A Bijection for Spanning Trees of Complete Multipartite Graphs

Ömer Eğecioğlu
Department of Computer Science
University of California
Santa Barbara, California 93106

Jeffrey B. Remmel*
Department of Mathematics,
University of California San Diego,
La Jolla, California 92093

Abstract

We construct a bijective proof for the number of spanning trees of complete multipartite graphs. The weight preserving properties of our bijection yields a 6–variate weight generating function which keeps track of various statistics on spanning trees. This bijection allows for the ranking and unranking of the spanning trees of an n-vertex complete multipartite graph in $O(n)$ time. As a further application, we compute the asymptotic distribution of leaves in these families of spanning trees.

Key Words: Spanning tree, multipartite graph, bijection, ranking.

1 Introduction

Let $K_{k_1,k_2,\ldots,k_p}$ denote the complete p-partite graph on vertex set $V = V_1 + V_2 + \cdots + V_p$, where $|V_i| = k_i$ for $i = 1, 2, \ldots, p$, and "+" denotes disjoint union. Put $s_0 = 0$ and define $s_t = k_1 + \cdots + k_t$ for $t = 1, 2, \ldots, p$. We assume that the total number of vertices is n $(= |V| = s_p)$ and the vertex set V_i consists of the integers in the half-open interval $(s_{i-1}, s_i]$. The edges in $K_{k_1,k_2,\ldots,k_p}$ are all pairs $\{i, j\}$ such that there is no t with $1 + s_t \leq i, j \leq s_{t+1}$. Let $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ denote the collection of spanning trees of $K_{k_1,k_2,\ldots,k_p}$. Using the matrix-tree theorem, Onodera [4] showed that

$$|\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})| = n^{p-2} \prod_{t=1}^{p} (n - k_t)^{k_t - 1}. \tag{1}$$

Our first aim is to construct a bijective proof of this formula by suitably interpreting the right hand side of (1) as the enumerator of a certain restricted class of functions mapping $\{2, 3, \ldots, n-1\}$ to $\{1, 2, \ldots, n\}$. The functional diagrams of these functions are then put in one to one correspondence with trees in $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$. This construction relies on a variant of the bijection for Cayley trees (i.e. the spanning trees $\mathcal{SP}_n(K_n)$ of the complete graph K_n) given by the authors in [2]. Indeed, the bijection presented here can be viewed as a generalization of of this bijection from complete graphs to complete multipartite graphs. For a generalization in a different direction, see [3].

*Partially supported by NSF grant No. DMS 87-02473

Our interpretation of the right hand side of (1) will allow us to rank and unrank the trees in $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ optimally in linear time. Furthermore, as is the case with most bijective proofs, we gain extra information about the underlying combinatorial objects by considering the special properties of the bijection constructed. Analogous to the Cayley tree case, our bijection for the spanning trees of complete multipartite graphs has a number of natural weight preserving properties. These allow for the derivation of various q-analogues of Onodera's result. For example, if we put $[0]=0$, and $[m]=1+q+\cdots+q^{m-1}$ for $m>0$, then

$$\sum_{T\in\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})} q^{\sum_i deg_T(i)} = q^{\binom{n+1}{2}-3}[n]^{p-2}\prod_{t=1}^{p}([s_t-1]+q^{s_t}[n-s_t])^{k_t-1},$$

where $deg_T(i)$ denotes the degree of vertex i in a tree T.

Further properties of our bijection allows for the computation of the asymptotic distribution of leaf nodes in $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ as well. More particularly, if the number of parts p is kept fixed and we let $n\to\infty$, then the asymptotic probability that a vertex v in $T\in\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ is a leaf is given by

$$e^{-\sum_{t=1}^{p}\frac{\alpha_t}{1-\alpha_t}}\sum_{i=1}^{p}\alpha_i\, e^{\frac{\alpha_i}{1-\alpha_i}}, \tag{2}$$

where $\alpha_i=\lim_{n\to\infty}\frac{k_i}{n}$.

The outline of this paper is as follows. In Section 2 we reproduce the θ_n bijection for the number of spanning trees of the complete graph K_n, which forms our point of departure. In Section 3, we construct our bijection for complete multipartite graphs. This is followed in Section 4 by weight-generating functions for spanning trees of complete multipartite graphs and various q-analogues of (1). In Section 5 we present ranking and unranking algorithms for $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ and the analysis of their time requirements. Finally, the asymptotic properties of $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ and the proof of (2) appears in Section 6.

2 The θ_n bijection for Cayley trees

For completeness we reproduce here the θ_n bijection for Cayley trees on n nodes that appears in [2], as it forms the basis for the bijection in the general case.

Denote by $\mathcal{C}_n$ the spanning trees $\mathcal{SP}_n(K_n)$, where we imagine each tree as rooted at the largest labeled node n. Furthermore, we orient each edge $\{i,j\}$ of a Cayley tree $T\in\mathcal{C}_n$ by directing it toward the root. Clearly, $|\mathcal{C}_n|=|\mathcal{SP}_n(K_n)|$. Next, let $\mathcal{F}_n$ denote the set of functions from $\{2,3,\ldots,n-1\}$ into $\{1,2,\ldots,n\}$. The bijection θ_n between $\mathcal{C}_n$ and $\mathcal{F}_n$ is most easily described by referring to an explicit example.

Suppose $n=21$ and $f\in\mathcal{F}_{21}$ is given by Table I

Table I

i	$f(i)$	i	$f(i)$	i	$f(i)$	i	$f(i)$
2	5	7	7	12	20	17	16
3	4	8	12	13	19	18	6
4	5	9	1	14	19	19	7
5	3	10	4	15	6	20	12
6	21	11	4	16	1		

We view f as a digraph with vertex set $\{1, 2, \ldots, 21\}$ by putting an edge from i to j if $f(i) = j$. For example, the digraph for f given above is pictured in Figure 1.

Figure 1

A moment's thought will convince one that in general, the digraph corresponding to an $f : \{2, 3, \ldots, n-1\} \to \{1, 2, \ldots, n\}$ will consist of two trees rooted at 1 and n, respectively, with all edges directed toward their roots plus a number of directed cycles of length ≥ 1 where for each vertex v on a given cycle, there is possibly a tree attached to v with v as the root and all edges directed toward v. Note that there are trees rooted at 1 and n due to the fact that 1 and n are not in the domain of f, and consequently there are no directed edges out of 1 or n. Note also that cycles of length one or loops simply correspond to fixed points of f.

As in Figure 1, we imagine the directed graph corresponding to $f \in \mathcal{F}_n$ is drawn so that

1. the trees rooted at 1 and n are drawn on the extreme left and extreme right respectively with their edges directed upwards,

2. the cycles are drawn so that their vertices form a directed path on the line between 1 and n with one back edge above the line and the tree attached to any vertex on a cycle is drawn below the line between 1 and n with edges directed upwards,

3. each cycle is arranged so that its smallest element is on the right and the cycles themselves are ordered from left to right by increasing smallest elements.

Once the directed graph for f is drawn as above, let us refer to the rightmost element in the i-th cycle as r_i and the leftmost element in the i-th cycle as l_i. Thus for the f given above, $l_1 = 4$, $r_1 = 3$, $l_2 = r_2 = 7$, $l_3 = 20$, and $r_3 = 12$. Once an $f \in \mathcal{F}_n$ is drawn in this manner, it is easy to describe the bijection $\theta_n(f)$. That is, if the directed graph of f has k cycles where $k > 0$, we simply eliminate the back edges $r_i \to l_i$ for $i = 1, 2, \ldots, k$ and add the edges $1 \to l_1$, $r_1 \to l_2$, $r_2 \to l_3, \ldots, r_k \to n$. For example, in Figure 1, we eliminate the back edges $3 \to 4$, $7 \to 7$, $12 \to 20$ and add the edges $1 \to 4$, $3 \to 7$, $7 \to 20$, and $12 \to 21$ which are dotted for emphasis. If there are no cycles in the directed graph of f, i.e., $k = 0$, then we simply add the edge $1 \to n$.

Note that it is immediate that θ_n is a bijection between $\mathcal{F}_n$ and $\mathcal{C}_n$ since given any Cayley tree $T \in \mathcal{C}_n$, we can easily recover the directed graph of $f \in \mathcal{F}_n$ such that $\theta_n(f) = T$. The key point here is that by our conventions for the ordering of the cycles of f, it is easy to recover the sequence of nodes $r_1, r_2, \ldots, r_k$ since r_1 is the smallest element on the path between 1 and n, r_2 is the smallest element on the path between r_1 and n, etc., and clearly, knowing $r_1, r_2, \ldots, r_k$ allows us to recover f from T.

Since $\theta_n : \mathcal{F}_n \to \mathcal{C}_n$ is a bijection, we arrive at Cayley's formula $n^{n-2} = |\mathcal{F}_n| = |\mathcal{C}_n| = |\mathcal{SP}_n(K_n)|$.

In the next section, we construct a variant of the bijection θ_n to set up a one-to-one correspondence between the spanning trees of $K_{k_1,k_2,\ldots,k_p}$ and a certain subset of functions in $\mathcal{F}_n$.

3 The bijection Ω_n for complete multipartite graphs

To prove (1) combinatorially, we interpret the right side of (1) in the following manner. First of all let $S_1 = V_1 \setminus \{1\}$, $S_p = V_p \setminus \{n\}$ and for $1 < t < p$ put $S_t = V_t$. Clearly $S_1, S_2, \ldots, S_p$ form a partition of $\{2, 3, \ldots, n-1\}$ with $|S_1| = k_1 - 1$, $|S_p| = k_p - 1$, and for $1 < t < p$, $|S_t| = k_t$. As before, we imagine each tree T in $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ as rooted at its largest labeled vertex n, and direct each edge $\{i, j\}$ in T towards the root. Let $\mathcal{F}_{k_1,k_2,\ldots,k_p}$ denote the set of functions $f \in \mathcal{F}_n$ with the following properties:

1. $f(S_1) \subseteq V_2 \cup V_3 \cup \cdots \cup V_p$, and $f(S_p) \subseteq V_1 \cup V_2 \cup \cdots \cup V_{p-1}$,
2. For $1 < t < p$ and $i \in S_t$, $f(i) \in V_1 \cup \cdots \cup V_{t-1} \cup \{i\} \cup V_{t+1} \cup \cdots \cup V_p$,
3. f has at most one fixed point on each S_t for $1 < t < p$.

Clearly we have

$$|\mathcal{F}_{k_1,k_2,\ldots,k_p}| = \prod_{i=1}^{p} |\mathcal{F}_i| \tag{3}$$

where $\mathcal{F}_1 = \{f : S_1 \to \bigcup_{k=2}^{p} V_k\}$, $\mathcal{F}_p = \{f : S_p \to \bigcup_{k=1}^{p-1} V_k\}$, and for $1 < t < p$, $\mathcal{F}_t$ is the set of functions $f : S_t \to \bigcup_{k=1}^{p} S_k$ satisfying conditions (2) and (3) above. Note that for $1 < t < p$, there are $(n - k_t)^{k_t}$ functions in $\mathcal{F}_t$ with no fixed points and $(n - k_t)^{k_t - 1}$ functions in $\mathcal{F}_t$ with fixed point i for any given $i \in S_t$. Thus for $1 < t < p$,

$$|\mathcal{F}_t| = (n - k_t)^{k_t} + k_t(n - k_t)^{k_t - 1} = n(n - k_t)^{k_t - 1}. \tag{4}$$

Now it is easy to see that $|\mathcal{F}_1| = (n-k_1)^{k_1-1}$ and $|\mathcal{F}_p| = (n-k_p)^{k_p-1}$, so that

$$\begin{aligned} |\mathcal{F}_{k_1,k_2,\ldots,k_p}| &= (n-k_1)^{k_1-1}(n-k_p)^{k_p-1}\prod_{t=2}^{p-1} n(n-k_t)^{k_t-1} \\ &= n^{p-2}\prod_{t=1}^{p}(n-k_t)^{k_t-1}. \end{aligned} \tag{5}$$

To define $\Omega_n(f)$ where $f \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$, we first draw the digraph of f in the manner of the θ_n bijection. Thus the trees rooted at 1 and n are drawn on the extreme left and extreme right respectively, and the cycles of f are arranged with their smallest element r_i on the right with a single back edge, ordered from left to right between 1 and n by increasing r_i. After the digraph of f is drawn in this manner, we further rearrange the cycles which are *fixed points* of f in the following manner. Suppose a cycle C_i is a fixed point of f. Then $r_i = l_i$ of C_i belongs to a set S_j for some j, $1 < j < p$. Note that by the definition of the class of functions $\mathcal{F}_{k_1,k_2,\ldots,k_p}$, f does not have any other fixed points on the set S_j. Let now t be the smallest index such that $r_t \in S_j$ and $r_t < r_i$. We then place C_i immediately before the cycle which has this r_t as its smallest element. Denote by $R(f)$ the functional diagram that results after the cycles corresponding to fixed points of f are rearranged in this manner. The bijection Ω_n is constructed from $R(f)$ exactly as in θ_n bijection. Namely, we connect the cycles in $R(f)$ by adding edges directed from left to right, and then we break the back edge in each cycle. As an example, consider the following function $f \in \mathcal{F}_{3,4,7,7}$ given in Table II:

Table II

i	$f(i)$	i	$f(i)$	i	$f(i)$	i	$f(i)$
2	21	7	13	12	12	17	6
3	11	8	4	13	19	18	1
4	8	9	20	14	3	19	9
5	2	10	6	15	8	20	13
6	6	11	3	16	12		

Here $V_1 = \{1,2,3\}$, $V_2 = \{4,5,6,7\}$, $V_3 = \{8,9,10,11,12,13,14\}$, and $V_4 = \{15,16,17,18,19,20,21\}$, with $S_1 = V_1 \setminus \{1\}$, $S_2 = V_2$, $S_3 = V_3$, $S_4 = V_4 \setminus \{21\}$. The numbers s_i are given by $s_1 = 3$, $s_2 = 7$, $s_3 = 14$, and $s_4 = 21$. When we order the cycles of f in the manner of the θ_n bijection, we obtain the digraph in Figure 2.

Figure 2

Next, rearranging the positions of the two fixed points of f results in $R(f)$ and $\Omega_{21}(f)$ depicted in Figure 3.

Figure 3

In this case $\Omega_{21}(f)$ is the spanning tree T of $K_{3,4,7,7}$ pictured in Figure 4.

To see that $\Omega_n(f) \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$, first observe that our definition of $\mathcal{F}_{k_1,k_2,\ldots,k_p}$ ensures that the only edges $i \to j$ in the digraph of f where both i and j lie in some V_t is if $i = j$ so that the edge $i \to j$ is a back edge associated to some fixed point of $R(f)$. Assume $R(f)$ has k cycles and let r_i and l_i denote the right and left hand endpoints of the i-th cycle respectively. Since all back edges of $R(f)$ are eliminated in $\Omega_n(f)$, it follows that the only edges $i \to j$ of $\Omega_n(f)$ which could be such that both i and j are in some V_t are among the newly added edges $1 \to l_1$, $r_1 \to l_2, \ldots, r_k \to n$. Of course if there are no cycles, i.e. $k = 0$, then we simply add the edge $1 \to n$ in which case we automatically have that $\Omega_n(f) \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$. Otherwise, consider r_1. If $r_1 \in S_1$, it follows that $l_1 \notin V_1$ since $f(S_1) \subseteq \bigcup_{t=2}^{p} V_t$. If $r_1 \in S_t$, where $t > 1$, then since r_1 is the smallest element in its cycle, it must be the case that $l_1 \in V_t \cup \cdots \cup V_p$. Thus in either case $l_1 \notin V_1$, and the edge $1 \to l_1$ does not have both of its end points in some V_t. Next consider r_k. It cannot be the case that $r_k \in S_p$ because $f(S_p) \subseteq \bigcup_{t=1}^{p-1} V_t$ and hence $l_k \in \bigcup_{t=1}^{p-1} V_t$. But then $l_k < r_k$, violating the fact that r_k is the smallest element in its cycle. Thus $r_k \notin V_p$ and the edge $r_k \to n$ does not have both of its end points in some V_t. Finally, consider two consecutive cycles in $R(f)$ with end points l_i, r_i and l_{i+1}, r_{i+1}. We shall show that $r_i \in S_u$ and $l_{i+1} \in S_v$ where $u < v$ so that the edge $r_i \to l_{i+1}$ in $\Omega_n(f)$ does not connect two points in some V_t. There are four cases to consider:

(i) *Neither cycle is a fixed point of f:* In this case $r_i < r_{i+1}$ so that $r_{i+1} \in S_w$ where $w \geq u$. Because $l_{i+1} \neq r_{i+1}$, it follows that $l_{i+1} \notin S_w$. But r_{i+1} is the smallest element in its cycle, so we must have that $l_{i+1} \in S_v$ where $v > w \geq u$. Thus $u \neq v$.

(ii) *Both cycles are fixed points of f:* Since f has at most one fixed point on any one of the sets S_j, it follows that r_i and r_{i+1} belong to S_u and S_v, respectively, with $u < v$.

Figure 4

(iii) *The first cycle is a fixed point of f:* Note that either $r_i < r_{i+1}$ or if $r_i > r_{i+1}$, then it must be the case that r_{i+1} is the least right hand endpoint of a cycle with $r_i, r_{i+1} \in S_u$. In either case, we can conclude that $r_{i+1} \in S_w$ where $v \leq w$. Then just as in case (i), we can argue that $l_{i+1} \in S_v$ where $v > w \geq u$.

(iv) *The second cycle is a fixed point of f:* Note that in the construction of $R(f)$, r_{i+1} is placed preceding the cycle with the smallest r_t with $r_t, r_i \in S_j$ and $r_t < r_i$. Therefore after the rearrangement r_i and r_{i+1} cannot belong to the same S_j. Thus if $r_i \in S_u$, it follows that $l_{i+1} = r_{i+1} \in S_v$ where $v > u$.

Thus we have shown that $\Omega_n(f) \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ for all $f \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$. Let us make one more observation about the map Ω_n. Let $c_1 < c_2 < \cdots < c_k$ denote the right hand endpoints of the cycles of the digraph of f before we move the fixed points to produce $R(f)$. Suppose that c_j is the first fixed point among $c_1, \ldots, c_k$ and i is the least index $m \leq j$ such that c_m and c_j lie in some S_t. Now if $i < j$, then keeping our notation above we will have $r_t = c_t$ for $t < i$, $r_i = c_j$, $r_t = c_{t-1}$ for $i < t \leq j$. Then we can recover $r_1, r_2, \ldots, r_j$ from $\Omega_n(f)$ as follows. Just as in the θ_n bijection, $r_1 = c_1$ is the least element on the path from 1 to n, $r_2 = c_2$ is the least element on the path from r_1 to n, ... , $r_{i-1} = c_{i-1}$ is the least element on the path from r_{i-2} to n. Now if we consider the least element on the path from r_{i-1} to n, this element is $r_{i+1} = c_i$. However, when we try to recover the cycle starting with r_{i+1} as in the θ_n bijection, we would try to draw the back edge $r_{i+1} \rightarrow r_i$. Of course, we would then recognize that the edge $r_{i+1} \rightarrow r_i$ cannot be an edge in the digraph of f because the only edges in the digraph of f which have both endpoints in some V_t are loops. Thus we know that r_i must be a fixed point of f and the back edge from r_{i+1} should go to the element immediately following r_i on the path from 1 to n. Then $r_{i+2} = c_{i+1}$ is the least element on the path from r_{i+1} to n, ... , $r_j = c_{j-1}$ is the least element on the path from r_{j-1} to n. Finally we observe that all the elements on the path from r_j to n are greater than r_i. Of course in the case where $j = i$, so that we did not need to move the fixed point c_j, we can recover $r_1, \ldots, r_j$ just as in the θ_n bijection. By using the same procedure on the elements which lie on the path from r_j to n, we can recover all the cycles of the digraph of f up to the next fixed point of f, etc..

It follows that given any spanning tree $T \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$, we can recover the digraph of $f_T \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$ such that $\Omega_n(f_T) = T$. To see this, consider the sequence of nodes $u_1, u_2, \ldots, u_k$ where u_1 is the smallest node on the path from 1 to n in T, u_2 is the smallest element on the path from u_1 to n, etc., exactly as in the θ_n bijection. Consider the left hand endpoints $v_1, v_2, \ldots, v_k$ determined by this sequence of nodes $u_1 < u_2 < \cdots < u_k$. For example if we start with the tree T pictured in Figure 3, $u_1 = 3$, $v_1 = 11$, $u_2 = 4$, $v_2 = 6$, $u_3 = 9$, and $v_3 = 12$. We then eliminate the edges $1 \to v_1$, $u_1 \to v_2, \ldots, u_{n-1} \to v_n$, $u_n \to n$ and attempt to draw the back edges $u_i \to v_i$ to complete the cycles of f. If u_i and v_i are in different parts of the partition $S_1 + S_2 + \cdots + S_p$, we keep the cycle. If for some t, $u_i, v_i \in S_t$, we declare v_i to be a fixed point of f_T and let w_i be the element which follows v_i on the path from 1 to n. Note that since $v_i \to w_i$ in T, we must have $w_i \notin S_t$. We then eliminate the edge $v_i \to w_i$ and draw the back edges $v_i \to v_i$ and $u_i \to w_i$ to give two cycles. We claim that this procedure always produces the digraph of a function $f_T \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$. Clearly, there is no difficulty with the edges which lie both in T and the digraph of f_T. The only problem can come from the back edges where we must show that there are no fixed points of f_T in S_1 or S_p and that there is at most one fixed point of f_T in S_t for $1 < t < p$. First we claim that there is no fixed point of f_T in S_1. That is, suppose v_i is a fixed point of f_T and $v_i \in S_1$. Then since $u_i \le v_i$, we must have $u_i \in S_1$. But then either $i = 1$ and the edge $1 \to v_1$ is in T, or $i > 1$ and the edge $u_{i-1} \to v_i$ is in T. In the latter case, $u_{i-1} < u_i$ and $u_i \in S_1$ implies $u_{i-1} \in S_1$. Thus in either case, we would get an edge in T connecting two points of V_1 which is impossible. Similarly, suppose v_i is a fixed point of f_T and $v_i \in S_p$. But then $u_i \in S_p$ and since $u_i < u_{i+1} < \cdots < u_k$, we would have $u_k \in S_p$. This is impossible because then the edge $u_k \to n$ would be in T and would connect two points in V_p. Finally, suppose there are indices $i < j$ where v_i and v_j are fixed points of f_T and $v_i, v_j \in S_t$ for some t. Thus it must be the case that u_i and u_j are in S_t. But then $u_i \le u_{j-1} < u_j$ so that $u_{j-1} \in S_t$. This is impossible because then $u_{j-1} \to v_j$ would be an edge in T connecting two vertices in V_t. Thus $f_T \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$ whenever $T \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$. Thus we have shown that Ω_n is a bijection between $\mathcal{F}_{k_1,k_2,\ldots,k_p}$ and $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$. Hence by (5) we have

$$|\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})| = n^{p-2} \prod_{t=1}^{p} (n - k_t)^{k_t - 1}.$$

4 Statistics on spanning trees

Consider $T \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$, where as before we consider T as rooted at its largest labeled node n, and direct each edge back toward the root. We call a directed edge

$i \to j$ a *rise* if $i < j$ and a *fall* if $i > j$. We assign a monomial weight

$$\omega(i \to j) = \begin{cases} xq^i t^j & \text{if } i > j, \\ yp^i s^j & \text{if } i < j. \end{cases} \tag{6}$$

We then define the weight of $T = (V, E) \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ by setting

$$\omega(T) = \prod_{e \in E} \omega(e). \tag{7}$$

For example, if T is the tree pictured in Figure 5, then the weight of the edge $7 \to 3$ is xq^7t^3, the weight of the edge $5 \to 8$ is yp^5s^8, and the weight of T itself is

$$\begin{aligned} \omega(T) &= (yps^4)(yp^2s^4)(yp^3s^8)(yp^4s^7)(yp^5s^8)(xq^6t^3)(xq^7t^3) \\ &= x^2y^5p^{15}s^{31}q^{13}t^6. \end{aligned}$$

Similarly, for $f \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$, define

$$\omega(f) = \prod_{i=2}^{n-1} \omega(f, i),$$

where

$$\omega(f, i) = \begin{cases} xq^i t^j & \text{if } f(i) = j \text{ and } i > j, \\ yp^i s^j & \text{if } f(i) = j \text{ and } i \le j. \end{cases} \tag{8}$$

Consider the weight generating functions

$$\mathbf{G}\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p}) = \sum_{T \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})} \omega(T), \tag{9}$$

$$\mathbf{G}(\mathcal{F}_{k_1,k_2,\ldots,k_p}) = \sum_{T \in \mathcal{F}_{k_1,k_2,\ldots,k_p}} \omega(f). \tag{10}$$

It is easy to see that

$$\mathbf{G}(\mathcal{F}_{k_1,k_2,\ldots,k_p}) = \mathbf{G}(\mathcal{F}_1) \times \mathbf{G}(\mathcal{F}_p) \times \prod_{t=2}^{p-1} \mathbf{G}(\mathcal{F}_t),$$

in which

$$\mathbf{G}(\mathcal{F}_i) = \sum_f \omega(f),$$

where the sum is over all functions f in $\mathcal{F}_i$ and $\mathcal{F}_1, \ldots, \mathcal{F}_p$ are defined as in Section 3. We find that

$$\mathbf{G}(\mathcal{F}_1) = \prod_{i=2}^{s_1} (yp^i(s^{1+s_1} + \cdots + s^n)), \tag{11}$$

$$\mathbf{G}(\mathcal{F}_p) = \prod_{i=1+s_{p-1}}^{n-1} (xq^i(t + \cdots + t^{s_{p-1}})), \tag{12}$$

and for $t = 2, \ldots, p-1$,

$$\begin{aligned} \mathbf{G}(\mathcal{F}_t) &= \prod_{i=1+s_{t-1}}^{s_t} (yp^i(s^{1+s_t} + \cdots + s^n) + xq^i(t + \cdots + t^{s_{t-1}})) + \\ &\quad + \sum_{i=1+s_{t-1}}^{s_t} yp^i s^i \prod_{\substack{j=1+s_{t-1} \\ j \neq i}}^{s_t} (yp^j(s^{1+s_t} + \cdots + s^n) + xq^j(t + \cdots + t^{s_{t-1}})). \end{aligned} \tag{13}$$

We have

Theorem 1 $\mathbf{G}\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p}) = yps^n \mathbf{G}(\mathcal{F}_{k_1,k_2,\ldots,k_p})$.

Proof We shall prove the theorem by showing that

$$\omega(\Omega_n(f)) = yps^n \omega(f) \tag{14}$$

for each $f \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$. To prove (14), note that our definitions ensure that if $f(i) = j$ and $i \to j$ remains a directed edge in both the directed graph of f and the directed graph of $T = \Omega(f)$, then $\omega(f,i) = \omega(i \to j)$. Thus in the case where the directed graph of f has no cycles, (14) is clear since in this case T is obtained from f by adding the edge $1 \to n$ to the digraph of f, and the contribution of this edge to the weight is yps^n. If the directed graph $R(f)$ which is obtained from the digraph of f after ordering the cycles and reordering the fixed points of f according to the definition of Ω has k cycles with $k > 0$, then we follow our conventions from Section 3 and let l_i and r_i denote the left and right hand points of the i-th cycle in the digraph of f. Note that the only difference between the weights of f and T are due to the difference between the weights of the edges $r_1 \to l_1, \ldots, r_k \to l_k$ which are deleted from the graph of f, and the weights of the new set of edges $S = \{1 \to l_1, r_1 \to l_2, \ldots, r_{k-1} \to l_k, r_k \to n\}$ added to the resulting digraph. Since r_i is the smallest element in the i-th cycle of f, we know that $l_i = f(r_i) \geq r_i$ for $i = 1, \ldots, k$. Thus we must have $\omega(f, r_i) = yp^{r_i}s^{l_i}$. It follows that

$$\begin{aligned} \omega(f) &= (yp^{r_1}s^{l_1})(yp^{r_2}s^{l_2}) \cdots (yp^{r_k}s^{l_k}) \prod_{i \notin \{r_1,\ldots,r_k\}} \omega(f,i) \\ &= y^k \, p^{\sum_i r_i} \, s^{\sum_i l_i} \prod_{i \notin \{r_1,\ldots,r_k\}} \omega(f,i). \end{aligned} \tag{15}$$

Now if $T = (V, E)$ then

$$\prod_{i \notin \{r_1,\ldots,r_k\}} \omega(f,i) = \prod_{i \to j \notin E \setminus S} \omega(i \to j),$$

since if $f(i) = j$ and $i \notin \{r_1, \ldots, r_k\}$, then $i \to j$ is an edge in both the directed graph of f and the directed graph of T. We claim that each of the edges in S are rise edges. It is clear that $1 \to l_1$ and $r_k \to n$ are rise edges. In Section 3, we proved that if $r_i \in S_u$ then $l_{i+1} \in S_v$ where $u < v$. Thus $r_i < l_{i+1}$ and $r_i \to l_{i+1}$ is a rise edge for every edge in S. Thus

$$\omega(T) = \prod_{i \to j \in S} \omega(i \to j) \prod_{i \to j \notin E \setminus S} \omega(i \to j)$$

$$
\begin{aligned}
&= (yps^{l_1})(yp^{r_1}s^{l_1})\cdots(yp^{r_{k-1}}s^{l_k})(yp^{r_k}s^n)\prod_{i\to j\notin E\setminus S}\omega(i\to j)\\
&= (yps^n)y^k\, p^{\sum_i r_i}\, s^{\sum_i l_i}\prod_{i\notin\{r_1,\ldots,r_k\}}\omega(f,i).
\end{aligned}
$$

Thus by (15)

$$\omega(T)=yps^n\omega(f),$$

and the Theorem follows. □

Due to the definition of the family of functions $\mathcal{F}_{k_1,k_2,\ldots,k_p}$, the weight generating function $\mathbf{G}\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ is not in a particularly simple form. However for certain interesting statistics on trees, $\mathbf{G}\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ specializes to a much nicer product form. We give an example: For $T\in\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$, let $\delta(T)=\sum_{i\in T} id_T(i)$, where $d_T(i)$ is the degree of the vertex i in T. Now if we set $x=y=1$ and $p,s,$ and t equal to q in the monomial weight of T, then each vertex i will contribute a factor of q^i to the resulting weight of T every time vertex i is either a right or left endpoint of a directed edge in T. Hence the resulting weight of T with those substitutions will be precisely $q^{\sum_i id_T(i)}=q^{\delta(T)}$. Note that the δ weight is independent of the fact that we regard T as rooted at vertex n. As a corollary of Theorem 1 we obtain the following q-analogue of (1):

Corollary 1

$$\sum_{T\in\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})} q^{\delta(T)} = q^{\binom{n+1}{2}-3}[n]^{p-2}\prod_{t=1}^{p}([s_{t-1}]+q^{s_t}[n-s_t])^{k_t-1},$$

where $s_t=k_1+\cdots+k_t$, $[0]=0$, *and* $[m]=1+q+..+q^{m-1}$ *for* $m>0$.

Proof It is easy to see that when we set $x=y=1$ and $p,s,$ and t equal to q in the monomial weight of T, then the partial weight generating functions $\mathbf{G}(\mathcal{F}_1)$, and $\mathbf{G}(\mathcal{F}_p)$ in (11) and (12) specialize to

$$q^{e_1}(q^{s_1}[n-s_1])^{k_1-1},\quad \text{and}\quad q^{e_p}[s_{p-1}]^{k_p-1},$$

respectively, where

$$e_1=\sum_{i=2}^{s_1}(i+1),\quad e_p=\sum_{i=1+s_{p-1}}^{n-1}(i+1).$$

For $t=2,\ldots,p-1$, put

$$e_t=\sum_{i=1+s_{t-1}}^{s_t}(i+1).$$

For these values of t, $\mathbf{G}(\mathcal{F}_t)$ given in (13) specializes to

$$
\begin{aligned}
&q^{e_t}([s_{t-1}]+q^{s_t}[n-s_t])^{k_t}+\sum_{i=1+s_{t-1}}^{s_t}q^{e_t+i-1}([s_{t-1}]+q^{s_t}[n-s_t])^{k_t-1}\\
&= q^{e_t}([s_{t-1}]+q^{s_t}[n-s_t]+q^{s_{t-1}}[k_t])([s_{t-1}]+q^{s_t}[n-s_t])^{k_t-1}\\
&= q^{e_t}[n]([s_{t-1}]+q^{s_t}[n-s_t])^{k_t-1}.
\end{aligned}
$$

Thus we have

$$\sum_{T\in\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})} q^{\delta(T)} = q^{e_1+e_2+\cdots+e_p}[n]^{p-2}\prod_{t=1}^{p}([s_{t-1}]+q^{s_t}[n-s_t])^{k_t-1}.$$

It can be easily verified that $e_1+e_2+\cdots+e_p = \sum_{i=2}^{n-1}(i+1) = \binom{n+1}{2}-3$ as claimed. □

We end this section with a brief outline of how we can also obtain a weight generating function similar to (10) for spanning trees of $K_{k_1,k_2,\ldots,k_p}$ which are rooted at vertex 1 instead of vertex n. That is, root each $T\in\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ at vertex 1 and direct each edge back toward the root. Define the weight of directed edge $\omega(i\to j)$ and the weight of a tree $T\in\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ by (6) and (7), respectively. Then define

$$\overline{\mathbf{G}\mathcal{SP}}_n(K_{k_1,k_2,\ldots,k_p}) = \sum_{\substack{T\in\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})\\ T\ rooted\ at\ 1}} \omega(T).$$

Next, define the weight of a function $f\in\mathcal{F}_{k_1,k_2,\ldots,k_p}$ by

$$\overline{\omega}(f) = \prod_{i=2}^{n-1}\overline{\omega}(f,i), \tag{16}$$

where

$$\overline{\omega}(f,i) = \begin{cases} xq^it^j & \text{if } f(i)=j \text{ and } i\geq j,\\ yp^is^j & \text{if } f(i)=j \text{ and } i<j.\end{cases}$$

Note that the only difference between (8) and (16) is the weight of the fixed points of f. Then let

$$\overline{\mathbf{G}}(\mathcal{F}_{k_1,k_2,\ldots,k_p}) = \sum_{T\in\mathcal{F}_{k_1,k_2,\ldots,k_p}} \overline{\omega}(f).$$

Again it is easy to see that

$$\overline{\mathbf{G}}(\mathcal{F}_{k_1,k_2,\ldots,k_p}) = \overline{\mathbf{G}}(\mathcal{F}_1)\times\overline{\mathbf{G}}(\mathcal{F}_p)\times\prod_{t=2}^{p-1}\overline{\mathbf{G}}(\mathcal{F}_t),$$

in which

$$\overline{\mathbf{G}}(\mathcal{F}_i) = \sum_f \overline{\omega}(f),$$

and $\mathcal{F}_i$ are defined as in Section 3. Then it is easy to check that $\overline{\mathbf{G}}(\mathcal{F}_1) = \mathbf{G}(\mathcal{F}_1)$, $\overline{\mathbf{G}}(\mathcal{F}_p) = \mathbf{G}(\mathcal{F}_p)$, and for $t=2,\ldots,p-1$

$$\begin{aligned}\overline{\mathbf{G}}(\mathcal{F}_t) &= \prod_{i=1+s_{t-1}}^{s_t}(yp^i(s^{1+s_t}+\cdots+s^n)+xq^i(t+\cdots+t^{s_{t-1}}))+\\ &+\sum_{i=1+s_{t-1}}^{s_t} xq^it^i\prod_{\substack{j=1+s_{t-1}\\ j\neq i}}^{s_t}(yp^j(s^{1+s_t}+\cdots+s^n)+xq^j(t+\cdots+t^{s_{t-1}})). \end{aligned}\tag{17}$$

Next we shall describe how we can modify the Ω_n bijection to produce a bijection $\overline{\Omega}_n : \mathcal{F}_{k_1,k_2,\ldots,k_p} \rightarrow \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ where for each $f \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$, $\overline{\Omega}_n(f)$ is a spanning tree rooted at 1 such that

$$xtq^n \,\overline{\omega}(f) = \omega(\,\overline{\Omega}_n(f)).$$

Thus $\overline{\Omega}_n$ will show that

$$\overline{\mathbf{G}\mathcal{SP}}_n(K_{k_1,k_2,\ldots,k_p}) = xtq^n \,\overline{\mathbf{G}}(\mathcal{F}_{k_1,k_2,\ldots,k_p}).$$

Now given an $f \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$, we draw the digraph of f in much the same way as in the first step of the Ω_n bijection except that we

1. put the tree rooted at 1 on the extreme right,
2. put the tree rooted at n on the extreme left,
3. draw the cycles so that the largest element is on the right,
4. and finally order the cycles from left to right by decreasing largest elements.

For example, the digraph of the function f in Figure 2 would be drawn as in Figure 6.

Figure 6

Next we rearrange the cycles corresponding to the fixed points of f. Suppose the i-th cycle C_i is a fixed point of f on S_t. Let C_j be the first cycle preceding C_i whose right hand endpoint is in S_t. If there is no such index j, then we do not move C_i. Otherwise we place C_i immediately before C_j. For example, for the function of Figure 6 where $V_1 = \{1,2,3\}$, $V_2 = \{4,5,6,7,8,9\}$, $V_3 = \{10,11,12,13,14\}$, and $V_4 = \{15,16,17,18,19,20,21\}$, the fixed point 6 is moved.

Now let $\overline{R}(f)$ denote the digraph of f after we have rearranged the cycles in the manner described above. Let r_i and l_i denote the right and left hand points of the $i-th$ cycle of $\overline{R}(f)$ reading from left to right. Then we obtain $\overline{\Omega}_n(f)$ from $\overline{R}(f)$ just as before, i.e., we eliminate the back edges $r_i \rightarrow l_i$ for $k = 1,\ldots,k$ where k is the number of cycles of $\overline{R}(f)$ and add the edges $n \rightarrow l_1$, $r_1 \rightarrow l_2, \ldots, r_{k-1} \rightarrow l_k$, and $r_k \rightarrow 1$. If there are no cycles, we just add the edge $n \rightarrow 1$. See Figure 7, for example.

The weight preserving properties of the $\overline{\Omega}_n$ bijection and the fact that $\overline{\Omega}_n \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ follow from an analysis very similar to the one given in Section 3 which shows that

1. $l_1 \notin S_p$,
2. $r_k \notin S_1$,

Figure 7

3. for each $i = 1, \ldots, k-1$, $r_i \in S_{u_i}$ and $l_{i+1} \in S_{v_{i+1}}$ where $u_i > v_{i+1}$.

Thus in particular, all the edges $n \to l_1$, $r_1 \to l_2, \ldots, r_{k-1} \to l_k$, and $r_k \to 1$ are falls.

Given a tree $T \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ rooted at 1, we can show that we can recover the function $f_T \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$ such that $\overline{\Omega}_n(f_T) = T$ as follows. Consider the path from n to 1 in T. Let u_1 be the largest element on the path from n to 1, u_2 be the largest element on the path from u_1 to 1, u_3 be the largest element on the path from u_2 to 1, etc. We use $u_1 > u_2 > \cdots > u_k$ to determine the cycles of f_T just as we do in reversing the Ω_n bijection. In other words, let $v_1, v_2, \ldots, v_k$ be the left hand points of the cycles determined by $u_1, u_2, \ldots, u_k$ respectively. Then we eliminate the edges $n \to v_1$, $u_1 \to v_2, \ldots, u_{k-1} \to v_k, u_k \to 1$. If u_i and v_i are in different parts of the partition $S_1 \cup \cdots \cup S_p$, we add the back edge $u_i \to v_i$. Otherwise we let w_i be the element immediately following v_i on the path from n to 1. We then eliminate the edge $v_i \to w_i$ and form two cycles by adding the back edges $v_i \to v_i$ and $u_i \to w_i$.

5 Ranking and unranking $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$

In a number of settings it is required to generate random combinatorial structures (k-subsets of an n-set, permutations, partitions, compositions, trees, planar graphs, Hamiltonian cycles, etc.), or random objects from a subclass of such an underlying family having a particular property, usually drawn from a uniform distribution. Efficient ranking is one of the obvious ways of achieving this. A collection of ranking and unranking algorithms for combinatorial structures of a diverse nature can be found in Nijenhaus and Wilf [6], and Reingold, Nievergelt, and Deo [7].

Colbourn, Day, and Nel [1] provided an $O(n^3)$ ranking and unranking algorithm for spanning trees of an arbitrary n-vertex graph G. This makes it possible to generate a random spanning tree of a given connected n-vertex graph in time $O(n^3)$. The bijection

Ω_n allows us to rank and unrank spanning trees of $K_{k_1,k_2,\ldots,k_p}$ in linear time by ranking and unranking the functions $\mathcal{F}_{k_1,k_2,\ldots,k_p}$.

5.1 The procedures $UNRANK(r)$ and $RANK(T)$

Given r with $0 \le r < |\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})|$, we construct in stages, an $f \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$. We first determine the values of f on S_1 and S_p as follows. Let

$$\begin{aligned} r &= q_1(n-k_1)^{k_1-1} + r_1, \\ q_1 &= q_p(n-k_p)^{k_p-1} + r_p, \end{aligned} \tag{18}$$

with $0 \le r_1 < (n-k_1)^{k_1-1}$, and $0 \le r_p < (n-k_p)^{k_p-1}$. Base $(n-k_1)$ expansion of r_1

$$r_1 = a_0 + a_1(n-k_1) + \cdots + a_{k_1-2}(n-k_1)^{k_1-2} \tag{19}$$

defines a partial function f in $\mathcal{F}_{k_1,k_2,\ldots,k_p}$, mapping S_1 to $S_2 \cup \cdots \cup S_p$, by suitably shifting the digits a_i of r_1 in (19). More precisely, we let $f(i) = k_1 + 1 + a_{i-2}$ for $i \in S_1$. Similarly, base $n - k_p$ expansion of r_p

$$r_p = b_0 + b_1(n-k_p) + \cdots + b_{k_p-2}(n-k_p)^{k_p-2} \tag{20}$$

defines f on S_p by the recipe $f(i) = 1 + b_{i-n+k_p-1}$. Thus $1 + b_0$ is the image of the smallest element in S_p under f, $1 + b_1$ the image of the second smallest, and so on.

Next, we define f on the sets S_t, for $1 < t < p$. Before doing this however, we first determine two integers Q and R from q_p by

$$q_p = Qn^{p-2} + R \tag{21}$$

where q_p is as found in (18) and $0 \le R < n^{p-2}$. If we now consider the base n expansion of the remainder R in (21),

$$R = n_2 + n_3 n + \cdots + n_{p-1} n^{p-3}, \tag{22}$$

we obtain $n-2$ integers n_t, $0 \le n_t < n$. We will use these numbers to interpret the factor n that appears under the product sign in (4) in deciding whether or not f should have a fixed point on S_t, $1 < t < p$. More precisely, there are two cases to consider. If $n_t \in \{0, 1, \ldots, k_t - 1\}$, we interpret this to mean that $(1 + n_t)$-th smallest element in S_t shall be a fixed point of f. Otherwise n_t takes on one of the $n - k_t$ values in the set $\{k_t, k_t + 1, \ldots, n-1\}$. In this case, we consider the unique order preserving bijection between $\{k_t, k_t+1, \ldots, n-1\}$ and the $n - k_t$ integers $S_1 \cup \cdots \cup S_{t-1} \cup S_{t+1} \cup \cdots \cup S_p$, i.e.,

1	2	.	.	s_{t-1}	s_t+1	.	.	n
k_t	k_t+1	.	.	s_t-1	s_t	.	.	$n-1$

The value of n_t is then used to define the image of the function f on the *smallest* element in S_t via this bijection. After this phase of the procedure, for every S_t with $1 < t < p$, either the unique fixed point of f on S_t, or the value of f on the smallest element in S_t is determined.

Next we need to define f on the remaining $k_t - 1$ elements of the sets S_t for $1 < t < p$. We do this for $t = 2, 3, \ldots, p-1$, in that order. To define f on S_2, consider the remainder r_2 in base $n - k_2$ expansion of the quotient Q obtained above

$$Q = q_2(n - k_2)^{k_2 - 1} + r_2, \tag{23}$$

with $0 \leq r_2 < (n - k_2)^{k_2 - 1}$. Now assume

$$r_2 = c_0 + c_1(n - k_2) + \cdots + c_{k_2 - 2}(n - k_2)^{k_2 - 2}. \tag{24}$$

First the digits $c_0, c_1, \ldots, c_{k_2 - 2}$ are assigned to the $k_2 - 1$ elements for which f has not yet been defined in S_2, from left to right in increasing order. It is easy to see that after this by a suitable translation, each c_i can be used to define the corresponding value of f via the unique order preserving bijection between $S_1 \cup S_3 \cup \cdots \cup S_p$ and $\{0, 1, \ldots, n - k_2 - 1\}$.

Now to define f on the remaining $k_3 - 1$ elements of S_3, we consider the base $n - k_3$ expansion of the remainder r_3 in

$$q_2 = q_3(n - k_3)^{k_3 - 1} + r_3, \tag{25}$$

and so on. Note that $q_{p-1} = 0$ and $q_{p-2} = r_{p-1}$.

After the function f corresponding to the given r is constructed in this manner, we set $UNRANK(r) = \Omega_n(f)$. Similarly, for a given $T \in \mathcal{SP}_n(K_{k_1, k_2, \ldots, k_p})$, we compute $RANK(T)$ by first constructing $f = \Omega_n^{-1}(T)$, and then reversing our steps above.

5.2 Analysis of $UNRANK(r)$ and $RANK(T)$

Now we consider the number of operations required for the procedures $UNRANK$ and $RANK$. Here we represent $T \in \mathcal{SP}_n(K_{k_1, k_2, \ldots, k_p})$ as an array $T[1], T[2], \ldots, T[n-1]$ where $T[i] = j$ iff the edge $\{i, j\}$ is oriented from vertex i to vertex j, when we consider each edge of T as oriented towards the root n. Similarly, an $f \in \mathcal{F}_n$ will be represented as an array of values $f[2], f[3], \cdots, f[n-1]$ of length $n - 2$. It is not difficult to see that with these representations of trees and functions, the computation of $\Omega_n(f)$ and $\Omega_n^{-1}(T)$ require only $O(n)$ operations.

For procedure $UNRANK$, we first need to compute $(n - k_i)^{k_i - 1}$ for $i = 1, 2, \ldots, p$, and also n^{p-2}. This requires a total of $O(\log p + \sum_i \log k_i) = O(n)$ arithmetic operations. Note that this computation is preprocessing, and is needed to be performed only once for $k_1, k_2, \ldots, k_p$ fixed.

Next, the computation of r_i and its base $n - k_i$ expansion requires k_i operations for $i = 1, 2, \ldots, p$. The computation of Q and R can be performed with $p + 1$ arithmetic operations. Once the expansions of the various r_i are known, $f[2], f[3], \ldots, f[n-1]$ can be found in time proportional to n. Thus the total time to compute f from r is $O(k_1 + k_2 + \cdots + k_p) = O(n)$. The application of Ω_n to f requires an additional $O(n)$ steps. Thus we conclude that with $O(n)$ preprocessing cost, each $UNRANK$ operation requires linear time to complete.

In computing $RANK(T)$ from the array representation of T, we first find the corresponding function $f = \Omega_n^{-1}(T)$ in $O(n)$ operations. By Horner's rule, each r_i can be computed with $O(k_i)$ arithmetic steps. Similarly, the computation of R and Q will require $O(p)$ arithmetic operations. Thus the computation of $RANK(T)$ requires $O(n)$ time as well.

In particular, if $R(n)$ denotes the optimal number of operations required to generate a random integer r in the range $0 \le r < |\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})|$, then $UNRANK(r)$ generates a random spanning tree of $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ optimally in $O(R(n)+n)$ time.

6 Asymptotic distribution of leaves in $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$

It easily follows from the Prüfer bijection [5] that the asymptotic probability that a vertex is a leaf (i.e., has degree one) in a Cayley tree $T \in \mathcal{SP}_n(K_n)$ is e^{-1}, where e is the base of natural logarithms. In this section, as another application of the bijection Ω, we compute the asymptotic distribution of leaves in $\mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ where we keep the number of parts p fixed and let n tend to infinity.

It is easy to see that a vertex v is a leaf in $T = \Omega(f)$ if and only if v has no preimage under $f \in \mathcal{F}_{k_1,k_2,\ldots,k_p}$. Let $\mathcal{F}^v_{k_1,k_2,\ldots,k_p}$ denote the collection of functions in $\mathcal{F}_{k_1,k_2,\ldots,k_p}$ in which v has no preimage. By a straightforward counting argument using the definition (3) we obtain

Lemma 1 *If $v \in V_t$, $t \in \{1,2,\ldots,p\}$, then*

$$|\mathcal{F}^v_{k_1,k_2,\ldots,k_p}| = (n-1)^{p-2}\frac{(n-k_t)^{k_t-1}}{(n-k_t-1)^{k_t-1}}\prod_{i=1}^{p}(n-k_i-1)^{k_i-1}.$$

Now assume that $\lim_{n\to\infty}\frac{k_i}{n} = \alpha_i$, for $i = 1,\ldots,p$. Thus $\alpha_1+\alpha_2+\cdots+\alpha_p = 1$.

Theorem 2 *The asymptotic probability that a vertex v in $T \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ is a leaf is given by*

$$e^{-\sum_{t=1}^{p}\frac{\alpha_t}{1-\alpha_t}}\sum_{i=1}^{p}\alpha_i\, e^{\frac{\alpha_i}{1-\alpha_i}}. \tag{26}$$

Proof Given that $v \in V_t$, by Lemma 1, the probability that v is a leaf is

$$\lim_{n\to\infty}\frac{|\mathcal{F}^v_{k_1,k_2,\ldots,k_p}|}{|\mathcal{F}_{k_1,k_2,\ldots,k_p}|} = \lim_{n\to\infty}(1-\frac{1}{n})^{p-2}\lim_{n\to\infty}\frac{\prod_{\substack{i=1\\ i\neq t}}^{p}(n-k_i-1)^{k_i-1}}{\prod_{\substack{i=1\\ i\neq t}}^{p}(n-k_i)^{k_i-1}}. \tag{27}$$

Using the fact that $\lim_{n\to\infty}(1+\frac{x}{n})^n = e^x$, we obtain that (27) is equal to

$$\begin{aligned}\lim_{n\to\infty}\prod_{\substack{i=1\\ i\neq t}}^{p}(1-\frac{1}{n-k_i})^{k_i-1} &= \prod_{\substack{i=1\\ i\neq t}}^{p}e^{-\lim_{n\to\infty}\frac{k_i-1}{n-k_i}} = e^{-\sum_{\substack{i=1\\ i\neq t}}^{p}\lim_{n\to\infty}\frac{k_i-1}{n-k_i}}\\ &= e^{-\frac{\alpha_1}{1-\alpha_1}-\cdots-\frac{\alpha_{t-1}}{1-\alpha_{t-1}}-\frac{\alpha_{t+1}}{1-\alpha_{t+1}}-\cdots-\frac{\alpha_p}{1-\alpha_p}}.\end{aligned}$$

Since the probability that $v \in V_i$ is α_i for $i = 1,\ldots,p$, the Theorem follows. □

From Theorem 2, we obtain the following corollary:

Corollary 2

(i) Consider the complete p-partite graph $K_{k,k,\ldots,k}$*. The asymptotic probability that a vertex v in* $T \in \mathcal{SP}_{pk}(K_{k,k,\ldots,k})$ *is a leaf is* e^{-1}*, independently of p.*

(ii) Let $\lim_{n\to\infty} \frac{k_i}{n} = \alpha_i$ *with* $0 < \alpha_i < 1$ *for* $i = 1, 2, \ldots, p$*. Then the asymptotic probability that a vertex v in* $T \in \mathcal{SP}_n(K_{k_1,k_2,\ldots,k_p})$ *is a leaf satisfies the inequality*

$$e^{-\sum_{t=1}^{p} \frac{\alpha_t}{1-\alpha_t}} \sum_{i=1}^{p} \alpha_i \, e^{\frac{\alpha_i}{1-\alpha_i}} \;\geq\; e^{-1}, \tag{28}$$

with equality iff $\alpha_1 = \alpha_2 = \cdots = \alpha_p = \frac{1}{p}$*.*

Proof For part (i), $\alpha_i = \frac{1}{p}$ for $i = 1, \ldots, p$. The result now follows from specializing (26) with these values of the α_i. For part (ii), note that (28) is equivalent to

$$e^{-\sum_{t=1}^{p} \frac{\alpha_t}{1-\alpha_t}} \sum_{i=1}^{p} \alpha_i \, e^{\frac{1}{1-\alpha_i}} \;\geq\; 1.$$

Since $\alpha_i > 0$ and $\alpha_1 + \alpha_2 + \cdots + \alpha_p = 1$, Part (ii) is a consequence of Jensen's inequality in the form

$$e^{\sum_{i=1}^{p} \alpha_i y_i} \;\leq\; \sum_{i=1}^{p} \alpha_i \, e^{y_i}$$

with $y_i = (1 - \alpha_i)^{-1}$. □

It is interesting to note that by Corollary 2, the asymptotic probability that a given vertex is a leaf in a spanning tree of a complete multipartite graph takes its minimum value e^{-1} for regular complete multipartite graphs.

Acknowledgement: We would like to thank Glenn Hurlbert for reference [**?**].

References

[1] C.J Colbourn, R.P.J. Day, and L.D. Nel, *Unranking and Ranking Spanning Trees of a Graph*, J. of Algorithms, 10 (1989), pp. 271–286.

[2] Ö. Eğecioğlu and J. Remmel, *Bijections for Cayley trees, spanning trees, and their q-analogues,* J. Combinatorial Theory, 42 (1986), pp. 15–30.

[3] Ö. Eğecioğlu and L.P. Shen, *A Bijective Proof for the Number of Labeled q-trees,* Ars Combinatoria, 25B (1988), pp. 3–30.

[4] R. Onodera, *Number of trees in the complete N-partite graph,* RAAG Res. Notes (3), No. 192 (1973), pp. i + 6.

[5] H. Prüfer, *Never Beweis eines Satzes uber Permutationen,* Arch. Math. Phys. Sci., 27 (1918), pp. 742–744.

[6] A. Nijenhaus and H.S. Wilf, "Combinatorial Algorithms", 2nd. ed., Academic Press, New York, 1978.

[7] E.M. Reingold, J. Nievergelt, and N. Deo, "Combinatorial Algorithms: Theory and Practice," Prentice Hall, Englewood Cliffs, NJ, 1977.