

The sum of powers of subtrees sizes for random trees

Svante Janson

70 Years of Percolation
and Geoffrey Grimmett
Cambridge, July, 2023

References

This talk is mainly based on joint work with

Jim Fill: *Electronic Journal of Probability* 27 (2022);

Jim Fill and Stephan Wagner: [arXiv:2212.10871](https://arxiv.org/abs/2212.10871).

General problem

Additive functional: Let $f(T)$ (**the toll function**) be a given functional of rooted trees, and define

$$F(T) := \sum_{v \in T} f(T_v),$$

where T_v is the fringe tree rooted at v , i.e. the subtree consisting of v and all its descendants.

Problem: Study asymptotics of $F(\mathcal{T}_n)$ (mean, variance, distribution, ...) when \mathcal{T}_n is some random tree of “size” n , and $n \rightarrow \infty$.

Today, the random tree \mathcal{T}_n will be a conditioned Galton–Watson tree with $|\mathcal{T}_n| = n$ (the number of vertices); the offspring distribution ξ will be critical with finite variance $0 < \sigma^2 < \infty$. (Higher moments assumed only occasionally.)

The toll function will be simply

$$f_\alpha(T) := |T|^\alpha$$

for a constant α .

Examples.

$\alpha = 0$ is trivial: $F_0(T) = |T|$.

(The derivative at 0 is the “shape functional”. No time today.)

$\alpha = 1$ gives $F_1(T) =$ the total pathlength.

We allow α to be complex, and we consider $F_\alpha(T)$ as a function of $\alpha \in \mathbb{C}$. We write

$$X_n(\alpha) := F_\alpha(\mathcal{T}_n) = \sum_{v \in \mathcal{T}_n} |(\mathcal{T}_n)_v|^\alpha$$

$$\tilde{X}_n(\alpha) := X_n(\alpha) - \mathbb{E} X_n(\alpha).$$

Remark

Why complex α ?

- ▶ Useful in proofs (also for real α) since powerful methods of analytic functions can be used.

Remark

Why complex α ?

- ▶ Useful in proofs (also for real α) since powerful methods of analytic functions can be used.
- ▶ Gives us new problems to study. How do the phase transitions look in the complex plane?

There are two phase transitions for real α : $\alpha = 0$ and $\alpha = \frac{1}{2}$.

Thus three phases in the complex plane:

$$\operatorname{Re}(\alpha) < 0, \quad 0 < \operatorname{Re}(\alpha) < \frac{1}{2}, \quad \operatorname{Re}(\alpha) > \frac{1}{2}.$$

There are two phase transitions for real α : $\alpha = 0$ and $\alpha = \frac{1}{2}$.

Thus three phases in the complex plane:

$$\operatorname{Re}(\alpha) < 0, \quad 0 < \operatorname{Re}(\alpha) < \frac{1}{2}, \quad \operatorname{Re}(\alpha) > \frac{1}{2}.$$

What happens at the boundaries $\operatorname{Re}(\alpha) = 0$ and $\operatorname{Re}(\alpha) = \frac{1}{2}$?

Let $\mathcal{T}_{n,*}$ be a random fringe tree, i.e. $(\mathcal{T}_n)_v$ for a random vertex $v \in \mathcal{T}_n$. Then

$$\mathbb{E} X_n(\alpha) = \sum_{k=0}^{\infty} k^\alpha n \mathbb{P}(|\mathcal{T}_{n,*}| = k).$$

Let \mathcal{T} be an (unconditioned) Galton–Watson tree with the given offspring distribution. Then $\mathcal{T}_{n,*}$ has asymptotically the distribution of \mathcal{T} (Aldous, 1991). Recall that

$$\mathbb{P}(|\mathcal{T}| = k) \sim ck^{-3/2}.$$

Consequently, the number of fringe trees of size k in \mathcal{T}_n is $\approx cnk^{-3/2}$.

Hence, $\mathbb{E} X_n(\alpha)$ is dominated by small fringe trees for $\operatorname{Re} \alpha < 1/2$, and by large fringe trees for $\operatorname{Re} \alpha > 1/2$.

Let $\mathcal{T}_{n,*}$ be a random fringe tree, i.e. $(\mathcal{T}_n)_v$ for a random vertex $v \in \mathcal{T}_n$. Then

$$\mathbb{E} X_n(\alpha) = \sum_{k=0}^{\infty} k^\alpha n \mathbb{P}(|\mathcal{T}_{n,*}| = k).$$

Let \mathcal{T} be an (unconditioned) Galton–Watson tree with the given offspring distribution. Then $\mathcal{T}_{n,*}$ has asymptotically the distribution of \mathcal{T} (Aldous, 1991). Recall that

$$\mathbb{P}(|\mathcal{T}| = k) \sim ck^{-3/2}.$$

Consequently, the number of fringe trees of size k in \mathcal{T}_n is $\approx cnk^{-3/2}$.

Hence, $\mathbb{E} X_n(\alpha)$ is dominated by small fringe trees for $\operatorname{Re} \alpha < 1/2$, and by large fringe trees for $\operatorname{Re} \alpha > 1/2$.

Similarly, the variance, and distribution, are dominated by small fringe trees for $\operatorname{Re} \alpha < 0$, and by large fringe trees for $\operatorname{Re} \alpha > 0$.

Let

$$\mu(\alpha) := \mathbb{E} |\mathcal{T}|^\alpha = \sum_{k=1}^{\infty} k^\alpha \mathbb{P}(|\mathcal{T}| = k).$$

This converges for $\operatorname{Re}(\alpha) < \frac{1}{2}$, and defines an analytic function in this half-plane.

However,

$$\mu(\alpha) \rightarrow \infty \quad \text{as } \alpha \nearrow \frac{1}{2}.$$

Theorem

(i). If $\operatorname{Re}(\alpha) < \frac{1}{2}$, then

$$\mathbb{E} X_n(\alpha) = \mu(\alpha)n + o(n)$$

(ii). If $\operatorname{Re}(\alpha) > \frac{1}{2}$, then

$$\mathbb{E} X_n(\alpha) = \frac{1}{\sqrt{2}\sigma} \frac{\Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} n^{\alpha + \frac{1}{2}} + o(n^{\alpha + \frac{1}{2}})$$

(iii). If $\alpha = \frac{1}{2}$, then

$$\mathbb{E} X_n(1/2) = \frac{1}{\sqrt{2\pi}\sigma^2} n \log n + o(n \log n).$$

Critical line $\operatorname{Re}(\alpha) = \frac{1}{2}$

Recall that $\mu(\alpha) \rightarrow \infty$ as $\alpha \nearrow \frac{1}{2}$.

Theorem

The function $\mu(\alpha)$ has a continuous extension to $\operatorname{Re}(\alpha) = \frac{1}{2}$, $\alpha \neq \frac{1}{2}$.

Theorem

If $-\frac{1}{2} < \operatorname{Re}(\alpha) \leq \frac{1}{2}$ and $\alpha \neq \frac{1}{2}$, then

$$\mathbb{E} X_n(\alpha) = \mu(\alpha)n + \frac{1}{\sqrt{2\sigma}} \frac{\Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} n^{\alpha + \frac{1}{2}} + o(n^{(\operatorname{Re} \alpha)_+ + \frac{1}{2}}).$$

Let ξ be the offspring distribution.

Theorem

Suppose that $\mathbb{E} \xi^{2+\delta} < \infty$ where $0 < \delta \leq 1$.

- (i). Then $\mu(\alpha)$ can be analytically continued to a meromorphic function in $\operatorname{Re}(\alpha) < \frac{1}{2} + \frac{\delta}{2}$ with a simple pole at $\frac{1}{2}$.
- (ii). Moreover, the estimate above

$$\mathbb{E} X_n(\alpha) = \mu(\alpha)n + \frac{1}{\sqrt{2}\sigma} \frac{\Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} n^{\alpha + \frac{1}{2}} + o(n^{(\operatorname{Re} \alpha)_+ + \frac{1}{2}})$$

holds for $-\frac{1}{2} < \operatorname{Re}(\alpha) < \frac{1}{2} + \frac{\delta}{2}$ with $\alpha \neq \frac{1}{2}$.

Let ξ be the offspring distribution.

Theorem

Suppose that $\mathbb{E} \xi^{2+\delta} < \infty$ where $0 < \delta \leq 1$.

- (i). Then $\mu(\alpha)$ can be analytically continued to a meromorphic function in $\operatorname{Re}(\alpha) < \frac{1}{2} + \frac{\delta}{2}$ with a simple pole at $\frac{1}{2}$.
- (ii). Moreover, the estimate above

$$\mathbb{E} X_n(\alpha) = \mu(\alpha)n + \frac{1}{\sqrt{2}\sigma} \frac{\Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} n^{\alpha + \frac{1}{2}} + o(n^{(\operatorname{Re} \alpha) + \frac{1}{2}})$$

holds for $-\frac{1}{2} < \operatorname{Re}(\alpha) < \frac{1}{2} + \frac{\delta}{2}$ with $\alpha \neq \frac{1}{2}$.

With more moments, $\mu(\alpha)$ can be extended further. In particular, if $\mathbb{E} \xi^r < \infty$ for all $r > 0$, then $\mu(\alpha)$ is meromorphic in \mathbb{C} , with (simple) poles only at $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$

The additional moment assumption is really needed here.

Theorem

There exists ξ with $\mathbb{E} \xi = 1$ and $\mathbb{E} \xi^2 < \infty$ such that $\mu(\alpha)$ cannot be extended analytically across $\operatorname{Re}(\alpha) = \frac{1}{2}$ (at any point).

Asymptotic distribution

Recall that $X_n(\alpha) := \tilde{X}_n(\alpha) + \mathbb{E} X_n(\alpha)$.

Hence it suffices to consider $\tilde{X}_n(\alpha)$ and then combine with the results above for $\mathbb{E} X_n(\alpha)$.

$\operatorname{Re}(\alpha) < 0$

Let $H_- := \{\alpha : \operatorname{Re}(\alpha) < 0\}$.

Theorem

- ▶ *There exists a random analytic function $\tilde{X}(\alpha)$, $\alpha \in H_-$, such that, as $n \rightarrow \infty$,*

$$n^{-1/2} \tilde{X}_n(\alpha) \xrightarrow{d} \tilde{X}(\alpha)$$

for each fixed $\alpha \in H_-$, and uniformly on each compact subset of H_- . (I.e., in the space $\mathcal{H}(H_-)$ of analytic functions on H_- .)

- ▶ *$\tilde{X}(\alpha)$ is a complex centred Gaussian, for every fixed $\alpha \in H_-$. Also jointly.*
- ▶ *The covariance matrix of $\tilde{X}(\alpha)$ depends on the offspring distribution.*

$\operatorname{Re}(\alpha) < 0$

Let $H_- := \{\alpha : \operatorname{Re}(\alpha) < 0\}$.

Theorem

- ▶ *There exists a random analytic function $\tilde{X}(\alpha)$, $\alpha \in H_-$, such that, as $n \rightarrow \infty$,*

$$n^{-1/2} \tilde{X}_n(\alpha) \xrightarrow{d} \tilde{X}(\alpha)$$

for each fixed $\alpha \in H_-$, and uniformly on each compact subset of H_- . (I.e., in the space $\mathcal{H}(H_-)$ of analytic functions on H_- .)

- ▶ *$\tilde{X}(\alpha)$ is a complex centred Gaussian, for every fixed $\alpha \in H_-$. Also jointly.*
- ▶ *The covariance matrix of $\tilde{X}(\alpha)$ depends on the offspring distribution.*

In this case $X_n(\alpha) = F_\alpha(\mathcal{T}_n)$ is dominated by the many small fringe trees. Hence normality, but not universality.

$\operatorname{Re}(\alpha) > 0$

Let $H_+ := \{\alpha : \operatorname{Re}(\alpha) > 0\}$. (No problems for $\operatorname{Re}(\alpha) = \frac{1}{2}$.)

Theorem

- ▶ *There exists a random analytic function $\tilde{Y}(\alpha)$, $\alpha \in H_+$, such that, as $n \rightarrow \infty$,*

$$\tilde{Y}_n(\alpha) := n^{-\alpha - \frac{1}{2}} \tilde{X}_n(\alpha) \xrightarrow{d} \sigma^{-1} \tilde{Y}(\alpha)$$

for each fixed $\alpha \in H_+$, and uniformly on each compact subset of H_+ . (I.e., in the space $\mathcal{H}(H_+)$ of analytic functions on H_+ .)

- ▶ $\tilde{Y}(\alpha)$ is not Gaussian.
- ▶ $\tilde{Y}(\alpha)$ does not depend on the offspring distribution.

$\operatorname{Re}(\alpha) > 0$

Let $H_+ := \{\alpha : \operatorname{Re}(\alpha) > 0\}$. (No problems for $\operatorname{Re}(\alpha) = \frac{1}{2}$.)

Theorem

- ▶ *There exists a random analytic function $\tilde{Y}(\alpha)$, $\alpha \in H_+$, such that, as $n \rightarrow \infty$,*

$$\tilde{Y}_n(\alpha) := n^{-\alpha - \frac{1}{2}} \tilde{X}_n(\alpha) \xrightarrow{d} \sigma^{-1} \tilde{Y}(\alpha)$$

for each fixed $\alpha \in H_+$, and uniformly on each compact subset of H_+ . (I.e., in the space $\mathcal{H}(H_+)$ of analytic functions on H_+ .)

- ▶ $\tilde{Y}(\alpha)$ is not Gaussian.
- ▶ $\tilde{Y}(\alpha)$ does not depend on the offspring distribution.

In this case $\tilde{X}_n(\alpha) = F_\alpha(\mathcal{T}_n) - \mathbb{E} F_\alpha(\mathcal{T}_n)$ is dominated by the large fringe trees. Therefore universality but not normality.

Critical line $\operatorname{Re}(\alpha) = 0$

Theorem

Assume $\mathbb{E} \xi^{2+\delta} < \infty$ for some $\delta > 0$. (Conjecture: not needed.)

- ▶ For every real $t \neq 0$, as $n \rightarrow \infty$,

$$\frac{\tilde{X}_n(it)}{\sqrt{n \log n}} \xrightarrow{d} \sigma^{-1} \tilde{Z}(it),$$

where $\tilde{Z}(it)$ is a symmetric complex normal variable with variance

$$\mathbb{E} |\tilde{Z}(it)|^2 = \frac{1}{\sqrt{\pi}} \operatorname{Re} \frac{\Gamma(it - \frac{1}{2})}{\Gamma(it)} > 0. \quad (1)$$

- ▶ $\tilde{Z}(it)$ thus does not depend on the offspring distribution.
- ▶ The convergence holds jointly for any finite number of t , with independent limits $\tilde{Z}(it)$ for all $t > 0$.
- ▶ Thus no convergence to a continuous random function on $i\mathbb{R}$.

Without centring

Let, for $\operatorname{Re}(\alpha) > 0$ and $\alpha \neq \frac{1}{2}$,

$$Y(\alpha) := \tilde{Y}(\alpha) + \frac{1}{\sqrt{2}\sigma} \frac{\Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)}.$$

Theorem

(i). If $\operatorname{Re}(\alpha) > \frac{1}{2}$, then

$$Y_n(\alpha) := n^{-\alpha - \frac{1}{2}} X_n(\alpha) \xrightarrow{d} \sigma^{-1} Y(\alpha).$$

(ii). If $0 < \operatorname{Re}(\alpha) \leq \frac{1}{2}$ and $\alpha \neq \frac{1}{2}$, then

$$n^{-\alpha - \frac{1}{2}} [X_n(\alpha) - n\mu(\alpha)] \xrightarrow{d} \sigma^{-1} Y(\alpha).$$

Moment convergence

Theorem

All moments converge in the limit theorems above. If $\operatorname{Re}(\alpha) > 0$ and $\alpha \neq \frac{1}{2}$, then the limiting moments $\kappa_\ell := \mathbb{E} Y(\alpha)^\ell$ satisfy the recursion

$$\kappa_1 = \frac{\Gamma(\alpha - \frac{1}{2})}{\sqrt{2}\Gamma(\alpha)},$$

and, for $\ell \geq 2$, with $\alpha' := \alpha + \frac{1}{2}$,

$$\begin{aligned} \kappa_\ell &= \frac{\ell\Gamma(\ell\alpha' - 1)}{\sqrt{2}\Gamma(\ell\alpha' - \frac{1}{2})}\kappa_{\ell-1} \\ &\quad + \frac{1}{4\sqrt{\pi}} \sum_{j=1}^{\ell-1} \binom{\ell}{j} \frac{\Gamma(j\alpha' - \frac{1}{2})\Gamma((\ell-j)\alpha' - \frac{1}{2})}{\Gamma(\ell\alpha' - \frac{1}{2})} \kappa_j \kappa_{\ell-j}. \end{aligned}$$

Proofs by singularity analysis of generating functions, using properties of Hadamard products.

Disclaimer. For $\alpha = \frac{1}{2}$, our proof requires that the offspring distribution ξ satisfies $\mathbb{E} \xi^{2+\delta} < \infty$ for some $\delta > 0$.

Brownian excursion, $\operatorname{Re} \alpha > 1$

Let \mathbf{e} be a standard Brownian excursion. Recall that this is a random continuous function $[0, 1] \rightarrow [0, \infty)$.

For a function g and $s < t$, define

$$m(g; s, t) := \inf_{u \in [s, t]} g(u).$$

Theorem

If $\operatorname{Re} \alpha > 1$, we can represent the limit $Y(\alpha)$ as

$$Y(\alpha) = 2\alpha(\alpha - 1) \iint_{0 < s < t < 1} (t - s)^{\alpha - 2} m(\mathbf{e}; s, t) ds dt.$$

Proof. If we replace \mathbf{e} by a suitably scaled version of the contour process of \mathcal{T}_n , then a calculation shows that the integral equals $n^{-\alpha - \frac{1}{2}} X_n(\alpha) + o(1)$. The contour process converges to \mathbf{e} (Aldous, 1993), and the integral is a continuous functional.

Brownian excursion, $\operatorname{Re} \alpha > 1/2$

Theorem

If $\operatorname{Re} \alpha > 1/2$, we can represent the limit $Y(\alpha)$ as

$$Y(\alpha) = 2\alpha \int_0^1 t^{\alpha-1} \mathbf{e}(t) dt \\ - 2\alpha(\alpha - 1) \iint_{0 < s < t < 1} (t - s)^{\alpha-2} [\mathbf{e}(t) - m(\mathbf{e}; s, t)] ds dt.$$

Example. $\alpha = 1$ (total pathlength) yields

$$Y(1) = 2 \int_0^1 \mathbf{e}(t) dt,$$

the *Brownian excursion area*. This case was proved by Aldous (1993).

Proof: Tightness

Lemma

- (i). If $\operatorname{Re} \alpha < 0$, then $\mathbb{E} |\tilde{X}_n(\alpha)|^2 \leq C(\alpha)n$.
- (ii). If $\operatorname{Re} \alpha > 0$, then $\mathbb{E} |\tilde{X}_n(\alpha)|^2 \leq C(\alpha)n^{2\operatorname{Re} \alpha + 1}$, and thus $\mathbb{E} |\tilde{Y}_n(\alpha)|^2 \leq C(\alpha)$.

In both cases $C(\alpha) = O(1 + |\alpha|^{-2})$.

This shows tightness at each fixed α .

Proof: Magic of analytic functions

Lemma

Let D be a domain in \mathbb{C} and let $(Y_n(z))$ be a sequence of random analytic functions in $\mathcal{H}(D)$. Suppose that there exists a function $\gamma : D \rightarrow (0, \infty)$, bounded on each compact subset of D , such that

$$\mathbb{E} |Y_n(z)| \leq \gamma(z)$$

for every $z \in D$. Then the sequence (Y_n) is tight in the space $\mathcal{H}(D)$ of analytic functions on D .

Proof. Cauchy's integral formula, together with $\mathbb{E} \int = \int \mathbb{E}$.

Proof: Magic of analytic functions

Lemma

Let D be a domain in \mathbb{C} and let $(Y_n(z))$ be a sequence of random analytic functions in $\mathcal{H}(D)$. Suppose that there exists a function $\gamma : D \rightarrow (0, \infty)$, bounded on each compact subset of D , such that

$$\mathbb{E} |Y_n(z)| \leq \gamma(z)$$

for every $z \in D$. Then the sequence (Y_n) is tight in the space $\mathcal{H}(D)$ of analytic functions on D .

Proof. Cauchy's integral formula, together with $\mathbb{E} \int = \int \mathbb{E}$.

Hence, the random functions $\tilde{Y}_n(\alpha)$ are tight in $\mathcal{H}(H_+)$.

More magic of analytic functions

Lemma

Let D be a domain in \mathbb{C} and let E be a subset of D that has a limit point in D . (I.e., there exists a sequence $z_n \in E$ of distinct points and $z_\infty \in D$ such that $z_n \rightarrow z_\infty$.) Suppose that (Y_n) is a tight sequence of random elements of $\mathcal{H}(D)$ and that there exists a family of random variables $\{Y_z : z \in E\}$ such that for each $z \in E$, $Y_n(z) \xrightarrow{d} Y_z$ and, moreover, this holds jointly for any finite set of $z \in E$. Then $Y_n \xrightarrow{d} Y$ in $\mathcal{H}(D)$, for some random function $Y(z) \in \mathcal{H}(D)$.

Proof. Subsequences converge, and limits are determined by the restriction to E , and therefore unique.

More magic of analytic functions

Lemma

Let D be a domain in \mathbb{C} and let E be a subset of D that has a limit point in D . (I.e., there exists a sequence $z_n \in E$ of distinct points and $z_\infty \in D$ such that $z_n \rightarrow z_\infty$.) Suppose that (Y_n) is a tight sequence of random elements of $\mathcal{H}(D)$ and that there exists a family of random variables $\{Y_z : z \in E\}$ such that for each $z \in E$, $Y_n(z) \xrightarrow{d} Y_z$ and, moreover, this holds jointly for any finite set of $z \in E$. Then $Y_n \xrightarrow{d} Y$ in $\mathcal{H}(D)$, for some random function $Y(z) \in \mathcal{H}(D)$.

Proof. Subsequences converge, and limits are determined by the restriction to E , and therefore unique.

Hence, the random functions $\tilde{Y}_n(\alpha)$ converge in distribution in $\mathcal{H}(H_+)$.

Problem: Brownian excursion, $\operatorname{Re} \alpha \leq 1/2$

For $\operatorname{Re} \alpha > 1/2$, we have seen above explicit representations of $\tilde{Y}(\alpha)$ using a Brownian excursion $\mathbf{e}(t)$.

We know that almost surely, this extends to an analytic function in the halfplane $H_+ = \{\alpha : \operatorname{Re} \alpha > 0\}$.

Problem: Brownian excursion, $\operatorname{Re} \alpha \leq 1/2$

For $\operatorname{Re} \alpha > 1/2$, we have seen above explicit representations of $\tilde{Y}(\alpha)$ using a Brownian excursion $\mathbf{e}(t)$.

We know that almost surely, this extends to an analytic function in the halfplane $H_+ = \{\alpha : \operatorname{Re} \alpha > 0\}$.

It follows, using general measure theory, that there exist a measurable function $\psi : C[0, 1] \rightarrow \mathcal{H}(H_+)$ such that

$$Y = \psi(\mathbf{e}).$$

Thus there exists a measurable function $\Psi : H_+ \times C[0, 1] \rightarrow \mathbb{C}$ such that

$$Y(\alpha) = \Psi(\alpha, \mathbf{e}), \quad \operatorname{Re} \alpha > 0.$$

However, this is only an existence statement, and we do not know any explicit representation when $\operatorname{Re} \alpha \leq 1/2$.

Is there an explicit formula giving $Y(\alpha)$ in terms of $\mathbf{e}(t)$ also for $0 < \operatorname{Re} \alpha < \frac{1}{2}$?

THE END