Ann. Comb.

https://doi.org/10.1007/s00026-024-00711-4

Annals of Combinatorics

The Likely Maximum Size of Twin Subtrees in a Large Random Tree

Miklós Bóna, Ovidiu Costin and Boris Pittel

Abstract. We call a pair of vertex-disjoint, induced subtrees of a rooted tree twins if they have the same counts of vertices by out-degrees. The likely maximum size of twins in a uniformly random, rooted Cayley tree of size $n \to \infty$ is studied. It is shown that the expected number of twins of size $(2+\delta)\sqrt{\log n \cdot \log\log n}$ approaches zero, while the expected number of twins of size $(2-\delta)\sqrt{\log n \cdot \log\log n}$ approaches infinity.

1. Introduction and the Main Result

For a given combinatorial structure S, there is intrinsic interest in finding the largest substructure of a given kind that is contained in S, or a set of two or more identical substructures of maximum size that are contained in S, or the size of the largest common substructure in two random structures S_1 and S_2. For instance, finding the length of a longest increasing subsequence of a random permutation is the famous Ulam problem (Sect. 6.5 of [4] is a basic survey of this subject). An early collection of problems featuring the largest common substructures in two combinatorial objects can be found in [1]. In a pioneering paper [5], Bryant, McKenzie, and Steel proved that the maximum size of a common subtree for two independent copies of a uniformly random binary tree with n leaves is likely to be of order $O(n^{1/2})$. Interestingly, the proof is based on a classic result that the length of the increasing subsequence of a uniformly random permutation is likely to be of order $O(n^{1/2})$. And a recent result of Aldous [2] is that the maximum size of this largest common subtree (maximum agreement subtree) of two independent uniform random binary trees on n leaves is likely to be of order at least n^{β}, where $\beta = (\sqrt{3}-1)/2 \approx 0.366$. Pittel [9] proved the bound $O(n^{1/2})$ for the maximum size of a subtree common to two independent copies of a random rooted tree, namely the terminal tree of a critical Galton–Watson branching process conditioned on the total number of leaves being n.

Published online: 27 July 2024

In [3], Bóna and Flajolet used a bivariate generating function approach to compute the probability that two randomly selected phylogenetic trees are isomorphic as unlabeled trees. The requested probability was found to be asymptotic to a decreasing exponential modulated by a polynomial factor. This line of research involved the study of the expected number of symmetries in such trees, which was initiated by McKeon in [8]; in [3], this number in large phylogenetic trees was found to obey a limiting Gaussian distribution.

Motivated by the cited work, we study the likely maximum size of two, almost identical, rooted (fringe) subtrees of a random rooted tree, rather than a pair of random trees. (A subtree rooted at a vertex v is defined as a subtree induced by v and the set of all the *descendants* of v.) Formally, we call two rooted trees almost identical, twins, if they have the same counts of vertices by their out-degrees. The twin subtrees are automatically vertex-disjoint. In this paper, the tree in question is the uniformly random (Cayley) rooted tree on vertex set $[n]$. Our main result is

Theorem 1.1. *Let $m_n(k)$ denote the expected number of twin pairs of subtrees, with k vertices each.* **(a)** *If $\delta > 0$ and*

$$K_n = \left\lfloor \exp\left((2+\delta)\sqrt{\log n \cdot \log(\log n)}\right)\right\rfloor,$$

then $\sum_{k\geq K_n} m_n(k) \to 0$. Consequently, with high probability, the largest size of a pair of twins is below K_n. **(b)** *If $\delta \in (0,2)$ and*

$$k_n = \left\lfloor \exp\left((2-\delta)\sqrt{\log n \cdot \log(\log n)}\right)\right\rfloor,$$

then $m_n(k_n) \to \infty$.

Note. **(i)** In fact, our proof of **(b)** consists of showing that even a *smaller* expected number of twins of maximum degree $\lfloor 2\log k_n/\log\log k_n\rfloor$ grows indefinitely as $n \to \infty$. **(ii)** Going out on a limb, we conjecture that the logarithm of the maximum twin size scaled by $\sqrt{\log n \cdot \log(\log n)}$ converges, in probability, to 2.

In conclusion, a reviewer indicated to us that the our problem is reminiscent of the study of distinct rooted subtrees in Flajolet, Sipala, and Steyaert [7] and the study of repeated subtrees in Ralaivaosaona and Wagner [10].

2. The Proof of Theorem 1.1

Given k, $\mathbf{r} = (r_0, r_1, \dots)$ is a sequence of counts of vertices by out-degree of a rooted tree on $[k]$ if and only if

$$\sum_i r_i = k, \quad \sum_i i r_i = k-1, \tag{2.1}$$

and $M(\mathbf{r})$, the number of such trees, is given by

$$M(\mathbf{r}) = \frac{(k-1)!}{\prod\limits_{j\geq 0} (j!)^{r_j}} \binom{k}{r_0, r_1, \dots}, \tag{2.2}$$

Stanley [11]. It follows that the total number of pairs of rooted subtrees, on two respective vertex-disjoint vertex sets, each of cardinality k, that have the same counts of vertices by out-degrees, is

$$N(k) = \sum_{\mathbf{r} \text{ meets } (2.1)} M^2(\mathbf{r}). \tag{2.3}$$

Note. If we consider only subtrees of maximum out-degree (strictly) below $d\,(\geq 2)$ then the index i in the condition (2.1) ranges from 0 to $d-1$.

Let $S_n(k)$ stand for the total number of pairs of vertex-disjoint rooted subtrees, each with k vertices, in all the rooted trees on $[n]$ that share the counts of vertices by out-degree. Then

$$S_n(k) = \binom{n}{k}\binom{n-k}{k} N(k)(n-2k)^{n-2k-1}(n-2k)^2. \tag{2.4}$$

And the expected number of the pairs of twins of size k each is given by $m_n(k) = \frac{S_n(k)}{n^{n-1}}$.

The challenge is to sharply estimate $N(k)$ given by (2.3).

Lemma 2.1. *Let $H(z) := \sum_{r\geq 0} \frac{z^r}{(r!)^2}$. Then for all $x_s > 0$, we have*

$$N(k) \leq \frac{\big(k!(k-1)!\big)^2}{(2\pi)^2 x_1^k x_2^{k-1}} \int\limits_{\xi_s \in(-\pi,\pi)} \prod_{j\geq 0} \left| H\Big(\tfrac{z_1 z_2^j}{(j!)^2}\Big)\right| d\xi_1 d\xi_2,\ \ z_s = x_s e^{i\xi_s}. \tag{2.5}$$

Proof. Using (2.2) for $M(\mathbf{r})$, and a generating function $H(z) := \sum_{r\geq 0} \frac{z^r}{(r!)^2}$, we evaluate

$$\begin{aligned}
N(k) = \sum_{\mathbf{r} \text{ meets } (2.1)} M^2(\mathbf{r}) &= \big(k!(k-1)!\big)^2 \sum_{\mathbf{r} \text{ meets } (2.1)} \Bigg(\prod_{j\geq 0} \frac{1}{(j!)^{r_j} r_j!}\Bigg)^2 \\
&= \big(k!(k-1)!\big)^2 [x_1^k x_2^{k-1}] \sum_{\mathbf{r}\geq \mathbf{0}} x_1^{\sum_\alpha r_\alpha} x_2^{\sum_\beta \beta\cdot r_\beta} \Bigg(\prod_{j\geq 0} \frac{1}{(j!)^{r_j} r_j!}\Bigg)^2 \\
&= \big(k!(k-1)!\big)^2 [x_1^k x_2^{k-1}] \sum_{\mathbf{r}\geq \mathbf{0}} \prod_{j\geq 0} \frac{(x_1 x_2^j)^{r_j}}{\big((j!)^{r_j} r_j!\big)^2} \\
&= \big(k!(k-1)!\big)^2 [x_1^k x_2^{k-1}] \prod_{j\geq 0} \sum_{r_j\geq 0} \frac{(x_1 x_2^j)^{r_j}}{\big((j!)^{r_j} r_j!\big)^2} \\
&= \big(k!(k-1)!\big)^2 [x_1^k x_2^{k-1}] \prod_{j\geq 0} H\big(\tfrac{x_1 x_2^j}{(j!)^2}\big).
\end{aligned} \tag{2.6}$$

If we count only the twins of maximum out-degree below d, then in the above product j ranges from 0 to $d-1$.

What can we get from (2.6) with a bare-minimum analytical work? One possibility is to use a Chernoff-type inequality implied by (2.6). Note that for a power series $p(x, y) = \sum_{i,j} a_{i,j} x^i y^j$ with nonnegative real coefficients, the inequality $a_{i,j} x^i y^j \leq p(x, y)$ holds for all (i, j) and all (x, y) where $x > 0$ and

$y > 0$. Dividing by $x^i y^j$ and taking the infimum on the right-hand side, we get that

$$a_{i,j} \le \inf_{x>0,y>0} \frac{p(x,y)}{x^i y^j}.$$

Applying this inequality for the power series defined by the last line of (2.6), we get

$$N(k) \le \inf_{x_1>0,x_2>0} \tfrac{(k!(k-1)!)^2}{x_1^k x_2^{k-1}} \prod_{j\ge 0} H\big(\tfrac{x_1 x_2^j}{(j!)^2}\big). \tag{2.7}$$

Observe that by the elementary inequality $\binom{2r}{r} \le 2^{2r}$, for $x > 0$

$$H(x) \le \sum_{r\ge 0} \tfrac{x^r 2^{2r}}{(2r)!} \le \sum_{\ell\ge 0} \tfrac{(2x^{1/2})^\ell}{\ell!} = \exp(2x^{1/2}). \tag{2.8}$$

So, (2.7) yields

$$\begin{aligned} N(k) &\le \inf_{x_1>0,x_2>0} \tfrac{(k!(k-1)!)^2}{x_1^k x_2^{k-1}} \exp\Bigg(2\sum_{j\ge 0} \tfrac{(x_1 x_2^j)^{1/2}}{j!}\Bigg) \\ &= \inf_{x_1>0,x_2>0} \tfrac{(k!(k-1)!)^2}{x_1^k x_2^{k-1}} \exp\big(2x_1^{1/2} e^{x_2^{1/2}}\big). \end{aligned}$$

A straightforward computation shows that the k-dependent factor in the RHS function attains its minimum at $x_1 = k^2 \exp\big(-\frac{2(k-1)}{k}\big)$, $x_2 = \big(\frac{k-1}{k}\big)^2$, and consequently

$$N(k) = O\big(\tfrac{(k!)^4 e^{4k}}{k^{2(k+1)}}\big). \tag{2.9}$$

The last estimate is too crude, but the above values of x_1 and x_2 will be used when we switch from Chernoff's bound to a bivariate Cauchy integral, taken on the Cartesian product of two cycles, i.e. counterclock-wise oriented boundaries of two circles, of radii x_1 and x_2, respectively. Essentially we will use a bivariate version of the saddle point method, with the above discussion allowing us to guess that (x_1, x_2) is a promising approximation of the saddle point itself.

Here is the switch. By (2.6) and the Cauchy formula we have

$$N(k) = \frac{\big(k!(k-1)!\big)^2}{(2\pi i)^2} \oint_{C_1\times C_2} \tfrac{1}{z_1^{k+1} z_2^k} \prod_{j\ge 0} H\Big(\tfrac{z_1 z_2^j}{(j!)^2}\Big)\, dz_1 dz_2, \tag{2.10}$$

$C_s = \big\{z_s \in \mathbb{C} : z_s = x_s e^{i\xi_s},\ \xi_s \in (-\pi,\pi)\big\}$, $s = 1, 2$. It follows that for all $x_s > 0$, the claim of the lemma holds. $\square$

Let z be a complex number so that $\arg(z) \in (-\pi, \pi]$, and set $z^{1/2} := |z|^{1/2} \exp\big(\frac{1}{2}\arg(z)\big)$.

Lemma 2.2. *For $H(z) = \sum_{r\ge 0} \frac{z^r}{(r!)^2}$, with z and $z^{1/2}$ as above, the inequality*

$$|H(z)| \le \tfrac{|\exp(2z^{1/2})|}{\max(1,\alpha|z|^{1/4})}$$

holds, where $\alpha > 0$ is an absolute constant.

Proof. Introduce $I(z) := \sum_{r=0}^{\infty} \frac{z^{2r}}{(r!)^2}$, $\arg(z) \in (-\pi/2, \pi/2]$, so that $H(z^2) = I(z)$. It suffices to prove that, for $\arg(z) \in (-\pi/2, \pi/2]$,

$$|I_0(2z)| \le \frac{\left|\exp(2z)\right|}{\max(1, \alpha|z|^{1/2})}, \quad I_0(\xi) = \sum_{j \ge 0} \frac{\left(\frac{\xi^2}{4}\right)^j}{(j!)^2}, \tag{2.11}$$

where $I_0(\xi)$ is the zero-order, modified, Bessel function, since $H(z^2) = I_0(2z)$. Here is a surprisingly simple integral formula for $I_0(2z)$ that does the job:

$$I_0(2z) = \pi^{-1} \exp(2z) \int_0^4 \frac{\exp(-zy)}{\sqrt{y(4-y)}}\, dy, \quad \arg(z) \in (-\pi/2, \pi/2]. \tag{2.12}$$

Equation (2.12) follows from identity 10.32.2 in [6], that is,

$$I_0(2z) = \tfrac{1}{\pi} \int_{-1}^{1} (1-t^2)^{-1/2} e^{-2tz}\, dt,$$

via substitution $y = 2t + 2$. (Our original proof of (2.12) was from scratch. We are grateful to the reviewer for the above reference.) Define

$$G(z) = \int_0^4 \frac{e^{-zy}}{\sqrt{y(4-y)}} dy, \tag{2.13}$$

so that $I_0(2z) = \pi^{-1} \exp(2z) G(z)$, whence $|I_0(2z)| = \pi^{-1} |\exp(2z)| \cdot |G(z)|$. Here for $\Re z \ge 0$

$$|G(z)| \le \int_0^4 \frac{1}{\sqrt{y(4-y)}}\, dy = \pi.$$

So, by (2.12), $|I_0(2z)| \le \pi^{-1} |\exp(2z)| \pi = |\exp(2z)|$. Next, let us show that for $g(z) := |G(z)|/G(|z|)$ we have $\sup_{\mathbb{H}} g(z) < \infty$, where $\mathbb{H} := \{z : \Re z \ge 0\}$. The function g is well-defined and continuous in $\mathbb{H}$, since $G(z)$ is continuous in $\mathbb{H}$ and $G(|z|) > 0$.

We are left with showing $\sup_{z \in \mathbb{H}: |z| \ge 1} g(z) < \infty$. Noting that $G(\overline{z}) = \overline{G(z)}$, it suffices to show this boundedness in the closed fourth quadrant Q_{IV}. For such z we have

$$G(z) = \int_{\mathcal{C}_a} \frac{e^{-zy}}{\sqrt{y(4-y)}} dy, \tag{2.14}$$

where $a > 0$ and $\mathcal{C}_a$ is the oriented polygonal line joining the points $y_1 = 0, y_2 = a\sqrt{i}, y_3 = a\sqrt{i} + 4$, and $y_4 = 4$. Since $|e^{-zy}| \le 1$ in $\mathbb{H}$ and $\min\{|y|, |4-y|\} \ge \Im y = 2^{-1/2} a$, the integral over the segment (y_2, y_3) is bounded by $4\sqrt{2}/a$, hence it goes to zero as $a \to \infty$. Passing to the limit $a \to \infty$, we thus get

$$\begin{aligned} G(z) &= \int_0^{\infty e^{i\pi/4}} \frac{e^{-zy}}{\sqrt{y(4-y)}} dy - \int_4^{\infty e^{i\pi/4} + 4} \frac{e^{-zy}}{\sqrt{y(4-y)}} dy \\ &= \int_0^{\infty e^{i\pi/4}} \left(\frac{e^{-zu}}{\sqrt{u(4-u)}} - i \frac{e^{-4z - zu}}{\sqrt{(4+u)u}} \right) du. \end{aligned} \tag{2.15}$$

In the last integral we have $|\sqrt{u(4-u)}| \geq 2^{3/4}\sqrt{|u|}$, $|\sqrt{u(4+u)}| \geq 2\sqrt{|u|}$, $|e^{-4z}| \leq 1$ and (since $z \in Q_{IV}$ and $\arg u = \frac{\pi}{4}$), $|e^{-zu}| \leq e^{-2^{-1/2}|z||u|}$. Since $2^{-3/4} + 2^{-1} < 2$, we get

$$|G(z)| \leq 2\int_0^\infty u^{-1/2}e^{-2^{-1/2}|z|u}\,du = 2^{5/4}\sqrt{\pi}|z|^{-1/2}. \tag{2.16}$$

Importantly, we need this bound only for $|z| \geq 1$. For those z's, we have

$$G(|z|) \geq \tfrac{1}{2}\int_0^4 \frac{e^{-|z|y}}{\sqrt{y}}dy \geq \tfrac{1}{2}|z|^{-1/2}\int_0^4 \eta^{-1/2}e^{-\eta}\,d\eta. \tag{2.17}$$

Combining (2.16) and (2.17), we see that $\sup_{z\in\mathbb{H}:|z|\geq 1}\frac{|G(z)|}{G(|z|)} < \infty$. Therefore

$$|I_0(2z)| = O\big(|\exp(2z)|\cdot G(|z|)\big) = O\big(|\exp(2z)|\cdot|z|^{-1/2}\big),$$

uniformly for $z \in \{\mathbb{H} : |z| \geq 1\}$. This bound and the inequality $|I_0(2z)| \leq |\exp(2z)|$, $\Re z \geq 0$, complete the proof of the lemma. $\square$

2.1. Proof of Theorem 1.1, Part (a)

We use the lemma 2.2 to prove that, for

$$\delta > 0, \quad K_n = \left\lfloor \exp\big((2+\delta)\sqrt{\log n\cdot\log(\log n)}\big)\right\rfloor,$$

we have $\sum_{k\geq K_n} m_n(k) \to 0$, where $m_n(k)$ is the expected number of pairs of twins of size k.

We turn back to the integrand in (2.5). With $z_s = x_s e^{i\xi_s}$, and $X_j := \frac{x_1 x_2^j}{(j!)^2}$, by Lemma 2.2 and (2.8), we have

$$\prod_{j\geq 0}\left|H\left(\tfrac{z_1 z_2^j}{(j!)^2}\right)\right| \leq \prod_{j\geq 0}\frac{\exp(2X_j^{1/2})}{\max(1,\alpha|X_j|^{1/4})}\exp\Big(X_j\big(\cos(\xi_1+j\xi_2)-1\big)\Big). \tag{2.18}$$

Let us look closely at $\{X_j\}$. Using $j! \leq \left(\frac{j+1}{2}\right)^j$, $j \geq 1$, we have

$$X_j \geq x_1\cdot\frac{x_2^j}{\left(\frac{j+1}{2}\right)^{2j}} = x_1\exp\Big(2j\log\frac{2x_2^{1/2}}{j+1}\Big) = x_1\exp\Big(2j\log\frac{2^{\frac{k-1}{k}}}{j+1}\Big).$$

The function $\eta\log\frac{2^{\frac{k-1}{k}}}{\eta+1}$ is decreasing for $\eta \geq 1$. Pick $\varepsilon \in (0,1/2)$ and introduce $J = \left\lfloor\varepsilon\frac{\log x_1}{\log(\log x_1)}\right\rfloor$. Then, uniformly for $j \leq J$,

$$\begin{aligned} X_j &\geq x_1\exp\left(2J\log\frac{2^{\frac{k-1}{k}}}{J+1}\right) = x_1\exp(-2J\log J + O(J)) \\ &= \exp\big((1-2\varepsilon)\log x_1 + o(\log x_1)\big) \to \infty. \end{aligned} \tag{2.19}$$

Further

$$\sum_{j=1}^{J-1} 2j\log\frac{2^{\frac{k-1}{k}}}{j+1} \geq 2\int_1^J \eta\log\frac{2^{\frac{k-1}{k}}}{\eta+1}\,d\eta \geq -J^2\log J;$$

the last inequality follows easily from integration by parts. Therefore

$$\begin{aligned}\prod_{j<J} X_j &\geq x_1^J \exp(-J^2 \log J) \geq \exp\big[J(\log x_1 - J\log J)\big]\\ &\geq \exp\Big(\varepsilon'(1-\varepsilon')\tfrac{\log^2 x_1}{\log(\log x_1)}\Big)\\ &= x_1^{\varepsilon'(1-\varepsilon')\frac{\log x_1}{\log(\log x_1)}}, \quad \forall \varepsilon' < \varepsilon. \end{aligned}\tag{2.20}$$

The point of this bound is that the product in question exceeds x_1 raised to a sub-logarithmic power, still approaching infinity when k does. It follows from (2.20) that

$$\begin{aligned}\prod_{j\leq J} \tfrac{\exp(2X_j^{1/2})}{\max(1,\alpha|X_j|^{1/4})} &\leq \exp\Big(-\varepsilon'(1-\varepsilon')\tfrac{\log^2 x_1}{4\log(\log x_1)} + J\log\alpha\Big)\exp\Big(2\sum_{j\leq J} X_j^{1/2}\Big)\\ &= \exp\Big(-\varepsilon'(1-\varepsilon')\tfrac{\log^2 x_1}{4\log(\log x_1)} + O(\varepsilon\log x_1)\Big)\exp\Big(2\sum_{j<J} X_j^{1/2}\Big).\end{aligned}$$

Since $\frac{\exp(2X_j^{1/2})}{\max(1,\alpha|X_j|^{1/4})} \leq \exp(2X_j^{1/2})$ for all j, the above inequality implies that for $\varepsilon'' < \varepsilon'$,

$$\begin{aligned}\prod_{j\geq 0} \tfrac{\exp(2X_j^{1/2})}{\max(1,\alpha|X_j|^{1/4})} &\leq \exp\Big(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 x_1}{4\log(\log x_1)}\Big)\exp\Big(2\sum_{j\geq 0} X_j^{1/2}\Big)\\ &= \exp\Big(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 x_1}{4\log(\log x_1)}\Big)\exp\big(2x_1^{1/2}\exp(x_2^{1/2})\big)\\ &\leq e^{2k}\cdot\exp\Big(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 k}{\log(\log k)}\Big),\end{aligned}\tag{2.21}$$

as $x_1 = k^2\exp\big(-\frac{2(k-1)}{k}\big)$, $x_2 = \big(\frac{k-1}{k}\big)^2$.

Dropping the factors $\exp\Big(X_j\big(\cos(\xi_1 + j\xi_2) - 1\big)\Big)(\leq 1)$ with $j > 1$ in (2.18), and using (2.21), we obtain

$$\begin{aligned}\prod_{j\geq 0}\Big|H\Big(\tfrac{z_1 z_2^j}{(j!)^2}\Big)\Big| &\leq e^{2k}\cdot\exp\Big(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 k}{\log(\log k)}\Big)\\ &\quad\times\exp\Big(x_1(\cos\xi_1 - 1) + x_1x_2\big(\cos(\xi_1+\xi_2)-1\big)\Big),\\ &\quad \xi_j \in (-\pi,\pi).\end{aligned}\tag{2.22}$$

To use this last estimate on the right-hand side of (2.5), we need to find an upper bound for the *integral* of the exponential factor above. Let $I = (-\pi,\pi)$. To obtain the needed upper bound, let us set

$$f(\xi_1,\xi_2) := x_1(\cos(\xi_1) - 1) + x_1x_2(\cos(\xi_1+\xi_2) - 1).$$

Clearly, $f(\xi_1,\xi_2) \leq 0$, with equality holding if and only if $\cos(\xi_1) = 1$, and $\cos(\xi_1+\xi_2) = 1$, where $\xi_i \in I$, or equivalently if and only if $\xi_1 = \xi_2 = 0$.

For $\max\{|\xi_1|,|\xi_2|\} \geq \pi/2$, we have

$$f(\xi_1,\xi_2) \leq -ak^2 < 0,\tag{2.23}$$

for some positive constant a. For $\max\{|\xi_1|,|\xi_2|\} \le \pi/2$, using the Taylor expansion of cos, we obtain

$$f(\xi_1,\xi_2) \le -b[x_1\xi_1^2 + x_1x_2(\xi_1+\xi_2)^2] \le -ck^2[\xi_1^2+\xi_2^2]. \tag{2.24}$$

Combining (2.23) and (2.24), we get

$$f(\xi_1,\xi_2) \le -dk^2[\xi_1^2+\xi_2^2], \tag{2.25}$$

for $\xi_i \in I$, for $i = 1, 2$, and some fixed constant $d > 0$.

Therefore,

$$\begin{aligned}
&\int_{I\times I} \exp(f(\xi_1,\xi_2))\; d\xi_1 d\xi_2 \\
&\quad = \int\limits_{I\times I} \exp[x_1(\cos(\xi_1)-1) + x_1x_2(\cos(\xi_1+\xi_2)-1)]\, d\xi_1 d\xi_2 \\
&\quad \le \int\limits_{I\times I} \exp(-dk^2(\xi_1^2+\xi_2^2))\; d\xi_1 d\xi_2 \\
&\quad = \int\limits_{I} \exp(-dk^2\xi_1^2)\, d\xi_1 \cdot \int\limits_{I} \exp(-dk^2\xi_2^2)\, d\xi_2 \le \frac{\pi}{dk^2}.
\end{aligned} \tag{2.26}$$

Putting together (2.5), (2.22) and the last inequality, we conclude that

$$\begin{aligned}
N(k) &= \exp\Big(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 k}{\log(\log k)}\Big) O\Big(\frac{(k!)^4 e^{2k}}{x_1^k x_2^{k-1} k^4}\Big) \\
&= \exp\Big(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 k}{\log(\log k)}\Big) O\Big(\tfrac{(k!)^4 e^{4k}}{k^{2k+4}}\Big).
\end{aligned}$$

This bound and (2.4) imply that

$$\begin{aligned}
S_n(k) &= O\left(\binom{n}{k}\binom{n-k}{k}(n-2k)^{n-2k+1}N(k)\right) \\
&= \exp\Big(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 k}{\log(\log k)}\Big) O\Big(\tfrac{n!(n-2k)^{n-2k+1}(k!)^2 k^{-2k-4} e^{4k}}{(n-2k)!}\Big).
\end{aligned}$$

Using Stirling's formula for factorials we obtain that

$$S_n(k) = \exp\Big(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 k}{\log(\log k)}\Big) O\Big(\tfrac{n^{n+1}}{k^3}\Big).$$

Dividing this bound by the number n^{n-1} of all rooted Cayley trees, we get that the expected number of the pairs of twins of size k, i.e. $m_n(k) = \frac{S_n(k)}{n^{n-1}}$, is

$$m_n(k) = O\bigg(\tfrac{n^2}{k^3} \exp\Big(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 k}{\log(\log k)}\Big)\bigg), \quad \forall \varepsilon'' < \varepsilon < 1/2. \tag{2.27}$$

Since the series $\sum_{k\ge 1} k^{-3}$ converges, it follows that for $k_n=\exp\Bigl((2+\delta)\sqrt{\log n\cdot \log(\log n)}\Bigr)$ and $\delta>0$ we have

$$\sum_{k\ge k_n} m_n(k)=O\Bigl(n^2\exp\Bigl(-\tfrac{\varepsilon''(1-\varepsilon'')\log^2 k_n}{\log(\log k_n)}\Bigr)\Bigr)\sum_{k\ge 1}k^{-3}$$
$$=O\exp\Bigl(2\bigl[1-(2+\delta)^2\varepsilon''(1-\varepsilon'')\bigr]\log n+o(\log n)\Bigr), \quad (2.28)$$

uniformly for all $\varepsilon''<1/2$. Picking ε'' sufficiently close to $1/2$, we conclude that $\lim_{n\to\infty}\sum_{k\ge k_n} m_n(k)=0$.

2.2. Proof of Theorem 1.1, Part (b)

The claim is: If $\delta\in(0,2)$, $k_n:=\Bigl\lfloor\exp\Bigl((2-\delta)\sqrt{\log n\cdot\log(\log n)}\Bigr)\Bigr\rfloor$, then $m_n(k_n)\to\infty$.

Proof. From here on we can, and will consider the twins of maximum out-degree $<d=d_n$, with d_n yet to be chosen. Let us upper-bound the contribution to the RHS of (2.5) for $k=k_n$, coming from $(z_1=x_1e^{i\xi_1}, z_2=x_2e^{i\xi_2})$, where $x_1=k^2\exp\bigl(-\frac{2(k-1)}{k}\bigr)$, $x_2=\bigl(\frac{k-1}{k}\bigr)^2$, and also $\|\boldsymbol{\xi}\|\le k^{-1/2+\varepsilon}$, $\varepsilon\in(0,1/2)$. By Lemma 2.2,

$$\prod_{j<d}\Bigl|H\Bigl(\tfrac{z_1z_2^j}{(j!)^2}\Bigr)\Bigr|\le\exp\Bigl(2x_1^{1/2}\sum_{j<d}\tfrac{x_2^{j/2}}{j!}\Re\bigl(\exp(i(\xi_1/2+j\xi_2/2)\bigr)\Bigr)$$
$$\le\exp\Bigl(2x_1^{1/2}\sum_{j<d}\tfrac{x_2^{j/2}}{j!}\Re\bigl(\exp(i(\xi_1/2+j\xi_2/2)\bigr)+O(x_1^{1/2}/d!)\Bigr)$$
$$=\exp\bigl(2x_1^{1/2}W(\xi_1,\xi_2)+O(x_1^{1/2}x_2^{d/2}/d!)\bigr),$$

where

$$W(\xi_1,\xi_2):=\cos\bigl(\tfrac{\xi_1}{2}+x_2^{1/2}\sin\tfrac{\xi_2}{2}\bigr)\exp\bigl(x_2^{1/2}\cos\tfrac{\xi_2}{2}\bigr).$$

Since $x_1=\Theta(k^2)$, $x_2=\Theta(1)$, the big-O term is $o(1)$, if $d=\lfloor 2\log k/\log\log k\rfloor$, which we assume from now.

Clearly, $W(\xi_1,\xi_2)\le\exp(x_2^{1/2})=W(0,0)$, and it is easy to check that $(0,0)$ is a single stationary point of $W(\xi_1,\xi_2)$ in $[-\pi,\pi]^2$, whence it is a unique maximum point of $W(\xi_1,\xi_2)$ in this square. In addition,

$$W''_{\xi_1}(0,0)=-\tfrac{e^{x_2^{1/2}}}{4},\ W''_{\xi_2}(0,0)=-\tfrac{(x_2+x_2^{1/2})e^{x_2^{1/2}}}{4},\ W''_{\xi_1,\xi_2}(0,0)=-\tfrac{x_2^{1/2}e^{x_2^{1/2}}}{4},$$

so that $W''_{\xi_1}(0,0), W''_{\xi_2}(0,0)<0$, and

$$W''_{\xi_1}(0,0)\cdot W''_{\xi_2}(0,0)-\bigl(W''_{\xi_1,\xi_2}(0,0)\bigr)^2=\frac{x_2^{1/2}e^{2x_2^{1/2}}}{16}>0.$$

This implies that $W(\xi_1,\xi_2)$ is strictly concave in the vicinity of $(0,0)$, and moreover $W(\xi_1,\xi_2)\le W(0,0)-\beta(\xi_1^2+\xi_2^2)$, $\xi_s\in[-\pi,\pi]$, for a constant $\beta>0$ as $k\to\infty$ because $x_2=\Theta(1)$. So,

$$\prod_{j<d}\Bigl|H\Bigl(\tfrac{z_1z_2^j}{(j!)^2}\Bigr)\Bigr|\le\exp\bigl(2x_1^{1/2}e^{x_2^{1/2}}-\Theta(k\|\boldsymbol{\xi}\|^2)+o(1)\bigr).$$

Therefore,

$$\int\limits_{\|\xi\|\ge k^{-1/2+\varepsilon}} \tfrac{1}{z_1^{k+1}z_2^k}\prod_{j<d} H\Big(\tfrac{z_1z_2^j}{(j!)^2}\Big)\,dz_1dz_2 = O\bigg(\tfrac{\exp\big(2x_1^{1/2}e^{x_2^{1/2}}\big)}{x_1^kx_2^{k-1}}\int\limits_{\|\boldsymbol{\xi}\|\ge k^{-1/2+\varepsilon}} e^{-\Theta(k\|\xi\|^2)}\,d\xi_1d\xi_2\bigg)$$
$$= O\bigg(\tfrac{\exp\big(2k-\Theta(k^{2\varepsilon})\big)}{x_1^kx_2^{k-1}}\bigg). \tag{2.29}$$

It remains to sharply evaluate the contribution to the Cauchy integral coming from $\boldsymbol{\xi}$'s with $\|\boldsymbol{\xi}\|\le k^{-1/2+\varepsilon}$. By Lemma 2.2, we have

$$H(z)=\pi^{-1}\exp(2\sqrt{z})\int_0^4 \tfrac{\exp(-\sqrt{z}y)}{\sqrt{y(4-y)}}\,dy,\quad \arg(z)\in[-\pi,\pi].$$

And, for the values of z_1,z_2, and j in question, $\arg(z_1z_2^j)=\xi_1+j\xi_2=O(k^{-1/2+\varepsilon}d)=o(1)$. Consequently,

$$\tfrac{1}{z_1^kz_2^{k-1}}\prod_{j<d}H\Big(\tfrac{z_1z_2^j}{(j!)^2}\Big)=\tfrac{\pi^{-d}}{z_1^kz_2^{k-1}}\prod_{j<d}\exp\Big(\tfrac{2z_1^{1/2}z_2^{j/2}}{j!}\Big)\int_0^4\frac{\exp\Big(-\frac{z_1^{1/2}z_2^{j/2}}{j!}y_j\Big)}{\sqrt{y_j(4-y_j)}}dy_j. \tag{2.30}$$

Here

$$\tfrac{1}{z_1^kz_2^{k-1}}\prod_{j<d}\exp\Big(\tfrac{2z_1^{1/2}z_2^{j/2}}{j!}\Big)=\tfrac{\exp\big(2z_1^{1/2}e^{z_2^{1/2}}\big)+O(x_2^{d/2}/d!)}{z_1^kz_2^{k-1}}, \tag{2.31}$$

and

$$\begin{aligned}2z_1^{1/2}e^{z_2^{1/2}} &= 2x_1^{1/2}\exp(i\xi_1/2+x_2^{1/2}e^{i\xi_2/2})\\ &= 2x_1^{1/2}e^{x_2^{1/2}}\exp\big(i\xi_1/2+x_2^{1/2}(e^{i\xi_2/2}-1)\big)\\ &= 2x_1^{1/2}e^{x_2^{1/2}}\big[1+(i\xi_1/2+x_2^{1/2}(e^{i\xi_2/2}-1))\\ &\quad+\tfrac12(i\xi_1/2+x_2^{1/2}(e^{i\xi_2/2}-1))^2+O(\|\boldsymbol{\xi}\|^3)\big]\\ &= 2x_1^{1/2}e^{x_2^{1/2}}\big[1+i(\xi_1/2+x_2^{1/2}\xi_2/2)\\ &\quad-\xi_1^2/8-(x_2+x_2^{1/2})\xi_2^2/8-\xi_1\xi_2x_2^{1/2}/4+O(\|\boldsymbol{\xi}\|^3)\big].\end{aligned}$$

Therefore,

$$\begin{aligned}2z_1^{1/2}e^{z_2^{1/2}}-k\log z_1-(k-1)\log z_2 &= 2x_1^{1/2}e^{x_2^{1/2}}-k\log x_1-(k-1)\log x_2\\ &\quad+i\xi_1(x_1^{1/2}e^{x_2^{1/2}}-k)\\ &\quad+i\xi_2(x_1^{1/2}e^{x_2^{1/2}}x_2^{1/2}-(k-1))\\ &\quad-\tfrac14x_1^{1/2}e^{x_2^{1/2}}(\xi_1^2+(x_2+x_2^{1/2})\xi_2^2-2\xi_1\xi_2x_2^{1/2})\\ &\quad+O(k\|\boldsymbol{\xi}\|^3).\end{aligned}$$

Recalling that $\|\boldsymbol{\xi}\|\le k^{-1/\varepsilon}$, we see that $O(k\|\boldsymbol{\xi}\|^3)=O(k^{-1/2+3\varepsilon})\to 0$, if $\varepsilon<1/6$, which we assume from now. Furthermore, the linear combination of $i\xi_1$ and $i\xi_2$ in the above sum disappears, thanks to

$$x_1^{1/2}e^{x_2^{1/2}}-k=0,\quad x_1^{1/2}e^{x_2^{1/2}}x_2^{1/2}-(k-1)=0,$$

the conditions we were led to in our preliminary attempt to bound $N(k)$. So, we have

$$2z_1^{1/2}e^{z_2^{1/2}} - k\log z_1 - (k-1)\log z_2 = 2k - k\log x_1 - (k-1)\log x_2 \\ -\tfrac{k}{4}\Big(\xi_1^2 + (x_2 + x_2^{1/2})\xi_2^2 \\ -2\xi_1\xi_2 x_2^{1/2}\Big) + O(k\|\boldsymbol{\xi}\|^3).$$

Therefore, Eq. (2.31) becomes

$$\tfrac{1}{z_1^k z_2^{k-1}}\prod_{j<d}\exp\big(\tfrac{2z_1^{1/2}z_2^{j/2}}{j!}\big) = \tfrac{e^{2k}}{x_1^k x_2^{k-1}}\exp\Big[-\tfrac{k}{4}\Big(\xi_1^2 + (x_2 + x_2^{1/2})\xi_2^2 \\ -2\xi_1\xi_2 x_2^{1/2}\Big) + O(k\|\boldsymbol{\xi}\|^3 + e^{-k})\Big]. \tag{2.32}$$

The quadratic form is negative definite, since $x_2 = \Theta(1)$. We still have to evaluate

$$\prod_{j<d}\int_0^4 \frac{\exp\big(-Z_j y_j\big)}{\sqrt{y_j(4-y_j)}}\,dy_j, \\ Z_j = X_j\exp(i(\xi_1/2 + j\xi_2/2)),\quad X_j := \frac{k\exp\Big(-\frac{1}{2}\big(\frac{k-1}{k}\big)^2\Big)}{j!}\big(\tfrac{k-1}{k}\big)^j, \tag{2.33}$$

where $d = \lfloor 2\log k/\log\log k\rfloor$, $\|\boldsymbol{\xi}\| \le k^{-1/2+\varepsilon}$, $\varepsilon < 1/6$.

(a) Suppose that $j < j(k) := \lfloor\frac{\log k}{\log\log k}\rfloor$. Then

$$\tfrac{k}{j!} \ge \frac{k}{j(k)!} = \exp\Big(\Theta\big(\tfrac{\log k\cdot\log\log\log k}{\log\log k}\big)\Big) \to \infty.$$

So, using

$$I_0(z) \sim \tfrac{e^z}{(2\pi z)^{1/2}}\big(1 + O(z^{-1})\big),\quad |\arg(z)| \le \tfrac{\pi}{2} - \delta,\ \delta\in(0,\pi/2),$$

(see [6]), we have

$$\begin{aligned}\int_0^4 \frac{\exp\big(-Z_j y_j\big)}{\sqrt{y_j(4-y_j)}}\,dy_j &= \pi\exp(-2Z_j)I_0(2Z_j)\\ &= \pi\exp(-2Z_j)\tfrac{e^{2Z_j}}{(4\pi Z_j)^{1/2}}\big(1 + O(Z_j^{-1})\big)\\ &= \tfrac{\pi}{(4\pi Z_j)^{1/2}}\big(1 + O(Z_j^{-1})\big)\\ &= \tfrac{\pi}{(4\pi X_j)^{1/2}}\cdot\Big[1 + O\Big(\exp\Big(-\Theta\big(\tfrac{\log k\cdot\log\log\log k}{\log\log k}\big)\Big)\Big)\Big],\end{aligned}$$

implying that

$$\prod_{j<j(k)}\int_0^4 \frac{\exp\big(-Z_j y_j\big)}{\sqrt{y_j(4-y_j)}}\,dy_j = (1 + O(\Delta_k))\prod_{j<j(k)}\tfrac{\pi}{(4\pi X_j)^{1/2}}, \\ \Delta_k := \tfrac{\log k}{\log\log k}\exp\Big(-\Theta\big(\tfrac{\log k\cdot\log\log\log k}{\log\log k}\big)\Big).$$

Here, by (2.33), $X_j \le \frac{k}{j!}$. So, using Stirling's formula with the remainder term, namely $j! = (2\pi j)^{1/2}\left(\frac{j}{e}\right)^j[1+O(1/(j+1))]$, we obtain

$$\begin{aligned}\prod_{j<j(k)} \tfrac{\pi}{(4\pi X_j)^{1/2}} &= \exp[O(j(k)\log j(k))]\cdot k^{-j(k)/2}\prod_{j<j(k)}\left(\tfrac{j}{e}\right)^{j/2}\\ &= \exp[O(j(k)\log j(k))]\cdot k^{-j(k)/2}\\ &\quad\cdot\exp\left(\tfrac{1}{2}\int_0^{j(k)} x\log(x/e)\,dx\right)\\ &= \exp\Big(-\tfrac{\log^2 k}{4\log\log k}+O\big(\tfrac{\log^2 k}{\log^2(\log k)}\big)\Big). \qquad (2.34)\end{aligned}$$

(b) Suppose that $j \ge j(k)$. Then, since $j<d$,

$$\begin{aligned}&\tfrac{k}{j!}\times\big|\exp(i(\xi_1/2+j\xi_2/2))-1\big| = O\Big(\tfrac{k^{1/2+\varepsilon}d}{(j(k)/e)^{j(k)}}\Big)\\ &\quad= O\big(\exp(-\Theta(1/2-\varepsilon)\log k)\big).\end{aligned}$$

Here and below we use the notation $\Theta(1/2-\varepsilon)$ as a shorthand for a quantity which is bounded below by $1/2-\varepsilon$ times an absolute constant. So, we have

$$Z_j y_j = X_j y_j + O\big(d\|\boldsymbol{\xi}\|X_j y_j\big).$$

Hence,

$$\begin{aligned}\sum_{j(k)\le j<d} Z_j y_j &= \sum_{j(k)\le j<d} X_j y_j + O\big(d^2\|\boldsymbol{\xi}\|\tfrac{k}{j(k)!}\big)\\ &= \sum_{j(k)\le j<d} X_j y_j + O\big(\exp(-\Theta(1/2-\varepsilon)\log k)\big),\end{aligned}$$

implying that

$$\begin{aligned}\prod_{j(k)\le j<d}\int_0^4 \tfrac{\exp(-Z_j y_j)}{\sqrt{y_j(4-y_j)}}\,dy_j = &\prod_{j(k)\le j<d}\int_0^4 \tfrac{\exp(-X_j y_j)}{\sqrt{y_j(4-y_j)}}\,dy_j\\ &\cdot\big[1+O\big(\exp(-\Theta(1/2-\varepsilon)\log k)\big)\big].\end{aligned}$$

Here, using the convexity of the exponential function,

$$\pi^{-1}\int_0^4 \tfrac{\exp(-X_j y_j)}{\sqrt{y_j(4-y_j)}}\,dy_j \ge \exp\left(-\tfrac{X_j}{\pi}\int_0^4\sqrt{\tfrac{y}{4-y}}\,dy\right) = e^{-2X_j}.$$

So,

$$\begin{aligned}\prod_{j(k)\le j<d}\pi^{-1}\int_0^4 \tfrac{\exp(-X_j y_j)}{\sqrt{y_j(4-y_j)}}\,dy_j &\ge \exp\Big(-2\sum_{j\ge j(k)} X_j\Big)\\ &= \exp\big(O(k/j(k)!)\big)\\ &= k^{O\left(\frac{\log\log\log k}{\log\log k}\right)}. \qquad (2.35)\end{aligned}$$

Combining (2.34) and (2.35) we arrive at

$$\prod_{j<d}\int_0^4 \tfrac{\exp(-Z_j y_j)}{\sqrt{y_j(4-y_j)}}\,dy_j = \exp\Big(-\tfrac{\log^2 k}{4\log\log k}+O\big(\tfrac{\log^2 k}{\log^2(\log k)}\big)\Big). \qquad (2.36)$$

Thus, Eqs. (2.32) and (2.36) transform Eq. (2.30) into

$$\frac{1}{z_1^k z_2^{k-1}} \prod_{j<d} H\Big(\frac{z_1 z_2^j}{(j!)^2}\Big) = \frac{e^{2k}}{x_1^k x_2^{k-1}} \exp\Big(-\frac{\log^2 k}{4\log\log k} + O\big(\frac{\log^2 k}{\log^2(\log k)}\big)\Big) \\ \times \exp\big[-\tfrac{k}{4}\big(\xi_1^2 + (x_2 + x_2^{1/2})\xi_2^2 - 2\xi_1\xi_2 x_2^{1/2}\big)\big], \tag{2.37}$$

uniformly for $\boldsymbol{\xi}$ with $\|\boldsymbol{\xi}\| \le k^{-1/2+\varepsilon}$, $\varepsilon \in (0, 1/6)$.

By (2.37), the contribution to the Cauchy integral on the RHS of (2.10) is

$$\frac{e^{2k}}{x_1^k x_2^{k-1}} \exp\Big(-\frac{\log^2 k}{4\log\log k} + O\big(\frac{\log^2 k}{\log^2(\log k)}\big)\Big) \\ \times \int_{\|\boldsymbol{\xi}\| \le k^{-1/2+\varepsilon}} \exp\big[-\tfrac{k}{4}\big(\xi_1^2 + (x_2 + x_2^{1/2})\xi_2^2 - 2\xi_1\xi_2 x_2^{1/2}\big)\big]\, d\xi_1 d\xi_2 \\ = \frac{e^{2k}}{x_1^k x_2^{k-1}} \exp\Big(-\frac{\log^2 k}{4\log\log k} + O\big(\frac{\log^2 k}{\log^2(\log k)}\big)\Big).$$

In combination with (2.29), this yields that

$$\frac{1}{(2\pi i)^2} \oint_{C_1\times C_2} \prod_{j\ge 0} \frac{1}{z_1^{k+1} z_2^k} H\Big(\frac{z_1 z_2^j}{(j!)^2}\Big)\, dz_1 dz_2 = \frac{e^{4k}}{x_1^k x_2^{k-1}} \exp\Big(-\frac{\log^2 k}{4\log\log k} + O\big(\frac{\log^2 k}{\log^2(\log k)}\big)\Big) \\ = \frac{e^{4k}}{k^{2k}} \exp\Big(-\frac{\log^2 k}{4\log\log k} + O\big(\frac{\log^2 k}{\log^2(\log k)}\big)\Big).$$

Therefore, by (2.10), $N(k)$ (the total number of pairs of twin trees, each with k vertices, and maximum degree $\lfloor 2\log k/\log\log k\rfloor$ is given by

$$N(k) = (k!)^4 \cdot \frac{e^{4k}}{k^{2k}} \exp\Big(-\frac{\log^2 k}{4\log\log k} + O\big(\frac{\log^2 k}{\log^2(\log k)}\big)\Big).$$

Combining this with (2.4), we obtain that the expected number of twin trees of size $k = o(n)$ in the random Cayley tree with n vertices is at least

$$\frac{1}{n^{n-1}} \binom{n}{k}\binom{n-k}{k} (n-2k)^{n-2k+1} \cdot N(k) \\ = n^2 \exp\Big(-\frac{\log^2 k}{4\log\log k} + O\big(\frac{\log^2 k}{\log^2(\log k)}\big)\Big).$$

Observe that this lower bound depends on k almost like the upper bound (2.27), since ε'' can be chosen arbitrarily close to $1/2$ from *below*.) And the lower bound diverges to infinity if

$$k \le \Big\lfloor \exp\Big((2-\delta)\sqrt{\log n \cdot \log(\log n)}\Big)\Big\rfloor.$$

□

Acknowledgements

We sincerely thank the hard-working referees for helping us to improve the paper, both mathematically and stylistically.

Data availability No additional data was produced during the research that led to this article.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] D. Aldous, Largest common substructures in probabilistic combinatorics. Preprint, available at https://www.stat.berkeley.edu/~aldous/Research/OP/common_subs.pdf.

[2] D. Aldous, On the largest common subtree of random leaf-labeled binary trees *SIAM J. Discrete Math.* **36** (2022), no. 1, 299–314.

[3] M. Bóna, P. Flajolet, Isomorphism and symmetries in random phylogenetic trees. *J. Appl. Probab.* **46** (2009), no. 4, 1005–1019.

[4] M. Bóna, Combinatorics of Permutations, third edition. *CRC Press*, 2022.

[5] D. Bryant, A. McKenzie, M. Steel, The size of a maximum agreement subtree for random binary trees, in *BioConsensus*, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. **61** AMS, Providence, RI 2003, 55–65.

[6] Digital Library of Mathematical Functions, 10.32.2.

[7] P. Flajolet, P. Sipala, and J. M. Steyaert, Analytic variations on the common subexpression problem, in *Proceedings of the* 17*-th International Colloquium on Automata, Languages and Programming, ICALP 1990*, **443**, Lecture Notes in Computer Science, 220–234.

[8] K. A. McKeon, The expected number of symmetries in locally-restricted trees I, in: *Graph Theory and Combinatorics, and Applications*, Y. Alavi et al., eds., 849–860.

[9] B. Pittel, Expected number of induced subtrees shared by two independent copies of a random tree. *SIAM J. Discrete Math.* **37** (2023), no. 1, 1–16.

[10] D. Ralaivaosaona, S. G. Wagner, Repeated fringe subtrees in random rooted trees, in *Proceedings of the* 12*-th Workshop on Analytic Algorithms and Combinatorics*, ANALCO 2015, 78–88.

[11] R. Stanley, Enumerative Combinatorics, v. ii. *Cambridge University Press*, 1999.

Miklós Bóna
Department of Mathematics
University of Florida
Gainesville
FL 32611
USA
e-mail: bona@ufl.edu

Ovidiu Costin and Boris Pittel
Department of Mathematics
Ohio State, University
Columbus
OH 43210
USA
e-mail: costin.9@osu.edu

Boris Pittel
e-mail: pittel.1@osu.edu

Communicated by Frédérique Bassino
Received: 28 November 2023.
Accepted: 15 July 2024.