# Coupures et records sur les arbres browniens

Patrick Hoscheit[1]

[1] CERMICS, MAPMO, SMILE

Journée de l'ANR MANEGE
École Polytechnique, 14 novembre 2012

# Plan de l'exposé

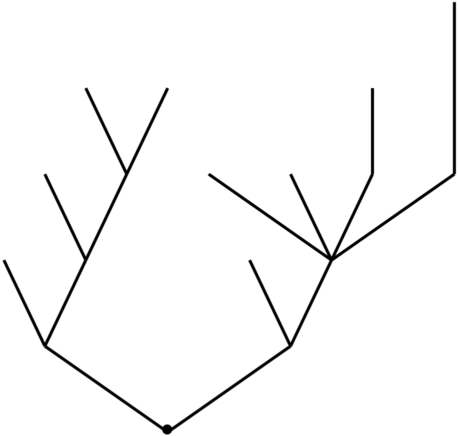
1. Coupures d'arbres aléatoires
   - Découpage d'arbres aléatoires discrets
   - Arbre aléatoire continu d'Aldous
   - Fragmentation d'Aldous-Pitman
2. Une limite presque sûre
   - Records sur le CRT
   - Convergence presque sûre
   - Esquisse de preuve
3. Fluctuations autour de la limite
   - Théorème principal
   - Une autre martingale !

# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

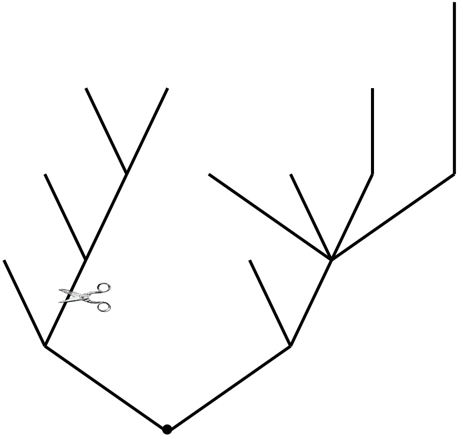


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

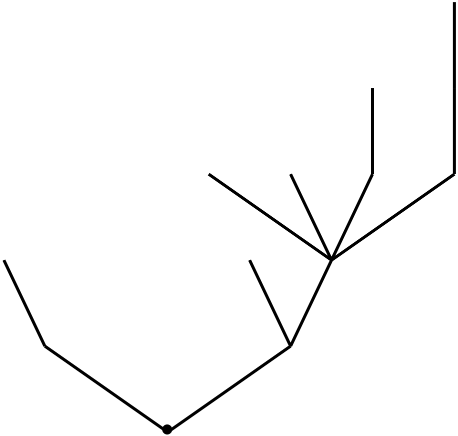


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

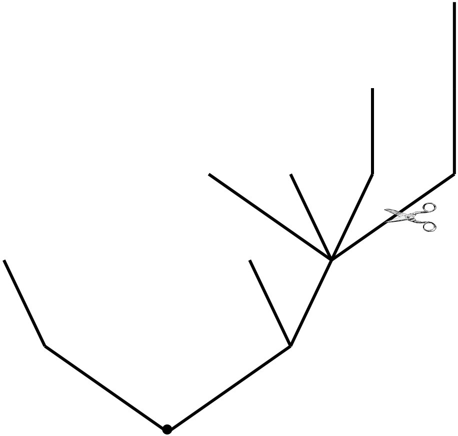


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

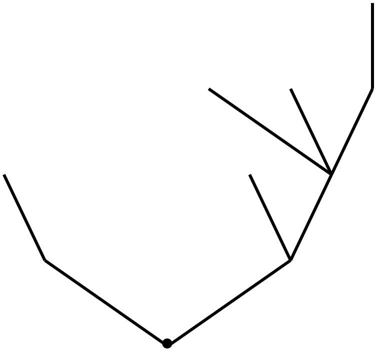


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

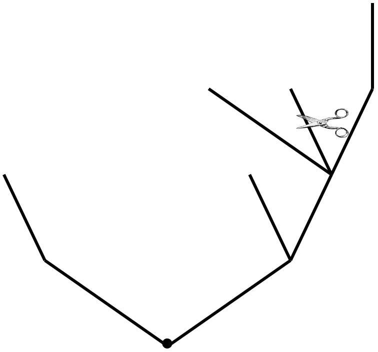


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

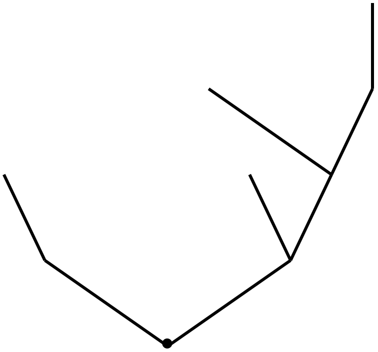


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

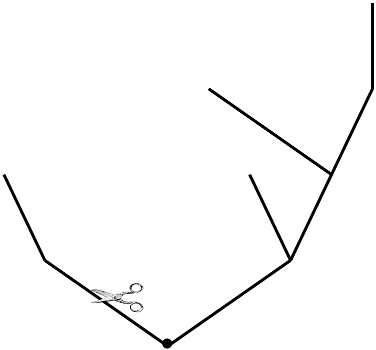


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

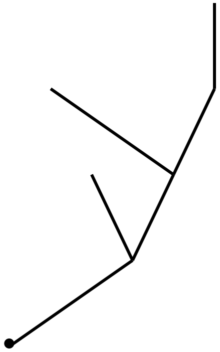


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

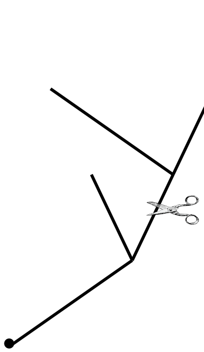


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

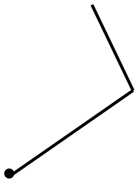


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

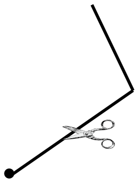


# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine



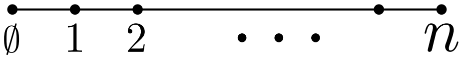
# Comment découper des arbres discrets

Soit $T_n$ un arbre discret enraciné, à $n$ arêtes.

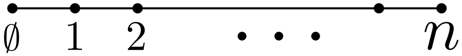
1. On choisit une arête uniformément au hasard
2. On supprime cette arête, ainsi que la c.c. ne contenant pas la racine
3. On itère jusqu'à ce que la racine soit isolée

$X(T_n)$ = nombre d'étapes nécessaires à l'isolation de la racine

# Un cas extrême

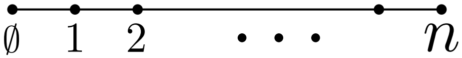


# Un cas extrême



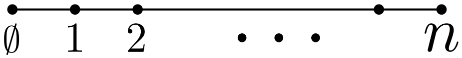

- Pour $1 \leq j \leq n$, $I_j = \mathbf{1}_{\{\text{La } j\text{-ème arête est coupée}\}}$, de telle sorte que $X(T_n) = \sum_{j=1}^{n} I_j$

# Un cas extrême



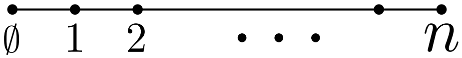

- Pour $1 \leq j \leq n$, $I_j = \mathbf{1}_{\{\text{La } j\text{-ème arête est coupée}\}}$, de telle sorte que $X(T_n) = \sum_{j=1}^{n} I_j$
- On a $\mathbb{P}(I_j = 1) = 1/j$

# Un cas extrême




- Pour $1 \leq j \leq n$, $I_j = \mathbf{1}_{\{\text{La } j\text{-ème arête est coupée}\}}$, de telle sorte que $X(T_n) = \sum_{j=1}^{n} I_j$
- On a $\mathbb{P}(I_j = 1) = 1/j$
- Les $(I_j,\ 1 \leq j \leq n)$ sont indépendantes

# Un cas extrême




- Pour $1 \leq j \leq n$, $I_j = \mathbf{1}_{\{\text{La } j\text{-ème arête est coupée}\}}$, de telle sorte que $X(T_n) = \sum_{j=1}^{n} I_j$
- On a $\mathbb{P}(I_j = 1) = 1/j$
- Les $(I_j,\ 1 \leq j \leq n)$ sont indépendantes

Ainsi $\mathbb{E}[X(T_n)] \sim \log n$ et $Var(X(T_n)) \sim \log n$. Par le théorème de Lindeberg–Feller,

$$\frac{X(T_n) - \log n}{\sqrt{\log n}} \xrightarrow{(d)} Z,$$

où $Z$ est une v.a. gaussienne standard.

# Arbres aléatoires I

Étude de $X(T_n)$ où $T_n$ est aléatoire et $\mathrm{diam}(T_n) \sim \log n$

- Drmota–Iksanov–Möhle–Rösler : arbres aléatoires récursifs
- Holmgren : “split trees”, dont arbres phylogénétiques
  - Arbres de recherche binaires ($\beta = 0$ dans le modèle d’Aldous)
  - Arbres préfixe ($\beta = \infty$)

# Arbres aléatoires I

**Theorem (DIMR '08, Holmgren '11)**

*Il existe des constantes $\alpha, \mu, \zeta$ telles que*

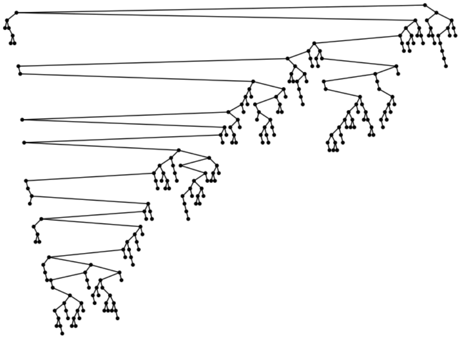
$$\frac{X(T_n) - \frac{\alpha n}{\mu \log n} - \frac{\alpha n \log\log n}{\mu \log^2 n} - \frac{\zeta n}{\mu \log^2 n}}{\frac{\alpha n}{\mu^2 \log^2 n}} \xrightarrow{(d)} Z,$$

*où $Z$ est de loi stable d'indice 1.*

# Arbres aléatoires II

Étude de $X(T_n)$ où $T_n$ est aléatoire et $\mathrm{diam}(T_n) \sim \sqrt{n}$

- Meir–Moon, Chassaing–Marchand : arbres de Cayley
- Panholzer, Janson : arbres de Galton-Watson conditionnés ($\beta = -3/2$ dans le modèle d'Aldous)



## Théorème (Janson '06)

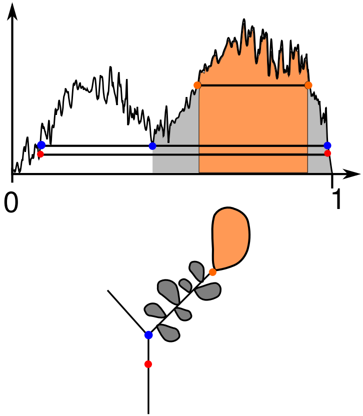
*Soit $\mu$ une mesure de probabilité critique sur $\mathbb{N}$, de variance $\sigma^2 < \infty$. Soit $T_n$ un arbre de Galton-Watson, de loi de reproduction $\mu$, conditionné à avoir n arêtes. Alors,*

$$\frac{X(T_n)}{\sigma\sqrt{n}} \xrightarrow{(d)} R,$$

*où $R$ est de loi Rayleigh $\mathbb{P}(R \in dx) = x \exp(-x^2/2)\, dx$.*

- Pourquoi une loi de Rayleigh ?
- Que se passe-t-il sur le CRT ?

# CRT d'Aldous (ou arbre brownien)



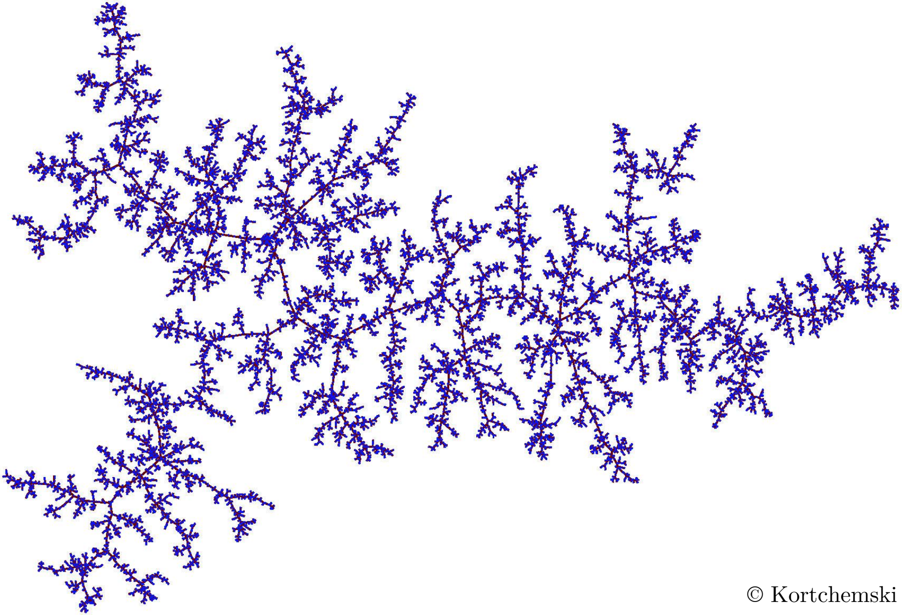

$B^{ex} : [0,1] \to \mathbb{R}_+$ excursion normalisée du MB.
Relation d'équivalence sur $[0,1]$ :
$x \sim y$ ssi

$$B^{ex}(x) = B^{ex}(y) = \min_{z \in [x,y]} B^{ex}(z).$$

Le CRT est un espace métrique aléatoire, défini par :

$$\mathcal{T}_{2B^{ex}} = [0,1]_{/\sim}$$

avec la distance $d_{2B^{ex}}(x,y) = 2B^{ex}(x) + 2B^{ex}(y) - 2m(x,y)$



© Kortchemski

# Propriétés du CRT

- Espace métrique compact, enraciné en $\varnothing = \bar{0}$
- Structure de $\mathbb{R}$-arbre
- Points de branchement binaires seulement
- Mesure de “Lebesgue” sur le squelette de l’arbre : $\ell(ds)$ ($\sigma$-finie)
- Mesure “uniforme” sur les feuilles de l’arbre : $\mathbf{m}(ds)$ (mesure de probabilité, image de la mesure de Lebesgue sur $[0, 1]$)
- Limite d’échelle d’arbres de Galton-Watson conditionnés

# Propriétés du CRT

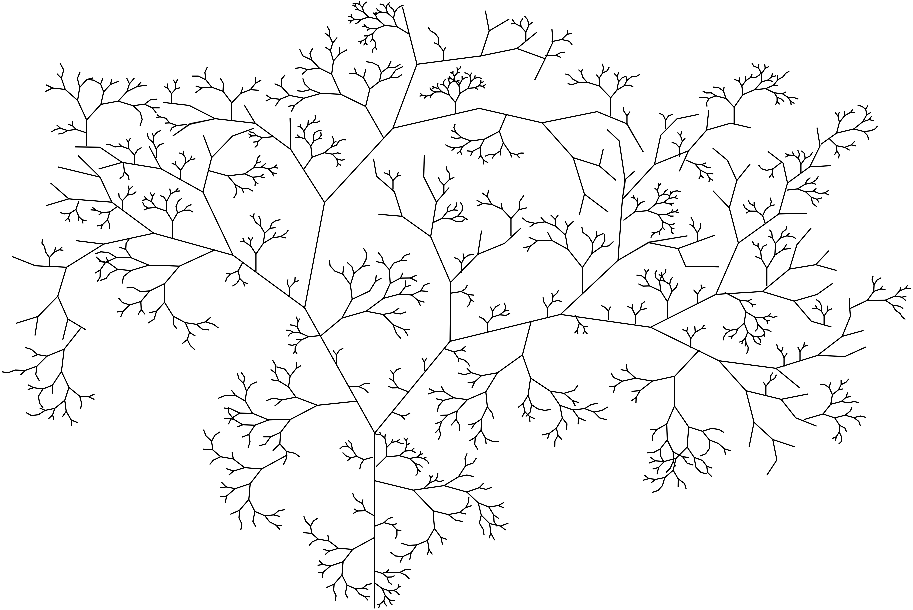
- Espace métrique compact, enraciné en $\varnothing = \overline{0}$
- Structure de $\mathbb{R}$-arbre
- Points de branchement binaires seulement
- Mesure de “Lebesgue” sur le squelette de l’arbre : $\ell(ds)$ ($\sigma$-finie)
- Mesure “uniforme” sur les feuilles de l’arbre : $\mathbf{m}(ds)$ (mesure de probabilité, image de la mesure de Lebesgue sur $[0, 1]$)
- Limite d’échelle d’arbres de Galton-Watson conditionnés

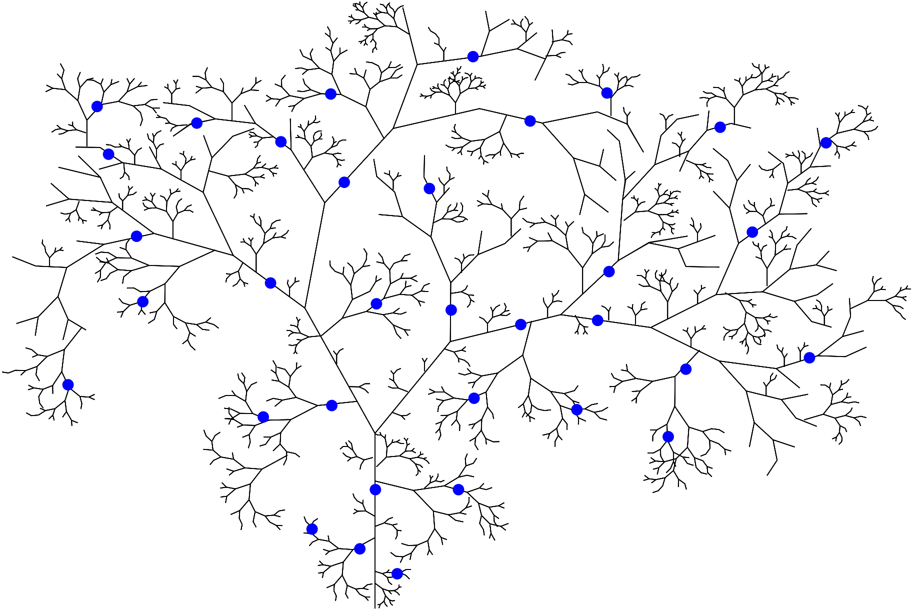
Un fait intéressant : $\int_{\mathcal{T}} d(\varnothing, s)\, \mathbf{m}(ds)$ est de loi Rayleigh...

# Fragmentation d'Aldous-Pitman

Soit $\mathcal{T}$ un arbre Brownien et $\mathcal{M}(dx, dt)$ une mesure ponctuelle de Poisson sur $\mathcal{T} \times \mathbb{R}_+$, d'intensité $\ell(dx) \otimes dt$.

- Marques "uniformes" $x$ sur le squelette de l'arbre, avec une étiquette $t$ (temps d'apparition)
- Au temps $t$, les atomes de $\mathcal{M}(dx, [0, t])$ séparent l'arbre en une infinité de composantes connexes
- Processus de fragmentation auto-similaire binaire, d'indice 1/2 (pas de perte de masse)





Peut-on retrouver une v.a. de Rayleigh à partir de ce découpage "continu" ?

Oui !

Oui !

- Addario-Berry–Broutin–Holmgren (2011) : procédure de reconstruction qui recycle les fragments du CRT ne contenant pas la racine. Ils construisent un autre CRT, pour lequel la hauteur d'une feuille uniforme correspond au “temps d'isolation” de la racine.
- Bertoin–Miermont (2012) : construction de deux arbres, $\mathrm{cut}(T_n)$ et $\mathrm{cut}(\mathcal{T})$ (même loi que le CRT), qui encodent respectivement le découpage aléatoire de $T_n$ et la fragmentation d'Aldous-Pitman de $\mathcal{T}$. Alors

$$\left(\frac{\sigma}{\sqrt{n}}T_n, \frac{1}{\sigma\sqrt{n}}\mathrm{cut}(T_n)\right) \xrightarrow{(d)} (\mathcal{T}, \mathrm{cut}(\mathcal{T})).$$

Approche différente, due à Abraham–Delmas (2011).

# Plan de l'exposé


1. Coupures d'arbres aléatoires
   - Découpage d'arbres aléatoires discrets
   - Arbre aléatoire continu d'Aldous
   - Fragmentation d'Aldous-Pitman

2. Une limite presque sûre
   - Records sur le CRT
   - Convergence presque sûre
   - Esquisse de preuve

3. Fluctuations autour de la limite
   - Théorème principal
   - Une autre martingale !

# Temps de séparation

**Definition**

Si $s \in \mathcal{T}$, on définit :

$$\theta(s) = \inf\{t \geq 0,\ \mathcal{M}(\llbracket \varnothing, s \rrbracket \times [0, t]) \geq 1\}$$

le temps de séparation de $s$ et de la racine $\varnothing$.

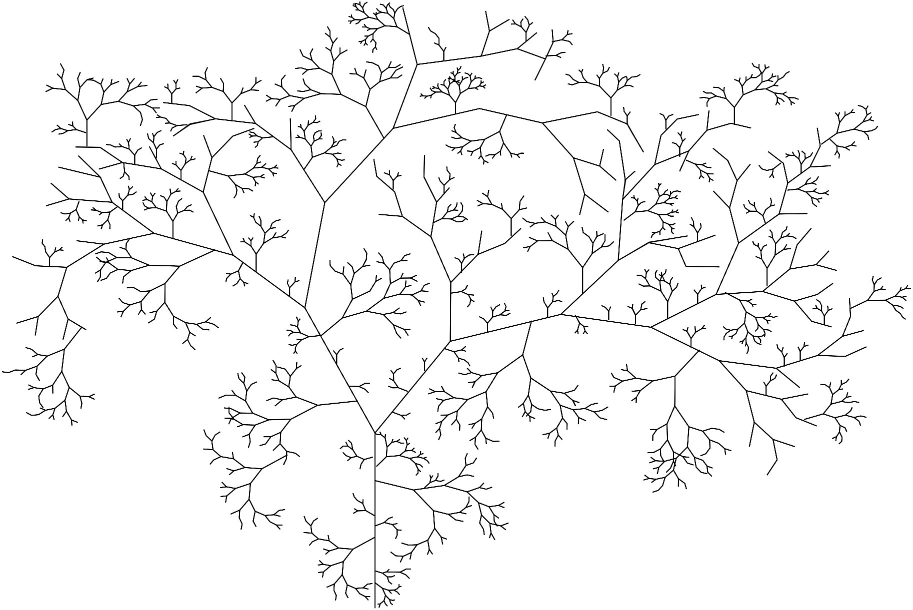
- Si $s \in \mathcal{T}$, $\theta(s)$ est de loi exponentielle, de paramètre $d(\varnothing, s)$
- Le processus $(\theta(s),\ s \in \mathcal{T})$ est appelé *processus de records*. Il est décroissant lorsqu'on s'éloigne de la racine, et $\theta(\varnothing) = \infty$.
- On note $\mathbb{P}_\infty$ la loi d'un CRT $\mathcal{T}$, muni d'un processus de records $\theta$.

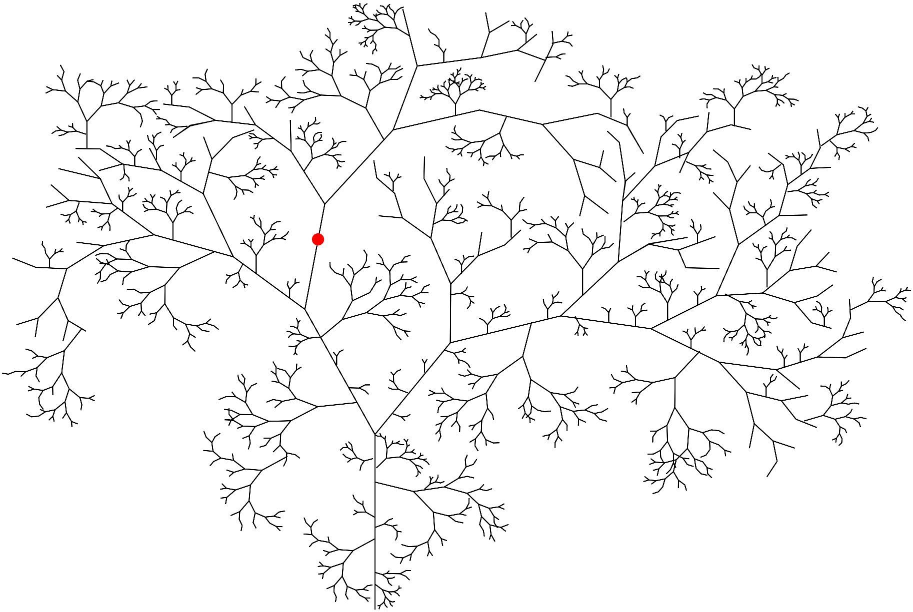
# Quelques définitions

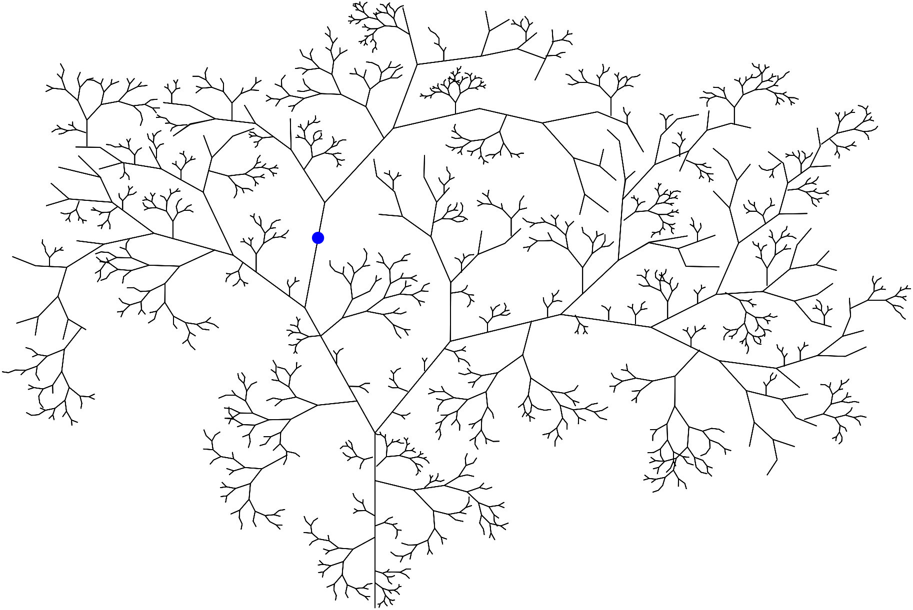
É tant donné un CRT $\mathcal{T}$ et un processus de records $\theta$, on choisit des feuilles iid $(L_n,\ n \geq 1)$ de loi $\mathbf{m}$. Si $n \geq 1$, on note $\mathcal{T}_n$ le sous-arbre de $\mathcal{T}$ engendré par les feuilles $(L_1, \ldots, L_n)$ et la racine.

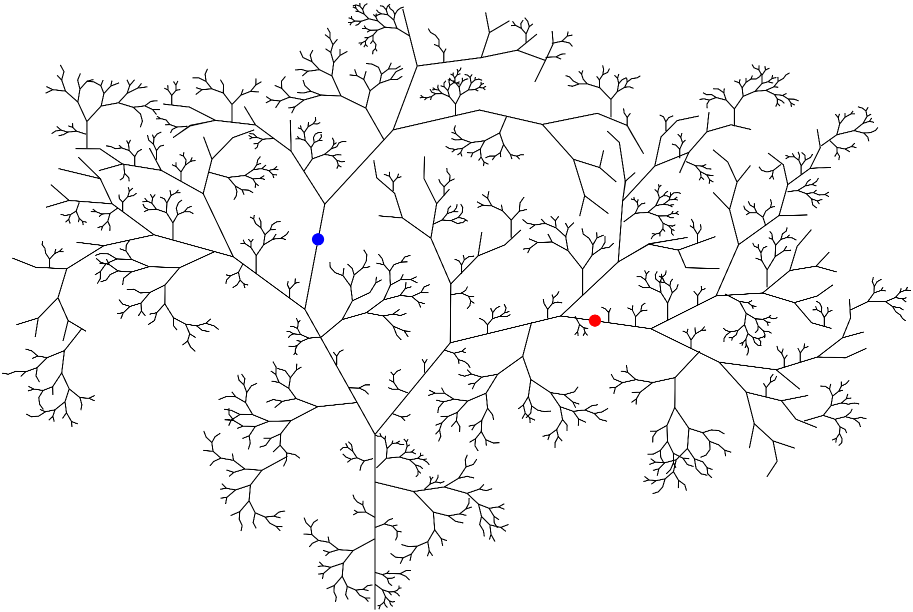
- L'arbre $\mathcal{T}_n$ est un arbre binaire (enraciné, ordonné) à $n$ feuilles uniforme, de longueurs de branche échangeables (Aldous)
- On note $x_{\emptyset,n}$ le premier point de branchement
- On définit $h_{\emptyset,n} = d(\emptyset, x_{\emptyset,n})$ et $\mathcal{T}_n^* = \mathcal{T}_n \setminus [\![\emptyset, x_{\emptyset,n}[\![$
- $X_n^*$ = nombre de *sauts* de $\theta$ sur $\mathcal{T}_n^*$.

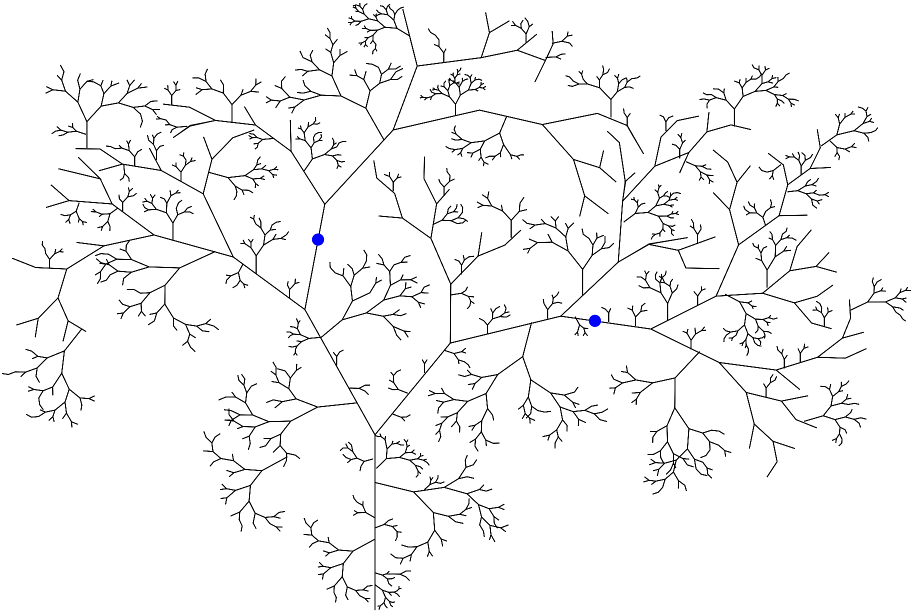
$X_n^*$ est le nombre de fragmentations nécessaires pour isoler la racine dans $\mathcal{T}_n$.

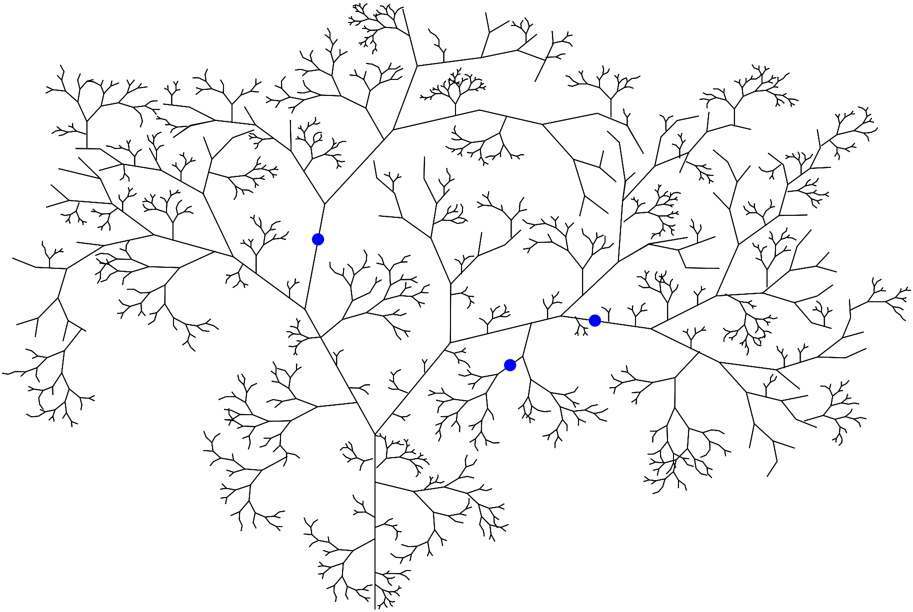


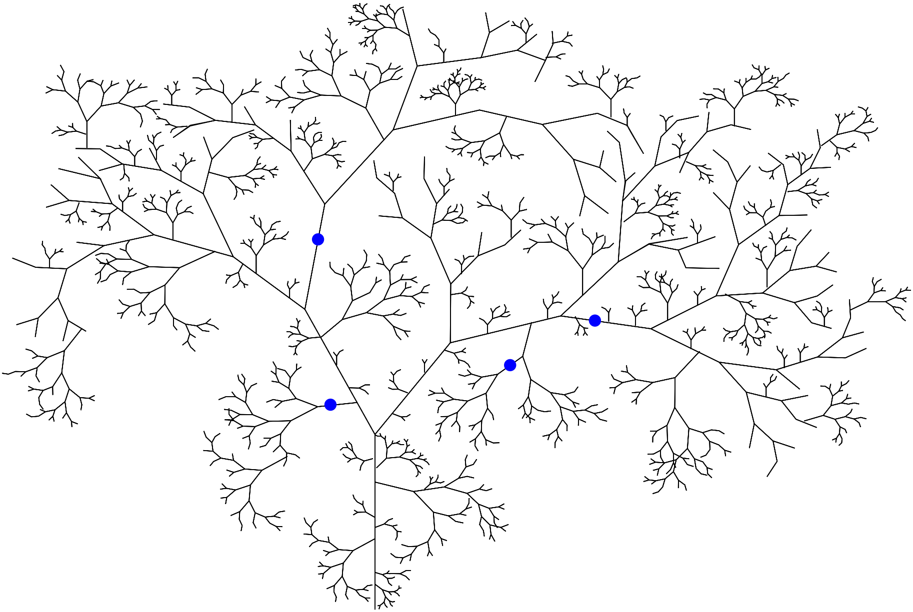


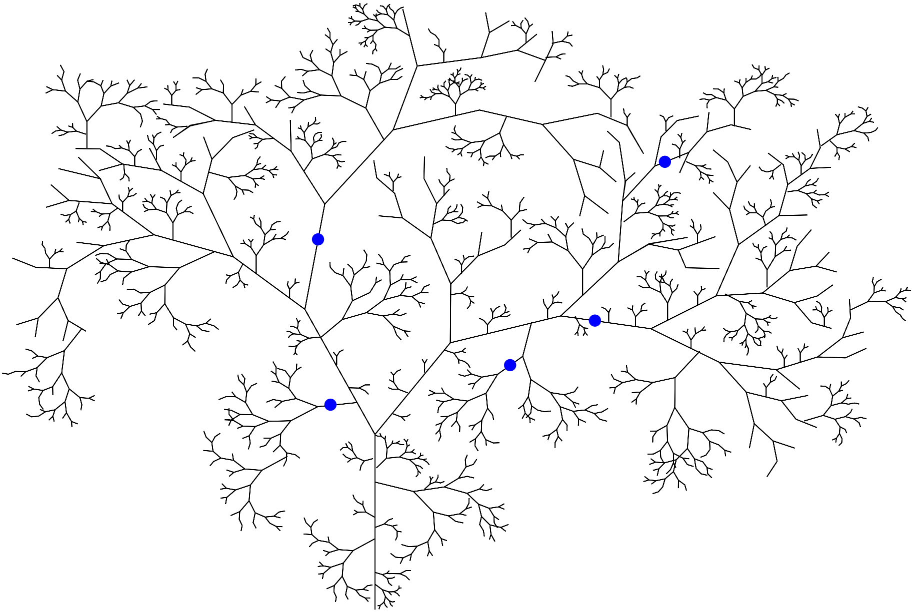


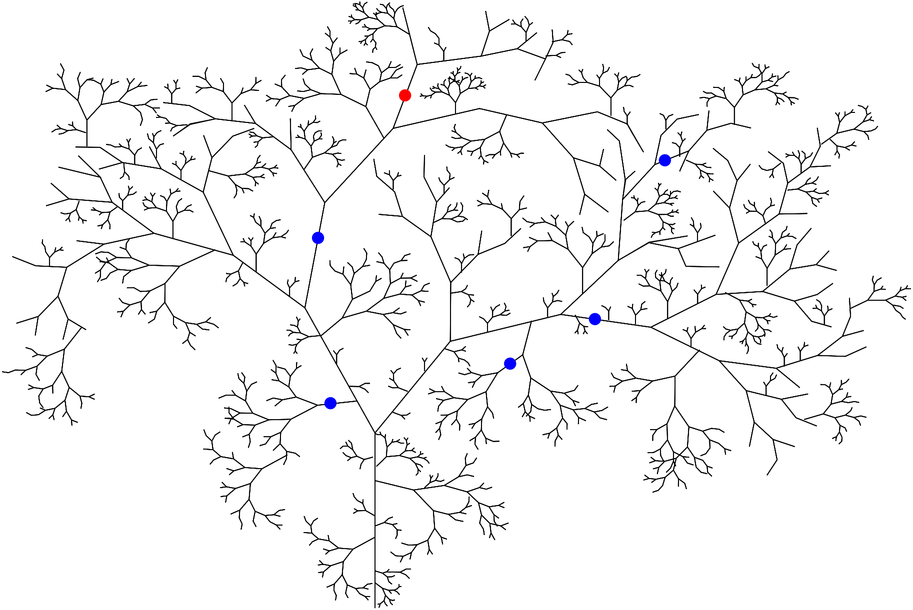


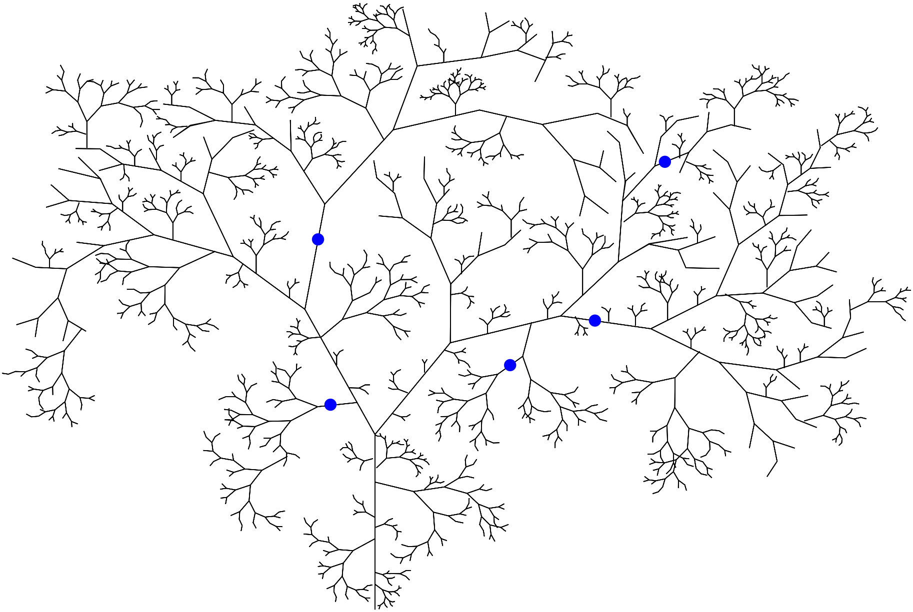


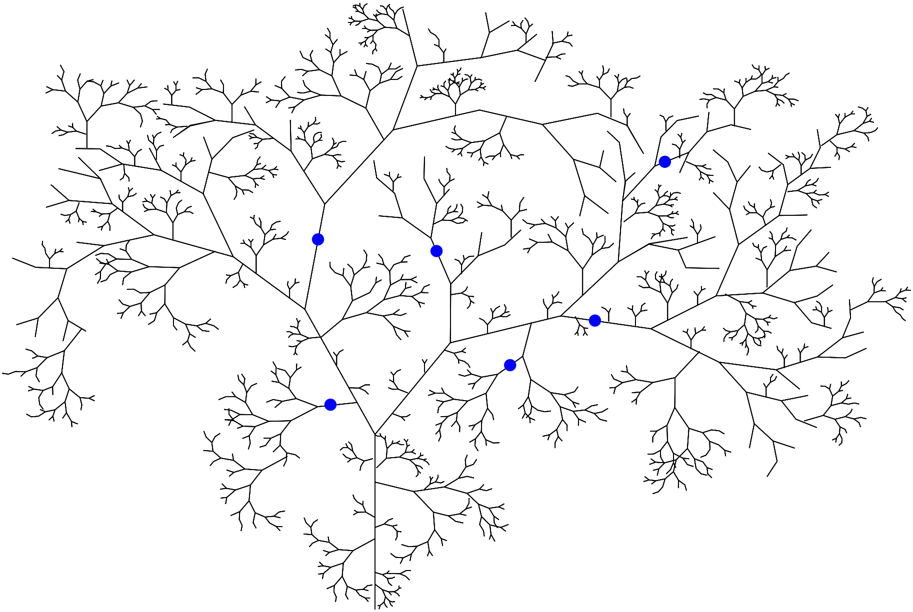


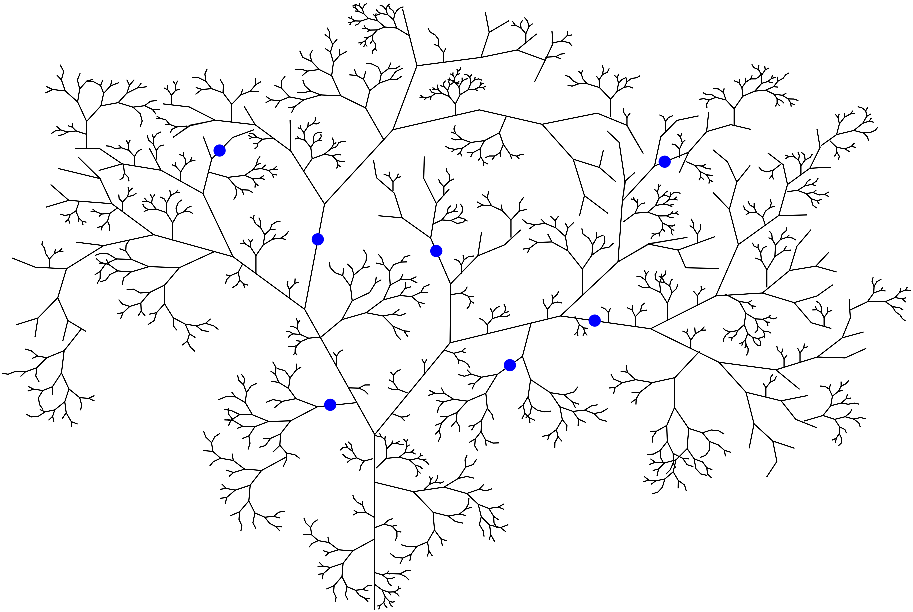


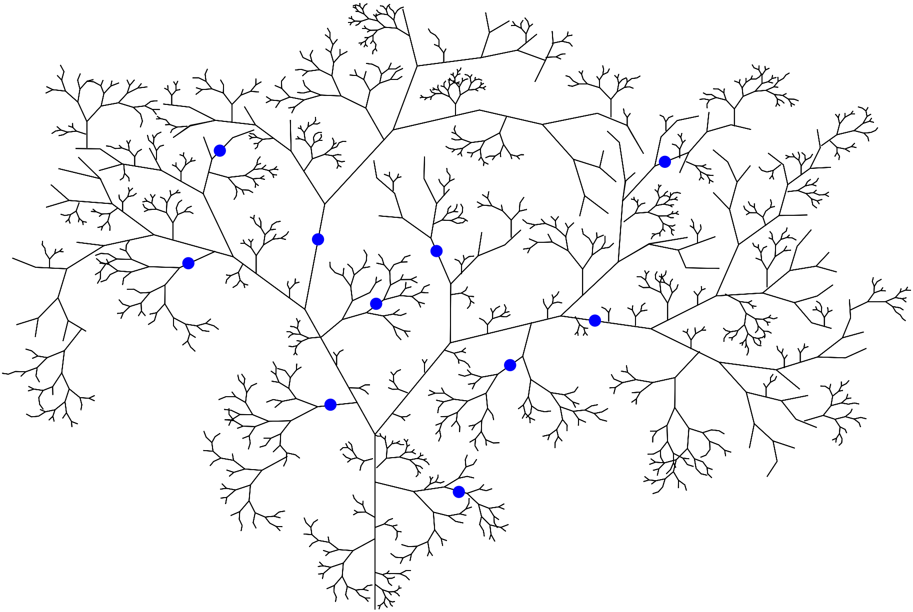


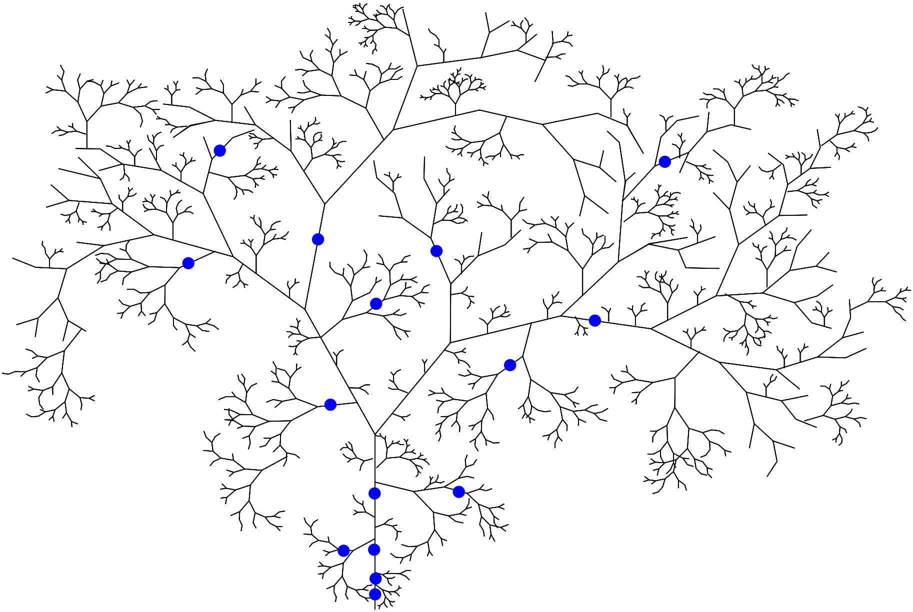


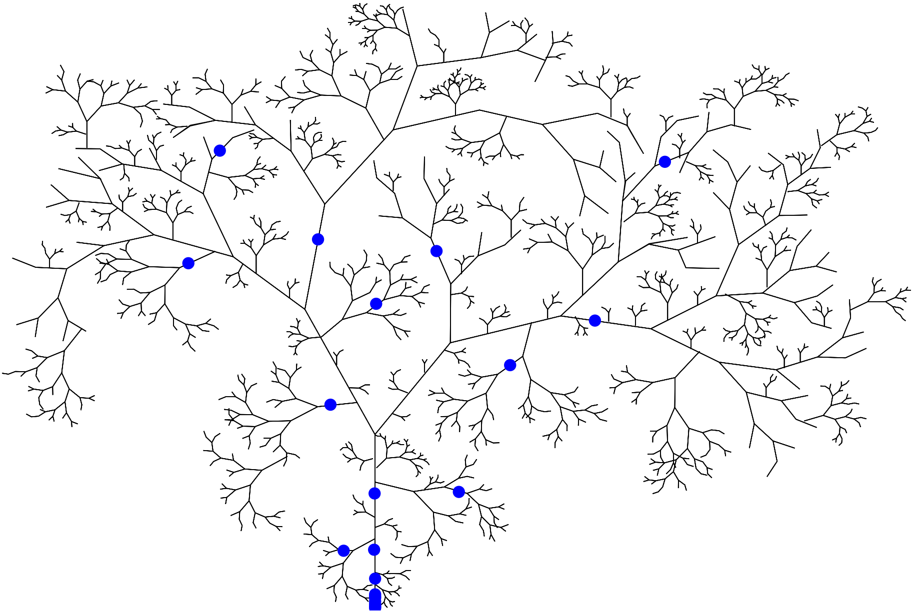


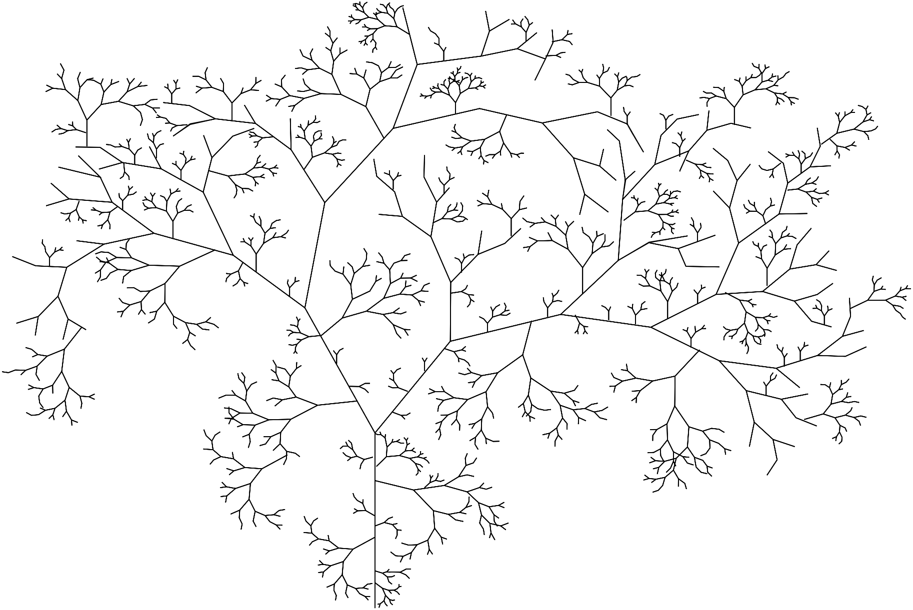


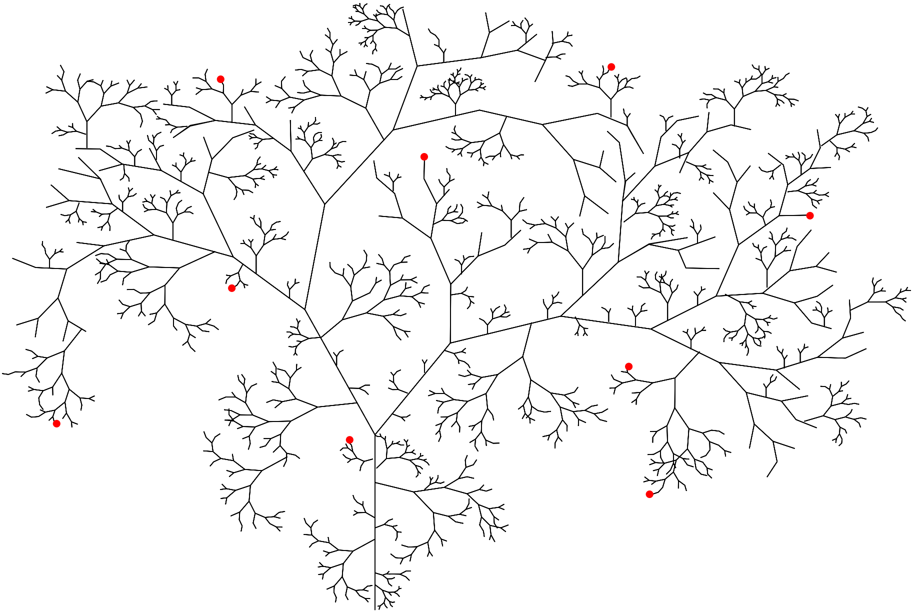


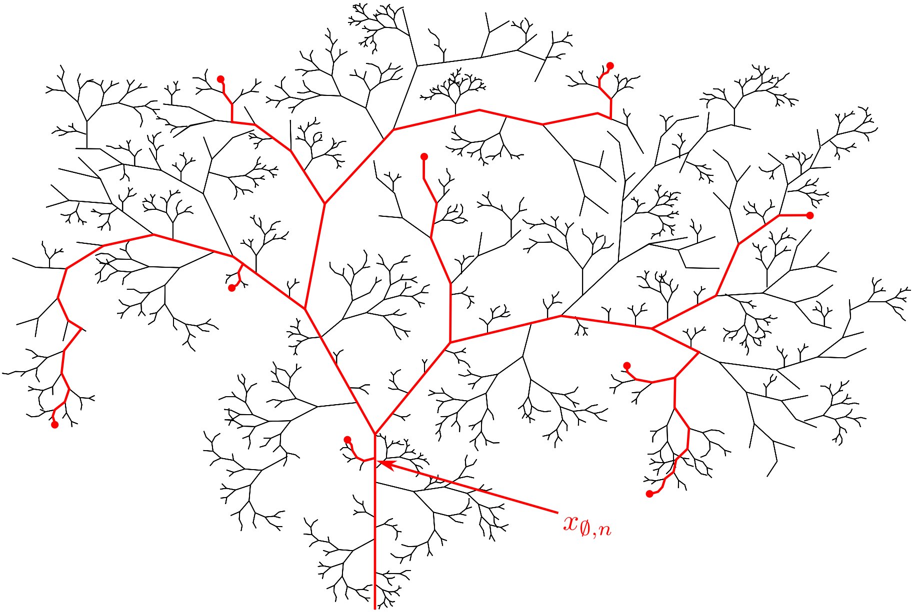


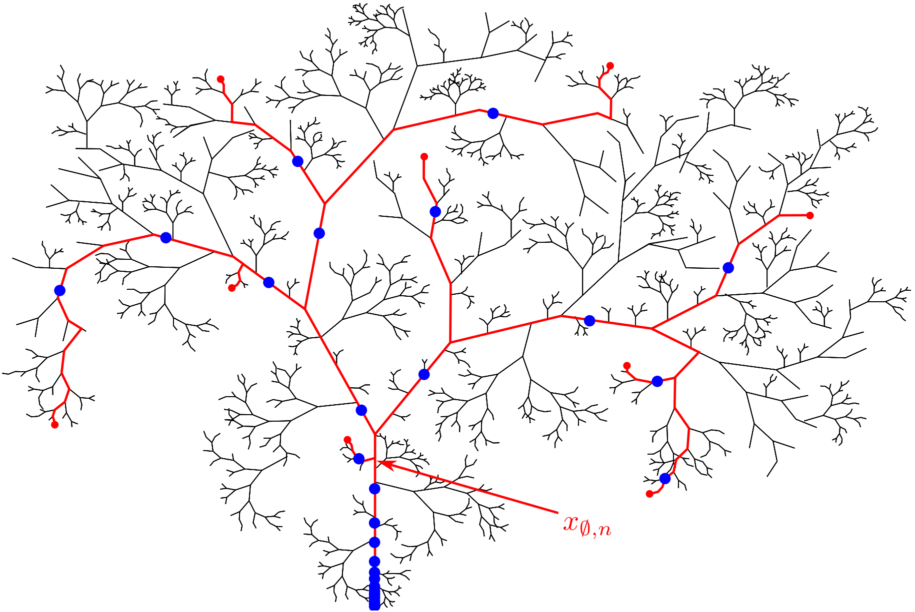









$x_{\emptyset,n}$


$x_{\emptyset,n}$

# Convergence presque sûre

**Théorème (Abraham–Delmas '11)**

*On a $\mathbb{P}_\infty$-p.s.*

$$\lim_{n\to\infty} \frac{X_n^*}{\sqrt{2n}} = \int_{\mathcal{T}} \theta(s)\,\mathbf{m}(ds) := \Theta.$$

*De plus, $\Theta$ est de loi Rayleigh sous $\mathbb{P}_\infty$.*

- Preuve : Introduction de martingales sur $\mathcal{T}$ associées au processus de fragmentation.

# Convergence presque sûre

**Théorème (Abraham–Delmas '11)**

*On a $\mathbb{P}_\infty$-p.s.*

$$\lim_{n\to\infty} \frac{X_n^*}{\sqrt{2n}} = \int_{\mathcal{T}} \theta(s)\, \mathbf{m}(ds) := \Theta.$$

*De plus, $\Theta$ est de loi Rayleigh sous $\mathbb{P}_\infty$.*

- Preuve : Introduction de martingales sur $\mathcal{T}$ associées au processus de fragmentation.
- Variations autour de cette limite ? Vitesse de convergence ?

# Processus de records sur $\mathbb{R}$

Sur $\mathbb{R}_+$, processus de records issu de $q \in (0, +\infty]$ :

- Défini grâce à un processus ponctuel de Poisson $\mathcal{M}_q(dx, dt)$ d'intensité $dx \otimes dt\ \mathbf{1}_{[0,q]}(t)$
- Tel que $\theta(0) = q$ p.s.
- Sachant $\theta(x) = t \in (0, q]$, le premier temps de saut de $\theta$ après $x$ sera de loi exponentielle, de paramètre $t$.

# Processus de records sur $\mathbb{R}$

Sur $\mathbb{R}_+$, processus de records issu de $q \in (0, +\infty]$ :

- Défini grâce à un processus ponctuel de Poisson $\mathcal{M}_q(dx, dt)$ d'intensité $dx \otimes dt\, \mathbf{1}_{[0,q]}(t)$
- Tel que $\theta(0) = q$ p.s.
- Sachant $\theta(x) = t \in (0, q]$, le premier temps de saut de $\theta$ après $x$ sera de loi exponentielle, de paramètre $t$.

## Proposition

*Le processus ponctuel de records sur $\mathbb{R}_+$ défini par $\sum_{s \in \mathbb{R}_+} \delta_s \mathbf{1}_{\theta(s-) > \theta(s)}$ est d'intensité (aléatoire) $\theta(s)\, ds$.*

# Deux martingales

Conséquences de la proposition précédente : pour tout $q \in (0, \infty)$,

- $\left(N_t = X([0,t]) - \int_0^t \theta(s)\, ds,\ t \geq 0\right)$ est une martingale
- $\left(N_t^2 - \int_0^t \theta(s)\, ds,\ t \geq 0\right)$ est *aussi* une martingale

# Deux martingales

Conséquences de la proposition précédente : pour tout $q \in (0, \infty)$,

- $\left(N_t = X([0,t]) - \int_0^t \theta(s)\, ds,\ t \geq 0\right)$ est une martingale
- $\left(N_t^2 - \int_0^t \theta(s)\, ds,\ t \geq 0\right)$ est *aussi* une martingale

Sur $\mathcal{T}_n^*$, on a

$$\mathbb{E}[X_n^* | \mathcal{T}_n^*] = \mathbb{E}\left[\int_{\mathcal{T}_n^*} \theta(s)\ \ell(ds) \middle| \mathcal{T}_n^*\right]$$

et

$$\mathbb{E}\left[\left(X_n^* - \int_{\mathcal{T}_n^*} \theta(s)\ \ell(ds)\right)^2 \middle| \mathcal{T}_n^*\right] = \mathbb{E}\left[\int_{\mathcal{T}_n^*} \theta(s)\ \ell(ds) \middle| \mathcal{T}_n^*\right].$$

Ceci permet de montrer que

$$\lim_{n\to\infty} \frac{X_n^*}{\sqrt{2n}} = \lim_{n\to\infty} \frac{1}{\sqrt{2n}} \int_{\mathcal{T}_n^*} \theta(s)\ \ell(ds)$$

Ceci permet de montrer que

$$\lim_{n\to\infty} \frac{X_n^*}{\sqrt{2n}} = \lim_{n\to\infty} \frac{1}{\sqrt{2n}} \int_{\mathcal{T}_n^*} \theta(s)\ \ell(ds)$$

On prouve alors que

$$\frac{1}{\sqrt{2n}} \int_{\mathcal{T}_n^*} \theta(s)\ell(ds) \simeq \mathbb{E}[\Theta|\mathcal{F}_n],$$

où $\mathcal{F}_n = \sigma(\{\mathcal{T}_n, (\theta(s),\ s \in \mathcal{T}_n)\})$, en utilisant une description de $(\mathcal{T}, \theta)$ conditionnellement à $\mathcal{F}_n$.

Ceci permet de montrer que

$$\lim_{n\to\infty} \frac{X_n^*}{\sqrt{2n}} = \lim_{n\to\infty} \frac{1}{\sqrt{2n}} \int_{\mathcal{T}_n^*} \theta(s)\ \ell(ds)$$

On prouve alors que

$$\frac{1}{\sqrt{2n}} \int_{\mathcal{T}_n^*} \theta(s)\ell(ds) \simeq \mathbb{E}[\Theta|\mathcal{F}_n],$$

où $\mathcal{F}_n = \sigma(\{\mathcal{T}_n, (\theta(s),\ s \in \mathcal{T}_n)\})$, en utilisant une description de $(\mathcal{T}, \theta)$ conditionnellement à $\mathcal{F}_n$.
Finalement, $\mathbb{E}[\Theta|\mathcal{F}_n] \to \Theta$, p.s. $\square$

# Plan de l'exposé


1. Coupures d'arbres aléatoires
   - Découpage d'arbres aléatoires discrets
   - Arbre aléatoire continu d'Aldous
   - Fragmentation d'Aldous-Pitman
2. Une limite presque sûre
   - Records sur le CRT
   - Convergence presque sûre
   - Esquisse de preuve
3. Fluctuations autour de la limite
   - Théorème principal
   - Une autre martingale !

# Théorème central limite

## Théorème (H '12)

*On a la convergence en loi suivante :*

$$\lim_{n\to\infty} \frac{X_n^* - \sqrt{2n}\,\Theta}{n^{1/4}} = 2^{1/4}\sqrt{\Theta}G,$$

*où G est une v.a. gaussienne indépendante.*

- Mélange de v.a. gaussiennes, avec une variance de loi Rayleigh (pas une loi connue).

Comment prouve-t-on une convergence vers une limite gaussienne ?

- Avec de l'indépendance (théorème de Lindeberg-Feller) : pas possible à cause de la structure d'arbre

Comment prouve-t-on une convergence vers une limite gaussienne ?

- Avec de l'indépendance (théorème de Lindeberg-Feller) : pas possible à cause de la structure d'arbre
- Trouver une martingale et utiliser l'indépendance conditionnelle :

### Theorem (Hall–Heyde)

*Soit $(M_n,\ n \geq 1)$ une $\mathcal{F}_n$-martingale de carré intégrable, telle qu'il existe une suite $a_n \uparrow \infty$ et une v.a. $\eta^2$, $\mathcal{F}_\infty$-measurable, finie p.s., telle que*

1. $a_n^{-2} \sum_{k=1}^{n} \mathbb{E}[(M_{k+1} - M_k)^2 | \mathcal{F}_k] \xrightarrow{\mathbb{P}} \eta^2$
2. *Pour tout $\varepsilon > 0$,*
   $a_n^{-2} \sum_{k=1}^{n} \mathbb{E}[(M_{k+1} - M_k)^2 \mathbf{1}_{|M_{k+1} - M_k| > \varepsilon a_n} | \mathcal{F}_k] \xrightarrow{\mathbb{P}} 0.$

*Alors, $a_n^{-1} M_n$ converge en loi vers une v.a. $Z$, dont la fonction caractéristique est $\mathbb{E}[\exp(-t^2\eta^2/2)]$.*

# Une autre martingale

Choix évident :

$$M_n^* = X_n^* - \int_{\mathcal{T}_n^*} \theta(s)\ell(ds)$$

Problème : $(M_n^*,\ n \geq 1)$ pas une martingale dans les filtrations utiles.
On considère la filtration

$$\mathcal{G}_n = \sigma(\{\mathcal{T},\ (\theta(s),\ s \in \mathcal{T}_n)\})$$

et le processus

$$\widehat{M_n} = \sum_{s \in \mathcal{T}_n \setminus \mathcal{T}_1} \mathbf{1}_{\theta(s-) > \theta(s)} - \int_{\mathcal{T}_n \setminus \mathcal{T}_1} \theta(s)\ell(ds),$$

qui est une $\mathcal{G}_n$-martingale.

# Convergence de la variance conditionnelle

Sur la branche $B_{k+1}$, conditionnellement à $\mathcal{G}_k$,

- $\widehat{M}_{k+1} - \widehat{M_k}$ est distribué comme $N_{\ell(B_{k+1})}$, pour un processus de records sur $\mathbb{R}_+$ issu de $\theta(s_{k+1})$, donc

$$\mathbb{E}\left[ (\widehat{M}_{k+1} - \widehat{M}_k)^2 \middle| \mathcal{G}_k \right] = \mathbb{E}\left[ \int_{B_{k+1}} \theta(s)\ell(ds) \middle| \mathcal{G}_k \right]$$

# Convergence de la variance conditionnelle

Sur la branche $B_{k+1}$, conditionnellement à $\mathcal{G}_k$,

- $\widehat{M}_{k+1} - \widehat{M_k}$ est distribué comme $N_{\ell(B_{k+1})}$, pour un processus de records sur $\mathbb{R}_+$ issu de $\theta(s_{k+1})$, donc

$$\mathbb{E}\left[ (\widehat{M}_{k+1} - \widehat{M}_k)^2 \middle| \mathcal{G}_k \right] = \mathbb{E}\left[ \int_{B_{k+1}} \theta(s)\ell(ds) \middle| \mathcal{G}_k \right]$$

- Par la Loi des Grands Nombres,

$$n^{-1/2} \sum_{k=1}^{n} \mathbb{E}\left[ \int_{B_{k+1}} \theta(s)\ell(ds) \middle| \mathcal{G}_k \right] \simeq n^{-1/2} \sum_{k=1}^{n} \int_{B_{k+1}} \theta(s)\ell(ds)$$

# Convergence de la variance conditionnelle

Sur la branche $B_{k+1}$, conditionnellement à $\mathcal{G}_k$,

- $\widehat{M}_{k+1} - \widehat{M_k}$ est distribué comme $N_{\ell(B_{k+1})}$, pour un processus de records sur $\mathbb{R}_+$ issu de $\theta(s_{k+1})$, donc

$$\mathbb{E}\left[ (\widehat{M}_{k+1} - \widehat{M}_k)^2 \middle| \mathcal{G}_k \right] = \mathbb{E}\left[ \int_{B_{k+1}} \theta(s)\ell(ds) \middle| \mathcal{G}_k \right]$$

- Par la Loi des Grands Nombres,

$$n^{-1/2} \sum_{k=1}^{n} \mathbb{E}\left[ \int_{B_{k+1}} \theta(s)\ell(ds) \middle| \mathcal{G}_k \right] \simeq n^{-1/2} \sum_{k=1}^{n} \int_{B_{k+1}} \theta(s)\ell(ds)$$

- $n^{-1/2} \sum_{k=1}^{n} \int_{B_{k+1}} \theta(s)\ell(ds) \to \Theta$ (Abraham–Delmas)

Donc, on obtient $n^{-1/2} \sum_{k=1}^{n} \mathbb{E}[(\widehat{M}_{k+1} - \widehat{M}_k)^2 | \mathcal{G}_k] \xrightarrow{\mathbb{P}} \Theta$

# Petitesse asymptotique

Pour montrer que, pour tout $\varepsilon > 0$,

$$n^{-1/2} \sum_{k=1}^{n} \mathbb{E}\left[ (\widehat{M}_{k+1} - \widehat{M}_k)^2 \, \mathbf{1}_{|\widehat{M}_{k+1} - \widehat{M}_k| > \varepsilon n^{1/4}} \middle| \mathcal{G}_k \right] \xrightarrow{\mathbb{P}} 0,$$

on utilise un argument de Liapounov, en bornant les quatrièmes moments des incréments.
Donc, on obtient une limite gaussienne pour $n^{-1/4}\widehat{M}_n$, de variance aléatoire $\Theta$.

# Nettoyage

- Comportement de $M_n^* = X_n^* - \int_{\mathcal{T}_n^*} \theta(s)\ell(ds)$ très semblable à $\widehat{M}_n$ (variance de l'épine dorsale négligeable)
- Mêmes approximations que dans [Abraham–Delmas] pour réduire le problème à
$$n^{1/4}(\mathbb{E}[\Theta|\mathcal{F}_n] - \Theta) \xrightarrow{(d)} 0$$
- Pour montrer cela, théorèmes limites de martingales + description conditionnelle de $\mathcal{T}$ sachant $\mathcal{T}_n$.

# Nettoyage

- Comportement de $M_n^* = X_n^* - \int_{\mathcal{T}_n^*} \theta(s)\ell(ds)$ très semblable à $\widehat{M}_n$ (variance de l'épine dorsale négligeable)
- Mêmes approximations que dans [Abraham–Delmas] pour réduire le problème à
$$n^{1/4}(\mathbb{E}[\Theta|\mathcal{F}_n] - \Theta) \xrightarrow{(d)} 0$$
- Pour montrer cela, théorèmes limites de martingales + description conditionnelle de $\mathcal{T}$ sachant $\mathcal{T}_n$.

Merci de votre attention !