

# Fonctionnelles additives sur les grands arbres de Galton-Watson

Michel Nassif  
avec R. Abraham and J.-F. Delmas

CERMICS (ENPC)

12 novembre 2020

# Table des matières

## 1 Introduction

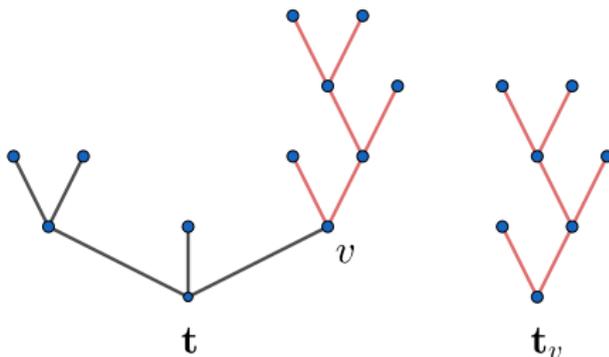
## 2 Limites d'échelle d'arbres de Galton-Watson

## 3 Asymptotiques des fonctionnelles additives

# Notations pour les arbres finis

Soit  $\mathbf{t}$  un arbre fini enraciné.

- Pour tout noeud  $v \in \mathbf{t}$ ,  $\mathbf{t}_v$  est le sous-arbre de  $\mathbf{t}$  au-dessus de  $v$ .



- $|\mathbf{t}|$  est sa taille (= nombre de noeuds),  $h(\mathbf{t})$  sa hauteur. Ici  $|\mathbf{t}| = 12$  et  $h(\mathbf{t}) = 4$ .
- $\mathbf{t}^\circ$  est l'ensemble des noeuds internes de  $\mathbf{t}$ .

# Arbre de Galton-Watson

$\xi$  une mesure de probabilité sur  $\mathbb{N}$ , la **loi de reproduction**.  
On construit  $\tau$  un **arbre de GW**( $\xi$ ) :

# Arbre de Galton-Watson

$\xi$  une mesure de probabilité sur  $\mathbb{N}$ , la **loi de reproduction**.

On construit  $\tau$  un **arbre de GW**( $\xi$ ) :

- on commence avec un individu (la racine),

# Arbre de Galton-Watson

$\xi$  une mesure de probabilité sur  $\mathbb{N}$ , la **loi de reproduction**.

On construit  $\tau$  un **arbre de GW**( $\xi$ ) :

- on commence avec un individu (la racine),
- chaque individu a un nombre aléatoire d'enfants de loi  $\xi$ , indépendamment des autres.

# Arbre de Galton-Watson

$\xi$  une mesure de probabilité sur  $\mathbb{N}$ , la **loi de reproduction**.

On construit  $\tau$  un **arbre de GW**( $\xi$ ) :

- on commence avec un individu (la racine),
- chaque individu a un nombre aléatoire d'enfants de loi  $\xi$ , indépendamment des autres.

**Lien avec le processus** : le nombre d'individus  $Z_n$  à la génération  $n$  est un processus de GW( $\xi$ ).

## Proposition

*Si  $\xi$  est (sous-)critique  $\sum_{n=1}^{\infty} n\xi(n) \leq 1$ , alors  $\tau$  est p.s. fini.*

# Arbre de Galton-Watson conditionné

Dans la suite,  $\xi$  est une **loi de reproduction critique**. On note  $\tau^n$  un **GW( $\xi$ ) conditionné à avoir  $n$  noeuds**.

# Arbre de Galton-Watson conditionné

Dans la suite,  $\xi$  est une **loi de reproduction critique**. On note  $\tau^n$  un **GW( $\xi$ ) conditionné à avoir  $n$  noeuds**.

## Exemple

- Si  $\xi$  est la **loi géométrique** de paramètre 1,  $\tau^n$  suit la **loi uniforme** sur l'ensemble des **arbres ordonnés à  $n$  noeuds**.

# Arbre de Galton-Watson conditionné

Dans la suite,  $\xi$  est une **loi de reproduction critique**. On note  $\tau^n$  un **GW( $\xi$ ) conditionné à avoir  $n$  noeuds**.

## Exemple

- Si  $\xi$  est la **loi géométrique** de paramètre 1,  $\tau^n$  suit la **loi uniforme** sur l'ensemble des **arbres ordonnés à  $n$  noeuds**.
- Si  $\xi(0) = \xi(2) = 1/2$ ,  $\tau^n$  suit la **loi uniforme** sur l'ensemble des **arbres binaires complets à  $n$  noeuds**. C'est le modèle de Catalan.

# Fonctionnelles additives

Une fonctionnelle  $F$  sur l'ensemble des arbres finis est dite **additive** si elle satisfait la relation

$$F(\mathbf{t}) = \sum_{i=1}^d F(\mathbf{t}_i) + f(\mathbf{t}),$$

où les  $\mathbf{t}_1, \dots, \mathbf{t}_d$  sont les sous-arbres de  $\mathbf{t}$  au-dessus de la racine et  $f$  est une fonction donnée.

# Fonctionnelles additives

Une fonctionnelle  $F$  sur l'ensemble des arbres finis est dite **additive** si elle satisfait la relation

$$F(\mathbf{t}) = \sum_{i=1}^d F(\mathbf{t}_i) + f(\mathbf{t}),$$

où les  $\mathbf{t}_1, \dots, \mathbf{t}_d$  sont les sous-arbres de  $\mathbf{t}$  au-dessus de la racine et  $f$  est une fonction donnée.

**Remarque.**

$$F(\mathbf{t}) = \sum_{w \in \mathbf{t}} f(\mathbf{t}_w),$$

où  $\mathbf{t}_w$  est le sous-arbre de  $\mathbf{t}$  au-dessus de  $w$ .

# Exemples de fonctionnelles additives

- La longueur de cheminement total

$$P(\mathbf{t}) = \sum_{v \in \mathbf{t}} d(\emptyset, v) = \sum_{w \in \mathbf{t}} |\mathbf{t}_w| - |\mathbf{t}|.$$

- L'indice de Wiener

$$W(\mathbf{t}) = \sum_{u, v \in \mathbf{t}} d(u, v) = 2|\mathbf{t}| \sum_{w \in \mathbf{t}} |\mathbf{t}_w| - 2 \sum_{w \in \mathbf{t}} |\mathbf{t}_w|^2.$$

- L'indice  $B_1$  de Shao et Sokal (1990)

$$B_1(\mathbf{t}) = \sum_{\substack{w \in \mathbf{t}^\circ \\ w \neq \emptyset}} \frac{1}{h(\mathbf{t}_w)},$$

où  $\mathbf{t}^\circ$  est l'ensemble des noeuds internes  $\mathbf{t}$ .

# Objectif

L'objectif est d'étudier le comportement asymptotique de

$$\sum_{w \in \tau^n} f(|\tau_w^n|, \mathfrak{h}(\tau_w^n)),$$

et plus particulièrement de

$$\sum_{w \in \tau^n} |\tau_w^n|^\alpha \mathfrak{h}(\tau_w^n)^\beta,$$

quand la loi de reproduction  $\xi$

- a une variance finie  $\sigma^2$ ,
- plus généralement, quand elle appartient au domaine d'attraction d'une loi stable d'indice  $\gamma \in (1, 2]$ .

# Résultats connus quand $\sigma^2 < \infty$

Sur les fonctionnelles ne dépendant que de la taille.

# Résultats connus quand $\sigma^2 < \infty$

Sur les fonctionnelles ne dépendant que de la taille.

- **Fill et Kapur** (2003) ont montré que dans le modèle de Catalan (i.e.  $\xi(0) = \xi(2) = 1/2$ ), pour  $\alpha > 0$ ,

$$n^{-(\alpha+1/2)} \sum_{w \in \mathcal{T}^n} |\tau_w^n|^\alpha \xrightarrow[n \rightarrow \infty]{(d)} Z_\alpha,$$

où  $Z_\alpha$  est caractérisé par ses moments.

# Résultats connus quand $\sigma^2 < \infty$

Sur les fonctionnelles ne dépendant que de la taille.

- **Fill et Kapur** (2003) ont montré que dans le modèle de Catalan (i.e.  $\xi(0) = \xi(2) = 1/2$ ), pour  $\alpha > 0$ ,

$$n^{-(\alpha+1/2)} \sum_{w \in \mathcal{T}^n} |\tau_w^n|^\alpha \xrightarrow[n \rightarrow \infty]{(d)} Z_\alpha,$$

où  $Z_\alpha$  est caractérisé par ses moments.

- **Fill et Janson** (2007) obtiennent une représentation de  $Z_\alpha$  en fonction de l'excursion Brownienne pour  $\alpha > 1/2$ , et plus tard généralisent le résultat pour des lois de reproduction quelconque à variance finie.

# Résultats connus quand $\sigma^2 < \infty$

Sur les fonctionnelles ne dépendant que de la taille.

- **Fill et Kapur** (2003) ont montré que dans le modèle de Catalan (i.e.  $\xi(0) = \xi(2) = 1/2$ ), pour  $\alpha > 0$ ,

$$n^{-(\alpha+1/2)} \sum_{w \in \mathcal{T}^n} |\tau_w^n|^\alpha \xrightarrow[n \rightarrow \infty]{(d)} Z_\alpha,$$

où  $Z_\alpha$  est caractérisé par ses moments.

- **Fill et Janson** (2007) obtiennent une représentation de  $Z_\alpha$  en fonction de l'excursion Brownienne pour  $\alpha > 1/2$ , et plus tard généralisent le résultat pour des lois de reproduction quelconque à variance finie.
- Pour  $\alpha < 0$ , on a un TCL avec la vitesse standard  $\sqrt{n}$ .

**Remarque.** 2 transitions de phase

# Table des matières

1 Introduction

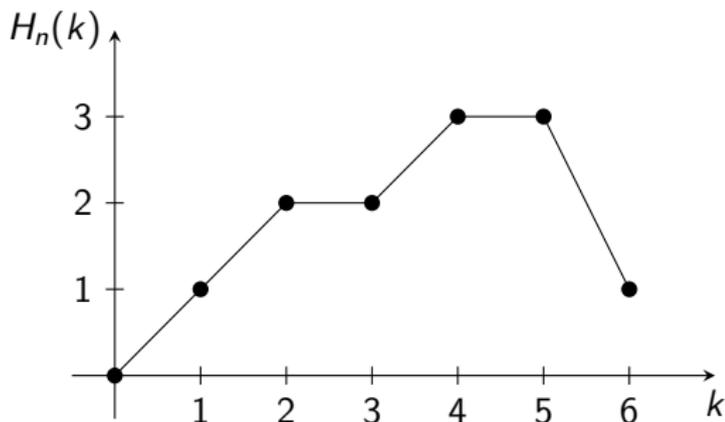
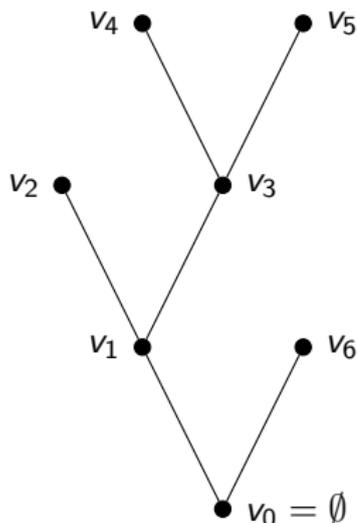
2 Limites d'échelle d'arbres de Galton-Watson

3 Asymptotiques des fonctionnelles additives

# Codage d'un arbre fini

Soit  $\tau^n$  un  $\text{GW}(\xi)$ . On liste les noeuds  $v_0 = \emptyset, v_1, \dots, v_{n-1}$  dans l'ordre correspondant au **Depth-First Search**. On définit

$$H_n(k) = d(\emptyset, v_k), \quad 0 \leq k \leq n-1.$$



# Convergence vers l'excursion brownienne

Soit  $\tau^n$  un  $\text{GW}(\xi)$  de loi de reproduction  $\xi$  critique de variance finie  $\sigma^2$ .

## Théorème

On a la convergence en loi dans l'espace  $C([0, 1])$

$$\left( \frac{1}{\sqrt{n}} H_n(nt), 0 \leq t \leq 1 \right) \xrightarrow[n \rightarrow \infty]{(d)} \left( \frac{2}{\sigma} B^{\text{ex}}(t), 0 \leq t \leq 1 \right)$$

où  $B^{\text{ex}}$  est l'excursion brownienne normalisée.

Rappel :  $B^{\text{ex}}$  est un mouvement brownien partant de 0, conditionné à prendre la valeur 0 au temps 1 et à rester positif sur  $(0, 1)$ .

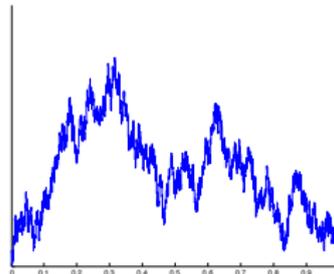


FIGURE – Simulation due à Igor Kortchemski

# Arbre réel mesuré

Un  $\mathbb{R}$ -**arbre** est un espace métrique complet  $(T, d)$  tel que pour tout  $x, y \in T$ ,

- il existe une unique isométrie  $\varphi: [0, d(x, y)] \rightarrow T$  tel que  $\varphi(0) = x$  et  $\varphi(d(x, y)) = y$ ,
- pour tout  $s \in (0, d(x, y))$ ,  $x$  et  $y$  sont dans des composantes connexes différentes de  $T \setminus \{\varphi(s)\}$ .

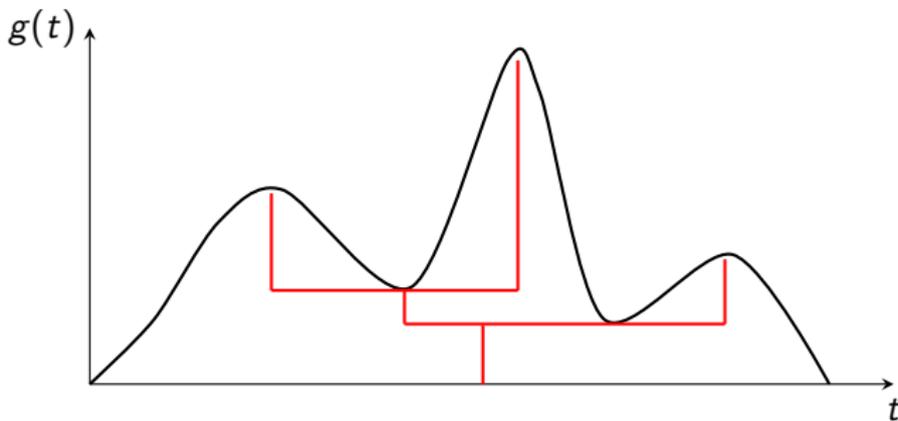
On considère des  $\mathbb{R}$ -arbres  $(T, d)$  avec une racine  $\emptyset \in T$  et une mesure finie  $\mu$ .

# Codage d'un arbre réel

Soit  $g: [0, 1] \rightarrow [0, \infty)$  une fonction continue telle que  $g(0) = g(1) = 0$ .

$$m_g(s, t) = \min_{[s, t]} g, \quad d_g(s, t) = g(s) + g(t) - 2m_g(s, t).$$

On définit une relation d'équivalence  $s \sim_g t$  ssi  $d_g(s, t) = 0$  ssi  $g(s) = g(t) = m_g(s, t)$ .



## Codage d'un arbre réel (2)

### Proposition

$T_g = [0, 1] / \sim_g$  muni de  $d_g$  est un  $\mathbb{R}$ -arbre **compact**.

Si  $p: [0, 1] \rightarrow T_g$  est la projection canonique,  $T_g$  est enraciné en  $p(0)$  et muni d'une mesure de probabilité  $\mu = \text{Leb} \circ p^{-1}$ .

# Distance de Gromov-Hausdorff-Prokhorov

Soient  $(T, d, \mu, \emptyset), (T', d', \mu', \emptyset')$  deux  $\mathbb{R}$ -arbres compacts mesurés. On définit

$$d_{\text{GHP}}(T, T') := \inf \{ \delta(\varphi(\emptyset), \varphi'(\emptyset')) \vee \delta_{\text{H}}(\varphi(T), \varphi'(T')) \vee \delta_{\text{P}}(\mu \circ \varphi^{-1}, \mu' \circ \varphi'^{-1}) \},$$

où l'infimum est pris sur toutes les isométries  $\varphi: T \rightarrow E$  et  $\varphi': T' \rightarrow E$  dans un espace métrique commun  $(E, \delta)$ .

# Distance de Gromov-Hausdorff-Prokhorov

Soient  $(T, d, \mu, \emptyset)$ ,  $(T', d', \mu', \emptyset')$  deux  $\mathbb{R}$ -arbres compacts mesurés. On définit

$$d_{\text{GHP}}(T, T') := \inf \left\{ \delta(\varphi(\emptyset), \varphi'(\emptyset')) \vee \delta_{\text{H}}(\varphi(T), \varphi'(T')) \vee \delta_{\text{P}}(\mu \circ \varphi^{-1}, \mu' \circ \varphi'^{-1}) \right\},$$

où l'infimum est pris sur toutes les isométries  $\varphi: T \rightarrow E$  et  $\varphi': T' \rightarrow E$  dans un espace métrique commun  $(E, \delta)$ .

## Proposition

*L'ensemble des  $\mathbb{R}$ -arbres compacts mesurés muni de la distance  $d_{\text{GHP}}$  est un espace métrique complet.*

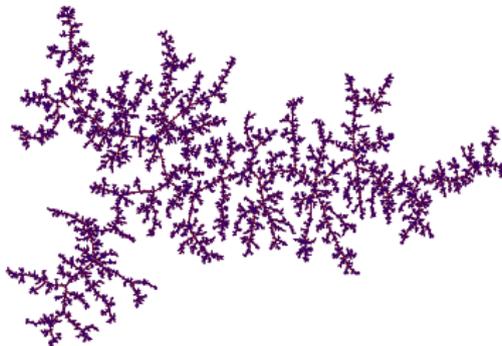
# Convergence d'arbres

## Théorème

*On a la convergence en loi*

$$\frac{1}{\sqrt{n}} \tau^n \xrightarrow[n \rightarrow \infty]{(d)} \frac{2}{\sigma} T_{B^{\text{ex}}},$$

*où  $T_{B^{\text{ex}}}$  est l'arbre codé par  $B^{\text{ex}}$ .*



# Table des matières

1 Introduction

2 Limites d'échelle d'arbres de Galton-Watson

3 Asymptotiques des fonctionnelles additives

# Convergence des polynômes en la taille et la hauteur

On suppose que  $\xi$  est critique et de variance finie  $\sigma^2$ .

## Théorème

- Si  $2\alpha + \beta > 1$ , on a la convergence en loi et du premier moment

$$\frac{1}{n^{\alpha+(\beta+1)/2}} \sum_{w \in \mathcal{T}^{n,0}} |\tau_w^n|^\alpha \mathfrak{h}(\tau_w^n)^\beta \xrightarrow[n \rightarrow \infty]{(d)+\text{mean}} \left(\frac{2}{\sigma}\right)^{\beta+1} Z_{\alpha,\beta}.$$

- Si  $2\alpha + \beta \leq 1$ , on a

$$\frac{1}{n^{\alpha+(\beta+1)/2}} \sum_{w \in \mathcal{T}^{n,0}} |\tau_w^n|^\alpha \mathfrak{h}(\tau_w^n)^\beta \xrightarrow[n \rightarrow \infty]{(d)+\text{mean}} \left(\frac{2}{\sigma}\right)^{\beta+1} Z_{\alpha,\beta} = \infty.$$

## Quelques commentaires

- Transition de phase en  $2\alpha + \beta = 1$  entre régime global et régime intermédiaire.

## Quelques commentaires

- Transition de phase en  $2\alpha + \beta = 1$  entre régime global et régime intermédiaire.
- Compensation entre masse et hauteur correspond au scaling de l'excursion : la hauteur (= le maximum) d'une excursion brownienne de longueur  $\zeta$  se comporte comme  $\sqrt{\zeta}$  :

$$\mathbf{N}[\zeta \in dx] \sim \frac{dx}{x^{3/2}}, \quad \mathbf{N}[h \in dx] \sim \frac{dx}{x^2}$$

où  $\mathbf{N}$  est la mesure d'excursion du mouvement brownien. (Aussi la hauteur de  $\tau^n$  est en  $\sqrt{n}$ ).

## Quelques commentaires

- Transition de phase en  $2\alpha + \beta = 1$  entre régime global et régime intermédiaire.
- Compensation entre masse et hauteur correspond au scaling de l'excursion : la hauteur (= le maximum) d'une excursion brownienne de longueur  $\zeta$  se comporte comme  $\sqrt{\zeta}$  :

$$\mathbf{N}[\zeta \in dx] \sim \frac{dx}{x^{3/2}}, \quad \mathbf{N}[h \in dx] \sim \frac{dx}{x^2}$$

où  $\mathbf{N}$  est la mesure d'excursion du mouvement brownien. (Aussi la hauteur de  $\tau^n$  est en  $\sqrt{n}$ ).

- On peut ajouter les feuilles quand cela a un sens ( $\beta \geq 0$ ).

# Représentation de la limite

$Z_{\alpha,\beta}$  admet la représentation suivante en termes de l'arbre brownien  
 $T = T_{B^{\text{ex}}}$  :

$$Z_{\alpha,\beta} = \int_T \mathfrak{m}(\mathcal{T}_y)^\alpha \mathfrak{h}(\mathcal{T}_y)^\beta \ell(dy),$$

où  $\ell$  est "la mesure de Lebesgue" sur  $T$ . Ici  $\mathcal{T}_y$  est le sous-arbre de  $T$  au-dessus de  $y$ ,  $\mathfrak{m}(\mathcal{T}_y)$  est sa masse et  $\mathfrak{h}(\mathcal{T}_y)$  sa hauteur.

# Idée de preuve

- Le bon objet à regarder est

$$\frac{1}{n^{3/2}} \sum_{w \in \mathcal{T}^{n,0}} |\tau_w^n| f \left( \frac{|\tau_w^n|}{n}, \frac{\mathfrak{h}(\tau_w^n)}{\sqrt{n}} \right).$$

Noter que  $|\tau^n| = n$  et  $\mathfrak{h}(\tau^n) \sim \sqrt{n}$ .

- On sort  $|\tau_w^n|$  pour avoir une mesure finie : à la limite,  $\ell(dy)$  est infinie mais  $\mathfrak{m}(T_y) \ell(dy)$  est finie.

# Idée de preuve (2)

## Étape 1. Approcher

$$\frac{1}{n^{3/2}} \sum_{w \in \mathcal{T}^{n,0}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{\mathfrak{h}(\tau_w^n)}{\sqrt{n}}\right) \approx \int_{T^n} \mathfrak{m}(T_y^n) f(\mathfrak{m}(T_y^n), \mathfrak{h}(T_y^n)) \ell_{T^n}(dy)$$

où  $T^n = \frac{1}{\sqrt{n}} \mathcal{T}^n$ .

# Idée de preuve (2)

**Étape 1.** Approcher

$$\frac{1}{n^{3/2}} \sum_{w \in \mathcal{T}^{n,0}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{\mathfrak{h}(\tau_w^n)}{\sqrt{n}}\right) \approx \int_{T^n} \mathfrak{m}(T_y^n) f(\mathfrak{m}(T_y^n), \mathfrak{h}(T_y^n)) \ell_{T^n}(dy)$$

où  $T^n = \frac{1}{\sqrt{n}} \mathcal{T}^n$ .

**Étape 2.** Montrer que

$$T \mapsto \int_T \mathfrak{m}(T_y) f(\mathfrak{m}(T_y), \mathfrak{h}(T_y)) \ell_T(dy)$$

est continue. On en déduit la convergence en loi quand  $f$  est continue bornée.

## Idée de preuve (3)

Et les fonctions qui explosent (dans notre cas les puissances négatives) ?

## Idée de preuve (3)

Et les fonctions qui explosent (dans notre cas les puissances négatives) ?  
Utiliser un résultat de type : si des mesures aléatoires  $\mu_n$  sur  $[0, 1]$  convergent en loi vers une mesure aléatoire  $\mu$  et s'il existe  $\varepsilon_0 > 0$  tel que

$$\sup_n \mathbb{E} [\mu_n(x^{-\varepsilon_0})] < \infty,$$

alors  $\mu_n(x^{-\varepsilon}) \xrightarrow{(d)} \mu(x^{-\varepsilon})$  pour tout  $\varepsilon < \varepsilon_0$ .

**Étape 3.** Donc il suffit de montrer que pour tout  $2\alpha + \beta > 1$ ,

$$\frac{1}{n^{3/2}} \sup_n \mathbb{E} \left[ \sum_{w \in \mathcal{T}^{n,0}} |\tau_w^n| \left( \frac{|\tau_w^n|}{n} \right)^{\alpha-1} \left( \frac{\mathfrak{h}(\tau_w^n)}{\sqrt{n}} \right)^\beta \right] < \infty.$$

# Espérance de $Z_{\alpha,\beta}$

## Proposition

Pour tous  $\alpha, \beta \in \mathbb{R}$  tels que  $2\alpha + \beta > 1$ , on a

$$\mathbb{E}[Z_{\alpha,\beta}] = \pi^{(\beta-1)/2} \xi(\beta) \mathbf{B}(\alpha + \beta/2 - 1/2, 1/2),$$

où  $\xi$  est la fonction xi de Riemann défini par

$$\xi(s) = \frac{1}{2} s(s-1) \pi^{-s/2} \Gamma(s/2) \zeta(s), \quad \forall s \in \mathbb{C}.$$

# Récap

On suppose que  $\xi$  est critique de variance finie. Comportement asymptotique de

$$\sum_{w \in \mathcal{T}^{n,0}} |\tau_w^n|^{\alpha} \mathfrak{h}(\tau_w^n)^{\beta}.$$

- $2\alpha + \beta > 1$  (**global**) : convergence en loi vers  $Z_{\alpha,\beta}$  (fonctionnelle de l'excursion brownienne) après renormalisation par  $n^{\alpha+(1+\beta)/2}$ .
- $2\alpha + \beta < 0$  (**local**) : TCL après recentrage et renormalisation par  $\sqrt{n}$ .
- $0 < 2\alpha + \beta < 1$  (**intermédiaire**) : connu seulement pour  $\beta = 0$ , pas de représentation de la limite.

**Remarque.** L'indice  $B_1$  de Shao et Sokal ( $\alpha = 0, \beta = -1$ ) se trouve dans la région  $2\alpha + \beta < 0$ .

# Convergence quand $\alpha \rightarrow \infty$

## Théorème

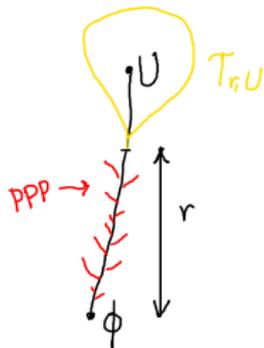
$$\sqrt{\alpha} Z_\alpha \xrightarrow[n \rightarrow \infty]{(d)} \int_0^\infty e^{-S_t} dt$$

où  $S$  est un subordonateur stable d'indice  $1/2$ .

## Idée de preuve

Soit  $U$  une feuille choisie au hasard selon  $\mu$ .

$$\begin{aligned}
 \sqrt{\alpha} \int_T \mu(dx) \int_0^{B^{\text{ex}}(x)} m(T_{r,x})^\alpha dr &= \sqrt{\alpha} \int_0^{B^{\text{ex}}(U)} m(T_{r,U})^\alpha dr \\
 &\approx \sqrt{\alpha} \int_0^\varepsilon m(T_{r,U})^\alpha dr &= \sqrt{\alpha} \int_0^\varepsilon (1 - S_r)^\alpha dr \\
 &= \int_0^{\sqrt{\alpha}\varepsilon} (1 - S_{t/\sqrt{\alpha}})^\alpha dt &\stackrel{(d)}{=} \int_0^{\sqrt{\alpha}\varepsilon} \left(1 - \frac{S_t}{\alpha}\right)^\alpha dt \\
 &\approx \int_0^{\sqrt{\alpha}\varepsilon} e^{-S_t} dt &\approx \int_0^\infty e^{S_t} dt.
 \end{aligned}$$



# Références

## Convergence des arbres de GW, arbres brownien et stable.

- D. Aldous. *The continuum random tree I.* (1991)
- T. Duquesne. *A limit theorem for the contour process of conditioned Galton-Watson trees.* (2003)
- T. Duquesne and J.-F. Le Gall. *Random trees, Lévy processes and spatial branching processes.* (2002)

## Fonctionnelles additives sur les arbres de GW.

- J. A. Fill and N. Kapur. *Limiting distributions for additive functionals on Catalan trees.* (2004)
- J.-F. Delmas, J.-S. Dhersin and M. Sciaudeau. *Cost functionals for large (uniform and simply generated) random trees.* (2018)
- R. Abraham, J.-F. Delmas and M.N. *Global regime for general additive functionals of conditioned Bienaymé-Galton-Watson trees.* (2020)

Merci..

pour votre attention !