Scaling limits of inhomogeneous trees and graphs

Part-II. Scaling limits of critical Galton–Watson trees

Minmin Wang (University of Bath)

BUC IX, Guanajuato, 11-15 December 2017

Outline

Aim: study the scaling limits of critical Galton–Watson trees.

Main steps:

- Encode Galton–Watson trees with stochastic processes.

  Based on the survey article of Le Gall '05, *Random trees and applications*.

- Construct potential limit processes for the discrete coding functions.

  Based on the monograph of Duquesne & Le Gall '02, *Random trees, Lévy processes and spatial branching processes*, Chapter 1.
  See also Le Gall & Le Jan '98, *Branching processes in Lévy processes: the exploration process*.

- Convergence of the discrete coding functions to the continuous ones.

  Based on the monograph of Duquesne & Le Gall '02, *Random trees, Lévy processes and spatial branching processes*, Chapter 2.

Outline

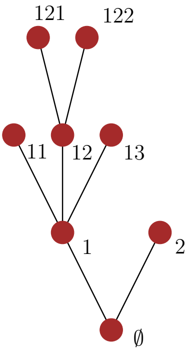
- Coding functions of Galton–Watson trees
- Construction of the potential limit processes
- Convergence of the coding functions of Galton–Watson forests

Galton–Watson trees

- Let $\mathcal{T}$ be a Galton–Watson tree with offspring distribution μ.
- We suppose that

$$\text{(Criticality)} \qquad \sum_k k\,\mu(k) = 1.$$

- **Ulam–Harris tree**: we identify $\mathcal{T}$ with the set of its node-labels and refer to a node of $\mathcal{T}$ by its label.



Lukasiewicz walk of a Galton–Watson tree

- We sort the nodes of $\mathcal{T}$ using the lexicographic order:

  e.g. $\varnothing < 1 < 11 < 12 < \cdots < 2 < \cdots$

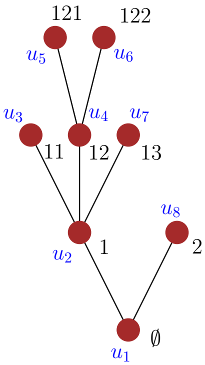
  Denote by $(u_i)_{1\leq i\leq |\mathcal{T}|}$ the sequence of the nodes of $\mathcal{T}$ ranked in the lexicographic order.

- For a node $u \in \mathcal{T}$, denote by k_u the number of children of u.

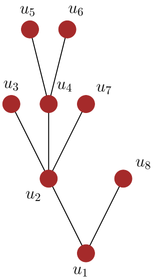
- Define $S = (S_i)_{0\leq i\leq |\mathcal{T}|}$ by

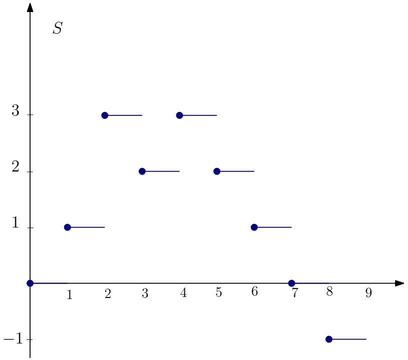
  $S_0 = 0 \quad$ and for $1 \leq n \leq |\mathcal{T}|$: $\quad S_i - S_{i-1} = k_{u_i} - 1$

  S is called the **Lukasiewicz walk** of $\mathcal{T}$.



Lukasiewicz walk: combinatorial properties





Proposition

$S_i \geq 0$ for all $0 \leq i < |\mathcal{T}|$ and $S_{|\mathcal{T}|} = -1$.

Lukasiewicz walk: distributional properties

Proposition

For a Galton–Watson tree with offspring distribution μ, $(k_{u_n})_{1\leq n\leq|\mathcal{T}|}$ is an i.i.d. sequence with common distribution μ.

Corollary

$$S \overset{\text{(d)}}{=} (V_i)_{0\leq i\leq\tau_1},$$

where

- $(V_i)_{i\geq 0}$ is a random walk where $V_i - V_{i-1}$ has distribution ν given by
$$\nu(k) = \mu(k+1), \quad k = -1, 0, \cdots ;$$
- $\tau_x = \inf\{i \geq 0 : V_i \leq -x\}$, $x \geq 0$.

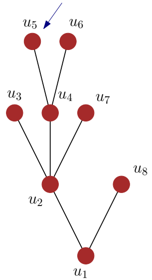
Rem. The above Proposition is not so obvious as it might seem: (u_n) is a (random) ordering which depends on $\mathcal{T}$ itself! See Le Gall's survey for a careful proof.

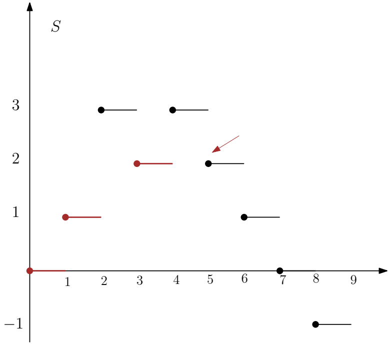
A key observation

Question: how can we recover $\mathcal{T}$ from S?

A key observation

Question: how can we recover $\mathcal{T}$ from S?





Theorem

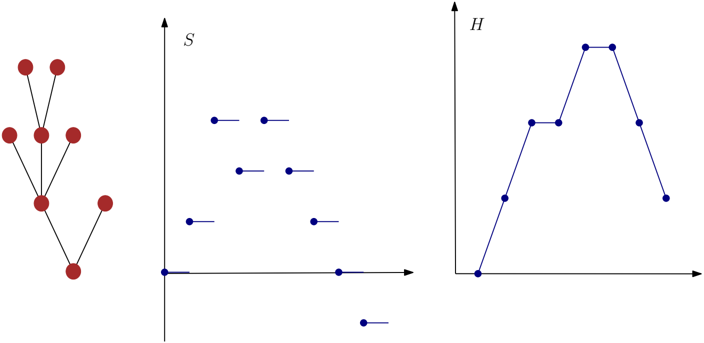
Denote by $h(u_i)$ = height of the node u_i in $\mathcal{T}$. For all $1 \leq i \leq |\mathcal{T}|$, we have

$$h(u_i) = \#\{j \in \{0, 1, \ldots, i-1\} : S_j = \min_{j \leq l \leq i} S_l\}.$$

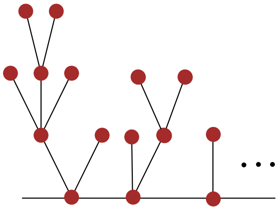
Height process of a Galton–Watson tree

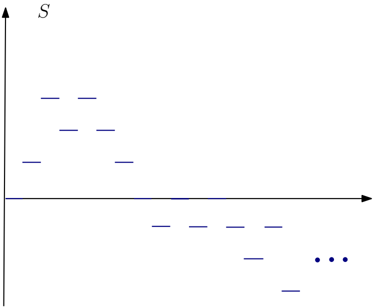
- Set $H_0 = 0$ and $H_i = h(u_i)$, for $i = 1, 2, \dots |\mathcal{T}|$.
- Let $H = (H_t)_{t \geq 0}$ be the linear interpolation of $(H_i)_{0 \leq i \leq |\mathcal{T}|}$.

The process H is called the **height process** of $\mathcal{T}$.



Lukasiewicz walk of a Galton–Watson forest

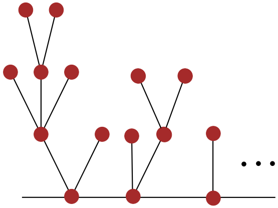


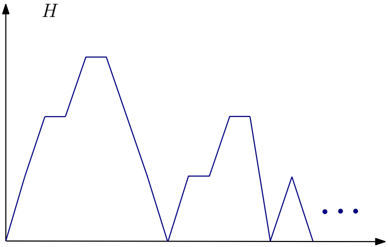


Let $(\mathcal{T}_i)_{i\geq 1}$ be i.i.d. copies of $\mathcal{T}$. Let S be the concatenation of the Lukasiewicz walks of $\mathcal{T}_1, \mathcal{T}_2, \ldots$. Then we have

- S is a random walk with increment distribution ν, where $\nu(k+1) = \mu(k)$, $k \geq -1$.
- Set $\underline{S}_i = \inf_{j\leq i} S_j$. Then $-\underline{S}_i =$ index of the tree we are exploring at time i.
- Let $e^{(k)} = k$-th excursion of $S - \underline{S} =$ portion of $S - \underline{S}$ between τ_{k-1} and τ_k. Then $e^{(k)} =$ Lukasiewicz walk of $\mathcal{T}_k$.

Height process of a Galton–Watson forest





Recall that $\mathcal{T}_i, i \geq 1$ are i.i.d. copies of $\mathcal{T}$. Let H be the concatenation of the height processes of $\mathcal{T}_1, \mathcal{T}_2, \ldots$.

- The previous relation between H and S can be extended to the forest case:

$$\forall\, i \geq 1 : \quad H_i = \#\big\{ j \in \{0, 1, \ldots, i-1\} : S_j = \min_{j \leq l \leq i} S_l \big\}$$

- The k-th excursion of H above 0 is the height process of $\mathcal{T}_k$.
  In particular, this entails the excursion intervals of H coincide with those of $S - \underline{S}$.

Outline

- Coding functions of Galton–Watson trees
- Construction of the potential limit processes
- Convergence of the coding functions of Galton–Watson forests

Limit process of Lukasiewicz walk

Question: What will be the analogue of S in the limit?

Recall that S is a spectrally positive random walk. Its scaling limit X will be a spectrally positive Lévy process.

We assume that

- X is a spectrally positive Lévy process;
- X is centred;
- sample paths of X have infinite variation.

Under these assumptions, the Laplace exponent ϕ of X has the following form:

$$\forall\,\lambda \geq 0: \quad \phi(\lambda) = \tfrac{1}{2}\beta\lambda^2 + \int (e^{-\lambda x} - 1 + \lambda x)\pi(dx),$$

where the Gaussian coefficient $\beta \in \mathbb{R}_+$ and the Lévy measure π is supported on $(0, \infty)$ and satisfying $\int (x \wedge x^2)\pi(dx) < \infty$. Moreover, either $\beta > 0$ or $\int_{(0,1)} x\,\pi(dx) = \infty$.

Limit process of height process

Question: What is the analogue of H in the limit?

Idea: Extend the relation

$$H_t = \#\big\{s \in \{0, 1, \ldots, t-1\} : S_s = \min_{s \leq u \leq t} S_u\big\}$$

to the Lévy process X. Intuitively, this amounts to putting an "appropriate" measure on the set

$$\big\{s < t : X_s = \inf_{s \leq u \leq t} X_u\big\}$$

Try with

- counting measure: $\#\big\{s < t : X_s = \inf_{s \leq u \leq t} X_u\big\} = \infty$, a.s.
- Lebesgue measure ℓ of $\mathbb{R}$: $\ell(\big\{s < t : X_s = \inf_{s \leq u \leq t} X_u\big\}) = 0$, a.s.

It turns out that here, one should put the local time measure.

Local times of the reflected processes

- Set $\overline{X}_t = \sup_{s \le t} X_s$. $X - \overline{X}$ is a strong Markov process and 0 is a regular point for this process.

  In particular, $\#\{s < t : X_s = \overline{X}_s\} = \infty$, while $\ell(\{s < t : X_s = \overline{X}_s\}) = 0$.

- Roughly speaking, the local time L_t for $X - \overline{X}$ measures the set $\{s < t : X_s = \overline{X}_s\}$.

- **Approximation of local times**: For each $t \ge 0$, the following limit exists a.s.

$$L_t := \lim_{\epsilon \to 0} \frac{1}{\epsilon} \ell\left(\left\{s \in [0, t] : X_s \ge \overline{X}_s - \epsilon\right\}\right) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t \mathbf{1}_{\{X_s \ge \overline{X}_s - \epsilon\}} ds.$$

  Moreover, $t \mapsto L_t$ is non decreasing and continuous. We refer to $L = (L_t)_{t \ge 0}$ as the local time process of $X - \overline{X}$.

- Let dL be the Stieltjes measure associated with L. Namely, dL is a measure on the Borel sets of $\mathbb{R}_+$ such that $dL((a, b]) = L_b - L_a$, $\forall\, a < b$. Then a.s. dL is supported on $\{t \ge 0 : X_t = \overline{X}_t\}$.

Height process of X

- **Time-reversal process**: fix $t > 0$, define $\hat{X}^{(t)}$ by

$$\hat{X}_s^{(t)} = X_t - X_{(t-s)-}, \quad 0 \le s \le t.$$

Law of a Lévy process is invariant under time-reversal: $\hat{X}^{(t)} \overset{(\mathrm{d})}{=} (X_s)_{s \le t}$. So the local time $L(\hat{X}^{(t)})$ is well-defined for the time-reversal process $\hat{X}^{(t)}$.

- For each $t \ge 0$, set

$$H_t := L_t(\hat{X}^{(t)}) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t \mathbf{1}_{\{\hat{X}_s^{(t)} \ge \overline{\hat{X}}_s^{(t)} - \epsilon\}} ds = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t \mathbf{1}_{\{X_s \le \inf_{s \le u \le t} X_u + \epsilon\}} ds.$$

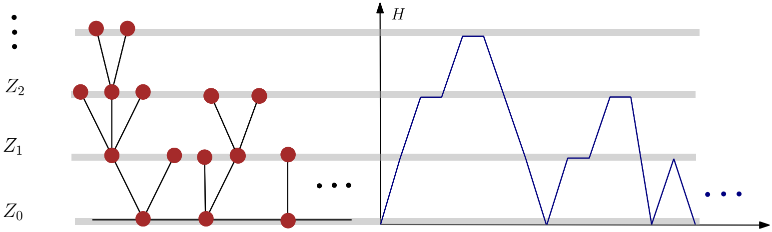
- It turns out that $H = (H_t)_{t \ge 0}$ has a continuous version if and only if

$$\text{(Grey)} \qquad \int^{\infty} \frac{d\lambda}{\phi(\lambda)} < \infty \,.$$

H is called the **height process** of X. It is non negative and each excursion of H above 0 encodes a "tree-like" metric space, called the **Lévy tree**.

Special case: if $X =$ Brownian Motion, then $H =$ reflected Brownian Motion.

Lévy trees as genealogy trees of CSBP



Analogously, we can define the "mass" of the Lévy forest at a given level a.

- There exists $(Z_t^a, t \geq 0, a \geq 0)$ satisfying for all ≥ 0, $t \mapsto Z_t^a$ is non decreasing and

$$Z_t^a = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t \mathbf{1}_{\{a \leq H_s \leq a+\epsilon\}} ds \quad \text{in probability.}$$

- Moreover, $\{Z_{\tau_x}^a : a \geq 0\}$ is distributed as a continuous-state branching process with branching mechanism ϕ and starting from x.
- Let $\zeta = \inf\{a > 0 : Z_{\tau_x}^a = 0\}$. Then under (Grey) condition, $\zeta < \infty$ a.s.

Outline

- Coding functions of Galton–Watson trees
- Construction of the potential limit processes
- Convergence of the coding functions of Galton–Watson forests

Convergence of Lukasiewicz walks

Notation: for each n, let

- μ_n = a critical offspring distribution;
- $\{\mathcal{T}_i^{(n)} : i \geq 1\}$ = independent GW-trees with common offspring distribution μ_n;
- $S^{(n)}$ = Lukasiewicz walk of $\{\mathcal{T}_i^{(n)} : i \geq 1\}$;
- $H^{(n)}$ = height process of $\{\mathcal{T}_i^{(n)} : i \geq 1\}$.

Recall that $S^{(n)}$ is a centred random walk with increment distribution

$$\nu_n(k) = \mu_n(k+1), \quad k = -1, 0, \ldots.$$

Recall that X is a spectrally positive and centred Lévy process with Laplace exponent

$$\phi(\lambda) = \tfrac{1}{2}\beta\lambda^2 + \int (e^{-\lambda x} - 1 + \lambda x)\pi(dx), \quad \lambda \geq 0.$$

In the previous session, we have seen sufficient and necessary conditions for

$$(\mathrm{CV_S}) \qquad a_n \to \infty, \qquad \left(\tfrac{1}{a_n} S^{(n)}_{\lfloor nt \rfloor} : t \geq 0\right) \xrightarrow{d} X \quad \text{in } \mathbb{D}.$$

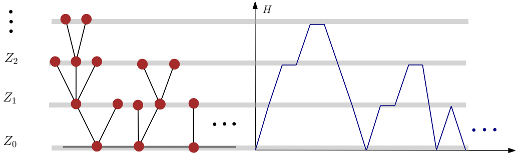
In what follows, we assume that $(\mathrm{CV_S})$ takes place.

Grimvall's Theorem

Question: What is the scaling for $H^{(n)}$?

Grimvall's Theorem

Question: What is the scaling for $H^{(n)}$?



Theorem (Grimvall '74)

- For $n \geq 1$, let $(Z_k^{(n)})_{k \geq 0}$ be a Galton–Watson branching process with offspring distribution μ_n starting from $Z_0^{(n)} = a_n$.
- Let $Z = (Z_t)_{t \geq 0}$ be a continuous-state branching process with branching mechanism ϕ starting rom $Z_0 = 1$.

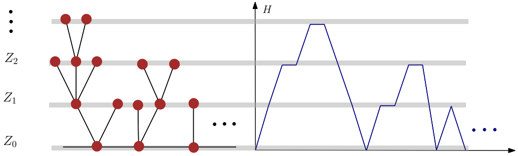
Then the following statements are equivalent

$$(\mathrm{CV_S}) \qquad \left(\frac{1}{a_n} S^{(n)}_{\lfloor nt \rfloor} : t \geq 0\right) \xrightarrow{d} X \quad \text{in } \mathbb{D}.$$

$$(\mathrm{CV_Z}) \qquad \left(\frac{1}{a_n} Z^{(n)}_{\lfloor nt/a_n \rfloor} : t \geq 0\right) \to Z \quad \text{in } \mathbb{D}.$$

This suggests that the scaling for $H^{(n)}$ should be $\frac{a_n}{n} H^{(n)}_{\lfloor nt \rfloor}$.

Marginal convergence of height processes



Proposition

Suppose that $(\mathrm{CV_S})$ takes place. Then for all $k \geq 1$ and $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k$,

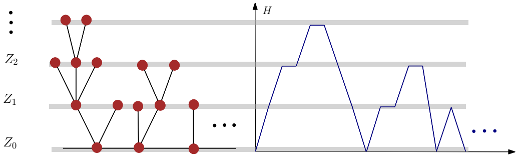
$$\left(\frac{a_n}{n} H^{(n)}_{\lfloor nt_1 \rfloor}, \frac{a_n}{n} H^{(n)}_{\lfloor nt_2 \rfloor}, \ldots, \frac{a_n}{n} H^{(n)}_{\lfloor nt_k \rfloor}\right) \xrightarrow{d} \left(H_{t_1}, H_{t_2}, \ldots, H_{t_k}\right).$$

Proof idea: recall that $H_t = L_t(\hat{X}^{(t)})$. Similarly, one can write $H^{(n)}_k$ as the discrete local time of the time-reversal process of $S^{(n)}$ at k. Then the proof for $k = 1$ reduces to proving the convergence of local times.

Question: Can we improve this into a uniform convergence?

Need to control fluctuations of $H^{(n)}$.

Marginal convergence of height processes



Proposition

Suppose that $(\mathrm{CV_S})$ takes place. Then for all $k \geq 1$ and $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k$,

$$\left(\frac{a_n}{n} H^{(n)}_{\lfloor nt_1 \rfloor}, \frac{a_n}{n} H^{(n)}_{\lfloor nt_2 \rfloor}, \ldots, \frac{a_n}{n} H^{(n)}_{\lfloor nt_k \rfloor}\right) \xrightarrow{d} \left(H_{t_1}, H_{t_2}, \ldots, H_{t_k}\right).$$

Proof idea: recall that $H_t = L_t(\hat{X}^{(t)})$. Similarly, one can write $H^{(n)}_k$ as the discrete local time of the time-reversal process of $S^{(n)}$ at k. Then the proof for $k = 1$ reduces to proving the convergence of local times.

Question: Can we improve this into a uniform convergence?

Need to control extinction times of branching processes $Z^{(n)}$.

Functional convergence of height processes

Theorem (Duquesne & Le Gall, '02)

Suppose that $(\mathrm{CV_S})$ holds. Let $\zeta_n = \inf\{k : Z_k^{(n)} = 0\}$. Suppose that $\exists\, \delta > 0$ s.t.

$$(\mathrm{Extinction}) \qquad \liminf_{n\to\infty} \mathbb{P}(\zeta_n \le \delta n/a_n) > 0.$$

Then we have

$$(\mathrm{CV_H}) \qquad \left(\tfrac{a_n}{n} H^{(n)}_{\lfloor nt\rfloor} : t \ge 0\right) \xrightarrow{d} H \quad \text{uniformly} .$$

Rem. It is easy to see that $(\mathrm{Extinction})$ is a necessary condition for $(\mathrm{CV_H})$. On the other hand, it is easy to construct examples where $(\frac{1}{a_n} Z^{(n)}_{\lfloor nt/a_n\rfloor})_{t\ge 0}$ converges but $(\mathrm{Extinction})$ does not hold.

Rem. If $\mu_n =$ some μ for all n and $(\mathrm{CV_S})$ holds, then (Extinction) is automatically satisfied.

Scaling limits of Galton–Watson trees: general case

For the critical case, namely,

- offspring distribution of the Galton–Watson trees has mean 1;
- limit Lévy process is centred;

Scaling limits of Galton–Watson trees: general case

For the subcritical case, namely,

- offspring distribution of the Galton–Watson trees has mean ≤ 1;
- limit Lévy process has negative drift;

one can check that

- the Lukasiewicz walks and height processes are also well-defined;
- the definition of the height process can be easily extended;
- same statement for the convergence of the discrete height processes towards the limit ones.

On the other hand, in the supercritical case, Galton–Watson trees have positive probabilities to be infinite. In that case, one can truncate Galton–Watson trees (resp. Lévy trees) at a given level and consider the corresponding coding processes; see Abraham & Delmas '12, He & Luan '13.

Summary

- Beautiful combinatorial encodings of Galton–Watson trees:

  Lukasiewicz walk, local-time type relation between the Lukasiewicz walk and height process, ...

- Application of powerful probabilistic tools:

  construction of local times, connection between spectrally positive Lévy processes and continuous-state branching processes, construction of height process, convergence of Lukasiewicz walk, ...

¡Muchas Gracias!