Marche aléatoire dans $\mathbb{Z}^d$ indexée par un arbre & son nombre de points visités

Shen LIN
en collaboration avec Jean-François LE GALL

Université Paris-Sud

Colloque Jeunes Probabilistes et Statisticiens
8 avril 2014

Introduction

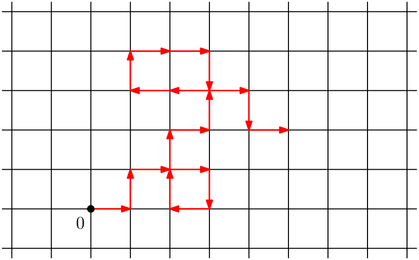
Soit $(X_n)_{n\geq 0}$ une marche aléatoire dans $\mathbb{Z}^d$, $X_0 = 0$,

$$X_n = Y_1 + Y_2 + \cdots + Y_n$$

où $Y_1, Y_2, \ldots$ sont indépendantes à valeurs dans $\mathbb{Z}^d$, et de même loi θ. Alors

$$\mathrm{R}_n := \#\{X_0, X_1, \ldots, X_n\}$$

est le nombre de points distincts dans $\mathbb{Z}^d$ visités par la marche jusqu'au temps n.



Ici, $\mathrm{R}_{19} = 18$

Théorème de Dvoretzky-Erdös

Dvoretzky et Erdös (1951) ont étudié le comportement asymptotique de R_n quand $n \to \infty$, dans le cas de la marche aléatoire simple (i.e. si θ est uniforme sur les voisins de 0) :

- si $d \geq 3$,

$$\frac{1}{n}\mathrm{R}_n \xrightarrow[n\to\infty]{\text{p.s.}} q_d > 0\,,$$

- si $d = 2$,

$$\frac{\log n}{n}\mathrm{R}_n \xrightarrow[n\to\infty]{\text{p.s.}} \pi\,,$$

- si $d = 1$,

$$\frac{\mathrm{R}_n}{\sqrt{n}} \xrightarrow[n\to\infty]{\text{(d)}} \sup_{0\leq t\leq 1} B_t - \inf_{0\leq t\leq 1} B_t\,,$$

où q_d est la probabilité que la marche ne revienne plus à son point de départ, et $(B_t)_{t\geq 0}$ est un mouvement brownien linéaire standard.

Appliquer le théorème ergodique sous-additif

Quand $d \geq 3$, il y a une preuve courte (Kesten-Spitzer-Whitman) qui utilise le théorème ergodique sous-additif de Kingman. Notons que, pour tous $m, n \geq 0$,

$$\mathrm{R}_{n+m} \leq \mathrm{R}_n + \mathrm{R}_m \circ \tau_n,$$

où τ_n est l'opérateur de shift sur les trajectoires : $\tau_n \circ X_k = X_{n+k} - X_n$.

Appliquer le théorème ergodique sous-additif

Quand $d \geq 3$, il y a une preuve courte (Kesten-Spitzer-Whitman) qui utilise le théorème ergodique sous-additif de Kingman. Notons que, pour tous $m, n \geq 0$,

$$\mathrm{R}_{n+m} \leq \mathrm{R}_n + \mathrm{R}_m \circ \tau_n,$$

où τ_n est l'opérateur de shift sur les trajectoires : $\tau_n \circ X_k = X_{n+k} - X_n$.

Le théorème ergodique sous-additif de Kingman donne immédiatement

$$\frac{1}{n}\mathrm{R}_n \xrightarrow[n\to\infty]{\text{p.s.}} q_d \, ,$$

où $q_d = \lim_{n\to\infty} \frac{1}{n} E[\mathrm{R}_n] = P(X_1 \neq 0, X_2 \neq 0, \ldots)$.

Appliquer le théorème ergodique sous-additif

Quand $d \geq 3$, il y a une preuve courte (Kesten-Spitzer-Whitman) qui utilise le théorème ergodique sous-additif de Kingman. Notons que, pour tous $m, n \geq 0$,

$$\mathrm{R}_{n+m} \leq \mathrm{R}_n + \mathrm{R}_m \circ \tau_n,$$

où τ_n est l'opérateur de shift sur les trajectoires : $\tau_n \circ X_k = X_{n+k} - X_n$.

Le théorème ergodique sous-additif de Kingman donne immédiatement

$$\frac{1}{n}\mathrm{R}_n \xrightarrow[n\to\infty]{\text{p.s.}} q_d \,,$$

où $q_d = \lim_{n\to\infty} \frac{1}{n} E[\mathrm{R}_n] = P(X_1 \neq 0, X_2 \neq 0, \ldots)$.

Cette méthode s'applique à toutes les marches aléatoires, et

$$\boxed{q_d > 0 \text{ si et seulement si } X \text{ est transitoire.}}$$

Marche aléatoire indexée par un arbre

Question (Itai Benjamini) : a-t-on un résultat analogue pour une marche aléatoire indexée par un arbre ?

Marche aléatoire indexée par un arbre

Question (Itai Benjamini) : a-t-on un résultat analogue pour une marche aléatoire indexée par un arbre ?

Considérons

- un arbre (aléatoire) $\mathcal{T}_n$ discret enraciné avec n sommets ;
- conditionné à $\mathcal{T}_n$, une collection $(Y_e)_{e\in\mathcal{E}(\mathcal{T}_n)}$ des v.a. indépendantes distribuées selon θ, indexées par l'ensemble des arêtes de $\mathcal{T}_n$.

Marche aléatoire indexée par un arbre

Question (Itai Benjamini) : a-t-on un résultat analogue pour une marche aléatoire indexée par un arbre ?

Considérons

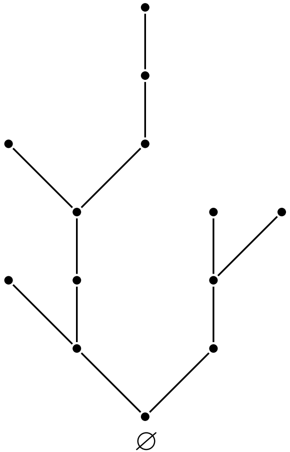
- un arbre (aléatoire) $\mathcal{T}_n$ discret enraciné avec n sommets ;
- conditionné à $\mathcal{T}_n$, une collection $(Y_e)_{e \in \mathcal{E}(\mathcal{T}_n)}$ des v.a. indépendantes distribuées selon θ, indexées par l'ensemble des arêtes de $\mathcal{T}_n$.

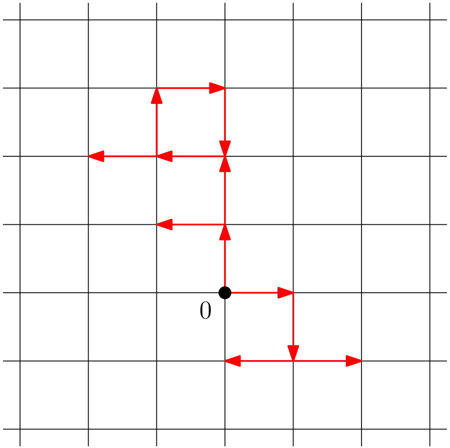
Pour chaque sommet v de $\mathcal{T}_n$, soit

$$X_v = \sum_{e:\ \varnothing \to v} Y_e$$

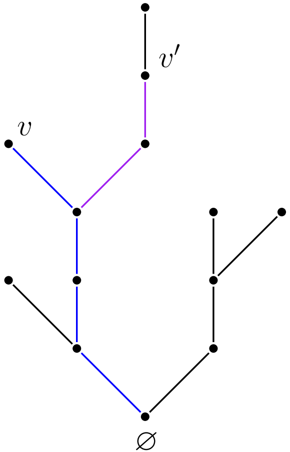
où l'on somme sur toutes les arêtes le long du chemin simple de la racine $\varnothing$ à v.

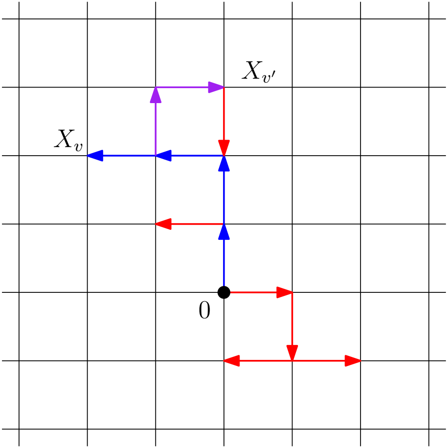
Marche aléatoire indexée par un arbre





Marche aléatoire indexée par un arbre





Arbres de Galton-Watson

Soit μ une mesure de probabilité sur $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ telle que :

- $\displaystyle\sum_{k=0}^{\infty} k\,\mu(k) = 1$ (μ est critique),
- $\displaystyle\sum_{k=0}^{\infty} k^2\mu(k) < \infty$ (μ de variance finie).

Arbres de Galton-Watson

Soit μ une mesure de probabilité sur $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ telle que :

- $\sum_{k=0}^{\infty} k\,\mu(k) = 1 \quad$ (μ est critique),
- $\sum_{k=0}^{\infty} k^2\mu(k) < \infty \quad$ (μ de variance finie).

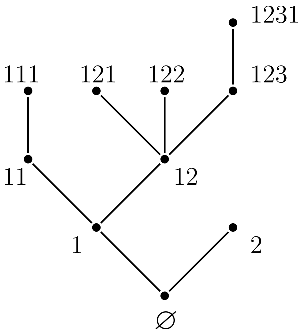
L'arbre de Galton-Watson $\mathcal{T}$ de loi de reproduction μ décrit la généalogie d'un processus de Galton-Watson de même loi de reproduction :

- le processus part d'un ancêtre à la génération 0 ;
- chaque individu a, indépendamment des autres, k enfants avec probabilité $\mu(k)$.

Comme μ est critique, l'arbre $\mathcal{T}$ est p.s. fini.

Notation : $\#\mathcal{T}$ = le nombre de sommets de $\mathcal{T}$.

Arbres de Galton-Watson



Un arbre planaire $\mathcal{T}$ (enraciné et ordonné)

Arbres de GW conditionnés par la taille

Soit $\mathcal{T}$ un arbre de Galton-Watson de loi de reproduction μ.
Pour chaque $n \geq 1$ tel que $P(\#\mathcal{T} = n) > 0$, soit

$$\mathcal{T}_n \overset{\text{(d)}}{=} \mathcal{T} \text{ conditionné à } \#\mathcal{T} = n.$$

Arbres de GW conditionnés par la taille

Soit $\mathcal{T}$ un arbre de Galton-Watson de loi de reproduction μ.
Pour chaque $n \geq 1$ tel que $P(\#\mathcal{T} = n) > 0$, soit

$$\mathcal{T}_n \overset{\text{(d)}}{=} \mathcal{T} \text{ conditionné à } \#\mathcal{T} = n.$$

Plusieurs classes combinatoires d'arbres aléatoires à n sommets retrouvées :

- Si $\mu(k) = 2^{-k-1}$, $\mathcal{T}_n$ est uniformément distribué sur l'ensemble des arbres planaires à n sommets ;
- Si $\mu(0) = \mu(2) = 1/2$, $\mathcal{T}_n$ est uniformément distribué sur l'ensemble des arbres binaires à n sommets (n impair) ;
- Si $\mu = \text{Poisson}(1)$, $\mathcal{T}_n$ est uniformément distribué sur l'ensemble des arbres de Cayley à n sommets.

Arbres de GW conditionnés par la taille

Soit $\mathcal{T}$ un arbre de Galton-Watson de loi de reproduction μ.
Pour chaque $n \geq 1$ tel que $P(\#\mathcal{T} = n) > 0$, soit

$$\mathcal{T}_n \overset{(\mathrm{d})}{=} \mathcal{T} \text{ conditionné à } \#\mathcal{T} = n.$$

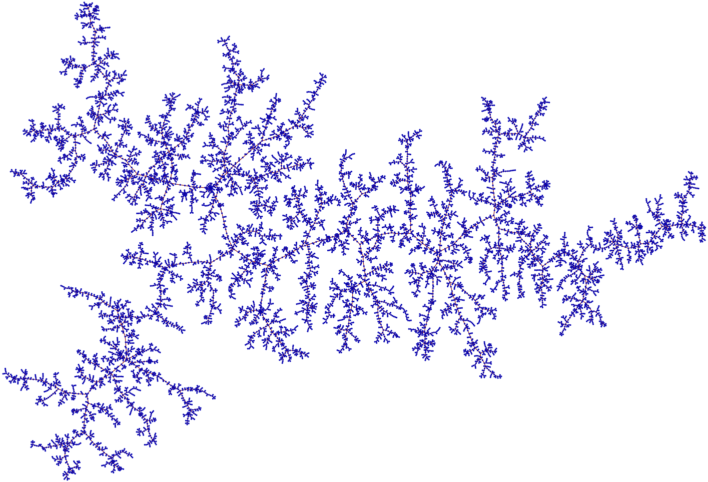
Plusieurs classes combinatoires d'arbres aléatoires à n sommets retrouvées :

- Si $\mu(k) = 2^{-k-1}$, $\mathcal{T}_n$ est uniformément distribué sur l'ensemble des arbres planaires à n sommets ;
- Si $\mu(0) = \mu(2) = 1/2$, $\mathcal{T}_n$ est uniformément distribué sur l'ensemble des arbres binaires à n sommets (n impair) ;
- Si $\mu = \text{Poisson}(1)$, $\mathcal{T}_n$ est uniformément distribué sur l'ensemble des arbres de Cayley à n sommets.

Théorème d'Aldous (1993) : Soit $\rho^2 = \operatorname{var} \mu$,

$$\left(\mathcal{T}_n, \frac{\rho}{2\sqrt{n}} d_{\mathrm{gr}}\right) \xrightarrow[n \to \infty]{(\mathrm{d})} (\mathbf{T_e}, d_{\mathbf{e}}).$$

Grand arbre planaire



Un arbre planaire uniforme à 27179 sommets (©Igor Kortchemski)

Nombre de points visités par une marche aléatoire indexée par $\mathcal{T}_n$

- $\mathcal{T}_n$ est un arbre de GW de loi de reproduction μ, conditionné à avoir n sommets ;
- Conditionnellement à $\mathcal{T}_n$, $(X_v)_{v\in\mathcal{V}(\mathcal{T}_n)}$ est une marche aléatoire de loi de saut θ, indexée par $\mathcal{T}_n$, et $\mathcal{R}_n = \#\{X_v : v \in \mathcal{V}(\mathcal{T}_n)\}$.

Nombre de points visités par une marche aléatoire indexée par $\mathcal{T}_n$

- $\mathcal{T}_n$ est un arbre de GW de loi de reproduction μ, conditionné à avoir n sommets ;
- Conditionnellement à $\mathcal{T}_n$, $(X_v)_{v \in \mathcal{V}(\mathcal{T}_n)}$ est une marche aléatoire de loi de saut θ, indexée par $\mathcal{T}_n$, et $\mathcal{R}_n = \#\{X_v : v \in \mathcal{V}(\mathcal{T}_n)\}$.

Théorème

Supposons que θ est centrée, non portée sur un sous-groupe de $\mathbb{Z}^d$ + conditions supplémentaires sur les moments.

Nombre de points visités par une marche aléatoire indexée par $\mathcal{T}_n$

- $\mathcal{T}_n$ est un arbre de GW de loi de reproduction μ, conditionné à avoir n sommets ;
- Conditionnellement à $\mathcal{T}_n$, $(X_v)_{v \in \mathcal{V}(\mathcal{T}_n)}$ est une marche aléatoire de loi de saut θ, indexée par $\mathcal{T}_n$, et $\mathcal{R}_n = \#\{X_v : v \in \mathcal{V}(\mathcal{T}_n)\}$.

Théorème

Supposons que θ est centrée, non portée sur un sous-groupe de $\mathbb{Z}^d$ + conditions supplémentaires sur les moments.

1. *Si $d \geq 5$, $n^{-1}\mathcal{R}_n \xrightarrow[n\to\infty]{(\mathrm{P})} c_{\mu,\theta} > 0$;*

Nombre de points visités par une marche aléatoire indexée par $\mathcal{T}_n$

- $\mathcal{T}_n$ est un arbre de GW de loi de reproduction μ, conditionné à avoir n sommets ;
- Conditionnellement à $\mathcal{T}_n$, $(X_v)_{v\in\mathcal{V}(\mathcal{T}_n)}$ est une marche aléatoire de loi de saut θ, indexée par $\mathcal{T}_n$, et $\mathcal{R}_n = \#\{X_v : v \in \mathcal{V}(\mathcal{T}_n)\}$.

Théorème

Supposons que θ est centrée, non portée sur un sous-groupe de $\mathbb{Z}^d$ + conditions supplémentaires sur les moments.

1. *Si $d \geq 5$, $n^{-1}\mathcal{R}_n \xrightarrow[n\to\infty]{(\mathrm{P})} c_{\mu,\theta} > 0$;*
2. *Si $d = 4$ et $\mu(k) = 2^{-k-1}$, $\frac{\log n}{n}\mathcal{R}_n \xrightarrow[n\to\infty]{\mathrm{L}^2} 8\,\pi^2\,\sigma^4$, où $\sigma^2 = (\det(\operatorname{cov}\theta))^{1/4}$;*

Nombre de points visités par une marche aléatoire indexée par $\mathcal{T}_n$

- $\mathcal{T}_n$ est un arbre de GW de loi de reproduction μ, conditionné à avoir n sommets ;
- Conditionnellement à $\mathcal{T}_n$, $(X_v)_{v\in\mathcal{V}(\mathcal{T}_n)}$ est une marche aléatoire de loi de saut θ, indexée par $\mathcal{T}_n$, et $\mathcal{R}_n = \#\{X_v : v \in \mathcal{V}(\mathcal{T}_n)\}$.

Théorème

Supposons que θ est centrée, non portée sur un sous-groupe de $\mathbb{Z}^d$ + conditions supplémentaires sur les moments.

1. *Si $d \geq 5$, $n^{-1}\mathcal{R}_n \xrightarrow[n\to\infty]{(\mathrm{P})} c_{\mu,\theta} > 0$;*
2. *Si $d = 4$ et $\mu(k) = 2^{-k-1}$, $\frac{\log n}{n}\mathcal{R}_n \xrightarrow[n\to\infty]{\mathrm{L}^2} 8\,\pi^2\,\sigma^4$, où $\sigma^2 = (\det(\operatorname{cov}\theta))^{1/4}$;*
3. *Si $d \leq 3$ et $\sum e^{\lambda k}\mu(k) < \infty$, $n^{-d/4}\mathcal{R}_n \xrightarrow[n\to\infty]{(\mathrm{d})} c_{\mu,\theta}\,\mathrm{Leb}(\mathrm{supp}(\mathcal{I}))$,*
   où $\mathcal{I}$ est la mesure aléatoire ISE sur $\mathbb{R}^d$.
   (ISE = integrated Super-Brownian excursion)

La croissance linéaire de $\mathcal{R}_n$

La convergence

$$\frac{\mathcal{R}_n}{n} \underset{n\to\infty}{\longrightarrow} c_{\mu,\theta} \geq 0$$

pourrait aussi se déduire du théorème ergodique sous-additif de Kingman.
(Trouver une application de shift τ sur un espace d'arbres convenable, et une mesure de probabilité invariante et ergodique sous τ.)

La croissance linéaire de $\mathcal{R}_n$

La convergence

$$\frac{\mathcal{R}_n}{n} \underset{n\to\infty}{\longrightarrow} c_{\mu,\theta} \geq 0$$

pourrait aussi se déduire du théorème ergodique sous-additif de Kingman. (Trouver une application de shift τ sur un espace d'arbres convenable, et une mesure de probabilité invariante et ergodique sous τ.)

On suppose que θ est centrée et a des moments d'ordre $d-1$ finis.

- Si μ est de variance finie, alors $c_{\mu,\theta} > 0$ quand $d \geq 5$;
- Si μ appartient au domaine d'attraction d'une loi stable d'indice $\alpha \in (1,2)$, alors $c_{\mu,\theta} > 0$ quand $d > \frac{2\alpha}{\alpha-1}$.

Les dimensions sous-critiques $d \leqslant 3$

$$\boxed{\frac{\mathcal{R}_n}{n^{d/4}} \xrightarrow[n\to\infty]{\text{(d)}} c_{\mu,\theta}\,\mathrm{Leb}(\mathrm{supp}(\mathcal{I}))}$$

Remarque. la variable $\mathrm{Leb}(\mathrm{supp}(\mathcal{I}))$ est non-dégénérée si et seulement si $d \leq 3$.

Les dimensions sous-critiques $d \leqslant 3$

$$\boxed{\frac{\mathcal{R}_n}{n^{d/4}} \xrightarrow[n\to\infty]{\text{(d)}} c_{\mu,\theta}\,\mathrm{Leb}(\mathrm{supp}(\mathcal{I}))}$$

Remarque. la variable $\mathrm{Leb}(\mathrm{supp}(\mathcal{I}))$ est non-dégénérée si et seulement si $d \leq 3$.

Hypothèses :

- $\sum_{k=0}^{\infty} e^{\lambda k}\mu(k) < \infty$, pour un $\lambda > 0$;
- θ est centrée, et elle a des moments d'ordre 4 finis.

Janson et Marckert (2005) :

$$cn^{-1/4}(X_v)_{v\in\mathcal{V}(\mathcal{T}_n)} \xrightarrow[n\to\infty]{} \text{Mouvement brownien indexé par } \mathbf{T_e},$$

où $c = \frac{1}{\sigma}\sqrt{\frac{\rho}{2}}$ ($\rho^2 = \mathrm{var}\,\mu$ et supposons que $\mathrm{cov}\,\theta = \sigma^2\,\mathrm{Id}$).

Les dimensions sous-critiques $d \leqslant 3$

En particulier,

$$\frac{1}{n}\sum_{v\in\mathcal{V}(\mathcal{T}_n)}\delta_{\frac{cX_v}{n^{1/4}}}$$

Les dimensions sous-critiques $d \leqslant 3$

En particulier,

$$\frac{1}{n} \sum_{v \in \mathcal{V}(\mathcal{T}_n)} \delta_{\frac{cX_v}{n^{1/4}}} \xrightarrow[n\to\infty]{\text{(d)}} \mathcal{I} \qquad \text{dans } \mathcal{M}_1(\mathbb{R}^d).$$

La mesure ISE $\mathcal{I}$ se définit comme la mesure d'occupation du mouvement brownien indexé par $\mathbf{T_e}$.

Les dimensions sous-critiques $d \leqslant 3$

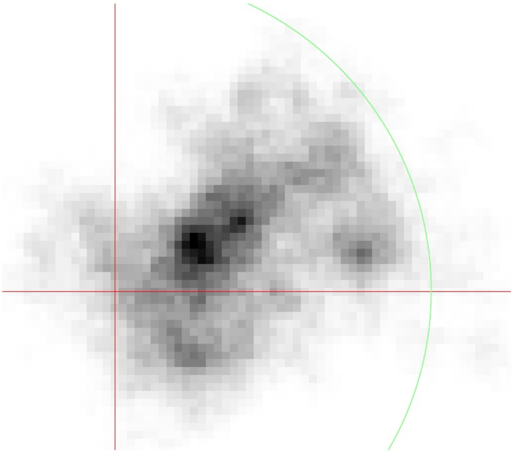
En particulier,

$$\frac{1}{n} \sum_{v \in \mathcal{V}(\mathcal{T}_n)} \delta_{\frac{cX_v}{n^{1/4}}} \xrightarrow[n\to\infty]{\text{(d)}} \mathcal{I} \qquad \text{dans } \mathcal{M}_1(\mathbb{R}^d).$$

La mesure ISE $\mathcal{I}$ se définit comme la mesure d'occupation du mouvement brownien indexé par $\mathbf{T_e}$.

Pour conclure :
l'existence de densité pour la mesure $\mathcal{I}$ + la convergence des temps locaux...

Distribution de Masse



Une simulation de $\sum_{v \in \mathcal{V}(\mathcal{T}_{\boldsymbol{n}})} \delta_{X_{\boldsymbol{v}}}$ avec $n \approx 96000$. (©Gordon Slade)

La dimension critique $d = 4$

Hypothèses :

- $\mu(k) = 2^{-k-1}$ (géométrique) ;
- θ est symétrique, et elle a des moments exponentiels finis.

$$\boxed{\frac{\log n}{n} \mathcal{R}_n \xrightarrow[n\to\infty]{\mathrm{L}^2} 8\,\pi^2\,\sigma^4}$$

IDÉE : Utiliser une chaîne de Markov à valeurs dans un espace de trajectoires, appelée le serpent discret.

PROBLÈME :
Si μ n'est pas géométrique, cette approche par l'utilisation du serpent discret ne fonctionne plus, et la situation devient plus compliquée !

Référence

J.-F. Le Gall & L. (2013)
The range of tree-indexed random walk
http://arxiv.org/abs/1307.5221

J.-F. Le Gall & L. (2014)
The range of tree-indexed random walk in low dimensions
http://arxiv.org/abs/1401.7830