Fragmentations and Random Real Trees

Bénédicte HAAS

Université Paris-Dauphine

YEP VI, Eindhoven 2009

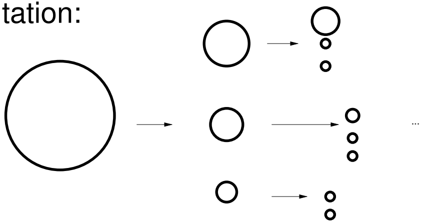
Outline

1. Self-similar fragmentations
2. Random real trees
3. Fragmentation trees
4. Scaling limits of consistent Markov branching models

Historical backgrounds

Natural phenomena involving fragmentation:

- degradation of polymer chains
- liquid droplet break up
- earthquakes
- grinding of rocks...



Probabilistic models

• 1941: Kolmogorov: discrete time process, all particles evolve according to the same dynamic

• 1961: Filippov: splitting rates depend on the masses of particles

$\updownarrow$ Brennan-Durrett 1986-87, Aldous-Pitman 1998

• 2001: Bertoin's self-similar fragmentations

Fragmentation equation

describes the evolution of the average number per unit volume of particles with a given mass (Melzak 1957)

extensive literature in physics and mathematics (since the 80ths-90ths)

Self-similar fragmentations

Random models for the evolution of particles (or objects) that split as time goes on.

Dynamic of break-up based on two main rules:

- a branching property: different particles evolve independently
- a self-similar property: the rate at which a particle splits is proportional to a power of its mass.

Outline:

1. Mass fragmentations, characterization
2. The tagged fragment
3. Formation of dust and extinction

Bibliography

- J. BERTOIN, Homogeneous fragmentation processes, *PTRF*, 2001
- J. BERTOIN, Self-similar fragmentations, *Ann. IHP*, 2002
- J. BERTOIN, The asymptotic behavior of fragmentation processes, *J. Eur. Math. Soc.*, 2003
- J. BERTOIN, Random fragmentation and coagulation processes, Cambridge UP, 2006
- J. BERESTYCKI, Ranked fragmentations, *ESAIM Prob. Stat.*, 2002

Mass fragmentations

Focus on masses of particles (not on their form, spatial position, etc.).

Starting from a particle with mass 1, a natural state space is:

$$\mathcal{S}^{\downarrow} := \left\{ \mathbf{s} = (s_1 \geq s_2 \geq ... \geq 0) : \sum_{i \geq 1} s_i \leq 1 \right\},$$

endowed with the uniform distance $d_{\mathcal{S}^{\downarrow}}(\mathbf{s}, \mathbf{s}') := \max_{i \geq 1} |s_i - s'_i|$.

Triangular extraction $\rightarrow (\mathcal{S}^{\downarrow}, d_{\mathcal{S}^{\downarrow}})$ is compact.

Includes cases where some mass is reduced to a *dust of* 0*-mass particles:*

$$\text{total mass of dust} = 1 - \sum_{i \geq 1} s_i.$$

Mass fragmentation

$(F(t), t \geq 0)$: càdlàg $\mathcal{S}^{\downarrow}$-valued Markov process.

$\mathbb{P}_m$: law of F starting from $(m, 0, ...)$, $0 \leq m \leq 1$.

Definition

F is a self-similar process with index $\alpha \in \mathbb{R}$ if

1. (self-similarity) for all $m \in [0, 1]$,

   the distribution of $(mF(m^{\alpha}t), t \geq 0)$ under $\mathbb{P}_1$ is $\mathbb{P}_m$.

2. (branching property) $\forall t_0 \geq 0$, conditional on $F(t_0) = (s_j)_{j\geq 1}$,

$$(F(t_0 + t))_{t\geq 0} \overset{\text{law}}{=} (\{F_i^{(j)}(t), i, j \geq 1\}^{\downarrow})_{t\geq 0}$$

   where $F^{(j)} \sim \mathbb{P}_{s_j}$ for all j, and the $F^{(j)}$s are independent.

$\{u_i, i \geq 1\}^{\downarrow}$: decreasing reordering of the u_is, $i \geq 1$

BERESTYCKI 02: F is Fellerian

Mass fragmentation

In the following, we work under $\mathbb{P}_1$

Influence of α on the speed of fragmentation:

- $\alpha = 0 \Rightarrow$ all masses split at the same speed (the fragmentation is said *homogeneous*);
- $\alpha > 0 \Rightarrow$ large masses split faster and so the break-up rates decrease as time goes on;
- $\alpha < 0 \Rightarrow$ small masses split faster and the break-up rates increase as time goes on.

Characterization

Theorem (Bertoin 02, Berestycki 02)

The law of a self-similar mass fragmentation F is characterized by 3 parameters:

1. the index of self-similarity $\alpha \in \mathbb{R}$
2. an *erosion coefficient* $c \geq 0$: codes a continuous melt of the fragments
3. a *dislocation measure* ν (measure on $\mathcal{S}^{\downarrow}$ such that $\int_{\mathcal{S}^{\downarrow}}(1-s_1)\nu(\mathrm{d}\mathbf{s}) < \infty$ and $\nu((1,0,...)) = 0$): discribes the sudden dislocations of fragments, i.e. the jumps of F.

T_{in} : “first dislocation time" of the process F, i.e.

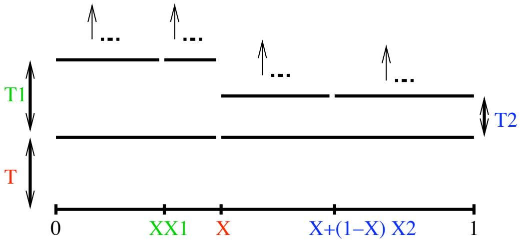
$$T_{\text{in}} := \inf\{t \geq 0 : F(t) \neq (1,0,...)\}.$$

The Blumenthal $0-1$ law $\Rightarrow$ either $T_{\text{in}} > 0$ a.s. or $T_{\text{in}} = 0$ a.s.

Examples.

Ex. 1. X r.v. with values in $(0,1)$ and $T \sim \exp(\lambda)$, independent of X

- at time T, $(0,1)$ splits in $(0,X)$ and $(X,1)$
- (X_1,T_1), (X_2,T_2) independent copies of (X,T), indep. of (X,T)
  - at time $T+T_1$, $(0,X) \to (0,XX_1) \cup (XX_1,X)$
  - at time T_2, $(X,1) \to (X, X+(1-X)X_2) \cup (X+(1-X)X_2, 1)$
- and so on...



$F(t)$: decreasing sequence of lengths of intervals present at time t.

F fragmentation with parameters $\alpha = 0$, $c = 0$, and

$$\int_{\mathcal{S}^{\downarrow}} f(\mathbf{s})\nu(\mathrm{d}\mathbf{s}) = \lambda \mathbb{E}[f(\max(X,1-X),\min(X,1-X),0,...)].$$

Examples.

Ex. 2. Poissonian rain

N standard Poisson process and $(U_i, i \geq 1)$ i.i.d. uniform on $(0,1)$, independent of N:

U_i "falls" in $(0,1)$ at the i^{th} jump time of N.

$F(t)$: $\downarrow$ lengths of connected components of $(0,1)\backslash\{U_1, U_2 ..., U_{N(t)}\}$ (=$(1,0,..)$ if $N(t)=0$).

Fragmentation with parameters $\alpha = 1$, $c = 0$ and

$$\int_{\mathcal{S}^{\downarrow}} f(\mathbf{s})\nu(\mathrm{d}\mathbf{s}) = \mathbb{E}\left[f(\max(U, 1-U), \min(U, 1-U), 0, ...)\right].$$

Cases where $T_{\text{in}} > 0$

Markov property of $F \Rightarrow T_{\text{in}} \sim \exp(\lambda)$ for a $\lambda > 0$

ν defined by $\nu(B) = \lambda \mathbb{P}(F(T_{\text{in}}) \in B)$ $\forall$ borelian B of $\mathcal{S}^{\downarrow}$

A mass m then evolves as follows:

(i) it has a life-time $\sim \exp(m^{\alpha}\lambda)$

(ii) when it dies, it splits in masses $mS_1, mS_2, ...$ where

$$(S_1, S_2, ...) \sim \lambda^{-1}\nu$$

and this sequence is independent of the life-time of the mass.

Here we set: $c = 0$.

Reciprocally, given $\alpha \in \mathbb{R}$ an a finite dislocation measure ν,

- start with a mass $m = 1$ that follows the dynamic (i)-(ii) above
- then, conditionally on $S_1, S_2, ...$, these masses evolves independently, according to the dynamic (i)-(ii). And so on.

$\Rightarrow$ gives a $(\alpha, c = 0, \nu)$-fragmentation.

General cases

1) Homogeneous fragmentations (BERTOIN 01, BERESTYCKI 02)

• F : $(0,0,\nu)$-fragmentation: ν infinite $\leftrightarrow$ fragments “split immediately".

Formalization via Poisson point processes (PPP):

$\exists$ $((\Delta(t),k(t)),t\geq 0)$ a $\mathcal{S}^{\downarrow}\times\mathbb{N}$-valued PPP with intensity $\nu\otimes\sharp$ (where $\sharp(A)=\mathrm{Card}(A)$, for $A\subset\mathbb{N}$), such that

jump times of F $=$ times of occurrence of the PPP

For such a time t,

- $F_{k(t)}(t-)$ is replaced by $F_{k(t)}(t-)\Delta(t)$
- $F(t)$: $\downarrow$ reordering of $F_i(t-), i\neq k(t)$, and $F_{k(t)}(t-)\Delta_i(t)$, $i\geq 0$.

Ex. $F(t-)=(1/2,1/3,1/6,0,..)$, $k(t)=2$ and $\Delta(t)=(3/4,1/4,0,..)$
$\Rightarrow F_{k(t)}(t-)\Delta(t)=(1/4,1/12,0,..)$ and $F(t)=(1/2,1/4,1/6,1/12,..)$.

Remark: ν finite: recovers the exponential life-times description.

General cases

• $(0, c, \nu)$-fragmentations: *c* rate of erosion, *deterministic phenomenon*

For all $(0, 0, \nu)$-fragmentation F and all $c \geq 0$:

$$(\exp(-ct)F(t), t \geq 0) \text{ is also a homogeneous fragmentation.}$$

Each homogeneous fragmentation can be contructed like this: its law is characterized by a pair (c, ν).

Remark. Similar to the Lévy-Itô decomposition of a subordinator.

2) Self-similar fragmentations (BERTOIN 02, BERESTYCKI 02)
α fixed:

one to one correspondence between the laws of α-self-similar fragmentations and those of homogeneous fragmentations

Idea: use interval representations + time-changes.

Interval fragmentations

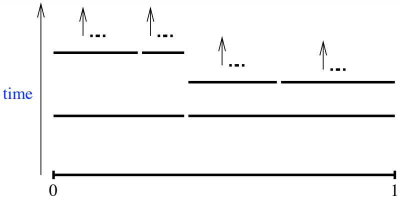
$\mathcal{O}_{(0,1)}$: set of open subsets of $(0,1)$

(topology: $d_{\mathcal{O}_{(0,1)}}(O,O') = d_{\mathrm{H}}([0,1]\backslash O, [0,1]\backslash O') \to$ compact)

Definition

A $\mathcal{O}_{(0,1)}$-valued process $(I(t), t \geq 0)$ is an interval fragmentation if

$$I(t) \subset I(s) \text{ whenever } 0 \leq s \leq t.$$



Evolution of $I(t)$, $t \geq 0$

Theorem (Bertoin 02, Berestycki 02)

Each (mass) α-self-similar fragmentation can be constructed from a α-self-similar interval fragmentation, by ranking the lengths of its interval components in the decreasing order.

Self-similar interval fragmentations

I: càdlàg Markov interval fragmentation.

For, $O \in \mathcal{O}_{(0,1)}$, $\mathbf{P}_O$: law of I starting from O.

I is α-self-similar if:

- branching property: if $O = \cup_{i \in \mathbb{N}} J_i$ (interval decomposition) and if $I^{(i)}$ i.i.d with respective laws $\mathbf{P}_{J_i}$, $i \geq 1$, then
$$\left(\cup_{i \in \mathbb{N}} I^{(i)}(t), \quad t \geq 0\right) \text{ has law } \mathbf{P}_O$$
- self-similarity: for all $0 \leq a < b \leq 1$:
$$\text{the law of } (g_{(a,b)}(I((b-a)^\alpha t), t \geq 0) \text{ under } \mathbf{P}_{(0,1)} \text{ is } \mathbf{P}_{(a,b)}$$
where $g_{(a,b)}(x) = a + (b-a)x$.

Time-changes

$(I^{(\alpha)}(t), t \geq 0)$: interval-fragmentation with index α

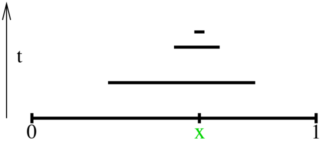


Figure: Evolution of $I_x(t), t \geq 0$

For $x \in (0,1)$, $I_x^{(\alpha)}(t)$ interval component of $I^{(\alpha)}(t)$ containing x $(= \emptyset$ if $x \notin I^{(\alpha)}(t))$

$|I_x^{(\alpha)}(t)|$: lentgh of $I_x^{(\alpha)}(t)$

Time-change depending on the past of the fragment:

$$T_x^{(\beta)}(t) = \inf\left\{u \geq 0 : \int_0^u \left(|I_x^{(\alpha)}(r)|\right)^{-\beta} \mathrm{d}r > t\right\}$$

Remark: $\beta \leq 0 \Leftrightarrow T_x^{(\beta)}(t) \geq t \Leftrightarrow$ time acceleration

Proposition (Bertoin 02)

$$J_x(t) := I_x^{(\alpha)}(T_x^{(\beta)}(t)), \quad t \geq 0, \quad x \in (0,1)$$

defines an interval fragmentation with index $\alpha + \beta$.

Remarks

1. If $c > 0$ or $\nu(\mathcal{S}^{\downarrow}) > 0$ $(\forall \alpha \in \mathbb{R})$, $F(t) \to (0, 0, ...)$ as $t \to 0$

   Proof:

   - when $\alpha = 0$ and $c = 0$, use the PPP construction

   - when $\alpha \neq 0$, use time-changes and construction via a homogeneous fragmentation

2. Connection with historical models

   - Kolmogorov: $\alpha = 0$, ν finite, $\nu(s_i > 0 \forall i) = 0$, $c = 0$

   - Filippov: $\alpha \in \mathbb{R}$, ν finite, $c = 0$ (and also non-self-similar frag.)

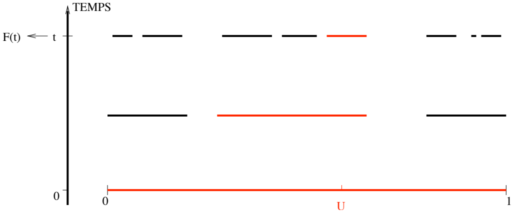
   - Brennan-Durrett: $\alpha \in \mathbb{R}$, $\nu = \delta_{(X,1-X,0,..)}$, X r.v. on [1/2,1), $c = 0$

   - Aldous-Pitman: $\alpha = 1/2$, ν infinite (see later), $c = 0$

The tagged fragment

The structure of a fragmentation is complex ▶ focus on one particle tagged at random.

- $(I(t), t \geq 0)$: interval representation of the fragmentation
- tagged point: U uniformly distributed on $(0,1)$, independent of I.



Mass of the tagged fragment at time t:

$$\lambda_{\text{tag}}(t) = |I_U(t)|$$

The tagged fragment: homogeneous cases

F : $(0, c, \nu)$-fragmentation. The PPP construction $\Rightarrow$

$$\xi(t) := -\ln(\lambda_{\mathrm{tag}}(t)), \quad t \geq 0,$$

is a subordinator (increasing process with independent and stationary increments).

The law of ξ is characterized by its Laplace exponent:

$$\mathbb{E}\left[\exp(-q\xi_t) = \exp(-t\phi(q))\right], \quad q \geq 0$$

where

$$\phi(q) = c(q+1) + \int_{\mathcal{S}^{\downarrow}} \left(1 - \sum_{i \geq 1} s_i^{q+1} \nu(\mathrm{d}\mathbf{s})\right), \tag{1}$$

$$= c(q+1) + \int_0^{\infty} (1 - \exp(-qx)) L(\mathrm{d}x),$$

i.e. $L(\mathrm{d}x) = \exp(-x) \sum_i \nu(-\ln(s_i) \in \mathrm{d}x) \mathbf{1}_{0<x<\infty}$
$+ \int_{\mathcal{S}^{\downarrow}} (1 - \sum_i s_i) \nu(\mathrm{d}\mathbf{s}) \delta_{\infty}(\mathrm{d}x).$

The tagged fragment: self-similar cases

$F : (\alpha, c, \nu)$ fragmentation. Time-changes + homogeneous case $\Rightarrow$

$$(\lambda_{\mathrm{tag}}(t), t \geq 0) \overset{\text{law}}{=} (-\exp(-\xi_{\rho(t)}), t \geq 0)$$

where ξ is a subordinator with Laplace exponent (1) and

$$\rho(t) := \inf\left\{u \geq 0 : \int_0^u \exp(\alpha \xi_r)\mathrm{d}r > t\right\}$$

with $\inf\{\emptyset\} := \infty$.

(Time-changed) subordinators have been studied in great details:

tagging a fragment at random is a powerfull tool to obtain some information on the fragmentation.

The tagged fragment

Limit of this approach: insufficient to recover the whole fragmentation.

Ex.: $(0, 0, \nu_1)$ and $(0, 0, \nu_2)$ fragmentations, with

$$\nu_1(\mathrm{d}s) = \delta_{(\frac{1}{2},\frac{1}{2},0,..)} + \delta_{(\frac{1}{2},\frac{1}{4},\frac{1}{4},0,..)}; \quad \nu_2(\mathrm{d}s) = \frac{3}{2}\delta_{(\frac{1}{2},\frac{1}{2},0,..)} + \frac{1}{2}\delta_{(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4},0,..)}$$

have identically distributed tagged fragments, with Laplace exponent

$$\phi(q) = 2 - 3\left(\frac{1}{2}\right)^{q+1} - 2\left(\frac{1}{4}\right)^{q+1}, \quad q \geq 0.$$

However, if $\nu(s_1 + s_2 < 1) = 0$ (binary), ϕ characterizes ν, since

$$s_1 \geq 1/2 \text{ and } s_2 = 1 - s_1 \ \nu\text{−a.e.}$$

$$\Rightarrow \mathbf{1}_{0<x\leq \ln(2)} L(\mathrm{d}x) \text{ gives } \nu(s_1 \in \mathrm{d}x) \text{ and then } \nu.$$

Formation of dust and extinction

Definition

A fragmentation F *produces some dust* if

$$\sum_{i=1}^{\infty} F_i(t) < 1 \text{ for some } t > 0.$$

The mass of dust at time t is then equal to $1 - \sum_{i=1}^{\infty} F_i(t)$.

There are three ways to produce some dust:

1. instantaneously when a particle splits when $\nu(\sum_i s_i < 1) > 0$
2. continuously by erosion when $c > 0$
3. by intensive splitting of small particles when $\alpha < 0$.

We focus on this last phenomenon and suppose that

$$c = 0 \qquad \nu\left(\sum_{i \geq 1} s_i < 1\right) = 0.$$

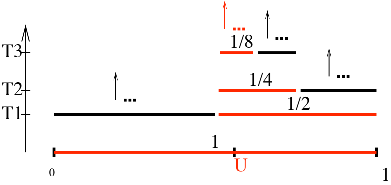
Formation of dust and extinction

Ex. $\alpha < 0$, $c = 0$ and $\nu(\mathrm{d}\mathbf{s}) = \delta_{(1/2,1/2,0,..)}$.

Follow a tagged particle:

$T_1 < T_2 < .. < T_n < ..$ its splitting times

$\lambda_{\mathrm{tag}}(T_n) = 2^{-n}$



$T_\infty := \lim_{n\to\infty} T_n$, $\lambda_{\mathrm{tag}}(T_\infty) = \lim_{n\to\infty} 2^{-n} = 0$.

But $T_1 \sim \exp(1)$ and $T_{n+1} - T_n \sim \exp(2^{-n\alpha})$ and these r.v. are independent,

$\Rightarrow$ $T_\infty \overset{\text{law}}{=} \sum_{n\geq 1} 2^{n\alpha}\mathbf{e}^{(i)}$, with $\mathbf{e}^{(i)}$s i.i.d, $\sim \exp(1)$

$\Rightarrow$ $\mathbb{E}[T_\infty] = \sum_{n\geq 1} 2^{n\alpha} < \infty$

$\Rightarrow$ $T_\infty < \infty$ a.s.

Hence, since $T_\infty > t \Leftrightarrow U \notin$ dust,

$$\mathbb{E}\left[\sum_{i\geq 1} F_i(t)\right] = \mathbb{E}\left[\mathbb{P}(T_\infty > t|F)\right] = \mathbb{P}(T_\infty > t) \to 0 \text{ as } t \to \infty.$$

Formation of dust and extinction

More generally, when $c = 0$ and $\nu\left(\sum_i s_i < 1\right) = 0$, the tagged fragment “dies" at time

$$D_{\text{tag}} := \inf\{t \geq 0 : \lambda_{\text{tag}}(t) = 0\} = \inf\{t : \rho(t) = \infty\} = \int_0^\infty \exp(\alpha\xi_r)\mathrm{d}r.$$

(such r.v. is called “exponential functional of a subordinator")

When $\alpha < 0$,

$$\mathbb{E}[D_{\text{tag}}] = \int_0^\infty \exp(-\phi(-\alpha)r)\mathrm{d}r = 1/\phi(-\alpha) < \infty$$

and more precisely, D_{tag} has positive exponential moments and

$$\mathbb{E}\left[D_{\text{tag}}^n\right] = n!/\prod_{i=1}^n \phi(-\alpha i), \quad n \in \mathbb{N}^*.$$

(see Carmona-Petit-Yor (1997) for this last result and more on exponential functional of subordinators)

Formation of dust and extinction

Theorem (BERTOIN 03)

F: (α, c, ν)-fragmentation such that $c = 0$, $\nu\left(\sum_i s_i < 1\right) = 0$. Then,

$$\text{some dust is produced} \Leftrightarrow \alpha < 0.$$

Moreover, when $\alpha < 0$,

$$\zeta := \inf\left\{t \geq 0 : \sum_{i=1}^{\infty} F_i(t) = 0\right\} < \infty \text{ a.s.}$$

Open problem. Find the distribution of ζ.

Proof. 1) FORMATION OF DUST.

$$\mathbb{E}\left[\sum_i F_i(t)\right] = \mathbb{E}\left[\mathbb{P}(D_{\text{tag}} > t | F)\right] = \begin{cases} \mathbb{P}(D_{\text{tag}} > t) = 1 \text{ when } \alpha \geq 0 \\ \mathbb{P}(D_{\text{tag}} > t) \underset{t\to\infty}{\to} 0 \text{ when } \alpha < 0 \end{cases}$$

Formation of dust and extinction

2) EXTINCTION: $\alpha < 0$, use time-changes:
$I^{(\alpha)}$ interval fragmentation, $I^{(0)}$ corresponding homogeneous frag.

$$I_x^{(\alpha)}(t) = I_x^{(0)}(T_x^{(\alpha)}(t)) \quad \forall x \in (0,1),\, t \geq 0$$

Lemma (BERTOIN 03)

$\exists c > 0$, a random $C < \infty$ a.s.: $I_x^{(0)}(t) \leq C\exp(-ct), \forall x \in (0,1), t \geq 0.$

$\Rightarrow \int_0^\infty |I_x^{(0)}(r)|^{-\alpha} \mathrm{d}r \leq C^{-\alpha}/c|\alpha|$

$\Rightarrow T_x^{(\alpha)}(C^{-\alpha}/c|\alpha|) = \infty$

$\Rightarrow I_x^{(\alpha)}(C^{-\alpha}/c|\alpha|) = 0, \quad \forall x \in (0,1).$ □

Sketch of proof of the lemma: $F^{(0)}$ mass frag. related to $I^{(0)}$. $\forall \gamma \geq 1$:
$\exp(\phi(\gamma-1)t)(I_x^{(0)}(t))^\gamma \leq \exp(\phi(\gamma-1)t)\sum_i(F_i^{(0)}(t))^\gamma := M_\gamma(t)$

But: $\mathbb{E}\left[\sum_i(F_i^{(0)}(t))^\gamma\right] = E[\lambda_{\mathrm{tag}}(t)^{\gamma-1}] = \exp(\phi(\gamma-1)t)$
+ $\alpha = 0$ + branching property $\Rightarrow M_\gamma$: martingale.

Moreover, $M_\gamma(t) \geq 0, \forall t$. Hence $\lim_{t\to\infty} M_\gamma(t)$ exists and is finite a.s. □

Formation of dust and extinction

Proposition (HAAS SPA 03)

When $c=0$, $\nu(\sum_i s_i < 1)=0$ and $\alpha<0$, $\exists C>0$ such that

$$\mathbb{P}(\zeta > t) \leq \exp(-Ct), \forall t \geq 0.$$

More precise results are available. Ex. $\phi(t) \underset{\infty}{\sim} Ct^{-\beta}$, $\beta \in (0,1)$,

$$\Rightarrow \exp(-Bt^{1/(1-\beta)}) \leq \mathbb{P}(\zeta > t) \leq \exp(-At^{1/(1-\beta)}).$$

- $\alpha = -1$**: critical index**

Proposition (BERTOIN 03)

When $\alpha \leq -1$, for all $t \geq 0$,

$$\mathrm{Card}\{i \geq 1 : F_i(t) > 0\} < \infty \quad a.s.$$

But: there may exist *random time t* s.t. $\mathrm{Card}\{i \geq 1 : F_i(t) > 0\} = \infty$.

When $-1 < \alpha < 0$, under some mild assumptions on ν, for a.e. $t > 0$,

$$\mathrm{Card}\{i \geq 1 : F_i(t) > 0\} = \infty \text{ a.s.}$$

Complements

1. *large-time asymptotics*: when $\alpha \geq 0$ (BERESTYCKI 03, BERTOIN 03, BERTOIN-GNEDIN 04, BERTOIN-ROUAULT 05, BRENNAN-DURRETT 86-87, FILIPPOV 61, KRELL 08), when $\alpha < 0$ (FILIPPOV 61, BERTOIN 03, HAAS 03, GOLDSCHMIDT-HAAS 09)
2. *duality with certain models of coalescence*: (ALDOUS-PITMAN 98, BERTOIN 04, BERTOIN-GOLDSCHMIDT 04, PITMAN 99)
3. *extensions of the model*: non-self-similar fragmentations, non-conservative fragmentations, fragmentations inhomogeneous in time, ordered fragmentations, coagulation-fragmentation processes, fragmentations with immigration, etc.

Random real trees

1. Real trees, Gromov-Hausdorff distance
2. Continuum Random Trees
3. Ex.: Brownian CRT, stables trees and associated fragmentations

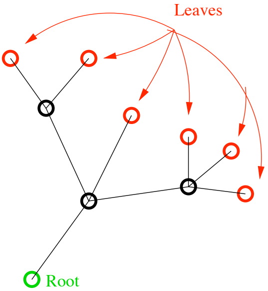
Random real trees: bibliography

- Aldous's Continuum Random Trees
  - D. ALDOUS. The continuum random tree I, *Ann. Probab.*, 1991
  - D. ALDOUS. Continuum Random Trees II. An overview. *London Math. Soc. Lecture Note Ser.*, 1991
  - D. ALDOUS. Continuum Random Trees III, *Ann. Probab.*, 1993
- Real trees and Random real trees
  - M. GROMOV. Metric structure for Riemannian and non-Riemannian spaces. *Progress in Mathematics 152*, 1999
  - S. EVANS. J. PITMAN, A. WINTER. Rayleight processes, real trees, and root growth with re-grafting. *PTRF*, 2006
  - S. EVANS A. WINTER. Subtree prune and regraft: a reversible real tree-valued Markov process. *Ann. Probab.*, 2006
  - S. EVANS. Probability and Real Trees. *Saint-Flour 2005*, 2008
  - A. GREVEN, P. PFAFFELHUBER, A. WINTER. Convergence in distribution of random metric measure spaces: (Lambda-coalescent measure trees) to appear in PTRF

Random real trees: bibliography

- Real trees and Lévy trees
  - T. DUQUESNE, J.F.-LE GALL. Random Trees, Lévy processes and spatial branching processes. *Astérisque*, 2002
  - T. DUQUESNE, J.F.-LE GALL. Probabilistic and fractal aspects of Lévy trees. *PTRF*, 2005
  - J.F.-LE GALL. Random trees and applications. *Probab. Survey*, 2005

Discrete trees



Discrete tree: connected acyclic finite graph (with delabeled vertices)

Our trees are *rooted*: a vertex is distinguished as the *root*.

Degree of a vertex: number of neighbors

The 1-degree vertices are called the *leaves* (except the root, if of degree 1).

Edge-lengths: each edge of the tree is equipped with a length> 0.

For a tree T, aT $=$same tree with all edge-lengths multiplied by $a > 0$

When nothing is specified: the edge-lengths are all equal to 1.

Real trees: metric space with the "tree" property: $\exists!$ path $x \to y$

Definition

A $\mathbb{R}$-*tree* $(\mathcal{T}, d)$ is a complete metric space such that, $\forall x, y \in \mathcal{T}$,

1. $\exists!$ isometry $\varphi_{x,y} : [0, d(x,y)] \to \mathcal{T}$ such that $\varphi_{x,y}(0) = x$ and $\varphi_{x,y}(d(x,y)) = y$
2. for every injective continuous map $c : [0,1] \to \mathcal{T}$ with $c(0) = x$, $c(1) = y$, one has $c([0,1]) = \varphi_{x,y}([0, d(x,y)])$.

Remark: (1) is satisfied by $\mathbb{R}^d$ with its usual metric.

- *vertices*: elements of $\mathcal{T}$
- for us, $\mathcal{T}$ is *rooted*: a vertex ρ is distinguished as the root
- $d(\rho, x) =$ *height* of x
- *leaf*: vertex x such that $\mathcal{T} \backslash \{x\}$ is connected
- $\mathcal{L}(\mathcal{T})$=set of leaves of $\mathcal{T}$
- $\mathcal{S}(\mathcal{T}) := \mathcal{T} \backslash \mathcal{L}(\mathcal{T})$: *squeletton* of the tree
- Degree of a vertex x: number of connected components of $\mathcal{T} \backslash \{x\}$.

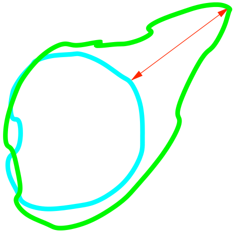
Gromov-Hausdorff distance

Goal: measure how far two metric spaces are from being isometric

Hausdorff distance: C, C' closed, non-empty, subsets of a metric space (E, d_E)

$$d_{\mathrm{H}}(C, C') = \inf\{\varepsilon > 0 : C \subset C'_\varepsilon \text{ and } C' \subset C_\varepsilon\}$$

where for $A \subset E$, $A_\varepsilon := \{x \in E : d_E(x, A) < \varepsilon\}$.



Gromov-Hausdorff distance:

Gromov-Hausdorff distance

Definition (Gromov-Hausdorff distance)

Let (X, x) and (Y, y) be pointed compact metric spaces,

$$d_{\mathrm{GH}}((X,x),(Y,y)) := \inf\left(d_{\mathrm{H}}(\varphi_1(X),\varphi_2(Y)) \vee d_Z(\varphi_1(x),(\varphi_2(y))\right)$$

where the infimum is over all isometric embeddings $\varphi_1 : X \hookrightarrow Z$ and $\varphi_2 : Y \hookrightarrow Z$ in a common metric space (Z, d_Z).

Two compact rooted $\mathbb{R}$-trees are *equivalent* if $\exists$ an isometry preserving the root mapping one tree onto the other.

$\mathbb{T}$: set of equivalence classes

Theorem (Evans, Pitman, Winter 06)

$(\mathbb{T}, d_{GH})$ is a Polish space (complete, separable).

Continuum Random trees (Aldous CRT III 93)

Definition (continuum tree)

Pair $(\mathcal{T}, \mu)$, $\mathcal{T}$ real tree and μ probability on the Borel σ-field $\mathcal{B}(\mathcal{T})$:

1. $\mu(\mathcal{L}(\mathcal{T})) = 1$
2. μ has no atom
3. $\forall x \in \mathcal{S}(\mathcal{T})$, $\mu(\{y : x \in \text{ path: root } \to y\} \backslash \{x\}) > 0$.

$\Rightarrow \mathcal{L}(\mathcal{T})$ uncountable, without isolated points.

How defining “continuum random trees" ? (which σ-algebra ?)

► use Evans-Winter Gromov-Hausdorff-Prokhorov distance: for (X, μ) and (Y, ν) compact metric spaces equipped with probability measures,

$$d_{\mathrm{GHP}}((X,\mu),(Y,\nu)) := \inf\left(d_{\mathrm{H}}(\varphi_1(X),\varphi_2(Y)) \vee d_P(\mu\circ\varphi_1^{-1},\nu\circ\varphi_2^{-1})\right)$$

where the infimum is over all isometric embeddings in a common metric space and d_P denotes the Prokhorov distance

$$d_P(m,m') = \inf\{\varepsilon > 0 : m(B) \leq m'(B_\varepsilon)+\varepsilon; m'(B) \leq m(B_\varepsilon)+\varepsilon,\ \forall B \text{ borelian}\}$$

Continuum Random trees (Aldous CRT III 93)

- Aldous' approach: trees embedded in (ℓ_1, d_{ℓ_1}):

$\ell_1 = \{\mathbf{x} = (x_1, x_2, ..), x_i \in \mathbb{R} : \sum_{i \geq 1} |x_i| < \infty\}$

$d_{\ell_1}(\mathbf{x}, \mathbf{x}') = \sum_{i \geq 1} |x_i - x'_i|$

Continuum tree $(\mathcal{T}, \mu)$:

1. $\mathcal{T}$: $\mathbb{R}$-tree, closed subset of ℓ_1, containing 0 (0=root)
2. μ probability measure on ℓ_1 satisfying the 3 points of the previous definition.

Hausdorff metric for closed sets + metric inducing weak convergence for proba. measures $\rightarrow$ product σ-algebra

$\forall$ CRT $(\mathcal{T}, \mu)$, $\mathcal{T}$: projective limit of its *marginals* $(R(k), k \geq 1)$:

- conditional on $(\mathcal{T}, \mu)$, let $(L_1, L_2, ...)$ i.i.d $\sim \mu$
- $R(k)$ subtree generated by the root + leaves $L_1, ..., L_k$, i.e.

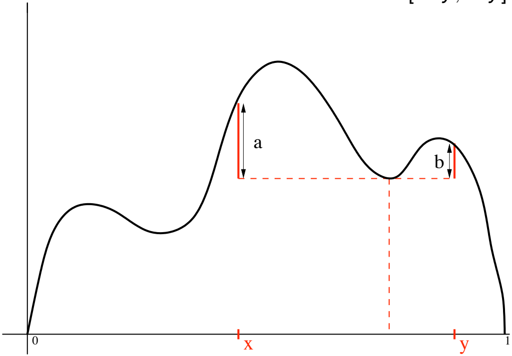
$$R(k) = \cup_{1 \leq i \leq k} [[0, L_i]], \quad [[0, L_i]] = \text{path } 0 \rightarrow L_i.$$

Coding via continuous functions

$g : [0, 1] \to [0, \infty)$ continuous: $g(0) = g(1) = 0$ and $g(x) > 0$ on $(0, 1)$.

Define a pseudo-distance on $[0, 1]$ by

$$\tilde{d}_g(x, y) = g(x) + g(y) - 2 \min_{z \in [x \wedge y, x \vee y]} g(z), \ x, y \in [0, 1].$$



$$\tilde{d}_g(x, y) = a + b$$

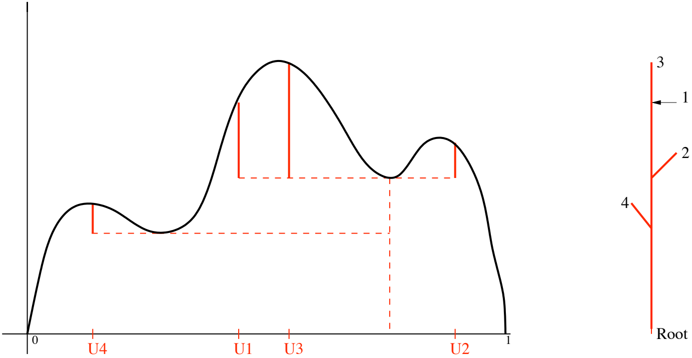
Equivalence relation $x \sim y \Leftrightarrow \tilde{d}_g(x, y) = 0$

$\tilde{d}_g$ induces a distance d_g on the quotient space $\mathcal{T}_g := [0, 1]/ \sim$.

Coding via continuous functions

Theorem (DUQUESNE-LE GALL 05)

(i) The metric space $(\mathcal{T}_g, d_g)$ is a compact $\mathbb{R}$-tree.
(ii) If $g_n \to g$ uniformly on $[0, 1]$, then $\mathcal{T}_{g_n} \to \mathcal{T}_g$ for the GH topology

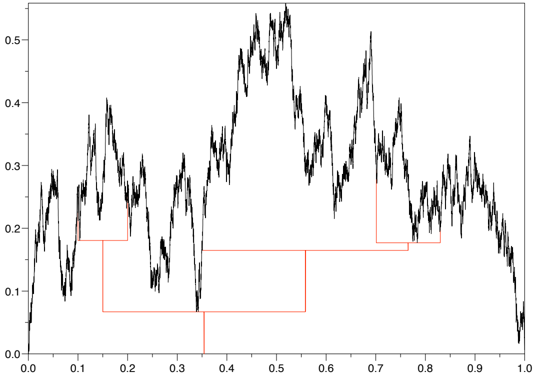


A natural probability measure on T_g is obtained as the image by the projection on $\mathcal{T}_g$ of the Lebesgue measure on $[0, 1]$.

The Brownian CRT and Brownian fragmentation

Normalized brownian excursion: $\mathbf{e} : [0,1] \to [0,\infty)$

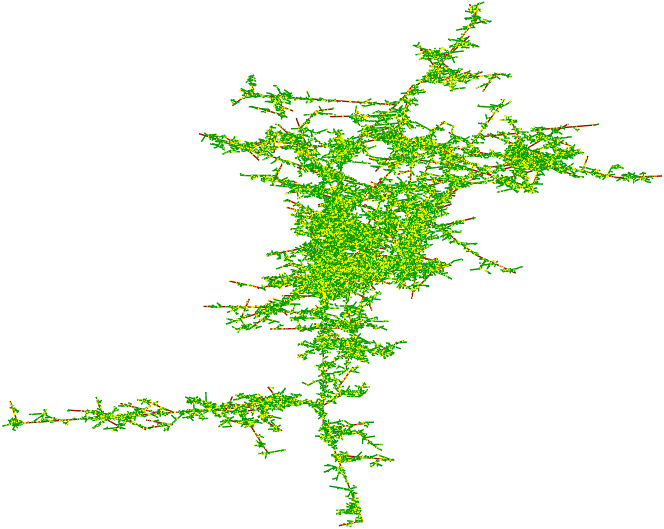
$(\mathcal{T}_{\mathbf{e}}, d_{\mathbf{e}})$: compact $\mathbb{R}$-tree constructed from $\mathbf{e}$



$\mu_{\mathbf{e}}$: mass measure obtained by projection of the Lebesgue measure.

Local minima a.s. distincts $\to \mathcal{T}_{\mathbf{e}}$ is *binary*: degrees vertices $\in \{1,2,3\}$.

The Brownian CRT: a picture in 3D



(by G.Miermont)

The Brownian CRT: scaling limit of Galton-Watson trees.

T: GW tree with offspring distribution η with mean 1, variance $\sigma < \infty$ and $\eta(1) < 1$.

T_n: r.v. with distribution that of T conditioned to have n vertices

μ_n: empirical measure assigning weight $1/n$ to all vertices

Theorem (ALDOUS 93)

$$\left(\frac{\sigma}{2\sqrt{n}} T_n, \mu_n\right) \stackrel{\text{law}}{\to} (\mathcal{T}_{\mathbf{e}}, \mu_{\mathbf{e}})$$

Brownian fragmentation

$\mathcal{L}_t(\mathcal{T}_{\mathbf{e}})$: set of leaves of height $> t$

$F_{\mathbf{e}}(t) := \downarrow$ sequence of μ-masses of connected components of $\mathcal{L}_t(\mathcal{T}_{\mathbf{e}})$

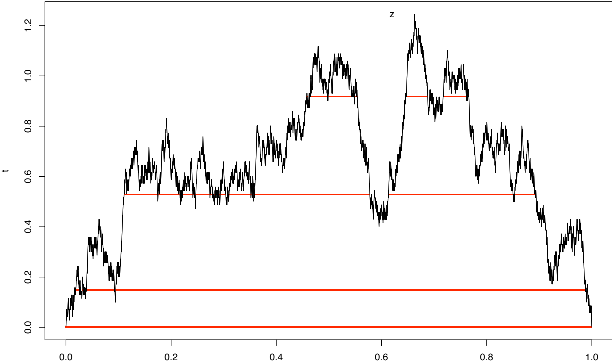


Figure: The Brownian interval fragmentation with the open intervals which constitute the state at times $t = 0, 0.15, 0.53$ and 0.92 indicated.

Brownian fragmentation

Theorem (Bertoin 02)

$(F_{\mathbf{e}}(t), t \geq 0)$ is a fragmentation with parameters $\alpha = -1/2$, $c = 0$, $\nu_{Br}(s_1 + s_2 < 1) = 0$ and

$$\nu_{Br}(s_1 \in \mathrm{d}x) = \frac{\sqrt{2}}{\sqrt{\pi x^3 (1-x)^3}} \mathrm{d}x, x \in [1/2, 1).$$

Proof. Tag a point at random: $U \sim$ uniform on $(0, 1)$.

Reduced to dust at time $D_{\mathrm{tag}} = \mathbf{e}(U)$ ▶ Rayleigh distribution

$$\mathbb{P}(2D_{\mathrm{tag}} \in \mathrm{d}t) = t \exp(-t^2/2)\mathrm{d}t, \quad t \geq 0.$$

$$\Rightarrow \mathbb{E}\left[D_{\mathrm{tag}}^n\right] = 2^{-n/2}\Gamma(1 + n/2) = n!/\prod_{i=1}^{n} \phi(n/2)$$

$\Rightarrow$ gives ϕ and then ν_{Br} (since binary). □

Stable trees $(\mathcal{T}_\beta, \mu_\beta)$, $1 < \beta \leq 2$ (Duquesne and Le Gall 02, 05)

- if $\beta = 2$: Brownian CRT
- if β, $1 < \beta < 2$,

$\mathcal{T}_\beta$: limit as $n \to \infty$ of critical GW trees T_n,

- T_n conditioned to have n vertices
- offspring distribution $(\eta_k, k \geq 0)$ s.t. $\eta_k \sim Ck^{-1-\beta}$ as $k \to \infty$
- edge-lengths $n^{1/\beta-1}$

μ_β: limit of empirical measures on the vertices of T_n.

Belong to the family of *Lévy trees* of Duquesne, Le Gall and Le Jan, introduced to code the genealogy of continuous state branching processes.

Stable fragmentations

$F_\beta(t)$:=ranked sequence of μ_β-masses of connected components of $\{$vertices of $\mathcal{T}_\beta$ with height $> t$ $\}$

Theorem (Miermont PTRF 03)

For $1 < \beta < 2$, F_β is a fragmentation with parameters $\alpha = 1/\beta - 1$, $c = 0$ and a dislocation measure ν_β

$$\int_{\mathcal{S}^{\downarrow}} f(\mathbf{s})\nu_\beta(\mathrm{d}\mathbf{s}) = \frac{\beta^2\Gamma(2-1/\beta)}{\Gamma(2-\beta)}\mathbb{E}\left[Tf\left(\frac{\Delta_1}{T}, \frac{\Delta_2}{T}, ...\right)\right]$$

Here $\Delta_1 > \Delta_2 > \cdots$ are the points of a Poisson process on $(0,\infty)$ with intensity $(\beta\Gamma(1-1/\beta))^{-1}x^{-1/\beta-1}\mathrm{d}x$ and $T = \sum_{i=1}^{\infty}\Delta_i$.

Each particle splits in an infinite number of fragments.

Other fragmentations of Lévy trees

1. Aldous-Pitman fragmentation
   obtained by logging the Brownian CRT along its skeletton: $(1/2, 0, \nu_{Br})$-fragmentation

   exponential time reversal $\rightarrow$ standard additive coalescent.

   Ref. ALDOUS-PITMAN, Ann. Probab. 98

2. Stable fragmentations with positive index:
   $(1/\beta, 0, \nu_\beta)$-fragmentation obtained by splitting the β-stable trees at nodes, $1 < \beta < 2$.

   Ref. MIERMONT, PTRF 05

3. Fragmentations of Lévy trees

   Ref. ABRAHAM-DELMAS, PTRF 08.

Fragmentation trees

Goal: generalize the construction of Brownian and stable fragmentations via compact real trees to fragmentations with a negative index

1. Construction of fragmentation trees
2. Some geometric and fractal properties

Bibliography

- B. HAAS, G. MIERMONT, The genealogy of self-similar fragmentations with negative index as a continuum random tree, EJP, 2004
- B. HAAS, J. PITMAN, M. WINKEL, Spinal partitions and invariance under re-rooting of CRT, Ann. Probab. (to appear)

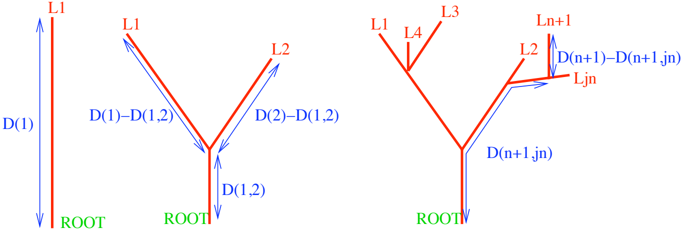
Construction and first properties

Hypothesis

F: (α, c, ν)-fragmentation s.t. $\alpha < 0$, $c = 0$ and $\nu(\sum_i s_i < 1) = 0$

I: interval version of F; $(U_i, i \geq 1)$: i.i.d, uniform on $(0,1)$, indep. of F

Recursive construction of $(T_n^{(F)}, n \geq 1)$, $T_n^{(F)}$: rooted tree with n leaves coding the genealogy of fragments containing $\{U_1, ..., U_n\}$

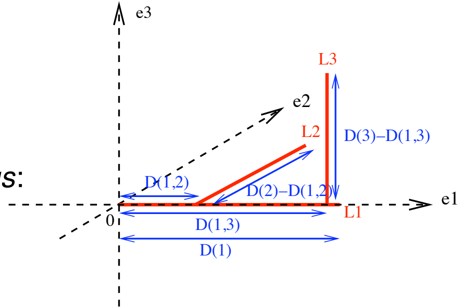


Labels: $L_i \leftrightarrow U_i$

$d(L_i, L_j) = D(i) + D(j) - 2D(i,j)$.

- $D(i) := \inf\{t : |I_{U_i}(t)| = 0\}$, death time of U_i
- $D(i,j) := \inf\{t : I_{U_i}(t) \neq I_{U_j}(t)\}$
- $j_n := \inf\{k \in \{1, ..., n\} : D(n+1, k) = \max_{1 \leq i \leq n} D(i,n)\}$

Construction and first properties



Projective limit of $T_n^{(F)}$ as $n \to \infty$?

Use *sequential construction of Aldous*:

$(\mathbf{e_1}, \mathbf{e_2}, ...)$ unit vector basis of ℓ_1

1. $L_1 := D(1)\mathbf{e_1}$
2. $L_2 := D(1,2)\mathbf{e_1} + (D(2) - D(1,2))\mathbf{e_2}$
3. $L_{n+1} := L_{j_n} - (D(j_n) - D(n+1, j_n))\mathbf{e}_{j_n} + (D(n+1) - D(n+1, j_n))\mathbf{e}_{n+1}$

- $T_n^{(F)} := \cup_{1 \le i \le n} [[0, L_i]]_{\mathrm{sp}}$, where $[[0, x]]_{\mathrm{sp}}$ piecewise linear path:

$$0 \to (x_1, 0, ..) \to (x_1, x_2, 0..) \to (x_1, x_2, x_3, 0...), \text{ etc.}$$

- $\mathcal{T}_F$: closure of $\cup_{1 \le i < \infty} [[0, L_i]]_{\mathrm{sp}}$

Construction and first properties

Lemma

$(T_n^{(F)}, n \geq 1)$ is

1. *strongly sampling consistent*: for all $1 \leq k \leq n$, picking k leaves at random in $T_n^{(F)}$ + the root generates a subtree distributed as $T_k^{(F)}$
2. *leaf-tight*: $\sup_{2 \leq j \leq k} d_{\ell_1}(L_1, L_j) \xrightarrow{\mathbb{P}} 0$

Then, by Aldous 93:

- a.s., $\mathcal{T}_F$ is a real tree + $\exists$ random μ_F : the sequence of empirical distributions of $(L_1, ..., L_n)$ converges to μ_F and $(\mathcal{T}_F, \mu_F)$ is a CRT
- a.s., $\forall t \geq 0$, $F(t) :=$ranked sequence of μ_F-masses of connected components of $\{\text{vertices of } \mathcal{T}_F \text{ with height} > t\}$
- if, conditioned on $(\mathcal{T}_F, \mu_F)$, $(Z_1, Z_2, ..)$ i.i.d. $\sim \mu_F$, then, the subtree generated by the root + $Z_1, ..., Z_n$ is distributed as $T_n^{(F)}$
- $\mathcal{S}(\mathcal{T}_F) = \cup_{1 \leq i < \infty}[[0, L_i]]_{\mathrm{sp}} \backslash \{L_i\}$

Construction and first properties

First properties

With probability one,

1. $\mathcal{T}_F$ is compact
2. $\mathcal{L}(\mathcal{T}_F)$ is dense in $\mathcal{T}_F \Leftrightarrow \nu(\mathcal{S}^{\downarrow}) = \infty$.

Proof. *- compacness*:

1) N_t^{ε}: nb. of fragments present at t not entirely reduced to dust at $t+\varepsilon$

$F_i(t) > 0$: dies at time $t + (F_i(t))^{-\alpha}\zeta^{(i)}$, with $\zeta^{(i)} \overset{\text{law}}{=} \zeta$ and indep. of $F(t)$

$\sum_i \mathbb{P}(F_i^{-\alpha}(t)\zeta^{(i)} > \varepsilon | F(t)) \leq \varepsilon^{1/\alpha} \sum_i F_i(t)\mathbb{E}[\zeta^{-1/\alpha}] < \infty$ a.s.

Borel-Cantelli $\Rightarrow N_t^{\varepsilon} < \infty$ a.s.

2) $\forall \ell \in \mathbb{N}$, $\forall$ fragment present at $\ell\varepsilon$, not entirely dead at $(\ell+1)\varepsilon$ and having lost some mass at $(\ell+2)\varepsilon$, choose a representative $U_{\sigma(\ell,i)}$ in this fragment, so that $D(\sigma(\ell,i)) < (\ell+2)\varepsilon$

Total number of $U_{\sigma(\ell,i)} \leq \sum_{\ell=0}^{[\zeta/l]} N_{\ell\varepsilon}^{\varepsilon} < \infty$ a.s.

Construction and first properties

3) $\forall j$, $\exists \ell$: U_j alive at time $\ell\varepsilon$, dead at time $(\ell+1)\varepsilon$

$\exists \sigma((\ell-1)\vee 0, i)$: $U_j, U_{\sigma((\ell-1)\vee 0,i)} \in$ same fragment at time $(\ell-1)\vee 0$

$\Rightarrow d_{\ell_1}(L_j, L_{\sigma((\ell-1)\vee 0,i)}) \leq 4\varepsilon$

$\Rightarrow \exists$ covering of $\{L_i, i \geq 1\}$ with a finite nb. of balls of radius $\leq 4\varepsilon$

$\Rightarrow \exists$ covering of $\cup_{1\leq i<\infty}[[0, L_i]]_{\mathrm{sp}}$ with a finite nb. of balls of radius $\leq 4\varepsilon$

$\Rightarrow \mathcal{T}_F$ is compact.

- leaf-density: • if $\nu(\mathcal{S}^{\downarrow}) < \infty$, the root $\notin \overline{\mathcal{L}(\mathcal{T}_F)}$

• if $\nu(\mathcal{S}^{\downarrow}) = \infty$: for $v \in [[0, L_1]]_{\mathrm{sp}} \backslash \{L_1\}$, $v = x\mathbf{e}_1$, PPP construction $\Rightarrow$ the fragment containing U_1 splits infinitely often between times x and $x+\varepsilon$

+ $N^{\varepsilon}_{x+\varepsilon} < \infty \Rightarrow \exists j$: $U_j \in I_{U_1}(x)$ that dies between times x and $x+2\varepsilon$

$\Rightarrow d_{\ell_1}(v, L_j) < 2\varepsilon$

Similarly for $x \in [[0, L_i]]_{\mathrm{sp}} \backslash \{L_i\}$, $i \geq 2$. $\square$

Hausdorff dimension

Measures the “size" of a sub-space X of a metric space: informally, $N(\varepsilon)$: minimal nb. of open balls with radius $\leq \varepsilon$ needed to cover X

$$N(\varepsilon) \sim C\varepsilon^{-\gamma} \text{ as } \varepsilon \downarrow 0 \leftrightarrow \text{dimension of } X \text{ is } \gamma.$$

Definition

Let X be a sub-space of a metric space M. Then

$$\dim_{\mathrm{H}}(X) := \inf\{\gamma > 0 : H_\gamma(X) = 0\},$$

where

$$H_\gamma(X) := \liminf_{\varepsilon \downarrow 0} \left\{ \sum_i |B_i|^\gamma : (B_i, i \geq 1) \in C_\varepsilon(X) \right\},$$

$C_\varepsilon(X)$ set of coverings of X by a countable nb. of balls with radius $\leq \varepsilon$.

Ex:

1. $\dim_{\mathrm{H}}(\mathbb{R}^n) = n$
2. $\dim_{\mathrm{H}}(\mathcal{C}) = \ln 2 / \ln 3$ where $\mathcal{C}$: standard Cantor set
3. countable sets have Hausdorff dimension 0.

Hausdorff dimension

Theorem (HAAS-MIERMONT 04)

Assume that $\int_{\mathcal{S}^{\downarrow}}(s_1^{-1}-1)\nu(\mathrm{d}\mathbf{s})<\infty$. Then,

$$\dim_{\mathrm{H}}\left(\mathcal{L}(\mathcal{T}_F)\right)=|\alpha|^{-1}.$$

The assumption on ν is not very restrictive. E.g., if $\nu(s_{N+1}>0)=0$,

$$\nu\text{-a.e.: } 1=s_1+...+s_N\leq Ns_1\Rightarrow 1/s_1-1\leq N(1-s_1)\Rightarrow \text{ok.}$$

Remark: $\dim_{\mathrm{H}}(X\cup Y)=\max\left(\dim_{\mathrm{H}}(X),\dim_{\mathrm{H}}(Y)\right)$

$$\Rightarrow \dim_{\mathrm{H}}(\mathcal{S}(\mathcal{T}_F))=1$$

$$\Rightarrow \dim_{\mathrm{H}}\left(\mathcal{T}_F\right)=\max(1,|\alpha|^{-1}).$$

-1: critical value for α.

Hausdorff dimension

Corollary (β-stable tree)

$\dim_{\mathrm{H}}(\mathcal{T}_\beta) = \beta/(\beta-1)$ a.s.

Ideas for the computation of the Hausdorff dimension.

- *Lower bound*: use Frostman lemma:

$\int_{\mathcal{L}(\mathcal{T}_F)} \int_{\mathcal{L}(\mathcal{T}_F)} d(x,y)^{-\gamma} \eta(\mathrm{d}x)\eta(\mathrm{d}y) < \infty$ for some non-trivial measure η,
$\Rightarrow \dim_H(\mathcal{L}(\mathcal{T}_F)) \geq \gamma$

Here: $\mathbb{E}\left[\int_{\mathcal{L}(\mathcal{T}_F)} \int_{\mathcal{L}(\mathcal{T}_F)} d(x,y)^{-\gamma} \mu_F(\mathrm{d}x)\mu_F(\mathrm{d}y)\right] = \mathbb{E}[d(L_1,L_2)^{-\gamma}]$,
where, conditional on $(\mathcal{T}_F, \mu_F)$, L_1, L_2 i.i.d $\sim \mu_F$

Then $d(L_1,L_2) = \lambda_1^{-\alpha}\tilde{D}(1) + \lambda_2^{-\alpha}\tilde{D}(2)$, where:
- λ_i=mass of the subtree containing L_i branched on the branchpoint of L_1 and L_2
- $\tilde{D}(i) \overset{\mathrm{law}}{=} D_{\mathrm{tag}}$, independent of (λ_1, λ_2)

$\Rightarrow \mathbb{E}[d(L_1,L_2)^{-\gamma}] \leq \mathbb{E}[\lambda_1^{\gamma\alpha}\mathbf{1}_{\lambda_1 \geq \lambda_2}]\mathbb{E}[D_{\mathrm{tag}}^{-\gamma}] \quad (2)$

Hausdorff dimension

- ▷ when $\nu(\mathcal{S}^{\downarrow}) < \infty$ and $\nu(s_{N+1} > 0) = 0$ for some N
  $\Rightarrow$ both expectations in (2) are finite for $\gamma < -1/\alpha$
  $\Rightarrow \dim_{\mathrm{H}}(\mathcal{L}(\mathcal{T}_F)) \geq -1/\alpha$

- ▷ in other cases: for $\varepsilon > 0$ and $N \in \mathbb{N}$, consider the subtree of $\mathcal{T}_F$ obtained by keeping at each branch point
  - either uniquely the largest subtree if it has a relative mass $> 1 - \varepsilon$
  - otherwise the N largest subtrees

- ▷ lower bound for the Hausdorff dim. $\rightarrow -1/\alpha$ as $\varepsilon \downarrow 0$ and $N \uparrow \infty$

- *Upper bound*: find a nice covering

Coding via continuous functions

Theorem (HAAS-MIERMONT 04)

Assume $\nu(\mathcal{S}^{\downarrow}) = \infty$.

1. Then $(\mathcal{T}_F, \mu_F)$ can be constructed via a continuous function on $[0,1]$, denoted by H_F.
2. Set,

$$\vartheta_\ell := \quad \sup\left\{b > 0 : \lim_{x \downarrow 0} x^b \nu(s_1 < 1 - x) = \infty\right\},$$

$$\vartheta_{\mathrm{u}} := \quad \inf\left\{b > 0 : \lim_{x \downarrow 0} x^b \nu(s_1 < 1 - x) = 0\right\}.$$

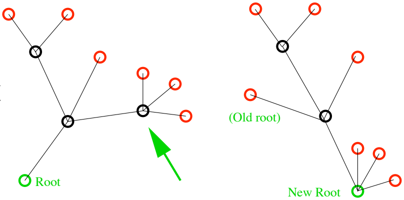
Then H_F is a.s. Hölder of order γ, $\forall \gamma < \vartheta_\ell \wedge |\alpha|$, and, provided that $\int_S (s_1^{-1} - 1)\nu(\mathrm{d}\mathbf{s}) < \infty$, a.s. not Hölder of order γ, for $\gamma > \vartheta_{\mathrm{u}} \wedge |\alpha|$.

Coding via $H_F \to$ “natural" interval representation.

Stable cases: recover Duquesne-Le Gall’s results ($\vartheta_\ell = \vartheta_{\mathrm{u}} = 1 - 1/\beta$).

Invariance under uniform re-rooting: discrete setting

1. Choose a vertex uniformly at random in a random rooted discrete tree
2. consider the tree rooted at this vertex
3. if old-root tree $\overset{\text{law}}{=}$ new-root tree, the (law of) the tree is *invariant under uniform re-rooting*.



A Galton-Watson tree invariant under uniform re-rooting:

Offspring distribution $\eta \sim \mathcal{P}(1)$ ($\eta(k) = e^{-1}/k!$, $k \in \mathbb{N}$)

T_n: such GW tree conditioned to have n vertices + labels $1, ..., n$ added uniformly at random

- T_n: uniform law over the set of rooted trees with n labeled vertices.

$\Rightarrow$ T_n is *invariant under uniform re-rooting*

Invariance under uniform re-rooting: continuous setting

$(\mathcal{T}, \mu)$: CRT rooted at ρ

Conditionally on $(\mathcal{T}, \mu)$: $(L_1, L_2, ...)$ i.i.d. sample of leaves $\sim \mu$.

$\mathcal{R}(\mathcal{T}, L_1, ..., L_n)$ subtree of $\mathcal{T}$ spanned by ρ and $L_1, ..., L_n$

$\mathcal{T}^{[L_1]}$: tree $\mathcal{T}$ re-rooted at L_1.

Definition

$(\mathcal{T}, \mu)$ is invariant under uniform re-rooting if for all $n \geq 1$,

$$\mathcal{R}(\mathcal{T}^{[L_1]}, \rho, L_2, ..., L_n) \overset{\text{law}}{=} \mathcal{R}(\mathcal{T}, L_1, L_2, ..., L_n).$$

Theorem

1. (ALDOUS 91, DUQUESNE-LE GALL 05) *For all* $\beta \in (1, 2]$, *the stable tree* $(\mathcal{T}_\beta, \mu_\beta)$ *is invariant under uniform re-rooting.*
2. (HAAS-PITMAN-WINKEL 09) *Let* $(\mathcal{T}, \mu)$ *be a* (α, ν)*-fragmentation tree, invariant under uniform re-rooting. Then* $\exists \beta \in (1, 2]$ *and* $C > 0$ *such that* $\nu = C\nu_\beta$ *and* $\alpha = 1/\beta - 1$.

Invariance under uniform re-rooting of fragmentation trees

Sketch of proof. I interval representation of the fragmentation, $U_i, i \geq 1$ i.i.d $\sim U(0,1)$, indep. of I

$\mathcal{C}_n = (C_{n,1}, ..., C_{n,k_n})$: ordered partition of $\{2, 3, ..., n+1\}$:

$i, j \in$ same block $\Leftrightarrow$ U_i and U_j separates from U_1 at the same time

order: genealogical order

But: $|I_{U_1}(t)| = \exp(-\xi_{\rho(t)}), t \geq 0$, time-change subordinator

$\Rightarrow \mathcal{C}_n$ obtained by throwing balls $\tilde{U}_2, ..., \tilde{U}_{n+1}$ in the "boxes" of $[0,1] \backslash \mathcal{R}$, where $\mathcal{R}$: closed range of $1 - \exp(-\xi)$, $\tilde{U}_i$ iid, indep. $\mathcal{R}$

• Invariance under re-rooting $\Rightarrow T^{(F)}_{n+1}$ is invariant under re-rooting, $\forall n$

$\Rightarrow (C_{n,1}, ..., C_{n,k_n}) \overset{\text{law}}{=} (C_{n,k_n}, ..., C_{n,1}), \forall n$

$\underset{\text{Gnedin-Pitman Ann.Probab.05}}{\Leftrightarrow}$ the Lévy measure of ξ is proportional to $(1 - \exp(-x))^{-b-1} \exp(-bx) \mathrm{d}x, x > 0$ for some $0 < b < 1$

Invariance under uniform re-rooting of fragmentation trees

• when $b = 1 - 1/\beta$, $\to$ Lévy measure of a β-stable fragmentation + $\nexists$ fragmentation tree with such Lévy measure when $b > 1/2$ (just use the construction of the Lévy measure from ν)

Conclusion: Invariance under re-rooting $\Rightarrow$ tagged fragment process $\overset{\text{law}}{=}$ β-stable tree's tagged fragment process

A priori insufficient to characterize ν.

But: invariance under-re-rooting + law of tagged fragment $\Rightarrow \nu$.

Last, using time-changes in the fragmentation +

$$D(1,2) \overset{\text{law}}{=} D(1) - D(1,2) \overset{\text{law}}{=} D(2) - D(1,2)$$

$\Rightarrow$ equation on α with a unique solution: $\alpha = 1/\beta - 1$. □

Scaling limits of consistent Markov branching models

Goal: exhibit a class of fragmentation trees which are scaling limits of *consistent Markov branching models*, i.e. increasing family of trees with

- a branching property
- a consistency property.

Outline

1. Consistent Markov branching models, examples
2. Characterization via homogeneous fragmentations
3. Scaling limits

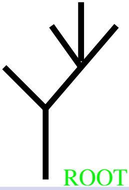
Bibliography

- Examples of consistent Markov branching models:
  - D. ALDOUS, Probability distributions on cladograms, *IMA Vol. Math. Appl.* ***76***, 1993
  - D. FORD, Probabilities on cladograms: introduction to the alpha model, *arXiv*
  - P. MARCHAL, A note on the fragmentation of a stable tree, *DMTCS*, 2009
  - D. CHEN, D. FORD, M. WINKEL, a new family of markov branching trees: the alpha-gamma model, *EJP*, 2009
- Scaling limits:
  - B. HAAS, G. MIERMONT, J. PITMAN, M. WINKEL, Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models, *Ann. Probab.*, 2008

Markov branching models

T_n: rooted tree with n leaves, *no degree-2 vertices*
+ the root has only one neighbor
+ edge-lengths 1

$\mathcal{D}(T_n)$: distribution of T_n



Definition (Markov branching property)

$(\mathcal{D}(T_n), n \geq 1)$ has the *Markov branching property* if,

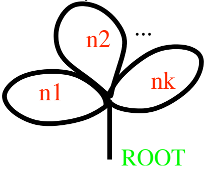
given that the branchpoint neighboring the root in T_n connects k subtrees with $n_1 \geq ... \geq n_k$ leaves,

these subtrees are independent with distributions that of $T_{n_1}, ..., T_{n_k}$.

$\Rightarrow$ Sequence characterized by the splitting rules $(q_n, n \geq 2)$,

$q_n(n_1, ..., n_k)$: probability that the branchpoint neighboring the root in T_n connects k subtrees with $n_1 \geq ... \geq n_k$ leaves

$$\forall (n_1 \geq ... \geq n_k), k \geq 2 : \sum_{i=1}^{k} n_i = n$$



Consistent Markov branching models

$T_{n+1}^{(-1)}$: obtained by removing 1 leaf at random in T_{n+1} (+ adjacent edge)

Definition (Consistent Markov branching model)

$(\mathcal{D}(T_n), n \geq 1)$ is a *consistent Markov branching model* if it has the Markov branching property and $\forall n \geq 2$,

$$T_{n+1}^{(-1)} \stackrel{\text{law}}{=} T_n.$$

$\rightarrow (q_n, n \geq 2)$: consistent splitting rules

$\Rightarrow$ recursive connection between the splitting rules q_ns, $n \geq 2$.

Definition

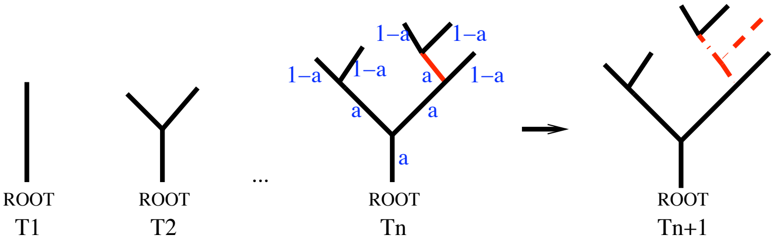
We say that $(T_n), n \geq 1)$ is a consistent Markov branching model if $(\mathcal{D}(T_n), n \geq 1)$ is a consistent Markov branching model

Consistent Markov branching models: examples

1. Ford's alpha model. $a \in [0, 1]$. Recursive construction:

weight $1 - a$ on edges neighboring a leaf
a on others egdes.

Pick an edge a random according to its weight an add a new edge + leaf.

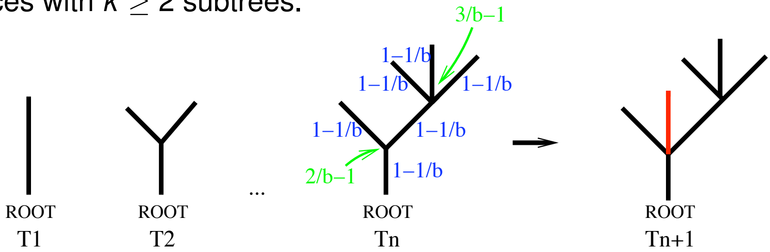


$a = 1/2$: uniform law on the set of binary trees with n labeled leaves

$a = 0$: Yule model (binary pure-birth)

Consistent Markov branching models: examples

2. Marchal's model. $b \in (1, 2]$. Weight $1 - 1/b$ on edges; $k/b - 1$ on vertices with $k \geq 2$ subtrees.



Marchal's $b = 2 \leftrightarrow$ Ford's $a = 1/2$

3. Chen-Ford-Winkel's alpha-gamma model. $a \in [0, 1]$, $\gamma \in [0, a]$. Weight $1 - a$ on edges neighboring a leaf; weight γ on other edges; weight $(k - 1)a - \gamma$ on vertices with $k \geq 2$ subtrees.

$\gamma = a$ is Ford's model

$\gamma = 1 - a$, $a \in [1/2, 1)$ is Marchal's model $b = 1/a$.

Consistent Markov branching models: examples

4. Aldous's beta-splitting model. $-2 < \beta < \infty$. Binary trees.
For $1 \leq k < n/2$,

$$q(n-k,k) = Z_n^{-1} 2 \int_0^1 \binom{n}{k} x^{k+\beta}(1-x)^{n-k+\beta}\mathrm{d}x$$

and for even n,

$$q(n/2,n/2) = Z_n^{-1} \int_0^1 \binom{n}{n/2} x^{n/2+\beta}(1-x)^{n/2+\beta}\mathrm{d}x.$$

$\beta = -3/2 \leftrightarrow$ Ford's $a = 1/2$ (uniform); $\beta = 0 \leftrightarrow$ Ford's $a = 0$ (Yule)

5. Discrete fragmentation trees. F: (α, c, ν)-fragmentation
$(T_n^{(F)}, n \geq 1)$ is consistent Markov branching
(replacing natural edge-lengths with edge lengths 1 $\rightarrow$ no influence of α on the splitting rules)

Characterization via homogeneous fragmentations

Theorem (HMPW 08)

Given $(T_n, n \geq 1)$ a consistent Markov branching model, $\exists$ a $(0, c, \nu)$-fragmentation F such that

$$T_n \overset{\text{law}}{=} T_n^{(F)}, \quad \forall n \geq 1.$$

For all $r > 0$, $(0, c, \nu)$ and $(0, rc, r\nu)$ correspond to the same splitting rules.

$\Rightarrow$ one-to-one correspondence between the set of consistent splitting rules and the set of characteristic triplets $(0, c, \nu)$ of homogeneous fragmentations (up to multiplicative constants).

Ex. When $c = 0$, $\nu(\sum_{i \geq 1} s_i < 1) = 0$,

$$q^{(c,\nu)}(n_1, \ldots, n_r) = \frac{1}{Z_n} \binom{n}{n_1, \ldots, n_r} \int_{\mathcal{S}^{\downarrow}} \sum_{i_1 \neq \ldots \neq i_r \geq 1} \prod_{j=1}^{r} s_{i_j}^{n_j} \, \nu(\mathrm{d}\mathbf{s}).$$

Characterization: examples

1. Ford's model. When $0 \leq a < 1$: $c = 0$, $\nu(s_1 + s_2 < 1) = 0$ and

$$\nu(s_1 \in \mathrm{d}x) = \left(\frac{a}{2}(x(1-x))^{-a-1} + (1-2a)(x(1-x))^{-a}\right)\mathrm{d}x,$$

$$(1/2 \leq x < 1).$$

When $a = 1$: pure erosion: $c > 0$ and $\nu(\mathcal{S}^{\downarrow}) = 0$.

2. Marchal's model: $c = 0$, $\nu = \nu_b$ (stable fragmentation with index b)

Characterization: examples

3. Chen-Ford-Winkel's model. When $0 < a < 1$ and $0 \leq \gamma \leq a$: $c = 0$ and for all test functions f

$$\int_{\mathcal{S}^{\downarrow}} f(\mathbf{s})\nu(\mathrm{d}\mathbf{s}) = \mathbb{E}\left[\sigma_1^{a+\gamma}\left(\gamma + (1-a-\gamma)\sum_{i \neq j} \Delta_i\sigma\Delta_j\sigma\right) f(\Delta\sigma_{[0,1]}/\sigma_1)\right]$$

where σ: stable subordinator with Laplace exponent $\lambda \to \lambda^a$
$\Delta\sigma_{[0,1]} = (\Delta_i\sigma, i \geq 1)$ decreasing sequence of jumps $(\Delta\sigma_t, t \in [0,1])$.

When $a = 0$: Ford's model with $a = 0$

When $a = 1$ and $0 < \gamma < 1$: $c = 0$ and

$$\int_{\mathcal{S}^{\downarrow}} f(\mathbf{s})\nu(\mathrm{d}\mathbf{s}) = \int_0^1 f(s_1, 0, ...)(\gamma(1-s_1)^{-1-\gamma}\mathrm{d}s_1 + \delta_0(\mathrm{d}s_1))$$

When $a = \gamma = 1$: $c = 0$ and $\nu(\mathrm{d}\mathbf{s}) = \delta_{(0,0.,..)}$.

4. Aldous's model. When $-2 < \beta < \infty$: $c = 0$, $\nu(s_1 + s_2 < 1) = 0$

$$\nu(s_1 \in \mathrm{d}x) = x^{\beta}(1-x)^{\beta}\mathrm{d}x, \quad 1/2 \leq x < 1.$$

Scaling limits of $(T_n, n \geq 1)$, consistent Markov branching model

Assumptions:

A1. $c = 0$ and $\nu(\sum_i s_i < 1) = 0$

A2. $\nu(s_1 < 1 - \varepsilon) \underset{\varepsilon \to 0}{\sim} C\varepsilon^{-\gamma_\nu}$, where $0 < C < \infty$ and $\gamma_\nu \in (0,1)$
($\Rightarrow \nu$ is infinite)

A3. $\int_{\mathcal{S}^\downarrow} \sum_{i \geq 2} s_i |\ln(s_i)|^\rho \nu(\mathrm{d}\mathbf{s}) < \infty$ for some $\rho > 0$
(satisfied e.g. when $\nu(s_{N+1} > 0) = 0$ for some $N > 0$).

Theorem (HMPW 08)

$\mathcal{T}_{(-\gamma_\nu, \nu)}$: fragmentation tree with index $-\gamma_\nu$ and dislocation measure ν.
Then,

$$\frac{T_n}{C n^{\gamma_\nu} \Gamma(1 - \gamma_\nu)} \underset{n \to \infty}{\to} \mathcal{T}_{(-\gamma_\nu, \nu)}$$

in law, with respect to the Gromov-Hausdorff metric.

Remark. A2 extends to "$\nu(s_1 < 1 - \varepsilon)$ varies regularly as $\varepsilon \to 0$".

Scaling limits: examples where A1, A2 and A3 hold

Ford's models for $0 < a < 1$. Then $\gamma_\nu = a$.

Marchal's model for $1 < b \leq 2$. Then $\gamma_\nu = 1 - 1/b$.
(Recover standard results on asymptotics as $n \to \infty$ of GW trees conditioned to have n vertices, except that here the trees are conditioned to have n leaves).

Chen-Ford-Winkel's model for $0 < a < 1$, $0 < \gamma \leq a$. Then, $\gamma_\nu = \gamma$.

Aldous's model for $-2 < \beta < -1$. Then $\gamma_\nu = -\beta - 1$.

Scaling limits: outline proof.

Sufficient to prove the result for $(T_n^{(F)}, n \geq 1)$, $T_n^{(F)}$ *with edge lengths 1*, F fragmentation with parameters satisfying A1, A2, A3.

Rk. A2 $\Rightarrow$ the associated Lévy measure L is such that

$$\int_x^\infty L(\mathrm{d}y) \underset{x \to 0}{\sim} Cx^{-\gamma_\nu}$$

- asymptotic height of a random leaf: using the interval construction: the height of leaf 1 in $T_{n+1}^{(F)} = k_n =$number of boxes of $\mathcal{R}$ occupied by at least one ball among $\tilde{U}_2, ..., \tilde{U}_{n+1}$ (notations slide 62)

  Gnedin-Pitman-Yor Ann. Probab. 2006: a.s.,

  $$\Rightarrow k_n \sim C\Gamma(1-\gamma_\nu) n^{\gamma_\nu} \int_0^\infty \exp(-\gamma_\nu \xi_r)\mathrm{d}r$$

- asymptotic behavior of subtrees spanned by k random leaves: use height of one leaf + branching + self-similar properties
- tightness