Leaf multiplicity in a Bienaymé–Galton–Watson tree

Anna M. Brandenberger, Luc Devroye, Marcel K. Goh, and Rosie Y. Zhao

School of Computer Science, McGill University

Abstract. This note defines a notion of multiplicity for nodes in a rooted tree and presents an asymptotic calculation of the maximum multiplicity over all leaves in a Bienaymé–Galton–Watson tree with critical offspring distribution ξ, conditioned on the tree being of size n. In particular, we show that if S_n is the maximum multiplicity in a conditioned Bienaymé–Galton–Watson tree, then $S_n = \Omega(\log n)$ asymptotically in probability and under the further assumption that $\mathbf{E}\{2^\xi\} < \infty$, we have $S_n = O(\log n)$ asymptotically in probability as well. Explicit formulas are given for the constants in both bounds. We conclude by discussing links with an alternate definition of multiplicity that arises in the root-estimation problem.

Keywords. Multiplicity, Bienaymé–Galton–Watson trees, automorphisms, Rényi entropy.

1. Introduction

Equivalence between two distinct mathematical objects is a far-reaching concept in mathematics. When two structures are similar, one may define a relation under which they are regarded as one and the same. The term "multiplicity" is often used to indicate the extent to which an object is, in some sense, not structurally unique (or how often it is repeated in a suitably-defined multiset). Towards a concept of the multiplicity of a node in a tree, consider the small example depicted in Fig. 1.

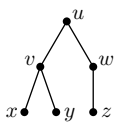

Fig. 1. A rooted tree with six nodes. The pair x and y are similar, but x and z are not.

Definitions and notation. Consider a tree T rooted at a node u. For a node v in the tree, we let T_v denote the subtree rooted at v. Let v and w be nodes in the tree and let $v = v_1, v_2, \ldots, v_n = u$ and $w = w_1, w_2, \ldots, w_m = u$ be the paths from v and w, respectively, to the root. We say that v and w are *identical* and write $v \equiv w$ if the paths have the same length and T_{v_j} and T_{w_j} are isomorphic as rooted ordered trees for $1 \le j \le n$. For example, in Fig. 1, the nodes x and y are identical in this sense, but different from node z.

It is clear that $\equiv$ defines an equivalence relation on the set of nodes in the tree, so we may now define the *multiplicity* $\sigma(v)$ of a node v to be the size of the equivalence class $[v]$ under the relation. The *leaf multiplicity* (or simply *multiplicity*, when no confusion can arise) $S(T)$ of a rooted tree T is the maximum value of $\sigma(v)$, taken over all nodes v of T. The name "leaf multiplicity" is motivated by the fact that the function σ increases monotonically away from the root, so that $S(T)$ remains the same when the maximum is only computed over the set of *leaves* of T.

Note that $\equiv$ is not the only structural equivalence relation one can define on the set of nodes in a tree, and thus σ is only one of many possible notions of leaf multiplicity. Towards the end of this paper, we will explore an alternate definition μ of multiplicity which extends well to free trees as well as rooted trees, and discuss the relationship between the two functions σ and μ.

The Bienaymé–Galton–Watson model. For a random variable ξ taking values in the nonnegative integers, a *Bienaymé–Galton–Watson tree* [2] is a rooted ordered tree in which every node has i children with probability $p_i = \mathbf{P}\{\xi = i\}$. We say that ξ is the *offspring distribution* of the tree. Trees arising from this process are often called *Galton–Watson trees*, but we include the name of I.-J. Bienaymé [3], since his work predates (and is more mathematically precise than) the analysis of F. Galton and H. W. Watson [7]; see [9] for an extended account of the history of branching processes. We shall deal with critical Bienaymé–Galton–Watson trees; that is, trees whose offspring distributions satisfy $\mathbf{E}\{\xi\} = 1$ and $\mathbf{V}\{\xi\} \in (0, \infty)$. This restriction on the variance ensures that $p_i \neq 1$, so that the tree is finite almost surely. The Bienaymé–Galton–Watson trees that we shall study are *conditioned Bienaymé–Galton–Watson trees* T_n. Such trees are conditioned on $|T| = n$, where $|T|$ is the number of nodes in the tree.

Rényi entropy. It will be convenient to simplify our notation with some information-theoretic definitions. Letting $p_i = \mathbf{P}\{\xi = i\}$, for $\alpha > 1$ we define the *Rényi entropy of order* α [17] (see also [8]) to be the value

$$\mathrm{H}_\alpha(\xi) = \frac{1}{\alpha - 1} \log_2 \frac{1}{\sum_{i \geq 0} {p_i}^\alpha}. \tag{1}$$

As $\alpha \to 1$, this approaches the *binary (Shannon) entropy* [18]

$$\mathrm{H}(\xi) = \sum_{i \geq 0} p_i \log_2 \frac{1}{p_i}.$$

Since ξ will be fixed throughout the paper, for brevity we will let $\mathrm{H}_\alpha = \mathrm{H}_\alpha(\xi)$ and $\mathrm{H} = \mathrm{H}(\xi)$.

Fix an offspring distribution ξ with mean 1 and nonzero finite variance; let T_n be a conditioned Bienaymé–Galton–Watson tree of size n with this offspring distribution. The leaf multiplicity $S(T_n)$ of this tree is a random variable, and will be denoted S_n. The main result of this paper gives bounds on S_n that are obeyed asymptotically in probability.

Theorem 1. *Let ξ be an offspring distribution with $\mathbf{E}\{\xi\} = 1$ and $\mathbf{V}\{\xi\} \in (0, \infty)$. If S_n is the multiplicity of a conditioned Bienaymé–Galton–Watson tree of size n with offspring distribution ξ, then letting*

$$\gamma = \max_{k \geq 2} {p_0}^k {p_k}^{k/(k-1)}, \tag{2}$$

we have for all $\epsilon > 0$,

$$\mathbf{P}\left\{ S_n \geq (1 - \epsilon) \frac{\log_2 n}{\log_2(1/\gamma)} \right\} \to 1$$

as $n \to \infty$, and under the further assumption that $\mathbf{E}\{2^\xi\} < \infty$, we have the upper bound

$$\mathbf{P}\left\{ S_n \leq (1 + \epsilon) \frac{2 \log_2 n}{\mathrm{H}_2} \right\} \to 1$$

as $n \to \infty$, where H_2 is the Rényi entropy of order 2 of the random variable ξ. ∎

This theorem will be proved as two separate lemmas in the next section.

2. Asymptotics of the leaf multiplicity

In this section we derive asymptotic upper and lower bounds on S_n. Before we begin, we observe that if $p_{\max} = \max_{i \geq 0} p_i$ and $1 < \alpha < \beta < \infty$, we have the inequalities

$$e^{-\mathrm{H}} \leq \Big(\sum_{i \geq 0} {p_i}^\alpha \Big)^{1/(\alpha-1)} \leq \Big(\sum_{i \geq 0} {p_i}^\beta \Big)^{1/(\beta-1)} \leq p_{\max} \leq \Big(\sum_{i \geq 0} {p_i}^\beta \Big)^{1/\beta} \leq \Big(\sum_{i \geq 0} {p_i}^\alpha \Big)^{1/\alpha} \leq 1, \tag{3}$$

and defining $\mathrm{H}_\infty = \log_2(1/p_{\max})$, we have the equivalent chain of inequalities

$$\mathrm{H} \geq \mathrm{H}_\alpha \geq \mathrm{H}_\beta \geq \mathrm{H}_\infty \geq \frac{\beta-1}{\beta}\mathrm{H}_\beta \geq \frac{\alpha-1}{\alpha}\mathrm{H}_\alpha \geq 0. \tag{4}$$

In particular, because we have assumed that $\mathbf{V}\{\xi\} \neq 0$, we have the strict inequality

$$\Big(\sum_{i\geq 0} {p_i}^k\Big)^{1/k} < \Big(\sum_{i\geq 0} {p_i}^2\Big)^{1/2} \tag{5}$$

for all $k > 2$.

We will prove the upper bound and lower bound for S_n as two separate lemmas.

Lemma 2. *Let ξ be an offspring distribution with mean 1 and nonzero finite variance σ^2. Suppose further that $\mathbf{E}\{2^\xi\}$ is finite. If S_n is the multiplicity of a conditioned Bienaymé–Galton–Watson tree of size n with offspring distribution ξ, then*

$$\mathbf{P}\{S_n > (1+\epsilon)2\log_2 n/\mathrm{H}_2\} \to 0 \tag{6}$$

for all $\epsilon > 0$, where H_2 is the Rényi entropy of order 2 of the random variable ξ.

Proof. For $1 \leq i \leq n$, let ξ_i denote the degree of the ith node in preorder in the tree T_n. For all $1 \leq t < n$, the partial sum $\sum_{i=1}^t \xi_i > t-1$ and $\sum_{i=1}^n \xi_i = n-1$. We will concentrate on the least common ancestor of the nodes in the largest equivalence class of T_n. This node, call it w, has the property that the nodes in the equivalence class belong to $k \geq 2$ different subtrees rooted at the children of w. The node w has (random) degree D, which we will deal with by summing over all possible degrees d. Let $\mathcal{A}_{wk}$ denote the collection of all subsets of size k of the children of w (naturally, this collection is empty if w has fewer than k children). For $x > 0$, a node w, integers $2 \leq k \leq d$, and a set $A \in \mathcal{A}_{wk}$, we let $E(x,w,k,A)$ be the event that all the nodes in A are identical and their subtree sizes are at least x/k. Now for integers $s \geq x/k$, we let $E'(x,k,d,s,A)$ be the event that a randomly chosen node w of the tree T_n has degree d and the leftmost k children of w are identical, with subtrees of size s. We have, by the union bound,

$$\mathbf{P}\{S_n \geq x\} \leq \mathbf{P}\Big\{\bigcup_{w\in T_n}\bigcup_{k\geq 2}\bigcup_{A\in\mathcal{A}_{wk}} E(x,w,k,A)\Big\} \leq n\sum_{k\geq 2}\sum_{d\geq k}\binom{d}{k}\sum_{s\geq x/k}\mathbf{P}\{E'(x,k,d,s,A)\}. \tag{7}$$

Supposing that w is the jth node in preorder, $E'(x,k,d,s,A)$ is the event that $\xi_j = d$, $(\xi_{j+1},\ldots,\xi_{j+s})$ forms a tree, and $(\xi_{j+rs+1,\ldots,j+rs+s}) = (\xi_{j+1},\ldots,\xi_{j+s})$ for all $1 \leq r < k$. Let us say that an integer j is "good" if these conditions hold when addition on the indices is done modulo n. Clearly, there are more good j than j satisfying the above conditions. Let G be the event that an index j chosen uniformly at random from $\{1,\ldots,n\}$ is "good"; let B be the event that $(\xi_2,\ldots,\xi_s)$ forms a tree and $(\xi_{rs+2},\ldots,\xi_{(r+1)s+1}) = (\xi_{j+1},\ldots,\xi_{j+s})$ for all $1 \leq r < k$. By a rotational argument due to Dwass [5],

$$\begin{aligned}
\mathbf{P}\left\{G \,|\, (\xi_1,\ldots,\xi_n) \text{ forms a tree}\right\} &= \mathbf{P}\Big\{G \,\Big|\, \sum_{i=1}^n \xi_i = n-1\Big\} \\
&= \frac{\mathbf{P}\{\xi_1 = d,\, B,\, \sum_{i=1}^n \xi_i = n-1\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}} \\
&= \frac{\mathbf{P}\{\xi_1 = d,\, B,\, \sum_{i=\lfloor 1+ks\rfloor+1}^n \xi_i = (n-1)-d-k(s-1)\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}},
\end{aligned}$$

so letting

$$R = \frac{\mathbf{P}\{\sum_{i=1}^{n-(1-ks)} \xi_i = \big(n-(1-ks)-1\big)+(k+1-d)\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}},$$

we have

$$\mathbf{P}\left\{G \,|\, (\xi_1,\ldots,\xi_n) \text{ forms a tree}\right\} = p_d\,\mathbf{P}\{B\}R. \tag{8}$$

Letting $\lambda = \gcd\{i : i \geq 1, p_i > 0\}$, a lemma of Kolchin [13] states that uniformly in y,

$$\mathbf{P}\Big\{\sum_{i=1}^n \xi_i = n-y\Big\} = \begin{cases} \lambda/\big(\sqrt{2\pi n}\sigma^2\big)\exp\big(-y^2/2n\sigma^2\big) + o(1)/\sqrt{n}, & \text{if } n \bmod \lambda = 0;\\ 0, & \text{if } n \bmod \lambda \neq 0.\end{cases} \tag{9}$$

As the $o(1)$ term does not depend on y, we find that

$$R = \sqrt{\frac{n-1}{n-(1-ks)-1+(k+1-d)}}\exp\Big(-\frac{(1-ks+d-k)^2}{2\big(n-(1-ks+d-k)\big)}\sigma^2\Big) + o(1),$$

where the $o(1)$ term depends only on n. Assuming that $ks+d \leq n/2$, we have $R \leq \sqrt{2}+o(1)$. Hence

$$\mathbf{P}\left\{G \,|\, (\xi_1,\ldots,\xi_n) \text{ forms a tree}\right\} \leq \big(\sqrt{2}+o(1)\big)p_d\,\mathbf{P}\{B\} \tag{10}$$

whenever $ks+d \leq n/2$. We now compute a bound on $\mathbf{P}\{B\}$. We have

$$\mathbf{P}\{B \,|\, \xi_2,\ldots,\xi_{1+s}\} = (p_{\xi_2}\cdots p_{\xi_{1+s}})^{k-1}$$

and therefore, by independence of the ξ_i,

$$\begin{aligned}\mathbf{P}\{B\} &= \mathbf{E}\left\{(p_{\xi_2}\cdots p_{\xi_{1+s}})^{k-1}\,\mathbf{1}_{[(\xi_2,\ldots,\xi_{1+s})\text{ forms a tree}]}\right\}\\ &\leq \prod_{i=2}^{1+s}\mathbf{E}\left\{(p_{\xi_i})^{k-1}\right\} = \Big(\sum_{i\geq 0}{p_i}^k\Big)^s.\end{aligned} \tag{11}$$

We can now combine all of these bounds. Substituting everything into (7), we have

$$\begin{aligned}\mathbf{P}\{S_n \geq x\} &\leq n\sum_{k\geq 2}\sum_{d\geq k}\binom{d}{k}\sum_{s\geq x/k}\big(\sqrt{2}+o(1)\big)p_d\Big(\sum_{i\geq 0}{p_i}^k\Big)^s\\ &\leq \big(\sqrt{2}+o(1)\big)n\sum_{k\geq 2}\sum_{d\geq k}p_d\binom{d}{k}\Big(\sum_{i\geq 0}{p_i}^k\Big)^{x/k}\frac{1}{1-\sum_{i\geq 0}{p_i}^k}.\end{aligned}$$

Since the inequality (5) was strict, there exists $0<\theta<1$ such that

$$\begin{aligned}\mathbf{P}\{S_n \geq x\} &\leq \frac{\sqrt{2}+o(1)}{1-\sum_{i\geq 0}{p_i}^2}\Big(n\sum_{d\geq 2}p_d\binom{d}{2}\Big(\sum_{i\geq 0}{p_i}^2\Big)^{x/2} + n\sum_{k\geq 3}\sum_{d\geq k}p_d\binom{d}{k}\Big(\sum_{i\geq 0}{p_i}^2\Big)^{x/2}\theta^x\Big)\\ &\leq \frac{\sqrt{2}+o(1)}{1-\sum_{i\geq 0}{p_1}^2}n(\sigma^2+1)\Big(\sum_{i\geq 0}{p_i}^2\Big)^{x/2} + n\Big(\sum_{i\geq 0}{p_i}^2\Big)^{x/2}\theta^x\sum_{k\geq 3}\sum_{d\leq k}p_d\binom{d}{k}.\end{aligned} \tag{12}$$

Since

$$\sum_{d\geq 2}p_d\sum_{k=3}^{d}\binom{d}{k} \leq \sum_{d\geq 3}p_d 2^d \leq \mathbf{E}\{2^\xi\},$$

we have

$$\mathbf{P}\{S_n \geq x\} \leq n\frac{\sqrt{2}(\sigma^2+1)}{1-\sum_{i\geq 0}{p_i}^2}\Big(\sum_{i\geq 0}{p_i}^2\Big)^{x/2}\big(1+o(1)\big), \tag{13}$$

provided that $\mathbf{E}\{2^\xi\} < \infty$. Setting

$$x = (1+\epsilon)\frac{2\log_2 n}{\log_2\big(1/\sum_{i\geq 0}{p_i}^2\big)} = (1+\epsilon)\frac{2\log_2 n}{\mathrm{H}_2},$$

we find that $\mathbf{P}\{S_n \geq x\} \to 0$ as $n \to \infty$. ∎

The next lemma presents a lower bound for S_n.

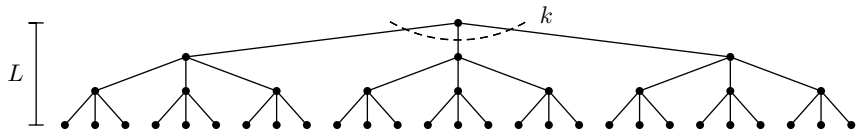

Fig. 2. The construction in the proof of Lemma 3. In this example, both k and L are equal to 3.

Lemma 3. *Let ξ be an offspring distribution with mean 1 and nonzero finite variance σ^2. If S_n is the multiplicity of a conditioned Bienaymé–Galton–Watson tree of size n with offspring distribution ξ, then*

$$\mathbf{P}\left\{S_n < (1-\epsilon)\frac{\log_2 n}{\log_2(1/\gamma)}\right\} \to 0 \tag{14}$$

for all $0 < \epsilon < 1$, where $\gamma = \max_{k\geq 2} {p_0}^k {p_k}^{k/(k-1)}$.

Proof. Consider a complete k-ary tree of height L. This tree has k^L leaves and $1 + k + \cdots + k^{L-1} = (k^L - 1)/(k-1)$ internal nodes, all of degree k. The probability that an unconditioned Bienaymé–Galton–Watson tree takes this shape is

$${p_0}^{k^L} {p_k}^{(k^L-1)/(k-1)};$$

call this probability q. For any real number x, the statement $S_n < x$ implies that no node in the tree can have the given k-ary tree as a subtree for any $k^L \geq x$, as the multiplicity of the k-ary tree is k^L. Fix $k \geq 2$ for now, let L be the first integer for which $k \geq x$, and let $y = k^L$. Observe that $y \leq kx$. Denote the size of the k-ary tree by $z = y + (y-1)/(k-1)$.

We now consider the indices $1, 1+z, 1+2z, 1+3z, \ldots$ in $\{1, \ldots, n-z\}$. Let Y_i be the event (and ${Y_i}^c$ its complement) that $(\xi_i, \ldots, \xi_{i+z-1})$ defines precisely the k-ary tree, where i is in the set of indices defined above, which has size $\lfloor (n-z)/z \rfloor$. Note that

$$\mathbf{P}\{S_n < x\} \leq \mathbf{P}\{S_n < y\} = \mathbf{P}\left\{\bigcap_{i=1}^{n-z} {Y_i}^c \,\middle|\, (\xi_1, \ldots, \xi_n) \text{ defines a tree}\right\}.$$

By Dwass' cycle lemma [5], the probability that $(\xi_1, \ldots, \xi_n)$ defines a tree is $\Theta(n^{3/2})$, so

$$\begin{aligned}
\mathbf{P}\{S_n < x\} &\leq \Theta(n^{3/2})\,\mathbf{P}\left\{\bigcap_{i=1}^{n-z} {Y_i}^c\right\} \\
&= \Theta(n^{3/2})\,\mathbf{P}\{{Y_i}^c\}^{\lfloor (n-z)/z \rfloor} \\
&= \Theta(n^{3/2})(1-q)^{\lfloor (n-z)/z \rfloor} \\
&\leq \Theta(n^{3/2}) \exp\left(-\left\lfloor \frac{n-z}{z} \right\rfloor {p_0}^y {p_k}^{(y-1)/(k-1)}\right) \\
&\leq \Theta(n^{3/2}) \exp\left(-\left\lfloor \frac{n-z}{z} \right\rfloor {p_0}^{kx} {p_k}^{(kx-1)/(k-1)}\right) \\
&\leq \Theta(n^{3/2}) \exp\left(-\Omega(1)\left\lfloor \frac{n-z}{z} \right\rfloor \left({p_0}^k {p_k}^{(k-1)/(k-1)}\right)^x\right) \\
&\leq \Theta(n^{3/2}) \exp\left(-\Omega(1)\left\lfloor \frac{n-z}{z} \right\rfloor \gamma^x\right).
\end{aligned} \tag{15}$$

Substituting $(1-\epsilon)\log_2 n/\log_2(1/\gamma)$ for x, and noting that $z = \Theta(\log n)$, we observe that this bound tends to 0. ∎

3. The maximal leaf-degree

Let T_n be a random critical Bienaymé–Galton–Watson tree of size n. We let ξ_u be the degree of the node u and let λ_u be the number of children of u that are leaves in T_n, i.e., the *leaf-degree* of u. We denote by L_n the random variable $\max_{u\in T_n}\lambda_u$; it is clear that the multiplicity S_n satisfies $M_n \geq L_n$. The next lemma shows that when the tail of the offspring distribution ξ decays at a rate slower than exponential, the ratio $L_n/\log n \to \infty$ in probability. So while our condition in the upper bound that $\mathbf{E}\{2^\xi\}$ be finite might have seemed somewhat artificial at first glance, we essentially cannot do without it.

Lemma 4. *Let $\mathbf{E}\{\xi\} = 1$, $\mathbf{V}\{\xi\} = \sigma^2 \in (0,\infty)$, and suppose that $\mathbf{E}\{\rho^\xi\} = \infty$ for every $1 < \rho < \infty$. Let L_n be the maximal leaf-degree in T_n, the Bienaymé–Galton–Watson tree induced by ξ, of size n. Then*

$$\frac{L_n}{\log n} \to \infty$$

in probability along a subsequence, as $n \to \infty$.

The proof of this lemma uses Kesten's limit tree [12] for the offspring distribution ξ, whose construction we briefly recall here (see also [14]). Kesten's infinite tree, denoted T_∞, is obtained by iterating the following step. Let the root of T_∞ be marked. A marked node has ζ children, where $\mathbf{P}\{\zeta = i\} = ip_i$ and $p_i = \mathbf{P}\{\xi = i\}$. Observe that $\zeta \geq 1$ and $\mathbf{E}\{\zeta\} = \mathbf{E}\{\xi^2\} = \sigma^2+1$. Of these ζ children, a random child is marked and all others are unmarked. The unmarked nodes are roots of independent (unconditioned) Bienaymé–Galton–Watson trees. The procedure is then repeated for the sole marked node.

Proof. We argue by coupling T_n with T_∞. Let (T_n, k) and (T_∞, k) denote the truncations of T_n and T_∞, respectively, to the nodes at distance $\leq k$ from the root. Then, denoting the total variation distance by TV, it is known that

$$\mathrm{TV}\big((T_n,k_n),(T_\infty,k_n)\big) = o(1)$$

if the sequence (k_n) is $o(\sqrt{n})$ (see, e.g., [11] and [19]). We couple (T_n, k_n) and (T_∞, k_n) such that

$$\mathbf{P}\big\{(T_n,k_n) \neq (T_\infty,k_n)\big\} = \mathrm{TV}\big((T_n,k_n),(T_\infty,k_n)\big) \to 0.$$

To show that $L_n/\log n \to \infty$ in probability, it suffices to show this for L'_n, the maximal leaf-degree among all marked nodes of (T_∞, k_n) at distance $< k_n$ from the root. Let $\zeta_0, \zeta_1, \ldots, \zeta_{k_n-1}$ be the degrees of the marked nodes in (T_∞, k_n), indexed by their distance from the root, let λ_i be the leaf-degree corresponding to ζ_i. Now, fix a constant c and let A_i be the event that $\lambda_i \leq c\log n$; we have

$$\begin{aligned}\mathbf{P}\{L'_n \leq c\log n\} &\leq \mathbf{P}\Big\{\bigcap_{i=0}^{k_n-1} A_i\Big\}\\ &= \mathbf{P}\{A_0\}^{k_n-1}\\ &\leq \exp\big(-(k_n-1)\,\mathbf{P}\{\lambda_0 > c\log n\}\big).\end{aligned} \tag{16}$$

Setting $k_n = \lceil n^{1/3}\rceil + 1$, we have

$$\mathbf{P}\{L'_n \leq c\log n\} \leq \exp\big(-\,n^{1/3}\,\mathbf{P}\{\lambda_0 > c\log n\}\big). \tag{17}$$

Note that $\lambda_0 \sim \mathrm{binomial}(\zeta_0 - 1, p_0)$, so that $\mathbf{P}\{\lambda_0 \leq p_0\zeta_0/2 \,|\, \zeta_0\} \leq 1/2$ for ζ_0 large enough, by the law of large numbers. Therefore, for n large enough, we have

$$\mathbf{P}\{\lambda_0 > c\log n\} \geq \mathbf{P}\big\{\lambda_0 \geq \frac{p_0\zeta_0}{2} > c\log n\big\} \geq \frac{1}{2}\,\mathbf{P}\Big\{\zeta_0 > \frac{2c}{p_0}\log n\Big\}. \tag{18}$$

To conclude the proof, we must show that $n^{1/3}\,\mathbf{P}\{\zeta_0 > 2c\log n/p_0\} \to \infty$ along a subsequence of n. Note that if $\mathbf{E}\{\rho^\xi\} = \infty$, then $\int_0^\infty \mathbf{P}\{\rho^\xi > x\}\,dx = \infty$, and thus

$$\sum_{\ell=1}^{\infty} 2^\ell\,\mathbf{P}\left\{\xi > \frac{\ell}{\log_2 \rho}\right\} \geq \sum_{\ell=1}^{\infty} 2^\ell\,\mathbf{P}\{\rho^\xi > 2^\ell\} \geq \sum_{\ell=1}^{\infty} \int_{2^\ell}^{2^{\ell+1}} \mathbf{P}\{\rho^\xi > x\}\,dx = \infty,$$

and consequently, $\mathbf{P}\{\xi > \ell/\log_2 \rho\} \geq \ell^{-2}2^{-\ell}$ for infinitely many $\ell \in \mathbf{N}$. As

$$\mathbf{P}\left\{\zeta > \frac{\ell}{\log_2 \rho}\right\} \geq \frac{\ell}{\log_2 \rho}\,\mathbf{P}\left\{\xi > \frac{\ell}{\log_2 \rho}\right\},$$

we see that

$$\mathbf{P}\left\{\zeta > \frac{\ell}{\log_2 \rho}\right\} \geq \frac{1}{\log_2 \rho \cdot \ell 2^\ell} \tag{19}$$

for infinitely many ℓ. Setting $\ell = (2c/p_0)\log n \log_2 \rho$, we have,

$$n^{1/3}\,\mathbf{P}\left\{\zeta > \frac{2c}{p_0}\log n\right\} \geq n^{1/3} \cdot \frac{1}{2^{2c\log n \log_2 \rho/p_0}} \cdot \frac{1}{\log_2 \rho \cdot 2c\log n \log_2 \rho/p_0} \tag{20}$$

for infinitely many n provided that

$$\frac{2c}{p_0}\log 2\log_2 \rho \leq \frac{1}{6},$$

which is possible by making $\rho > 1$ small enough. Thus, for every $c > 0$,

$$\limsup_{n\to\infty} \mathbf{P}\{L'_n > c\log n\} = 1,$$

which is what we wanted to show. ∎

Note that if for every $\rho < 1$, $p_n > \rho^n$ for all n large enough, then $L_n/\log n \to \infty$ in probability (instead of just along a subsequence).

4. Examples

There exists an important link between certain offspring distributions of conditioned Bienaymé–Galton–Watson trees and families of "simply-generated trees" [15]. In this section we examine several important families of trees in the Bienaymé–Galton–Watson context, and give explicit asymptotic upper and lower bounds for the multiplicity. In each case, the two important parameters will be

$$\gamma = \max_{k\geq 2} {p_0}^k {p_k}^{k/(k-1)} \qquad \text{and} \qquad \mathrm{H}_2 = \log_2 \frac{1}{\sum_{i\geq 0} {p_i}^2}.$$

We must also verify that $\mathbf{E}\{2^\xi\}$ is finite, if the upper bound is to hold. In particular, this latter condition always holds if ξ is bounded. A summary of this section is displayed in Table 1.

Full binary trees. These are trees in which every node must have exactly zero or two children, and arise from the distribution $p_0 = p_2 = 1/2$. We compute $\gamma = 1/16$ and $\mathrm{H}_2 = 1$, so that

$$(1-\epsilon)\frac{\log_2 n}{4} \leq S_n \leq (1+\epsilon)2\log_2 n \tag{21}$$

asymptotically in probability. Because the multiplicity in a full binary tree must be a power of 2, in essence this means that there exists a sequence of integers (a_n) such that

$$\mathbf{P}\big\{S_n \in \{2^{a_n}, 2^{a_n+1}, 2^{a_n+2}, 2^{a_n+3}\}\big\} \to 1.$$

Table 1. LEAF MULTIPLICITIES OF CERTAIN FAMILIES OF TREES

Family	γ	H_2	Lower bound	Upper bound
Full binary (uniform$\{0,2\}$)	$\frac{1}{16}$	1	$\frac{\log_2 n}{4}$	$2\log_2 n$
Flajolet t-ary ($p_0 = 1-1/t;\ p_t = 1/t$)	$e^{-1+o_{t\to\infty}(1)-\log t}$	$\log_2 \frac{1}{1-2/t-t/t^2}$	$\frac{\log_2 n}{\log_2 e + \log_2 t + o_{t\to\infty}(1)}$	$\sim_{t\to\infty} t\log n$
Cayley (Poisson(1))	$\frac{1}{4e^4}$	$\log_2\left(\frac{e^2}{I_0(2)}\right)$	$\frac{\log_2 n}{2+4\log_2 e}$	$\frac{2\log_2 n}{\log_2(e^2/(I_0(2))}$
Catalan (binomial$(2,1/2)$)	$\frac{1}{256}$	$\log_2(8/3)$	$\log_{256} n$	$\frac{2\log_2 n}{\log_2(8/3)}$
Binomial (binomial$(d,1/d)$)	$\frac{1}{4}\left(1-\frac{1}{d}\right)^{4d-2}$	$\log_2\left(\frac{e^2}{I_0(2)}\right)+o_{d\to\infty}(1)$	$\frac{\log_2 n}{2-\log_2((1-1/d)^{4d-2})}$	$\frac{2\log_2 n}{\log_2(e^2/(I_0(2))+o_{d\to\infty}(1)}$
Motzkin (uniform$\{0,1,2\}$)	$\frac{1}{81}$	$\log_2 3$	$\log_{81} n$	$2\log_3 n$
Planted plane (geometric$(1/2)$)	$\frac{1}{256}$	—	$\log_{256} n$	—

In other words, in general one cannot improve the ratio between the upper and lower bounds in Theorem 1 to a factor of less than $8+\epsilon$.

Flajolet t-ary trees. Full binary trees are a special case of a Flajolet t-ary tree for $t=2$ (see [6], p. 68). In general, these are trees whose non-leaf nodes each have t children, and they arise from the finite distribution $p_0 = (t-1)/t$, $p_t = 1/t$. We have

$$\begin{aligned}\gamma &= {p_0}^t {p_t}^{t/(t-1)} \\ &= \left(1-\frac{1}{t}\right)^t \left(\frac{1}{t}\right)^{t/(t-1)} \\ &= \exp\bigl(-1+o_t(1)-\log t\bigr),\end{aligned} \tag{22}$$

so $\log_2(1/\gamma) = \log_2 e + \log_2 t + o(1)$ as $t\to\infty$. On the other hand,

$$\mathrm{H}_2 = \log_2 \frac{1}{{p_0}^2 + {p_t}^2} = \log_2 \frac{1}{1-2/t+2/t^2}, \tag{23}$$

so $\mathrm{H}_2 \sim 2\log_2 e/t$ as $t\to\infty$. This means that as t gets large, the ratio between the upper and lower bound grows as $t\log t$.

Cayley trees. These trees arise from a Poisson(1) distribution, where $p_i = 1/(e\cdot i!)$ for $i\ge 0$. We verify first that

$$\mathbf{E}\{2^\xi\} = \sum_{i=0}^\infty \frac{2^i}{ei!} = e < \infty,$$

and then work out that $\gamma = 1/(4e^4)$. Letting

$$I_0(z) = \sum_{i=0}^\infty \frac{(i^2/4)^k}{i!\cdot\Gamma(z+1)} = \frac{1}{\pi}\int_0^\pi e^{z\cos\theta}\,d\theta$$

be the modified Bessel function of the first kind (see [1], p. 376), we find that

$$\sum_{i=0}^\infty {p_i}^2 = \frac{1}{e^2}\sum_{i=0}^\infty \frac{1}{(i!)^2} = \frac{1}{e^2} I_0(2), \tag{24}$$

meaning that $\mathrm{H}_2 = 2\log_2 e - \log_2\big(I_0(2)\big)$. Putting everything together, the lower and upper bounds in probability for S_n are, respectively,

$$\frac{\log_2 n}{2 + 4\log_2 e} \approx \frac{\log_2 n}{7.771} \qquad \text{and} \qquad \frac{2\log_2 n}{2\log_2 e - \log_2\left((1/\pi)\int_0^\pi e^{2\cos\theta}\,d\theta\right)} \approx \frac{\log_2 n}{0.8483}. \tag{25}$$

Catalan trees. When we set $p_0 = p_2 = 1/4$ and $p_1 = 1/2$, we obtain a family of trees often called Catalan trees, since the number of such trees on n nodes is $\binom{2n}{n}/(n+1)$. There is a one-to-one correspondence between Catalan trees on n nodes and full binary trees on $2n+1$ nodes, since one obtains a full binary tree from a Catalan tree by adding artificial external nodes to every empty slot, and this procedure is reversed by removing all leaves from a full binary tree. It is easy to see that the leaf multiplicity of a full binary tree is exactly double the multiplicity of its corresponding Catalan tree. By plugging in $d = 2$ above, the lower bound given by Lemma 3 is $\log_2 n/8$, which makes sense since the correspondence with full binary trees tells us that the lower bound on the Catalan trees should be similar to $\log_2(2n+1)/8$. We calculate $\mathrm{H}_2 = \log_2(8/3)$ and the upper bound is $2\log_2 n/\log_2(8/3)$, so the ratio between the upper and lower bounds is $16/\log_2(8/3)$.

Binomial trees. Catalan trees are a special case of a binomial tree. For integer parameter $d \geq 2$, nodes in these trees have d "slots" that may or may not contain a child; so there are $\binom{d}{i}$ ways for a node to have i children, for $0 \leq i \leq d$. These trees correspond to a binomial$(d, 1/d)$ distribution. We compute

$$\gamma = (p_0 p_2)^2 = \left(\left(\frac{d-1}{d}\right)^d \cdot \frac{d(d-1)}{2} \cdot \frac{(d-1)^{d-2}}{d^d}\right)^2 = \frac{1}{4}\left(1 - \frac{1}{d}\right)^{4d-2}. \tag{26}$$

Note that taking the limit $d \to \infty$, the binomial$(d, 1/d)$ distributions approach a Poisson(1) distribution. Thus we see from our earlier discussion on the Cayley trees that $\mathrm{H}_2 = \log_2\big(e^2/(I_0(2)\big) + o_{d\to\infty}(1)$. This gives the respective lower and upper bounds

$$\frac{\log_2 n}{2 - (4d-2)\log_2(1-1/d)} \qquad \text{and} \qquad \frac{2\log_2 n}{\log_2\big(e^2/(I_0(2)\big) + o_{d\to\infty}(1)}. \tag{27}$$

The lower bound tends to $\log_2 n/(2 + 4\log_2 e)$ as $d \to \infty$, matching the lower bound we obtained for Cayley trees above.

Motzkin trees. Also known as unary-binary trees, these are trees in which each non-leaf node can have either one or two children. They correspond to the distribution with $p_0 = p_1 = p_2 = 1/3$. We easily compute $\gamma = 1/81$ and $\mathrm{H}_2 = \log_2 3$, which yields an asymptotic lower bound of $\log_2 n/(\log_2 81) = \log_{81} n$ and an asymptotic upper bound of $2\log_2 n/\log_2 3 = 2\log_3 n$. The ratio between the upper and lower bounds is 8.

Planted plane trees. These are trees with ordered children, so that each can be embedded in the plane in a unique way. They correspond to a geometric(1/2) distribution, with $p_i = 1/2^{i+1}$ for all i. We find that $\gamma = 1/256$, so we have the asymptotic lower bound $S_n \geq \log_2 n/8$. Unfortunately, we have $\mathbf{E}\{2^\xi\} = \sum_{i\geq 0} 1/2 = \infty$, so Lemma 2 cannot be applied to give an upper bound here. We note that the maximal degree Δ_n of T_n satisfies $\Delta_n/\log_2 n \to 1$ in probability (see, e.g., [4], Lemma 6). However, this does not imply that $S_n = O(\log n)$ in probability.

5. Automorphic multiplicity

The multiplicity of a tree does not have a natural extension to unrooted trees, because whether or not two nodes are identical depends crucially on their position in relation to a distinguished root node u. In this

section we briefly investigate an alternate notion of multiplicity that does extend nicely to free trees. It arises in the problem of root estimation in Galton–Watson trees described [4]. We briefly recall some definitions. Let T be a rooted tree. By disregarding the parent-child directions of the edges, we obtain a free tree T_{F}. Conversely, if we start with a free tree T_{F} and any node u, we can define a *rooting of* T_{F} *at* u to be the rooted tree T_u obtained by fixing u as the root. This does not give rise to a unique tree in general, because children of a given node may hang on the wall in an arbitrary left-to-right order, but our new notion of multiplicity will treat all of these possible ordered trees the same.

Let $\mathrm{Aut}(T_{\mathrm{F}})$ be the group of all graph automorphisms of T_{F}, that is, bijections f from the set of vertices T_{F} to itself such that for vertices u and v, $f(u)$ is adjacent to $f(v)$ whenever u is adjacent to v. We can then define an *automorphism* of T_u to be a graph automorphism of T_{F} such that the root u stays fixed. By a slight abuse of notation, we denote the set of these rooted-tree automorphisms by $\mathrm{Aut}(T_u)$; formally this is the *stabilizer subgroup*

$$\mathrm{Stab}(u) = \{g \in \mathrm{Aut}(T_{\mathrm{F}}) : g \cdot u = u\}$$

of $\mathrm{Aut}(T_{\mathrm{F}})$. We will say that two nodes v and w in T_u are *congruent* and write $v \sim_u w$ if v and w belong to the same orbit under the action of $\mathrm{Aut}(T_u)$. This means that there exists an element f of $\mathrm{Aut}(T_u)$ such that $f(v) = w$. It is clear that this gives us an equivalence relation on the set of all nodes of T_u, and the *automorphic multiplicity* of a node v, denoted $\mu(u, v)$, is the size of the equivalence class of v under this relation. Since any node can be mapped to itself under an automorphism, $\mu(u, v) \geq 1$ for all v.

In fact, one can define the relation $\sim_u$, and consequently the function μ, purely in terms of the relation $\equiv$. We have $v \sim_u w$ if and only if there exists a permutation for every node in T_u such that applying each permutation to the left-to-right ordering of its respective node's children results in a tree in which $v \equiv w$. The analogue of S in this setting is the *automorphic (leaf) multiplicity* $M(T)$ of a rooted tree T. If o is the root of the tree T, then $M(T)$ is the maximum value of $\mu(o, v)$ over all nodes v in the tree T.

Fig. 3. Different leaf multiplicities but the same automorphic leaf multiplicity.

Fig. 3 illustrates the distinction between the automorphic and non-automorphic multiplicity. We have $S(T_1) = M(T_1) = 4$, since the two non-leaf children of the root have identical (and therefore congruent) subtrees. In T_2, on the other hand, these subtrees are congruent but not identical, so that $M(T_2) = 4$ but the non-automorphic multiplicity of T_2 is only 2.

This definition is still somewhat at odds with the notion of multiplicity that arises in the root estimation problem from [4]. In that setting, one considers all graph automorphisms of the free tree, not just ones that fix the root. We will call the size of the orbit of a node under this larger action the *free multiplicity* μ_{F} and if two nodes u and v are congruent under an arbitrary graph automorphism, then we write $u \sim_{\mathrm{F}} v$ and say that the two nodes are *free-congruent*. We also let $M_{\mathrm{F}}(T)$ denote the *free (leaf) multiplicity*, the maximum value of μ_{F} over all nodes in the free tree T_{F}.

Fig. 4 shows the relation between the automorphic multiplicity of a rooted tree and the free multiplicity its free-tree counterpart. Note that $M(T) \leq M_{\mathrm{F}}(T)$ for any rooted tree T, since we have $\mu(u) \leq \mu_{\mathrm{F}}(u)$ for every node u. We shall spend the rest of this section showing that this inequality can more or less be reversed. First, we need three lemmas, and in their statements and proofs, we shall understand "multiplicity" to mean "free multiplicity". In the following proof, we also write $[G : H]$ to denote the *index* of a subgroup H in a larger group G; that is, the cardinality of the coset space G/H.

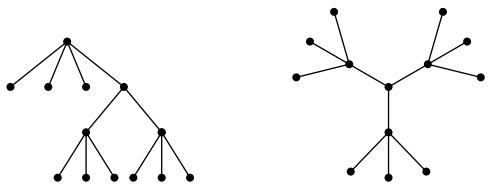

Fig. 4. A rooted tree T with $M(T) = 6$ and $M_{\mathrm{F}}(T) = 9$.

Lemma 5. *If u and v are adjacent nodes in a finite free tree T, then either $\mu_{\mathrm{F}}(u)$ is an integer multiple of $\mu_{\mathrm{F}}(v)$ or the other way around.*

Proof. We may reduce to the case where one of u or v is a leaf. This is because if neither is a leaf, then it is not in the orbit of any leaf by graph automorphism, so we can remove all the leaves from the tree T without changing either of $\mu_{\mathrm{F}}(u)$ or $\mu_{\mathrm{F}}(v)$. This is done finitely many times since T is finite and always contains at least one leaf.

Now without loss of generality, suppose u is the leaf and v is its unique neighbour. By the orbit-stabilizer theorem,

$$\big|\mathrm{Aut}(T)\big| = \mu_{\mathrm{F}}(u)\big|\mathrm{Stab}(u)\big| = \mu_{\mathrm{F}}(v)\big|\mathrm{Stab}(v)\big|, \tag{28}$$

where stabilizers are taken with respect to the group $\mathrm{Aut}(T)$. Every automorphism fixing u must permute its neighbours, but since u only has one neighbour, we have $\mathrm{Stab}(u) \subseteq \mathrm{Stab}(v)$. Thus

$$\begin{aligned}\mu_{\mathrm{F}}(u) &= \frac{\mu_{\mathrm{F}}(v)\big|\mathrm{Stab}(v)\big|}{\big|\mathrm{Stab}(u)\big|} \\ &= \frac{\mu_{\mathrm{F}}(v)\big[\mathrm{Stab}(v) : \mathrm{Stab}(u)\big]\big|\mathrm{Stab}(u)\big|}{\big|\mathrm{Stab}(u)\big|} \\ &= \mu_{\mathrm{F}}(v)\big[\mathrm{Stab}(v) : \mathrm{Stab}(u)\big],\end{aligned} \tag{29}$$

proving the lemma. ∎

The next lemma formalizes the intuitive notation that in a free tree, the multiplicities are in some sense smaller towards the centre of the tree.

Lemma 6. *Let u — v — w be neighbouring nodes in a free tree T with v being the central node. Then v cannot have strict maximal free multiplicity among the three nodes; that is, $\mu_{\mathrm{F}}(v) \leq \mu_{\mathrm{F}}(u)$ or $\mu_{\mathrm{F}}(v) \leq \mu_{\mathrm{F}}(w)$.*

Proof. Suppose for contradiction that $\mu_{\mathrm{F}}(v) > \mu_{\mathrm{F}}(u)$ and $\mu_{\mathrm{F}}(v) > \mu_{\mathrm{F}}(w)$. Then, for each of the pairs of neighbours u — v and v — w, the multiplicity of one of the nodes must be an integer multiple of the multiplicity of the other, by the previous lemma. So there must be integers $r, s > 1$ such that

$$\mu_{\mathrm{F}}(v) = r\mu_{\mathrm{F}}(w) \quad \text{and} \quad \mu_{\mathrm{F}}(v) = s\mu_{\mathrm{F}}(u). \tag{30}$$

The situation is illustrated in Fig. 5. Since $\mu_{\mathrm{F}}(v) = s\mu_{\mathrm{F}}(u)$, u must have $s-1$ children in the orbit of v and thus have subtree rooted at each of these children be isomorphic to B. Similarly, since $\mu_{\mathrm{F}}(v) = r\mu_{\mathrm{F}}(w)$, w must have $r-1$ child subtrees isomorphic to A.

We note that in order to satisfy the $r, s > 1$ requirements, we must have

$$|A| \geq (s-1)|B| + 2 \quad \text{and} \quad |B| \geq (r-1)|A| + 2, \tag{31}$$

where the additional $+2$ terms come respectively from nodes u and v (for $|A|$) or v and w (for $|B|$). This implies that

$$|A| \geq (s-1)(r-1)|A| + 2s,$$

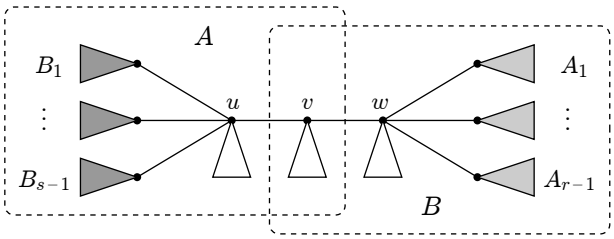

Fig. 5. Three adjacent nodes and their subtrees.

which is impossible if $|A| \geq 1$ and $r, s > 1$. The contradiction tells us that v cannot have strict maximal multiplicity among the three nodes. ▮

We have established that if we embed a free tree into the (x, y)-plane and then lift the nodes up by setting each node's z-coordinate to its multiplicity, then the result is a convex, spidery bowl or valley. This is illustrated in Fig. 6.

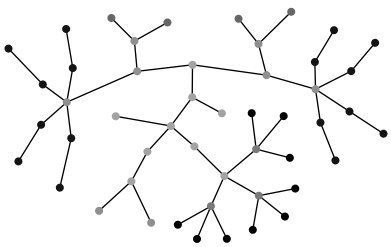

Fig. 6. Darker shades of grey indicate higher multiplicities in this free tree.

On a path between any two endpoints, the multiplicities decrease monotonically towards the centre of the tree before increasing monotonically towards the endpoint. There is a central connected core of nodes of minimal multiplicity and we are able to show that this minimal multiplicity cannot be greater than 2.

Lemma 7. *If $F = (V, E)$ is a finite free tree, then the node of minimal multiplicity in F has multiplicity 1 or 2.*

Proof. The proof is by contraposition. Let $u \in V(F)$ be a node of minimal multiplicity and suppose $\mu_{\mathrm{F}}(u) > 2$. Let C_u be the orbit of u. There is a subtree F' whose endpoints are the members of C_u; since $m > 2$ and the graph is connected, there is necessarily at least one node $v \in F' \setminus C_u$. By Lemma 6, we have $\mu_{\mathrm{F}}(v) \leq \mu_{\mathrm{F}}(u)$ but by minimality of $\mu_{\mathrm{F}}(u)$, we know that $\mu_{\mathrm{F}}(v) = \mu_{\mathrm{F}}(u)$. So we can repeat the argument with C_v to find that the tree is infinite (at each step we are removing $\mu_{\mathrm{F}}(u)$ nodes from the free tree, but the process never terminates).

Note that this argument does not work when $\mu_{\mathrm{F}}(u) = 2$ because F' may simply consist of two nodes connected by one edge. ▮

Theorem 8. *Let T be a rooted tree with n nodes; let $M(T)$ and $M_{\mathrm{F}}(T)$ be the automorphic multiplicity and free multiplicity of T, respectively. We have the inequality*

$$M_{\mathrm{F}}(T) \leq 2M(T),$$

and this bound is the best possible.

Proof. Suppose first that $n \geq 3$. Let v be a leaf of maximal automorphic multiplicity in T_{F}, and let $[v]$ denote the set of nodes that are free-congruent to v (so $\big|[v]\big| = M_{\mathrm{F}}(T)$). By Lemma 7, a node s of minimal

automorphic multiplicity either has $\mu_{\mathrm{F}}(s) = 1$ or $\mu_{\mathrm{F}}(s) = 2$, and since we assumed that $n \geq 3$, we can require that s not be a leaf.

If $\mu_{\mathrm{F}}(s) = 1$, then $M(T_s) = M_{\mathrm{F}}(T)$, since any automorphism of T_{F} already fixes s. The nodes in $[v]$ all lie in some subtrees of s, and without loss of generality, we may assume that they do not all lie in the same subtree, since if s' is the only child of s whose subtree contains nodes of $[v]$, we can reroot the tree T_s at s' instead without changing the maximum automorphic multiplicity. There are $d \geq 2$ children of the root whose subtrees contain elements of $[v]$; each one contains an equal proportion of these nodes, so d divides $M_{\mathrm{F}}(T)$. If we reroot the tree at any node outside these subtrees, then the automorphic multiplicity of the tree does not change. If, on the other hand, we choose a node in one of these subtrees, then there are still $(d-1)M_{\mathrm{F}}(T)/d$ leaves that can still be shuffled amongst themselves, so the maximum automorphic multiplicity is $(d-1)M_{\mathrm{F}}(T)/d \geq M_{\mathrm{F}}(T)/2$.

If $\mu_{\mathrm{F}}(s) = 2$, there is a node s' that is free-congruent to s, and there is mirror symmetry in the graph. This means that there is a way to split the graph along an edge such that the two sides have the exact same shape, one contains s, and the other contains s'. The side containing s has $M_{\mathrm{F}}(T)/2$ members of $[v]$; call this half $[v]_s$ and the other half $[v]_{s'}$. When the tree is rooted at s, we find that $M(T_s) = M_{\mathrm{F}}(T)/2$, since any two members of $[v]_s$ can be exchanged and any two members of $[v]_{s'}$ can be exchanged (but exchanges cannot happen between the two subtrees). And rerooting the tree at an arbitrary node, it is clear that the automorphic multiplicity of the tree will not decrease.

When $n = 1$ the statement is trivial, and taking $n = 2$ shows that the bound is the best possible, because if T is the tree with a root and a single (leaf) child, then $M_{\mathrm{F}}(T) = 2$ and $M(T) = 1$. ∎

This theorem tells us that the asymptotics of the free multiplicity are the same as the asymptotics of the automorphic multiplicity, up to a fudge factor of 2.

Because congruence of two nodes is immediately implied by their being identical under $\equiv$, we have $S(T) \leq M(T)$ for all rooted trees T. Thus if $M_n = M(T_n)$ and $F_n = M_{\mathrm{F}}(T_n)$, where T_n is a conditioned Galton–Watson tree of size n, then if γ is as defined in Lemma 3, the inequality

$$F_n \geq M_n \geq (1-\epsilon)\frac{\log_2 n}{\log_2(1/\gamma)} \tag{32}$$

holds with probability tending to 1.

Acknowledgements

We thank the anonymous referee for numerous insightful comments that substantially improved the the readability and rigour of the paper. We also thank Jonah Saks for helping us find a clean proof of Lemma 5.

References

[1] Milton Abramowitz and Irene Ann Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Washington: U.S. Government Printing Office, 1972).

[2] Krishna Athreya and Peter Ney, *Branching Processes* (Berlin: Springer Verlag, 1972).

[3] Irénée-Jules Bienaymé, "De la loi de multiplication et de la durée des familles," *Société Philomathique Paris Extraits* **5** (1845), 37–39.

[4] Anna M. Brandenberger, Luc Devroye, and Marcel K. Goh, "Root estimation in Galton-Watson trees," *arXiv preprint 2007.05681* (2020).

[5] Meyer Dwass, "The total progeny in a branching process," *Journal of Applied Probability* **6** (1969), 682–686.

[6] Philippe Flajolet and Robert Sedgewick, *Analytic Combinatorics* (New York: Cambridge University Press, 2009).

[7] Francis Galton and Henry William Watson, "On the probability of extinction of families," *Journal of the Anthropological Institute of Great Britian and Ireland* **4** (1874), 138–144.

[8] Philippe Jacquet and Wojciech Szpankowski, "Entropy computations via analytic depoissonization," *IEEE Transactions on Information Theory* **45** (1999), 1072–1081.

[9] Peter Jagers, *Some Notes on the History of Branching Processes, from my Perspective* (Lecture at the Oberwolfach Symposium on "Random Trees", 2009).

[10] Svante Janson, "Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees," *Random Structures and Algorithms* **48** (2016), 57–101.

[11] Götz Kersting, "On the height profile of a conditioned Galton–Watson tree," *arXiv preprint 1101.3656* (1998).

[12] Harry Kesten, "Subdiffusive behavior of a random walk on a random cluster," *Annales de l'Institut Henri Poincaré Probabilité et Statistiques* **22** (1986), 425–487.

[13] Valentin Fedorovich Kolchin, *Random Mappings* (New York: Optimisation Software Inc., 1986).

[14] Russell Lyons and Yuval Peres, *Probability on Trees and Networks* (New York: Cambridge University Press, 2016).

[15] Amram Meir and John Wesley Moon, "On the altitude of nodes in random trees," *Canadian Journal of Mathematics* **30** (1978), 997–1015.

[16] Richard Otter, "The number of trees," *Annals of Mathematics* **49** (1948), 583–599.

[17] Alfréd Rényi, "On measures of entropy and information," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (1961), 547–561.

[18] Claude Elwood Shannon, "A mathematical theory of communication," *Bell System Technical Journal* **27** (1948), 379–423.

[19] Benedikt Stufler, "Local limits of large Galton–Watson trees rerooted at a random vertex," *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **55** (2019), 155–183.

[20] Le Yu, *Automorphisms of Random Trees*, Ph.D. thesis (Drexel University, Philadelphia, Pa., 2012).