# Scaling limits of random trees and graphs 

## Christina Goldschmidt

https://www.stats.ox.ac.uk/~goldschm/ALEAminicourse.html

degariment of STATISTICS

## 1. INTRODUCTION: BINARY TREES



## Binary trees

- Let $\mathbb{T}_{n}$ be the set of planted plane binary trees with $n$ leaves.
- Note that every element of $\mathbb{T}_{n}$ has $n-1$ internal vertices and $2 n-1$ edges.



## Binary trees

- Let $\mathbb{T}_{n}$ be the set of planted plane binary trees with $n$ leaves.
- Note that every element of $\mathbb{T}_{n}$ has $n-1$ internal vertices and $2 n-1$ edges.
- Let $\mathbb{T}_{n}^{*}$ be the set of planted plane binary leaf-labelled trees with $n$ labelled leaves.
- The root, labelled 0 is, by convention, not a leaf.



## Binary trees

$\mathbb{T}_{n}$ is the set of planted plane binary trees with $n$ leaves. $\mathbb{T}_{n}^{*}$ is the set of planted plane binary leaf-labelled trees with $n$ labelled leaves.

$$
\left|\mathbb{T}_{n}\right|=\frac{1}{n}\binom{2 n-2}{n-1} \quad \text { (Catalan numbers), } \quad\left|\mathbb{T}_{n}^{*}\right|=n!\left|\mathbb{T}_{n}\right|
$$

## Uniform binary plane trees

$$
\left|\mathbb{T}_{n}\right|=\frac{1}{n}\binom{2 n-2}{n-1} \sim \frac{4^{n-1}}{n^{3 / 2} \sqrt{\pi}} \quad \text { as } n \rightarrow \infty
$$

## Uniform binary plane trees

$$
\left|\mathbb{T}_{n}\right|=\frac{1}{n}\binom{2 n-2}{n-1} \sim \frac{4^{n-1}}{n^{3 / 2} \sqrt{\pi}} \quad \text { as } n \rightarrow \infty
$$

Our first object of interest is a uniform random element of $\mathbb{T}_{n}$.

## Uniform binary plane trees

$$
\left|\mathbb{T}_{n}\right|=\frac{1}{n}\binom{2 n-2}{n-1} \sim \frac{4^{n-1}}{n^{3 / 2} \sqrt{\pi}} \quad \text { as } n \rightarrow \infty
$$

Our first object of interest is a uniform random element of $\mathbb{T}_{n}$. It will be more convenient to work with $\mathbb{T}_{n}^{*}$ and then ignore the leaf labels.

## Uniform binary plane trees

$$
\left|\mathbb{T}_{n}\right|=\frac{1}{n}\binom{2 n-2}{n-1} \sim \frac{4^{n-1}}{n^{3 / 2} \sqrt{\pi}} \quad \text { as } n \rightarrow \infty
$$

Our first object of interest is a uniform random element of $\mathbb{T}_{n}$. It will be more convenient to work with $\mathbb{T}_{n}^{*}$ and then ignore the leaf labels.

Rémy's algorithm recursively constructs a sequence $\left(T_{n}\right)_{n \geq 1}$ of trees such that $T_{n}$ is uniform on $\mathbb{T}_{n}^{*}$ for each $n$.

## Rémy's algorithm

- Start from a single edge with endpoints labelled 0 and 1.
- At step $n \geq 2$, pick an edge uniformly at random, divide it into two edges, insert a new vertex in the middle and attach to that vertex a new edge with a leaf labelled $n$ at its other end, chosen to point in one of the two possible directions each with probability $1 / 2$.

Rémy's algorithm


Rémy's algorithm


Rémy's algorithm


Rémy's algorithm


Rémy's algorithm


Rémy's algorithm


Rémy's algorithm


## Rémy's algorithm

Claim: for each $n, T_{n}$ is a uniform element of $\mathbb{T}_{n}^{*}$.

[J.-L. Rémy, Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire, RAIRO. Informatique théorique 19:2 (1985), pp.179-195]

## Taking limits

Vague question: what can we say about $T_{n}$ as $n \rightarrow \infty$ ?

## Taking limits

Vague question: what can we say about $T_{n}$ as $n \rightarrow \infty$ ?
Concrete first question: as $n \rightarrow \infty$, how does the distance between 0 and 1 behave?

## An urn in Rémy's algorithm

The total number of edges present at step $n$ is equal to $2 n-1$.

Consider the number of edges in the path between 0 and 1 :

- If we add our new leaf somewhere along that path, it gets longer by 1.
- If we add our new leaf anywhere else, the length of the path remains the same.


## An urn in Rémy's algorithm

We have an urn process with two colours, say black and white, where each black ball represents an edge in the path between 0 and 1, and each white ball represents an edge elsewhere.

When we pick a black ball, we replace it in the urn together with one black and one white ball.

When we pick a white ball, we replace it in the urn together with two new white balls.


We start with a single black ball. At step $n$, we always have $2 n-1$ balls present.

## An urn in Rémy's algorithm

Let $B_{n}$ be the number of black balls at step $n$.

## An urn in Rémy's algorithm

Let $B_{n}$ be the number of black balls at step $n$.
We have $B_{1}=1$.

For $n \geq 1$,

$$
\mathbb{E}\left[B_{n+1} \mid \mathcal{F}_{n}\right]=\frac{B_{n}}{2 n-1}\left(B_{n}+1\right)+\frac{2 n-1-B_{n}}{2 n-1} B_{n}
$$

## An urn in Rémy's algorithm

Let $B_{n}$ be the number of black balls at step $n$.
We have $B_{1}=1$.

For $n \geq 1$,

$$
\mathbb{E}\left[B_{n+1} \mid \mathcal{F}_{n}\right]=\frac{B_{n}}{2 n-1}\left(B_{n}+1\right)+\frac{2 n-1-B_{n}}{2 n-1} B_{n}=\frac{2 n}{2 n-1} B_{n} .
$$

## An urn in Rémy's algorithm

Let $B_{n}$ be the number of black balls at step $n$.

We have $B_{1}=1$.

For $n \geq 1$,

$$
\mathbb{E}\left[B_{n+1} \mid \mathcal{F}_{n}\right]=\frac{B_{n}}{2 n-1}\left(B_{n}+1\right)+\frac{2 n-1-B_{n}}{2 n-1} B_{n}=\frac{2 n}{2 n-1} B_{n} .
$$

Define a sequence by $b_{1}=1$ and $b_{n+1}=\frac{2^{2 n}(n!)^{2}}{(2 n)!}$ for $n \geq 1$. Then

$$
b_{n+1}=\frac{2 n}{2 n-1} b_{n} .
$$

## An urn in Rémy's algorithm

Let $B_{n}$ be the number of black balls at step $n$.
We have $B_{1}=1$.

For $n \geq 1$,

$$
\mathbb{E}\left[B_{n+1} \mid \mathcal{F}_{n}\right]=\frac{B_{n}}{2 n-1}\left(B_{n}+1\right)+\frac{2 n-1-B_{n}}{2 n-1} B_{n}=\frac{2 n}{2 n-1} B_{n} .
$$

Define a sequence by $b_{1}=1$ and $b_{n+1}=\frac{2^{2 n}(n!)^{2}}{(2 n)!}$ for $n \geq 1$. Then

$$
b_{n+1}=\frac{2 n}{2 n-1} b_{n} .
$$

Then we have that

$$
\left(\frac{B_{n}}{b_{n}}\right)_{n \geq 1} \text { is a non-negative martingale. }
$$

## Martingale limit

$\left(B_{n} / b_{n}\right)_{n \geq 1}$ is also bounded in $L^{2}$, hence uniformly integrable, and so it has an almost sure limit by the martingale convergence theorem.

Since

$$
b_{n+1}=\frac{2^{2 n}(n!)^{2}}{(2 n)!} \sim \sqrt{\pi n}
$$

we get that

$$
\frac{B_{n}}{\sqrt{2 n}} \rightarrow L \quad \text { a.s. as } n \rightarrow \infty
$$

for some limit random variable $L$.
[P. Marchal, A note on the fragmentation of the stable tree, Fifth Colloquium on Mathematics and Computer Science, DMTCS (2008), pp.489-500]

## Limiting distribution for the length

It also turns out (using a generating function argument) that the law of $B_{n+1}$ is explicit:

$$
\mathbb{P}\left(B_{n+1}=k\right)=\frac{k-1}{n} 2^{k-1} \frac{\binom{2 n-k}{n-1}}{\binom{2 n}{n}}
$$

and so

$$
\mathbb{P}\left(B_{n+1}=\lfloor x \sqrt{2 n}\rfloor\right) \sim \frac{x}{\sqrt{2 n}} e^{-x^{2} / 2}, \quad x>0
$$

## Limiting distribution for the length

It also turns out (using a generating function argument) that the law of $B_{n+1}$ is explicit:

$$
\mathbb{P}\left(B_{n+1}=k\right)=\frac{k-1}{n} 2^{k-1} \frac{\binom{2 n-k}{n-1}}{\binom{2 n}{n}}
$$

and so

$$
\mathbb{P}\left(B_{n+1}=\lfloor x \sqrt{2 n}\rfloor\right) \sim \frac{x}{\sqrt{2 n}} e^{-x^{2} / 2}, \quad x>0 .
$$

In other words, we get

$$
\frac{B_{n}}{\sqrt{2 n}} \rightarrow L \quad \text { a.s. as } n \rightarrow \infty
$$

where the limit $L$ has the Rayleigh distribution, with density $x e^{-x^{2} / 2}$ on $\mathbb{R}_{+}$.

## Consequences

The distance between 0 and 1 varies as $\sqrt{2 n}$, with a nice almost sure limit. What can we say about the distances between the other leaves as $n \rightarrow \infty$ ?

## Consequences

The distance between 0 and 1 varies as $\sqrt{2 n}$, with a nice almost sure limit. What can we say about the distances between the other leaves as $n \rightarrow \infty$ ?

For example, let's think about the distance from 2 to the path between 0 and 1, and the position along that path at which it branches off.


## More urns: self-similarity

At step 2 of Rémy's algorithm, we always have


## More urns: self-similarity

At step 2 of Rémy's algorithm, we always have


Each of the three parts here behaves precisely as a little copy of Rémy's algorithm, although the numbers of leaves we add to each copy are dependent.

## More urns: self-similarity

Each of the three parts here behaves precisely as a little copy of Rémy's algorithm, although the numbers of leaves we add to each copy are dependent.

A useful consequence is that given the three sets of leaves, these three trees are themselves uniform binary plane trees.


## More urns: self-similarity

The numbers of vertices in each of the three little trees evolve according to a variant of Pólya's urn with three colours, red, green and blue. We start with one ball of each colour. We pick a ball at random and replace it in the urn with two more of the same colour. Let $R_{n}, G_{n}, B_{n}$ be the numbers of red, green and blue balls respectively at step $n$.


## More urns: self-similarity

It is then standard that

$$
\frac{1}{2 n+3}\left(R_{n}, G_{n}, B_{n}\right) \rightarrow\left(\Delta_{1}, \Delta_{2}, \Delta_{3}\right) \quad \text { a.s. as } n \rightarrow \infty
$$

where $\left(\Delta_{1}, \Delta_{2}, \Delta_{3}\right) \sim \operatorname{Dirichlet}(1 / 2,1 / 2,1 / 2)$.

## More urns: self-similarity

It is then standard that

$$
\frac{1}{2 n+3}\left(R_{n}, G_{n}, B_{n}\right) \rightarrow\left(\Delta_{1}, \Delta_{2}, \Delta_{3}\right) \quad \text { a.s. as } n \rightarrow \infty
$$

where $\left(\Delta_{1}, \Delta_{2}, \Delta_{3}\right) \sim \operatorname{Dirichlet}(1 / 2,1 / 2,1 / 2)$.
The Dirichlet distribution with parameters $\alpha_{1}, \alpha_{2}, \ldots, \alpha_{k}>0$ has density

$$
\frac{\Gamma\left(\sum_{i=1}^{k} \alpha_{i}\right)}{\prod_{i=1}^{k} \Gamma\left(\alpha_{i}\right)} x_{1}^{\alpha_{1}-1} \ldots x_{k}^{\alpha_{k}-1}
$$

with respect to Lebesgue measure on

$$
\left\{\mathbf{x}=\left(x_{1}, \ldots, x_{k}\right) \in \mathbb{R}_{+}^{k}: \sum_{i=1}^{k} x_{i}=1\right\} .
$$

## More urns: self-similarity

The numbers of leaves in each of the three subtrees are given by

$$
N_{n}^{R}=\left(R_{n}+1\right) / 2, \quad N_{n}^{G}=\left(G_{n}+1\right) / 2, \quad N_{n}^{B}=\left(B_{n}+1\right) / 2 .
$$

So we have

$$
\frac{1}{n}\left(N_{n}^{R}, N_{n}^{G}, N_{n}^{B}\right) \rightarrow\left(\Delta_{1}, \Delta_{2}, \Delta_{3}\right) \quad \text { a.s. }
$$

## More urns: self-similarity

The numbers of leaves in each of the three subtrees are given by

$$
N_{n}^{R}=\left(R_{n}+1\right) / 2, \quad N_{n}^{G}=\left(G_{n}+1\right) / 2, \quad N_{n}^{B}=\left(B_{n}+1\right) / 2 .
$$

So we have

$$
\frac{1}{n}\left(N_{n}^{R}, N_{n}^{G}, N_{n}^{B}\right) \rightarrow\left(\Delta_{1}, \Delta_{2}, \Delta_{3}\right) \quad \text { a.s. }
$$

Writing $L_{n}^{R}, L_{n}^{G}, L_{n}^{B}$ for the lengths of the three paths at step $n$, we see that they look like small copies of the first urn model run for numbers of steps which are approximately $n \Delta_{1}, n \Delta_{2}$ and $n \Delta_{3}$.

## More urns: self-similarity

The numbers of leaves in each of the three subtrees are given by

$$
N_{n}^{R}=\left(R_{n}+1\right) / 2, \quad N_{n}^{G}=\left(G_{n}+1\right) / 2, \quad N_{n}^{B}=\left(B_{n}+1\right) / 2 .
$$

So we have

$$
\frac{1}{n}\left(N_{n}^{R}, N_{n}^{G}, N_{n}^{B}\right) \rightarrow\left(\Delta_{1}, \Delta_{2}, \Delta_{3}\right) \quad \text { a.s. }
$$

Writing $L_{n}^{R}, L_{n}^{G}, L_{n}^{B}$ for the lengths of the three paths at step $n$, we see that they look like small copies of the first urn model run for numbers of steps which are approximately $n \Delta_{1}, n \Delta_{2}$ and $n \Delta_{3}$. It follows that

$$
\frac{1}{\sqrt{2 n}}\left(L_{n}^{R}, L_{n}^{G}, L_{n}^{B}\right) \rightarrow\left(\sqrt{\Delta_{1}} L_{1}, \sqrt{\Delta_{2}} L_{2}, \sqrt{\Delta_{3}} L_{3}\right) \quad \text { a.s. }
$$

where $L_{1}, L_{2}, L_{3}$ are i.i.d. Rayleigh random variables, independent of $\left(\Delta_{1}, \Delta_{2}, \Delta_{3}\right)$.

## Limiting subtree lengths

An elementary distributional calculation yields that

$$
\left(\sqrt{\Delta_{1}} L_{1}, \sqrt{\Delta_{2}} L_{2}, \sqrt{\Delta_{3}} L_{3}\right) \stackrel{d}{=} \sqrt{\Gamma_{2}} \times \operatorname{Dir}(1,1,1)
$$

where $\Gamma_{2} \sim \operatorname{Gamma}(2,1 / 2)$ and the two factors are independent.

## Limiting subtree lengths

An elementary distributional calculation yields that

$$
\left(\sqrt{\Delta_{1}} L_{1}, \sqrt{\Delta_{2}} L_{2}, \sqrt{\Delta_{3}} L_{3}\right) \stackrel{d}{=} \sqrt{\Gamma_{2}} \times \operatorname{Dir}(1,1,1)
$$

where $\Gamma_{2} \sim \operatorname{Gamma}(2,1 / 2)$ and the two factors are independent.
More generally, if we consider the subtree spanned by 0 and the leaves labelled $1,2, \ldots, k$, we get $2 k-1$ edges whose lengths are distributed as

$$
\sqrt{\Gamma_{k}} \times \operatorname{Dir}(\underbrace{1,1, \ldots, 1}_{2 k-1})
$$

where again $\Gamma_{k} \sim \operatorname{Gamma}(k, 1 / 2)$ and the two factors are independent.

## Limiting subtree lengths

An elementary distributional calculation yields that

$$
\left(\sqrt{\Delta_{1}} L_{1}, \sqrt{\Delta_{2}} L_{2}, \sqrt{\Delta_{3}} L_{3}\right) \stackrel{d}{=} \sqrt{\Gamma_{2}} \times \operatorname{Dir}(1,1,1)
$$

where $\Gamma_{2} \sim \operatorname{Gamma}(2,1 / 2)$ and the two factors are independent.
More generally, if we consider the subtree spanned by 0 and the leaves labelled $1,2, \ldots, k$, we get $2 k-1$ edges whose lengths are distributed as

$$
\sqrt{\Gamma_{k}} \times \operatorname{Dir}(\underbrace{1,1, \ldots, 1}_{2 k-1}),
$$

where again $\Gamma_{k} \sim \operatorname{Gamma}(k, 1 / 2)$ and the two factors are independent.
(Note that the $k=1$ case fits into this pattern, since Rayleigh $\stackrel{d}{=} \sqrt{\Gamma_{1}}$. )

## A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

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


## A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

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


A useful way of constructing this is to let $E_{1}, E_{2}, \ldots$ be i.i.d.
$\operatorname{Exp}(1 / 2)$ and set $C_{i}=\sqrt{\sum_{j=1}^{i} E_{j}}$.

## A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

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


A useful way of constructing this is to let $E_{1}, E_{2}, \ldots$ be i.i.d. $\operatorname{Exp}(1 / 2)$ and set $C_{i}=\sqrt{\sum_{j=1}^{i} E_{j}}$.

- Consider the line-segments $\left[0, C_{1}\right),\left[C_{1}, C_{2}\right), \ldots$


## A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

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


A useful way of constructing this is to let $E_{1}, E_{2}, \ldots$ be i.i.d. $\operatorname{Exp}(1 / 2)$ and set $C_{i}=\sqrt{\sum_{j=1}^{i} E_{j}}$.

- Consider the line-segments $\left[0, C_{1}\right),\left[C_{1}, C_{2}\right), \ldots$.
- Start from $\left[0, C_{1}\right)$ and proceed inductively.


## A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

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


A useful way of constructing this is to let $E_{1}, E_{2}, \ldots$ be i.i.d. $\operatorname{Exp}(1 / 2)$ and set $C_{i}=\sqrt{\sum_{j=1}^{i} E_{j}}$.

- Consider the line-segments $\left[0, C_{1}\right),\left[C_{1}, C_{2}\right), \ldots$
- Start from $\left[0, C_{1}\right)$ and proceed inductively.
- For $i \geq 2$, attach $\left[C_{i-1}, C_{i}\right.$ ) at a random point chosen uniformly over the existing tree.


## Line-breaking construction



## Line-breaking construction



## Line-breaking construction



## Line-breaking construction



## Line-breaking construction



## Line-breaking construction



## Line-breaking construction



## Why is this the right limit?

Claim: this gives the almost sure limit of the subtree spanned by 0 and the leaves $1,2, \ldots, k$ in the rescaled version of Rémy's algorithm.

- The tree at step $k \geq 1$ has total length

$$
C_{k}=\sqrt{\sum_{i=1}^{k} E_{i}} \stackrel{d}{=} \sqrt{\Gamma_{k}}
$$

where $\Gamma_{k} \sim \operatorname{Gamma}(k, 1 / 2)$.

- The combinatorics of the attachment mechanism are exactly the same as in Rémy's algorithm - so the underlying binary leaf-labelled tree has the right distribution.
- A calculation shows that the cut-points and attachment points split up the interval $\left[0, C_{k}\right)$ uniformly.


## The line-breaking definition of the Brownian CRT

- Start from $\left[0, C_{1}\right)$ and proceed inductively.
- For $i \geq 1$, sample $B_{i}$ uniformly from $\left[0, C_{i}\right.$ ) and attach [ $C_{i}, C_{i+1}$ ) at the corresponding point of the tree constructed so far (this is a point chosen uniformly at random over the existing tree).


## The line-breaking definition of the Brownian CRT

- Start from $\left[0, C_{1}\right)$ and proceed inductively.
- For $i \geq 1$, sample $B_{i}$ uniformly from $\left[0, C_{i}\right.$ ) and attach [ $C_{i}, C_{i+1}$ ) at the corresponding point of the tree constructed so far (this is a point chosen uniformly at random over the existing tree).

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

This procedure gives (somewhat informally expressed) definition of the Brownian continuum random tree (CRT) which is a key object in this minicourse.

## The line-breaking definition of the Brownian CRT


[Picture by Igor Kortchemski]

## The scaling limit of the uniform binary plane tree

In the next section, we will make sense of the following statement.
Theorem. (Marchal (2003), Curien and Haas (2013))
As $n \rightarrow \infty$,

$$
\frac{1}{\sqrt{2 n}} T_{n} \rightarrow \mathcal{T} \quad \text { a.s. }
$$

where $\mathcal{T}$ is the Brownian CRT.

## The scaling limit of the uniform binary plane tree

In the next section, we will make sense of the following statement.
Theorem. (Marchal (2003), Curien and Haas (2013))
As $n \rightarrow \infty$,

$$
\frac{1}{\sqrt{2 n}} T_{n} \rightarrow \mathcal{T} \quad \text { a.s. }
$$

where $\mathcal{T}$ is the Brownian CRT.

We need to know what sort of objects we're really dealing with, and what is the topology in which the convergence occurs. (Also: why is $\mathcal{T}$ Brownian?!)

Before we do that, let's record an immediate consequence of the theorem.

## Self-similarity of the Brownian CRT

Recall: we split our uniform binary plane tree into three little uniform binary plane trees of random sizes.


This property passes to the limit, and so the Brownian CRT can be split into three randomly rescaled Brownian CRTs. In particular, the Brownian CRT is a random fractal.

## 2. $\mathbb{R}$-TREES AND CONVERGENCE

Key reference:

Jean-François Le Gall, Random trees and applications,
Probability Surveys 2 (2005) pp.245-311.


## Continuous trees

We want a continuous notion of a tree. We don't really care about vertices: the important aspects are the shape of the tree and the distances. So it makes sense to think in terms of metric spaces.

## $\mathbb{R}$-trees

Definition. A compact metric space $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree if for all $x, y \in \mathcal{T}$,

- There exists a unique shortest path $[[x, y]]$ from $x$ to $y$ (of length $d(x, y))$.
- The only non-self-intersecting path from $x$ to $y$ is $[[x, y]]$.


## $\mathbb{R}$-trees

Definition. A compact metric space $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree if for all $x, y \in \mathcal{T}$,

- There exists a unique shortest path [[x,y]] from $x$ to $y$ (of length $d(x, y)$ ). (There is a unique isometric map $f_{x, y}$ from $[0, d(x, y)]$ into $\mathcal{T}$ such that $f(0)=x$ and $f(d(x, y))=y$. We write $f_{x, y}([0, d(x, y)])=[[x, y]]$.)
- The only non-self-intersecting path from $x$ to $y$ is $[[x, y]]$. (If $g$ is a continuous injective map from $[0,1]$ into $\mathcal{T}$, such that $g(0)=x$ and $g(1)=y$, then $g([0,1])=[[x, y]]$.


## $\mathbb{R}$-trees

Definition. A compact metric space $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree if for all $x, y \in \mathcal{T}$,

- There exists a unique shortest path $[[x, y]]$ from $x$ to $y$ (of length $d(x, y)$ ). (There is a unique isometric map $f_{x, y}$ from $[0, d(x, y)]$ into $\mathcal{T}$ such that $f(0)=x$ and $f(d(x, y))=y$. We write $f_{x, y}([0, d(x, y)])=[[x, y]]$.)
- The only non-self-intersecting path from $x$ to $y$ is $[[x, y]]$. (If $g$ is a continuous injective map from $[0,1]$ into $\mathcal{T}$, such that $g(0)=x$ and $g(1)=y$, then $g([0,1])=[[x, y]]$.

An element $v \in \mathcal{T}$ is called a vertex.
A rooted $\mathbb{R}$-tree has a distinguished vertex $\rho$ called the root.
The height of a vertex $v$ is its distance $d(\rho, v)$ from the root.
A leaf is a vertex $v$ such that $v \notin[[\rho, w]]$ for any $w \neq v$.

## Coding $\mathbb{R}$-trees

Let $h:[0,1] \rightarrow \mathbb{R}^{+}$be an excursion, that is a continuous function such that $h(0)=h(1)=0$ and $h(x)>0$ for $x \in(0,1)$.


## Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...


## Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...


## Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...


## Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...


## Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...


## Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...


## Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...


## Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together to get a tree.


## Coding $\mathbb{R}$-trees

Formally, use $h$ to define a distance:

$$
d_{h}(x, y)=h(x)+h(y)-2 \inf _{x \wedge y \leq z \leq x \vee y} h(z) .
$$



## Coding $\mathbb{R}$-trees

Let $y \sim y^{\prime}$ if $d_{h}\left(y, y^{\prime}\right)=0$ and take the quotient $\mathcal{T}_{h}=[0,1] / \sim$.


## Coding $\mathbb{R}$-trees

Theorem. For any excursion $h,\left(\mathcal{T}_{h}, d_{h}\right)$ is an $\mathbb{R}$-tree.

Write $\pi_{h}:[0,1] \rightarrow \mathcal{T}_{h}$ for the projection map.

We will often root $\mathcal{T}_{h}$ at $\rho=\pi_{h}(0)=\pi_{h}(1)$.

## A natural measure

We will want to be able to sample random points in our trees. There is a natural "uniform" measure $\mu_{h}$ which is the push-forward of the Lebesgue measure on $[0,1]$ onto $\mathcal{T}_{h}$.

## A natural measure

We will want to be able to sample random points in our trees.
There is a natural "uniform" measure $\mu_{h}$ which is the push-forward of the Lebesgue measure on $[0,1]$ onto $\mathcal{T}_{h}$.

To pick a point of $\mathcal{T}_{h}$ according to $\mu_{h}$, we simply sample $U \sim U[0,1]$ and then take our point to be $\pi_{h}(U)$.

## A natural measure

We will want to be able to sample random points in our trees.
There is a natural "uniform" measure $\mu_{h}$ which is the push-forward of the Lebesgue measure on $[0,1]$ onto $\mathcal{T}_{h}$.

To pick a point of $\mathcal{T}_{h}$ according to $\mu_{h}$, we simply sample $U \sim \mathrm{U}[0,1]$ and then take our point to be $\pi_{h}(U)$.

We will typically think of our continuous trees as triples $\left(\mathcal{T}_{h}, d_{h}, \mu_{h}\right)$.

## Topological considerations

Let $\mathbb{M}$ be the space of compact metric spaces endowed with a Borel probability measure, up to measure-preserving isometry.

We will define a metric $\mathrm{d}_{\mathrm{GHP}}$, the Gromov-Hausdorff-Prokhorov distance on $\mathbb{M}$.

## Topological considerations

Suppose that $(X, d)$ and $\left(X^{\prime}, d^{\prime}\right)$ are compact metric spaces.


## Topological considerations

Suppose that $(X, d)$ and $\left(X^{\prime}, d^{\prime}\right)$ are compact metric spaces.

A correspondence $R$ is a subset of $X \times X^{\prime}$ such that for every $x \in X$, there exists $x^{\prime} \in X^{\prime}$ with $\left(x, x^{\prime}\right) \in R$ and vice versa.


## Topological considerations

Suppose that $(X, d)$ and $\left(X^{\prime}, d^{\prime}\right)$ are compact metric spaces.

A correspondence $R$ is a subset of $X \times X^{\prime}$ such that for every $x \in X$, there exists $x^{\prime} \in X^{\prime}$ with $\left(x, x^{\prime}\right) \in R$ and vice versa.


## Topological considerations

The distortion of $R$ is

$$
\operatorname{dis}(R)=\sup \left\{\left|d(x, y)-d^{\prime}\left(x^{\prime}, y^{\prime}\right)\right|:\left(x, x^{\prime}\right),\left(y, y^{\prime}\right) \in R\right\} .
$$



## Topological considerations

Suppose that $\mu$ is a Borel probability measure on $(X, d)$ and that $\mu^{\prime}$ is a Borel probability measure on $\left(X^{\prime}, d^{\prime}\right)$.

A measure $\nu$ on $X \times X^{\prime}$ is a coupling of $\mu$ and $\mu^{\prime}$ if $\nu\left(\cdot, X^{\prime}\right)=\mu(\cdot)$ and $\nu(X, \cdot)=\mu^{\prime}(\cdot)$.

## Topological considerations

Suppose that $\mu$ is a Borel probability measure on $(X, d)$ and that $\mu^{\prime}$ is a Borel probability measure on $\left(X^{\prime}, d^{\prime}\right)$.

A measure $\nu$ on $X \times X^{\prime}$ is a coupling of $\mu$ and $\mu^{\prime}$ if $\nu\left(\cdot, X^{\prime}\right)=\mu(\cdot)$ and $\nu(X, \cdot)=\mu^{\prime}(\cdot)$.

Idea: find a correspondence and a coupling such that the correspondence has small distortion and the coupling "lines up" well with the correspondence i.e. if $\left(V, V^{\prime}\right) \sim \nu$ then $\mathbb{P}\left(\left(V, V^{\prime}\right) \in R\right)=\nu(R)$ is close to 1 .

## Topological considerations

The Gromov-Hausdorff-Prokhorov distance between $(X, d, \mu)$ and ( $X^{\prime}, d^{\prime}, \mu^{\prime}$ ) is defined to be

$$
\mathrm{d}_{\mathrm{GHP}}\left((X, d, \mu),\left(X^{\prime}, d^{\prime}, \mu^{\prime}\right)\right)=\frac{1}{2} \inf _{R, \nu} \max \left\{\operatorname{dis}(R), \nu\left(R^{c}\right)\right\} .
$$

## Topological considerations

The Gromov-Hausdorff-Prokhorov distance between ( $X, d, \mu$ ) and ( $X^{\prime}, d^{\prime}, \mu^{\prime}$ ) is defined to be

$$
\mathrm{d}_{\mathrm{GHP}}\left((X, d, \mu),\left(X^{\prime}, d^{\prime}, \mu^{\prime}\right)\right)=\frac{1}{2} \inf _{R, \nu} \max \left\{\operatorname{dis}(R), \nu\left(R^{c}\right)\right\} .
$$

Theorem. ( $\mathbb{M}, \mathrm{d}_{\mathrm{GHP}}$ ) is a complete separable metric space.
[S. Evans, J. Pitman and A. Winter, Rayleigh processes, real trees, and root growth with re-grafting, Probability Theory and Related Fields 134 (2006) pp.81-126.]
[R. Abraham, J.-F. Delmas and P. Hoscheit, A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces, Electronic Journal of Probability 18 (2013), no. 14.]

## The Brownian CRT

Definition. The Brownian continuum random tree is $\left(\mathcal{T}_{2 e}, d_{2 e}, \mu_{2 e}\right)$, where $e$ is a standard Brownian excursion.



## A planar ordering




Observe that the excursion comes with more information than the the tree: if $s<t$ and $\pi_{2 e}(s)$ and $\pi_{2 e}(t)$ are leaves, it is natural to think of $\pi_{2 e}(s)$ appearing to the left of $\pi_{2 e}(t)$ (c.f. Rémy's algorithm).

## Discrete trees as metric spaces

We want to think of ( $T_{n}, n \geq 1$ ) as metric spaces.

The vertices of $T_{n}$ (labelled and unlabelled) come equipped with a natural metric: the graph distance $d_{n}$.


We sometimes write $a T_{n}$ for the metric space $\left(T_{n}, a d_{n}\right)$ given by the vertices of $T_{n}$ with the graph distance scaled by $a$.

## Uniform measure



We will endow $T_{n}$ with $\mu_{n}$, the measure which puts mass $1 /(2 n)$ on each of the $2 n$ vertices.

## Convergence

Theorem. As $n \rightarrow \infty$,

$$
\left(T_{n}, \frac{d_{n}}{\sqrt{2 n}}, \mu_{n}\right) \rightarrow\left(\mathcal{T}_{2 e}, d_{2 e}, \mu_{2 e}\right) \quad \text { a.s. }
$$

with respect to the Gromov-Hausdorff-Prokhorov topology.
[P. Marchal, Constructing a sequence of random walks strongly converging to Brownian motion, Discrete Mathematics and Theoretical Computer Science, 2003, pp.181-190.]
[N. Curien \& B. Haas, The stable trees are nested, Probability Theory and Related Fields 157, 2013, pp.847-883.]

## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with $n$ leaves and lattice excursions with $2 n$ steps.

Start every excursion with a +1 step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is +1 if the vertex is not a leaf and -1 if the vertex is a leaf.


## Binary trees and lattice excursions

To go back the other way, it's easy to recover the tree:


## Binary trees and lattice excursions

To go back the other way, it's easy to recover the tree:


## Binary trees and lattice excursions



Since our trees are uniform, so are the lattice excursions. In other words, they are excursions of simple random walk away from 0 .

## Binary trees and lattice excursions



Since our trees are uniform, so are the lattice excursions. In other words, they are excursions of simple random walk away from 0 . So (at least in distribution), it's clear that, suitably rescaled, they should converge to a Brownian excursion.

## Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.


## Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.


## Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.


## Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.


## Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.


## Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.


## Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.


## Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.


## Binary trees and lattice excursions

Let $\left(E_{n}\right)_{n \geq 1}$ be the sequence of lattice excursions.
Theorem. (Marchal (2003))
As $n \rightarrow \infty$, we have

$$
\frac{1}{\sqrt{2 n}}\left(E_{n}(\lfloor 2 n t\rfloor), 0 \leq t \leq 1\right) \rightarrow(e(t), 0 \leq t \leq 1)
$$

uniformly on $[0,1]$, almost surely.

## Convergence of the trees

This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.


## Convergence of the trees

This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.


## Convergence of the trees

This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.


## Convergence of the trees

This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.


Write $H_{n}(k)$ for the distance from the root to the vertex visited at time $k$. Then

$$
H_{n}(k)=\left|\left\{0 \leq i \leq k-1: E_{n}(i)=\min _{i \leq j \leq k} E_{n}(k)\right\}\right| .
$$

## Convergence of the trees

This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.


Write $H_{n}(k)$ for the distance from the root to the vertex visited at time $k$. Then

$$
H_{n}(k)=\left|\left\{0 \leq i \leq k-1: E_{n}(i)=\min _{i \leq j \leq k} E_{n}(k)\right\}\right| .
$$

It turns out that $H_{n}(k) \approx 2 E_{n}(k)$.

## Convergence of the trees

Theorem. As $n \rightarrow \infty$,

$$
\frac{1}{\sqrt{2 n}}\left(H_{n}(\lfloor 2 n t\rfloor), 0 \leq t \leq 1\right) \rightarrow(2 e(t), 0 \leq t \leq 1)
$$

uniformly on $[0,1]$, almost surely.
[J.-F. Marckert \& A. Mokkadem, The depth first processes of Galton-Watson trees converge to the same Brownian excursion, Annals of Probability, 31(3), pp.1655-1678, 2003.]

## Convergence of the trees

Let's call the vertices be $v_{0}, v_{1}, \ldots, v_{2 n-1}$ in the order we visit them, where $v_{0}$ is the root.

## Convergence of the trees

Let's call the vertices be $v_{0}, v_{1}, \ldots, v_{2 n-1}$ in the order we visit them, where $v_{0}$ is the root.

By definition,

$$
d_{n}\left(v_{0}, v_{k}\right)=H_{n}(k)
$$

## Convergence of the trees

Let's call the vertices be $v_{0}, v_{1}, \ldots, v_{2 n-1}$ in the order we visit them, where $v_{0}$ is the root.

By definition,

$$
d_{n}\left(v_{0}, v_{k}\right)=H_{n}(k)
$$

More generally, for $0 \leq i<j \leq 2 n-1$, write $v_{i} \wedge v_{j}$ for the most recent common ancestor of $v_{i}$ and $v_{j}$ in the tree. Then

$$
d_{n}\left(v_{i}, v_{j}\right)=d_{n}\left(v_{0}, v_{i}\right)+d_{n}\left(v_{0}, v_{j}\right)-2 d_{n}\left(v_{0}, v_{i} \wedge v_{j}\right)
$$

## Convergence of the trees

$$
d_{n}\left(v_{0}, v_{i} \wedge v_{j}\right)= \begin{cases}\min _{i \leq k \leq j} H_{n}(k)-1 & \text { if } v_{i} \text { not an ancestor of } v_{j} \\ \min _{i \leq k \leq j} H_{n}(k)=H_{n}(i) & \text { if } v_{i} \text { an ancestor of } v_{j} .\end{cases}
$$



So

$$
\left|d_{n}\left(v_{0}, v_{i} \wedge v_{j}\right)-\min _{i \leq k \leq j} H_{n}(k)\right| \leq 1
$$

## A correspondence

Define a correspondence $R_{n}$ between $\left\{v_{0}, v_{1}, \ldots, v_{2 n-1}\right\}$ and $[0,1]$ by declaring $\left(v_{i}, s\right) \in R_{n}$ if $i=\lfloor 2 n s\rfloor$.

Endow $[0,1]$ with the pseudo-metric $d_{2 e}$. We will bound $\operatorname{dis}\left(R_{n}\right)$.

## A correspondence

Define a correspondence $R_{n}$ between $\left\{v_{0}, v_{1}, \ldots, v_{2 n-1}\right\}$ and $[0,1]$ by declaring $\left(v_{i}, s\right) \in R_{n}$ if $i=\lfloor 2 n s\rfloor$.

Endow $[0,1]$ with the pseudo-metric $d_{2 e}$. We will bound $\operatorname{dis}\left(R_{n}\right)$.
Suppose that $\left(v_{i}, s\right),\left(v_{j}, t\right) \in R_{n}$ with $s \leq t$. Then

$$
\begin{aligned}
& \left|d_{n}\left(v_{i}, v_{j}\right)-d_{2 e}(s, t)\right| \\
& \leq \left\lvert\, \frac{1}{\sqrt{2 n}}\left(H_{n}(\lfloor 2 n s\rfloor)+H_{n}(\lfloor 2 n t\rfloor)-2 \min _{s \leq u \leq t} H_{n}(\lfloor 2 n u\rfloor)\right)\right. \\
& \\
& \quad-\left(2 e(s)+2 e(t)-4 \min _{s \leq u \leq t} e(u)\right) \left\lvert\,+\frac{2}{\sqrt{2 n}} .\right.
\end{aligned}
$$

## A correspondence

Define a correspondence $R_{n}$ between $\left\{v_{0}, v_{1}, \ldots, v_{2 n-1}\right\}$ and $[0,1]$ by declaring $\left(v_{i}, s\right) \in R_{n}$ if $i=\lfloor 2 n s\rfloor$.

Endow $[0,1]$ with the pseudo-metric $d_{2 e}$. We will bound $\operatorname{dis}\left(R_{n}\right)$.
Suppose that $\left(v_{i}, s\right),\left(v_{j}, t\right) \in R_{n}$ with $s \leq t$. Then

$$
\begin{aligned}
& \left|d_{n}\left(v_{i}, v_{j}\right)-d_{2 e}(s, t)\right| \\
& \leq \left\lvert\, \frac{1}{\sqrt{2 n}}\left(H_{n}(\lfloor 2 n s\rfloor)+H_{n}(\lfloor 2 n t\rfloor)-2 \min _{s \leq u \leq t} H_{n}(\lfloor 2 n u\rfloor)\right)\right. \\
& \\
& \quad-\left(2 e(s)+2 e(t)-4 \min _{s \leq u \leq t} e(u)\right) \left\lvert\,+\frac{2}{\sqrt{2 n}} .\right.
\end{aligned}
$$

The right-hand side converges to 0 uniformly in $s, t \in[0,1]$. So

$$
\operatorname{dis}\left(R_{n}\right) \rightarrow 0 \quad \text { a.s. }
$$

## A coupling

Recall that $\mu_{n}$ is the measure which puts mass $1 /(2 n)$ on each of the vertices $v_{0}, v_{1}, \ldots, v_{2 n-1}$. Then we may couple $\mu_{n}$ and $\mu_{2 e}$ by taking $U \sim \mathrm{U}[0,1]$ and taking $\nu$ to be the law of the pair

$$
\left(v_{\lfloor 2 n U\rfloor}, \pi_{2 e}(U)\right)
$$

This is precisely the natural coupling $\nu_{n}$ induced by the correspondence $R_{n}$, and so $\nu_{n}\left(R_{n}^{c}\right)=0$.

## GHP convergence

But then

$$
\begin{aligned}
& d_{\mathrm{GHP}}\left(\left(T_{n}, \frac{d_{n}}{\sqrt{2 n}}, \mu_{n}\right),\left(\mathcal{T}_{2 e}, d_{2 e}, \mu_{2 e}\right)\right) \\
& \leq \frac{1}{2} \max \left\{\operatorname{dis}\left(R_{n}\right), \nu_{n}\left(R_{n}^{c}\right)\right\} \rightarrow 0
\end{aligned}
$$

almost surely as $n \rightarrow \infty$.

## 3. UNIVERSALITY

Key reference:

Jean-François Le Gall, Random trees and applications, Probability Surveys 2 (2005) pp.245-311.


## Branching processes

A Bienaymé-Galton-Watson (BGW) branching process $\left(Z_{n}\right)_{n \geq 0}$ describes the size of a population which evolves as follows:

- Start with a single individual.
- This individual has a number of children distributed according to the offspring distribution $p$, where $p(k)$ gives the probability of $k$ children, $k \geq 0$.
- Each child reproduces as an independent copy of the original individual.
$Z_{n}$ gives the number of individuals in generation $n$ (in particular, $Z_{0}=1$ ).


## BGW trees

A BGW tree is the family tree arising from a BGW branching process. We will think of this as a rooted ordered tree.

## BGW trees

A BGW tree is the family tree arising from a BGW branching process. We will think of this as a rooted ordered tree.

Consider the case where the offspring distribution $p$ is critical i.e.

$$
\sum_{k=1}^{\infty} k p(k)=1
$$

This ensures, in particular, that the resulting tree, $T$, is finite.

## Combinatorial trees in disguise

Let $T$ be a BGW tree with offspring distribution $p$ and total progeny $N$.

- If $p(0)=1 / 2$ and $p(2)=1 / 2$ then, conditional on $N=2 n-1$, the tree is uniform on the set of rooted (unplanted!) plane binary trees with $n$ leaves.
- If $p(k)=2^{-k-1}, k \geq 0$ (i.e. Geometric( $1 / 2$ ) offspring distribution) then conditional on $N=n$, the tree is uniform on the set of plane trees with $n$ vertices.
- If $p(k)=\frac{e^{-1}}{k!}, k \geq 0$ (i.e. Poisson(1) offspring distribution) then conditional on $N=n$, if we assign the vertices labels chosen uniformly at random from $\{1,2, \ldots, n\}$ and then forget the ordering and the root, we obtain a labelled tree $\tilde{T}$ which is uniform on the set of possibilities.

The last example will be particularly important in Lecture 2.

## A universal scaling limit

Let $T_{n}$ be the family tree of a BGW process with critical offspring distribution of variance $\sigma^{2} \in(0, \infty)$, conditioned to have total progeny $n$. Let $d_{n}$ be the graph distance on $T_{n}$ and let $\mu_{n}$ be the uniform measure on the vertices.

Theorem. (Aldous (1993), Le Gall (2005))
As $n \rightarrow \infty$,

$$
\left(T_{n}, \frac{\sigma}{\sqrt{n}} d_{n}, \mu_{n}\right) \xrightarrow{d}\left(\mathcal{T}_{2 e}, d_{2 e}, \mu_{2 e}\right),
$$

where convergence is in the Gromov-Hausdorff-Prokhorov sense.

## Two ways of encoding a tree

As we have seen, it is convenient to encode our trees in terms of discrete functions which are easier to manipulate.

We will do this is two different ways:

- the height function
- the depth-first walk (or Łukasiewicz path).


## Height function

Suppose that our tree has $n$ vertices. Let them be $v_{0}, v_{1}, \ldots, v_{n-1}$, listed in depth-first order.

## Height function

Suppose that our tree has $n$ vertices. Let them be $v_{0}, v_{1}, \ldots, v_{n-1}$, listed in depth-first order.

Then the height function is defined by

$$
H(k)=d_{n}\left(v_{0}, v_{k}\right), \quad 0 \leq k \leq n-1 .
$$

## Height function




## Height function




## Height function




## Height function



## Height function



## Height function



## Height function




## Height function




We can easily recover the tree from its height function.

## Depth-first walk

Let $c(v)$ be the number of children of $v$, and that $v_{0}, v_{1}, \ldots, v_{n-1}$ is a list of the vertices in depth-first order.

## Depth-first walk

Let $c(v)$ be the number of children of $v$, and that $v_{0}, v_{1}, \ldots, v_{n-1}$ is a list of the vertices in depth-first order.

Define

$$
\begin{aligned}
& X(0)=0 \\
& X(i)=\sum_{j=0}^{i-1}\left(c\left(v_{j}\right)-1\right), \text { for } 1 \leq i \leq n .
\end{aligned}
$$

## Depth-first walk

Let $c(v)$ be the number of children of $v$, and that $v_{0}, v_{1}, \ldots, v_{n-1}$ is a list of the vertices in depth-first order.

Define

$$
\begin{aligned}
& X(0)=0 \\
& X(i)=\sum_{j=0}^{i-1}\left(c\left(v_{j}\right)-1\right), \text { for } 1 \leq i \leq n .
\end{aligned}
$$

In other words,

$$
X(i+1)=X(i)+c\left(v_{i}\right)-1, \quad 0 \leq i \leq n-1
$$

## Depth-first walk



## Depth-first walk




## Depth-first walk



## Depth-first walk




## Depth-first walk




## Depth-first walk




## Depth-first walk




## Depth-first walk




## Depth-first walk




Proposition. For $0 \leq i \leq n-1$,

$$
H(i)=\#\left\{0 \leq j \leq i-1: X(j)=\min _{j \leq k \leq i} X(k)\right\}
$$

The depth-first walk of a BGW tree is a stopped random walk

Recall that $p$ is a distribution on $\mathbb{Z}_{+}$such that $\sum_{k=1}^{\infty} k p(k)=1$.

Recall that $p$ is a distribution on $\mathbb{Z}_{+}$such that $\sum_{k=1}^{\infty} k p(k)=1$.
Proposition. Let $(R(k), k \geq 0)$ be a random walk with initial value 0 and step distribution $\nu(k)=p(k+1), k \geq-1$. Set

$$
M=\inf \{k \geq 0: R(k)=-1\}
$$

Recall that $p$ is a distribution on $\mathbb{Z}_{+}$such that $\sum_{k=1}^{\infty} k p(k)=1$.
Proposition. Let $(R(k), k \geq 0)$ be a random walk with initial value 0 and step distribution $\nu(k)=p(k+1), k \geq-1$. Set

$$
M=\inf \{k \geq 0: R(k)=-1\}
$$

Now suppose that $T$ is a BGW tree with offspring distribution $p$ and total progeny $N$. Then

$$
(X(k), 0 \leq k \leq N) \stackrel{d}{=}(R(k), 0 \leq k \leq M)
$$

Recall that $p$ is a distribution on $\mathbb{Z}_{+}$such that $\sum_{k=1}^{\infty} k p(k)=1$.
Proposition. Let $(R(k), k \geq 0)$ be a random walk with initial value 0 and step distribution $\nu(k)=p(k+1), k \geq-1$. Set

$$
M=\inf \{k \geq 0: R(k)=-1\}
$$

Now suppose that $T$ is a BGW tree with offspring distribution $p$ and total progeny $N$. Then

$$
(X(k), 0 \leq k \leq N) \stackrel{d}{=}(R(k), 0 \leq k \leq M)
$$

[Careful proof: see Le Gall (2005).]

## BGW trees conditioned on their total progeny

Suppose now that we have offspring variance $\sigma^{2}:=\sum_{k=1}^{\infty}(k-1)^{2} p(k) \in(0, \infty)$.

The depth-first walk $X$ is a random walk with step mean 0 and variance $\sigma^{2}$, stopped at the first time it hits -1 . The underlying random walk has a Brownian motion as its scaling limit, by Donsker's theorem.

The total progeny $N$ is equal to $\inf \{k \geq 0: X(k)=-1\}$. We want to condition on the event $\{N=n\}$.

## BGW trees conditioned on their total progeny

Suppose now that we have offspring variance $\sigma^{2}:=\sum_{k=1}^{\infty}(k-1)^{2} p(k) \in(0, \infty)$.

The depth-first walk $X$ is a random walk with step mean 0 and variance $\sigma^{2}$, stopped at the first time it hits -1 . The underlying random walk has a Brownian motion as its scaling limit, by Donsker's theorem.

The total progeny $N$ is equal to $\inf \{k \geq 0: X(k)=-1\}$. We want to condition on the event $\{N=n\}$.

Standing assumption: $\mathbb{P}(N=n)>0$ for all $n$ sufficiently large.

## BGW trees conditioned on their total progeny

Write $\left(X_{n}(k), 0 \leq k \leq n\right)$ for the depth-first walk conditioned on $\{N=n\}$. Then there is a conditional version of Donsker's theorem.

Theorem. As $n \rightarrow \infty$,

$$
\frac{1}{\sigma \sqrt{n}}\left(X_{n}(\lfloor n t\rfloor), 0 \leq t \leq 1\right) \xrightarrow{d}(e(t), 0 \leq t \leq 1),
$$

where ( $e(t), 0 \leq t \leq 1$ ) is a standard Brownian excursion.
[W.D. Kaigh, An invariance principle for random walk conditioned by a late return to zero, Annals of Probability 4, 1976, pp.115-121.]

## Height process

Let $\left(H_{n}(i), 0 \leq i \leq n\right)$ be the height process of a critical BGW tree with offspring variance $\sigma^{2} \in(0, \infty)$, conditioned to have total progeny $n$, so that

$$
H_{n}(i)=\#\left\{0 \leq j \leq i-1: X_{n}(j)=\min _{j \leq k \leq i} X_{n}(k)\right\}
$$

Theorem. As $n \rightarrow \infty$,

$$
\left.\frac{\sigma}{\sqrt{n}}\left(H_{n}(\lfloor n t\rfloor), 0 \leq t \leq 1\right) \xrightarrow{d} 2(e(t), 0 \leq t \leq 1)\right)
$$

where $(e(t), 0 \leq t \leq 1)$ is a standard Brownian excursion.

## Convergence to the Brownian CRT

The convergence

$$
\left(T_{n}, \frac{\sigma}{\sqrt{n}} d_{n}, \mu_{n}\right) \xrightarrow{d}\left(\mathcal{T}_{2 e}, d_{2 e}, \mu_{2 e}\right),
$$

now follows by applying Skorokhod's theorem (in order to work on a probability space where the height process converges almost surely) and then using the same proof that we had in the case of binary trees.

## Universality

The universality class of the Brownian CRT is, in fact, even larger. Some other examples of trees (and graphs!) with the Brownian CRT as their scaling limit are:

- uniform unordered unlabelled rooted trees
- uniform unordered unlabelled unrooted trees
- critical multi-type BGW trees
- random trees with a prescribed degree sequence satisfying certain conditions
- random dissections
- random graphs from subcritical classes.

