ON THE PROFILE OF RANDOM FORESTS

BERNHARD GITTENBERGER*

Abstract. An approach via generating functions is used to derive multivariate asymptotic distributions for the number of nodes in strata of random forests. For a certain range for the strata numbers we obtain a weak limit theorem to Brownian motion as well. Moreover, a moment convergence theorem for the width of random forests is derived.

1. Introduction

We consider the set $F(n,N)$ of random forests consisting of n vertices and N rooted trees which can be viewed as realizations of Galton-Watson branching processes with N initial particles and conditioned to have total progeny n. Such forests consist of simply generated trees according to Meir and Moon [20] and therefore they can easily be described by generating functions: Let $b(z)=\sum_{n\geq 0} b_{n,N} z^n$ denote the generating function for those forests. Then we have $b(z)=a(z)^N$ with $a(z)=z\varphi(a(z))$. Here $a(z)$ is the generating function for a single tree and $\varphi(t)=\sum_{n\geq 0}\varphi_n t^n$ is the generating function of an arbitrary sequence $(\varphi_k)_{k\geq 0}$ of nonnegative numbers with $\varphi_0>0$. In this setting $b_{n,N}$ can be viewed as the number of forests in $F(n,N)$, weighted according to the probability on $F(n,N)$, i.e., to each forest F is assigned a weight

$$\omega(F)=\prod_{k\geq 0}\varphi_k^{n_k(F)}$$

where $n_k(F)$ is the number of nodes with out-degree k. The φ_k are related to the offspring distribution ξ via $\mathbf{P}\{\xi=k\}=\tau^k\varphi_k/\varphi(\tau)$, with a positive number τ within the circle of convergence of $\varphi(t)$. This means that the probability that the realization CGW of a conditioned Galton-Watson process as described above (offspring ξ, N initial particles, and conditioned to total progeny n) equals a given forest $f\in F(n,N)$ is proportional to the weight of f, precisely, we have

$$\mathbf{P}\{\mathrm{CGW}=f\}=\omega(f)\Big/\sum_{f\in F(n,N)}\omega(f).$$

Without loss of generality we may assume $\mathbf{E}\xi=1$ which equivalently means that τ satisfies $\tau\varphi'(\tau)=\varphi(\tau)$. Then the variance of ξ can also be expressed in terms of $\varphi(t)$ and is given by

$$\sigma^2=\frac{\tau^2\varphi''(\tau)}{\varphi(\tau)}. \tag{1.1}$$

The height of a vertex x is defined by the number of edges comprising the unique path which connects x with the root of the tree containing x. We are interested in the profile of random forests, thus we define $L_{n,N}(k)$ to be the number of vertices at height k in a random forest in $F(n,N)$. First, let us mention that the average height of a random forest in $F(n,N)$ is proportional to $\sqrt{n}$ as $n\to\infty$ and $N=O(\sqrt{n})$, see [21] and [23] for special cases and [24] for general simply generated forests. Thus the most interesting range is $k=O(\sqrt{n})$. Pavlov [22] derived distributional results for various ranges of n,N,k for labeled trees. Different tree classes are treated in [4, 10] and results for other ranges can be found in [26]. For a survey of results on random forests we refer the reader to [25]. Theorems 5 and 6 in [22] give a formula for the limiting distribution as integral with respect to a two-dimensional probability distribution with explicit Fourier transform for the ranges $k/\sqrt{n}\to\alpha>0$ and $N=o(\sqrt{n})$ and $N\sim\sqrt{n}$ (cf. [19] for the random tree analogue). These theorems have been generalized by Pitman [27] who related the profile of simply generated

Date: June 19, 2002.

* Department of Geometry, Technische Universität Wien, Wiedner Hauptstraße 8-10/113, A-1040 Wien, Austria.

random forests in the above mentioned range for n, N, k to stochastic differential equations and obtained a weak limit theorem

$$\left(\frac{2}{\sigma\sqrt{n}}L_{n,N}\left(\frac{2\kappa\sqrt{n}}{\sigma}\right), \kappa \geq 0\right) \xrightarrow{d} (X_{\alpha,\kappa}, \kappa \geq 0) \tag{1.2}$$

if $2N/\sigma\sqrt{n} \to \alpha$, where $X_{\alpha,\kappa}$ can be characterized by a stochastic differential equation: Let β denote a Brownian motion and set

$$u(X) = \inf\left\{v : \int_0^v X_s\, ds = 1\right\}.$$

Then Pitman [27] showed that for each $\alpha > 0$ there exists a unique strong solution of the Itô SDE

$$X_0 = \alpha, \quad dX_\kappa = \delta_\kappa(X)\, dv + 2\sqrt{X\kappa}\, d\beta_\kappa; \qquad \kappa \in [0, u(X)), \quad X_\kappa \equiv 0 \text{ for } \kappa \geq u(X)$$

with

$$\delta_\kappa(X) = 4 - X_\kappa^2\left(1 - \int_0^\kappa X_s\, ds\right)^{-1}.$$

This process can be identified as total local time of a Brownian bridge B of length one conditioned to have total local time α at level 0 (see [27]),

$$X_{\alpha,v} \stackrel{d}{=} (\ell_v(B) | \ell_0(B) = \alpha), \tag{1.3}$$

which coincides with a Brownian excursion local time if $\alpha = 0$ (cf. the analogous results for random trees, see [6] for the combinatorial setting and [27] for the stochastic calculus setting).

In this paper we are interested in the behavior of $L_{n,N}(k)$ in low strata of random forests. Starting point is the following central limit theorem (see [22, 10, 4]):

Theorem 1.1. *Let $n \to \infty$, $N = O(\sqrt{n})$, and $k = o(N)$. Then*

$$\mathbf{P}\left\{\frac{L_{n,N}(k) - N}{\sigma\sqrt{Nk}} \leq x\right\} \to \frac{1}{2\pi}\int_{-\infty}^x e^{-u^2/2}\, du$$

for any fixed x.

In order to simplify the proofs in the following, let us define $L_{n,N}(t)$ also for noninteger t by linear interpolation:

$$L_n(t) = (\lfloor t \rfloor + 1 - t)L_n(\lfloor t \rfloor) + (t - \lfloor t \rfloor)L_n(\lfloor t \rfloor + 1), \quad t \geq 0.$$

This does not change the limit of the finite-dimensional distributions and simplifies the proof of tightness significantly, since we are dealing with continuous functions.

Theorem 1.1 suggests the convergence to a Gaussian limiting process. In fact we will show the following theorem.

Theorem 1.2. *Let $\varphi(t)$ be a generating function associated to a family of simply generated trees. Assume that $\varphi(t)$ has a positive or infinite radius of convergence R and $\zeta = \gcd\{i | \varphi_i > 0\} = 1$. Suppose that the equation $t\varphi'(t) = \varphi(t)$ has a minimal positive solution $\tau < R$ and that σ^2 defined by (1.1) is finite. Furthermore, let (c_n) be an arbitrary sequence satisfying $c_n \to \infty$ and $c_n = o(N)$. Moreover, assume $N = O(\sqrt{n})$. Then*

$$\left(\frac{1}{\sigma\sqrt{Nc_n}}\left(L_{n,N}(tc_n) - N\right), t \geq 0\right) \xrightarrow{d} W_t$$

where W_t is a standard Brownian motion.

The proof of this theorem is done by first deriving a limit theorem for the finite-dimensional distributions which is done in the next section. This will be established by describing the joint distribution by means of a suitable generating function (see [9] or [14] for a general background) and then determining an asymptotic formula (and thus the limiting distribution) by complex contour integration. Afterwards we have to prove tightness which is done in Section 3. Section 4 is devoted to higher strata of random forests, i.e., the case $c_n/\sqrt{n} \to \eta > 0$. The limiting process for this case has been completely characterized by Pitman [27] (see (1.3)), however, using the combinatorial scheme of Section 2 we can give more explicit expressions for the finite-dimensional distributions in terms of integral transforms for the characteristic functions. Moreover, due to a tight bound derived in Section 3, it is also possible to derive a moment convergence theorem for the node numbers at this range as well as for the width of random forests, which complements the weak limit theorem of Pitman [27] (cf. also [3] and [7] for the corresponding results for trees). In fact, we will show

Theorem 1.3. *Set $M_\alpha := \sup_{v\geq 0} X_{\alpha,v}$ and $w_{n,N} := \sup_k 2L_{n,N}(k)/\sigma\sqrt{n}$. If $n, N \to \infty$ such that $2N/\sigma\sqrt{n} \to \alpha > 0$, then we have for every $d > 0$*

$$\mathbf{E}w_{n,N}^d \to \mathbf{E}M_\alpha^d \text{ and } \mathbf{E}\left(\frac{2}{\sigma\sqrt{n}}L_{n,N}\left(\frac{2\kappa\sqrt{n}}{\sigma}\right)\right)^d \to \mathbf{E}X_{\alpha,\kappa}^d$$

2. The finite-dimensional distributions

We have to compute the joint distribution of $L_{n,N}(k_1), \dots, L_{n,N}(k_d)$. This can be done by determining the quotient

$$\mathbf{P}\{L_n(k_1) = m_1, \dots, L_n(k_d) = m_d\} = \frac{b_{k_1,m_1,k_2,m_2,\dots,k_d,m_d,n,N}}{b_{n,N}},$$

where $b_{k_1,m_1,k_2,m_2,\dots,k_d,m_d,n,N}$ is the (weighted) number of forests in $F(n,N)$ with m_i nodes in stratum k_i for $i = 1, \dots, d$. Therefore define first the generating function (see [6] for a more detailed description)

$$\sum_{m_1,\dots,m_d,n\geq 0} a_{k_1m_1k_2m_2\cdots k_dm_dn}u_1^{m_1}\cdots u_d^{m_d}z^n = y_{k_1}\left(z, u_1 y_{k_2-k_1}\left(z, \dots y_{k_d-k_{d-1}}\left(z, u_d a(z)\right)\dots\right)\right),$$

where $a_{k_1m_1k_2m_2\cdots k_dm_dn}$ is the number of single trees with the above property and

$$y_0(z,u) = u, \qquad y_{i+1}(z,u) = z\varphi(y_i(z,u)), \quad i \geq 0.$$

Forests consisting of N trees can now be described by the Nth power of this function and thus the characteristic function of the joint distribution of $\frac{1}{\sigma\sqrt{Nc_n}}L_{n,N}(k_1), \dots, \frac{1}{\sigma\sqrt{Nc_n}}L_{n,N}(k_d)$ is given by the coefficient

$$\phi_{k_1,\cdots,k_d,n,N}(t_1,\dots,t_d) = \frac{1}{b_{n,N}}[z^n]y_{k_1}\left(z, e^{it_1/\sigma\sqrt{Nc_n}}y_{k_2-k_1}\left(z,\dots y_{k_d-k_{d-1}}\left(z, e^{it_d/\sigma\sqrt{Nc_n}}a(z)\right)\dots\right)^N \tag{2.1}$$

where $[z^n]f(z)$ denotes the coefficient of z^n in the power series of $f(z)$.

In order to extract the desired coefficient we will need some lemmas. First we need the tree function and related functions (see [20] or [11]).

Lemma 2.1. *Let $z_0 = 1/\varphi'(\tau)$ be the point on the circle of convergence of $a(z)$ which lies on the positive real axis. Set $\alpha(z) = z\varphi'(a(z))$ and $\beta(z) = z\varphi''(a(z))$ and assume $\arg(z - z_0) \neq 0$. Then the following local expansions hold:*

$$a(z) = \tau - \frac{\tau\sqrt{2}}{\sigma}\sqrt{1-\frac{z}{z_0}} + O\left(\left|1-\frac{z}{z_0}\right|\right) \qquad \text{as } z \to z_0$$

$$\alpha(z) = 1 - \sigma\sqrt{2}\sqrt{1-\frac{z}{z_0}} + O\left(\left|1-\frac{z}{z_0}\right|\right) \qquad \text{as } z \to z_0$$

$$\beta(z) = \frac{\sigma^2}{\tau} + O\left(\sqrt{1-\frac{z}{z_0}}\right) \qquad \text{as } z \to z_0.$$

The previous two lemmas immediately imply

$$b_{n,N} = \frac{N\tau^N}{\sigma z_0^n\sqrt{2\pi n^3}}\left(\exp\left(-\frac{N^2}{2n\sigma^2}\right) + O\left(\frac{1}{\sqrt{n}}\right)\right) \tag{2.2}$$

We will need an expansion of the bivariate generating function $y_k(z,u)$ as well (see [15], cf. also [6, Lemmas 2.1 and 3.1]).

Lemma 2.2. *Set $w = u - a(z)$, If $w \to 0$ and $z - z_0 \to 0$ in such a way that* $\arg(z - z_0) \neq 0$ *and* $|1 - \sqrt{z - z_0}| \leq 1 + O\left(n^{-1/2}\right)$, *then $y_k(z,u)$ admits the local representation*

$$y_k(z,u) = a(z) + \frac{\alpha^k(z)w}{1 - \frac{\beta(z)}{2\alpha(z)}\frac{1-\alpha^k(z)}{1-\alpha(z)}w + O\left(\left|\frac{1-\alpha^{2k}(z)}{1-\alpha^2(z)}\right||w|^2\right)} \tag{2.3}$$

uniformly for $k = O(1/|w|)$.

With the help of these asymptotic expansions we can prove the convergence of the finite-dimensional distributions to a Gaussian limiting distribution now.

Theorem 2.1. *Let $n \to \infty$, $N = O(\sqrt{n})$ and $c_n \to \infty$ such that $c_n = o(N)$. Moreover, set*

$$X_\kappa := \frac{L_{n,N}(\kappa c_n) - N}{\sigma\sqrt{N c_n}}.$$

Then the joint distribution of $X_{\kappa_1}, \ldots, X_{\kappa_d}$ converges to a centered Gaussian distribution with covariance $\mathbf{Cov}(X_s, X_t) = \min(s,t)$.

Proof. We have to show that the characteristic function of the centered joint distribution of $\frac{1}{\sigma\sqrt{Nc_n}}L_n(k_1), \ldots, \frac{1}{\sigma\sqrt{Nc_n}}L_n(k_d)$ for $k_j = \lfloor \kappa_j c_n \rfloor$, $j = 1, \ldots, d$ satisfies

$$\lim_{n\to\infty} \exp\left(-i\frac{\sqrt{N}}{\sigma\sqrt{c_n}}\sum_{j=1}^d t_j\right)\phi_{k_1,\cdots,k_d,n,N}(t_1,\ldots,t_d) = \exp\left(-\sum_{j=1}^d \frac{\kappa_j t_j^2}{2} - \sum_{j,\ell=1;j<\ell}^d \kappa_j t_j t_\ell\right) \tag{2.4}$$

Therefore we apply Cauchy's integral formula on (2.1) with the integration contour $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4$ where

$$\begin{aligned} \Gamma_1 &= \left\{ z = z_0\left(1 + \frac{x}{n}\right) \middle| \Re x \leq 0 \text{ and } |x| = 1 \right\} \\ \Gamma_2 &= \left\{ z = z_0\left(1 + \frac{x}{n}\right) \middle| \Im x = 1 \text{ and } 0 \leq \Re x \leq n^{1/3} \right\}, \qquad \Gamma_3 = \overline{\Gamma}_2 \\ \Gamma_4 &= \left\{ z : |z| = z_0\left|1 + \frac{\log^2 n + i}{n}\right| \text{ and } \arg\left(1 + \frac{\log^2 n + i}{n}\right) \leq |\arg(z)| \leq \pi \right\}. \end{aligned} \tag{2.5}$$

Let us first study the contribution of $\gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ which will turn out to be the main term. For notational convenience, let us abbreviate the second term in (2.3) by

$$R_k := R_k(z,u) = \frac{\alpha^k w}{1 - \frac{\beta}{\alpha}\frac{1-\alpha^k}{1-\alpha}w + O\left(\left|\frac{1-\alpha^{2k}}{1-\alpha^2}\right||w|^2\right)} \tag{2.6}$$

and let us omit the function arguments z, u and so forth whenever there is no ambiguity. Furthermore, set $u_j = e^{it_j/\sigma\sqrt{Nc_n}}$ and $w_j = (u_j - 1)a$. Since on γ the equation $|1 - \sqrt{-x/n}| = 1 + O\left(n^{-2/3}\right)$ is valid, the assumptions of Lemma 2.2 are fulfilled. Thus we have on γ

$$\begin{aligned} y_{k_1}^N &= y_{k_1}\left(z, u_1 y_{k_2-k_1}\left(z, \ldots y_{k_d-k_{d-1}}\left(z, u_d a(z)\right)\ldots\right)^N\right. \\ &= a^N\left(1 + \frac{\alpha^{k_1}(w_1 + u_1 R_{k_2-k_1})/a}{1 - \frac{\beta}{\alpha}\frac{1-\alpha^{k_1}}{1-\alpha}(w_1 + u_1 R_{k_2-k_1}) + O\left(\frac{1}{N}\right)}\right)^N \end{aligned} \tag{2.7}$$

Here $R_{k_2-k_1} = R_{k_2-k_1}\left(z, u_2 y_{k_2-k_1}\left(z, u_3 y_{k_3-k_2}(z, \ldots y_{k_d-k_{d-1}}\left(z, u_d a(z)\right)\ldots\right)\right.$. Expanding the second factor and using the asymptotic relations

$$u_j = 1 + \frac{it_j}{\sigma\sqrt{Nc_n}} + O\left(\frac{1}{Nc_n}\right) \tag{2.8}$$

$$w_j = \tau\left(\frac{it_j}{\sigma\sqrt{Nc_n}} - \frac{t_j^2}{2Nc_n\sigma^2}\right) + O\left(\frac{1}{N^{3/2}c_n^{3/2}}\right) \tag{2.9}$$

$$\alpha^k = 1 + O\left(\frac{k\sqrt{|x|}}{\sqrt{n}}\right) = 1 + O\left(\frac{c_n\sqrt{|x|}}{\sqrt{n}}\right) \qquad \text{for } k = O(c_n) \tag{2.10}$$

as well as those in Lemma 2.1 yield

$$\begin{aligned}
y_{k_1}^N &= a^N \exp\left(N\alpha^{k_1}\frac{w_1}{a} + N\frac{\alpha^{k_1}u_1R_{k_2-k_1}}{a}\right.\\
&\quad \left.+N\left(\alpha^{k_1-1}\beta\frac{1-\alpha^{k_1}}{1-\alpha} - \frac{\alpha^{2k_1}}{2a^2}\right)\left(w_1^2 + 2u_1w_1R_{k_2-k_1} + u_1^2R_{k_2-k_1}^2\right) + O\left(k_1Nw_1^3\right)\right)\\
&= a^N \exp\left(it_1\sqrt{\frac{N}{c_n\sigma^2}} - \frac{t_1^2}{2c_n\sigma^2} + N\left(1+\frac{it_1}{\sigma\sqrt{Nc_n}}\right)\frac{R_{k_2-k_1}}{\tau}\right.\\
&\quad + \left(\frac{k_1\sigma^2}{2}-\frac{1}{2}\right)\left(-\frac{t_1^2}{c_n\sigma^2} + 2it_1\sqrt{\frac{N}{c_n\sigma^2}}\frac{R_{k_2-k_1}}{\tau} + N\left(1+\frac{it_1}{\sigma\sqrt{Nc_n}}\right)\frac{R_{k_2-k_1}^2}{\tau^2}\right)\\
&\quad \left.+O\left(\frac{k_1}{\sqrt{Nc_n^3}}\right) + O\left(\frac{k_1\sqrt{N}}{\sqrt{nc_n}}\sqrt{|x|}\right)\right) \qquad (2.11)
\end{aligned}$$

Now observe that for $\ell \le d$ we have by (2.6) and (2.7) as well as the asymptotic expansions (2.8)–(2.10)

$$\begin{aligned}
\frac{R_{k_\ell-k_{\ell-1}}}{\tau} &= \frac{\alpha^{k_\ell-k_{\ell-1}}(w_\ell + u_{\ell+1}R_{k_{\ell+1}-k_\ell})/\tau}{1-\frac{\beta}{\alpha}\frac{1-\alpha^{k_\ell-k_{\ell-1}}}{1-\alpha}\left(w_\ell + u_\ell R_{k_{\ell+1}-k_\ell}\right) + O\left(\frac{k}{Nc_n}\right)}\\
&= \left(\frac{it_\ell}{\sigma\sqrt{Nc_n}} - \frac{t_\ell^2}{2Nc_n\sigma^2} + \left(1+\frac{it_\ell}{\sigma\sqrt{Nc_n}}\right)\frac{R_{k_{\ell+1}-k_\ell}}{\tau} + O\left(\frac{1}{\sqrt{N^3c_n^3}}\right) + O\left(\sqrt{\frac{|x|}{n}}\right)\right)\\
&\quad\times\left(1+\frac{(k_\ell-k_{\ell-1})\sigma^2}{2}\left(\frac{it_2}{\sigma\sqrt{Nc_n}} + \frac{R_{k_{\ell+1}-k_\ell}}{\tau}\right) + O\left(\frac{1}{N}\right)\right)
\end{aligned}$$

and $R_{k_\ell-k_{\ell-1}} \equiv 0$ for $\ell > d$. Thus $R_{k_\ell-k_{\ell-1}} = O\left(1/\sqrt{Nc_n}\right)$ and in particular

$$R_{k_d-k_{d-1}} = \left(\frac{it_d\tau}{\sigma\sqrt{Nc_n}} - \frac{t_d^2\tau}{2Nc_n\sigma^2}\right)\left(1+\frac{it_d(k_d-k_{d-1})\sigma^2}{2\sigma\sqrt{Nc_n}}\right) + O\left(\frac{1}{\sqrt{N^3c_n}}\right) + O\left(\sqrt{\frac{|x|}{n}}\right).$$

Plugging the expressions for $R_{k_2-k_1}$ into (2.11) yields

$$\begin{aligned}
y_{k_1}^N &= a^N\exp\left(it_1\sqrt{\frac{N}{c_n\sigma^2}} - \frac{t_1^2}{2c_n\sigma^2} + it_2\sqrt{\frac{N}{c_n\sigma^2}} - \frac{t_2^2}{2c_n\sigma^2} - \frac{t_1t_2}{\sigma^2c_n}\right.\\
&\quad + N\left(1+\frac{it_1}{\sigma\sqrt{Nc_n}} + \frac{it_2}{\sigma\sqrt{Nc_n}}\right)\frac{R_{k_3-k_2}}{\tau} + N\frac{(k_2-k_1)\sigma^2}{2}\left(\frac{it_2}{\sigma\sqrt{Nc_n}} + \frac{R_{k_3-k_2}}{\tau}\right)^2\\
&\quad + \left(\frac{k_1\sigma^2}{2}-\frac{1}{2}\right)\left(-\frac{t_1^2}{c_n\sigma^2} - \frac{2t_1t_2}{c_n\sigma^2} + 2it_1\sqrt{\frac{N}{c_n\sigma^2}}\frac{R_{k_3-k_2}}{\tau} + N\left(\frac{it_2}{\sigma\sqrt{Nc_n}} + \frac{R_{k_3-k_2}}{\tau}\right)^2\right)\\
&\quad \left.+O\left(\frac{1}{\sqrt{Nc_n}}\right)\right) + O\left(\sqrt{\frac{Nc_n}{n}}\sqrt{|x|}\right)
\end{aligned}$$

and then by substituting $R_{k_3-k_2}$, $R_{k_4-k_3}$, …, step by step we arrive at

$$\begin{aligned}
y_{k_1}^N &= a^N\exp\left(i(t_1+\cdots+t_d)\sqrt{\frac{N}{c_n\sigma^2}} - \frac{k_1t_1^2+\cdots+k_dt_d^2}{2c_n} - \sum_{j_1=1}^{d-1}\sum_{j_2=j_1+1}^{d}\frac{k_{j_1}t_{j_1}t_{j_2}}{c_n}\right.\\
&\quad \left.+O\left(\frac{1}{\sqrt{Nc_n}}\right) + O\left(\sqrt{\frac{Nc_n}{n}}\sqrt{|x|}\right)\right) \qquad (2.12)
\end{aligned}$$

If we substitute $z = z_0(1+x/n)$ on γ, we get

$$\frac{1}{b_{n,N}}[z^n]y_{k_1}^N = \frac{1}{2\pi i b_{n,N}z_0^n n}\left(\int_\gamma y_{k_1}^N e^{-x}\,dx\left(1+O\left(\frac{1}{n^{1/3}}\right)\right)\right) + O\left(\frac{1}{b_{n,N}}\int_{\Gamma_4} y_{k_1}^N\frac{dz}{z^{n+1}}\right). \quad (2.13)$$

Moreover, observe that for any $M > 0$ we have

$$\frac{1}{2\pi i}\int_{\gamma'} e^{-\lambda\sqrt{-x}-x}\,dx = \frac{\lambda}{2\sqrt{\pi}}e^{-\lambda^2/4} + O\left(e^{-M}\right)$$

for $\gamma' = \{x:\ |x|=1, \Re x \le 0\} \cup \{x : 0 \le \Re x \le M, \Im x = \pm 1\}$, as can be easily seen by substituting $u^2 = x$. Thus, since $\sqrt{Nc_n/n} = o(1)$, the error terms in (2.12) are negligibly small, and hence the first term in (2.13) in conjunction with (2.12) yields (2.4).

So let us estimate the second term in (2.13). By Taylor's theorem we have

$$y_{k_d-k_{d-1}}(z, ua(z))^N = a^N\left(1 + \alpha(z)^{k_d-k_{d-1}}(u_d-1)a(z) + O\left((u_d-1)^2\right)\right)^N.$$

Since we required $\zeta = 1$ (see Theorem 1.2), we get

$$\max_{z\in\Gamma_4}|\alpha(z)| = |\alpha(\tilde z)| \quad \text{and} \quad \max_{z\in\Gamma_4}|a(z)| = |a(\tilde z)|,$$

where $\tilde z \in \gamma \cap \Gamma_4$. There the local expansions of Lemma 2.1 are still valid and hence $|\alpha| < 1$ and $|a| < \tau$. Consequently, with $u = e^{it_d/\sigma\sqrt{Nc_n}}$ we get

$$\left|y_{k_d-k_{d-1}}(z, ua(z))^N\right| = \tau^N \exp\left(\sqrt{\frac{N}{c_n}} + O\left(\frac{1}{c_n^2}\right)\right).$$

Inserting this into $y_{k_\ell-k_{\ell-1}}$, $\ell = 2,\dots,d-1$, and y_{k_1} and arguing as above, we get the same estimate for $y_{k_1}^N$. Finally, using $|z|^{-n} \sim z_0^{-n}\exp\left(-n^{1/3}\right)$ implies the existence of some positive constant C such that

$$\int_{\Gamma_4}\left|y_{k_1}^N\right|\frac{|dz|}{|z^{n+1}|} = O\left(\exp\left(-n^{1/3} + Cn^{1/4}\right)\right)$$

which is exponentielly small compared to the integral over γ and the proof is complete. □

3. Tightness

In order to complete the proof of Theorem 1.2 we have to show that the sequence of random variables $L_{n,N}(c_n t)/\sigma\sqrt{Nc_n}$, $t \ge 0$, is tight in C$[0,\infty)$. By [18, Theorem 4.10] it suffices to establish tightness in C$[0,T]$. Thus by [1, Theorem 12.3] we only have to show that $L_{n,N}(0)$ is tight, which is obviously true, and that there exist constants $\alpha > 1$, $\beta \ge 0$, and $C > 0$ such that

$$\mathbf{P}\left\{|L_{n,N}(\rho c_n) - L_{n,N}((\rho+\theta)c_n)| \ge \varepsilon\sigma\sqrt{Nc_n}\right\} \le C\frac{\theta^\alpha}{\varepsilon^\beta} \tag{3.1}$$

This inequality follows from the following theorem.

Theorem 3.1. *There exists a constant $C > 0$ such that for all $r, h \ge 0$ and for $N = O(\sqrt{n})$ the following inequality holds:*

$$\mathbf{E}\left|L_{n,N}(r+h) - L_{n,N}(r)\right|^4 \le CN^2h^2 \tag{3.2}$$

In order to show this inequality we will investigate a more general situation. First, observe that the left-hand side can be represented by the coefficient of a proper generating function. In fact we have

$$\mathbf{E}\left(L_{n,N}(r) - L_{n,N}(r+h)\right)^4 = \frac{1}{b_{n,N}}[z^n]H_{rh}^{(4)}(z),$$

where

$$\begin{aligned} H_{rh}^{(4)}(z) &= \left(u\frac{\partial}{\partial u}\right)^4 y_r(z, uy_h(z, u^{-1}a(z)))^N\Big|_{u=1} \\ &= \left[\left(\frac{\partial}{\partial u} + 7\frac{\partial^2}{\partial u^2} + 6\frac{\partial^3}{\partial u^3} + \frac{\partial^4}{\partial u^4}\right) y_r(z, uy_h(z, u^{-1}a(z)))^N\right]_{u=1}. \end{aligned} \tag{3.3}$$

Since $b_{n,N} \sim (N\tau^N/\sqrt{2\pi\sigma^2})z_0^{-n}n^{-3/2}\exp(-N^2/2n\sigma^2)$ (see (2.2)), (3.2) is valid if

$$[z^n]H_{rh}^{(4)}(z) = O\left(\frac{N^3\tau^N h^2}{z_0^n n^{3/2}}\right) \tag{3.4}$$

holds uniformly for $r, h \geq 0$. We will estimate this coefficient by analyzing the function $H_{rh}(z)$ and using Flajolet and Odlyzko's [13] transfer lemma:

Lemma 3.1. *Let $F(z)$ be analytic in Δ defined by*

$$\Delta = \{z : |z| < z_0 + \eta,\ |\arg(z - z_0)| > \vartheta\},$$

where z_0 and η are positive real numbers and $0 < \vartheta < \pi/2$. Furthermore suppose that there exists a real number β such that

$$F(z) = O\left((1 - z/z_0)^{-\beta}\right) \qquad (z \in \Delta).$$

Then

$$[z^n]F(z) = O\left(z_0^{-n} n^{\beta-1}\right).$$

Set $Y_{rh}(z, u) = y_r(z, u y_h(z, u^{-1}a(z)))$. We analyze the derivatives of $Y_{rh}(z, u)$ with respect to u in the next lemma.

Lemma 3.2. *Let Δ be the domain defined in Lemma 3.1. Then there exists a finite index set I and functions $\alpha_{i\ell rh}(z)$ such that for all $\ell > 0$*

$$\frac{\partial^\ell}{\partial u^\ell} Y_{rh}(z, 1) = \sum_{i \in I} \alpha_{i\ell rh}(z), \tag{3.5}$$

where the functions $\alpha_{ilrh}(z)$ satisfy for $z \in \Delta$

$$\alpha_{i\ell rh}(z) = O\left(|\alpha^r| \left|\frac{1-\alpha^r}{1-\alpha}\right|^{\mu_1} \left|\frac{1-\alpha^h}{1-\alpha}\right|^{\mu_2} |1-\alpha|^{\mu_3}\right)$$

for some nonnegative integers μ_1, μ_2, μ_3 with $\mu_1 + \mu_2 - \mu_3 \leq \ell - 1$.

Proof. First compute the first few derivatives of $\frac{\partial^\ell}{\partial u^\ell} y_r(z, a(z))$,

$$\frac{\partial y_r}{\partial u}(z, a(z)) = \alpha^r, \quad \frac{\partial^2 y_r}{\partial u^2}(z, a(z)) = \frac{\beta}{\alpha}\alpha^r \frac{1-\alpha^r}{1-\alpha},$$
$$\frac{\partial^3 y_r}{\partial u^3}(z, a(z)) = \frac{\tilde\beta}{\alpha}\alpha^r \frac{1-\alpha^{2r}}{1-\alpha^2} + 3\frac{\beta^2}{\alpha}\alpha^r \frac{(1-\alpha^r)(1-\alpha^{r-1})}{(1-\alpha)(1-\alpha^2)},$$

where $\beta = z\varphi''(a(z))$ and $\tilde\beta = z\varphi'''(a(z))$. Noticing that Faà di Bruno's formula (see e.g. [5]) gives

$$\frac{\partial^\ell y_r}{\partial u^\ell}(z, 1) = \sum_{\sum_{i=1}^{\ell-1} i k_i = \ell} \frac{\ell!}{k_1! \cdots k_{\ell-1}!} z\varphi^{(k_1+\cdots+k_{\ell-1})}(a(z)) \prod_{j=1}^{\ell-1} \left(\frac{1}{j!}\frac{\partial^j y_{r-1}}{\partial u^j}\right)^{k_j} + \alpha(z)\frac{\partial^\ell y_{r-1}}{\partial u^\ell}(z, 1).$$

and that hence $\frac{\partial^\ell}{\partial u^\ell} y_r(z, 1)$ is the solution of an inhomogeneous first order linear recurrence, the estimate

$$\frac{\partial^\ell}{\partial u^\ell} y_r(z, 1) = O\left(|\alpha^r| \left|\frac{1-\alpha^r}{1-\alpha}\right|^{l-1}\right) \tag{3.6}$$

is now easily proved by induction.

Now, employing again Faà di Bruno's formula, this time to Y_{rh}, yields

$$\begin{aligned}\frac{\partial^\ell Y_{rh}}{\partial u^\ell}(z, 1) &= \sum_{\sum i k_i = \ell} \frac{\ell!}{k_1! \cdots k_\ell!} \left(\frac{\partial}{\partial u}\right)^{k_1+\cdots+k_\ell} y_r(z, a(z)) \prod_{j=1}^{\ell} \left(\frac{1}{j!}\left(\frac{\partial}{\partial u}\right)^j \left(u y_h\left(z, \frac{a(z)}{u}\right)\right)\right)^{k_j} \\ &= \sum_{\sum i k_i = \ell} \frac{\ell!}{k_1! \cdots k_\ell!} \left(\frac{\partial}{\partial u}\right)^{k_1+\cdots+k_\ell} y_r(z, a(z)) \prod_{i=1}^{\ell} (A_i - A_{i-1})^{k_i}\end{aligned}$$

where

$$A_i = (-1)^i \sum_{\sum j m_j = i} \frac{i!}{m_1! \cdots m_i!} \left(\frac{\partial}{\partial u}\right)^{m_1+\cdots+m_i} y_h(z, 1).$$

Since

$$\frac{\partial}{\partial u} u y_h(z, a(z)/u)\bigg|_{u=1} = 1 - \alpha^h$$

we get by (3.6)

$$\left(\frac{\partial}{\partial u}\right)^{k_1+\cdots+k_\ell} y_r(z,1)\left(\frac{\partial}{\partial u} u y_h(z,a(z)/u)\bigg|_{u=1}\right)^{k_1}\prod_{j=2}^{\ell}\left(\frac{1}{j!}\left(\frac{\partial}{\partial u}\right)^j u y_h(z,a(z)/u)\bigg|_{u=1}\right)^{k_j}$$
$$= O\left(|\alpha^r|\left|\frac{1-\alpha^r}{1-\alpha}\right|^{k_1+\cdots+k_\ell-1}|1-\alpha|^{k_1}\left|\frac{1-\alpha^h}{1-\alpha}\right|^{k_1+\sum(i-1)k_i}\right).$$

Note that we omitted a factor α^h coming from y_h. This is justified since $|\alpha| < 1$ in Δ. So we could also neglect the factor α^r but this one is needed in the sequel.

If we set $\mu_1 = \sum_i k_i - 1$, $\mu_2 = k_1 + \sum_i (i-1)k_i$, and $\mu_3 = k_1$, then obviously $\mu_1 + \mu_2 - \mu_3 = \ell - 1$ which yields (3.5) and completes the proof. □

For the tightness inequality (3.2) we need the derivatives of Y_{rh}^N. These are investigated in the next lemma.

Lemma 3.3. *There exist bounded functions $\beta_{\ell,M}$ on Δ such that*

$$\left(\frac{\partial}{\partial u}\right)^M Y_{rh}^N = \sum_{\ell=1}^{M}\beta_{\ell,M}(z)N^\ell a(z)^{n-\ell}\prod_i\left(\frac{\partial^i Y_{rh}}{\partial u^i}\right)^{c_i}$$

where the c_i satisfy $\sum_i (i-1)c_i \le M - \ell$.

Proof. Faà di Bruno's formula yields

$$\left(\frac{\partial}{\partial u}\right)^M Y_{rh}^N = \sum_{\sum ik_i = M}\frac{M!}{k_1!\cdots k_M!}N(N-1)\cdots\left(N - \sum_i k_i + 1\right)Y_{rh}^{N-\sum_i k_i}\prod_{j=1}^{M}\left(\frac{1}{j!}\frac{\partial^j Y_{rh}}{\partial u^j}\right)^{k_j}$$

and because of $Y_{rh}(z,1) = a(z)$ and $\sum_i (i-1)k_i = M - \ell$ we are done. □

Now we are able to prove Theorem 3.1:

Proof of Theorem 3.1. By Lemmas 3.2 and 3.3, all terms of $H_{rh(z)}^{(4)}$ are bounded by functions of the form

$$N^\ell|\alpha^r||a|^N\left|\frac{1-\alpha^r}{1-\alpha}\right|^{\mu_1}\left|\frac{1-\alpha^h}{1-\alpha}\right|^{\mu_2}|1-\alpha|^{\mu_3}$$
$$= O\left(N^\ell\tau^N|\alpha^r||1-\alpha^r|^{\mu_1}|1-\alpha^h|^{\mu_2-d}h^d|1-\alpha|^{\mu_3-\mu_1-\mu_2+d}\right), \tag{3.7}$$

where $\mu_3 - \mu_1 - \mu_2 \ge 2d - \ell$ and $d = 2$. Thus by Lemma 3.1 and the fact that $N = O(\sqrt{n})$ we get

$$[z^n]H_{rh}(z) = O\left(\frac{\tau^N N^\ell h^2}{z_0^n n^{1+(\ell-2)/2}}\right) = O\left(\frac{\tau^N N^3 h^2}{z_0^n n^{3/2}}\right)$$

as desired. □

4. The profile in the range $c_n/\sqrt{n} \to \eta > 0$ and the width of random forests

Equations (1.2) and (1.3) characterize the distributions of $L_{n,N}(k)$ in the range $k \approx \sqrt{n}$ by a limiting process given implicitly by a stochastic differential equation and by conditioning a well known process. The same ideas as in Section 2 allow us to make the distributions more explicit, leading to a representation in terms of an integral transform for the characteristic functions of the finite-dimensional distributions. Starting again with (2.3) we get as above

$$y_{k_1}^N = a^N\left(1 + \frac{\alpha^{k_1}(w_1 + u_1 R_{k_2-k_1})/a}{1 - \frac{\beta}{\alpha}\frac{1-\alpha^{k_1}}{1-\alpha}(w_1 + u_1 R_{k_2-k_1}) + O\left(\frac{1}{N}\right)}\right)^N.$$

Insert the asymptotic approximations for $u_j = e^{2it_j/\sigma\sqrt{n}}$, w_j, β,

$$a^N = \tau^N \exp\left(-\frac{N\sqrt{-2x}}{\sigma\sqrt{n}} + O\left(\frac{N}{n}\right)\right),$$
$$\alpha^k = \exp\left(-2\kappa\sqrt{-2x} + O\left(\frac{|x|}{\sqrt{n}}\right)\right),$$

for $k = 2\kappa\sqrt{n}/\sigma$ we finally arrive at

$$y_{k_1}^N = \tau^N \exp\Bigg(-\frac{N\sqrt{-2x}}{\sigma\sqrt{n}} + \frac{N\sqrt{-x}\exp\left(-\kappa_1\sqrt{-2x}\right)(2it_1/\sigma\sqrt{n} + R_{k_2-k_1}/\tau)}{\sqrt{-x}\exp\left(\kappa_1\sqrt{-2x}\right) - (it_1\sqrt{2} + (\sigma\sqrt{n}/\tau\sqrt{2})R_{k_2-k_1}\sinh\left(\kappa_1\sqrt{-2x}\right)} + O\left(\frac{N}{n}\right)\Bigg)$$

Error estimation for Γ_4 works similar as in Section 2 and therefore we get the following theorem.

Theorem 4.1. *Assume $2N/\sigma\sqrt{n} \to \alpha > 0$. Furthermore, let $k_j = 2\kappa_j\sqrt{n}/\sigma$. Then the characteristic function of the joint distribution of $\frac{2}{\sigma\sqrt{n}}L_{n,N}(k_1), \ldots, \frac{2}{\sigma\sqrt{n}}L_{n,N}(k_d)$, satisfies*

$$\tilde{\phi}_{k_1,\cdots,k_d,n,N}(t_1,\ldots,t_d) = \frac{\sqrt{2}}{i\alpha\sqrt{\pi}}\int_{\gamma'} \exp\Big(-x - \alpha\sqrt{-x/2} + \Psi_{\kappa_1}(x, it_1 + \Psi_{\kappa_2-\kappa_1},(\ldots\Psi_{\kappa_{p-1}-\kappa_{p-2}}(x, it_{p-1} + \Psi_{\kappa_d-\kappa_{d-1}}(x, it_d))\cdots)\Big)\,dx$$

with

$$\Psi_\kappa(x,t) = \frac{\alpha t\sqrt{-x}e^{-\kappa\sqrt{-2x}}}{\sqrt{-x}e^{\kappa\sqrt{-2x}} - t\sqrt{2}\sinh\left(\kappa\sqrt{-2x}\right)}.$$

Now we turn to the width: The structure of the functions in the previous section allows us to prove an even tighter bound for the moments of $L_{n,N}(r+h) - L_{n,N}(r)$ with the help of the following lemma (cf. [7] and [16, Lemma 3.5]).

Lemma 4.1. *Let $f(z)$ and $g(z)$ be analytic functions in Δ which satisfy*

$$|f(z)| \le \exp\left(-C\sqrt{\left|1 - \frac{z}{z_0}\right|}\right), \quad z \in \Delta,$$
$$g(z) = 1 - D\sqrt{1 - \frac{z}{z_0}} + O\left(1 - \frac{z}{z_0}\right), \quad z \in \Delta,$$

for some positive constants C, D. Then for any fixed ℓ there exists a constant $C' > 0$ such that

$$[z^n]\frac{f(z)^r}{(1-g(z))^\ell} = O\left(e^{-C'r/\sqrt{n}}n^{(\ell-2)/2}\right)$$

uniformly for all $r, n \ge 0$.

Theorem 4.2. *For every fixed positive integer d there exist constants c_1, c_2 such that for every $r, h > 0$*

$$\mathbf{E}\left|L_{n,N}(r) - L_{n,N}(r+h)\right|^{2d} \le c_1 e^{-c_2 r/\sqrt{n}} h^d n^{d/2}. \tag{4.1}$$

The constants c_1 and c_2 are independent of n and N, provided that $N = O\left(\sqrt{n}\right)$

Proof. Since

$$\left(u\frac{\partial}{\partial u}\right)^{2d} = \sum_{k=1}^{2d} s_{2d,k}u^k\left(\frac{\partial}{\partial u}\right)^k,$$

where $s_{n,k}$ are the Stirling numbers of the second kind, we can apply Lemmas 3.2 and 3.3 directly to $H_{rh}^{(2d)}$ and get (3.7). Keep in mind that $\alpha(z)$ admits a representation like $g(z)$ in Lemma 4.1

due to Lemma 2.1 and thus there exists a constant $C > 0$ such that in Δ the inequality $|\alpha(z)| \le \exp\left(-C\sqrt{|1-z/z_0|}\right)$ holds. Hence we obtain

$$\mathbf{E}\,|L_{n,N}(r) - L_{n,N}(r+h)|^{2d} = \frac{1}{b_{n,N}}[z^n]\left(u\frac{\partial}{\partial u}\right)^{2d} Y_{rh}(z,1) = O\left(\frac{1}{b_{n,N}} e^{-c_2 r/\sqrt{n}} \frac{N\tau^N h^d n^{(d-3)/2}}{z_0^n}\right)$$

and this immediately implies (4.1). □

By [8, Theorem 1], this property in conjunction with the fact that there exists a $t \ge 0$ (in fact we can choose $t = 0$) such that $|\sup_n \mathbf{E}(L_{n,N}(t)/\sqrt{n})^k| < \infty$ for all $k \ge 0$ and for $2N/\sigma\sqrt{n} \in [\alpha - f_n, \alpha + f_n]$ with $\alpha > 0$, $f_n \to 0$ implies that the sequence $(L_{n,N}(\cdot)/\sqrt{n})_{n>0}$ is polynomially convergent in the sense of Drmota and Marckert [8], i.e.,

$$\mathbf{E}F\left(\left(\frac{2}{\sigma\sqrt{n}}L_{n,N}\left(\frac{2\sqrt{n}\kappa}{\sigma}\right)_{\kappa\ge 0}\right)\right) \to \mathbf{E}F((X_{\alpha,\kappa})_{\kappa\ge 0}) \qquad n \to \infty$$

for every functional F satisfying $|F(f)| \le C(1+\|f\|_\infty)^k$ for some constants $C, k > 0$. If we choose, in particular, $F(f) = \|f\|_\infty^d$ and $F(f) = f$, respectively, then Theorem 1.3 is proved.

Remark 1. Note that Drmota and Marckert [8] studied the concept of polynomial convergence only for processes with compact support. But since the height of random forests is asymptotically a.s. bounded by $c\sqrt{n}$, it suffices to study truncated processes with arbitrarily large but compact support and then argue in the same way as in [7, Lemma 5]. Thus [8, Theorem 1] is applicable in this case as well.

Remark 2. Note that the connection between random forests and conditioned Brownian bridge can be used to compute functionals of the latter one. Recently, this has been done for the conditioned Brownian bridge area, see [12, 17, 2]. Thus this approach can help us to derive expressions for the moments of $X_{\alpha,v}$. This will be done in a forthcoming paper with G. Louchard.

References

[1] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.

[2] P. Chassaing and G. Louchard *Reflected Brownian bridge area conditioned on its local time at the origin*, submitted, 2001, `http://www.ulb.ac.be/di/mcs/louchard/`.

[3] P. Chassaing and J.-F. Marckert, Parking functions, empirical processes, and the width of rooted labeled trees, *Electronic J. Combinatorics 8(1)*, 2001, R14.

[4] I. A. Cheplyukova, Limit distributions for the number of vertices in layers of a random forest, *Diskret. Mat. 9*, No. 4, 150–157, 1997.

[5] L. Comtet, *Advanced Combinatorics*, Reidel, Dordrecht, Netherlands, 1974.

[6] M. Drmota and B. Gittenberger, On the profile of random trees, *Random Structures and Algorithms 10*, 421–451, 1997.

[7] M. Drmota and B. Gittenberger, The width of Galton-Watson trees, *submitted*, see `http://www.geometrie.tuwien.ac.at/drmota/width.ps.gz`

[8] M. Drmota and J. F. Marckert, Reinforced weak convergence of stochastic processes, submitted, 2001, see `http://www.geometrie.tuwien.ac.at/drmota/reinf.ps.gz`.

[9] M. Drmota and M. Soria, Marking in combinatorial constructions: generating functions and limiting distributions, *Theoretical Comp. Science 144*, 67–99, 1995.

[10] I. A. Egorova, The distribution of vertices in strata of plane planted forests, in *Probabilistic methods in discrete mathematics (Petrozavodsk, 1996)*, 179–188, VSP, Utrecht, 1997.

[11] P. Flajolet and A. M. Odlyzko, Random mapping statistics, In *Advances in Cryptology* (1990), J.-J. Quisquater and J. Vandewalle, Eds., vol. 434 of *Lecture Notes in Computer Science,* Springer Verlag, pp. 329-354. Proceedings of Eurocrypt '89, Houtalen, Belgium, April 1989.

[12] P. Flajolet, P. Poblete, and A. Viola, On the analysis of linear probing hashing, *Algorithmica 22,* 490–515, 1998.

[13] P. Flajolet and A. M. Odlyzko, Singularity analysis of generating functions, *SIAM J. on Discrete Math. 3,* No. 2, 216–240, 1990.

[14] P. Flajolet and J. S. Vitter, Average-Case Analysis of Algorithms and Data Structures, in *Handbook of Theoretical Computer Science,* J. van Leeuwen, Ed., vol. A: Algorithms and Complexity. North Holland, 1990, ch. 9, pp. 431–524.

[15] B. Gittenberger, Convergence of branching processes to the local time of a Bessel process, *Random Structures and Algorithms 13*, 423–438, 1998.

[16] B. Gittenberger, On the contour of random trees, *SIAM J. Discrete Math. 12,* No. 4, 434–458, 1999.

[17] S. Janson, Asymptotic distributions for the cost of linear probing hashing, *Random Structures and Algorithms,* to appear, `http://www.math.uu.se/svante/papers/`.

[18] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus,* Springer, New York, 1988.

[19] V. F. Kolchin, *Random Mappings,* Optimization Software, New York, 1986.

[20] A. Meir and J. W. Moon, On the altitude of nodes in random trees, *Canadian Journal of Mathematics 30*, 997–1015, 1978.

[21] Yu. L. Pavlov, Limit distributions of the height of a random forest, *Theory Probab. Appl. 28*, 471–480, 1983.

[22] Yu. L. Pavlov, Distributions of the number of vertices in strata of a random forest, *Theory Probab. Appl. 33*, 96–104, 1988.

[23] Yu. L. Pavlov, Limit distributions of the height of a random forest consisting of plane rooted trees, *Discr. Math. Appl. 4*, 73–88, 1994 (translated from *Diskret. Mat. 6*, 137–154, 1994).

[24] Yu. L. Pavlov, Asymptotic behavior of the height of a random forest. (Russian) *A collection of papers, No. 1* (Russian), 4–17, 139, Ross. Akad. Nauk, Karel. Nauchn. Tsentr, Petrozavodsk, 1994.

[25] Yu. L. Pavlov, Random forests, in *Probabilistic methods in discrete mathematics (Petrozavodsk, 1996)*, 11–18, VSP, Utrecht, 1997.

[26] Yu. L. Pavlov and I. A. Cheplyukova, Limit distributions of the number of vertices in strata of a simply generated forest, *Discr. Math. Appl. 9*, 137–154, 1999 (translated from *Diskret. Mat. 11*, 97–112, 1999).

[27] J. Pitman, The SDE solved by local times of a Brownian excursion or bridge derived from the height profile of a random tree, *Ann. Probab. 27*, 261–283, 1999.