Local limit of massive spanning forests on the complete graph

Matteo D'Achille[1], Nathanaël Enriquez[1,2], Paul Melotti[1]

[1]Laboratoire de Mathématiques d'Orsay, CNRS, Université Paris-Saclay, 91405, Orsay, France
[2]DMA, Ecole Normale Supérieure – PSL, 45 rue d'Ulm, F-75230 Cedex 5 Paris, France
{matteo.dachille,nathanael.enriquez,paul.melotti}@universite-paris-saclay.fr

March 18, 2024

Abstract

We identify the local limit of massive spanning forests on the complete graph. This generalizes a well-known theorem of Grimmett on the local limit of uniform spanning trees on the complete graph.

1 Introduction and main result

Let G be a graph without self-loops. A ***spanning tree*** of G is a connected subgraph of G without cycles and containing all the vertices of G. A ***spanning forest*** of G is a subgraph of G without cycles and containing all the vertices of G. Denote by $\mathcal{F}(G)$ the set of spanning forests of G and for $k \geq 1$ the subset $\mathcal{F}_k(G)$ of spanning forests with k connected components (so that $\mathcal{F}_1(G)$ is the set of spanning trees of G). A ***rooted spanning forest*** is a spanning forest with one marked vertex per connected component, also called the ***root*** of that tree. The set of rooted spanning forests is denoted by $\mathcal{F}^\bullet(\mathcal{G})$, and the subset of those made of k trees is denoted by $\mathcal{F}_k^\bullet(\mathcal{G})$. Note that each forest $f \in \mathcal{F}_k(\mathcal{G})$ with connected components $t_1, \ldots, t_k$ corresponds to $\prod_{i=1}^k |t_i|$ rooted forests, where $|t|$ stands for the number of vertices in t.

For $\lambda > 0$, we introduce the random ***λ-massive spanning forest*** $\lambda\mathbf{SF}(G)$ whose distribution is defined, for all $f \in \mathcal{F}(G)$ having a set of connected components denoted by $C(f)$, by

$$\mathbf{P}\left(\lambda\mathbf{SF}(G) = f\right) \propto \lambda^{|C(f)|} \prod_{t \in C(f)} |t| \,. \tag{1}$$

As $\lambda \downarrow 0$, $\lambda\mathbf{SF}(G)$ converges in distribution to $\mathbf{UST}(G)$. Let $\mathbf{K}_n$ be the complete graph with n vertices and with a distinguished vertex $\mathbf{o}$. Grimmett [Gri80, Theorem 3] proved that the local limit of $\mathbf{UST}(\mathbf{K}_n)$ is a Bienaymé-Galton-Watson tree with reproduction law Poisson(1), denoted by BGWP(1), conditioned to survive forever. The resulting tree $\mathcal{T}_0$ coincides with the random tree obtained by growing a sequence of independent unconditioned BGWP(1) on each vertex of the semi-infinite line rooted at $\mathbf{o}$. The result of Grimmett was recently extended by Nachmias–Peres [NP22,

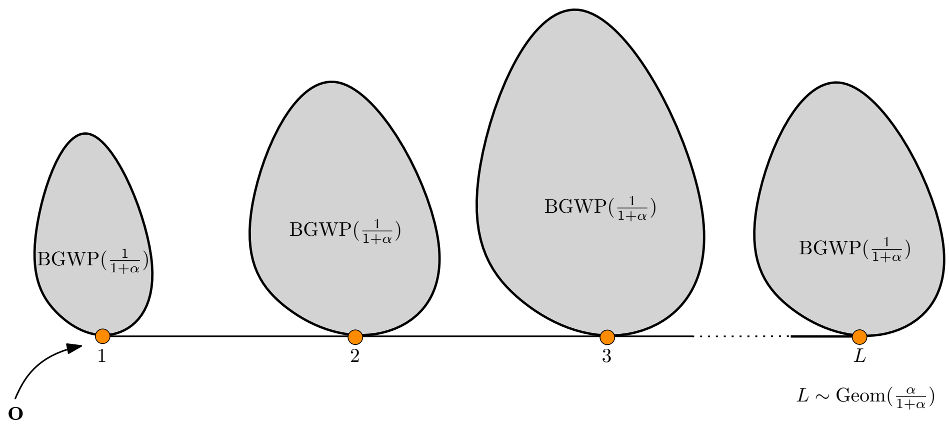

Figure 1: The distribution of the local limit $\mathcal{T}_\alpha$.

Theorem 1.1] from $\mathbf{UST}(\mathbf{K}_n)$ to $\mathbf{UST}(G)$, where G is any simple, connected, regular graph with divergent degree.

On the other hand, massive spanning forests have attracted recent attention, in relation with graph spectra in a series of work by Avena and Gaudillière [AG18a, AG18b], or with dimers and limits of near-critical trees in a work by Rey [Rey24]. A related model of *cycle rooted spanning forests* introduced by Kenyon [Ken11] is also investigated on increasing sequences of graph in a recent work of Constantin [Con23].

Our aim in this paper is to generalize the result of Grimmett to $\lambda\mathbf{SF}(\mathbf{K}_n)$. In order to state our result, we have to introduce the random tree $\mathcal{T}_\alpha$: for $\alpha > 0$, it is obtained by growing a sequence of independent unconditioned BGWP($\frac{1}{1+\alpha}$) on each vertex of a spine rooted at $\mathbf{o}$ with size given by an independent random variable of law Geom($\frac{\alpha}{1+\alpha}$) (see Fig. 1 for a portrait).

Let $T_{n,\lambda}$ be the connected component of $\mathbf{o}$ in $\lambda\mathbf{SF}(\mathbf{K}_n)$. For a given labeled tree T, we denote by Shape$(T, \mathbf{o})$ the non-planar unlabeled tree obtained from T and rooted at $\mathbf{o}$. We endow the set of all locally finite (but possibly infinite) rooted trees with the topology inherited from the product topology. Then the convergence in distribution is called *local convergence*, and coincides with the convergence of probabilities that the tree cut at a finite height is equal to a given pattern. We then have the following local convergence result, for λ_n depending on n as $n \to \infty$:

Theorem 1 (Local limit of $\lambda\mathbf{SF}(K_n)$)**.** *The following local convergence holds:*

$$\mathrm{Shape}(T_{n,\lambda_n}, \mathbf{o}) \xrightarrow[n\to\infty]{(\mathrm{d})} \begin{cases} \mathcal{T}_0 & \text{if } \lambda_n = o(n) \\ \mathcal{T}_\alpha & \text{if } \lambda_n \sim \alpha n \\ \{\mathbf{o}\} & \text{if } \lambda_n \gg n \,. \end{cases}$$

Remark 1.1. A corollary of Theorem 1 is that $\mathcal{T}_\alpha$ is unimodular.

Remark 1.2. As shown in [CAGM18, Eq. 19], the number of connected components has mean $\frac{\lambda+1}{\lambda+n}n$ and is concentrated. When $\lambda_n \sim \alpha n$, this gives asymptotically $\frac{\alpha}{1+\alpha}n$. This result can be informally recovered by the computation of the expectation of the inverse of the total progeny of $\mathcal{T}_\alpha$, which is equal to $\frac{\alpha}{1+\alpha}$.

Acknowledgements. We thank Eleanor Archer and Lucas Rey for insightful discussions. The work of M.A. was supported by the ERC Consolidator Grant SuperGRandMA (Grant No. 101087572). N.E. was partially supported by the CNRS grant RT 2179, MAIAGES. M.A. et N.E. were partially supported by the CNRS grant RT 2173 Mathématiques et Physique, MP.

2 Determinantal properties of massive spanning forests on a graph

Let H be an undirected graph without self-loops but with possibly multiple edges, denote by $V(H)$ and $E(H)$ respectively the vertex set and the edge set of H. Let also $n = |V(H)|$ be the ***size*** of the graph. Given two vertices $u, v \in V(H)$, we shall denote by $u \sim v$ the fact that there exists an edge connecting u to v, and by $u \sim_e v$ the fact that a specific edge $e \in E(H)$ connects u to v. For a vertex $v \in V(H)$, we shall denote by $d(v)$ the degree of v. Given some labeling of the vertices $(v_i)_{i=1}^n$, the ***adjacency matrix*** A_H is defined by

$$\forall i, j = 1, \ldots, n \quad (A_H)_{ij} \stackrel{\text{def}}{=} |\{e \in E(H)\,;\; v_i \sim_e v_j\}| \tag{2.1}$$

and the ***graph Laplacian*** Δ_H is a symmetric, positive semi-definite matrix defined by

$$\forall i, j = 1, \ldots, n \quad (\Delta_H)_{ij} \stackrel{\text{def}}{=} d(v_i)\delta_{ij} - (A_H)_{ij} \ . \tag{2.2}$$

We denote by P_H the ***characteristic polynomial*** of Δ_H defined by

$$P_H(\lambda) \stackrel{\text{def}}{=} \det(\Delta_H + \lambda I_n) \ . \tag{2.3}$$

This characteristic polynomial turns out be the partition function of spanning forests according to their number of connected components, as shown by the following celebrated Kirchhoff's matrix-forest theorem.

Theorem 2.1 (Kirchhoff's or Matrix-Forest Theorem [Kir47])**.** *Let H be a graph of size n without self-loops. Then*

$$P_H(\lambda) = \sum_{k=1}^{n} \lambda^k |\mathcal{F}_k^\bullet(H)| = \sum_{k=1}^{n} \lambda^k \left[\sum_{f \in \mathcal{F}_k(H)} \prod_{t \in C(f)} |t| \right] . \tag{2.4}$$

Kenyon generalized this result in [Ken11] where he described the partition function of cycle rooted spanning forests in terms of bundle Laplacian.

A consequence of this fundamental theorem is that the uniform spanning trees, and more generally the λ**SF** model of spanning forests, are determinantal. This was first discovered by Burton and Pemantle in [BP93] for the uniform spanning tree case, see also the book by Lyons and Peres [LP16, Chapter 4] for an entire chapter devoted to the topic. As for spanning forests, determinantal formulas for correlations seem to have appeared sporadically. They are described in the PhD thesis of Chang [Cha13, Section 5.2] who considered the equivalent framework of spanning trees rooted at a cemetery point. Concerning the set of roots, determinantal formulas are stated by Avena and Gaudillière in [AG18b]. In this work, we need an expression of presence or absence of edges which is given by a determinant, generalizing [BP93].

For $\lambda \notin \mathrm{Sp}(\Delta_H)$, consider the *resolvent* matrix (also called *massive Green's function*)

$$R_\lambda = (\Delta_H + \lambda I_n)^{-1}. \tag{2.5}$$

Using this operator on $V(H)$, we define another operator on $E(H)$ by first fixing an arbitrary orientation of the edges, so that every edge $e \in E(H)$ has an origin vertex e_- and a target vertex e_+. Then for two edges $e, f \in E(H)$ we define

$$K_\lambda(e,f) = R_\lambda(e_-, f_-) + R_\lambda(e_+, f_+) - R_\lambda(e_-, f_+) - R_\lambda(e_+, f_-). \tag{2.6}$$

Seen as a matrix, K is called ***transfer current matrix***. This statement says that the λ**SF** model is determinantal with kernel K_λ.

Proposition 2.1. *Let $e_1, \ldots, e_k, e_{k+1}, \ldots, e_p$ be p distinct edges of H, let $\lambda > 0$. Then*

$$\mathbf{P}(e_1, \ldots, e_k \in \lambda\mathbf{SF}(H),\ e_{k+1}, \ldots, e_p \notin \lambda\mathbf{SF}(H)) = \det M \tag{2.7}$$

where M is the $p \times p$ matrix with entries

$$M_{i,j} = \begin{cases} K_\lambda(e_i, e_j) & \text{if } i \le k \\ \delta_{i,j} - K_\lambda(e_i, e_j) & \text{if } k < i \le p. \end{cases} \tag{2.8}$$

Proof. We start with the case $k = 0$, that is, we want to compute $\mathbf{P}(e_1, \ldots, e_p \notin \lambda\mathbf{SF}(H))$. Let $H \setminus \{e_1, \ldots, e_p\}$ be the graph obtained by deleting the edges $\{e_1, \ldots, e_p\}$ from $E(H)$. By Theorem 2.1, this may be expressed as

$$\frac{P_{H\setminus\{e_1,\ldots,e_p\}}(\lambda)}{P_H(\lambda)} = \frac{\det\left(\Delta_{H\setminus\{e_1,\ldots,e_p\}} + \lambda I_n\right)}{\det\left(\Delta_H + \lambda I_n\right)}. \tag{2.9}$$

The two involved matrices can be related in the following way. Recall that we fixed an arbitrary orientation of every edge. For any $1 \le i \le p$, let b_i be the vector of dimension $n = |V(H)|$ with entries -1 at the origin of e_i, $+1$ at the tip of e_i, and 0 otherwise. Then we have

$$\Delta_{H\setminus\{e_1,\ldots,e_p\}} = \Delta_H - \sum_{i=1}^{p} b_i b_i^T = \Delta_H - BB^T \tag{2.10}$$

where B is the $n \times p$ matrix with columns b_i. Therefore, using the Weinstein–Aronszajn identity,

$$\begin{aligned} \det\left(\Delta_{H\setminus\{e_1,\ldots,e_p\}}\right) &= \det\left(\Delta_H + \lambda I_n - BB^T\right) \\ &= \det\left(\Delta_H + \lambda I_n\right)\det\left(I_n - R_\lambda BB^T\right) \\ &= \det\left(\Delta_H + \lambda I_n\right)\det\left(I_p - B^T R_\lambda B\right) \end{aligned} \tag{2.11}$$

The entries of $B^T R_\lambda B$ can be seen to be equal to that of K_λ by a direct computation. This concludes the proof in the case $k = 0$.

For $k \ge 1$, we can write the probability of the event $\{e_1, \ldots, e_k \in \lambda\mathbf{SF}(H),\ e_{k+1}, \ldots, e_p \notin \lambda\mathbf{SF}(H)\}$ using inclusion-exclusion as a combination of probabilities of $\{e_{i_1}, \ldots, e_{i_l}, e_{k+1}, \ldots, e_p \notin \lambda\mathbf{SF}(H)\}$, and those can be expressed via the first case. It is direct to check that multi-linearity of the determinants implies the generic formula. □

3 A formula for the inclusion probability of a tree on $\mathbf{K}_n$

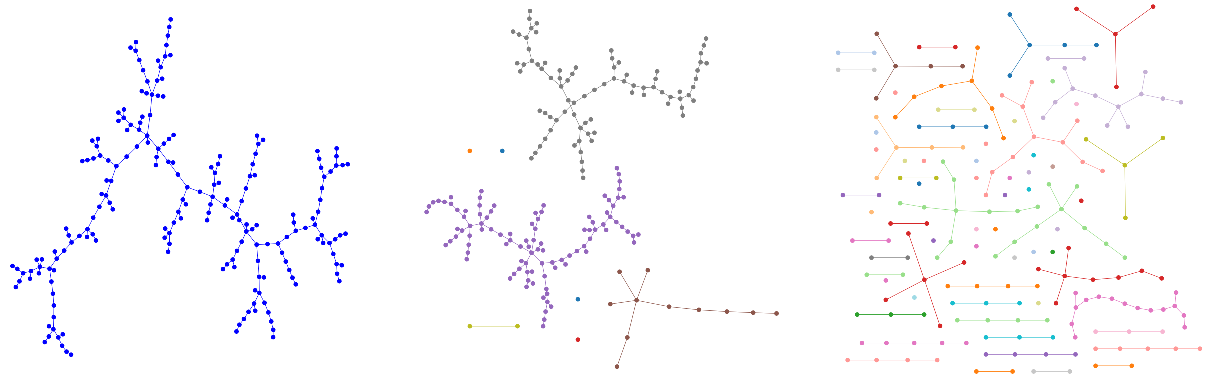

Figure 2: Samples of $\lambda\mathbf{SF}(\mathbf{K}_{200})$ obtained via Wilson's algorithm with a uniform, positive killing rate $r = 1 - \mu = \frac{\lambda}{\lambda+n-1}$. From left to right: $r = 0$ (i.e. a sample of $\mathbf{UST}(\mathbf{K}_{200})$), $r = \frac{1}{40}$ and $r = \frac{1}{3}$.

Let $\mathbf{K}_n$ denote the complete graph on n vertices.

Proposition 3.1 (Resolvent and transfer current matrix)**.** *The resolvent and transfer current matrix of $\mathbf{K}_n$ are given by*

$$R_\lambda(i,j) = \begin{cases} \frac{1+\lambda}{\lambda(n+\lambda)} & \text{if } i = j \\ \frac{1}{\lambda(n+\lambda)} & \text{if } i \neq j \end{cases} \tag{3.1}$$

and

$$K_\lambda(e,f) = \frac{1}{(n+\lambda)}\left(\mathbf{1}_{e_-=f_-} + \mathbf{1}_{e_+=f_+} - \mathbf{1}_{e_-=f_+} - \mathbf{1}_{e_+=f_-}\right) \tag{3.2}$$

Proof. We start with a preparatory computation of the Green's function of a uniform random walk X_k^μ on $\mathbf{K}_n$ killed at some positive rate $1-\mu$[1], which is equal to $G^\mu(x,y) = \sum_{k\geq 0} \mu^k \mathbf{P}_x(X_k = y)$, where X_k denotes the uniform RW. Using the notation of Eq. (2.5), and the fact that the transition matrix of the random walk P is related to the laplacian via $\Delta = (n-1)(I-P)$, we get

$$R_\lambda(x,y) = \frac{\mu}{n-1} G^\mu(x,y) \quad \text{where } \mu = \frac{n-1}{n-1+\lambda}\ . \tag{3.3}$$

By symmetry of the problem, we only need to compute $G^\mu(1,1)$ and $G^\mu(1,2)$. We need now to determine the distribution of X_k conditional on $X_0 = 1$. Denoting by $x_k \overset{\text{def}}{=} \mathbf{P}_1(X_k = 1)$ and $y_k \overset{\text{def}}{=} \mathbf{P}_1(X_k = 2)$, usual Markov properties show that for $k \geq 0$,

$$\begin{cases} x_{k+1} = y_k, \\ y_{k+1} = \frac{n-2}{n-1} y_k + \frac{1}{n-1} x_k, \end{cases}$$

with $(x_0, y_0) = (1, 0)$. From this recursion we deduce that, for all $k \geq 0$, $y_{k+1} - y_k = -(-\frac{1}{n-1})^{k+1}$ and using Abel transform, we get $G^\mu(1,2) = \sum_{k\geq 1} \mu^k y_k = \frac{n-1+\lambda}{\lambda(n+\lambda)}$, giving, via Eq. (3.3), $R_\lambda(1,2) = \frac{1}{\lambda(n+\lambda)}$.

[1] This also provides an easy recipe to sample $\lambda\mathbf{SF}(\mathbf{K}_n)$, see Figure 2.

Since $x_{k+1} = y_k$, $G^\mu(1,1) = 1 + \mu G^\mu(1,2)$, and $R_\lambda(1,1) = \frac{\mu}{n-1} + \mu\, R_\lambda(1,2)$, which gives the formulas for R_λ.

From Eq. (2.6), we can write

$$K_\lambda(e,f) = \begin{cases} 2\left(R_\lambda(x,x) - R_\lambda(x,y)\right) \text{ if } e = f \\ \left(R_\lambda(x,x) - R_\lambda(x,y)\right)\left(\mathbf{1}_{e_-=f_-} + \mathbf{1}_{e_+=f_+} - \mathbf{1}_{e_-=f_+} - \mathbf{1}_{e_+=f_-}\right) \text{ if } e \neq f \end{cases} \tag{3.4}$$

which leads to the formulas for K_λ.

□

We can now compute the probability that a given tree is included in $\lambda\mathbf{SF}(\mathbf{K}_n)$:

Lemma 3.2 (Inclusion probability)**.** *Let t be a tree with vertices labeled by distinct integers in $\{1,\dots,n\}$, then*

$$\mathbf{P}\left(t \subset \lambda\mathbf{SF}(\mathbf{K}_n)\right) = \frac{|t|}{(n+\lambda)^{|t|-1}}\,.$$

Proof. Let $k = |t|$. Fix an orientation of t. From Proposition 2.1, $\mathbf{P}\left(t \subset \lambda\mathbf{SF}(\mathbf{K}_n)\right) = \det M_\lambda$, where M_λ is matrix of size $(k-1)\times(k-1)$ indexed by the edges of t, whose entries are given by Proposition 3.1. One can check directly that

$$M_\lambda = \frac{1}{n+\lambda}\mathcal{A}\mathcal{A}^T,$$

where $\mathcal{A}$ is the (rectangular) oriented incidence matrix of t: it has size $(k-1)\times k$, with rows indexed by the edges of t, columns indexed by the vertices of t, and whose nonzero entries are $+1$ (resp. -1) when the vertex is the tip (resp. origin) of the edge. Then, by the Cauchy-Binet formula,

$$\det M_\lambda = \frac{1}{(n+\lambda)^{k-1}}\sum_{j=1}^{k}\left(\det \mathcal{A}_j\right)^2$$

where $\mathcal{A}_j$ is obtained from $\mathcal{A}$ by deletion of column j. Then one checks directly that $\det \mathcal{A}_j = \pm 1$, for instance by expanding the determinant over permutations and seeing that only one permutation contributes. □

4 Identification of the local limit: Proof of Theorem 1

The proof comes in two steps. First, we identify the limit law of the shape of $\lambda\mathbf{SF}(\mathbf{K}_n)$ at a given distance h from $\mathbf{o}$. Second, we show that it coincides with the distribution of $\mathcal{T}_\alpha$ restricted at distance h from its root. These steps are the content of the following two propositions.

For a given rooted unlabeled tree T and an integer h, we denote by:

- T_h the set of vertices of T at distance exactly h from the root;
- $T_{\leq h}$ the tree given by the intersection of T with the closed ball of radius h centered at the root;
- $T_{<h}$ the tree given by the intersection of T with the closed ball of radius $h-1$ centered at the root;

- $\mathrm{Aut}(T)$ the set of graph automorphisms of T preserving its root;
- $\mathrm{d}(u,v)$ the graph distance between vertices u and v.

Proposition 4.1. *Let $h \in \mathbb{N}$. Let t be an unlabeled tree of height $\leq h$. Then*

$$\mathbf{P}\left(\mathrm{Shape}(\lambda\mathbf{SF}(\mathbf{K}_n),\mathbf{o})_{\leq h} = t\right) = \frac{1}{|\mathrm{Aut}(t)|}\frac{(n-1)!}{(n-|t|)!}\frac{1}{(n+\lambda)^{|t|}}\left(n|t_h| + \lambda|t|\right)\left(1-\frac{|t_{<h}|}{n+\lambda}\right)^{n-|t|-1} .$$

When $\lambda = \lambda_n$,

$$\lim_{n\to\infty}\mathbf{P}\left(\mathrm{Shape}(\lambda_n\mathbf{SF}(\mathbf{K}_n),\mathbf{o})_{\leq h} = t\right) = \begin{cases} \frac{|t_h|}{|\mathrm{Aut}(t)|}\exp\left(-|t_{<h}|\right) & \text{if } \lambda_n = o(n) \\ \frac{\left(|t_h|+\alpha|t|\right)}{|\mathrm{Aut}(t)|}\frac{1}{(1+\alpha)^{|t|}}\exp\left(-\frac{|t_{<h}|}{1+\alpha}\right) & \text{if } \lambda_n \sim \alpha n \\ \delta_{t=\{\mathbf{o}\}} & \text{if } \lambda_n \gg n . \end{cases}$$

Proof. Let $T^{(n)}_{\mathrm{lab}}$ be the connected component of $\lambda\mathbf{SF}(\mathbf{K}_n)$ containing $\mathbf{o}$, $T^{(n)}$ its unlabeled version, t as in the statement and t_{lab} a labeled tree whose shape is t. Denote the size of t by $|t|(=|t_{\mathrm{lab}}|)$. We wish to compute the probability of the event $T^{(n)}_{\mathrm{lab},\leq \mathrm{h}} = t_{\mathrm{lab}}$. This event coincides with the presence of all the edges of t_{lab} in $T^{(n)}_{\mathrm{lab}}$ and with the absence of all the edges joining a vertex of $t_{\mathrm{lab},<\mathrm{h}}$ to a vertex which does not belong to t_{lab} (of which there are $n-|t|$). Therefore there are $q := |t_{\mathrm{lab},<\mathrm{h}}|\cdot(n-|t|)$ such absent edges, call them $e_1,\ldots,e_q$. By the inclusion-exclusion principle, we can express the probability of this event only in terms of inclusion probabilities and use Lemma 3.2. Namely,

$$\begin{aligned}\mathbf{P}\left(T^{(n)}_{\mathrm{lab},\leq\mathrm{h}} = t_{\mathrm{lab}}\right) &= \mathbf{P}\left(t_{\mathrm{lab}} \subset T^{(n)}_{\mathrm{lab},\leq\mathrm{h}}, e_1 \notin T^{(n)}_{\mathrm{lab}},\ldots,e_q \notin T^{(n)}_{\mathrm{lab}}\right) \\ &= \sum_{k=0}^{q}(-1)^k \sum_{\substack{J\subset\{1,\ldots,q\}\\|J|=k}} \mathbf{P}\left(t_{\mathrm{lab}} \subset T^{(n)}_{\mathrm{lab}},\ \forall j\in J\ e_j \in T^{(n)}_{\mathrm{lab}}\right) .\end{aligned} \tag{4.1}$$

The probability appearing in the second sum on the rhs is equal to $\frac{|t|+k}{(n+\lambda)^{|t|-1+k}}$ if $t_{\mathrm{lab}} \cup \{e_j\,;\ j\in J\}$ is still a tree, and 0 otherwise. A cycle is formed if and only if at least two edges among $\{e_j\,;\ j\in J\}$ are incident at the same vertex not in t_{lab}. Therefore, as soon as $k > n-|t|$, this probability is zero. Otherwise, there are $\binom{n-|t|}{k}$ choices of "target" vertices for the tips of $\{e_j\,;\ j\in J\}$ and the roots of $\{e_j\,;\ j\in J\}$ can be chosen freely inside $t_{\mathrm{lab},<\mathrm{h}}$. We thus get, after straightforward computation,

$$\begin{aligned}\mathbf{P}\left(T^{(n)}_{\mathrm{lab},\leq\mathrm{h}} = t_{\mathrm{lab}}\right) &= \sum_{k=0}^{n-|t|}(-1)^k\binom{n-|t|}{k}|t_{<h}|^k\frac{|t|+k}{(n+\lambda)^{|t|-1+k}} \\ &= \frac{1}{(n+\lambda)^{|t|}}\left(n|t_h|+\lambda|t|\right)\left(1-\frac{|t_{<h}|}{n+\lambda}\right)^{n-|t|-1} .\end{aligned} \tag{4.2}$$

Lastly, we want to compute $\mathbf{P}\left(\mathrm{Shape}(\lambda\mathbf{SF}(\mathbf{K}_n),\mathbf{o})_{\leq h} = t\right)$. For each labeling $\mathbf{l}$ of t, there are exactly $|\mathrm{Aut}(t)|$ labelings $\mathbf{l}'$ for which the events $\{T^{(n)}_{\mathrm{lab},\leq h} = t_{\mathbf{l}}\}$ and $\{T^{(n)}_{\mathrm{lab},\leq h} = t_{\mathbf{l}'}\}$ are the same. On the other hand, there are $\frac{(n-1)!}{(n-|t|)!}$ possible labelings. This concludes the proof for n fixed, and the limits are straightforward.

□

We turn to the second step, where we study the limiting random tree $\mathcal{T}_\alpha$. We start with a preliminary lemma on standard subcritical BGWP(β).

Lemma 4.2. *Let $\beta < 1$, let $h \in \mathbb{N}$. Let t be an unlabeled tree of height $\leq h$.*

$$\mathbf{P}\left(\mathrm{BGWP}(\beta)_{\leq h} = t\right) = \frac{1}{|\mathrm{Aut}(t)|}\beta^{|t|-1}e^{-\beta|t_{<h}|} \, .$$

Proof. We compare the recursion relations satisfied by both quantities. We group the children of the root in equivalence classes: two children are in the same class if and only if they are the root of isomorphic subtrees. We denote now by ℓ the number of equivalence classes, and by $n_1, \dots, n_l$ their cardinality. Since the number of children of the root is Poisson(β) distributed, we can write

$$\begin{aligned}\mathbf{P}\left(\mathrm{BGWP}(\beta)_{\leq h} = t\right) &= \frac{\beta^{n_1+\cdots+n_l}e^{-\beta}}{(n_1+\cdots+n_l)!}\binom{n_1+\cdots+n_l}{n_1,\dots,n_l}\prod_{i=1}^{l}\mathbf{P}\left(\mathrm{BGWP}(\beta)_{\leq h-1} = t_i\right)^{n_i} \\ &= \frac{\beta^{n_1+\cdots+n_l}e^{-\beta}}{n_1!\cdots n_l!}\prod_{i=1}^{l}\mathbf{P}\left(\mathrm{BGWP}(\beta)_{\leq h-1} = t_i\right)^{n_i}\end{aligned} \tag{4.3}$$

where t_i is the subtree of t emanating from a child of the root belonging to the i-th equivalence class. On the other hand, $|\mathrm{Aut}(t)|$ satisfies the recursion

$$|\mathrm{Aut}(t)| = \prod_{i=1}^{l} n_i! |\mathrm{Aut}(t_i)|^{n_i}$$

hence if we denote by f the function on finite trees which maps t onto $f(t) = \frac{1}{|\mathrm{Aut}(t)|}\beta^{|t|-1}e^{-\beta|t_{<H(t)}|}$, where $H(t)$ is the height of t,

$$\frac{\beta^{n_1+\cdots+n_l}e^{-\beta}}{n_1!\cdots n_l!}\prod_{i=1}^{l} f(t_i)^{n_i} = \frac{\beta^{n_1+\cdots+n_l}e^{-\beta}}{n_1!\cdots n_l!}\prod_{i=1}^{l}\left(\frac{1}{|\mathrm{Aut}(t_i)|}\beta^{|t_i|-1}e^{-\beta|t_{i,<H(t_i)}|}\right)^{n_i} = f(t) \, . \tag{4.4}$$

Equations (4.3) and (4.4) show that both quantities of the statement satisfy the same recursion. □

We can now identify the distribution of Shape($\mathcal{T}_\alpha, \mathbf{o}$):

Proposition 4.3. *Let $h \in \mathbb{N}$. Let t be an unlabeled tree of height $\leq h$. Then*

$$\mathbf{P}\left(\mathrm{Shape}(\mathcal{T}_\alpha, \mathbf{o})_{\leq h} = t\right) = \frac{|t_h| + \alpha|t|}{|\mathrm{Aut}(t)|}\frac{1}{(1+\alpha)^{|t|}}\exp\left(-\frac{|t_{<h}|}{1+\alpha}\right) \, .$$

Proof. We decompose the probability according to the position of the vertex v of the spine which is the farthest from $\mathbf{o}$. Let us denote by $[v]$ its equivalence class in the quotient of t by the action of $\mathrm{Aut}(t)$. The vertex v might be at distance h or not; in any case, the probability given that v satisfies our requirement depends only on $[v]$. Calling L the cardinality of the spine, we will use the relation $\mathbf{P}(L \geq h) = \left(\frac{1}{1+\alpha}\right)^h$ when $v \in t_h$ and the relation $\mathbf{P}(L = k+1) = \frac{\alpha}{(1+\alpha)^{k+1}}$ when $\mathrm{d}(v, \mathbf{o}) = k$. Calling $t_{i,v}$ the tree growing from the i-th vertex of the spine, we have

$$\begin{aligned}\mathbf{P}\left(\mathrm{Shape}(\mathcal{T}_\alpha, \mathbf{o})_{\leq h} = t\right) = &\sum_{[v]\in t_h/\mathrm{Aut}(t)}\left[\prod_{i=0}^{h}\mathbf{P}\left(\mathrm{BGWP}\left(\frac{1}{1+\alpha}\right)_{h-i+1} = t_{i,v}\right)\right]\left(\frac{1}{1+\alpha}\right)^h + \\ &+ \sum_{[v]\in t_{<h}/\mathrm{Aut}(t)}\left[\prod_{i=0}^{\mathrm{d}(v,\mathbf{o})}\mathbf{P}\left(\mathrm{BGWP}\left(\frac{1}{1+\alpha}\right)_{h-i+1} = t_{i,v}\right)\right]\frac{\alpha}{(1+\alpha)^{\mathrm{d}(v,\mathbf{o})+1}} \, .\end{aligned} \tag{4.5}$$

We use Lemma 4.2 with $\beta = \frac{1}{1+\alpha}$. Note that

$$(1+\alpha)^{|t_{1,v}|-1} \cdots (1+\alpha)^{|t_{h,v}|-1} (1+\alpha)^{\mathrm{d}(v,\mathbf{o})} = (1+\alpha)^{|t|},$$

which gives

$$\begin{aligned}\mathbf{P}\left(\mathrm{Shape}(\mathcal{T}_\alpha, \mathbf{o})_{\leq h} = t\right) = &\sum_{[v] \in t_h/\mathrm{Aut}(t)} \frac{\exp\left(-\frac{|t_{<h}|}{1+\alpha}\right)}{(1+\alpha)^{|t|}} \prod_{i=0}^{h} \frac{1}{|\mathrm{Aut}(t_{i,v})|} + \\ &+ \frac{\alpha}{1+\alpha} \sum_{[v] \in t_{<h}/\mathrm{Aut}(t)} \frac{\exp\left(-\frac{|t_{<h}|}{1+\alpha}\right)}{(1+\alpha)^{|t|}} \prod_{i=0}^{\mathrm{d}(v,\mathbf{o})} \frac{1}{|\mathrm{Aut}(t_{i,v})|}\,.\end{aligned} \tag{4.6}$$

To conclude, we use the Burnside lemma for the action of Aut on t_h and on $t_{<h}$, giving

$$|t_h| = \sum_{[v] \in t_h/\mathrm{Aut}(t)} |\mathrm{Aut}(t)| \prod_{i=0}^{h} \frac{1}{|\mathrm{Aut}(t_{i,v})|} \text{ and } |t_{<h}| = \sum_{[v] \in t_{<h}/\mathrm{Aut}(t)} |\mathrm{Aut}(t)| \prod_{i=0}^{\mathrm{d}(v,\mathbf{o})} \frac{1}{|\mathrm{Aut}(t_{i,v})|},$$

and the statement follows.

□

References

[AG18a] Luca Avena and Alexandre Gaudillière. A proof of the transfer-current theorem in absence of reversibility. *Statistics & Probability Letters*, 142:17–22, 2018.

[AG18b] Luca Avena and Alexandre Gaudillière. Two Applications of Random Spanning Forests. *Journal of Theoretical Probability*, 31(4):1975–2004, December 2018.

[BP93] Robert Burton and Robin Pemantle. Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *The Annals of Probability*, 21(3):1329–1371, 1993.

[CAGM18] Fabienne Castell, Luca Avena, Alexandre Gaudilliere, and Clothilde Melot. Random Forests and Networks Analysis. *Journal of Statistical Physics*, August 2018.

[Cha13] Yinshan Chang. *Contribution à l'étude des lacets markoviens.* Phd thesis, Université Paris Sud - Paris XI, June 2013.

[Con23] Héloïse Constantin. Sampling random cycle-rooted spanning forests on infinite graphs. *arXiv preprint arXiv:2308.09425*, 2023.

[Gri80] Geoffrey R. Grimmett. Random labelled trees and their branching networks. *Journal of the Australian Mathematical Society*, 30(2):229–237, 1980.

[Ken11] Richard Kenyon. Spanning forests and the vector bundle Laplacian. *The Annals of Probability*, 39(5):1983 – 2017, 2011.

[Kir47] Gustav R. Kirchhoff. Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Annalen der Physik*, 148:497–508, 1847.

[LP16] Russell Lyons and Yuval Peres. *Probability on Trees and Networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York, 2016. Available at https://rdlyons.pages.iu.edu/.

[NP22] Asaf Nachmias and Yuval Peres. The local limit of uniform spanning trees. *Probability Theory and Related Fields*, 182(3-4):1133–1161, 2022.

[Rey24] Lucas Rey. The Doob transform and the tree behind the forest, with application to near-critical dimers. *arXiv preprint arXiv:2401.13599*, 2024.