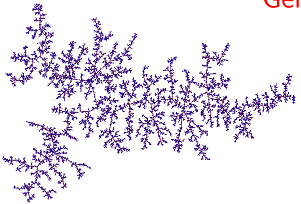


German Probability and Statistics Days 2018,
Freiburg im Breisgau

Voronoi cells in the Brownian continuum random tree

Christina Goldschmidt (Oxford)

Joint work with Louigi Addario-Berry (McGill), Omer Angel (UBC), Guillaume Chapuy (Paris 7) and Éric Fusy (École polytechnique)

Voronoi tessellations in the CRT and continuum random maps of finite excess, *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2018)*, pp.933-946.

Voronoi cells in the Brownian continuum random tree

Voronoi tessellations in the CRT and continuum random maps of finite excess, *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2018)*, pp.933-946.

Part I: Voronoi tessellations

Voronoi cells in a metric space

Let (M, d) be a metric space.

Fix $k \geq 1$ and let $S = \{x_i : 1 \leq i \leq k\}$ be a collection of points in M, the centres.

For $1 \leq i \leq k$, the Voronoi cells are

$$V_i = \{y \in M : d(y, S) = d(y, x_i)\}.$$

(Note that the Voronoi cells are not necessarily disjoint.)

Voronoi cells in a metric space

Let (M, d) be a metric space.

Fix $k \geq 1$ and let $S = \{x_i : 1 \leq i \leq k\}$ be a collection of points in M, the centres.

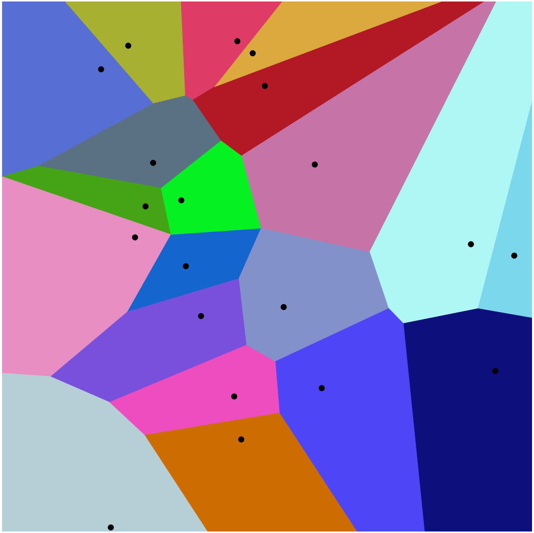
For $1 \leq i \leq k$, the Voronoi cells are

$$V_i = \{y \in M : d(y, S) = d(y, x_i)\}.$$

(Note that the Voronoi cells are not necessarily disjoint.)

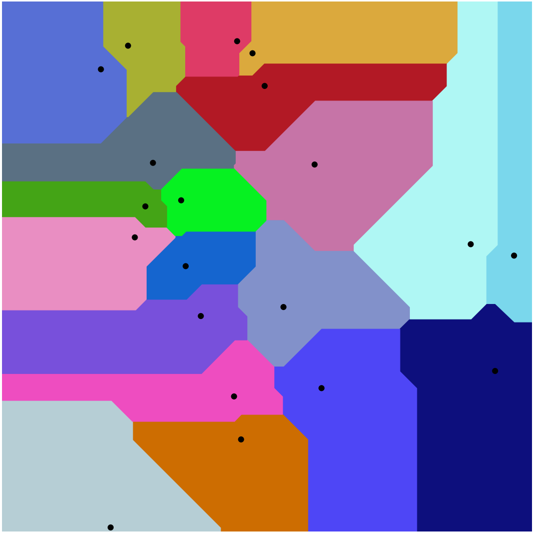
Standard example: Voronoi cells in $\mathbb{R}^2$

Euclidean distance

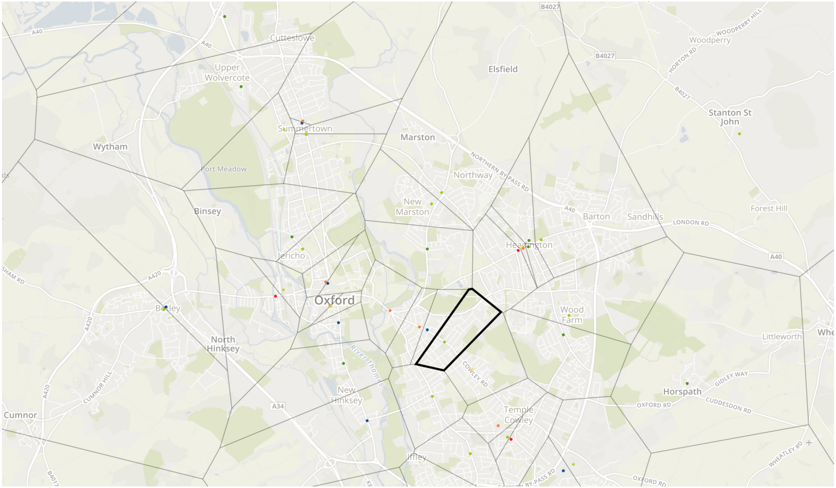


Standard example: Voronoi cells in $\mathbb{R}^2$

Manhattan distance

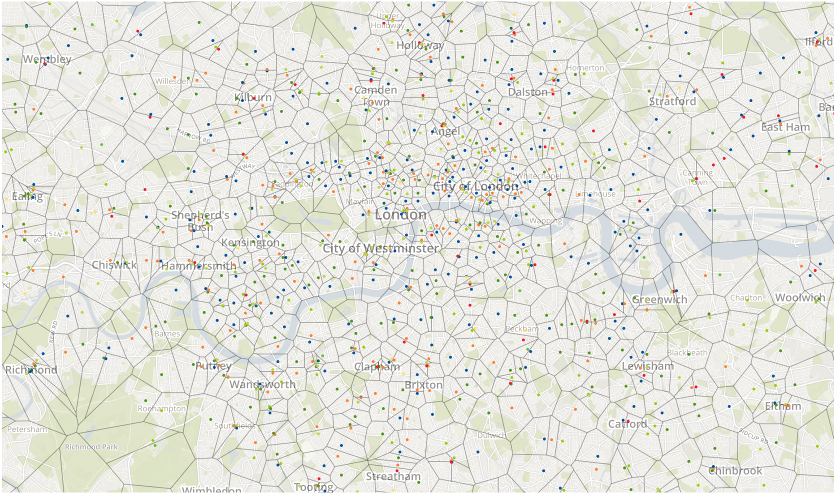


Voronoi supermarkets



See `https://chriszetter.com/voronoi-map/examples/uk-supermarkets/`

Voronoi supermarkets



See `https://chriszetter.com/voronoi-map/examples/uk-supermarkets/`

General set-up: Voronoi cells in a metric space

Let (M, d) be a metric space endowed with a Borel probability measure μ.

Fix $k \geq 1$ and let $S = \{x_i : 1 \leq i \leq k\}$ be a collection of points in M, the centres. Typically these will be random and i.i.d. samples from μ.

For $1 \leq i \leq k$, the Voronoi cells are

$$V_i = \{y \in M : d(y, S) = d(y, x_i)\}.$$

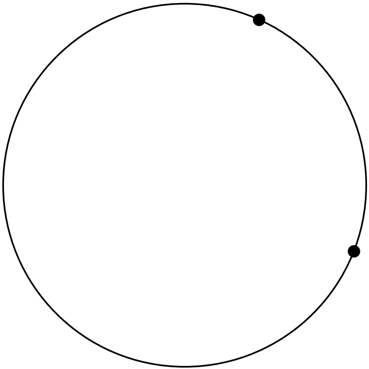
(Note that the Voronoi cells are not necessarily disjoint.)

We will be interested in the "masses" of these cells, as measured by μ, i.e.

$$(\mu(V_1), \mu(V_2), \dots, \mu(V_k)).$$

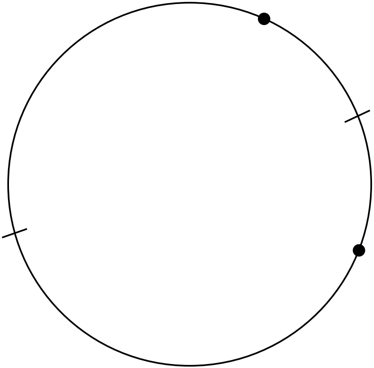
Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Any two points.



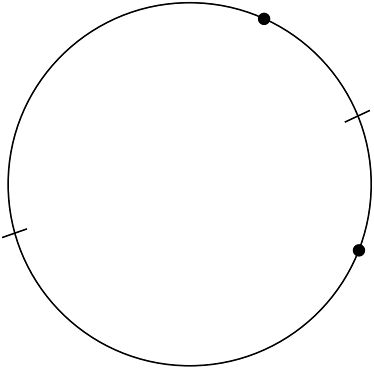
Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Any two points.



Warm-up: circle

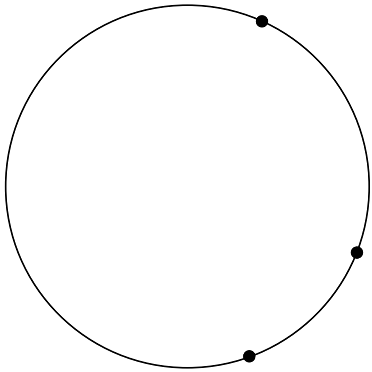
Circle of circumference 1, Euclidean distance, Lebesgue measure.
Any two points.



$(\mu(V_1), \mu(V_2)) = (1/2, 1/2)$.

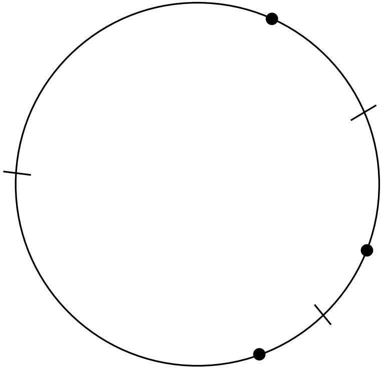
Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Three uniform points.



Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Three uniform points.



Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Three uniform points.

Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Three uniform points.

The lengths of these intervals are uniform on the 2-dimensional simplex i.e. have $\mathrm{Dir}(1,1,1)$ distribution.

Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Three uniform points.

The lengths of these intervals are uniform on the 2-dimensional simplex i.e. have Dir$(1, 1, 1)$ distribution.

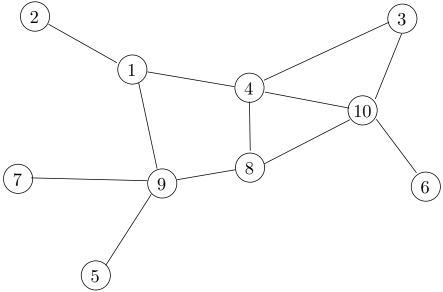
We get that the Lebesgue measures of the Voronoi cells are

$$(\mu(V_1), \mu(V_2), \mu(V_3)) = \left(\tfrac{1}{2} U_{(2)}, \tfrac{1}{2}\left(1 - U_{(1)}\right), \tfrac{1}{2}\left(1 - U_{(1)} - U_{(2)}\right)\right)$$

(exchangeable with marginals distributed as $\frac{1}{2}$Beta$(2, 1)$).

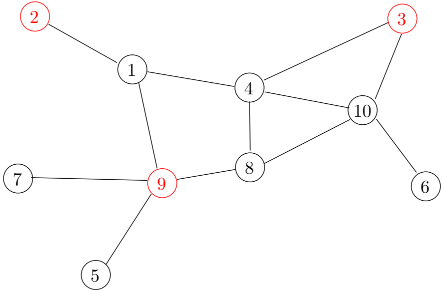
Voronoi cells in graphs

A very simple example of a metric space is a connected graph: the vertices are the points of the metric space and we use the graph distance for the metric.



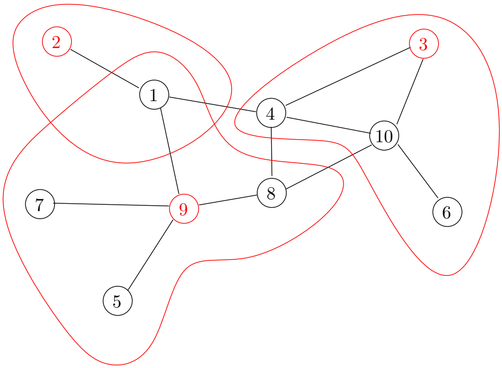
Voronoi cells in graphs

A very simple example of a metric space is a connected graph: the vertices are the points of the metric space and we use the graph distance for the metric.



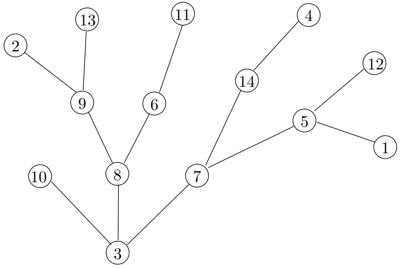
Voronoi cells in graphs

A very simple example of a metric space is a connected graph: the vertices are the points of the metric space and we use the graph distance for the metric.



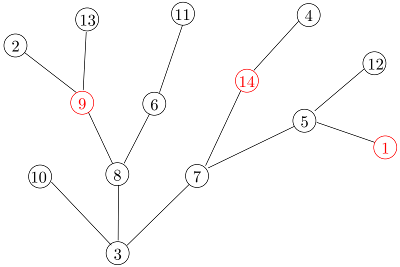
Setting of interest: random trees

Let T_n be a uniform random labelled tree on n vertices.



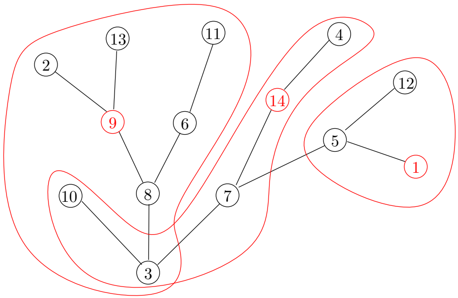
Setting of interest: random trees

Let T_n be a uniform random labelled tree on n vertices. Pick 3 uniform points (with replacement).



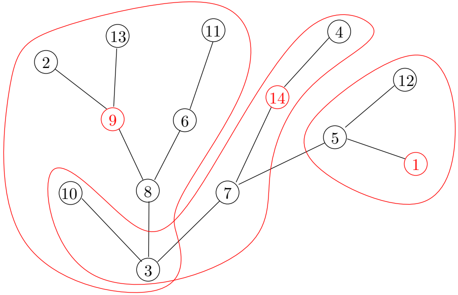
Setting of interest: random trees

Let T_n be a uniform random labelled tree on n vertices. Pick 3 uniform points (with replacement).



Setting of interest: random trees

Let T_n be a uniform random labelled tree on n vertices. Pick 3 uniform points (with replacement).



Question. If we sample k uniform points in T_n, how large are the Voronoi cells?

Voronoi mass-partition in random trees

Theorem. (Addario-Berry, Angel, Chapuy, Fusy & G, 2018)
Let T_n be a uniform random tree on n labelled vertices. Fix $k \geq 2$ and let $X_1^n, X_2^n, \ldots, X_k^n$ be independent uniform points. Let $V_1^n, V_2^n, \ldots, V_k^n$ be the corresponding Voronoi cells. Then

$$\frac{1}{n}\left(|V_1^n|, |V_2^n|, \ldots, |V_k^n|\right) \xrightarrow{d} \mathrm{Dir}(1, 1, \ldots, 1)$$

as $n \to \infty$.

Voronoi mass-partition in random trees

Theorem. (Addario-Berry, Angel, Chapuy, Fusy & G, 2018)
Let T_n be a uniform random tree on n labelled vertices. Fix $k \geq 2$ and let $X_1^n, X_2^n, \ldots, X_k^n$ be independent uniform points. Let $V_1^n, V_2^n, \ldots, V_k^n$ be the corresponding Voronoi cells. Then

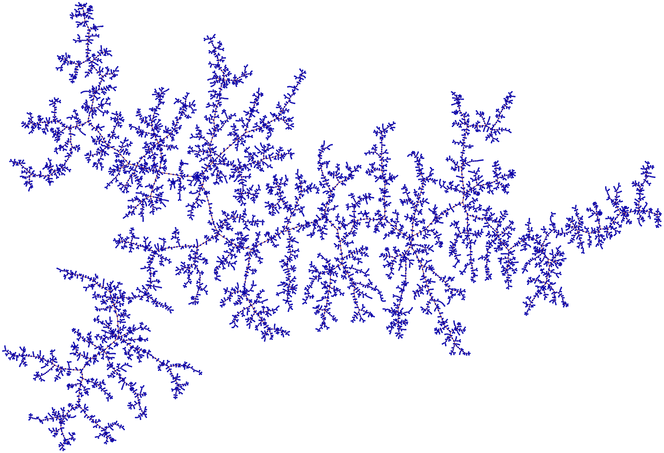
$$\frac{1}{n}\left(|V_1^n|, |V_2^n|, \ldots, |V_k^n|\right) \xrightarrow{d} \mathrm{Dir}(1, 1, \ldots, 1)$$

as $n \to \infty$.

If you want to chop up a random tree in a uniform manner, pick uniform points and find their Voronoi cells.

The Brownian continuum random tree

The neatest formulation (and proof) are for the scaling limit, the Brownian continuum random tree (CRT).



Picture by Igor Kortchemski

Part II: The Brownian CRT

Convergence to the Brownian CRT

Recall that T_n is a uniform random labelled tree on n vertices. Write d_n for the graph distance in T_n and μ_n for the uniform measure on the vertices.

Theorem. (Aldous, Le Gall)
We have

$$\left(T_n, \frac{1}{\sqrt{n}} d_n, \mu_n\right) \xrightarrow{d} (\mathcal{T}, d, \mu),$$

as $n \to \infty$, where $(\mathcal{T}, d, \mu)$ is the Brownian CRT.

$(\mathcal{T}, d)$ is a random path metric space. μ is usually referred to as its mass measure.

Convergence to the Brownian CRT

Recall that T_n is a uniform random labelled tree on n vertices. Write d_n for the graph distance in T_n and μ_n for the uniform measure on the vertices.

Theorem. (Aldous, Le Gall)
We have

$$\left(T_n, \frac{1}{\sqrt{n}} d_n, \mu_n\right) \xrightarrow{d} (\mathcal{T}, d, \mu),$$

as $n \to \infty$, where $(\mathcal{T}, d, \mu)$ is the Brownian CRT.

$(\mathcal{T}, d)$ is a random path metric space. μ is usually referred to as its mass measure.

The convergence occurs in the sense of the Gromov–Hausdorff–Prokhorov topology. This is essentially strong enough to deduce the convergence of the μ_n-masses of the Voronoi cells.

Universality

The convergence to the Brownian CRT holds, in fact, for a much more general class of trees. We may take T_n to be any Galton–Watson tree with offspring distribution of mean 1 and finite variance $\sigma^2 > 0$, conditioned to have precisely n vertices. Then

$$\left(T_n, \frac{\sigma}{\sqrt{n}} d_n, \mu_n\right) \xrightarrow{d} (\mathcal{T}, d, \mu),$$

as $n \to \infty$.

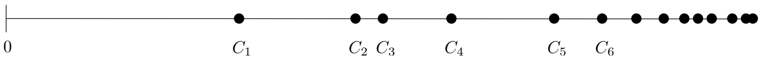
This class contains, for example,

- uniform random labelled trees
- uniform random plane trees
- uniform random binary trees.

Our theorem on the Voronoi mass-partition holds in these settings also.

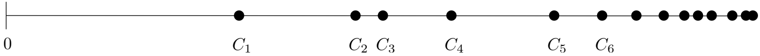
Construction 1: line-breaking

Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.



Construction 1: line-breaking

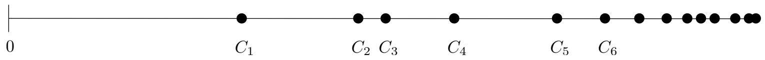
Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.



Consider the line-segments $[0, C_1)$, $[C_1, C_2), \ldots$.

Construction 1: line-breaking

Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.

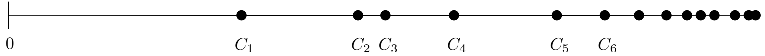


Consider the line-segments $[0, C_1)$, $[C_1, C_2)$, $\ldots$.

Start from $[0, C_1)$ and proceed inductively.

Construction 1: line-breaking

Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.

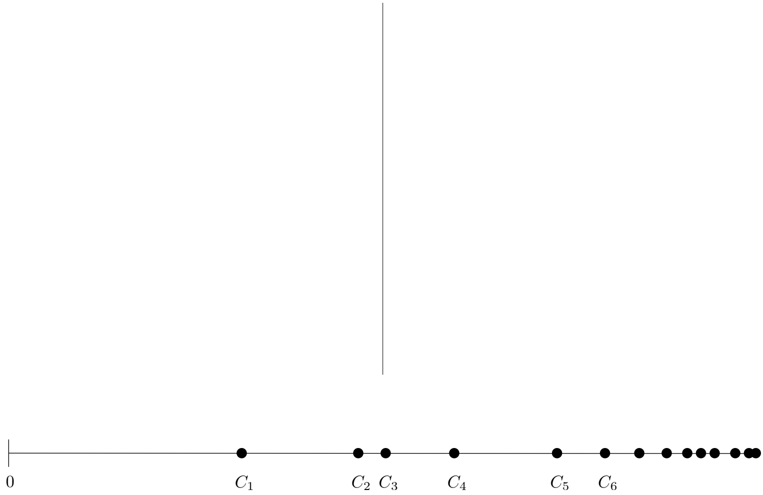


Consider the line-segments $[0, C_1)$, $[C_1, C_2), \ldots$.

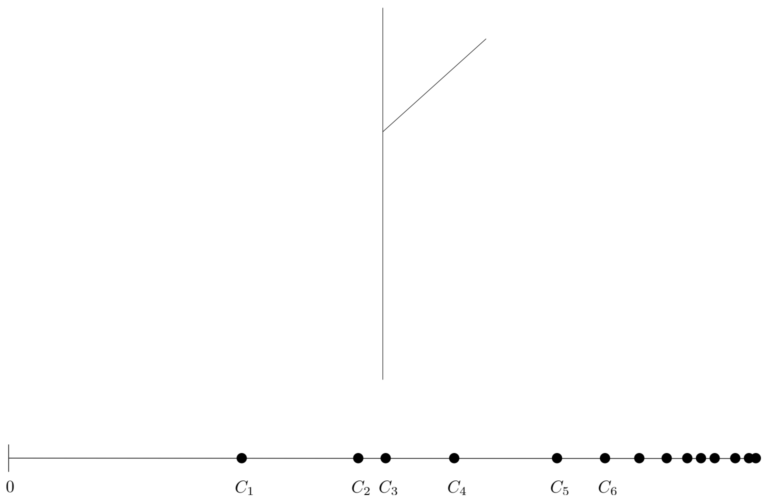
Start from $[0, C_1)$ and proceed inductively.

For $i \geq 2$, attach $[C_{i-1}, C_i)$ at a random point chosen uniformly over the existing tree.

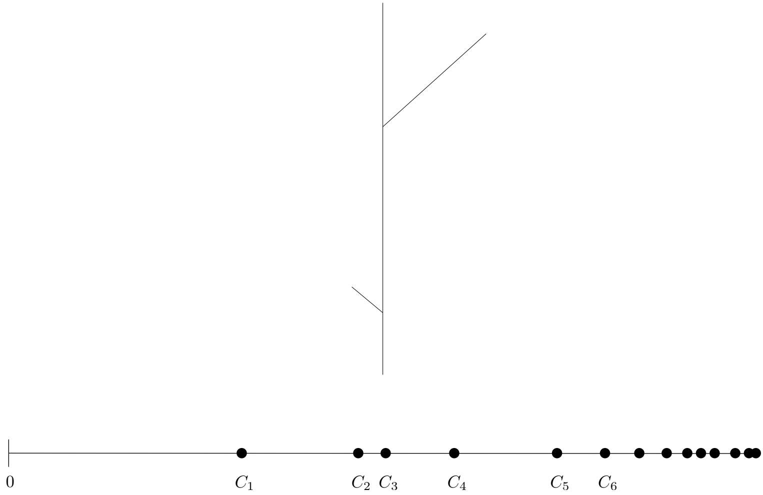
Construction 1: line-breaking



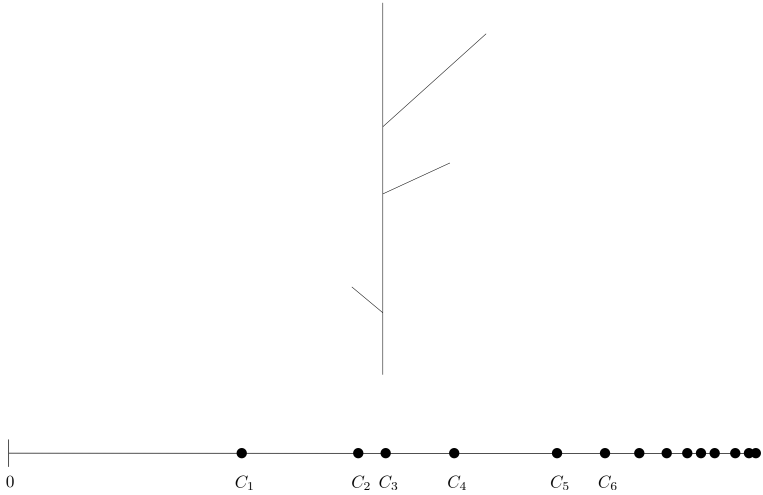
Construction 1: line-breaking



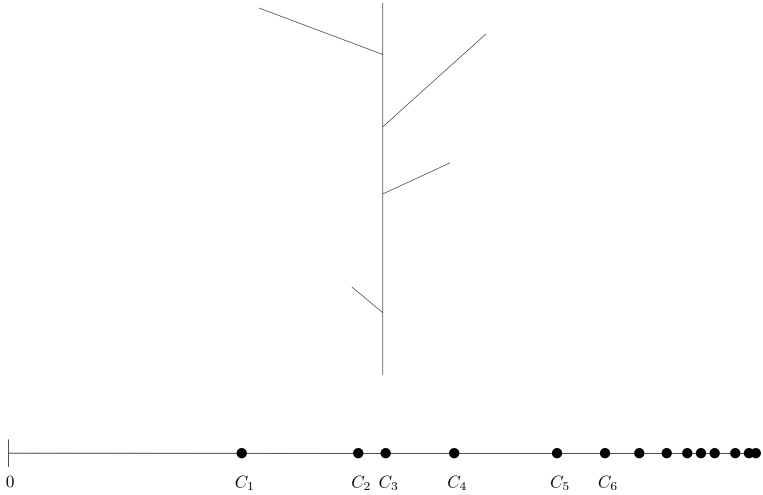
Construction 1: line-breaking



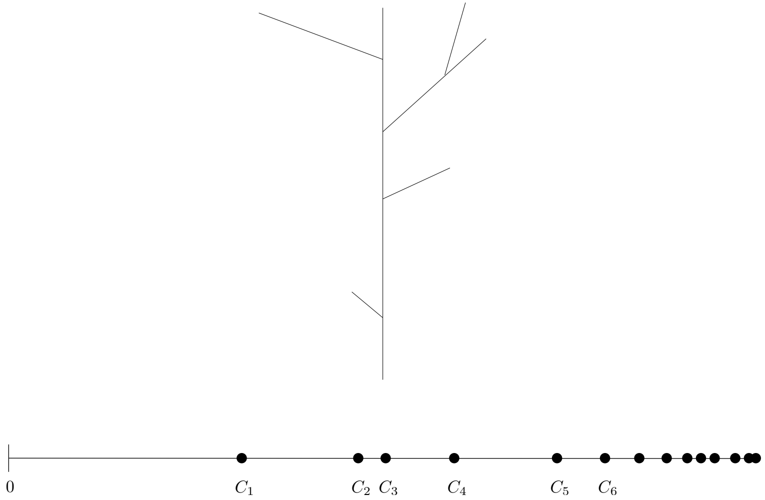
Construction 1: line-breaking



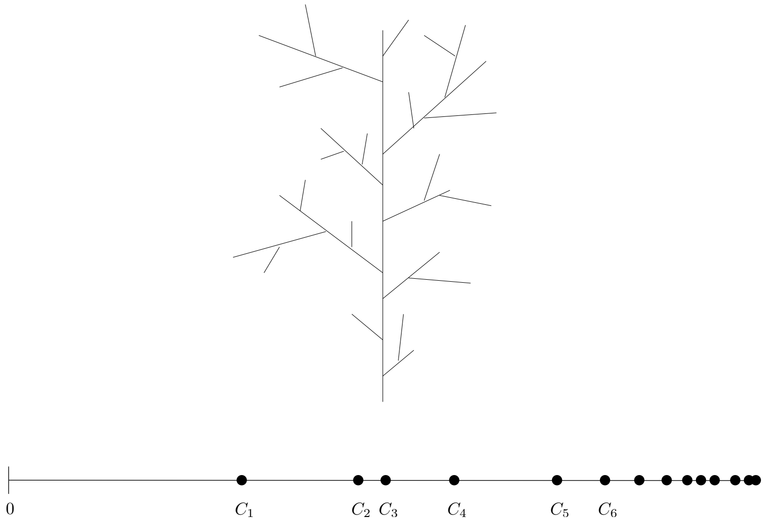
Construction 1: line-breaking



Construction 1: line-breaking



Construction 1: line-breaking



Construction 1: line-breaking

Start from $[0, C_1)$ and proceed inductively.

For $i \geq 2$, attach $[C_{i-1}, C_i)$ at a random point chosen uniformly over the existing tree.

Take the union of all the branches, thought of as a metric space, and then take its completion.

Construction 1: line-breaking

Start from $[0, C_1)$ and proceed inductively.

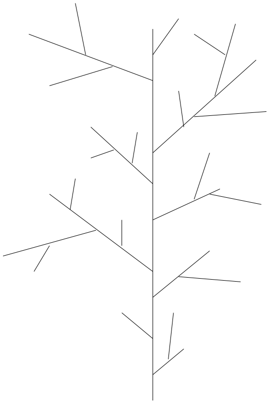
For $i \geq 2$, attach $[C_{i-1}, C_i)$ at a random point chosen uniformly over the existing tree.

Take the union of all the branches, thought of as a metric space, and then take its completion.

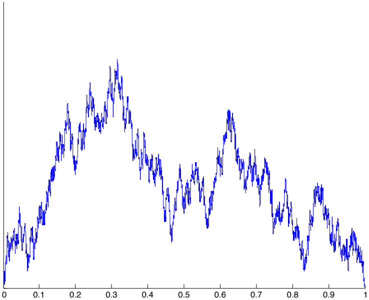
(Technical point: it will be useful later on to think of branches as having two sides, and we glue each new branch with probability $1/2$ on the left side and with probability $1/2$ on the right side. This endows the branches with a planar order.)

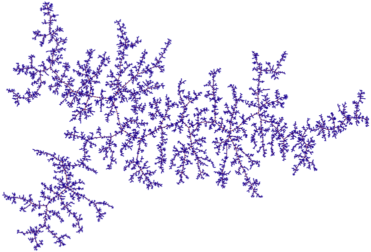
Construction 1: line-breaking

Write $\hat{\mu}_m$ for the empirical measure on the leaves after m steps. It turns out that this converges as $m \to \infty$ to a limiting probability measure μ.



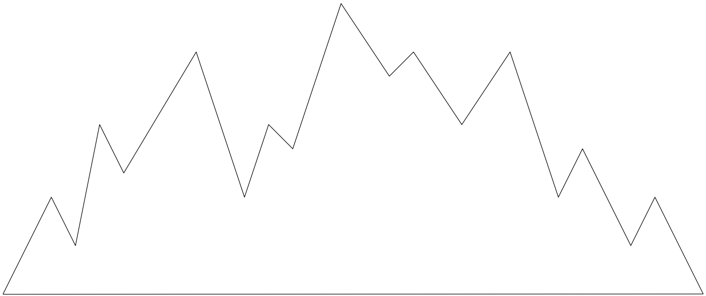
Construction 2: from a Brownian excursion



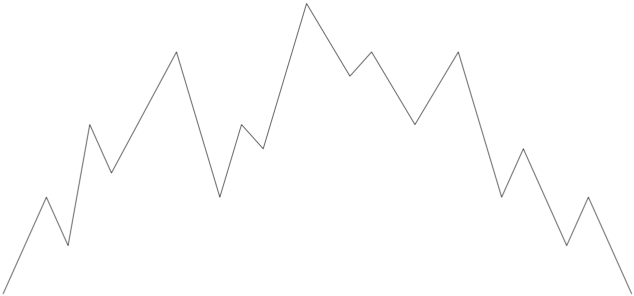


Picture by Igor Kortchemski

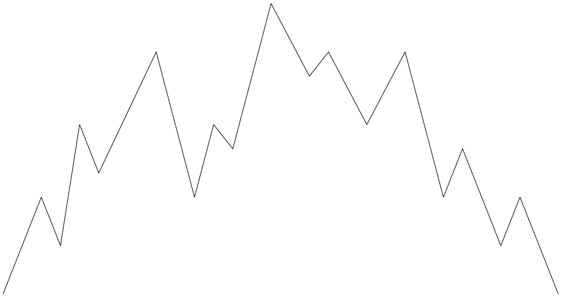
Construction 2: from a Brownian excursion



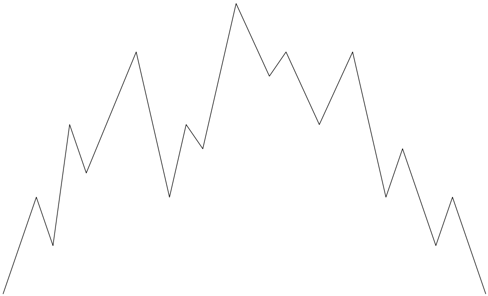
Construction 2: from a Brownian excursion



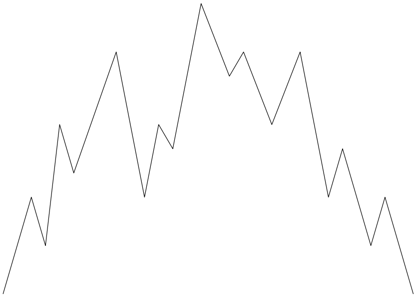
Construction 2: from a Brownian excursion



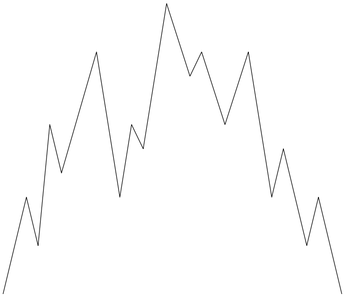
Construction 2: from a Brownian excursion



Construction 2: from a Brownian excursion



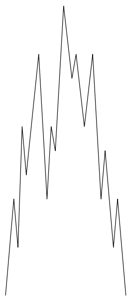
Construction 2: from a Brownian excursion



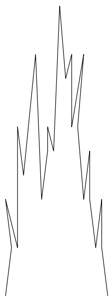
Construction 2: from a Brownian excursion

Construction 2: from a Brownian excursion

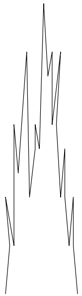
Construction 2: from a Brownian excursion



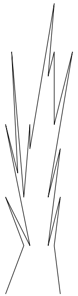
Construction 2: from a Brownian excursion



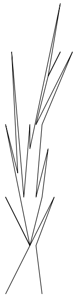
Construction 2: from a Brownian excursion



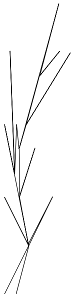
Construction 2: from a Brownian excursion



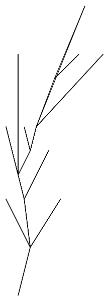
Construction 2: from a Brownian excursion



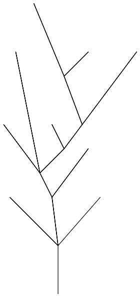
Construction 2: from a Brownian excursion



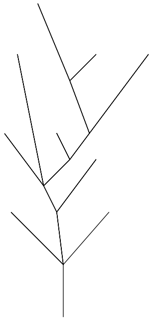
Construction 2: from a Brownian excursion



Construction 2: from a Brownian excursion

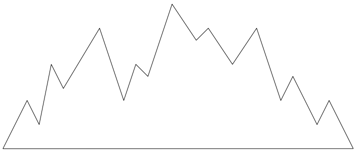


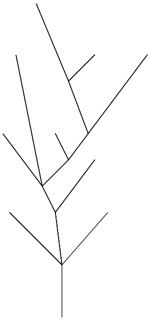
Construction 2: from a Brownian excursion



(Interpret distances vertically)

Construction 2: from a Brownian excursion





Local minima correspond to branch-points in the tree. These are a.s. unique, so the tree is binary.

μ is the push-forward of the Lebesgue measure on $[0,1]$ onto $(\mathcal{T}, d)$. It turns out that if $\mathcal{L}$ is the set of leaves of $\mathcal{T}$ then $\mu(\mathcal{L}) = 1$. The root is a uniform sample from μ.

Voronoi mass-partition in the Brownian CRT

Theorem. (Addario-Berry, Angel, Chapuy, Fusy & G, 2018)
Let $(\mathcal{T}, d, \mu)$ be the Brownian CRT. Fix $k \geq 2$ and let $X_1, X_2, \ldots, X_k$ be i.i.d. samples from μ. Let $V_1, V_2, \ldots, V_k$ be the corresponding Voronoi cells. Then

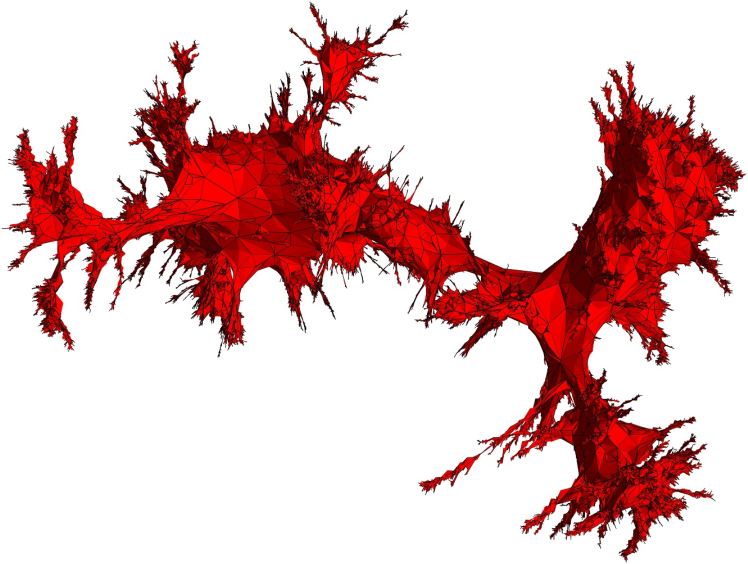
$$(\mu(V_1), \mu(V_2), \ldots, \mu(V_k)) \sim \mathrm{Dir}(1, 1, \ldots, 1).$$

Our original motivation

Conjecture. (Chapuy, 2016)
Let $(\mathcal{B}, d, \mu)$ be the Brownian map (or Brownian surface of genus $g \geq 0$). Let $X_1, X_2, \ldots, X_k$ be i.i.d. points sampled from μ and $V_1, V_2, \ldots, V_k$ be the corresponding Voronoi cells. Then

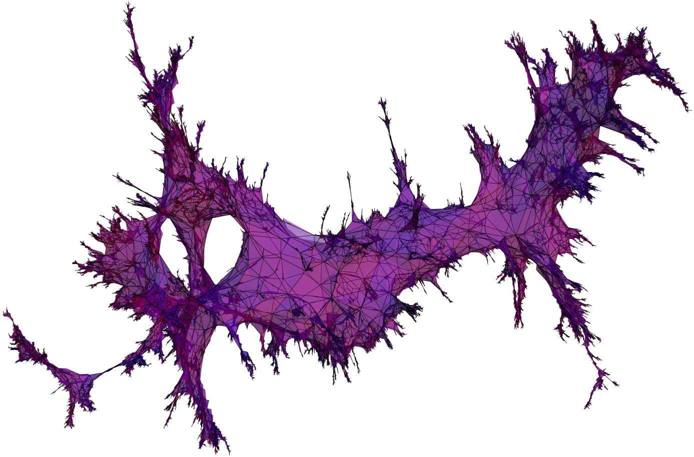
$$(\mu(V_1), \mu(V_2), \ldots, \mu(V_k)) \sim \mathrm{Dir}(1, 1, \ldots).$$

The Brownian map (sphere)



Picture by Jérémie Bettinelli

The Brownian double torus



Picture by Jérémie Bettinelli

Brownian surfaces

Conjecture. (Chapuy, 2016)
Let $(\mathcal{B}, d, \mu)$ be the Brownian map (or Brownian surface of genus $g \geq 0$). Let $X_1, X_2, \ldots, X_k$ be i.i.d. points sampled from μ and $V_1, V_2, \ldots, V_k$ be the corresponding Voronoi cells. Then

$$(\mu(V_1), \mu(V_2), \ldots, \mu(V_k)) \sim \mathrm{Dir}(1, 1, \ldots).$$

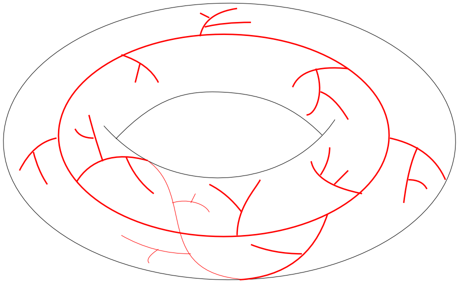
Proved for $g = 0, k = 2$ by Emmanuel Guitter (but proof does not generalise).

Unicellular random maps

Let $\mathcal{S}$ be an arbitrary compact surface without boundary. Let M_n be a uniform random map drawn on $\mathcal{S}$ with n vertices and a single face. (M_n is unicellular.) If $\mathcal{S}$ is the sphere then M_n is a uniform random plane tree.

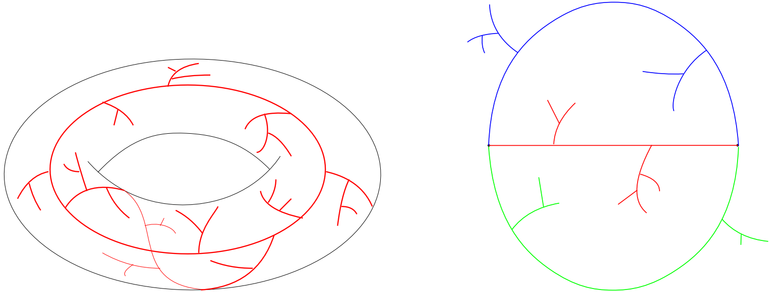
Then there is a scaling limit: as $n \to \infty$,

$$\left(M_n, \frac{1}{\sqrt{n}} d_n, \mu_n\right) \xrightarrow{d} (\mathcal{M}, d, \mu).$$



Unicellular random maps

In the case of the torus, as a graph we have

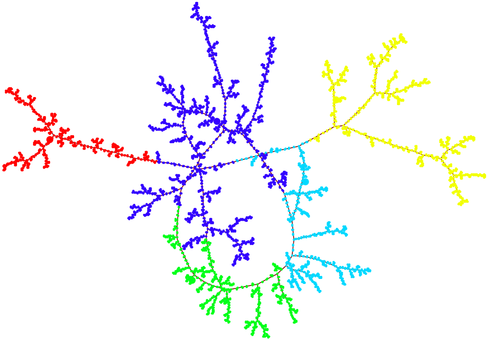


From Addario-Berry, Broutin & G. (2010, 2012), we may deduce that the scaling limit of a graph conditioned to have a theta-shaped kernel may be constructed out of three independent randomly rescaled Brownian CRT's.

A generalisation of our result to unicellular random maps

Theorem. (Addario-Berry, Angel, Chapuy, Fusy & G, 2018+)
For any compact surface $\mathcal{S}$ without boundary, $(\mathcal{M}, d, \mu)$ has uniform Voronoi mass-partitions.

$k = 5$, genus 2:



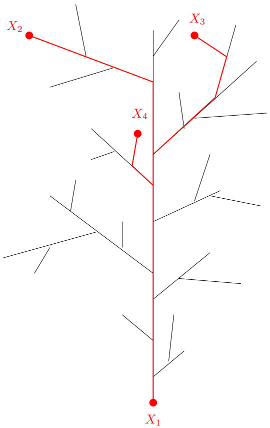
Picture by Igor Kortchemski

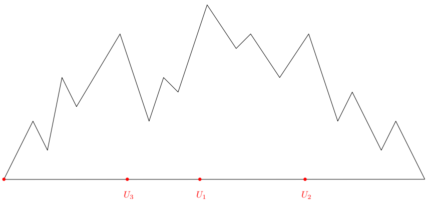
Open problem. Which properties of a random metric space give rise to uniform Voronoi mass-partitions?

Part III: proof of Brownian CRT case

Useful properties: sampling uniform points

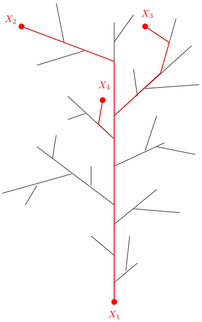
Sampling $X_1, \ldots, X_k$ i.i.d. points from μ is easy: use the excursion construction, and take the push-forwards of 0 and $U_1, U_2, \ldots, U_{k-1} \overset{\text{i.i.d.}}{\sim} \mathrm{U}[0,1]$.

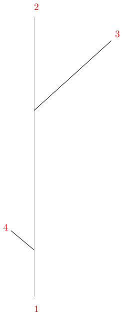




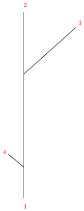
Useful properties: sampling uniform points

Non-trivial fact: the subtree spanned by $X_1, X_2, \ldots, X_k$ has the same distribution as the tree obtained after $k-1$ steps of the line-breaking construction.





Useful properties: sampling uniform points



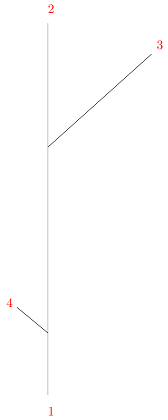
Non-trivial fact: the subtree spanned by $X_1, X_2, \ldots, X_k$ has the same distribution as the tree obtained after $k-1$ steps of the line-breaking construction.

This subtree is a uniform binary leaf-labelled plane tree whose $2k-3$ edge-lengths are exchangeable with

$$(L_1, L_2, \ldots, L_{2k-3}) \sim \sqrt{\Gamma_k} \times \mathrm{Dir}(1, 1, \ldots, 1),$$

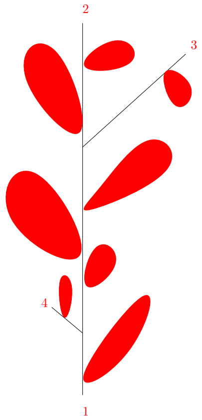
where the two factors on the right-hand side are independent and $\Gamma_k \sim \mathrm{Gamma}(k-1, 1/2)$.

Useful properties: reconstructing the whole tree



Suppose we start from the subtree spanned by $X_1, \ldots, X_k$.

Useful properties: reconstructing the whole tree



Suppose we start from the subtree spanned by $X_1, \ldots, X_k$. In order to get back to the whole tree, we need to take i.i.d. copies of the Brownian CRT, randomly rescaled by an exchangeable vector with sum 1, and glued onto the subtree at i.i.d. uniform positions.

Useful properties: the Dirichlet distribution

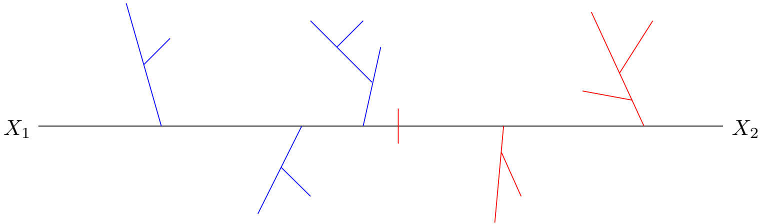
Let $E_1, E_2, \ldots, E_m$ be i.i.d. Exp(1). Then

$$\frac{1}{\sum_{i=1}^m E_i}(E_1, E_2, \ldots, E_m) \sim \text{Dir}(1, 1, \ldots, 1),$$

and is independent of $\sum_{i=1}^m E_i$.

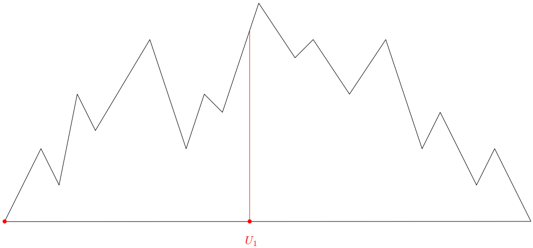
Base case: $k = 2$

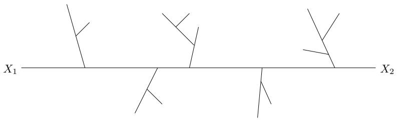
The proof goes via induction, with the base case being $k = 2$.



We wish to find the masses of the blue and red parts.

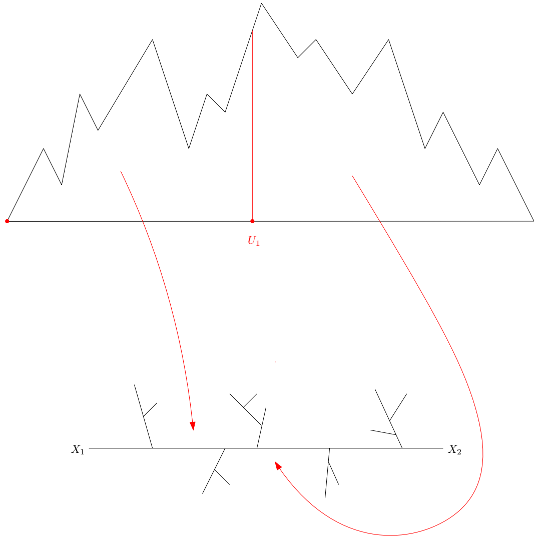
$k = 2$: an observation





Call the masses above and below the backbone the contour cells.

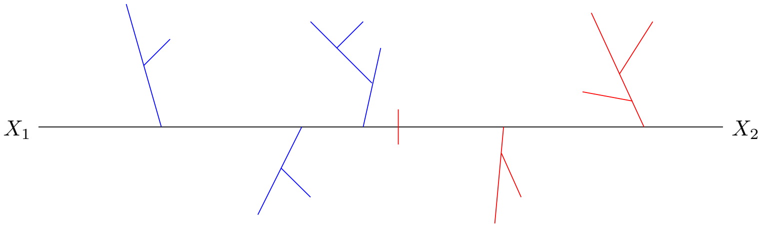
$k = 2$: an observation



Call the masses above and below the backbone the contour cells. These are equal to U_1 and $1 - U_1$, with $U_1 \sim \mathrm{U}[0, 1]$. The little trees attached to the backbone have exchangeable masses.

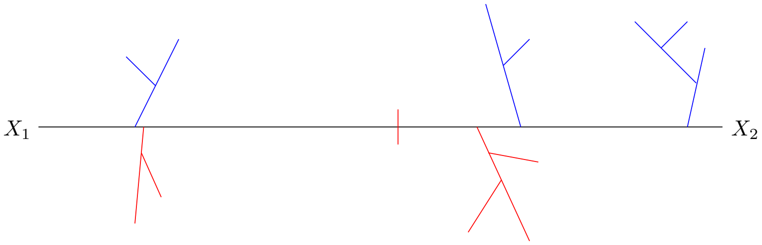
$k = 2$: a bijection

We may convert the Voronoi cells into the contour cells of a different tree:



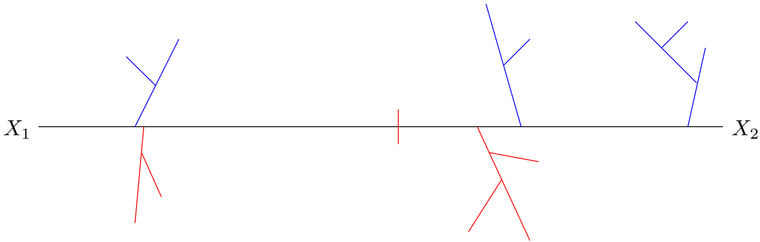
$k = 2$: a bijection

We may convert the Voronoi cells into the contour cells of a different tree:



$k = 2$: a bijection

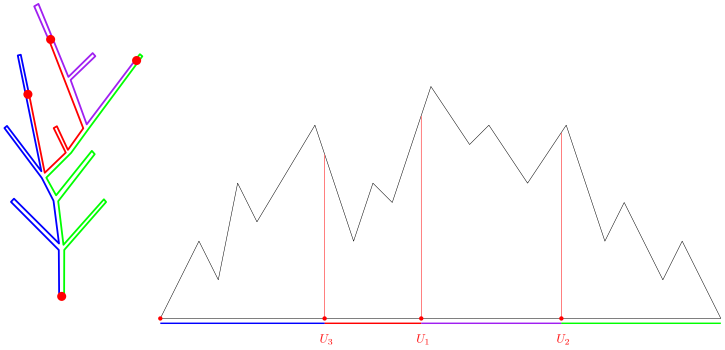
We may convert the Voronoi cells into the contour cells of a different tree:



Since the subtree masses are exchangeable, the new tree is again a Brownian CRT. But the contour cells in a Brownian CRT have $(U, 1-U)$ mass split, so the same must be true for the Voronoi cells. (This may be read off from results of Lévy (1939) or Bertoin and Pitman (1994).)

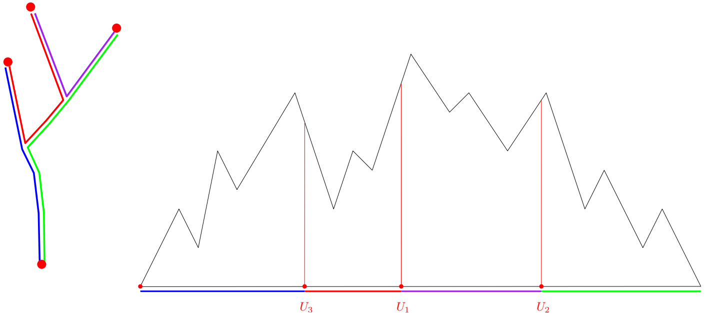
Reductions for $k \geq 3$: contour cells

Consider the subtree spanned by our uniform points.



Reductions for $k \geq 3$: contour cells

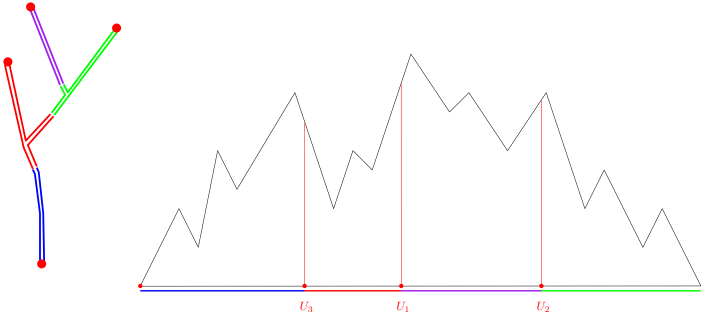
Consider the subtree spanned by our uniform points.



We will show that the lengths of the coloured intervals (the contour intervals) have the same joint law as the lengths of the Voronoi cells in the subtree.

Reductions for $k \geq 3$: contour cells

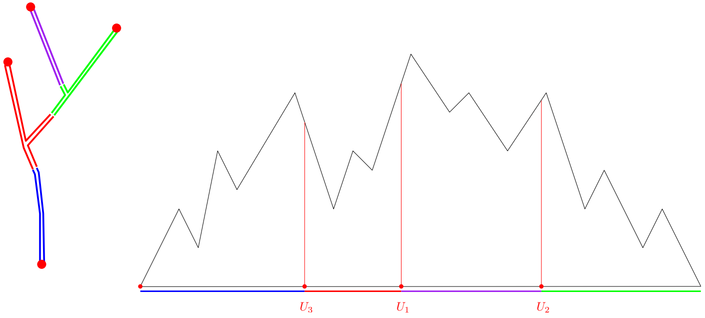
Consider the subtree spanned by our uniform points.



We will show that the lengths of the coloured intervals (the contour intervals) have the same joint law as the lengths of the Voronoi cells in the subtree.

Reductions for $k \geq 3$: contour cells

Consider the subtree spanned by our uniform points.



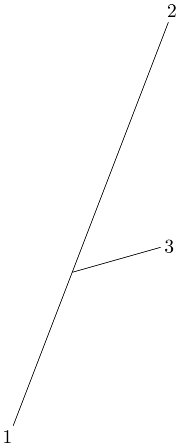
We will show that the lengths of the coloured intervals (the contour intervals) have the same joint law as the lengths of the Voronoi cells in the subtree. Since the mass attached to the contour intervals yields a uniform split of unity, the same must then be true for the Voronoi cells.

Reductions for $k \geq 3$: scaling

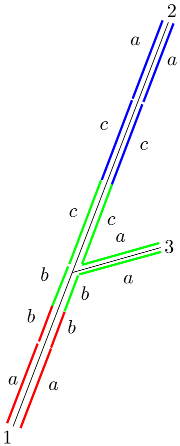
Since we're now only interested in showing that two vectors of lengths have the same distribution, it makes no difference if we rescale the whole tree.

So by the properties of the Brownian CRT, we may take the edge-lengths in the subtree spanned by our uniform points to be i.i.d. Exp(1).

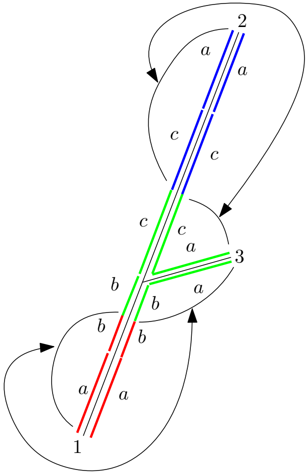
$k = 3$: contour lengths $\leftrightarrow$ Voronoi lengths



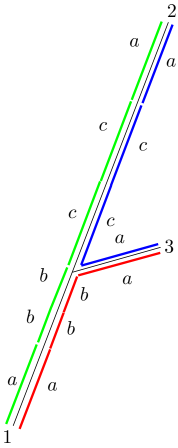
$k = 3$: contour lengths ↔ Voronoi lengths



$k = 3$: contour lengths $\leftrightarrow$ Voronoi lengths



$k = 3$: contour lengths $\leftrightarrow$ Voronoi lengths

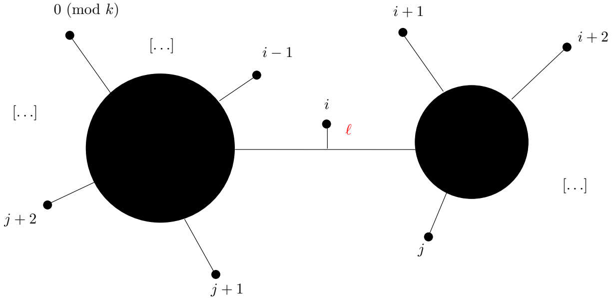


So again we have a bijection between the contour lengths and the Voronoi lengths.

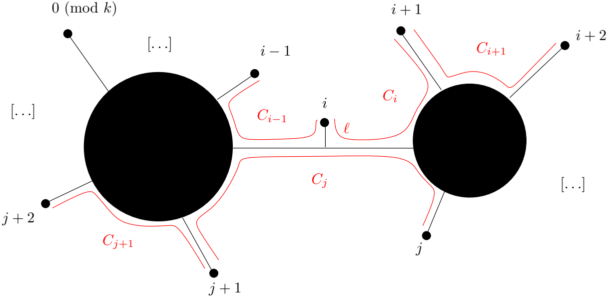
General $k \geq 3$: by induction

Suppose the result is true for all smaller k. We start with a uniform binary plane leaf-labelled tree with i.i.d. Exp(1) edge-lengths.

Start from the shortest branch incident to a leaf. This branch is uniform among all those incident to leaves. Call its leaf i and its length ℓ. Call the "opposite leaf" j.



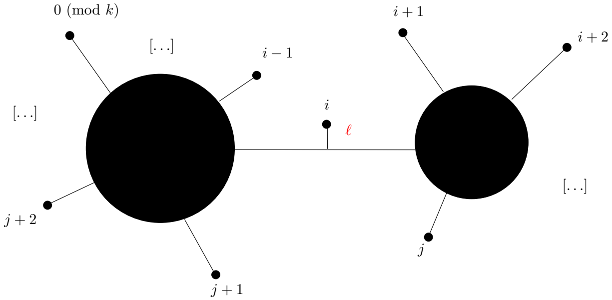
General $k \geq 3$: by induction



Voronoi lengths: $(L_0, L_1, \ldots, L_{k-1})$
Contour lengths: $(C_0, C_1, \ldots, C_{k-1})$.

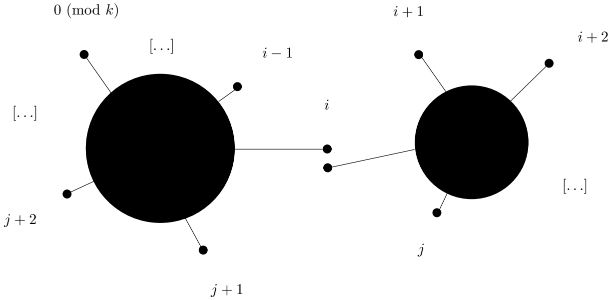
General $k \geq 3$: by induction

Now burn in from every leaf to remove length ℓ:



General $k \geq 3$: by induction

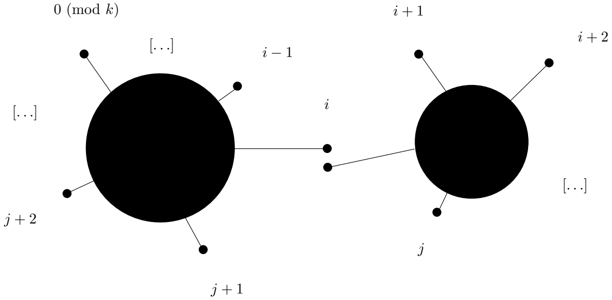
Now burn in from every leaf to remove length ℓ:



By the memoryless property of the exponential, and the uniformity of the shortest leaf, we split into two uniform binary leaf-labelled trees with i.i.d. exponential edge-lengths, each with $< k$ leaves.

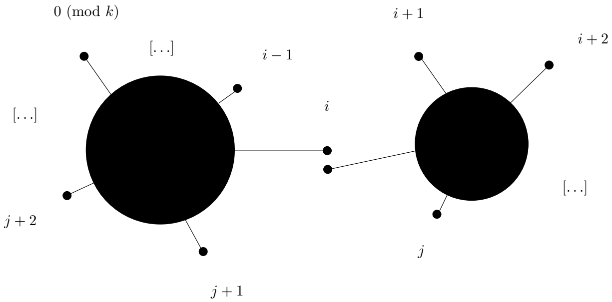
So, by the induction hypothesis, the Voronoi and contour lengths have the same laws in each of the subtrees.

General $k \geq 3$: by induction



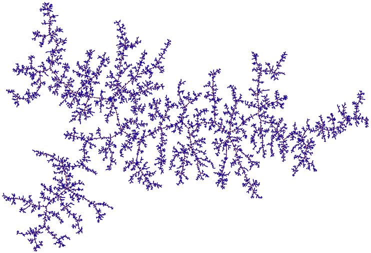
- For each leaf other than j, we can get back the original contour length C_r from r to $r+1$ by simply adding 2ℓ to the contours in the smaller problems.
- For the contour from j to $j+1$, we must add two contours together and add 2ℓ.

General $k \geq 3$: by induction



- For the Voronoi cells, add 2ℓ to the new lengths of the cells to get L_r, $r \neq i$.
- For the cell of i, add two Voronoi cells from the smaller trees, plus 2ℓ.

By induction, the vectors of lengths therefore have the same law. $\square$



Thank you!

Voronoi tessellations in the CRT and continuum random maps of finite excess, *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2018)*, pp.933-946.