremal. Writing $(T_n^\infty)_{n\in\mathbb{N}}$ for the infinite Rémy bridge associated with y, this is equivalent to showing that the tail σ-field of $(T_n^\infty)_{n\in\mathbb{N}}$ is $\mathbb{P}$-a.s. trivial. By Proposition 5.19, this is further equivalent to establishing that the exchangeable random didendritic system $\mathbf{D}$ associated with $(T_n^\infty)_{n\in\mathbb{N}}$ is ergodic, which as we observed in Remark 5.18, is the same as proving that the exchangeable random didendritic systems $\mathbf{D}$ induces on disjoint (finite) subsets of $\mathbb{N}$ are independent (recall from Remark 5.18 our comment about generalizing the notion of a didendritic system from the setting where the underlying set is $\mathbb{N}$ to the setting where the underlying set is an arbitrary countable set).

Recall that $(T_1^{\mathbf{t}_p}, \dots, T_{m(\mathbf{t}_p)}^{\mathbf{t}_p})$ denotes the Rémy bridge to $\mathbf{t}_p$. For any $\ell\in\mathbb{N}$, $T_\ell^{\mathbf{t}_p}$ converges in distribution to T_ℓ^∞ as $p\to\infty$. We can build a labeled version $(\tilde{T}_1^{\mathbf{t}_p}, \dots, \tilde{T}_{m(\mathbf{t}_p)}^{\mathbf{t}_p})$ of $(T_1^{\mathbf{t}_p}, \dots, T_{m(\mathbf{t}_p)}^{\mathbf{t}_p})$ in much the same way that we built a labeled version of an infinite Rémy bridge: $\tilde{T}_{m(\mathbf{t}_p)}^{\mathbf{t}_p}$ consists of the binary tree $T_{m(\mathbf{t}_p)}^{\mathbf{t}_p} = \mathbf{t}_p$ with its $m(t_p)+1$ leaves labeled uniformly at random with $[m(t_p)+1]$ and the backward evolution of such a labeled finite Rémy bridge is the same as that of the labeled infinite Rémy bridge. It is clear that $\tilde{T}_\ell^{\mathbf{t}_p}$ converges in distribution to $\tilde{T}_\ell^\infty$

as $p \to \infty$ for all $\ell \in \mathbb{N}$: indeed, $\tilde{T}_\ell^{\mathbf{t}_p}$ and $\tilde{T}_\ell^\infty$ are just $T_\ell^{\mathbf{t}_p}$ and T_ℓ^∞, respectively, equipped with uniform random labelings of their $\ell + 1$ leaves by $[\ell + 1]$.

Suppose that $\ell \le m(\mathbf{t}_p)$. The labeled binary tree $\tilde{T}_\ell^{\mathbf{t}_p}$ (resp., $\tilde{T}_{m(\mathbf{t}_p)}^{\mathbf{t}_p}$) can be coded bijectively by an exchangeable random didendritic system $\mathbf{D}_{\ell,p}$ (resp., $\mathbf{D}_p$) on the finite set $[\ell + 1]$ (resp., $[m(\mathbf{t}_p) + 1]$), and $\mathbf{D}_{\ell,p}$ is the didendritic system on $[\ell + 1]$ induced by $\mathbf{D}_p$. The labeled tree $\tilde{T}_\ell^\infty$ can be coded bijectively by an exchangeable random didendritic system $\mathbf{D}_{\ell,\infty}$ on the finite set $[\ell + 1]$, and $\mathbf{D}_{\ell,\infty}$ is the didendritic system on $[\ell + 1]$ induced by $\mathbf{D}$. It follows from the convergence in distribution of $\tilde{T}_\ell^{\mathbf{t}_p}$ to $\tilde{T}_\ell^\infty$ as $p \to \infty$ for all $\ell \in \mathbb{N}$ that $\mathbf{D}_{\ell,p}$ converges in distribution to $\mathbf{D}_{\ell,\infty}$ as $p \to \infty$ for all $\ell \in \mathbb{N}$.

Let $\mathcal{I}$ denote the set of twelve possible isomorphism types for a labeled binary tree with three leaves. We know from Lemma 5.13 that $\mathbf{D}$ can be coded bijectively by a jointly exchangeable random array $\mathbf{Z}_\infty$, say, indexed by $\{(i, j, k) : i, j, k \in \mathbb{N}, i, j, k \text{ distinct}\}$ with values in $\mathcal{I}$. Similarly, $\mathbf{D}_{\ell,p}$, $\mathbf{D}_p$ and $\mathbf{D}_{\ell,\infty}$ can be coded bijectively by arrays that we denote by $\mathbf{Z}_{\ell,p}$, $\mathbf{Z}_p$ and $\mathbf{Z}_{\ell,\infty}$. The array $\mathbf{Z}_{\ell,p}$ (resp., $\mathbf{Z}_{\ell,\infty}$) is just the subarray of $\mathbf{Z}_p$ (resp., $\mathbf{Z}_\infty$) consisting of the entries indexed by $\{(i, j, k) : i, j, k \in [\ell + 1], i, j, k \text{ distinct}\}$. It follows from the convergence of $\mathbf{D}_{\ell,p}$ in distribution to $\mathbf{D}_{\ell,\infty}$ that $\mathbf{Z}_{\ell,p}$ converges in distribution to $\mathbf{Z}_{\ell,\infty}$ as $p \to \infty$ for all $\ell \in \mathbb{N}$.

Suppose that $H_1, \ldots, H_s$ are disjoint finite subsets of $\mathbb{N}$. We need to show that the exchangeable random didendritic systems that $\mathbf{D}$ induces on these sets are independent. This is equivalent to establishing that the subarrays of $\mathbf{Z}_\infty$ consisting of entries indexed by $\{(i, j, k) : i, j, k \in H_r, i, j, k \text{ distinct}\}$, $1 \le r \le s$, are independent. Taking ℓ so that $H_1 \sqcup \cdots \sqcup H_s \subseteq [\ell + 1]$, this is the same as proving that the subarrays of $\mathbf{Z}_{\ell,\infty}$ consisting of entries indexed by these same sets of triples are independent.

We can build the array $\mathbf{Z}_{\ell,p}$ using the binary tree $\mathbf{t}_p$ and random variables $\xi_1, \ldots, \xi_{\ell+1}$ that form a sequence of uniform random draws without replacement from the leaves of $\mathbf{t}_p$: the (i, j, k) entry of the array is the isomorphism type of the subtree of $\mathbf{t}_p$ spanned by the leaves ξ_i, ξ_j, ξ_k. Let $\dagger$ be an element not in $\mathcal{I}$, take $\zeta_1, \ldots, \zeta_{\ell+1}$ to be independent uniform random draws (with replacement) from the leaves of $\mathbf{t}_p$, and define an array $\mathbf{Z}_{\ell,p}^\dagger$ with the same index set as $\mathbf{Z}_{\ell,p}$ but with values in $\mathcal{I} \sqcup \{\dagger\}$ by letting the (i, j, k) entry of the array be the isomorphism type of the subtree of $\mathbf{t}_p$ spanned by the leaves $\zeta_i, \zeta_j, \zeta_k$ if $\zeta_i, \zeta_j, \zeta_k$ are distinct and $\dagger$ otherwise. A familiar coupling argument shows that it is possible to construct $\xi_1, \ldots, \xi_{\ell+1}$ and $\zeta_1, \ldots, \zeta_{\ell+1}$ on the same probability space in such a way that $\mathbb{P}\{\exists 1 \le i \le \ell + 1 : \xi_i \ne \zeta_i\}$ depends on $\mathbf{t}_p$ only through $m(\mathbf{t}_p)$ and converges to zero as $m(\mathbf{t}_p) \to \infty$; more specifically, we first construct $\zeta_1, \ldots, \zeta_{\ell+1}$, set $(\xi_1, \ldots, \xi_{\ell+1}) = (\zeta_1, \ldots, \zeta_{\ell+1})$ if $\zeta_1, \ldots, \zeta_\ell$ are distinct and let $(\xi_1, \ldots, \xi_\ell)$ be some other independent sequence of uniform draws without replacement from the leaves of $\mathbf{t}_p$ otherwise. Thus, $\mathbb{P}\{\mathbf{Z}_{\ell,p} \ne \mathbf{Z}_{\ell,p}^\dagger\}$ depends on t_p only through $m(t_p)$

and converges to zero as $m(t_p) \to \infty$. The subarrays of $\mathbf{Z}^{\dagger}_{\ell,p}$ consisting of entries indexed by $\{(i,j,k) : i,j,k \in H_r, i,j,k \text{ distinct}\}$, $1 \le r \le s$, are obviously independent because they are built from the binary tree $\mathbf{t}_p$ and the disjoint collections of random variables $\{\zeta_i : i \in H_r\}$, $1 \le r \le s$.

Combining the convergence in distribution of $\mathbf{Z}_{\ell,p}$ to $\mathbf{Z}_{\ell,\infty}$ as $p \to \infty$, the convergence to zero as $p \to \infty$ of the total variation distance between the distribution of $\mathbf{Z}_{\ell,p}$ and the distribution of $\mathbf{Z}^{\dagger}_{\ell,p}$, and the observation that the subarrays of $\mathbf{Z}^{\dagger}_{\ell,p}$ consisting of entries indexed by $\{(i,j,k) : i,j,k \in H_r, i,j,k \text{ distinct}\}$, $1 \le r \le s$, are independent, it is clear that the subarrays of $\mathbf{Z}_{\ell,\infty}$ indexed by these same sets of triples are independent, as required. $\square$

6. A real tree associated with an extremal infinite Rémy bridge. With Corollary 5.21 in hand, the task of identifying the Doob–Martin boundary of the Rémy chain reduces to characterizing the extremal infinite Rémy bridges, where we stress that such a characterization will also determine the topological structure of the boundary because convergence of boundary points is equivalent to convergence of finite-dimensional distributions of the corresponding infinite Rémy bridges.

The construction of Section 5 used the labeled version of an infinite Rémy bridge (equivalently, an exchangeable random didendritic system) to provide an embedding of $\mathbb{N}$ as the leaves of a tree-like combinatorial object whose vertices correspond to equivalence classes of the didendritic system's equivalence relation. In this section, we embed this tree-like object into an $\mathbb{R}$-tree by constructing a metric on the set of equivalence classes. We assume throughout this section that $(\equiv, \langle\cdot,\cdot\rangle, <_L, <_R)$ is an ergodic exchangeable random didendritic system and that $(\tilde{T}^{\infty}_n)_{n\in\mathbb{N}}$ is the labeled version of the associated extremal infinite Rémy bridge.

Consider $i,j \in \mathbb{N}$. For $p \in \mathbb{N}$ set

$$I_p := \mathbb{1}\{\langle i,j\rangle \le p\} \tag{6.1}$$

(recall our convention of writing p for the equivalence class $\langle p,p\rangle$).

By construction, the sequence of random variables $(I_p)_{p>i\vee j}$ is exchangeable. Hence, by de Finetti's theorem and the strong law of large numbers,

$$d(i,j) := \lim_{n\to\infty} \frac{1}{n}\sum_{p=1}^{n} I_p \tag{6.2}$$

exists almost surely.

LEMMA 6.1. *Almost surely, d is a ultrametric on $\mathbb{N}$. That is, almost surely the following hold*:

- *For all $i,j \in \mathbb{N}$, $d(i,j) \ge 0$, and $d(i,j) = 0$ if and only if $i = j$.*
- *For all $i,j \in \mathbb{N}$, $d(i,j) = d(j,i)$.*

- *For all* $i, j, k \in \mathbb{N}$, $d(i,k) \le d(i,j) \vee d(j,k)$.

A fortiori, d *is almost surely a metric on* $\mathbb{N}$.

PROOF. We first show for fixed distinct $i, j \in \mathbb{N}$ that $d(i,j) > 0$ almost surely. By exchangeability, de Finetti's theorem and the strong law of large numbers, the event $\{d(i,j) = 0\}$ coincides almost surely with the event $\{I_p = 0\ \forall p \notin \{i,j\}\} = \{\nexists p \notin \{i,j\} : \langle i,j\rangle \le p\}$. For $i, j \in [n+1]$, the event $\{I_p = 0\ \forall p \in [n+1], p \notin \{i,j\}\} = \{\nexists p \in [n+1] \setminus \{i,j\} : \langle i,j\rangle \le p\}$ is the event that in the representation of $\tilde{T}_n^\infty$ as a subset of $\{0,1\}^*$ labeled by $[n+1]$, there is an interior vertex $u_1 \cdots u_\ell$ such that i labels $u_1 \cdots u_\ell 0$ and j labels $u_1 \cdots u_\ell 1$ or *vice versa* (i.e., the two leaves of $\tilde{T}_n^\infty$ labeled by i and j are siblings and form what is often called a "cherry"). Now, the number of cherries in $\tilde{T}_n^\infty$ is at most $\lfloor \frac{n+1}{2} \rfloor$, and so the probability that i and j label the leaves of the same cherry is at most $2\lfloor \frac{n+1}{2} \rfloor \frac{1}{n+1}\frac{1}{n}$. Thus,

$$\mathbb{P}\{d(i,j) = 0\} = \lim_{n\to\infty} \mathbb{P}\{I_p = 0\ \forall p \in [n+1], p \notin \{i,j\}\} = 0.$$

It is clear that almost surely $d(i,i) = 0$ and $d(i,j) = d(j,i)$.

Lastly, for $i, j, k \in \mathbb{N}$ we have that $\langle i,j\rangle = \langle j,k\rangle \le \langle k,i\rangle$ or one of the two other similar inequalities obtained by cyclically permuting i, j, k holds. Therefore, $d(k,i) \le d(i,j) = d(j,k)$ almost surely or one of the two other similar inequalities obtained by cyclically permuting i, j, k holds. $\square$

For $t \in \mathbb{R}_+$ define an equivalence relation $\sim_t$ on $\mathbb{N}$ by declaring that $i \sim_t j$ if and only if $d(i,j) \le t$. Note that we can identify $\mathbb{N}$ with the equivalence classes of $\sim_0$. We now extend the metric d to a metric on the set $\mathbf{U}^o$ of pairs of the form (B,t), where $t \in \mathbb{R}_+$ and B is an equivalence class of $\sim_t$. Given an equivalence class A of $\sim_s$ and an equivalence class B of $\sim_t$, set

$$H\big((A,s),(B,t)\big) := \inf\{u \ge s \vee t : k \sim_u \ell\ \forall k \in A, \ell \in B\}$$

and

$$d\big((A,s),(B,t)\big) := \frac{1}{2}\big(\big[H\big((A,s),(B,t)\big) - s\big] + \big[H\big((A,s),(B,t)\big) - t\big]\big).$$

For $i, j \in \mathbb{N}$, we have $H((\{i\},0),(\{j\},0)) = d(i,j)$ and so $d((\{i\},0),(\{j\},0)) = d(i,j)$, confirming that we have an extension of the original definition of d. It is straightforward to check that this extension of d is a metric on $\mathbf{U}^o$ that satisfies the *four-point condition*; that is, for 4 elements $w, x, y, z \in \mathbf{U}^0$ at least one of the following conditions holds:

- $d(w,x) + d(y,z) \le d(w,y) + d(x,z) = d(w,z) + d(x,y)$,
- $d(w,z) + d(x,y) \le d(w,x) + d(y,z) = d(w,y) + d(x,z)$,
- $d(w,y) + d(x,z) \le d(w,z) + d(x,y) = d(w,x) + d(y,z)$.

It is, moreover, not difficult to show that the metric space $(\mathbf{U}^o, d)$ is connected and hence it is an $\mathbb{R}$-tree (see [13], Example 3.41, for more details). The completion $(\mathbf{U}, d)$ of $(\mathbf{U}^o, d)$ is also an $\mathbb{R}$-tree that is complete and separable.

There is a natural partial order on the $\mathbb{R}$-tree $(\mathbf{U}^o, d)$ defined by the requirement that the pair (A, s) precedes the pair (B, t) if $A \supseteq B$ and $s > t$. If we consider the subtree of $(\mathbf{U}, d)$ [equivalently, of $(\mathbf{U}^o, d)$] spanned by the set $\{(\{i\}, 0) : i \in [n+1]\}$, then combinatorially we have a leaf-labeled tree. The vertices of this combinatorial tree correspond to pairs of the form $(B_{ij}, d(i, j))$, $i, j \in [n+1]$, where B_{ij} is the equivalence class $\{k \in \mathbb{N} : d(i,k) \le d(i, j)\} = \{k \in \mathbb{N} : d(j,k) \le d(i, j)\}$. Moreover, the combinatorial tree inherits the partial order from $(\mathbf{U}^o, d)$ and the vertex $(B_{ij}, d(i, j))$ is the most recent common ancestor of the leaves $(\{i\}, 0)$ and $(\{j\}, 0)$ in this partial order.

We claim that this leaf-labeled tree with its partial order is isomorphic to $\tilde{T}_n^\infty$, with the vertex $(B_{ij}, d(i, j))$ corresponding to the vertex $[i, j]_n$ and, in particular, the leaf $(\{i\}, 0)$ corresponding to the leaf i. This is equivalent to showing the following.

LEMMA 6.2. *For distinct $i, j, k \in [n+1]$, $[i,k]_n = [j,k]_n < [i, j]_n$ if and only if $d(i,k) = d(j,k) > d(i, j)$.*

PROOF. It suffices to show that $[i,k]_n = [j,k]_n$ if and only if $d(i,k) = d(j,k)$ and $[j,k]_n < [i, j]_n$ if and only if $d(j,k) > d(i, j)$. Note that $[i,k]_n = [j,k]_n$ if and only if it is not the case that $[i,k]_n < [j,k]_n$ or $[i,k]_n > [j,k]_n$. Similarly, $d(i,k) = d(j,k)$ if and only if it is not the case that $d(i,k) > d(j,k)$ or $d(i,k) < d(j,k)$. It will thus further suffice to show for distinct $i, j, k \in [n+1]$ that $[j,k]_n < [i, j]_n$ if and only if $d(j,k) > d(i, j)$.

It is clear that if $d(j,k) > d(i, j)$, then $\langle j,k\rangle < \langle i, j\rangle$, and hence $[j,k]_n < [i, j]_n$. For the reverse implication, we certainly have that $[j,k]_n < [i, j]_n$ (and hence $\langle j,k\rangle < \langle i, j\rangle$) implies that $d(j,k) \ge d(i, j)$, and thus we only need to rule out the possibility of equality.

By exchangeability, de Finetti's theorem and the strong law of large numbers, the event $\{\langle j,k\rangle < \langle i, j\rangle, d(j,k) = d(i, j)\}$ coincides almost surely with the event

$$\{\langle j,k\rangle < \langle i, j\rangle\} \cap \{\nexists p \in \mathbb{N} \setminus \{k\} : \langle j,k\rangle \le p, \langle i, j\rangle \not\le p\}.$$

In order to show that the probability of the latter event is zero, it suffices to show that for $m \ge n$ the probability of the event

$$\{[j,k]_m < [i, j]_m\} \cap \{\nexists p \in [m+1] \setminus \{k\} : [j,k]_m \le p, [i, j]_m \not\le p\}$$

converges to zero as $m \to \infty$. In words, the last event occurs when the sibling of the most recent common ancestor in $\tilde{T}_m^\infty$ of the leaves labeled i and j is a leaf and that leaf is labeled by k. If we condition on T_m^∞ and the locations of the leaves labeled i and j, then the conditional probability of the last event is either $\frac{1}{m-1}$

or 0, depending on whether the sibling of the most recent common ancestor of the leaves labeled i and j is a leaf, and so the (unconditional) probability of the last event certainly converges to zero as $m \to \infty$. $\square$

Write $\mathbf{T}^o$ for the subtree of $\mathbf{U}^o$ (and hence of $\mathbf{U}$) spanned by the set $\{(\{i\}, 0) : i \in \mathbb{N}\}$ and let $\mathbf{T}$ be the closure of $\mathbf{T}^o$ in $\mathbf{U}$. We denote the restriction of the metric d to $\mathbf{T}$ also by d. From the above considerations, we infer immediately the following.

PROPOSITION 6.3. *There is an injective mapping from the set of equivalence classes $\langle i, j\rangle$, $i, j \in \mathbb{N}$, of the ergodic didendritic system $(\equiv, \langle\cdot,\cdot\rangle, <_L, <_R)$ into the complete, separable $\mathbb{R}$-tree $\mathbf{T}$ constructed above such that the distance $d(i, j)$ defined by (6.2) coincides with the distance in $\mathbf{T}$ between the images of equivalence classes $\langle i, i\rangle$ and $\langle j, j\rangle$.*

From now on we will, with a slight abuse of notation, think of the equivalence classes $\langle i, j\rangle$, $i, j \in \mathbb{N}$, (including the leaves $i = \langle i, i\rangle$, $i \in \mathbb{N}$) as being elements of the $\mathbb{R}$-tree $(\mathbf{T}, d)$.

REMARK 6.4. Consider two equivalence classes $\langle h, i\rangle$ and $\langle j, k\rangle$. Recall that the most recent common ancestor of $\langle h, i\rangle$ and $\langle j, k\rangle$ is of the form $\langle \ell, m\rangle$, where $\ell \in \{h, i\}$ and $m \in \{j, k\}$. In terms of the metric d, ℓ and m are any such pair for which $d(\ell, m) = d(h, j) \vee d(h, k) \vee d(i, j) \vee d(i, k)$. We therefore have

$$\begin{aligned} d\big(\langle h, i\rangle, \langle j, k\rangle\big) &= \frac{1}{2}\big(\big[d(\ell, m) - d(h, i)\big] + \big[d(\ell, m) - d(j, k)\big]\big) \\ &= d(h, j) \vee d(h, k) \vee d(i, j) \vee d(i, k) - \frac{1}{2}\big(d(h, i) + d(j, k)\big). \end{aligned}$$

In particular,

$$d\big(i, \langle i, j\rangle\big) = \frac{1}{2} d(i, j),$$

as we would expect.

REMARK 6.5. It follows from the construction of $\mathbf{T}$ that $\max\{d(x, y) : x, y \in \mathbf{T}\} \le 1$. For $n \in \mathbb{N}$, let ρ_n be the most recent common ancestor of $1, 2, \dots, n+1$ with respect to the partial order $<$. Note that $\rho_n = \langle i, j\rangle \in \mathbf{T}$ for distinct $i, j \in [n+1]$. The successive points $\rho_1, \rho_2, \dots$ are linearly ordered along a geodesic ray in $\mathbf{T}$. Because $\mathbf{T}$ is a complete separable $\mathbb{R}$-tree with a finite diameter, it follows that $(\rho_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, and hence convergent to a point $\rho \in \mathbf{T}$. We can, as with any rooted $\mathbb{R}$-tree, define a partial order on $\mathbf{T}$ by declaring that x precedes y if and only if $x \neq y$ and x belongs to the geodesic segment $[\rho, y]$ between ρ and y (equivalently, $[\rho, x] \subsetneq [\rho, y]$).

The following result is now immediate.

PROPOSITION 6.6. *The partial order on* $\mathbf{T}$ *defined by the root* ρ *extends the partial order* $<$ *on the equivalence classes* $\{\langle i, j\rangle : i, j \in \mathbb{N}\}$, *and the most recent common ancestor of* $\langle h, i\rangle$ *and* $\langle j, k\rangle$ *is the equivalence class* $\langle \ell, m\rangle$ *such that* $[\rho, \langle \ell, m\rangle] = [\rho, \langle h, i\rangle] \cap [\rho, \langle j, k\rangle]$.

EXAMPLE 6.7. Consider the infinite Rémy bridge in Example 4.1. A concrete realization of the $\mathbb{R}$-tree $(\mathbf{T}, d)$ can be constructed as follows. Let $(U_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables that each have the uniform distribution on $[0, 1]$. Take the interval $[0, \frac{1}{2}]$ and build an $\mathbb{R}$-tree by, for each $n \in \mathbb{N}$, attaching one end of a closed line segment of length $\frac{1}{2}U_n$ to the point $\frac{1}{2}U_n \in [0, \frac{1}{2}]$ and labeling the other end of the line segment with n. The distance between i and j in the resulting $\mathbb{R}$-tree is then

$$\left|\frac{1}{2}U_i - \frac{1}{2}U_j\right| + \frac{1}{2}U_i + \frac{1}{2}U_j = U_i \vee U_j.$$

For $i \neq j$, we can identify $\langle i, j\rangle$ with $\frac{1}{2}(U_i \vee U_j) \in [0, \frac{1}{2}]$. For $i \neq j$ and $k \neq \ell$, we have $\langle i, j\rangle < k$ if $U_i \vee U_j > U_k$ and $\langle i, j\rangle < \langle k, \ell\rangle$ if $U_i \vee U_j > U_k \vee U_\ell$. Note that, as required, the distance between i and j is

$$U_i \vee U_j = \lim_{n\to\infty} \frac{1}{n} \sum_{p=1}^{n} \mathbb{1}\{U_i \vee U_j \geq U_p\} = \lim_{n\to\infty} \frac{1}{n} \sum_{p=1}^{n} \mathbb{1}\{\langle i, j\rangle \leq p\}.$$

The root ρ is the point $\frac{1}{2}$ in the interval $[0, \frac{1}{2}]$.

EXAMPLE 6.8. Consider the infinite Rémy bridge in Example 4.2. A concrete realization of the $\mathbb{R}$-tree $(\mathbf{T}, d)$ can be constructed as follows. Take the complete binary tree $\{0, 1\}^*$ and join two elements of the form $v_1 \cdots v_k$ and $v_1 \cdots v_k v_{k+1}$ with a segment of length $\frac{1}{2^{k+2}}$. This gives an $\mathbb{R}$-tree such that if $u_1 \cdots u_m$ and $v_1 \cdots v_n$ are elements of $\{0, 1\}^*$ for which $p = \max\{j : u_j = v_j\}$, then the distance between the corresponding points in the $\mathbb{R}$-tree is

$$\left(\frac{1}{2^{p+2}} + \frac{1}{2^{p+3}} + \cdots + \frac{1}{2^{m+1}}\right) + \left(\frac{1}{2^{p+2}} + \frac{1}{2^{p+3}} + \cdots + \frac{1}{2^{n+1}}\right).$$

We can identify $(\mathbf{T}, d)$ with the completion of this $\mathbb{R}$-tree. There is a bijective correspondence between $\{0, 1\}^\infty$ and the points "added" in passing to the completion. The distance between the points in the completion corresponding to $u_1 u_2 \cdots$ and $v_1 v_2 \cdots$ in $\{0, 1\}^\infty$ with $p = \max\{j : u_j = v_j\}$ is

$$\left(\frac{1}{2^{p+2}} + \frac{1}{2^{p+3}} + \cdots\right) + \left(\frac{1}{2^{p+2}} + \frac{1}{2^{p+3}} + \cdots\right) = \frac{1}{2^p}.$$

7. The sampling measure on the real tree. Throughout this section, let $(\mathbf{T}, d)$ be the $\mathbb{R}$-tree constructed in Section 6 from an ergodic exchangeable random didendritic system $\mathbf{D} = (\equiv, \langle\cdot,\cdot\rangle, <_L, <_R)$ [equivalently, from the labeled version $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ of an extremal infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$]. Recall from Proposition 6.6 that we can extend the partial order $<$ to all of $\mathbf{T}$.

DEFINITION 7.1. Suppose that $\mathbf{V}$ is a complete separable $\mathbb{R}$-tree with finite diameter. A *leaf* of $\mathbf{V}$ is a point $x \in \mathbf{V}$ such that there do not exist two points $y, z \in \mathbf{V}$ for which x is in the interior of the segment between y and z. The $\mathbb{R}$-tree $\mathbf{V}$ is spanned by its set of leaves.

An *isolated leaf* of a complete separable $\mathbb{R}$-tree $\mathbf{V}$ is a leaf $x \in \mathbf{V}$ such that for some ε the open ball of radius ε centered at x is a half-open line segment with x at the closed end of the segment. There is a maximal such ε and we write $[x, \Pi(x))$ for the corresponding half-open line segment. For a leaf x that is not isolated, we set $\Pi(x) := x$.

The *core* of $\mathbf{V}$ is the subtree $\Gamma(\mathbf{V})$ spanned by the set of points of the form $\Pi(x)$ as x ranges over the leaves of $\mathbf{V}$. It is not hard to show that $\Gamma(\mathbf{V})$ is a closed $\mathbb{R}$-tree and that $\Pi(x)$ is the unique point of $\Gamma(\mathbf{V})$ that is closest to the leaf x and so we think of $\Pi(x)$ as the *point of attachment* of x to the core. Also, if for a leaf $x \in \mathbf{V}$ we let $\mathbf{V}^x$ be the closure of the subtree of $\mathbf{V}$ spanned by the leaves of $\mathbf{V}$ other than x, then $\Gamma(\mathbf{V}) = \bigcap_x \mathbf{V}^x$.

LEMMA 7.2. (a) *The core of* $\mathbf{T}$ *is the closure of the subtree spanned by the set* $\{\langle i, j\rangle : i, j \in \mathbb{N}, i \neq j\}$.

(b) *For all* $i \in \mathbb{N}$,

$$\begin{aligned} d\bigl(i, \Pi(i)\bigr) &= \inf\bigl\{d\bigl(i, \langle i, j\rangle\bigr) : j \in \mathbb{N}, j \neq i\bigr\} \\ &= \inf\bigl\{d\bigl(j, \langle i, j\rangle\bigr) : j \in \mathbb{N}, j \neq i\bigr\} \\ &= \frac{1}{2}\inf\bigl\{d(i, j) : j \in \mathbb{N}, j \neq i\bigr\} \end{aligned}$$

and if $(j_n)_{n\in\mathbb{N}}$ *is any sequence in* $\mathbb{N} \setminus \{i\}$ *such that* $\lim_{n\to\infty} d(i, \langle i, j_n\rangle) = d(i, \Pi(i))$, *then* $\Pi(i) = \lim_{n\to\infty}\langle i, j_n\rangle$.

(c) *For* $i \in \mathbb{N}$, $\Pi(i) \le i$.

(d) *For* $i, j \in \mathbb{N}$ *with* $i \neq j$, $\Pi(i) \neq \Pi(j)$.

(e) *For* $i, j \in \mathbb{N}$ *with* $i \neq j$, *the most recent common ancestor of* $\Pi(i)$ *and* $\Pi(j)$ *in the partial order that the core* $\Gamma(\mathbf{T})$ *inherits from* $\mathbf{T}$ *is* $\langle i, j\rangle$ *and*

$$d(i, j) = \lim_{n\to\infty} \frac{1}{n} \sum_{p=1}^{n} \mathbb{1}\bigl\{\langle i, j\rangle \le \Pi(p)\bigr\}.$$

(f) *Under our standing ergodicity assumption, the isometry class of* $\Gamma(\mathbf{T})$ *together with the partial order on* $\Gamma(\mathbf{T})$ *inherited from the partial order* $<$ *is constant almost surely.*

PROOF. Parts (a), (b) and (c) are straightforward and are left to the reader. For part (d), suppose that $\Pi(i) = \Pi(j)$ for $i \neq j$. By part (c), $\Pi(i) = \Pi(j) \leq \langle i, j \rangle$. Thus, by part (b), $\Pi(i) = \Pi(j) = \langle i, j \rangle$ and $d(i, \Pi(i)) = d(i, \Pi(j)) = d(j, \Pi(i)) = d(j, \Pi(j)) = d(i, \langle i, j \rangle) = d(j, \langle i, j \rangle) = \frac{1}{2} d(i, j)$. This is not possible unless i and j are both isolated. By the definition of $d(i, j)$, there are infinitely many $p \in \mathbb{N} \setminus \{i, j\}$ such that $\langle i, j \rangle \leq p$. For any such p we must have either $\langle i, j \rangle < \langle i, p \rangle$ or $\langle i, j \rangle < \langle j, p \rangle$, so that $d(i, \langle i, p \rangle) < d(i, \langle i, j \rangle)$ or $d(j, \langle j, p \rangle) < d(j, \langle i, j \rangle)$, but this contradicts $\Pi(i) = \Pi(j) = \langle i, j \rangle$.

Part (e) is also clear and is left to the reader.

For part (f), note first of all that if σ is a permutation of $\mathbb{N}$ such that $\sigma(i) = i$ for all but finitely many $i \in \mathbb{N}$ and $(\equiv^\sigma, \langle \cdot, \cdot \rangle^\sigma, <_L^\sigma, <_R^\sigma)$ is the random didendritic system defined in Definition 5.11, then the isometry class of $\Gamma(\mathbf{T})$ as a random complete separable metric space is unchanged if we replace $(\equiv, \langle \cdot, \cdot \rangle, <_L, <_R)$ by $(\equiv^\sigma, \langle \cdot, \cdot \rangle^\sigma, <_L^\sigma, <_R^\sigma)$. Our standing ergodicity assumption gives that the isometry class of $\Gamma(\mathbf{T})$ is constant almost surely.

The root ρ defined in Remark 6.5 is an element of $\Gamma(\mathbf{T})$. It is clear that the location of ρ is unchanged if we replace $\mathbf{D} = (\equiv, \langle \cdot, \cdot \rangle, <_L, <_R)$ by $\mathbf{D}^\sigma = (\equiv^\sigma, \langle \cdot, \cdot \rangle^\sigma, <_L^\sigma, <_R^\sigma)$, and so the restriction of the random partial order $<$ to $\Gamma(\mathbf{T})$ is also constant. $\square$

EXAMPLE 7.3. Consider the $\mathbb{R}$-tree $\mathbf{T}$ constructed in Example 6.7 from the infinite Rémy bridge introduced in Example 4.1. The core of $\mathbf{T}$ is the interval $[0, \frac{1}{2}]$.

Consider the maps $\kappa_- : \mathbb{N} \to \mathbb{N}$ and $\kappa_+ : \mathbb{N} \to \mathbb{N}$ given by $\kappa_-(n) = 2n - 1$ and $\kappa_+(n) = 2n$, $n \in \mathbb{N}$. Define the exchangeable random didendritic systems $\mathbf{D}_- = (\equiv_-, \langle \cdot, \cdot \rangle_-, <_{-,L}, <_{-,R})$ and $\mathbf{D}_+ = (\equiv_+, \langle \cdot, \cdot \rangle_+, <_{+,L}, <_{+,R})$ by

$$(h, i) \equiv_- (j, k) \quad \Longleftrightarrow \quad \bigl(\kappa_-(h), \kappa_-(i)\bigr) \equiv \bigl(\kappa_-(j), \kappa_-(k)\bigr)$$

and

$$(h, i) \equiv_+ (j, k) \quad \Longleftrightarrow \quad \bigl(\kappa_+(h), \kappa_+(i)\bigr) \equiv \bigl(\kappa_+(j), \kappa_+(k)\bigr),$$

$$\langle i, j \rangle_- \text{ is the } \equiv_- \text{ equivalence class of } (i, j)$$

and

$$\langle i, j \rangle_+ \text{ is the } \equiv_+ \text{ equivalence class of } (i, j),$$

$$\langle h, i \rangle_- <_{-,L} \langle j, k \rangle_- \quad \Longleftrightarrow \quad \bigl\langle \kappa_-(h), \kappa_-(i) \bigr\rangle <_L \bigl\langle \kappa_-(j), \kappa_-(k) \bigr\rangle$$

and

$$\langle h, i \rangle_+ <_{+,L} \langle j, k \rangle_+ \quad \Longleftrightarrow \quad \bigl\langle \kappa_+(h), \kappa_+(i) \bigr\rangle <_L \bigl\langle \kappa_+(j), \kappa_+(k) \bigr\rangle,$$

and

$$\langle h, i \rangle_- <_{-,R} \langle j, k \rangle_- \quad \Longleftrightarrow \quad \bigl\langle \kappa_-(h), \kappa_-(i) \bigr\rangle <_R \bigl\langle \kappa_-(j), \kappa_-(k) \bigr\rangle$$

and

$$\langle h,i\rangle_+ <_{+,R} \langle j,k\rangle_+ \iff \langle \kappa_+(h),\kappa_+(i)\rangle <_R \langle \kappa_+(j),\kappa_+(k)\rangle.$$

Define the partial orders $<_-$ and $<_+$ on $\{\langle i,j\rangle_+ : i,j\in\mathbb{N}\}$, respectively, by declaring that

$$\langle h,i\rangle_- <_- \langle j,k\rangle_- \iff \langle \kappa_-(h)\kappa_-(i)\rangle < \langle \kappa_-(j),\kappa_-(k)\rangle$$

and

$$\langle h,i\rangle_+ <_- \langle j,k\rangle_+ \implies \langle \kappa_+(h)\kappa_+(i)\rangle < \langle \kappa_+(j),\kappa_+(k)\rangle,$$

or, equivalently,

$$\langle h,i\rangle_- <_- \langle j,k\rangle_- \iff \langle h,i\rangle_- <_{-,L} \langle j,k\rangle_- \quad\text{or}\quad \langle h,i\rangle_- <_{-,R} \langle j,k\rangle_-$$

and

$$\langle h,i\rangle_+ <_+ \langle j,k\rangle_+ \iff \langle h,i\rangle_+ <_{+,L} \langle j,k\rangle_+ \quad\text{or}\quad \langle h,i\rangle_+ <_{+,R} \langle j,k\rangle_+.$$

By exchangeability, the random didendritic systems $\mathbf{D}_-$, $\mathbf{D}_+$ and $\mathbf{D}$ have the same distribution. By the ergodicity of $\mathbf{D}$, the random didendritic systems $\mathbf{D}_-$ and $\mathbf{D}_+$ are independent. Construct random partially ordered $\mathbb{R}$-trees $(\mathbf{T}^-, <_-)$ and $(\mathbf{T}^+, <_+)$ from $\mathbf{D}_-$ and $\mathbf{D}_+$ in the same manner that $(\mathbf{T}, <)$ was constructed from $\mathbf{D}$.

By de Finetti's theorem and the strong law of large numbers,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\sum_{p=1}^{n}\mathbb{1}\{\langle i,j\rangle \le p\} &= \lim_{n\to\infty}\frac{1}{n}\sum_{p=1}^{n}\mathbb{1}\{\langle i,j\rangle \le 2p-1\}\\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{p=1}^{n}\mathbb{1}\{\langle i,j\rangle \le 2p\}\end{aligned}$$

for any $i,j\in\mathbb{N}$. Therefore, the distance between k and ℓ in $\mathbf{T}^-$ (resp. $\mathbf{T}^+$) is the same as the distance between $\kappa_-(k)$ and $\kappa_-(\ell)$ [resp. $\kappa_+(k)$ and $\kappa_+(\ell)$] in $\mathbf{T}$, and hence we may (and will) identify $\mathbf{T}^-$ and $\mathbf{T}^+$ with the closures in $\mathbf{T}$ of the respective sets $2\mathbb{N}-1$ and $2\mathbb{N}$.

The set $\Gamma(\mathbf{T})$ is the closure of the subtree spanned by the set of attachment points $\{\Pi(i) : i\in\mathbb{N}\}$ and, by part (a) of Lemma 7.2, also the closure of the subtree spanned by the set of points $\{\langle j,k\rangle : j,k\in\mathbb{N},\ j\neq k\}$. It is clear that $\Gamma(\mathbf{T}^-)\subseteq\Gamma(\mathbf{T})$ and $\Gamma(\mathbf{T}^+)\subseteq\Gamma(\mathbf{T})$. It follows from Remark 11.2 and the second proof of Proposition 3.16 in [17] that almost surely for any $\delta>0$ and $i\in\mathbb{N}$, there exists $j,k\in 2\mathbb{N}-1$ (resp. $j,k\in 2\mathbb{N}$) with $d(\Pi(i),\langle j,k\rangle)<\delta$ and hence $\Gamma(\mathbf{T}^-)=\Gamma(\mathbf{T}^+)=\Gamma(\mathbf{T})$.

For the benefit of the reader, we sketch the argument from [17] in our notation. Fix $\epsilon>0$ and a deterministic sequence $0<h_1^{(\epsilon)}<h_2^{(\epsilon)}<\dots\uparrow\infty$ such that $h_1^{(\epsilon)}<\epsilon$, $h_{n+1}^{(\epsilon)}-h_n^{(\epsilon)}<\epsilon$, and $\mathbb{P}\{d(i,\Pi(i))=h_n^{(\epsilon)}\}=0$ for all $i,n\in\mathbb{N}$. Set $h_0^{(\epsilon)}=0$.

Define an exchangeable equivalence relation $\sim^\epsilon$ on $\mathbb{N}$ by declaring that $i \sim^\epsilon j$, $i, j \in \mathbb{N}$, $i \neq j$, if $d(i, \Pi(i)), d(j, \Pi(j)), d(i, \langle i, j\rangle) = d(j, \langle i, j\rangle) \in [h_{n-1}^{(\epsilon)}, h_n^{(\epsilon)})$ for some $n \in \mathbb{N}$. Note that if $i \sim^\epsilon j$, $i, j \in \mathbb{N}$, $i \neq j$, then

$$d\big(\Pi(i), \langle i, j\rangle\big) \vee d\big(\Pi(j), \langle i, j\rangle\big) < \epsilon.$$

The crucial observation, stated and proved as (11.4) of [17], is that almost surely the exchangeable equivalence relation $\sim^\epsilon$ does not have any singleton equivalence classes. It then follows from Kingman's paintbox construction of exchangeable equivalence relations that almost surely for any $i \in \mathbb{N}$ there exists $j, k \in 2\mathbb{N} - 1$ (respectively, $j, k \in 2\mathbb{N}$) with i, j, k distinct and $i \sim^\epsilon j \sim^\epsilon k$, and hence

$$d\big(\Pi(i), \langle j, k\rangle\big) \le d\big(\Pi(i), \langle i, j\rangle\big) + d\big(\langle i, j\rangle, \Pi(j)\big) + d\big(\Pi(j), \langle j, k\rangle\big) < 3\epsilon.$$

Let Π_- and Π_+ be the analogues of Π for $(\mathbf{T}^-, <_-)$ and $(\mathbf{T}^+, <_+)$. For $i \in \mathbb{N}$, we have that $\Pi_+(i)$, the closest point in $\Gamma(\mathbf{T}^+)$ to i (where we stress that i labels an element of $\mathbf{T}^+$), is an element of $\Gamma(\mathbf{T}^-) = \Gamma(\mathbf{T}^+)$. It follows from the exchangeability inherent in our construction that $(\Pi_+(i))_{i\in\mathbb{N}}$ is an exchangeable sequence of random elements of $\Gamma(\mathbf{T}^-)$. By our standing ergodicity assumption and de Finetti's theorem, the random elements in this sequence are independent and identically distributed, and it is a consequence of part (d) of Lemma 7.2 that their common distribution is, prefiguring the notation in the statement of Proposition 7.4 below, a diffuse probability measure μ on the $\mathbb{R}$-tree $\mathbf{S} := \Gamma(\mathbf{T})$ that is contained in $\mathbf{T}$ and rooted in $\theta := \rho$. The probability measure μ and the $\mathbb{R}$-tree $\mathbf{S}$ are the objects addressed in this section's title.

For $i, j \in \mathbb{N}$ with $i \neq j$, part (e) of Lemma 7.2 gives that $\langle i, j\rangle_+$ is the most recent common ancestor of $\Pi_+(i)$ and $\Pi_+(j)$ in the partial order $<_+$. Moreover, $\langle i, j\rangle_+ \le_+ \Pi_+(p)$ if and only if $[\rho, \Pi_+(i)] \cap [\rho, \Pi_+(j)] \subseteq [\rho, \Pi_+(i)] \cap [\rho, \Pi_+(p)]$ or $[\rho, \Pi_+(i)] \cap [\rho, \Pi_+(j)] \subseteq [\rho, \Pi_+(j)] \cap [\rho, \Pi_+(p)]$, where $[\rho, x]$ is the geodesic segment between ρ and x in $\Gamma(\mathbf{T}) = \Gamma(\mathbf{T}^-) = \Gamma(\mathbf{T}^+)$, and so if we write $d_+(i, j)$ for the distance between $i, j \in \mathbb{N}$, $i \neq j$, in $\mathbf{T}^+$, we have from part (e) of Lemma 7.2 that

$$\begin{aligned} d_+(i, j) = \lim_{n\to\infty} \frac{1}{n} \sum_{p=1}^{n} \mathbb{1}\big(&\big\{[\rho, \Pi_+(i)] \cap [\rho, \Pi_+(j)] \subseteq [\rho, \Pi_+(i)] \cap [\rho, \Pi_+(p)]\big\} \\ &\cup \big\{[\rho, \Pi_+(i)] \cap [\rho, \Pi_+(j)] \subseteq [\rho, \Pi_+(j)] \cap [\rho, \Pi_+(p)]\big\}\big). \end{aligned}$$

Because $(\mathbf{T}^+, <_+)$ has the same distribution as $(\mathbf{T}, <)$, we have established the following result.

PROPOSITION 7.4. *Suppose that* $\equiv$, $\langle\cdot,\cdot\rangle$ *and* $<$ *are the equivalence relation on* $\mathbb{N} \times \mathbb{N}$, *the equivalence classes and the partial order on those equivalence classes arising from an ergodic exchangeable random didendritic system* (*equivalently, from the labeled version of an extremal infinite Rémy bridge*). *There is a* *complete separable* $\mathbb{R}$*-tree* $\mathbf{S}$, *a point* $\theta \in \mathbf{S}$, *and a diffuse probability measure* μ *on* $\mathbf{S}$ *such that the following hold. Let* $\xi_1, \xi_2, \dots$ *be i.i.d. random elements of* $\mathbf{S}$ *with common distribution* μ. *Define a random equivalence relation* $\equiv_\#$ *on* $\mathbb{N} \times \mathbb{N}$ *by declaring that* $(i, i) \equiv_\# (k, \ell)$ *if and only if* $(i, i) = (k, \ell)$, *and* $(i, j) \equiv_\# (k, \ell)$ *for* $i \neq j$ *and* $k \neq \ell$ *if and only if* $[\theta, \xi_i] \cap [\theta, \xi_j] = [\theta, \xi_k] \cap [\theta, \xi_\ell]$, *where* $[\theta, x]$ *is the geodesic segment between* θ *and* x, *in* $\mathbf{S}$. *Denote the equivalence class containing* $(i, j) \in \mathbb{N} \times \mathbb{N}$ *by* $\langle i, j\rangle_\#$. *Define a partial order* $<_\#$ *on the set of equivalence classes by declaring that* $\langle i, j\rangle_\# <_\# \langle i, i\rangle_\#$ *for all* $i \neq j$ *and that* $\langle i, j\rangle_\# <_\# \langle k, \ell\rangle_\#$ *for* $i \neq j$ *and* $k \neq \ell$ *if* $[\theta, \xi_i] \cap [\theta, \xi_j] \subsetneq [\theta, \xi_k] \cap [\theta, \xi_\ell]$. *The object* $(\equiv_\#, \langle \cdot, \cdot\rangle_\#, <_\#)$ *has the same distribution as* $(\equiv, \langle \cdot, \cdot\rangle, <)$.

8. Distinguishing between left and right. Throughout this section, let $(\mathbf{T}, d)$ be the $\mathbb{R}$-tree constructed in Section 6 from an ergodic exchangeable random didendritic system $\mathbf{D} = (\equiv, \langle \cdot, \cdot\rangle, <_L, <_R)$ [equivalently, from the labeled version $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ of an extremal infinite Rémy bridge $(T_n^\infty)_{n\in\mathbb{N}}$].

Let $\mathbf{S}$, θ and μ be the objects described in Proposition 7.4. Thus, $\mathbf{S}$ is a complete separable $\mathbb{R}$-tree, θ is an element of $\mathbf{S}$, and μ is a diffuse probability measure on $\mathbf{S}$. Further, let $\xi_1, \xi_2, \dots$ be i.i.d. random elements of $\mathbf{S}$ with common distribution μ. We may suppose that $(i, i) \equiv (k, \ell)$ if and only if $(i, i) = (k, \ell)$ and that $(i, j) \equiv (k, \ell)$ for $i \neq j$ and $k \neq \ell$ if and only if $[\theta, \xi_i] \cap [\theta, \xi_j] = [\theta, \xi_k] \cap [\theta, \xi_\ell]$, where we recall that $[\theta, x]$ is the geodesic segment between θ and x in $\mathbf{S}$.

Recall from Lemma 5.17 that if we know the equivalence relation $\equiv$ and the partial order $<$ of the didendritic system $\mathbf{D}$, then the partial orders $<_L$ and $<_R$ (and hence the didendritic system) is uniquely determined by the specification for all distinct $i, j \in \mathbb{N}$ whether $\langle i, j\rangle <_L i$ and $\langle i, j\rangle <_R j$ or $\langle i, j\rangle <_R i$ and $\langle i, j\rangle <_L j$.

Put

$$J_{ij} := \mathbb{1}\{\langle i, j\rangle <_L i, \langle i, j\rangle <_R j\}$$

for $(i, j) \in \mathbb{N} \times \mathbb{N} \setminus \delta$, where $\delta := \{(k, k) : k \in \mathbb{N}\}$. Note for all $(i, j) \in \mathbb{N} \times \mathbb{N} \setminus \delta$ that $J_{ij} = 1$ if and only if $J_{ji} = 0$.

It follows from the exchangeability and ergodicity of $\mathbf{D}$ that the random array J is jointly exchangeable and ergodic and, indeed, the random array $(\xi_i, \xi_j, J_{ij})_{(i,j)\in\mathbb{N}\times\mathbb{N}\setminus\delta}$ is also jointly exchangeable and ergodic. Therefore, by the Aldous–Hoover–Kallenberg theory of such random arrays, we may suppose that on some extension of our underlying probability space there exist i.i.d. random variables $(U_i)_{i\in\mathbb{N}}$, and $(U_{ij})_{i,j\in\mathbb{N},i<j}$ that are uniform on $[0, 1]$ and a function $F : (\mathbf{S} \times [0, 1])^2 \times [0, 1] \to \{0, 1\}$ such that

$$J_{ij} = F(\xi_i, U_i, \xi_j, U_j, U_{ij}),$$

where $U_{ij} = U_{ji}$ for $i > j$ (here $<$ is the usual order on $\mathbb{N}$) (see [19], Theorem 7.22, Lemma 7.35). Because $J_{ij} = 1 - J_{ji}$, the function F has the property $F(y, v, x, u, w) = 1 - F(x, u, y, v, w)$.

If $i, j, k \in \mathbb{N}$ are distinct, we have $J_{ij} = J_{ik}$ on the event $\{\langle i, j\rangle = \langle i, k\rangle < \langle j, k\rangle\} = \{[\theta, \xi_i] \cap [\theta, \xi_j] = [\theta, \xi_i] \cap [\theta, \xi_k] \subsetneq [\theta, \xi_j] \cap [\theta, \xi_k]\}$. That is, $F(\xi_i, U_i, \xi_j, U_j, U_{ij}) = J_{ij} = J_{ik} = F(\xi_i, U_i, \xi_k, U_j, U_{ik}) \in \{0, 1\}$ on the latter event. Similarly, $1 - F(\xi_i, U_i, \xi_j, U_j, U_{ij}) = J_{ji} = J_{jk} = F(\xi_j, U_j, \xi_k, U_k, U_{jk}) \in \{0, 1\}$ on the event $\{\langle j, i\rangle = \langle j, k\rangle < \langle i, k\rangle\} = \{[\theta, \xi_j] \cap [\theta, \xi_i] = [\theta, \xi_j] \cap [\theta, \xi_k] \subsetneq [\theta, \xi_i] \cap [\theta, \xi_k]\}$.

By Lemma 6.1,

$$\mathbb{P}\{\langle k, i\rangle = \langle k, j\rangle < \langle i, j\rangle \ \forall k \notin \{i, j\}\} = 0,$$

and Lemma 8.1 below then gives that

$$J_{ij} = W(\xi_i, U_i, \xi_j, U_j) \tag{8.1}$$

almost surely for some Borel function $W : (\mathbf{S} \times [0, 1])^2 \to \{0, 1\}$.

The intuition for (8.1) being true is firstly that if we had $\langle k, i\rangle = \langle k, j\rangle < \langle i, j\rangle$ for all $k \notin \{i, j\}$, so that $\{i, j\}$ is a cherry, then there could be a need to use the randomization provided by U_{ij} to build J_{ij}; that is, $F(\xi_i, U_i, \xi_j, U_j, U_{ij})$ could depend on U_{ij}. However, cherries do not occur with positive probability because of Lemma 6.1. Moreover, on the event where $\langle i, j\rangle = \langle i, k\rangle < \langle j, k\rangle$ for at least one, and hence infinitely many, $k \notin \{i, j\}$, it follows that $F(\xi_i, U_i, \xi_j, U_j, U_{ij}) = F(\xi_i, U_i, \xi_k, U_k, U_{ik})$ for all such k, which cannot happen if $(x, u, y, v, w) \mapsto F(x, u, y, v, w)$ has a genuine functional dependence on (v, w), and hence only the extra randomization provided by U_i (rather than that provided by U_j and U_{ij}) might be needed to build J_{ij}. Similarly, on the event where $\langle i, j\rangle = \langle j, k\rangle < \langle i, k\rangle$ for at least one, and hence infinitely many, $k \notin \{i, j\}$, it follows that $F(\xi_i, U_i, \xi_j, U_j, U_{ij}) = F(\xi_k, U_k, \xi_j, U_j, U_{kj})$ for all such k and only the extra randomization provided by U_j (rather than that provided by U_i and U_{ij}) might be needed to build J_{ij}. Lastly, on the event where $\langle i, k\rangle > \langle i, j\rangle = \langle k, \ell\rangle = \langle i, \ell\rangle = \langle k, j\rangle < \langle j, \ell\rangle$ for at least one, and hence infinitely many, pairs $k, \ell \notin \{i, j\}$, it follows that $F(\xi_i, U_i, \xi_j, U_j, U_{ij}) = F(\xi_k, U_k, \xi_\ell, U_\ell, U_{k,\ell})$ for all such pairs k, ℓ and there is no need for the extra randomization provided by U_i, U_j or U_{ij} to build J_{ij}.

We now give the promised formal argument for (8.1).

LEMMA 8.1. *Consider on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ independent random elements X_1, X_2, X_3 of some Borel space $(D, \mathcal{D})$ and Y_{12}, Y_{13}, Y_{23} of some Borel space $(E, \mathcal{E})$. Suppose that X_1, X_2, X_3 have the same diffuse probability distribution α and that Y_{12}, Y_{13}, Y_{23} have the same diffuse probability distribution β. Write B for the subset of D^3 that consists of triplets with distinct entries. Given an ordered listing i, j, k of $\{1, 2, 3\}$ and a set $C \subseteq B$, put $C_{ijk} := \{(x_1, x_2, x_3) \in B : (x_i, x_j, x_k) \in C\}$. Suppose that there is a set $A \in \mathcal{D}^3$ such*:

- $A_{123} = A_{132}$,
- $A_{213} = A_{231}$,

- *$A_{312} = A_{321}$,*
- *these 3 sets are pairwise disjoint and their union is B.*

Suppose further that

$$\begin{aligned}
\alpha^{\otimes 2}&\{(x_1, x_2) \in D^2 : \alpha\{x_3 \in D : (x_1, x_2, x_3) \in B \setminus A_{312}\} = 0\} \\
&= \alpha^{\otimes 2}\{(x_1, x_2) \in D^2 : \alpha\{x_3 \in D : (x_1, x_2, x_3) \in B \setminus A_{321}\} = 0\} \\
&= 0.
\end{aligned}$$

Consider a Borel function $H : D^2 \times E \to \{0, 1\}$ such that:

- *$H(X_1, X_2, Y_{12}) = H(X_1, X_3, Y_{13})$ on the event $\{(X_1, X_2, X_3) \in A\} = \{(X_1, X_3, X_2) \in A\}$,*
- *$H(X_1, X_2, Y_{12}) = 1 - H(X_2, X_3, Y_{23})$ on the event $\{(X_2, X_1, X_3) \in A\} = \{(X_2, X_3, X_1) \in A\}$.*

Then there exists a Borel function $K : D^2 \to \{0, 1\}$ such that $H(X_1, X_2, Y_{12}) = K(X_1, X_2)$ almost surely.

PROOF. For $(x_1, x_2, y_{12}) \in D^2 \times E$ with $x_1 \neq x_2$, we have

$$\begin{aligned}
\int_{D\times E} & H(x_1, x_2, y_{12}) \mathbb{1}_{A_{123}}(x_1, x_2, x_3) \alpha \otimes \beta(d(x_3, y_{13})) \\
&= \int_{D\times E} H(x_1, x_3, y_{13}) \mathbb{1}_{A_{123}}(x_1, x_2, x_3) \alpha \otimes \beta(d(x_3, y_{13}))
\end{aligned}$$

and

$$\begin{aligned}
\int_{D\times E} & H(x_1, x_2, y_{12}) \mathbb{1}_{A_{213}}(x_1, x_2, x_3) \alpha \otimes \beta(d(x_3, y_{23})) \\
&= \int_{D\times E} (1 - H(x_2, x_3, y_{23})) \mathbb{1}_{A_{213}}(x_1, x_2, x_3) \alpha \otimes \beta(d(x_3, y_{23})).
\end{aligned}$$

Thus,

$$\begin{aligned}
H(x_1, x_2, y_{12}) & \int_D \mathbb{1}_{A_{123}}(x_1, x_2, x_3) \alpha(dx_3) \\
&= \int_{D\times E} H(x_1, x_3, y_{13}) \mathbb{1}_{A_{123}}(x_1, x_2, x_3) \alpha \otimes \beta(d(x_3, y_{13}))
\end{aligned}$$

and

$$\begin{aligned}
H(x_1, x_2, y_{12}) & \int_D \mathbb{1}_{A_{213}}(x_1, x_2, x_3) \alpha(dx_3) \\
&= \int_{D\times E} (1 - H(x_2, x_3, y_{23})) \mathbb{1}_{A_{213}}(x_1, x_2, x_3) \alpha \otimes \beta(d(x_3, y_{23})).
\end{aligned}$$

The last two equations specify the value of $H(x_1, x_2, y_{12})$ as a quantity depending on (x_1, x_2) alone except for those pairs (x_1, x_2) such that

$$\int_D \mathbb{1}_{A_{123}}(x_1, x_2, x_3)\alpha(dx_3) = \int_D \mathbb{1}_{A_{213}}(x_1, x_2, x_3)\alpha(dx_3) = 0,$$

but the set of such pairs has zero $\alpha^{\otimes 2}$-measure by assumption. □

The function W is not arbitrary; it must satisfy some obvious consistency conditions. For example, for distinct $i, j, k \in \mathbb{N}$ when $\langle i, j\rangle = \langle i, k\rangle < \langle j, k\rangle$, it must be the case that $\langle i, j\rangle <_L i$ if and only if $\langle i, k\rangle <_L i$, and this translates into the requirement that when $[\rho, \xi_i] \cap [\rho, \xi_j] = [\rho, \xi_i] \cap [\rho, \xi_k] \subsetneq [\rho, \xi_j] \cap [\rho, \xi_k]$ it must be the case that $W(\xi_i, U_i, \xi_j, U_j) = 1$ if and only if $W(\xi_i, U_i, \xi_k, U_k) = 1$, which in turn translates into the requirement that for $(\mu \otimes \lambda)^{\otimes 3}$-a.e. $((x, u), (y, v), (z, w)) \in (\mathbf{S} \times [0, 1])^3$ when $[\theta, x] \cap [\theta, y] = [\theta, x] \cap [\theta, z] \subsetneq [\theta, y] \cap [\theta, z]$ it must be the case that $W((x, u), (y, v)) = W((x, u), (z, w))$.

The next result specifies fully these consistency conditions and combines, without the need for significant further argument, the development leading to Proposition 7.4 with the considerations so far in this section about resolving "left-vs.-right" to give a complete characterization of the family of ergodic exchangeable random didendritic systems, and hence a concrete description of the family of extremal infinite Rémy bridges. The result is thus an explicit determination of the Doob–Martin boundary of the Rémy chain. The only point that deserves some added explanation is the claim of ergodicity in the statement of the converse; however, this follows from the observation made in Remark 5.18 that ergodicity of an exchangeable random didendritic system (on $\mathbb{N}$) is equivalent to the independence of the exchangeable random didendritic systems it induces on disjoint subset of $\mathbb{N}$.

THEOREM 8.2. *Consider a complete separable $\mathbb{R}$-tree $\mathbf{S}$, a point $\theta \in \mathbf{S}$, a diffuse probability measure μ on $\mathbf{S}$, and a Borel function $W : (\mathbf{S} \times [0, 1])^2 \to \{0, 1\}$. Let λ be Lebesgue measure on $[0, 1]$. Suppose that the following hold*:

- *For $\mu^{\otimes 3}$-a.e. $(x, y, z) \in \mathbf{S}^3$, two of the three geodesic segments $[\theta, x] \cap [\theta, y]$, $[\theta, x] \cap [\theta, z]$, $[\theta, y] \cap [\theta, z]$ are equal and these two are strictly contained in the third.*
- *For $(\mu \otimes \lambda)^{\otimes 3}$-a.e. $((x, u), (y, v), (z, w)) \in (\mathbf{S} \times [0, 1])^3$, $[\theta, x] \cap [\theta, y] = [\theta, x] \cap [\theta, z] \subsetneq [\theta, y] \cap [\theta, z]$ implies that $W((x, u), (y, v)) = W((x, u), (z, w))$.*
- *For $(\mu \otimes \lambda)^{\otimes 2}$-a.e. $((x, u), (y, v)) \in (\mathbf{S} \times [0, 1])^2$, $W((x, u), (y, v)) = 1 - W((y, v), (x, u))$.*

Let $(\xi_1, U_1), (\xi_2, U_2), \ldots$ be i.i.d. random elements of $\mathbf{S} \times [0, 1]$ with common distribution $\mu \otimes \lambda$. There is an ergodic exchangeable random didendritic system $(\equiv, \langle\cdot, \cdot\rangle, <_L, <_R)$ defined as follows. The random equivalence relation $\equiv$ on $\mathbb{N} \times \mathbb{N}$ is given by declaring that

$$(i, i) \equiv (k, \ell) \quad \text{if and only if} \quad (i, i) = (k, \ell),$$

and

$$(i, j) \equiv (k, \ell), i \neq j, k \neq \ell, \quad \textit{if and only if} \quad [\theta, \xi_i] \cap [\theta, \xi_j] = [\theta, \xi_k] \cap [\theta, \xi_\ell].$$

The random partial orders $<_L$ and $<_R$ on the corresponding set of equivalence classes $\{\langle i, j\rangle : i, j \in \mathbb{N}\}$ are specified by declaring for $i, j \in \mathbb{N}$, $i \neq j$, that

$$\langle i, j\rangle <_L i \textit{ and } \langle i, j\rangle <_R j \quad \textit{if and only if} \quad W(\xi_i, U_i, \xi_j, U_j) = 1.$$

Conversely, any ergodic exchangeable random didendritic system has the same probability distribution as one constructed in this manner for $\mathbf{S}$, θ, μ, W *satisfying the assumptions above.*

EXAMPLE 8.3. Recall the infinite Rémy bridge of Example 4.1. We know from Example 6.7 that we may take:

- $\mathbf{S}$ to be $[0, 1]$ equipped with the usual metric,
- θ to be the point $0 \in [0, 1]$,
- μ to be Lebesgue measure on $[0, 1]$.

We may then take

$$W(x, u, y, v) = \begin{cases} 1, & \text{if } x < y \text{ and } u < \dfrac{1}{2}, \\ 0, & \text{if } x < y \text{ and } u > \dfrac{1}{2}, \\ 1, & \text{if } y < x \text{ and } v < \dfrac{1}{2}, \\ 0, & \text{if } y < x \text{ and } v > \dfrac{1}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

EXAMPLE 8.4. Now consider the infinite Rémy bridge of Example 4.2. Here, $\mathbf{S}$ is the $\mathbb{R}$-tree $\mathbf{T}$ of Example 6.8. The leaves of $\mathbf{S}$ are in a bijective correspondence with $\{0, 1\}^\infty$ and μ may be identified with the fair coin-tossing measure κ on $\{0, 1\}^\infty$ introduced in Example 3.4. There is no need for genuine randomization in this case. Indeed, for $\mu^{\otimes 2}$-a.e. (ξ_1, ξ_2) we have either $W(\xi_1, u_1, \xi_2, u_2) = 0$ for $\lambda^{\otimes 2}$-a.e. (u_1, u_2) or $W(\xi_1, u_1, \xi_2, u_2) = 1$ for $\lambda^{\otimes 2}$-a.e. (u_1, u_2). That is, we can just take the $\mathbb{R}$-tree $\mathbf{S}$ and augment it with *deterministic* left–right choices because in this case for any i, j we have $\langle i, j\rangle = \langle k, \ell\rangle$ for infinitely many other k, ℓ. The resulting representation of the infinite Rémy bridge coincides with the one given in Example 4.2.

REMARK 8.5. As we remarked in the Introduction, the distribution of the limit in the Doob–Martin topology of the Rémy chain (i.e., the probability measure that appears in the Poisson boundary of the Rémy chain) is concentrated on points that can be represented in terms of ensembles $\mathbf{S}, \theta, \mu, W$ such that W takes values in $\{0, 1\}$.

REMARK 8.6. Theorem 8.2 gives a concrete characterization of the family of ergodic exchangeable random didendritic systems or, equivalently, the family of extremal infinite Rémy bridges. Consequently, it gives an explicit description of the points in the Doob–Martin boundary of the Rémy chain. Of course, the ingredients appearing in the representation afforded by the result are not unique. Also, the Doob–Martin boundary is not just a set: it carries a metrizable topological structure. However, a sequence of representations corresponds to a convergent sequence of boundary points if and only if the restrictions of the associated exchangeable random didendritic systems to finite subsets of $\mathbb{N}$ converge in distribution. That is, a sequence of representations corresponds to a convergent sequence of boundary points if and only if for all m the sequence of random binary trees built by sampling $m+1$ points according to the associated sampling measure and determining left-versus-right orderings using extra randomness as necessary converges in distribution.

Acknowledgments. Rudolf Grübel thanks Julian Gerstenberg for many stimulating discussions. We thank Stephan Gufler for pointing out a gap in a previous version of Section 7 and showing us how to repair it.

REFERENCES

[1] ALDOUS, D. (1991). The continuum random tree. I. *Ann. Probab.* **19** 1–28. MR1085326

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

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

[4] ALONSO, L. and SCHOTT, R. (1995). *Random Generation of Trees*. *Random Generators in Computer Science*. Kluwer Academic, Boston, MA. MR1331596

[5] AUSTIN, T. (2008). On exchangeable random variables and the statistics of large graphs and hypergraphs. *Probab. Surv.* **5** 80–145. MR2426176

[6] BLACKWELL, D. and MACQUEEN, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355. MR0362614

[7] BORGS, C., CHAYES, J. and LOVÁSZ, L. (2010). Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal.* **19** 1597–1619. MR2594615

[8] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2008). Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** 1801–1851. MR2455626

[9] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2012). Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math.* (2) **176** 151–219. MR2925382

[10] CHEN, B. and WINKEL, M. (2013). Restricted exchangeable partitions and embedding of associated hierarchies in continuum random trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 839–872. MR3112436

[11] DIACONIS, P. and JANSON, S. (2008). Graph limits and exchangeable random graphs. *Rend. Mat. Appl.* (7) **28** 33–61. MR2463439

[12] DOOB, J. L. (1959). Discrete potential theory and boundaries. *J. Math. Mech.* **8** 433–458; erratum 993. MR0107098

[13] EVANS, S. N. (2008). *Probability and Real Trees. Lecture Notes in Math.* **1920**. Springer, Berlin. MR2351587

[14] EVANS, S. N., GRÜBEL, R. and WAKOLBINGER, A. (2012). Trickle-down processes and their boundaries. *Electron. J. Probab.* **17** no. 1, 58. MR2869248

[15] FÖLLMER, H. (1975). Phase transition and Martin boundary. In *Séminaire de Probabilités, IX* (*Seconde Partie, Univ. Strasbourg, Strasbourg, Années Universitaires* 1973/1974 *et* 1974/1975). *Lecture Notes in Math.* **465** 305–317. Springer, Berlin. MR0426176

[16] GRÜBEL, R. (2015). Persisting randomness in randomly growing discrete structures: Graphs and search trees. *Discrete Math. Theor. Comput. Sci.* **18** 1–23.

[17] GUFLER, S. (2016). A representation for exchangeable coalescent trees and generalized tree-valued Fleming–Viot processes. Preprint. Available at arXiv:1608.08074.

[18] JANSON, S. (2011). Poset limits and exchangeable random posets. *Combinatorica* **31** 529–563. MR2886098

[19] KALLENBERG, O. (2005). *Probabilistic Symmetries and Invariance Principles*. Springer, New York. MR2161313

[20] KEMENY, J. G., SNELL, J. L. and KNAPP, A. W. (1976). *Denumerable Markov Chains*, 2nd ed. Springer, New York. MR0407981

[21] LE GALL, J.-F. (1999). *Spatial Branching Processes, Random Snakes and Partial Differential Equations*. Birkhäuser, Basel. MR1714707

[22] LOVÁSZ, L. (2012). *Large Networks and Graph Limits. American Mathematical Society Colloquium Publications* **60**. Amer. Math. Soc., Providence, RI. MR3012035

[23] LOVÁSZ, L. and SZEGEDY, B. (2006). Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** 933–957. MR2274085

[24] LUCZAK, M. and WINKLER, P. (2004). Building uniformly random subtrees. *Random Structures Algorithms* **24** 420–443. MR2060629

[25] MARCHAL, P. (2003). Constructing a sequence of random walks strongly converging to Brownian motion. In *Discrete Random Walks* (*Paris*, 2003). *Discrete Math. Theor. Comput. Sci. Proc., AC* 181–190 (electronic). Assoc. Discrete Math. Theor. Comput. Sci., Nancy. MR2042386

[26] NOAH, C., HAULK, C. and PITMAN, J. (2011). A representation of exchangeable hierarchies by sampling from real trees. Preprint. Available at arXiv:1101.5619.

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

[28] RÉMY, J.-L. (1985). Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire. *RAIRO. Inform. Théor.* **19** 179–195. MR0803997

[29] REVUZ, D. (1975). *Markov Chains. North-Holland Mathematical Library* **11**. North-Holland, Amsterdam. MR0415773

[30] ROGERS, L. C. G. and WILLIAMS, D. (2000). *Diffusions, Markov Processes, and Martingales. Cambridge Mathematical Library* **1**. Cambridge Univ. Press, Cambridge. Foundations, Reprint of the second (1994) edition. MR1796539

[31] SAWYER, S. A. (1997). Martin boundaries and random walks. In *Harmonic Functions on Trees and Buildings* (*New York*, 1995). *Contemp. Math.* **206** 17–44. Amer. Math. Soc., Providence, RI. MR1463727

[32] STANLEY, R. P. (1997). *Enumerative Combinatorics. Vol.* 1. *Cambridge Studies in Advanced Mathematics* **49**. Cambridge Univ. Press, Cambridge. MR1442260

[33] TAO, T. (2007). A correspondence principle between (hyper)graph theory and probability theory, and the (hyper)graph removal lemma. *J. Anal. Math.* **103** 1–45. MR2373263

[34] VON WEIZSÄCKER, H. (1983). Exchanging the order of taking suprema and countable intersections of σ-algebras. *Ann. Inst. H. Poincaré Sect. B* (*N.S.*) **19** 91–100. MR0699981

[35] WOESS, W. (2009). *Denumerable Markov Chains. Generating Functions, Boundary Theory, Random Walks on Trees*. European Mathematical Society (EMS), Zürich. MR2548569

S. N. EVANS
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
367 EVANS HALL #3860
BERKELEY, CALIFORNIA 94720-3860
USA
E-MAIL: evans@stat.berkeley.edu

R. GRÜBEL
INSTITUT FÜR MATHEMATISCHE STOCHASTIK
LEIBNIZ UNIVERSITÄT HANNOVER
POSTFACH 6009
30060 HANNOVER
GERMANY
E-MAIL: rgrubel@stochastik.uni-hannover.de

A. WAKOLBINGER
INSTITUT FÜR MATHEMATIK
GOETHE-UNIVERSITÄT
60054 FRANKFURT AM MAIN
GERMANY
E-MAIL: wakolbin@math.uni-frankfurt.de