Statistics of Parking Functions and Labeled Forests

Stephan Wagner
Institute of Discrete Mathematics, TU Graz, Austria
Department of Mathematics, Uppsala University, Sweden

Mei Yin
Department of Mathematics, University of Denver, CO, USA

Abstract

In this paper we obtain some new results on the enumeration of parking functions and labeled forests. We introduce new statistics both for parking functions and for labeled forests that are connected to each other by means of a bijection. We determine the joint distribution of two statistics on parking functions and their counterparts on labeled forests. Our results on labeled forests also serve to explain the mysterious equidistribution between two seemingly unrelated statistics in parking functions recently identified by Stanley and Yin and give an explicit bijection between the two statistics.

2012 ACM Subject Classification Mathematics of computing → Enumeration; Mathematics of computing → Generating functions

Keywords and phrases parking function, labeled forest, generating function, Pollak's circle argument, bijection

Digital Object Identifier 10.4230/LIPIcs.AofA.2024.29

Funding *Stephan Wagner*: Supported by the Swedish research council (VR), grant 2022-04030.
Mei Yin: Supported by the University of Denver's Professional Research Opportunities for Faculty Fund 80369-145601.

Acknowledgements The authors benefited from participation in the Workshop on Analytic and Probabilistic Combinatorics at the Banff International Research Station in November 2022 and also from participation in the 34th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms in Taipei in June 2023.

1 Introduction

In this paper we obtain some new results on the enumeration of parking functions and labeled forests. We introduce new statistics both for parking functions and for labeled forests and connect them by means of a bijection. We determine the joint distribution of two statistics on parking functions and their counterparts on labeled forests. Our enumerative results for parking functions and for labeled forests inform each other. In particular, our results on labeled forests serve to explain the mysterious equidistribution between two seemingly unrelated statistics in parking functions recently identified by Stanley and Yin [16] and give an explicit bijection between the two statistics.

We begin with the necessary definitions. In the parking function scenario due to Konheim and Weiss [9], there are n parking spaces on a one-way street, labeled $1, 2, \ldots, n$ in consecutive order. A line of $m \leq n$ cars enters the street, one by one. The ith car drives to its preferred spot π_i and parks there if possible; if the spot is already occupied then the car parks in the first available spot after that. The list of preferences $\pi = (\pi_1, \ldots, \pi_m)$ is called a *parking function* if all cars successfully park. (The parking function is called "classical" when $m = n$.) We denote the set of parking functions by $\mathrm{PF}(m, n)$, where m is the number of cars and n is the number of parking spots. Using the pigeonhole principle, we see that a parking function

35th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2024).
Editors: Cécile Mailler and Sebastian Wild; Article No. 29; pp. 29:1–29:14
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

$\pi \in \mathrm{PF}(m,n)$ must have at most one value $= n$, at most two values $\geq n-1$, and for each k at most k values $\geq n-k+1$, and any such function is a parking function. Equivalently, π is a parking function if and only if

$$\#\{k : \pi_k \leq i\} \geq m-n+i, \quad \text{for } i = n-m+1, \ldots, n. \tag{1}$$

We make two immediate observations from (1). The first observation is that parking functions are invariant under the action of the symmetric group $\mathfrak{S}_m$ permuting the m cars, that is, permuting the list of preferences π. The second observation is that when some π_i takes values in the set $\{1,2,\ldots,n-m+1\}$, changing π_i to any other value in the set $\{1,2,\ldots,n-m+1\}$ has no effect on π being a parking function.

One of the most fundamental results on parking functions is that the number of parking functions is $|\mathrm{PF}(m,n)| = (n-m+1)(n+1)^{m-1}$. A famous combinatorial proof in the classical case was given by Pollak (unpublished but recounted in [5] and [12]). See also Pitman and Stanley [15] for a generalization of Pollak's circle argument. The combinatorial argument boils down to the following easily verified statement: Let G denote the group of all m-tuples $(a_1,\ldots,a_m) \in [n+1]^m$ with componentwise addition modulo $n+1$. Let H be the subgroup of G generated by $(1,1,\ldots,1)$. Then every coset of H contains exactly $n-m+1$ parking functions. Interpreted probabilistically, the combinatorial operation involves assigning m cars on a circle with $n+1$ spots and recording those car assignments where spot $n+1$ is left empty after circular rotation. Since there are $n-m+1$ missing spaces for the assignment of any preference sequence, any preference sequence π has $n-m+1$ rotations that are valid parking functions. Our parking function proofs will be based on refinements of Pollak's proof technique, where we investigate the individual parking statistics for each car the moment it is parked on the circle.

This new line of approach first appeared in a paper by Stanley and Yin [16] and is extended in this paper, where we introduce the new statistics *leading elements* and *size of level set* on parking functions $\pi \in \mathrm{PF}(m,n)$ and examine their joint distributions. The statistic leading elements was introduced in [16] earlier for classical parking functions $\pi \in \mathrm{PF}(n,n)$ and counts the total number of cars whose desired spot is the same as that of the first car. It was shown in [16, Theorem 4.2] via a generating function approach that for classical parking functions the leading elements statistic is equidistributed with the widely-studied 1's statistic that counts the total number of cars whose desired spot is spot 1. This feature of parking functions is quite mysterious as these two parking function statistics seem unrelated and are not of the same nature. While the leading elements statistic is invariant under circular rotation, it does not satisfy permutation symmetry as permuting the entries might change the first element. On the other hand, though the 1's statistic is invariant under permuting all the entries, it does not exhibit circular rotation invariance. Indeed, only one out of $n+1$ rotations of an assignment of n cars on a circle with $n+1$ spots gives a valid parking function. It is thus intriguing what is hidden behind the pair of statistics (leading elements, 1's). By casting parking functions in the context of labeled forests, this question will be answered in our paper. We explain the gist of our argument below.

Let $[n] = \{1,2,\ldots,n\}$ and $[n]_0 = \{0,1,\ldots,n\}$. Let $\mathcal{T}(n)$ denote the set of all *rooted trees* T on the vertex set $[n]_0$ with root 0. More generally, let $\mathcal{F}(m,n)$ denote the set of all *rooted forests* F with $n+1$ vertices and m edges (equivalently, $n-m+1$ distinct tree components) such that a specified set of $n-m+1$ vertices are the roots of the different trees. We label the roots of F by $\{01, 02, \ldots, 0(n-m+1)\}$ and the non-root vertices by $\{1,2,\ldots,m\}$. The fact that the cardinalities of classical parking functions and of rooted trees are the same, i.e.,

$$|\mathrm{PF}(n,n)| = |\mathcal{T}(n)|,$$

and more generally

$$|\operatorname{PF}(m,n)| = |\mathcal{F}(m,n)|,$$

has motivated much work in the study of connections between the two combinatorial structures. One bijective construction between parking functions and labeled forests goes back to Foata and Riordan [5]. Their construction is for the special case $m = n$ and is referred to as a breadth first search (with a queue) on rooted trees. See Yan [17, Section 1.2.3] and also Chassaing and Marckert [1]. We will show that under the bijective correspondence induced by breadth first search, the seemingly unrelated leading elements statistic and the 1's statistic for classical parking functions both become degree statistics. One of them (the root degree) is classical, while the other (degree of the parent of a fixed vertex) appears to be new.

Generally, statistics based on degrees and other aspects of labeled trees have been studied extensively. A well-known generalization of Cayley's tree theorem includes the degrees of all vertices as additional statistics [7, 14]. Another interesting example is the enumeration of labeled trees with respect to their indegrees: there are two versions of defining indegrees, both leading to the same enumeration formula. In the global orientation (see e.g. [14]), all edges are oriented towards the root; in the local orientation (see [13]), they are oriented towards the higher label. For many more interesting statistics of labeled trees, see for instance [3] (descents), [11] (inversions, which were also connected to parking functions in [6]), or [10] (runs).

This paper is organized as follows. In Section 2 we extend the statistic leading elements (denoted $\operatorname{lel}(\pi)$) for classical parking functions that was studied by Stanley and Yin [16] to general parking functions $\pi \in \operatorname{PF}(m,n)$. We also introduce a new statistic, size of level set (denoted $\operatorname{slev}(\pi)$), for parking functions $\pi \in \operatorname{PF}(m,n)$ that extends the notion of the 1's statistic (denoted $\operatorname{ones}(\pi)$) for classical parking functions and study the joint distribution of the level set statistic and the leading elements statistic. We establish the generating function for the pair of statistics $(\operatorname{slev}(\pi), \operatorname{lel}(\pi))$ using variations of Pollak's argument.

In Section 3, we apply the aforementioned bijection between parking functions and labeled forests that is based on breadth first search. By means of this bijection, we find that the pair of statistics $(\operatorname{slev}(\pi), \operatorname{lel}(\pi))$ on the set of parking functions $\operatorname{PF}(m,n)$ is equidistributed with the pair of statistics $(\deg(0), \deg(p))$ on the set of rooted forests $\mathcal{F}(m,n)$, where $\deg(0)$ is the root degree of a rooted forest (the total number of children of all roots $01, 02, \dots, 0(n-m+1)$), and $\deg(p)$ is the number of children of the parent p of the vertex labeled 1 (by degree, we generally mean more precisely the number of children of a vertex in a rooted tree, which is 1 less than the degree in the graph-theoretical sense for non-root vertices).

The pair of statistics $(\deg(0), \deg(p))$ is further considered in Section 4. In particular, we directly prove a formula for the number of rooted forests in $\mathcal{F}(m,n)$ for which $\deg(0)$ and $\deg(p)$ take on given values by means of a combinatorial argument. In the special case $m = n$ (i.e., for labeled trees), the two statistics $\deg(0)$ and $\deg(p)$ also have the same distribution. We provide an explicit bijection for this fact.

In Section 5, we finally examine the asymptotic properties of the statistics investigated in our paper using standard probabilistic tools.

2 Statistics on parking functions

In this section we investigate the joint distribution of the pair of statistics $(\operatorname{slev}(\pi), \operatorname{lel}(\pi))$ on parking functions $\pi \in \operatorname{PF}(m,n)$. The precise definitions of the individual statistics read as follows:

- Leading elements: total number of cars whose desired spot is the same as that of the first car, denoted $\mathrm{lel}(\pi)$. This statistic was recently introduced (in the special case $m = n$) by Stanley and Yin [16].
- Size of level set: total number of cars whose desired spot is in the range $\{1, 2, \ldots, n-m+1\}$, denoted $\mathrm{slev}(\pi)$. When $m = n$, the level set statistic reduces to the 1's statistic for classical parking functions, which counts the total number of cars whose desired spot is spot 1, often denoted $\mathrm{ones}(\pi)$. The level set statistic $\mathrm{slev}(\pi)$ has not been considered before, but the 1's statistic $\mathrm{ones}(\pi)$ has been widely studied.

Our results for $\mathrm{PF}(m, n)$ are extensions of corresponding results for classical parking functions $\mathrm{PF}(n, n)$ in [16]. As mentioned in the introduction, we will expand upon Pollak's ingenious circle argument [5] for the street parking model to derive our results.

The following lemma was proven before using other means, see for example Kenyon and Yin [8, Corollary 3.4]. Our direct combinatorial argument below will shed light on the structure of parking functions and will be useful in the proof of Theorem 4. As implied by the necessary and sufficient condition (1), changing $\pi_1 = 1$ to whichever value below $n-m+1$ will still keep π a parking function. The number of parking functions $\pi \in \mathrm{PF}(m, n)$ with $\pi_1 \in \{1, 2, \ldots, n-m+1\}$ is thus $n-m+1$ times the number of parking functions $\pi \in \mathrm{PF}(m, n)$ with $\pi_1 = 1$.

▶ **Lemma 1.** *We have*

$$\#\{\pi \in \mathrm{PF}(m, n) : \pi_1 = 1\} = (n-m+2)(n+1)^{m-2},$$

which implies that

$$\#\{\pi \in \mathrm{PF}(m, n) : \pi_1 \in \{1, 2, \ldots, n-m+1\}\} = (n-m+1)(n-m+2)(n+1)^{m-2}.$$

Proof. The statement is trivially true for $m = 1$. For $m \geq 2$, we assign cars $2, \ldots, m$ independently on a circle of length $n+1$. Taking circular rotation into consideration, the car assignments give rise to $(n-m+2)(n+1)^{m-2}$ valid parking functions. Note that car 1 will always be able to park if its desired spot is spot 1. Our conclusion readily follows. ◀

The following lemma allows us to split a parking function in $\mathrm{PF}(m, n)$ into an arbitrary map whose range is precisely the set $\{1, 2, \ldots, n-m+1\}$ (that is relevant for the statistic slev) and a parking function on a smaller domain. It will be very useful in proving our results on the distribution of statistics on $\mathrm{PF}(m, n)$.

▶ **Lemma 2.** *Consider a function $\pi : [m] \to [n]$. Fix the elements of π that are equal to one of $1, 2, \ldots, n-m+1$, and suppose that there are $s \geq 0$ such elements (this is precisely $\mathrm{slev}(\pi)$). Let the other elements be $\pi_{j_1}, \pi_{j_2}, \ldots, \pi_{j_{m-s}}$, and define a new function $\tilde{\pi} : [m-s] \to [m-1]$ by $\tilde{\pi}_i = \pi_{j_i} - (n-m+1)$. Then π is a parking function in $\mathrm{PF}(m, n)$ if and only if $\tilde{\pi}$ is a parking function in $\mathrm{PF}(m-s, m-1)$.*

▶ Remark 3. For $s = 0$, there is no valid parking function in view of (1). This is consistent with the fact that $\mathrm{PF}(m, m-1)$ is (trivially) empty.

Proof. We make use of the characterization (1) of parking functions. For any $i > m-n$, we have

$$\#\{k : \pi_k \leq i\} = s + \#\{k : \tilde{\pi}_k \leq i - (n-m+1)\}.$$

So the condition in (1) is equivalent to

$$\#\{k : \tilde{\pi}_k \leq i - (n-m+1)\} \geq m - n + i - s$$

for $i = m-n+1, m-n+2, \dots, n$. Substituting $h = i-(n-m+1)$, this becomes

$$\#\{k : \tilde{\pi}_k \le h\} \ge h+1-s \tag{2}$$

for $h = 0,1,\dots,m-1$. This is precisely the condition for a parking function in $\mathrm{PF}(m-s, m-1)$, except for one detail: the conditions start at $h=0$ rather than $h=s$. However, for $h<s$, (2) is trivially satisfied. This completes the proof. ◀

Lemma 2 means that every parking function in $\mathrm{PF}(m,n)$ can be uniquely decomposed into an arbitrary function $\pi_a : A \to [n-m+1]$ on a set $A \subseteq [m]$ of cardinality s and a function that is equivalent to a parking function π_p in $\mathrm{PF}(m-s, m-1)$.

As a consequence of the decomposition in Lemma 2, we will now be able to prove results on the distributions of the statistics $\mathrm{slev}(\pi)$ and $\mathrm{lel}(\pi)$. We start with the joint distribution of $\mathrm{slev}(\pi)$ and $\mathrm{lel}(\pi)$, which is determined by the following theorem.

▶ **Theorem 4.** *Let $s,t \ge 1$. We have*

$$\begin{aligned}\#\{\pi \in \mathrm{PF}(m,n)\colon\ &\mathrm{slev}(\pi) = s \text{ and } \mathrm{lel}(\pi) = t\} \\ &= \binom{m-2}{s-1, t-1, m-s-t}(n-m+1)^s(m-1)^{m-s-t+1} \\ &\qquad + \binom{m-1}{t-1, s-t, m-s} s(n-m+1)(n-m)^{s-t} m^{m-s-1}.\end{aligned}$$

We will revisit this formula later in the context of rooted forests. The generating function for the joint distribution of $\mathrm{slev}(\pi)$ and $\mathrm{lel}(\pi)$ is obtained in a straightforward fashion by summing over all s and t.

▶ **Corollary 5.**

$$\begin{aligned}\sum_{\pi \in \mathrm{PF}(m,n)} x^{\mathrm{slev}(\pi)} y^{\mathrm{lel}(\pi)} = (n-m+1)xy\,\big[&(m-1)((n-m+1)x+y+m-1)^{m-2} \\ &+(xy+(n-m)x+1)(xy+(n-m)x+m)^{m-2}\big]. \end{aligned} \tag{3}$$

Proof of Theorem 4. Let us count parking functions $\pi \in \mathrm{PF}(m,n)$ for which $\mathrm{slev}(\pi) = s$. Lemma 2 shows that we can decompose π into an arbitrary function π_a from A to $[n-m+1]$, where $|A| = s$ and $A \subseteq [m]$, and a (function equivalent to a) parking function π_p in $\mathrm{PF}(m-s, m-1)$. For $\mathrm{lel}(\pi)$, we distinguish two cases:

- The spot of the first car does not lie in $\{1,2,\dots,n-m+1\}$. In this case, $1 \notin A$, and the value of $\mathrm{lel}(\pi)$ is determined by the function π_p. Recall that by Pollak's argument, for every possible map from $[m-s]$ to $[m]$, there are s possible rotations that will turn it into a parking function in $\mathrm{PF}(m-s, m-1)$. Thus in a randomly chosen parking function, each car (other than the first) takes the same spot as car 1 with the same probability $\frac{1}{m}$, and all the cars are independent. So $\mathrm{lel}(\pi)$ follows a binomial distribution in this case, and there are $s\binom{m-s-1}{t-1}(m-1)^{m-s-t}$ possibilities for π_p such that $\mathrm{lel}(\pi) = t$. Since there are $\binom{m-1}{s}$ choices for the set A (the domain of π_a) and $(n-m+1)^s$ choices for the function π_a itself, we obtain a total contribution of

$$\begin{aligned}\binom{m-1}{s} &\cdot (n-m+1)^s \cdot s\binom{m-s-1}{t-1}(m-1)^{m-s-t} \\ &= \binom{m-2}{s-1, t-1, m-s-t}(n-m+1)^s(m-1)^{m-s-t+1}.\end{aligned}$$

The corresponding generating function is

$$\sum_{s=1}^{m}\binom{m-1}{s}(n-m+1)^s ys(m-1+y)^{m-s-1}x^s$$
$$= (n-m+1)xy(m-1)((n-m+1)x+y+m-1)^{m-2}.$$

- The spot of the first car lies in $\{1,2,\ldots,n-m+1\}$. Then $1\in A$, and $\mathrm{lel}(\pi)$ is completely determined by the function π_a. Each element of $A\backslash\{1\}$ has the same probability $\frac{1}{n-m+1}$ of being mapped to the same element as car 1 by π_a, and all these elements are independent. So given s, $\mathrm{lel}(\pi)$ follows a binomial distribution in this case as well, and given $\pi_a(1)$, there are $\binom{s-1}{t-1}(n-m)^{s-t}$ possibilities for the map π_a such that $\mathrm{lel}(\pi)=t$. There are now $\binom{m-1}{s-1}$ choices for the set A, $n-m+1$ choices for the spot of the first car, and sm^{m-s-1} possible choices for π_p, so this case yields a contribution of

$$\binom{s-1}{t-1}(n-m)^{s-t}\cdot\binom{m-1}{s-1}(n-m+1)\cdot sm^{m-s-1}$$
$$=\binom{m-1}{t-1,s-t,m-s}s(n-m+1)(n-m)^{s-t}m^{m-s-1}.$$

The generating function associated with this case is

$$\sum_{s=1}^{m}\binom{m-1}{s-1}y(n-m+1)(n-m+y)^{s-1}sm^{m-s-1}x^s$$
$$=(n-m+1)xy(xy+(n-m)x+m)^{m-2}(xy+(n-m)x+1).$$

Combining the two contributions, we obtain Theorem 4 and Corollary 5. ◀

Specializing the generating function by setting $x=1$ or $y=1$, we immediately obtain the distributions of $\mathrm{lel}(\pi)$ and $\mathrm{slev}(\pi)$. These are given in the following two corollaries.

▶ **Corollary 6.** *Taking $x=1$ in (3), we have*

$$\sum_{\pi\in\mathrm{PF}(m,n)} y^{\mathrm{lel}(\pi)}=(n-m+1)y(y+n)^{m-1}.$$

▶ **Corollary 7.** *Taking $y=1$ in (3), we have*

$$\sum_{\pi\in\mathrm{PF}(m,n)} x^{\mathrm{slev}(\pi)}=(n-m+1)x\left((n-m+1)x+m\right)^{m-1}.$$

Observe that the generating functions in Corollaries 6 and 7 are identical for $m=n$ (up to renaming the variable). In this case, $\mathrm{slev}(\pi)$ becomes the statistic $\mathrm{ones}(\pi)$ (number of 1's in the parking function).

3 Breadth first search

In this section we explore the implications of the breadth first search (BFS) algorithm connecting parking functions $\mathrm{PF}(m,n)$ and rooted forests $\mathcal{F}(m,n)$. This allows us to transfer results from parking functions to forests. Our construction will extend the corresponding construction between classical parking functions $\mathrm{PF}(n,n)$ and rooted trees $\mathcal{T}(n)$ by Foata and Riordan [5, Section 3]. That our construction is a bijection may be similarly argued as in [5],

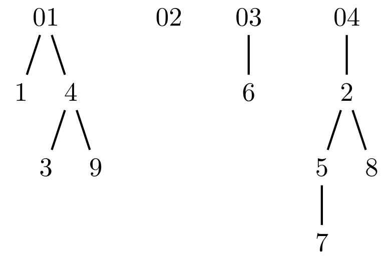

Figure 1 Rooted spanning forest.

with minor adaptations. We will not go over all the technical details here, but will provide the explicit formulas for the generalized construction and illustrate the correspondence with a concrete example.

A forest $F \in \mathcal{F}(m,n)$ may be represented by an acyclic function f, where for a non-root vertex i, $f_i = j$ indicates that vertex j is the parent of vertex i in a tree component of the forest. Take $m = 9$ and $n = 12$. See Figure 1 representing an element of $F \in \mathcal{F}(9,12)$, which corresponds to the acyclic function f given below:

$$\begin{array}{ccccccccccc} i & = & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ f_i & = & 01 & 04 & 4 & 01 & 2 & 03 & 5 & 2 & 4 \end{array}. \tag{4}$$

We read the vertices of the forest in breadth first search (BFS) order. That is, read root vertices in order first, then all vertices at level 1 (children of a root), then those at level 2 (distance 2 from a root), and so on, where vertices at a given level are naturally ordered in order of increasing predecessor, and, if they have the same predecessor, increasing order. See [17, Section 1.2.3] for a description of this graph searching algorithm in the language of computer science. Applying BFS to the forest F in Figure 1, we have

$$v_{01}, \dots, v_{04}, v_5, \dots, v_{13} = 01, 02, 03, 04, 1, 4, 6, 2, 3, 9, 5, 8, 7.$$

We let σ_f^{-1} be the vertex ordering once we remove the root vertices and σ_f be the *inverse order permutation* of σ_f^{-1}.

$$\begin{array}{ccccccccccc} i & = & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \sigma_f^{-1}(i) & = & 1 & 4 & 6 & 2 & 3 & 9 & 5 & 8 & 7 \\ \sigma_f(i) & = & 1 & 4 & 5 & 2 & 7 & 3 & 9 & 8 & 6 \end{array}.$$

We further let $t(f) = (r_1, \dots, r_{12})$ with r_i recording the degree of v_i, starting with v_{01} and ending with v_{12} (ignoring the final vertex v_{13}), that is,

$$t(f) = (2, 0, 1, 1, 0, 2, 0, 2, 0, 0, 1, 0).$$

The sequence $t(f)$ is referred to as the *forest specification* of F.

Via the breadth first search, a generic forest $F \in \mathcal{F}(m,n)$ may thus be uniquely characterized by its associated specification $t(f)$ and order permutation σ_f. Furthermore, the pair $(t(f), \sigma_f)$ must satisfy certain *balance* and *compatibility* conditions. For exact definitions of these conditions, see [8, Section 2.2] and the references therein. Indeed, if we let $\mathcal{C}(m,n)$ be the set of all feasible pairs, then $\mathcal{C}(m,n)$ is in one-to-one correspondence with $\mathcal{F}(m,n)$. It turns out that $\mathcal{C}(m,n)$ is also in one-to-one correspondence with the set of parking functions $\mathrm{PF}(m,n)$, which we now describe.

For a parking function $\pi \in \mathrm{PF}(m,n)$, the associated *specification* is $s(\pi) = (r_1, \ldots, r_n)$, where $r_k = \#\{i : \pi_i = k\}$ records the number of cars whose parking preference is spot k. The *order permutation* $\tau_\pi \in \mathfrak{S}_m$, on the other hand, is defined by $\tau_\pi(i) = \#\{j : \pi_j < \pi_i, \text{ or } \pi_j = \pi_i \text{ and } j \leq i\}$, and so is the permutation that orders the list, without switching elements that are the same. In words, $\tau_\pi(i)$ is the position of the entry π_i in the non-decreasing rearrangement of π. For example, for $\pi = (3,1,3,1)$, $\tau_\pi(1) = 3$, $\tau_\pi(2) = 1$, $\tau_\pi(3) = 4$, and $\tau_\pi(4) = 2$. We can easily recover a parking function π by replacing i in τ_π with the ith smallest term in the sequence $1^{r_1} \ldots n^{r_n}$. As in the case of rooted forests, all feasible pairs $(s(\pi), \tau_\pi)$ for parking functions constitute the set $\mathcal{C}(m,n)$.

Combining the above perspectives, we see that the breadth first search algorithm bijectively connects parking functions and rooted forests, where $(t(f), \sigma_f) = (s(\pi), \tau_\pi)$. Continuing with our earlier example, for the forest $F \in \mathcal{F}(9,12)$ in Figure 1 with acyclic function representation f given by (4), we have

$$s(\pi) = (2,0,1,1,0,2,0,2,0,0,1,0),$$

and

$$\begin{array}{ccccccccccc} i & = & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \tau_\pi^{-1}(i) & = & 1 & 4 & 6 & 2 & 3 & 9 & 5 & 8 & 7 \\ \tau_\pi(i) & = & 1 & 4 & 5 & 2 & 7 & 3 & 9 & 8 & 6 \end{array}.$$

We form the non-decreasing rearrangement sequence associated with $s(\pi)$:

$$1^2, 3^1, 4^1, 6^2, 8^2, 11^1 = 1,1,3,4,6,6,8,8,11.$$

Replacing i in τ_π with the ith smallest term in this sequence yields the corresponding parking function $\pi \in \mathrm{PF}(9,12)$ given below:

$$\begin{array}{ccccccccccc} i & = & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \pi_i & = & 1 & 4 & 6 & 1 & 8 & 3 & 11 & 8 & 6 \end{array}.$$

This bijective construction between parking functions and rooted forests has some interesting implications that are listed in the following theorem.

▶ **Theorem 8.** *The following statistics are equidistributed:*

- *The number of times π_i appears in a parking function $\pi \in \mathrm{PF}(m,n)$ equals the degree of the parent of vertex i in the corresponding forest $F \in \mathcal{F}(m,n)$.*
- *The number of times $1, 2, \ldots, n-m+1$ appears in a parking function $\pi \in \mathrm{PF}(m,n)$ respectively equals the degree of the root vertex $01, 02, \ldots, 0(n-m+1)$ in the corresponding forest $F \in \mathcal{F}(m,n)$.*

Proof. This is due to our specific construction. From a forest F to a parking function π, we have

$$\pi_i = \begin{cases} j & \text{if } f_i = 0j \text{ for some } j = 1, 2, \ldots, n-m+1, \\ (n-m+1) + \sigma_f(f_i) & \text{otherwise.} \end{cases}$$

Conversely, from a parking function π to a forest F, we have

$$f_i = \begin{cases} 0j & \text{if } \pi_i = j \text{ for some } j = 1, 2, \ldots, n-m+1, \\ \tau_\pi^{-1}(\pi_i - (n-m+1)) & \text{otherwise.} \end{cases}$$

The second claim is clear. For the first claim, we note that $\pi_i = \pi_j$ corresponds to $f_i = f_j$, i.e., vertices i and j have the same parent.

In our illustrative example, the number of times each entry of π appears is given by the vector $\vec{w} = (2, 1, 2, 2, 2, 1, 1, 2, 2)$, whose entries coincide with the respective degree of the parent of vertex i for $i \in \{1, 2, \dots, 9\}$. An immediate consequence of this fact is that

$$\sum_{i=1}^{9} w_i = \sum_{i=1}^{4} \deg^2(0i) + \sum_{i=5}^{13} \deg^2(i) = 15.$$

We also observe that the number of times $1, 2, 3, 4$ appears in π respectively agrees with the degrees of the roots $01, 02, 03, 04$ in the forest F, both yielding the vector $(2, 0, 1, 1)$. ◀

4 Statistics on trees and forests

In the bijection described in the previous section, the number of times π_1 occurs in the parking function (the statistic $\mathrm{lel}(\pi)$) corresponds to the degree $\deg(p)$ of the parent p of vertex 1 (see Theorem 8). Moreover, the total number of times $1, 2, \dots, n - m + 1$ occur in the parking function (our statistic $\mathrm{slev}(\pi)$) corresponds to the total root degree $\deg(0)$, i.e., the sum of the degrees of all roots. Hence, the pair $(\deg(0), \deg(p))$ follows the same joint distribution as the pair $(\mathrm{slev}(\pi), \mathrm{lel}(\pi))$. The following generating function identity is therefore an automatic consequence of Corollary 5.

▶ **Theorem 9.**

$$\sum_{F \in \mathcal{F}(m,n)} x^{\deg(0)} y^{\deg(p)} = (n - m + 1)xy \big[(m - 1)((n - m + 1)x + y + m - 1)^{m-2} \\ + (xy + (n - m)x + 1)(xy + (n - m)x + m)^{m-2}\big]. \quad (5)$$

When $m = n$, the rooted spanning forest $F \in \mathcal{F}(m, n)$ reduces to a rooted tree $T \in \mathcal{T}(n)$. We recognize from Corollaries 6 and 7 that $\mathrm{ones}(\pi)$ and $\mathrm{lel}(\pi)$ are equidistributed. The breadth first search algorithm maps them to $\deg(0)$ and $\deg(p)$, respectively, so those are equidistributed as well. In words, the distribution of the number of children of the root 0 of a rooted labeled tree follows the same distribution as the number of children of the parent of vertex 1 (or indeed by symmetry the parent of any fixed vertex). The property of being parent of a specific vertex induces a bias towards higher degrees, which turns out to be equivalent to the bias induced by being the root (which necessarily has at least one child). The following procedure provides an explicit bijection for the equidistribution of $\deg(0)$ and $\deg(p)$.

1. Remove the edge connecting vertices 1 and p.
2. Connect vertices 0 and 1 by an edge.
3. Interchange vertices 0 and p.

See Figure 2 for the general procedure and Figure 3 for an example. This map has the extra benefit of being an involution. Moreover, the degrees of all vertices except 0 and p are preserved. Some nice features are hence introduced in the corresponding parking function bijection, where in our example

$$\pi = (8, 4, 5, 1, 2, 1, 1, 5, 6) \leftrightarrow \pi' = (5, 8, 2, 2, 1, 5, 5, 4, 9).$$

We see that $\mathrm{ones}(\pi)$ and $\mathrm{lel}(\pi)$ are switched, but the frequencies of the non-1 and non-leading elements are preserved up to permutation. In the example, the non-1 and non-leading elements in π are 5 (occurring twice), 2, 4, and 6. Those in π' are 2 (occurring twice), 4, 8, and 9.

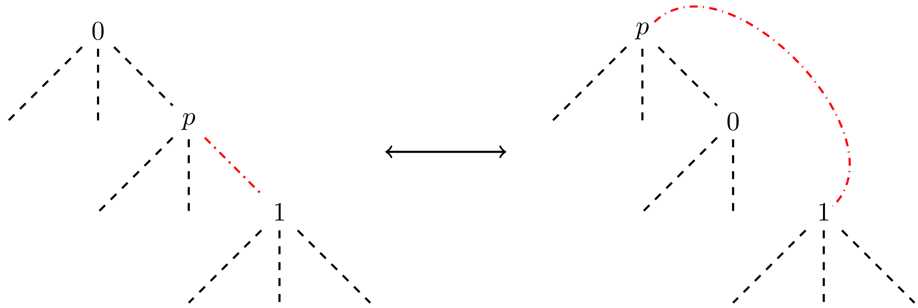

Figure 2 A bijective map between $\deg(0)$ and $\deg(p)$ (illustration).

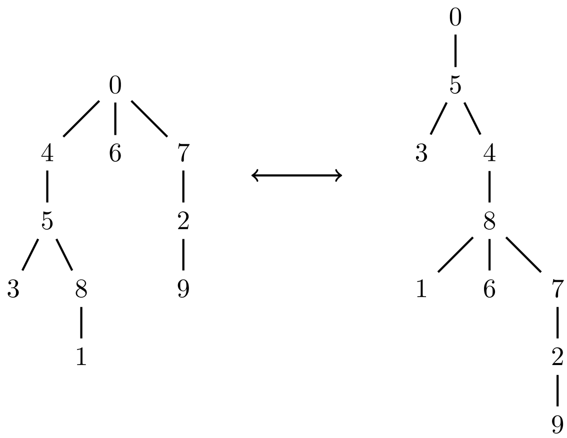

Figure 3 A bijective map between $\deg(0)$ and $\deg(p)$ (example).

The counting formula in our next proposition is equivalent to Theorem 4 in view of the bijection between parking functions and forests. In the following, we also illustrate it with a combinatorial proof of the statement in the setting of forests.

▶ **Proposition 10.** *Let $s, t \geq 1$. We have*

$$\begin{aligned}&\#\{F \in \mathcal{F}(m,n) \colon \deg(0) = s \text{ and } \deg(p) = t\}\\ &= \binom{m-2}{s-1, t-1, m-s-t}(n-m+1)^s(m-1)^{m-s-t+1}\\ &\qquad\qquad\qquad + \binom{m-1}{t-1, s-t, m-s} s(n-m+1)(n-m)^{s-t}m^{m-s-1}.\end{aligned}$$

Proof. As a starting point, we recall the well-known result that the number of rooted forests with vertex set $[a]$ and b components whose root labels are given is ba^{a-b-1} (see [14]). Thus the number of such rooted forests *with one distinguished vertex* (possibly one of the roots) is ba^{a-b}, and by symmetry the number of such rooted forests where a vertex in the first component is distinguished must be a^{a-b}. Now in order to prove our formula, we distinguish two cases:

Case 1. The parent p of vertex 1 is not one of the roots. A forest F with this property as well as $\deg(0) = s$ and $\deg(p) = t$ can be uniquely constructed as follows:

- Choose a label $r \in [m] \setminus \{1\}$ ($m-1$ possibilities).
- Choose two disjoint sets of labels $\{x_1, x_2, \dots, x_{s-1}\}$ and $\{y_1, y_2, \dots, y_{t-1}\}$ from $[m]\setminus\{1, r\}$ (for which there are $\binom{m-2}{s-1, t-1, m-s-t}$ possibilities).
- Choose a rooted forest on $[m] \setminus \{1\}$ with root labels $r, x_1, x_2, \dots, x_{s-1}, y_1, y_2, \dots, y_{t-1}$ and a distinguished vertex p in the first component (there are $(m-1)^{m-s-t}$ possibilities, as explained above). Note that potentially $p = r$.
- Split the distinguished vertex p into two vertices, labeled p and 1 respectively, where p is the parent and 1 is the child, and all former children of p now become children of 1.
- Add roots $01, 02, \dots, 0(n-m+1)$, and connect each of the vertices $r, x_1, x_2, \dots, x_{s-1}$ with one of these roots by an edge. There are $(n-m+1)^s$ possibilities for this.
- Add edges between vertex p and $y_1, y_2, \dots, y_{t-1}$. Note that p is indeed the parent of 1 in this construction, and that $\deg(0) = s$ as well as $\deg(p) = t$.

It is easy to reverse the procedure given a forest with $\deg(0) = s$ and $\deg(p) = t$ for which p is not one of the roots. So to summarize, we have

$$\binom{m-2}{s-1, t-1, m-s-t}(n-m+1)^s(m-1)^{m-s-t+1}$$

possible forests in this case, which accounts for the first term in our formula. Case 1 is illustrated in the special case $n = m = 13$, $s = 3$ and $t = 2$ in Figure 4. The choice of root labels is $r = 5$, $\{x_1, x_2\} = \{3, 13\}$ and $\{y_1\} = \{9\}$.

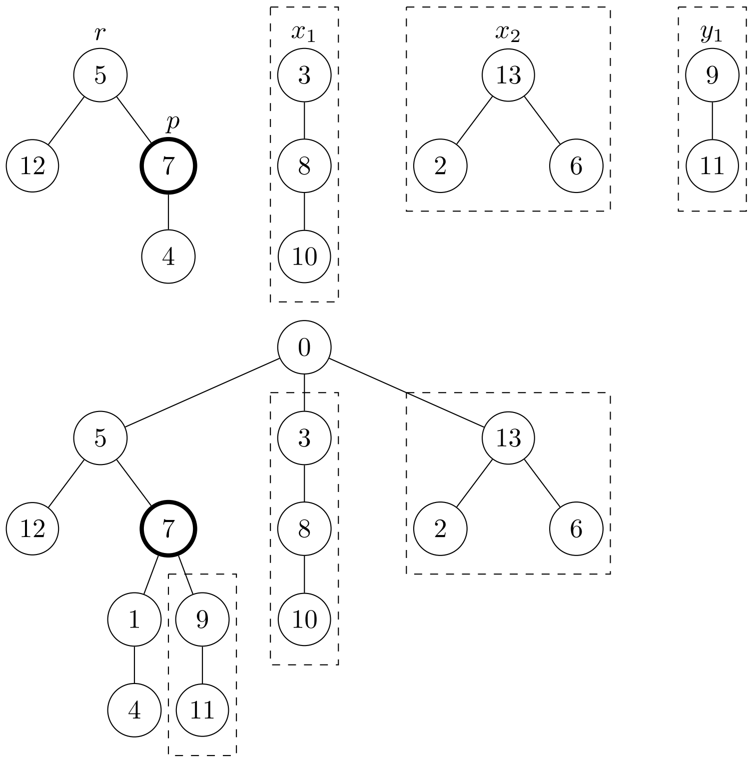

Figure 4 Illustration of the procedure: the rooted forest (top) with the distinguished vertex p indicated by a thick node, and the final tree (bottom).

Case 2. The parent p of vertex 1 is one of the roots. Again, there is a unique way to construct all these forests:

- Select a set of labels $\{x_1, x_2, \dots, x_{s-1}\}$ from $[m] \setminus \{1\}$ in $\binom{m-1}{s-1}$ ways.
- Construct a rooted forest with vertex set $[m]$ and root labels $1, x_1, x_2, \dots, x_{s-1}$ (there are sm^{m-s-1} possibilities).
- Among the labels $x_1, x_2, \dots, x_{s-1}$, choose the siblings of vertex 1; there are $\binom{s-1}{t-1}$ possible choices.
- Pick one of the $n-m+1$ roots $01, 02, \dots, 0(n-m+1)$ as the parent p of vertex 1, and connect it and all the siblings chosen in the previous step to it by an edge.
- Lastly, pick one of the other $n-m$ roots as parent for each of the remaining vertices with label in the set $\{x_1, x_2, \dots, x_{s-1}\}$. This step yields $(n-m)^{s-t}$ possibilities.

Putting these together gives us the second term and thus completes the proof. ◀

5 Limit distributions

In this section we conclude the paper with some asymptotic observations on the statistics investigated in our paper. We consider the scenario where m is a linear function of n, and determine the limit distributions of the statistics lel and slev (for parking functions) and $\deg(0)$ as well as $\deg(p)$ (for rooted forests), respectively.

▶ **Proposition 11.** *Let s be a fixed positive integer, and take $m = cn$ for some $0 < c \le 1$ as $n \to \infty$. Consider the parking preference $\pi \in \mathrm{PF}(m,n)$ chosen uniformly at random, and let $\mathrm{lel}_s(\pi)$ be the number of cars with the same preference as car s. Then $\mathrm{lel}_s(\pi) - 1 \xrightarrow{\mathrm{d}} \mathrm{Poisson}(c)$, i.e., for every fixed nonnegative integer j,*

$$\mathbb{P}\left(\mathrm{lel}_s(\pi) = 1 + j \mid \pi \in \mathrm{PF}(m,n)\right) \sim \frac{c^j e^{-c}}{j!}.$$

Proof. By permutation symmetry, we only need to prove this result for $s = 1$, where $\mathrm{lel}_s(\pi)$ is exactly $\mathrm{lel}(\pi)$. We divide both sides of the generating function for $\mathrm{lel}(\pi)$ in Corollary 6 through by $(n-m+1)(n+1)^{m-1}$. The right side becomes the probability generating function of $S(m,n) = 1 + \sum_{i=1}^{m-1} X_i$, where the X_i are independent Bernoulli random variables:

$$X_i = \begin{cases} 0, & \text{probability } n/(n+1), \\ 1, & \text{probability } 1/(n+1). \end{cases}$$

Hence we have a standard case of the law of rare events, leading to a Poisson limit distribution. ◀

▶ **Corollary 12.** *Let s be a fixed positive integer, and take $m = cn$ for some $0 < c \le 1$ as $n \to \infty$. Consider the labeled forest $F \in \mathcal{F}(m,n)$ chosen uniformly at random, and let $\deg(p_s)$ be the degree of the parent p_s of vertex s. Then $\deg(p_s) - 1 \xrightarrow{\mathrm{d}} \mathrm{Poisson}(c)$.*

Proof. By permutation symmetry, we only need to prove this result for $s = 1$, where $\deg(p_s)$ is exactly $\deg(p)$. Since $\mathrm{lel}(\pi)$ and $\deg(p)$ are equidistributed (by Theorem 8), the statement follows from Proposition 11. ◀

▶ **Proposition 13.** *Take $m = cn$ for some $0 < c < 1$ as $n \to \infty$. Consider the parking preference $\pi \in \mathrm{PF}(m,n)$ chosen uniformly at random. Then we have*

$$\frac{\mathrm{slev}(\pi) - c(1-c)n}{\sqrt{c^2(1-c)n}} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1).$$

Proof. We proceed as in the proof of Proposition 11 and divide both sides of the generating function for $\mathrm{slev}(\pi)$ in Corollary 7 through by $(n-m+1)(n+1)^{m-1}$. The right side becomes the probability generating function of $S(m,n) = 1 + \sum_{i=1}^{m-1} X_i$, where the X_i are independent Bernoulli random variables:

$$X_i = \begin{cases} 0, & \text{probability } m/(n+1), \\ 1, & \text{probability } (n-m+1)/(n+1). \end{cases}$$

The probabilities converge to c and $1-c$ respectively, and the standard central limit theorem applies. This means that $\mathrm{slev}(\pi)$ may be approximated by $\mathcal{N}(0,1)$ after standardization. ◀

▶ **Corollary 14.** *Take $m = cn$ for some $0 < c < 1$ as $n \to \infty$. Consider the labeled forest $F \in \mathcal{F}(m,n)$ chosen uniformly at random. Then we have*

$$\frac{\deg(0) - c(1-c)n}{\sqrt{c^2(1-c)n}} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1).$$

Proof. Since $\mathrm{slev}(\pi)$ and $\deg(0)$ are equidistributed (again by Theorem 8), this is the same proof as for Proposition 13. ◀

▶ Remark 15. The asymptotic analysis of the special case $\mathrm{ones}(\pi)$ for $\pi \in \mathrm{PF}(n,n)$ was conducted by Diaconis and Hicks [2]. The limit distribution of the root degree in labeled trees is classical [4, Example IX.6]. Both $\mathrm{ones}(\pi) - 1$ and $\deg(0) - 1$ can be approximated by Poisson(1).

References

1 Philippe Chassaing and Jean-François Marckert. Parking functions, empirical processes, and the width of rooted labeled trees. *Electron. J. Combin.*, 8(1):Research Paper 14, 19, 2001. doi:10.37236/1558.

2 Persi Diaconis and Angela Hicks. Probabilizing parking functions. *Adv. in Appl. Math.*, 89:125–155, 2017. doi:10.1016/j.aam.2017.05.004.

3 Ömer Eğecioğlu and Jeffrey B. Remmel. Bijections for Cayley trees, spanning trees, and their q-analogues. *J. Combin. Theory Ser. A*, 42(1):15–30, 1986. doi:10.1016/0097-3165(86)90004-X.

4 Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009.

5 Dominique Foata and John Riordan. Mappings of acyclic and parking functions. *Aequationes Math.*, 10:10–22, 1974. doi:10.1007/BF01834776.

6 António Guedes de Oliveira and Michel Las Vergnas. Parking functions and labeled trees. *Sém. Lothar. Combin.*, 65:Art. B65e, 10, 2010/12.

7 Frank Harary and Edgar M. Palmer. *Graphical enumeration.* Academic Press, New York-London, 1973.

8 Richard Kenyon and Mei Yin. Parking functions: from combinatorics to probability. *Methodol. Comput. Appl. Probab.*, 25(1):Paper No. 32, 30, 2023. doi:10.1007/s11009-023-10022-5.

9 Alan G. Konheim and Benjamin Weiss. An occupancy discipline and applications. *SIAM J. Appl. Math.*, 14:1266–1274, 1966.

10 Marie-Louise Lackner and Alois Panholzer. Runs in labelled trees and mappings. *Discrete Math.*, 343(9):111990, 16, 2020. doi:10.1016/j.disc.2020.111990.

11 C. L. Mallows and John Riordan. The inversion enumerator for labeled trees. *Bull. Amer. Math. Soc.*, 74:92–94, 1968. doi:10.1090/S0002-9904-1968-11888-9.

12 John Riordan. Ballots and trees. *J. Combinatorial Theory*, 6:408–411, 1969.

13 Heesung Shin and Jiang Zeng. A bijective enumeration of labeled trees with given indegree sequence. *J. Combin. Theory Ser. A*, 118(1):115–128, 2011. doi:10.1016/j.jcta.2010.07.001.

14 Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. doi:10.1017/CBO9780511609589.

15 Richard P. Stanley and Jim Pitman. A polytope related to empirical distributions, plane trees, parking functions, and the associahedron. *Discrete Comput. Geom.*, 27(4):603–634, 2002. doi:10.1007/s00454-002-2776-6.

16 Richard P. Stanley and Mei Yin. Some enumerative properties of parking functions, 2023. arXiv: 2306.08681.

17 Catherine H. Yan. Parking functions. In *Handbook of enumerative combinatorics*, Discrete Math. Appl. (Boca Raton), pages 835–893. CRC Press, Boca Raton, FL, 2015.