Functional limit theorems for the profile of random recursive trees

Alexander Iksanov, Kyiv, Ukraine

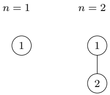
based on joint work with Zakhar Kabluchko (Münster)

Conference ‘Probability and Analysis 2019’, May 20-24, 2019, Będlewo, Poland

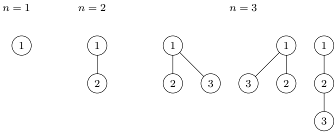
Non-plane recursive trees:
deterministic trees

$n = 1$

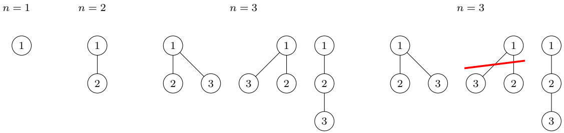
Non-plane recursive trees:
deterministic trees



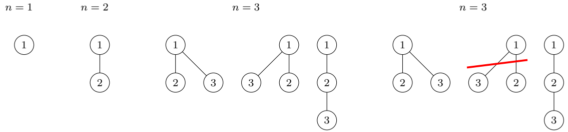
Non-plane recursive trees:
deterministic trees



Non-plane recursive trees:
deterministic trees



Non-plane recursive trees:
deterministic trees



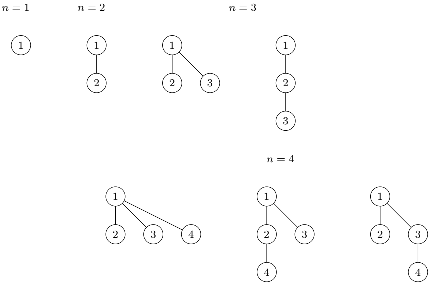
Denote by a_n the number of non-plane deterministic recursive trees on n vertices. Then

$$a_{n+1} = n a_n, \quad a_1 = 1,$$

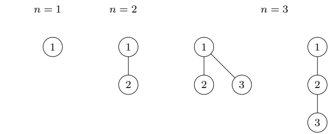
whence

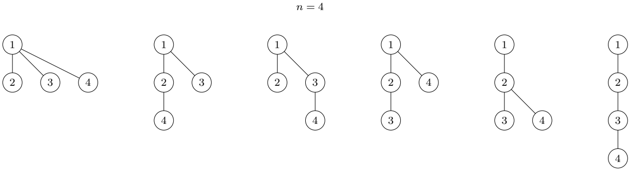
$$a_n = (n-1)!, \quad n \in \mathbb{N}.$$

Non-plane recursive trees:
deterministic trees



Non-plane recursive trees:
deterministic trees





Non-plane recursive trees:
random trees

A random object $\mathcal{R}_n$ is called random recursive tree with n vertices if

$$\mathbb{P}\{\mathcal{R}_n = R_{i,n}\} = \frac{1}{(n-1)!}, \quad i = 1, 2, \ldots, (n-1)!,$$

where $R_{1,n}$, $R_{2,n}, \ldots, R_{(n-1)!,n}$ are deterministic recursive trees with n vertices.

Non-plane recursive trees:
random trees

A random object $\mathcal{R}_n$ is called random recursive tree with n vertices if

$$\mathbb{P}\{\mathcal{R}_n = R_{i,n}\} = \frac{1}{(n-1)!}, \quad i = 1, 2, \ldots, (n-1)!,$$

where $R_{1,n}$, $R_{2,n}, \ldots, R_{(n-1)!,n}$ are deterministic recursive trees with n vertices.

A way to generate random recursive trees :

At time 0 start with a tree consisting of the sole root.

At time n, given a random recursive tree with $n+1$ vertices, pick a vertex uniformly at random and attach to this vertex the new vertex $n+2$.

Non-plane recursive trees:
random trees

A random object $\mathcal{R}_n$ is called random recursive tree with n vertices if

$$\mathbb{P}\{\mathcal{R}_n = R_{i,n}\} = \frac{1}{(n-1)!}, \quad i = 1, 2, \ldots, (n-1)!,$$

where $R_{1,n}$, $R_{2,n}, \ldots, R_{(n-1)!,n}$ are deterministic recursive trees with n vertices.

A way to generate random recursive trees :

At time 0 start with a tree consisting of the sole root.

At time n, given a random recursive tree with $n+1$ vertices, pick a vertex uniformly at random and attach to this vertex the new vertex $n+2$.

A random tree obtained at time n has the same distribution as $\mathcal{R}_{n+1}$!

Non-plane recursive trees:
profile

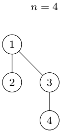
For $k \in \mathbb{N}_0$, denote by $X_n(k)$ the number of vertices at level k in a recursive tree with $n+1$ vertices (rather than n vertices!). Note that $X_n(0) = 1$.

The function $k \mapsto X_n(k)$ for $k = 0, 1, \ldots, n$ is called profile of the tree.

Non-plane recursive trees:
profile

For $k \in \mathbb{N}_0$, denote by $X_n(k)$ the number of vertices at level k in a recursive tree with $n+1$ vertices (rather than n vertices!). Note that $X_n(0) = 1$.

The function $k \mapsto X_n(k)$ for $k = 0, 1, \ldots, n$ is called profile of the tree.



$X_3(0) = 1,\ X_3(1) = 2,\ X_3(2) = 1,\ X_3(3) = 0$

Non-plane recursive trees:
profile

For $k \in \mathbb{N}_0$, denote by $X_n(k)$ the number of vertices at level k in a recursive tree with $n+1$ vertices (rather than n vertices!).

Pittel (1994) proved that

$$\lim_{n\to\infty} \frac{\max\{k \in \mathbb{N} : X_n(k) \neq 0\}}{\log n} = e \quad \text{a.s.},$$

that is, random recursive trees are of logarithmic height.

Asymptotics of the profile of random recursive trees:
low levels profile

For $u \geq 0$ and $n \in \mathbb{N}$, $X_{\lfloor n^u \rfloor}(k)$ is the number of vertices at level k in a random recursive tree with $\lfloor n^u \rfloor + 1$ vertices.

Theorem (I. & Kabluchko (2018))
The following functional limit theorem holds for the low levels profile of a random recursive tree:

$$\left(\frac{(k-1)!\big(X_{\lfloor n^u \rfloor}(k) - (u\log n)^k/k!\big)}{(\log n)^{k-1/2}}\right)_{k\in\mathbb{N}}$$

$$\Rightarrow \quad \left(\int_{[0,\,u]} (u-y)^{k-1}\mathrm{d}B(y)\right)_{k\in\mathbb{N}} =: (R_k(u))_{k\in\mathbb{N}}$$

in the product J_1-topology on $D^{\mathbb{N}}$, where $(B(u))_{u\geq 0}$ is a standard Brownian motion and $D = D[0,\infty)$ is the Skorokhod space.

Asymptotics of the profile of random recursive trees:
low levels profile

For $u \geq 0$ and $n \in \mathbb{N}$, $X_{\lfloor n^u \rfloor}(k)$ is the number of vertices at level k in a random recursive tree with $\lfloor n^u \rfloor + 1$ vertices.

Theorem (I. & Kabluchko (2018))
The following functional limit theorem holds for the low levels profile of a random recursive tree:

$$\left(\frac{(k-1)!\big(X_{\lfloor n^u \rfloor}(k) - (u \log n)^k / k!\big)}{(\log n)^{k-1/2}} \right)_{k \in \mathbb{N}}$$

$$\Rightarrow \quad \left(\int_{[0,\, u]} (u-y)^{k-1} \mathrm{d}B(y) \right)_{k \in \mathbb{N}} =: (R_k(u))_{k \in \mathbb{N}}$$

in the product J_1-topology on $D^{\mathbb{N}}$, where $(B(u))_{u \geq 0}$ is a standard Brownian motion and $D = D[0, \infty)$ is the Skorokhod space.

$R_1(u) = B(u)$, $R_2(u) = \int_0^u B(y)\mathrm{d}y$, $R_3(u) = 2\int_0^u \int_0^s B(y)\mathrm{d}y\mathrm{d}s, \ldots$

Asymptotics of the profile of random recursive trees:
intermediate levels

What about the allocation of vertices of a random recursive tree with $n+1$ vertices over the levels $\lfloor k_n u \rfloor$, $u > 0$, where $k_n \to \infty$ and $k_n = o(\log n)$ as $n \to \infty$?

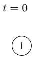
For $k \in \mathbb{N}_0$ and $n \in \mathbb{N}$, $X_n(k)$ is the number of vertices at level k in a random recursive tree with $n+1$ vertices.

Theorem (I. & Kabluchko (2018))
Let $(k_n)_{n \in \mathbb{N}}$ be a sequence of positive numbers satisfying $k_n \to \infty$ and $k_n = o(\log n)$ as $n \to \infty$. The following limit theorem holds for the intermediate levels of a random recursive tree with $n+1$ vertices:

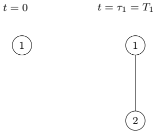
$$\left(\frac{\lfloor k_n \rfloor^{1/2} (\lfloor k_n u \rfloor - 1)! \big(X_n(\lfloor k_n u \rfloor) - (\log n)^{\lfloor k_n u \rfloor} / \lfloor k_n u \rfloor ! \big)}{(\log n)^{\lfloor k_n u \rfloor - 1/2}} \right)_{u>0} \overset{\text{f.d.d.}}{\Rightarrow} \left(\int_{[0,\infty)} e^{-uy} \mathrm{d}B(y) \right)_{u>0},$$

where $(B(u))_{u \ge 0}$ is a standard Brownian motion.

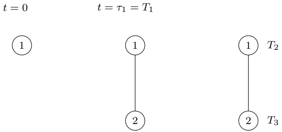
Idea of proofs:
random recursive trees in continuous time



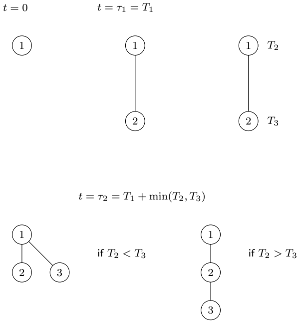
Idea of proofs:
random recursive trees in continuous time



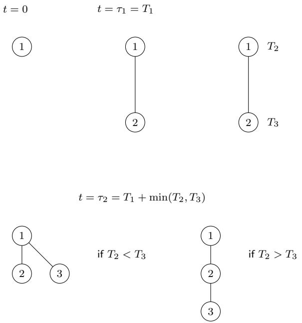
Idea of proofs:
random recursive trees in continuous time



Idea of proofs:
random recursive trees in continuous time



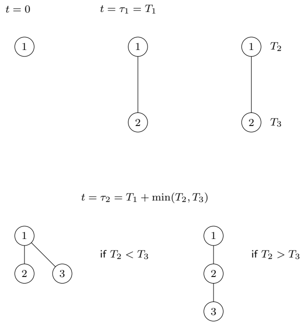
Idea of proofs:
random recursive trees in continuous time



Denote by τ_n the time at which the vertex $n+1$ is added to the tree.

A random tree obtained at time τ_n has the same distribution as $\mathcal{R}_{n+1}$.

Idea of proofs:
random recursive trees in continuous time

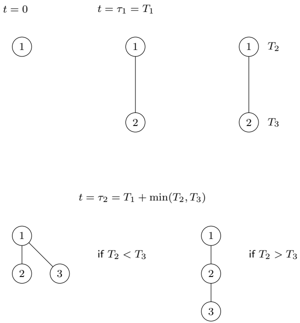


Denote by τ_n the time at which the vertex $n+1$ is added to the tree.

A random tree obtained at time τ_n has the same distribution as $\mathcal{R}_{n+1}$.

Observe that $\tau_1, \tau_2 - \tau_1, \ldots, \tau_n - \tau_{n-1}$ are independent, and $\tau_j - \tau_{j-1}$ has an exponential distribution of mean $1/j$.

Idea of proofs:
random recursive trees in continuous time



Denote by τ_n the time at which the vertex $n+1$ is added to the tree.

A random tree obtained at time τ_n has the same distribution as $\mathcal{R}_{n+1}$.

Observe that $\tau_1, \tau_2 - \tau_1, \ldots, \tau_n - \tau_{n-1}$ are independent, and $\tau_j - \tau_{j-1}$ has an exponential distribution of mean $1/j$.

The construction was used by Pittel (1994).

Idea of proofs:
CMJ-branching process generated by standard random walk

Let $\xi_1, \xi_2, \ldots$ be i.i.d. positive random variables. Put

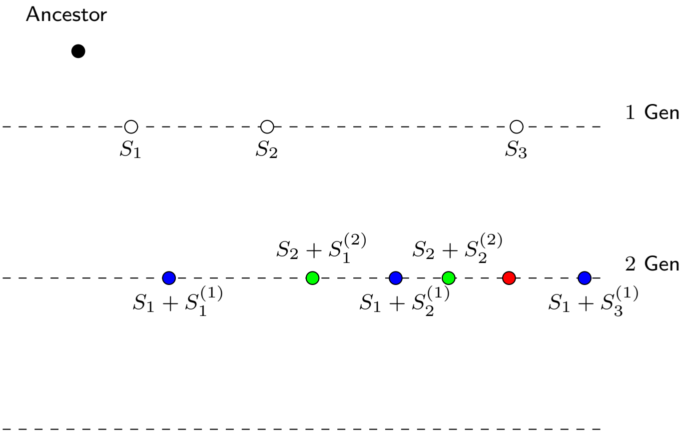
$$S_n := \xi_1 + \ldots + \xi_n, \quad n \in \mathbb{N}; \quad N(t) := \sum_{n \geq 1} \mathbb{1}_{\{S_n \leq t\}}, \quad t \geq 0.$$

$S := (S_n)_{n \in \mathbb{N}}$ – standard random walk

$(N(t))_{t \geq 0}$ – renewal process

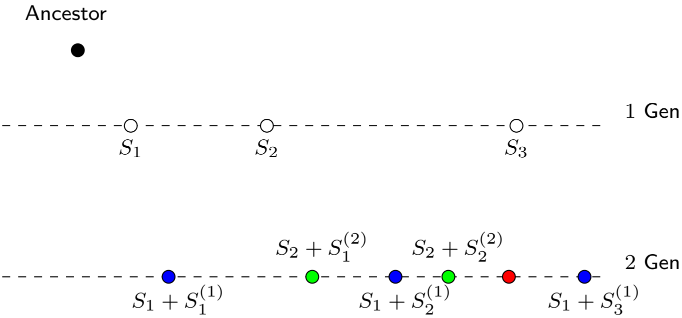
Idea of proofs:
CMJ-branching process generated by standard random walk

Crump-Mode-Jagers branching process generated by the standard random walk S



Crump-Mode-Jagers branching process generated by the standard random walk S

For $k \in \mathbb{N}$, $Y_k(t)$ – the number of the kth generation individuals with birth times $\leq t$. In particular, $(Y_1(t))_{t\geq 0} = (N(t))_{t\geq 0}$ is a renewal process.



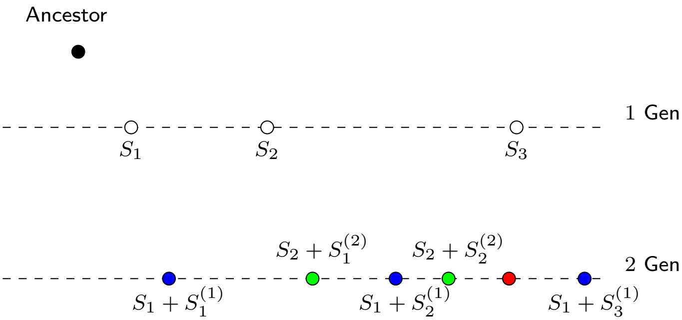
Idea of proofs:

CMJ-branching process generated by standard random walk

Let $\xi_1, \xi_2, \ldots \sim \mathrm{Exp}(1)$

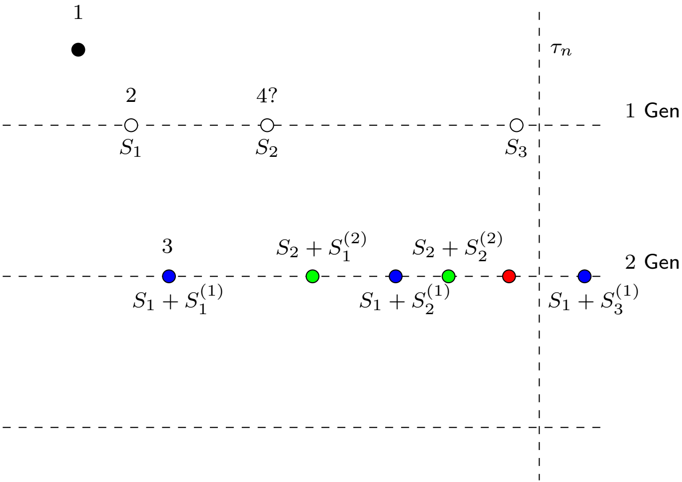
$S_1, S_2, S_3, \ldots$ – the times when the vertices of infinite random recursive tree are attached to the root

Important connection: If the time is infinite the kth generation of the branching process corresponds to the kth level of infinite recursive tree.



Idea of proofs:

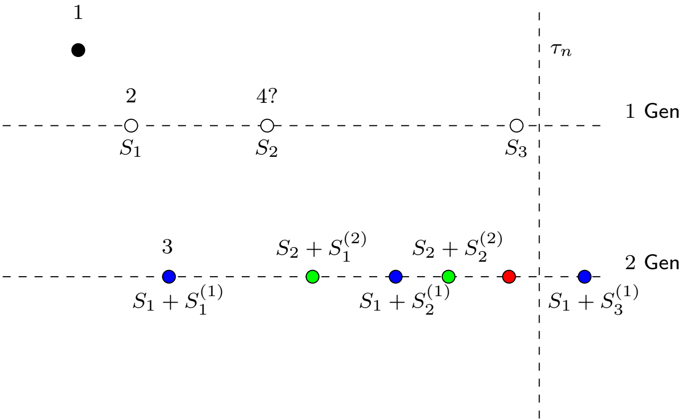
CMJ-branching process generated by standard random walk



Idea of proofs:
CMJ-branching process generated by standard random walk

Basic observation:

$$(X_{\lfloor n^u \rfloor}(k))_{u\geq 0, k\in\mathbb{N}} \overset{\mathrm{d}}{=} (Y_k(\tau_{\lfloor n^u \rfloor}))_{u\geq 0, k\in\mathbb{N}}.$$



Idea of proofs:
CMJ-branching process generated by standard random walk

Basic observation:

$$(X_{\lfloor n^u \rfloor}(k))_{u\geq 0, k\in\mathbb{N}} \overset{\mathrm{d}}{=} (Y_k(\tau_{\lfloor n^u \rfloor}))_{u\geq 0, k\in\mathbb{N}} \overset{\mathrm{d}}{\approx} (Y_k(u\log n))_{u\geq 0, k\in\mathbb{N}}.$$

Why? $(\tau_n - (1+1/2+\ldots+1/n))_{n\in\mathbb{N}}$ is a square-integrable martingale. Hence, for all $T>0$,

$$\lim_{n\to\infty} \sup_{u\in[0,\,T]} |\tau_{\lfloor n^u \rfloor} - \log\lfloor n^u \rfloor| = \sup_{j\geq 1} |\tau_j - \log j| < \infty \quad \text{a.s.}$$

Idea of proofs:
CMJ-branching process generated by standard random walk

Basic observation:

$$(X_{\lfloor n^u \rfloor}(k))_{u\geq 0, k\in\mathbb{N}} \overset{\mathrm{d}}{=} (Y_k(\tau_{\lfloor n^u \rfloor}))_{u\geq 0, k\in\mathbb{N}}$$

$$\overset{\mathrm{d}}{\approx} (Y_k(u\log n))_{u\geq 0, k\in\mathbb{N}}.$$

Why? $(\tau_n - (1+1/2+\ldots+1/n))_{n\in\mathbb{N}}$ is a square-integrable martingale. Hence, for all $T>0$,

$$\lim_{n\to\infty}\sup_{u\in[0,\,T]}|\tau_{\lfloor n^u\rfloor} - \log\lfloor n^u\rfloor| = \sup_{j\geq 1}|\tau_j - \log j| < \infty \quad \text{a.s.}$$

and thereupon

$$\lim_{n\to\infty}\sup_{u\in[0,\,T]}\left|\frac{\tau_{\lfloor n^u\rfloor}}{\log n} - u\right| = 0 \quad \text{a.s.}$$

Idea of proofs:
CMJ-branching process generated by standard random walk

Basic observation:

$$(X_{\lfloor n^u \rfloor}(k))_{u\geq 0, k\in\mathbb{N}} \overset{\mathrm{d}}{=} (Y_k(\tau_{\lfloor n^u \rfloor}))_{u\geq 0, k\in\mathbb{N}}$$
$$\overset{\mathrm{d}}{\approx} (Y_k(u\log n))_{u\geq 0, k\in\mathbb{N}}.$$

Theorem (I. & Kabluchko (2018))
Suppose that $\sigma^2 := \mathrm{Var}\,\xi \in (0,\infty)$. Then the following functional limit theorem holds for counting process in the CMJ-branching process:

$$\left(\frac{(k-1)!\big(Y_k(ut)-(ut)^k/k!\big)}{\sqrt{\sigma^2\mu^{-2k-1}t^{2k-1}}}\right)_{k\in\mathbb{N}}$$
$$\Rightarrow \quad \left(\int_{[0,\,u]}(u-y)^{k-1}\mathrm{d}B(y)\right)_{k\in\mathbb{N}}$$

in the product J_1-topology on $D^{\mathbb{N}}$, where $(B(u))_{u\geq 0}$ is a standard Brownian motion and $\mu := \mathbb{E}\xi < \infty$.

Idea of proofs:
CMJ-branching process generated by standard random walk

How to prove the latter theorem? Content myself with checking

$$\frac{Y_2(ut)-(ut)^2/2}{\text{const}\, t^{3/2}} \overset{\mathrm{J}_1}{\Rightarrow} R_2(u)=\int_{[0,\,u]}(u-y)\mathrm{d}B(y), \quad t\to\infty$$

For $i\in\mathbb{N}$, consider the 1st generation individual born at time S_i and denote by $Y_1^{(i)}(t)$ the number of her offspring in the 2nd generation with birth times $\leq t+S_i$.

By the branching property, $(Y_1^{(1)}(t))_{t\geq 0}$, $(Y_1^{(2)}(t))_{t\geq 0}, \ldots$ are independent copies of $(Y_1(t))_{t\geq 0}$ (renewal process).

Idea of proofs:
CMJ-branching process generated by standard random walk

How to prove the latter theorem? Content myself with checking

$$\frac{Y_2(ut)-(ut)^2/2}{\text{const}\, t^{3/2}} \overset{\mathrm{J}_1}{\Rightarrow} R_2(u)=\int_{[0,\,u]}(u-y)\mathrm{d}B(y), \quad t\to\infty$$

For $i\in\mathbb{N}$, consider the 1st generation individual born at time S_i and denote by $Y_1^{(i)}(t)$ the number of her offspring in the 2nd generation with birth times $\leq t+S_i$.

By the branching property, $(Y_1^{(1)}(t))_{t\geq 0}$, $(Y_1^{(2)}(t))_{t\geq 0},\ldots$ are independent copies of $(Y_1(t))_{t\geq 0}$ (renewal process).

$$Y_2(t)=\sum_{i\geq 1}Y_1^{(i)}(t-S_i), \quad t\geq 0$$

This process is a particular instance of a random process with immigration at the epochs of a renewal process.

How to prove the latter theorem? Content myself with checking

$$\frac{Y_2(ut)-(ut)^2/2}{\text{const}\, t^{3/2}} \overset{J_1}{\Rightarrow} R_2(u)=\int_{[0,\,u]}(u-y)\mathrm{d}B(y),\quad t\to\infty$$

For $i\in\mathbb{N}$, consider the 1st generation individual born at time S_i and denote by $Y_1^{(i)}(t)$ the number of her offspring in the 2nd generation with birth times $\le t+S_i$.

$$Y_2(t)=\sum_{i\ge1}Y_1^{(i)}(t-S_i),\quad t\ge0$$

This process is a particular instance of a random process with immigration at the epochs of a renewal process.

- I. (2013)
- I., Marynych & Meiners (2014, 2017a, 2017b)
- I., Kabluchko & Marynych (2016)
- I. (2016) – survey

Relevant literature

A one dimensional version of our theorems reads as follows.

Theorem (Fuchs, Hwang & Neininger (2006))
Let $k = k_n = o(\log n)$ as $n \to \infty$. Then

$$\frac{(k-1)!(2k-1)(X_n(k) - (\log n)^k/k!)}{(\log n)^{k-1/2}} \stackrel{\text{d}}{\to} \text{Normal}\,(0,1), \quad n \to \infty,$$

where $X_n(k)$ is the number of vertices at level k of a random recursive tree with $n+1$ vertices.

Relevant literature

A one dimensional version of our theorems reads as follows.

Theorem (Fuchs, Hwang & Neininger (2006))
Let $k = k_n = o(\log n)$ as $n \to \infty$. Then

$$\frac{(k-1)!(2k-1)(X_n(k) - (\log n)^k/k!)}{(\log n)^{k-1/2}} \stackrel{\mathrm{d}}{\to} \mathrm{Normal}\,(0,1), \quad n \to \infty,$$

where $X_n(k)$ is the number of vertices at level k of a random recursive tree with $n+1$ vertices.

Backhausz & Móri (2012) – the only other article known to us which treats the low levels of random recursive trees

Relevant literature

A one dimensional version of our theorems reads as follows.

Theorem (Fuchs, Hwang & Neininger (2006))
Let $k = k_n = o(\log n)$ as $n \to \infty$. Then

$$\frac{(k-1)!(2k-1)(X_n(k) - (\log n)^k/k!)}{(\log n)^{k-1/2}} \overset{\mathrm{d}}{\to} \mathrm{Normal}\,(0,1), \quad n \to \infty,$$

where $X_n(k)$ is the number of vertices at level k of a random recursive tree with $n+1$ vertices.

Backhausz & Móri (2012) – the only other article known to us which treats the low levels of random recursive trees

What about levels close to $\log n$?

- Fuchs, Hwang & Neininger (2006)
- Schopp (2010)
- Drmota (2010)
- Kabluchko, Marynych & Sulzbach (2016)

Relevant literature

The methods exploited in these works are

- contraction method in combination with singularity analysis of generating functions;
- method of moments;
- martingale techniques;
- novel method invented by Kabluchko, Marynych & Sulzbach (2016) which is based on asymptotic Edgeworth-like expansions for the profile

Relevant literature

The methods exploited in these works are

- contraction method in combination with singularity analysis of generating functions;
- method of moments;
- martingale techniques;
- novel method invented by Kabluchko, Marynych & Sulzbach (2016) which is based on asymptotic Edgeworth-like expansions for the profile

Theorem (Fuchs, Hwang & Neininger (2006))
Let $k_n \sim \alpha \log n$ as $n \to \infty$ for some $\alpha \in (0, e)$. Then

$$\frac{X_n(\lfloor k_n \rfloor)}{\mathbb{E} X_n(\lfloor k_n \rfloor)} \stackrel{\mathrm{d}}{\to} X(\alpha), \quad n \to \infty,$$

where $X_n(k)$ is the number of vertices at level k of a random recursive tree with $n+1$ vertices. Here,

$$X(\alpha) \stackrel{\mathrm{d}}{=} \alpha U^{\alpha} X'(\alpha) + (1-U)^{\alpha} X''(\alpha),$$

where $X'(\alpha)$ and $X''(\alpha)$ are independent copies of $X(\alpha)$ which are also independent of a random variable U having a uniform distribution on $(0, 1)$.

THANK YOU FOR YOUR ATTENTION!