DISTANCES BETWEEN PAIRS OF VERTICES AND VERTICAL PROFILE IN CONDITIONED GALTON–WATSON TREES

LUC DEVROYE AND SVANTE JANSON

Abstract. We consider a conditioned Galton–Watson tree and prove an estimate of the number of pairs of vertices with a given distance, or, equivalently, the number of paths of a given length.

We give two proofs of this result, one probabilistic and the other using generating functions and singularity analysis.

Moreover, the latter proof yields a more general estimate for generating functions, which is used to prove a conjecture by Bousquet–Mélou and Janson [5], saying that the vertical profile of a randomly labelled conditioned Galton–Watson tree converges in distribution, after suitable normalization, to the density of ISE (Integrated Superbrownian Excursion).

1. Introduction and results

Let T_n be a conditioned Galton–Watson tree, i.e., the random rooted tree $\mathcal{T}$ obtained as the family tree of a Galton–Watson process with some given offspring distribution ξ, conditioned on the number of vertices $|\mathcal{T}| = n$. We will always assume that

$$\mathbb{E}\,\xi = 1 \qquad \text{and} \qquad 0 < \sigma^2 := \operatorname{Var}\xi < \infty. \tag{1.1}$$

In other words, the Galton–Watson process is critical and with finite variance, and $\mathbb{P}(\xi = 1) < 1$. (Note that this entails $0 < \mathbb{P}(\xi = 0) < 1$.) It is well-known that this assumption is without essential loss of generality, and that the resulting random trees are essentially the same as the simply generated families of trees introduced by Meir and Moon [13]. The importance of this construction lies in that many combinatorially interesting random trees are of this type, for example (uniformly chosen) random plane (= ordered) trees, random unordered labelled trees (Cayley trees), random binary trees, and (more generally) random d-ary trees. For further examples see e.g. Aldous [1] and Devroye [6].

Date: December 17, 2008; revised September 28, 2009.

2000 *Mathematics Subject Classification.* 60C05; 05C05.

Key words and phrases. Random Galton–Watson tree, paths of given length, vertical profile, probabilistic analysis of algorithms, branching process.

CR Categories. 3.74, 5.25, 5.5.

The first author's research was sponsored by NSERC Grant A3456.

We consider only n such that T_n exists, i.e., such that $\mathbb{P}(|\mathcal{T}| = n) > 0$. The *span* of ξ is defined to be the largest integer d such that $\xi \in d\mathbb{Z}$ a.s. If the span of ξ is d, then T_n exists only for $n \equiv 1 \pmod d$, and it exists for all large such n.

For our main theorem, we assume that we are given a further random variable η. Given a rooted tree τ, we take an independent copy η_e of η for every edge $e \in \tau$. We give each vertex v the label L_v obtained by summing η_e for all e in the path from the root o to v. (Thus, $L_o = 0$.) We assume that

$$\mathbb{E}\,\eta = 0 \qquad \text{and} \qquad 0 < \sigma_\eta^2 := \operatorname{Var}\eta < \infty. \tag{1.2}$$

We further assume that

$$\eta \text{ is integer valued and with span 1;} \tag{1.3}$$

thus all labels are integers, and all integers are possible labels.

We let $X(j;\tau)$ be the number of vertices in τ with label j; the sequence $(X(j;\tau))_{j=-\infty}^{\infty}$ is the *vertical profile* of the labelled tree.

For the random tree T_n, we assume that the variables η_e are independent of T_n. The vertical profile $(X(j;T_n))$ then is a random function defined for $j \in \mathbb{Z}$. We write $X_n(j) := X(j;T_n)$, and extend the domain of X_n to $\mathbb{R}$ by linear interpolation between the integer points; thus X_n is a random continuous function on $\mathbb{R}$. Our main theorem says that this function X_n, suitable normalized, converges in distribution in the space $C_0(\mathbb{R})$ of continuous functions on $\mathbb{R}$ that tend to 0 at $\pm\infty$; we equip $C_0(\mathbb{R})$ with the usual uniform topology defined by the supremum norm. Let, further, f_{ISE} denote the density of the random measure ISE introduced by Aldous [3]; f_{ISE} is a random continuous function with (random) compact support, see Bousquet–Mélou and Janson [5, Theorem 2.1].

Theorem 1.1. *With the assumptions* (1.1), (1.2), *and* (1.3), *let* $\gamma := \sigma_\eta^{-1}\sigma^{1/2}$. *Then, as* $n \to \infty$,

$$\frac{1}{n}\gamma^{-1}n^{1/4}X_n\bigl(\gamma^{-1}n^{1/4}\cdot\bigr) \xrightarrow{\mathrm{d}} f_{\mathrm{ISE}}(\cdot), \tag{1.4}$$

in the space $C_0(\mathbb{R})$ *with the usual uniform topology. Equivalently,*

$$n^{-3/4}X_n\bigl(n^{1/4}\cdot\bigr) \xrightarrow{\mathrm{d}} \gamma f_{\mathrm{ISE}}(\gamma\,\cdot). \tag{1.5}$$

Hence, if $n \to \infty$ *and* $j_n/n^{1/4} \to x$, *where* $-\infty < x < \infty$, *then*

$$n^{-3/4}X(j_n;T_n) \xrightarrow{\mathrm{d}} \gamma f_{\mathrm{ISE}}(\gamma x). \tag{1.6}$$

Note that the random functions on the left and right hand sides of (1.4) and (1.5) are density functions, i.e., non-negative functions with integral 1. The limit law in (1.6) is characterized in [4] by a formula for its Laplace transform.

Theorem 1.1 was conjectured in [5], and proved there in two special cases, viz. when ξ has the Geometric distribution Ge(1/2) and thus T_n is a random ordered tree, and η is uniformly distributed on either $\{-1,1\}$ or $\{-1,0,1\}$. Moreover, it was shown there [5, Remark 3.7] that the proof given in [5] applies generally under the assumptions above, provided the following estimate holds. (We let in this paper $C_1, C_2, \dots$ and $c_1, c_2, \dots$ denote various positive constants that may depend on (the distribution of) ξ, and sometimes on η, but not on n, k and other variables unless explicitly stated.)

Lemma 1.2. *Under the assumptions above, there exists a constant C_1 such that for all $n \geq 1$ and $t \in [-\pi, \pi]$,*

$$\mathbb{E}\left|\frac{1}{n}\sum_j X(j;T_n)e^{\mathrm{i}jt}\right|^2 \leq \frac{C_1}{1+nt^4}. \tag{1.7}$$

We prove Lemma 1.2, and thus Theorem 1.1, in Section 3, using Theorem 1.6 below. Before presenting that theorem, we consider a simpler version, and further results that we prove by the same method.

For an arbitrary rooted tree τ, let $P_k(\tau)$, $k \geq 1$, be the number of (unordered) pairs of vertices $\{v, w\}$ in τ such that the distance $d(v,w) = k$; equivalently, $P_k(\tau)$ is the number of paths of length k in τ. Our next result is an estimate, uniform in all k and n, of the expectation of this number $P_k(T_n)$ for a conditioned Galton–Watson tree T_n.

Recall that we tacitly assume (1.1). (But no stronger moment condition.)

Theorem 1.3. *There exists a constant C_2 such that for all $k \geq 1$ and $n \geq 1$, $\mathbb{E}\, P_k(T_n) \leq C_2 nk$.*

One way to interpret this result is that the expected number of vertices of distance k from a randomly chosen vertex in T_n is $O(k)$. In other words, if T_n^* is T_n randomly rerooted, and $Z_k(\tau)$ is the number of vertices of distance k from the root in a rooted tree τ, then the following holds.

Corollary 1.4. $\mathbb{E}\, Z_k(T_n^*) = O(k)$*, uniformly in all $k \geq 1$ and $n \geq 1$.*

This can be compared to [10, Theorem 1.13], which shows that

$$\mathbb{E}\, Z_k(T_n) = O(k), \tag{1.8}$$

again uniformly in k and n. Note that in the special case when T_n is a random (unordered) labelled tree, T_n^* has the same distribution, so Corollary 1.4 reduces to (1.8). However, in general, a randomly rerooted conditioned Galton–Watson tree is not a conditioned Galton–Watson tree.

Remark. The emphasis is on uniformity in both k and n. If we, on the contrary, fix k and consider limits as $n \to \infty$, we have $\mathbb{E}\, Z_k(T_n) \to 1 + k\sigma^2$, see Meir and Moon [13] and Janson [10; 11]. It is shown in [11] that the sequence $\mathbb{E}\, Z_k(T_n)$ is not always monotone in n.

We give a probabilistic proof of Theorem 1.3, and thus of Corollary 1.4 too, in Section 4.

We also give another proof by first proving a corresponding estimate for the generating function. (We present two different proofs, since we find both methods interesting, and both methods yield as intermediary steps in the proofs other results that we find interesting.) Let $f_n(z)$ be the generating function defined by

$$f_n(z) := \sum_{k=1}^{\infty} \mathbb{E}\, P_k(T_n)\, z^k.$$

We will use standard singularity analysis, see e.g. Flajolet and Sedgewick [9], and define the domain, for $0 < \beta < \pi/2$ and $\delta > 0$,

$$\Delta(\beta,\delta) := \{z \in \mathbb{C} : |z| < 1+\delta,\ z \neq 1,\ |\arg(z-1)| > \pi/2 - \beta\}.$$

Note that $|\arg(z-1)| > \pi/2-\beta$ is equivalent to $|\arg(1-z)| < \pi/2+\beta$.

Theorem 1.5. *For every ξ there exist positive constants C_3, β, δ such that for all $n \geq 1$, f_n extends to an analytic function in $\Delta(\beta,\delta)$ with*

$$|f_n(z)| \leq C_3 n |1-z|^{-2}, \qquad z \in \Delta(\beta,\delta). \tag{1.9}$$

By standard singularity analysis (i.e., estimate of the Taylor coefficients of $f_n(z)$ using Cauchy's formula and a suitable contour in $\Delta(\beta,\delta)$), (1.9) implies $\mathbb{E}\, P_k(T_n) = O(nk)$, see Flajolet and Sedgewick [9], Theorem VI.3 and (for the uniformity in n) Lemma IX.2 (applied to the family $\{f_n(z)/n\}$). Hence, Theorem 1.3 follows from Theorem 1.5.

For each pair of vertices v, w in a rooted tree, the path from v to w consists of two (possibly empty) parts, one going from v towards the root, ending at the last common ancestor $v \wedge w$ of v and w, and another part going from $v \wedge w$ to w in the direction away from the root. We will also prove extensions of the results above for $P_k(T_n)$, where we consider separately the lengths of these two parts. Define the corresponding bivariate generating function (now considering ordered pairs v, w)

$$h_n(x,y) := \mathbb{E} \sum_{v,w \in T_n} x^{d(v,v\wedge w)} y^{d(w,v\wedge w)}. \tag{1.10}$$

Theorem 1.6. *For every ξ there exist positive constants C_4, β, δ such that for all $n \geq 1$,*

$$|h_n(x,y)| \leq C_4 n |1-x|^{-1} |1-y|^{-1}, \qquad x, y \in \Delta(\beta,\delta).$$

Note that, by (1.10) and (1.9)

$$h_n(z,z) = \mathbb{E} \sum_{v,w \in T_n} z^{d(v,w)} = n + 2 f_n(z).$$

Hence Theorem 1.5 follows from Theorem 1.6, and thus Theorem 1.3 follows too from it. We prove Theorem 1.6 in Section 2.

If we define $\widetilde{P}_{\ell,m}(\tau) := \#\{(v,w) \in \tau : d(v, v\wedge w) = \ell,\ d(w, v \wedge w) = m\}$, then singularity analysis as above (but twice) shows that Theorem 1.6 implies the following. (Since $P_k = \frac{1}{2}\sum_{\ell=0}^{k} \widetilde{P}_{\ell,k-\ell}$, this too implies Theorem 1.3.)

Theorem 1.7. *There exists a constant C_5 such that for all $\ell, m \geq 0$ and $n \geq 1$, $\mathbb{E}\, \widetilde{P}_{\ell,m}(T_n) \leq C_5 n$.*

Acknowledgement. This research was mainly done during a workshop at Bellairs Research Institute in Barbados, March 2006, and completed during a visit of SJ to Centre de recherches mathématiques, Université de Montréal, October 2008. We thank two anonymous referees for helpful suggestions.

2. Proof of Theorem 1.6

We use some further generating functions. Recall that $\mathcal{T}$ is the (unconditioned) Galton–Watson tree with offspring distribution ξ, and define

$$\begin{aligned}
\Phi(z) &:= \mathbb{E}\, z^{\xi}, \\
F(z) &:= \mathbb{E}\, z^{|\mathcal{T}|}, \\
G(z,x) &:= \mathbb{E}\Bigl(z^{|\mathcal{T}|} \sum_{v\in\mathcal{T}} x^{d(v,o)}\Bigr), \\
H(z,x,y) &:= \mathbb{E}\Bigl(z^{|\mathcal{T}|} \sum_{v,w\in\mathcal{T}} x^{d(v,v\wedge w)} y^{d(w,v\wedge w)}\Bigr) = \sum_{n=1}^{\infty} \mathbb{P}(|\mathcal{T}|=n) h_n(x,y) z^n.
\end{aligned}$$

These functions are defined and analytic at least for $|z|,|x|,|y|<1$.

Let us condition on the degree d_o of the root of $\mathcal{T}$, recalling that $d_o \overset{\mathrm{d}}{=} \xi$. If $d_o = \ell$, then $\mathcal{T}$ has ℓ subtrees $\mathcal{T}_1,\dots,\mathcal{T}_\ell$ at the root o, and conditioned on $d_o=\ell$, these are independent and with the same distribution as $\mathcal{T}$; we denote their roots (the neighbours of o), by $o_1,\dots,o_\ell$.

Assume $d_o=\ell$, and let $|z|,|x|,|y|<1$. First, $|\mathcal{T}| = 1+\sum_{i=1}^{\ell} |\mathcal{T}_i|$ and thus $z^{|\mathcal{T}|} = z\prod_{i=1}^{\ell} z^{|\mathcal{T}_i|}$. Taking the expectation, we obtain, as is well-known, first

$$\mathbb{E}\bigl(z^{|\mathcal{T}|} \mid d_o=\ell\bigr) = z\,\mathbb{E}\prod_{i=1}^{\ell} z^{|\mathcal{T}_i|} = zF(z)^{\ell},$$

and then

$$F(z) = \mathbb{E}\bigl(z^{|\mathcal{T}|}\bigr) = z\sum_{\ell=0}^{\infty} \mathbb{P}(\xi=\ell)F(z)^{\ell} = z\Phi(F(z)).$$

Similarly, separating the cases $v\in\mathcal{T}_i$, $i=1,\dots,\ell$, and $v=o$,

$$\sum_{v\in\mathcal{T}} x^{d(v,o)} = \sum_{i=1}^{\ell}\sum_{v\in\mathcal{T}_i} x^{d(v,o_i)+1} + 1.$$

Hence,

$$\begin{aligned}
\mathbb{E}\bigl(z^{|\mathcal{T}|}\sum_{v\in\mathcal{T}} x^{d(v,o)} \mid d_o=\ell\bigr) &= \mathbb{E}\sum_{i=1}^{\ell} zz^{|\mathcal{T}_i|}\sum_{v\in\mathcal{T}_i} x^{d(v,o_i)+1}\prod_{j\neq i} z^{|\mathcal{T}_j|} + \mathbb{E}\Bigl(z\prod_{i=1}^{\ell} z^{|\mathcal{T}_i|}\Bigr) \\
&= \ell zxG(z,x)F(z)^{\ell-1} + zF(z)^{\ell}
\end{aligned}$$

and

$$G(z,x)=\sum_{\ell=0}^{\infty}\mathbb{P}(\xi=\ell)\ell zxG(z,x)F(z)^{\ell-1}+F(z)=zx\Phi'(F(z))G(z,x)+F(z)$$

which gives

$$G(z,x)=\frac{F(z)}{1-zx\Phi'(F(z))}. \tag{2.1}$$

Similarly,

$$\begin{aligned}\mathbb{E}\Bigl(z^{|\mathcal{T}|}\sum_{v,w\in\mathcal{T}}x^{d(v,v\wedge w)}y^{d(w,v\wedge w)}\mid d_o=\ell\Bigr)=\\
\mathbb{E}\sum_{i=1}^{\ell}zz^{|\mathcal{T}_i|}\sum_{v,w\in\mathcal{T}_i}x^{d(v,v\wedge w)}y^{d(w,v\wedge w)}\prod_{j\neq i}z^{|\mathcal{T}_j|}\\
+\mathbb{E}\sum_{i\neq j}zz^{|\mathcal{T}_i|}\sum_{v\in\mathcal{T}_i}x^{d(v,o_i)+1}z^{|\mathcal{T}_j|}\sum_{w\in\mathcal{T}_j}y^{d(w,o_j)+1}\prod_{k\neq i,j}z^{|\mathcal{T}_k|}\\
+\mathbb{E}\sum_{i=1}^{\ell}zz^{|\mathcal{T}_i|}\sum_{v\in\mathcal{T}_i}x^{d(v,o_i)+1}\prod_{k\neq i}z^{|\mathcal{T}_k|}\\
+\mathbb{E}\sum_{j=1}^{\ell}zz^{|\mathcal{T}_j|}\sum_{w\in\mathcal{T}_j}y^{d(w,o_j)+1}\prod_{k\neq j}z^{|\mathcal{T}_k|}+\mathbb{E}\Bigl(z\prod_{i=1}^{\ell}z^{|\mathcal{T}_i|}\Bigr)\end{aligned}$$

leading to

$$\begin{aligned}H(z,x,y)=z\Phi'(F(z))H(z,x,y)+zxy\Phi''(F(z))G(z,x)G(z,y)\\
+zx\Phi'(F(z))G(z,x)+zy\Phi'(F(z))G(z,y)+F(z)\end{aligned}$$

which gives

$$H(z,x,y)=\\
\frac{zxy\Phi''(F(z))G(z,x)G(z,y)+z\Phi'(F(z))\bigl(xG(z,x)+yG(z,y)\bigr)+F(z)}{1-z\Phi'(F(z))} \tag{2.2}$$

Assume now for simplicity that ξ has span 1. (The case when the span is $d>1$ is treated similarly with the standard modification that we have to give special treatment to neighbourhoods of the d:th unit roots.) Then, by [10, Lemma A.2], for some $\delta>0$ and $\beta\leq\pi/4$, F extends to an analytic function in $\Delta(\beta,\delta)$ with $|F(z)|<1$ for $z\in\Delta(\beta,\delta)$ and

$$F(z)=1-\sqrt{2}\sigma^{-1}\sqrt{1-z}+o\bigl(|z-1|^{1/2}\bigr),\qquad\text{as } z\to1 \text{ with } z\in\Delta(\beta,\delta). \tag{2.3}$$

We will prove the following companion results.

Lemma 2.1. *If ξ has span 1, then there exists $\beta, \delta > 0$ such that F extends to an analytic function in $\Delta(\beta, \delta)$ and, for some $c_1, c_2 > 0$, if $x, z \in \Delta(\beta, \delta)$, then*

$$|1 - z\Phi'(F(z))| \geq c_1 |1 - z|^{1/2}, \tag{2.4}$$

$$|1 - xz\Phi'(F(z))| \geq c_2 |1 - x|. \tag{2.5}$$

Consequently, $G(z, x)$ and $H(z, x, y)$ extend to analytic functions of $x, y, z \in \Delta(\beta, \delta)$, and, for all $x, y, z \in \Delta(\beta, \delta)$,

$$|G(z, x)| \leq C_6 |1 - x|^{-1}, \tag{2.6}$$

$$|H(z, x, y)| \leq C_7 |1 - z|^{-1/2} |1 - x|^{-1} |1 - y|^{-1}. \tag{2.7}$$

Standard singularity analysis [9, Lemma IX.2] applied to (2.7) yields

$$|\,\mathbb{P}(|\mathcal{T}| = n) h_n(x, y)| \leq C_8 n^{-1/2} |1 - x|^{-1} |1 - y|^{-1}, \qquad x, y \in \Delta(\beta, \delta),$$

which proves Theorem 1.6 because, as is well known, a singularity analysis of (2.3) yields

$$\mathbb{P}(|\mathcal{T}| = n) \sim c_3 n^{-3/2}.$$

It thus remains only to prove Lemma 2.1.

Proof of Lemma 2.1. Since $\mathbb{E}\,\xi^2 < \infty$, Φ' and Φ'' extend to continuous functions on the closed unit disc with $\Phi'(1) = \mathbb{E}\,\xi = 1$ and $\Phi''(1) = \mathbb{E}\,\xi(\xi - 1) = \sigma^2$. Hence, (2.3) yields, for $z \in \Delta(\beta, \delta)$,

$$\begin{aligned}\Phi'(F(z)) &= \Phi'(1) + \Phi''(1)\bigl(F(z) - 1\bigr) + o(|F(z) - 1|) \\ &= 1 - \sqrt{2}\sigma\sqrt{1 - z} + o\bigl(|z - 1|^{1/2}\bigr)\end{aligned}$$

and

$$z\Phi'(F(z)) = \Phi'(F(z)) + O(|z - 1|) = 1 - \sqrt{2}\sigma\sqrt{1 - z} + o\bigl(|z - 1|^{1/2}\bigr). \tag{2.8}$$

Let $B(1, \varepsilon) := \{z : |z - 1| < \varepsilon\}$, and take $\beta < \pi/4$. Since $z \in \overline{\Delta(\beta, \delta)} \setminus \{1\}$ entails $|\arg(1 - z)| \leq \pi/2 + \beta$ and thus $|\arg\sqrt{1 - z}| \leq \pi/4 + \beta/2$, it follows from (2.8) that, for some small $\varepsilon > 0$, if $z \in \overline{\Delta(\beta, \delta) \cap B(1, \varepsilon)}$ with $z \neq 1$, then (2.4) holds, $\bigl|z\Phi'(F(z)) - 1\bigr| = O(\varepsilon^{1/2})$,

$$\begin{aligned}\bigl|\arg\bigl(z\Phi'(F(z)) - 1\bigr)\bigr| &> |\arg(-\sqrt{1 - z})| - \beta/2 \\ &\geq \pi - (\pi/4 + \beta/2) - \beta/2 = 3\pi/4 - \beta,\end{aligned} \tag{2.9}$$

and consequently, since $3\pi/4 - \beta > \pi/2$, if ε is small enough,

$$|z\Phi'(F(z))| < 1. \tag{2.10}$$

Similarly, if $x \in \Delta(\beta, \delta)$, then $|\arg(1 - x)| < \pi/2 + \beta$ and

$$x^{-1} = \bigl(1 - (1 - x)\bigr)^{-1} = 1 + (1 - x) + o(|1 - x|), \qquad x \to 1,$$

so if $\varepsilon > 0$ is small enough, then, for $x \in \Delta(\beta, \delta) \cap B(1, \varepsilon)$,

$$|\arg(x^{-1} - 1)| < \pi/2 + 2\beta. \tag{2.11}$$

If we choose $\beta \le \pi/16$, it follows from (2.9) and (2.11) that the triangle with vertices in 1, x^{-1} and $z\Phi'(F(z))$ has an angle at least $\pi/4 - 3\beta \ge \pi/16$ at 1, and thus by elementary trigonometry (the sine theorem),

$$|x^{-1} - z\Phi'(F(z))| \ge c_4|x^{-1} - 1|,$$

and so (2.5) holds, when $z, x \in \Delta(\beta,\delta) \cap B(1,\varepsilon)$, provided $\beta, \delta, \varepsilon$ are small enough.

It remains to treat the case when x or z does not belong to $B(1,\varepsilon)$, i.e., $|x-1| \ge \varepsilon$ or $|z-1| \ge \varepsilon$. We do this by compactness arguments.

First, let

$$A := \{z\Phi'(F(z)) : z \in \Delta(\beta,\delta) \cap B(1,\varepsilon)\}$$
$$B_\rho := \{x^{-1} : x \in \overline{\Delta(\beta,\rho)} \setminus B(1,\varepsilon),\ |x| \ge 1/2\}.$$

Then $B := \bigcap_{\rho>0} B_\rho \subset \{\zeta : |\zeta| \ge 1\} \setminus \{1\}$, and it follows from (2.10) that $\overline{A} \cap B = \emptyset$. Since $\overline{A}$ and all B_ρ are compact, it follows that $\overline{A} \cap B_\rho = \emptyset$ for some $\rho > 0$, and thus, if $z \in \Delta(\beta,\delta) \cap B(1,\varepsilon)$ and $x \in \Delta(\beta,\rho) \setminus B(1,\varepsilon)$ with $|x| \ge 1/2$, then $|x^{-1} - z\Phi'(F(z))| \ge c_5$ for some $c_5 > 0$, which implies (2.5) for such z and x. Moreover, if $z \in \Delta(\beta,\delta) \cap B(1,\varepsilon)$ and $|x| < 1/2$, (2.10) shows that $|1 - xz\Phi'(F(z))| \ge 1 - |x| \ge 1/2$, so (2.5) then holds if $c_2 \le 1/3$.

Finally, if $z \in \Delta(\beta,\delta)$, then $|F(z)| < 1$ [10, Lemma A.2] as stated above, and thus $|\Phi'(F(z))| < 1$. If $0 < \beta_1 < \beta$ and $0 < \delta_1 < \delta$, then $\overline{\Delta(\beta_1,\delta_1)} \subset \Delta(\beta,\delta) \cup \{1\}$, and thus by compactness

$$C_\varepsilon := \sup\bigl\{|\Phi'(F(z))| : z \in \overline{\Delta(\beta_1,\delta_1)} \setminus B(1,\varepsilon)\bigr\} < 1.$$

Consequently, if $\delta_2 \le \delta_1$ is small enough and $x, z \in \Delta(\beta_1,\delta_2)$ with $|z-1| \ge \varepsilon$, then

$$|xz\Phi'(F(z))| \le (1+\delta_2)^2 C_\varepsilon < 1.$$

Hence (2.5) holds in this case too for some $c_2 > 0$, and similarly (2.4) holds for $z \in \Delta(\beta_1,\delta_2) \setminus B(1,\varepsilon)$.

This completes the proof of (2.4) and (2.5), for some new $\beta, \delta > 0$ (viz., β_1 and $\min(\delta_2,\rho)$). $G(z,x)$ now can be defined for all $x, z \in \Delta(\beta,\delta)$ by (2.1), and (2.6) holds by (2.5). Similarly, $H(z,x,y)$ can be defined for all $x, y, z \in \Delta(\beta,\delta)$ by (2.2), and (2.7) holds by (2.4), (2.6), and the fact that Φ' and Φ'' are bounded on the unit disc. (Recall that $|F(z)| < 1$ for $z \in \Delta(\beta,\delta)$.)

This completes the proof of Lemma 2.1, and thus of Theorem 1.6. As said above, Theorems 1.5 and 1.3 follows. □

3. Proof of Lemma 1.2 and Theorem 1.1

Denote the left hand side of (1.7) by $\Psi(n,t)$. Since $\sum_j X(j;T_n)e^{\mathrm{i}jt} = \sum_{v\in T_n} e^{\mathrm{i}tL_v}$, we have

$$\Psi(n,t) = n^{-2}\,\mathbb{E}\Bigl|\sum_{v\in T_n} e^{\mathrm{i}tL_v}\Bigr|^2 = n^{-2}\,\mathbb{E}\sum_{v,w\in T_n} e^{\mathrm{i}t(L_v-L_w)}. \tag{3.1}$$

Condition on T_n and consider two vertices v and w in T_n. If $v \wedge w$ is the last common ancestor of v and w, then $L_v - L_{v\wedge w}$ and $L_w - L_{v\wedge w}$ are (conditionally, given T_n) independent sums of $d(v, v\wedge w)$ and $d(w, v\wedge w)$ copies of η, respectively. Consequently, letting $\varphi_\eta(t) := \mathbb{E}\, e^{\mathrm{i}t\eta}$ be the characteristic function of η,

$$\begin{aligned}\mathbb{E}\big(e^{\mathrm{i}t(L_v-L_w)} \mid T_n\big) &= \mathbb{E}\big(e^{\mathrm{i}t(L_v-L_{v\wedge w})} \mid T_n\big)\, \mathbb{E}\big(e^{-\mathrm{i}t(L_w-L_{v\wedge w})} \mid T_n\big) \\ &= \varphi_\eta(t)^{d(v,v\wedge w)}\, \overline{\varphi_\eta(t)}^{d(w,v\wedge w)}.\end{aligned}$$

Hence, by (3.1) and (1.10),

$$\Psi(n,t) = n^{-2}\,\mathbb{E}\sum_{v,w\in T_n} \varphi_\eta(t)^{d(v,v\wedge w)}\overline{\varphi_\eta(t)}^{d(w,v\wedge w)} = n^{-2} h_n\big(\varphi_\eta(t), \overline{\varphi_\eta(t)}\big)$$

and Theorem 1.6 yields

$$\Psi(n,t) \le C_4 n^{-1} |1-\varphi_\eta(t)|^{-2}. \tag{3.2}$$

Since $\mathbb{E}\,\eta = 0$ and $\mathbb{E}\,\eta^2 = \sigma_\eta^2 < \infty$, we have $\varphi_\eta(t) = \exp\big(-\frac12\sigma_\eta^2 t^2 + o(t^2)\big)$ for small $|t|$; moreover, since η has span 1, $\varphi_\eta(t) \neq 1$ for $0 < |t| \le \pi$. It follows that $\psi(t) := (1-\varphi_\eta(t))/t^2$ is a continuous non-zero function on $[-\pi,\pi]$ (with $\psi(0) := \frac12\sigma_\eta^2$); hence, by compactness, $|\psi(t)| \ge c_6$ for some $c_7 > 0$, and thus

$$|1-\varphi_\eta(t)| \ge c_7 t^2, \qquad |t| \le \pi.$$

It now follows from (3.2), and the obvious fact that $\Psi(n,t) \le 1$, that

$$(1+nt^4)\Psi(n,t) \le 1 + nt^4\Psi(n,t) \le 1 + C_4\frac{t^4}{|1-\varphi_\eta(t)|^2} \le C_9.$$

This proves Lemma 1.2, which as remarked in Section 1 implies Theorem 1.1 by [5, Remark 3.7].

4. Probabilistic proof of Theorem 1.3

In a rooted tree τ, let $Q_k(\tau)$, $k \ge 1$, denote the number of (unordered) pairs of vertices at path distance k from each other such that the path between them visits the root, and let $Q'_k(\tau)$ be the number of such pairs where the root cannot be one of the two vertices in the pair; thus $Q_k(\tau) = Q'_k(\tau) + Z_k(\tau)$. Then, in the Galton–Watson tree $\mathcal{T}$, if ξ is the number of children of the root, and the subtrees rooted at these children are denoted $\mathcal{T}_1, \dots, \mathcal{T}_\xi$,

$$Q'_k(\mathcal{T}) = \sum_{(r,s):1\le r<s\le\xi}\ \sum_{j=0}^{k-2} Z_j(\mathcal{T}_r) Z_{k-2-j}(\mathcal{T}_s) \tag{4.1}$$

and thus, since we assume $\mathcal{T}$ to be critical, i.e., $\mathbb{E}\,\xi = 1$, so $\mathbb{E}\,Z_k(\mathcal{T}) = 1$ for every k,

$$\mathbb{E}\,Q_k(\mathcal{T}) = \mathbb{E}\,Z_k(\mathcal{T}) + \mathbb{E}\,Q'_k(\mathcal{T}) = 1 + \mathbb{E}\,\frac{\xi(\xi-1)}{2}(k-1) = 1 + (k-1)\frac{\sigma^2}{2}. \tag{4.2}$$

Let $\widehat{T}_n$ denote the random subtree of T_n rooted at a uniformly selected random vertex. (Note the difference from T_n^* in Corollary 1.4; in T_n^* we keep all n vertices, but in $\widehat{T}_n$ we keep only the vertices below the new root.) Then, clearly,

$$\mathbb{E}\{P_k(T_n)\} = n\,\mathbb{E}\{Q_k(\widehat{T}_n)\}.$$

Consequently, Theorem 1.3 is equivalent to:

Theorem 4.1. *There exists a constant C_{10} such that for all $k \geq 1$ and $n \geq 1$, $\mathbb{E}\,Q_k(\widehat{T}_n) \leq C_{10}k$.*

In order to prove this, we will need a related, but different, estimate for the conditioned Galton–Watson tree T_n.

Theorem 4.2. *There exists a constant C_{11} such that for all $k \geq 1$ and $n \geq 1$, $\mathbb{E}\,Q_k(T_n) \leq C_{11}k\sqrt{n}$.*

It is easy to see $\mathbb{E}\,Q_k(T_n) \geq c_8 n^{3/2}$ when $k \sim \sqrt{n}$, so the estimate in Theorem 4.2 then is of the right order; in particular, the estimate in Theorem 4.1 for $\widehat{T}_n$ does *not* hold for T_n.

To prove these theorems we use a few more or less standard estimates.

Lemma 4.3. *Assume, as above, (1.1), and let d be the span of ξ. Let $S_n := \sum_{i=1}^n \xi_i$, where ξ_i are independent copies of ξ. Then, for $n \equiv 1 \pmod d$,*

$$\mathbb{P}(|\mathcal{T}| = n) = \frac{1}{n}\,\mathbb{P}(S_n = n-1) \sim \frac{d}{\sigma\sqrt{2\pi}\,n^{3/2}} \quad \text{as } n \to \infty. \tag{4.3}$$

More generally, let $W_\ell := \sum_{i=1}^{\ell} |\mathcal{T}_i|$ be the size of the union of ℓ independent copies of $\mathcal{T}$, or equivalently, the total progeny of a Galton–Watson process started with ℓ individuals, with offspring distribution ξ. Then, for all $\ell \geq 1$ and $n \geq 1$,

$$\mathbb{P}(W_\ell = n) = \frac{\ell}{n}\,\mathbb{P}(S_n = n-\ell) \leq C_{12}\ell n^{-3/2}\exp(-c_9\ell^2/n). \tag{4.4}$$

In particular,

$$\mathbb{P}(W_\ell = n) \leq C_{13}n^{-1}. \tag{4.5}$$

Proof. The identity in (4.4) is well-known, see e.g., Dwass [8], Kolchin [12, Lemma 2.1.3, p. 105] and Pitman [14]. The identity in (4.3) is the special case $\ell = 1$, and the well-known tail estimate in (4.3) then follows by the local central limit theorem, see, e.g., Kolchin [12, Lemma 2.1.4, p. 105].

Similarly, the inequality in (4.4) follows by the estimate $\mathbb{P}(S_n = n-\ell) \leq C_{12}n^{-1/2}\exp(-c_9\ell^2/n)$ from [10, Lemma 2.1]. The inequality $e^{-x} \leq x^{-1/2}$ yields (4.5). ∎

Lemma 4.4. *For every $r > 0$ there is a constant $C_{14}(r)$ such that for all $k \geq 0$ and $n \geq 1$, $\mathbb{E}\,Z_k(T_n)^r \leq C_{14}(r)n^{r/2}$.*

Proof. For any rooted tree T, let $T^{(k)}$ be the tree pruned at height k, i.e., the subtree consisting of all vertices of distance at most k from the root. Let τ be a given rooted tree of height k, and let $m := Z_k(\tau)$, the number of leaves at maximal depth. Note that if $\tau = T^{(k)}$ for some tree T, then $|T| = n$ if and only if T has $n - |\tau|$ vertices at greater depth than k, and thus $N := n - |\tau| + m$ vertices at depth k or greater. Hence, with W_m as in Lemma 4.3 and using (4.3) and (4.4), for any $r > 0$ and assuming $N > 0$ (otherwise the probability is 0),

$$\begin{aligned}\mathbb{P}\big(T_n^{(k)} = \tau\big) &= \frac{\mathbb{P}\big(\mathcal{T}^{(k)} = \tau,\ |\mathcal{T}| = n\big)}{\mathbb{P}(|\mathcal{T}| = n)} = \frac{\mathbb{P}(\mathcal{T}^{(k)} = \tau)\,\mathbb{P}(W_m = N)}{\mathbb{P}(|\mathcal{T}| = n)} \\ &\le C_{15} n^{3/2}\,\mathbb{P}(\mathcal{T}^{(k)} = \tau) m N^{-3/2} e^{-c_9 m^2/N} \\ &\le C_{16}(r) n^{3/2}\,\mathbb{P}(\mathcal{T}^{(k)} = \tau) m N^{-3/2} (m^2/N)^{-r/2} \\ &= C_{16}(r) n^{3/2} N^{r/2-3/2} m^{1-r}\,\mathbb{P}(\mathcal{T}^{(k)} = \tau).\end{aligned}$$

If $r \ge 3$, this yields, since $N \le n$, the estimate

$$\mathbb{P}(T_n^{(k)} = \tau) \le C_{16}(r) n^{r/2} m^{1-r}\,\mathbb{P}(\mathcal{T}^{(k)} = \tau),$$

and summing over all τ of height k with $Z_k(\tau) = m$ we obtain

$$\mathbb{P}(Z_k(T_n) = m) \le C_{16}(r) n^{r/2} m^{1-r}\,\mathbb{P}(Z_k(\mathcal{T}) = m).$$

Consequently,

$$\begin{aligned}\mathbb{E}\, Z_k(T_n)^r &= \sum_{m=1}^{\infty} m^r\,\mathbb{P}(Z_k(T_n) = m) \\ &\le C_{16}(r) n^{r/2} \sum_{m=1}^{\infty} m\,\mathbb{P}(Z_k(\mathcal{T}) = m) \\ &= C_{16}(r) n^{r/2}\,\mathbb{E}\, Z_k(\mathcal{T}) = C_{16}(r) n^{r/2}.\end{aligned}$$

This proves the result for $r \ge 3$, and the result for $0 < r < 3$ follows by Lyapounov's (or Hölder's) inequality. □

Lemma 4.5. *For all $k \ge 1$ and $n \ge 1$, $\mathbb{E}\, Z_k(T_n) \le C_{17}(k \wedge \sqrt{n})$. Equivalently, for all $k \ge 0$ and $n \ge 1$, $\mathbb{E}\, Z_k(T_n) \le C_{18}((k+1) \wedge \sqrt{n})$.*

Proof. The estimate $\mathbb{E}\, Z_k(T_n) \le C_{19} k$ is (1.8), which is proved in [10, Theorem 1.13]. The estimate $\mathbb{E}\, Z_k(T_n) \le C_{20}\sqrt{n}$ is proved in Lemma 4.4. □

Remark. The estimate $\mathbb{E}\, Z_k(T_n) \le C_{21}\sqrt{n}$ was proved by Drmota and Gittenberger [7], assuming that ξ has an exponential moment; in fact, they then prove the stronger bound $\mathbb{E}\, Z_k(T_n) \le C_{22}\sqrt{n}\exp(-c_{10} k/\sqrt{n})$. The bound in Lemma 4.5 can be further improved to $\mathbb{E}\, Z_k(T_n) \le C_{23} k \exp(-c_{11} k^2/n)$, but we do not know a reference for this estimate. (Details may appear elsewhere.)

Remark. Note that Lemma 4.4 yields an estimate $O(n^{r/2})$ of the rth moment of $Z_k(T_n)$ for an arbitrary r assuming only a second moment of ξ. This is in contrast to the estimate (1.8), where the corresponding estimate $\mathbb{E}\, Z_k(T_n)^r = O(k^r)$ is valid (for integer $r \geq 1$ at least) if ξ has a finite $r+1$:th moment, but not otherwise (not even for a fixed $k \geq 2$); one direction is by Theorem 1.13 in [10], and the converse follows from the discussion after Lemma 2.1 in [10].

Proof of Theorem 4.2. We have $Q_k(T_n) = Q'_k(T_n) + Z_k(T_n)$, and $\mathbb{E}\, Z_k(T_n) \leq C_{17} k$ by Lemma 4.5, so it suffices to show that $\mathbb{E}\, Q'_k(T_n) \leq C_{24} k\sqrt{n}$.

We use (4.1), condition on $|\mathcal{T}| = n$ and take expectations. Using the symmetry and recalling that $\mathcal{T}_1, \dots, \mathcal{T}_\xi$ are independent and $(\mathcal{T}_i \mid |\mathcal{T}_i| = n_i) \overset{\mathrm{d}}{=} (\mathcal{T} \mid |\mathcal{T}| = n_i) \overset{\mathrm{d}}{=} T_{n_i}$ for any n_i, we obtain, with $p_\ell := \mathbb{P}(\xi = \ell)$ and $q_m := \mathbb{P}(|\mathcal{T}| = m)$,

$$
\begin{aligned}
\mathbb{E}\left\{Q'_k(T_n)\right\} &= \frac{\mathbb{E}\left\{\mathbf{1}_{[\xi\geq 2, |\mathcal{T}|=n]} \sum_{1\leq r<s\leq \xi} \sum_{j=0}^{k-2} Z_j(\mathcal{T}_r) Z_{k-2-j}(\mathcal{T}_s)\right\}}{\mathbb{P}\{|\mathcal{T}| = n\}} \\
&= \frac{\mathbb{E}\left\{\mathbf{1}_{[|\mathcal{T}|=n]} \binom{\xi}{2} \sum_{j=0}^{k-2} Z_j(\mathcal{T}_1) Z_{k-2-j}(\mathcal{T}_2)\right\}}{\mathbb{P}\{|\mathcal{T}| = n\}} \\
&= q_n^{-1} \sum_{\ell=2}^{\infty} p_\ell \binom{\ell}{2} \sum_{n_1, n_2 \geq 1} q_{n_1} q_{n_2}\, \mathbb{P}\left(\sum_{i=3}^{\ell} |\mathcal{T}_i| = n - 1 - n_1 - n_2\right) \\
&\qquad \times \sum_{j=0}^{k-2} \mathbb{E}\left\{Z_j(T_{n_1})\right\} \mathbb{E}\left\{Z_{k-2-j}(T_{n_2})\right\}.
\end{aligned}
$$

We begin with the inner sum over j, $\Sigma_1(n_1, n_2)$ say. By symmetry, we consider only $n_1 \leq n_2$, and then we obtain from Lemma 4.5 the estimates $\mathbb{E}\, Z_{k-2-j}(T_{n_2}) \leq C_{18}((k-1-j) \wedge n_2^{1/2}) \leq C_{18}(k \wedge n_2^{1/2})$ and

$$
\sum_{j=0}^{k-2} \mathbb{E}\left\{Z_j(T_{n_1})\right\} \leq \begin{cases} \sum_{j=0}^{k-2} C_{18}(j+1) \leq C_{18} k^2, \\ \mathbb{E} \sum_{j=0}^{\infty} Z_j(T_{n_1}) = n_1. \end{cases}
$$

Hence

$$
\Sigma_1(n_1, n_2) \leq C_{25}(k^2 \wedge n_1)(k \wedge n_2^{1/2}). \tag{4.6}
$$

Let $m := n_1 + n_2$ and sum over n_1, n_2 with a given sum m. We have by (4.6) and (4.3),

$$\begin{aligned}
\Sigma_2(m) &:= \sum_{n_1+n_2=m} q_{n_1} q_{n_2} \Sigma_1(n_1, n_2) \\
&\le 2 \sum_{n_1=1}^{m/2} q_{n_1} q_{m-n_1} C_{25} (k^2 \wedge n_1)(k \wedge (m-n_1)^{1/2}) \\
&\le C_{26} \sum_{n_1=1}^{m/2} n_1^{-3/2} (m-n_1)^{-3/2} (k^2 \wedge n_1)(k \wedge (m-n_1)^{1/2}) \\
&\le C_{27} \frac{k \wedge m^{1/2}}{m^{3/2}} \sum_{n_1=1}^{m/2} \frac{k^2 \wedge n_1}{n_1^{3/2}} \\
&\le C_{28} \frac{k \wedge m^{1/2}}{m^{3/2}} (k \wedge m^{1/2}) = C_{28} \frac{k^2 \wedge m}{m^{3/2}}.
\end{aligned} \tag{4.7}$$

We define further

$$\Sigma_3(\ell) := \sum_{m=2}^{n-1} \Sigma_2(m) \, \mathbb{P}\Big(\sum_{i=3}^{\ell} |\mathcal{T}_i| = n-1-m \Big); \tag{4.8}$$

thus

$$\mathbb{E}\, Q_k'(T_n) = q_n^{-1} \sum_{\ell=2}^{\infty} p_\ell \binom{\ell}{2} \Sigma_3(\ell) \le C_{29} n^{3/2} \sum_{\ell=2}^{\infty} p_\ell \ell^2 \Sigma_3(\ell). \tag{4.9}$$

We will show that $\Sigma_3(\ell) \le C_{30} k/n$, uniformly in $\ell \ge 2$, and the result follows by (4.9), recalling that $\sum_\ell p_\ell \ell^2 = \mathbb{E}\,\xi^2 < \infty$. (The proof can be simplified in the case $\mathbb{E}\,\xi^3 < \infty$, when it suffices to show that $\Sigma_3(\ell) \le C_{31} k\ell/n$.)

First, if $\ell = 2$, the only non-zero term in (4.8) is for $m = n-1$, which yields, by (4.7),

$$\Sigma_3(2) = \Sigma_2(n-1) \le C_{32} \frac{k^2 \wedge n}{n^{3/2}} \le C_{32} \frac{k\sqrt{n}}{n^{3/2}}.$$

For $\ell > 2$, we split the sum in (4.8) into two parts, with $m \le n/2$ and $m > n/2$. We have, by (4.7),

$$\begin{aligned}
\sum_{m=n/2}^{n-1} \Sigma_2(m) \, \mathbb{P}\Big(\sum_{i=3}^{\ell} |\mathcal{T}_i| = n-1-m \Big) &\le C_{33} \frac{k^2 \wedge n}{n^{3/2}} \, \mathbb{P}\Big(\sum_{i=3}^{\ell} |\mathcal{T}_i| \le n/2 \Big) \\
&\le C_{33} \frac{k^2 \wedge n}{n^{3/2}} \le C_{33} \frac{k}{n}.
\end{aligned}$$

Similarly, using (4.7) and (4.5) (with ℓ replaced by $\ell-2$),

$$\sum_{m=1}^{n/2}\Sigma_2(m)\,\mathbb{P}\Bigl(\sum_{i=3}^{\ell}|\mathcal{T}_i|=n-1-m\Bigr)\le C_{34}\sum_{m=1}^{n/2}\frac{k^2\wedge m}{m^{3/2}}\cdot\frac{1}{n}$$
$$\le\frac{C_{34}}{n}\sum_{m=1}^{\infty}\frac{k^2\wedge m}{m^{3/2}}\le C_{35}\frac{k}{n}.$$

Thus $\Sigma_3(\ell)\le C_{36}k/n$, and the theorem follows by (4.9). □

Proof of Theorems 4.1 and 1.3. Aldous [2] has studied the behavior of a random subtree $\widehat{T}_n$ in a conditional Galton–Watson tree T_n. In particular, he has the following identity [2, p. 242], for any fixed ordered tree τ of order at most n (provided that the probabilities in the denominators are nonzero):

$$\frac{\mathbb{P}\{\widehat{T}_n=\tau\}}{\mathbb{P}\{\mathcal{T}=\tau\}}=\frac{(n-|\tau|+1)\,\mathbb{P}\{|\mathcal{T}|=n-|\tau|+1\}}{n\,\mathbb{P}\{|\mathcal{T}|=n\}}\cdot\frac{\gamma}{p_0},$$

where γ is the expected proportion of leaves in $\widehat{T}_{n-|\tau|+1}$ and $p_0=\mathbb{P}\{\xi=0\}$. We will simply bound $\gamma\le 1$, but it is well-known that as $n-|\tau|+1\to\infty$, $\gamma\to p_0$, see e.g. Kolchin [12, Theorem 2.3.1, p. 113]. Thus, using the well-known tail estimate (4.3), for all (permitted) n and τ

$$\frac{\mathbb{P}\{\widehat{T}_n=\tau\}}{\mathbb{P}\{\mathcal{T}=\tau\}}\le C_{37}\frac{(n-|\tau|+1)\,\mathbb{P}\{|\mathcal{T}|=n-|\tau|+1\}}{n\,\mathbb{P}\{|\mathcal{T}|=n\}}\le C_{38}\sqrt{\frac{n}{n-|\tau|+1}}.$$

Hence,

$$\begin{aligned}\mathbb{E}\{Q_k(\widehat{T}_n)\}&=\sum_{\tau}Q_k(\tau)\,\mathbb{P}\{\widehat{T}_n=\tau\}\\&\le C_{38}\sum_{\tau}Q_k(\tau)\sqrt{\frac{n}{n-|\tau|+1}}\,\mathbb{P}\{\mathcal{T}=\tau\}\\&=C_{38}\,\mathbb{E}\left(Q_k(\mathcal{T})\sqrt{\frac{n}{n-|\mathcal{T}|+1}}\right)\\&\le C_{39}\,\mathbb{E}\{Q_k(\mathcal{T})\}+C_{38}\sum_{n\ge\ell>n/2}\sqrt{\frac{n}{n-\ell+1}}\,\mathbb{E}\left\{\mathbf{1}_{[|\mathcal{T}|=\ell]}Q_k(\mathcal{T})\right\}.\end{aligned}$$

We have $\mathbb{E}\{Q_k(\mathcal{T})\}\le C_{40}k$ by (4.2), and, using (4.3) and Theorem 4.2,

$$\mathbb{E}\left\{\mathbf{1}_{[|\mathcal{T}|=\ell]}Q_k(\mathcal{T})\right\}=\mathbb{P}\{|\mathcal{T}|=\ell\}\,\mathbb{E}\left\{Q_k(T_\ell)\right\}\le C_{41}\ell^{-3/2}k\ell^{1/2}=C_{41}k/\ell.$$

Consequently,

$$
\begin{aligned}
\mathbb{E}\{Q_k(\widehat{T}_n)\} &\le C_{40}k + C_{42} \sum_{n \ge \ell > n/2} \frac{k}{\ell} \sqrt{\frac{n}{n-\ell+1}} \\
&\le C_{40}k + C_{43} \frac{k}{n} \sum_{j=1}^{n} \sqrt{\frac{n}{j}} \\
&\le C_{44}k.
\end{aligned}
$$

This proves Theorem 4.1, which by the argument at the beginning of the section yields Theorem 1.3. □

References

[1] D. Aldous, The continuum random tree II: an overview. *Stochastic Analysis (Proc., Durham, 1990)*, 23–70, London Math. Soc. Lecture Note Ser. 167, Cambridge Univ. Press, Cambridge, 1991.

[2] D. Aldous, Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.* **1**, no. 2, 228–266 (1991).

[3] D. Aldous, Tree-based models for random distribution of mass. *J. Statist. Phys.* **73**, 625–641 (1993).

[4] M. Bousquet-Mélou, Limit laws for embedded trees. Applications to the integrated superBrownian excursion. *Random Struct. Alg.* **29**, no. 4, 475–523 (2006).

[5] M. Bousquet-Mélou & S. Janson, The density of the ISE and local limit laws for embedded trees. *Ann. Appl. Probab.* **16**, no. 3, 1597–1632 (2006).

[6] L. Devroye, Branching processes and their applications in the analysis of tree structures and tree algorithms. *Probabilistic Methods for Algorithmic Discrete Mathematics*, 249–314, eds. M. Habib et al., Algorithms Combin. 16, Springer-Verlag, Berlin, 1998.

[7] M. Drmota & B. Gittenberger, The width of Galton–Watson trees conditioned by the size. *Discrete Mathematics and Theoretical Computer Science*, **6**, 387–400 (2004).

[8] M. Dwass, The total progeny in a branching process and a related random walk. *J. Appl. Probab.* **6** (1969), 682–686.

[9] P. Flajolet & R. Sedgewick, *Analytic Combinatorics.* Cambridge Univ. Press, Cambridge, 2008.

[10] S. Janson, Random cutting and records in deterministic and random trees. *Random Struct. Alg.* **29**, no. 2, 139–179 (2006).

[11] S. Janson, Conditioned Galton–Watson trees do not grow. *Proceedings, Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities (Nancy, 2006)*, DMTCS Proceedings AG, 331–334 (2006).

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

[13] A. Meir & J.W. Moon, On the altitude of nodes in random trees. *Canad. J. Math.* **30**, 997–1015 (1978).

[14] J. Pitman, Enumerations of trees and forests related to branching processes and random walks. *Microsurveys in Discrete Probability (Princeton, NJ, 1997)*, 163–180, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 41, Amer. Math. Soc., Providence, RI, 1998.

School of Computer Science, McGill University, 3480 University Street, Montreal, Canada H3A 2K6

E-mail address: `lucdevroye@gmail.com`
URL: `http://cg.scs.carleton.ca/~luc/`

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: `svante.janson@math.uu.se`
URL: `http://www.math.uu.se/~svante/`