THE Continuum Random Tree

Emmanuel Schertzer & Anita Winter

Genealogies of interacting particle systems

University Singapore, July 29$^{\text{th}}$ 2017

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

PROBABILISTIC STRUCTURES IN EVOLUTION
DFG SPP 1590

Outline

This learning session has three parts.

- **Part I: Definitions and representations of the CRT**
  - a random distance matrix
    - $\rightsquigarrow$ stick-breaking
  - tree below a random excursion

Outline

This learning session has three parts.

- **Part I: Definitions and representations of the CRT**
  - a random distance matrix
    - $\rightsquigarrow$ stick-breaking
  - tree below a random excursion
- **Part II: CRT as limit of finite tree models**
  - (sub-)critical, finite variance GW trees
  - giant component of the Erdös-Reny random graph
    - $\rightsquigarrow$ the multiplicative coalescent

Outline

This learning session has three parts.

- **Part I: Definitions and representations of the CRT**
  - a random distance matrix

    $\rightsquigarrow$ stick-breaking
  - tree below a random excursion
- **Part II: CRT as limit of finite tree models**
  - (sub-)critical, finite variance GW trees
  - giant component of the Erdös-Reny random graph

    $\rightsquigarrow$ the multiplicative coalescent
- **Part III: CRT is invariant under certain tree operation**
  - cutting down trees
  - the cut tree = the genealogy of cutting

    $\rightsquigarrow$ the additive coalescent

Part I

Outline: Part I
Definitions and representations of the CRT [3]

1. Aldous's Continuum Random Tree (CRT) is a random variable in the space of **continuum metric measure trees**.

Outline: Part I
Definitions and representations of the CRT [3]

1. Aldous's Continuum Random Tree (CRT) is a random variable in the space of **continuum metric measure trees**.
2. The distribution of such random variables is determined by the evaluation of all **polynomials**. Originally, the **CRT is defined through the densities** of all polynomials.

Outline: Part I
Definitions and representations of the CRT [3]

1. Aldous's Continuum Random Tree (CRT) is a random variable in the space of **continuum metric measure trees**.
2. The distribution of such random variables is determined by the evaluation of all **polynomials**. Originally, the **CRT is defined through the densities** of all polynomials.
3. Such a random variable was constructed via **stick-breaking**. This construction yields a representation of the **tree-lengths vector**.

Outline: Part I
Definitions and representations of the CRT [3]

1. Aldous's Continuum Random Tree (CRT) is a random variable in the space of **continuum metric measure trees**.
2. The distribution of such random variables is determined by the evaluation of all **polynomials**. Originally, the **CRT is defined through the densities** of all polynomials.
3. Such a random variable was constructed via **stick-breaking**. This construction yields a representation of the **tree-lengths vector**.
4. Aldous' celebrated result shows that **GW trees** with critical offspring distribution of finite variance **converge to the CRT**.

Outline: Part I
Definitions and representations of the CRT [3]

1. Aldous's Continuum Random Tree (CRT) is a random variable in the space of **continuum metric measure trees**.
2. The distribution of such random variables is determined by the evaluation of all **polynomials**. Originally, the **CRT is defined through the densities** of all polynomials.
3. Such a random variable was constructed via **stick-breaking**. This construction yields a representation of the **tree-lengths vector**.
4. Aldous' celebrated result shows that **GW trees** with critical offspring distribution of finite variance **converge to the CRT**.
5. A source of examples for continuum metric measure trees come from encoding them by **excursions**.

Outline: Part I
Definitions and representations of the CRT [3]

1. Aldous's Continuum Random Tree (CRT) is a random variable in the space of **continuum metric measure trees**.
2. The distribution of such random variables is determined by the evaluation of all **polynomials**. Originally, the **CRT is defined through the densities** of all polynomials.
3. Such a random variable was constructed via **stick-breaking**. This construction yields a representation of the **tree-lengths vector**.
4. Aldous' celebrated result shows that **GW trees** with critical offspring distribution of finite variance **converge to the CRT**.
5. A source of examples for continuum metric measure trees come from encoding them by **excursions**.
6. In Part 2 Emmanuel will show that the CRT is the tree associated with twice the **standard Brownian excursion**.

Outline: Part I
Definitions and representations of the CRT [3]

1. Aldous's Continuum Random Tree (CRT) is a random variable in the space of **continuum metric measure trees**.
2. The distribution of such random variables is determined by the evaluation of all **polynomials**. Originally, the **CRT is defined through the densities** of all polynomials.
3. Such a random variable was constructed via **stick-breaking**. This construction yields a representation of the **tree-lengths vector**.
4. Aldous' celebrated result shows that **GW trees** with critical offspring distribution of finite variance **converge to the CRT**.
5. A source of examples for continuum metric measure trees come from encoding them by **excursions**.
6. In Part 2 Emmanuel will show that the CRT is the tree associated with twice the **standard Brownian excursion**.

What are continuum trees? $\mathbb{R}$-trees

$\leadsto$ THE continuum random tree = random metric measure tree

Definition ([Tits '77], [Dress: *T-theory*], [Chiswell '01])

A metric space (T, d) is an **$\mathbb{R}$-tree** iff

- (T, d) is **connected**.
- (T, d) satisfies the **4-point condition**, i.e., $\forall\, x_1, x_2, x_3, x_4 \in T$,

$$d(x_1, x_2) + d(x_3, x_4) \leq \max\left\{d(x_1, x_3) + d(x_2, x_4), d(x_1, x_4) + d(x_2, x_3)\right\}.$$

$\rightsquigarrow$ THE continuum random tree = random metric measure tree

Definition ([Tits '77], [Dress: *T-theory*], [Chiswell '01])

A metric space (T, d) is an **$\mathbb{R}$-tree** iff

- (T, d) is **connected**.
- (T, d) satisfies the **4-point condition**, i.e., $\forall\, x_1, x_2, x_3, x_4 \in T$,

  $$d(x_1, x_2) + d(x_3, x_4) \leq \max\big\{d(x_1, x_3) + d(x_2, x_4), d(x_1, x_4) + d(x_2, x_3)\big\}.$$

$\rightsquigarrow$ $\mathbb{R}$-trees are understood as **continuum trees** and have the following *intrinsic property*

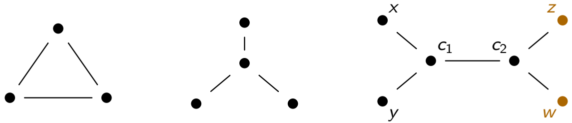
For $x_1, x_2, x_3 \in T$ there **exists a unique branch point** $c(x_1, x_2, x_3) \in T$ with

$$\big[x_1, x_2\big] \cap \big[x_2, x_3\big] \cap \big[x_1, x_3\big] = \big\{c(x_1, x_2, x_3)\big\},$$

where

$$[x, y] := \big\{z \in T :\ r(x, z) + r(z, y) = r(x, y)\big\}.$$

$\rightsquigarrow$ *drop connectedness*

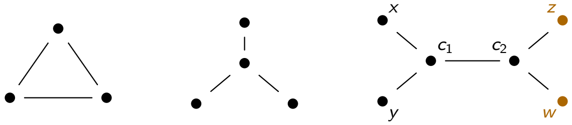


I am not a tree :-(I am a tree :-)

Definition ([Athreya, Löhr, W. (2016)])

A metric space (T, d) is a **metric tree** if it is (isometric to) a subset of an $\mathbb{R}$-tree with $c(x, y, z) \in T$ for all $x, y, z \in T$.

$\rightsquigarrow$ *drop connectedness*



I am not a tree :-(I am a tree :-)

Definition ([Athreya, Löhr, W. (2016)])

A metric space (T, d) is a **metric tree** if it is (isometric to) a subset of an $\mathbb{R}$-tree with $c(x, y, z) \in T$ for all $x, y, z \in T$.

Examples.

- Continuum trees/$\mathbb{R}$-trees are metric trees.
- Graph-theoretical trees are (discrete) metric trees.

The lengths measure

$\rightsquigarrow$ In this lecture we consider **compact** metric spaces only.

Compact metric spaces are separable.

Let (T, r) be a metric tree, $\rho \in T$ a distinguished point, and $T' \subseteq T$ countably dense.

- Denote by $T^o := \bigcup_{x \in T'} (\rho, x)$ the *skeleton* of T, and by **iso**(T) the set of *isolated leaves* of T.
- There is a unique σ-finite measure $\ell^{(T,r,\rho)}$ on $\mathrm{iso}(T) \uplus T^o$ with $\ell^{(T,r,\rho)}\big(T \setminus (\mathrm{iso}(T) \uplus T^o)\big) = 0$ and

$$\ell^{(T,r,\rho)}\big((\rho, x]\big) := r(\rho, x), \quad x \in T'.$$

$\rightsquigarrow$ On continuum trees the length measure does not depend on the choice of the root, and be considered as a generalization of the *Lebesgue measure*.

Metric measure spaces and Gromov-weak topology

In this lecture we consider **compact** metric spaces only.

- (X, r, μ) is a *metric measure space* if (X, r) is a compact metric space and μ a probability measure on $\mathcal{B}(X)$ with $\mathrm{supp}(\mu) = X$.
- We call (T, r, μ) a **measure $\mathbb{R}$-tree** or a **metric measure tree** iff (T, r) is a $\mathbb{R}$-tree resp. a metric tree.
- We call (X, r, μ) and (X', r', μ') *equivalent* iff there exists a measure preserving isometry $\phi : X \to X'$.
- Denote by

$$\mathbb{X},\ \bar{\mathbb{T}} \text{ and } \mathbb{T}$$

the spaces of equivalence classes of metric measure spaces, measure $\mathbb{R}$-trees resp. metric measure trees.

Convergence in the space of metric measure spaces

Gromov-weak convergence

-

Convergence in the space of metric measure spaces

Gromov-weak convergence

- Let $(X, r, \mu) \in \mathbb{X}$ (μ to *sample* a random *finite subspace*) and $X_1, X_2, \ldots \in X$ independent, μ-distributed random variables.
- We require for all $m \in \mathbb{N}$ *convergence in distribution* of the random (pseudo-)metric $r(i,j) := d(X_i, X_j)$ on $\{1, \ldots, m\}$.

Gromov-weak convergence

- Let $(X, r, \mu) \in \mathbb{X}$ (μ to *sample* a random *finite subspace*) and $X_1, X_2, \ldots \in X$ independent, μ-distributed random variables.
- We require for all $m \in \mathbb{N}$ *convergence in distribution* of the random (pseudo-)metric $r(i,j) := d(X_i, X_j)$ on $\{1, \ldots, m\}$.
- Different formulation: Require convergence of **polynomials** :

$$\Phi(\mathcal{X}) := \int_{T^m} \phi\big(r(x_i, x_j)_{1 \le i,j \le m}\big)\mu^{\otimes m}(\mathrm{d}\underline{u}), \quad (\phi \in \mathcal{C}_b(\mathbb{R}_+^{m \times m}))$$

Gromov-weak convergence

- Let $(X, r, \mu) \in \mathbb{X}$ (μ to *sample* a random *finite subspace*) and $X_1, X_2, \ldots \in X$ independent, μ-distributed random variables.
- We require for all $m \in \mathbb{N}$ *convergence in distribution* of the random (pseudo-)metric $r(i,j) := d(X_i, X_j)$ on $\{1, \ldots, m\}$.
- Different formulation: Require convergence of **polynomials** :

$$\Phi(\mathcal{X}) := \int_{T^m} \phi\big(r(x_i, x_j)_{1 \le i,j \le m}\big) \mu^{\otimes m}(\mathrm{d}\underline{u}), \quad (\phi \in \mathcal{C}_b(\mathbb{R}_+^{m \times m}))$$

- (*Vershik's reconstruction theorem*) If $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{X}$ such that

$$\Phi(\mathcal{X}_1) = \Phi(\mathcal{X}_2) \quad \text{for all polynomials} \;\Leftrightarrow\; \mathcal{X}_1 = \mathcal{X}_2.$$

Convergence in the space of metric measure spaces

Gromov-weak convergence

- Let $(X, r, \mu) \in \mathbb{X}$ (μ to *sample* a random *finite subspace*) and $X_1, X_2, \ldots \in X$ independent, μ-distributed random variables.
- We require for all $m \in \mathbb{N}$ *convergence in distribution* of the random (pseudo-)metric $r(i,j) := d(X_i, X_j)$ on $\{1, \ldots, m\}$.
- Different formulation: Require convergence of **polynomials** :

$$\Phi(\mathcal{X}) := \int_{T^m} \phi\big(r(x_i, x_j)_{1 \le i,j \le m}\big)\mu^{\otimes m}(\mathrm{d}\underline{u}), \quad (\phi \in \mathcal{C}_b(\mathbb{R}_+^{m \times m}))$$

- (*Depperschmidt, Greven & Pfaffelhuber 2012, [2]*, *Löhr '13[4]*) If $(\mathcal{X}_n)_{n \in \mathbb{N}}$ and $\mathcal{X}$ are **random elements** in $\mathbb{X}$, then

$$\mathbb{E}\big[\Phi(\mathcal{X}_n)\big] \to \mathbb{E}\big[\Phi(\mathcal{X})\big] \quad \text{for all polynomials} \;\Leftrightarrow\; \mathcal{X}_n \Rightarrow \mathcal{X}.$$

THE Continuum Random Tree (CRT)

- For $m \geq 2$, we consider **binary trees** with m **leaves labelled** $\{1, 2, ..., m\}$ and positive edge lengths $\{l_e;\ e \text{ edges}\}$.
- Each such tree has $2m - 3$ edges. When edge lengths are ignored, there are $\prod_{i=1}^{m-2}(2i - 3)$ many possible shapes $\hat{t}$ for the tree.

Definition (CRT, [3])

The CRT is the random (equivalence class of the) continuum measure tree (T, r, μ) such that for each $m \in \mathbb{N}$ the distribution of the vector of the shape together with the tree lengths of the subtree spanned by a μ-sample of size m has density:

$$f\big(\mathrm{t}, l_1, ..., l_{2m-3}\big) = s \cdot \exp\big(-s^2/2\big)\mathrm{d}l_1 ... \mathrm{d}l_{2m-3},$$

where $s := \sum_{i=1}^{2m-3} l_i$.

CRT: A few remarks

$$\mathbb{P}\big(\text{shape}(\mathcal{R}(k)) = \hat{t}, L_1 \in \mathrm{d}l_1, ..., L_{2k-3} \in \mathrm{d}l_{2k-3}\big)$$
$$= s \cdot \exp\big(-s^2/2\big)\mathrm{d}l_1 ... \mathrm{d}l_{2k-3}, \quad s := \sum_{i=1}^{2k-3} l_i.$$

Not hard to show that this is a probability density function.

Remarks.

1. The shape is uniform on the set of possible shapes, the edge lengths are independent of the shape and edge lengths are exchangeable.
2. The above defines a distance matrix distribution. If $m = 2$, the subtree has 2 leaves, 1 possible shape, 1 edge, no internal node. The single edge's length is **Rayleigh distributed**, i.e.,

$$\mathbb{P}\big(L \in \mathrm{d}l\big) = l \cdot \exp\big(-l^2/2\big)\mathrm{d}l.$$

Note that if X *is mean* 1 *exponential*, then $\mathbb{P}(\sqrt{2X} \geq t) = e^{-t^2/2}$, i.e., $\sqrt{2X}$ is Rayleigh distributed.

CRT: The stick-breaking construction

1. Let $(C_1, C_2, C_3, \dots)$ be the times of a **non-homogeneous Poisson point process** with **rate** $r(t) = t$, i.e., for example,

$$\mathbb{P}\{C_1 > t\} = \mathbb{P}\{\text{no point in } [0,t]\} = e^{-\int_0^t \mathrm{d}s\, r(s)} = e^{-\frac{t^2}{2}}.$$

CRT: The stick-breaking construction

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a **non-homogeneous Poisson point process** with **rate** $r(t) = t$, i.e., for example,

$$\mathbb{P}\{C_1 > t\} = \mathbb{P}\{\text{no point in } [0, t]\} = e^{-\int_0^t \mathrm{d}s r(s)} = e^{-\frac{t^2}{2}}.$$

2. Let $\mathcal{R}(1)$ be a **line of length** C_1 from a root to leaf 1.

CRT: The stick-breaking construction

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a **non-homogeneous Poisson point process** with **rate** $r(t) = t$, i.e., for example,

$$\mathbb{P}\{C_1 > t\} = \mathbb{P}\{\text{no point in } [0,t]\} = e^{-\int_0^t \mathrm{d}s r(s)} = e^{-\frac{t^2}{2}}.$$

2. Let $\mathcal{R}(1)$ be a **line of length** C_1 from a root to leaf 1.
3. Inductively, **obtain** $\mathcal{R}(m+1)$ **from** $\mathcal{R}(m)$ **by attaching an edge of length** $C_{m+1} - C_m$ **to a uniform random point of** $\mathcal{R}(m)$ (i.e., sampled with respect to the normalized Lebesgue measure on the edges), labeling a new leaf $m+1$.

CRT: The stick-breaking construction

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a **non-homogeneous Poisson point process** with **rate** $r(t) = t$, i.e., for example,

$$\mathbb{P}\{C_1 > t\} = \mathbb{P}\{\text{no point in } [0, t]\} = e^{-\int_0^t \mathrm{d}s\, r(s)} = e^{-\frac{t^2}{2}}.$$

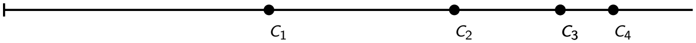
2. Let $\mathcal{R}(1)$ be a **line of length** C_1 from a root to leaf 1.
3. Inductively, **obtain** $\mathcal{R}(m+1)$ **from** $\mathcal{R}(m)$ **by attaching an edge of length** $C_{m+1} - C_m$ **to a uniform random point of** $\mathcal{R}(m)$ (i.e., sampled with respect to the normalized Lebesgue measure on the edges), labeling a new leaf $m + 1$.

Proposition

For each $m \in \mathbb{N}$, $\mathcal{R}(m)$ has the "CRT-density".

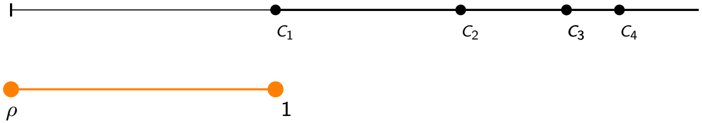
Stick-breaking: an illustration

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.



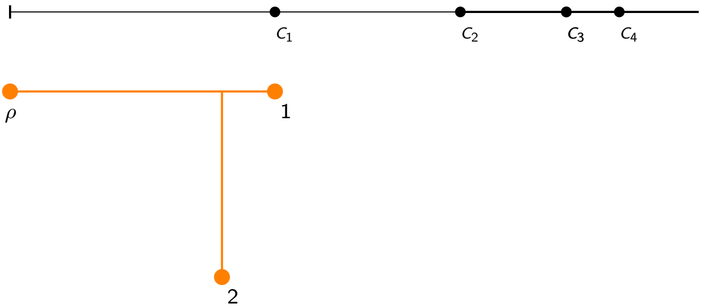
Stick-breaking: an illustration

1. Let $(C_1, C_2, C_3, ...)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.
2. Let $\mathcal{R}(1)$ be a line of length C_1 from a root to leaf 1.



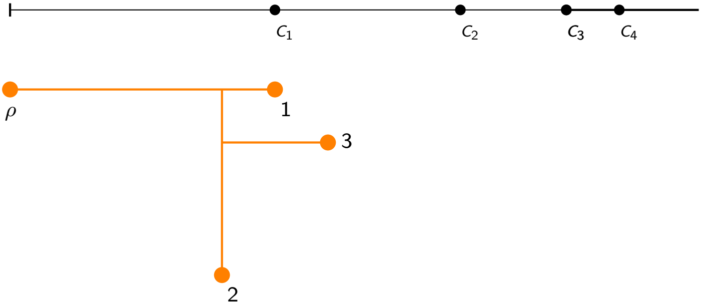
Stick-breaking: an illustration

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.
2. Let $\mathcal{R}(1)$ be a line of length C_1 from a root to leaf 1.
3. Inductively, obtain $\mathcal{R}(m+1)$ from $\mathcal{R}(m)$ by attaching an edge of length $C_{m+1} - C_m$ to a uniform random point of $\mathcal{R}(m)$ labeling a new leaf $m+1$.

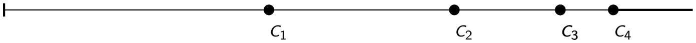


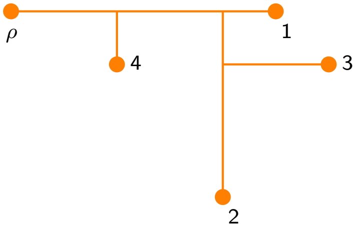
Stick-breaking: an illustration

1. Let $(C_1, C_2, C_3, ...)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.
2. Let $\mathcal{R}(1)$ be a line of length C_1 from a root to leaf 1.
3. Inductively, obtain $\mathcal{R}(m+1)$ from $\mathcal{R}(m)$ by attaching an edge of length $C_{m+1} - C_m$ to a uniform random point of $\mathcal{R}(m)$ labeling a new leaf $m+1$.



1. Let $(C_1, C_2, C_3, ...)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.
2. Let $\mathcal{R}(1)$ be a line of length C_1 from a root to leaf 1.
3. Inductively, obtain $\mathcal{R}(m+1)$ from $\mathcal{R}(m)$ by attaching an edge of length $C_{m+1} - C_m$ to a uniform random point of $\mathcal{R}(m)$ labeling a new leaf $m+1$.





CRT: analysing the stick-breaking density

- We have seen that the density of C_1 is the right *Rayleigh* distribution.

CRT: analysing the stick-breaking density

- We have seen that the density of C_1 is the right *Rayleigh* distribution.
- We proceed by induction. Let $(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*)$ be a binary tree with $k+1$ leaves, shape $\mathfrak{t}$ and $2k+1$ edge lengths x_1^*, ..., x_{2k+1}^*, and Let $(\mathfrak{t}, x_1, ..., x_{2k-1})$ be the associated binary tree spanned by the leaves $\{1, 2, ..., k\}$.

CRT: analysing the stick-breaking density

- We have seen that the density of C_1 is the right *Rayleigh* distribution.
- We proceed by induction. Let $(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*)$ be a binary tree with $k+1$ leaves, shape $\mathfrak{t}$ and $2k+1$ edge lengths x_1^*, ..., x_{2k+1}^*, and Let $(\mathfrak{t}, x_1, ..., x_{2k-1})$ be the associated binary tree spanned by the leaves $\{1, 2, ..., k\}$.
- By construction, $\mathfrak{t}^*$ is obtained from $\mathfrak{t}$ by splitting an edge x_j for some $j = 1, ..., 2k-1$ into two edges of lengths $x_{j_1}^*$ and $x_{j_2}^*$ with $x_j = x_{j_1}^* + x_{j_2}^*$, and joining leaf $k+1$ to that new internal vertex by an edge $x_{j_3}^* = s^* - s$, say.

CRT: analysing the stick-breaking density

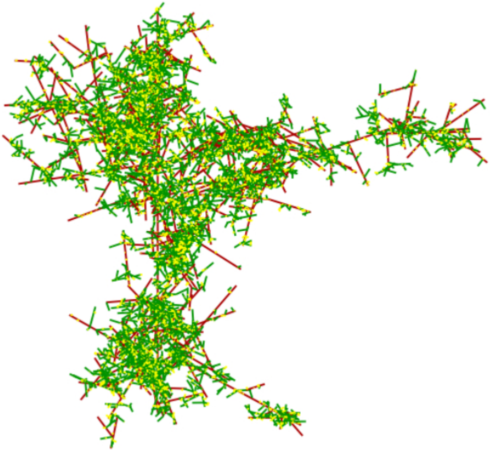
- We have seen that the density of C_1 is the right *Rayleigh* distribution.
- We proceed by induction. Let $(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*)$ be a binary tree with $k+1$ leaves, shape $\mathfrak{t}$ and $2k+1$ edge lengths x_1^*, ..., x_{2k+1}^*, and Let $(\mathfrak{t}, x_1, ..., x_{2k-1})$ be the associated binary tree spanned by the leaves $\{1, 2, ..., k\}$.
- By construction, $\mathfrak{t}^*$ is obtained from $\mathfrak{t}$ by splitting an edge x_j for some $j = 1, ..., 2k-1$ into two edges of lengths $x_{j_1}^*$ and $x_{j_2}^*$ with $x_j = x_{j_1}^* + x_{j_2}^*$, and joining leaf $k+1$ to that new internal vertex by an edge $x_{j_3}^* = s^* - s$, say.
- That is,

$$f\left(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*\right) = f\left(\mathfrak{t}, x_1, ..., x_{2k-1}\right) s^* \cdot e^{-\frac{1}{2}((s^*)^2 - s^2)} \cdot s^{-1}$$
$$= e^{-\frac{s^2}{2}} s^* \cdot e^{-\frac{1}{2}((s^*)^2 - s^2)} \cdot s^{-1} = s^* \cdot e^{-\frac{1}{2}(s^*)^2},$$

where s^{-1} is the probability density that the $(k+1)^{\text{st}}$ edge is attached at a particular place in the existing tree.

Simulations are often based on stick-breaking construction

⇝ Several simulations can be found on the home page of Grégory Miermont, e.g.,



The random tree-lengths vector

Let (T, r, μ) be the CRT, and X_1, X_2, ... independent and identically μ-distributed. Denote by Θ_n the random length of the subtree spanned by the first n-leaves.

- It follows from the stick-breaking construction that

$$\mathbb{P}\{\Theta_k > x\} = \mathbb{P}\{N(\frac{x^2}{2}) < k\},$$

  where $N(\lambda)$ denotes a *Poisson* variable with intensity λ.

- It follows that Θ_k has a **Chi distribution** with parameter $2k$, i.e., with density

$$f_{\Theta_k}(x) = \frac{2^{-(k-1)} x^{2k-1}}{(k-1)!} \exp\big(-x^2/2\big), \quad x > 0.$$

- Moreover, one can easily show that for all $n \in \mathbb{N}$,

$$(\Theta_1, \Theta_2, ..., \Theta_n) \stackrel{d}{=} \big(\sqrt{2X_1}, \sqrt{2(X_1 + X_2)}, ..., \sqrt{2(X_1 + ... + X_n)}\big),$$

  where X_1, X_2, ... are independent rate 1 exponentially distributed.

A problem concerning tree-lengths

Let (T, r, μ) be a *metric measure tree*, and X_1, X_2, ... independent and identically μ-distributed. Denote by Θ_n the random length of the subtree spanned by the first n-leaves. We refer to the random vector $(\Theta_1, \Theta_2, \ldots)$ as the **tree-length vector**.

- $\rightsquigarrow$ It is shown in Greven, Pfaffelhuber & Winter (2013) that in the space of *ultra-metric* measure trees, the distribution of the tree-length vector determines (T, r, μ) uniquely.

Open problem

Under which assumption is the distribution of the tree-length vector convergence determining?

The CRT as uniform $\mathbb{R}$-tree

The main result of [3] is the following **invariance principle**.

Theorem

Let $\mathcal{X}_N$ be the Galton-Watson tree conditioned on total population size N and with critical offspring distribution of finite variance $\sigma^2 > 0$. If $\hat{x}_N$ is $\mathcal{X}_N$ with edge lengths rescaled by $\frac{\sigma}{\sqrt{N}}$ and equipped with the uniform measure on the leaves, then

$$\hat{\mathcal{X}}_N \xrightarrow[N\to\infty]{\mathrm{w}} \mathrm{CRT}.$$

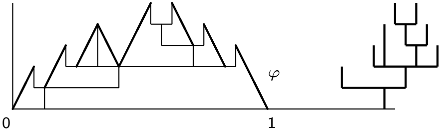
$\rightsquigarrow$ Such a invariance principle will be the link to an equivalent definition of the CRT.

Correspondence to excursions

- A (continuous) **excursion** is a function $\varphi \in C([0,1])$ with $\varphi\big|_{\{0,1\}} = 0$ and $\varphi\big|_{(0,1)} > 0$.
- With every excursion φ we associate a **pseudo-metric on** $[0,1]$**:**

$$r_\varphi(s,t) := \varphi(s) + \varphi(t) - 2 \cdot \inf_{u \in [s,t]} \varphi(u).$$

- Let μ be the image measure of the Lebesgue measure on $[0,1]$ under the map which sends a point of $[0,1]$ to a point in the tree.
Fact. $T\big|_\varphi = [0,1]_{/\sim_\varphi}, \mu$ is a rooted, measure $\mathbb{R}$-tree with **root** 0.



$\rightsquigarrow$ Convergence of excursions w.r.t. uniform convergence of continuous functions implies Gromov-weak convergence of the associated trees. Emmanuel will use this to argue in the second part that the CRT is the tree associated with $2\cdot$ **Brownian excursion**.

Proof of the invariance principle in a nutshell

- Let $\mathcal{T}$ be an unconditioned (ordered) Galton-Watson tree, and $\mathfrak{t}$ be a discrete tree with k leaves labelled $1, \ldots, k$ and $k-1$ unlabeled branch points. Then

$$\begin{aligned}&\mathbb{E}\big[\#\text{ subtrees of } \mathcal{T} \text{ with } k \text{ leaves isomorphic to } \mathfrak{t}\big]\mathbf{1}\{\mathcal{T}=n\}\\&=\big(\tfrac{\sigma^2}{2}\big)^{k-1}\mathbb{P}\big\{S_{L(t)+k}=n-\#\mathfrak{t}-L(t)\big\},\end{aligned}$$

where

$$L(t)=\sum_{i=1}^{\#\mathfrak{t}-(2k-1)}\hat{\xi}_i+\sum_{i=1}^{k-1}\tilde{\xi}_i,\quad S_n:=\sum_{i=1}^{m}X_i,$$

and $(\hat{\xi}_i,\tilde{\xi}_i,X_i)$ are i.i.d. with $\hat{\xi}$ having the "size"-biased offspring distribution $\big(\mathbb{P}(\hat{\xi}=i)=(i+1)\mathbb{P}(\xi=i)$; note that $\mathbb{E}[\hat{\xi}=\sigma^2]\big)$, $\tilde{\xi}$ having the "size-size"-biased $\big(\mathbb{P}(\tilde{\xi}=i)=\sigma^{-2}(i+2)(i+1)\mathbb{P}(\xi=i+2)\big)$ offspring distribution and X being distributed as $\#\mathcal{T}$.

Proof of the invariance principle in a nutshell

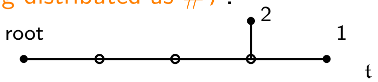
- Let $\mathcal{T}$ be an unconditioned (ordered) Galton-Watson tree, and $\mathfrak{t}$ be a discrete tree with k leaves labelled $1, \ldots, k$ and $k-1$ unlabeled branch points. Then

$$\mathbb{E}\big[\# \text{ subtrees of } \mathcal{T} \text{ with } k \text{ leaves isomorphic to } \mathfrak{t}\big]\mathbf{1}\{\mathcal{T}=n\}$$
$$= \big(\tfrac{\sigma^2}{2}\big)^{k-1}\mathbb{P}\big\{S_{L(t)+k} = n - \#\mathfrak{t} - L(t)\big\},$$

where

$$L(t) = \sum_{i=1}^{\#\mathfrak{t}-(2k-1)} \hat{\xi}_i + \sum_{i=1}^{k-1} \tilde{\xi}_i, \quad S_n := \sum_{i=1}^{m} X_i,$$

and $(\hat{\xi}_i, \tilde{\xi}_i, X_i)$ are i.i.d. with $\hat{\xi}$ having the "size"-biased offspring distribution ($\mathbb{P}(\hat{\xi}=i) = (i+1)\mathbb{P}(\xi=i)$; note that $\mathbb{E}[\hat{\xi}=\sigma^2]$), $\tilde{\xi}$ having the "size-size"-biased ($\mathbb{P}(\tilde{\xi}=i) = \sigma^{-2}(i+2)(i+1)\mathbb{P}(\xi=i+2)$) offspring distribution and X being distributed as $\#\mathcal{T}$.



Proof of the invariance principle in a nutshell

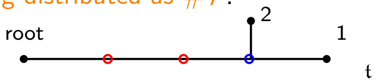
- Let $\mathcal{T}$ be an unconditioned (ordered) Galton-Watson tree, and $\mathfrak{t}$ be a discrete tree with k leaves labelled $1, \ldots, k$ and $k-1$ unlabeled branch points. Then

$$\mathbb{E}\big[\# \text{ subtrees of } \mathcal{T} \text{ with } k \text{ leaves isomorphic to } \mathfrak{t}\big]\mathbf{1}\{\mathcal{T}=n\}$$
$$=\big(\tfrac{\sigma^2}{2}\big)^{k-1}\mathbb{P}\big\{S_{L(t)+k}=n-\#\mathfrak{t}-L(t)\big\},$$

where

$$L(t)=\sum_{i=1}^{\#\mathfrak{t}-(2k-1)}\hat{\xi}_i+\sum_{i=1}^{k-1}\tilde{\xi}_i, \quad S_n:=\sum_{i=1}^{m}X_i,$$

and $(\hat{\xi}_i,\tilde{\xi}_i,X_i)$ are i.i.d. with $\hat{\xi}$ having the "size"-biased offspring distribution $\big(\mathbb{P}(\hat{\xi}=i)=(i+1)\mathbb{P}(\xi=i)$; note that $\mathbb{E}[\hat{\xi}=\sigma^2]\big)$, $\tilde{\xi}$ having the "size-size"-biased $\big(\mathbb{P}(\tilde{\xi}=i)=\sigma^{-2}(i+2)(i+1)\mathbb{P}(\xi=i+2)\big)$ offspring distribution and X being distributed as $\#\mathcal{T}$.



Proof of the invariance principle in a nutshell

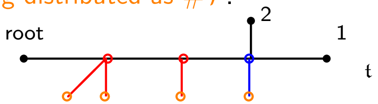
- Let $\mathcal{T}$ be an unconditioned (ordered) Galton-Watson tree, and $\mathfrak{t}$ be a discrete tree with k leaves labelled $1, \dots, k$ and $k-1$ unlabeled branch points. Then

$$\mathbb{E}\big[\#\text{ subtrees of } \mathcal{T} \text{ with } k \text{ leaves isomorphic to } \mathfrak{t}\big]\mathbf{1}\{\mathcal{T}=n\}$$
$$= \big(\tfrac{\sigma^2}{2}\big)^{k-1}\mathbb{P}\big\{S_{L(t)+k} = n - \#\mathfrak{t} - L(t)\big\},$$

where

$$L(t) = \sum_{i=1}^{\#\mathfrak{t}-(2k-1)} \hat{\xi}_i + \sum_{i=1}^{k-1} \tilde{\xi}_i, \quad S_n := \sum_{i=1}^{m} X_i,$$

and $(\hat{\xi}_i, \tilde{\xi}_i, X_i)$ are i.i.d. with $\hat{\xi}$ having the "size"-biased offspring distribution ($\mathbb{P}(\hat{\xi}=i) = (i+1)\mathbb{P}(\xi=i)$; note that $\mathbb{E}[\hat{\xi}=\sigma^2]$), $\tilde{\xi}$ having the "size-size"-biased ($\mathbb{P}(\tilde{\xi}=i) = \sigma^{-2}(i+2)(i+1)\mathbb{P}(\xi=i+2)$) offspring distribution and X being distributed as $\#\mathcal{T}$.



Proof of the invariance principle in a nutshell

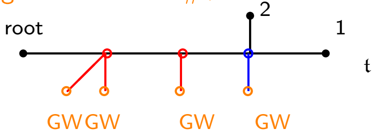
- Let $\mathcal{T}$ be an unconditioned (ordered) Galton-Watson tree, and $\mathfrak{t}$ be a discrete tree with k leaves labelled $1, \ldots, k$ and $k-1$ unlabeled branch points. Then

$$\begin{aligned}&\mathbb{E}\big[\#\text{ subtrees of } \mathcal{T} \text{ with } k \text{ leaves isomorphic to } \mathfrak{t}\big]\mathbf{1}\{\mathcal{T}=n\}\\ &=\big(\tfrac{\sigma^2}{2}\big)^{k-1}\mathbb{P}\big\{S_{L(t)+k}=n-\#\mathfrak{t}-L(t)\big\},\end{aligned}$$

where

$$L(t)=\sum_{i=1}^{\#\mathfrak{t}-(2k-1)}\hat{\xi}_i+\sum_{i=1}^{k-1}\tilde{\xi}_i,\quad S_n:=\sum_{i=1}^{m}X_i,$$

and $(\hat{\xi}_i, \tilde{\xi}_i, X_i)$ are i.i.d. with $\hat{\xi}$ having the "size"-biased offspring distribution ($\mathbb{P}(\hat{\xi}=i)=(i+1)\mathbb{P}(\xi=i)$; note that $\mathbb{E}[\hat{\xi}=\sigma^2]$), $\tilde{\xi}$ having the "size-size"-biased ($\mathbb{P}(\tilde{\xi}=i)=\sigma^{-2}(i+2)(i+1)\mathbb{P}(\xi=i+2)$) offspring distribution and X being distributed as $\#\mathcal{T}$.



- Consequently, if $\mathcal{T}$ is the (ordered) Galton-Watson tree conditioned on size n, and $\mathcal{R}(k,n)$ denotes the random subtree spanned by a sample of size k, then

$$\mathbb{P}\{\mathcal{R}(k,n) \textbf{ is isomorphic to } \mathfrak{t}\}$$
$$= \frac{n!}{(n-k)!}\big(\tfrac{\sigma^2}{2}\big)^{k-1}\frac{\mathbb{P}\{S_{L(\mathfrak{t})+k} = n - \#\mathfrak{t} - L(\mathfrak{t})\}}{\mathbb{P}\{\#\mathcal{T} = n\}}.$$

Proof of the invariance principle in a nutshell

- Consequently, if $\mathcal{T}$ is the (ordered) Galton-Watson tree conditioned on size n, and $\mathcal{R}(k,n)$ denotes the random subtree spanned by a sample of size k, then

$$\begin{aligned}&\mathbb{P}\{\mathcal{R}(k,n) \textbf{ is isomorphic to } \mathfrak{t}\}\\ &= \frac{n!}{(n-k)!}\left(\frac{\sigma^2}{2}\right)^{k-1}\frac{\mathbb{P}\{S_{L(t)+k} = n - \#\mathfrak{t} - L(t)\}}{\mathbb{P}\{\#\mathcal{T} = n\}}.\end{aligned}$$

- Applying the local central limit theorem gives us

$$\begin{aligned}&\mathbb{P}\{\mathcal{R}(k,n) \textbf{ is isomorphic to } \mathfrak{t}\}\\ &= \frac{n!}{(n-k)!}\left(\frac{\sigma^2}{2}\right)^{k-1}\frac{\mathbb{P}\{S_{L(t)+k} = n - \#\mathfrak{t} - L(t)\}}{\mathbb{P}\{\#\mathcal{T} = n\}}\\ &\sim 2^{k-1}\left(\frac{\sigma}{\sqrt{n}}\right)^{2k}\#\mathfrak{t}\exp\left(-\frac{\#\mathfrak{t}^2\sigma^2}{2n}\right), \quad \#\mathfrak{t} = \mathcal{O}(\sqrt{n}).\end{aligned}$$

Literature: Part I

- David Aldous (1993). The continuum random tree III, Annals of Probability.
- Andre Depperschmidt, Andreas Greven and Peter Pfaffelhuber (2011), Marked metric measure spaces, ECP. Tree-valued Fleming-Viot dynamics, PTRF.
- Andreas Greven, Peter Pfaffelhuber and Anita Winter (2013), Tree-valued Fleming-Viot dynamics, PTRF.
- Wolfgang Löhr (2013), Equivalence of Gromov-Prohorov- and Gromov's Box-Metric on the Space of Metric Measure Spaces, ECP.

Part II

Outline: Part II
CRT as scaling limit

Outline

1. Convergence of large critical Galton Watson (GW) trees to the CRT.
2. Along the way, we will introduce some tools (Depth-first search tree, Lukasiewicz path) that will be not only useful to study large GW trees. Indeed we will explore two applications of the approach presented in this section:
   (2.1) A (short) detour: Lévy trees.
   (2.2) Erdös Rényi Graph near criticality.

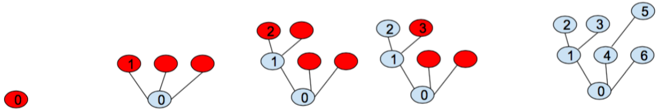
Ordered Galton Watson trees

- Consider a GW tree with offspring distribution μ with the condition

$$\underbrace{\langle \mu, x \rangle = 1}_{\text{critical}}, \quad \underbrace{\langle \mu, x^2 \rangle - 1 = \sigma^2 < \infty}_{\text{finite variance}}$$

- Instead of considering a plain GW trees, we endow the tree with an ordering of the nodes.

Ordered trees



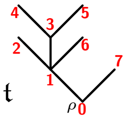
- *Step* 0: Label the root of the tree by 0. 0 belongs to the stack.
- *Step* k: Remove node k from the stack. Node $k+1$ is chosen according to the following rule.
  - If k has $n_k > 0$ children. Add all the nodes to the stack. Pick one of the children uniformly at random, label it $k+1$.
  - If $n_k = 0$, pick the highest node available in the stack and label it $k+1$. If there is no more individual in the stack, then the exploration is over.

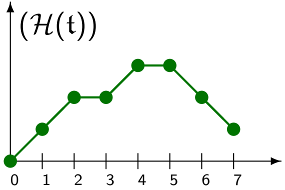
Depth-First Search Algorithm

- Related to the *depth-first-search algorithm*.
- Start the exploration at the root (Step 0).
- At a step k, among all the vertices which have been discovered so far, explore the offspring of the heighest individual
- Labels in the tree correspond to the order of exploration.
- The labelling of the nodes induce a natural encoding of the tree by the *height process* $\mathcal{H}$

$$\mathcal{H}(k) \ = \ \text{graph distance of node } k \text{ from the root}$$

$$\mathcal{H}(k) \;=\; \text{graph distance of node } k \text{ from the root}$$





Theorem (Aldous (93) Marckert, Mokkadem (03) for a stronger version)

Let $\mathcal{T}$ be a critical ordered GW tree with $\mu(1) \in (0,1)$ conditioned on the event size$(\mathcal{T}) = n$. Then

$$\left(\frac{1}{\sqrt{n}}\mathcal{H}([nt]); t \in [0,1]\right) \Longrightarrow \frac{2}{\sigma}\mathbf{e}$$

where $\mathbf{e}$ *is the Brownian excursion of length* 1 *and the convergence is meant in the weak toplogy.*

- Applications in Combinatorics. After conditioning on the event $\{\text{size}(\mathcal{T}) = n\}$:
  - $\mu(k) = \frac{1}{2^{k+1}}$ u.m. on rooted, ordered trees of size n.
  - $\mu = \frac{1}{2}(\delta_0 + \delta_2)$ u.m. on rooted, ordered, binary trees of size n.
  - $\mu(k) = \exp(-1)/k!$: related to rooted Cayley trees (uniformed labelled (not ordered) trees).
- *Universality principle*: diameter of those combinatorial objects of size n is of the order $\sqrt{n}$.

Strategy of the proof

- We start with an "unconditioned version" of Aldous result.
- *Ordered infinite GW forest* : Start with a single labelled GW tree. Let N_1 its size so that its nodes are labelled from 0 to $N_1 - 1$.
- Label the root of the second tree with N_1 and so on.
- Define the height of the n^{th} node as the distance to the floor of the forest.

Proposition (Variation from Aldous result)

$$\frac{1}{\sqrt{n}} \mathcal{H}([n\cdot]) \Longrightarrow \left(\frac{2}{\sigma} \left(w(t) - \inf_{[0,t]} w \right) ; \; t \geq 0 \right)$$

where w is a std BM.

- Excursions of the RHS away from 0 encode the large trees of the underlying random forest.

The Lukasiewicz path S

Definition

The Lukasiewicz path is the integer valued process $S(k) = \rho_k - n_k$ where

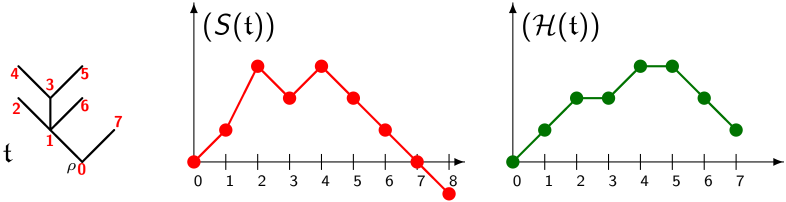
- ρ_k is the number of elements in the stack at time k,
- n_k is the label of the tree visited at k (labelled from 1 to ∞).

Lemma

$$\Delta S(k) := S(k+1) - S(k) = X_k - 1$$

where X_k is the number of children of k. In particular, S is a critical and spectrally positive random walk starting at 0.

$$\begin{array}{rc} \text{if } X_k > 0 & \Delta\rho_k = X_k - 1,\ \Delta n_k = 0 \\ \text{if } X_k = 0, \rho_k = 1 & \Delta\rho_k = -1 + 1,\ \Delta n_k = 1 \\ \text{if } X_k = 0, \rho_k > 1 & \Delta\rho_k = -1,\ \Delta n_k = 0 \end{array}$$



Question: Relation between $\mathcal{H}$, S and the underlying tree ?

- Lukasiewicz path: $S(k) = \rho_k - n_k$
- The Lukasiewicz path provides a direct information about the size of the trees. If

$$\tau = \inf\{n \ : \ S(n) = -1\}$$

Then $\tau = \text{size}(\mathcal{T}_1)$ where $\mathcal{T}_1$ is the first tree in the forest.

Proposition

If $\mu(1) > 0$ (aperiodicity condition on S) then

$$P\left(\text{size}(\mathcal{T}_1) = n\right) \sim \frac{c}{n^{3/2}} \ \text{ as } n \to \infty$$

- *More generally*, the lengths of the successive excursions of the reflected process

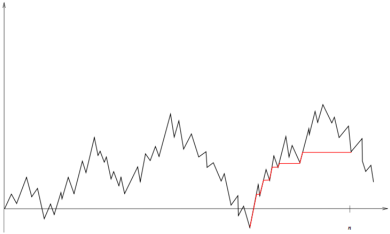
$$\left(S(t) - \inf_{[0,t]} S; t \geq 0\right)$$

away from 0 coincide with the size of successive trees.

Relation between S and $\mathcal{H}$

Lemma

$$\mathcal{H}(p) \;=\; \#\{1 \le i < p \;:\; \inf_{u \in \{i, \cdots, p\}} S(u) \;=\; S(i)\}$$



- *Spine decomposition*: sufficient to show that the ancestors of p are provided by the set above.
- Show that the father of p is the greatest element of the previous set. Then proceed by induction.

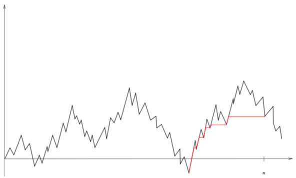
Duality Principle

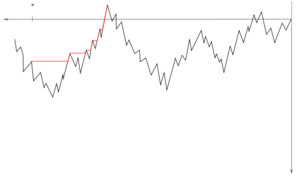
- Define the dual walk at p (Geometrically: flip the picture by 180°):

$$\forall k \le p \ : \ \hat{S}^p(k) \ = \ S(p) - S(p-k)$$

- *$\hat{S}^p$ is distributed as the original walk.*
- Straightforward manipulations yield

$$\mathcal{H}(p) = \#\{1 \le i \le p \ : \ \hat{S}^p(i) = \max_{u \in \{0,\cdots,i\}} \hat{S}^p\}$$





Ladder height process

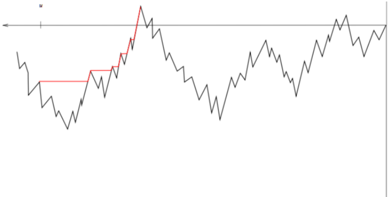
- Let S be the Lukasiewicz path.
- Set $\tau_0 = 0$ and for any $k \geq 0$

$$\tau_{k+1} \;=\; \inf\{j > \tau_k \;:\; S(j) \geq S(\tau_k), \quad O_{k+1} = S(\tau_{k+1}) \;-\; S(\tau_k)\}$$

the sequence of (weak) record times of S and the corresponding overshoots upon reaching those maxima.

Lemma

- $(\tau_{k+1} - \tau_k, O_{k+1})_{k \geq 0}$ *is a sequence of i.i.d r.v.'s.*
- $E\left(O_1\right) = \sigma^2/2$ *(relies on the fact that the walk is spectrally positive).*



A nice asymptotic relation

- Recall that $\mathcal{H}(p) = \#\{1 \leq i \leq p \ : \ \hat{S}^p(i) = \max_{u \in \{1,\cdots,i\}} \hat{S}^p(u)\}$. When p is large, we claim that

$$\frac{\sigma^2}{2} \#\{1 \leq i \leq p \ : \ \hat{S}^p(i) = \max_{u \in \{1,\cdots,p\}} \hat{S}^p(u)\} \ \approx \ \max_{u \in \{1,\cdots,p\}} \hat{S}^p(u)$$

- *RHS is the sum of the overshoots of the dual walk in* $[p]$ ($\max_{u \in \{0,\cdots,p\}} S(u) = \sum_{\tau_k \leq p} O_k$)
- Since the overshoots are i.i.d., the latter approximation is a direct consequence of the L.L.N.
- Finally, straightforward manipulations yield that

$$\mathcal{H}(p) \ \approx \ \frac{2}{\sigma^2} \left(S(p) - \inf_{\{0,\cdots,p\}} S \right), \ \text{ so that}$$

$$\frac{1}{\sqrt{n}} \mathcal{H}([n \cdot]) \Longrightarrow \left(\frac{2}{\sigma} \left(w(t) - \inf_{[0,t]} w \right) ; \ t \geq 0 \right) \quad ()$$

Conditioning

- With some extra work (Marckert Mokkadem (03)): there exists $\alpha > 0$

$$P\left(\sup_{[0,1]} \frac{1}{\sqrt{n}}|\mathcal{H}([nt]) - \frac{2}{\sigma^2}\left(S([nt]) - \inf_{[0,nt]} S\right)| > \frac{1}{n^{1/8}}\right) \leq \exp(-n^{\alpha})$$

- Since $P\left(\underbrace{\text{size}(\mathcal{T}_1) = n}_{=A_n}\right) \approx \frac{c}{n^{3/2}}$, there exists $0 < \alpha' < \alpha$ s.t.

$$P\left(\sup_{[0,1]} \frac{1}{\sqrt{n}}|\mathcal{H}([nt]) - \frac{2}{\sigma^2}\left(S([nt]) - \inf_{[0,nt]} S\right)| > \frac{1}{n^{1/8}} \mid A_n\right) \leq \exp(-n^{\alpha'})$$

- On A_n, $S([n\cdot])$ makes an excursion away from -1, and thus

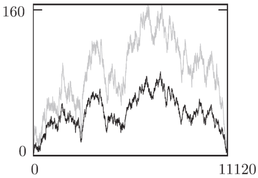
$$P\left(\sup_{[0,1]} \frac{1}{\sqrt{n}}|\mathcal{H}([nt]) - \frac{2}{\sigma^2} S([nt])| > \frac{1}{n^{1/8}} \mid A_n\right) \leq \exp(-n^{\alpha'})$$

- Since $\frac{1}{\sqrt{n}} S([n\cdot]) \mid A_n \Longrightarrow \sigma \mathbf{e}$ ...

Joint convergence of the Lukasiewicz and height processes.

Theorem (Marckert, Mokkadem (03))

$$\frac{1}{\sqrt{n}}\left(S([nt]), \mathcal{H}([nt]); t \in [0,1]\right) \Longrightarrow \left(\mathbf{e}, \frac{2}{\sigma}\mathbf{e}\right)$$



Lévy tree (Le Gall, Le Jan (98), Dusquesne, Legall (02))

- What about the scaling limit of large trees with $\langle \mu, x \rangle = 1$ but $\langle \mu, x^2 \rangle = \infty$?
- Recall the formula :

$$\mathcal{H}(p) \;=\; \#\{1 \leq i \leq p \;:\; \max_{\{0,\cdots,i\}} \hat{S}^p \;-\; S^p(i) = 0\}$$

- (Under mild assumptions), there exists $\epsilon_p \to 0$ such that

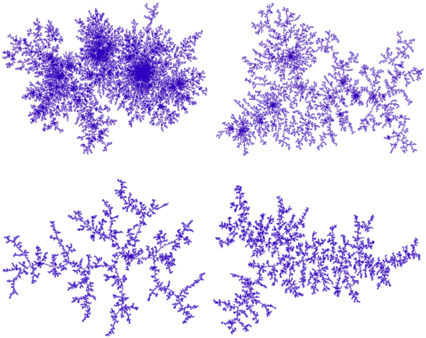
$$(\epsilon_p S([pt]); t \geq 0) \implies X$$

  where X is a spectrally positive Lévy process with inifinite variation, and which does not drift to $+\infty$.
- Define L_t the local time at 0 of the process $\sup_{[0,t]} X - X_t$ and define

$$\mathcal{H}^\infty(t) \;=\hat{L}_t^{(t)}$$

  where $\hat{L}^{(t)}$ is the local time at 0 for the dual process $(X(t) - X(t-s); s \in [0,t])$
- There exists a continuous extension of $\mathcal{H}^\infty$ on $\mathbb{R}^+$.

Figure from Igor Kortchemski ($\alpha = 1.1, 1.5, 1.9, 2$)



(3) Near critical Erdös-Rényi graph

- *Phase transition*. If the probability of connectivity between two vertices is c/n then

$$\begin{cases} |L_n| = O(n) & \text{if } c > 1 \\ |L_n| = O(log(n)) & \text{if } c < 1 \end{cases}$$

  where L_n is the largest connected component.
- $G(n, \frac{1}{n} + \frac{\lambda}{n^{4/3}})$: ER graph of size n and parameter $\frac{1}{n} + \frac{\lambda}{n^{4/3}}$.
- *Near critical random walk*. Consider a sequence of random walks with

$$E(\Delta S^{(p)}[pt]) = \frac{c(t)}{\sqrt{p}} + o(1/\sqrt{p}) \text{ and } Var(\Delta S) = \sigma^2 + o(1)$$

  Then (with some extra conditions for tightness)

$$\left(\frac{1}{\sqrt{p}} S^{(p)}([pt]); t \geq 0\right) \implies \left(\sigma w(t) + \int_0^t c(s)ds; t \geq 0\right)$$

Depth-first spanning forest

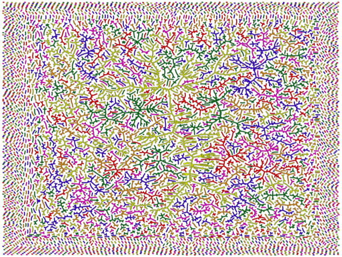


Figure: from Broutin's random gallery

- Explore the ER graph sequentially using the depth-first algorithm.
- This generates a random (ordered) spanning forest of the graph (each tree corresponding to the search tree of a cluster).
- Let $S^{(n)}$ be the Lukasiewicz path of the forest.

Theorem (Aldous (97))

$$\left(\frac{1}{n^{1/3}}S^{(n)}([n^{2/3}t];\ t\geq 0\right)\Longrightarrow\left(B^{\lambda}(t):=w(t)+t\lambda-\frac{t^2}{2};\ t\geq 0\right)$$

- Largest excursions of B^{λ} above its past infimum is finite.
- Lengths of the successive excursions of the reflected $S^{(n)}$ coincide with the size of the clusters in the ER graph.
- At fixed n, let c_i^n be the size of the i^{th} cluster (ranked in decreasing order).

Corollary

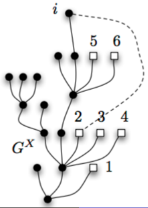
$\frac{1}{n^{2/3}}(c_i^n; i\geq 0) \Longrightarrow (c_i^{\infty}; i\geq 0)$ (in $l^2_{\downarrow}$), where c^{∞} is the sequence of (ranked) excursion lengths of $B^{\lambda}-\inf_{[0,\cdot]}B^{\lambda}$.

Proof of the theorem

- $\Delta S^{(n)}(p)$ is distributed as

$$\text{Binomial}\left(n - p - \bar{S}^{(n)}(p), \frac{1}{n} + \frac{\lambda}{n^{4/3}}\right) - 1$$

  where $\bar{S}^{(n)}$ is the reflection of $S^{(n)}$ above its past infimum. (p terms have been fully explored; *terms in the stack are not eligible to avoid cycles* when constructing the depth-search spanning forest).
- Consider the walk $S'^{(n)}$ with $\Delta S'^{(n)}$ distributed as $\text{Binomial}\left(n - p, \frac{1}{n} + \frac{\lambda}{n^{4/3}}\right) - 1$



Proof of the theorem

- $\Delta S'^{(n)} \sim \text{Binomial}\left(n-p, \frac{1}{n}+\frac{\lambda}{n^{4/3}})\right) - 1$, and thus

$$\begin{aligned} E(\Delta S'^{(n)}([n^{2/3}t])) &= (\frac{1}{n}+\frac{\lambda}{n^{4/3}})\left(n-[n^{2/3}t]\right) - 1 \approx \frac{\lambda - t}{n^{1/3}}, \\ \text{Var}(\Delta S'^{(n)}([n^{2/3}t])) &\approx 1. \end{aligned}$$

  i.e., S'^{n} is a *near-critical random walk*, and thus

$$\left(\frac{1}{n^{1/3}} S'^{(n)}([n^{2/3}t]); t \geq 0\right) \Longrightarrow \left(w(t) + \lambda t - \frac{t^2}{2};\ t \geq 0\right).$$

- Recall that
  $\Delta S^{(n)}(p) \sim \text{Binomial}\left(n - p - \bar{S}^{(n)}(p), \frac{1}{n}+\frac{\lambda}{n^{4/3}}\right) - 1.$
- *Question* : how good is the approximation of S by S' ?

- $S'^{(n)}$ stochastically dominates $S^{(n)}$. Consider the natural coupling such that $S'^{(n)} \geq S^{(n)}$, i.e., couple the increments $X_p^{(n)}$ and $X_p'^{(n)}$ such that

$$Y_p^{(n)} = X_p'^{(n)} - X_p^{(n)}$$

is identical in law to *Binomial* $\left(\bar{S}(p), \frac{1}{n} + \frac{\lambda}{n^{4/3}}\right)$.

- Under this coupling

$$0 \leq S'^{(n)}(p) - S^{(n)}(p) \; = \; \sum_{i=1}^{p} Y_i^{(n)}.$$

- The invariance principle on $S'^{(n)}$ shows that $\bar{S}^{(n)}([n^{2/3}t]) = O(n^{1/3})$ and thus

$$\sum_{i=1}^{[n^{2/3}t]} Y_p^{(n)} \; = \; n^{2/3} O\left(\frac{1}{n} n^{1/3}\right) = O(1)$$

Cycles

- With a little bit of extra work,

$$\left(\sum_{i=1}^{[n^{2/3}t]} Y_p^{(n)}, \frac{1}{n^{1/3}} S^{(n)}([n^{2/3}t]);\ t \geq 0\right) \implies \left(\mu^\infty, B^\lambda\right)$$

where conditional on B^λ, μ^∞ is a PPP with intensity measure

$$\left(B^\lambda(t) - \inf_{[0,t]} B^\lambda\right) dt$$

- *Interpretation of μ^∞*: time at which a cycle occurs as we explore the ER graph. (In generating S', we pick an ineligible edge in the stack).

Aldous result (97)

- Take a near-critical random graph $G(n, \frac{1}{n} + \frac{\lambda}{n^{4/3}})$
- At fixed n, let $(\frac{1}{n^{2/3}} c_i^n, s_i^n)_i$ be the sequences of cluster sizes and # of surplus edges (where sizes are ranked in decreasing order).
- *Continuum object*: $B^\lambda(t) = w(t) + \lambda t - \frac{t^2}{2}$ and $\bar{B}^\lambda$ reflection above the past infimum. Let μ_∞ be the random point measure such that given $\bar{B}^\lambda$, μ_∞ is a PPP on $\mathbb{R}^+$ with intensity measure

$$\bar{B}^\lambda(t) dt.$$

Define $(c_i^\infty)_i$ the sequence of excursion lengths and s_i^∞ the # of marks under the corresponding excursion.

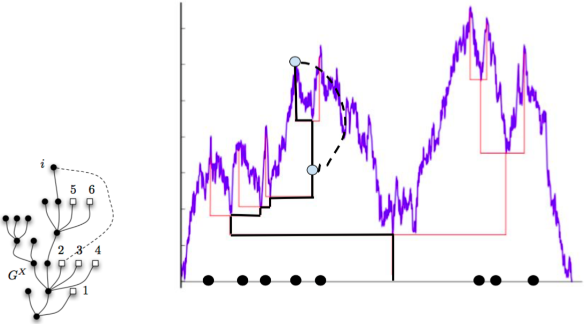
Theorem

$$\left(\frac{1}{n^{2/3}} c_i^n, s_i^n; i \geq 0\right) \implies (c_i^\infty, s_i^\infty; i \geq 0) \text{ (in } l^2_\downarrow).$$

What about the geometry of the ER graph at the limit ? (Addario-Berry, Broutin, Goldsmidt (10))

- For random GW with second finite moment, the Lukasiewicz path is asymptotically equal (up to rescaling) to the height process.
- Intuition behind Addario-Berry, Broutin, Goldsmith (10): in random graphs, *the Lukasiewicz path almost coincides with the height process of the (depth-first search) spanning forest.*
- Excursion of $\bar{B}^{\lambda}$ encodes the trees of the spanning forests and the PPP indicates the occurence of cycles.
- *Extra edges are added on top the continuum trees in order to take into account the existence of cycles.*

The limit of the critical ER (seen as metric spaces)



- Poisson times : leaves at the end point of an edge creating a cycle
- Other end point: chosen uniformly at random along the ancestral line.

Two natural dynamics on the ER graph (I)

- *See λ as a "time" parameter.*
- Natural coupling: i.i.d U_e Uniform([0,1]) r.v. on every edge. Declare an edge to be open at λ iff

$$U_e \leq \frac{1}{n} + \frac{\lambda}{n^{4/3}}$$

- As λ increases, connected components coalesce.
- Two components with macroscopic size x_1 and x_2 will coalesce in a time window $\Delta\lambda$ with probability

$$\underbrace{\frac{\Delta\lambda}{n^{4/3}}}_{\text{proba of becoming open}} \underbrace{c_1 n^{2/3} c_2 n^{2/3}}_{\text{number of closed edges}} = \Delta\lambda x_1 x_2.$$

- The process recording cluster sizes evolve according to a *multiplicative coalescent* (original motivation of Aldous (97))

Two natural dynamics on the ER graph (I)

Theorem (Aldous - CV of 1-d marginals)

Start the multiplicative coalescent with initial condition $(\underbrace{1, \cdots, 1}_{n \text{ times}}, 0, 0, \cdots)$. *Let* $(\tilde{c}^n(\lambda))$ *be the sequences of cluster sizes at time* $\lambda + n^{1/3}$. *Then for every* $\lambda > 0$:

$$\frac{1}{n^{2/3}}(\tilde{c}^n(\lambda)) \Longrightarrow (c^\infty(\lambda)) \quad \text{in } l^2_\downarrow$$

where $(c_\infty(\lambda))$ *is the sequence of excursion lengths of* $\bar{B}^\lambda$.

Two natural dynamics on the ER graph (I)

- Further improvement by Broutin, Marckert (2015)), using the Prim's ordering to generate the spanning forest (minimal spanning forest).
- For every n, $(\tilde{c}^n(\lambda), \lambda \in \mathbb{R})$ is valued in $D(\mathbb{R}, l^2_\downarrow)$.
- Let w be a standard BM and for every λ, use w to define

$$B^\lambda(t) = w(t) + \lambda t - \frac{t^2}{2}$$

- For every time λ, let $c^\infty(\lambda)$ be the sequence of excursion lengths. $(c^\infty(\lambda); \lambda \in \mathbb{R})$ defines a coalescent process.

Theorem (Broutin, Marckert (15))

$$\frac{1}{n^{2/3}}(\tilde{c}^n(\lambda); \lambda \in \mathbb{R}) \Longrightarrow (c^\infty(\lambda); \lambda \in \mathbb{R}) \quad \textit{in } D(\mathbb{R}, l^2_\downarrow).$$

See also Bhamadi, Budhiraja, Wang (2013).

Two natural dynamics on the ER graph (II)

- Poisson clock on every edge: at every clock ring, set the edge open with probability $1/n$.
- Rate of the Poisson clock
  1. *Rate 1*: Roberts and Sengul (17) studied the set of exceptional times at which an anomalous component appears. More precisely, there exists $\beta > 0$ such that

     $$P\left(|L_n|/n^{2/3} ln(n)^{1/3} > \beta\right) \to 1, \; L_n = \text{largest component}$$

  2. *Rate* $1/n^{1/3}$: limiting fragmentation-coagulation process (Rossignol, in progress).

Open questions

- In both dynamics, Addario-Berry et al (10) describe the one-dimensional marginal of the two previous dynamics in terms of the marked excursions of the process

$$\bar{B}^{\lambda} = B^{\lambda}(t) - \inf_{[0,t]} B^{\lambda}, \quad B^{\lambda}(t) = w(t) + \lambda t - \frac{t^2}{2}$$

- For the multiplicative coalescent, Broutin Marckert (15) describe the evolution of cluster sizes. What about the geometry of the clusters ? Does there exist a multi-dimensional version of the construction of Addario-Berry et al (10) ?
- Same question for the second dynamics.

Open questions

- Main issue with Broutin Marckert (15): coding is done using the Prim's order.
- To construct dynamics (I): assign i.i.d U_e Uniform([01]) r.v. at every edge $e \in [n] \times [n]$. Declare an edge to be open at time λ iff $U_e \leq \frac{1}{n} + \frac{\lambda}{n^{4/3}}$.
- Starting from v_1, perform invasion percolation on the complete graph using the weights (U_e) and order vertices according to their order of visit.
- Perform the exploration by always exploring the "smallest" vertex available in the stack (in contrast with the deepest vertex available).
- The length of excursions of the Lukasiewicz path still correspond to the size of the clusters. But no obvious way to recover the geometry !

Open questions

- Same question for dynamics (II). At every time t, the geometry can be described in terms of the excursions of

$$\bar{B}^0 = B^0(t) - \inf_{[0,t]} B^0, \quad B^0(t) = w(t) - \frac{t^2}{2}$$

- At the limit, there should exists a field $(w(\sigma, t))$ such that for every σ, $t \to w(\sigma, t)$ is a standard Brownian motion. Can we describe the structure of this field ? Is it Gaussian ?
- The lenghs of the excursions above the past infimum of $t \to w(\sigma, t) - \frac{t^2}{2}$ would provide a description of cluster sizes at "dynamical" time σ.
- Which ordering should we pick to encode the lengths of the excursion ? Depth ? Prim's ordering ?
- What about the geometry at different times ?

- A large class of random graph exhibiting a phase transition are believed to behave as the ER in the critical window.
- Poissonian random graph (Norros-Reittul model). Every vertex i is assigned an attractiveness $w_i^{(n)}$. Assume that

$$\frac{1}{n}\sum_{i=1}^{n}\delta_{w_i^{(n)}} \Longrightarrow W$$

  Given $(w_i^{(n)})$ set (i,j) open with probability $1-\exp\left(-w_i^{(n)}w_j^{(n)}/\sum w_k^{(n)}\right)$.
- Define $\nu = \frac{E(W^2)}{E(W)}$.
- The model exhibits a phase transition at $\nu = 1$ (Bollobàs, Janson,Riordan (07)).

Universality class of the Erös Rényi graph and beyond

- Bhamidi, Sen, Wang (14): assuming $E(W^{6+\epsilon}) < \infty$, the geometric structure is similar to ER at criticallity, i.e., there is convergence to random metric space described by Addario-Berry et al. See also Bhamidi, Broutin, Sen, Wang (13).
- Bhamidi, van der Hofstad, van Leeuwaarden (10): when $E(W^3) < \infty$, then the maximal component is of the order $n^{2/3}$.
- Bhamidi, van der Hofstad, Sen (17): when $E(W^3) = \infty$, the size and geometry of the clusters are dramatically different. The limitimg structure can be described in terms Lévy trees.
- As before: is there a natural dynamics at the continuum for those quantities.

Literature: Part II

- Louigi Addario-Berry, Nicolas Broutin and Christina Goldschmidt, Critical random graphs: limiting constructions and distributional properties, Electronic Journal of Probability.
- Louigi Addario-Berry, Nicolas Broutin and Christina Goldschmidt, The continuum limit of critical random graphs, Probability Theory and Related Fields.
- David Aldous (1993). The continuum random tree III, Annals of Probability.
- David Aldous (1997). Brownian excursions, critical random graphs and the multiplicative coalescent, Annals of probability.
- Jürgen Bennies & Götz Kersting (2000). A random walk approach to Galton-Watson trees, Journal of theoretical probability.
- Jean-François Marckert & Abdelkader Mokkadem (2003). The depth first process of Galton-Watson trees converge to the same Brownian excursion, Annals of Probabiity.
- Nicolas Broutin, Jean-François Marckert (2015). A new encoding of coalescent processes. Applications to the additive and multiplicative casesProb. Theory Rel. Fields.
- Mattew Roberts, Bait Sengul (2017). Exceptional times of the critical dynamical Erdös-Rényi graph. Available on Arxiv.

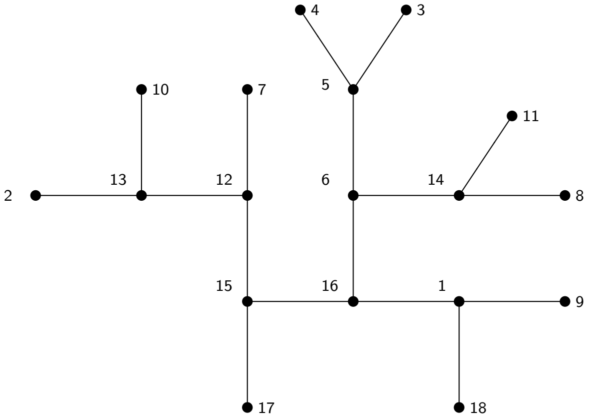
Part III

Classical problem: Cutting down (graph-theoretical) trees

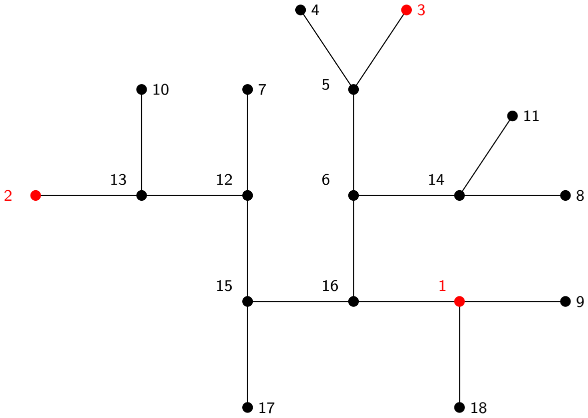
1. Given a finite tree $\mathcal{T} = (T, E)$, distinguish k vertices $\{x_1, ..., x_k\}$.
2. *Remove an edge uniformly at random*, and independent of the k distinguished vertices.

   $\rightsquigarrow$ This disconnects into two subtrees.
3. If one of the subtrees does not contain any of the distinguished points, destroy this subtree. Else we keep two subtrees.
4. We iterate until each distinguished point has been isolated.

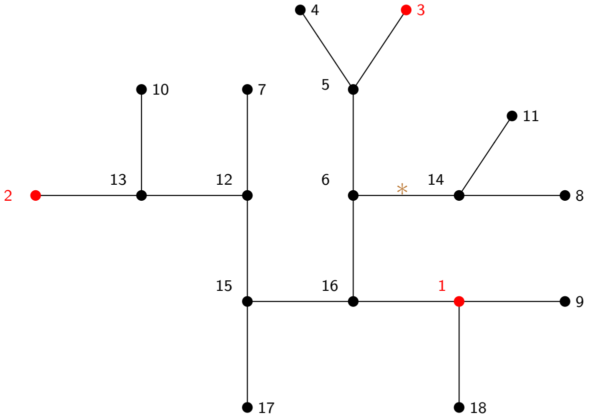
Classical problem: Cutting down (graph-theoretical) trees



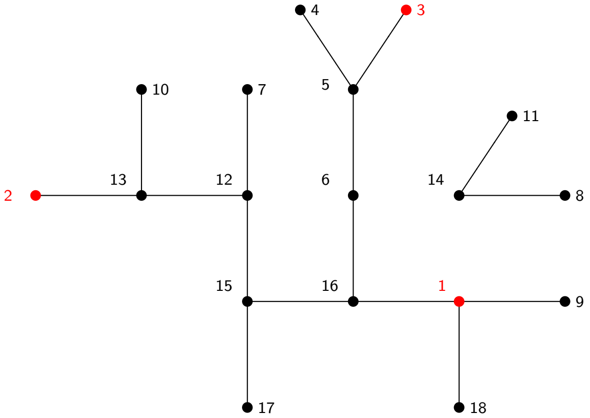
Classical problem: Cutting down (graph-theoretical) trees



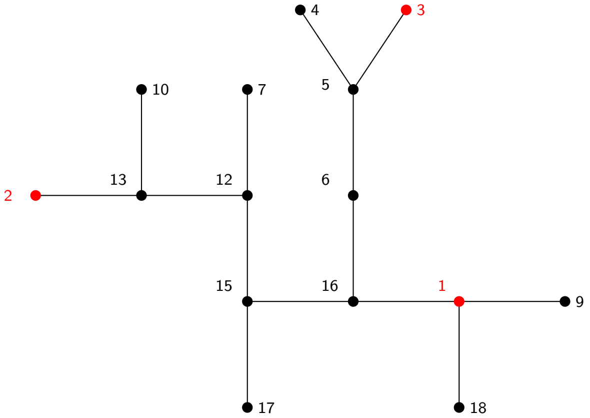
Classical problem: Cutting down (graph-theoretical) trees



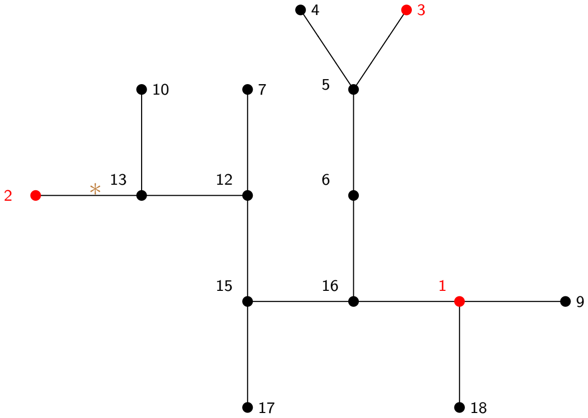
Classical problem: Cutting down (graph-theoretical) trees



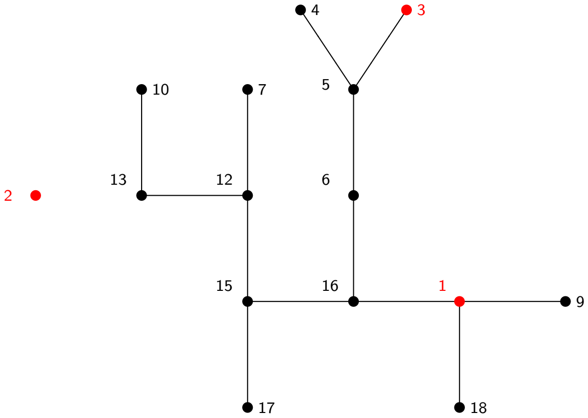
Classical problem: Cutting down (graph-theoretical) trees



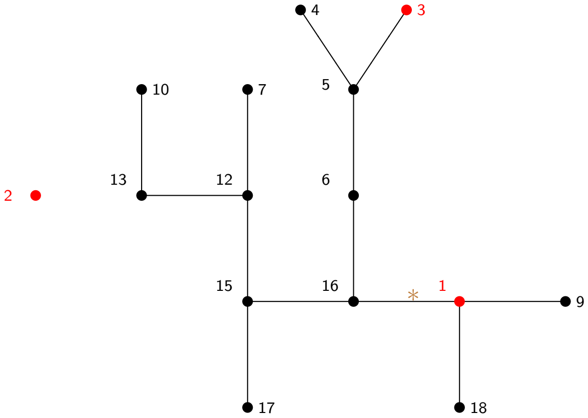
Classical problem: Cutting down (graph-theoretical) trees



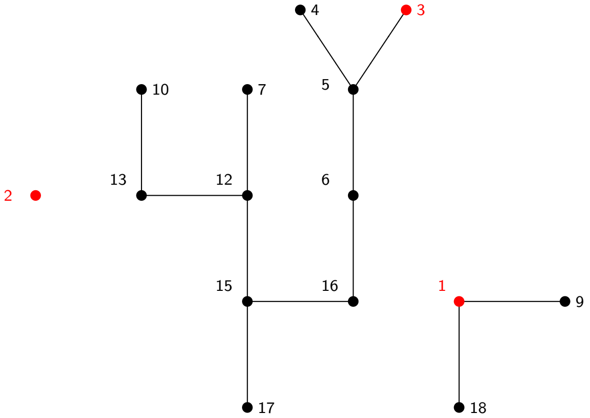
Classical problem: Cutting down (graph-theoretical) trees



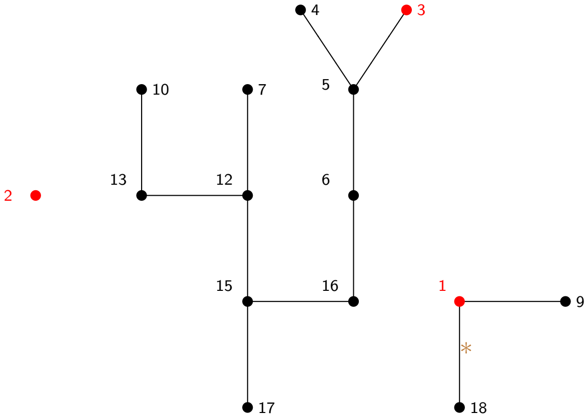
Classical problem: Cutting down (graph-theoretical) trees



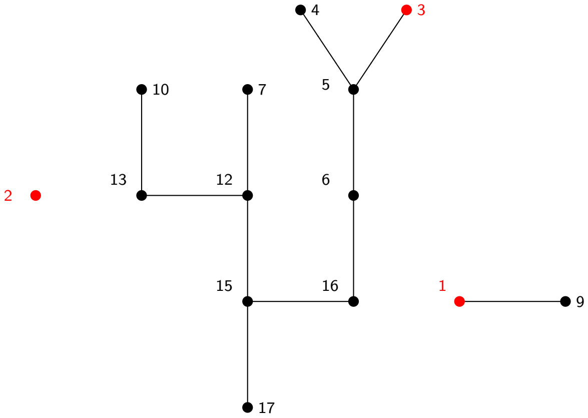
Classical problem: Cutting down (graph-theoretical) trees



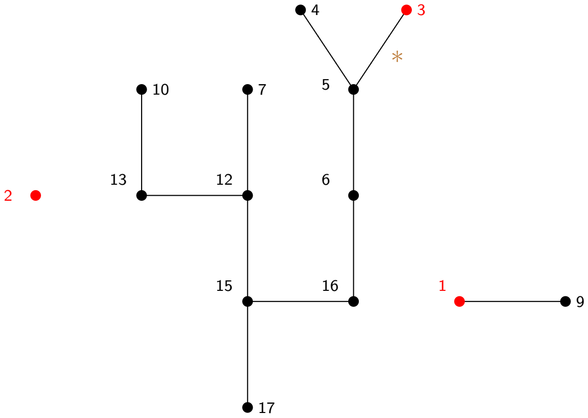
Classical problem: Cutting down (graph-theoretical) trees



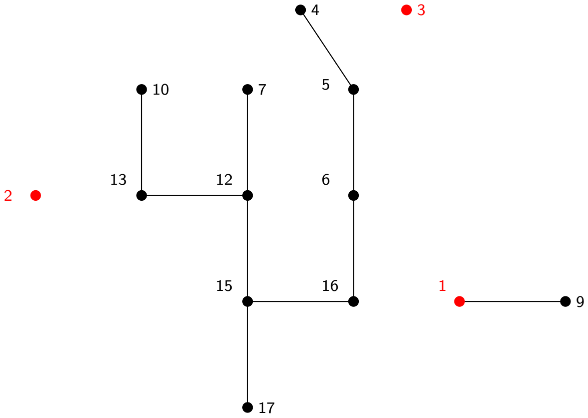
Classical problem: Cutting down (graph-theoretical) trees



Classical problem: Cutting down (graph-theoretical) trees



Classical problem: Cutting down (graph-theoretical) trees



Classical problem: Cutting down (graph-theoretical) trees

Classical problem: Cutting down (graph-theoretical) trees

Classical problem: Cutting down (graph-theoretical) trees

Classical problem: Cutting down (graph-theoretical) trees

1. Given a finite tree $\mathcal{T} = (T, E)$, distinguish k vertices $\{x_1, ..., x_k\}$.
2. *Remove an edge uniformly at random*, and independent of the k distinguished vertices.
   - $\rightsquigarrow$ This disconnects into two subtrees.
3. If one of the subtrees does not contain any of the distinguished points, destroy this subtree. Else we keep two subtrees.
4. We iterate until each distinguished point has been isolated.

1. Denote by $Y(\mathcal{T}, \{x_1, ..., x_k\})$ the (random) number of cuts needed to isolate k random points.

Question:

What can we say about the distribution of $Y(\mathcal{T}, \{x_1, ..., x_k\})$?

The number of cuts needed to isolate k-points

Theorem (Bertoin and Miermont (2013), [6] (see also [8, 1]))

Let $\mathcal{G}_n$ be the GW-tree with critical offspring distribution of finite variance $\sigma^2 > 0$ conditioned to have n vertices. Then for a random sample $\{X_1, ..., X_k\}$ of size k

$$\frac{1}{\sigma\sqrt{n}} Y\left(\mathcal{G}_n, \{X_1, ..., X_k\}\right) \xrightarrow[n\to\infty]{\text{w}} \chi(2k),$$

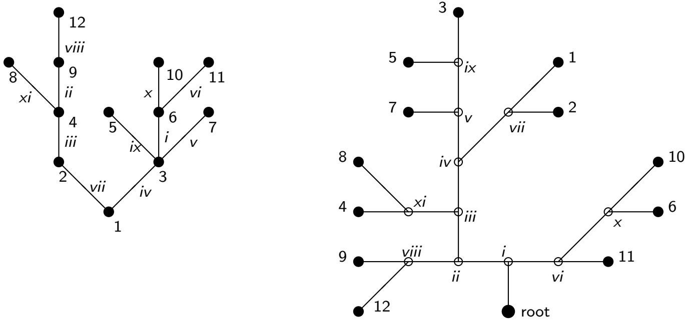
where $\chi(2k)$ is **Chi distributed** *with parameter $2k$.*

The length of the subtree of the CRT spanned by k randomly sampled points is **Chi distributed with parameter $2k$!**

Question. Is this by accident?

$\rightsquigarrow$ genealogy of edge-deletion procedure

- Let (T, E) be a *graph-theoretical tree* with $\#T = n$ (and thus $\#E = n - 1$). Moreover, let $\Pi : E \to \{1, 2, ..., n-1\}$ a random labeling of E indicating in which order the edges are chosen.
- The *fragmentation tree* is a rooted, binary tree with n leaves (other than the root) and such that the **distance from a leaf to the root equals the number of cuts needed to isolate** the original vertex corresponding to that leaf in the fragmentation tree.



$\rightsquigarrow$ To define the cutting procedure on arbitrary trees, we **erase points (on the skeleton)** rather than whole edges.

- Given a a measure $\mathbb{R}$-tree $\mathcal{X} = (T, r, \mu)$ with $\mathrm{supp}(\mu) = T$, we let *cut points rain down on the tree at unit rate per unit length*.
- Each time a cut point hits the tree, it is taken away and thereby splitting one of the connected components into two.
- For each $t \geq 0$, put

$$\mathcal{C}_t := \textbf{ the set of all connected components at time } t,$$

and

$$\mathcal{T}^x(t) := \textbf{ the unique } C \in \mathcal{C}_t \textbf{ with } x \in C$$

.

The fragmentation tree

Given a measure $\mathbb{R}$-tree $x = (T, r, \mu)$ with $\mathrm{supp}(\mu) = T$, *cut points rain down on the tree at unit rate per unit length*.

$\rightsquigarrow$ Once more we keep track of the genealogy of this *fragmentation*.

- Let $\mathcal{T}$ be a random measure $\mathbb{R}$-tree and Π *a Poisson point process on* $T \times \mathbb{R}_+$ *with intensity measure* $\ell^{(T,r)} \otimes \mathrm{d}t$.

The fragmentation tree

Given a measure $\mathbb{R}$-tree $x = (T, r, \mu)$ with $\text{supp}(\mu) = T$, *cut points rain down on the tree at unit rate per unit length*.

$\rightsquigarrow$ Once more we keep track of the genealogy of this *fragmentation*.

- Let $\mathcal{T}$ be a random measure $\mathbb{R}$-tree and Π *a Poisson point process on* $T \times \mathbb{R}_+$ *with intensity measure* $\ell^{(T,r)} \otimes \mathrm{d}t$.
- The fragmentation tree $\text{frag}((T, r, \mu), \Pi) = (\hat{T}, \hat{r}_{\text{frag}}, \hat{\mu}, \rho)$ is the random rooted metric measure tree defined as follows:
  - Put $\hat{T} := \cup_{t \geq 0} C_t$, and $\rho := T$.
  - For $A, B \in \hat{T}$, let $\tau_A := \inf\{t \geq 0 : \exists x \in A \text{ s.t. } (x, t) \in \Pi\}$ and $A \wedge B := \sup\{t \geq 0 : \exists C \in \hat{A} \text{ s.t. } A \cup B \subseteq C\}$. Put

$$\hat{r}_{\text{frag}}(A, B) := (\tau_A - \tau_{A\wedge B}) + (\tau_B - \tau_{A\wedge B}).$$

  - Denote by $S^A := \{A' \in \hat{T} : A' \subseteq A\}$ the *subtree above* $A \in \hat{T}$. There is a unique probability measure $\hat{\mu}$ with

$$\hat{\mu}(S^A) := \mu(A).$$

The fragmentation tree

Given a measure $\mathbb{R}$-tree $x = (T, r, \mu)$ with $\mathrm{supp}(\mu) = T$, *cut points rain down on the tree at unit rate per unit length*.

$\rightsquigarrow$ Once more we keep track of the genealogy of this *fragmentation*.

- Let $\mathcal{T}$ be a random measure $\mathbb{R}$-tree and Π *a Poisson point process on $T \times \mathbb{R}_+$ with intensity measure $\ell^{(T,r)} \otimes \mathrm{d}t$*.
- The fragmentation tree $\mathrm{frag}((T, r, \mu), \Pi) = (\hat{T}, \hat{r}_{\mathrm{frag}}, \hat{\mu}, \rho)$ is a random rooted metric measure tree of infinite height.
  - $\rightsquigarrow$ encodes a fragmentation process

The fragmentation tree

Given a measure $\mathbb{R}$-tree $x = (T, r, \mu)$ with $\mathrm{supp}(\mu) = T$, *cut points rain down on the tree at unit rate per unit length*.

$\rightsquigarrow$ Once more we keep track of the genealogy of this *fragmentation*.

- Let $\mathcal{T}$ be a random measure $\mathbb{R}$-tree and Π *a Poisson point process on $T \times \mathbb{R}_+$ with intensity measure $\ell^{(T,r)} \otimes \mathrm{d}t$*.
- The fragmentation tree $\mathrm{frag}((T, r, \mu), \Pi) = (\hat{T}, \hat{r}_{\mathrm{frag}}, \hat{\mu}, \rho)$ is a random rooted metric measure tree of infinite height.

  $\rightsquigarrow$ encodes a fragmentation process

The fragmentation tree

Given a measure $\mathbb{R}$-tree $\mathcal{X} = (T, r, \mu)$ with $\mathrm{supp}(\mu) = T$, *cut points rain down on the tree at unit rate per unit length*.

$\rightsquigarrow$ Once more we keep track of the genealogy of this *fragmentation*.

- Let $\mathcal{T}$ be a random measure $\mathbb{R}$-tree and Π *a Poisson point process on $T \times \mathbb{R}_+$ with intensity measure $\ell^{(T,r)} \otimes \mathrm{d}t$*.
- The fragmentation tree $\mathrm{frag}((T, r, \mu), \Pi) = (\hat{T}, \hat{r}_{\mathrm{frag}}, \hat{\mu}, \rho)$ is a random rooted metric measure tree of infinite height.

  $\rightsquigarrow$ encodes a fragmentation process

As it takes infinitely many cuts to isolate leaves, we are rather interested in the **rescaled** time needed to isolate points.

$\rightsquigarrow$ Compress distances

How do we compress distances? A martingale argument

- Let (T_N, r_N, μ_N) be a finite graph-theoretical tree with $\#T_N = N$, r_N the graph distance and $\mu_N := \frac{1}{N}\sum_{x \in T_N} \delta_x$.

How do we compress distances? A martingale argument

- Let (T_N, r_N, μ_N) be a finite graph-theoretical tree with $\#T_N = N$, r_N the graph distance and $\mu_N := \frac{1}{N}\sum_{x\in T_N}\delta_x$.
- Assume that there exists a measure $\mathbb{R}$-tree (T, r, μ) and $f : \mathbb{N} \to \mathbb{N}$ such that $\left(T_N, \frac{1}{f(N)} r_N, \mu_N\right) \xrightarrow{\text{Gw}} (T, r, \mu)$.

How do we compress distances? A martingale argument

- Let (T_N, r_N, μ_N) be a finite graph-theoretical tree with $\# T_N = N$, r_N the graph distance and $\mu_N := \frac{1}{N}\sum_{x\in T_N}\delta_x$.
- Assume that there exists a measure $\mathbb{R}$-tree (T, r, μ) and $f : \mathbb{N} \to \mathbb{N}$ such that $\big(T_N, \frac{1}{f(N)} r_N, \mu_N\big) \xrightarrow{\mathrm{Gw}} \big(T, r, \mu\big)$.
- Further, let the edge-deletion process run **at rate** $\frac{f(N)}{N}$ and denote by Y_N^x the number of edges that have been removed by time t from the connected component containing x.

How do we compress distances? A martingale argument

- Let (T_N, r_N, μ_N) be a finite graph-theoretical tree with $\# T_N = N$, r_N the graph distance and $\mu_N := \frac{1}{N}\sum_{x\in T_N}\delta_x$.
- Assume that there exists a measure $\mathbb{R}$-tree (T, r, μ) and $f : \mathbb{N} \to \mathbb{N}$ such that $\big(T_N, \frac{1}{f(N)} r_N, \mu_N\big) \xrightarrow{\mathrm{Gw}} \big(T, r, \mu\big)$.
- Further, let the edge-deletion process run **at rate** $\frac{f(N)}{N}$ and denote by Y_N^x the number of edges that have been removed by time t from the connected component containing x.
- Obviously, for all $x \in T_N$, $Y_N^x(t) \to Y((T_N, r_N), \{x\})$ a.s.

How do we compress distances? A martingale argument

- Let (T_N, r_N, μ_N) be a finite graph-theoretical tree with $\#T_N = N$, r_N the graph distance and $\mu_N := \frac{1}{N}\sum_{x\in T_N}\delta_x$.
- Assume that there exists a measure $\mathbb{R}$-tree (T, r, μ) and $f : \mathbb{N} \to \mathbb{N}$ such that $\left(T_N, \frac{1}{f(N)} r_N, \mu_N\right) \xrightarrow{\mathrm{Gw}} (T, r, \mu)$.
- Further, let the edge-deletion process run **at rate** $\frac{f(N)}{N}$ and denote by Y_N^x the number of edges that have been removed by time t from the connected component containing x.
- Obviously, for all $x \in T_N$, $Y_N^x(t) \to Y((T_N, r_N), \{x\})$ a.s.
- Since edges are removed at rate $\frac{f(N)}{N}$ independently, the process

$$M(t) := Y_N^x(t) - f(N)\int_0^t \tfrac{1}{N}\ell^{(T_N,r_N)}(\mathcal{T}^x(s))\mathrm{d}s, \quad t \ge 0,$$

is a purely discontinuous martingale and

$$\mathbb{E}\Big[\Big(\tfrac{1}{f(N)}Y((T_N, r_N), \{x\}) - \int_0^\infty \mu_N(\mathcal{T}^x(s))\mathrm{d}s\Big)^2\Big] = \tfrac{1}{f(N)}\mathbb{E}\Big[\int_0^\infty \mu_N(\mathcal{T}^x(s))\mathrm{d}s\Big].$$

The cut tree

- Let $\mathcal{T}$ be a random measure $\mathbb{R}$-tree and Π a Poisson point process on $T \times \mathbb{R}_+$ with intensity measure $\ell^{(T,r)} \otimes \mathrm{d}t$.

The cut tree

- Let $\mathcal{T}$ be a random measure $\mathbb{R}$-tree and Π a Poisson point process on $T \times \mathbb{R}_+$ with intensity measure $\ell^{(T,r)} \otimes \mathrm{d}t$.
- The *fragmentation tree*

$$(\hat{T}, \hat{r}_{\mathrm{frag}}, \hat{\mu}, \rho) = \mathrm{frag}((T, r, \mu), \Pi)$$

is a random rooted metric $\mathbb{R}$-tree which encodes the genealogies of the cutting procedure.

The cut tree

- Let $\mathcal{T}$ be a random measure $\mathbb{R}$-tree and Π a Poisson point process on $T \times \mathbb{R}_+$ with intensity measure $\ell^{(T,r)} \otimes \mathrm{d}t$.
- The *fragmentation tree*

$$(\hat{T}, \hat{r}_{\mathrm{frag}}, \hat{\mu}, \rho) = \mathrm{frag}((T, r, \mu), \Pi)$$

  is a random rooted metric $\mathbb{R}$-tree which encodes the genealogies of the cutting procedure.
- The *cut tree*

$$(\hat{T}, \hat{r}_{\mathrm{cut}}, \hat{\mu}, \rho) = \mathrm{cut}((T, r, \mu), \Pi)$$

  has the same tree topology as the fragmentation tree but with compressed distances

$$\hat{r}_{\mathrm{cut}}(A, B) = \int_{\tau_{A \wedge B}}^{\tau_A} \mu(A|_s)\, \mathrm{d}s + \int_{\tau_{A \wedge B}}^{\tau_B} \mu(B|_s)\, \mathrm{d}s,$$

  where $A|_s \in \hat{T}$ denotes the unique $A' \supset A$ with $\hat{r}_{\mathrm{frag}}(A, A') = s$.

$\rightsquigarrow$ encodes additive coalescent

The rescaled number of cuts needed

$$\mathbb{E}\Big[\Big(\tfrac{1}{f(N)}Y((T_N,r_N),\{x\})-\int_0^\infty \mu_N(\mathcal{T}^x(s))\mathrm{d}s\Big)^2\Big]=\tfrac{1}{f(N)}\mathbb{E}\Big[\int_0^\infty \mu_N(\mathcal{T}^x(s))\mathrm{d}s\Big].$$

Theorem

Let (T_N, r_N, μ_N) be a finite graph-theoretical tree with $\# T_N = N$, r_N the graph distance and $\mu_N := \frac{1}{N}\sum_{x\in T_N}\delta_x$. Assume that there exists a measure $\mathbb{R}$-tree (T, r, μ) and $f:\mathbb{N}\to\mathbb{N}$ such that $(T_N, \frac{1}{f(N)}r_N, \mu_N) \xrightarrow{\mathrm{Gw}} (T, r, \mu)$. Then if

$$\sup_{n\in\mathbb{N}} \frac{1}{f(N)}\mathbb{E}\Big[\int_0^\infty \mu_N(\mathcal{T}^x(s))\mathrm{d}s\Big] < \infty,$$

then

$$\frac{Y((T_N,r_N),\{x\})}{f(N)} \xrightarrow[N\to\infty]{\mathrm{w}} \int_0^\infty \mu(\mathcal{T}^x(s))\mathrm{d}s, \quad a.s.$$

Convergence to the cut tree

Theorem

Let (T_N, r_N, μ_N) be a finite graph-theoretical tree with $\#T_N = N$, r_N the graph distance and $\mu_N := \frac{1}{N}\sum_{x\in T_N}\delta_x$. Assume that there exists a measure $\mathbb{R}$-tree (T, r, μ) and $f : \mathbb{N} \to \mathbb{N}$ such that $\big(T_N, \frac{1}{f(N)} r_N, \mu_N\big) \xrightarrow{\text{Gw}} \big(T, r, \mu\big)$. Then if

$$\sup_{n\in\mathbb{N}} \tfrac{1}{f(N)}\mathbb{E}\Big[\int_0^\infty \mu_N\big(\mathcal{T}^x(s)\big)\mathrm{d}s\Big] < \infty,$$

then

$$\tfrac{Y((T_N, r_N), \{x\})}{f(N)} \xrightarrow[N\to\infty]{\text{w}} \int_0^\infty \mu\big(\mathcal{T}^x(s)\big)\mathrm{d}s, \quad a.s.$$

$\rightsquigarrow$ It is probably not hard to show that the **cut tree map** which sends a metric measure tree together with a PPP of unit intensity per unit length to the cut tree **is continuous**.

If so, even the following stronger result holds:

$$\Big(\big(T_N, \tfrac{r_N}{f(N)}, \mu_N\big), \mathrm{cut}\big(\big(T_N, \tfrac{r_N}{f(N)}, \mu_N\big), \Pi_N\big)\Big) \xrightarrow[N\to\infty]{\text{w}} \Big(\big(T, r, \mu\big), \mathrm{cut}\big(\big(T, r, \mu\big), \Pi\big)\Big).$$

The cut tree of the CRT is the CRT

$\rightsquigarrow$ To prove Bertoin and Miermont's statement, it remains to prove that the cut tree of the CRT is the CRT.

The cut tree of the CRT is the CRT

$\rightsquigarrow$ To prove Bertoin and Miermont's statement, it remains to prove that the cut tree of the CRT is the CRT.

Bertoin and Miermont's proof relies on the characterization of the Brownian fragmentation (see, Bertoin 2002 [5]).

The cut tree of the CRT is the CRT

$\rightsquigarrow$ To prove Bertoin and Miermont's statement, it remains to prove that the cut tree of the CRT is the CRT.

Bertoin and Miermont's proof relies on the characterization of the Brownian fragmentation (see, Bertoin 2002 [5]).

Proposition (Total rate affecting a random leaf component is Rayleigh)

If $\mathcal{X} = (T, r, \mu)$ is the CRT and $X \in T$ is a random leaf, then

$$h^{\mathcal{X}}(X) := \int_0^\infty \mu(\mathcal{T}_{\mathcal{X}}^X(s))\mathrm{d}s$$

is Rayleigh distributed.

The cut tree of the CRT is the CRT

- $\rightsquigarrow$ To prove Bertoin and Miermont's statement, it remains to prove that the cut tree of the CRT is the CRT.

Bertoin and Miermont's proof relies on the characterization of the Brownian fragmentation (see, Bertoin 2002 [5]).

Proposition (Total rate affecting a random leaf component is Rayleigh)

If $\mathcal{X} = (T, r, \mu)$ is the CRT and $X \in T$ is a random leaf, then

$$h^{\mathcal{X}}(X) := \int_0^\infty \mu(\mathcal{T}_{\mathcal{X}}^X(s))\mathrm{d}s$$

is Rayleigh distributed.

- $\rightsquigarrow$ The proof of f.d.d.-convergence goes by analogous arguments.

Convergence to the Rayleigh distribution

Lemma (Total rate affecting a random leaf component is Rayleigh)

If $\mathcal{X}$ is the CRT and $X \in T$ is a random leaf, then

$$h^{\mathcal{X}}(X) := \int_0^\infty \mu(\mathcal{T}_{\mathcal{X}}^X(s))\mathrm{d}s$$

is Rayleigh distributed.

Proof. The proof relies on the following identity in law:

$$\left(\mu(\mathcal{T}_{\mathcal{X}}^X(s));\, s \geq 0\right) \stackrel{\mathrm{d}}{=} \left(\tfrac{1}{1+\tau^0(s)};\, s \geq 0\right),$$

where $(\tau^0(s);\, s \geq 0)$ is the *inverse local time process* of reflected BM at level 0. Once this is proven, we obtain

$$\int_0^\infty \mu(\mathcal{T}_{\mathcal{X}}^X(s))\mathrm{d}s \stackrel{d}{=} \int_0^\infty \tfrac{1}{1+\tau^0(s)}\mathrm{d}s =: C(\tau^0),$$

which is the Cauchy transform of $(\tau^0(s),\, s \geq 0)$.

Bertoin showed in [4] that $\mathbb{P}\{C(\tau^0) \leq t\} = 1 - e^{-\frac{t^2}{2}}$, which gives the Rayleigh distribution.

Sketch of proof of the duality to stable subordinators

- Consider the Galton-Watson tree with Poisson offspring conditioned to have n vertices. It is known that this is the uniform unordered labeled tree (with labels ignored).

  The number of all unrooted unordered labeled trees of size n equals n^{n-2}.

- Assume we are taking away an edge sampled uniformly. We are interested in the size distribution of the component Y^X containing a randomly sampled leaf.
- Then applying *Stirling formula*

$$\mathbb{P}\{\#Y^X = k\} = \frac{\binom{n-1}{k-1} k^{k-2}(n-k)^{n-k-2} k(n-k)}{(n-1)n^{n-2}}$$
$$\sim n^{-\frac{3}{2}}(2\pi)^{-\frac{1}{2}} y^{-\frac{1}{2}}(1-y)^{-\frac{3}{2}}, \quad \frac{k}{n} \to y.$$

- One can check that the latter equals the density of $\frac{1}{1+\tau^0(1)}$.
- The general result can be derived by scaling, and checking that $S(t) := \frac{1}{Y^X(t)} - 1$ has the same jump rate densities as $\tau^0(t)$.

Open problems for possible discussion

It is probably not hard to show that the map which sends a metric measure tree together with a PPP of unit intensity per unit length to the cut tree is continuous.

$\rightsquigarrow$ Alternative we cut show that the cut tree of the CRT is the CRT by arguing along the discrete trees.

Open question

Are you aware of a tree models on trees with n vertices on the one hand and binary, rooted trees on the other which satisfies:

1. Both tree models can be rescaled to the CRT.
2. The image of the first tree model together with a random permutation of $\{1, \ldots, n\}$ under this **fragmentation map** is the second model.

Literature: Part III

- Louigi Addario-Berry, Nicolas Broutin and Cecilia Holmgren (2014), Cutting down trees with a Markov chainsaw, Annals of Probability.
- Marie Albenque and Christina Goldschmidt (2015), The Brownian continuum random tree as the unique solution to a fixed point equation, ECP.
- David Aldous and Jim Pitman (1998). The standard additive coalescent. Annals of Probability.
- Jean Bertoin (1997) Cauchy's principal value of local time of Levy processes with no negative jumps via continuous branching processes, EJP.
- Jean Bertoin (2002) Self-similar fragmentation, Poincaré.
- Jean Bertoin & Grégory Miermont (2013). The cut-tree of large Galton-Watson trees and the Brownian CRT, Annals of Applied Probability.
- Benedict Haas and Gregory Miermont (2004), The genealogy of self-similar fragmentations with negative index as a continuum tree, EJP.
- Svantje Janson (2006). Random cuttings and records in deterministic and random trees, Random Structures and Algorithms.
- Ju L. Pavlov (1977). Limit theorems for the number of trees of a given size in a random forest (1977), Math. USSR, Subornik.
- Jim Pitman (1999), Coalescent Random Forests, Journal of Combinatorial Theory.