Processus de largeur et de contour des arbres continus inhomogènes

PTRF **129**, 182–218 (2004)

David Aldous, Grégory Miermont & Jim Pitman

Dept. of Statistics, UC Berkeley

& DMA, École Normale Supérieure

Paris, France

Journées M.A.S. de la S.M.A.I., 6–8 septembre 2004

Introduction: arbre brownien

- On considère $n \in \mathbb{N}$, et on note $\mathbf{T}_n$ l'ensemble des arbres à sommets dans $[n] := \{1, \ldots, n\}$, enracinés. Soit $\mathcal{T}^n \in \mathbf{T}_n$ pris uniformément parmi les n^{n-1} possibles.

Introduction: arbre brownien

- On considère $n \in \mathbb{N}$, et on note $\mathbf{T}_n$ l'ensemble des arbres à sommets dans $[n] := \{1, \ldots, n\}$, enracinés. Soit $\mathcal{T}^n \in \mathbf{T}_n$ pris uniformément parmi les n^{n-1} possibles.
- La distance "typique" de deux points pris au hasard sur $\mathcal{T}^n$ est $O(\sqrt{n})$. On assigne donc à chaque arête de $\mathcal{T}^n$ a une longueur $1/\sqrt{n}$, ce qui en fait un espace métrique (encore appelé $\mathcal{T}^n$).

Introduction: arbre brownien

- On considère $n \in \mathbb{N}$, et on note $\mathbf{T}_n$ l'ensemble des arbres à sommets dans $[n] := \{1, \ldots, n\}$, enracinés. Soit $\mathcal{T}^n \in \mathbf{T}_n$ pris uniformément parmi les n^{n-1} possibles.
- La distance "typique" de deux points pris au hasard sur $\mathcal{T}^n$ est $O(\sqrt{n})$. On assigne donc à chaque arête de $\mathcal{T}^n$ a une longueur $1/\sqrt{n}$, ce qui en fait un espace métrique (encore appelé $\mathcal{T}^n$).
- On fait tendre $n \to \infty$. Alors $\mathcal{T}^n$ converge en loi vers $\mathcal{T}$, l'arbre continu d'Aldous (1991).

Arbre brownien, excursion brownienne

Soit e (deux fois) l'excursion brownienne standard. On note

$$d(s,s') = e(s) + e(s') - 2 \inf_{s \wedge s' \leq u \leq s \vee s'} e(u),$$

et $s \sim s' \iff d(s,s') = 0$.

Arbre brownien, excursion brownienne

Soit e (deux fois) l'excursion brownienne standard. On note

$$d(s,s') = e(s) + e(s') - 2 \inf_{s \wedge s' \leq u \leq s \vee s'} e(u),$$

et $s \sim s' \iff d(s,s') = 0$. Alors $([0,1]/\sim, d)$ est un espace métrique "égal en loi" à $\mathcal{T}$ (Aldous, Le Gall, 1993).

Arbre brownien, excursion brownienne

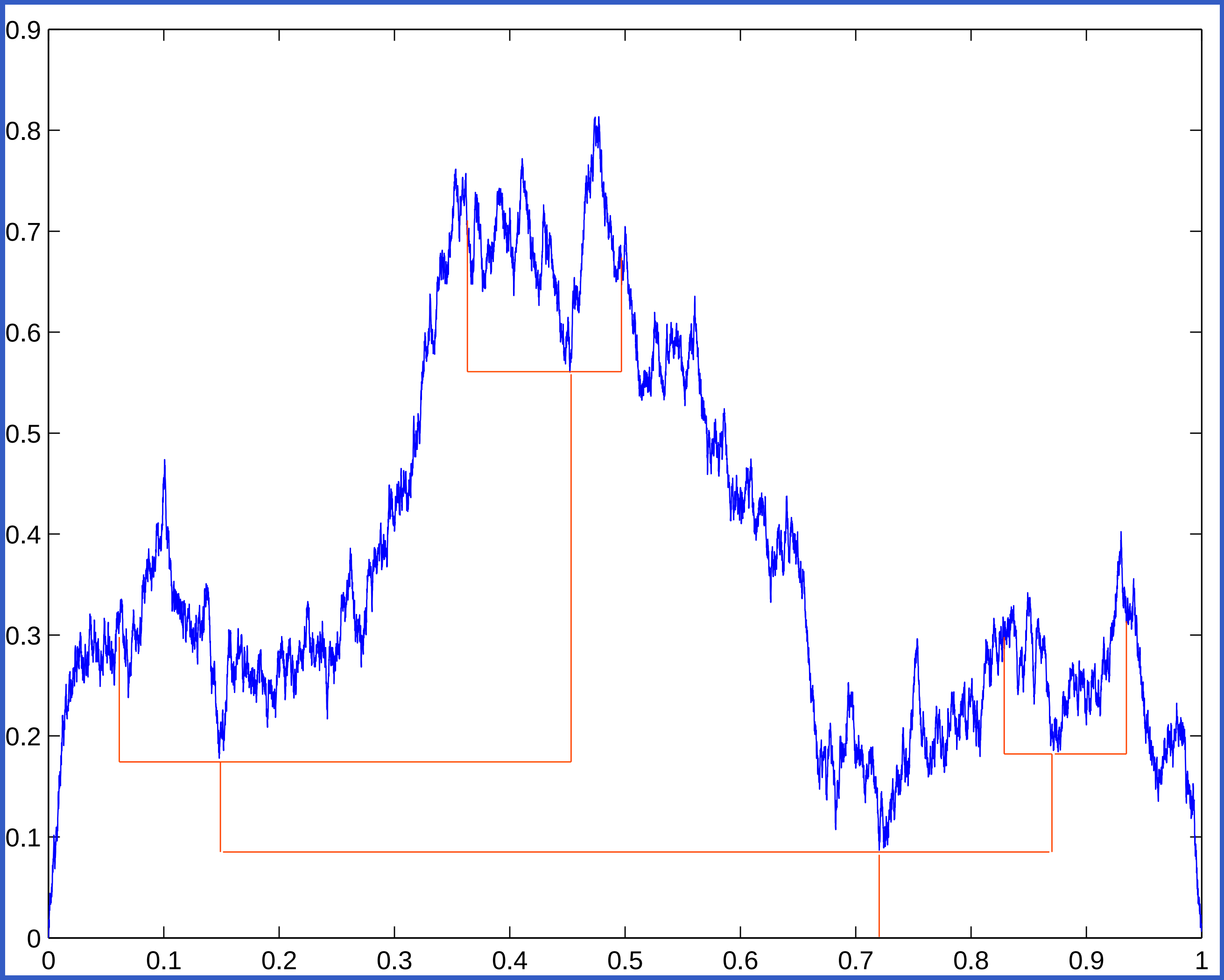
Soit e (deux fois) l'excursion brownienne standard. On note

$$d(s,s') = e(s) + e(s') - 2 \inf_{s \wedge s' \leq u \leq s \vee s'} e(u),$$

et $s \sim s' \iff d(s,s') = 0$. Alors $([0,1]/\sim, d)$ est un espace métrique “égal en loi” à $\mathcal{T}$ (Aldous, Le Gall, 1993).

On dit alors que e est le processus de contour de $\mathcal{T}$.

Un dessin



Motivation

- On s'intéresse à une généralisation de $\mathcal{T}$, appelée ICRT (Arbre Continu Aléatoire Inhomogène), obtenu comme limite d'arbres plus généraux que $\mathcal{T}^n$, appelés $\mathbf{p}$-arbres, où $\mathbf{p} = (p_1, \ldots, p_n)$ est une probabilité sur $[n]$.
- Le $\mathbf{p}$-arbre $\mathcal{T}^{\mathbf{p}}$ est déterminé par la loi

$$P(\mathcal{T}^{\mathbf{p}} = \mathbf{t}) = \prod_{i=1}^{n} p_i^{c_i(\mathbf{t})}, \quad \mathbf{t} \in \mathbf{T}_n$$

($c_i(\mathbf{t}) = \#\{$enfants de i dans $\mathbf{t}\}$).

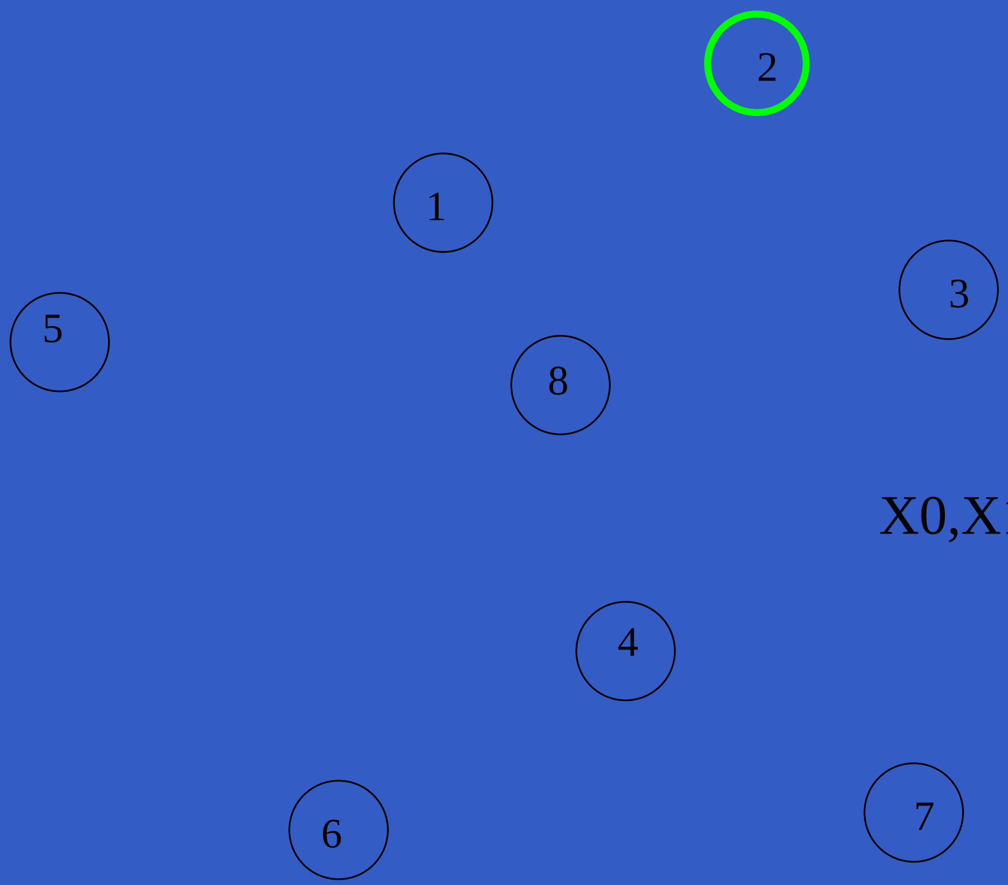
p-arbres

- Ce sont des objets naturels en combinatoire, leur loi est motivée par la formule de Cayley:

$$(x_1 + \ldots + x_n)^{n-1} = \sum_{\mathbf{t} \in \mathbf{T}_n} x_i^{c_i(\mathbf{t})}.$$

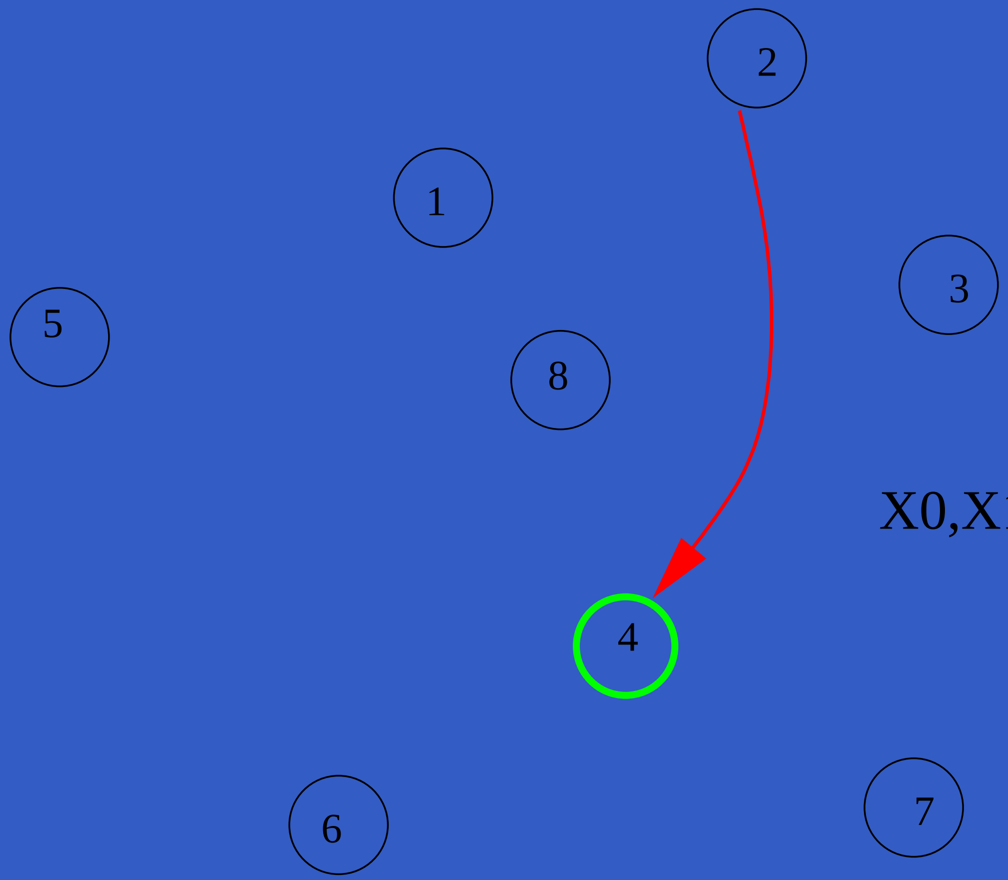
- On les retrouve dans l'étude du coalescent stochastique additif (Pitman 1996, Aldous-Pitman 1998, 1999), et dans l'étude de certains modèles naturels d'applications de $[n]$ dans $[n]$ aléatoires (Aldous-Miermont-Pitman 2004).

Construction d'un p-arbre



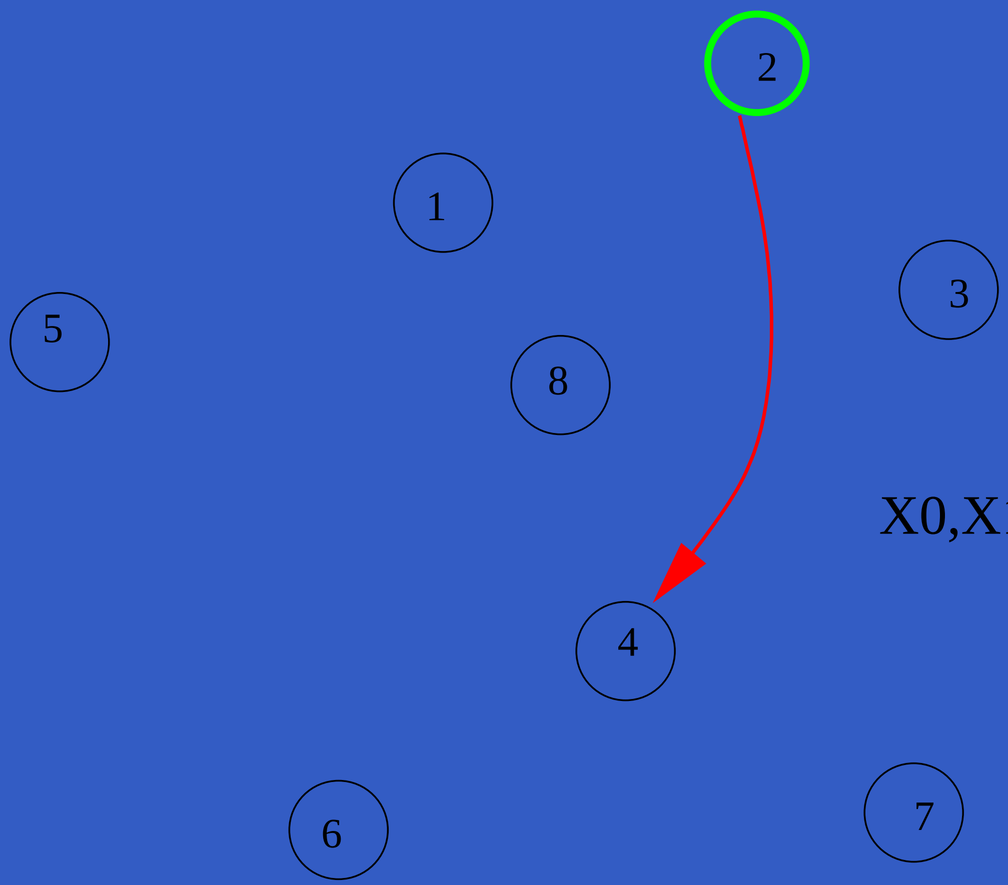
X0,X1,...=2,4,2,1,3,4,6,1,5,6,8,4,7,...

Construction d'un p-arbre



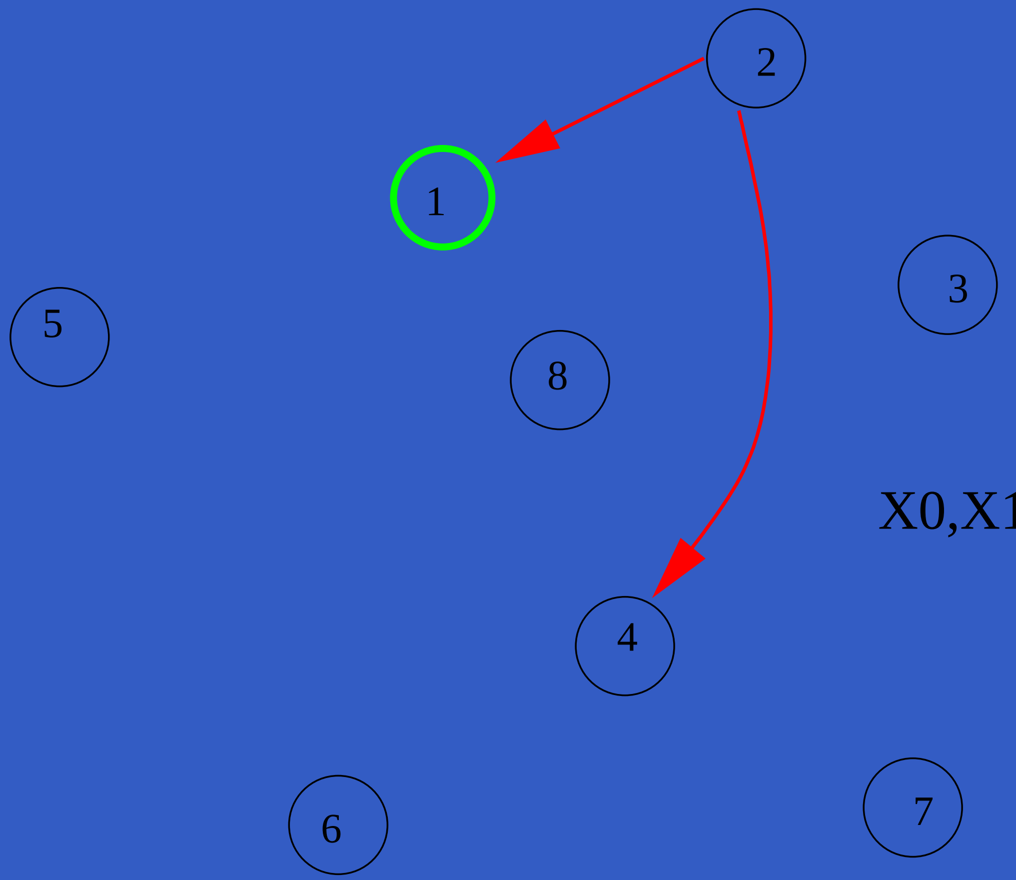
X0,X1,...=2,4,2,1,3,4,6,1,5,6,8,4,7,...

Construction d’un p-arbre



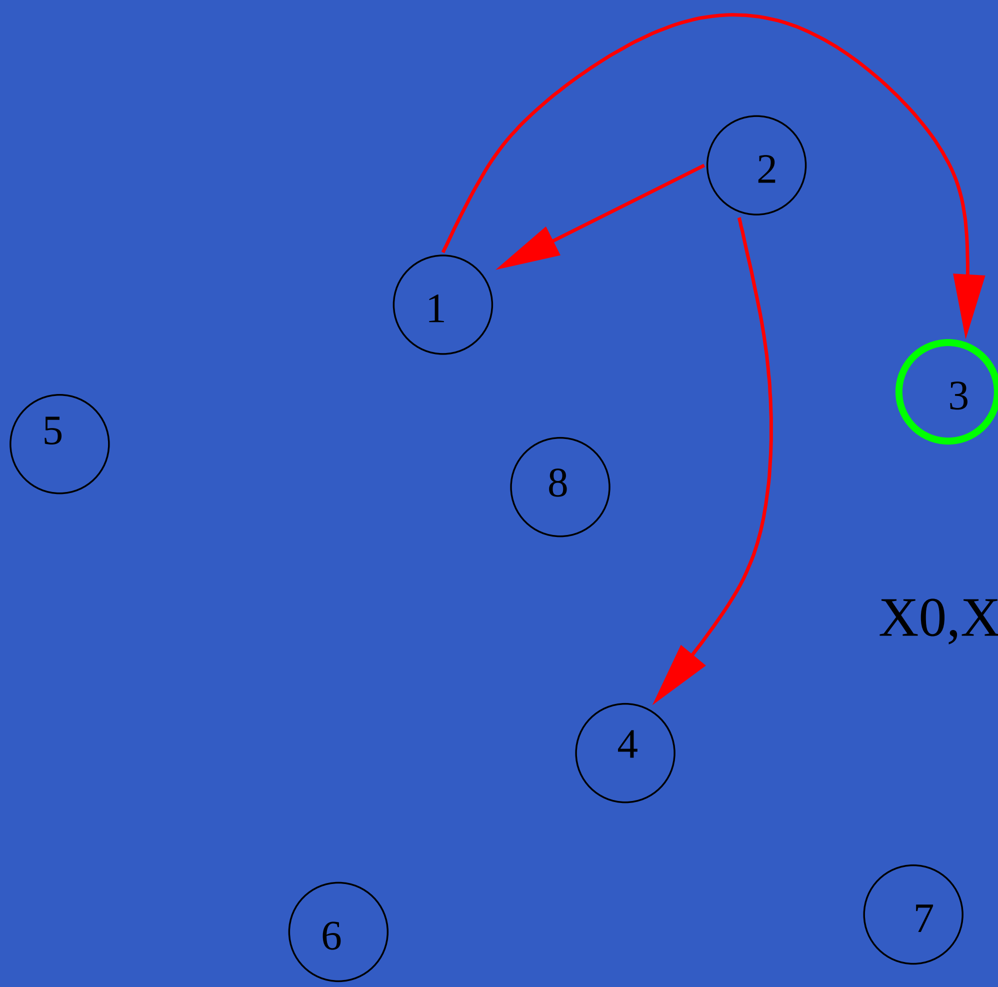
$X_0, X_1, \ldots = 2,4,2,1,3,4,6,1,5,6,8,4,7,\ldots$

Construction d'un p-arbre



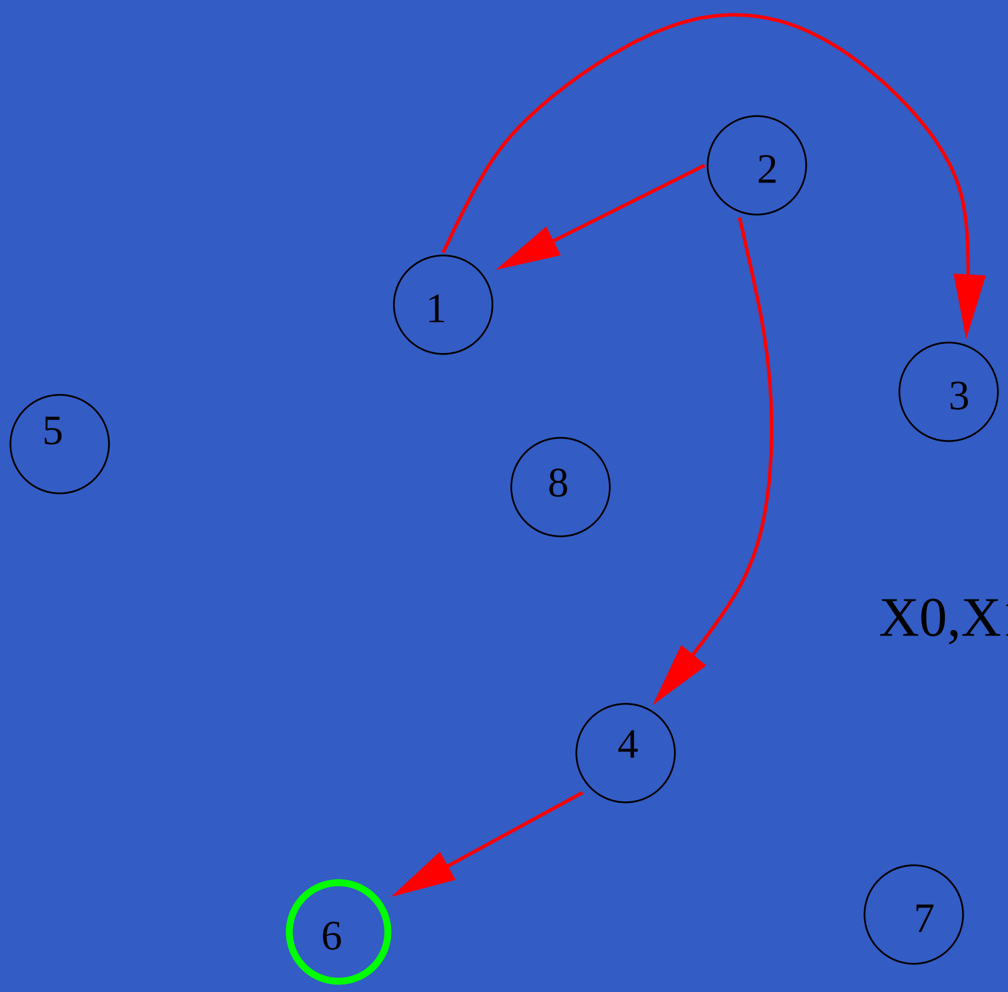
X0,X1,...=2,4,2,1,3,4,6,1,5,6,8,4,7,...

Construction d’un p-arbre



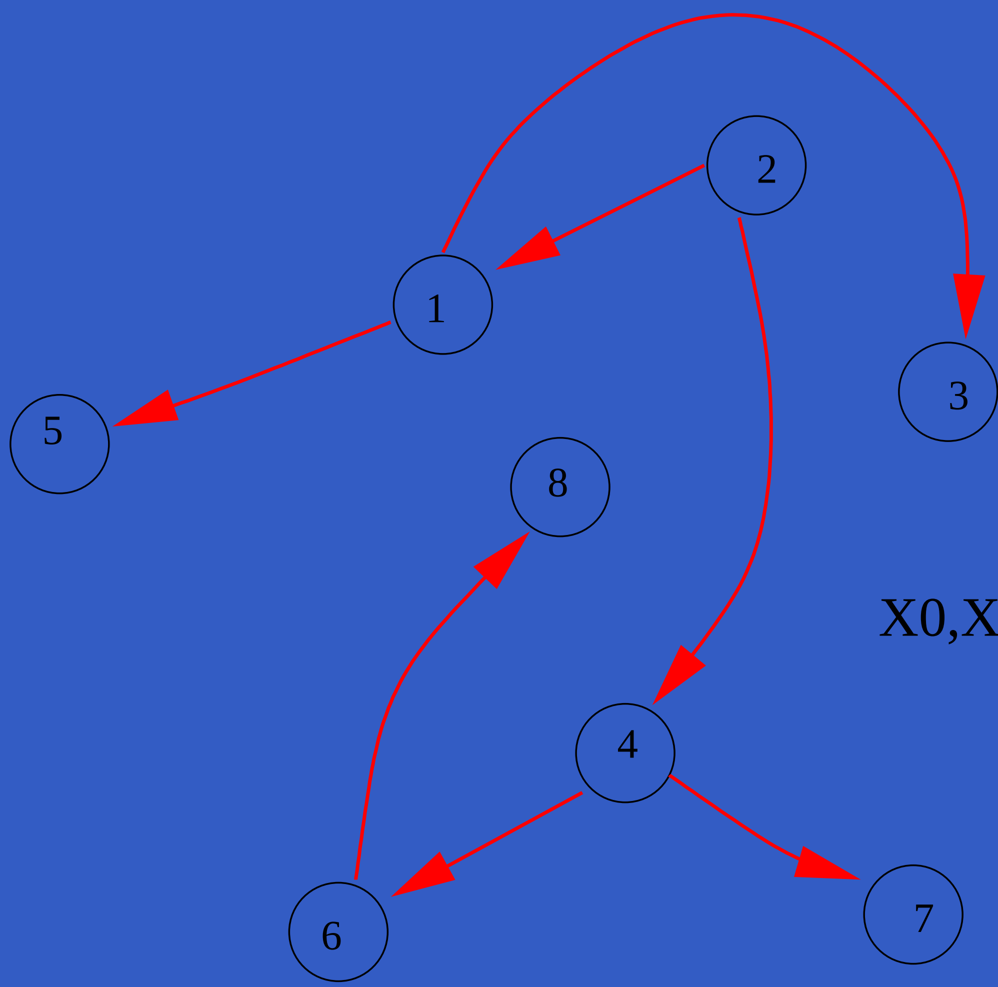
X0,X1,...=2,4,2,1,3,4,6,1,5,6,8,4,7,...

Construction d'un p-arbre



X0,X1,...=2,4,2,1,3,4,6,1,5,6,8,4,7,...

Construction d'un p-arbre



X0,X1,...=2,4,2,1,3,4,6,1,5,6,8,4,7,...

Les ICRT

Théorème (Camarri-Pitman 1997)
Soit $\sigma(\mathbf{p}) = \sqrt{\sum_{i=1}^{n} p_i^2}$. On suppose que, quand $n \to \infty$,

$$\max_i p_i \to 0 \quad , \quad \frac{p_i}{\sigma(\mathbf{p})} \to \theta_i, i \geq 1,$$

où $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots)$ vérifie $\sum_{i\geq 1} \theta_i^2 := 1 - \theta_0^2 \leq 1$.

Les ICRT

Théorème (Camarri-Pitman 1997)
Soit $\sigma(\mathbf{p}) = \sqrt{\sum_{i=1}^n p_i^2}$. On suppose que, quand $n \to \infty$,

$$\max_i p_i \to 0 \quad , \quad \frac{p_i}{\sigma(\mathbf{p})} \to \theta_i, i \geq 1,$$

où $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots)$ vérifie $\sum_{i\geq 1} \theta_i^2 := 1 - \theta_0^2 \leq 1$. Alors, l'arbre $\mathcal{T}^{\mathbf{p}}$ où chaque arête a pour longueur $\sigma(\mathbf{p})$, converge vers un arbre continu $\mathcal{T}^{\theta}$ (en particulier $\mathcal{T}^0 = \mathcal{T}$).

Questions

1. Peut-on trouver le processus de contour $(\mathcal{E}^\theta(s), 0 \leq s \leq 1)$ de $\mathcal{T}^\theta$?
2. Si oui, pour $h \geq 0$, soit

$$\bar{\mathcal{W}}^\theta(h) = \text{Leb}\{s \in [0,1] : \mathcal{E}^\theta(s) \leq h\}$$

   la “masse” des individus en-dessous du niveau h.
   A-t-on $\mathrm{d}\bar{\mathcal{W}}^\theta(h) = \mathcal{W}^\theta(h)\mathrm{d}h$? On appelle alors $\mathcal{W}^\theta$ le processus de largeur.

Dans le cas brownien...

Dans le cas $\theta = 0$ (arbre brownien), c'est un théorème de Jeulin (1985), ou théorème de Ray-Knight conditionné, reliant les temps locaux d'une excursion brownienne à une excursion brownienne

Dans le cas brownien...

Dans le cas $\theta = 0$ (arbre brownien), c'est un théorème de Jeulin (1985), ou théorème de Ray-Knight conditionné, reliant les temps locaux d'une excursion brownienne à une excursion brownienne changée de temps

Dans le cas brownien...

Dans le cas $\theta = 0$ (arbre brownien), c'est un théorème de Jeulin (1985), ou théorème de Ray-Knight conditionné, reliant les temps locaux d'une excursion brownienne à une excursion brownienne changée de temps (un processus de branchement quadratique)

Dans le cas brownien...

Dans le cas $\theta = 0$ (arbre brownien), c'est un théorème de Jeulin (1985), ou théorème de Ray-Knight conditionné, reliant les temps locaux d'une excursion brownienne à une excursion brownienne changée de temps (un processus de branchement quadratique conditionné à avoir une population totale 1)

Dans le cas brownien...

Dans le cas $\theta = 0$ (arbre brownien), c'est un théorème de Jeulin (1985), ou théorème de Ray-Knight conditionné, reliant les temps locaux d'une excursion brownienne à une excursion brownienne changée de temps (un processus de branchement quadratique conditionné à avoir une population totale 1, ou un Bessel$^2(0)$ conditionné).

Réponses: “oui” et oui

Réponses: “oui” et oui

- Soit $(U_i, i \geq 1)$ indépendantes uniformes$[0, 1]$,

Réponses: “oui” et oui

- Soit $(U_i, i \geq 1)$ indépendantes uniformes$[0,1]$, B^{br} un pont brownien standard,

Réponses: “oui” et oui

- Soit $(U_i, i \geq 1)$ indépendantes uniformes$[0,1]$, B^{br} un pont brownien standard, et

$$X_t^{\mathrm{br},\theta} = \theta_0 B_t^{\mathrm{br}} + \sum_{i=1}^{\infty} \theta_i (\mathbb{1}_{\{U_i \leq t\}} - t), \quad 0 \leq t \leq 1,$$

un pont à accroissements échangeables.

Réponses: “oui” et oui

- Soit $(U_i, i \geq 1)$ indépendantes uniformes$[0, 1]$, B^{br} un pont brownien standard, et

$$X_t^{\mathrm{br},\theta} = \theta_0 B_t^{\mathrm{br}} + \sum_{i=1}^{\infty} \theta_i (\mathbb{1}_{\{U_i \leq t\}} - t), \quad 0 \leq t \leq 1,$$

- Soit

$$X_t^{\theta} = X_{t+t_{\min}}^{\mathrm{br},\theta} - X_{t_{\min}}^{\mathrm{br},\theta}, \quad 0 \leq t \leq 1$$

Réponses: "oui" et oui

- Soit $(U_i, i \geq 1)$ indépendantes uniformes$[0,1]$, B^{br} un pont brownien standard, et

$$X_t^{\mathrm{br},\theta} = \theta_0 B_t^{\mathrm{br}} + \sum_{i=1}^{\infty} \theta_i (\mathbb{1}_{\{U_i \leq t\}} - t), \quad 0 \leq t \leq 1,$$

- Soit

$$X_t^{\theta} = X_{t+t_{\min}}^{\mathrm{br},\theta} - X_{t_{\min}}^{\mathrm{br},\theta}, \quad 0 \leq t \leq 1$$

où $t_{\min}$ est tel que $X^{\mathrm{br},\theta}(t_{\min}) = \inf X^{\mathrm{br},\theta}$

Réponses: “oui” et oui

- Soit $(U_i, i \geq 1)$ indépendantes uniformes$[0,1]$, B^{br} un pont brownien standard, et

$$X_t^{\mathrm{br},\theta} = \theta_0 B_t^{\mathrm{br}} + \sum_{i=1}^{\infty} \theta_i (\mathbb{1}_{\{U_i \leq t\}} - t), \quad 0 \leq t \leq 1,$$

- Soit

$$X_t^{\theta} = X_{t+t_{\min}}^{\mathrm{br},\theta} - X_{t_{\min}}^{\mathrm{br},\theta}, \quad 0 \leq t \leq 1$$

sa transformée de Vervaat.

Réponse au second problème

Théorème
On suppose que $\theta_0 > 0$ ou $\sum_i \theta_i = \infty$. Le processus des hauteurs $\mathcal{W}^\theta$ existe. Sa loi est caractérisée par une transformation de type Lamperti de X^θ:

$$\mathcal{W}^\theta \circ (\bar{\mathcal{W}}^\theta)^{-1} \stackrel{\mathcal{L}}{=} X^\theta.$$

Ou encore

$$\mathcal{W}^\theta \stackrel{\mathcal{L}}{=} X^\theta \circ \rho, \quad \rho(h) = \inf\{r \geq 0 : \int_0^r \frac{\mathrm{d}s}{X^\theta(s)} > h\}.$$

Idée de la démonstration

On “approxime” les ICRT par les $\mathbf{p}$-arbres.
Soient $U_i, i \geq 1$ i.i.d. uniformes $[0,1]$. On pose

$$F^{\mathbf{p}}_s = \sum_{i=1}^{n} p_i(\mathbb{1}_{\{s \geq U_i\}} - s),$$

et $F^{\mathrm{exc},\mathbf{p}}$ sa transformée de Vervaat.

Idée de la démonstration

On “approxime” les ICRT par les $\mathbf{p}$-arbres.
Soient $U_i, i \geq 1$ i.i.d. uniformes $[0,1]$. On pose

$$F_s^{\mathbf{p}} = \sum_{i=1}^{n} p_i (\mathbb{1}_{\{s \geq U_i\}} - s),$$

et $F^{\mathrm{exc},\mathbf{p}}$ sa transformée de Vervaat.
Soient $0 = V_0 \leq V_1 \leq V_2, \ldots$ les instants de saut de $F^{\mathrm{exc},\mathbf{p}}$, de sorte que $\Delta F^{\mathrm{exc},\mathbf{p}}(V_k) = p_{\pi(k)}$.

On pose $u^{\mathbf{p}}(k) = p_{\pi(0)} + \ldots + p_{\pi(k-1)}$.

Idée de la démonstration

On construit un arbre comme suit: les enfants de i sont les j tels que $V_{\pi^{-1}(j)} \in]u^{\mathbf{p}}(\pi^{-1}(i)), u^{\mathbf{p}}(\pi^{-1}(i)+1)]$.
C’est un $\mathbf{p}$-arbre.

Idée de la démonstration

On construit un arbre comme suit: les enfants de i sont les j tels que $V_{\pi^{-1}(j)} \in]u^{\mathbf{p}}(\pi^{-1}(i)), u^{\mathbf{p}}(\pi^{-1}(i)+1)]$.
C'est un $\mathbf{p}$-arbre. De plus, si $\mathrm{ht}(i)$ est la hauteur de i dans cet arbre, et si

- $\bar{W}^{\mathbf{p}}(k) = \sum_{i=1}^{n} p_i \mathbb{1}_{\{\mathrm{ht}(i) \leq k-1\}}$
- $W^{\mathbf{p}}(k) = \sum_{i=1}^{n} p_i \mathbb{1}_{\{\mathrm{ht}(i) = k\}}$

alors $F^{\mathrm{exc},\mathbf{p}}(\bar{W}^{\mathbf{p}}(k)) = W^{\mathbf{p}}(k)$.

Idée de la démonstration

On utilise alors un théorème limite de Kallenberg (1974): sous les hypothèses sur $\mathbf{p}$,

$$\sigma(\mathbf{p})F^{\mathbf{p}} \xrightarrow{\mathcal{L}} X^{\mathrm{br},\theta},$$

et $\bar{W}^{\mathbf{p}}([\sigma(\mathbf{p})^{-1}h]) \to \bar{\mathcal{W}}^{\theta}(h)$, de sorte qu'à la limite

$$X(\bar{\mathcal{W}}^{\theta}(h)) = \mathcal{W}^{\theta}(h),$$

où $X =^{\mathcal{L}} X^{\theta}$.

Et pour la question 1 ?

On suppose $\theta_0 > 0$ et $\theta_I = 0$ pour un I assez grand. On définit un processus

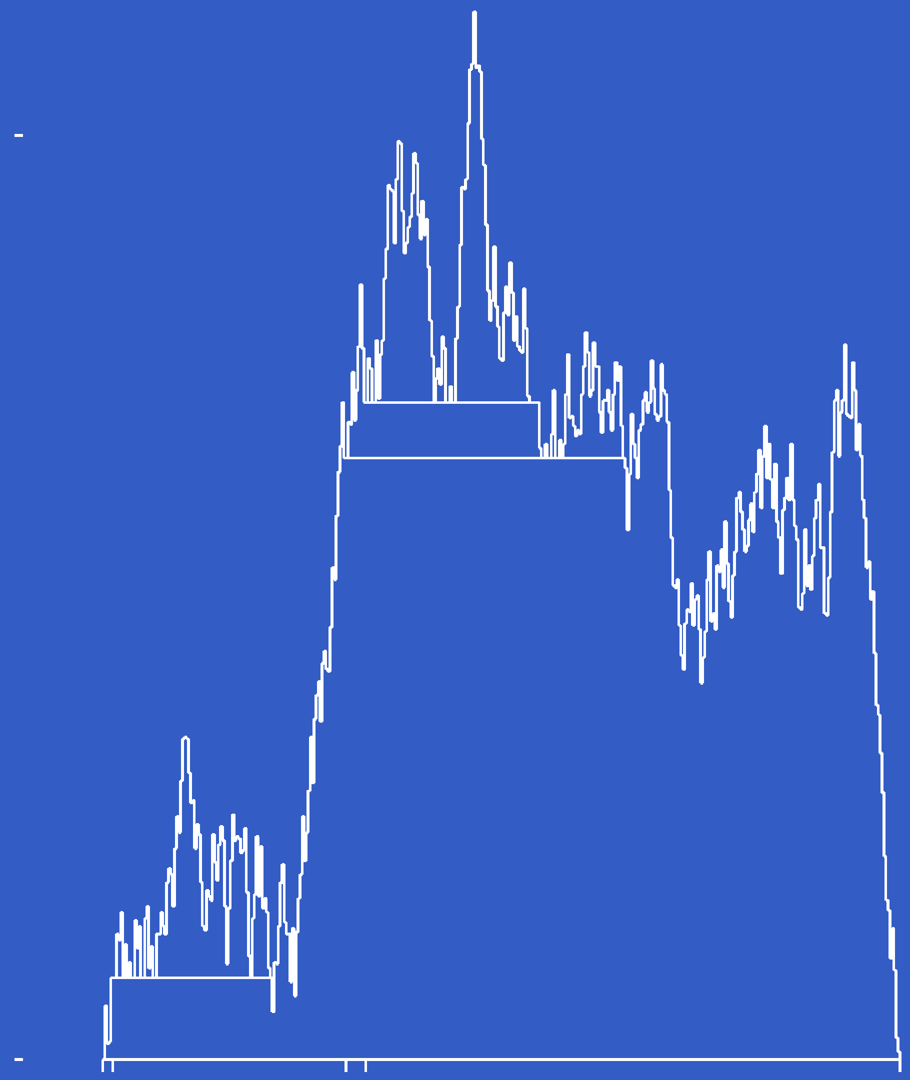
$$Y^\theta(s) = \text{Leb}\left\{\inf_{0\le u\le r} X^\theta(u), 0 \le r \le s\right\}.$$

Plus visuellement, on “gomme” les sauts de X^θ par un procédé de “réflexion”.

Théorème

Le processus de contour de $\mathcal{T}^\theta$ existe et vaut $\frac{2}{\theta_0^2} Y^\theta$.

Un dessin



Idée de la démonstration

On réinterprète la fonction $F^{\mathrm{exc},\mathbf{p}}$ de façon différente pour construire un autre $\mathbf{p}$-arbre. Au lieu d'un parcours en largeur, on fait un parcours en profondeur:

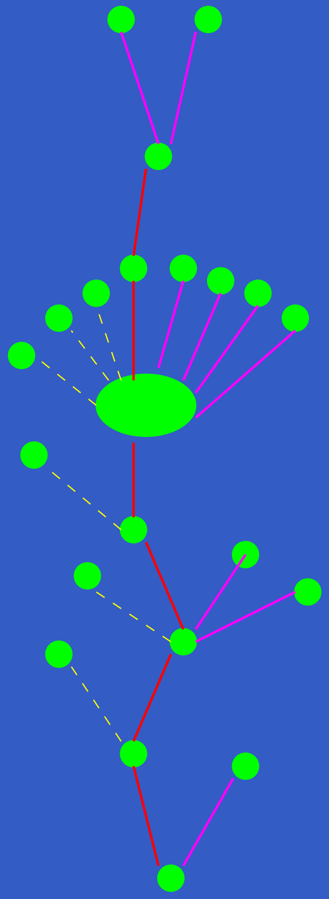
Les fils de i sont les j tels que $V_{\pi^{-1}(j)}$ tombe dans un intervalle $]e(i), e(i) + p_i]$ de longueur p_i. Après cet intervalle, on explore le premier de ces fils, ou le premier noeud non exploré parmi ceux découverts auparavant. L'arbre $\mathcal{T}^{\mathbf{p}}$ ainsi construit est un $\mathbf{p}$-arbre.

Idée de la démonstration

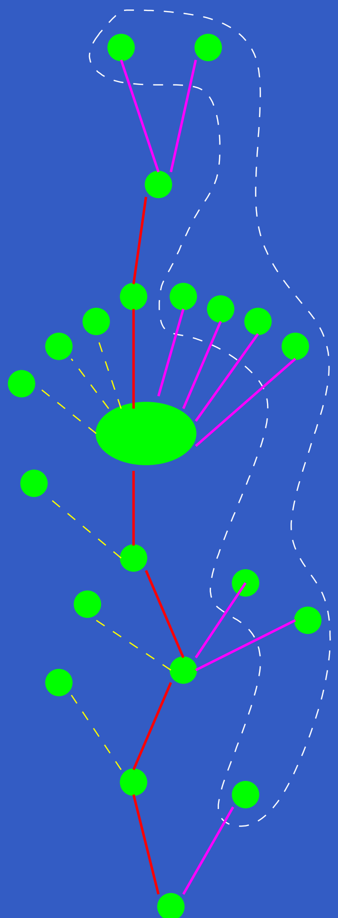
On vérifie alors que $F^{\mathrm{exc},\mathbf{p}}(e(i)) = \sum_{j \in \mathcal{N}(i)} p_j$, où

$$\mathcal{N}(i) = \{\text{fils de } i \text{ et fils non encore explorés des ancêtres de } i\}.$$

Un dessin



Un dessin



Idée de la démonstration

- On choisit une suite approximante $\mathbf{p}$ telle que $p_i \sim 1/n, i > I$ et $p_i \sim \theta_i/\sqrt{n}, 1 \leq i \leq I$.

Idée de la démonstration

- On choisit une suite approximante $\mathbf{p}$ telle que $p_i \sim 1/n, i > I$ et $p_i \sim \theta_i/\sqrt{n}, 1 \leq i \leq I$.
- Une loi faible des grands nombres permet de montrer que la $\mathbf{p}$-masse de $\mathcal{N}(i)$ est proche de la moitié de celle de $\mathcal{A}(i)$, ensemble des ancêtres de i, pour un i typique.

Idée de la démonstration

- On choisit une suite approximante $\mathbf{p}$ telle que $p_i \sim 1/n, i > I$ et $p_i \sim \theta_i/\sqrt{n}, 1 \leq i \leq I$.
- Une loi faible des grands nombres permet de montrer que la $\mathbf{p}$-masse de $\mathcal{N}(i)$ est proche de la moitié de celle de $\mathcal{A}(i)$, ensemble des ancêtres de i, pour un i typique.
- On obtient donc $\mathrm{ht}(i)/n$, plus des termes d'ordre $1/\sqrt{n}$ correspondant aux “gros” nœuds, ayant beaucoup de fils dans $\mathcal{T}^{\mathbf{p}}$. Négliger leurs poids revient à l'opération de gommage des sauts.

Mais encore

- Les détails de la méthode utilisée permet d'obtenir des théorèmes de convergence du processus de hauteur des $\mathbf{p}$-arbres vers $\mathcal{E}^\theta$.
- On peut traiter un cas un peu plus général que $\boldsymbol{\theta}$ presque nulle.
- Liens avec les arbres Lévy : les ICRT sont les "briques élémentaires" des arbres Lévy, au même sens qu'un pont de processus de Lévy est un mélange de ponts de type $X^{\mathrm{br},\theta}$.

Fin

(...)