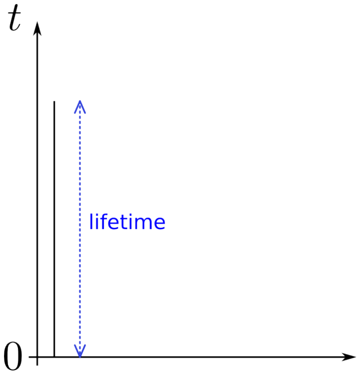
Invariance principle for the marked coalescent point process

Cécile Delaporte

LPMA, Paris 6 - SMILE, CIRB Collège de France

1er avril 2013

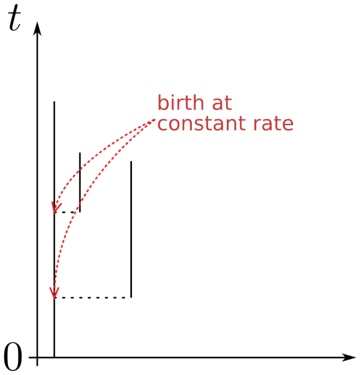
Splitting trees (without mutations)



Individuals

- behave independently from one another,
- have i.i.d. life durations (with general distribution),
- give birth at constant rate during their lifetime.

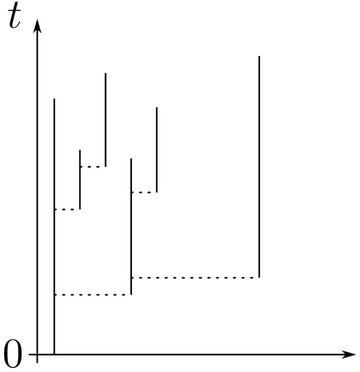
Splitting trees (without mutations)



Individuals

- behave independently from one another,
- have i.i.d. life durations (with general distribution),
- give birth at constant rate during their lifetime.

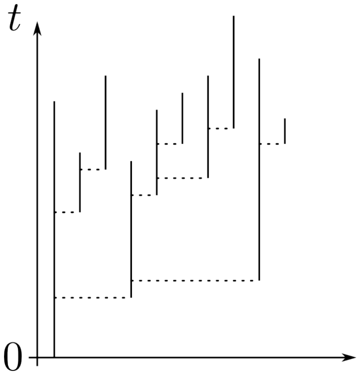
Splitting trees (without mutations)



Individuals

- behave independently from one another,
- have i.i.d. life durations (with general distribution),
- give birth at constant rate during their lifetime.

Splitting trees (without mutations)



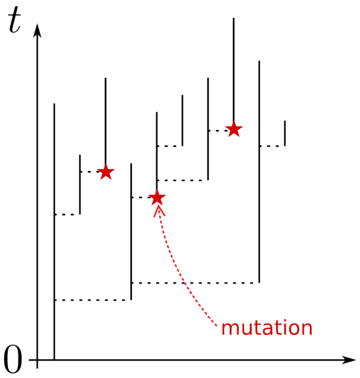
Individuals

- behave independently from one another,
- have i.i.d. life durations (with general distribution),
- give birth at constant rate during their lifetime.

A splitting tree is characterized by a σ-finite measure Λ on $(0,\infty)$ satisfying $\int_{(0,\infty)}(1 \wedge r)\Lambda(\mathrm{d}r) < \infty$ (the *lifespan measure*).

Example : if Λ is finite with mass b, individuals give birth at rate b and have life durations distributed as $\Lambda(\cdot)/b$.

Marked splitting trees

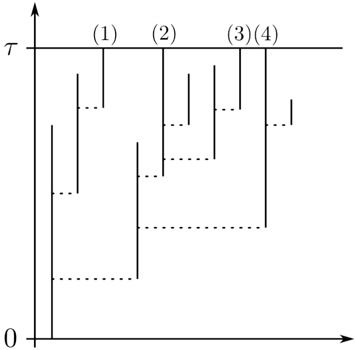


- Individuals carry clonally inherited types,
- Neutral mutations may happen along the birth events : every newborn is affected by a mutation with probability θ.

A *marked splitting tree* is characterized by its lifespan measure Λ and its *mutation parameter* θ.

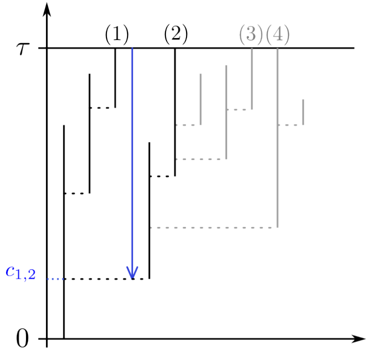
The coalescent point process

From now on, we fix $\tau > 0$. The *coalescent point process* (CPP) characterizes the genealogy of the individuals alive at τ :



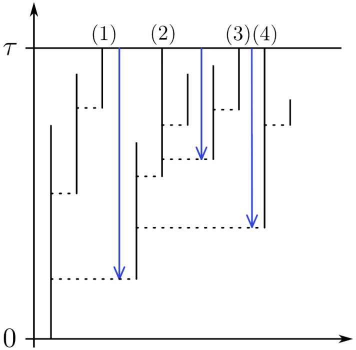
The coalescent point process

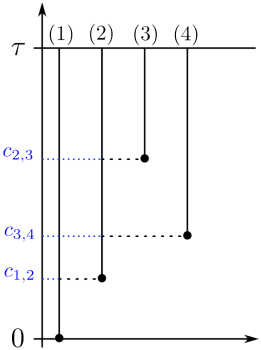
From now on, we fix $\tau > 0$. The *coalescent point process* (CPP) characterizes the genealogy of the individuals alive at τ :



The coalescent point process

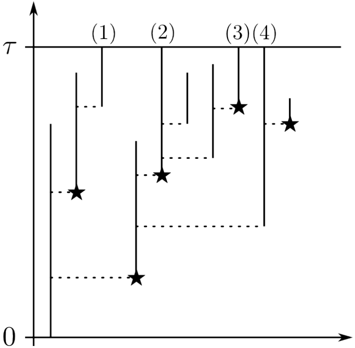
From now on, we fix $\tau > 0$. The *coalescent point process* (CPP) characterizes the genealogy of the individuals alive at τ :

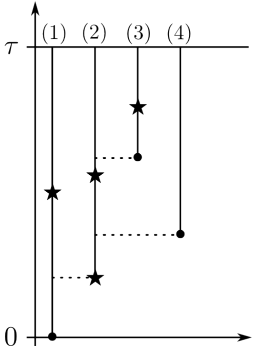




The marked coalescent point process

The *marked coalescent point process* characterizes the genealogy of the individuals alive at τ, enriched with the history of the mutations that appeared over time :





Goal : getting asymptotic results for the marked coalescent point process when the population size is large and mutations rare.

Contour of a splitting tree (without mutations)

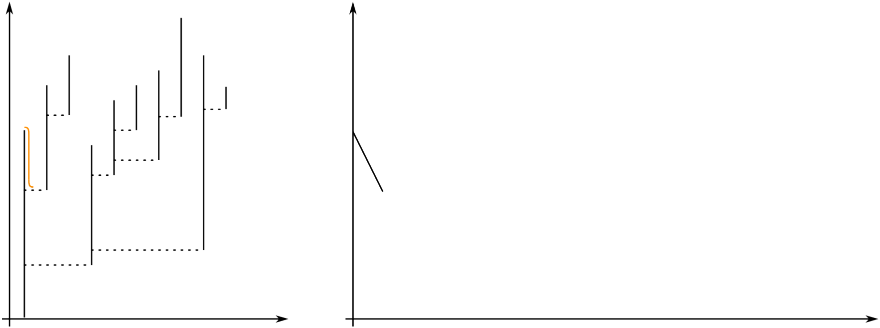


Figure: Example of a finite splitting tree and its contour process.

Contour of a splitting tree (without mutations)

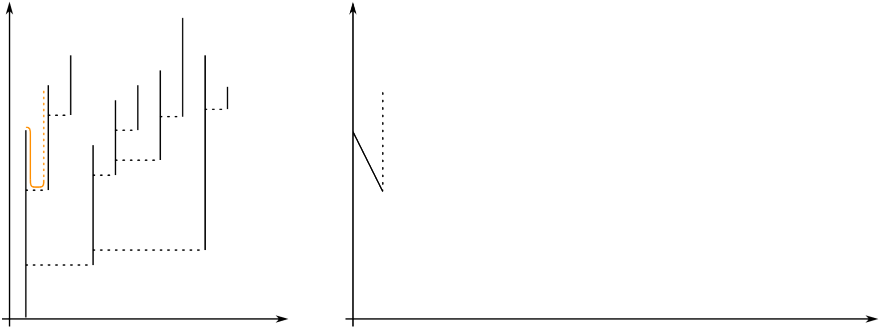


Figure: Example of a finite splitting tree and its contour process.

Contour of a splitting tree (without mutations)

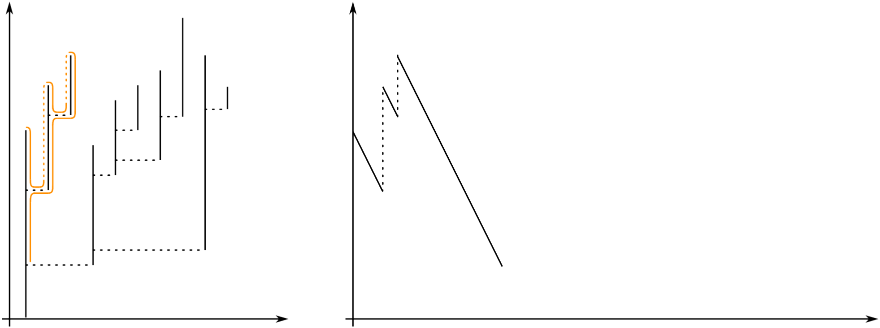


Figure: Example of a finite splitting tree and its contour process.

Contour of a splitting tree (without mutations)

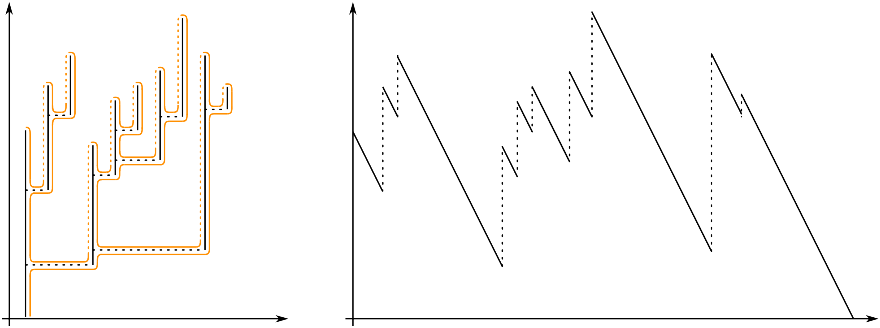
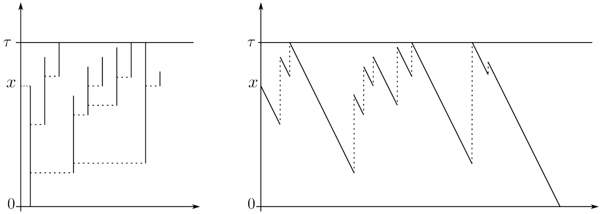


Figure: Example of a finite splitting tree and its contour process.

Contour of a splitting tree (without mutations)

Consider :

- $\mathbb{T}$ a splitting tree with lifespan measure Λ,
- $\mathbb{T}^\tau$ its truncation up to level τ
- Z a finite variation Lévy process with Lévy measure Λ and drift -1.



Theorem (A. Lambert '10)

Conditional on the first individual of $\mathbb{T}$ to have life span x, the contour of $\mathbb{T}^\tau$ is distributed as Z, starting at $x \wedge \tau$, reflected below τ and killed upon hitting 0.

Generalization to marked splitting trees

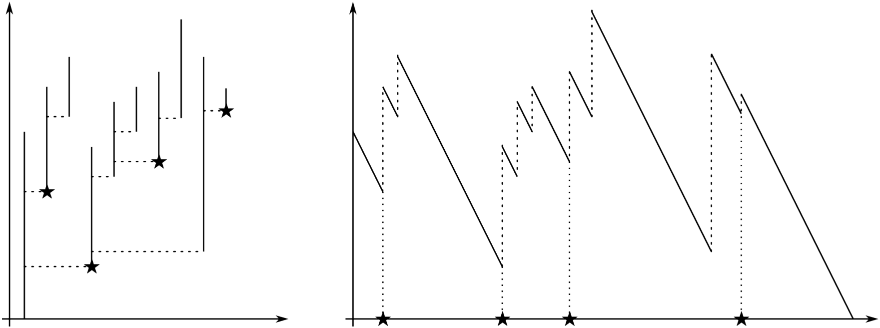


Figure: Example of a finite marked splitting tree and its contour process.

Generalization to marked splitting trees

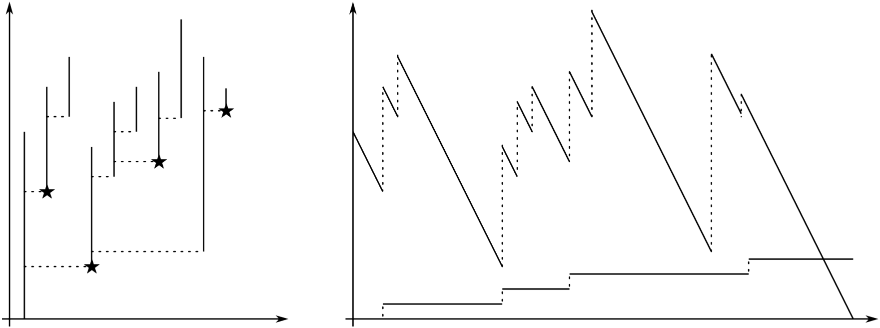


Figure: Example of a finite marked splitting tree and its contour process.

Rescaling the populations

Let $(\mathbb{T}_n)_{n\geq 1}$ be a sequence of marked splitting trees :

$$\mathbb{T}_n \text{ has } \begin{cases} \text{lifespan measure } \Lambda_n \\ \text{mutation parameter } \theta_n. \end{cases}$$

Consider **the rescaled marked splitting trees** $\tilde{\mathbb{T}}_n$ obtained from $\mathbb{T}_n$ by rescaling the branch lengths by a factor $\frac{1}{n}$.

Convergence assumptions

Let $(d_n)_{n\geq 1}$ be a sequence of positive real numbers, and
Z_n be a finite variation Lévy process with Lévy measure Λ_n and drift -1.

Define

$$\tilde{Z}_n := \left(\frac{1}{n} Z_n(d_n t)\right)_{t\geq 0}.$$

The Lévy process $\tilde{Z}_n$ has drift $-\frac{d_n}{n}$ and Lévy measure $d_n \Lambda_n(n\,\cdot\,)$.

$$\begin{array}{lcl} \text{Contour of } \mathbb{T}_n & \leftrightarrow & Z_n \\ \text{Contour of } \tilde{\mathbb{T}}_n & \leftrightarrow & \tilde{Z}_n \end{array}$$

Convergence assumptions

Condition the population of $\tilde{\mathbb{T}}_n$ on having I_n individuals alive at τ, where $I_n \underset{n\to\infty}{\sim} \frac{d_n}{n}$.

Assumption A

As $n \to \infty$, $\tilde{Z}_n = \left(\frac{1}{n} Z_n(d_n t)\right)_{t\geq 0}$ converges in distribution towards a Lévy process Z with infinite variation.

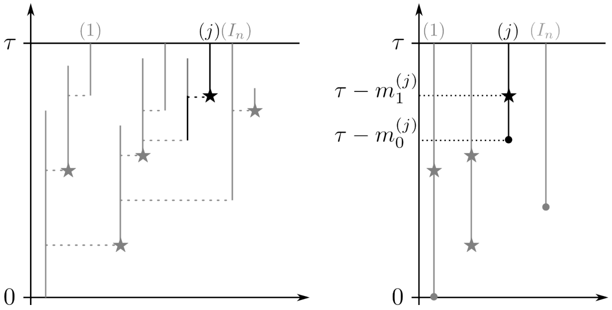
Remark : $\frac{d_n}{n} \to \infty$.

Assumption B

As $n \to \infty$, $\frac{d_n}{n}\theta_n$ converges to some finite real number θ.

Marked coalescent point process of the rescaled population

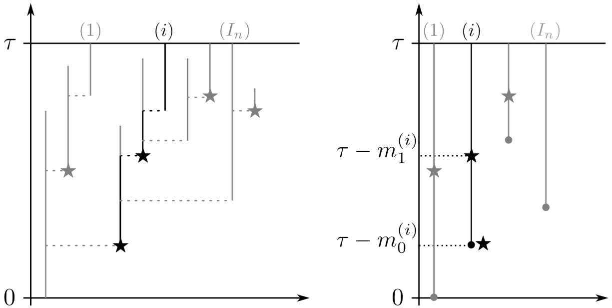
For $j \in \{1, ..., I_n\}$ we define $\sigma_n^{(j)}$ as follows :



$$\sigma_n^{(j)} = \{(m_0^{(j)}, 0), (m_1^{(j)}, 1)\}$$

Marked coalescent point process of the rescaled population

For $j \in \{1, ..., I_n\}$ we define $\sigma_n^{(j)}$ as follows :



$$\sigma_n^{(i)} = \{(m_0^{(i)}, 0), (m_0^{(i)}, 1), (m_1^{(i)}, 1)\}$$

Marked coalescent point process of the rescaled population

Define the random point measure :

$$\Sigma_n = \sum_{1<i\leq I_n} \delta_{\{\frac{in}{d_n},\sigma_n^{(i)}\}}$$

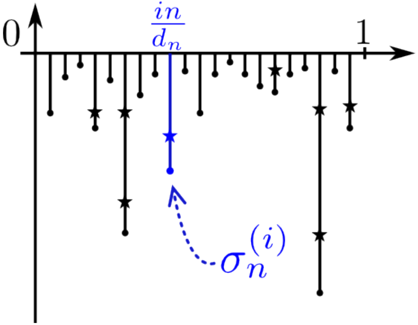
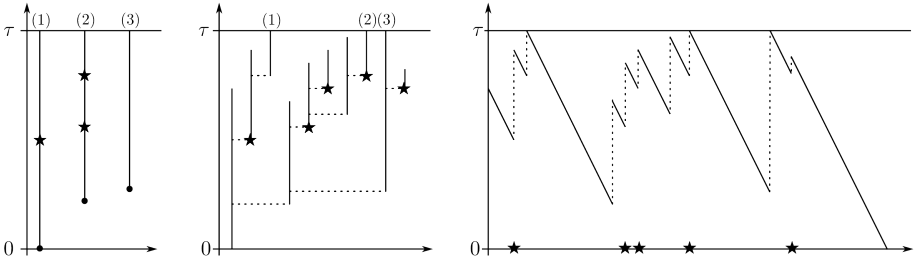


Figure: A graphical representation of Σ_n

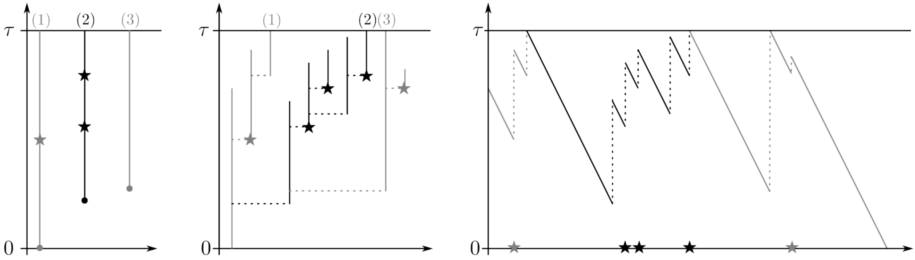
How to characterize the law of $(\sigma_n^{(i)})_{1<i\leq I_n}$?

The marked CPP of $\tilde{\mathbb{T}}_n^\tau$, the marked splitting tree $\tilde{\mathbb{T}}_n^\tau$, and its contour process :



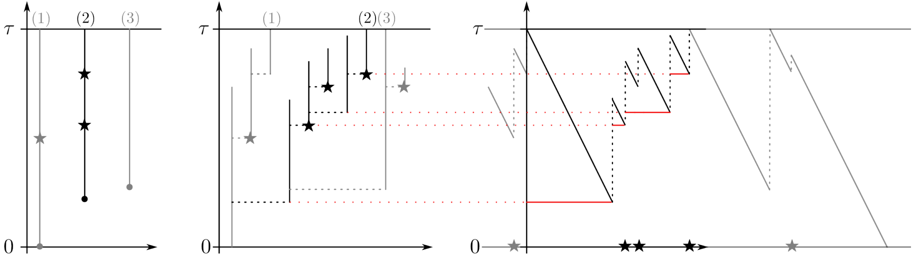
How to characterize the law of $(\sigma_n^{(i)})_{1<i\leq I_n}$?

The marked CPP of $\tilde{\mathbb{T}}_n^\tau$, the marked splitting tree $\tilde{\mathbb{T}}_n^\tau$, and its contour process :



How to characterize the law of $(\sigma_n^{(i)})_{1<i\leq I_n}$?

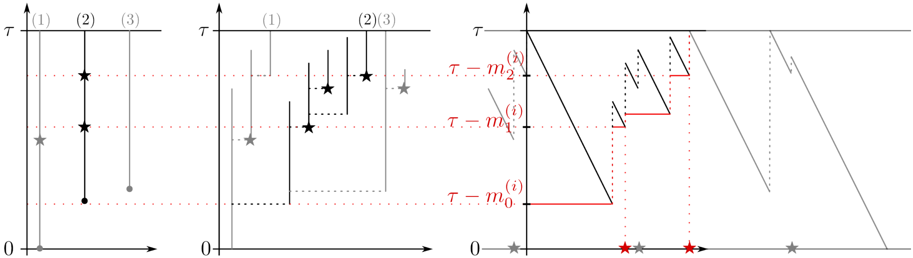
The marked CPP of $\tilde{\mathbb{T}}_n^{\tau}$, the marked splitting tree $\tilde{\mathbb{T}}_n^{\tau}$, and its contour process :



$\rightarrow$ The r.v. $(\sigma_n^{(i)})_{1<i\leq I_n}$ are i.i.d.

How to characterize the law of $(\sigma_n^{(i)})_{1<i\leq I_n}$?

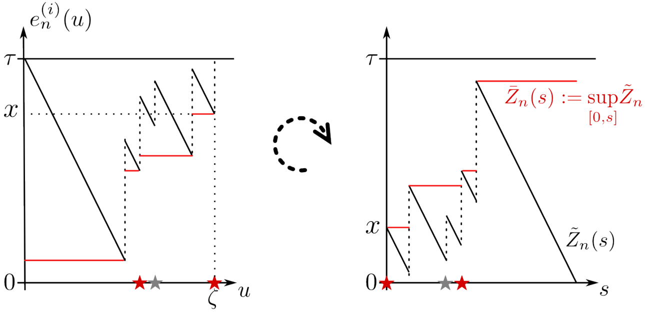
The marked CPP of $\tilde{\mathbb{T}}_n^\tau$, the marked splitting tree $\tilde{\mathbb{T}}_n^\tau$, and its contour process :



→ The r.v. $(\sigma_n^{(i)})_{1<i\leq I_n}$ are i.i.d.

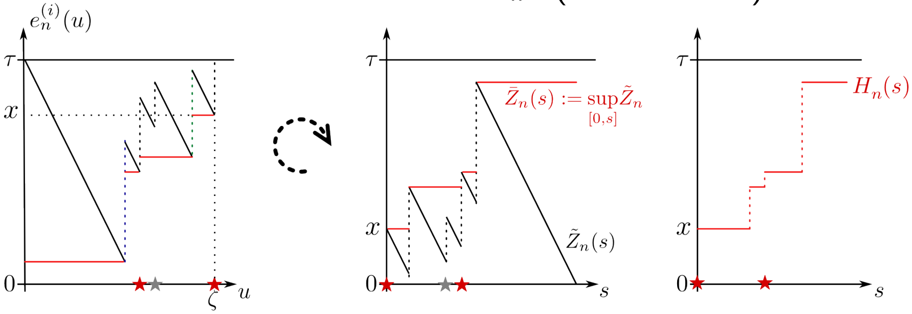
→ For fixed i :
$\sigma_n^{(i)} = \{(m_0^{(i)}, 0), (m_0^{(i)}, 1), (m_1^{(i)}, 1)\}$ can be described from the future infimum of the i-th excursion under τ of the contour process, and the set of its marked jump times.

How to characterize the law of $\sigma_n^{(i)}$ (for a fixed i) ?



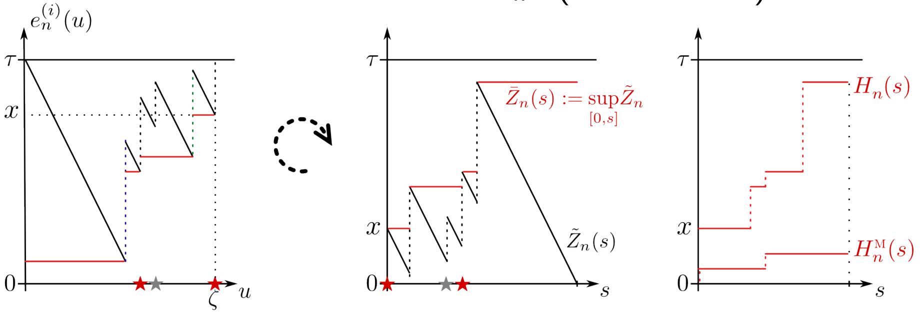
Conditional on $e_n^{(i)}(\zeta-) = x$, the reversed excursion $(\tau - e_n^{(i)}((\zeta - t)-),\ 0 \leq t < \zeta)$ is distributed as $\tilde{Z}_n(t)$ starting at x, hitting 0 before (τ, ∞) and killed when hitting 0.

How to characterize the law of $\sigma_n^{(i)}$ (for a fixed i) ?



Define H_n the ladder height process of $\tilde{Z}_n$: $H_n = \bar{Z}_n \circ L_n^{-1}$, where L_n is a local time at 0 of $\bar{Z}_n - \tilde{Z}_n$, and L_n^{-1} its inverse.

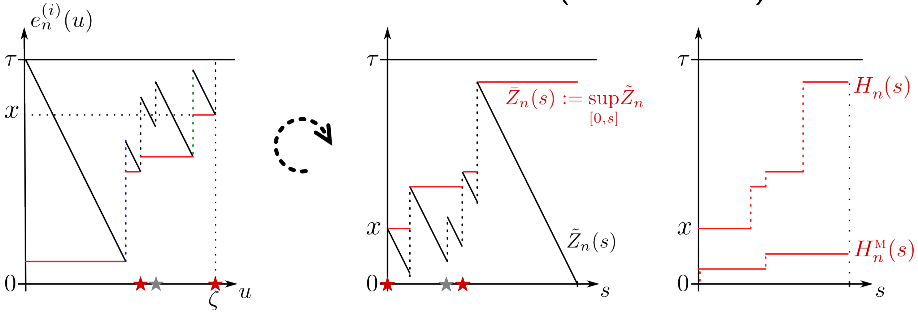
How to characterize the law of $\sigma_n^{(i)}$ (for a fixed i) ?



Define H_n the ladder height process of $\tilde{Z}_n$: $H_n = \bar{Z}_n \circ L_n^{-1}$, where L_n is a local time at 0 of $\bar{Z}_n - \tilde{Z}_n$, and L_n^{-1} its inverse.

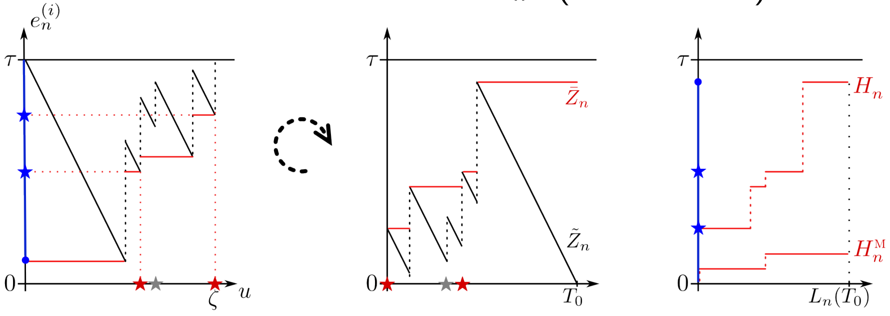
We mark the jumps of H_n in accordance with the marks of $\tilde{Z}_n$, and denote by H_n^{M} the counting process of these marks.

How to characterize the law of $\sigma_n^{(i)}$ (for a fixed i) ?



(H_n, H_n^{M}) is a (possibly killed) bivariate subordinator.
We call it the *marked ladder height process*.

How to characterize the law of $\sigma_n^{(i)}$ (for a fixed i) ?



The law of $\sigma_n^{(i)}$ can be described from

- the image of the jump times of H_n^{M} by H_n,
- and $H_n(L_n(T_0)-)$ (the terminal value of H_n in the picture).

Convergence of the marked LHP

Under assumptions A and B :

Lemma

the sequence of bivariate subordinators (H_n, H_n^{M}) converges weakly in law to a (possibly killed) subordinator (H, H^{M}), where

- *H and H^{M} are independent,*
- *H is the ladder height process of Z,*
- *H^{M} is a Poisson process with parameter θ.*

Recall the definition of the random point measure Σ_n :

$$\Sigma_n = \sum_{1<i\leq I_n} \delta_{\{\frac{in}{d_n},\sigma_n^{(i)}\}}$$

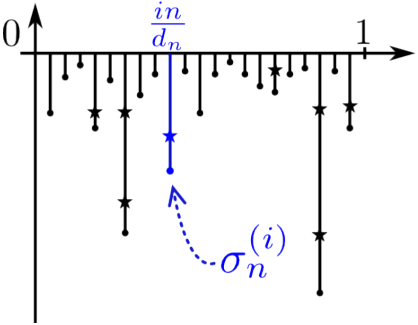


Figure: A graphical representation of Σ_n

Define

- (Θ_i) : the jump times of an indep. Poisson process with parameter θ
- $J := \inf\{i \geq 0,\ \Theta_i > L(T_0)\}$
- $\sigma = \{(H(\Theta_0), 1), \ldots, (H(\Theta_{J-1}), 1), (H(L(T_0)-), 0)\}$

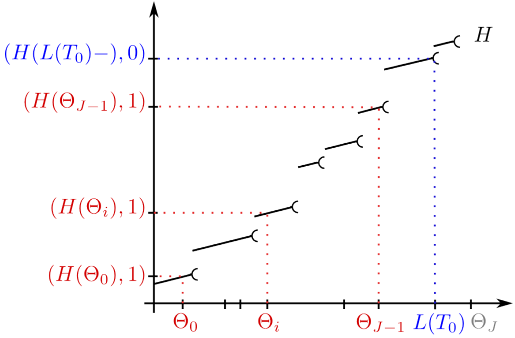


Figure: A graphical representation of σ

Define

- (Θ_i) : the jump times of an indep. Poisson process with parameter θ
- $J := \inf\{i \geq 0,\ \Theta_i > L(T_0)\}$
- $\sigma = \{(H(\Theta_0), 1), \dots, (H(\Theta_{J-1}), 1), (H(L(T_0)-), 0)\}$

Theorem

As $n \to \infty$, (Σ_n) converges in distribution towards a Poisson point measure with intensity

$$Leb|_{[0,1]} \cdot N(\sigma \in \cdot\,,\ \sup \epsilon < \tau),$$

where N is the excursion measure of Z away from 0.

The Brownian case

$\mathbb{T}_n$ is a critical branching process such that :

$$\mathbb{T}_n \text{ has } \begin{cases} \text{exponential lifespan measure } \Lambda_n(\mathrm{d}r) = e^{-r}\mathbb{1}_{r\geq 0}\mathrm{d}r \\ \text{mutation parameter } \theta_n = \frac{\beta}{n} \text{ for some } \beta \in [0,1]. \end{cases}$$

- $\tilde{Z}_n \Rightarrow B$ (B the standard Brownian motion) (Ass. A)
- $\frac{d_n}{n} \Rightarrow \theta = \frac{\beta}{2}$ (Ass. B)

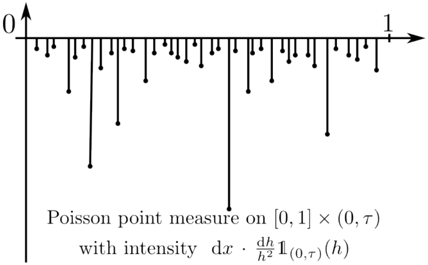
The ladder height process of B is $H(t) = 2t$.

The Brownian case

The limiting Poisson point measure Σ has intensity

$$\mathsf{Leb}|_{[0,1]} \cdot N(\Theta|_{[0,\sup \epsilon]} \in \cdot\,,\ \sup \epsilon < \tau),$$

where Θ is an independent Poisson process with parameter β and $\Theta|_{[0,T]}$ denotes its restriction to $[0, T]$.



The Brownian case

The limiting Poisson point measure Σ has intensity

$$\mathrm{Leb}|_{[0,1]} \cdot N(\Theta|_{[0,\sup\epsilon]} \in \cdot\,,\ \sup\epsilon < \tau),$$

where Θ is an independent Poisson process with parameter β and $\Theta|_{[0,T]}$ denotes its restriction to $[0, T]$.

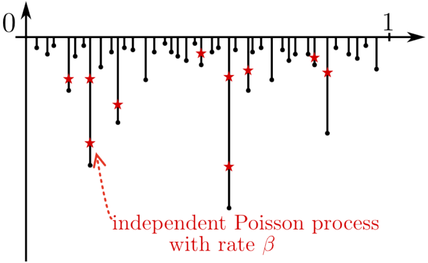


Figure: A graphical representation of Σ