MONOTONE TREE STRUCTURES REVISITED

KATHERINE MORRIS AND HELMUT PRODINGER

ABSTRACT. Let T be a rooted tree structure with n nodes $t_1, \ldots, t_n$. A function $f : \{t_1, \ldots, t_n\}$ into $\{1 < \ldots < k\}$ is monotone if whenever t_i is a descendant of t_j then $f(t_i) \geq f(t_j)$. This paper deals with classes of trees where the nodes are labelled with elements from the set $\{1, 2, \ldots, k\}$ such that the sequence of labels connecting the root with a leaf is weakly monotone. Statistics concerning distances between p randomly chosen nodes in these monotonic trees are computed. The parameters under consideration are the *size of the ancestor tree* and the *Steiner distance*.

1. Introduction

Let T be a rooted tree structure with n nodes $t_1, \ldots, t_n$. A function $f : \{t_1, \ldots, t_n\}$ into $\{1 < \ldots < k\}$ is monotone if whenever t_i is a descendant of t_j then $f(t_i) \geq f(t_j)$. Also, if $k = n$ and f is a bijection then it is called a monotone bijection. In their paper [11], Prodinger and Urbanek have determined the average number of such bijections for several classes of trees. The average height of the j-th leaf of monotonic trees with n leaves has been considered by Kirschenhofer in [5].

In this paper we analyse two parameters, namely the *size of the ancestor tree* and the *Steiner distance* for the classes of monotonic trees introduced in [11]. Consider a tree, T, and select p random nodes in it. The ancestor tree is the subtree of T which is spanned by the root and the p chosen nodes and hence it is defined as the tree containing all ascendants of the p given nodes. Moreover, the Steiner distance is the size of the subtree spanned by these p nodes. If all of the p selected nodes lie in one subtree of the root then the Steiner distance is smaller than the size of the ancestor tree. Otherwise the two parameters are the same. The notation used is consistent with that of [11] and [5].

For the convenience of the reader we present a result from [3], which is used throughout this paper.

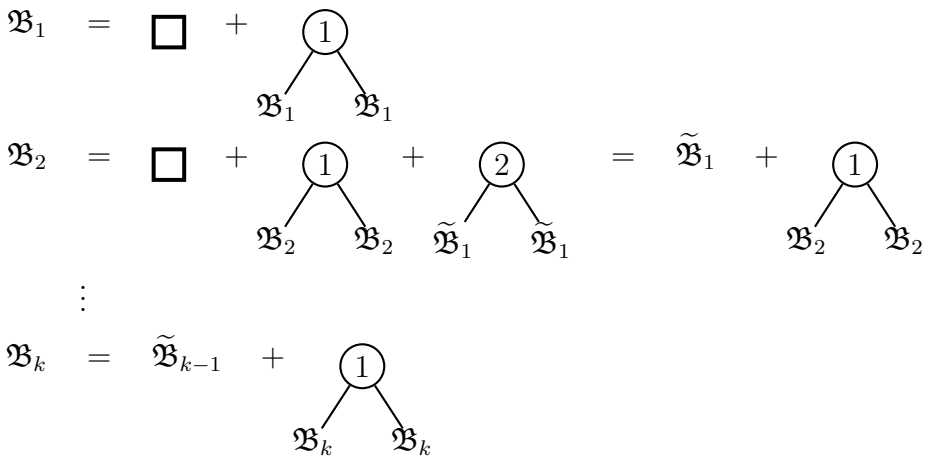
Proposition 1.1. *[Flajolet, Odlyzko] The binomial coefficients expressing* $[z^n](1-z)^\alpha = \binom{n+\alpha-1}{n}$ *have an asymptotic expansion as* $n \to \infty$, $[z^n](1-z)^\alpha \sim \frac{n^{-\alpha-1}}{\Gamma(-\alpha)}\Big(1 + \sum_{k\geq 1} \frac{e_k^{(\alpha)}}{n^k}\Big), \quad \alpha \notin \{0, 1, 2, \cdots\}$, *where* $e_k^{(\alpha)} = \sum_{l=k}^{2k} (-1)^l \lambda_{k,l} (\alpha+1)(\alpha+2)\cdots(\alpha+l)$, *with* $\sum_{k,l\geq 0} \lambda_{k,l} v^k t^l = e^t(1+vt)^{-1-1/v}$.

2. Binary Trees

We consider the binary model proposed in [11]: the class, $\mathfrak{B}_k$, of binary trees whose nodes are monotonically labelled with $1, 2, \ldots, k$, with corresponding generating function $y_k(z) =$

Date: February 3, 2005.

$\sum_{n\geq 0} b_n^{(k)}\, z^n$ where $b_n^{(k)}$ denotes the number of trees in $\mathfrak{B}_k$ with n nodes. Moreover, we let $\widetilde{\mathfrak{B}_k}$ be the class of binary trees whose nodes are monotonically labelled with $2,\dots,k+1$. The generating functions of $\mathfrak{B}_k$ and $\widetilde{\mathfrak{B}}_k$ coincide. Using the terminology from Flajolet [2] and [11] we obtain that



Thus the generating functions are $y_k(z) = y_{k-1}(z) + z y_k^2(z)$, $k \geq 1$, $y_0(z) = 1$. It was established in [11] that $y_k = \frac{1}{2q_k} - \frac{c_k}{2q_k}\sqrt{q_k - z} + \mathcal{O}(q_k - z)$ for $z \to q_k$. Other constants needed for our results are presented in the table below. Exact values are given for the first few cases of k.

constant	$k=1$	$k=2$	$k=3$
$c_k = 2(r_2 \cdots r_k)^{-1/4}$	2	$2\sqrt{2}$	$8/\sqrt{5}$
$p_k = c_k/(4\sqrt{\pi q_k})$	$1/\sqrt{\pi}$	$4/\sqrt{6\pi}$	$32/\sqrt{195\pi}$
$q_k = (1-r_k)/4$; $q_{k+1} = q_k(1-q_k)$	1/4	3/16	39/256
$r_k = 1 - 4q_k$	0	1/4	25/64

2.1. **Size of the ancestor tree.** It is perhaps easiest to begin the section with a result relating to the size of ancestor tree in binary trees and then use it to derive the corresponding one for monotonic binary trees.

Theorem 2.1. *The generating function for the size of the ancestor tree in binary trees satisfies*

$$B(z,u,v) = zv(1+u)B^2(z,u,v) - zvT^2(z) + T(z)\,, \tag{1}$$

where $T(z) = \frac{1-\sqrt{1-4z}}{2z}$ *is the generating function for binary trees.*

Proof: Let ϕ_p be the size of the ancestor tree in a binary tree of size n with p nodes selected at random. Then ϕ_p satisfies the recurrence

$$\phi_p(z,v) = zv\sum_{i=0}^{p-1}\phi_i(z,v)\,\phi_{p-1-i}(z,v) + zv\sum_{i=0}^{p}\phi_i(z,v)\,\phi_{p-i}(z,v)\,,\quad p\geq 1\,, \tag{2}$$

where the first term comes from selecting the root and the second term arises when the root is not selected. Now we multiply both sides with u^p and sum over p to obtain

$$\sum_{p\geq 1} u^p\,\phi_p(z,v) = zv\sum_{p\geq 1} u^p \sum_{i=0}^{p-1} \phi_i(z,v)\,\phi_{p-1-i}(z,v) + zv\sum_{p\geq 1} u^p \sum_{i=0}^{p} \phi_i(z,v)\,\phi_{p-i}(z,v)\,,\ p\geq 1\,. \tag{3}$$

It follows that

$$\begin{aligned}\sum_{p\geq 0} u^p\,\phi_p(z,v) - \phi_0(z,v) = zv\sum_{p\geq 0} u^p \sum_{i=0}^{p-1} \phi_i(z,v)\,\phi_{p-1-i}(z,v) - zv\phi_0(z,v)\phi_0(z,v)\\ + zv\sum_{p\geq 0} u^p \sum_{i=0}^{p} \phi_i(z,v)\,\phi_{p-i}(z,v)\,,\ p\geq 0\,.\end{aligned} \tag{4}$$

We let $B(z,u,v) = \sum_{p\geq 0} u^p\,\phi_p(z,v)$ and $T(z) = \phi_0(z,v)$ and obtain

$$B(z,u,v) - \phi_0(z,v) = zvuB^2(z,u,v) + zvB^2(z,u,v) - zv\phi_0^2(z,v)\,. \tag{5}$$

■

One can easily modify (1) to obtain the equations of the size of the ancestor tree for the binary trees given in the Prodinger-Urbanek model, by replacing $T(z)$ with y_k and $B(z,u,v)$ with $B_k(z,u,v)$.

Theorem 2.2. *The generating functions definining the size of the ancestor tree in monotonic binary trees are determined by*

$$B_k(z,u,v) = B_{k-1}(z,u,v) + zv(1+u)B_k^2(z,u,v) + z(1-v)y_k^2(z)\,,\ B_0(z,u,v) = 1\,. \tag{6}$$

■

Our main aim is to produce the expectation and variance for the size of the ancestor tree and this is done via the usual method of differentiating the generating function. Firstly, we find the expectation by differentiating B_k with respect to v and then evaluating at $v=1$. To simplify our notation we denote this first derivative by $\beta_k(z,u)$ and let $B_k(z,u,1) = y_k(z(1+u)) = \bar{y}_k(z)$ which yields

$$(1-2z(1+u)\bar{y}_k(z))\beta_k(z,u) = \beta_{k-1}(z,u) + \bar{y}_k(z) - \bar{y}_{k-1}(z) - y_k(z) + y_{k-1}(z)\,. \tag{7}$$

Lemma 2.3. *The solution of $\beta_k(z,u)$ is given by*

$$\beta_k(z,u) = \sum_{j=1}^{k}\Big[\prod_{i=j}^{k}(1-2z(1+u)\bar{y}_i(z))\Big]^{-1}\big(\bar{y}_j(z) - \bar{y}_{j-1}(z) - y_j(z) + y_{j-1}(z)\big)\,, \tag{8}$$

with $\beta_0(z,u) = 0$. Hence the dominant term in this solution is

$$\beta_k(z,u) \sim \frac{\bar{y}_k(z) - y_k(z)}{1-2z(1+u)\bar{y}_k(z)} \sim \frac{\sqrt{q_k - z} - \sqrt{q_k - z(1+u)}}{2q_k\sqrt{q_k - z(1+u)}}\,. \tag{9}$$

Proof: Both sides of (7) are multiplied with the product $\prod_{i=1}^{k-1}(1-2z(1+u)\bar{y}_i(z))$ and summed over the k's

$$\begin{aligned}&\Big[\prod_{i=1}^{k-1}(1-2z(1+u)\bar{y}_i(z))\Big](1-2z(1+u)\bar{y}_k(z))\beta_k(z,u)\\&=\sum_{j=0}^{k-1}\Big[\prod_{i=1}^{j}(1-2z(1+u)\bar{y}_i(z))\Big]\big(\bar{y}_j(z)-\bar{y}_{j-1}(z)-y_j(z)+y_{j-1}(z)\big)\,,\end{aligned}\tag{10}$$

which leads to the result upon division by $\prod_{i=1}^{k}(1-2z(1+u)\bar{y}_i(z))$. Now the dominant term occurs for $i=j=k$. Since the singularities of $y_{k-1}(z)$ and $\bar{y}_{k-1}(z)$ are further away from the origin than those of $y_k(z)$ and $\bar{y}_k(z)$ they will not contribute to the main term in $\beta_k(z,u)$ and we do not include them. Thus the dominant term in $\beta_k(z,u)$ is as given in the statement of the theorem. ■

For our analysis we need to know the asymptotic behaviour of the coefficients $[u^p]\beta_k(z,u)$ near the dominant singularity q_k.

Lemma 2.4. *As $z\to q_k$, the coefficients of u^p in $\beta_k(z,u)$ have the form*

$$[u^p]\beta_k(z,u)\sim\frac{1}{2^{2p+1}\,q_k}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p}.\tag{11}$$

Proof: We look at the coefficient of u^p in (9)

$$\begin{aligned}&[u^p]\frac{\sqrt{q_k-z}}{2q_k\sqrt{q_k-z(1+u)}}=\frac{\Big(1-\frac{z}{q_k}\Big)^{1/2}}{2q_k}[u^p]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{1/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{1/2}}\\&=\frac{1}{2q_k}\Big(-\frac{z}{q_k}\Big)^p\binom{-\frac12}{p}\Big(1-\frac{z}{q_k}\Big)^{-p}=\frac{z^p}{2^{2p+1}}q_k^{p+1}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p}\\&\sim\frac{1}{2^{2p+1}q_k}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p},\ z\to q_k\,.\end{aligned}\tag{12}$$

■

In order to obtain the first moment the coefficients of z^n in $\beta_k(z,u)$ are also needed. This is done by using Proposition 1.1 with $\alpha=-p$, which yields $[z^nu^p]\beta_k(z,u)\sim\frac{1}{2^{2p+1}\,q_k^{n+1}}\binom{2p}{p}\frac{n^{p-1}}{\Gamma(p)}$. Finally, these coefficients will be normalised by $p_kq_k^{-n}n^{-3/2}\binom{n}{p}\sim p_kq_k^{-n}n^{-3/2}\frac{n^p}{\Gamma(p+1)}$.

Theorem 2.5. *The expectation for the size of the ancestor tree in monotonic binary trees is*

$$E_{n,p}^{(k)}\sim\frac{p}{2^{2p+1}\,p_k\,q_k}\binom{2p}{p}\sqrt{n}\,,\ n\to\infty\,,\ \textit{fixed } k\,.\tag{13}$$

Higher order statistics can be computed in the same fashion, so for the variance of the size of the ancestor tree, B_k is differentiated twice with respect to v and evaluated at $v=1$. For convenience, we denote this derivative by $\Theta_k(z,u)$ which gives

$$\big(1-2z(1+u)\bar{y}_k(z)\big)\Theta_k(z,u)=\Theta_{k-1}(z,u)+4z(1+u)\bar{y}_k(z)\beta_k(z,u)+2z(1+u)\beta_k^2(z,u)\,.\tag{14}$$

Lemma 2.6. *The solution to $\Theta_k(z,u)$ has the closed form*

$$\Theta_k(z,u)=\sum_{j=1}^{k}\Big[\prod_{i=j}^{k}(1-2z(1+u)\bar{y}_i(z))\Big]^{-1}\Big(4z(1+u)\bar{y}_j(z)\beta_j(z,u)+2z(1+u)\beta_j^2(z,u)\Big)\,, \quad (15)$$

and the dominant term is asymptotically

$$\Theta_k(z,u)\sim\frac{4z(1+u)\bar{y}_k\,\beta_k+2z(1+u)\beta_k^2(z,u)}{1-2z(1+u)\bar{y}_k}\,. \quad (16)$$

Proof: The proof is omitted as it is similar to Lemma 2.6. ■

Using the known local asymptotic expansions for $\bar{y}_k$ and y_k, we can write (16) as

$$\Theta_k(z,u)\sim\frac{zu}{2c_kq_k(q_k-z(1+u))^{3/2}}\,, \quad (17)$$

and this enables us to look at the asymptotic behaviour of the coefficients.

Lemma 2.7. *The coefficients of u^p in $\Theta_k(z,u)$ have the form*

$$[u^p]\Theta_k(z,u)\sim\frac{p}{2^{2p}\,c_k\,q_k^{3/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ c_k=4p_k\sqrt{\pi\,q_k}\,,\ z\to q_k\,. \quad (18)$$

Proof: The coefficients of u^p are computed from the term in (17), as follows

$$\begin{aligned}[u^p]\frac{zu}{2c_kq_k(q_k-z(1+u))^{3/2}}&=\frac{z}{2c_kq_k^{5/2}}[u^{p-1}]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{3/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{3/2}}\\ &=\frac{pz^p}{2^{2p}c_kq_k^{p+3/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2}\sim\frac{p}{2^{2p}\,c_k\,q_k^{3/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ z\to q_k\,.\end{aligned} \quad (19)$$

■

Once again we apply Proposition 1.1 with $\alpha=-p-1/2$ to get the coefficients of z^n: $[z^nu^p]\Theta_k(z,u)\sim\frac{p}{2^{2p+2}\sqrt{\pi}\,p_k\,q_k^{n+2}}\binom{2p}{p}\frac{n^{p-1/2}}{\Gamma(p+\frac12)}$ and use them to compute the variance.

Theorem 2.8. *The variance for the size of the ancestor tree in monotonic binary trees is*

$$V_{n,p}^{(k)}=\frac{[z^nu^p]\Theta_k(z,u)}{p_kq_k^{-n}\frac{n}{\Gamma(p+1)}n^{-3/2}}+E_{n,p}^{(k)}-(E_{n,p}^{(k)})^2=\frac{p}{4p_k^2q_k^2}\Big(\frac{1}{\pi}-\frac{p}{2^{4p}}\binom{2p}{p}^2\Big)n+\mathcal{O}(\sqrt{n}\,)\,, \quad (20)$$

for $n\to\infty$ and fixed k.

One notes that $\frac{1}{\pi}-\frac{p}{2^{4p}}\binom{2p}{p}^2\to0$ as $p\to\infty$.

2.2. **The limiting distribution for the size of the ancestor tree.** Let $X_{n,p}$ be the random variable which counts the size of the ancestor tree of p randomly chosen nodes in a monotonically labelled binary tree of size n. We study the probabilities $\mathbb{P}\{X_{n,p}=m\}=\frac{[z^nu^pv^m]B_k(z,u,v)}{b_n^{(k)}\binom{n}{p}}$, where $b_n^{(k)}\sim p_kq_k^{-n}n^{-3/2}$ is the number of monotonically labelled binary trees which was computed in [11]. The first step in obtaining our limiting distribution is to find

all the moments for B_k. If we differentiate B_k M times with respect to v then using Leibniz's rule we obtain

$$\begin{aligned}\frac{\partial^M B_k(z,u,v)}{\partial v^M} &= \frac{\partial^M B_{k-1}(z,u,v)}{\partial v^M}\\ &\quad + Mz(1+u)\sum_{i=0}^{M-1}\binom{M-1}{i}\frac{\partial^i B_k(z,u,v)}{\partial v^i}\frac{\partial^{M-i-1}B_k(z,u,v)}{\partial v^{M-i-1}}\\ &\quad + zv(1+u)\sum_{i=0}^{M}\binom{M}{i}\frac{\partial^i B_k(z,u,v)}{\partial v^i}\frac{\partial^{M-i}B_k(z,u,v)}{\partial v^{M-i}}\,,\ M\geq 2\,,\end{aligned} \tag{21}$$

and it follows that

$$\begin{aligned}&\frac{\partial^M B_k(z,u,v)}{\partial v^M}(1-2zv(1+u)B_k(z,u,v))\\ &= \frac{\partial^M B_{k-1}(z,u,v)}{\partial v^M} + Mz(1+u)\sum_{i=0}^{M-1}\binom{M-1}{i}\frac{\partial^i B_k(z,u,v)}{\partial v^i}\frac{\partial^{M-i-1}B_k(z,u,v)}{\partial v^{M-i-1}}\\ &+ zv(1+u)\sum_{i=1}^{M-1}\binom{M}{i}\frac{\partial^i B_k(z,u,v)}{\partial v^i}\frac{\partial^{M-i}B_k(z,u,v)}{\partial v^{M-i}}\,,\ M\geq 2\,.\end{aligned} \tag{22}$$

This recursive relation is now used to solve for the M-th derivative of B_k and the latter is evaluated at a particular point.

Lemma 2.9. *The M-th derivative with respect to v in $B_k(z,u,v)$, evaluated at $v=1$ has the form*

$$\begin{aligned}\frac{\partial^M B_k(z,u,v)}{\partial v^M}\Big|_{v=1} &= \sum_{j=1}^{k}\Big[\prod_{l=j}^{k}(1-2z(1+u)\bar{y}_l(z))\Big]^{-1}\times\\ &\quad\Big[Mz(1+u)\sum_{i=0}^{M-1}\binom{M-1}{i}\frac{\partial^i B_j(z,u,v)}{\partial v^i}\Big|_{v=1}\frac{\partial^{M-i-1}B_j(z,u,v)}{\partial v^{M-i-1}}\Big|_{v=1}\\ &\quad + z(1+u)\sum_{i=1}^{M-1}\binom{M}{i}\frac{\partial^i B_j(z,u,v)}{\partial v^i}\Big|_{v=1}\frac{\partial^{M-i}B_j(z,u,v)}{\partial v^{M-i}}\Big|_{v=1}\Big]\,,\ M\geq 2\,,\end{aligned} \tag{23}$$

and its dominant term is

$$\begin{aligned}\frac{\partial^M B_k(z,u,v)}{\partial v^M}\Big|_{v=1} &\sim \Big[(1-2z(1+u)\bar{y}_k(z))\Big]^{-1}\times\\ &\quad\Big[Mz(1+u)\sum_{i=0}^{M-1}\binom{M-1}{i}\frac{\partial^i B_k(z,u,v)}{\partial v^i}\Big|_{v=1}\frac{\partial^{M-i-1}B_k(z,u,v)}{\partial v^{M-i-1}}\Big|_{v=1}\\ &\quad + z(1+u)\sum_{i=1}^{M-1}\binom{M}{i}\frac{\partial^i B_k(z,u,v)}{\partial v^i}\Big|_{v=1}\frac{\partial^{M-i}B_k(z,u,v)}{\partial v^{M-i}}\Big|_{v=1}\Big]\,.\end{aligned} \tag{24}$$

Proof: We multiply both sides of (22) with $\prod_{l=1}^{k-1}(1-2z(1+u)\bar{y}_l(z))$ and proceed with iterating the result. This proof is similar to the one for Lemma 2.6 so we do not give the details here. ■

The required moments are produced by analysing the behaviour of the coefficients in (24) near their dominant singularities q_k. For the coefficients of u^p we find that

$$[u^p]\frac{\partial^M B_k(z,u,v)}{\partial v^M}\Big|_{v=1} \sim \frac{\Gamma(2p+M-1)}{p2^{2p}q_k^{(M+1)/2}c_k^{M-1}(\Gamma(p))^2}\Big(1-\frac{z}{q_k}\Big)^{-p-(M-1)/2}. \tag{25}$$

Moreover, the coefficients of z^n are

$$[z^n u^p]\frac{\partial^M B_k(z,u,v)}{\partial v^M}\Big|_{v=1} \sim \frac{2^{M-2}\Gamma(p+\frac{M}{2})}{p\sqrt{\pi}q_k^{(M+1)/2+n}c_k^{M-1}(\Gamma(p))^2}n^{p-(3-M)/2}, \tag{26}$$

which (after normalisation) lead to the next result.

Theorem 2.10. *The M-th moments, $\mathbb{E}(X_{n,p}^M)$, of the size of the ancestor tree of p randomly chosen nodes in a monotonically labelled binary tree of size n are asymptotically given by*

$$\mathbb{E}(X_{n,p}^M) = \frac{[z^n u^p]\frac{\partial^M B_k(z,u,v)}{\partial v^M}\big|_{v=1}}{p_k q_k^{-n}\frac{n^p}{\Gamma(p+1)}n^{-3/2}} = \frac{2^{M-2}}{\sqrt{\pi}c_k^{M-1}q_k^{(M+1)/2}p_k}\frac{\Gamma\left(p+\frac{M}{2}\right)}{\Gamma(p)}\,n^{M/2}\Big(1+\mathcal{O}\Big(\frac{1}{\sqrt{n}}\Big)\Big), \tag{27}$$

for $M\geq 2$, fixed $p\geq 1$ and $n\to\infty$, where $c_k = 2\Big(\prod_{i=2}^{k}(1-4q_i)\Big)^{-1/4}$ and $p_k = \frac{c_k}{4\sqrt{\pi q_k}}$.

Proof: Induction on M is employed. Let $M=2$, then the left hand side of (27) is

$$\begin{aligned}\frac{[z^n u^p]\frac{\partial^2 B_k(z,u,v)}{\partial v^2}\big|_{v=1}}{p_k q_k^{-n}p_k\frac{n^p}{\Gamma(p+1)}n^{-3/2}} &= \frac{[z^n u^p]\frac{\left(4z(1+u)\bar{y}_k(z)\frac{\partial B_k(z,u,v)}{\partial v}\Big|_{v=1}+2z(1+u)\left(\frac{\partial B_k(z,u,v)}{\partial v}\Big|_{v=1}\right)^2\right)}{1-2z(1+u)\bar{y}_k(z)}}{p_k q_k^{-n}\frac{n^p}{\Gamma(p+1)}n^{-3/2}} \\ &= \frac{p}{4\pi p_k^2 q_k^2}+\mathcal{O}\Big(\frac{1}{\sqrt{n}}\Big),\end{aligned} \tag{28}$$

as computed earlier, which also equals the right hand side of (27). Now, assume (27) is true for M and prove it for $M+1$. We have

$$\begin{aligned}&\frac{1}{p_k q_k^{-n}\frac{n^p}{\Gamma(p+1)}n^{-3/2}}[z^n u^p]Mz(1+u)\sum_{i=0}^{M-1}\binom{M-1}{i}\frac{\partial^i B_k(z,u,v)}{\partial v^i}\Big|_{v=1}\frac{\partial^{M-i-1}B_k(z,u,v)}{\partial v^{M-i-1}}\Big|_{v=1} \\ &\sim \frac{2^{M-5}\Gamma(M+1)}{\pi c_k^{M-3}q_k^{(M-1)/2}p_k^2(\Gamma(p))^2}\sum_{i=0}^{M-1}\frac{\Gamma(p+\frac{i}{2})\Gamma(p+\frac{M-i-1}{2})}{\Gamma(i+1)\Gamma(M-i)},\end{aligned} \tag{29}$$

as well as

$$\begin{aligned}&\frac{1}{p_k q_k^{-n}\frac{n^p}{\Gamma(p+1)}n^{-3/2}}[z^n u^p]z(1+u)\sum_{i=1}^{M-1}\binom{M}{i}\frac{\partial^i B_k(z,u,v)}{\partial v^i}\Big|_{v=1}\frac{\partial^{M-i}B_k(z,u,v)}{\partial v^{M-i}}\Big|_{v=1} \\ &\sim \frac{2^{M-4}\Gamma(M+1)}{\pi c_k^{M-2}q_k^{M/2}p_k^2(\Gamma(p))^2}\sum_{i=1}^{M-1}\frac{\Gamma(p+\frac{i}{2})\Gamma(p+\frac{M-i}{2})}{\Gamma(i+1)\Gamma(M-i+1)}.\end{aligned} \tag{30}$$

■

It is well-known that the density function of the generalised Gamma distribution is $g(a,h,A;x) = \frac{|h|}{\Gamma(a)A}\left(\frac{x}{A}\right)^{ah-1}e^{-\left(\frac{x}{A}\right)^h}$, for $x>0$ and it will be used for the derivation of our limiting distribution.

Theorem 2.11. *The probability, $\mathbb{P}\{X_{n,p} = m\}$, that the size of the ancestor tree of p randomly selected nodes in a monotonically labelled binary tree of size n is equal to m is asymptotically*

$$\mathbb{P}\{X_{n,p} = m\} = m^{p-1}\frac{q_k^{p/2}}{n^p(p-1)!}e^{-q_k^{1/2}\frac{m}{n}}\Big(1+\mathcal{O}\Big(\frac{1}{\sqrt{n}}\Big)+\mathcal{O}\Big(\frac{n}{m}\Big)+\mathcal{O}\Big(\frac{1}{m}\Big)\Big)\,, \qquad (31)$$

for $p \geq 1$, $n \to \infty$ and $m \leq K\sqrt{n}$ ($K > 0$ arbitrary but fixed).

The above result leads to the following theorem which characterises the limiting distribution for the size of the ancestor trees in monotonic binary trees.

Theorem 2.12. *For fixed $p \geq 1$, $x > 0$, $m = x\sqrt{n}$ we have*

$$\sqrt{n}\,\mathbb{P}\{X_{n,p} = m\} \sim x^{p-1}\frac{q_k^{p/2}}{(p-1)!}e^{-q_k^{1/2}x} = g\Big(p, 1, \frac{1}{\sqrt{q_k}}; x\Big)\,. \qquad (32)$$

Thus the limiting distribution of the normalised random variable $\frac{X_{n,p}}{\sqrt{n}}$ is asymptotically, for fixed $p \geq 1$ and $n \to \infty$, a generalised Gamma distribution with parameters $\Big(p, 1, \frac{1}{\sqrt{q_k}}\Big)$.

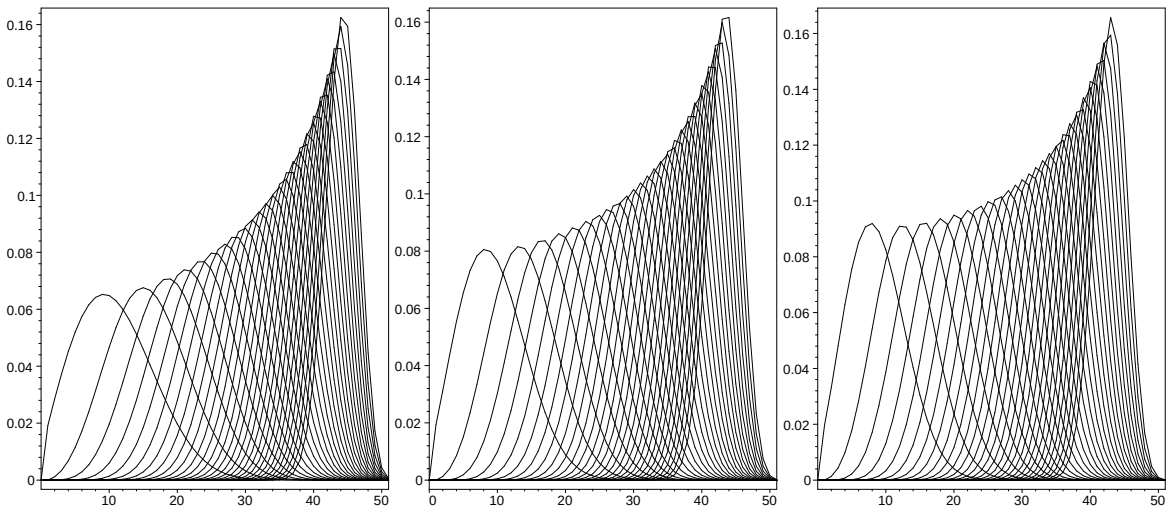


FIGURE 1. The probability distributions for the size of the ancestor tree in monotonically labelled binary trees for $n = 51$, $p = 1 \ldots 30$ and $k = 1, 2, 3$.

2.3. **The Steiner distance.** Our second parameter of interest is now analysed. The generating function of the Steiner distance in ordinary binary trees is presented first and then modified suitably for the monotonic case.

Theorem 2.13. *The generating function for the Steiner distance in binary trees is*

$$C(z,u,v) = \frac{zv(1+u)B^2(z,u,v) - 2zv\,T(z)\,B(z,u,v) + zT^2(z)(v-2) + T(z)}{1-2z\,T(z)}\,. \qquad (33)$$

Proof: The proof is similar to the one in Theorem 2.1. One starts with the recurrence

$$\psi_p(z,v) = zv\sum_{i=0}^{p-1}\phi_i(z,v)\,\phi_{p-1-i}(z,v) + zv\sum_{i=1}^{p-1}\phi_i(z,v)\,\phi_{p-i}(z,v) + 2z\,\psi_0(z,v)\,\psi_p(z,v)\,, p \geq 1\,, \qquad (34)$$

and then sums over p on both sides. ■

This theorem is adapted (by replacing $B(z,u,v)$ with $B_k(z,u,v)$ and $T(z)$ with $y_k(z)$) and gives the corresponding result for monotonic binary trees.

Theorem 2.14. *The generating function for the Steiner distance in monotonic binary trees satisfies the equation*

$$C_k(z,u,v) = \frac{zv(1+u)B_k^2(z,u,v) - 2zvy_k(z)B_k(z,u,v) + zy_k^2(z)(v-2) + y_k(z)}{1-2zy_k(z)}. \tag{35}$$

The expectation of the Steiner distance is computed first. After denoting the derivative of C_k with respect to v (and evaluated at 1) by $\Omega_k(z,u)$ and making the usual substitutions we obtain

$$\begin{aligned}\Omega_k(z,u) &= \frac{1}{1-2z\,y_k(z)}\Big(z(1+u)\,\bar{y}_k^2(z) + 2z(1+u)\,\bar{y}_k(z)\,\beta_k(z,u)\\ &\quad - 2z\,y_k(z)\,\bar{y}_k(z) - 2z\,y_k(z)\,\beta_k(z,u) + z\,y_k^2(z)\Big)\\ &\sim \frac{\sqrt{q_k-z}}{q_k\sqrt{q_k-z(1+u)}} - \frac{zu}{2q_k\sqrt{q_k-z}\sqrt{q_k-z(1+u)}}.\end{aligned} \tag{36}$$

In order to proceed with our analysis, we require the asymptotic behaviour of the coefficients in Ω_k.

Lemma 2.15. *As $z \to q_k$, the coefficients of u^p in $\Omega_k(z,u)$ have the form*

$$[u^p]\Omega_k(z,u) \sim \frac{p-1}{2^{2p}\,(2p-1)\,q_k}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p}. \tag{37}$$

Proof: The coefficient in the first term of (36) has been computed already

$$[u^p]\frac{\sqrt{q_k-z}}{q_k\sqrt{q_k-z(1+u)}} \sim \frac{1}{2^{2p}q_k}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p}. \tag{38}$$

Then the coefficient in the second term is as follows

$$\begin{aligned}&-[u^p]\frac{zu}{2q_k\sqrt{q_k-z}\sqrt{q_k-z(1+u)}} = -\frac{z}{2q_k^2}[u^{p-1}]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{1/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{1/2}}\\ &= -\frac{z^p p}{2^{2p}q_k^{p+1}(2p-1)}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p} \sim -\frac{p}{2^{2p}q_k(2p-1)}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p},\ z\to q_k\,.\end{aligned} \tag{39}$$

Finally, adding the above coefficients gives the required result. ■

Next, the coefficients of z^n in $\Omega_k(z,u)$ turn out to be $[z^p u^p]\Omega_k(z,u) \sim \frac{p-1}{(2p-1)\,2^{2p}\,q_k^{n+1}}\binom{2p}{p}\frac{n^{p-1}}{\Gamma(p)}$ and their normalisation leads to the first moment.

Theorem 2.16. *The expectation for the Steiner distance in monotonic binary trees is*

$$E_{n,p}^{(k)} \sim \frac{(p-1)\,p}{(2p-1)\,2^{2p}\,p_k\,q_k}\binom{2p}{p}\sqrt{n}\,,\ n\to\infty\,,\ \textit{fixed } k. \tag{40}$$

We move on to finding the next moment for the Steiner distance. The second derivative of C_k with respect to v (evaluated at $v=1$), denoted by $\Lambda_k(z,u)$, is

$$\begin{aligned}\Lambda_k(z,u) &= \frac{1}{1-2z\,y_k(z)}\Big(4z\,(1+u)\,\bar{y}_k(z)\,\beta_k(z,u)+2z\,(1+u)\,\beta_k^2(z,u)\\ &\quad +2z\,(1+u)\,\bar{y}_k(z)\,\Theta_k(z,u)-4z\,y_k(z)\,\beta_k(z,u)-2z\,y_k(z)\,\Theta_k(z,u)\Big) \qquad (41)\\ &\sim \frac{3zu}{2c_kq_k(q_k-z(1+u))^{3/2}}-\frac{q_k-z}{c_kq_k(q_k-z(1+u))^{3/2}}\,.\end{aligned}$$

This simplified form of the derivative is used for finding the behaviour of the coefficients of u^p and z^n.

Lemma 2.17. *The coefficients of u^p in $\Lambda_k(z,u)$ are*

$$[u^p]\Lambda_k(z,u)\sim\frac{(p-1)}{2^{2p}\,c_k\,q_k^{3/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ z\to q_k\,. \qquad (42)$$

Proof: The coefficients of u^p in the first term of (41) is

$$\begin{aligned}&[u^p]\frac{3zu}{2c_kq_k(q_k-z(1+u))^{3/2}}=\frac{3z}{2c_kq_k^{5/2}}[u^{p-1}]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{3/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{3/2}}\\ &=\frac{3pz^p}{2^{2p}c_kq_k^{p+3/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2}\sim\frac{3p}{2^{2p}c_kq_k^{3/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ z\to q_k\,,\end{aligned} \qquad (43)$$

and similarly

$$-[u^p]\frac{q_k-z}{c_kq_k(q_k-z(1+u))^{3/2}}\sim-\frac{2p+1}{2^{2p}c_kq_k^{3/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ z\to q_k\,. \qquad (44)$$

Thus combining these coefficients gives the result in the lemma. ∎

Next, we find that the coefficients of z^n are: $[z^nu^p]\,\Lambda_k(z,u)\sim\frac{p-1}{2^{2p+2}\sqrt{\pi}\,p_k\,q_k^{n+2}}\binom{2p}{p}\frac{n^{p-1/2}}{\Gamma(p+\frac12)}$, which give the following result.

Theorem 2.18. *The variance for the Steiner distance in monotonic binary trees is*

$$V_{n,p}^{(k)}=\frac{(p-1)}{p_k^2q_k^2}\Big(\frac{1}{4\pi}-\frac{(p-1)\,p^2}{(2p-1)^2\,2^{4p}}\binom{2p}{p}^2\Big)n+\mathcal{O}(\sqrt{n}\,)\,, \qquad (45)$$

as $n\to\infty$ and fixed k.

It is interesting to note that, by Stirling's formula, $\frac{1}{4\pi}-\frac{(p-1)\,p^2}{(2p-1)^2\,2^{4p}}\binom{2p}{p}^2$ goes to zero as $p\to\infty$.

2.4. **The limiting distribution for the Steiner distance.** Let $Y_{n,p}$ be the random variable which counts the size of the Steiner distance for p randomly chosen nodes in a monotonically labelled binary tree of size n. So we are interested in the probabilities $\mathbb{P}\{Y_{n,p}=m\}=\frac{[z^nu^pv^m]C_k(z,u,v)}{b_n^{(k)}\binom{n}{p}}$.

Theorem 2.19. *The M-th moments, $\mathbb{E}(Y_{n,p}^M)$, of the Steiner distance for p randomly chosen nodes in a monotonically labelled binary tree of size n are asymptotically given by*

$$\mathbb{E}(Y_{n,p}^M) = \frac{[z^n u^p]\frac{\partial^M C_k(z,u,v)}{\partial v^M}\big|_{v=1}}{p_k q_k^{-n}\frac{n^p}{\Gamma(p+1)}n^{-3/2}} = \frac{2^{M-2}}{\sqrt{\pi}c_k^{M-1}q_k^{(M+1)/2}p_k}\frac{\Gamma\left(p+\frac{M}{2}-1\right)}{\Gamma(p-1)}\,n^{M/2}\Big(1+\mathcal{O}\Big(\frac{1}{\sqrt{n}}\Big)\Big), \tag{46}$$

for fixed $p \geq 2$ and $n \to \infty$, where $c_k = 2\Big(\prod_{i=2}^{k}(1-4q_i)\Big)^{-1/4}$ and $p_k = \frac{c_k}{4\sqrt{\pi q_k}}$.

Once again the density function of the generalised Gamma distribution is used to derive the limiting distribution.

Theorem 2.20. *The probability, $\mathbb{P}\{Y_{n,p} = m\}$, that the Steiner distance for p randomly selected nodes in a monotonically labelled binary tree of size n is equal to m is asymptotically*

$$\mathbb{P}\{Y_{n,p} = m\} \sim m^{p-2}\frac{q_k^{(p-1)/2}}{n^{p-1}(p-2)!}e^{-q_k^{1/2}\frac{m}{n}}\Big(1+\mathcal{O}\Big(\frac{1}{\sqrt{n}}\Big)+\mathcal{O}\Big(\frac{n}{m}\Big)+\mathcal{O}\Big(\frac{1}{m}\Big)\Big), \tag{47}$$

for $p \geq 2$, $n \to \infty$ and $m \leq K\sqrt{n}$ ($K > 0$ arbitrary but fixed).

Now we are in a position to find the limiting behaviour of the Steiner distance for our monotonic binary trees.

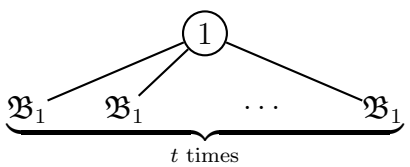
Theorem 2.21. *For fixed $p \geq 2$, $x > 0$, $m = x\sqrt{n}$ we have*

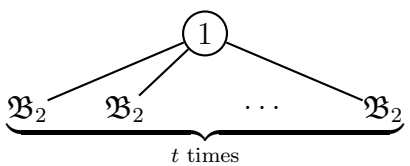
$$\sqrt{n}\,\mathbb{P}\{Y_{n,p} = m\} \sim x^{p-2}\frac{q_k^{(p-1)/2}}{(p-2)!}e^{-q_k^{1/2}x} = g\Big(p-1, 1, \frac{1}{\sqrt{q_k}}; x\Big). \tag{48}$$

Thus the limiting distribution of the normalised random variable $\frac{Y_{n,p}}{\sqrt{n}}$ is asymptotically, for fixed $p \geq 2$ and $n \to \infty$, a generalised Gamma distribution with parameters $\Big(p-1, 1, \frac{1}{\sqrt{q_k}}\Big)$.

3. T-ary Trees

In this section we generalise the binary tree model and consider the class of monotonically labelled t-ary trees, whose defining equations are given by

$\mathfrak{B}_1 = \square +$ 

$\mathfrak{B}_2 = \widetilde{\mathfrak{B}}_1 +$ 

$\vdots$

$\mathfrak{B}_k = \widetilde{\mathfrak{B}}_{k-1} + \cdots$

Thus the generating functions are $y_k(z) = y_{k-1}(z) + z\,y_k^t(z)$ with $y_0(z) = 1$. From [11] we know that $y_k \sim y_k(q_k) - C_k(t)\sqrt{q_k - z} = \frac{1}{r_k} - C_k(t)\sqrt{q_k - z}$ as z tends to the singularity q_k, where $C_k(t)$ are constants as defined below.

Lemma 3.1. *The constants $C_k(t)$ are given by*

$$\Big[2\Big(\sum_{i=1}^{k-1}\Big(\frac{(y_k(q_k)-y_{k-1}(q_k))}{z}\prod_{j=i}^{k-1}\frac{1}{1-t\frac{(y_j(q_k)-y_{j-1}(q_k))}{y_j(q_k)}}\Big)+\frac{(y_k(q_k)-y_{k-1}(q_k))}{z}\Big)\frac{y_k(q_k)}{t-1}\Big]^{1/2}. \tag{49}$$

3.1. **Size of the ancestor tree.** We begin this section by first presenting the generating function for t-ary trees (which is a generalisation of the result in the binary trees section).

Theorem 3.2. *The generating function for the size of the ancestor tree in t-ary trees is*

$$G(z,u,v) = zv(1+u)G^t(z,u,v) - zv\,M^t(z) + M(z)\,. \tag{50}$$

This result is adjusted accordingly to take into account the monotonic nature of the trees and leads to the next theorem.

Theorem 3.3. *The generating functions definining the size of the ancestor tree in monotonic t-ary trees are*

$$G_k(z,u,v) = G_{k-1}(z,u,v) + zv(1+u)G_k^t(z,u,v) + (1-v)zy_k^t(z)\,,\ G_0(z,u,v) = 1\,. \tag{51}$$

The first step in getting the variance for the size of the ancestor tree is to differentiate G_k with respect to v and evaluate it at $v = 1$. Then we denote it by $g_k(z,u)$ and set $G_k(z,u,1) = \bar{y}_k(z)$

$$(1 - tz(1+u)\bar{y}_k^{t-1}(z))g_k(z,u) = g_{k-1}(z,u) + \bar{y}_k(z) - \bar{y}_{k-1}(z) - y_k(z) + y_{k-1}(z)\,, \tag{52}$$

since $zy_k^t(z) = y_k(z) - y_{k-1}(z)$ and $z(1+u)y_k^t(z(1+u)) = \bar{y}_k(z) - \bar{y}_{k-1}(z)$.

Lemma 3.4. *The solution of $g_k(z,u)$ is*

$$g_k(z,u) = \sum_{j=1}^{k}\Big[\prod_{i=j}^{k}(1 - tz(1+u)\bar{y}_i^{t-1}(z))\Big]^{-1}\big(\bar{y}_j(z) - \bar{y}_{j-1}(z) - y_j(z) + y_{j-1}(z)\big)\,, \tag{53}$$

and its dominant term has the form

$$g_k(z,u) \sim \frac{\bar{y}_k(z) - y_k(z)}{1 - t\frac{(\bar{y}_k(z)-\bar{y}_{k-1}(z))}{\bar{y}_k(z)}}\,. \tag{54}$$

Proof: Both sides of (52) are multiplied with the product $\prod_{i=1}^{k-1}(1 - tz(1+u)\bar{y}_i^{t-1}(z))$. Then we sum over the k's and obtain

$$\begin{aligned}&\Big[\prod_{i=1}^{k-1}(1 - tz(1+u)\bar{y}_i^{t-1}(z))\Big](1 - tz(1+u)\bar{y}_k^{t-1}(z))g_k(z,u)\\ &= \sum_{j=0}^{k-1}\Big[\prod_{i=1}^{j}(1 - tz(1+u)\bar{y}_i^{t-1}(z))\Big]\big(\bar{y}_j(z) - \bar{y}_{j-1}(z) - y_j(z) + y_{j-1}(z)\big)\,,\end{aligned} \tag{55}$$

which yields the required result upon division with $\prod_{i=1}^{k}(1-tz(1+u)\bar{y}_i^{t-1}(z))$. The dominant term in this occurs when $i=j=k$. Since the singularities of $y_{k-1}(z)$ and $\bar{y}_{k-1}(z)$ are further away from the origin than those of $y_k(z)$ and $\bar{y}_k(z)$ they will not contribute to the main term in $g_k(z,u)$ and we do not include them. Thus the dominant term in $g_k(z,u)$ is as given in the statement of the theorem. ■

Now we substitute the known asymptotic expansions in (54) and obtain

$$g_k(z,u) \sim \frac{\sqrt{q_k - z}}{(t-1)(tq_k)^{1/(t-1)}\sqrt{q_k - z(1+u)}}\,. \tag{56}$$

which enables us to look at the behaviour of the coefficients in g_k.

Lemma 3.5. *The coefficients of u^p in $g_k(z,u)$ have the form*

$$[u^p]g_k(z,u) \sim \frac{1}{2^{2p}q_k^{1/(t-1)}(t-1)t^{1/(t-1)}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p}\,. \tag{57}$$

Proof: The coefficients of u^p in (56) are computed, as follows

$$\begin{aligned}[u^p]\frac{\sqrt{q_k - z}}{(t-1)(tq_k)^{1/(t-1)}\sqrt{q_k - z(1+u)}} &= \frac{\Big(1-\frac{z}{q_k}\Big)}{(t-1)(tq_k)^{1/(t-1)}}[u^p]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{1/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{1/2}}\\
&= \frac{z^p}{2^{2p}(t-1)(tq_k)^{1/(t-1)}q_k^p}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p} \sim \frac{1}{2^{2p}(t-1)(tq_k)^{1/(t-1)}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p},\end{aligned} \tag{58}$$

as z tends to the dominant singularity q_k. ■

Finally, we use Proposition 1.1, with $\alpha=-p$, to obtain the asymptotic expansion for the coefficients of z^n: $[z^n u^p]g_k(z,u) \sim \frac{1}{2^{2p}q_k^{n/(t-1)}(t-1)t^{1/(t-1)}}\binom{2p}{p}\frac{n^{p-1}}{\Gamma(p)}$.

Theorem 3.6. *The expectation for the size of the ancestor tree in monotonic t-ary trees is*

$$E_{n,p}^{(k)} \sim \frac{p}{2^{2p}q_k^{1/(t-1)}p_k(t-1)t^{1/(t-1)}}\binom{2p}{p}\sqrt{n}\,,\ n\to\infty\,,\ \textit{fixed } k\,. \tag{59}$$

One can proceed with finding the variance for the size of the ancestor tree. The second derivative of G_k with respect to v (evaluated at 1) is denoted by $h_k(z,u)$ and gives

$$\begin{aligned}(1-tz(1+u)\bar{y}_k^{t-1}(z))h_k(z,u) = h_{k-1}(z,u) &+ 2tz(1+u)\bar{y}_k^{t-1}(z)g_k(z,u)\\ &+ tz(1+u)\bar{y}_k^{t-2}(z)g_k^2(z,u)(t-1)\,.\end{aligned} \tag{60}$$

Lemma 3.7. *The solution of $h_k(z,u)$ has the closed form*

$$\begin{aligned}h_k(z,u) = \sum_{j=1}^{k}\Big\{\Big(\prod_{i=j}^{k}(1-tz(1+u)\bar{y}_i^{t-1}(z))\Big)^{-1}\big[&2tz(1+u)\bar{y}_j^{t-1}(z)g_j(z,u)\\ &+ tz(1+u)\bar{y}_j^{t-2}(z)g_j^2(z,u)(t-1)\big]\Big\}\,,\end{aligned} \tag{61}$$

and the dominant term is

$$h_k(z,u) \sim \frac{2tz(1+u)\bar{y}_k^{t-1}(z)g_k(z,u) + tz(1+u)\bar{y}_k^{t-2}(z)g_k^2(z,u)(t-1)}{1-tz(1+u)\bar{y}_k^{t-1}(z))}\,. \tag{62}$$

Proof: The proof is omitted as it is similar to the one for Lemma 3.4. ■

Furthermore, if we substitute $\bar{y}_{k-1}(z)$ and $\bar{y}_k(z)$ with their known asymptotic expressions then $h_k(z,u)$ simplifies to

$$h_k(z,u) \sim \frac{zu}{C_k(t)(t-1)^2(q_k - z(1+u))^{3/2}(tq_k)^{2/(t-1)}}\,. \tag{63}$$

Our analysis requires us to find the asymptotic behaviour of the coefficients of u^p and z^n in the simplified version of h_k.

Lemma 3.8. *The coefficients of u^p in $h_k(z,u)$ have the asymptotic expansion*

$$[u^p]h_k(z,u) \sim \frac{pq_k^{-(3+t)/(2(t-1))}}{2^{2p-1}C_k(t)(t-1)^2t^{2/(t-1)}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2}. \tag{64}$$

Proof: The coefficient of u^p in (63) gives the result of the lemma

$$\begin{aligned}
&[u^p]\frac{zu}{C_k(t)(t-1)^2(q_k - z(1+u))^{3/2}(tq_k)^{2/(t-1)}}\\
&= \frac{z}{C_k(t)(t-1)^2q_k^{3/2}(tq_k)^{2/(t-1)}}[u^{p-1}]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{3/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{3/2}}\\
&= \frac{pz^p}{2^{2p-1}C_k(t)(t-1)^2q_k^{p+1/2}(tq_k)^{2/(t-1)}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2}\\
&\sim \frac{p}{2^{2p-1}C_k(t)(t-1)^2q_k^{1/2}(tq_k)^{2/(t-1)}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ z\to q_k\,.
\end{aligned} \tag{65}$$

■

Proposition 1.1, with $\alpha = -p-1/2$, is then used to find the asymptotic expansion for the coefficients of z^n: $[z^nu^p]h_k(z,u) \sim \frac{pq_k^{-(3+t)/(2(t-1))-n}}{2^{2p-1}C_k(t)(t-1)^2t^{2/(t-1)}}\binom{2p}{p}\frac{n^{p-1/2}}{\Gamma(p+\frac12)}$, which will be normalised and give the second moment.

Theorem 3.9. *The variance for the size of the ancestor tree in monotonic t-ary trees is*

$$V_{n,p}^{(k)} = \frac{p}{p_k(t-1)^2t^{2/(t-1)}}\Big(\frac{2}{\sqrt{\pi}C_k(t)q_k^{(3+t)/(2(t-1))}} - \frac{p}{2^{4p}q_k^{2/(t-1)}p_k}\binom{2p}{p}^2\Big)n + \mathcal{O}(\sqrt{n}\,), \tag{66}$$

for $n\to\infty$ and k fixed.

3.2. **The Steiner distance.** We begin by presenting the generating function of ordinary t-ary trees which gives us a starting point for developing the corresponding result for monotonic t-ary trees.

Theorem 3.10. *The generating function for the Steiner distance in t-ary trees is*

$$F(z,u,v) = \frac{zv(1+u)G^t(z,u,v) - tzvM^{t-1}(z)G(z,u,v) + zM^t(z)(v-t) + M(z)}{1-tzM^{t-1}(z)}. \quad (67)$$

This theorem is modified, by replacing $M(z)$ with $y_k(z)$, so that it gives the Steiner distance for the t-ary tree model in [11].

Theorem 3.11. *The Steiner distance for monotonic t-ary trees satisfies*

$$F_k(z,u,v) = \frac{zv(1+u)G_k^t(z,u,v) - tzvy_k^{t-1}(z)G_k(z,u,v) + zy_k^t(z)(v-t) + y_k(z)}{1-tzy_k^{t-1}(z)}. \quad (68)$$

The first moment is computed by differentiating F_k with respect to v and evaluating at $v=1$. Then the known asymptotic formulas are substituted in to simplify the derivative

$$\begin{aligned}\frac{\partial F_k(z,u,v)}{\partial v}\Big|_{v=1} &= \frac{1}{1-tzy_k^{t-1}(z)}\Big(z(1+u)\bar{y}_k^t(z) + tz(1+u)\bar{y}_k^{t-1}(z)g_k(z,u)\\ &\quad - tzy_k^{t-1}(z)\bar{y}_k(z,u,1) - tzy_k^{t-1}(z)g_k(z,u) + zy_k^t(z)\Big)\\ &\sim \frac{2\sqrt{q_k-z}}{(t-1)(tq_k)^{1/(t-1)}\sqrt{q_k-z(1+u)}}\\ &\quad - \frac{zu}{(t-1)(tq_k)^{1/(t-1)}\sqrt{q_k-z}\sqrt{q_k-z(1+u)}},\end{aligned} \quad (69)$$

and this enables us to consider the behaviour of the coefficients of u^p around their dominant singularity q_k.

Lemma 3.12. *The coefficients of u^p in $\frac{\partial F_k(z,u,v)}{\partial v}\big|_{v=1}$ are*

$$[u^p]\frac{\partial F_k(z,u,v)}{\partial v}\Big|_{v=1} \sim \frac{p-1}{2^{2p-1}(2p-1)(t-1)(tq_k)^{1/(t-1)}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p}. \quad (70)$$

Proof: The coefficient of u^p from the first term in (69) is

$$\begin{aligned}[u^p]\frac{2\sqrt{q_k-z}}{(t-1)(tq_k)^{1/(t-1)}\sqrt{q_k-z(1+u)}} &= \frac{2\Big(1-\frac{z}{q_k}\Big)^{1/2}}{(t-1)(tq_k)^{1/(t-1)}}[u^p]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{1/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{1/2}}\\ = \frac{z^p}{2^{2p-1}(t-1)(tq_k)^{1/(t-1)}q_k^p}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p} &\sim \frac{1}{2^{2p-1}(t-1)(tq_k)^{1/(t-1)}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p},\end{aligned} \quad (71)$$

for $z \to q_k$. Moreover the second term yields

$$\begin{aligned}&-[u^p]\frac{zu}{(t-1)(tq_k)^{1/(t-1)}\sqrt{q_k-z}\sqrt{q_k-z(1+u)}}\\ &\sim -\frac{p}{2^{2p-1}(2p-1)(t-1)(tq_k)^{1/(t-1)}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p},\end{aligned} \quad (72)$$

around the dominant singularity q_k. Adding these two coefficients gives the result of the lemma. ■

Next, the coefficients of z^n are computed: $[z^n u^p]\frac{\partial F_k(z,u,v)}{\partial v}\big|_{v=1} \sim \frac{(p-1)}{2^{2p-1}(2p-1)(t-1)(tq_k)^{1/(t-1)}q_k^n}\binom{2p}{p}\frac{n^{p-1}}{\Gamma(p)}$, which will be normalised and give the next result.

Theorem 3.13. *The expectation of the Steiner distance in monotonic t-ary trees is*

$$E_{n,p}^{(k)} \sim \frac{p(p-1)}{2^{2p-1}(2p-1)(t-1)(tq_k)^{1/(t-1)}p_k}\binom{2p}{p}\sqrt{n}\,,\ n\to\infty\,,\ \textit{fixed } k\,. \tag{73}$$

We proceed by computing the next moment for the Steiner distance. After differentiating F_k twice with respect to v (and evaluating at 1) we get

$$\begin{aligned}\frac{\partial^2 F_k(z,u,v)}{\partial v^2}\Big|_{v=1} &= \frac{1}{1-tzy_k^{t-1}(z)}\Big(2tz(1+u)\bar{y}_k^{t-1}(z)g_k(z,u)+tz(1+u)\bar{y}_k^{t-2}(z)g_k^2(z,u)(t-1)\\ &+tz(1+u)\bar{y}_k^{t-1}(z)h_k(z,u)-2tzy_k^{t-1}(z)g_k(z,u)-tzy_k^{t-1}(z)h_k(z,u)\Big)\\ &\sim \frac{zu}{C_k(t)(t-1)^2(tq_k)^{2/(t-1)}(q_k-z(1+u))^{3/2}}-\frac{2}{C_k(t)(t-1)^2(tq_k)^{2/(t-1)}\sqrt{q_k-z(1+u)}}\,.\end{aligned} \tag{74}$$

This asymptotic expression makes it easy for us to analyse the behaviour of its coefficients.

Lemma 3.14. *The coefficients of u^p in $\frac{\partial^2 F_k(z,u,v)}{\partial v^2}\big|_{v=1}$ are*

$$[u^p]\frac{\partial^2 F_k(z,u,v)}{\partial v^2}\Big|_{v=1} \sim \frac{p-1}{2^{2p-1}C_k(t)(t-1)^2t^{2/(t-1)}q_k^{(3+t)/(2(t-1))}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2}. \tag{75}$$

Proof: The coefficients of u^p from the terms in (74) are extracted as follows

$$\begin{aligned}&[u^p]\frac{zu}{C_k(t)(t-1)^2(tq_k)^{2/(t-1)}(q_k-z(1+u))^{3/2}}\\ &=\frac{z}{C_k(t)(t-1)^2(tq_k)^{2/(t-1)}q_k^{3/2}}[u^{p-1}]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{3/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{3/2}}\\ &=\frac{pz^p}{2^{2p-1}C_k(t)(t-1)^2(tq_k)^{2/(t-1)}q_k^{p+1/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2}\\ &\sim\frac{p}{2^{2p-1}C_k(t)(t-1)^2(tq_k)^{2/(t-1)}q_k^{1/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ z\to q_k\,,\end{aligned} \tag{76}$$

and similarly

$$\begin{aligned}&-[u^p]\frac{2}{C_k(t)(t-1)^2(tq_k)^{2/(t-1)}\sqrt{q_k-z(1+u)}}\\ &\sim\frac{1}{2^{2p-1}C_k(t)(t-1)^2(tq_k)^{2/(t-1)}q_k^{1/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ z\to q_k\,.\end{aligned} \tag{77}$$

By adding these two coefficients we obtain the result of the lemma. ■

The last required coefficients are:

$$[z^n u^p]\frac{\partial^2 F_k(z,u,v)}{\partial v^2}\Big|_{v=1} \sim \frac{p-1}{2^{2p-1}C_k(t)(t-1)^2 t^{2/(t-1)} q_k^{(3+t)/(2(t-1))} q_k^n}\binom{2p}{p}\frac{n^{p-1/2}}{\Gamma(p+\frac{1}{2})}\,,$$

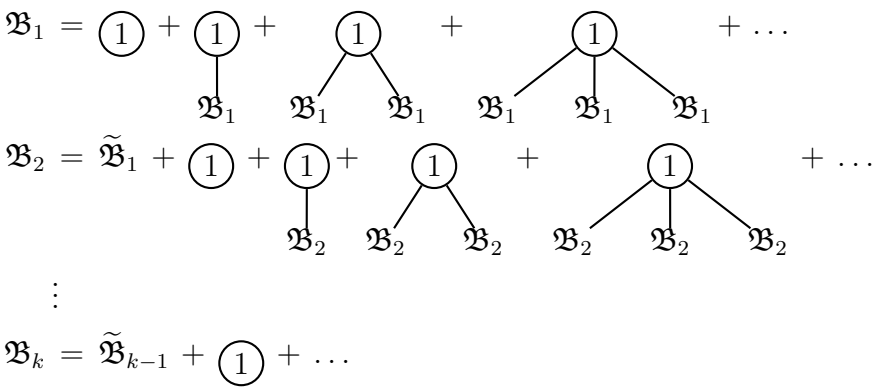
and after normalisation they give rise to the second moment.

Theorem 3.15. *The variance for the Steiner distance in monotonic t-ary trees is*

$$\begin{aligned} V_{n,p}^{(k)} = \frac{(p-1)}{p_k(t-1)^2 t^{2/(t-1)}}\Big(&\frac{2}{\sqrt{\pi}C_k(t)q_k^{(3+t)/(2(t-1))}p_k} \\ &- \frac{p^2(p-1)}{2^{4p-2}(2p-1)^2 q_k^{2/(t-1)}p_k}\binom{2p}{p}^2\Big)n + \mathcal{O}(\sqrt{n}\,),\ n\to\infty\,,\ \textit{fixed } k\,. \end{aligned} \tag{78}$$

4. Ordered Trees

Finally, we look at the class, $\mathfrak{B}_k$, of ordered trees whose nodes are monotonically labelled with $1,2,\ldots,k$ and we let $\widetilde{\mathfrak{B}_k}$ be the class of ordered trees whose nodes are monotonically labelled with $2,\ldots,k+1$. The defining equations for the classes $\mathfrak{B}_k$ are given by



and the generating functions are $y_k(z) = y_{k-1}(z) + \frac{z}{1-y_k(z)}$ with $y_0(z) = 1$.

4.1. **Size of the ancestor tree.** It is convenient to state the following result which can be adjusted accordingly in order to produce the equivalent one for monotonic ordered trees.

Theorem 4.1. *The size of the ancestor tree in ordered trees satisfies the functional equation*

$$P(z,u,v) = \frac{zv(1+u)}{1-P(z,u,v)} + (1-v)N(z)\,, \tag{79}$$

where $N(z) = \frac{1-\sqrt{1-4z}}{2}$ is the generating function for ordered trees.

The result corresponding to monotonic ordered trees arises after replacing $P(z,u,v)$ by $P_k(z,u,v)$ and $N(z)$ by $y_k(z)$.

Theorem 4.2. *The generating functions for the size of the ancestor tree in monotonic ordered trees are*

$$P_k(z,u,v) = P_{k-1}(z,u,v) + \frac{z\,v\,(1+u)}{1-P_k(z,u,v)} + \frac{z\,(1-v)}{1-y_k(z)}\,, \quad P_0(z,u,v) = 1\,. \tag{80}$$

The expectation for the size of the ancestor tree is the first quantity of interest. We denote the derivative of P_k with respect to v (evaluated at 1) by $A_k(z,u)$ and let $P_k(z,u,1) = y_k(z(1+u)) = \bar{y}_k(z)$. This gives

$$\Big(1 - \frac{z\,(1+u)}{(1-\bar{y}_k(z))^2}\Big) A_k(z,u) = A_{k-1}(z,u) + \frac{z\,(1+u)}{1-\bar{y}_k(z)} - \frac{z}{1-y_k(z)}\,. \tag{81}$$

Lemma 4.3. *The solution of $A_k(z,u)$ is*

$$A_k(z,u) = \sum_{j=1}^{k} \Big[\prod_{i=1}^{j} \Big(1 - \frac{z\,(1+u)}{(1-\bar{y}_i(z))^2}\Big)^{-1} \Big(\frac{z\,(1+u)}{1-\bar{y}_j(z)} - \frac{z}{1-y_j(z)} \Big) \Big]\,, \tag{82}$$

with $A_0(z,u) = 0$, hence the dominant term has the asymptotic form

$$A_k(z,u) \sim \Big(1 - \frac{z(1+u)}{(1-\bar{y}_k(z))^2}\Big)^{-1} \Big(\frac{z(1+u)}{1-\bar{y}_k(z)} - \frac{z}{1-y_k(z)} \Big)\,. \tag{83}$$

Proof: We multiply both sides of (81) with the product $\prod_{i=0}^{k-1} \big(1 - \frac{z\,(1+u)}{(1-\bar{y}_i(z))^2}\big)$ then sum over the k's and obtain

$$\begin{aligned} &\prod_{i=0}^{k-1} \Big(1 - \frac{z\,(1+u)}{(1-\bar{y}_i(z))^2}\Big) \Big(1 - \frac{z\,(1+u)}{(1-\bar{y}_k(z))^2}\Big) A_k(z,u) \\ &= \sum_{j=0}^{k-1} \Big[\prod_{i=0}^{j} \Big(1 - \frac{z\,(1+u)}{(1-\bar{y}_i(z))^2}\Big) \Big(\frac{z\,(1+u)}{1-\bar{y}_j(z)} - \frac{z}{1-y_j(z)} \Big) \Big]\,. \end{aligned} \tag{84}$$

The result of the lemma is obtained by dividing this with $\prod_{i=0}^{k} \big(1 - \frac{z\,(1+u)}{(1-\bar{y}_i(z))^2}\big)$ and the dominant term occurs when $i = j = k$. ■

Now from [11] one knows the following: $y_k(z) = y_k(q_k) - d_k\sqrt{q_k - z} + \mathcal{O}(q_k - z) = 1 - \sqrt{q_k} - d_k\sqrt{q_k - z} + \mathcal{O}(q_k - z)$ as $z \to q_k$, $y_{k-l}(q_k) = 1 - r_l w$ for $w = \frac{1}{r_l} = \sqrt{q_l}$ and $y_{k-1}(q_k) = 1 - r_1 w = 1 - 2\sqrt{q_k}$. We observe that $\frac{z}{1-y_k(z)} = y_k(z) - y_{k-1}(z)$ (and similarly for $\bar{y}_k(z)$), so these are substituted into (83)

$$A_k(z,u) \sim \frac{\sqrt{q_k}\sqrt{q_k - z}}{2\sqrt{q_k - z(1+u)}}\,. \tag{85}$$

This expression for the first derivative makes it possible for us to find the asymptotic behaviour of the coefficients of u^p and z^n around their dominant singularities q_k.

Lemma 4.4. *The coefficients of u^p in $A_k(z,u)$ are*

$$[u^p] A_k(z,u) \sim \frac{q_k^{1/2}}{2^{2p+1}} \binom{2p}{p} \Big(1 - \frac{z}{q_k}\Big)^{-p}\,. \tag{86}$$

Proof: We proceed by finding the coefficient of u^p from (85), which proves the result

$$[u^p]\frac{\sqrt{q_k}\sqrt{q_k-z}}{2\sqrt{q_k-z(1+u)}} = \frac{q_k^{1/2}\Big(1-\frac{z}{q_k}\Big)^{1/2}}{2}[u^p]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{1/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{1/2}} \tag{87}$$
$$= \frac{q_k^{1/2-p}z^p}{2^{2p+1}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p} \sim \frac{q_k^{1/2}}{2^{2p+1}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p},\ z\to q_k\,.$$

■

Next, the coefficients of z^n are computed: $[z^nu^p]A_k(z,u) \sim \frac{q_k^{1/2-n}}{2^{2p+1}}\binom{2p}{p}\frac{n^{p-1}}{\Gamma(p)}$, which will be normalised in order to find the first moment.

Theorem 4.5. *The expectation for the size of the ancestor tree in monotonic ordered trees is asymptotically given by*

$$E_{n,p}^{(k)} \sim \frac{p\sqrt{q_k}}{2^{2p+1}\,p_k}\binom{2p}{p}\sqrt{n}\,,\ n\to\infty\,,\ \textit{fixed } k. \tag{88}$$

Now we move on and compute the variance. For convenience, the second derivative of P_k with respect to v, evaluated at $v=1$, is denoted by $B_k(z,u)$, and it follows that

$$\Big(1-\frac{z\,(1+u)}{(1-\bar{y}_k(z))^2}\Big)B_k(z,u) = B_{k-1}(z,u) + \frac{2z\,(1+u)\,A_k(z,u)}{(1-\bar{y}_k(z))^2} + \frac{2z\,(1+u)\,A_k^2(z,u)}{(1-\bar{y}_k(z))^3}\,. \tag{89}$$

Lemma 4.6. *The solution to $B_k(z,u)$ is*

$$B_k(z,u) = \sum_{j=1}^{k}\Big[\prod_{i=1}^{j}\Big(1-\frac{z\,(1+u)}{(1-\bar{y}_i(z))^2}\Big)^{-1}\Big(\frac{2z\,(1+u)\,A_j(z,u)}{(1-\bar{y}_j(z))^2} + \frac{2z\,(1+u)\,A_j^2(z,u)}{(1-\bar{y}_j(z))^3}\Big)\Big], \tag{90}$$

with dominant term

$$B_k(z,u) \sim \Big(1-\frac{z(1+u)}{(1-\bar{y}_k(z))^2}\Big)^{-1}\Big(\frac{2z(1+u)A_k(z,u)}{(1-\bar{y}_k(z))^2} + \frac{2z(1+u)A_k^2(z,u)}{(1-\bar{y}_k(z))^3}\Big)\,. \tag{91}$$

Proof: The proof is similar to Lemma 4.3 so we omit it. ■

After substituting in the known asymptotic expansions, (91) simplifies to the following

$$B_k(z,u) \sim \frac{zuq_k}{4d_k(q_k-z(1+u))^{3/2}}\,, \tag{92}$$

and this expression is used in analysing the behaviour of its coefficients.

Lemma 4.7. *The coefficients of u^p in $B_k(z,u)$ are*

$$[u^p]B_k(z,u) \sim \frac{p\,q_k^{1/2}}{d_k\,2^{2p+1}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ d_k = \frac{2\sqrt{\pi}p_k}{\sqrt{q_k}}\,. \tag{93}$$

Proof: The extraction of the coefficient of u^p in (92) gives the following

$$\begin{aligned}[u^p]\frac{zuq_k}{4d_k(q_k-z(1+u))^{3/2}} &= \frac{z}{4d_kq_k^{1/2}}[u^{p-1}]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{3/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{3/2}}\\ &= \frac{pz^p}{2^{2p+1}d_kq_k^{p-1/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2} \sim \frac{pq_k^{1/2}}{2^{2p+1}d_k}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2},\ z\to q_k\,.\end{aligned} \tag{94}$$

■

Moreover, the coefficient of z^n is: $[z^nu^p]B_k(z,u) \sim \frac{p\,q_k^{1/2-n}}{d_k\,2^{2p+1}}\binom{2p}{p}\frac{n^{p-1/2}}{\Gamma(p+\frac12)} = \frac{p\,q_k^{1-n}}{\sqrt{\pi}p_k2^{2p+2}}\binom{2p}{p}\frac{n^{p-1/2}}{\Gamma(p+\frac12)}$, since $d_k=\frac{2\sqrt{\pi}p_k}{\sqrt{q_k}}$ (see [11]) and it yields the second moment upon normalisation.

Theorem 4.8. *The variance for the size of the ancestor tree in monotonic ordered trees is asymptotically*

$$V_{n,p}^{(k)} \sim \frac{pq_k}{p_k^2}\Big(\frac{1}{4\pi}-\frac{p}{2^{4p+2}}\binom{2p}{p}^2\Big)n+\mathcal{O}(\sqrt{n}\,),\ n\to\infty,\ \textit{fixed } k\,, \tag{95}$$

and $\frac{1}{4\pi}-\frac{p}{2^{4p+2}}\binom{2p}{p}^2\to 0$ *as* $p\to\infty$.

4.2. **The Steiner distance.** Once again, the derivation of the generating function for this parameter is most convenient when we use the defining equation of the Steiner distance in ordered trees as a starting point.

Theorem 4.9. *The Steiner distance in ordered trees has the generating function*

$$S(z,u,v)=\frac{P(z,u,v)(1-\frac{zv}{(1-N(z))^2})-\frac{z(1-v)N(z)}{(1-N(z))^2}}{1-\frac{z}{(1-N(z))^2}}\,. \tag{96}$$

Then, one can modify the above to produce an equivalent result for monotonic ordered trees.

Theorem 4.10. *The generating function for the Steiner distance in monotonic ordered trees is given by*

$$S_k(z,u,v)=\frac{P_k(z,u,v)\,\Big(1-\frac{z\,v}{(1-y_k(z))^2}\Big)-\frac{z\,(1-v)\,y_k(z)}{(1-y_k(z))^2}}{1-\frac{z}{(1-y_k(z))^2}}\,. \tag{97}$$

To compute the expectation, we differentiate $S_k(z,u,v)$ with respect to v and evaluate at 1. Then we substitute in the known asymptotic expressions and get

$$\begin{aligned}\frac{\partial S_k(z,u,v)}{\partial v}\Big|_{v=1} &= \frac{A_k(z,u)\,\Big(1-\frac{z}{(1-y_k(z))^2}\Big)-\frac{z\bar{y}_k(z)}{(1-y_k(z))^2}+\frac{z\,y_k(z)}{(1-y_k(z))^2}}{1-\frac{z}{(1-y_k(z))^2}}\\ &\sim \frac{\sqrt{q_k}\sqrt{q_k-z}}{\sqrt{q_k-z(1+u)}}-\frac{zu\sqrt{q_k}}{2\sqrt{q_k-z}\sqrt{q_k-z(1+u)}}\,.\end{aligned} \tag{98}$$

This simplified expansion helps us with finding the asymptotic behaviour of the coefficients in the first derivative of S_k.

Lemma 4.11. *The coefficients of u^p in $\frac{\partial S_k(z,u,v)}{\partial v}\big|_{v=1}$ have the form*

$$[u^p]\frac{\partial S_k(z,u,v)}{\partial v}\Big|_{v=1} \sim \frac{q_k^{1/2}(p-1)}{(2p-1)2^{2p}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p}. \tag{99}$$

Proof: We compute the coefficients of u^p in (98). The first term gives the following

$$\begin{aligned}[u^p]\frac{\sqrt{q_k}\sqrt{q_k-z}}{\sqrt{q_k-z(1+u)}} &= q_k^{1/2}\Big(1-\frac{z}{q_k}\Big)^{1/2}[u^p]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{1/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{1/2}}\\ &= \frac{z^p}{2^{2p}q_k^{p-1/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p} \sim \frac{q_k^{1/2}}{2^{2p}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p},\ z\to q_k\,.\end{aligned} \tag{100}$$

Finally, for the second term in (98) we find

$$-[u^p]\frac{zu\sqrt{q_k}}{2\sqrt{q_k-z}\sqrt{q_k-z(1+u)}} \sim -\frac{pq_k^{1/2}}{2^{2p}(2p-1)}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p},\ z\to q_k\,. \tag{101}$$

The result of the lemma is obtained by adding these two coefficients. ■

Now it is easy to find the other coefficients: $[z^n u^p]\frac{\partial S_k(z,u,v)}{\partial v}\big|_{v=1} \sim \frac{q_k^{1/2}\,(p-1)}{(2p-1)\,2^{2p}}\binom{2p}{p}q_k^{-n}\frac{n^{p-1}}{\Gamma(p)}$ and they are used to find the first moment.

Theorem 4.12. *The expectation for the Steiner distance in monotonic ordered trees is*

$$E_{n,p}^{(k)} \sim \frac{q_k^{1/2}\,p\,(p-1)}{p_k\,(2p-1)\,2^{2p}}\binom{2p}{p}\sqrt{n}\,,\ n\to\infty\,,\ \textit{fixed } k. \tag{102}$$

Finally, the variance of the Steiner distance can be analysed. The second derivative of S_k with respect to v (evaluated at $v=1$) yields

$$\begin{aligned}\frac{\partial^2 S_k(z,u,v)}{\partial v^2}\Big|_{v=1} &= \frac{B_k(z,u)\left(1-\frac{(y_k(z)-y_{k-1}(z))}{1-y_k(z)}\right)-\frac{2A_k(z,u)(y_k(z)-y_{k-1}(z))}{1-y_k(z)}}{1-\frac{(y_k(z)-y_{k-1}(z))}{1-y_k(z)}}\\ &\sim \frac{3zuq_k}{4d_k(q_k-z(1+u))^{3/2}}-\frac{q_k(q_k-z)}{2d_k(q_k-z(1+u))^{3/2}}\,.\end{aligned} \tag{103}$$

This asymptotic expression is used to find the behaviour of the coefficients of u^p and z^n in the above derivative of S_k.

Lemma 4.13. *The coefficients of u^p in $\frac{\partial^2 S_k(z,u,v)}{\partial v^2}\big|_{v=1}$ are*

$$[u^p]\,\frac{\partial^2 S_k(z,u,v)}{\partial v^2}\Big|_{v=1} \sim \frac{q_k^{1/2}\,(p-1)}{d_k\,2^{2p+1}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2}. \tag{104}$$

Proof: One proceeds by computing the coefficients of u^p in (103). For the first term we find

$$[u^p]\frac{3zuq_k}{4d_k(q_k-z(1+u))^{3/2}} = \frac{3z}{4d_kq_k^{1/2}}[u^{p-1}]\frac{1}{\Big(1-\frac{z}{q_k}\Big)^{3/2}\Big(1-\frac{\frac{zu}{q_k}}{1-\frac{z}{q_k}}\Big)^{3/2}}$$
$$= \frac{3pz^p}{2^{2p+1}q_k^{p-1/2}}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2} \sim \frac{3pq_k^{1/2}}{2^{2p+1}d_k}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2}\,,\ z\to q_k\,. \tag{105}$$

For the second term in (103) we obtain

$$-[u^p]\frac{q_k(q_k-z)}{2d_k(q_k-z(1+u))^{3/2}} \sim -\frac{(2p+1)q_k^{1/2}}{2^{2p+1}d_k}\binom{2p}{p}\Big(1-\frac{z}{q_k}\Big)^{-p-1/2}\,,\ z\to q_k\,. \tag{106}$$

By adding these two coefficients we get the result of the lemma. ∎

Then one finds that the coefficients of z^n are: $[z^nu^p]\,\frac{\partial^2 S_k(z,u,v)}{\partial v^2}\big|_{v=1} \sim \frac{q_k(p-1)}{\sqrt{\pi}p_k2^{2p+2}}\binom{2p}{p}q_k^{-n}\frac{n^{p-1/2}}{\Gamma(p+\frac12)}$, which will be normalised to produce the second moment.

Theorem 4.14. *The variance for the Steiner distance in monotonic ordered trees is*

$$V_{n,p}^{(k)} \sim \frac{(p-1)q_k}{p_k^2}\Big(\frac{1}{4\pi}-\frac{p^2(p-1)}{2^{4p}(2p-1)^2}\binom{2p}{p}^2\Big)n+\mathcal{O}(\sqrt{n}\,)\,,\ n\to\infty\,,\ \textit{fixed } k\,, \tag{107}$$

where $\frac{1}{4\pi}-\frac{p^2(p-1)}{2^{4p}(2p-1)^2}\binom{2p}{p}^2\to 0$ *as* $p\to\infty$.

References

[1] N. G. De Bruijn. Asymptotic Methods in Analysis. North-Holland, 1972.
[2] P. Flajolet. Analyse d'algorithmes de manipulation d'arbres et des fichiers. Thèse Paris-Sud-Orsay. 1979
[3] P. Flajolet and A. Odlyzko. Singularity analysis of generating functions. *SIAM Journal of Discrete Mathematics*, 3:216–240, 1990.
[4] F. Harary and E. M. Palmer. Graphical Enumeration. Academic Press, 1973.
[5] P. Kirschenhofer. On the average shape of monotonically labelled tree structures. *Discrete Applied Mathematics*, 7:161–181, 1984.
[6] P. Kirschenhofer and H. Prodinger. On the average oscillation of the contour of monotonically labelled ordered trees. *Proceedings of the Second World Conference on Mathematics at the Service of Man (Las Palmas)*, 394–399, 1982.
[7] D. E. Knuth. The Art of Computer Programming, Fundamental Algorithms, Volume 1. Addison-Wesley, 1998.
[8] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canadian Journal of Mathematics*, 5:997–1015, 1978
[9] A. Panholzer. The distribution of the size of the ancestor tree and of the induced spanning subtree for random trees. *Random Structures and Algorithms*, 25:179–207, 2004.
[10] A. Panholzer and H. Prodinger. Spanning tree size in random binary search trees, *Annals of Applied Probability*, 14:718–733, 2004.
[11] H. Prodinger and F. J. Urbanek. On monotone functions of tree structures. *Discrete Applied Mathematics*, 5:223-239, 1983.

Katherine Morris, PO Box 1926, Cresta 2118, Johannesburg, South Africa

E-mail address: katherine.morris@gmail.com

Helmut Prodinger, Mathematics Department, University of Stellenbosch, 7602 Stellenbosch, South Africa

E-mail address: hproding@sun.ac.za