Cost functionals for large random trees

J.-F. DELMAS
Joint work with J.-S. DHERSIN and M. SCIAUVEAU

http://cermics.enpc.fr/~delmas

Ecole des Ponts, Univ. Paris-Est

AofA 2016 (Krakow, Poland)

Rooted ordered trees

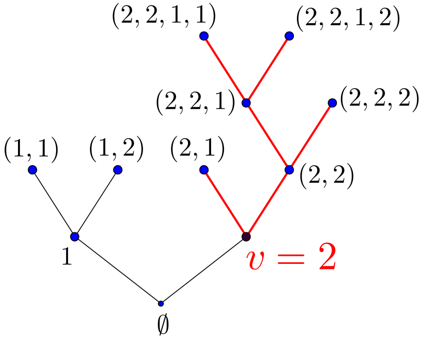


Figure: Binary tree with 5 internal nodes

- $\mathbf{t}$ rooted ordered tree.
- Cardinal of $\mathbf{t}$: $|\mathbf{t}|$.
- Leaves of $\mathbf{t}$: $\mathcal{L}(\mathbf{t})$.
- Sub-tree of $\mathbf{t}$ above $v \in \mathbf{t}$: $\mathbf{t}_v$.
- If $\mathbf{t}$ is also full binary then:
  - $|\mathbf{t}| = 2|\mathcal{L}(\mathbf{t})| - 1$.
  - Left-sub-tree of $\mathbf{t}$: $\mathbf{t}_1$.
  - Right-sub-tree of $\mathbf{t}$: $\mathbf{t}_2$.

Additive functionals and applications

- Unnormalized measure: $\forall f$ continuous defined on $[0, 1]$,

$$\mathcal{A}_{\mathbf{t}}^*(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v| \; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right).$$

- Link with known functionals
  - Total path length: $P(\mathbf{t}) = \sum_{w\in\mathbf{t}} d(\emptyset, w)$. As $d(\emptyset, w) = \sum_{v\in\mathbf{t}} \mathbf{1}_{\{v\prec w\}}$, we get:

$$P(\mathbf{t}) = \sum_{v\in\mathbf{t}} \sum_{w\in\mathbf{t}} \mathbf{1}_{\{v\prec w\}} = \sum_{v\in\mathbf{t}} (|\mathbf{t}_v| - 1) = \mathcal{A}_{\mathbf{t}}^*(1) - |\mathbf{t}|.$$

  - Wiener index: $W(\mathbf{t}) = \sum_{u,w\in\mathbf{t}} d(u, w)$. Use:

$$d(u, w) = \sum_{v\in\mathbf{t}} \mathbf{1}_{\{v\preccurlyeq u\}} + \mathbf{1}_{\{v\preccurlyeq w\}} - 2\mathbf{1}_{\{v\preccurlyeq u,\, v\preccurlyeq w\}}$$

  to get that:

$$W(\mathbf{t}) = 2|\mathbf{t}| \left(\mathcal{A}_{\mathbf{t}}^*(1) - \mathcal{A}_{\mathbf{t}}^*(x)\right).$$

  - Sackin index: $S(\mathbf{t}) = \sum_{w\in\mathcal{L}(\mathbf{t})} d(\emptyset, w)$. If $\mathbf{t}$ is full binary, then:

$$S(\mathbf{t}) = (\mathcal{A}_{\mathbf{t}}^*(1) - 1)/2.$$

Additive functionals and applications

- Unnormalized measure: $\forall f$ continuous defined on $[0, 1]$,

$$\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v \in \mathbf{t}} |\mathbf{t}_v| \; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right).$$

- Link with known functionals
  - Total path length: $P(\mathbf{t}) = \sum_{w \in \mathbf{t}} d(\emptyset, w)$. As $d(\emptyset, w) = \sum_{v \in \mathbf{t}} \mathbf{1}_{\{v \prec w\}}$, we get:

$$P(\mathbf{t}) = \sum_{v \in \mathbf{t}} \sum_{w \in \mathbf{t}} \mathbf{1}_{\{v \prec w\}} = \sum_{v \in \mathbf{t}} (|\mathbf{t}_v| - 1) = \mathcal{A}^*_{\mathbf{t}}(1) - |\mathbf{t}|.$$

  - Wiener index: $W(\mathbf{t}) = \sum_{u,w \in \mathbf{t}} d(u, w)$. Use:

$$d(u, w) = \sum_{v \in \mathbf{t}} \mathbf{1}_{\{v \preccurlyeq u\}} + \mathbf{1}_{\{v \preccurlyeq w\}} - 2\mathbf{1}_{\{v \preccurlyeq u,\, v \preccurlyeq w\}}$$

    to get that:

$$W(\mathbf{t}) = 2|\mathbf{t}| \left(\mathcal{A}^*_{\mathbf{t}}(1) - \mathcal{A}^*_{\mathbf{t}}(x)\right).$$

  - Sackin index: $S(\mathbf{t}) = \sum_{w \in \mathcal{L}(\mathbf{t})} d(\emptyset, w)$. If $\mathbf{t}$ is full binary, then:

$$S(\mathbf{t}) = (\mathcal{A}^*_{\mathbf{t}}(1) - 1)/2.$$

Additive functionals and applications

- Unnormalized measure: $\forall f$ continuous defined on $[0,1]$,

$$\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v \in \mathbf{t}} |\mathbf{t}_v| \; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right).$$

- Link with known functionals
  - Total path length: $P(\mathbf{t}) = \sum_{w \in \mathbf{t}} d(\emptyset, w)$. As $d(\emptyset, w) = \sum_{v \in \mathbf{t}} \mathbf{1}_{\{v \prec w\}}$, we get:

$$P(\mathbf{t}) = \sum_{v \in \mathbf{t}} \sum_{w \in \mathbf{t}} \mathbf{1}_{\{v \prec w\}} = \sum_{v \in \mathbf{t}} (|\mathbf{t}_v| - 1) = \mathcal{A}^*_{\mathbf{t}}(1) - |\mathbf{t}|.$$

  - Wiener index: $W(\mathbf{t}) = \sum_{u,w \in \mathbf{t}} d(u,w)$. Use:

$$d(u,w) = \sum_{v \in \mathbf{t}} \mathbf{1}_{\{v \preccurlyeq u\}} + \mathbf{1}_{\{v \preccurlyeq w\}} - 2\mathbf{1}_{\{v \preccurlyeq u,\, v \preccurlyeq w\}}$$

    to get that:

$$W(\mathbf{t}) = 2|\mathbf{t}| \left(\mathcal{A}^*_{\mathbf{t}}(1) - \mathcal{A}^*_{\mathbf{t}}(x)\right).$$

  - Sackin index: $S(\mathbf{t}) = \sum_{w \in \mathcal{L}(\mathbf{t})} d(\emptyset, w)$. If $\mathbf{t}$ is full binary, then:

$$S(\mathbf{t}) = (\mathcal{A}^*_{\mathbf{t}}(1) - 1)/2.$$

Additive functionals and applications

- Unnormalized measure: $\forall f$ continuous defined on $[0,1]$,

$$\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v| \; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right).$$

- Link with known functionals
  - Total path length: $P(\mathbf{t}) = \sum_{w\in\mathbf{t}} d(\emptyset, w)$. As $d(\emptyset, w) = \sum_{v\in\mathbf{t}} \mathbf{1}_{\{v\prec w\}}$, we get:

$$P(\mathbf{t}) = \sum_{v\in\mathbf{t}}\sum_{w\in\mathbf{t}} \mathbf{1}_{\{v\prec w\}} = \sum_{v\in\mathbf{t}}(|\mathbf{t}_v|-1) = \mathcal{A}^*_{\mathbf{t}}(1) - |\mathbf{t}|.$$

  - Wiener index: $W(\mathbf{t}) = \sum_{u,w\in\mathbf{t}} d(u,w)$. Use:

$$d(u,w) = \sum_{v\in\mathbf{t}} \mathbf{1}_{\{v\preccurlyeq u\}} + \mathbf{1}_{\{v\preccurlyeq w\}} - 2\mathbf{1}_{\{v\preccurlyeq u,\; v\preccurlyeq w\}}$$

    to get that:

$$W(\mathbf{t}) = 2|\mathbf{t}| \left(\mathcal{A}^*_{\mathbf{t}}(1) - \mathcal{A}^*_{\mathbf{t}}(x)\right).$$

  - Sackin index: $S(\mathbf{t}) = \sum_{w\in\mathcal{L}(\mathbf{t})} d(\emptyset, w)$. If $\mathbf{t}$ is full binary, then:

$$S(\mathbf{t}) = (\mathcal{A}^*_{\mathbf{t}}(1) - 1)/2.$$

Cost functionals for binary trees ($\beta > 1/2$)

- $\mathbf{t}$: full binary rooted ordered tree.
- Cost functional defined by recursive equation (“divide and conquer”):

$$F_\beta(\mathbf{t}) = F_\beta(\mathbf{t}_1) + F_\beta(\mathbf{t}_2) + b_\beta(|\mathbf{t}|),$$

with toll function $b_\beta(n) = n^\beta$.

- We have:

$$F_\beta(\mathbf{t}) = \sum_{v \in \mathbf{t}} |\mathbf{t}_v|^\beta = |\mathbf{t}|^{\beta-1} \mathcal{A}^*_{\mathbf{t}}(x^{\beta-1}).$$

- **Fill** and **Kapur** (2003) on Catalan model: T_n uniformly chosen among the full binary rooted ordered trees with n internal nodes (and $2n+1$ nodes):

$$|\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(x^{\beta-1}) = |\mathrm{T}_n|^{-\beta-1/2} F_\beta(\mathrm{T}_n) \xrightarrow[n\to+\infty]{(d)} 2Z_\beta.$$

  Fill and **Janson** (2007) announced a formula for Z_β using Brownian excursion. (Case $\beta = 1$, see **Aldous** (1993) or **Takàcs** (1994).)

Cost functionals for binary trees ($\beta > 1/2$)

- $\mathbf{t}$: full binary rooted ordered tree.
- Cost functional defined by recursive equation ("divide and conquer"):

$$F_\beta(\mathbf{t}) = F_\beta(\mathbf{t}_1) + F_\beta(\mathbf{t}_2) + b_\beta(|\mathbf{t}|),$$

with toll function $b_\beta(n) = n^\beta$.

- We have:

$$F_\beta(\mathbf{t}) = \sum_{v \in \mathbf{t}} |\mathbf{t}_v|^\beta = |\mathbf{t}|^{\beta-1} \mathcal{A}^*_{\mathbf{t}}(x^{\beta-1}).$$

- **Fill** and **Kapur** (2003) on Catalan model: T_n uniformly chosen among the full binary rooted ordered trees with n internal nodes (and $2n+1$ nodes):

$$|\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(x^{\beta-1}) = |\mathrm{T}_n|^{-\beta-1/2} F_\beta(\mathrm{T}_n) \xrightarrow[n\to+\infty]{(d)} 2Z_\beta.$$

Fill and **Janson** (2007) announced a formula for Z_β using Brownian excursion. (Case $\beta = 1$, see **Aldous** (1993) or **Takàcs** (1994).)

Cost functionals for binary trees ($\beta > 1/2$)

- $\mathbf{t}$: full binary rooted ordered tree.
- Cost functional defined by recursive equation ("divide and conquer"):

$$F_\beta(\mathbf{t}) = F_\beta(\mathbf{t}_1) + F_\beta(\mathbf{t}_2) + b_\beta(|\mathbf{t}|),$$

  with toll function $b_\beta(n) = n^\beta$.
- We have:

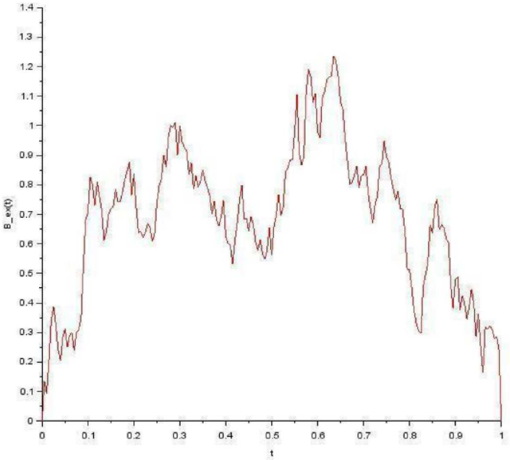
$$F_\beta(\mathbf{t}) = \sum_{v \in \mathbf{t}} |\mathbf{t}_v|^\beta = |\mathbf{t}|^{\beta-1} \mathcal{A}^*_{\mathbf{t}}(x^{\beta-1}).$$

- **Fill** and **Kapur** (2003) on Catalan model: T_n uniformly chosen among the full binary rooted ordered trees with n internal nodes (and $2n+1$ nodes):

$$|\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(x^{\beta-1}) = |\mathrm{T}_n|^{-\beta-1/2} F_\beta(\mathrm{T}_n) \xrightarrow[n \to +\infty]{\text{(d)}} 2Z_\beta.$$

  Fill and **Janson** (2007) announced a formula for Z_β using Brownian excursion. (Case $\beta = 1$, see **Aldous** (1993) or **Takàcs** (1994).)

Binary trees in the Brownian excursion (1)

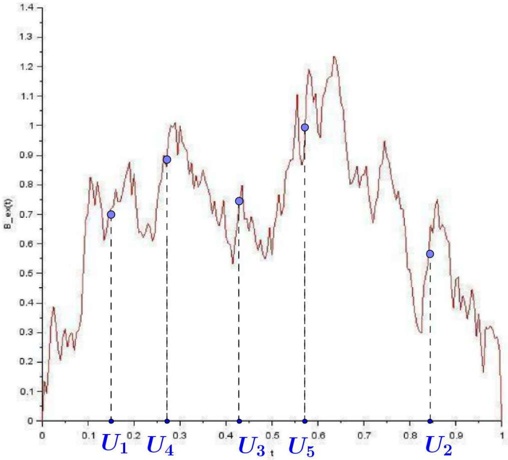


- Normalized Brownian excursion: B.
- $(U_i)_{1 \leq i \leq 5}$ indep. uniform r.v. on $[0, 1]$ and indep. of B.
- $(V_i)_{1 \leq i \leq 4}$ such that:

$$B(V_i) = \min_{u \in [U_{(i)}, U_{(i+1)}]} B(u).$$

with $U_{(1)} < \cdots < U_{(n+1)}$ order statistics.

Binary trees in the Brownian excursion (1)

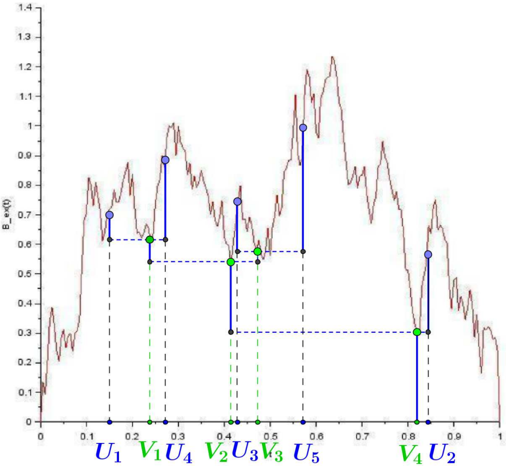


- Normalized Brownian excursion: B.
- $(U_i)_{1\leq i\leq 5}$ indep. uniform r.v. on $[0,1]$ and indep. of B.
- $(V_i)_{1\leq i\leq 4}$ such that:

$$B(V_i) = \min_{u\in[U_{(i)},U_{(i+1)}]} B(u).$$

with $U_{(1)} < \cdots < U_{(n+1)}$ order statistics.

Binary trees in the Brownian excursion (1)



- Normalized Brownian excursion: B.
- $(U_i)_{1\leq i\leq 5}$ indep. uniform r.v. on $[0, 1]$ and indep. of B.
- $(V_i)_{1\leq i\leq 4}$ such that:

  $$B(V_i) = \min_{u\in[U_{(i)},U_{(i+1)}]} B(u).$$

  with $U_{(1)} < \cdots < U_{(n+1)}$ order statistics.

Binary trees in the Brownian excursion (2)

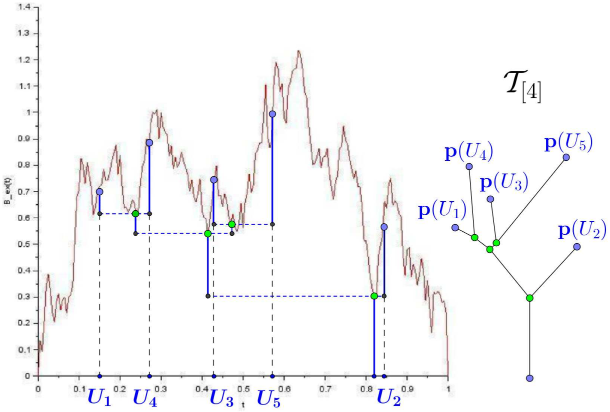


Figure: The Brownian excursion and continuous binary tree $\mathcal{T}_{[n]}$ (for $n+1=5$ leaves and $n=4$ internal nodes).

Binary trees in the Brownian excursion (2)

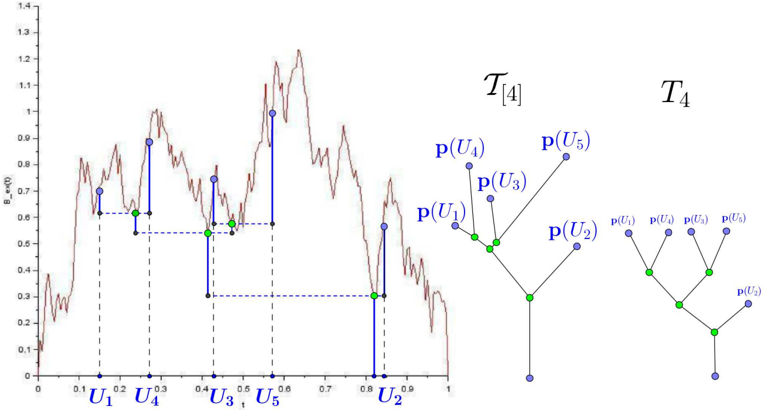


Figure: The Brownian excursion, $\mathcal{T}_{[n]}$ and T_n (for $n = 4$). $|\mathrm{T}_n|$ has n internal nodes and is uniformly chosen among the binary trees with n internal nodes.

Main result: convergence of $\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v| \; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right)$

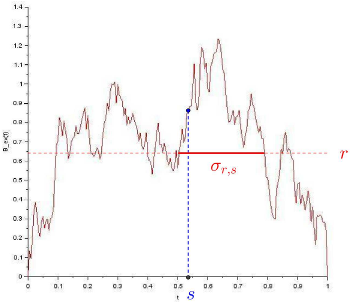


Figure: $\sigma_{r,s}$ = length of the excursion of B above level r straddling s

- $\sigma_{r,s} = \int_0^1 dt\, \mathbf{1}_{\{\min_{[s,t]} B \geq r\}}$
- Measure Φ_B: $\forall f$ continuous on $[0, 1]$

$$\Phi_B(f) = \int_0^1 ds \int_0^{B(s)} dr\, f(\sigma_{r,s}).$$

- **Main result**: A.s. $\forall f$ continuous:

$$|\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(f) \xrightarrow[n\to+\infty]{} 2\Phi_B(f).$$

- ($\Rightarrow$ joint convergence of the indexes.)

Main result: convergence of $\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v| \; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right)$

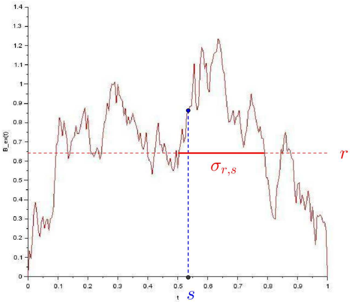


Figure: $\sigma_{r,s}$ = length of the excursion of B above level r straddling s

- $\sigma_{r,s} = \int_0^1 dt\, \mathbf{1}_{\{\min_{[s,t]} B \geq r\}}$
- Measure Φ_B: $\forall f$ continuous on $[0, 1]$

$$\Phi_B(f) = \int_0^1 ds \int_0^{B(s)} dr\, f(\sigma_{r,s}).$$

- **Main result**: A.s. $\forall f$ continuous:

$$|\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(f) \xrightarrow[n\to+\infty]{} 2\Phi_B(f).$$

- ($\Rightarrow$ joint convergence of the indexes.)

Main result: convergence of $\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v|\; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right)$

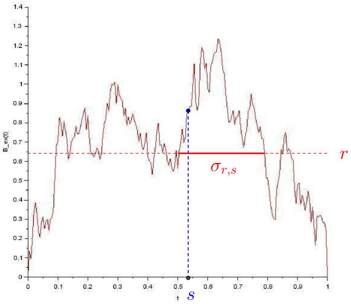


Figure: $\sigma_{r,s}$ = length of the excursion of B above level r straddling s

- $\sigma_{r,s} = \int_0^1 dt\, \mathbf{1}_{\{\min_{[s,t]} B \geq r\}}$
- Measure Φ_B: $\forall f$ continuous on $[0,1]$

$$\Phi_B(f) = \int_0^1 ds \int_0^{B(s)} dr\, f(\sigma_{r,s}).$$

- **Main result**: A.s. $\forall f$ continuous:

$$|\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(f) \xrightarrow[n\to+\infty]{} 2\Phi_B(f).$$

- ($\Rightarrow$ joint convergence of the indexes.)

Main result: convergence of $\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v| \; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right)$

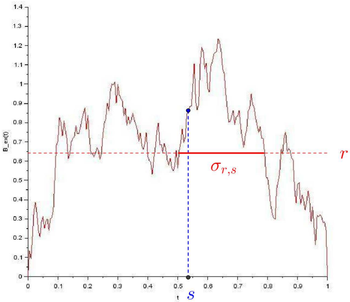


Figure: $\sigma_{r,s}$ = length of the excursion of B above level r straddling s

- $\sigma_{r,s} = \int_0^1 dt\, \mathbf{1}_{\{\min_{[s,t]} B \geq r\}}$
- Measure Φ_B: $\forall f$ continuous on $[0, 1]$

$$\Phi_B(f) = \int_0^1 ds \int_0^{B(s)} dr\, f(\sigma_{r,s}).$$

- **Main result**: A.s. $\forall f$ continuous:

$$|\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(f) \xrightarrow[n\to+\infty]{} 2\Phi_B(f).$$

- ($\Rightarrow$ joint convergence of the indexes.)

Other results for $\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v| \; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right)$

A.s. $\forall f$ continuous on $[0, 1]$:

$$|\mathrm{T}_n|^{-3/2}\mathcal{A}^*_{\mathrm{T}_n}(f) \xrightarrow[n\to+\infty]{} 2\Phi_B(f). \qquad (1)$$

- Eq. (1) is also true for $f(x) = x^{\beta-1}$ and $\beta > 1/2$ (“global limit” case).
- We have the corresponding fluctuations for all f Lipschitz and $f(x) = x^{\beta-1}$ with $\beta \geq 1$.
- If T_n is replaced by simply generated random trees τ_p with p nodes (with weight = a critical probability distribution in the domain of attraction of stable law), then there exists a sequence $(a_p)_{p\geq 1}$ and a random function H such that:

$$\frac{a_p}{|\tau_p|^2}\mathcal{A}^*_{\tau_p}(f) \xrightarrow[n\to+\infty]{(\mathrm{d})} \Phi_H(f).$$

The convergences hold jointly for all continuous function f. The process H is the so-called height process of a Lévy continuum tree. (Finite variance case: $a_p = \sigma\sqrt{p}/2$ and $H = B$.)

Other results for $\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v|\ f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right)$

A.s. $\forall f$ continuous on $[0,1]$:

$$|\mathrm{T}_n|^{-3/2}\mathcal{A}^*_{\mathrm{T}_n}(f) \xrightarrow[n\to+\infty]{} 2\Phi_B(f). \qquad (1)$$

- Eq. (1) is also true for $f(x) = x^{\beta-1}$ and $\beta > 1/2$ ("global limit" case).
- We have the corresponding fluctuations for all f Lipschitz and $f(x) = x^{\beta-1}$ with $\beta \geq 1$.
- If T_n is replaced by simply generated random trees τ_p with p nodes (with weight = a critical probability distribution in the domain of attraction of stable law), then there exists a sequence $(a_p)_{p\geq 1}$ and a random function H such that:

$$\frac{a_p}{|\tau_p|^2}\mathcal{A}^*_{\tau_p}(f) \xrightarrow[n\to+\infty]{(d)} \Phi_H(f).$$

  The convergences hold jointly for all continuous function f. The process H is the so-called height process of a Lévy continuum tree. (Finite variance case: $a_p = \sigma\sqrt{p}/2$ and $H = B$.)

Other results for $\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v| \, f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right)$

A.s. $\forall f$ continuous on $[0,1]$:

$$|\mathrm{T}_n|^{-3/2}\mathcal{A}^*_{\mathrm{T}_n}(f) \xrightarrow[n\to+\infty]{} 2\Phi_B(f). \quad (1)$$

- Eq. (1) is also true for $f(x) = x^{\beta-1}$ and $\beta > 1/2$ ("global limit" case).
- We have the corresponding fluctuations for all f Lipschitz and $f(x) = x^{\beta-1}$ with $\beta \geq 1$.
- If T_n is replaced by simply generated random trees τ_p with p nodes (with weight = a critical probability distribution in the domain of attraction of stable law), then there exists a sequence $(a_p)_{p\geq 1}$ and a random function H such that:

$$\frac{a_p}{|\tau_p|^2}\mathcal{A}^*_{\tau_p}(f) \xrightarrow[n\to+\infty]{(d)} \Phi_H(f).$$

  The convergences hold jointly for all continuous function f. The process H is the so-called height process of a Lévy continuum tree. (Finite variance case: $a_p = \sigma\sqrt{p}/2$ and $H = B$.)

Elementary proof of (1)

- Leaves of $\mathbf{t}$: $\mathcal{L}(\mathbf{t})$. Recall $|\mathbf{t}| = 2|\mathcal{L}(\mathbf{t})| - 1$.

$$A_n := |\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(f) = |\mathrm{T}_n|^{-3/2} \sum_{v \in \mathbf{t}} |\mathrm{T}_{n,v}| \; f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right)$$

$$\simeq 2|\mathrm{T}_n|^{-3/2} \sum_{v \in \mathbf{t}} |\mathcal{L}(\mathrm{T}_{n,v})| \; f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right)$$

$$= 2|\mathrm{T}_n|^{-3/2} \sum_{u \in \mathcal{L}(\mathrm{T}_n)} \sum_{v \preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) \mathbf{1}$$

$$\simeq 2|\mathrm{T}_n|^{-3/2} \sum_{u \in \mathcal{L}(\mathrm{T}_n)} \sum_{v \preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) 2\sqrt{2n} \; h_v,$$

with h_v the length of the branch below v in $\mathcal{T}_{[n]}$ (with mean $\simeq 1/2\sqrt{2n}$).

Elementary proof of (1)

- Leaves of $\mathbf{t}$: $\mathcal{L}(\mathbf{t})$. Recall $|\mathbf{t}| = 2|\mathcal{L}(\mathbf{t})| - 1$.

$$
\begin{aligned}
A_n := |\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(f) &= |\mathrm{T}_n|^{-3/2} \sum_{v \in \mathbf{t}} |\mathrm{T}_{n,v}| \; f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|} \right) \\
&\simeq 2|\mathrm{T}_n|^{-3/2} \sum_{v \in \mathbf{t}} |\mathcal{L}(\mathrm{T}_{n,v})| \; f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|} \right) \\
&= 2|\mathrm{T}_n|^{-3/2} \sum_{u \in \mathcal{L}(\mathrm{T}_n)} \sum_{v \preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|} \right) \mathbf{1} \\
&\simeq 2|\mathrm{T}_n|^{-3/2} \sum_{u \in \mathcal{L}(\mathrm{T}_n)} \sum_{v \preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|} \right) 2\sqrt{2n}\; h_v,
\end{aligned}
$$

with h_v the length of the branch below v in $\mathcal{T}_{[n]}$ (with mean $\simeq 1/2\sqrt{2n}$).

Elementary proof of (1)

- Leaves of $\mathbf{t}$: $\mathcal{L}(\mathbf{t})$. Recall $|\mathbf{t}| = 2|\mathcal{L}(\mathbf{t})| - 1$.

$$
\begin{aligned}
A_n := |\mathrm{T}_n|^{-3/2}\mathcal{A}^*_{\mathrm{T}_n}(f) &= |\mathrm{T}_n|^{-3/2}\sum_{v\in\mathbf{t}} |\mathrm{T}_{n,v}|\; f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right)\\
&\simeq 2|\mathrm{T}_n|^{-3/2}\sum_{v\in\mathbf{t}} |\mathcal{L}(\mathrm{T}_{n,v})|\; f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right)\\
&= 2|\mathrm{T}_n|^{-3/2}\sum_{u\in\mathcal{L}(\mathrm{T}_n)}\sum_{v\preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) \mathbf{1}\\
&\simeq 2|\mathrm{T}_n|^{-3/2}\sum_{u\in\mathcal{L}(\mathrm{T}_n)}\sum_{v\preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) 2\sqrt{2n}\; h_v,
\end{aligned}
$$

with h_v the length of the branch below v in $\mathcal{T}_{[n]}$ (with mean $\simeq 1/2\sqrt{2n}$).

Elementary proof of (1)

- Leaves of $\mathbf{t}$: $\mathcal{L}(\mathbf{t})$. Recall $|\mathbf{t}| = 2|\mathcal{L}(\mathbf{t})| - 1$.

$$
\begin{aligned}
A_n := |\mathrm{T}_n|^{-3/2}\mathcal{A}^*_{\mathrm{T}_n}(f) &= |\mathrm{T}_n|^{-3/2}\sum_{v\in\mathbf{t}} |\mathrm{T}_{n,v}|\ f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right)\\
&\simeq 2|\mathrm{T}_n|^{-3/2}\sum_{v\in\mathbf{t}} |\mathcal{L}(\mathrm{T}_{n,v})|\ f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right)\\
&= 2|\mathrm{T}_n|^{-3/2}\sum_{u\in\mathcal{L}(\mathrm{T}_n)}\sum_{v\preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right)\mathbf{1}\\
&\simeq 2|\mathrm{T}_n|^{-3/2}\sum_{u\in\mathcal{L}(\mathrm{T}_n)}\sum_{v\preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) 2\sqrt{2n}\ h_v,
\end{aligned}
$$

with h_v the length of the branch below v in $\mathcal{T}_{[n]}$ (with mean $\simeq 1/2\sqrt{2n}$).

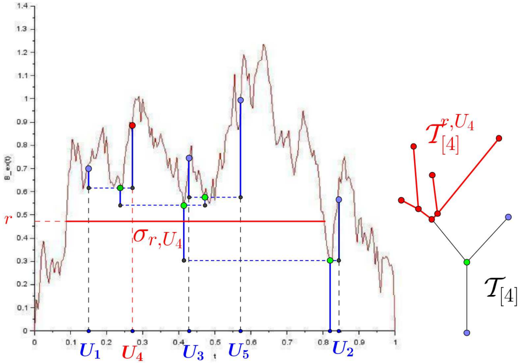
Elementary proof of (1)

- Leaves of $\mathbf{t}$: $\mathcal{L}(\mathbf{t})$. Recall $|\mathbf{t}| = 2|\mathcal{L}(\mathbf{t})| - 1$.

$$\begin{aligned} A_n := |\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(f) &= |\mathrm{T}_n|^{-3/2} \sum_{v \in \mathbf{t}} |\mathrm{T}_{n,v}| \; f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) \\ &\simeq 2|\mathrm{T}_n|^{-3/2} \sum_{v \in \mathbf{t}} |\mathcal{L}(\mathrm{T}_{n,v})| \; f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) \\ &= 2|\mathrm{T}_n|^{-3/2} \sum_{u \in \mathcal{L}(\mathrm{T}_n)} \sum_{v \preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) \mathbf{1} \\ &\simeq 2|\mathrm{T}_n|^{-3/2} \sum_{u \in \mathcal{L}(\mathrm{T}_n)} \sum_{v \preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) 2\sqrt{2n}\; h_v, \end{aligned}$$

with h_v the length of the branch below v in $\mathcal{T}_{[n]}$ (with mean $\simeq 1/2\sqrt{2n}$).

Elementary proof of (1)



$$A_n \simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{u\in\mathcal{L}(\mathrm{T}_n)} \sum_{v\preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) h_v$$

$$= \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr\, f\left(\frac{|\mathcal{T}_{[n]}^{r,U_k}|}{|\mathrm{T}_n|}\right)$$

$$\simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr\, f\left(\sigma_{r,U_k}\right),$$

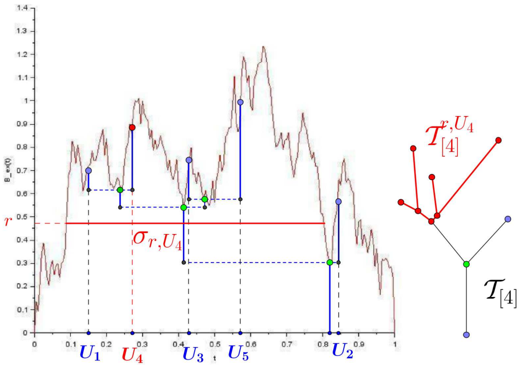
with $\mathcal{T}_{[n]}^{r,U_k}$ the subtree above level r containing U_k, and we use $|\mathcal{T}_{[n]}^{r,U_k}| \sim 2\mathcal{B}(n,\sigma_{r,U_k})+1$, that is

$$|\mathcal{T}_{[n]}^{r,U_k}|/|\mathrm{T}_n| \simeq \sigma_{r,U_k}.$$

Using the SLLN (cond. on B), we get:

$$A_n \xrightarrow[n\to+\infty]{\text{a.s.}} 2\int_0^1 ds \int_0^{B(s)} dr\, f(\sigma_{r,s}).$$

Elementary proof of (1)



$$A_n \simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{u\in\mathcal{L}(\mathrm{T}_n)} \sum_{v\preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) h_v$$

$$= \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr f\left(\frac{|\mathcal{T}_{[n]}^{r,U_k}|}{|\mathrm{T}_n|}\right)$$

$$\simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr f\left(\sigma_{r,U_k}\right),$$

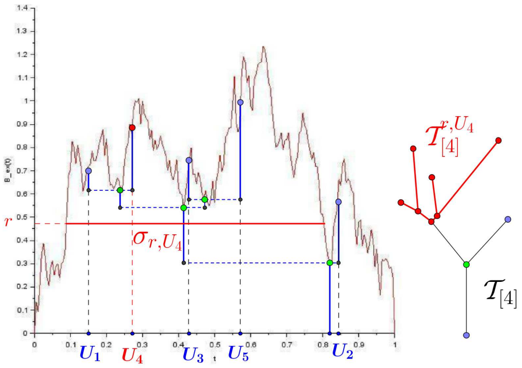
with $\mathcal{T}_{[n]}^{r,U_k}$ the subtree above level r containing U_k, and we use $|\mathcal{T}_{[n]}^{r,U_k}| \sim 2\mathcal{B}(n, \sigma_{r,U_k}) + 1$, that is

$$|\mathcal{T}_{[n]}^{r,U_k}|/|\mathrm{T}_n| \simeq \sigma_{r,U_k}.$$

Using the SLLN (cond. on B), we get:

$$A_n \xrightarrow[n\to+\infty]{\text{a.s.}} 2\int_0^1 ds \int_0^{B(s)} dr f(\sigma_{r,s}).$$

Elementary proof of (1)



$$A_n \simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{u \in \mathcal{L}(\mathrm{T}_n)} \sum_{v \preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) h_v$$

$$= \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr\, f\left(\frac{|\mathcal{T}^{r,U_k}_{[n]}|}{|\mathrm{T}_n|}\right)$$

$$\simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr\, f\left(\sigma_{r,U_k}\right),$$

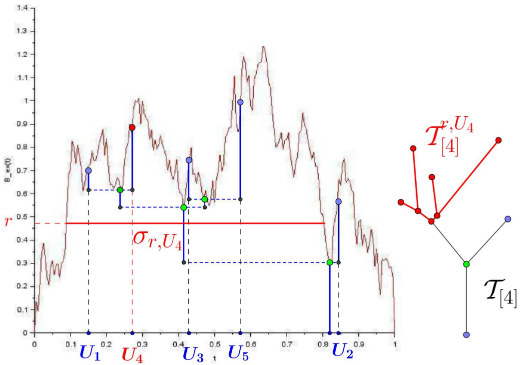
with $\mathcal{T}^{r,U_k}_{[n]}$ the subtree above level r containing U_k, and we use $|\mathcal{T}^{r,U_k}_{[n]}| \sim 2\mathcal{B}(n, \sigma_{r,U_k}) + 1$, that is

$$|\mathcal{T}^{r,U_k}_{[n]}|/|\mathrm{T}_n| \simeq \sigma_{r,U_k}.$$

Using the SLLN (cond. on B), we get:

$$A_n \xrightarrow[n\to+\infty]{\text{a.s.}} 2\int_0^1 ds \int_0^{B(s)} dr\, f(\sigma_{r,s}).$$

Elementary proof of (1)



$$A_n \simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{u\in\mathcal{L}(\mathrm{T}_n)} \sum_{v\preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) h_v$$

$$= \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr f\left(\frac{|\mathcal{T}_{[n]}^{r,U_k}|}{|\mathrm{T}_n|}\right)$$

$$\simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr f\left(\sigma_{r,U_k}\right),$$

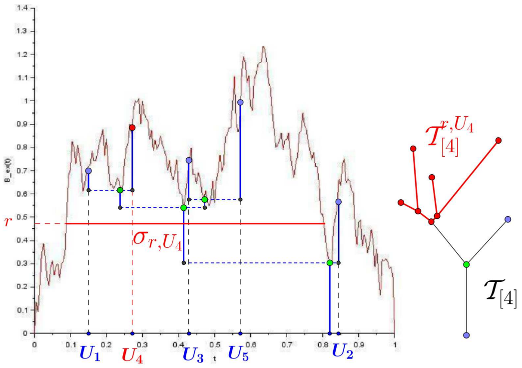
with $\mathcal{T}_{[n]}^{r,U_k}$ the subtree above level r containing U_k, and we use $|\mathcal{T}_{[n]}^{r,U_k}| \sim 2\mathcal{B}(n, \sigma_{r,U_k}) + 1$, that is

$$|\mathcal{T}_{[n]}^{r,U_k}|/|\mathrm{T}_n| \simeq \sigma_{r,U_k}.$$

Using the SLLN (cond. on B), we get:

$$A_n \xrightarrow[n\to+\infty]{\text{a.s.}} 2\int_0^1 ds \int_0^{B(s)} dr f(\sigma_{r,s}).$$

Elementary proof of (1)



$$A_n \simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{u\in\mathcal{L}(\mathrm{T}_n)} \sum_{v\preccurlyeq u} f\left(\frac{|\mathrm{T}_{n,v}|}{|\mathrm{T}_n|}\right) h_v$$

$$= \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr f\left(\frac{|\mathcal{T}_{[n]}^{r,U_k}|}{|\mathrm{T}_n|}\right)$$

$$\simeq \frac{4\sqrt{2n}}{|\mathrm{T}_n|^{3/2}} \sum_{k=1}^{n+1} \int_0^{B(U_k)} dr f\left(\sigma_{r,U_k}\right),$$

with $\mathcal{T}_{[n]}^{r,U_k}$ the subtree above level r containing U_k, and we use $|\mathcal{T}_{[n]}^{r,U_k}| \sim 2\mathcal{B}(n, \sigma_{r,U_k}) + 1$, that is

$$|\mathcal{T}_{[n]}^{r,U_k}|/|\mathrm{T}_n| \simeq \sigma_{r,U_k}.$$

Using the SLLN (cond. on B), we get:

$$A_n \xrightarrow[n\to+\infty]{\text{a.s.}} 2\int_0^1 ds \int_0^{B(s)} dr f(\sigma_{r,s}).$$

Conclusion: convergence of $\mathcal{A}^*_{\mathbf{t}}(f) = \sum_{v\in\mathbf{t}} |\mathbf{t}_v| \; f\left(\frac{|\mathbf{t}_v|}{|\mathbf{t}|}\right)$

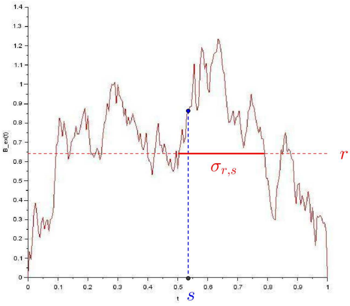


Figure: $\sigma_{r,s}$ = length of the excursion of B above level r straddling s

- $\sigma_{r,s} = \int_0^1 dt\, \mathbf{1}_{\{\min_{[s,t]} B \geq r\}}$
- Measure Φ_B: $\forall f$ continuous on $[0, 1]$

$$\Phi_B(f) = \int_0^1 ds \int_0^{B(s)} dr\, f(\sigma_{r,s}).$$

- **Main result**: A.s. $\forall f$ continuous on $(0, 1]$ s.t. there exists $a > 1/2$ and $\lim_{x\downarrow 0} x^a f(x) = 0$, we have:

$$\boxed{|\mathrm{T}_n|^{-3/2} \mathcal{A}^*_{\mathrm{T}_n}(f) \xrightarrow[n\to+\infty]{} 2\Phi_B(f)}$$

with T_n unif. among the full binary rooted ordered trees with n internal nodes.