Convex minorant trees associated with Brownian paths and the continuum limit of the minimum spanning tree

Nicolas Broutin [*] Jean-François Marckert [†]

July 21, 2023

Abstract

We give an explicit construction of the scaling limit of the minimum spanning tree of the complete graph. The limit object is described using a recursive construction involving the convex minorants of a Brownian motion with parabolic drift (and countably many i.i.d. uniform random variables); we call it the Brownian parabolic tree.

Aside from the new representation, this point of view has multiple consequences. For instance, it permits us to prove that its Hausdorff dimension is almost surely 3. It also intrinsically contains information related to some underlying dynamics: one notable by-product is the construction of a standard metric multiplicative coalescent which couples the scaling limits of random graphs at different points of the critical window in terms of the same simple building blocks.

The above results actually fit in a more general framework. They result from the introduction of a new family of continuum random trees associated with functions via their convex minorants, that we call convex minorant trees. We initiate the study of these structures in the case of Brownian-like paths. In passing, we prove that the convex minorant tree of a Brownian excursion is a Brownian continuum ranndom tree, and that it provides a coupling between the Aldous–Pitman fragmentation of the Brownian continuum random tree and its representation by Bertoin.

1 Introduction

1.1 Main results

For a connected graph $G = (V, E)$, together with distinct positive weights associated to the edges, the minimum weight spanning tree is the unique connected spanning subgraph of G that minimizes the total sum of the edge weights. The classical random model consists in taking the complete graph on $[n] := \{1, 2, \dots, n\}$ and independent and identically distributed (i.i.d.) random weights w_e, $e \in \binom{[n]}{2}$, uniform on $[0, 1]$. Then let M_n denote the corresponding minimum spanning tree (MST) rooted at $\rho_n = 1$. It has been proved by Addario-Berry, Broutin, Goldschmidt, and Miermont [4] that, seen as a metric space, M_n admits a scaling limit in the following sense: Let d_n be the graph distance on M_n, let μ_n be the counting measure on $[n]$. Then, there exists a (non-trivial) compact measured metric space $(\mathscr{M}, d)$, a point $\rho \in \mathscr{M}$, and a Borel probability measure μ on $(\mathscr{M}, d)$ such that

$$(M_n, n^{-1/3} d_n, n^{-1} \mu_n, \rho_n) \xrightarrow[n\to\infty]{} (\mathscr{M}, d, \mu, \rho) \tag{1}$$

in distribution, in the sense of Gromov–Hausdorff–Prokhorov. The main result of this paper is to provide an explicit representation of the measured metric space $(\mathscr{M}, d, \mu)$ using a Brownian motion, and a countable collection of i.i.d. uniform random variables, and to initiate the study of some of its properties and consequences. To do so, we introduce a new general class of tree-like structures constructed from functions in a way that differs from the classical contour function encoding.

[*]LPSM, Sorbonne Université, 4 Place Jussieu, 75005 Paris and Institut Universitaire de France (IUF)
[†]Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, 33400 Talence, France

The study of trees and their encoding has a long history. A prominent example is the now classical encoding of trees from a height or contour function which defines a tree-like metric d from a continuous function using the recursive structure of its level sets. The representation is intimately related to branching processes and fragmentations related to heights, and thus to the process of local times of the height function [47, 49, 54]. Notable examples include the Brownian continuum random tree [11] seen as encoded by a Brownian excursion [12, 48], and Lévy trees [49].

Our construction differs radically. The tree will be associated to a continuous function ω defined on an interval $D \subseteq \mathbb{R}$ using the tree-like structure of the family of greatest convex minorants of the graph of ω on the intervals $[0, x]$, $x \in D$. Furthermore, while the classical height function encoding provides a metric d_ω that is continuous on D^2, the metrics we construct are discontinuous at every local minimum, and the information contained in the encoding function is greatly shuffled when ω is irregular. We nonetheless hope to demonstrate that the proposed construction provides a convincing point of view for a number of natural problems involving dynamics, in particular those related to remarkable coalescent and fragmentation processes.

In the present paper, we focus on the very specific case of Brownian like functions, but the reader will easily be convinced that the procedure should apply more generally to càdlàg functions that have only positive jumps such as spectrally positive Lévy processes which will be studied elsewhere. We will in particular define a *convex minorant tree* $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ from a Brownian excursion $\mathrm{e} = (\mathrm{e}_s)_{s\in[0,1]}$ and an independent family of uniform random variables $\mathbf{U} = (U_i)_{i\geq 1}$. Formally, $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ will be a compact pointed measure metric space that we initially define together with a metric d on $[0, 1]$.

Defining $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ is an important building block towards the definition of our main object of interest, where we replace the Brownian excursion e by another Brownian-like path. Let $(W_s)_{s\geq 0}$ be a standard (linear) Brownian motion on $\mathbb{R}_+$ and for $\lambda \in \mathbb{R}$, and $s \geq 0$, define the Brownian motion with parabolic drift by

$$X_s^\lambda = W_s - \frac{s^2}{2} + \lambda s. \tag{2}$$

We usually write $X := X^0$ when $\lambda = 0$. Our main result is the following:

Theorem 1.1. *As $n \to \infty$, we have the following convergence in distribution for the Gromov–Hausdorff–Prokohorov topology:*

$$(M_n, n^{-1/3} d_n, n^{-1}\mu_n, \rho_n) \xrightarrow[n\to\infty]{} \mathrm{CMT}(X, \mathbf{U})\,.$$

We call $\mathrm{CMT}(X, \mathbf{U})$ the Brownian parabolic tree. In particular, the limit appearing in (1) is such that

$$(\mathscr{M}, d, \mu, \rho) \overset{d}{=} \mathrm{CMT}(X, \mathbf{U})\,.$$

Its structure and properties provide a way to make explicit computations. For instance, the Hausdorff dimension of $(\mathscr{M}, d)$ was still unknown, and we show directly

Theorem 1.2. *Almost surely, the space $\mathrm{CMT}(X, \mathbf{U})$ is compact and has Hausdorff dimension 3.*

$\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ is one of the central objects of this paper, together with its variants. The following results can be seen as consequences of Theorem 1.1. For a natural number $s \geq 0$, we let C_n^s denote a uniformly random connected graph on $[n]$ with $n - 1 + s$ edges. Assuming that the edge weights on this component are i.i.d. uniform on $[0, 1]$, C_n^s possesses an a.s. unique minimum spanning tree that we denote by T_n^s. It is a consequence of [4] that, for any $s \geq 0$, the graphs T_n^s considered as metric spaces equipped with the graph distance d_n^s and the counting measure on the nodes μ_n^s have a limit when suitably rescaled. The following theorem provides an explicit representation of these limits. For $s \geq 0$, let $\mathrm{e}^{(s)}$ be a process on $[0, 1]$ whose distribution is characterized by (for all $f : \mathcal{C}([0, 1]) \to \mathbb{R}$ bounded continuous)

$$\mathbf{E}[f(\mathrm{e}^{(s)})] = \frac{\mathbf{E}[f(\mathrm{e}) \cdot (\int_0^1 \mathrm{e}(u)du)^s]}{\mathbf{E}[(\int_0^1 \mathrm{e}(u)du)^s]}\,,$$

where e is a standard normalized Brownian excursion.

Theorem 1.3. *For any natural number $s \geq 0$, we have the following convergence in distribution for the Gromov–Hausdorff–Prokhorov topology:*

$$(T_n^s, n^{-1/2} d_n^s, \mu_n^s) \xrightarrow[n\to\infty]{} \mathrm{CMT}(\mathrm{e}^{(s)}, \mathbf{U})\,.$$

In particular, for $s = 0$, this implies that $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ is a Brownian continuum random tree.

The last claim when $s = 0$ follows from simple observations: first $\mathrm{e}^{(0)}$ is simply a standard Brownian excursion; second $T_n^0 = C_n^0$ since the latter is already a tree, which must then be uniform, and it is well-known that such trees converge to the Brownian continuum random tree [11, 12, 48].

Let us to back to the case of the Brownian motion with parabolic drift X. The construction of the convex minorant tree inherently captures some hidden dynamics. The explanation shall come later, and we will for now only present some facts. For $\lambda \in \mathbb{R}$ and $t \geq 0$, let

$$B_t^\lambda := X_t^\lambda - \underline{X}_t^\lambda \qquad \text{and} \qquad Z^\lambda = \{s \in \mathbb{R}_+ : B_s^\lambda = 0\}\,.$$

The process $(Z^\lambda)_{\lambda\in\mathbb{R}}$ is non-increasing for the inclusion, and therefore induces a coalescent of $\mathbb{R}_+$: the intervals of $\mathbb{R}_+ \setminus Z^\lambda$ can be a.s. indexed in decreasing order of their lengths as $\boldsymbol{\gamma}^\lambda = (\gamma_1^\lambda, \gamma_2^\lambda, \dots)$. It is known [17, 24] that the process of the lengths of the intervals $(|\boldsymbol{\gamma}^\lambda|)_{\lambda\in\mathbb{R}}$ is the standard multiplicative coalescent constructed by Aldous [14]. However, the space $\mathrm{CMT}(X, \mathbf{U})$ being constructed as $(\mathscr{M}, d, \mu, \rho)$ from the completion of a random metric d on $\mathbb{R}_+$, it comes with a canonical injection $\pi : \mathbb{R}_+ \to \mathscr{M}$ that allows to transport Z^λ into $\mathscr{M}$. As a consequence, as λ varies, the points of Z^λ actually also induce a coalescent/fragmentation of $\mathrm{CMT}(X, \mathbf{U})$ in the sense that $\pi(Z^\lambda)$ is a non-increasing set of points in $\mathscr{M}$. We shall now explore more precisely this process.

In the construction of $\mathrm{CMT}(X, \mathbf{U})$, the entries in $\mathbf{U}$, which are i.i.d. uniform random variables, are assigned to the local minima of X. For an interval $I \subseteq \mathbb{R}_+$, let $\mathbf{U}|_I$ denote the sequence of those entries that are assigned to local minima lying in I (in the same order as in $\mathbf{U}$). For each $i \geq 1$, let

$$\tilde{e}_i^\lambda(s) := B^\lambda(s + \inf \gamma_i^\lambda)\mathbf{1}_{0\leq s\leq |\gamma_i^\lambda|}\,.$$

Let $\mathfrak{F}^\lambda = (\mathrm{CMT}(\tilde{e}_i^\lambda, \mathbf{U}|_{\gamma_i^\lambda}), i \geq 1)$ be the collection of convex minorant trees of the excursions $\tilde{e}_i^\lambda$, $i \geq 1$. Let now $\mathcal{S}$ be an independent Poisson point process with intensity a half on $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$. There exists a measurable function of $(X, \mathbf{U}, \mathcal{S})$ that yields, for each $\lambda \in \mathbb{R}$, a collection of measured metric spaces $\mathfrak{G}^\lambda$ obtained from $\mathfrak{F}^\lambda$ by identifying the points $\pi(x)$ and $\pi(y)$ for each $(x, y, t) \in \mathcal{S}$ such that $t \leq \lambda$ and no point of Z^t lies in the closed interval between x and y (i.e., x and y are in the same interval of the fragmentation at time t). Almost surely, there are only finitely many points of $\mathcal{S}$ satisfying these constraints for each $\lambda \in \mathbb{R}$ and $i \geq 1$.

We now define some discrete analogs, which are more classical. Let E^n denote $\binom{[n]}{2}$. For each $p \in [0, 1]$ write $E_p^n := \{e \in E^n : w_e \leq p\}$ so that the graph $G(n, p) = ([n], E_p^n)$ is a classical Erdős–Rényi random graph, and the process $(G_n^p)_{p\in[0,1]}$ is non-decreasing (in the sense of inclusion of edge sets). The regime of interest is the one when

$$p = p_n(\lambda) := \frac{1}{n} + \frac{\lambda}{n^{4/3}}. \tag{3}$$

Let $C_i^{n,\lambda}$ be the i-th largest connected component of $G(n, p_n(\lambda))$, breaking ties using the minimum label. Let $\mathfrak{G}^{n,\lambda} = (\mathfrak{G}_i^{n,\lambda}, i \geq 1)$, where

$$\mathfrak{G}_i^{n,\lambda} = (C_i^{n,\lambda}, n^{-1/3} d_i^{n,\lambda}, n^{-2/3} \mu_i^{n,\lambda})$$

denotes the corresponding measured metric space, where $d_i^{n,\lambda}$ is the graph distance, and $\mu_i^{n,\lambda}$ denote the counting measure on (the vertex set of) $C_i^{n,\lambda}$. One may similarly define the minimum spanning forest $\mathfrak{F}^{n,\lambda} = (\mathfrak{F}_i^{n,\lambda}, i \geq 1)$, where

$$\mathfrak{F}_i^{n,\lambda} = (C_i^{n,\lambda}, n^{-1/3} \delta_i^{n,\lambda}, n^{-2/3} \mu_i^{n,\lambda})$$

and $\delta_i^{n,\lambda}$ is the graph distance on the minimum spanning tree of $C_i^{n,\lambda}$ (constructed from the same collection of weights (w_e)).

Then the processes $(\mathfrak{F}^\lambda)_{\lambda\in\mathbb{R}}$ and $(\mathfrak{G}^\lambda)_{\lambda\in\mathbb{R}}$ enjoy some continuum Kruskal and Erdős–Rényi dynamics reflecting the evolution of $\mathfrak{F}^{n,\lambda}$ and $\mathfrak{G}^{n,\lambda}$, respectively, in following sense:

Theorem 1.4. *For each λ, and each $i \geq 1$, $\mathfrak{F}_i^\lambda$ is isometric to the subet of $\mathscr{M}$ induced by γ_i^λ. Furthermore we have, for any $k \geq 1$ and $\lambda_1 < \lambda_2 < \cdots < \lambda_k$, jointly*

$$(\mathfrak{G}^{n,\lambda_1}, \mathfrak{G}^{n,\lambda_2}, \dots, \mathfrak{G}^{n,\lambda_k}) \xrightarrow[n\to\infty]{} (\mathfrak{G}^{\lambda_1}, \mathfrak{G}^{\lambda_2}, \dots, \mathfrak{G}^{\lambda_k}) \quad \textit{and},$$
$$(\mathfrak{F}^{n,\lambda_1}, \mathfrak{F}^{n,\lambda_2}, \dots, \mathfrak{F}^{n,\lambda_k}) \xrightarrow[n\to\infty]{} (\mathfrak{F}^{\lambda_1}, \mathfrak{F}^{\lambda_2}, \dots, \mathfrak{F}^{\lambda_k}),$$

in distribution, where, in each case, the convergence holds with respect to the product Gromov–Hausdorff–Prokhorov topology on sequences of measured metric spaces.

Theorem 1.4 provides an explicit coupling for the standard metric coalescent dynamics constructed by Rossignol in [62] (see also [6]). In particular, this shows that $\mathrm{CMT}(X, \mathbf{U})$ is the right object to lift the multiplicative coalescent defined by Aldous to the level of metric spaces (as well as its augmented version [22]). We are not interested here in verifying that there is indeed a natural Markov semigroup acting on measured metric spaces that formalizes these dynamics; such a Markov processes is constructed and studied in [6] (see also [35]).

There is also an analog to Theorem 1.4 replacing $\mathrm{CMT}(X, \mathbf{U})$ by $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ which is relevant to the additive coalescent. The following notation are intentionally similar to that used previously; it shall always be clear to which case we refer. For each $\lambda \leq 0$, let $\mathrm{e}^\lambda(s) = \mathrm{e}(s) + \lambda s$, and $\underline{\mathrm{e}}^\lambda(s) = \inf\{\mathrm{e}^\lambda(r) : 0 \leq r \leq s\}$. Write $Z^\lambda = \{s \in [0,1] : \mathrm{e}^\lambda(s) = \underline{\mathrm{e}}^\lambda(s)\}$. The process Z^λ is non-increasing in λ and induces a coalescent of $[0,1]$, as λ varies in $(-\infty, 0]$. Let $(\gamma_i^\lambda)_{i\geq 1}$ denote the sequence of lengths of the intervals of $[0,1] \setminus Z^\lambda$, in decreasing order. It is known since the results of Bertoin [20] that the process $t \mapsto (\gamma_i^{-t})_{i\geq 1}$ for $t \geq 0$ is the fragmentation dual to the standard additive coalescent introduced by Aldous and Pitman [15]. Just as before $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ is a random measured real tree $\mathfrak{T} = (\mathscr{T}, d, \mu)$ defined through the completion of a random metric d on $[0,1]$, and we let $\pi : [0,1] \to \mathscr{T}$ denote the canonical injection. This allows one to transport Z^λ in $\mathscr{T}$ and therefore, to see $\mathscr{T} \setminus \pi(Z^{-t})$, $t \geq 0$, as a fragmentation of $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$.

Formally, for each $\lambda \leq 0$ and $i \geq 1$, let

$$\mathrm{e}_i^\lambda(s) := (\mathrm{e}^\lambda(\inf \gamma_i^\lambda + s) - \underline{\mathrm{e}}^\lambda(s))\mathbf{1}_{0\leq s\leq |\gamma_i^\lambda|}.$$

Let $\mathfrak{T}_i^\lambda := \mathrm{CMT}(\mathrm{e}_i, \mathbf{U}_i^\lambda) = (\mathscr{T}_i^\lambda, d_i^\lambda, \mu_i^\lambda)$. The following theorem provides an explicit coupling between the representations of the fragmentation that is dual to the additive coalescent due to Aldous & Pitman on the one hand [15], and to Bertoin [20] on the other. It also provides another point of view on some recent results of Kortchemski and Thévenin [45]. Define $\mathcal{P} := \{(\pi(x), -\lambda) : x \in Z^{\lambda-} \setminus Z^\lambda, \pi(x) \in \mathrm{Skel}(\mathscr{T}), x \in [0,1], \lambda \leq 0\}$, where $\mathrm{Skel}(\mathscr{T})$ is the skeleton of $\mathscr{T}$ that we define here as the set of points $u \in \mathscr{T}$ such that $\mathscr{T} \setminus \{u\}$ has at least two connected components.

Theorem 1.5 (Aldous–Pitman vs Bertoin)**.** *Let* e *be a normalized Brownian excursion and recall that $\mathfrak{T} = (\mathscr{T}, d, \mu) = \mathrm{CMT}(\mathrm{e}, \mathbf{U})$. Almost surely, for all $\lambda \leq 0$ and all $i \geq 1$, $\mathfrak{T}_i^\lambda$ is isometric to the subtree of $\mathscr{T}$ induced by γ_i^λ. Furthermore*

i) for each $\lambda \leq 0$ and each $i \geq 0$, $|\gamma_i^\lambda| = \mu(\mathscr{T}_i^\lambda)$,

ii) conditionally on $(\mathscr{T}, d)$, $\mathcal{P}$ is a Poisson point process of unit intensity on $\mathrm{Skel}(\mathscr{T}) \times \mathbb{R}_+$.

Hence, the process $(\boldsymbol{\gamma}^{-t})_{t\geq 0}$ is precisely the Aldous–Pitman fragmentation of the Brownian CRT $\mathfrak{T}$.

1.2 Motivation and history of related results

It was already known from the work of Addario-Berry, Broutin, Goldschmidt, and Miermont [4] that $\mathfrak{M}_n = (M_n, n^{-1/3}d_n, n^{-1}\mu_n, \rho_n)$ converges in distribution. The proof relies on a Cauchy sequence

argument for the distribution of $\mathfrak{M}_n$, and is thus essentially existential. In particular, it does not provide an explicit construction of the limit. The novelty of Theorem 1.1 lies in the identification of the limit as the convex minorant tree $\mathrm{CMT}(X, \mathbf{U})$. Note that, by results of Addario-Berry and Sen [8], this is also the scaling limit of random 3-regular graphs.

ABOUT THE SCALING LIMIT OF $\mathfrak{M}_n$. In order to understand the underlying issues, let us be more specific about the approach used in [4]. The general idea is to analyse the minimum spanning tree using Kruskal's algorithm [46]. This algorithm proceeds by adding the edges by increasing order of weights to an initially empty graph, provided doing so does not create a cycle. Since in the random setting, the order is uniformly random, the (conditional) distribution according to which the edges are added at each step is straightforward, and the difficulty consists in avoiding the cycles. So one may try to first add edges regardless of whether they create cycles or not, with the hope to be able to deal with that issue later on. One shall do this up to a threshold for the weights that ensures that there are not too many cycles (or dealing with them would be hard), but that the connected components are already fairly large (or we have basically gathered no information). These two competing constraints lead to the choice of keeping only edges with weight at most $p_n(\lambda) = \frac{1}{n} + \lambda n^{-4/3}$ with $\lambda \in \mathbb{R}$ large.

This $p_n(\lambda)$ happens to be precisely the critical window of the random graphs. The scaling limit of $G(n, p_n(\lambda))$, seen as the sequence of compact metric spaces $\mathfrak{G}^{n,p_n(\lambda)}$ is known from the results of Addario-Berry, Broutin, and Goldschmidt [2] who built on the pioneering work of Aldous who had previously obtained the scaling limit for the vector of the sizes of the connected components [14]. The analysis in [4] relies on the fact that, (1) given a connected component of the random graph, one may obtain a tree distributed like its minimum spanning tree by breaking cycles randomly (removing uniformly edges, unless they disconnect the component), and (2) that a similar procedure works on the scaling limit. This forward/backward procedure provides some geometric information but it is inherently tricky to track it precisely. This explains why it does not lead to an explicit construction of the limit in terms of simple building blocks, or also why the Hausdorff dimension (Proposition 5.5) remained unknown. Furthermore, this approach is fundamentally incapable of providing any result about the behavior at different times, since the cycle breaking procedure removes all cycles.

RELATED RESULTS ON THE MST. Let us mention that Angel and Senizergues are currently finishing a paper in which they study the scaling limit of the local limit of M_n, that was described by Addario-Berry [7]. As the local weak limit of M_n is an infinite tree, their object $\mathcal{M}$ is not compact; still $\mathcal{M}$ has Hausdorff dimension 3, and it also seems to be the local limit of $\mathrm{CMT}(X, \mathbf{U})$. It is our understanding that they also plan to study a "mesoscopic" limit that would be an analog to the self-similar CRT of Aldous [10] for $\mathfrak{M}_n$.

ABOUT THE SCALING LIMIT OF RANDOM GRAPHS. It is known that the critical random graphs have a scaling limit [2], which has been constructed for each $\lambda \in \mathbb{R}$ in [2] (see also [1]): for each $i \geq 1$, a connected component is built as the tree with height process $\tilde{e}_i^\lambda$, in which cycles are created by identifying pairs of points whose locations are given by a Poisson point process under $\tilde{e}_i^\lambda$. The tree is genuinely different from $\mathrm{CMT}(\tilde{e}_i^\lambda, \mathbf{U}|_{\gamma_i^\lambda})$ that we use here. However, the marginals described in Theorem 1.4 of course correspond. For instance, the number of pairs of points that are identified must have the same distribution conditionally on the excursion. One quickly verifies that (Lemma A.3), for $\lambda \in \mathbb{R}$ and $i \geq 1$, the average number of pairs given $\tilde{e}_i^\lambda$ (which also determines all the γ_j^r which are subsets of γ_i^λ), is

$$\int_{\gamma_i^\lambda} \tilde{e}_i^\lambda(s) ds = \frac{1}{2} \int_{-\infty}^{\lambda} \sum_{j \geq 1} |\gamma_j^r|^2 \mathbf{1}_{\gamma_j^r \subseteq \gamma_i^\lambda} dr \,.$$

The spanning subtree used in [2] is discovered by a depth-first search; quite recently, Miermont and Sen [55] have studied the construction of these scaling limits from a breadth-first exploration. The procedures used in [2, 14, 55] are not consistent as $\lambda \in \mathbb{R}$ varies, and the objects obtained for two different values of λ have no reason to be close, and do not relate simply to any dynamics.

ABOUT THE LIMIT ERDŐS–RÉNYI AND KRUSKAL DYNAMICS. Since there is an obvious process version for the entire structure at the discrete level, the question of the dynamics for the limit objects (continuum forests or graphs) is quite natural. First it is known from results of Armendariz [17] and Broutin and Marckert [24] that the process $X = X^0$ defined in (2) encodes the standard multiplicative coalescent, and thus permits to obtain a coupling of the limit of the sizes of the connected components (see also [51]). A minor modification also yields a coupling of both the sizes and the number of extra edges via an explicit construction of the augmented multiplicative coalescent constructed by Bhamidi, Budhiraja, and Wang [22] (see also the recent point of view by Corujo and Limic [28, 29]). The metrics require the new point of view of the convex minorant tree. We emphasize that what we mean here by dynamics is a process in λ whose marginals are the scaling limits for fixed λ, and that we do not consider the question of the existence of nice Markov semigroup acting on sequences of compact measured metric spaces; this question is addressed in [6, 62].

Let us now say a few words about the case of the (standard) additive coalescent. It was first introduced by Aldous and Pitman [15] as the time reversal of the fragmentation process where a Brownian continuum random tree is split as time goes using a Poisson point process. Bertoin [20] then observed that one obtains the same fragmentation process by cutting the unit interval at the times where a Brownian excursion plus an increasing linear drift touches its running infimum. These two constructions have been connected in a number of ways at the discrete level, starting with Chassaing and Louchard [26] who used a representation based on hashing with linear probing [43, 44]; the construction of $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ can be seen as a scaling limit for the tree appearing there. Broutin and Marckert [24] and Marckert and Wang [50] provide alternative approaches. Quite recently, the two processes have been coupled directly in the continuous by Kortchemski and Thévenin [45], just as Theorem 1.5. Let us also emphasize the fact that Theorem 1.5 is a by-product of the same construction used for Theorem 1.4: this shows that the standard additive and multiplicative coalescent are, even when considered at the enriched metric level, very strongly related since they are two versions of the same construction applied two different functions (e and X, respectively).

1.3 Intuition and techniques

We shall now try to convey the main ideas that underlie the construction of our scaling limits. The intuition comes from the discrete setting, and we shall explain why the relevant objects should have continuum analogs, and how these limits could be formally defined. There is no very simple axiomatic definition of the minimum spanning tree, at least none that seems suitable to a direct analysis, and one is lead to track the evolution of a construction algorithm in order to obtain the minimum spanning tree. While the construction in [4] relies on Kruskal's algorithm which grows a forest [46], our approach is based on a combination of algorithms by Kruskal and Prim [60], which grows a tree containing a given vertex.

Fix $n \geq 1$. Prim's algorithm proceeds as follows. Let $v_1 = 1$. We define the order of vertices $v_2, v_3, \ldots, v_n$ iteratively. For every $j = 1, \ldots, n$, we let $V_j = \{v_1, \ldots, v_j\}$. For $i = 2, \ldots, n$, let e_i be the edge between V_{i-1} and $[n] \setminus V_{i-1}$ that has the smallest weight. Write $e_i = \{u_i, v_i\}$ with $u_i \in V_{i-1}$ and $v_i \in [n] \setminus V_{i-1}$. Then the minimum spanning tree M_n is the graph on $[n]$ with edge set $\{e_2, \ldots, e_n\}$. The order $v_1, v_2, \ldots, v_n$ is called the Prim order. It turns out that, for any $p \in [0, 1]$, the connected components of $G(n, p) = ([n], E_p^n)$, where $E_p^n = \{e \in E^n : w_e \leq p\}$ are intervals in the Prim order (that is, the vertex set of each connected components is $\{v_a, v_{a+1}, \cdots, v_b\}$ for some $1 \leq a \leq b \leq n$). In particular, as p increases, only adjacent intervals may merge.

Now, consider the graph consisting of edges with weights (strictly) lower than w_{e_i}, $\{e : w_e < w_{e_i}\}$, that is just before the edge e_i is added. Let L_i be the connected component containing v_{i-1} in this graph. These are precisely the connected components that merge when e_i is added. For $p \in [0, 1]$, let $\mathcal{F}_p$ denote the sigma-algebra generated by the events $\{w_e \leq p, e \in E\}$. The following is straightforward:

Lemma 1.6. *For each $2 \leq i \leq n$, conditionally on $\mathcal{F}_{w_{e_i}-}$, the vertex u_i is uniformly random in L_i.*

In other words, in this discrete representation in which the vertices are placed in the Prim order $v_1, \cdots, v_n$, conditionally on the sequence of intervals that merge, the edges that are part of the minimum spanning tree precisely connect a uniform random vertex in the left interval to the left-most vertex in the right interval. Still in this discrete representation, determining the distribution of the sequence of pairs of intervals that merge together is not quite as easy any longer. Fortunately, in the limit, it is given explicitly by the rather nice process $\mathbb{R}_+ \setminus Z^\lambda$. One might thus hope that, in the limit, one should be able to construct the scaling limit of the minimum spanning tree as follows: for each $\lambda \in \mathbb{R}$, each interval γ_i^λ should be associated to a continuum random tree, and as λ increases, these trees should merge using an analog of the discrete dynamics: each time two intervals merge, a uniformly random point in the left interval and the left-most point of the right one should be identified; the minimum spanning tree should then be the limit as $\lambda \to \infty$ (which would indeed be a tree since $\gamma_1^\lambda \uparrow (0, \infty)$ as $\lambda \to \infty$).

While these dynamics are reasonable, they do not really provide a clear path towards a construction: while at the discrete level, the addition of edges does create some length, identifying points in the limit does not, and it remains to understand from what the length emerges. A natural idea consists in constructing the length using some kind of local time arising from the process $(Z^\lambda)_{\lambda \in \mathbb{R}}$. With this objective in mind, let us go back to the discrete setting. For any $i, j \in [n]$, we may find all the nodes on the path between i and j in the minimum spanning tree as follows. For some $p \in [0, 1]$, it is convenient to write $i \sim_{n,p} j$ if i and j lie in the same connected component of the graph with edges of weight at most p: let $i < j \in [n]$ and let $p(i, j) = \inf\{p : i \sim_{n,p} j\}$. The path between i and j must go through the unique edge $e_k = \{u_k, v_k\}$ with weight $p(i, j)$; then, at time $p(i, j)-$ we are left with two connected components, each containing a pair of points (u_k and i on the one hand, and v_k and j on the other) that should each be connected by a path. Proceeding recursively, the process eventually terminates and yields precisely the collection of nodes which are on the path between i and j, and the distance $d_n(i, j)$ is then simply the cardinality of that set (minus one).

This approach is amenable to an extension to the continuous setting, that we expose here informally. For $x, y \in \mathbb{R}_+$, let $x \sim_\lambda y$ if there is no point of Z^λ in the closed interval between x and y. Let $I^\lambda(x) := \{y \in \mathbb{R}_+ : x \sim_\lambda y\}$. Take now $x < y$ for convenience. Let $\lambda(x, y) = \inf\{\lambda \in \mathbb{R} : x \sim_\lambda y\}$. It turns out that $Z^{\lambda(x,y)}$ almost surely contains a single point in $[x, y]$, that we denote by $\kappa(x, y)$. Then, just before x and y get connected, we have two distinct intervals $I^{\lambda(x,y)}(x)$, and $I^{\lambda(x,y)}(y)$, which are separated by the point $\kappa(x, y)$. The discrete setting suggests that one should choose a uniformly random point $\eta(x, y)$ in $I^{\lambda(x,y)}$ (this is where the uniforms in $\mathbf{U}$ are used). Then, the two points $\eta(x, y)$ and $\kappa(x, y)$ should be the continuous analog of the extremities of the maximum weight edge on the path between x and y. Proceeding recursively by looking for the path between x and $\eta(x, y)$ in $I^\lambda(x, y)(x)$ on the left, and the path between $\kappa(x, y)$ and y in $I^{\lambda(x,y)}(y)$ on the right should yield a random subset of $\mathbb{R}_+$ containing all the points used to go from x to y, that should resemble some kind of random Cantor set, and the distance between x and y should be some Hausdorff measure of that set. Our main objective is now to verify that this intuition can be turned into formal definitions, but also that the objects constructed are indeed the ones we are looking for.

1.4 Organization of the paper

The paper is organized as follows. In Section 3, we discuss recursive convex minorants, the associated trees and their properties. In particular, it is there that we define the convex minorant trees $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ and $\mathrm{CMT}(X, \mathbf{U})$. In Section 4, we prove that the tree $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ is a Brownian CRT, and we exhibit the coupling mentioned above between the representations of the fragmentation dual to the additive coalescent by Adous–Pitman [15] on the one hand, and by Bertoin [20] on the other. In Section 5, we prove that $\mathrm{CMT}(X, \mathbf{U})$ is almost surely compact. In Section 6, we construct the mass measure and use it to lower bound the Hausdorff dimension. Finally, Section 7 is devoted to proving that the Brownian parabolic tree $\mathrm{CMT}(X, \mathbf{U})$ is distributed like scaling limit of the minimum spanning tree.

Contents

2 Notation

Let $\mathbf{W}$ be the Wiener measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$, the set of continuous functions $f : \mathbb{R}_+ \to \mathbb{R}$; this is the law of standard Brownian motion $(W_t)_{t\geq 0}$ starting at 0. For a continuous process $\omega = (\omega_t)_{t\geq 0}$, we let $\underline{\omega}$ and $\overline{\omega}$ denote respectively the running infimum and supremum processes: $\underline{\omega}_t := \inf\{\omega_s : 0 \leq s \leq t\}$ and $\overline{\omega}_t = \sup\{\omega_t : 0 \leq s \leq t\}$.

Let $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \mathbb{N} = \cup\{0\}$. Let $\mathcal{U} = \bigcup_{n\geq 0} \mathbb{N}^n$ be the set of finite words on $\mathbb{N}$. The empty word, denoted by $\varnothing$, is the only element of $\mathbb{N}^0$. We see the elements of $\mathcal{U}$ as words on $\mathbb{N}$. For $u \in \mathbb{N}^n$ and $i \in \mathbb{N}$, we let ui denote the element of $\mathbb{N}^{n+1}$ obtained by appending i after u, so if $u = (u_1, u_2, \dots, u_k)$, $ui = (u_1, \dots, u_n, i)$. We see $\mathcal{U}$ as a tree rooted at $\varnothing$, where the natural genealogical order denoted by $\preceq$ is such that we have $u \preceq v$ if u is a prefix of v, potentially $u = v$. Similarly, we let $\mathcal{U}_2 = \bigcup_{n\geq 0}\{0, 1\}^n$.

3 Recursive convex minorants and their associated trees

3.1 Convex minorants of continuous functions

Let $D \subseteq \mathbb{R}$ be an interval containing 0 that will in general be $[0,1]$ or $\mathbb{R}_+$ in the sequel. Let $\mathcal{C}(D,\mathbb{R})$ be the set of continuous functions on D equipped with the uniform distance. For $\omega \in \mathcal{C}(D,\mathbb{R})$ such that $\omega(0) = 0$ and $x \in D$, the (greatest) convex minorant of ω on $[0,x]$ is the maximum convex function $c_x(\cdot,\omega)$ defined on $[0,x]$ such that $c_x(t,\omega) \leq \omega(t)$ for all $t \in [0,x]$. We let $\mathcal{V}_x(\omega) = \{t \in [0,x] : c_x(t,\omega) = \omega(t)\}$, and call the elements of $\mathcal{V}_x(\omega) \setminus \{x\}$ the vertices of the convex minorant of ω on $[0,x]$. Observe that $0 \in \mathcal{V}_x(\omega)$. We shall see shortly why the extremity x ought to be treated differently.

Up to now, the literature has mainly focused on properties of the convex minorant of a function on a fixed domain (for a fixed x). Of prime importance to us, is instead the structure of the different convex minorants $c_x(\cdot,\omega)$ of a fixed function ω as $x \in D$ varies. We start with the following straightforward (deterministic) geometric observation:

Lemma 3.1. *Let $\omega \in \mathcal{C}(D,\mathbb{R})$ be such that $\omega(0) = 0$ and let $x, y \in D$ with $0 \leq y < x$. For any $t \in \mathcal{V}_x(\omega) \cap [0,y]$, we have $\mathcal{V}_y(\omega) \cap [0,t] = \mathcal{V}_x(\omega) \cap [0,t]$.*

In other words, traversing them from the left to the right, the convex minorants $c_x(\cdot,\omega)$ and $c_y(\cdot,\omega)$ on $[0,x]$ and $[0,y]$ coincide on a non-empty closed interval, and then split for good. This induces a natural branching structure for $\{\mathcal{V}_x(\omega), x \in D\}$ that is depicted in Figure 1 that is central to the paper. This also justifies that for $t \in \mathcal{V}_x(\omega)$, the slope of the convex minorant $c_x(\cdot,\omega)$ to the left of t, defined by

$$\boldsymbol{s}(t,\omega) = \sup\left\{\frac{c_x(t) - c_x(t-s)}{s} : t - s \in \mathcal{V}_x(\omega)\right\},$$

is well-defined intrinsically since for any x' such that $t \in \mathcal{V}_{x'}(\omega)$ would yield the same value. If $t \in \mathcal{V}_x(\omega)$ for some x, let

$$\boldsymbol{r}(t,\omega) = \inf\{s > t : \omega(s) \leq \omega(t) + \boldsymbol{s}(t,\omega)(s-t)\}.$$

We call $\boldsymbol{r}(t,\omega)$ the *intercept associated to* t; this is defined independently of the choice of x for which $t \in \mathcal{V}_x(\omega)$. (The notation $\boldsymbol{r}(t,\omega)$ comes from "right".) The following is clear by construction:

Lemma 3.2. *Suppose that $t \in \mathcal{V}_x(\omega)$ for some $x \in D$.*
i) If $y \in [t, \boldsymbol{r}(t,\omega)]$, then $t \in \mathcal{V}_y(\omega)$.
i) If $y < t$ or $y > \boldsymbol{r}(t,\omega)$ then $t \notin \mathcal{V}_y(\omega)$.

Observe that if $\mathscr{L}(\omega)$ denotes the set of local minima of ω, then $\boldsymbol{s}(t,\omega)$ and $\boldsymbol{r}(t,\omega)$ are well-defined for every $t \in \mathscr{L}(\omega) \setminus \{0\}$.

3.2 Convex minorants of Brownian paths

We are interested in convex minorants of various Brownian-like paths such as Brownian motion or the Brownian excursion, the latter being more essential because of classical path decompositions. Such convex minorants have been studied for instance by Groeneboom [38] and Pitman [59]; in the following, we will mostly rely on the work of Pitman and Ross [58] that provides means to do explicit calculations; more information about related studies and references can be found there. We therefore now focus on these cases.

Since we should focus on the structure of $\mathcal{V}_x(\omega)$ as x varies, the following lemma is crucial.

Lemma 3.3 (No exceptional point)**.** *There exists a Borel set $\Omega^\star$ of $\mathcal{C}(\mathbb{R}_+,\mathbb{R})$ with $\mathbf{W}(\Omega^\star) = 1$, such that if $\omega \in \Omega^\star$, then for every $x \in \mathbb{R}_+$, it holds that:*
i) $\mathcal{V}_x(\omega)$ is countable;

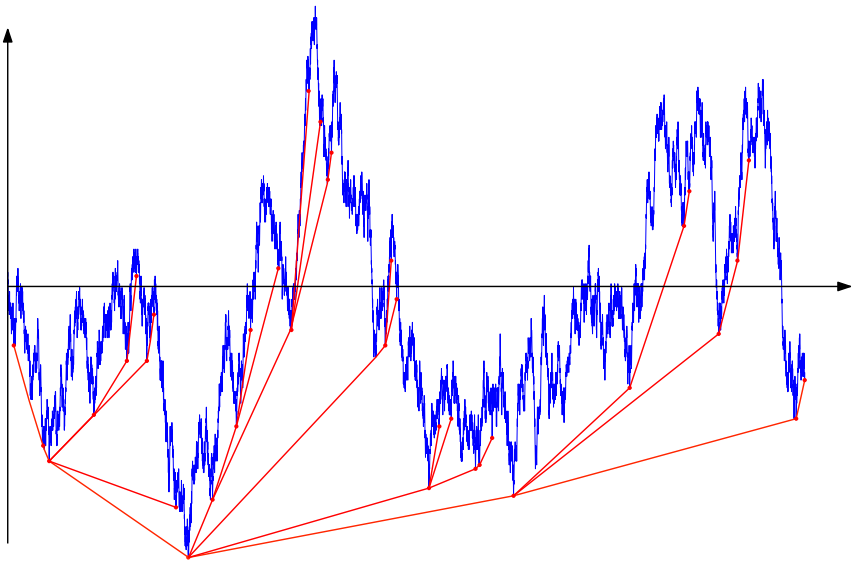

Figure 1: *The tree-like structure of convex minorants: in red, portions of the convex minorants of a fixed given function on finitely many intervals* $[0, x]$*, for* $x \in \mathbb{R}_+$*.*

ii) $\mathcal{V}_x(\omega)$ *has no accumulation point in* $(0, x)$*;*
iii) the elements of $\mathcal{V}_x(\omega) \setminus \{0, x\}$ *are all local minima;*
iv) the slopes $\mathfrak{s}(t, \omega)$ *at the points* $t \in \mathcal{V}_x(\omega) \setminus \{0, x\}$ *are all distinct.*

Remark 3.4. Note that a version of Lemma 3.3 holds for a Brownian excursion e on $[0, 1]$ instead of a Brownian motion: in this case, the claims in *i)–iv)* also hold, even with $[0, x)$ instead of $(0, x)$ in *ii)*.

Let $(W_t)_{t\geq 0}$ be a Brownian motion on $\mathbb{R}_+$. The results of [38, 58] imply that for W, the set of exceptional points x, for which one of the properties in *i)–iii)* might fail has Lebesgue measure zero. We verify that with probability one, there is no exceptional point by showing that if there were an exceptional point, then a.s. the set of such points would be of positive Lebesgue measure. In the following, we drop the dependence in W.

Proof of Lemma 3.3. *i)* Suppose that, with positive probability, there is some x is such that $\mathcal{V}_x$ is uncountable. Since there are countably many intervals $[a, b]$ with rational endpoints $0 \leq a < b < x$, one of them must be such that $\mathcal{V}_x \cap [a, b]$ is uncountable. By Lemma 3.1, for any point $y \in [b, x]$ we have $\mathcal{V}_x \cap [a, b] \subseteq \mathcal{V}_y$, so that the set of exceptional points would have positive Lebesgue measure.

ii) Suppose that, with positive probability, there exists some $x \in \mathbb{R}_+$ such that $\mathcal{V}_x$ has a an accumulation point y in $(0, x)$. By construction, for any $w \in (y, x]$, we have $\mathcal{V}_x \cap [0, y] \subseteq \mathcal{V}_w \cap [0, y]$, so that the set of points w for which *ii)* fails has positive Lebesgue measure, a contradition.

iii) Finally, suppose that there is some $x \in \mathbb{R}_+$ and $t \in \mathcal{V}_x \setminus \{0, x\}$ which is not a local minimum. Then, by Lemma 3.1, for any $y \in [t, x]$ we have $t \in \mathcal{V}_y$, and the proof is complete since $t < x$.

iv) If the slopes at t and t' such that $t' < t < x$ are identical, then the same holds for the convex minorants c_y on the intervals $[0, y]$ for every $y \in [t, x]$, so that the set of exceptional points has positive Lebesgue measure. □

The typical situation is that $\mathcal{V}_x(W)$ has accumulation points at both 0 and x, but it may also happen that x is not an accumulation point: this happens for instance when x is a local minimum or $x = \mathfrak{r}(t, W)$ for some local minimum t. Let $(t_i)_{i\in\mathbb{Z}} = (t_i(x))_{i\in\mathbb{Z}}$ denote the vertices in $\mathcal{V}_x(W) \setminus \{x\}$, indexed in such a way that $t_i \leq t_{i+1}$ and $t_0 = \arg\min\{W_s : s \in [0, x]\}$. In the case where x is not an accumulation point of $\mathcal{V}_x(W)$, it is understood that the sequence is only defined for $i \leq k$ for some $k \geq 0$. The intervals

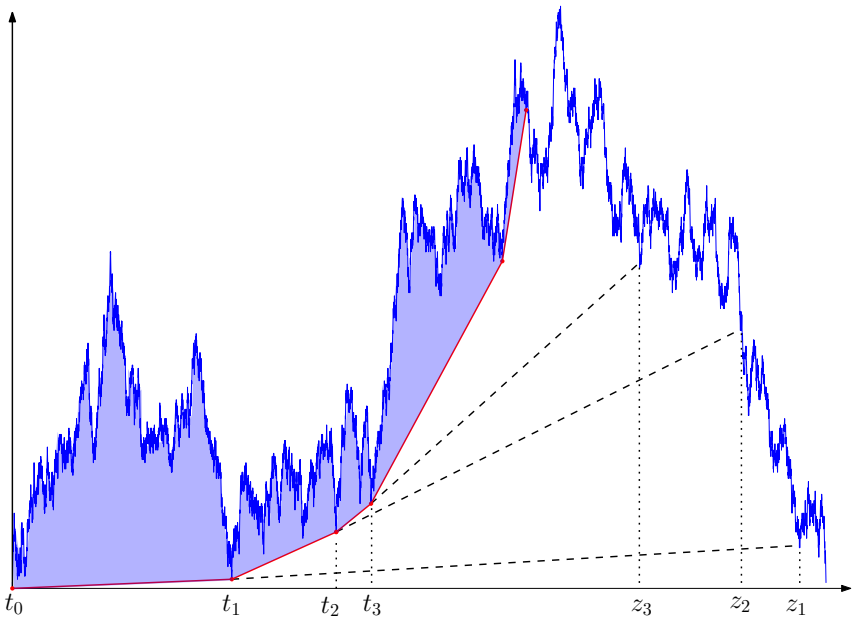

Figure 2: *The greatest convex minorant of a Brownian excursion* e *on some interval* $[0, x]$*, the sequence of vertices* $t_0 = 0 < t_1 < \cdots < x$ *and intercepts* $z_1 > z_2 > \cdots > x$*. The distribution of these quantities are given in Lemmas 3.6 and 3.8 when* x *is uniform in* $[0, 1]$.

$[t_i, t_{i+1}]$ where the slope of $c_x(W)$ is constant are called the *faces* of the convex minorant. Let γ_i denote the slope of the convex minorant on $[t_i, t_{i+1}]$, and z_i be the intercept associated to t_i:

$$\gamma_i = \boldsymbol{s}(t_{i+1}, W) = \frac{W(t_{i+1}) - W(t_i)}{t_{i+1} - t_i} \quad \text{and} \quad z_i = \boldsymbol{r}(t_i, W) = \inf\{s > t_i : W_s \leq W_{t_i} + \gamma_{i-1}(s - t_i)\}\,.$$

It is possible that $z_i = \infty$ for some $i \leq 0$, but a.s. $z_i < \infty$ for $i \geq 1$. For every $i \geq 1$, such that $z_i < \infty$, and for $s \geq 0$ let

$$g_i(s) := (W_{t_i+s} - W_{t_i} - \gamma_i s)\mathbf{1}_{t_i+s \leq t_{i+1}} \quad \text{and} \quad h_i(s) = (W_{t_i+s} - W_{t_i} - \gamma_{i-1} s)\mathbf{1}_{t_i+s \leq z_i}\,. \tag{4}$$

Let $\mathbf{n}_\sigma$ be the law of a Brownian excursion of duration $\sigma > 0$. The following decomposition lemma is straighforward from Theorem 2.2 of [38] (see also Theorem 2 and Corollary 2 of [58]):

Lemma 3.5. *For any* $i \in \mathbb{Z}$ *such that* $z_i < \infty$*, conditionally on* $(t_j, \gamma_j)_{j<i}$*, and* (t_i, z_i)*, the collection of functions* g_j*,* $j < i$*, and* h_i *form an independent family with law given respectively by* $\mathbf{n}_{t_{j+1}-t_j}$*,* $j < i$*, and* $\mathbf{n}_{z_i - t_i}$.

Together with the previous considerations about the decomposition, we are thus let to studying convex minorants of Brownian excursions, which is the subject of the next section.

3.3 Convex minorants of a Brownian excursion

In this section, we consider a Brownian excursion e on $[0, 1]$. We use the notation of the previous section with $\omega = \mathrm{e}$, up to the obvious modifications: The vertices of the convex minorant of e on $[0, x]$ are denoted by $\mathcal{V}_x = \mathcal{V}_x(\mathrm{e})$, and can be enumerated in increasing order as $(t_i)_{i \geq 0}$ with $t_0 = 0$, where $t_i = t_i(x) = t_i(x, \mathrm{e})$. The slopes $\gamma_i = \gamma_i(x, \mathrm{e})$ are defined as before. For each $i \geq 1$, let $z_i = z_i(x) = z_i(x, \mathrm{e}) = \inf\{s > t_i : \mathrm{e}_s = \mathrm{e}_{t_i} + \gamma_{i-1}(s - t_i)\}$. We define $z_0 = 1$ for convenience. See Figure 2.

A simple induction yields the description of the restriction of the excursion e to the interval $[0, x]$ as a collection of Brownian excursions above the graph of $c_x(\cdot, \mathrm{e})$.

Lemma 3.6. *Let* e *be a standard Brownian excursion on* $[0,1]$*, and let* $x = V$ *be an independent random variable uniform on* $[0,1]$*. Then consider the convex minorant* $c_V(\cdot, \mathrm{e})$ *of* e *on* $[0, V]$*, with vertices* $(t_i)_{i\geq 0}$*. Define the functions* e_0 *and* e_1 *by*

$$\mathrm{e}_0(s) := (\mathrm{e}(s) - s \cdot \mathrm{e}(t_1))\mathbf{1}_{s\leq t_1} \qquad \text{and} \qquad \mathrm{e}_1(s) := (\mathrm{e}(t_1+s) - \mathrm{e}(t_1) - s \cdot \mathrm{e}(t_1))\mathbf{1}_{t_1+s\leq z_1}\,.$$

Then $(t_1, z_1 - t_1, 1 - z_1)$ *is a Dirichlet*$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ *random vector, and conditionally on* (t_1, z_1)*,* $(\mathrm{e}_0, \mathrm{e}_1, V)$ *are independent,* e_0 *and* e_1 *are Brownian excursions of durations* t_1 *and* $z_1 - t_1$*, respectively, and* V *is uniform on* (t_1, z_1)*.*

Proof. The claimed properties follow from the decomposition of the Brownian excursion e using the line linking $(0,0)$ to $(t_1, \mathrm{e}(t_1))$ (notice that $0 < t_1 < 1$ a.s.). For $s \geq 0$, consider the straight line $\{(t, st) : t \in [0,1]\}$, and increase the value of s from 0 until the first value $s = \gamma_1$ at which the location of the first intersection $i(s) = \inf\{t \geq 0 : \mathrm{e}(t) = st\}$ is at most V: $\gamma_1 = \inf\{s \geq 0 : i(s) \leq V\}$, $t_1 = i(\gamma_1)$. Now, by strong Markov property, $(\mathrm{e}_{1-t}, z_1 \leq t \leq 1)$ is a Brownian meander of duration $1 - z_1$ conditioned to end at e_{z_1}; the path e_0 coincides with the path above the first face, and is a Brownian excursion; the path e_1 is the path above the same line, between t_1 and z_1. Let $\varphi_t(x) = e^{-x^2/(2t)}/\sqrt{t2\pi}$. The vector (γ_1, t_1, z_1, V) has a distribution which is absolutely continuous with respect to Lebesgue measure on the set $D := \mathbb{R}_+ \times \{(t,z,u) : 0 < t < v < z < 1\}$, with density $f(s,t,z,v)$ given by

$$\begin{aligned} f(s,x,z,v) &= \mathbf{1}_{(s,t,z,v)\in D} \cdot \frac{\varphi_t(ts)}{t} \cdot \frac{\varphi_{z-t}((z-t)s)}{z-t} \cdot \frac{\varphi_{1-z}(zs)zs}{1-z} \cdot t\sqrt{2\pi} \\ &= \mathbf{1}_{(s,t,z,v)\in D} \cdot \frac{sze^{-\frac{s^2}{2(1-z)/z}}}{1-z} \times \frac{1}{2\pi\sqrt{t(z-t)(1-z)}} \times \frac{1}{z-t}\,. \end{aligned}$$

Integrating for $s \in \mathbb{R}_+$, this yields the claimed distribution. □

Remark 3.7. We point out that the distribution of $(t_1, z_1 - t_1, 1 - z_1)$ may also be obtained, without any calculation, using the correspondence with the cut tree and the decomposition of a Brownian continuum random tree into tree pieces that is induced by removing the branch point at the intersection of the geodesics between three random points (see Section 4.2 and [13]).

A straightforward induction yields the distribution of the vector of lengths of the faces of the convex minorant $c_V(\cdot, \mathrm{e})$ of e on $[0, V]$ for an independent uniform point V in $[0,1]$:

Lemma 3.8. *Let* e *be a standard Brownian excursion on* $[0,1]$*, and let* $x = V$ *be an independent random variable uniform on* $[0,1]$*. Let* $(\Delta_{i,1}, \Delta_{i,2}, \Delta_{i,3})_{i\geq 0}$ *denote a family of independent Dirichlet*$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ *random vectors. Then, for* $(t_i)_{i\geq 0} = (t_i(V))_{i\geq 0}$ *the sequence of vertices, we have*

$$(t_i - t_{i-1})_{i\geq 1} \overset{d}{=} \left(\Delta_{i,1} \cdot \prod_{1\leq j<i} \Delta_{j,2}\right)_{i\geq 1}.$$

3.4 Recursive convex minorants of a Brownian excursion

The results of the previous section point out the recursive structure of convex minorants of a Brownian excursion. Here we will use it to construct the tree $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$. This is the first building block of our construction of $\mathrm{CMT}(X, \mathbf{U})$ the scaling limit of the minimum spanning tree, and it already reveals some of the main ingredients. Before proceeding to the details, let us explain roughly the strategy:

- for all $x, y \in [0,1]$, we define the set $[\![x, y]\!]$ which is meant to be the collection of points used to go from x to y (somewhat pre-arcs or pre-branches);
- we also show that it is possible to assign a "measure" $d(x,y)$ to $[\![x, y]\!]$ that induces a 0-hyperbolic metric space.

We will see that the metric space induced by d on $[0,1]$ is connected if we restrict our attention to points at finite distance from 0, so that the subset of $[0,1]$ with this property, endowed with d is thus an $\mathbb{R}$-tree (in the sense of Section 2.2 of [4]). Later on, we will show that the metric completion of $([0,1],d)$ is compact, so that no point is put aside.

The definition of $[\![x,y]\!]$ will be done in stages: first $[\![0,x]\!]$ with x restricted to some suitable dense subset of $[0,1]$; then, we extend the definition of $[\![0,x]\!]$ to all $x \in [0,1]$; finally, $[\![x,y]\!]$ is defined in Section 3.5 using a notion of common ancestor of x and y.

Remark 3.9. Let $\mathbf{n}_\sigma$ denote the law of a standard Brownian excursion of duration $\sigma > 0$. We will define $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ as a proper random variable for $\mathbf{n}_1$-almost every function ω, and almost all sequences $\mathbf{U} = (U_1, U_2, \dots)$ of independent random variables, uniform on $[0,1]$. For this, the components of $\mathbf{U}$ are associated to the local minima of e. This can be done by defining a canonical bijection between $\mathbb{N}$ and the set $\mathscr{L}(\mathrm{e})$ of local minima of e. For instance, consider an enumeration $I = (I_j, j \geq 0)$ of the (countable) set of all intervals with rational extremities on $[0,1]$. Since a.s. each local minimum of e is a global minimum on at least one interval of I, associate with each local minimum $t \in \mathscr{L}$ the index $j(t)$ of the first interval of I on which t is a global minimum; after that associate with t, the uniform random variable $U_{j(t)}$. In the sequel, j is called the association map of e. The proofs of convergence in Section 7 will need a different, more complex association, but we believe it is not necessary until then.

Let $\mathcal{U} = \bigcup_{n\geq 0} \mathbb{N}^n$ (see Section 2). For any $x \in \mathscr{L} = \mathscr{L}(\mathrm{e})$, we define recursively a collection $(t_u, \xi_u, \gamma_u, e_u)$, $u \in \mathcal{U}$, that a priori depends on x. Lemma 3.3 ensures that, with probability one, the following definition makes sense for all $x \in \mathscr{L}$.

We first let $t_\varnothing = 0$, $\xi_\varnothing = x$, $\gamma_\varnothing = 0$, and $e_\varnothing = \mathrm{e}$. Almost surely, there are only finitely many vertices of the convex minorant of e on $[0,x]$, and they are all elements of $\mathscr{L}$ and denoted by $t_0 = 0 < t_1 < t_2 < \cdots < t_k = x$ for some $k \in \mathbb{N}$. For each $i = 0, \dots, k-1$, let $\xi_i = t_i + U_{t_{i+1}}|t_{i+1} - t_i|$, $\gamma_i = \boldsymbol{s}(t_{i+1}, \mathrm{e}) = (\mathrm{e}(t_{i+1}) - \mathrm{e}(t_i))/|t_{i+1} - t_i|$ and let e_i be defined by, for $s \geq 0$,

$$e_i(s) = (\mathrm{e}(t_i + s) - \mathrm{e}(t_i) - s \cdot \gamma_i)\mathbf{1}_{t_i + s \leq t_{i+1}}\,.$$

More generally, suppose now that we have defined $(t_u, \xi_u, \gamma_u, e_u)$ for some $u \in \mathbb{N}^n$, $n \geq 1$. Let $\theta^u_0 = 0 < \theta^u_1 < \theta^u_2 < \dots$ be the vertices of the convex minorant of e_u on the interval $[0, \xi_u - t_u]$, and set $t_{ui} = t_u + \theta^u_i$ for all $i \geq 0$; observe that the $t_{ui} = t_u + \theta^u_i$, $i \geq 0$, are precisely the elements of $(t_j(\xi_u))_{j\geq 0}$ lying in $[t_u, \xi_u]$. Then let $\varphi^u_i = (e_u(\theta^u_{i+1}) - e_u(\theta^u_i))/|\theta^u_{i+1} - \theta^u_i|$ be the slope of the convex minorant of e_u on $[\theta^u_i, \theta^u_{i+1}]$. For each $i \in \mathbb{N}$, we let $m_{ui} = |\theta^u_{i+1} - \theta^u_i| = |t_{u(i+1)} - t_{ui}|$, $\xi_{ui} = t_{ui} + U_{t_{u(i+1)}} m_{ui}$, $\gamma_{ui} = (e(t_{u(i+1)}) - e(t_{ui}))/m_{ui} = \gamma_u + \varphi^u_i$ and define the function $e_{ui} : [0, m_{ui}] \to \mathbb{R}_+$ by

$$e_{ui}(s) = (\mathrm{e}(t_{ui} + s) - \mathrm{e}(t_{ui}) - s \cdot \gamma_{ui})\mathbf{1}_{t_{ui} + s \leq t_{u(i+1)}}\,.$$

We then define

$$[\![0,x]\!] := \{x\} \cup \bigcap_{n\geq 0} \overline{\bigcup_{|u|=n} [t_u, \xi_u]}\,, \tag{5}$$

which is then a non-empty closed subset of $[0,x]$. For each $n \geq 1$, we also let

$$d_n(0,x) := \sqrt{\frac{\pi}{2}} \cdot \sum_{|u|=n} m_u^{1/2} \qquad \text{and} \qquad d(0,x) := \limsup_{n\to\infty} d_n(0,x)\,. \tag{6}$$

Lemma 3.10. *For each $x \in \mathscr{L}$, the sequence $(d_n(0,x))_{n\geq 1}$ is a non-negative martingale. As a consequence, with probability one, the sequences $d_n(0,x)$ converge for all $x \in \mathscr{L}$ to finite limits $d(0,x)$.*

Proof. Fix $x \in \mathscr{L}$. Let $\mathcal{F}_n$ denote the sigma-algebra generated by the random variables $\{(t_u, \gamma_u), |u| \leq n\}$; in particular, $(m_u)_{|u|\leq n}$ is $\mathcal{F}_n$-measurable. Conditionally on $\mathcal{F}_n$, the functions e_u, $u \in \mathbb{N}^n$, are independent Brownian excursions of respective durations m_u. It follows that

$$d_{n+1}(0,x) = \mathbf{E}\left[\sum_{|u|=n+1} m_u^{1/2} \,\middle|\, \mathcal{F}_n\right] = \sum_{|u|=n} m_u^{1/2} \cdot \mathbf{E}\left[\sum_{i\geq 0} (m_{ui}/m_u)^{1/2} \,\middle|\, \mathcal{F}_n\right]\,. \tag{7}$$

Let $(\Delta_{i,1}, \Delta_{i,2}, \Delta_{i,3})_{i\geq 0}$ be iid Dirichlet$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ random vectors. Then, by Lemma 3.8, we have conditionally on m_u:

$$\left(\frac{m_{ui}}{m_u}\right)_{i\geq 0} \stackrel{d}{=} \left(\prod_{1\leq j<i} \Delta_{j,2}\cdot\Delta_{i,1}\right)_{i\geq 0} \tag{8}$$

From there, it is straightforward to verify by induction that, since $\mathbf{E}[\Delta_{i,1}^{1/2} + \Delta_{i,2}^{1/2}] = 1$, for each $i \geq 1$, the expectation of the square root of the right-hand side of (8) equals 2^{-i}. As a consequence, the conditional expectations in the right-hand side of (7) all equal one almost surely, so that $d_n(0,x)$ is indeed a martingale. Since $\mathscr{L}$ is countable, the convergence is almost surely for all $x \in \mathscr{L}$. □

The function $d(0,\cdot)$ can be extented to $[0,1]$ as follows. First let $[\![0,0]\!] = \{0\}$ and $d(0,0) = 0$. Then, for each point $x \in (0,1]$, any $t \in \mathcal{V}_x \setminus \{0,x\}$ is a local minimum, and therefore $[\![0,t]\!]$ and $d(0,t)$ has already been defined in (5) and (6), respectively. We rely on those to define, for $x \in (0,1]$,

$$[\![0,x]\!] = \overline{\bigcup_{t\in\mathcal{V}_x\cap\mathscr{L}} [\![0,t]\!]} \cup \{x\} \qquad \text{and} \qquad d(0,x) = \sup_{t\in\mathcal{V}_x\cap\mathscr{L}} d(0,t) \in \mathbb{R}_+ \cup \{\infty\}\,. \tag{9}$$

Remark 3.11. **i)** The slight subtlety in the definition in (9), where the union is taken on $t \in \mathcal{V}_x \cap \mathscr{L}$ rather than $\mathcal{V}_x \setminus \{x\}$ is to ensure that the definion in (9) is consistent with the one in (5) in the case that $x \in \mathscr{L}$. For instance, if $t \in \mathscr{L}$ and $x = \boldsymbol{r}(t)$ then $x \in \mathcal{V}_x$ but almost surely not in $\mathscr{L}$.
ii) The recursive construction yields a collection of "join points" associated to the local minima. First there is a well-defined face to the left of t: almost surely for $t \in [0,1]$, $\boldsymbol{\ell}(t) = \sup \mathcal{V}_t < t$, and $\boldsymbol{\ell}(t) \in \mathscr{L}$, so that $[\boldsymbol{\ell}(t), t]$ is a face of the convex minorant c_t. The slope $\boldsymbol{s}(t)$ is precisely the slope of this face. Furthermore, U_t is used to define a uniform random point in $[\boldsymbol{\ell}(t), t]$ that we denote by $\boldsymbol{j}(t)$. In the previous decomposition, for any $x \in [0,1]$ and any $u \in \mathcal{U}$, $i \in \mathbb{N}$ such that $t_{u(i+1)} = t$, we have $t_{ui} = \boldsymbol{\ell}(t)$ and $\xi_{ui} = \boldsymbol{j}(t)$.

Before going further, let us prove the following lemma, that will be useful later (Lemma 3.13). Observe first that (9) allows to extend the definition of $t_u(x), \xi_u(x), m_u(x)$, $u \in \mathcal{U}$, to all $x \in [0,1]$: let $t_u(x), \xi_u(x)$ and $m_u(x)$ coincide with $t_u(t_i), \xi_u(t_i), m_u(t_i)$ for all u of the form $u = jv$, with $j < i$ and $v \in \mathcal{U}$. When $\mathcal{V}_x$ is a finite set, it is understood $m_{jv}(x)$ is only defined for the relevant values of j.

Lemma 3.12. *Almost surely, for every $x \in [0,1]$ we have $\mathcal{V}_x \subseteq [\![0,x]\!]$ and furthermore:*
i) for every $u \in \mathcal{U}$, $t_u(x), \xi_u(x) \in [\![0,x]\!]$,
ii) $\sup_{|u|=n} m_u(x) \to 0$, as $n \to \infty$, and thus
iii) $[\![0,x]\!]$ is the closure of $\{t_u(x) : u \in \mathcal{U}\} \cup \{x\}$, in particular, if $x = \boldsymbol{r}(t)$, then $[\![0,x]\!] = [\![0,t]\!] \cup \{x\}$.

Proof. The first claim is clear from (9). We first prove *i)* for $x \in \mathscr{L}$. For $u \in \mathcal{U}$, $t_u \in [\![0,x]\!]$ by definition: indeed, for each $v \in \mathcal{U}$, $t_{v0} = t_v$, and thus $t_u \in \cap_{n\geq|u|} \cup_{|v|=n} [t_v, \xi_v]$. For ξ_u, note that, almost surely $\mathcal{V}_{\xi_u}$ has an accumulation point at ξ_u. It follows that ξ_u lies in the closure of $\{t_{uk} : k \geq 0\}$. By the previous argument, all these points lie in $[\![0,x]\!]$ which is closed, and thus $\xi_u \in [\![0,x]\!]$ as well. Now, since $\mathcal{U}$ is countable, this is true for every u, and, hence, for every $x \in \mathscr{L}$. Finally, this is true for all $x \in [0,1]$ by definition of $[\![0,x]\!]$ in (9).

ii) We restrict our attention to the set of probability one where $m_u(y) \to 0$ as $|u| \to \infty$ for all $y \in \mathscr{L}$. Fix any $x \in [0,1]$ and $\epsilon > 0$. There is an $i \in \mathbb{N}$ large enough that $\sup \mathcal{V}_x \cap \mathscr{L} \leq t_i + \epsilon$. Then, for any u of the form jv with $j \geq i$ and $v \in \mathcal{U}$ either $m_{jv}(x) \leq \epsilon$, or $m_{jv}(x)$ is not defined. On the other hand, for u of the form jv with $j < i$ and $v \in \mathcal{U}$, we have $m_{jv}(x) = m_{jv}(t_i)$. It follows that $\sup\{m_u(x) : |u| = n\} \leq \epsilon$, which completes the proof since $\epsilon > 0$ was arbitrary.

iii) follows readily from *i)*, *ii)* and the definition. □

3.5 The branching structure and the convex minorant tree

We now move on to the branching structure. Let $x, x' \in [0,1]$. With the ultimate objective of defining $d(x,x')$ we first define $x \wedge x' := \sup([\![0,x]\!] \cap [\![0,x']\!])$. It should be understood as the closest common ancestor of x and x', when 0 is seen as the root. It follows readily that the definition that the sets $[\![0,x]\!]$ enjoy the following restriction property:

Lemma 3.13. *Almost surely, for any $x \in [0,1]$ and $y \in [\![0,x]\!]$, we have $y = x \wedge y$, $[\![0,y]\!] = [\![0,x]\!] \cap [0,y]$, and $d(0,y) \leq d(0,x)$*

Proof. We restrict our attention to the set of probability one on which the events of Lemma 3.12 all occur. If $y = x$, the claim is clear, so suppose that $y < x$, which implies that $y \leq \sup(\mathcal{V}_x \cap \mathscr{L})$. If $y = \sup(\mathcal{V}_x \cap \mathscr{L})$ then $[\![0,y]\!]$ is the closure of $\bigcup_{i\geq 1} [\![0,t_i(x)]\!]$, so that the claim holds by (9). Finally consider the last case $y < \sup(\mathcal{V}_x \cap \mathscr{L})$, and let $t_u = t_u(x)$, and $m_u = m_u(x)$, $u \in \mathcal{U}$, defined in the previous section. For any $n \geq 1$ there exists some $u \in \mathcal{U}$ with $|u| = n$ such that $y \in [t_u, t_u + m_u]$. It follows easily that $[\![0,x]\!] \cap [0,t_u] \subseteq [\![0,y]\!]$. Since $m_u \to 0$ by Lemma 3.12, we have $[\![0,y]\!] = [\![0,x]\!] \cap [0,y]$. The claim about the distance follows readily. □

The extension of $[\![\cdot,\cdot]\!]$ and $d(\cdot,\cdot)$ to $[0,1]^2$ will require the following lemma, that will allow us to bring the (nice) points of $\mathscr{L}$ back in the game:

Lemma 3.14. *With probability one, for every $x, y \in [0,1]$ with $x > y$, there exists some $t \in \mathcal{V}_x \cap \mathscr{L}$ such that $x \wedge y = t \wedge y \in [\![0,t]\!]$.*

Proof. We work on set $\Omega^\star$ of probability one where all the events of Lemma 3.3 all occur. Since $x > y$, let $(t_i)_{i\geq 0}$ be the vertices of $\mathcal{V}_x \setminus \{x\}$, which might be a finite sequence. Then $[t_i, t_{i+1})$, $i \geq 0$, together with $[\sup_i t_i, x)$ forms a partition of $[0,x)$. On the event $\Omega^\star$, it suffices to consider the following two cases. (a) If $\sup_i t_i = x$, then there exists some $i \in \mathbb{N}$ for which $y \in [t_i, t_{i+1})$. By definition, $x \wedge y = t_{i+1} \wedge y \in [\![0,t_{i+1}]\!]$, and $t_{i+1} \in \mathscr{L}$. (b) Otherwise there are only finitely many vertices $t_0, t_1, \ldots, t_k$, all of which are in $\mathscr{L}$; then since $y < x$, we have $x \wedge y = t_k \wedge y \in [\![0,t_k]\!]$. □

The following lemma makes formal the branching structure of the sets $[\![0,x]\!]$, $x \in [0,1]$.

Lemma 3.15. *There exists a set of probability one on which for any $x, x' \in (0,1)$, $x \wedge x' > 0$ and*

$$[\![0,x]\!] \cap [0, x \wedge x'] = [\![0,x']\!] \cap [0, x \wedge x'] \qquad \text{and} \qquad [\![0,x]\!] \cap [\![0,x']\!] \cap (x \wedge x', 1] = \varnothing\,.$$

Proof. We work on a set $\Omega^\star$ of probability one where the events of Lemmas 3.3 and 3.12 all occur. Let $(t_u)_{u\in\mathcal{U}}$ and $(t'_u)_{u\in\mathcal{U}}$ denote the recursive collections of points introduced before, for the points x and x', respectively. For $n \geq 1$, define

$$x \wedge_n x' := \sup\{\{t_u : |u| \leq n\} \cap \{t'_u : |u| \leq n\}\} \geq 0\,.$$

Then Lemma 3.12 *iii)* implies that for every $n \geq 1$ and all $y < x \wedge_n x'$, we have $y \in [\![0,x]\!]$ if and only if $y \in [\![0,x']\!]$. The sequence $(x \wedge_n x')_{n\geq 1}$ is non-decreasing and taking the limit as $n \to \infty$, it follows that $[\![0,x]\!]$ and $[\![0,x']\!]$ coincide on $[0, \sup_n x \wedge_n x')$.

On the other hand, by definition of $x \wedge x' = \sup [\![0,x]\!] \cap [\![0,x']\!]$, the sets $[\![0,x]\!]$ and $[\![0,x']\!]$ are disjoint on $(x \wedge x', \infty)$. So to complete the proof, it suffices to prove that $\sup_n x \wedge_n x' = x \wedge x'$. For every $n \geq 1$, we have $x \wedge_n x' \in [\![0,x]\!] \cap [\![0,x']\!]$ so that $x \wedge_n x' \leq x \wedge x'$. To prove the converse inequality, consider an arbitrary point $y \in [\![0,x]\!] \cap [\![0,x']\!]$, and observe that for any $n \geq 1$, there exists $u, v \in \mathcal{U}$ with $|u| = |v| = n$ such that $y \in [t_u, \xi_u]$ and $y \in [t'_v, \xi'_v]$, and necessarily $t_u, t'_v \leq x \wedge_n x'$. It follows that

$$y \leq x \wedge_n x' + \sup\{m_u(x) : |u| = n\}\,.$$

It follows that from Lemma 3.12 that $y \leq \sup_n x \wedge_n x'$. Since $y \in [\![0,x]\!] \cap [\![0,x']\!]$ was arbitrary, we may take it as close to $x \wedge x'$ as we want, which proves that $\sup_n x \wedge_n x' = x \wedge x'$.

Finally, we show that $x \wedge x' > 0$. Without loss of generality, we assume that $x' < x$. If $x' \geq t_1$, then $x \wedge x' \geq t_1 > t_0 = 0$. More generally, for any $n \geq 1$, if $x' \geq t_{0^{(n)}1}$, then $x' \wedge x \geq t_{0^{(n)}1}$, where $0^{(n)}1$ is the sequence formed by n consecutive 0 followed by a 1. But for every $n \geq 1$, $t_{0^{(n)}1}$ is distributed like $m_1 \times \prod_{2 \leq i \leq n} \Delta_i$, where $(\Delta_i)_{i \geq 2}$ is a family of i.i.d. random variables with distribution Beta$(\frac{1}{2}, 1)$; as a consequence, $t_{0^{(n)}1} > 0$ a.s. for every $n \geq 1$ and $t_{0^{(n)}1} \leq \sup\{m_u(x) : |u| = n+1\} \to 0$ by Lemma 3.12. Since $x' > 0$, there is some $n \geq 1$ for which $0 < t_{0^{(n)}1} \leq x'$ which proves that $x \wedge x' > 0$. The latter decomposition depends on x, but either $x' \geq t_1(x) > 0$, or $x \wedge x' = t_1(x) \wedge x'$ so that it suffices to consider the decomposition at the set of local minima, which is countable; it follows that, almost surely, for every $x, x' \in (0,1)$, $x \wedge x' > 0$. □

We are now ready to define $[\![x, y]\!]$ and $d(x, y)$ for all $x, y \in [0, 1]$. Observe first that, by Lemma 3.14, almost surely, for all $x \neq y$, we have $x \wedge y \in [\![0, t]\!]$ for some $t \in \mathscr{L}$, so that $d(0, x \wedge y) < \infty$. Now, if $x = y$, we set $d(x, y) = 0$, and otherwise

$$d(x,y) := d(0,x) + d(0,y) - 2d(0, x \wedge y) \qquad \text{and} \qquad [\![x,y]\!] := ([\![0,x]\!] \cup [\![0,y]\!]) \cap [x \wedge y, 1]\,. \quad (10)$$

By the previous remark, both $d(\cdot,\cdot)$ and $[\![\cdot,\cdot]\!]$ are well-defined and symmetric on $[0,1]^2$. When necessary, we write $[\![x,y[\![= [\![x,y]\!] \setminus \{y\}$; $]\!]x,y]\!]$ and $]\!]x,y[\![$ are defined similarly.

Finally, we verify now that d induces a metric space that has the topology of a tree. In the following, we let $(x \cdot y)_0 := \frac{1}{2}(d(0,x) + d(0,y) - d(x,y))$. Observe that, by definition, we have $(x \cdot y)_0 = d(0, x \wedge y)$.

Lemma 3.16 (Triangle inequality and four-point condition)**.** *A.s., for every $x, y, z \in [0,1]$, we have*

i) $0 \leq d(x,y) \leq d(x,z) + d(z,y)$*, and*

ii) $(x \cdot y)_0 \geq \min\{(x \cdot z)_0, (z \cdot y)_0\}$.

Proof. We prove *i)* and *ii)* simultaneously. Note that $d(x,y) \geq 0$ by Lemma 3.13. By definition, $x \wedge z \in [\![0, x]\!]$. Suppose first that $x \wedge z \in [\![0, x \wedge y[\![$. Then, $z \wedge y \in [\![0, x \wedge y[\![$ as well by Lemma 3.15. It follows readily that $d(0, x \wedge y) \geq d(0, x \wedge z), d(0, y \wedge z)$. Furthermore, by definition,

$$d(x,y) = d(x, x \wedge y) + d(x \wedge y, y) \leq d(x, x \wedge z) + d(x \wedge y, y) \leq d(x,z) + d(z,y)\,.$$

If on the other hand, we have $x \wedge z \in [\![x \wedge y, x]\!]$, then Lemma 3.15 implies that $z \wedge y = x \wedge y$. As a consequence, we have $d(0, x \wedge y) = d(0, z \wedge y) = \min\{d(0, x \wedge z), d(0, z \wedge y)\}$. Moreover

$$d(x,y) = d(x, x \wedge y) + d(x \wedge y, y) = d(x, x \wedge z) + d(x \wedge z, x \wedge y) + d(x \wedge y, y)\,, \quad (11)$$

and

$$\begin{aligned} d(x,z) + d(z,y) &= d(x, x \wedge z) + d(x \wedge z, z) + d(z, y \wedge z) + d(y \wedge z, y) \\ &= d(x, x \wedge z) + d(x \wedge z, z) + d(z, x \wedge y) + d(x \wedge y, y)\,, \end{aligned}$$

which is easily seen to be at least as large as the right-hand side of (11). □

By Lemma 3.16, d satisfies the triangle inequality and thus induces a metric on the quotient space: Let $x \sim y$ if $d(x,y) = 0$. Let $\mathscr{T}^\circ := \{x \in [0,1] : d(0,x) < \infty\}$, and write $\mathscr{T}$ for the metric completion of the quotient $\mathscr{T}^\circ/_\sim$; we still write d for the induced metric on $\mathscr{T}$. Writing π for the canonical projection, we let $\rho = \pi(0)$ be the root of $\mathscr{T}$ and μ be the push-foward of the Lebesgue measure on $[0,1]$ by π. We define $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ as $\mathfrak{T} := (\mathscr{T}, d, \mu, \rho)$.

We will later on identify exactly the distribution of $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ the Brownian CRT (Theorem 1.3); however since the proof requires to introduce a number of additional concepts, it is interesting to first verify that:

Proposition 3.17. *With probability one, the metric space $(\mathscr{T}, d)$ is a real tree.*

Remark 3.18. i) We define $\mathscr{T}^\circ$ to ensure that $(\mathscr{T}, d)$ is connected. We will see later that a.s. $\mathscr{T}^\circ = [0,1]$. **ii)** It is plausible that $(\mathscr{T}^\circ, d)$ is already complete; we do not have a short argument for either direction, and we did not try to investigate further since there is no real influence on what follows.

Proof of Proposition 3.17. By the four-point condition in Lemma 3.16 *ii)* and Lemma 3.10 of [32], $\mathscr{T}$ is 0-hyperbolic [see also 27]. Then, by Theorem 3.40 of [32], it suffices to prove that $\mathscr{T}$ is connected to complete the proof.

We show that $\mathscr{T}$ is path-connected; this relies on the fact, proved in Section 3.6, that there exists a measure ℓ on $\mathscr{T}$ such that for all $x, y \in [0,1]$ we have $d(x,y) = \ell([\![x,y]\!])$. Let π denote the canonical projection from $[0,1]$ onto $\mathscr{T}$. For any $r \in [0, d(0,x)]$, let $x_r := \sup\{s \in [\![0,x]\!] : d(0,s) \leq r\}$. Then we claim that the map ϕ given by $\phi(r) = \pi(x_r)$ is an isometry from $[0, d(0,x)]$ to $\mathscr{T}$. To see this, note first that since $d(0,s)$ is non-decreasing for $s \in [\![0,x]\!]$, and the set $\{s \in [\![0,x]\!] : d(0,s) \leq r\}$ is closed, we have $d(0, \phi(r)) = d(0, x_r) \leq r$. On the other hand, for $s \in [\![0,x]\!]$, we have $d(0,s) = \ell([\![0,s]\!]) = \ell([\![0,x]\!] \cap [0,s])$; since $\ell([\![0,x]\!]) < \infty$ the right-hand side is continuous if we consider $s \in [0,1]$. It follows that $d(0, \phi(r)) = r$, and that ϕ is an isometry. Therefore, for every x, there is a geodesic from 0 to x, and $\mathscr{T}$ is path-connected and then connected. □

The following consistency property will be useful. It implies in particular that the pairwise distances may be defined using only certain suitable sub-excursions of e.

Lemma 3.19 (Restriction and consistency)**.** *For $x \in [0,1]$, let $(t_i)_{i \geq 0}$ denote the vertices of $\mathcal{V}_x \cap \mathscr{L}$, and $z_i = z_i(x)$, $i \geq 0$, the corresponding intercepts. Then*

i) for any $i \geq 0$, we have

$$[\![0,x]\!] = \bigcup_{0 \leq j < i} [\![t_j, t_{j+1}]\!] \cup [\![t_i, x]\!] \qquad \text{and} \qquad d(0,x) = \sum_{0 \leq j < i} d(t_j, t_{j+1}) + d(t_i, x)\,.$$

ii) for every $i \geq 0$ there exists a vector $\mathbf{U}_i$ constructed from $\mathbf{U}$ such that, almost surely, the restriction of $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ to $\pi([t_i, z_i])$ is isometric to $\mathrm{CMT}(h_i, \mathbf{U}_i)$, where h_i is the excursion defined in (4).

Proof. *i)* By definition, $t_i \in [\![0,x]\!]$ so that decomposing $[\![0,x]\!]$ on $[0,t_i] \cup [t_i, x]$, it follows immediately that $[\![0,x]\!] = [\![0,t_i]\!] \cup [\![t_i,x]\!]$. A straightforward induction yields the claim.

ii) By Lemma 3.2, $t_i \in \mathcal{V}_y$ so that $t_i \in [\![0,y]\!] \cap [\![0,x]\!]$. It follows that $t_i \leq x \wedge y$. For the interval $[t_i, z_i] \subseteq [0,1]$, we now define a sequence $\mathbf{U}_i$ from $\mathbf{U}$ as follows. Recall Remark 3.9 about the association map, and let $(I_j)_{j \geq 1}$ the enumeration of the intervals with rational end points there. Recall also the definition of h_i in (4). We denote by $j(\mathrm{e}, \cdot)$ and $j(h_i, \cdot)$ the association maps of e and h_i respectively. For every $t \in \mathscr{L}(\mathrm{e})$, then $t - t_i \in \mathscr{L}(h_i)$ (the set of local minima is a.s. preserved by the removal of a linear drift). Let

$$U_{i,k} := \begin{cases} U_{j(\mathrm{e}, t+t_i)} & \text{if } k = j(h_i, t) \text{ for some } t \in \mathscr{L}(h_i) \\ 0 & \text{otherwise}\,. \end{cases}$$

Let $\mathbf{U}_i = (U_{i,k})_{k \geq 1}$. Then, the restriction of $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ to $\pi([t_i, z_i])$ is isometric to $\mathrm{CMT}(h_i, \mathbf{U}_i)$. Note that the components of $\mathbf{U}_i$ that we have set to 0 above are never used in the construction; if one wants to enforce that $\mathbf{U}_i$ has the same distribution as $\mathbf{U}$, one can instead use independent uniform random variables to complete the definition of $\mathbf{U}_i$. □

3.6 Geodesics and the length measure

In this section, we show that the distance $d(x,y)$ is actually a measurable function of the set $[\![x,y]\!]$. Let ψ be the function defined by $\psi(r) = \sqrt{r |\log| \log r||}$ for $r > 0$, and let m^ψ denote the Hausdorff measure constructed on $\mathbb{R}$ using ψ as a gauge function. Recall that the ψ-Hausdorff measure m^ψ of a Borel set $E \subseteq \mathbb{R}$ is defined by [33, 34, 52]

$$m^\psi(E) := \lim_{\delta \to 0+} \inf \left\{ \sum_{i \geq 1} \psi(A_i) : E \subseteq \bigcup_{i \geq 1} A_i, |A_i| < \delta \right\},$$

where the A_i are intervals and $|A_i|$ are their lengths.

For any $x, y \in [0,1]$, the distances between pairs of points of $[\![x,y]\!]$ naturally define a measure as follows: for any $z,t \in [\![x,y]\!]$, we have $[\![z,t]\!] \subseteq [\![x,y]\!]$ and we let $\ell_{x,y}([\![z,t]\!]) = d(z,t)$. More generally, for any compact interval $A \subseteq [0,1]$, we let $\ell^\circ_{[\![x,y]\!]}(A) = \ell^\circ_{[\![x,y]\!]}([\![x,y]\!] \cap A) = d(\inf [\![x,y]\!] \cap A, \sup [\![x,y]\!] \cap A)$. This defines $\ell^\circ_{[\![x,y]\!]}$ uniquely as a Borel measure on $[0,1]$.

Lemma 3.20. *Let V be a random variable with uniform distribution independent of $(\mathrm{e}, \mathbf{U})$. There exists a constant $a > 0$ such that, with probability one, for any Borel set $A \subseteq [0,1]$, we have $\ell_{[\![0,V]\!]}(A) = a \cdot m^\psi(A \cap [\![0,V]\!])$. In particular, $d(0,V) = a \cdot m^\psi([\![0,V]\!])$.*

Remark 3.21. It would be possible to identify the constant a using Theorem 1 of Perkins [57] who strengthened the results of Taylor and Wendel [63] by (among others) identifying the multiplicative constant between the ψ-Hausdorff measure and the local time for the zero set of Brownian motion. However, we did not pursue this further.

Proof. For V uniform on $[0,1]$, the Cantor set $[\![0,V]\!]$ has a recursive structure that is tractable with the tools developed by Graf, Mauldin, and Williams [36] and Mauldin and Williams [53], which will allow us to compare $m^\psi([\![0,V]\!])$ and $d(0,V)$.

Let $t_1 = t_1(V)$ be the location of the first vertex of the convex minorant of e on the interval $[0,V]$, and let $z_1 = z_1(V)$. Then, by Lemma 3.6,

$$(t_1, z_1 - t_1, 1 - z_1) \sim \text{Dirichlet}(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2})\,,$$

so that, conditionally on (t_1, z_1), V is uniform in (t_1, z_1). On the other hand, the random jump $\xi_1 = \boldsymbol{j}(t_1)$ is uniform in $(0, t_1)$ and independent of the rest. This implies that the random Cantor set $[\![0,V]\!]$ has the same distribution as C constructed as follows. Let $\mathcal{U}_2 := \bigcup_{n\geq 0}\{1,2\}^n$, and let $(\Delta_1(u), \Delta_2(u), \Delta_3(u))$, $u \in \mathcal{U}_2$, be i.i.d. copies of a Dirichlet$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ random vector $(\Delta_1, \Delta_2, \Delta_3)$. Set $C_\varnothing = [0,1]$ and, for each $u \in \mathcal{U}_2$, let

$$C_{u1} := [\inf C_u, \inf C_u + |C_u| \cdot \Delta_1(u)], \qquad \text{and} \qquad C_{u2} := [\sup C_{u1}, \sup C_{u1} + |C_u| \cdot \Delta_2(u)]\,.$$

Observe that, for each u, C_{u1} and C_{u2} are two intervals in C_u, with disjoint interior. Then, for $n \geq 0$, we set $C^n = \bigcup_{|u|=n} C_u$ and $C = \bigcap_{n\geq 0} C^n$. Note in particular that no additional randomness is needed, that would correspond to the point V: with this definition $\sup C$ is uniformly distributed on $[0,1]$.

The law of (Δ_1, Δ_2) is explicit and its density $\rho(x_1, x_2)$ is given by

$$\rho(x_1, x_2) = x_1^{-1/2} x_2^{-1/2} (1 - x_1 - x_2)^{-1/2} \frac{\Gamma(3/2)}{\Gamma(1/2)^3} \cdot \mathbf{1}_{x_1 + x_2 \leq 1}\,.$$

Theorem 5.1 of [36] applies: one easily verifies that for $\alpha = 1/2$ we have $\mathbf{E}[\Delta_1^\alpha + \Delta_2^\alpha] = 1$, $\mathbf{P}(\Delta_1^\alpha + \Delta_2^\alpha = 1) = 0$, and $\mathbf{E}[1/\min\{\Delta_1^\nu, \Delta_2^\nu\}] \leq 2\mathbf{E}[\Delta_1^{-\nu}] < \infty$ for all $\nu \in (0, \frac{1}{2})$. Furthermore, Condition (5.1) of [36] is satisfied for the point $(x_1, x_2) = (\frac{1}{2}, \frac{1}{2})$, since the density ρ is bounded away from zero uniformly. It follows that with probability one, $m^\psi(C) \in (0, \infty)$. Now, Theorem 5.5 there does not directly apply since the C_{u1} and C_{u2} intersect for every u, but this is only at one point, and one easily verifies that the proof there still holds since m^ψ assigns measure zero to any countable collection of points. We conclude that there exists a constant $a > 0$ such that $d(0,V) = a \cdot m^\psi(C)$.

Furthermore, the measure $a \cdot m^\psi(\,\cdot \cap C)$ coincides with the construction measure ν of Mauldin and Williams [53], which is easily seen to correspond here to the measure $\ell^\circ_{[\![0,V]\!]}$. For an interval $A \subseteq [0,1]$, the sequence

$$\nu_n(A) := \sum_{|u|=n, C_u \cap A \neq \varnothing} |C_u|^{1/2}$$

almost surely converges to a limit value $\nu(A)$. This defines the Borel measure ν on $[0,1]$ of total mass $d(0,V)$. The fact that $\nu(A) = \ell^\circ_{[\![0,V]\!]}(A)$ should by now be straightforward. □

The measures $\ell^\circ_{[\![x,y]\!]}$, $x, y \in [0,1]$, are actually the restrictions of a general measure on $[0,1]$ which projects to the length measure on the convex minorant tree. There is a pre-skeleton on $[0,1]$ which is defined by $\mathrm{Skel}([0,1]) = \cup_{x\in\mathscr{L}}]\!]0,x[\![= \cup_{x,y\in\mathscr{L}}]\!]x,y[\![$. Let ℓ° be the Borel sigma-finite measure on $[0,1]$ uniquely defined by

i) $\ell^\circ(\mathrm{Skel}([0,1])^c) = 0$, and

ii) for every $x, y \in \mathscr{L}$, and every interval A of $[0,1]$, $\ell^\circ(A \cap [\![x,y]\!]) = \ell^\circ_{[\![x,y]\!]}(A)$.

Then the push-foward measure $\ell = \pi_*\ell^\circ$ is the length measure on the convex minorant tree $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$. Finally, we verify that this corresponds to the push-forward of m^ψ (up to a multiplicative constant). This is essentially just the fact that the skeleton $\mathrm{Skel}([0,1])$ is a countable union of segments, that we can rewrite in terms of a sequence of i.i.d. uniform points on $[0,1]$

Proposition 3.22. *Let $(V_i)_{i\geq 1}$ be i.i.d. uniform on $[0,1]$, also independent of $(\mathrm{e}, \mathbf{U})$. Then*
i) $\mathrm{Skel}([0,1]) \subseteq \cup_{i\geq 1}[\![0, V_i]\!]$, and
ii) with the constant $a > 0$ of Lemma 3.20, the measures ℓ° and $a \cdot m^\psi$ almost surely coincide.

Proof. *i)* Fix any $x \in \mathscr{L}$. Almost surely, there is an $i \in \mathbb{N}$, such that $x = t_i(x)$ and $z_i(x) > t_i(x)$. For any $y \in [t_i, z_i)$ we have $x = t_i(x) = t_i(y)$. In particular, a.s. there exist infinitely many $n \geq 1$ such that $x = t_i(V_n)$ for some $i \geq 1$. It follows that $\mathrm{Skel}([0,1]) \subset \cup_{n\geq 1}[\![0, V_n]\!]$.

ii) For each $n \geq 1$, let $B_n := [\![0, V_n]\!] \setminus \bigcup_{1\leq j<n}[\![0, V_j]\!]$. Then, $\mathrm{Skel}([0,1])$ is contained in the union of the B_n, $n \geq 1$, which are disjoint sets, and for any interval $A \subseteq [0,1]$, we have

$$\ell^\circ(A) = \sum_{n\geq 1} \ell^\circ(A \cap B_n) = \sum_{n\geq 1} a \cdot m^\psi(A \cap B_n) = m^\psi(A)\,,$$

so that the measures ℓ° and $a \cdot m^\psi$ indeed concide. □

Remark 3.23. We note the decomposition for the distance $d(0,V)$ identifies its distribution: indeed, by Brownian scaling if Ddenotes the random variable $d(0,V)$, and D_1 and D_2 are two independent copies of D, then we have

$$D \overset{d}{=} \sqrt{\Delta_1} D_1 + \sqrt{\Delta_2} D_2$$

which implies that, up to a deterministic multiplicative constant, D has the Rayleigh distribution (see for instance Proposition 2.1 of [9]). This can be seen as a first step towards the identification of the law of $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$; see Section 4 for a full proof of this fact.

3.7 Recursive convex minorants of Brownian motion with parabolic drift

We now move on to the definition of the main object of the paper, the tree $\mathrm{CMT}(X, \mathbf{U})$. A straightforward application of the Girsanov Theorem shows that, for any $x \in \mathbb{R}_+$, the law of $(X_{s\in[0,x]})$ is absolutely continuous with respect to that of $(W_s)_{s\in[0,x]}$. As a consequence, "local properties" that hold almost surely for W also hold almost surely for X as well. Since we are only interested in a definition in this section, we may focus on the case of a Brownian motion. In the following, we use the same notation as for the Brownian excursion, we believe that it should not cause any confusion.

The convex minorant tree associated with a Brownian motion. We consider $(W_s)_{s\geq 0}$ a standard Brownian motion. Fix $x \in \mathbb{R}_+$ and consider the recursive convex minorants of W on $[0,x]$. Recall that $\mathcal{V}_x(W)$ a.s. has an accumulation point at 0; let $(t_i)_{i\in\mathbb{Z}} = (t_i(x, W))_{i\in\mathbb{Z}}$ as defined in Section 3.2. The sequence $(t_i)_{i\in\mathbb{Z}}$ is bi-infinite, but one can write for a fixed i

$$[\![0,x]\!] := \{0\} \cup \bigcup_{j:j\leq i} [\![t_{j-1}, t_j]\!] \cup [\![t_i, x]\!]\,. \tag{12}$$

By Lemma 3.5, the sets $[\![t_{j-1}, t_j]\!]$, for j such that $j \leq i$ and $[\![t_i, x]\!]$ are well-defined by the construction of Section 3 (for the Brownian excursion). Furthermore, Lemma 3.19 ensures that the value of $[\![0, x]\!]$ is independent of $i \in \mathbb{Z}$, so that $[\![0, x]\!]$ is well-defined as well. To define the distance $d(0, x) = d_W(0, x)$, we shall verify that the sum of distances given by the decomposition in (12) converges (a priori, 0 could be at infinite distance from every point $x > 0$). Observe that, still from Lemma 3.5, conditionally on $(t_i)_{i\in\mathbb{Z}}$, for any $i \in \mathbb{Z}$, $d(t_i, t_{i+1})$ is distributed like $|t_{i+1} - t_i|^{1/2}$ times the distance between 0 and 1 in $\mathrm{CMT}(\mathbf{e}, \mathbf{U})$. Also by Theorem 1, Corollary 1 of [58] and Brownian scaling, $(|t_{i+1} - t_i|)_{i\in\mathbb{Z}}$ has the same distribution as $(x_i / \sum_{j\in\mathbb{Z}} x_j)_{i\in\mathbb{Z}}$, where $(x_j)_{j\in\mathbb{Z}}$ denote the points of a Poisson point process of intensity $e^{-x}/x dx$ on $\mathbb{R}_+$. Straightforward calculation shows that $\mathbf{P}(\sum_i x_i > 0) = 1$ and $\mathbf{E}[\sum_i x_i^p] = \int_0^{+\infty} x^{p-1} e^{-x} dx < \infty$ for $p \in (0, 1]$ so that a.s. $\sum_i \sqrt{x_i} < +\infty$ and $\sum_i x_i < +\infty$ which implies that $\sum_i |t_{i+1} - t_i|^{1/2} < \infty$ almost surely.

So, for the distance we may define $d(0, x) = d_W(0, x)$ by

$$d_W(0, x) := \sum_{j\leq i} d(t_{j-i}, t_j) + d(t_i, x) = \sum_{j\in\mathbb{Z}} d(t_{j-1}, t_j)\,,$$

which is almost surely finite for almost all $x \in \mathbb{R}_+$. In particular, with probability one $d_W(0, x) < \infty$ for every $x \in \mathscr{L}(W)$ by Lemma 3.3.

Finally, for any $x, y > 0$, we can define $d(x, y) = d_W(x, y)$ as follows. Writing $x \wedge y = \sup[\![0, x]\!] \cap [\![0, y]\!]$ as before, we have $x \wedge y \in [\![0, t]\!]$ for some $t \in \mathscr{L}$, by the obvious extension of Lemma 3.14 to the case of Brownian motion. Therefore, $d(0, x \wedge y) < \infty$. We may thus define

$$d_W(x, y) := d_W(0, x) + d_W(0, y) - 2d_W(0, x \wedge y)\,,$$

and we will prove that it is a.s. finite for all x, y. Note also that, assuming without loss of generality that $y < x$, there exists some $i \in \mathbb{Z}$ such that $y \in [t_i, t_{i+1})$. In particular, $x, y \in [t_i, z_i]$ and we may equivalently define $d(x, y)$ by

$$d(t_i, x) + d(t_i, y) - 2d(t_i, x \wedge y)\,,$$

and any i of which $t_i < y$ would yield the exact same value.

The convex minorant tree associated with X: The Brownian parabolic tree. At last, we consider X, the Brownian motion with parabolic drift. By absolute continuity, the sets $[\![x, y]\!] = [\![x, y]\!]_X$ and $d(x, y) = d_X(x, y)$ are also well-defined for every $x, y \in \mathbb{R}_+$. The triangle inequality and four-point condition are satisfied by construction (Lemma 3.16). Let $x \sim y$ if $d_X(x, y) = 0$, and let $(\mathscr{M}, d)$ denote the metric completion of the quotient metric space; define $\rho = \pi(0)$. For the mass measure, one needs some rescaling and we shall admit for now that the collection of measures $(x^{-1}\pi_* \operatorname{Leb}|_{[0,x]})_{x>0}$ converges weakly with probability one to a probability measure μ. The proof of this fact is the topic of Section 6.1. Finally, we let $\mathrm{CMT}(X, \mathbf{U})$ denote the pointed measured complete metric space $\mathfrak{M} := (\mathscr{M}, d, \mu, \rho)$, and we call it the Brownian parabolic tree.

4 A dynamic point of view and the law of $\mathrm{CMT}(\mathbf{e}, \mathbf{U})$

In this section, we study the convex minorant tree of a standard Brownian excursion. We prove Theorem 1.3 in the case where $s = 0$ which says that $\mathrm{CMT}(\mathbf{e}, \mathbf{U})$ is distributed like the Brownian continuum random tree, and Theorem 1.5 which relates $\mathrm{CMT}(\mathbf{e}, \mathbf{U})$ to the additive coalescent. We are interested here in the case of excursions, and the natural range of interest for $Z^\lambda(\omega)$ is then $\lambda \in (-\infty, 0]$, and we shall therefore rather work with $(Z^{-\tau}(\omega))_{\tau\geq 0}$, which also turns out to be a cadlag process (see Lemma A.1). We still occasionally use the parameterization with λ.

4.1 A fragmentation connected to Brownian motion

The properties of Z^λ is intimately related to the following operators. For $\lambda \in \mathbb{R}$, define the operator Ψ_λ as follows: for a function f continuous on an interval $D \subseteq \mathbb{R}_+$, and $t \in D$

$$\Psi_\lambda f(t) := f(t) + \lambda t - \inf\{f(s) + \lambda s : s \in D, s \leq t\}\,. \tag{13}$$

Then, $Z^\lambda(\omega) = \{s \in D : \Psi_\lambda \omega(s) = 0\}$. The family of operators $(\Psi_\lambda)_{\lambda\in\mathbb{R}}$ enjoys the following composition property, which is a straightforward reformulation of the arguments leading to Theorem 1 i) of [20]. For $t \geq 0$, let $\mathfrak{S}_t$ denote the shift operator defined by $\mathfrak{S}_t f(s) = f(t+s)$, for all $s \geq 0$.

Lemma 4.1. *Let f be a continuous function on $D \subseteq \mathbb{R}_+$ and suppose that, for some $\lambda \in \mathbb{R}$ and $t \in D$, we have $\Psi_\lambda f(t) = 0$. Then, for all $h, s \geq 0$ with $t + s \in D$ one has*

$$\mathfrak{S}_t \Psi_{\lambda-h} f(s) = \Psi_{\lambda-h} f(t+s) = \Psi_{-h} \mathfrak{S}_t \Psi_\lambda f(s)\,.$$

In particular, $\Psi_{\lambda-h} f(t) = 0$ for all $h \geq 0$.

We now go back to the case where $f = \mathrm{e}$ is a Brownian excursion and write $Z^\lambda = Z^\lambda(\mathrm{e})$ (in this case, $D = [0,1]$). Lemma 4.1 implies for instance that $Z^{-\tau} = Z^{-\tau}(\mathrm{e})$ is non-decreasing in τ for the inclusion, and thus induces a fragmentation in the sense that the connected components of its complement split as τ increases. For any $x \in [0,1)$ let $I^\tau(x)$ be the maximal interval of the form $[a,b)$ containing x such that for $(a,b) \cap Z^{-\tau} = \varnothing$. For $x, y \in [0,1)$, we let $x \sim_\tau y$ if $I^\tau(x) = I^\tau(y)$. Observe that, for every $\tau \geq 0$, the collection of $I^\tau(x)$ forms a partition of $[0,1)$.

By Lemma A.1, a.s, for every $\tau \geq 0$, $[0,1] \setminus Z^{-\tau}$ consists in countably many open intervals, whose lengths we denote by $F_1(\tau), F_2(\tau), \ldots$ in the decreasing order. Then, let $F(\tau) = (F_i(\tau))_{i\geq 1}$. The main result of Bertoin [20] is that the process $(F(\tau), \tau \geq 0)$ has the same distribution as another remarkable fragmentation introduced by Aldous and Pitman [15], where a Brownian continuum random tree is logged along its skeleton at the points of an (independent) Poisson point process of unit intensity; the process of interest is the sequence of sorted masses of the fragments. This shows in particular that, up to a time change, the time reversal of $(F(\tau))_{\tau\geq 0}$ is the classical standard additive coalescent.

Although there is no obvious coupling between the two representations directly in the continuous, this shows that the fragmentation of $[0,1]$ constructed by Bertoin corresponds to a fragmentation of a certain Brownian continuum random tree. This section will show that (one choice for) this tree is the convex minorant tree $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$. We will also identify the collection of points/times where/when it should be cut and thereby, provide a coupling between the two representations.

The rest of the section is organized as follows. In Section 4.2, we make explicit the correspondence between $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ and the dynamics related to the process $Z^{-\tau}$ described above. In Section 4.3, following Bertoin and Miermont [21], we introduce the cut tree which encodes the genealogy of the fragmentation $(F(\tau))_{\tau\geq 0}$. The cut tree is a crucial ingredient since it provides the link between the fragmentation and the recovery of "the tree what was logged" through the *inverse cut tree transform* that has been studied in [1, 25]. In Section 4.4, we make the connection between the cut tree, the inverse transform and $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ and complete the proofs of Theorem 1.3 (with $s = 0$) and Theorem 1.5.

4.2 Making the dynamics explicit

In this section, we provide another point of view on the convex minorant tree that makes explicit its relation with the fragmentation of $[0,1]$ induced by $Z^{-\tau} = Z^{-\tau}(\mathrm{e})$.

Lemma 4.2. *Almost surely, the following holds for every point $x \in [0,1]$. Let $(t_i)_{i\geq 0}$ and $(\gamma_i)_{i\geq 0}$ be the vertices and the slopes of the convex minorant of e on $[0,x]$. Then, setting $\gamma_{-1} = 0$ for convenience, we have for all $i \geq 0$,*

i) $\inf I^\tau(x) = t_i$ *for all* $\tau \in [\gamma_{i-1}, \gamma_i)$, *and*

ii) $\sup I^\tau(x) = z_i$ *for* $\tau = \gamma_{i-1}$.

Proof. We work on a set $\Omega^\star$ of probability one where all the events of Lemma 3.3 occur, in particular, the slopes $(\gamma_i)_{i\geq 0}$ are strictly increasing for every $x \in [0,1]$. The rest of the proof is deterministic, and we proceed by induction on $i \geq 0$.

Write e^λ for the function $s \mapsto \mathrm{e}(s) + \lambda s$. For $i = 0$, by construction of the convex minorant, for every $\tau \in [0, \gamma_0)$, $\mathrm{e}^{-\tau}$ is positive on $(0, x]$ and thus $\inf I^\tau(x) = 0 = t_0$. For $\tau = \gamma_0$ we have $\mathrm{e}^{-\gamma_0}(t_1) = \mathrm{e}^{-\gamma_0}(z_1) = 0$, and $\mathrm{e}^{-\gamma_0}(s) > 0$ for $s \in (t_1, z_1)$. It follows that $t_1, z_1 \in Z^{-\gamma_0}$ and that $\sup I^{\gamma_0}(x) = z_1$.

Suppose now that, for some $j \geq 0$, the claims in *i)* and *ii)* both hold for all $0 \leq i \leq j$, and that $t_{j+1}, z_{j+1} \in Z^{-\gamma_j}$. By expressing $\Psi_{-\gamma_j - h}\mathrm{e}$ for $h \geq 0$ in terms of $\Psi_{-\gamma_j}\mathrm{e}$, Lemma 4.1 allows us to proceed. First note that the vertices of the convex minorant of $\Psi_{-\gamma_j}\mathrm{e}$ on $[0, x]$ that are in $[t_{j+1}, 1]$ are precisely $(t_{j+k})_{k\geq 1}$ and the corresponding slopes are $(\gamma_{j+k} - \gamma_j)_{k\geq 1}$. The argument we have just used for $j = 0$ applies to $\mathfrak{S}_{t_{j+1}}\Psi_{-\gamma_j}\mathrm{e}$ and yields that for all $h \in [0, \gamma_{j+1} - \gamma_j)$, we have $\inf I^{\gamma_j + h}(x) = t_{i+1}$ and $\sup I^{\gamma_{i+1}}(x) = z_{i+1}$. Furthermore, for $h = \gamma_{j+1} - \gamma_j$, t_{j+2} and z_{j+2} are both zeros of $\Psi_{-\gamma_{j+1}}\mathrm{e}$, while the latter is positive on (t_{j+2}, z_{j+2}). This completes the proof. □

For $x, y \in [0, 1)$, define $\tau(x, y) = \sup\{\tau \geq 0 : x \sim_\tau y\}$. Note that, by the left-continuity of Z^λ, we have $Z^{-\tau(x,y)} \cap [x, y] \neq \varnothing$. Recall the definition of ξ_m from Section 3.4, which is also the point $\boldsymbol{\imath}(t_m)$ as defined in Remark 3.11.

Lemma 4.3. *Almost surely for every $x \neq y \in [0, 1)$, we have the following: let $(t_i)_{i\geq 0}$ be the vertices of the convex minorant of* e *on $[0, \max\{x, y\}]$. Then, $m := \min\{i \geq 1 : t_i > \min\{x, y\}\} < \infty$, and:*

i) $Z^{-\tau(x,y)} \cap [x, y]$ *consists of the single point $\kappa(x, y) = t_m$ that we call a cut point;*

ii) $\tau(x, y) = \gamma_{m-1}$;

iii) $I^{\tau(x,y)}(\min\{x, y\}) = [t_{m-1}, t_m)$ *and* $I^{\tau(x,y)}(\max\{x, y\}) = [t_m, z_m)$.

Furthermore, we let $\eta(x, y) = \xi_m = \boldsymbol{\imath}(t_m) \in (t_{m-1}, t_m)$; conditionally on $I^{\tau(x,y)}(\min\{x, y\}) = S$, $\eta(x, y)$ is uniformly distributed on S.

Proof. The set of probability one is $\Omega^\star$ where all the events of Lemma 3.3 occur for every point of $[0, 1]$. The points *i)* to *iii)* are straightforward consequences of Lemma 4.2, applied to the fragment containing $I^\tau(\max\{x, y\})$ until the time when it does not contain $\min\{x, y\}$ any longer. The statement concerning the distribution of $\eta(x, y)$ is a consequence of fact that $\eta(x, y)$ is then $\boldsymbol{\imath}(t_m)$, which is uniform in $[t_{m-1}, t_m]$ conditionally on t_{m-1}, t_m. □

Observe that Lemma 4.3 implies that, almost surely for every $x \neq y$, we have

$$I^{\tau(x,y)-}(x) = \bigcap_{\tau < \tau(x,y)} I^\tau(x) = I^{\tau(x,y)}(x) \sqcup I^{\tau(x,y)}(y)\,. \tag{14}$$

We are now ready to move on to the main objective of this section, namely proving that both $[\![x, y]\!]$ and $d(x, y)$ may be defined using an alternative binary decomposition where the intervals containing a pair of marked points are split at the corresponding cut point, just as in (14) above.

Let $\mathcal{U}_2 = \bigcup_{n\geq 0}\{0, 1\}^n$, where it is understood that $\{0, 1\}^0 = \{\varnothing\}$. Fix now $x, y \in (0, 1)$. We define recursively $(\Pi_u, \tau_u, \kappa_u, A_u, B_u)_{u\in\mathcal{U}_2}$, where Π_u is an interval, $A_u \leq B_u$ are two points in the closure of Π_u, and the values $\tau_u \in \mathbb{R}_+$, $\kappa_u \in [0, 1]$ are always such that $\tau_u = \tau(A_u, B_u)$, $\kappa_u = \kappa(A_u, B_u)$. It is understood that all these random variables depend on x, y, so we actually have $\Pi_u(x, y), \tau_u(x, y), \kappa_u(x, y)$, $A_u(x, y)$, $B_u(x, y)$, for $u \in \mathcal{U}_2$, but we usually omit the reference to x, y. Set $\Pi_\varnothing = (0, 1)$, $\tau_\varnothing(x, y) = \tau(x, y)$, $\kappa_\varnothing(x, y) = \kappa(x, y)$ and $A_\varnothing = \min\{x, y\}$, $B_\varnothing = \max\{x, y\}$. Let $\Pi_0 = I^{\tau_\varnothing}(A_\varnothing)$ and $\Pi_1 = I^{\tau_\varnothing}(B_\varnothing)$.

Assuming that we have defined $(\Pi_u, \tau_u, \kappa_u, A_u, B_u)$ for some $u \in \mathcal{U}_2$ we then set $\Pi_{u0} = I^{\tau_u}(A_u)$, $\Pi_{u1} = I^{\tau_u}(B_u)$, $A_{u0} = \min\{A_u, \eta(A_u, B_u)\}$, $B_{u0} = \max\{A_u, \eta(A_u, B_u)\}$ and $A_{u1} = \kappa_u = \inf \Pi_{u1}$, $B_{u1} = B_u$. We finally define $\tau_{ui} = \tau(A_{ui}, B_{ui})$ and $\kappa_{ui} = \kappa(A_{ui}, B_{ui})$ for $i \in \{0, 1\}$.

Lemma 4.4. *Almost surely for every $x, y \in [0,1]$, for every $u \in \mathcal{U}_2$, we have $A_u, B_u \in [\![0, x]\!] \cup [\![0, y]\!]$.*

Proof. This is a straightforward induction. For $u = \varnothing$, we have $\{A_\varnothing, B_\varnothing\} = \{x, y\}$ and the claim holds by definition. Assume now that it holds for some $u \in \mathcal{U}$. We have $\{A_{u0}, B_{u0}, A_{u1}, B_{u1}\} = \{A_u, B_u, \kappa(A_u, B_u), \eta(A_u, B_u)\}$. By Lemma 4.3, $\kappa(A_u, B_u)$ is a vertex on the convex minorant of e on the interval $[0, \max\{A_u, B_u\}]$ and thus lies in $[\![0, \max\{A_u, B_u\}]\!] \subseteq [\![0, x]\!] \cup [\![0, y]\!]$ by the induction hypothesis and Lemma 3.13. The same holds for $\eta(A_u, B_u)$ by Lemma 4.3 and the definition of $[\![\cdot, \cdot]\!]$. □

To avoid any difficulties, we define $\hat{d}_\lambda$ only for almost every pair of points. This will be enough to exhibit the dynamic properties we have in mind, and settle the foundations for the coupling of Section 7 that allows to identify the law of $\mathrm{CMT}(X, \mathbf{U})$. In the following, $\overline{\Pi}_u$ denotes the closure of Π_u. Let

$$\Pi(x, y) = \bigcap_{n \geq 0} \bigcup_{|u|=n} \overline{\Pi}_u \qquad \text{and} \qquad \hat{d}(x, y) = \limsup_{n \to \infty} \sum_{|u|=n} |\Pi_u|^{1/2} .$$

The set $\Pi(x, y)$ is well-defined and non-empty, but it is so far unclear whether $\hat{d}(x, y)$ is finite.

Proposition 4.5. *For any $x, y \in [0,1]$, we have almost surely*
i) $\Pi(x, y) = [\![x, y]\!]$, *and*
ii) $\hat{d}(x, y) = d(x, y)$.

Proof. *i)* Recall that, by definition, $[\![x, y]\!]$ is the union of $[\![0, x]\!] \cap [x \wedge y, 1]$ and $[\![0, y]\!] \cap [x \wedge y, 1]$. We follow the binary decomposition defining $\Pi(x, y)$; for each $n \geq 0$, let 0^n be the left-most node in $\mathcal{U}_2$ at level n; we agree that, in this context, $0^0 = \varnothing$. We show that, for each $n \geq 0$, the two sets $\Pi(x, y)$ and $[\![x, y]\!]$ coincide on $[\kappa_{0^n}, 1]$; we will then show that $\kappa_{0^n} = \kappa_{0^n} \downarrow x \wedge y$ as $n \to \infty$.

For $n = 0$, we have $\kappa_{0^0} = \kappa_\varnothing = \kappa(x, y)$, $A_\varnothing = \min\{x, y\}$ and $B_\varnothing = \max\{x, y\}$. By Lemma 4.3 *i)* and the definition of $[\![0, B_\varnothing]\!]$ in Equation (5), the set $\Pi(x, y) \cap [\kappa_{0^0}, 1]$ is contained in $[\![0, B_\varnothing]\!]$; furthermore, since $x \wedge y \leq \min\{x, y\} = A_\varnothing$, it is also the case that $\Pi(x, y) \cap [\kappa_{0^0}, 1]$ is contained in $[\![0, B_\varnothing]\!] \cap [x \wedge y, 1]$. On the other hand, by Lemma 4.4, $\kappa(x, y) \in [\![0, \max\{x, y\}]\!]$, and one easily sees that $\Pi(x, y) \cap [\kappa_{0^0}, 1] = [\![0, B_\varnothing]\!] \cap [\kappa_{0^0}, 1]$. Indeed, we may now expand Π_1 on the right using the recurrence relation: writing $0^i 1^j$ for the node at level $i + j$ in $\mathcal{U}_2$ obtained by walking i steps left, and then j steps right from the root, and it should be plain that the points κ_{1^k}, $k \geq 0$, are simply the vertices of the convex minorant of e on $[0, B_\varnothing]$ that are larger than $\kappa_\varnothing = \kappa_{0^0}$. It follows that the sets $\Pi_{10}, \Pi_{1^2 0}, \ldots, \Pi_{1^i 0}, \ldots$ all explicitly appear in the decomposition defining $[\![0, B_\varnothing]\!]$ on the interval $[\kappa_{0^0}, 1]$.

Now for any $n \geq 1$, assuming that we have treated the part of $\Pi(x, y)$ lying in $[\kappa_{0^n}, 1]$, we are left with the portion of $\Pi(x, y)$ that lies in $[0, \kappa_{0^n}]$, which is constructed from $\Pi_{0^{n+1}}$. By Lemma 4.4, we have $A_{0^{n+1}}, B_{0^{n+1}} \in [\![0, x]\!] \cup [\![0, y]\!]$, and we have $\kappa_{0^{n+1}} = \kappa(A_{0^{n+1}}, B_{0^{n+1}})$. To the right, we have the set $\Pi_{0^{n+1}1}$, that we may expand from the right using the recurrence relation. The arguments above imply that the $\kappa_{0^n 1^k}$, $k \geq 0$, are the vertices of the convex minorant on the interval $[0, B_{0^{n+1}}]$ that are at least $\kappa_{0^{n+1}}$. Therefore, $\Pi(x, y)$ and $[\![x, y]\!]$ coincide on $[\kappa_{0^{n+1}}, B_{0^{n+1}}]$ and thus on $[\kappa_{0^{n+1}}, \kappa_{0^n}]$, and in turn on $[\kappa_{0^{n+1}}, 1]$ by the induction hypothesis.

Then, note that for each $n \geq 0$, $x \wedge y \in \Pi_{0^n}$. To see this, it suffices to note that for each $n \geq 0$, one of (A_{0^n}, B_{0^n}) or (B_{0^n}, A_{0^n}) lies in $[\![0, x]\!] \times [\![0, y]\!]$. This is clearly true for $n = 0$, and carries on because at each step we replace $\max\{A_{0^n}, B_{0^n}\}$ by $\eta(A_{0^n}, B_{0^n})$ which lies in $[\![0, \max\{A_{0^n}, B_{0^n}\}]\!]$. Lemma 3.13 them implies that $\inf \Pi_{0^n} \in [\![0, x]\!] \cap [\![0, y]\!]$, which proves the claim. Since $\inf \Pi_{0^n}$ is non-decreasing, it would suffice to prove that $\mathrm{diam}(\Pi_{0^n}) = \mathrm{Leb}(\Pi_{0^n}) \to 0$ in order to show that $\kappa_{0^n} \to x \wedge y$, which would complete the proof of *i)*. So let us now this why $\mathrm{Leb}(\Pi_{0^n}) \to 0$. For every u, we have $A_u \in [\inf \Pi_u, \kappa_u]$. Then, either $A_u < \eta_u$ and $|\Pi_{u0}| \leq |\Pi_u| \cdot U$ where U is uniformly random on $[0, 1]$, or $A_u \geq \eta_u$, and then Π_{u0} contains two uniform random points so that, $|\Pi_{u00}| \leq |\Pi_u| \cdot M$, where M is a $\mathrm{Beta}(\frac{1}{2}, 1)$ random variable by Lemma 3.6. Since all the random variables are independent, it is straightforward that $|\Pi_{0^n}| \to 0$ with probability one as $n \to \infty$.

ii) The correspondence between the sets that are used to define $\Pi(x, y)$ and $[\![x, y]\!]$ in the proof of *i)*, also yields a way of rewriting the sums which proves that $\hat{d}(x, y) = d(x, y)$. We omit the details. □

4.3 The cut tree and the reconstruction problem

The fragmentation we have presented in Section 4.1 has a remarkable genealogy, which can be encoded into a *cut tree* introduced by Bertoin and Miermont [21], and which turns out to be distributed like a Brownian continuum random tree.

Let $(\zeta_i)_{i\geq 1}$ be i.i.d. uniform points in $[0,1]$, which are also independent of $(e,\mathbf{U})$. Almost surely, for all $i\neq j$, we have $i\sim_0 j$. Then, for distinct $i,j\geq 1$ let $\tau_{ij}=\inf\{\tau\geq 0 : \zeta_i \not\sim_\tau \zeta_j\}$ be the first time when ζ_i and ζ_j are separated by a point of $Z^{-\tau}$. Then, we define a function δ on $\mathbb{N}_0\times\mathbb{N}_0$ as follows:

$$\delta(0,i)=\int_0^\infty |I^\tau(\zeta_i)|d\lambda \qquad \text{and} \qquad \delta(i,j)=\int_{\tau_{ij}}^\infty |I^\tau(\zeta_i)|d\tau+\int_{\tau_{ij}}^\infty |I^\tau(\zeta_j)|d\tau\,, \tag{15}$$

where $|\cdot|$ denotes the Lebesgue measure on $[0,1]$. It is known that δ defines a real tree [21]: let $\mathscr{C}$ denote the completion of $\mathbb{N}_0$ with respect to δ, and let ν denote the weak limit of probability rescaled counting measure on $\{0,1,2,\dots,n\}$; then $(\mathscr{C},\delta,\nu,0)$ is a measured real tree rooted at 0 that we call the *cut tree*; $\mathbb{N}_0=\{0,1,2,\dots\}$ should be seen as a collection of marks in $\mathscr{C}$. The measured tree $(\mathscr{C},\delta,\nu,0)$ is distributed like a Brownian continuum random tree, and the collection of points $\mathbb{N}\subseteq\mathscr{C}$ is an i.i.d. sequence with common distribution ν [5, 21, 25].

For each $s\in[0,1]$, let $\Gamma_\tau(s):=\{i\in\mathbb{N}:\zeta_i\sim_\tau s\}$. Then, each $i\in\mathbb{N}$ is the image of ζ_i in the cut tree $\mathscr{C}$ in the sense that $\Gamma_\tau(\zeta_i)$ converges in $\mathscr{C}$ as $\tau\to\infty$ to the singleton $\{i\}$ (see [5]). Every branch point of $\mathscr{C}$ corresponds to a fragmentation event, just as reflected by the definition in (15). For any $i,j\in\mathbb{N}$, let $i\curlywedge j$ be the common ancestor of i and j in $\mathscr{C}$, that is the point at distance

$$\int_0^{\tau_{ij}} |I^\tau(\zeta_i)|d\tau=\int_0^{\tau_{ij}} |I^\tau(\zeta_j)|d\tau$$

from 0 on the paths between 0 and i, and between 0 and j. Here, $i\curlywedge_{\mathscr{C}} j$ corresponds to the (unique) fragmentation event that occurs at time τ_{ij}, and that separates ζ_i from ζ_j. Let $\mathscr{C}^i_{i\curlywedge j}$ and $\mathscr{C}^j_{i\curlywedge j}$ be the two subtrees of $\mathscr{C}$ above the point $i\curlywedge j$ that contain respectively i and j; then for every $k\in\mathbb{N}$ we have $\zeta_i\sim_{\tau_{ij}}\zeta_k$ precisely if $k\in\mathscr{C}^i_{i\curlywedge j}$. Furthermore, the interval $I^{\tau_{ij}-}(\zeta_i)=I^{\tau_{ij}-}(\zeta_j)$ which contains all the ζ_k for which $\zeta_k\sim_t\zeta_i$ for all $i<\tau_{ij}$ splits into the two intervals $I^{\tau_{ij}}(\zeta_i)$ and $I^{\tau_{ij}}(\zeta_j)$ by the removal of the unique point of $Z^{-\tau_{ij}}$ lying in the interior of $I^{\tau_{ij}-}(\zeta_i)$.

Observe that the cut tree is only constructed from the process of masses of the fragments containing a sequence of i.i.d. uniform points; this is crucial since the "identities" of the fragments seen as subsets of $[0,1]$ retain some information (for instance, only neighbouring intervals can merge). More precisely, we can do so using only the process of masses, by exchangeability of $(\zeta_i)_{i\geq 1}$.

If we see the fragmentation $(F(\tau))_{\tau\geq 0}$ from the point of view of Aldous and Pitman in [15], the cut tree is the genealogy of the fragmentation of a Brownian continuum random tree, and it is natural to try to ask whether one can recover the initial tree $(\mathscr{T},d,\mu)$ from $\mathfrak{C}=(\mathscr{C},\delta,\nu,0)$, or if not, what minimal additional information is necessary. This question has been studied by Broutin and Wang [25] and Addario-Berry, Dieuleveut, and Goldschmidt [5] (see also [3] for a partial result). Quite naturally, since the cut tree is constructed from the process of masses only, the locations of the cuts are lost, and reconstruction is impossible without additional information. The main result of [5, 25] is that these locations is the only information that is lost, and that one can recover $(\mathscr{T},d,\mu)$ from $(\mathscr{C},\delta,\nu)$ plus this additional information.

Since the fragmentation is binary, for every fragmentation event, there should correspond two points, one in each of the two fragments created. It turns out that these points can be given through their images in $\mathscr{C}$: the additional information comes in the form of a countable collection of marks in the cut tree, and the only relevant information to us is its distribution conditionally on $(\mathscr{C},\delta,\nu)$. Let $\mathrm{Br}(\mathscr{C})$ denote the set of branch points of $\mathscr{C}$. Almost surely, for each $b\in\mathrm{Br}(\mathscr{C})$, there are precisely three connected components to $\mathscr{C}\setminus\{b\}$, and we denote by $\mathscr{C}'_b$ and $\mathscr{C}''_b$ the two which are not containing 0, agreeing that $\nu(\mathscr{C}'_b)>\nu(\mathscr{C}''_b)$. Let $\mathbf{V}=\{(V'_b,V''_b):b\in\mathrm{Br}(\mathscr{C})\}$ be an independent family of random variables such

that, for each $b \in \mathrm{Br}(\mathscr{C})$, (V_b', V_b'') has distribution

$$\frac{\nu(\cdot \cap \mathscr{C}_b')}{\nu(\mathscr{C}_b')} \otimes \frac{\nu(\cdot \cap \mathscr{C}_b'')}{\nu(\mathscr{C}_b'')} \,. \tag{16}$$

The inverse cut tree transform then goes as follows: there exists a (measurable) map Φ that associates, to a pair $(\mathfrak{C}, \mathbf{V})$ a measured real tree that is distributed like a Brownian CRT. We will verify that $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ turns out to be $\Phi(\mathfrak{C}, \mathbf{V})$ for a suitable collection $\mathbf{V}$, but for now, let us describe the procedure if $\mathbf{V}$ is given and has the distribution described above (this follows [5]).

For $i, j \in \mathbb{N}$, we can recursively identify a collection of branch points in $\mathscr{C}$, which are meant to correspond to the cut points on the path between ζ_i and ζ_j in $\mathscr{T}$. With this goal in mind, we now define a collection (C_u, p_u^0, p_u^1), $u \in \mathcal{U}_2$, where C_u is a subtree of $\mathscr{C}$, and $p_u^0, p_u^1 \in C_u$. First set $C_\varnothing = \mathscr{C}$ and let $p_\varnothing^0 = i$, $p_\varnothing^1 = j$. Then, given (C_u, p_u^0, p_u^1), and writing $b = p_u^0 \curlywedge p_u^1$, let C_{u0} (resp. C_{u1}) be the one among $\mathscr{C}_b'$ and $\mathscr{C}_b''$ which contains p_u^0 (resp. p_u^1). Let $p_{u0}^0 = p_u^0$, $p_{u1}^1 = p_u^1$ and then define p_{u0}^1 (resp. p_{u1}^0) be the one of V_b' and V_b'' that lies in C_{u0} (resp. C_{u1}). Then for each $n \geq 0$ define

$$Y_n(i,j) = \sqrt{\frac{\pi}{2}} \cdot \sum_{|u|=n} \nu(C_u)^{1/2} \,.$$

Almost surely for all i, j, $Y_n(i,j) \to Y(i,j)$ as $n \to \infty$. Then the collection of random variables $(Y(i,j) : i, j \in \mathbb{N})$ has the same distribution as $(\delta(i,j), i, j \in \mathbb{N})$. Seen as a matrix of pairwise distances, this defines uniquely an isometry class of a random compact real tree, which is a Brownian continuum random tree.

4.4 The convex minorant tree as the inverse cut tree transform

From the previous considerations, proving that $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ is indeed a Brownian CRT boils down to verifying that it can be seen as obtained from the inverse cut tree transform from $\mathfrak{C}$ using a certain collection of points that we will denote by $\{(\beta_b', \beta_b''), b \in \mathrm{Br}(\mathscr{C})\}$. Our collection is in part constructed as a measurable function of e alone, and the main task consists in verifying that it has indeed the same distribution as $\mathbf{V}$ defined above in (16).

We start with a canonical exploration of the fragmentation. We construct a process $(S_u, \tau_u, \epsilon_u)_{u \in \mathcal{U}}$ where S_u is a half-open interval of $[0,1)$, τ_u is the unique time when there exists $x \in [0,1]$ such that $S_u = I^{\tau_u}(x)$ (that is the interior of S_u is a connected component of $[0,1) \setminus Z^{-\tau_u}$); furthermore, writing $\ell_u = |S_u|$, ϵ_u is a continuous function on $[0, \ell_u]$ with $\epsilon_u(0) = \epsilon_u(\ell_u) = 0$ and $\epsilon_u(r) > 0$ on $[0, \ell_u]$. It will also be convenient to write $a_u = \inf S_u \in S_u$. The precise order in which the intervals and times are associated with the elements of $\mathcal{U}$ is key to control the independence structure which turns out to be crucial.

We first set $S_\varnothing = [0,1)$, $\tau_\varnothing = 0$, and $\epsilon_\varnothing = \mathrm{e}$; we then have $\ell_\varnothing = 1$ and $a_\varnothing = 0$. For $x \in [0,1)$ and $t \geq 0$, let $R_t(x) = \sup I^t(x)$. The process $(R_t(a_\varnothing))_{t \geq \tau_\varnothing} = (R_t(0))_{t \geq 0}$ has countably many negative jumps. We let $\ell_1 > \ell_2 > \cdots \geq 0$ denote their ranked sizes (in absolute value); then $\sum_i \ell_i = 1$ almost surely. For each $i \geq 1$, we let τ_i be the unique $t \geq \tau_\varnothing = 0$ with $R_{t-}(a_\varnothing) - R_t(a_\varnothing) = \ell_i$, and define $S_i = [R_{\tau_i}, R_{\tau_i -})$; one then has $a_i = R_{\tau_i}$. We then let $\epsilon_i : [0, \ell_i] \to \mathbb{R}_+$ be defined for $r \geq 0$ by

$$\epsilon_i(r) = e^{\tau_i}(a_i + r)\mathbf{1}_{0 \leq r \leq \ell_i} \,.$$

Now, for each $u \in \mathcal{U}$, given S_u and τ_u, let $(\tau_{ui}, \ell_{ui})_{i \geq 1}$ denote the times $\tau_{ui} \geq \tau_u$ and sizes of the jumps of the process $(R_t(a_u))_{t \geq \tau_u}$ sorted in such a way that $\ell_{u1} > \ell_{u2} > \cdots \geq 0$. Write $S_{ui} = [R_{\tau_{ui}}(a_u), R_{\tau_{ui}-}(a_u))$, $a_{ui} = \inf S_{ui}$, $\ell_{ui} = |S_{ui}|$ and define $\epsilon_{ui} : [0, \ell_{ui}] \to \mathbb{R}_+$ for $r \geq 0$ by

$$\epsilon_{ui}(r) = e^{\tau_{ui}}(a_{ui} + r)\mathbf{1}_{r \leq \ell_{ui}} = \epsilon_u^{\tau_{ui} - \tau_u}\left(\sum_{j \geq 1} \ell_{uj} \mathbf{1}_{a_{uj} < a_{ui}} + r \right) \mathbf{1}_{0 \leq r \leq \ell_{ui}} \,.$$

Let $\mathcal{F}_\varnothing$ be the sigma-algebra generated by $(R_t(0))_{t\geq 0}$. Then $(S_i, \ell_i, \tau_i, a_i)_{i\geq 1}$ is $\mathcal{F}_\varnothing$-measurable while, conditionally on $\mathcal{F}_\varnothing$, the $(\epsilon_i)_{i\geq 1}$ are independent Brownian excursions of durations $\ell_1 > \ell_2 > \cdots \geq 0$. More generally, let $\mathcal{F}_u$ be the sigma-algebra generated by $\{(R_t(a_v))_{t\geq \tau_v}, v \preceq u\}$. Then $(S_{vi}, \ell_{vi}, \tau_{vi}, a_{vi})_{v\preceq u, i\geq 1}$ is $\mathcal{F}_u$-measurable while, conditionally on $\mathcal{F}_u$, the functions $(\epsilon_{ui})_{i\geq 1}$ are independent Brownian excursions of durations $\ell_{u1} > \ell_{u2} > \cdots \geq 0$.

The recursive exploration $(S_u, \tau_u, \epsilon_u)_{u\in\mathcal{U}}$ we have just defined yields a canonical recursive spinal decomposition of the cut tree $\mathscr{C}$; by canonical we mean that the random points that are used are constructed from e only. We say that a point $s \in [0,1]$ has an image $x \in \mathscr{C}$ if $\Pi_t(s) = \overline{\{j \in \mathbb{N} : \zeta_j \sim_t s\}}$ decreases to the singleton $\{x\}$ as $t \to \infty$. We let $\mathscr{C}_\varnothing = \mathscr{C}$ and $b_\varnothing = 0$. Working towards the definition of (β'_b, β''_b), $b \in \mathrm{Br}(\mathscr{C})$, we start by defining a collection η_u, $u \in \mathcal{U}$. In the following, for $x, y \in \mathscr{C}$, $[\![x, y]\!]_\mathscr{C}$ denotes the range of the unique geodesic in $\mathscr{C}$ between x and y.

Lemma 4.6. *With probability one, the points $(a_u)_{u\in\mathcal{U}}$ have images in $\mathscr{C}$ that we denote by $(\eta_u)_{u\in\mathcal{U}}$. They are defined inductively and satisfy:*
- *$\eta_\varnothing$ is the image of $a_\varnothing$;*
- *given $\mathscr{C}_u$ and $\eta_u \in \mathscr{C}_u$ the points b_{ui} are the points of $[\![b_u, \eta_u]\!]_\mathscr{C}$ at distance $\int_{\tau_u}^{\tau_{ui}} |I^t(a_u)|dt$ from b_u;*
- *$\mathscr{C}_{ui}$ is the subtree of $\mathscr{C}_u \setminus \{b_{ui}\}$ which contains neither b_u nor η_u;*
- *η_{ui} is the image of a_{ui} in $\mathscr{C}$, which turns out to be in $\mathscr{C}_{ui}$.*

Furthermore, the family $(\eta_u)_{u\in\mathcal{U}}$ is independent and for each $u \in \mathcal{U}$, η_u has distribution $\nu(\,\cdot\cap\mathscr{C}_u)/\nu(\mathscr{C}_u)$.

Proof. The proof is by induction. It is proved in [20] that the process $(|I^t(0)|)_{t\geq 0}$ has the same distribution as the process $(|I^t(\zeta_1)|)_{t\geq 0}$. Therefore with $\Pi_t(0) := \overline{\{i \in \mathbb{N} : \zeta_i \in I^t(0)\}}$, we have almost surely $\sup\{\delta(i,j) : i, j \in \Pi_t(0)\} \to 0$ as $t \to \infty$ so that there is a limit point that we denote by $\eta_\varnothing$ such that $\Pi_t(0) \to \{\eta_\varnothing\}$; by definition $\eta_\varnothing$ is the image of $a_\varnothing$ in $\mathscr{C}$. It also follows that $\eta_\varnothing$ has distribution ν in $\mathscr{C}$, since 1, the image of ζ_1 in $\mathscr{C}$, does. The points b_i, $i \geq 1$, are precisely the branch points of $\mathscr{C}$ along the segment $[\![0, \eta_\varnothing]\!]_\mathscr{C}$, sorted in the decreasing order of the masses $\nu(\mathscr{C}_i) = \ell_i$ of the subtrees of $\mathscr{C}$ hanging from the segment.

Observe now that for $u \in \mathcal{U}$ and $i \geq 1$, conditionally on $\mathcal{F}_u$, the process $(I^{\tau_{ui}+t}(a_{ui}))_{t\geq 0}$ is precisely the process of masses of the fragment containing 0 in the fragmentation of the excursion ϵ_{ui}. As a consequence, the image η_{ui} of a_{ui} is well-defined. Furthermore, the distribution of η_{ui} is the rescaled mass measure ν in the image of S_{ui} in $\mathscr{C}$, which is precisely $\mathscr{C}_{ui}$. Finally, conditionally on $\mathcal{F}_u$, the functions $(\epsilon_{u_i})_{i\geq 1}$ are independent, and so are the $(\eta_{ui})_{i\geq 1}$: for any collection of bounded continuous functionals $(f_i)_{i\geq 1}$, we have

$$\mathbf{E}\left[\prod_{i\geq 1} f_i(\eta_{ui}) \,\middle|\, \mathcal{F}_u\right] = \prod_{i\geq 1} \int_{\mathscr{C}_{ui}} f_i(x_i) \frac{\nu(dx_i)}{\nu(\mathscr{C}_{ui})}\,.$$

The claim follows by induction. □

Remark 4.7. Observe that, $\mathscr{C}_u$ is the complete subtree of $\mathscr{C}$ induced by $\{i \in \mathbb{N} : \zeta_i \in S_u\}$. Furthermore, except for $b_\varnothing$, the b_u are all branch points in $\mathscr{C}$; more precisely b_{ui} is the common ancestor of η_u and η_{ui}, namely $b_{ui} = \eta_u \curlywedge \eta_{ui}$.

The collection $(\eta_u)_{u\in\mathcal{U}}$ only provides part of the marks we shall need in the cut tree $\mathscr{C}$. The remaining marks are the images of the random points constructed using the sequence of uniform random variables $\mathbf{U}$, and which are associated to the local minima of e.

Lemma 4.8. *There is a one-to-one correspondence between the local minima of* e *and the branch points of the cut tree $\mathscr{C}$: every branch point of $\mathscr{C}$ is of the form $\eta_u \curlywedge \eta_{ui}$ for $u \in \mathcal{U}$ and $i \geq 1$, and the corresponding local minimum is a_{ui}.*

Proof. For each $t \in \mathscr{L}(e)$, let $t_0 = 0 < t_1 < t_2 < \cdots < t_i = t$ be the vertices of the convex minorant of e on the interval $[0, t]$. Let $z_0 = 1 > z_1 > \cdots > z_k$ and $\gamma_1 < \gamma_2 < \cdots < \gamma_k$ be the corresponding

intercepts and slopes. Then, at time γ_k the interval $[t_{k-1}, z_k)$ is split into the pair $[t_{k-1}, t_k)$, $[t_k, z_k)$. Let $j_1 = \inf\{j \in \mathbb{N} : \zeta_{j_1} \in [t_{k-1}, t_k)\}$ and $j_2 = \inf\{j \in \mathbb{N} : \zeta_{j_2} \in [t_k, z_k)\}$. To make the correspondence more explicit, we exhibit the two points η_u and η_{ui} in $\mathscr{C}$ such that the branch point corresponding to t is $\eta_u \curlywedge \eta_{ui} = j_1 \curlywedge j_2$. It shall be noted that the branch point corresponding to $t = t_i$ is not the image of t in the cut tree $\mathscr{C}$, the latter being almost surely the leaf η_{ui} that we will exhibit. The path to follow in $\mathcal{U}$ is given by the convex minorant. Let $i_1 \geq 1$ be the unique index such that $\tau_{i_1} = \gamma_1$; then, let i_2 be the unique index such that $\tau_{i_1 i_2} = \gamma_2$, and so on which yields a point $u = i_1 i_2 \dots i_{k-1}$ with $\tau_u = \gamma_{k-1}$, and $a_u = t_{k-1}$. Finally, let i be the unique index such that $\tau_{ui} = \gamma_k$; then we have $a_u = t_k = t$ while $z_k = R_{\tau_{ui}-}$. The images η_u and η_{ui} of t_u and t_{ui} in $\mathscr{C}$ are such that $\eta_u \curlywedge \eta_{ui}$ is the branch point $j_1 \curlywedge j_2$.

Conversely, the sequence of sets $\{[\![0, \eta_u]\!]_{\mathscr{C}} : |u| \leq n\}$, $n \geq 1$, increases to $\mathscr{C}$ and thus exhausts all the branch points. In particular, every branch point b of $\mathscr{C}$ is of the form $\eta_u \curlywedge \eta_{ui}$ for some $u \in \mathcal{U}$ and $i \geq 1$. Now, for such a branch point, $a_{ui} \in [0, 1]$ is the local minimum of e that separates the points from $[a_u, a_{ui})$ from S_{ui} at time τ_{ui}. □

Finally, we complete the definition of the set of marks in the cut tree $\mathscr{C}$. Consider a branch point b of $\mathscr{C}$; by Lemma 4.8, it is of the form $\eta_u \curlywedge \eta_{ui}$ for some $(u, i) \in \mathcal{U} \times \mathbb{N}$ and a_{ui} is the corresponding local minimum. Recall now the join point $\boldsymbol{\jmath}(a_{ui})$ associated to $a_{ui} \in \mathscr{L}$ (Remark 3.11 on page 14). Observe that, by construction, at time τ_{ui}, the two intervals that get separated are S_{ui} to the right, and $[a_u, a_{ui})$, to the left. The subtree of $\mathscr{C}$ above the branch point b is therefore the completion of $\{i \in \mathbb{N} : \zeta_i \in [a_u, \sup S_{ui})\}$, and two intervals $[a_u, a_{ui})$ and S_{ui} correspond to the two subtrees of $\mathscr{C}$ above the branch point b, that we previously denoted by $\mathscr{C}'_b$ and $\mathscr{C}''_b$.

Recall that $\mathcal{F}_u$ is the sigma-algebra generated by $\{(R_t(a_v))_{t \geq \tau_v}, v \preceq u\}$, and that, as a consequence, a_u, a_{ui} and S_{ui} are $\mathcal{F}_u$-measurable. By Lemma 4.6, conditionally on $\mathcal{F}_u$, the image η_{ui} of a_{ui} in $\mathscr{C}$ is distributed like $\nu(\cdot \cap \mathscr{C}_{ui})/\nu(\mathscr{C}_{ui})$. Let $(\beta'_b, \beta''_b) \in \mathscr{C}'_b \times \mathscr{C}''_b$ be the pair of points formed by η_{ui} and the image of $\boldsymbol{\jmath}(a_{ui})$ in $\mathscr{C}$ (which a.s. exists since $\boldsymbol{\jmath}(a_{ui})$ is uniform in $[a_u, a_{ui})$). Then, conditionally on $\mathcal{F}_u$, and by Lemma 4.6, the collection (β'_b, β''_b), $b \in \mathrm{Br}(\mathscr{C})$ has the same distribution as $\mathbf{V}$:

Lemma 4.9. *The marked cut tree $(\mathscr{C}, \{\beta'_b, \beta''_b : b \in \mathrm{Br}(\mathscr{C}))$ is such that:*

i) $\{(\beta'_b, \beta''_b) : b \in \mathrm{Br}(\mathscr{C})\}$ are independent conditionally on $\mathscr{C}$, and

ii) for each b, (β'_b, β''_b) are independent random variables with distribution $\nu(\cdot \cap \mathscr{C}'_b)/\nu(\mathscr{C}'_b) \otimes \nu(\cdot \cap \mathscr{C}''_b)/\nu(\mathscr{C}''_b)$.

The points (β'_b, β''_b), $b \in \mathscr{C}$, now being defined, we are ready to verify that $\Phi(\mathfrak{C}, \boldsymbol{\beta})$ and $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ are almost surely isometric. The arguments above should already make this pretty clear: indeed, for each $b \in \mathrm{Br}(\mathscr{C})$, the set of marks $\{\beta'_b, \beta''_b\}$ is precisely the image in $\mathscr{C}$ the set of points $\{\boldsymbol{\jmath}(a_{ui}), a_{ui}\}$ which are identified at time τ_{ui}. To make this formal, fix any $i, j \in \mathbb{N}$, and consider $Y_n(i, j)$ and $\hat{d}_n(\zeta_i, \zeta_j)$. The choice of the marks (β'_b, β''_b), $b \in \mathrm{Br}(\mathscr{C})$, is precisely made so that, for every $n \geq 1$, sorting the sets $\{\nu(C_u), |u| = n\}$ and $\{|\Pi_u|, |u| = n\}$ in decreasing order yields the same sequence, and therefore

$$Y_n(i, j) = \sqrt{\frac{\pi}{2}} \sum_{|u|=n} \nu(C_u)^{1/2} = \sqrt{\frac{\pi}{2}} \sum_{|u|=n} |\Pi_u|^{1/2} = \hat{d}_n(\zeta_i, \zeta_j) = d_n(\zeta_i, \zeta_j)\,,$$

where the last step follows from Proposition 4.5. Taking the limit as $n \to \infty$, this implies that, for each $k \geq 1$ the metric spaces $(\{\zeta_i, 1 \leq i \leq k\}, d)$ and $([k], Y)$ are isometric (with the correspondence (i, ζ_i), $i \in [k]$). Since $(\{\zeta_i, 1 \leq i \leq k\}, d)$ increases to $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$ (Proposition 3.22), the claim follows by taking the limit as $k \to \infty$.

Finally, we are ready to prove Theorem 1.5 which shows that the convex minorant tree provides a coupling between the two classical constructions of the additive coalescent by Aldous & Pitman [15] on the one hand, and Bertoin [20] on the other. Let $\mathcal{P} = \{(\pi(x), \boldsymbol{\mathfrak{s}}(x)) : x \in \mathscr{L}(e)\}$.

Proof of Theorem 1.5. Observe that, by Lemma 4.8, with probability one, all the local minima of e are of the form $\eta_u \curlywedge \eta_{ui}$ defined in Lemma 4.3, and therefore, almost surely, $\mathcal{P} = \{(\pi(a_{ui}), \tau_{ui}) :$

$u \in \mathcal{U}, i \in \mathbb{N}\}$. Since, a.s. for all $u \in \mathcal{U}$, we have $\ell_u > 0$, this can equivalently be put as $\mathcal{P} = \{(\pi(\kappa(\zeta_i, \zeta_j)), \tau(\zeta_i, \zeta_j)) : i \neq j \in \mathbb{N}\}$. From there, the claim is an easy consequence of Theorem 16, and Corollaries 17-18 of [5] (it is even simpler since we do not need to infer the $\tau(\zeta_i, \zeta_j)$ from the cut tree, they can be read directly from the fragmentation). □

5 Compactness of the Brownian parabolic tree $\mathrm{CMT}(X, \mathbf{U})$

In this section, we prove the compactness of $\mathrm{CMT}(X, \mathbf{U})$ constructed in Section 3.7. The completeness is plain from the definition and we only need to verify that $(\mathscr{M}, d)$ is totally bounded.

We will proceed by controlling the growth of a well-chosen sequence of subspaces that increase to $\mathscr{M}$ using a chaining argument. This leads us to a process that is reminiscent to a certain version of Prim's exploration at the discrete level, and that also turns out to be crucial in the calculation of the Hausdorff dimension (see Section 6). The general strategy is inspired from the arguments of Curien and Haas [30] for the compactness of trees constructed by aggregation of segments.

5.1 The growth process

In the entire section, we consider the process X, and the random variables $Z^\lambda = Z^\lambda(X)$ refer to this case. We may see the metric space $(\mathscr{M}, d)$ as obtained from the coalescent process induced by Z^λ on $\mathbb{R}_+$, which turns out to be the standard multiplicative coalescent [17, 24]. In this process, fragments only merge by pairs, but obtaining fine quantitative estimates is delicate since for any $\lambda \in \mathbb{R}$ and $h > 0$, $Z^\lambda \setminus Z^{\lambda+h}$ is a.s. not contained in any compact interval. We shall thus track a single connected component as λ increases.

Remark 5.1. The most natural choice of a connected component to track is the largest one, but this leads to some inconvenient conditioning. One could also track a connected component containing a fixed point (at the discrete level), but, without any additional structure, such a node must be uniformly random, and thus the corresponding component would be too small to lead to anything interesting. Here, the structure imposed by the representation on $\mathbb{R}_+$ allows us to track any fixed point *among the ones that do matter* even though they are a negligible for the mass measure (that is, the nodes v_i for $\epsilon n^{2/3} \leq i \leq Cn^{2/3}$ for constants $0 < \epsilon < C$).

For each $\lambda \in \mathbb{R}$, let $L_\lambda = \sup Z^\lambda \cap [0, 1)$ and $R_\lambda = \inf Z^\lambda \cap (1, \infty)$, and define $H_\lambda = [L_\lambda, R_\lambda)$. So, up to inclusion of the left-most point, $H_\lambda \subseteq \mathbb{R}_+$ is the interval of $\mathbb{R}_+ \setminus Z^\lambda$ which contains the point 1; since $1 \notin \bigcup_\lambda Z^\lambda$ with probability one, this is well-defined for all $\lambda \in \mathbb{R}$. We have the following asymptotics, whose proofs are found in Section 5.2.

Lemma 5.2. *There exist constant $c > 0$ and x_0, λ_0 such that, for all $x > x_0$ and $\lambda \geq \lambda_0$ we have*

$$\mathbf{P}(L_\lambda \geq \tfrac{x}{\lambda^2}) \leq e^{-cx} \qquad \text{and} \qquad \mathbf{P}(|R_\lambda - 2\lambda| > 1) \leq e^{-c\lambda}\,.$$

It follows that $H_\lambda \uparrow (0, \infty)$ as $\lambda \to \infty$, so that H_λ provides a suitable increasing family of subspaces of $\mathscr{M}$. So for any $x \in \mathbb{R}_+ \setminus \{0, 1\}$, we let $\lambda(x) = \inf\{\lambda : x \in H_\lambda\}$ be the time at which x joins the connected component containing 1. The intervals that join H_λ play a different role depending on whether they lie to the left or to the right, and we define $\Lambda^R = \{\lambda \in \mathbb{R} : R_\lambda > R_{\lambda-}\}$ and $\Lambda^L := \{\lambda \in \mathbb{R} : L_\lambda < L_{\lambda-}\}$; then Λ^R and Λ^L are both countable. Furthermore, they are almost surely disjoint; this is because the standard multiplicative coalescent is binary [14], and could also be proved directly from the representation with X (see the proof of Proposition 7.7). We let $\Lambda = \Lambda^R \cup \Lambda^L$.

For $x, y \in \mathbb{R}_+$, we let $x \leftrightarrow y$ if $\lambda(x) = \lambda(y)$, and let q_λ, $\lambda \in \Lambda$, be the equivalence classes of this relation. For $\lambda \in \Lambda^L$, we have $q_\lambda = [L_\lambda, L_{\lambda-})$, while $q_\lambda = [R_{\lambda-}, R_\lambda)$ for $\lambda \in \Lambda^R$. Because of this, $(q_\lambda)_{\lambda \in \Lambda^L}$ and $(q_\lambda)_{\lambda \in \Lambda^R}$ respectively define partitions of $(0, 1)$ and $(1, \infty)$ into countably many disjoint intervals.

For an interval $I \subset \mathbb{R}_+$, we let $\mathscr{M}|_I$ be the *intrinsic* metric space induced by d on I: this is the metric space (I, d_I) where $d_I(x,y) = d(x,y)$ if $[\![x,y]\!] \subseteq I$ and $d_I(x,y) = +\infty$ otherwise. So in general, $\mathscr{M}|_I$ might be disconnected. Let $\mathscr{M}_\lambda := \mathscr{M}|_{H_\lambda}$; then $\mathscr{M}_\lambda$ is connected for all $\lambda \in \mathbb{R}$. For $\lambda \in \Lambda^R$ (resp. Λ^L), we also let $\mathscr{T}^R_\lambda := \mathscr{M}|_{q_\lambda}$ (resp. $\mathscr{T}^L_\lambda := \mathscr{M}_{q_\lambda}$). Quite plainly, and up to the metric completion, the metric space $\mathscr{M}$ is obtained by combining the $\mathscr{T}^R_\lambda$, $\lambda \in \Lambda^R$, and $\mathscr{T}^L_\lambda$, $\lambda \in \Lambda^L$, using the identifications performed during the construction (using the random points constructed from $\mathbf{U}$). This process actually turns out rather agreeable: the metric spaces $\mathscr{T}^R_\lambda$ and $\mathscr{T}^L_\lambda$ are easy to understand because they are small as $\lambda \to \infty$, and the way they are put together is also easy to control. Informally, the dynamics as λ increases are as follows:

- at time $\lambda \in \Lambda^R$, $\mathscr{T}^R_\lambda$ merges with $\mathscr{M}_{\lambda-}$ by identifying $\inf q_\lambda$ with a uniform point in $H_{\lambda-}$;
- at time $\lambda \in \Lambda^L$, $\mathscr{M}_{\lambda-}$ connects with $\mathscr{T}^L_\lambda$ by identifying $L_{\lambda-}$ with a uniform point in $\mathscr{T}^L_\lambda$.

The $(\mathscr{M}_{k^3})_{k\geq 1}$ is the convenient sequence of subspaces of $\mathscr{M}$ that we mentioned before. The following decomposition which takes advantage of these dynamics will be useful. What matters for now is the global picture, we will fill out the details later on.

- *The annuli of forests to the right.* For any $k \geq 1$ define $\Lambda_k := \{\lambda \in \Lambda^R : k^3 < \lambda \leq (k+1)^3\}$, and let $\mathscr{F}_k$ be the intrinsic metric space induced by $\mathscr{M}$ on $(R_{k^3}, R_{(k+1)^3}]$. For each $k \geq 1$, $\mathscr{F}_k$ is a forest consisting of infinitely many connected components obtained when only the identifications within $(R_{k^3}, R_{(k+1)^3}]$ are performed. Our aim is to bound the maximum diameter of the connected component $\mathscr{F}_k$ in order to control the worst case accumulation of length when putting all the $\mathscr{F}_k$ together. Formally, for $\lambda, \lambda' \in \Lambda_k$ we write $\lambda \equiv_k \lambda'$ if $R_{\lambda-} \wedge R_{\lambda'-} > R_{k^3}$, which implies that $\mathscr{T}^R_\lambda$ and $\mathscr{T}^R_{\lambda'}$ are connected within $\mathscr{F}_k$. We will prove that, almost surely, every equivalence class of this relation is finite. So for each $\lambda \in \Lambda_k$, we may define $\rho_\lambda := \min\{R_{\lambda'-} : \lambda' \equiv_k \lambda\}$ as the leftmost point of the connected component containing $\mathscr{T}^R_\lambda$ within $\mathscr{F}_k$. More generally, the equivalence relation $\equiv_k$ naturally extends as follows: for $x, y \in (R_{k^3}, R_{(k+1)^3}]$ we let $x \sim_k y$ if $x \in q_\lambda$, $y \in q_{\lambda'}$ and $\lambda \equiv_k \lambda'$. Defining the diameter of a potentially disconnected metric space as the supremum of the diameters of its connected components, we therefore have $\mathrm{diam}(\mathscr{F}_k) = \sup\{d(x,y) : x, y \in (R_{k^3}, R_{(k+1)^3}], x \equiv_k y\}$.
- *The chain of beads to the origin.* For any $\lambda \in \mathbb{R}$, we let $\mathscr{P}_\lambda$ denote the intrinsic metric space induced by $\mathscr{M}$ on $[0, L_\lambda)$. This metric is almost surely connected and has the structure of a "string of beads" that we now describe. For $a \in \mathbb{R}$, let $\Lambda^L_a := \Lambda^L \cap (a, \infty)$. Almost surely for any $a \in \mathbb{R}$, Λ^L_a only contains finitely many points in any compact interval of $(0, +\infty)$, so that we may enumerate its elements in increasing order as $(\lambda_i)_{i\geq 1}$ (Lemma 3.3). The metric space $\mathscr{P}_\lambda$ is obtained by putting together the metric spaces $\mathscr{T}^L_{\lambda_j}$, $j \geq 1$, into a chain by connecting $\mathscr{T}^L_{\lambda_i}$ to a uniform random point in $\mathscr{T}^L_{\lambda_{i+1}}$ for each $i \geq 1$.

By giving an estimate of the extent of H_λ, Lemma 5.2 provides an effective way to control the locations of the "gluing points" which lies at the core of the proofs of the compactness and of the computation of the Hausdorff dimensions. The contribution of $\mathscr{P}_\lambda$ is easily treated separately, and the crucial steps consists in controlling the diameters of the $\mathscr{F}_k$, $k \geq 1$.

Proposition 5.3 (Diameter of annuli forests)**.** *There exists $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$ we have*

$$\mathbf{P}\left(\mathrm{diam}(\mathscr{F}_k) > k^{-3/2}\right) \leq 11 k^{-5/4}\,.$$

Proposition 5.4 (Diameter of the string of beads)**.** *Almost surely,*

$$\lim_{\lambda\to\infty} \mathrm{diam}(\mathscr{P}_\lambda) = 0\,.$$

Taking Propositions 5.3 and 5.4 for granted for now, the proof of compactness is then straightforward.

Proposition 5.5 (Compactness of $\mathscr{M}$)**.** *The metric space $(\mathscr{M}, d)$ is almost surely compact.*

Proof. By Proposition 5.3 and the Borel–Cantelli lemma, with probability one, there exists an almost surely finite random variable K such that $\mathrm{diam}(\mathscr{F}_k) \le k^{-3/2}$ for all $k \ge K$. In particular, for all $k \ge K$,

$$\mathrm{d_H}(\mathscr{M}_{k^3}, \mathscr{M}) \le 2k^{-1/2} + \mathrm{diam}(\mathscr{P}_{k^3})\,.$$

Fix any $\epsilon > 0$. By Proposition 5.4, $\mathrm{diam}(\mathscr{P}_{k^3}) < \epsilon/3$ for all k large enough; it follows that there exists $k \ge K$ large enough such that $\mathrm{d_H}(\mathscr{M}_{k^3}, \mathscr{M}) < \epsilon/2$. Recall that, for any fixed $\lambda \in \mathbb{R}$, the restriction $\mathscr{M}_\lambda$ of $\mathscr{M}$ to H_λ is almost surely compact by absolute continuity with the Brownian continuum random tree. So there exists a cover of $\mathscr{M}_{k^3}$ by finitely many balls of radius $\epsilon/2$; increasing the radius of each ball to ϵ yields a finite cover of $\mathscr{M}$. We have thus proved that $\mathscr{M}$ is totally bounded. Since it is complete by definition, it is compact. □

The remainder of the section is devoted to the proof of Propositions 5.3 and 5.4. We first prove Lemma 5.2 in Section 5.2. The forests $\mathscr{F}_k$ are made of the trees $\mathscr{T}_\lambda^R$, for $\lambda \in \Lambda_k$. For all $k \ge 1$, Λ_k is infinite, which causes some difficulties. Still, we expect that, for large k, the components $\mathscr{T}_\lambda^R$ with $\lambda \in \Lambda_k$ should be rather small; Section 5.3 deals with the question of uniform bounds on distances in the $\mathscr{T}_\lambda^R$ in terms of the lengths $|q_\lambda|$. We then obtain in Section 5.4 the relevant statistics about the connected components $\mathscr{T}_\lambda^R$ for $\lambda \in \Lambda_k$ which includes information about the lengths $|q_\lambda|$ and their diameters. In Section 5.5 we put together all the pieces and prove Proposition 5.3 which essentially says that the diameter of $\mathscr{F}_k$ is comparable to the maximum diameter of the $\mathscr{T}_\lambda^R$, $\lambda \in \Lambda_k$. Finally, we prove Propostition 5.4 in Section 5.6.

5.2 The position of the component containing 1: Proof of Lemma 5.2

Recall that for a continuous process $\omega = (\omega_t)_{t\ge 0}$, we let $\underline{\omega}$ and $\overline{\omega}$ denote respectively the running infimum and supremum processes: $\underline{\omega}_t := \inf\{\omega_s : 0 \le s \le t\}$ and $\overline{\omega}_t = \sup\{\omega_t : 0 \le s \le t\}$.

Recall that $(W_t)_{t\ge 0}$ denotes a standard Brownian motion. We will use repeatedly the following simple fact (see, e.g., [41] page 96, consequence of the fact that $\overline{W}_t$ has same law as $|W_t|$, for a fixed t): for all $x \ge 0$, we have

$$\mathbf{P}(\overline{W}_t \ge x) = \mathbf{P}(\underline{W}_t \le -x) \le e^{-x^2/(2t)}\,. \tag{17}$$

Let $\lambda \ge 0$. Let us first focus on the upper bound. For any $t \in \mathbb{R}_+$, we have $R_\lambda > t$ if only if there is an excursion of X^λ above its running minimum that straddles both 1 and t, that is if $1 \sim_\lambda t$. Thus

$$\begin{aligned}\mathbf{P}(R_\lambda > 2\lambda + 1) &= \mathbf{P}(\underline{X}_1^\lambda = \underline{X}_{2\lambda+1}^\lambda)\\ &= \mathbf{P}(\underline{X}_1^\lambda \le \underline{X}_{2\lambda+1}^\lambda)\\ &\le \inf_x\{\mathbf{P}(\underline{X}_1^\lambda \le x) + \mathbf{P}(\underline{X}_{2\lambda+1}^\lambda > x)\}\,.\end{aligned} \tag{18}$$

A quick inspection of the expected values leads to the choice $x = -\frac{\lambda}{2}$. On the one hand $\underline{X}_1^\lambda \ge -\frac{1}{2} + \underline{W}_1$ so that, by (17), we obtain

$$\mathbf{P}(\underline{X}_1^\lambda \le -\tfrac{\lambda}{2}) \le \mathbf{P}(-\tfrac{1}{2} + \underline{W}_1 \le -\tfrac{\lambda}{2}) \le e^{-\lambda^2/9}\,, \tag{19}$$

for all λ large enough. On the other hand, $\underline{X}_{2\lambda+1}^\lambda \le X_{2\lambda+1}^\lambda = -\lambda - \frac{1}{2} + W_{2\lambda+1}$ which implies that

$$\mathbf{P}(\underline{X}_{2\lambda+1}^\lambda > -\tfrac{\lambda}{2}) \le \mathbf{P}(W_{2\lambda+1} > \tfrac{\lambda}{2}) \le e^{-\lambda/20}\,, \tag{20}$$

for all λ large enough. Putting together (18)–(20) completes the proof of the upper bound.

For the lower bound, observe that

$$\begin{aligned}\mathbf{P}\left(R_\lambda < 2\lambda - 1\right) &= \mathbf{P}(\underline{X}_1^\lambda > \underline{X}_{2\lambda-1}^\lambda)\\ &\le \inf_x\left\{\mathbf{P}(X_1^\lambda < x) + \mathbf{P}(\underline{X}_1^\lambda > \underline{X}_{2\lambda-1}^\lambda, X_1^\lambda \ge x)\right\}.\end{aligned} \tag{21}$$

Considering the fact that $X_1^\lambda = -\frac{1}{2} + \lambda + W_1$, we are lead to choosing $x = \frac{\lambda}{2}$. We have $\mathbf{P}(X_1^\lambda < \frac{\lambda}{2}) \leq e^{-\lambda^2/9}$ for all λ large enough. To deal with the second part of the right-hand side of (21), we use Markov's property at time 1 and observe that for $s \geq 1$, $X_s^\lambda - X_1^\lambda$ is distributed like $X_{s-1}^{\lambda-1}$:

$$\mathbf{P}(\underline{X}_1^\lambda > \underline{X}_{2\lambda-1}^\lambda \mid X_1^\lambda \geq \tfrac{\lambda}{2}) \leq \mathbf{P}(\underline{X}_{2\lambda-2}^{\lambda-1} < -\tfrac{\lambda}{2}) \,. \tag{22}$$

However, since $s \mapsto -\frac{s^2}{2} + (\lambda-1)s$ is non-negative on $[0, 2(\lambda-1)]$, we have

$$\underline{X}_{2\lambda-2}^{\lambda-1} = \inf\{-\tfrac{s^2}{2} + (\lambda-1)s + W_s : s \leq 2\lambda - 2\} \geq \underline{W}_{2\lambda-2} \,,$$

so that, by (17),

$$\mathbf{P}(\underline{X}_{2\lambda-2}^{\lambda-1} < -\tfrac{\lambda}{2}) \leq \mathbf{P}(\underline{W}_{2\lambda-2} < -\tfrac{\lambda}{2}) \leq e^{-\lambda/20} \,,$$

for all λ large enough. Putting this together with (21) and (22) yields the lower bound on R_λ.

Finally, we deal with the lower bound on L_λ. Observe that $\mathbf{E}[X_s^\lambda] = -\frac{s^2}{2} + \lambda s \geq s\lambda/2$ for all $s \in [0,1]$ provided that λ is large enough. So writing $t = a/\lambda^2$, we have, since $\underline{X}_t^\lambda \leq 0$,

$$\begin{aligned}\mathbf{P}(L_\lambda \geq t) = \mathbf{P}(\underline{X}_t^\lambda > \underline{X}_1^\lambda) &\leq \mathbf{P}(\exists s \in [t,1] : X_s^\lambda \leq 0) \\ &\leq \mathbf{P}(\exists s \in [t,1] : W_s \leq -\lambda s/2) \,.\end{aligned}$$

Observe that $\{(s,y) : s \in [t,1], y \leq -\lambda s/2\} \subseteq \bigcup_{i\geq 1}\{(s,y) : s \leq (i+1)t, y \leq -\lambda i t/2\}$. So writing $\tau(-x) := \inf\{s \geq 0 : W_s < -x\}$, we have

$$\begin{aligned}\mathbf{P}(\exists s \in [t,1] : W_s \leq -\lambda s/2) &\leq \sum_{i\geq 1} \mathbf{P}(\tau(-\lambda i t/2) \leq (i+1)t) \\ &\leq \sum_{i\geq 1} \mathbf{P}(\underline{W}_{(i+1)t} \leq -\lambda i \tfrac{t}{2})\end{aligned}$$

which, using (17) is at most $e^{-a/20}$ for all a large enough. This completes the proof of Lemma 5.2.

5.3 Distances in small aggregated components

In this section, we are interested in the intrinsic metric space induced by d on q_λ. Write $D_\lambda = \sup\{d(x,y) : x, y \in q_\lambda\}$ for the diameter of this metric space, and for ζ_λ a random variable with uniform distribution in q_λ, independent of everything else, let $Y_\lambda := d(l_\lambda, \zeta_\lambda)$ where $l_\lambda = \inf q_\lambda$.

Recall that $X_t^\lambda = W_t - \frac{t^2}{2} + \lambda t$ and $B_t^\lambda = X_t^\lambda - \underline{X}_t^\lambda$ is the process reflected in the running infimum. We define $m_\lambda = |q_\lambda|$ and $\varepsilon_\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$\varepsilon_\lambda(s) := B^{\lambda-h}(l_\lambda + s)\mathbf{1}_{0\leq s\leq m_\lambda} \,. \tag{23}$$

The excursion ε_λ has duration m_λ and encodes the metric space $\mathscr{T}_\lambda^R$ or $\mathscr{T}_\lambda^L$ supported by q_λ. For any fixed $\lambda' \in \mathbb{R}$, and q an interval of $\mathbb{R}_+ \setminus Z^{\lambda'}$, conditionally on $|q| = \sigma$, the distribution of the excursion of $B^{\lambda'}$ straddling q is given by

$$\tilde{\mathbf{n}}_\sigma(A) = \frac{\int \mathbf{1}_{\omega\in A} \exp(\int_0^\sigma \omega(r)dr)\mathbf{n}_\sigma(d\omega)}{\int \exp(\int_0^\sigma \omega(r)dr)\mathbf{n}_\sigma(d\omega)} \tag{24}$$

where $\mathbf{n}_\sigma$ is the law of a Brownian excursion of duration $\sigma > 0$. Consider now, $\lambda \in \Lambda$, which is random. For any $\lambda' < \lambda$, $l_\lambda \in Z^{\lambda'}$; furthermore, as $\lambda' \uparrow \lambda$, we have $\inf\{s > 0 : B^{\lambda'}(l_\lambda + s) = 0\} \uparrow m_\lambda$ and the excursion of $B^{\lambda'}$ starting at l_λ converges to ε_λ. It follows that, for a bounded continuous functional ϕ,

$$\mathbf{E}[\phi(\varepsilon_\lambda) \mid m_\lambda = \sigma] = \int \phi(\omega)\tilde{\mathbf{n}}_\sigma(d\omega) \,. \tag{25}$$

We note further that, by the strong Markov property, the excursions ε_λ, $\lambda \in \Lambda$ are mutually independent conditionally on m_λ, $\lambda \in \Lambda$.

Proposition 5.6 (Distances in small components)**.** *For any $\epsilon > 0$, there exist*

i) a sub-Gaussian random variable $D^\star$ such that, if $m_\lambda \in (0, \epsilon]$ then $m_\lambda^{-1/2} D_\lambda \leq_{st} D^\star$, and

ii) a random variable $Y^\star$ with $\mathbf{E}[1/Y^\star] < \infty$ and such that if $m_\lambda \in (0, \epsilon]$ then $Y^\star \leq_{st} m_\lambda^{-1/2} Y_\lambda$.

Proof. *i)* Let $(\omega, \mathbf{u}) \mapsto d_{\omega,\mathbf{u}}$ a deterministic measurable function which gives the metric of $\mathrm{CMT}(\omega, \mathbf{u})$ for $\mathbf{n} \otimes d\mathbf{u}$-a.e. $(\omega, \mathbf{u})$, and let $\bar{d}_{\omega,\mathbf{u}} = \sup_{x,y} d_{\omega,\mathbf{u}}(x, y)$ be the corresponding diameter. From the law of ε_λ in (24)–(25) and the Cauchy–Schwarz inequality, for any $r \in \mathbb{R}$,

$$\begin{aligned}
\mathbf{E}\left[e^{r m_\lambda^{-1/2} D_\lambda} \,\middle|\, m_\lambda = \sigma\right] &= \int \exp(r\sigma^{-1/2} \bar{d}_{\omega,\mathbf{u}}) \tilde{\mathbf{n}}_\sigma(d\omega) d\mathbf{u} \\
&= \int \exp(r\sigma^{-1/2} \bar{d}_{\omega,\mathbf{u}}) \frac{\exp(\int_0^\sigma \omega(s)ds)\mathbf{n}_\sigma(d\omega)}{\int \exp(\int_0^\sigma \omega(s)ds)\mathbf{n}_\sigma(d\omega)} d\mathbf{u} \\
&\leq \frac{\sqrt{\int \exp(2r\sigma^{-1/2} \bar{d}_{\omega,\mathbf{u}})\mathbf{n}_\sigma(d\omega)d\mathbf{u}} \cdot \sqrt{\int \exp(2\int_0^\sigma \omega(s)ds)\mathbf{n}_\sigma(d\omega)}}{\int \exp(\int_0^\sigma \omega(s)ds)\mathbf{n}_\sigma(d\omega)}. \qquad (26)
\end{aligned}$$

In the first factor above, the metric space is encoded by a Brownian excursion, and is therefore exactly distributed like a Brownian continuum random tree of mass σ (Theorem 1.3 with $s = 0$). It follows by Brownian scaling that, under $\mathbf{n}_\sigma \otimes d\mathbf{u}$, $\sigma^{-1/2} \bar{d}_{\omega,\mathbf{u}}$ is dominated by $2\|e\|$, twice the supremum of a standard Brownian excursion, which is well-known to be sub-Gaussian (see, for instance, [31, 42]). On the other hand, observe that $\psi(z) := \int \exp(z \int_0^1 \omega(s)ds)\mathbf{n}_1(d\omega) = \mathbf{E}[\exp(z \int_0^1 e(s)ds)]$, the Laplace transform of Brownian excursion area [40], is continuous and positive on any interval $[0, \epsilon]$. It follows that there exist constants $A = A_\epsilon$ and $v > 0$ such that, for any $\sigma \in [0, \epsilon]$,

$$\mathbf{E}\left[\exp(r m_\lambda^{-1/2} D_\lambda) \,\middle|\, m_\lambda = \sigma\right] \leq e^{r^2/(2v)} \cdot \sup_{\sigma \in [0,\epsilon]} \frac{\sqrt{\psi(2\sigma^{3/2})}}{\psi(\sigma^{3/2})} \leq A e^{r^2/(2v)}. \qquad (27)$$

One then easily constructs (the law of a variable) $D^\star$ by inverse transform. For $\sigma > 0$ let $F_\sigma(x) := \mathbf{P}(D_\lambda \leq x \mid m_\lambda = \sigma)$, and $F_\star(x) := \inf\{F_\sigma(x) : \sigma \leq \epsilon\}$. Then for U a $[0, 1]$-uniform random variable, $F_\star$ is the distribution function of a random variable $D^\star := F_\star^{-1}(U)$ that dominates all the $m_\lambda^{-1/2} D_\lambda \mathbf{1}_{m_\lambda \leq \epsilon}$. The random variables $F_\sigma^{-1}(U)$ are uniformly sub-Gaussian by (27), and so is $D^\star$. This completes the proof of *i)*.

The proof of *ii)* about the distance to a random point is similar: we only discuss the adaptation of the arguments in *i)* to bound $\mathbf{E}[m_\lambda^{1/2}/Y_\lambda \mid m_\lambda = \sigma]$. Instead of Cauchy–Schwarz, using Hölder's inequality (with exponents $3/2$ and 3) provides an upper bound similar to (26) where the main term involves the $d_{\omega,\mathbf{u}}(0, \xi)$ for ξ independent and uniform in $[0, \sigma]$ under $\mathbf{n}(d\omega)$: this is the distance between two random points in a unit mass Brownian CRT, which is a random variable Y with density $xe^{-x^2/2}dx$ on $\mathbb{R}_+$ [11], so that $\mathbf{E}[Y^{-3/2}] < \infty$. The multiplicative error term is bounded just as above, and the random variable $Y^\star$ is constructed similarly, using the supremum of the distribution functions instead of the infimum. We omit the details. □

5.4 Statistics of the aggregated components $\mathscr{T}_\lambda^R$

From the results of the previous section, especially the law of the ε_λ, it is crucial to understand the distribution of the m_λ. In the following we let $\mathscr{E}$ be the space of continuous excursions, that is, the functions $f \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$ with $f(0) = 0$ such that there exists $\sigma \in [0, \infty)$ such that $f(s) > 0$ for $s \in (0, \sigma)$ and $f(s) = 0$ for $s \geq \sigma$.

Fix any $\lambda \in \mathbb{R}$ and consider the excursions of X^λ away from its running infimum $\underline{X}^\lambda$. For $y \in \mathbb{R}_+$, let $\tau_y^\lambda := \inf\{s > 0 : -\underline{X}_s^\lambda > y\}$. Then, $\{(y, \tau_y^\lambda - \tau_{y-}^\lambda) : \tau_y^\lambda > \tau_{y-}^\lambda\}$ is a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}_+$ of intensity $dy\tilde{\varrho}_{\tau_{y-}}^\lambda$, where the inhomogeneous measure $\tilde{\varrho}_x^\lambda$ is defined by (see, e.g., [14], Section 5.2)

$$\tilde{\varrho}_x^\lambda(l, \infty) := \lim_{\epsilon \to 0} \frac{1}{\epsilon} \mathbf{P}\left(\inf\{s > 0 : X_{x+s}^\lambda \leq 0\} > l \,\middle|\, X_x^\lambda = \epsilon\right). \qquad (28)$$

Proposition 5.7. *The process* $\{(\lambda, \varepsilon_\lambda) : R_\lambda > R_{\lambda-}\}$ *is a Poisson point process on* $\mathbb{R} \times \mathscr{E}$ *of intensity* $R_{\lambda-} d\lambda \tilde{\varrho}^{\lambda}_{R_{\lambda-}}(d\sigma)\tilde{\mathbf{n}}_\sigma(d\omega)$.

Proof. The distribution of the excursions ε_λ conditionally on their duration is known from the previous section, and we only need to deal with the sizes of the jumps of $(R_\lambda)_{\lambda\in\mathbb{R}}$. These are formed by agglomeration of some of the excursion lengths of B^λ, for $\lambda \in \mathbb{R}$, which are described by the excursion length measure in (28): As λ increases, the excursions of B^λ away from 0 merge together until they eventually join the connected component containing 1.

We proceed geometrically using the process X^0 only. The excursions ε_λ are simply read from X^0: for any λ such that $R_\lambda > R_{\lambda-}$, ε_λ is obtained as

$$\varepsilon_\lambda(s) = (X^0(R_{\lambda-} + s) - X^0(R_{\lambda-}) + s\lambda)\mathbf{1}_{R_{\lambda-}+s\le R_\lambda} \qquad \text{for } s \ge 0.$$

Here, notice that ε_λ is indeed obtained from the agglomeration of countably many excursions of $X^{\lambda-h} - \underline{X}^{\lambda-h}$, which might be described using the process involving straight lines with slopes $\lambda - h$ (at time λ, the excursions of interest are those of X^0 above the process $s \mapsto \underline{X}^\lambda_s - \lambda s$). Now, knowing the intensity of jumps of τ^λ for each $\lambda \in \mathbb{R}$, it is routine to deduce the intensity of excursions of jumps of $(R_\lambda)_{\lambda\in\mathbb{R}}$: at time λ, we always have an excursion, the increase in local time is $d\lambda R_{\lambda-}$, and this gives rise to excursions whose durations are governed by $\tilde{\rho}^\lambda_{R_{\lambda-}}$. □

For each $k \ge 1$ we say $\lambda \in \Lambda_k$ is of level $i \ge 0$ and write $\lambda \in \Lambda_{k,i}$ if its duration satisfies $m_\lambda \in [k^{-6-i}, k^{-5-i})$. We define the $M_{k,i}$ total duration (mass) of excursions of level at least i

$$M_{k,i} = \sum_{j\ge i} \sum_{\lambda\in\Lambda_{k,j}} m_\lambda \,.$$

Lemma 5.8 (Statistics for fragments $\mathscr{T}^R_\lambda$, $\lambda \in \Lambda^R$)**.** *For any* $k \ge 1$ *and* $i \ge 0$*, there exists an event* $B_{k,i}$ *of probability at most* $8\min\{k^{-1-i/4}, k^{-5/4}\}$ *such that outside of* $B_{k,i}$ *we have*

i) the longest excursion: $\sup\{m_\lambda : \lambda \in \Lambda_k\} \le k^{-5}$
ii) total duration of excursions of level at least i: $M_{k,i} \le \min\{k^{7/2-i/4}, 7k^2\}$,
iii) number of excursions of level i: $\#\Lambda_{k,i} \le 7k^{8+i}$,
iv) maximum diameter of an excursion of level i: $D_{k,i} = \sup\{D_\lambda : \lambda \in \Lambda_{k,i}\} \le k^{-2-i/4}$.

In particular, $\bigcup_{i\ge0} B_{k,i}$ *occurs with probability at most* $10k^{-5/4}$.

The proof of Lemma 5.8 is based on upper bounds on the durations $|q_\lambda|$, $\lambda \in \Lambda_k$. The relevant calculations are simplified if we upper bound the Brownian with parabolic drift by a Brownian with a suitable linear drift. This is why the following is especially useful. In the following, we let ϱ^μ be the excursion length measure for Brownian motion with linear drift μ, $W^\mu := s \mapsto W_s + \mu s$. We are mostly interested in what happens at large positions, for which the drift is negative: we note that, for $\mu \ge 0$,

$$\varrho^{-\mu}(x,\infty) = \lim_{\epsilon\to0} \frac{1}{\epsilon}\mathbf{P}\left(\inf\{s > 0 : W^{-\mu}_s \le 0\} > x \,\middle|\, W^{-\mu}_0 = \epsilon\right) = \int_x^\infty \frac{e^{-\mu^2 r/2}}{\sqrt{2\pi r^3}} dr \,.$$

The following proof relies this and the fact that $X^{-\mu}_s \le W^{-\mu}_s$ for all $s \ge 0$ in the natural coupling.

Proof of Lemma 5.8. Let B_k be the event that $R_{k^3} > 2k^3 - 1$ or $R_{(k+1)^3} < 2(k+1)^3 + 1$. Then, by Lemma 5.2, $\mathbf{P}(B_k) \le \exp(-ck^3)$ for some constant $c > 0$. Now, on the one hand, the quantity of local time LT_k corresponding to Λ_k is

$$LT_k := \int_{k^3}^{(k+1)^3} R_{\lambda-} d\lambda = \int_{k^3}^{(k+1)^3} R_\lambda d\lambda \le R_{(k+1)^3}[(k+1)^3 - k^3] \,,$$

so that $LT_k \le 7k^5$ for all k large enough on the complement event B^c_k. On the other hand, $\sup\{\lambda - \inf q_\lambda : \lambda \in \Lambda_k\} \le (k+1)^3 - R_{k^3} \le -k^3/2$ for large k on B^c_k.

By Proposition 5.7, it follows that, on the event B_k^c, for every $x > 0$, $\#\{\lambda \in \Lambda_k : m_\lambda > x\}$ is stochastically dominated by a Poisson random variable with parameter $\Delta_k \varrho^{-\mu_k}(x, \infty)$ with $\Delta_k = 7k^5$ and $\mu_k = k^3/2$. The properties *i)* to *iv)* in the statement then follow easily.
i) Longest excursion. We have

$$\mathbf{E}\left[\sum_{\lambda\in\Lambda_k} \mathbf{1}_{m_\lambda > k^{-5}} \mathbf{1}_{B_k^c}\right] \leq 7k^5 \varrho^{-\mu_k}(k^{-5}, \infty) = 7k^5 \int_{k^{-5}}^{\infty} \frac{e^{-k^6 t/8}}{\sqrt{2\pi t^3}} dt \leq e^{-ck},$$

for some constant c and all k large enough. Markov's inequality then implies the claim.
ii) Total length of excursions of level i. We proceed similarly for the upper bound on $M_{k,i}$. We have

$$\mathbf{E}\left[\sum_{\lambda\in\Lambda_k} m_\lambda \mathbf{1}_{m_\lambda \leq k^{-5-i}} \mathbf{1}_{B_k^c}\right] \leq 7k^5 \int_0^{k^{-5-i}} x\varrho^{-\mu_k}(dx) \leq 7k^5 \int_0^{k^{-5-i}} \frac{xe^{-k^6 x/8}}{\sqrt{2\pi x^3}} dx \leq 7k^{5/2-i/2}.$$

Moreover, we also always have the bound $\sum_{\lambda\in\Lambda_k} m_\lambda \leq R_{(k+1)^3} - R_{k^3}$ which is at most $7k^2$ on B_k^c. Markov's inequality then yields

$$\mathbf{P}(M_{k,i} \geq k^{7/2-i/4}, B_k^c) \leq \min\{7k^{-1-i/4}, 7k^{-3/2+i/4}\}.$$

iii) Cardinality of $\Lambda_{k,i}$. This is a deterministic bound on B_k^c: we have $\#\Lambda_{k,i} \leq 7k^2 \cdot k^{6+i} = 7k^{8+i}$.
iv) Maximum diameter in level i. By Proposition 5.6, when the event of *i)* occurs, $D_{k,i}$ is stochastically dominated by

$$k^{-5/2-i/2} \cdot \max\{D_j^\star : 1 \leq j \leq \#\Lambda_{k,i}\},$$

where $D_j^\star$ are iid copies of the random variable $D^\star$ that is sub-Gaussian (choose $\epsilon = 1$). Using the bound for $\#\Lambda_{k,i}$ in *iii)* above, it follows easily that

$$\begin{aligned}\mathbf{P}(D_{k,i} \geq k^{-2-i/4}, B_k^c) &\leq 7k^{8+i} \cdot \mathbf{P}(D^\star \geq k^{1/2+i/4}) \\ &\leq 7k^{8+i} \cdot \exp(-\tfrac{k^{1+i/2}}{2v}),\end{aligned}$$

for some constant $v > 0$.

Finally, writing $B_{k,i}$ for the event that either B_k or any of the bad events in *i)–iv)* occur, we have $\mathbf{P}(B_{k,i}) \leq 8\min\{7k^{-1-i/4}, k^{-5/4}\}$ for large k (this is essentially limited by the event in *ii)*). The union bound yields the last claim. □

5.5 The accumulation of length in an annulus: Proof of Proposition 5.3

Each component $\mathscr{F}_\lambda$ gets connected to some point to its left (in $\mathbb{R}_+$), which falls within some $\mathscr{F}_{\lambda'}$, for some $\lambda' < \lambda$, and so on. The proof of Proposition 5.3 consists in bounding the accumulation of these lengths before a connection to $\mathscr{M}_{k^3}$ eventually occurs. We are only interested here in the points of $(1, \infty)$.

Recalling the notation $\boldsymbol{\jmath}(\cdot)$ from Remark 3.11, by construction, for each $\lambda \in \Lambda^R$, the point $l_\lambda = \inf q_\lambda$ is identified with $\boldsymbol{\jmath}(l_\lambda)$ in the metric space $(\mathscr{M}, d)$. Furthermore, $\boldsymbol{\jmath}(l_\lambda)$ is uniform in $H_{\lambda-}$. For any $x \in q_\lambda$, with $\lambda \in \Lambda^R$, the segment between x and 1 must contain the points $\boldsymbol{\jmath}(l_\lambda)$ that we may see as a projection of q_λ on $H_{\lambda-}$. With this in mind, we let $p(x) = \boldsymbol{\jmath}(l_\lambda)$ if $x \in q_\lambda$, $\lambda \in \Lambda^R$. Then, for any point $x > 1$, we consider the sequence of successive projections defined by $p^0(x) = x$, and provided that $p^n(x) > 1$, $p^{n+1}(x) = p(p^n(x))$, until we eventually find a point in $[0, 1]$.

Fix now some natural number $k \geq 1$ and let $A_k := (R_{k^3}, R_{(k+1)^3}]$ denote the set of points of the annulus of fragments $\mathscr{F}_\lambda$, for $\lambda \in \Lambda_k$. For each $\lambda \in \Lambda_k$, and $n \geq 0$, let λ_n be such that $p^n(l_\lambda) \in q_{\lambda_n}$. Recall that $D_\lambda = \sup\{d(x, y) : x, y \in q_\lambda\}$. Clearly, the distance from any point $x \in A_k$ to $\mathscr{M}_{k^3}$ is at most

$$\sup_{\lambda\in\Lambda_k} \sup_{x\in q_\lambda} d(x, \mathscr{M}_{k^3}) \leq \sup_{\lambda\in\Lambda_k} \sum_{1\leq n\leq N_\lambda} D_{\lambda_n},$$

where $N_\lambda = \inf\{n \geq 1 : p^n(l_\lambda) < R_{k^3}\}$. However, since Λ_k is infinite, we shall refine the analysis and rely on the decomposition into different levels introduced in the previous section.

Recall that we say that q_λ is an interval of level i, and write $\lambda \in \Lambda_{k,i}$ if $m_\lambda \in [k^{-6-i}, k^{-5-i})$; let $A_{k,i} = \bigcup_{\lambda \in \Lambda_{k,i}} q_\lambda$ be the subset of A_k consisting of the intervals of level i. For $i \geq 0$, and $\lambda \in \Lambda_{k,i}$, let $N_i(\lambda) := \inf\{n \geq 1 : p^n(l_\lambda) \notin A_{k,i}\}$ be the number of hops until hitting an interval of level lower than i, or exiting A_k altogether from the left. We then have

$$\begin{aligned}
\sup_{\lambda \in \Lambda_k} \sum_{1 \leq n \leq N_\lambda} D_{\lambda_n} &\leq \sum_{i \geq 0} \sup_{\lambda \in \Lambda_{k,i}} \sum_{1 \leq n \leq N_i(\lambda)} D_{\lambda_n} \\
&\leq \sum_{i \geq 0} \sup_{\lambda \in \Lambda_{k,i}} N_i(\lambda) \cdot \sup_{j \geq i} \sup_{\lambda \in \Lambda_{k,j}} D_\lambda \\
&\leq \sum_{i \geq 0} \sup_{\lambda \in \Lambda_{k,i}} N_i(\lambda) \cdot k^{-2-i/4}\,, \qquad (29)
\end{aligned}$$

provided that the event $\bigcap_{i \geq 0} B^c_{k,i}$ from Lemma 5.4 occurs. So it remains only to upper bound $N_{k,i} = \sup\{N_i(\lambda) : \lambda \in \Lambda_{k,i}\}$. We do this using the properties of the sequence of the projections.

Lemma 5.9. *For any $k \geq 1$ large enough, with probability at least $1 - 11k^{-5/4}$, we have for every $i \geq 0$,*

$$N_{k,i} := \sup_{\lambda \in \Lambda_{k,i}} N_\lambda < 20\,.$$

Proof. Let $\mathcal{G}$ be the sigma-algebra generated by $(X_s)_{s \geq 0}$. Then (q_λ), $\lambda \in \Lambda^R$ is $\mathcal{G}$-measurable while, conditionally on $\mathcal{G}$, the random variables $\xi(l_\lambda)$ are independent and uniform in $H_{\lambda-}$. Let B_k be the event that $R_{k^3} < 2k^3 - 1$ or $R_{(k+1)^3} > 2(k+1)^3 + 1$. For any $\lambda \in \Lambda_k$, $(1, R_{k^3}] \subseteq H_{\lambda-}$; therefore, on the event B^c_k for any Borel set A, $\mathbf{P}(\xi_\lambda \in A \mid B^c_k, \mathcal{G}) \leq \mathrm{Leb}(A)/k^3$. Furthermore, by Lemma 5.4, on $B^c_{k,i}$ we have $M_{k,i} \leq \max\{k^{7/2-i/4}, 7k^2\}$ and $\#\Lambda_{k,i} \leq 7k^{8+i}$. It follows by the union bound, that for any natural number $m \geq 1$, we have

$$\begin{aligned}
\mathbf{P}(N_{k,i} \geq m \mid B^c_{k,i}, B^c_k, \mathcal{G}) &\leq \#\Lambda_{k,i} \cdot \sup_{\lambda \in \Lambda_{k,i}} \mathbf{P}(N_\lambda \geq m \mid M_{k,i} \leq k^{7/2-i/4}) \\
&\leq 7k^{8+i} \cdot (k^{7/2-i/4}/k^3)^m \\
&\leq 7k^{8+i} \cdot k^{-m(i-2)/4}\,.
\end{aligned}$$

As a consequence, for $m = 20$, we obtain

$$\mathbf{P}(N_{k,i} \geq 20 \mid B^c_{k,i}, B^c_k, \mathcal{G}) \leq 7k^{18-4i}\,,$$

which will be good enough for $i \geq 5$. On the other hand, for $0 \leq i \leq 4$, the alternative bound $M_{k,i} \leq 7k^2$ yields a bound of $7k^{8+i-m} \leq 7k^{-8}$ for $m = 20$. Putting everything together, we have $N_{k,i} \geq 20$ for some $i \geq 0$ with probability at most $10k^{-5/4} + 28k^{-8} + 8k^{-2} \leq 11k^{-5/4}$ for all k large enough. □

Going back to (29), Lemma 5.9 implies that

$$\sup_{\lambda \in \Lambda_k} \sum_{1 \leq n \leq N_\lambda} D_{\lambda_n} \leq 20 \sum_{i \geq 0} k^{-2-i/4} \leq k^{-3/2}$$

with probability at least $1 - 11k^{-5/4}$ which completes the proof of Proposition 5.3.

5.6 The diameter of the string of beads: Proof of Proposition 5.4

By construction, for any $\lambda \in \mathbb{R}$, the diameter of $\mathscr{P}_\lambda$ is no greater than

$$\sum_{\lambda' > \lambda} \operatorname{diam}(\mathscr{T}^-_{\lambda'}) \mathbf{1}_{\lambda' \in \Lambda^L} \leq \sum_{\lambda' > \lambda} \sqrt{m_{\lambda'}} \cdot D^\star_\lambda \mathbf{1}_{\lambda' \in \Lambda^L}\,, \tag{30}$$

where $D^\star_\lambda$, $\lambda \in \Lambda^L$, are i.i.d. copies of the sub-Gaussian random variable whose existence is guaranteed by Proposition 5.6 with $\epsilon = 1$. We have already bounded a similar sum in Section 3.7; in particular, the arguments there show that almost surely

$$\sum_{\lambda' \in \Lambda^L} \sqrt{m_{\lambda'}} < \infty\,. \tag{31}$$

Finally, consider the process $M_\lambda(s)$ defined for $s \geq 0$ by

$$M_\lambda(s) := \sum_{\lambda < \lambda' \leq \lambda + s} \sqrt{m_{\lambda'}} \cdot (D^\star_{\lambda'} - \mathbf{E}[D^\star_{\lambda'}])\,.$$

Conditionally on the m_λ, $\lambda \in \Lambda^-$, $(M_\lambda(s))_{s \geq 0}$ is a martingale. Since $D^\star$ is sub-Gaussian, $M_\lambda(s)$ is bounded in L^2 and thus converges almost surely to a finite limit as $s \to \infty$. Putting this together with (31) shows that the right-hand side of (30) and hence $\operatorname{diam}(\mathscr{P}_\lambda)$ tends to zero as $\lambda \to \infty$, which completes the proof of Proposition 5.4.

6 The mass measure and Hausdorff dimension of $\mathrm{CMT}(X, \mathbf{U})$

In this section we prove the lower bound on the Hausdorff dimension of $\mathscr{M}$. We use the mass distribution principle using the mass measure μ that is defined in Section 6.1. The asymptotics for the μ-mass of small balls are provided in Section 6.2 and relies heavily on the growth process defined in Section 5.

6.1 The mass measure

We start with the construction of the mass measure μ on $\mathscr{M}$. The measures in this section will always be seen as Borel measures on $\mathscr{M}$, the completion of $\mathbb{R}_+$ with respect to d. For $t \in \mathbb{R}_+$ we let μ_t be the rescaled Lebesgue measure on $[0, t]$: $\mu_t(A) = t^{-1} \operatorname{Leb}(A \cap [0, t])$. For each t, μ_t is a probability measure on $\mathscr{M}$ which charges only a subtree containing the root 0 (it is easy to see that $[0, t]$ is connected). Let $\mathcal{L}$ be the set of leaves of $\mathscr{M}$, that is the set of points x such that $\mathscr{M} \setminus \{x\}$ is connected. Our aim in this section is the following

Proposition 6.1. *With probability one, as $t \to \infty$, μ_t converges weakly to a limit probability measure that we denote by μ and call the mass measure on $\mathscr{M}$. Furthermore $\mu(\mathbb{R}_+) = 0$ so $\mu(\mathcal{L}) = 1$.*

Recall the notation in Section 5.1. For a subset S of the tree $\mathscr{M}$ and $x \in \mathbb{R}_+$, we define the projection of x onto S as the point of S that is closest to x. Fix $\lambda_0 \in \mathbb{R}$. We are interested in the projection onto the subset of $\mathscr{M}$ consisting of the points $[0, R_{\lambda_0}]$. For $x \in \mathscr{M}$, we let $[x]_{\lambda_0}$ denote the corresponding point. Observe that, with the notation of the previous section, a.s. $[x]_{\lambda_0} = \sup\{p^n(x) \cap [0, R_{\lambda_0}], n \geq 0\}$.

Even though $[0, R_{\lambda_0}]$ is not closed in $\mathscr{M}$, we will always have $[x]_{\lambda_0} \in [0, R_{\lambda_0})$ for the points x we consider. Rather than working with the measures μ_t, $t \in \mathbb{R}_+$, it will be more convenient to work with $\bar{\mu}_\lambda := \mu_{R_\lambda}$ for $\lambda \in \mathbb{R}$; Lemma 5.2 which says that $R_\lambda \to \infty$ guarantees that taking the limits as $t \to \infty$ or $\lambda \to \infty$ is equivalent.

We define the following process: for a Borel set $S \subseteq [0, R_{\lambda_0}]$ and $\lambda \geq \lambda_0$,

$$M_\lambda = M_\lambda(S) := \bar{\mu}_\lambda(\{x \leq R_\lambda : [x]_{\lambda_0} \in S\})\,. \tag{32}$$

We will consider only the randomness coming from $\mathbf{U}$ and study M_λ conditionally on $\sigma(R_\lambda : \lambda \geq \lambda_0)$. We let $\mathbf{P}^\downarrow$ and $\mathbf{E}^\downarrow$ be the corresponding probability and expectation.

Lemma 6.2. *The process $(M_\lambda, \lambda \geq \lambda_0)$ is almost surely a supermartingale under $\mathbf{P}^\downarrow$.*

Proof. For each $\lambda \in \Lambda^R$, $\lambda > \lambda_0$, all the points $x \in q_\lambda$ have the same projection on $[0, R_{\lambda_0}]$ since $p(x) = \xi(l_\lambda)$. Furthermore, in order to determine where an interval q_λ projects onto H_{λ_0} it suffices to follow the sequence of random projections/jumps $p^n(x)$, $n \geq 1$. Almost surely, $\inf\{n \geq 0 : p^n(l_\lambda) \in H_{\lambda_0}\} < \infty$ and the point $[l_\lambda]_{\lambda_0}$ is uniformly random in H_{λ_0}. In the following, λ_0 being fixed, we use $\phi(\lambda)$ as a short-hand for $[l_\lambda]_{\lambda_0}$. The points $\phi(\lambda)$, $\lambda > \lambda_0$, are of course not independent because of the coalescence of the trajectories. Then,

$$M_\lambda = \frac{1}{R_\lambda}\left[M_{\lambda_0} \cdot R_{\lambda_0} + \sum_{\lambda_0 < \lambda' \leq \lambda} m_{\lambda'} \mathbf{1}_{\phi(\lambda') \in S}\right]. \tag{33}$$

It follows that, writing $\mathcal{F}_\lambda$ for the sigma-algebra generated by $(\not{z}(l_{\lambda'}) : \lambda' \leq \lambda)$, the random variable M_λ is bounded and $\mathcal{F}_\lambda$-measurable. The expression in (33) is amenable to a simple evaluation of the conditional expectations: for $h \geq 0$,

$$\begin{aligned}\mathbf{E}^\downarrow[M_{\lambda+h} \mid \mathcal{F}_\lambda] &= \frac{1}{R_{\lambda+h}} \cdot \mathbf{E}^\downarrow\left[M_{\lambda_0} \cdot R_{\lambda_0} + \sum_{\lambda_0 < \lambda' \leq \lambda+h} m_{\lambda'} \mathbf{1}_{\phi(\lambda') \in S} \,\middle|\, \mathcal{F}_\lambda\right] \\ &= \frac{1}{R_{\lambda+h}}\left(R_\lambda \cdot M_\lambda + \sum_{\lambda < \lambda' \leq \lambda+h} m_{\lambda'} \mathbf{P}^\downarrow(\phi(\lambda') \in S \mid \mathcal{F}_\lambda)\right).\end{aligned}$$

However, almost surely conditionaly on $\mathcal{F}_\lambda$, $\phi(\lambda') \in S$ if and only if the first point of the sequence $(p^i(l_{\lambda'}))_{i \geq 1}$ that falls in H_λ lies in some interval $q_{\lambda''}$ itself such that $\phi(\lambda'') \in S$. By definition, such a point is the projection of $l_{\lambda'}$ on H_λ, and is uniform in $H_\lambda \subset [0, R_\lambda]$ and therefore $\mathbf{P}^\downarrow(\phi(\lambda') \in S \mid \mathcal{F}_\lambda) \leq M_\lambda$. It follows that

$$\mathbf{E}^\downarrow[M_{\lambda+h} \mid \mathcal{F}_\lambda] \leq \frac{M_\lambda}{R_{\lambda+h}} \cdot \left(R_\lambda + \sum_{\lambda < \lambda' \leq \lambda+h} m_{\lambda'}\right) = M_\lambda\,,$$

which completes the proof. □

Proof of Proposition 6.1. Since $\mathscr{M}$ is compact by Proposition 5.5, the collection of measures $(\mu_t)_{t>0}$ is tight. We prove that it is Cauchy for the Prohorov metric using the super-martingales M_λ we have just introduced. Recall that, for two Borel measures ν and ν' on $(\mathscr{M}, d)$, the Prohorov distance is given by

$$\mathrm{d_P}(\nu, \nu') = \inf\{\epsilon > 0 : \nu(A) \leq \nu'(A^\epsilon) + \epsilon, \nu'(A) \leq \nu(A^\epsilon) + \epsilon \text{ for all Borel sets } A\}$$

where $A^\epsilon = \{x : d(x, A) < \epsilon\}$.

The arguments for compactness in Section 5 show that for any $\epsilon > 0$, there exists a $\lambda_0 \in \mathbb{R}_+$ such that $\sup_x d(x, [0, R_\lambda]) < \epsilon$ for all $\lambda \geq \lambda_0$. With this choice for λ_0, it follows that, for any $\lambda > \lambda_0$,

$$\mathrm{d_P}(\mu_\lambda, [\mu_\lambda]_{\lambda_0}) < \epsilon\,,$$

where $[\mu_\lambda]_{\lambda_0}$ denotes the image of μ_λ by the projection onto $[0, R_{\lambda_0}]$. Therefore, for any $\lambda, \lambda' \geq \lambda_0$,

$$\mathrm{d_P}(\mu_\lambda, \mu_{\lambda'}) \leq \mathrm{d_P}([\mu_\lambda]_{\lambda_0}, [\mu_{\lambda'}]_{\lambda_0}) + 2\epsilon\,.$$

To complete the proof, cover $[0, R_{\lambda_0}]$ with finitely many balls of diameter ϵ, say $B_1, B_2, \ldots, B_k$. Then, by definition of $M_\lambda(S)$, we can construct a coupling (X, Y) with $X \sim [\mu_\lambda]_{\lambda_0}$ and $Y \sim [\mu_{\lambda'}]_{\lambda_0}$ such that $(X, Y) \notin \cup_i B_i \times B_i$ with probability at most $\sum_i |M_\lambda(B_i) - M_{\lambda'}(B_i)|$. Since the diameter of B_i is at most ϵ, the cost of the coupling on $B_i \times B_i$ is at most ϵ. It follows that

$$\mathrm{d_P}([\mu_\lambda]_{\lambda_0}, [\mu_{\lambda'}]_{\lambda_0}) \leq \sum_{i=1}^k |M_\lambda(B_i) - M_{\lambda'}(B_i)| + \epsilon\,,$$

which is at most 2ϵ for all λ, λ' large enough because of the convergence of the mass super-martingales of Lemma 6.2. This completes the proof of convergence.

The two additional properties are straightforward from the definition. First, for any interval $[i, i+1)$, $\mu([i, i+1)) = \lim_\lambda \mu_\lambda([i, i+1]) = 0$ since $R_\lambda \to \infty$, and thus $\mu(\mathbb{R}_+) = 0$. The completion of $\mathbb{R}_+$ with respect to d only adds leaves, so $\mathscr{M} \setminus \mathbb{R}_+ \subseteq \mathcal{L}$, and therefore $\mu(\mathcal{L}) = 1$. □

6.2 The mass of balls around zero

It is proved in [4] that the Minkowski dimension of $\mathscr{M}$ is almost surely equal to 3, and we thus only need to find a lower bound. For this, we aim at using the mass distribution principle with the mass measure. In this direction, one needs to upper bound the μ-mass of balls centered at points with distribution μ. In general, this might be delicate since we need to identify the balls around these points, and they are almost surely not in $\mathbb{R}_+$ (Proposition 6.1). This is why the following result is crucial; the intuition should be intuitively clear from the discrete setting, where the point 1 can be replaced in Prim's algorithm by a uniformly random point in $[n]$ without altering the distributions. For a point $x \in \mathscr{M}$ and $r > 0$ we let $B_x(r)$ denote the open ball of radius r centered at x (for the metric d).

Lemma 6.3. *Let ζ be a point of $\mathscr{M}$ with distribution μ. Then, the processes $(\mu(B_0(r)))_{r\geq 0}$ has the same distribution as $(\mu(B_\zeta(r)))_{r\geq 0}$.*

Proof. For $n \geq 1$, recall from Section 1.2 that $v_1, v_2, \ldots, v_n$ denote the Prim order on $[n]$ on the complete graph with edge weights (w_e), $e \in E^n$; in what follows, we will occasionally write $v(i)$ instead of v_i. For $n \geq 1$, let V_λ^n be the collection of vertices connected to $v(\lfloor n^{2/3} \rfloor)$ in the random graph with edge weights at most $p_n(\lambda)$, and let H_λ^n denote the collection of their Prim ranks, and let $L_\lambda^n = \min H_\lambda^n$. Let $\hat{\mu}_\lambda^n$ denote the uniform probability distribution on H_λ^n. By Lemma 7.4, conditionally on V_λ^n, the vertex $v(L_\lambda^n)$ is uniformly random in V_λ^n, and independent of the random variables w_e associated to the edges with end points in V_λ^n. It follows in particular that, for any $r \geq 0$,

$$\mu_\lambda^n(\{u \in H_\lambda^n : d_\lambda^n(L_\lambda^n, u) \leq r n^{1/3}\}) \stackrel{d}{=} \mu_\lambda^n(\{u \in H_\lambda^n : d_\lambda^n(\zeta_\lambda^n, u) \leq r n^{1/3}\}), \tag{34}$$

where ζ_λ^n denotes an independent point with distribution $\hat{\mu}_\lambda^n$ (uniform in H_λ^n). By Proposition 7.13, $(H_\lambda^n, d_\lambda^n, \mu_\lambda^n, L_\lambda^n, \zeta_\lambda^n)$ converges in distribution in the sense of Gromov–Prokhorov to $(H_\lambda, d, \hat{\mu}_\lambda, L_\lambda, \zeta_\lambda)$, where ζ_λ is an independent point with distribution μ_λ. This implies the convergence of the random variables in (34) as $n \to \infty$ towards

$$\hat{\mu}_\lambda(\{u \in H_\lambda : d(L_\lambda, u) \leq r\}) \stackrel{d}{=} \hat{\mu}_\lambda(\{u \in H_\lambda : d(\zeta_\lambda, u) \leq r\}).$$

Now, $d(0, L_\lambda) \leq \mathrm{diam}(\mathscr{P}_\lambda) \to 0$ as $\lambda \to \infty$ by Proposition 5.4. Note also that a straightforward coupling yields $\mathrm{d_P}(\hat{\mu}_\lambda, \mu_\lambda) \leq L_\lambda / R_\lambda \leq L_\lambda$ (independently of the metric since we can match the points exactly on a set of probability $1 - L_\lambda/R_\lambda$). Taking the limit as $\lambda \to \infty$, Lemma 5.2 and Proposition 6.1 yield the claim for every fixed $r \geq 0$. This is easily extended to the joint convergence for finitely many values $r_1 < r_2 < \cdots < r_k$, which completes the proof. □

In order to upper bound $\mu(B_0(r))$ we will proceed in two steps: we will first upper bound $\bar{\mu}_\lambda(B_0(r))$ showing that it is of the correct order of magnitude, that is roughly r^3 for some well-chosen λ depending on r (λ of order $1/r$); we will then rely on the concentration for the mass supermartingales of the previous section, which controls the evolution of the mass as λ increases, to show that $\mu(B_0(r)) = \lim_\lambda \bar{\mu}_\lambda(B_0(r))$ remains of order r^3. Once we have the relevant upper bound for a fixed r, the proof is easily completed using routine arguments (taking a suitable subsequence and the Borel–Cantelli lemma).

Proposition 6.4. *There exists a constant $c > 0$ such that, for any $\epsilon \in (0, 1)$ and every $r > 0$ small enough,*

$$\mathbf{P}(\bar{\mu}_{r^{\epsilon-1}}(B_0(r)) > r^{3-\epsilon}) \leq r^{c\epsilon}.$$

Proof. Observe that if $d(0, H_\lambda) > r$ then no point of H_λ lies within $B_0(r)$, so that $\bar{\mu}_\lambda(B_0(r))$ is at most L_λ/R_λ. Using this with $\lambda = r^{\epsilon-1}$, it follows that

$$\begin{aligned}\mathbf{P}(\bar{\mu}_{r^{\epsilon-1}}(B_0(r)) > r^{3-6\epsilon}) &\leq \mathbf{P}\left(\bar{\mu}_{r^{\epsilon-1}}(B_0(r)) > \frac{L_{r^{\epsilon-1}}}{R_{r^{\epsilon-1}}}\right) + \mathbf{P}\left(\frac{L_{r^{\epsilon-1}}}{R_{r^{\epsilon-1}}} > r^{3-6\epsilon}\right) \\ &\leq \mathbf{P}(d(0, H_{r^{\epsilon-1}}) \leq r) + \mathbf{P}(L_{r^{\epsilon-1}} > r^{2-4\epsilon}) + \mathbf{P}(R_{r^{\epsilon-1}} \leq r^{-1+2\epsilon}) \\ &\leq \mathbf{P}(d(0, H_{r^{\epsilon-1}}) \leq r) + \exp(-r^{-\epsilon}), \end{aligned} \tag{35}$$

where the last line follows, for all $r > 0$ small enough, from the bounds in Lemma 5.2.

Most of the work now consists in bounding the first term in (35) above. Observe that the geodesic from $H_{r^{\epsilon-1}}$ to 0 must cross every single one of the metric spaces induced by $\mathscr{M}$ on the intervals $q_\lambda \subseteq [0,1]$ with $\lambda > r^{\epsilon-1}$. Furthermore, with the notation of Section 5.3, $d(0, H_{r^{\epsilon-1}})$ decomposes as follows: since here the portion of path in q_λ is precisely between $\boldsymbol{j}(\sup q_\lambda) \in q_\lambda$ a.s. and $\inf q_\lambda$; we have

$$d(0, H_{r^{\epsilon-1}}) = \sum_{\lambda \in \Lambda^L} Y_\lambda \mathbf{1}_{\lambda > r^{\epsilon-1}} \geq_{st} \sum_{\lambda \in \Lambda^L} Y_\lambda^\star \cdot m_\lambda^{1/2} \cdot \mathbf{1}_{\lambda > r^{\epsilon-1}},$$

where the last inequality is a stochastic minoration that relies on Proposition 5.6: the m_λ, $\lambda \in \Lambda^L$, are the sizes of the jumps of L_λ, and the $Y_\lambda^\star$ are conditionally independent from the entire collection $(m_\lambda : \lambda \in \Lambda^L)$. In order to lower bound $d(0, H_{r^{\epsilon-1}})$ it suffices to focus on a single term of the sum in the right-hand side: if any of those terms is greater than r, then $d(0, H_{r^{\epsilon-1}}) > r$ as well, thus

$$\begin{aligned}\mathbf{P}(d(0, H_{r^{\epsilon-1}}) \leq r) &\leq \mathbf{P}\left(\#\{\lambda > r^{1-\epsilon} : \lambda \in \Lambda^L, Y_\lambda^\star m_\lambda^{1/2} > r\} = 0\right) \\ &\leq \mathbf{P}\left(\sup\{m_\lambda : \lambda \in \Lambda^L, \lambda > r^{1-\epsilon}\} \leq r^{2-\epsilon}\right) + \mathbf{P}(Y^\star \leq r^{\epsilon/2}). \end{aligned} \tag{36}$$

The second term is at most $r^{c\epsilon}$ by Proposition 5.6 *ii)* and Markov's inequality. The (m_λ) are also the lengths of the faces of the convex minorant of X on the interval $[0,1]$ by Lemma 4.2, and to deal with the first term, we relate it to the convex minorant of a standard Brownian motion W.

Let $(Q_t)_{t \geq 0}$ be defined by

$$Q_t = \exp\left(\int_0^t s dW_s - \frac{1}{2}\int_0^t s^2 ds\right),$$

Then by the Cameron–Martin–Girsanov formula (Theorem 38.5 of [61]), the laws of X and W are related by a change of measure whose density is given by the martingale Q_t. For a function $\omega \in \mathcal{C}(\mathbb{R}_+, \mathbb{R})$, we consider the convex minorant of ω on $[0,1]$ and we let $\chi_r(\omega)$ denote the indicator that the longest face with slope (strictly) smaller than $-r^{\epsilon-1}$ has length at most $r^{2-\epsilon/2}$. Then, by Lemma 4.2 we have

$$\begin{aligned}\mathbf{P}\left(\sup\{m_\lambda : \lambda \in \Lambda^L, \lambda > r^{1-\epsilon}\} \leq r^{2-\epsilon}\right) &= \mathbf{E}[\chi_r(X^0)] \\ &= \mathbf{E}[\chi_r(W) \cdot Q_1] \\ &\leq \mathbf{E}[\chi_r(W)^2]^{1/2} \cdot \mathbf{E}[Q_1^2]^{1/2}, \end{aligned} \tag{37}$$

by the Cauchy–Schwarz inequality. Observe that in the right-hand side above, $\mathbf{E}[Q_1^2] \leq \mathbf{E}[\exp(2\overline{W}_1)]$ is finite and independent of r thanks to the Gaussian tails of $\overline{W}_1$.

The first factor in (37) can be estimated using the results of Pitman and Ross [58, Theorem 1] and Brownian scaling. Let $(x_i, s_i)_i$ be the points of a Poisson point process with intensity $(2\pi x)^{-1/2} \cdot \exp(-(2+s^2)x/2) dx ds$ on $\mathbb{R}_+ \times \mathbb{R}$. Then (x_i, s_i) are the lengths and slopes of the faces of the convex minorant of W on the interval $[0, E]$ where E is an independent exponential random variable with mean one, and here $E = \sum x_j$. Therefore, by Brownian scaling,

$$\begin{aligned}\mathbf{E}[\chi_r(W)] &= \mathbf{P}\left(\sup\{x_i : s_i\sqrt{E} < -r^{\epsilon-1}\} \leq E \cdot r^{2-\epsilon}\right) \\ &\leq \mathbf{P}(A^c) + \mathbf{P}\left(\sup\{x_i : s_i r^{1-5\epsilon/6} < -1\} \leq r^{2-4\epsilon/3}\right), \end{aligned}$$

where A denotes the event that $E = \sum x_j \in [r^{\epsilon/3}, r^{-\epsilon/3}]$. Since E is exponential with mean one, we have $\mathbf{P}(A^c) \leq 2r^{\epsilon/3}$ for all $r > 0$ small enough. We claim that there exists a constant $c > 0$ such that the second term above is no larger than $r^{c\epsilon}$, and in order the complete the proof, it suffices to justify that claim. We slightly change the scaling and write $\delta = r^{1-5\epsilon/6}$ and $\gamma = \epsilon/3$ to lighten the notation. Then $r^{2-4\epsilon/3} \leq \delta^{2+\gamma}$, and we focus on

$$\mathbf{P}\left(\sup\{x_i : s_i\delta < -1\} \leq \delta^{2+\gamma}\right) = \exp\left(-\iint \frac{e^{-(2+s^2)x/2}}{\sqrt{2\pi x}} \mathbf{1}_{x\geq\delta^{2+\gamma}, s\delta<-1} ds dx\right). \tag{38}$$

We just need to lower bound the integral in the right-hand side: consider the subregion Σ of $[\delta^{2+\gamma}, \infty) \times (-\infty, -1/\delta]$ where $u = xs^2 \leq 1$, that is $\Sigma := \{(x, s) : \delta^{2+\gamma} \leq x \leq \delta^2, -x^{-1/2} \leq s \leq -1/\delta\}$:

$$\begin{aligned}
\iint \frac{e^{-(2+s^2)x/2}}{\sqrt{2\pi x}} \mathbf{1}_{x\geq\delta^{2+\gamma}, s\delta<-1} ds dx &\geq \iint_\Sigma \frac{e^{-(s^2+2)x/2}}{\sqrt{2\pi x}} ds dx \\
&\geq e^{-\delta^2} \iint_\Sigma \frac{e^{-s^2x/2}}{\sqrt{2\pi x}} ds dx \\
&\geq \frac{e^{-\delta^2-1/2}}{\sqrt{2\pi}} \int_{\delta^{2+\gamma}}^{\delta^2} \left[\frac{1}{x} - \frac{1}{\delta\sqrt{x}}\right] dx \\
&\geq \frac{-\gamma \log \delta e^{-\delta^2-1/2}}{\sqrt{2\pi}},
\end{aligned}$$

for all $\delta > 0$ small enough. It follows easily that, there exists a constant $c > 0$ such that for all $\delta > 0$ small enough the right-hand side of (38) is at most $\delta^{c\gamma}$, which translated into the original parameters yields a bound of $r^{c'\epsilon}$ for the right-hand side of (37), and in turn for (36) and (35). This completes the proof. □

Finally, using Lemma 6.3, the following proposition completes the proof of Theorem 1.2

Proposition 6.5. *Let ζ be a point of $\mathscr{M}$ with distribution μ. Then for every $\epsilon \in (0, 1)$, almost surely, for all $r > 0$ small enough we have*

$$\mu(B_\zeta(r)) \leq r^{3-\epsilon}.$$

As a consequence $\dim_{\mathrm{H}}(\mathscr{M}) \geq 3$.

Proof. By Lemma 6.3, it suffices to prove the bound for $\mu(B_0(r))$. Fix $\lambda_0 = 1/r$, set $S = B_0(r)$ and recall the process $M_\lambda = \bar{\mu}_\lambda(\{x \leq R_\lambda : d(0, [x]_{\lambda_0}) \leq r\})$ of (32). Then, for any $\lambda \geq \lambda_0$, $\bar{\mu}_\lambda(B_0(r))$ is no larger than the $\bar{\mu}_\lambda$-mass of the excursions which are grafted within distance r or the origin: we have

$$\bar{\mu}_\lambda(B_0(r)) \leq M_\lambda.$$

Since M_λ is bounded, Lemma 6.2 implies that M_λ converges almost surely as $\lambda \to \infty$, but it also implies some concentration results since the increments of M_λ are bounded by the m_λ, $\lambda \in \Lambda^R$. By the Azuma–Hoeffding inequality [19, 23, 39], for any $x > 0$, we have, conditionally on $(m_\lambda, \lambda \in \Lambda)$,

$$\mathbf{P}\left(M_\lambda - M_{\lambda_0} > x \mid m_\lambda, \lambda > \lambda_0\right) \leq \exp\left(-\frac{x^2}{2\sum_{\lambda_0<\lambda'} m_{\lambda'}^2}\right). \tag{39}$$

Bounding the ℓ^2-norm of $(m_\lambda)_{\lambda>\lambda_0}$ is routine using Lemma 5.8. Indeed, for any $k \geq 1$,

$$\mathbf{E}\left[\sum_\lambda m_\lambda^2 \cdot \mathbf{1}_{\lambda\in\Lambda_k}\right] \leq 8k^5 \int_0^\infty \frac{t^2}{\sqrt{2\pi t^3}} e^{-k^6 t/8} dt = 64 \cdot k^{-4},$$

and Markov's inequality then yields, for some constant K,

$$\mathbf{P}\left(\sum_{\lambda:\lambda r>1} m_\lambda^2 \geq r^{3-\epsilon}\right) \leq r^{\epsilon-3}\cdot\mathbf{E}\left[\sum_\lambda m_\lambda^2\cdot\mathbf{1}_{\lambda>1/r}\right] \leq Kr^\epsilon\,. \tag{40}$$

Since $M_{\lambda_0}=\bar{\mu}_{\lambda_0}(B_0(r))$, it follows from Proposition 6.4 and (39)–(40) that

$$\begin{aligned}\mathbf{P}(\mu(B_0(r)\geq 2r^{3-\epsilon})) &\leq \mathbf{P}(\mu_{\lambda_0}(B_0(r))\geq r^{3-\epsilon}) + \mathbf{P}(M_\infty - M_{\lambda_0}\geq r^{3-\epsilon})\\ &\leq r^{c\epsilon}+\exp(-cr^{-\epsilon})+Kr^\epsilon\,.\end{aligned}$$

From there, completing the proof is standard: take a subsequence $r_i=2^{-i}$, $i\geq 1$; the Borel–Cantelli implies that for all but finitely many values of $i\geq 1$, we have $\mu(B_0(r_i))\leq 2r_i^{3-\epsilon}$, and thus $\mu(B_0(r))\leq 16r^{3-\epsilon}$ for all $r>0$ small enough. As a consequence, the mass distribution principle (see, e.g., Proposition 4.9 of [34]) implies that $\dim_{\mathrm{H}}(\mathscr{M})\geq 3-\epsilon$, which completes the proof since $\epsilon>0$ was arbitrary. □

7 Distances in the Brownian parabolic tree

All the proofs of convergence will be based on couplings with discrete objects. It would be possible to identify the distribution of $\mathrm{CMT}(X,\mathbf{U})$ as that of the scaling limit of the minimum spanning tree constructed in [4] directly in the continuum using the dynamics as λ evolves and the tools developed in [6]. However, since we need comparisons with discrete objects anyway for Theorems 1.3 and 1.4, we do not pursue this here. All the limit theorems essentially boil down to proving that, in a suitable coupling, and for every $\lambda\in\mathbb{R}$, the restriction of the metric space $\mathrm{CMT}(X,\mathbf{U})$ to any interval (a,b) of $\mathbb{R}_+\setminus Z^\lambda$ is the limit (in probability) of the minimum spanning tree of a connected component induced by a vertex set whose node have Prim ranks in an interval $\{a_n,a_n+1,\dots,b_n-1\}$ where $a_n\sim an^{2/3}$ and $b_n\sim bn^{2/3}$. Our coupling will be "global" in the sense that it allows a transparent application to any collection of times $\lambda_1<\lambda_2<\cdots<\lambda_k$ and any finite collection of intervals at these times.

7.1 Discrete preliminaries

In this section, we provide the discrete representation that we will use to prove our limit theorems. They all heavily rely on the Prim order introduced in [24] and its properties. We will in particular give a representation of the minimum spanning tree M_n, and of the random graph $G(n,p)$ that we will see as the union of a portion of the minimum spanning tree, the Kruskal forest denoted by $K(n,p)$, together with additional cyclic edges.

Recall the Prim algorithm and the Prim order $v_1,v_2,\dots,v_n$ discussed in Section 1.3. Recall also that $V_k=\{v_1,\dots,v_k\}$. For $k\in[n]$, let $N_k^{n,p}$ be the number of nodes in $[n]\setminus V_k$ which have a neighbour in V_k in the graph $G(n,p)$ whose edge set is $\{e:w_e\leq p\}$. For $\lambda\in\mathbb{R}$, set $p_n(\lambda)=1/n+\lambda n^{-4/3}$. Then define, for $t\geq 0$,

$$X_t^{n,\lambda}:=n^{-1/3}\left(N^{n,p_n(\lambda)}_{\lfloor tn^{2/3}\rfloor}-\#\{i\leq tn^{2/3}:N_i^{n,p_n(\lambda)}=0\}\right)\,. \tag{41}$$

Let $Z^{n,\lambda}$ be $n^{2/3}$ times the collection of instants when $X^{n,\lambda}$ reaches a new minimum. Then, the points of $Z^{n,\lambda}$ are the Prim ranks of the first vertices of the connected components of $G(n,p_n(\lambda))$. Furthermore, recalling the definitions in Section 1.3, the collection of the edges of the minimum spanning tree M_n are precisely $\{e_i=(u_i,v_i):2\leq i\leq n\}$. The identities of the nodes can of course not be recovered from $(X^{n,\lambda})_{\lambda\in\mathbb{R}}$, but one may use the Prim ranks to construct a graph on $[n]$ that is isomorphic to M_n using $(X^{n,\lambda})_{\lambda\in\mathbb{R}}$ only. However, the information about the location of the u_i vanishes in the limit, and we shall construct a graph that has the correct distribution of the left-end points u_i, conditionally on $(Z^{n,\lambda})_{\lambda\in\mathbb{R}}$.

We start with an encoding of the merges. Note that there are precisely $n-1$ jumps to the process $(Z^{n,\lambda})_{\lambda\in\mathbb{R}}$, each one corresponding to the appearance of one of the edges $e_i = (u_i, v_i)$ for some $2 \le i \le n$. Let $\boldsymbol{\ell}_n(i)$ and $\boldsymbol{r}_n(i)$ be respectively the Prim ranks of the left-most and right-most vertices of the connected component of v_i at time w_{e_i}; let also $\boldsymbol{s}_n(i) = (1 - nw_{e_i})n^{1/3}$ be the discrete analog of the slope of a point $t \in \mathscr{L}$ in the continuous setting. Then, the set

$$\mathrm{Merge}_n(X^n) := \{(n^{-2/3}\boldsymbol{\ell}_n(i), n^{-2/3}i, n^{-2/3}\boldsymbol{r}_n(i), -\boldsymbol{s}_n(i)) : 2 \le i \le n\} \tag{42}$$

contains all the information about the merges of connected components. We can rephrase the fact that the extremities u_i of the edges $e_i = (u_i, v_i)$ are uniform in the connected component containing v_{i-1} as follows. Let $(U_i)_{i\ge1}$ be i.i.d. uniform on $[0,1]$, also independent of $\mathrm{Merge}(X^n)$. For each i, let $\boldsymbol{j}_n(i) = \boldsymbol{\ell}_n(i) + \lfloor U_i(i - \boldsymbol{\ell}_n(i))\rfloor$. Then, $\boldsymbol{j}_n(i)$ is uniform in $\{\boldsymbol{\ell}_n(i), \boldsymbol{\ell}_n(i)+1, \dots, i-1\}$. The following lemma is a simple reformulation of Lemma 1.6.

Proposition 7.1 (A representation of the minimum spanning forest)**.** *Conditionally on* $\mathrm{Merge}_n(X^n)$*, the collection of Prim ranks of the nodes* $(u_i : 2 \le i \le n)$ *has the same distribution as* $(\boldsymbol{j}_n(i) : 2 \le i \le n)$*. In particular, up to a relabelling of the nodes of* M_n *using the Prim ranks:*

i) the graph on $[n]$ *with edges* $\{\boldsymbol{j}_n(i), i\}$*,* $2 \le i \le n$*, is distributed like* M_n*;*

ii) the graph on $[n]$ *with edges* $\{\boldsymbol{j}_n(i), i\}$*,* $2 \le i \le n$ *with* $-\boldsymbol{s}_n(i) \le \lambda$ *is distributed like the Kruskal forest* $K(n, p_n(\lambda))$*.*

We now move on the representation of the random graphs. We say that an edge is *cyclic* if it is the maximum weight edge of some cycle. For each $p \in [0,1]$, the graph $G(n,p)$ is formed of the portion of the minimum spanning tree consisting of the edges of weight at most p, together with the cyclic edges of weight at most p. Observe that while the edges of the minimum spanning tree are all a.s. a function of $(X^{n,\lambda})_{\lambda\in\mathbb{R}}$, this is not the case for the cyclic edges (with positive probability some information is lost, even at the discrete level). Again, rather than collecting the information from the random graph, it is more instructive to construct this information with the correct distribution conditionally on $(Z^{n,\lambda})_{\lambda\in\mathbb{R}}$.

Let $\{Y_{ij}, 1 \le i < j \le n\}$ be i.i.d. random variables uniform on $[0,1]$ and independent of everything else (namely $\mathrm{Merge}_n(X^n)$ and $(U_i)_{2\le i\le n}$). For each $1 \le i < j \le n$, let $\lambda^n_{ij} = (nY_{ij} - 1)n^{1/3}$. We store the information concerning cyclic edges in a point process. Define

$$\Xi_n := \left\{(in^{-2/3}, jn^{-2/3}, \lambda^n_{ij}) : 1 \le i < j \le n, Z^{n,\lambda^n_{ij}} \cap \{i+1, i+2, \dots, j\} = \varnothing\right\}, \tag{43}$$

so that Ξ_n is the collection of triples $(i/n^{2/3}, j/n^{2/3}, \lambda^n_{ij})$ for which v_i and v_j are in the same connected component of $G(n, p_n(\lambda^n_{ij}) - \delta)$ for $\delta > 0$ small enough. The total number of cyclic edges, sometimes called the surplus, of a connected component is also a quantity of interest, and can be expressed in terms of Ξ_n. Recall that $C^{n,\lambda}_i$, $i \ge 1$, denote the collection of vertex sets of the connected components of the random graph $G(n, p_n(\lambda))$, sorted in decreasing order of their sizes. For a discrete connected component $C^{n,\lambda}_j$, the number of surplus edges in $G(n, p_n(\lambda))$ is given by

$$\mathrm{surp}^{n,\lambda}_j = \#\{(x, y, \lambda') \in \Xi_n : xn^{2/3}, yn^{2/3} \in C^{n,\lambda}_j, \lambda' \le \lambda\}. \tag{44}$$

Proposition 7.2 (A representation of the random graph)**.** *Up to a relabelling of the nodes with the Prim ranks, the graph on* $[n]$ *with edge set consisting of the union of*

- *the edges* $\{\boldsymbol{j}_n(i), i\}$*,* $2 \le i \le n$ *such that* $\boldsymbol{s}_n(i) \ge -\lambda$*, and*
- *the edges* $\{i, j\}$*,* $1 \le i < j \le n$ *such that* $\lambda^n_{ij} \le \lambda$ *and* $Z^{n,\lambda^n_{ij}} \cap \{i+1, i+2, \dots, j\} = \varnothing$

has the same distribution as $G(n, p_n(\lambda))$*.*

Proof. Note first that we only care about the distribution of the edges of the second set that are not already in the first one. From Kruskal's algorithm, it is clear that, conditionally on the minimum spanning tree M_n, the weights of the edges in the complement are independent. Furthermore, for any fixed pair of

nodes $u, v \in [n]$ which are not adjacent in the minimum spanning tree, the weight of the edge between u and v is uniform, conditioned on being larger than the value $p_n(\lambda)$ at which u and v first become part of the same connected component. This is precisely what the second condition says when expressed in terms of the Prim ranks. □

Recall that $E_p^n = \{e \in E^n : w_e \leq p\}$ denotes edge set of the random graph $G(n,p)$. Define similarly $F_p^n \subseteq E_p^n$ the edge set of the minimum spanning forest $K(n,p)$, that is the collection of edges of the minimum spanning tree which have weight at most p.

Lemma 7.3. *Let $p \in [0,1]$.*
i) Conditionally on E_p^n, $(w_e : e \in E_p^n)$ is a family of i.i.d. uniform random variables (r.v.) on $[0,p]$;
ii) Conditionally on F_p^n, $(w_e : e \in F_p^n)$ is dominated by a family of i.i.d. uniform r.v. on $[0,p]$.

Proof. *i)* The first assertion is immediate since $\{w_e : e \in E_p^n\}$ is simply a collection of i.i.d. uniform r.v. on $[0,1]$ conditioned on being at most p. *ii)* The second claim is a consequence of Kruskal's algorithm: The set $F_p^n \subseteq E_p^n$ is obtained from E_p^n by iteratively removing the edge with maximum weight that belongs to a cycle, until there are no more cycles. The remaining edge have thus been selected for not being the maximum edge of any cycle; by *i)* the initial weights in E_p^n are i.i.d. uniform random variables on $[0,p]$, and the weights in F_p^n are therefore dominated by a collection of i.i.d. uniform r.v. on $[0,p]$. □

Lemma 7.4. *Fix any $\lambda \in \mathbb{R}$. Conditionally on $a, b \in [n]$, $a < b$, being two successive points of $Z^{n,\lambda}$,*
i) the vertices whose Prim ranks are in $\{a, a+1, \ldots, b-1\}$ form a connected component of $G(n, p_n(\lambda))$;
ii) conditionally on $S = \{v_a, \ldots, v_{b-1}\}$, the vertex v_a with Prim rank a is uniformly random in S, and independent of $G(n, p_n(\lambda))$.

Proof. The first claim is immediate from the definition of $X^{n,\lambda}$; see Section 4.1 of [24]. The second point is a consequence of the definition of the Prim order. Consider the time in Prim's algorithm when we decide who gets to have rank a: conditionally on the event in *i)*, this depends on an edge with weight (strictly) larger than $p_n(\lambda)$; conditionally on having its extremity in the set of vertices with Prim ranks $a, a+a, \ldots, b-1$, the end point is uniformly random, and declared to have Prim rank a. This completes the proof. □

7.2 Asymptotic properties of random graphs for $\lambda \to -\infty$

Recall that we identify the nodes with their Prim ranks, so v_i is simply denoted by i. For points $s_1, \ldots, s_k \in (0,\infty)$, let $\mathrm{Span}_n(s_1, \ldots, s_k)$ denote the collection of vertices that belong to one of the paths in the minimum spanning tree M_n between some $s_i^n = \lfloor s_i n^{2/3} \rfloor$ and $s_j^n = \lfloor s_j n^{2/3} \rfloor$. For $\lambda \in \mathbb{R}$ and $s_1, s_2, \ldots, s_k \in (0,\infty)$, let $J_\lambda^n(s_1, s_2, \ldots, s_k)$ be the collection of indices $j \geq 1$ such that $C_j^{n,\lambda}$ intersects $\mathrm{Span}_n(s_1, \ldots, s_k)$.

The following lemma shows that all the connected components containing part of the path in the minimum spanning tree between a collection of random points have a size of order $n^{2/3}$.

Proposition 7.5. *Let $I \subset (0,\infty)$ be any compact interval, and let $s_1, s_2, \ldots, s_k \in (0,\infty)$ be i.i.d. uniform in I. Then, for any $\epsilon > 0$ there exists $\lambda \in \mathbb{R}$, and $\delta > 0$, such that, with probability at least $1-\epsilon$, all the connected components of $G(n, p_n(\lambda))$ containing nodes of $\mathrm{Span}_n(s_1, \ldots, s_k)$ are trees and have size at least $\delta n^{2/3}$.*

Proof. *i)* We abbreviate $J_\lambda^n(s_1, \ldots, s_k)$ as J_λ^n. For any fixed λ, the collection of connected components containing any of the s_i^n, $1 \leq i \leq k$, have Prim ranks at most $n^{2/3} \sup I + |C_1^{n,\lambda}|$. With high probability, this is at most $tn^{2/3}$ for some fixed t for all $\lambda \leq 0$ (say). However, by the representations in [2] or [24], the number of surplus edges involving pairs of nodes with Prim rank at most $tn^{2/3}$ converges to a Poisson random variable whose parameter the $\int_0^t (X_s^\lambda - \underline{X}_s^\lambda) ds$, which tends to zero almost surely as $\lambda \to -\infty$.

Thus, for any $\epsilon > 0$, we can indeed choose λ small enough for all the $C_j^{n,\lambda}$, $j \in J_\lambda^n$ to be trees with probability at least $1 - \epsilon$.

ii) We shall prove that the family of random variables $\max\{n^{2/3}/|C_j^{n,\lambda}| : j \in J_\lambda^n\}$, $n \geq 1$, is tight. The arguments are all routine, and we only provide the main structure of the proof. Fix $\epsilon > 0$. First, let $\overline{\lambda}$ be large enough that $s_1^n, \ldots, s_k^n$ are all in the same connected component H^n which also contains the point $\lfloor n^{2/3} \rfloor$ with probability at least $1 - \epsilon$ (see for instance, Lemma 5.2).

By Lemma 7.3 *ii)*, when decreasing p from $p_n(\overline{\lambda})$ to $p_n(\lambda)$, each edge is removed with probability at most $(p_n(\overline{\lambda}) - p_n(\lambda))/p_n(\overline{\lambda}) \sim (\overline{\lambda} - \lambda)n^{-1/3}$ independently of the others. So, the number of edges removed on a prescribed path of length at most $Cn^{1/3}$ is dominated by a binomial random variable with parameters $Cn^{1/3}$ and $(\overline{\lambda} - \lambda)n^{-1/3}$ and is thus tight. Since $n^{-1/3}\operatorname{diam}(H^n) \leq n^{-1/3}\operatorname{diam}(M_n)$ which is tight ([1]), the same holds for the length of the path between any of two of the $\{s_1^n, \ldots, s_k^n\}$. This implies the tightness of $(|J_\lambda^n|)_{n\geq 1}$, for any $\lambda \in \mathbb{R}$. This also readily implies that $n^{1/3}$ divided by the smallest distance in the minimum spanning tree between any two removed edges is tight. On the other hand, the minimum distance between any two of the $\{s_1^n, \ldots, s_k^n\}$ is itself of order $n^{1/3}$ (this is lower bounded by the distance in the corresponding graph $G(n, p_n(\overline{\lambda}))$, and thus follows from the results in [2, 4]). It follows that the smallest portion of a path connecting the $\{s_1^n, \ldots, s_k^n\}$ in the Kruskal forest $K(n, p_n(\lambda))$ is also of order at least $n^{1/3}$.

Now, by *i)*, let λ be small enough that all the involved connected components are trees at time $p_n(\lambda)$ with probability at least $1 - \epsilon$. Conditionally on the number of its nodes being m, the diameter any such connected component is of order $m^{1/2}$ ([11, 12]). Putting this together with the facts that, for this value of λ, the number of portions of paths is tight J_λ^n, that each of the portions has length of order $n^{1/3}$, this implies that each of the portions is contained in a connected component whose size is indeed of order $n^{2/3}$ (and no smaller). □

The following folklore global asymptotic properties for the connected components will be useful.

Lemma 7.6. *For any $\epsilon, \delta, \delta' > 0$, there exists $\lambda \in \mathbb{R}$ such that, with probability at least $1 - \epsilon$,*
i) the largest connected component of $G(n, p_n(\lambda))$ contains at most $\delta n^{2/3}$ nodes;
ii) the maximum diameter of a connected component of $G(n, p_n(\lambda))$ is at most $\delta' n^{1/3}$.

Proof. For any $\alpha \leq \delta$, the probability that either *i)* or *ii)* fails is at most

$$\mathbf{P}(|C_1^{n,\lambda}| \geq \alpha n^{2/3}) + \mathbf{P}\left(\max_{j\geq 1} \operatorname{diam}(C_j^{n,\lambda}) \geq \delta' n^{1/3}, |C_1^{n,\lambda}| \leq \alpha n^{2/3}\right).$$

By Theorem 1.3 of [56], there exists $\alpha > 0$ small enough such that the second term is at most $\epsilon/2$. Then, choose λ small enough that the first term is also at most $\epsilon/2$. The fact that such a λ exists follows for instance from the results of [16] on the entrance boundary for the standard multiplicative (Theorem 4 there), and the relation between the random graph and the multiplicative coalescent in [14] (Proposition 4). □

7.3 A global coupling argument

Before actually proving the convergence of the trees or graphs seen as metric spaces, we verify that the main objects, on which the representations of the previous section rely, do converge. The objective is to eventually construct a rich enough probability space on which enough parameters converge almost surely, in order the make the final proof of convergence of the metric as easy as possible. The starting point is the process $(X^{n,\lambda})_{\lambda\in\mathbb{R}}$ introduced in (41). By Theorem 7 of [24], we have

$$(X^{n,\lambda})_{\lambda\in\mathbb{R}} \xrightarrow[n\to\infty]{} (X^\lambda)_{\lambda\in R}\,, \tag{45}$$

in distribution in $\mathbb{D}(\mathbb{R}, \mathbb{C}([0,\infty), \mathbb{R}))$. The first essential ingredient consists in proving that this implies that the macroscopic merges restricted to any compact region of time and space also converge. Define

$$\operatorname{Merge}(X) = \{(\boldsymbol{\ell}(t), t, \boldsymbol{r}(t), -\boldsymbol{s}(t)) : t \in \mathscr{L}(X)\}\,.$$

We say that $\mathrm{Merge}_n((X^{n,\lambda})_{\lambda\in\mathbb{R}}) \to \mathrm{Merge}(X)$ if for any compact intervals $I \subset (0,\infty)$ and $\Lambda \subset \mathbb{R}$, and any threshold $\epsilon > 0$, the subset of $\mathrm{Merge}_n((X^{n,\lambda})_{\lambda\in\mathbb{R}}))$ consisting of points $(l,t,r,-s)$ such that $r-t, t-l > \epsilon$, $l,t,r \in I$ and $-s \in \Lambda$ converges to the corresponding subset of $\mathrm{Merge}(X)$.

Proposition 7.7 (Convergence of large merges)**.** *Consider a probability space in which* $(X^{n,\lambda})_{\lambda\in\mathbb{R}} \to (X^\lambda)_{\lambda\in\mathbb{R}}$ *almost surely. Then, in probability,*

$$\mathrm{Merge}_n((X^{n,\lambda})_{\lambda\in\mathbb{R}}) \xrightarrow[n\to\infty]{} \mathrm{Merge}(X)\,.$$

Proof. We will use the following fact: a.s., there does not exist three local minima of $t, t', t'' \in \mathscr{L}(X)$ such that the points (t, X_t), $(t', X_{t'})$ and $(t'', X_{t''})$ all lie on the same line; using the representation in [24], this is essentially equivalent to the fact that the standard multiplicative coalescent is binary. To see that this is indeed the case, note that local minima of a continuous function are also global minima on an interval with rational extremities; then for, three disjoint intervals $[a,b]$, $[a',b']$ and $[a'',b'']$ with rational extremities, the local minima (t, X_t), $(t', X_{t'})$ and $(t'', X_{t''})$ on each of these intervals have a law which absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^2$, and then, are aligned with probability zero; the union of this countable number of zero probability events also has probability zero. Assume that $(a,b,c,\lambda) \in \mathrm{Merge}(X)$. In this case, b is a local minimum of X^λ, and X^λ is strictly above the (horizontal) line connecting $(a, X^\lambda_a), (b, X^\lambda_b), (c, X^\lambda_c)$ on $(a,b)\cup(b,c)$; because of the property recalled above, since (a, X^λ_a) and (b, X^λ_b) are local minima of X^λ, it is (a.s.) not the case for (c, X^λ_c), so that, $\inf\{X^\lambda_t : t \in (c, c+\eta)\} < 0$ for any $\eta > 0$.

Given compact intervals $I \subset (0,\infty)$, $\Lambda \subset \mathbb{R}$ and a threshold $\epsilon > 0$, there are only finitely many points in $\mathrm{Merge}(X) \cap I^3 \times \Lambda$ with the first three coordinates at least ϵ apart, and it suffices to consider each one separately. Take some $(a,b,c,\lambda) \in \mathrm{Merge}(X)$, so that, in particular $X^\lambda_a = X^\lambda_b = X^\lambda_c$. Consider, for $\varepsilon' > 0$, $\delta > 0$, the event

$$\begin{aligned}\mathcal{E}((a,b,c,\lambda);\varepsilon',\delta) = &\Big\{ \inf_{[0,a-\varepsilon']} X^\lambda \geq X^\lambda_a + \delta \Big\} \cap \Big\{ \inf_{[a+\varepsilon',b-\varepsilon']} X^\lambda > X^\lambda_a + \delta \Big\} \\ &\cap \Big\{ \inf_{[b+\varepsilon',c-\varepsilon']} X^\lambda > X^\lambda_a + \delta \Big\} \cap \Big\{ \inf_{[c,c+\varepsilon']} X^\lambda < X^\lambda_c - \delta \Big\}\,.\end{aligned}$$

Fix $\varepsilon' > 0$, $a,b,c \in (0,\infty)$ such that $|b-a|, |c-b| \geq 2\varepsilon'$, and $\lambda \in \mathbb{R}$. Using the properties of the local minima of the Brownian motion, for any $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\mathbf{P}(\mathcal{E}((a,b,c,\lambda);\varepsilon',\delta) \mid (a,b,c,\lambda) \in \mathrm{Merge}(X)) \geq 1-\epsilon.$$

Let us show that, the event $\mathcal{E}((a,b,c,\lambda);\varepsilon',\delta)$, for all n large enough, there must exist some vector (a_n,b_n,c_n,λ_n) close to (a,b,c,λ) such that $(a_n,b_n,c_n,\lambda_n) \in \mathrm{Merge}_n((X^{n,\lambda})_{\lambda\in\mathbb{R}})$. To prove this, it suffices to show that for some $\lambda' < \lambda$ close enough to λ there are points $(a_n,b_n,c_n) \in Z^{n,\lambda'}$ which are close to (a,b,c), while for some $\lambda'' > \lambda$ close enough to λ there are points $a''_n, b''_n \in Z^{n,\lambda''}$ with (a''_n, c''_n) close to (a,c) but no other point of $Z^{n,\lambda''}$ between a''_n and c''_n. We now proceed with the details.

On the one hand, for any $\lambda' < \lambda$ there exists $\delta' > 0$ small enough such that $X^\lambda_b < X^{\lambda'}_a - \delta'$ and $X^{\lambda'}_c < X^{\lambda'}_b - \delta'$. Taking λ' close enough to λ and $\delta' > 0$ even smaller, we may also ensure that $X^{\lambda'} > \delta'$ on $[a+\varepsilon', b-\varepsilon']$ and $[b+\varepsilon', c-\varepsilon']$. The convergence of $X^{n,\lambda'}$ to $X^{\lambda'}$ then ensures that we may find $a_n, b_n, c_n \in Z^{n,\lambda'}$ all within distance ε' of a, b or c.

On the other hand, for any $\lambda' > \lambda$, we have $X^{\lambda'} > \delta'$ on $(a+\varepsilon', c-\varepsilon')$; we may take λ' close enough to λ, and $\delta' > 0$ small enough such that we also have $\inf\{X^{\lambda'} : [c, c+\varepsilon']\} < \inf\{X^{\lambda'} : [0, a-\varepsilon']\} - \delta'$. The convergence of $X^{n,\lambda'}$ to X^λ then ensures that there exists $a_n \in [a-\varepsilon', a+\epsilon']$ and $c_n \in [c-\varepsilon', c+\varepsilon']$ such that $a_n, c_n \in Z^{n,\lambda'}$ and $Z^{n,\lambda}$ does not have any other point between a_n and c_n.

Consider now a accumulation point (a,b,c,λ) of $\mathrm{Merge}_n((X^{n,\lambda})_{\lambda\in\mathbb{R}})$ with $a < b < c$. Then, there exists a sequence (a_n,b_n,c_n,λ_n) converging to (a,b,c,λ) with $a_n,b_n,c_n \in Z^{n,\lambda_n}$. It follows that

$$X^{n,\lambda_n}_{a_n} = \underline{X}^{n,\lambda_n}_{a_n} = X^{n,\lambda_n}_{b_n} + n^{-1/3} = X^{n,\lambda_n}_{c_n} + 2n^{-1/3}\,.$$

The convergence of $X^{n,\lambda}$ to X^λ then implies that $\underline{X}^\lambda_a = X^\lambda_a = X^\lambda_b = X^\lambda_c$, so that $a, b, c \in Z^\lambda$. The path properties of X also imply that, for any $\lambda' > \lambda$, $(a, c) \cap Z^{\lambda'} = \varnothing$ so that $(a, b, c, \lambda) \in \mathrm{Merge}(X)$. This completes the proof. □

The next step concerns the convergence of the representation of the cyclic edges. Given X, let Ξ be a Poisson point process with intensity $\mathbf{1}_{x<y}\mathbf{1}_{(x,y)\cap Z^\lambda=\varnothing}dxdyd\lambda$ on $\mathbb{R}^2_+ \times \mathbb{R}$.

Proposition 7.8 (Convergence of cyclic edges)**.** *We have the convergence in distribution*

$$((X^{n,\lambda})_{\lambda\in\mathbb{R}}, \Xi_n) \xrightarrow[n\to\infty]{d} ((X^\lambda)_{\lambda\in\mathbb{R}}, \Xi)\,.$$

Furthermore, for any $a \in \mathbb{R}_+$ and $\lambda \in \mathbb{R}$, we have jointly,

$$\#\Xi_n \cap \Big([0, a]^2 \times (-\infty, \lambda]\Big) \xrightarrow[n\to\infty]{d} \#\Xi \cap \Big([0, a]^2 \times (-\infty, \lambda]\Big)\,.$$

Proof. Consider first the larger point process $\Xi^\circ_n = \{(in^{-2/3}, jn^{-2/3}, \lambda^n_{ij}) : 1 \le i < j \le n\}$. For every set $A \subset \{(x, y, \lambda) : a \le x < y \le b, \lambda_1 \le \lambda \le \lambda_2\}$, $|\Pi_n \cap A|$ is a binomial random variable with parameters asymptotic to $n^{4/3}(b-a)^2/2$ and $\mathbf{P}(\lambda^n_{ij} \in [\lambda_1, \lambda_2]) = (\lambda_2 - \lambda_1)n^{-4/3}$, and thus converges to a Poisson random variable P_A with parameter $(\lambda_2 - \lambda_1)(b-a)^2/2$.

Let now A and A' be any two disjoint such sets; we show that the distributional limits P_A and $P_{A'}$ are independent. If the space intervals $[a, b]$ and $[a', b']$ are disjoint, this is straightforward since the random variables (Y_{ij}) involved in the definitions of Π_n on A and A' are themselves independent. Otherwise the space intervals do intersect, and the time intervals must then be disjoint. Fix any $\epsilon > 0$. There is some constant K such that $\mathbf{P}(|\Pi_n \cap A| > K) < \epsilon$. Now, conditionaly on $\Pi_n \cap A$ and $|\Pi_n \cap A| \le K$, to estimate the distribution of $|\Pi_n \cap A'|$ we shall remove the pairs (i, j) which correspond to a point in $\Pi_n \cap A$, and correct the probability of every other to account for the fact that they did not occur in $[\lambda_1, \lambda_2]$ (for those that indeed intersect). This removes only at most K out of the $n^{4/3}(b'-a')^2/2$ pairs, and boots the probability of some of others by a factor $(1 - n^{-4/3}(\lambda_2 - \lambda_1))$. Overall, the limit remains Poisson random variable with the same distribution. Since $\epsilon > 0$ was arbitrary, this proves that Ξ°_n converges to a Poisson point process Ξ° with unit rate on $\{(x, z, \lambda) \in \mathbb{R}^2_+ \times \mathbb{R} : x < y\}$.

For the remainder of the proof, we consider now a probability space on which $(X^{n,\lambda})_{\lambda\in\mathbb{R}}$ converges almost surely to $(X^\lambda)_{\lambda\in\mathbb{R}}$. To complete the proof, it now suffices to show that, the set $A^n = \{(in^{-2/3}, jn^{-2/3}, \lambda) : Z^{n,\lambda} \cap \{i+1, \dots, j\} = \varnothing\}$ used to filter the points of Ξ°_n converges to $A = \{(x, y, \lambda) : Z^\lambda \cap (x, y) = \varnothing\}$ used to filter those of Ξ°, for the Hausdorff distance. Indeed, since $Z^{n,\lambda}$ and Z^λ are both decreasing in λ this would imply the convergence of their Lebesgue measures. Let $(x, y, \lambda) \in A$, then $(x, y) \cap Z^\lambda = \varnothing$, so that for any $\epsilon > 0$ small enough, $\inf\{X^\lambda_s - \underline{X}^\lambda_s : s \in (x+\epsilon, y-\epsilon)\} > 0$. It follows that $\inf\{X^{n,\lambda}_s - \underline{X}^{n,\lambda}_s : s \in (x+\epsilon, x-\epsilon)\} > 0$ as well for all n large enough, so that $(x + \epsilon, y - \epsilon, \lambda) \in A^n$. Similarly, if $(x, y, \lambda) \notin A$, then $\inf\{X^\lambda_s - \underline{X}^\lambda_s : s \in (x + \epsilon, y - \epsilon)\} \le 0$ for every $\epsilon > 0$ small enough, and therefore $\inf\{X^{\lambda-\epsilon}_s - \underline{X}^{\lambda-\epsilon}_s : s \in (x + \epsilon, y - \epsilon)\} < 0$ for some $\epsilon > 0$ small enough. The argument we used above implies that $(x+\epsilon, y-\epsilon, \lambda-\epsilon) \notin A^n$ for all n large enough.

For the second claim, one only needs the additional tightness of the number N_n of points of Ξ_n in $[0, a]^2 \times (-\infty, \lambda]$. Since the discrete representation of Ξ_n is delicate to handle, we shall change the point of view: By the exact distribution of N_n in Proposition 7.2, N_n is dominated by the number of surplus edges in the connected components that have nodes with Prim ranks at most $an^{2/3}$ at time $p_n(\lambda)$. The latter is known to be tight by the results in [24] (Corollary 20 and Section 7.2), which consider the alternative representation for the surplus edges using a Bernoulli pointset under the graph of the discrete reflected process $X^{n,\lambda} - \underline{X}^{n,\lambda}$ (just as in the results of Aldous in [14]). □

We may also recast the results of [24] about the convergence of surpluses of connected components in the Prim order in the present setting (see also, [14]). Recall the discrete defined in (44). There is a continuum analog to the surplus of a connected component, that can be defined in terms of Ξ. At time

$\lambda \in \mathbb{R}$, the connected components correspond to the intervals of $\mathbb{R}_+ \setminus Z^\lambda$, that are sorted in decreasing order as $(\gamma_j^\lambda)_{j\geq 1}$. We define

$$\mathrm{surp}_j^\lambda = \#\big\{(x,y,\lambda') \in \Xi : x,y \in \gamma_j^\lambda, \lambda' \leq \lambda\big\}\,. \tag{46}$$

Corollary 7.9. *Jointly with the convergence in Proposition 7.8, for any fixed $\lambda \in \mathbb{R}$, and any $j \geq 1$, we have*

$$\mathrm{surp}_j^{n,\lambda} \xrightarrow[n\to\infty]{d} \mathrm{surp}_j^\lambda\ .$$

We shall also want to couple the collection of uniform choices that are associated to each merge, and used to sample the edges in discrete and the point identifications in the limit. In the limit, merge events are characterized by some element (x,y,z,λ) where $0 < x < y < z$ and $\lambda \in \mathbb{R}$. Let $\Delta := \{(x,y,z) : 0 < x < y < z\} \subseteq \mathbb{R}_+^3$. We shall discretize the set of merge events. We decompose the collection of all potential merge triples Δ into countably many cells (up to a Lebesgue null set), so that each cell can contain at most one of the elements $(x,y,z) \in \Delta$ for which there exists some λ and $(x,y,z,\lambda) \in \mathrm{Merge}(X)$.

For $x > 0$, let $K(x)$ be the unique integer $k \in \mathbb{Z}$ such that $x \in [1/2^{k+1}, 1/2^k)$. For $k \in \mathbb{Z}$ and $(x,y,z) \in \Delta$ define

$$D(x,y,z;k) = \big(\lfloor x2^k\rfloor/2^k, \lfloor (y-x)2^k\rfloor/2^k, \lfloor (z-y)2^k\rfloor/2^k\big).$$

Then, the integer $M(x,y,z) = 1 + \max\{K(y-x), K(z-y)\}$ is used as the precision at which we shall encode the triple (x,y,z). For each $(x,y,z) \in \Delta$, define

$$\mathrm{Code}(x,y,z) = (M(x,y,z), D(x,y,z; M(x,y,z))\,,$$

and observe that Code takes its values in the countable set $\mathbb{N}\times\mathbb{Q}^3$. A subset $S \subset \Delta$ is called nested if for any $(x,y,z),(x',y',z') \in S$ we have either (a) $(x,z)\cap(x',z') = \varnothing$, or (b) (x',z') is contained in either (x,y) or (y,z), or (c) (x,z) is contained in either (x',y') or (y',z'). The set $\mathrm{Merge}(X)$ is such that its projection on the first three coordinates is nested. One easily verifies that, the map Code is injective on any nested subset $S \subset \Delta$.

Lemma 7.10. *Let $S \subseteq \Delta$ be nested. Then for any $(x,y,z),(x',y',z') \in S$ we have $\mathrm{Code}(x,y,z) \neq \mathrm{Code}(x',y',z')$.*

Proof. Suppose that $\mathrm{Code}(x_1,y_1,z_1) = \mathrm{Code}(x_2,y_2,z_2)$. Then, there is $k \in \mathbb{Z}$ and $(a,b,c) \in \mathbb{R}_+^3$ such that for $i \in \{1,2\}$, we have

i) $M(x_i,y_i,z_i) = k$, so that $\min\{|y_i - x_i|, |z_i - y_i|\} \in [1/2^k, 1/2^{k-1})$, and
ii) $x_i \in [a, a+1/2^k)$, $y_i - x_i \in [b, b+1/2^k)$ and $z - y \in [c, c+1/2^k)$.

Now, since S is nested, and there are two alternatives. Either $(x_1,z_1)\cap(x_2,z_2) = \varnothing$, and without loss of generality, $z_1 \leq x_2$, and because of i) we have $|x_1 - x_2| \geq 1/2^{k-1}$ which contradicts ii). Or, without of generality, $(x_1,z_1) \subset [x_2,y_2]$ or $(x_1,z_1) \subset [y_2,z_2]$, and either way, because of i), we cannot have simultaneously all the inequalities in ii). □

The following lemma is straightforward:

Lemma 7.11. *Suppose that $(a_n,b_n,c_n)_{n\geq 1}$ converges to $(a,b,c) \in \Delta$ such that neither a, $b-a$ nor $c-b$ are multiple of some 2^j, with $j \in \mathbb{Z}$. Then, for all n large enough $\mathrm{Code}(a_n,b_n,c_n) = \mathrm{Code}(a,b,c)$.*

From now on, we will see $\mathbf{U}$ as indexed by this $\mathrm{Code}(\Delta)$ (which is countable) rather than by $\mathbb{N}$.

7.4 A global coupling and ideal forests

We are now ready to define the probability space on which we will work. By iterative applications of Skorohod's representation theorem, we can find a probability space in which we have the following almost sure convergences, as $n \to \infty$:

- $(X^{n,\lambda})_{\lambda\in\mathbb{R}} \to (X^\lambda)_{\lambda\in\mathbb{R}}$;
- $\Xi_n \to \Xi$;
- $\mathrm{Merge}_n((X^{n,\lambda})_{\lambda\in\mathbb{R}}) \to \mathrm{Merge}(X)$.

Our starting point is the following: for the discrete objects, we consider the Kruskal forest $K(n,p)$ and random graphs $G(n,p)$ that are constructed from $(X^{n,\lambda})_{\lambda\in\mathbb{R}}$, $\mathbf{U}$, and Ξ_n, which also define the metric space (M_n, d_n). The continuous objects are those built from X, $\mathbf{U}$ and Ξ, in particular the metric d on $\mathbb{R}_+$. In the following, we shall identify the label and the Prim ranks to ease the discussion.

These objects are crucial for us, and we will show that their macroscopic structures are similar. Rather than trying to couple details at the scale $\delta n^{2/3}$ in the discrete, and δ in the continuous, we shall proceed as follows: both in the discrete and continuous setting, we can see the metric spaces at some given time λ (that is $K(n, p_n(\lambda))$ and $G(n, p_n(\lambda))$ in discrete, and the metric spaces induced by the intervals of $\mathbb{R}_+ \setminus Z^\lambda$ in continuous) as combining together the metric spaces that were already present far in the past, say at some time $\underline{\lambda} < \lambda$. We will never look any further in time, and replace the metrics in the connected components at time $\underline{\lambda}$ by some idealization. In general, the distribution will be incorrect, but we will ensure that $\underline{\lambda}$ can be chosen far enough for the distributions to be exact (or close enough) on an event of arbitrarily large probability. Observe that, this modification at time $\underline{\lambda}$ provides a coupling at all times $\lambda > \underline{\lambda}$ simultaneously.

Fix any two points $s_1, s_2 \in (0,\infty)$ and define $s_i^n = \lfloor s_i n^{2/3} \rfloor$, $i = 1, 2$. There is always some deterministic $\overline{\lambda}$ large enough such that s_1^n and s_2^n lie in the same connected component of $K(n, p_n(\overline{\lambda}))$ for n large enough with probability close to one; the path between s_1^n and s_2^n we refer to is the one in this connected component (and at any larger time). For any $\lambda \in \mathbb{R}$, let $J_\lambda^n(s_1, s_2)$ denote the collection of ranks of the connected components of $G(n, p_n(\lambda))$ that contain some node on the path between s_1^n and s_2^n; $j \in J_\lambda^n(s_1, s_2)$ means that the jth largest connected component of $K(n, p_n(\lambda))$ contains some node of the path between s_1^n and s_2^n. Similarly, let $J_\lambda(s_1, s_2)$ be the collection of indices of the intervals $(\gamma_j^\lambda)_{j\geq 1}$ obtained as $\mathbb{R}_+ \setminus Z^\lambda$ which contain part of $[\![s_1, s_2]\!]$. By construction, any of the connected components at time λ are traversersed by a single portion of the path, between two points that we will denote by a_j^n, b_j^n and a_j, b_j respectively. Then, we have the following exact decompositions:

$$d_n(s_1^n, s_2^n) = \sum_{j\in J_\lambda^n(s_1,s_2)} d_n(a_j^n, b_j^n) + \#J_\lambda^n(s_1, s_2) - 1\,, \tag{47}$$

and

$$d(s_1, s_2) = \sum_{j\in J_\lambda(s_1,s_2)} d(a_j, b_j)\,. \tag{48}$$

We will simply replace the distances in the components at time λ by what they should be in an ideal situation; for now, we are only interested in the definition, the verifications will come later. With this goal in mind, let us suppose our probability space contains the following sequences of random variables. Let $(V_j)_{j\geq 1}$ be i.i.d. random variables uniform on $[0, 1]$. For each $m \geq 1$, let F_m denote the distribution function of the distance D_m between two independent uniformly random points in a uniformly random labelled tree on m nodes. Then, for each $m \geq 1$, $\bar{D}_j(m) = m^{-1/2} F_m^{-1}(V_j)$. This provides a sequence of random variables where each term $\bar{D}_j(m)$ is distributed like $m^{-1/2} D_m$, and that converges almost surely as $m \to \infty$ to a Rayleigh random variable $\bar{R}_j$ with density $xe^{-x^2/2}$ on $\mathbb{R}_+$ (see, e.g., [11] for the convergence in distribution).

The objective is to control the matrix of pairwise distances between multiple points, and our new approximation of the distance will depend on the entire set of points. Let $\mathbf{s} = (s_1, s_2, \ldots, s_k) \in \mathbb{R}_+^k$. We

will only replace the distance in the connected components that only contain a single portion of paths between these points; in all the other components, which contain branch points of the collection of paths between the elements of $\mathbf{s}$, we will keep the distance unchanged. Let

$$J_\lambda^n(s_p, s_q; \mathbf{s}) = J_\lambda^n(s_p, s_q) \setminus \bigcup_{i<j, i\neq p, j\neq q} J_\lambda^n(s_i, s_j)\,,$$

and similarly, define the continuum analog by

$$J_\lambda(s_p, s_q; \mathbf{s}) = J_\lambda(s_p, s_q) \setminus \bigcup_{i<j, i\neq p, j\neq q} J_\lambda(s_i, s_j)\,.$$

Let $(n^{2/3}|\gamma_j^{n,\lambda}|)_{j\geq 1}$ denote the collection of sizes of the connected components at time λ, just as $(|\gamma_j^\lambda|)_{j\geq 1}$ denotes the Lebesgue measures in the continuous setting. Define the following approximations, for $1 \leq p < q \leq k$,

$$\begin{aligned}\tilde{d}_n(s_p^n, s_q^n) = n^{1/3} &\sum_{j\in J_\lambda^n(s_p,s_q;\mathbf{s})} |\gamma_j^{n,\lambda}|^{1/2} \bar{D}_j(n^{2/3}|\gamma_j^{n,\lambda}|) \\ &+ \sum_{j\in J_\lambda^n(s_p,s_1)\setminus J_\lambda^n(s_p,s_1;\mathbf{s})} d_n(a_j^n, b_j^n) + \#J_\lambda^n(s_1, s_2) - 1\,,\end{aligned} \tag{49}$$

and

$$\tilde{d}(s_p, s_q) = \sum_{j\in J_\lambda(s_p,s_q;\mathbf{s})} |\gamma_j^\lambda|^{1/2}\bar{R}_j + \sum_{j\in J_\lambda(s_p,s_q)\setminus J_\lambda(s_p,s_q;\mathbf{s})} d(a_j, b_j)\,. \tag{50}$$

We first verify that these provide a suitable coupling of the pairwise distances between the k points $s_1^n, s_2^n, \ldots, s_k^n$ and $s_1, s_2, \ldots, s_k$, respectively.

Proposition 7.12. *Fix some compact interval $I \in (0, \infty)$. For any $\epsilon > 0$, there exists $\lambda \in \mathbb{R}$ and an event A of probability at least $1 - \epsilon$, such that, for any k i.i.d. uniform points $s_1, s_2, \ldots, s_k \in I$, for all n large enough, on the event A, we have*

$$(\tilde{d}_n(s_p, s_q))_{1\leq p<q\leq k} \overset{d}{=} (d_n(s_p^n, s_q^n))_{1\leq p<q\leq k} \quad \textit{and} \quad (\tilde{d}(s_p, s_q))_{1\leq p<q\leq k} \overset{d}{=} (d(s_p, s_q))_{1\leq p<q\leq k}\,.$$

Proof. There exists a λ_1 large enough that I is contained in a single connected component at time λ_1: with Lemma 5.2 in mind, let λ_1 be the smallest λ for which $\sup I < 2\lambda + 1$ and $\exp(-c\lambda) < \epsilon/2$. This value being fixed, I is contained in $[0, 2\lambda_1 + 1]$ with probability at least $1 - \epsilon/2$.

Let $A = A_\lambda$ be the event that Ξ does not contain any point in $[0, 2\lambda_1 + 1]$ with time lower than λ. From the correspondence between the intensity of Ξ and the area of $X^\lambda - \underline{X}^\lambda$, there exists λ small enough that A has probability at least $1 - \epsilon$. The convergence of Ξ_n to Ξ implies that, on this event, for all n large enough, Ξ_n also has no point in I with times before λ (Proposition 7.8).

We may choose λ even smaller to ensure that, $|\gamma_1^\lambda| < \epsilon/(4k)$, so that, the probability that some point s_i, $1 \leq i \leq k$, falls in an interval of $\mathbb{R}_+ \setminus Z^\lambda$ that is not fully contained in I is at most $\epsilon/(2|I|)$. When this occurs, conditionally on $s_i \in \gamma_j^\lambda$, the position s_i is uniformly random in γ_j^λ. The same holds true for the discrete counterparts s_i^n for all n large enough.

Now, on the event A, for all n large enough, all the connected components of the random graph $G(n, p_n(\lambda))$ containing nodes with label at most $(2\lambda_1 + 1)n^{2/3}$ are all trees, which are uniformly random. These are thus identical to the components in the Kruskal forest $K(n, p_n(\lambda))$. By Proposition 7.1 and Lemma 7.4 the points a_j^n and b_j^n are independent and uniformly random (their actual labels!) and independent of the component. Since the end points s_i^n are themselves uniformly random in the connected component in which they lie, by the previous paragraph, this proves that, on A, the discrete approximation $\tilde{d}(s_p, s_q)$, $1 \leq p < q \leq k$, has the same distribution as $d(s_p, s_q)$, $1 \leq p < q \leq k$.

The continuous analog follows from the calculations in Section 2.1 of [1] saying that, conditionally on having no point under the curve, an excursion under $\tilde{\mathbf{n}}_\sigma$ is distributed according to $\mathbf{n}_\sigma$, and is thus exactly a Brownian excursion; Brownian scaling and Remark 3.23 saying that in $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$, the distance $d_{\mathrm{e}}(0, V)$ between 0 and a uniformly random point V is Rayleigh distributed, which completes the proof. □

7.5 Main proof of convergence

Finally, we are ready to prove that, in the probability space defined in the previous section, we have convergence in probability of the pairwise distance.

Proposition 7.13. *Fix I a compact interval of $(0, \infty)$, and let $s_1, s_2, \dots, s_k$ be i.i.d. uniform points in I. For any $\epsilon, \delta > 0$, there exists $\lambda \in \mathbb{R}$ such that,*

$$\limsup_{n\to\infty} \mathbf{P}\left(\sup_{1\le p<q\le k} \left| n^{-1/3} \tilde{d}_n(s_p^n, s_q^n) - \tilde{d}(s_p, s_q) \right| > \delta \right) \le \epsilon \,.$$

Proof. Fix any $\epsilon, \delta > 0$. Let us first deal with the portions of paths contained in connected components that are traversed by more than one path, and that we did not bother coupling. Consider the event A in Proposition 7.12 and the corresponding value λ_1 for λ which ensures that $\mathbf{P}(A^c) \le \epsilon/4$. By Lemma 7.6, there exists λ_2 such that, for all $\lambda \le \lambda_2$ and all n large enough, the probability that the maximum diameter of a connected component of $G(n, p_n(\lambda))$ is larger than $n^{1/3}\delta/(3k)$ is at most $\epsilon/4$. Furthermore, on the event A, each one of the $k-1$ portions of continuum paths in the intervals γ_j^λ which contain more than one portion has a length stochastically dominated by $|\gamma_j^\lambda|\bar{R}_j$ (or the diameter of the corresponding CRT). We can choose λ_3 small enough such that the probability that any of them is greater than $\delta/(3k)$ is at most $\epsilon/4$. Fix $\lambda = \min\{\lambda_1, \lambda_2, \lambda_3\}$. Finally, by Proposition 7.5, for this value of λ, there is some $\delta' > 0$ small enough such that, with probability at least $1 - \epsilon/4$ all the connected components $C_j^{n,\lambda}$, $j \in J_\lambda^n(s_p, s_q)$, $1 \le p < q \le k$, contain at least $\delta' n^{2/3}$ nodes. The probability that either of these bad events occur is at most ϵ, and we now suppose we work on the event A' that none occurs.

On the event A', we have from (49) and (50), for any $1 \le p < q \le k$,

$$\begin{aligned} \left| n^{-1/3} \tilde{d}_n(s_p^n, s_q^n) - \tilde{d}(s_p, s_q) \right| \le & \left| \sum_{j \in J_\lambda^n(s_p, s_q; \mathbf{s})} n^{-1/3} |\gamma_j^{n,\lambda}|^{1/2} \bar{D}_j(n^{2/3}|\gamma_j^{n,\lambda}|) - \sum_{j \in J_\lambda(s_p, s_q; \mathbf{s})} |\gamma_j^\lambda|^{1/2} \bar{R}_j \right| \\ & + n^{-1/3} \# J_\lambda^n(s_p, s_q) + 2\delta/3 \,. \end{aligned}$$

Since the $\# J_\lambda^n(s_p, s_q)$ are all tight by the proof of Proposition 7.5, we only need to deal with the first term the right-hand side above.

We claim that the fact that all discrete connected components $C_j^{n,\lambda}$, for $j \in J_\lambda^n(s_p, s_q)$ for some $1 \le p < q \le k$ contain at least $\delta' n^{2/3}$ nodes, the convergence of the merge events implies that, for all n large enough, we have $J_\lambda^n(s_p, s_q) = J_\lambda(s_p, s_q)$ for every $1 \le p < q \le k$. The reason is the following: (1) for all n large enough, for every i, if $s_i \in \gamma_j^\lambda$, then $s_i^n \in C_j^{n,\lambda}$, because $\{s_1, \dots, s_k\}$ and Z^λ are almost surely disjoint. (2) The merges of large connected components do converge because $\mathrm{Merge}((X^{n,\lambda})_{\lambda\in\mathbb{R}}) \to \mathrm{Merge}(X)$. (3) The points random points constructed in the discrete and continuous model for matching merges use the same uniforms by Lemma 7.11. It follows that, for n large enough, these points themselves end up in matching pair of discrete and continuum components. (4) The number of such merges is finite (the $\# J_\lambda^n(s_p, s_q)$, $J_\lambda(s_p, s_q)$ are tight). As a consequence, for all n large enough, we are lead to bounding

$$\mathbf{P}\left(\left| \sum_{j \in J_\lambda(s_p, s_q; \mathbf{s})} n^{-1/3} |\gamma_j^{n,\lambda}|^{1/2} \bar{D}_j(n^{2/3}|\gamma_j^{n,\lambda}|) - |\gamma_j^\lambda|^{1/2} \bar{R}_j \right| > \delta/3 \right) \,,$$

but we our coupling precisely ensures that every single term of the sum converges almost surely to zero. This completes the proof. □

7.6 Remaining proofs of convergence

Finally, we rely on the results of the previous section to complete the proofs of the remaining results, namely that of Theorem 1.3 about the MST of a connected graph with given surplus, and Theorem 1.4 about the dynamics for the limit random graph and Kruskal processes.

Before going further, let us discuss the types of convergence. Proposition 7.13 implies the convergence of the distribution of the matrix of pairwise distances between any finite number of points, and may thus be used to prove convergence in the Gromov–Prokhorov (GP) sense (Theorem 5 of [37]): indeed, for any λ, restriction of the d to any interval γ_j^λ, $j \geq 1$, is the limit of the metric of the discrete minimum on $C_j^{n,\lambda}$. The reason why this suffices to also prove convergence in the sense of Gromov–Hausdorff–Prokhorov (GHP) is that we actually already know that the sequences are tight for GHP ([2, 4]), and that the limit we construct has a mass measure which has full support because of Proposition 3.22 (see [18]). In the following, we thus only discuss GP convergence.

Proof of Theorem 1.4. *i)* Since the coupling is global, the proof of the joint convergence of the Kruskal forest $(\mathfrak{F}^{n,\lambda_1}, \dots, \mathfrak{F}^{n,\lambda_k})$ at times $\lambda_1 < \lambda_2 < \cdots < \lambda_k$ is an immediate consequence of Proposition 7.13, and the above discussion about the GHP versus GP convergence. The connected components at time λ_i correspond to the intervals of $\mathbb{R}_+ \setminus Z^{\lambda_i}$, equipped with the metric induced by d.
ii) For the same reason, the proof of the joint convergence $(\mathfrak{G}^{n,\lambda_1}, \dots, \mathfrak{G}^{n,\lambda_k})$ would be complete once we have an analog of Proposition 7.13 for the random graph at a fixed time. Proving this amounts to verifying that the joint convergence of the minimum spanning tree and of Ξ_n is sufficient to guarantee the convergence of the end points of every single surplus edge.

Once we have convergence of the end points of the edges, the techniques in [2] imply the convergence of the graph. Proving that we indeed have convergence of the locations of the end points of edges is not immediate because the function $d(x, y)$ is not continuous in either x or y. However, we can find a small $\underline{\lambda} \in \mathbb{R}$ such that the points appear between the correct connected components at time $\underline{\lambda}$ for all n large enough (almost surely, since the points have a diffuse distribution). Since the diameter of these components at time $\underline{\lambda}$ may be made arbitrarily small by choice of $\underline{\lambda}$, we do have convergence of the locations of the end points. This completes the proof of the sequence of graphs, in the product topology for a fixed λ. The extension to a vector of $(\lambda_1, \dots, \lambda_k)$ is immediate using the same arguments as above. □

Proof of Theorem 1.3. Consider the probability space from above, and fix some interval γ_i^λ of $\mathbb{R}_+ \setminus Z^\lambda$. Recall the discrete and continuum surplus defined in (44) and (46), respectively. Furthermore, surp_i^λ is a Poisson random variable with parameter the area of the process $X^\lambda - \underline{X}^\lambda$ on γ_i^λ. It thus follows from the calculations in Section 2.1 of [1] that,

$$\mathbf{E}\left[f((X^\lambda_{t_0+t} - \underline{X}^\lambda_{t_0+t})_{0\leq t\leq \sigma}) \;\middle|\; \gamma_i^\lambda = (t_0, t_0+\sigma), \mathrm{surp}_j^\lambda = s \right] = \frac{\mathbf{E}[f(\mathrm{e}_\sigma) \cdot (\int_0^\sigma \mathrm{e}_\sigma(u)du)^s]}{\mathbf{E}[(\int_0^\sigma \mathrm{e}_\sigma(u)du)^s]}$$

where e_σ is a Brownian excursion of duration σ. By definition, the right-hand side above is nothing else than $\mathbf{E}[f(\mathrm{e}_\sigma^{(s)})]$. Furthermore, on the event that $\mathrm{surp}_i^\lambda = s$, by Corollary 7.9, we have $\mathrm{surp}_i^{n,\lambda} = s$ for all n large enough. Therefore, up to a trivial relabelling, $C_i^{n,\lambda}$ is a uniformly random connected component with surplus s and size $\gamma_i^{n,\lambda}$. Since each of the values for surp_i^λ has positive probability, Theorem 1.3 follows from Proposition 7.13, and the discussion about the strengthening to Gromov–Hausdorff–Prokhorov convergence. □

Finally, we prove our main result about the entire minimum spanning tree. In [4], it is proved that the scaling limit of the minimum spanning tree can be constructed as the limit as $\lambda \to \infty$ of the scaling limit

of the minimum spanning tree of the largest connected component of the random graph at $p_n(\lambda)$. Here, we use the limit as $\lambda \to \infty$ of the connected component containing the vertex with Prim order $\lfloor n^{2/3} \rfloor$. We now verify that this coincides with our definition, which uses a connected component $H_\lambda$containing the point 1 and the measured metric space $(H_\lambda, d, \hat{\mu}_\lambda)$, $\hat{\mu}_\lambda$ is the (image of the) probability measure which is proportional to Lebesgue measure on H_λ. At this point, this should be essentially straightforward.

Proof of Theorem 1.1. Let $\mathcal{E}^\star_\lambda$ be the event that the largest connected component of $\mathbb{R}_+ \setminus Z^\lambda$ contains the point 1. Observe that, for all $\lambda \geq 2$, with $R_\lambda = \sup H_\lambda$,

$$\begin{aligned}\mathbf{P}(\mathcal{E}^\star_\lambda) &\geq \mathbf{P}(R_\lambda > 2\lambda - 1, |\gamma^\lambda_2| \leq 2) \\ &\geq 1 - \mathbf{P}(R_\lambda \leq 2\lambda - 1) - \mathbf{P}(|\gamma^\lambda_2| \geq 2).\end{aligned}$$

Lemma 5.2 implies that the first probability in the right-hand side above tends to zero as $\lambda \to \infty$. The same holds for the second one, see for instance, Proposition 5.3 of [6] which says that $|\gamma^\lambda_2|$ is $O(\lambda^{-2} \log \lambda)$ in probability. This also easily follows from Lemma 5.8 *i)*: indeed, for any natural number $i \geq 1$, on the event that $R_\lambda > i^3$, we have (with the notation of Section 5.4)

$$|\gamma^\lambda_2| \leq 1 + \sup_{k \geq i} \sup\{m_{\lambda'} : \lambda' \in \Lambda_k\},$$

which is at most $1 + i^{-5}$ with probability at least $1 - O(i^{-1/4})$. This implies $\mathbf{P}(|\gamma^\lambda_2| \geq 2) = O(\lambda^{-1/12})$, and in turn that $\mathbf{P}(\mathcal{E}^\star_\lambda) \to 1$ as $\lambda \to \infty$. By Proposition 7.13, $(H_\lambda, d, \hat{\mu}_\lambda)$ is the Gromov–Prokhorov limit (in distribution) of the minimum spanning tree of the connected component containing the vertex with Prim order $n^{2/3}$. However we know by the results of [7] that the sequence of rescaled minimum spanning trees converge for the Gromov–Hausdorff–Prokhorov topology, so that the convergence actually holds for GHP. Together with the fact that $\mathbf{P}(\mathcal{E}^\star_\lambda) \to 1$ as $\lambda \to \infty$, this proves that $(\mathscr{M}, d, \mu)$ has the same distribution as $(\mathscr{M}', d', \mu')$ constructed in [7]. □

References

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt. Critical random graphs: limiting constructions and distributional properties. *Electronic Journal of Probability*, 15:741–774, 2010.

[2] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probability Theory and Related Fields*, 152:367–406, 2012. doi: 10.1007/s00440-010-0325-4.

[3] L. Addario-Berry, N. Broutin, and C. Holmgren. Cutting down trees with a Markov chainsaw. *The Annals of Applied Probability*, 24:2297–2339, 2014.

[4] L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont. The scaling limit of the minimum spanning tree of the complete graph. *The Annals of Probability*, 45:3075–3144., 2017.

[5] L. Addario-Berry, D. Dieuleveut, and C. Goldschmidt. Inverting the cut-tree transform. *Annales de l'I. H. P. Probabilités et Statistiques*, 55:1349–1376, 2019.

[6] L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont. Continuum Erdős–Rényi and Kruskal dynamics. In preparation, 2023.

[7] Louigi Addario-Berry. The local weak limit of the minimum spanning tree of the complete graph. *arXiv preprint arXiv:1301.1667*, 2013.

[8] Louigi Addario-Berry and Sanchayan Sen. Geometry of the minimal spanning tree of a random 3-regular graph. *Probability Theory and Related Fields*, 180(3):553–620, 2021.

[9] M. Albenque and C. Goldschmidt. The Brownian continuum random tree as the unique solution to a fixed point equation. *Electronic Communications in Probability*, 20(61):1–14, 2015.

[10] D. Aldous. The continuum random tree II: an overview. In M.T. Barlow and N.H. Bingham, editors, *Stochastic Analysis*, pages 23–70. Cambridge University Press, 1991.

[11] D. Aldous. The continuum random tree. I. *The Annals of Probability*, 19:1–28, 1991.

[12] D. Aldous. The continuum random tree III. *The Annals of Probability*, 21:248–289, 1993.

[13] D. Aldous. Recursive self-similarity for random trees, random triangulations and Brownian excursion. *The Annals of Probability*, 22:527–545, 1994.

[14] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, 25:812–854, 1997.
[15] D. Aldous and J. Pitman. The standart additive coalescent. *The Annals of Probability*, 26:1703–1726, 1998.
[16] D.J. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. *Electronic Journal of Probability*, 3:1–59, 1998.
[17] I. Armendariz. *Brownian excursions and coalescing particle systems*. Phd thesis, New York University, 2001.
[18] S. Athreya, W. Löhr, and A. Winter. The gap between Gromov-vague and Gromov–Hausdorff-vague topology. *Stochastic Processes and their Applications*, 126:2527–2553, 2016.
[19] K. Azuma. Weighted sums of certain dependent random variables. *Tohoku Mathematical Journal*, 37:357–367, 1967.
[20] J. Bertoin. A fragmentation process connected to Brownian motion. *Probability Theory and Related Fields*, 117:289–301, 2000.
[21] J. Bertoin and G. Miermont. The cut-tree of large Galton-Watson trees and the Brownian CRT. *The Annals of Applied Probability*, 23:1469–1493, 2013.
[22] S. Bhamidi, A. Budhiraja, and X. Wang. The augmented multiplicative coaslescent, bounded size rules and critical dynamics of random graphs. *Probability Theory and Related Fields*, 160:733–796, 2014.
[23] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities - A nonasymptotic theory of independence*. Clarendon Press, Oxford, 2012.
[24] N. Broutin and J.-F. Marckert. A new encoding of coalescent processes. Applications to the additive and multiplicative cases. *Probability Theory and Related Fields*, 166:515–552, 2016.
[25] N. Broutin and M. Wang. Reversing the cut tree of the Brownian continuum random tree. *Electronic Journal of Probability*, 22(80):1–23, 2017.
[26] P. Chassaing and G. Louchard. Phase transition for parking blocks, Brownian excursion and coalescence. *Random Structures & Algorithms*, 21:76–119, 2002.
[27] I. Chiswell. *Introduction to Λ-trees*. World Scientific Publishing Company, Singapore, 2001.
[28] Josué Corujo and Vlada Limic. The standard augmented multiplicative coalescent revisited, 2023.
[29] Josué Corujo and Vlada Limic. A dynamical approach to spanning and surplus edges of random graphs, 2023.
[30] N. Curien and B. Haas. Random trees constructed by aggregation. *Annales de l'Institut Fourier*, 67:1963–2001, 2017.
[31] R.T. Durrett and D.L. Iglehart. Functionals of Brownian meander and Brownian excursion. *The Annals of Probability*, 5:130–135, 1977.
[32] S.N. Evans. *Probability and real trees, École d'Été de Probabilités de Saint-Flour XXXV-2005*, volume 1920 of *Lecture Notes in Mathematics*. Springer, 2005.
[33] K. J. Falconer. *The Geometry of Fractal Sets*, volume 85 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1986.
[34] Kenneth Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons Ltd., Chichester, 1990.
[35] N. Frilet. *Metric coalescence of homogeneous and inhomogeneous random graphs*. Phd thesis, Université Grenoble-Alpes, 2021. https://hal.inria.fr/tel-03667362v1.
[36] Siegfried Graf, R Daniel Mauldin, and Stanley C Williams. *The exact Hausdorff dimension in random recursive constructions*, volume 381. American Mathematical Soc., 1988.
[37] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (λ-coalescent measure trees). *Probability Theory and Related Fields*, 145(1):285–322, 2009.
[38] P. Groeneboom. The concave majorant of Brownian motion. *The Annals of Probability*, 11:1016–1027, 1983.
[39] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963.
[40] S. Janson. Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas. *Probability Surveys*, 4:80–145, 2007.
[41] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer, New York, 1988.
[42] D.P. Kennedy. The distribution of the maximum Brownian excursion. *Journal of Applied Probability*, 13: 371–376, 1976.
[43] D. E. Knuth. *The Art of Computer Programming: Sorting and Searching*, volume 3. Addison-Wesley, Reading, MA, 1973.
[44] Alan G Konheim and Benjamin Weiss. An occupancy discipline and applications. *SIAM Journal on Applied Mathematics*, 14(6):1266–1274, 1966.

[45] Igor Kortchemski and Paul Thévenin. Coupling Bertoin's and Aldous-Pitman's representations of the additive coalescent. *arXiv preprint arXiv:2301.01153*, 2023.
[46] J.B. Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, 2:48–50, 1956.
[47] J.-F. Le Gall. Brownian excursions, trees and measure-valued branching processes. *The Annals of Probability*, 19:1399–1439, 1991.
[48] J.-F. Le Gall. The uniform random tree in a Brownian excursion. *Probability Theory and Related Fields*, 96: 369–383, 1993.
[49] J.F. Le Gall and Y. Le Jan. Branching processes in Levy processes: Laplace functionals of snakes and superprocesses. *The Annals of Probability*, 26:1407–1432, 1998.
[50] Jean-François Marckert and Minmin Wang. A new combinatorial representation of the additive coalescent. *Random Structures & Algorithms*, 54(2):340–370, Apr 2018. doi: 10.1002/rsa.20775.
[51] James B. Martin and Balázs Ráth. Rigid representations of the multiplicative coalescent with linear deletion. *Electronic Journal of Probability*, 22(0), 2017. doi: 10.1214/17-ejp100.
[52] P. Mattila. *Geometry of sets and measures in Euclidean spaces: Fractals and rectifiability*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge Univ Press, 1999.
[53] R.D. Mauldin and S.C. Williams. Random recursive constructions: Asymptotic Geometric and Topological Properties. *Transactions of the American Mathematical Society*, 295:325–346, 1986.
[54] G. Miermont. Self-similar fragmentations derived from the stable tree I: Splitting at heights. *Probability Theory and Related Fields*, 127:423–454, 2003.
[55] Grégory Miermont and Sanchayan Sen. On breadth-first constructions of scaling limits of random graphs and random unicellular maps. *Random Structures & Algorithms*, 2022.
[56] A. Nachmias and Y. Peres. Critical random graphs: diameter and mixing time. *The Annals of Probability*, 36:1267–1286, 2008.
[57] E. Perkins. The exact Hausdorff measure of the level sets of Brownian motion. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 58:373–388, 1981.
[58] Jim Pitman and Nathan Ross. The greatest convex minorant of brownian motion, meander, and bridge. *Probability Theory and Related Fields*, 153(3-4):771–807, Aug 2011. doi: 10.1007/s00440-011-0385-0.
[59] J.W. Pitman. Remarks on the convex minorant of Brownian motion. In E. Çinlar, K.L. Chung, and Getoor R.K., editors, *Seminar on Stochastic Processes*, volume 5 of *Progress in Probability and Statistics*, pages 219–227, Boston, 1983. Birkhäuser.
[60] R. C. Prim. Shortest connection networks and some generalizations. *Bell Syst. Tech. J.*, 36:1389–1401, 1957.
[61] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Itô Calculus*, volume 2. Cambridge University Press, Cambridge, UK, 2 edition, 2000.
[62] R. Rossignol. Scaling limit of dynamical percolation on critical Erdős–Rényi random graphs. *The Annals of Probability*, 49:322–399, 2021.
[63] S.J. Taylor and J.G. Wendel. The exact Hausdorff measure of the zero set of a stable process. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 6:170–180, 1966.

A Auxiliary technical results

Lemma A.1. *Let $\omega \in \mathcal{C}(\mathbb{R}_+, \mathbb{R})$. Then*
i) the process $(Z^\lambda(\omega))_{\lambda\in\mathbb{R}}$ non-increasing and right-continuous with left-limits;
ii) for every $\lambda \in \mathbb{R}$, $Z^\lambda(\omega)$ and $Z^{\lambda-}$ are both closed;
iii) the set $\{\lambda \in \mathbb{R} : Z^{\lambda-}(\omega) \setminus Z^\lambda(\omega) \neq \varnothing\}$ is countable.

Proof. i) The monotony is a consequence of Lemma 4.1, this implies the existence of the left and right limits $\cap_{h>0} Z^{\lambda-h}$ and $\cup_{h>0} Z^{\lambda+h}$, respectively. The right-continuity follows by continuity of the maps $\lambda \mapsto \omega^\lambda$ and $\underline{\omega}^\lambda$: if $s \in Z^{\lambda-h}$ for all $h > 0$, then $\omega(s) + (\lambda - h)s = \inf\{\omega(r) + (\lambda - h)r : 0 \leq r \leq s\}$ for all $h > 0$, and thus this also holds for $h = 0$. *ii)* The fact that Z^λ is closed is an easy consequence of the continuity of ω. The monotony shows that $Z^{\lambda-}$ is a decreasing limit of closed sets, and is thus closed. *iii)* Since Z^λ and $Z^{\lambda-}$ are both closed for every $\lambda \in \mathbb{R}$, if λ is such that $Z^{\lambda-} \setminus Z^\lambda \neq \varnothing$, then

there exists $\epsilon > 0$ and $x = x_\lambda \in Z^{\lambda-}$ with $d(x, Z^\lambda) > \epsilon$. It follows that

$$\{\lambda \in \mathbb{R} : Z^{\lambda-} \setminus Z^\lambda \neq \varnothing\} = \bigcup_{n\geq 1} \{\lambda \in \mathbb{R} : \mathrm{d_H}(Z^{\lambda-}, Z^\lambda) > 1/n\}\,.$$

For each $n \geq 1$, there must exist for each λ a ball of radius $1/n$, and the collection of these balls must be disjoint. For each $n \geq 1$, any collection of open balls of radius $1/n$ must be countable, and therefore any set in the right-hand side above is countable. The claim follows. □

Lemma A.2 (Continuity properties of the metric d)**.** *Let $d(\cdot,\cdot)$ be the pseudo-metric on $[0,1]$ defined from the pair $(\mathrm{e}, \mathbf{U})$ used in the construction of $\mathrm{CMT}(\mathrm{e}, \mathbf{U})$. Almost surely,*

i) the map $d(0,\cdot)$ is continuous almost everywhere, but
ii) for every $x \in \mathscr{L}(\mathrm{e}) \cap (0,1)$, the map $d(0,\cdot)$ is not left-continuous at x, and
iii) for every $x \in \mathscr{L}(\mathrm{e}) \cap (0,1)$, the map $d(0,\cdot)$ is neither left- nor right-continuous at $\boldsymbol{\mathfrak{r}}(x)$.

Proof. *i)* Let x be uniformly random in $[0,1]$, then a.s. the vertices of the convex minorant of e on $[0,x]$, $(t_i(x))_{i\geq 0}$ and the corresponding intercepts $(z_i(x))_{i\geq 0}$ are such that $t_i(x) < x < z_i(x)$. Furthermore, there exists a sequence of local minima $z_n > x$ with $z_n \downarrow x$ such that the vertices of the convex minorant of e on $[0, z_n]$ are precisely $\mathcal{V}_{z_n} = \mathcal{V}_x \cup \{z_n\}$. Then, for any $i \geq 1$, $\sup\{d(x,t) : t \in (t_i, z_i)\} \leq D_i \cdot |t_i - z_i|^{1/2}$, where $(D_i)_{i\geq 1}$ are random variables distributed like the diameter of a continuum random tree of unit mass (which are not independent). It follows that

$$\begin{aligned}\mathbf{P}(\sup\{d(x,t) : t \in (t_i, z_i)\} > |z_i - t_i|^{1/4}) &\leq \mathbf{P}(D_i > |z_i - t_i|^{-1/4}) \\ &\leq \exp(-|z_i - t_i|^{-1/4}/2v)\,,\end{aligned}$$

for some constant $v > 0$. It follows by the Borel–Cantelli Lemma that a.s. $\sup\{d(x,t) : t \in (t_i, z_i)\} \leq |z_i - t_i|^{1/4}$ for all but finitely many values of i, so that $|d(0,t) - d(0,x)| \leq d(x,t) \to 0$ as $t \to x$.

ii) Let $x \in \mathscr{L} \cap (0,1)$; then a.s. there are only finitely many vertices in $\mathcal{V}_x$, and $x = t_i(x)$ for some $i \geq 1$. Let z_n be a sequence of local minima with $z_n \in (t_{i-1}, t_i)$ and $z_n \uparrow t_i$ as $n \to \infty$. Then, for all n_0 large enough, the vertices of the convex minorant of e on $[0, z_n]$ are exactly $\{t_j, j < i\} \cup \{z_n\}$. For each $n \geq n_0$, the point $\boldsymbol{\mathfrak{j}}(z_n)$ is uniform in (t_{i-1}, z_n) and $\boldsymbol{\mathfrak{j}}(x)$ is uniform in (t_{i-1}, t_i). With probability one, there exists a subsequence $(n_j)_{j\geq 1}$ such that $0 < \boldsymbol{\mathfrak{j}}(z_{n_j}) - t_{i-1} < \frac{1}{2}(\boldsymbol{\mathfrak{j}}(x) - t_{i-1})$. In particular, since $[\![0,x]\!]$ a.s. has an accumulation point at $\boldsymbol{\mathfrak{j}}(x)$, we have $\sup d(0, z_{n_j}) = \sup d(0, \boldsymbol{\mathfrak{j}}(z_{n_j})) < d(0, \boldsymbol{\mathfrak{j}}(x)) = d(0,x)$. It follows that, for any $\epsilon > 0$, $\inf\{d(0,s) : s \in (x-\epsilon, x)\} < d(0,x)$.

iii) For $x \in \mathscr{L} \cap (0,1)$, the point $\boldsymbol{\mathfrak{r}}(x)$ is some intercept, and the proof that $d(0,\cdot)$ is not left-continuous at $\boldsymbol{\mathfrak{r}}(x)$ is the same as in *ii)*. For the lack of right-continuity at $\boldsymbol{\mathfrak{r}}(x)$, this is also similar, but relies on the fact that one may find a sequence of local minima z_n in $(\boldsymbol{\mathfrak{r}}(x), 1)$ with $z_n \downarrow \boldsymbol{\mathfrak{r}}(x)$ such that $\mathcal{V}_{z_n} = \mathcal{V}_x \cup \{z_n\}$. The same argument as above can then be used by considering the random points $\boldsymbol{\mathfrak{j}}(z_n)$, which are independent, and uniform in $[x, z_n]$. □

Lemma A.3 (Surplus and area under the curve)**.** *Let* e *be a Brownian excursion. Consider the subset D of $[0,1]^2 \times \mathbb{R}$ of points (x,y,λ) such that $[x,y] \cap Z^\lambda(\mathrm{e}) = \varnothing$. Then, the 3-dimensional volume of D is equal to $\int_0^1 \mathrm{e}(x)dx$.*

Proof. Recall the recursive decomposition of Section 4.4. Then, the set D can be decomposed into countably many portions (with disjoint interior) D_u, $u \in \mathcal{U}$, as follows:

$$D_u := \{(x,y,\lambda) : a_u \leq x < y < R_{-\lambda}(a_u), \lambda \leq -\tau_u\}\,.$$

There is a corresponding decomposition of the set $\{(s,t) : s \in [0,1], 0 \leq t \leq \mathrm{e}(s)\}$ also into portions with disjoint interior, $E_u = \{(s, \mathrm{e}(a_u) - \lambda(s - a_u)) : a_u \leq s \leq R_{-\lambda}(a_u), -\lambda \geq \tau_u\}$, for $u \in \mathcal{U}$. We show that, for each $u \in \mathcal{U}$,

$$\int_{E_u} ds dt = \int_{D_u} dx dy d\lambda\,.$$

We treat the case $u = \varnothing$, the others are just the same, up to the more complicated notation. First observe that the left-hand side above with $u = \varnothing$ is precisely the area under the function f given by, for $i \geq 1$,

$$f(s) = \mathsf{e}(a_i) + \tau_i(s - a_i) \qquad a_i = R_{\tau_i} \leq s < R_{\tau_i -} \,.$$

Now, since each point $z = (s, f(s))$ can be represented in polar coordinates as $z = \rho(\theta)e^{i\theta}$, or alternatively by the pair $(-\lambda, R_{-\lambda}(0))$, where $-\lambda = f(s)/s$ is the slope of the line from 0 to z, we have

$$\int_{E_\varnothing} dsdt = \int_0^1 f(s)ds = \frac{1}{2}\int_0^{\pi/2} \rho(\theta)^2 d\theta = \int_{D_\varnothing} dxdyd\lambda \,.$$

The claim follows by summing the contributions for $u \in \mathcal{U}$. □