

Probabilistic Parking Functions

Irfan Durmić^a Alex Han^a Pamela E. Harris^b
Rodrigo Ribeiro^c Mei Yin^c

Submitted: Nov 3, 2022; Accepted: Mar 6, 2023; Published: Aug 11, 2023
© The authors. Released under the CC BY-ND license (International 4.0).

Abstract

We consider the notion of classical parking functions by introducing randomness and a new parking protocol, as inspired by the work presented in the paper “Parking Functions: Choose your own adventure,” (arXiv:2001.04817) by Carlson, Christensen, Harris, Jones, and Rodríguez. Among our results, we prove that the probability of obtaining a parking function, from a length n preference vector, is independent of the probabilistic parameter p . We also explore the properties of a preference vector given that it is a parking function and discuss the effect of the probabilistic parameter p . Of special interest is when $p = 1/2$, where we demonstrate a sharp transition in some parking statistics. We also present several interesting combinatorial consequences of the parking protocol. In particular, we provide a combinatorial interpretation for the array described in OEIS A220884 as the expected number of preference sequences with a particular property related to occupied parking spots. Lastly, we connect our results to other weighted phenomena in combinatorics and provide further directions for research.

Mathematics Subject Classifications: 05A19, 60C05, 05A16

1 Introduction

The classical parking process occurs on a one-way street with $n \in \mathbb{N} = \{1, 2, 3, \dots\}$ parking spots numbered in sequence from one to n . In this setting, n cars enter the street, from left to right, each having a parking preference, which we compile in a vector $\alpha = (a_1, a_2, \dots, a_n) \in [n]^n$, where $[n] = \{1, 2, \dots, n\}$. Each car enters the street and proceeds to their preferred parking spot. If they find it unoccupied they park there, else they continue down the street parking in the next available spot. If under this parking protocol all of the cars are able to park, then we say that the preference vector α is a parking function of length n . For example, any preference vector that is a permutation is a parking function as every car parks in its preferred parking spot. In addition the all

^aMathematics and Statistics, Williams College, Williamstown, MA, U.S.A.
(id5@williams.edu, alh5@williams.edu).

^bDepartment of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI, U.S.A.
(peharris@uwm.edu).

^cDepartment of Mathematics, University of Denver, Denver, CO, U.S.A. (rodrigo.ribeiro@du.edu, mei.yin@du.edu).

ones vector $\alpha = (1, 1, \dots, 1) \in [n]^n$ is also a parking function in which, for each $1 \leq i \leq n$, car i parks in spot i . However, not every parking preference is a parking function. Take for example the vector (n, n, \dots, n) in which case the first car parks in its preferred spot n , while the remaining $n - 1$ cars, having the same preference, find their spot occupied and then exit the street.

First explored by Konheim and Weiss in the 1960s [12], classical parking functions have received considerable attention by mathematical researchers. If PF_n denotes the set of all parking functions of length n , then it is well-known that $|\text{PF}_n| = (n + 1)^{n-1}$. Many generalizations of parking functions exist and include considering the case where there are more parking spots than cars, considering cases where cars prefer an interval of parking spots rather than a single parking spot, and others include changing the parking procedure and allowing cars to do something different whenever they find their preferred spot occupied. For an expository article with recent finding and many open problems related to parking functions we recommend [2].

We remark that although parking functions, and their generalizations, are well-studied there has been less work done on considering a probabilistic parking protocol. With this point of view in mind, we provide some initial results involving a probabilistic parking protocol. Our initial result considers the following probabilistic scenario: Fix $p \in [0, 1]$ and consider a coin which flips to heads with probability p and tails with probability $1 - p$. Then, our parking protocol proceeds as follows: If a car arrives at its preferred spot and finds it unoccupied it parks there. If instead the spot is occupied, then the driver tosses the biased coin. If the coin lands on heads, with probability p , the driver continues moving forward in the street. However, if the coin lands on tails, with probability $1 - p$, the car moves backwards and tries to find an unoccupied parking spot. Given this new probabilistic parking function protocol we determine the likelihood that a preference vector $\alpha \in [n]^n$ is a parking function. We also explore the properties of α given that it is a parking function and demonstrate the effect of the parameter p (and lack thereof) on the parking protocol. Before proceeding further, we note a parking symmetry imposed by this probabilistic parking protocol: Having n cars enter the street from left to right with preference vector $\alpha = (a_1, \dots, a_n)$ and park under protocol with parameter p depicts the same scenario as having n cars enter the street from right to left with preference vector $\alpha' = (n + 1 - a_1, \dots, n + 1 - a_n)$ and park under protocol with parameter $1 - p$.

Due to the probabilistic nature of our model, we offer a word of caution in interpreting our results. In the case of classical parking functions, a preference vector $\alpha \in [n]^n$ is either deterministically a parking function or not, and we can talk about the set of parking functions $\text{PF}_n \subseteq [n]^n$ and its cardinality $|\text{PF}_n|$. Contrarily, by incorporating a probabilistic parking protocol, all preference vectors $\alpha \in [n]^n$ have a positive probability of being a parking function. In stating our results, $\alpha \in \text{PF}_n$ depicts the situation that n cars with preference vector α park, and we are interested in the probability that this happens and its further implications.

Our main results are the following.¹

¹While writing a first draft of this work, we became aware of parallel, independent work by Tian and Trevino [17]. Aside from both introducing the probabilistic parking protocol, the results of the two papers are essentially disjoint, except that Theorem 1 appeared as Theorem 4 in [17] with a different proof.

Theorem 1. Consider the preference vector $\alpha \in [n]^n$, chosen uniformly at random. Then

$$P(\alpha \in \text{PF}_n | \alpha \in [n]^n) = \frac{(n+1)^{n-1}}{n^n}. \quad (1.1)$$

We emphasize that this result is rather surprising. In Example 10 we show that the probability of a particular preference vector $\alpha \in [n]^n$ being a parking function does depend on p . What is magical is that the probabilities of being a parking function for all preference vectors add up in a way so that dependence on p is canceled and there is invariance to the forward probability p for the randomly selected vector. In Section 2.6, we show that this invariance generalizes further for $m \in [n]$ cars on a one-way street with n spots. That is we establish the following.

Theorem 2. For $m \in [n]$, consider the preference vector $\alpha \in [n]^m$, chosen uniformly at random. Then

$$P(\alpha \in \text{PF}_n | \alpha \in [n]^m) = \frac{(n+1-m)(n+1)^{m-1}}{n^m}. \quad (1.2)$$

Sections 2.3 and 2.4 are devoted to the natural follow-up questions: *Does introducing randomness to the parking protocol achieve anything new? What kind of parking statistics depend on the probabilistic parameter p ?* In 2.3 we present one such parking statistic, namely the parking preference of the last car a_n , and explore its statistical properties. In Theorem 3 we give the distribution of a_n and in Theorem 4 we calculate the asymptotics of its expected value, and the formulas show explicit p dependence.

Theorem 3. Consider the preference vector $\alpha = (a_1, a_2, \dots, a_n) \in [n]^n$, chosen uniformly at random. Then given that $\alpha \in \text{PF}_n$,

$$P(a_n = j | \alpha \in \text{PF}_n) = \frac{2}{n+1} - \frac{1}{(n+1)^{n-1}} \left[p \sum_{s=n-j+1}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \right. \\ \left. + (1-p) \sum_{s=0}^{n-j-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \right]. \quad (1.3)$$

Theorem 4. Take n large. For preference vector $\alpha = (a_1, \dots, a_n) \in [n]^n$ chosen uniformly at random, we have

$$\mathbb{E}(a_n | \alpha \in \text{PF}_n) = \frac{n+1}{2} - (2p-1) \left[\frac{\sqrt{2\pi}}{4} n^{1/2} - \frac{7}{6} \right] + o(1). \quad (1.4)$$

Based on these results, in Section 2.4, we derive the rate of convergence of the distribution of a_n to uniform distribution on $[n]$ in total variation (TV) distance. In the classical parking model (corresponding to $p = 1$), this rate of convergence was shown to be upper bounded by $O(1/\sqrt{n})$ in [1]. We significantly extend this result and find the exact rate of convergence for any $p \in [0, 1]$. Specifically, we show that for generic $p \neq 1/2$ the convergence rate is $\Theta(1/\sqrt{n})$, while for $p = 1/2$ the convergence rate is $\Theta(1/n)$. Note that a consequence of Theorem 3 is that the conditional distribution of a_n for a generic p can be written as a convex combination of the cases $p = 0$ and $p = 1$. This convexity argument will be central to the proof of Theorem 6.

Theorem 5. Let $Q_{n,p}(\cdot)$ be $P(a_n = \cdot \mid \alpha \in PF_n)$ under the probabilistic parking protocol with parameter p . For $p = 1/2$, the following bound holds

$$\|Q_{n,1/2} - \text{Uni}_n\|_{\text{TV}} = \Theta\left(\frac{1}{n}\right). \quad (1.5)$$

Theorem 6. Let $Q_{n,p}(\cdot)$ be $P(a_n = \cdot \mid \alpha \in PF_n)$ under the probabilistic parking protocol with parameter p . For $p \neq 1/2$, we have that

$$\|Q_{n,p} - \text{Uni}_n\|_{\text{TV}} = \Theta\left(\frac{1}{\sqrt{n}}\right). \quad (1.6)$$

The much faster convergence to uniform when $p = 1/2$ is reflected in our simulations. Figure 1 shows a histogram of the values of a_n based on 100,000 random samples for $n = 100$, with varying probabilistic parameter p . We give some intuitive explanations for this interesting phenomenon. As p increases from 0 to $1/2$, the distribution of a_n places less and less mass on cars with large parking preferences, as cars are gradually losing their backwards moving bias when their desired spot is taken. Similarly, as p decreases from 1 to $1/2$, the distribution of a_n places less and less mass on cars with small parking preferences, as cars are gradually losing their forward moving bias when their desired spot is taken. For $p = 1/2$, the movement bias of a car when the desired spot is taken is gone because going left and right have the same probability, so the distribution of a_n is much closer to uniform distribution.

In Section 3, we then present several combinatorial consequences of the probabilistic parking protocol. In particular, we relate this to the notion of lucky cars in Section 3.1, inspired by Gessel and Seo [6, Page 1], and prove the following.

Theorem 7. Suppose n cars park on a circle with spots $[n + 1]$. The expected number of preference vectors in $[n + 1]^n$ containing $0 \leq k < n$ unlucky cars is

$$U_n(k) = (n + 1)! \sum_{C_k \vdash \{2, \dots, n\}} \prod_{c \in C_k} \frac{i - 1}{(n + 1) - (i - 1)}, \quad (1.7)$$

where C_k is a size- k subset of the cars $\{2, \dots, n\}$, and each car $i \in C_k$ is unable to park at its preferred spot.

Then, in Section 3.2, we further explore Theorem 7 and present an interesting connection to a problem proposed by Novelli and Thibon [14] involving the integer sequence OEIS [A220884](#), summarized by the following.

Theorem 8. Let $Q_n(q)$ be the generating polynomial $\prod_{k=2}^n [(n + 1 - k)q + k]$. Then

$$Q_n(q) = \sum_{k=0}^{n-1} U_n(k)q^k. \quad (1.8)$$

Finally, in Section 3.3, we present an interesting connection between Pascal's triangle and a recurrence relation which counts the expected number of preference sequences with a particular property related to occupied parking spots. This recurrence relation has the form:

$$E_n(i, k) = ((n + 1) - (i - 1)) \cdot E_n(i - 1, k) + (i - 1) \cdot E_n(i - 1, k - 1), \quad (1.9)$$

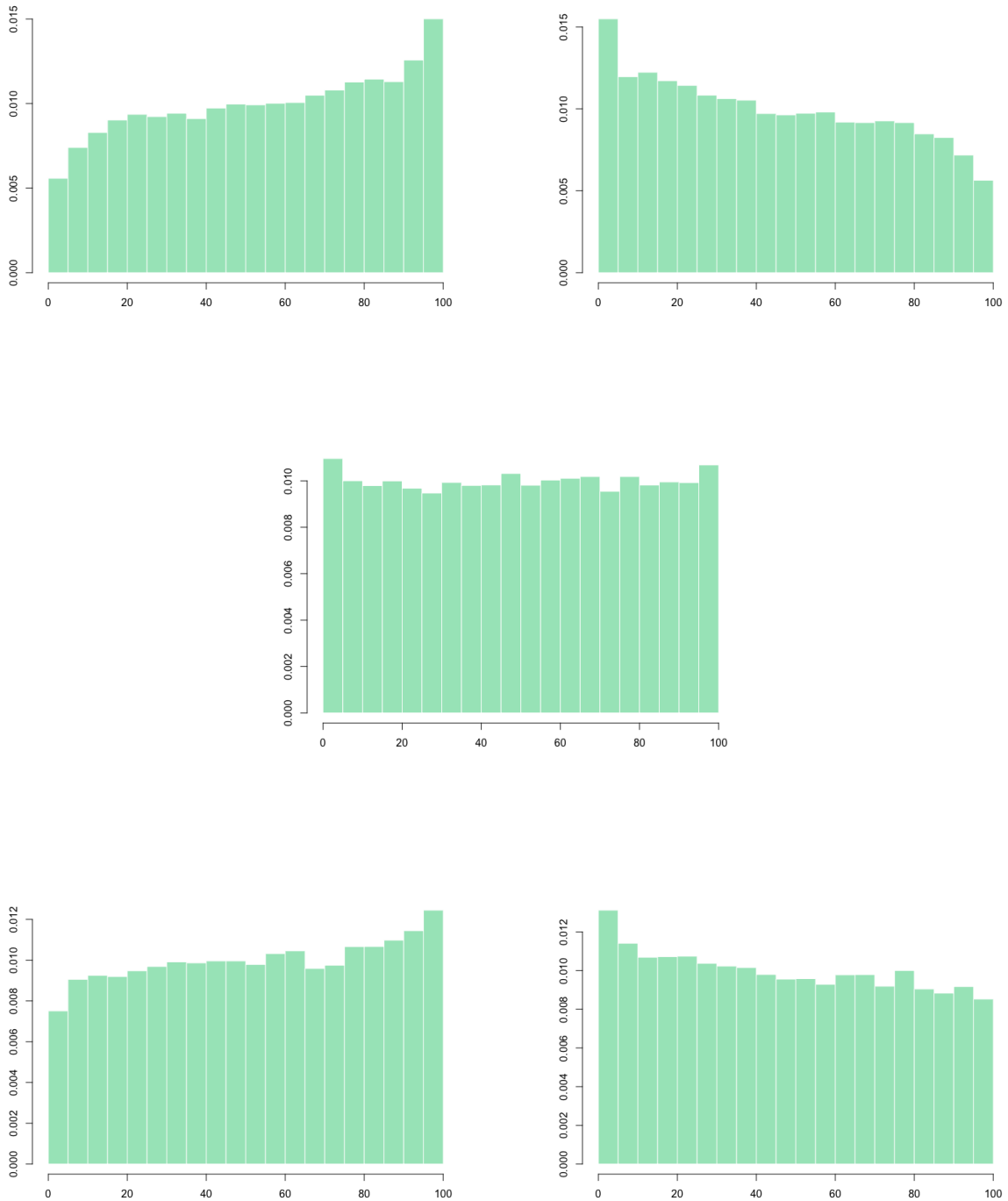


Figure 1: The conditional distribution of a_n (parking preference of the last car) in 100,000 samples of preference vectors of size 100 chosen uniformly at random. The upper left plot is for $p = 0$ and the upper right plot is for $p = 1$. The middle plot is for $p = 0.5$. The lower left plot is for $p = 0.25$ and the lower right plot is for $p = 0.75$. Note the preference symmetry between p and $1 - p$ as observed earlier in the introduction.

where $E_n(i, k)$ denotes the expected number of preference vectors in $[n + 1]^i$ containing k unlucky cars, for some $0 \leq k < i \leq n$. Note that Gessel and Seo [6] define a car as lucky if it is able to park at its preferred spot.

This paper is organized as follows. In Section 2, we prove that the probabilistic parameter p for the parking protocol involving a coin flip does not affect the probability that a preference vector $\alpha \in [n]^n$ is a parking function. We also explore the properties of α given that it is a parking function and discuss situations in which the effect of the probabilistic parameter p comes up. Of special interest is when $p = 1/2$, where we demonstrate a sharp transition in some parking statistics. In Section 3, we present several interesting combinatorial consequences of the parking protocol. We conclude with Section 4, where we offer directions for future work.

2 Introducing probability

We begin by stating an important result credited to Henry O. Pollak². We emphasize its importance in motivating the proof of Theorem 1, but direct the interested reader to [7, 8] for a detailed proof of counting the number of parking functions of length n .

Theorem 9. *If $n \geq 1$, then $|\text{PF}_n| = (n + 1)^{n-1}$.*

2.1 The Coin Problem

We now introduce the Coin Problem, which involves a new parking protocol, motivated by Carlson et al. in [2]. We have n cars in a queue to enter a one-way street with n parking spots numbered from 1 to n . Let a_i denote the preference of car i , and let $\alpha = (a_1, a_2, \dots, a_n) \in [n]^n$ denote the vector of all n cars' parking preferences. In the Coin Problem, each parking preference $a_i \in [n]$ is chosen uniformly at random, i.e., with probability $1/n$.

As in the classical parking process, each car enters the street from their preferred spot and directly parks there if possible. However, if a car's preferred spot is occupied, that driver tosses a weighted coin with probability p of landing heads. If the coin toss is heads, the driver proceeds to the end of the street, either parking at the next available spot, or exiting the street if none are available. But, if the coin toss is tails, the driver proceeds in reverse, toward the start of the street. Notice that $p = 1$ in the classical process, so it may be viewed as a special case of the Coin Problem. We henceforth refer to p as the "forward probability."

Then, given this new parking protocol, for a one-way street with n spots, what is the probability that a preference vector $\alpha \in \text{PF}_n$, chosen uniformly at random, allows all n cars to park on the street? Recall that under the classical parking protocol, such an answer is easy to obtain via the pigeonhole principle: $\alpha \in [n]^n$ allows all n cars to park if and only if

$$\#\{k : a_k \leq i\} \geq i, \quad \text{for all } i = 1, \dots, n, \quad (2.1)$$

and so in particular the probability of a preference vector being a parking function is invariant under the action of the symmetric group \mathfrak{S}_n by permuting cars. This nice and

²While there is no official paper written by Pollak on the topic, he has been credited for the proof by multiple authors, including Foata and Riordan [5].

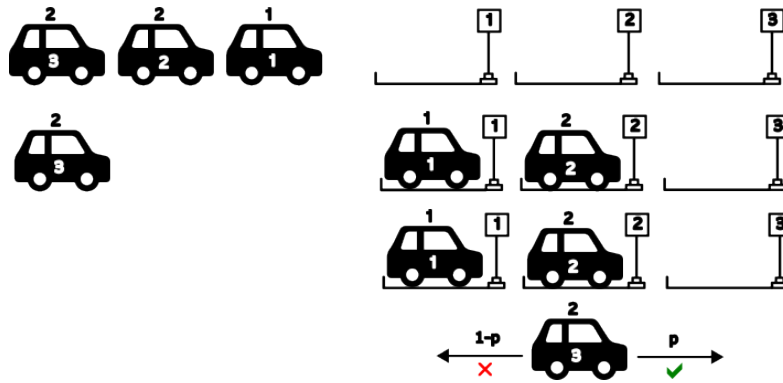


Figure 2: Parking scheme for the preference vector $(1, 2, 2)$. The number at the top of the car represents its preference. With probability p , all cars can park.

simple permutation symmetry however breaks under the probabilistic parking protocol, adding more delicacy to the answer, as we explain in the example below.

Example 10 (Lack of permutation symmetry under the probabilistic parking protocol). The preference vector $\alpha = (1, 2, 2) \in [3]^3$ is a parking function, or *parks*, with probability p . Cars numbered 1 and 2 will directly park at their preferred spot, whereas car 3 can only park at spot 3 as it is the only remaining available spot. This can happen only when the biased coin flips to heads, which occurs with probability p . The reader might find it instructive to check Figure 2 for an illustration of the above key parking steps. On the other hand, consider now the preference vector $\alpha' = (2, 2, 1)$, which is a permutation of $\alpha = (1, 2, 2)$. Car 1 will park at spot 2 directly. But since the desired spot for car 2 is now taken, instead of directly parking at spot 2, car 2 will have two choices. Either car 2 moves forward to park at spot 3, leaving spot 1 open for car 3, which happens to be car 3's desired spot. This parking situation happens with probability p . Or car 2 moves backwards to park at spot 1, which coincides with the desired spot for car 3, forcing car 3 to move forward to park when it enters the street. This parking situation happens with probability $(1 - p)p$. Altogether, α' is a parking function with probability $p + (1 - p)p$.

Given the preference vector $\alpha \in [3]^3$, Table 1 gives the probability that all three cars park under the probabilistic parking protocol. We note that each preference vector $\alpha \in [3]^3$ has probability $1/3^3 = 1/27$ of being chosen, and the sum of the probabilities in the table add up to $(3 + 1)^{3-1} = 16$. This result is especially interesting because it is independent of p , and suggests that the probability that a random preference vector $\alpha \in \text{PF}_n$ is a parking function is identical to that in the classical process (see Theorem 9).

$\alpha \in [3]^3$	(1, 1, 1)	(1, 1, 2)	(1, 1, 3)	(1, 2, 1)	(1, 2, 2)	(1, 2, 3)	(1, 3, 1)
$P(\alpha \in \text{PF}_3 \alpha)$	p^2	p^2	p	p	p	1	p
$\alpha \in [3]^3$	(2, 1, 1)	(2, 1, 2)	(2, 1, 3)	(2, 2, 1)	(2, 2, 2)	(2, 2, 3)	(2, 3, 1)
$P(\alpha \in \text{PF}_3 \alpha)$	p	p	1	$p + (1 - p)p$	$2p(1 - p)$	$p(1 - p) + (1 - p)$	1
$\alpha \in [3]^3$	(3, 1, 1)	(3, 1, 2)	(3, 1, 3)	(3, 2, 1)	(3, 2, 2)	(3, 2, 3)	(3, 3, 1)
$P(\alpha \in \text{PF}_3 \alpha)$	p	1	$1 - p$	1	$1 - p$	$1 - p$	$1 - p$
$\alpha \in [3]^3$	(1, 3, 2)	(1, 3, 3)	(2, 3, 2)	(2, 3, 3)	(3, 3, 2)	(3, 3, 3)	
$P(\alpha \in \text{PF}_3 \alpha)$	1	$1 - p$	$1 - p$	$1 - p$	$(1 - p)^2$	$(1 - p)^2$	

Table 1: All preference vectors for $n = 3$ and their probability of being a parking function.

2.2 Invariance to the forward probability

In this section, we generalize the results of Example 10 and Table 1, proving that the probability that a random preference vector is a parking function is identical for all forward probabilities $p \in [0, 1]$.

Theorem 1. Consider the preference vector $\alpha \in [n]^n$ chosen uniformly at random. Then

$$P(\alpha \in \text{PF}_n | \alpha \in [n]^n) = \frac{(n+1)^{n-1}}{n^n}. \quad (2.2)$$

In particular, despite the forward probability p , the probability that a random preference vector in $[n]^n$ is a parking function is identical to the classical case, in which cars may only search in the forward direction.

Similar to Pollak's proof for counting classical parking functions, as cited in Theorem 9, we consider the scenario in which n cars park on a circle with spots $[n+1]$ arranged clockwise, where each car may prefer any spot in $[n+1]$. Then, there are $(n+1)^n$ equally likely preference vectors. Our probabilistic parking protocol for one-way parking may be interpreted similarly in this circular parking situation. If a car arrives at its preferred spot and finds it unoccupied it parks there. If instead the spot is occupied, then the driver tosses the biased coin. If the coin lands on heads, with probability p , the driver continues moving clockwise in the street. However, if the coin lands on tails, with probability $1-p$, the car moves counterclockwise and tries to find an unoccupied parking spot.

For convenience, we define the following set.

Definition 11. For all $i \in [n+1]$, let

$$S_i = \{\alpha \in [n+1]^n : \text{spot } i \text{ is vacant after all } n \text{ cars have parked}\}.$$

In the classical case, any preference vector in S_{n+1} must also be a parking function. This is because the forward probability is $p = 1$, so cars may only search in the forward direction. Conversely, any preference vector which is a parking function must belong in S_{n+1} , because no car may be assigned to spot $(n+1)$. However, this relationship between S_{n+1} and PF_n is not limited to the classical case.

Lemma 12. *Given arbitrary preference vector $\alpha \in [n+1]^n$, then $\alpha \in \text{PF}_n$ if and only if $\alpha \in S_{n+1}$.*

Proof. In the forward direction, if $\alpha \in \text{PF}_n$, then none of the n cars park in spot $(n+1)$, which implies that $\alpha \in S_{n+1}$.

In the backward direction, if $\alpha \in S_{n+1}$, then none of the n cars prefer spot $(n+1)$. For contradiction, suppose some car i preferred spot $(n+1)$. Then car i either parks in spot $(n+1)$ directly, or it cannot, because that spot is already occupied. In either case, some car parks in spot $(n+1)$, and therefore $\alpha \notin S_{n+1}$. That is, all n cars prefer some spot from the original one-way street. And, because $\alpha \in S_{n+1}$, all cars are able to fit on the one-way street. Therefore, $\alpha \in \text{PF}_n$. \square

Next, recall that in Table 1, we saw that p did not affect the final probability $P(\alpha \in \text{PF}_3 | \alpha \in [3]^3)$. This was surprising because p appeared in the probabilities corresponding to all cars parking on the street, as shown in Example 10. This motivates analysis via conditional probability.

Definition 13. Given an arbitrary preference vector $\alpha \in [n + 1]^n$, let the conditional probability mass function $f_\alpha : [n + 1] \rightarrow [0, 1]$ denote the probability that $\alpha \in S_i$. That is, $f_\alpha(i) = P(\alpha \in S_i | \alpha)$.

In Definition 11, S_i denotes the event that all n cars have parked on the circular street with $n + 1$ spots, only leaving spot i vacant. The conditional sample space for some $\alpha \in [n + 1]^n$ is thus partitioned into the disjoint events $\{\alpha \in S_1, \alpha \in S_2, \dots, \alpha \in S_{n+1}\}$. This implies that $\sum_{i=1}^{n+1} f_\alpha(i) = 1$ by Definition 13.

Lemma 14. Consider any two preference vectors $\alpha, \alpha^* \in [n + 1]^n$ satisfying $\alpha^* = \alpha + (k * 1) \pmod{(n + 1)}$, for some $k \in [n]$. Let $j = (i + k) \pmod{(n + 1)}$. Then $f_\alpha(i) = f_{\alpha^*}(j)$.

Proof. This follows from the fact that the preference vectors themselves are congruent modulo $(n + 1)$. In particular, α^* is derived by adding $k \pmod{(n + 1)}$ to each preference in α , so the probability that $\alpha \in S_i$ is equivalent to the probability that $\alpha^* \in S_j$. \square

Definition 15. Let L_a denote the set of all preference vectors which have leading preference $a \in [n + 1]$. That is, $L_a = \{\alpha \in [n + 1]^n : a_1 = a\}$.

Remark 16. Notice that $|L_a| = (n + 1)^n / (n + 1) = (n + 1)^{n-1}$, where we divide by $(n + 1)$ because the first preference is fixed to be a .

Definition 17. For all cars $i \in [n + 1]$ and leading preferences $a \in [n + 1]$, let $F_a \in [0, n + 1]^{n+1}$ denote a length- $(n + 1)$ vector satisfying

$$F_a^{(i)} = \sum_{\forall \alpha \in L_a} f_\alpha(i). \tag{2.3}$$

Intuitively, if we simulated the parking procedure for all preference vectors $\alpha \in L_a$, then $F_a^{(i)}$ denotes the expected number of preference vectors which have leading preference a and leave spot i vacant. Notice that $F_a^{(a)} = 0$, because all $\alpha \in L_a$ satisfy $f_\alpha(a) = 0$. i.e., spot a will always be occupied by the leading car with preference a . Applying Definition 17, it follows from Lemma 14 that $F_a^{(i)} = F_b^{(j)}$, for all $a, b, i, j \in [n + 1]$ satisfying $b - a \equiv j - i \pmod{(n + 1)}$.

Next, purely as a visual aid, we construct a $(n + 1) \times (n + 1)$ matrix \mathbf{M} , where entry $\mathbf{M}_{a,i} = F_a^{(i)}$. Then $\mathbf{M}_{a,i} = \mathbf{M}_{a+1,i+1} = \dots = \mathbf{M}_{a+n,i+n}$, where the subscripts are modulo $(n + 1)$. In other words, each row in \mathbf{M} is the same as the previous row, but with entries shifted one cell to the right, modulo $(n + 1)$. That is,

$$\mathbf{M} = \begin{pmatrix} F_1^{(1)} & F_1^{(2)} & \dots & F_1^{(n+1)} \\ F_2^{(1)} & F_2^{(2)} & \dots & F_2^{(n+1)} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n+1}^{(1)} & F_{n+1}^{(2)} & \dots & F_{n+1}^{(n+1)} \end{pmatrix} = \begin{pmatrix} F_1^{(1)} & F_1^{(2)} & \dots & F_1^{(n+1)} \\ F_1^{(n+1)} & F_1^{(1)} & \dots & F_1^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ F_1^{(2)} & F_1^{(3)} & \dots & F_1^{(1)} \end{pmatrix}. \tag{2.4}$$

Equation (2.4) highlights the underlying symmetry via partitioning preference vectors based on leading preference. Note that while the simplified result is written in terms of entries from the first row, i.e., with subscript $a = 1$, similar equivalence holds for any

leading preference $a \in [n + 1]$. In particular, every column and row has the same sum. That is,

$$\sum_{a=1}^{n+1} F_a^{(1)} = \dots = \sum_{a=1}^{n+1} F_a^{(n+1)} = \sum_{i=1}^{n+1} F_1^{(i)} = \dots = \sum_{i=1}^{n+1} F_{n+1}^{(i)}. \quad (2.5)$$

This symmetry helps us compute the following probability:

$$\begin{aligned} P(\alpha \in S_{n+1}) &= \sum_{\forall \alpha} P(\alpha \in S_{n+1} | \alpha) P(\alpha) && \text{(By law of total probability)} \\ &= \sum_{a=1}^{n+1} \sum_{\forall \alpha \in L_a} P(\alpha \in S_{n+1} | \alpha) \frac{1}{(n+1)^n}. \end{aligned} \quad (2.6)$$

Notice that we are partitioning the sample space of all $\alpha \in [n + 1]^n$ into the disjoint sets L_1, L_2, \dots, L_{n+1} . And, since each α is chosen uniformly at random, we have that $P(\alpha) = 1/(n + 1)^n$. We now complete the computation:

$$\begin{aligned} P(\alpha \in S_{n+1}) &= \frac{1}{(n+1)^n} \sum_{a=1}^{n+1} \sum_{\forall \alpha \in L_a} f_\alpha(n+1) && \text{(By Definition 13)} \\ &= \frac{1}{(n+1)^n} \sum_{a=1}^{n+1} F_a^{(n+1)} && \text{(By Definition 17)} \\ &= \frac{1}{(n+1)^n} \sum_{i=1}^{n+1} F_1^{(i)} && \text{(By Equation (2.5))} \\ &= \frac{1}{(n+1)^n} \sum_{i=1}^{n+1} \sum_{\forall \alpha \in L_1} f_\alpha(i) && \text{(By Definition 17)} \quad (2.7) \\ &= \frac{1}{(n+1)^n} \sum_{\forall \alpha \in L_1} \sum_{i=1}^{n+1} f_\alpha(i) \\ &= \frac{1}{(n+1)^n} \sum_{\forall \alpha \in L_1} 1 && \text{(by Definition 13)} \\ &= \frac{(n+1)^{n-1}}{(n+1)^n}. && \text{(by Remark 16)} \end{aligned}$$

Recall that the relevant sample space for Theorem 1 is $[n]^n$, and not $[n + 1]^n$. That is, we want to find $P(\alpha \in S_{n+1} | \alpha \in [n]^n)$. Since this is a conditional probability, it is useful to consider the intersection of the two events. From Lemma 12, we know that $\alpha \in S_{n+1}$ implies that $\alpha \in [n]^n$. However, the converse is not necessarily true. As a counterexample, suppose all n cars prefer spot 1. Then, car 1 parks at spot 1, and car 2 may park at spot $(n + 1)$ with probability $1 - p$, since its preferred spot is occupied. Therefore, $P(\alpha \in S_{n+1} \cap \alpha \in [n]^n) = P(S_{n+1})$.

Now, we may complete the derivation of Theorem 1:

$$\begin{aligned}
 P(\alpha \in S_{n+1} | \alpha \in [n]^n) &= \frac{P(\alpha \in S_{n+1} \cap \alpha \in [n]^n)}{P(\alpha \in [n]^n)} \\
 &= \frac{P(\alpha \in S_{n+1})}{P(\alpha \in [n]^n)} \\
 &= \frac{(n+1)^{n-1} / (n+1)^n}{n^n / (n+1)^n} \\
 &= \frac{(n+1)^{n-1}}{n^n}.
 \end{aligned} \tag{2.8}$$

Therefore, the probability that an arbitrary preference vector in $[n]^n$ assigns all n cars to a unique spot on a one-way street with n spots is completely independent of the forward probability p .

2.3 Effect of the forward probability

In the previous section, we showed that the probability that a random preference vector $\alpha \in [n]^n$ is a parking function does not depend on the forward probability p . Thus, a natural question to ask is: what kind of parking statistics does depend on p ? In this section, we present one such parking statistic and explore its statistical properties.

Our investigations in this section rely on a combinatorial construction which we term a *parking function shuffle* and Abel's extension of the binomial theorem. These concepts were first discussed in Diaconis and Hicks [3] and later extended further in Kenyon and Yin [9]. Some asymptotic expansion formulas will also prove useful. We provide some background on these concepts first.

Definition 18. Let $1 \leq k \leq n$. Say that (a_1, \dots, a_{n-1}) is a parking function shuffle of the parking function $\alpha \in \text{PF}_{k-1}$ and $\beta \in \text{PF}_{n-k}$ if a_1, \dots, a_{n-1} is any permutation of the union of the two words α and $\beta + (k, \dots, k)$.

Example 19. Take $n = 8$ and $k = 4$. Take $\alpha = (2, 1, 2) \in \text{PF}_3$ and $\beta = (1, 2, 4, 3) \in \text{PF}_4$. Then $(2, \underline{5}, 2, \underline{8}, \underline{7}, 1, \underline{6})$ is a shuffle of the two words $(2, 1, 2)$ and $(5, 6, 8, 7)$.

Theorem 20 (Abel's extension of the binomial theorem, derived from Pitman [15] and Riordan [16]). *Let*

$$A_n(x, y; p, q) = \sum_{s=0}^n \binom{n}{s} (x+s)^{s+p} (y+n-s)^{n-s+q}. \tag{2.9}$$

Then

$$A_n(x, y; p, q) = A_n(y, x; q, p), \tag{2.10}$$

$$A_n(x, y; p, q) = A_{n-1}(x, y+1; p, q+1) + A_{n-1}(x+1, y; p+1, q), \tag{2.11}$$

$$A_n(x, y; p, q) = \sum_{s=0}^n \binom{n}{s} s! (x+s) A_{n-s}(x+s, y; p-1, q). \tag{2.12}$$

Moreover, the following special instances hold via the basic recurrences listed above:

$$A_n(x, y; -1, -1) = (x^{-1} + y^{-1})(x + y + n)^{n-1}, \quad (2.13)$$

$$A_n(x, y; -1, 0) = x^{-1}(x + y + n)^n, \quad (2.14)$$

$$A_n(x, y; -1, 1) = x^{-1} \sum_{s=0}^n \binom{n}{s} (x + y + n)^s (y + n - s)(n - s)!, \quad (2.15)$$

$$A_n(x, y; 0, 0) = \sum_{s=0}^n \binom{n}{s} (x + y + n)^s (n - s)!. \quad (2.16)$$

Lemma 21. Let X_1, X_2, \dots be iid Poisson(1) random variables. Then

$$P(X_1 + \dots + X_n \leq n) = \frac{1}{2} + \frac{2}{3} \frac{1}{\sqrt{2\pi n}} + o\left(\frac{1}{\sqrt{n}}\right). \quad (2.17)$$

Proof. We apply the continuity-corrected Edgeworth expansion as in Esseen [4] and Kollasa and McCullagh [11], so that

$$\begin{aligned} P(X_1 + \dots + X_n \leq n + x\sqrt{n}) &= P\left(\frac{X_1 + \dots + X_n - n}{\sqrt{n}} \leq x\right) \\ &= \Phi(0) + \frac{\exp(-x^2/2)}{\sqrt{2\pi n}} \left(\frac{\mu_3}{6\sigma^3}(1 - x^2) + \frac{1}{\sigma}D(x\sigma\sqrt{n})\right) + o\left(\frac{1}{\sqrt{n}}\right), \end{aligned} \quad (2.18)$$

where σ is the standard deviation and μ_3 is the third central moment of X_1 , $\Phi(x)$ is the distribution function of the standard normal, and $D(x) = [x] - x + 1/2$ is a periodic function that addresses the discontinuity in the distribution function of the discrete random variable $(X_1 + \dots + X_n - n)/\sqrt{n}$. \square

We are now ready to show that the distribution of a_n , the parking preference of the last car, depends on p and give explicit formulas.

Theorem 3. Consider the preference vector $\alpha \in [n]^n$, chosen uniformly at random. Then given that $\alpha \in \text{PF}_n$,

$$\begin{aligned} P(a_n = j | \alpha \in \text{PF}_n) &= \frac{2}{n+1} - \frac{1}{(n+1)^{n-1}} \left[p \sum_{s=n-j+1}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \right. \\ &\quad \left. + (1-p) \sum_{s=0}^{n-j-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \right], \end{aligned} \quad (2.19)$$

where a_n denotes the parking preference of the last car.

Remark 22. Note the parking symmetry as observed in the introduction: $P(a_n = j | \alpha \in \text{PF}_n)$ under protocol with parameter p equals $P(a_n = n+1-j | \alpha \in \text{PF}_n)$ under protocol with parameter $1-p$. See Figure 1 and Table 1.

Proof of Theorem 3. Cars $1, \dots, n-1$ have all parked along the one-way street before the n th car enters, leaving only one open spot k for the n th car to park. Since a car cannot jump over an empty spot, the parking protocol implies that (a_1, \dots, a_{n-1}) is a

parking function shuffle of $\alpha \in \text{PF}_{k-1}$ and $\beta \in \text{PF}_{n-k}$, and α and β do not interact with each other. This open spot k could be either the same as j , the preference of the last car, in which case the car parks directly. Or, k could be bigger than or less than j , in which case whether the last car parks or not depends on the outcome of the biased coin flip, as it will dictate the car to go forward or backward. Using Theorem 1, we have $\mathbb{P}(a_n = j | \alpha \in \text{PF}_n)$

$$= \frac{1}{(n+1)^{n-1}} \left[(1-p) \sum_{k=1}^{j-1} \binom{n-1}{n-k} k^{k-2} (n-k+1)^{n-k-1} + \binom{n-1}{n-j} j^{j-2} (n-j+1)^{n-j-1} + p \sum_{k=j+1}^n \binom{n-1}{n-k} k^{k-2} (n-k+1)^{n-k-1} \right] \quad (2.20)$$

$$= \frac{1}{(n+1)^{n-1}} \left[\sum_{k=1}^n \binom{n-1}{n-k} k^{k-2} (n-k+1)^{n-k-1} - p \sum_{k=1}^{j-1} \binom{n-1}{n-k} k^{k-2} (n-k+1)^{n-k-1} - (1-p) \sum_{k=j+1}^n \binom{n-1}{n-k} k^{k-2} (n-k+1)^{n-k-1} \right] \quad (2.21)$$

$$= \frac{1}{(n+1)^{n-1}} \left[\sum_{s=0}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} - p \sum_{s=n-j+1}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} - (1-p) \sum_{s=0}^{n-j-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \right] \quad (2.22)$$

$$= \frac{2}{n+1} - \frac{1}{(n+1)^{n-1}} \left[p \sum_{s=n-j+1}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} + (1-p) \sum_{s=0}^{n-j-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \right]. \quad (2.23)$$

Here the binomial coefficient $\binom{n-1}{n-k}$ accounts for the shuffling of the length $k-1$ subsequence and the length $n-k$ subsequence of parking preferences among the 1st through the $(n-1)$ st cars. We used a simple change of variables $s = n-k$ going from (2.21) to (2.22). From (2.22) to (2.23), we applied Abel's binomial identity (2.13): $A_{n-1}(1, 1; -1, -1) = 2(n+1)^{n-2}$. \square

Theorem 4. Take n large. For preference vector $\alpha \in [n]^n$ chosen uniformly at random, we have

$$\mathbb{E}(a_n | \alpha \in \text{PF}_n) = \frac{n+1}{2} - (2p-1) \left[\frac{\sqrt{2\pi}}{4} n^{1/2} - \frac{7}{6} \right] + o(1). \quad (2.24)$$

Remark 23. Note the parking symmetry as observed in the introduction: the sum of $\mathbb{E}(a_n | \alpha \in \text{PF}_n)$ under protocol with parameter p and $\mathbb{E}(a_n | \alpha \in \text{PF}_n)$ under protocol with parameter $1-p$ is $n+1$. The lower order correction terms from $(n+1)/2$ for $\mathbb{E}(a_n | \alpha \in \text{PF}_n)$ vanish completely under protocol with parameter $p = 1/2$, and $\mathbb{E}(a_n | \alpha \in \text{PF}_n) = (n+1)/2$ exactly. See Figure 3.

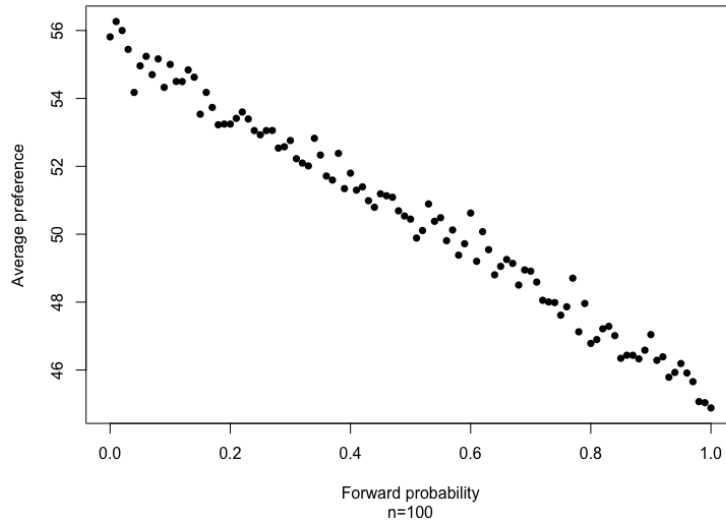


Figure 3: The conditional expectation of a_n (parking preference of the last car) given $\alpha \in \text{PF}_n$ under the probabilistic parking protocol with parameter p . Here size $n = 100$ and 100,000 samples are drawn uniformly at random.

Proof of Theorem 4. By Theorem 3,

$$\begin{aligned}
 \mathbb{E}(a_n | \alpha \in \text{PF}_n) &= \sum_{j=1}^n j P(a_n = j | \alpha \in \text{PF}_n) \\
 &= n - \frac{1}{(n+1)^{n-1}} \sum_{j=1}^n j \left[p \sum_{s=n-j+1}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \right. \\
 &\quad \left. + (1-p) \sum_{s=0}^{n-j-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \right] \\
 &= n - \frac{1}{(n+1)^{n-1}} \sum_{s=0}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \left[p \sum_{j=n-s+1}^n j + (1-p) \sum_{j=1}^{n-s-1} j \right] \\
 &= n - \frac{1}{2(n+1)^{n-1}} \left[\sum_{s=0}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \cdot \right. \\
 &\quad \left. \cdot [(1-p)(n-s)^2 + p(n-s)(s+1) + p(n+1)(s+1) - (n-s) - p(n+1)] \right] \tag{2.25} \\
 &= n - \frac{1}{2(n+1)^{n-1}} \left[(1-p)A_{n-1}(1, 1; -1, 1) + pA_{n-1}(1, 1; 0, 0) \right. \\
 &\quad \left. + (pn + p - 1)A_{n-1}(1, 1; -1, 0) - p(n+1)A_{n-1}(1, 1; -1, -1) \right], \tag{2.26}
 \end{aligned}$$

where Abel's binomial theorem is used multiple times from (2.25) to (2.26).

We proceed to estimate (2.26) asymptotically. Using (2.13), $A_{n-1}(1, 1; -1, -1) =$

$2(n+1)^{n-2}$ and using (2.14), $A_{n-1}(1, 1; -1, 0) = (n+1)^{n-1}$. By (2.15) and (2.16),

$$\begin{aligned} & (1-p)A_{n-1}(1, 1; -1, 1) + pA_{n-1}(1, 1; 0, 0) \\ &= (1-p) \sum_{s=0}^{n-1} \binom{n-1}{s} (n+1)^s (n-s)(n-1-s)! + p \sum_{s=0}^{n-1} \binom{n-1}{s} (n+1)^s (n-1-s)! \\ &= (n-np+p)(n+1)^{n-1} + (2p-1)(n-1)! \sum_{s=0}^{n-2} \frac{(n+1)^s}{s!}. \end{aligned} \tag{2.27}$$

From Stirling's formula,

$$(n-1)! \sim \sqrt{2\pi(n-1)} e^{-(n-1)} (n-1)^{n-1} \left[1 + \frac{1}{12(n-1)} \right]. \tag{2.28}$$

We also recognize that

$$e^{-(n+1)} \sum_{s=0}^{n-2} \frac{(n+1)^s}{s!} \tag{2.29}$$

equals the probability that the sum of $n+1$ iid Poisson(1) random variables is less than or equal to $n-2$, and is asymptotic to

$$\frac{1}{2} + \frac{2}{3} \frac{1}{\sqrt{2\pi(n+1)}} - 3e^{-(n+1)} \frac{(n+1)^{n-1}}{(n-1)!} \tag{2.30}$$

from the Edgeworth expansion in Lemma 21. Combining the above,

$$\begin{aligned} \mathbb{E}(a_n | \alpha \in \text{PF}_n) &= \frac{n+1}{2} - (2p-1) \frac{(n-1)!}{2(n+1)^{n-1}} \sum_{s=0}^{n-2} \frac{(n+1)^s}{s!} \\ &= \frac{n+1}{2} - (2p-1) \left[\frac{\sqrt{2\pi}}{4} n^{1/2} - \frac{7}{6} \right] + o(1). \end{aligned} \tag{2.31}$$

□

2.4 Why the forward probability $p = 1/2$ is special

The conditional distribution of a_n for the classical model is close to a uniform distribution over $[n]$ in a precise sense given by the total variation distance. Building upon the results in the last section, in this section we show that the presence of p has a consequential impact on this rate of convergence to uniform. Specifically, for $p = 1/2$ the symmetry provided by p plays an important role of speeding up the rate of convergence. This explains why in the histogram of Figure 1 for $p = 1/2$ we already see a distribution that looks uniform, whereas for other values of p the presence of more mass on small or large values of $[n]$ is quite apparent.

In order to better discuss our results in this section, let us introduce a few definitions and set our notation. For two distributions P and Q over $[n]$, their total variation distance is given by

$$\|P - Q\|_{\text{TV}} := \frac{1}{2} \sum_{j=1}^n |P(j) - Q(j)|. \tag{2.32}$$

For $p \in [0, 1]$, for notational convenience, let $Q_{n,p}$ be the conditional probability of a_n given $\alpha \in PF_n$ under the probabilistic parking protocol with parameter p . That is, $Q_{n,p}$ is the distribution over $[n]$ established in Theorem 3: $Q_{n,p}(j) = P(a_n = j | \alpha \in PF_n)$. Additionally, let Uni_n be the uniform distribution over $[n]$: $\text{Uni}_n(j) = 1/n$ for all $j \in [n]$.

In [3] Diaconis and Hicks showed that $\|Q_{n,1/2} - \text{Uni}_n\|_{\text{TV}}$ goes to zero as n goes to infinity. They also conjectured that the rate of convergence is $O(1/\sqrt{n})$. In [1] Bellin confirmed that their conjecture is indeed true.

Theorem 5. Let $Q_{n,p}(\cdot)$ be $P(a_n = \cdot | \alpha \in PF_n)$ under the probabilistic parking protocol with parameter p . For $p = 1/2$, the following bound holds

$$\|Q_{n,1/2} - \text{Uni}_n\|_{\text{TV}} = \Theta\left(\frac{1}{n}\right). \quad (2.33)$$

Proof. We use the definition of total variation distance together with Theorem 3 as follows. For $j \in [n]$, Theorem 3 gives us

$$Q_{n,1/2}(j) = \frac{2}{n+1} - \frac{1}{(n+1)^{n-1}} \left[\frac{1}{2} \sum_{s=n-j+1}^{n-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} + \frac{1}{2} \sum_{s=0}^{n-j-1} \binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1} \right]. \quad (2.34)$$

Adding and subtracting $\binom{n-1}{n-j} j^{j-2} (n-j+1)^{n-j-1} / 2(n+1)^{n-1}$ on the right-hand side of (2.34), and recalling that $A_{n-1}(1, 1; -1, -1) = 2(n+1)^{n-2}$ yields

$$\begin{aligned} Q_{n,1/2}(j) &= \frac{2}{n+1} - \frac{1}{(n+1)^{n-1}} \left[\frac{1}{2} A_{n-1}(1, 1; -1, -1) - \frac{1}{2} \binom{n-1}{n-j} j^{j-2} (n-j+1)^{n-j-1} \right] \\ &= \frac{1}{n+1} + \frac{\binom{n-1}{n-j} j^{j-2} (n-j+1)^{n-j-1}}{2(n+1)^{n-1}}. \end{aligned} \quad (2.35)$$

Thus, substituting the identity in (2.35) into (2.32), we obtain

$$\begin{aligned} \|Q_{n,1/2} - \text{Uni}_n\|_{\text{TV}} &= \frac{1}{2} \sum_{j=1}^n \left| \frac{1}{n+1} + \frac{\binom{n-1}{n-j} j^{j-2} (n-j+1)^{n-j-1}}{2(n+1)^{n-1}} - \frac{1}{n} \right| \\ &= \frac{1}{2} \sum_{j=1}^n \left| \frac{\binom{n-1}{n-j} j^{j-2} (n-j+1)^{n-j-1}}{2(n+1)^{n-1}} - \frac{1}{n(n+1)} \right| \\ &= \frac{1}{2} \sum_{s=0}^{n-1} \left| \frac{\binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1}}{2(n+1)^{n-1}} - \frac{1}{n(n+1)} \right| \\ &\leq \frac{1}{2} \sum_{s=0}^{n-1} \left[\frac{\binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1}}{2(n+1)^{n-1}} + \frac{1}{n(n+1)} \right] \\ &= \frac{A_{n-1}(1, 1; -1, -1)}{4(n+1)^{n-1}} + \frac{1}{2(n+1)} = \frac{1}{n+1} \sim \frac{1}{n}. \end{aligned} \quad (2.36)$$

To obtain a lower bound of the same order, notice that

$$\begin{aligned} \|Q_{n,1/2} - \text{Uni}_n\|_{\text{TV}} &= \frac{1}{2} \sum_{s=0}^{n-1} \left| \frac{\binom{n-1}{s} (n-s)^{n-s-2} (s+1)^{s-1}}{2(n+1)^{n-1}} - \frac{1}{n(n+1)} \right| \\ &\geq \frac{1}{2} \left| \frac{\binom{n-1}{n-1} (1)^{-1} n^{n-2}}{2(n+1)^{n-1}} - \frac{1}{n(n+1)} \right| \\ &\geq \frac{n^{n-2}}{4(n+1)^{n-1}} - \frac{1}{n(n+1)} \sim \frac{1}{4e} \frac{1}{n}, \end{aligned} \tag{2.37}$$

which proves that $\|Q_{n,1/2} - \text{Uni}_n\|_{\text{TV}} = \Theta(1/n)$. \square

Next we show that indeed the case $p = 1/2$ is special when considering the rate of convergence of $\|Q_{n,p} - \text{Uni}_n\|_{\text{TV}}$ to zero. We show that for any p other than $p = 1/2$, the sequence $(\|Q_{n,p} - \text{Uni}_n\|_{\text{TV}})_{n \geq 1}$ cannot converge to zero much faster than the classical model given by $p = 1$.

Proposition 24. *For any $p \neq 1/2$, the following bound holds*

$$|2p - 1| \|Q_{n,1} - \text{Uni}_n\|_{\text{TV}} \leq \|Q_{n,p} - \text{Uni}_n\|_{\text{TV}} \leq \|Q_{n,1} - \text{Uni}_n\|_{\text{TV}}. \tag{2.38}$$

Remark 25. Proposition 24 shows that the rate of convergence is of the same order for any $p \neq 1/2$. This important fact will be used to establish the exact rate of convergence for generic $p \neq 1/2$ in Theorem 6.

Proof of Proposition 24. Recall that by Theorem 3, for any $p \in [0, 1]$ and $n \in \mathbb{N}$, $Q_{n,p}$ is the convex combination of $Q_{n,1}$ and $Q_{n,0}$. More precisely,

$$Q_{n,p} = pQ_{n,1} + (1-p)Q_{n,0}. \tag{2.39}$$

Moreover, since $Q_{n,1}(j) = Q_{n,0}(n+1-j)$ (see Remark 22), it follows that

$$\|Q_{n,1} - \text{Uni}_n\|_{\text{TV}} = \|Q_{n,0} - \text{Uni}_n\|_{\text{TV}}. \tag{2.40}$$

Let us assume $p > 1/2$ for the moment. Via the above symmetry, the reader will notice that the case for $p < 1/2$ follows a similar argument. From the definition of the total variation distance combined with (2.39) and that $|a+b| \geq |a| - |b|$, we obtain

$$\begin{aligned} \|Q_{n,p} - \text{Uni}_n\|_{\text{TV}} &= \|pQ_{n,1} + (1-p)Q_{n,0} - \text{Uni}_n\|_{\text{TV}} \\ &= \frac{1}{2} \sum_{j=1}^n \left| pQ_{n,1}(j) + (1-p)Q_{n,0}(j) - \frac{1}{n} \right| \\ &= \frac{1}{2} \sum_{j=1}^n \left| pQ_{n,1}(j) - \frac{p}{n} + (1-p)Q_{n,0}(j) - \frac{(1-p)}{n} \right| \\ &\geq \frac{1}{2} \sum_{j=1}^n \left| pQ_{n,1}(j) - \frac{p}{n} \right| - \frac{1}{2} \sum_{j=1}^n \left| (1-p)Q_{n,0}(j) - \frac{(1-p)}{n} \right| \\ &= p \|Q_{n,1} - \text{Uni}_n\|_{\text{TV}} - (1-p) \|Q_{n,0} - \text{Uni}_n\|_{\text{TV}} \\ &= (2p - 1) \|Q_{n,1} - \text{Uni}_n\|_{\text{TV}}. \end{aligned} \tag{2.41}$$

For the upper bound, we instead apply the triangle inequality to the third line of (2.41):

$$\begin{aligned} \|Q_{n,p} - \text{Uni}_n\|_{\text{TV}} &= \frac{1}{2} \sum_{j=1}^n \left| pQ_{n,1}(j) - \frac{p}{n} + (1-p)Q_{n,0}(j) - \frac{(1-p)}{n} \right| \\ &\leq \frac{1}{2} \sum_{j=1}^n \left| pQ_{n,1}(j) - \frac{p}{n} \right| + \frac{1}{2} \sum_{j=1}^n \left| (1-p)Q_{n,0}(j) - \frac{(1-p)}{n} \right| \\ &= p\|Q_{n,1} - \text{Uni}_n\|_{\text{TV}} + (1-p)\|Q_{n,0} - \text{Uni}_n\|_{\text{TV}} \\ &= \|Q_{n,1} - \text{Uni}_n\|_{\text{TV}}. \quad \square \end{aligned}$$

For the next result we use the following alternative definition for the total variation distance.

Proposition 26 (Proposition 4.5 in Levin and Peres [13]). *Let P and Q be two distributions over $[n]$, then*

$$\|P - Q\|_{\text{TV}} = \frac{1}{2} \sup \left\{ \sum_{j=1}^n f(j)P(j) - \sum_{j=1}^n f(j)Q(j) : \max_j |f(j)| \leq 1 \right\}. \quad (2.42)$$

With the alternative definition in hand, we can show the right order of the rate of convergence of $\|Q_{n,p} - \text{Uni}_n\|_{\text{TV}}$ for $p \neq 1/2$ which is our next result.

Theorem 6. Let $Q_{n,p}(\cdot)$ be $\mathbb{P}(a_n = \cdot \mid \alpha \in \text{PF}_n)$ under the probabilistic parking protocol with parameter p . For $p \neq 1/2$, we have that

$$\|Q_{n,p} - \text{Uni}_n\|_{\text{TV}} = \Theta \left(\frac{1}{\sqrt{n}} \right). \quad (2.43)$$

Remark 27. The above result for generic p is in sharp contrast with the behavior of the model at $p = 1/2$ for which the rate of convergence is $1/n$ by Theorem 5. Before we start the proof of the theorem, let us explain the flow of our argument. In Theorem 1 of [1], the author proved that for the classical model (corresponding to $p = 1$), the rate of convergence of $(\|Q_{n,1} - \text{Uni}_n\|_{\text{TV}})_{n \geq 1}$ is $O(1/\sqrt{n})$. This result together with the second inequality of Proposition 24 shows that in the probabilistic model under consideration, the rate of convergence is actually $O(1/\sqrt{n})$ for any $p \neq 1/2$. We then go one step further and establish that $1/\sqrt{n}$ is indeed the right order for the rate of convergence. Besides the first inequality of Proposition 24, a central ingredient in the proof is a clever choice of the test function f as in Equation (2.42).

Proof. For the upper bound, notice that by Proposition 24 the rate of convergence in the general case is bounded from above by the corresponding rate in the classical case $p = 1$. Then, taking $k = 1$ in Theorem 1 of [1], we obtain that

$$\|Q_{n,p} - \text{Uni}_n\|_{\text{TV}} = O \left(\frac{1}{\sqrt{n}} \right). \quad (2.44)$$

For the lower bound, define the function f over $[n]$ such that $f(j) = j/n$ for all $j \in [n]$. Notice that by the definition of f we have that

$$\sum_{j=1}^n f(j)\text{Uni}_n(j) = \frac{n+1}{2n}. \quad (2.45)$$

Whereas, by Theorem 4 we have

$$\sum_{j=1}^n f(j)Q_{n,p}(j) = \frac{\mathbb{E}(a_n|\alpha \in \text{PF}_n)}{n} = \frac{n+1}{2n} - (2p-1) \left[\frac{\sqrt{2\pi}}{4\sqrt{n}} - \frac{7}{6n} \right] + \frac{o(1)}{n}. \quad (2.46)$$

Thus, for $p = 1$, substitution (2.45) and (2.46) into the identity given by Proposition 26 we have that

$$\begin{aligned} \|Q_{n,1} - \text{Uni}_n\|_{\text{TV}} &\geq \frac{1}{2} \left[\sum_{j=1}^n f(j)\text{Uni}_n(j) - \sum_{j=1}^n f(j)Q_{n,p}(j) \right] \\ &= \frac{1}{2} \left[\frac{\sqrt{2\pi}}{4\sqrt{n}} - \frac{7}{6n} + \frac{o(1)}{n} \right]. \end{aligned} \quad (2.47)$$

Finally, by Proposition 24 we know that

$$\|Q_{n,p} - \text{Uni}_n\|_{\text{TV}} \geq |2p-1| \|Q_{n,1} - \text{Uni}_n\|_{\text{TV}} \sim |2p-1| \frac{\sqrt{2\pi}}{8} \frac{1}{\sqrt{n}}, \quad (2.48)$$

which proves the theorem. \square

2.5 A statistical interpretation

Let us contextualize the Coin Problem with a more formal statistical framework. We again consider the problem of parking n cars on a circle with $(n+1)$ spots. Suppose we have the random vector $\alpha = (a_1, a_2, \dots, a_n)$, where each a_i is drawn with replacement from $[n+1]$. That is, α denotes a random sample of parking preferences. Then, consider the following random variables:

- Let $Y_i \sim g(\alpha|a_1, \dots, a_{i-1}; p)$ be a random variable denoting the spot at which car i parks. Notice that this distribution is parameterized by the forward probability $p \in [0, 1]$, and furthermore, is conditional on previous preferences.
- Let T denote the vacant spot on the circle, after all n cars have parked, similar to Definition 11. Since $[n+1]$ denotes the spots on the circle, we have $T = \sum_{i=1}^{n+1} i - \sum_{j=1}^n Y_j$.
- Finally, let $Z = \mathbb{I}(\alpha \in S_{n+1}) = \mathbb{I}(T_i \neq n+1)$, where \mathbb{I} is the indicator function. That is, Z checks whether α is a parking function on the one-way street.

Reframing the results of Theorem 1, we have that $Z \sim \text{Bernoulli}(\theta)$, for $\theta = (n+1)^{n-1}/n^n$. Similar to Theorem 1, the sampling distribution of Z is independent of p . Then, if we applied the probabilistic parking protocol to each preference vector in $[n+1]^n$, treating each instance as an independent Bernoulli trial, we expect $(n+1)^{n-1}$ of those preference vectors to be parking functions. Finally, we emphasize that we precisely derived this Bernoulli distribution, without knowledge of the underlying conditional distribution $g(\cdot)$, which determines the cars' individual parking outcomes.

2.6 The Coin Problem for $m \leq n$ cars

Having formalized the symmetric properties of parking on a circle in Section 2.2, we now prove that Theorem 1 generalizes to the problem of parking $m \in [n]$ cars on a one-way street with n spots. Here, we slightly adjust the definition of a parking function to mean any preference vector in $[n]^m$ which assigns all m cars to a unique spot on the one-way street with n spots.

Theorem 2. For $m \in [n]$, consider the preference vector $\alpha \in [n]^m$, chosen uniformly at random. Then

$$P(\alpha \in \text{PF}_n | \alpha \in [n]^m) = \frac{(n+1-m)(n+1)^{m-1}}{n^m}. \quad (2.49)$$

Proof. With $m \leq n$ cars, it is possible that more than 1 spot is vacant after all m cars park on a circle with spots $[n+1]$. However, the general principles of symmetry, formalized by Lemma 14 and Equation (2.5), still apply. i.e., each preference vector in $[n+1]^m$ is modularly congruent to n other preference vectors in $[n+1]^m$, namely those which are obtained by shifting all m preferences by some $k \in [n]$. Then, the expected number of preference vectors in $[n+1]^m$ which leave spot i vacant is equivalent for all $i \in [n+1]$, despite the fact that there are many more possible combinations of vacant spots.

Next, by symmetry, each of the $n+1$ spots has equal probability of being vacant. Since there are always $n+1-m$ vacant spots after all m cars have parked, we know that the probability that an arbitrary spot in $[n+1]$ is vacant is $(n+1-m)/(n+1)$. Therefore, the expected number of preference vectors in $[n+1]^m$ which are parking functions is

$$\mathbb{E}[\#\{\alpha \in [n+1]^m \cap \alpha \in \text{PF}_n\}] = \frac{n+1-m}{n+1} \cdot (n+1)^m = (n+1-m)(n+1)^{m-1}. \quad (2.50)$$

But, generalizing the result of Lemma 12, we know that if a preference vector in $[n+1]^m$ is a parking function, then none of the m cars may prefer spot $n+1$. Therefore, our effective sample space is the set of preference vectors in $[n]^m$. Since we choose uniformly at random, we divide the expectation in (2.50) by n^m , the size of this sample space, to obtain the probability in question. That is,

$$P(\alpha \in \text{PF}_n | \alpha \in [n]^m) = \frac{(n+1-m)(n+1)^{m-1}}{n^m}.$$

□

3 Related combinatorial results

In this section, we present various combinatorial properties of parking functions.

3.1 Lucky and unlucky cars

Gessel and Seo [6] defined a car being “lucky” if it is able to park in its preferred spot. Similarly, we say that a car is “unlucky” if it cannot park in its preferred spot.

Theorem 7. Suppose n cars park on a circle with spots $[n+1]$. The expected number of preference vectors in $[n+1]^n$ containing $0 \leq k < n$ unlucky cars is

$$U_n(k) = (n+1)! \sum_{C_k + \{2, \dots, n\}} \prod_{i \in C_k} \frac{i-1}{(n+1) - (i-1)}, \quad (3.1)$$

where C_k is a size- k subset of the cars $\{2, \dots, n\}$, and each car $i \in C_k$ is unable to park at its preferred spot.

Proof. We start by recalling that the first car is always lucky, since it enters an empty street. Therefore, we will fix the leading preference as spot 1.

Let us first consider the case for $k = 0$, i.e., no cars are unlucky. Here, there is only one unique approach to choose 0 unlucky cars among the remaining set of cars $\{2, \dots, n\}$. For car 2 to be lucky, it must prefer one of the n vacant spots. Similarly, for car 3 to be lucky, it must prefer one of the $n - 1$ vacant spots. Then, in general, there are $n(n - 1) \cdots 2$ combinations of preferences which allow all cars to be lucky. Since the leading preference has been fixed at spot 1, we have $U_n(0) = (n + 1) \times n! = (n + 1)!$, where the $n + 1$ factor accounts for the $n + 1$ possibilities for the leading preference.

Next, still keeping the leading preference fixed at 1, let $C_k \vdash \{2, \dots, n\}$ denote an arbitrary combination of k cars designated as unlucky. When each car $i \in C_k$ enters the street, since it is unlucky, it must prefer one of the $i - 1$ occupied spots. That is, in the descending product $n!$, for each $i \in C_k$, we replace the term $(n + 1) - (i - 1)$ with $i - 1$. This is done algebraically by dividing $n!$ by $(n + 1) - (i - 1)$, then multiplying by $i - 1$. Summing over all possible combinations C_k yields the expected number of preference vectors in $[n + 1]^n$ containing exactly k unlucky cars and whose leading preference is 1. We then get $U_n(k)$ after multiplication by $n + 1$. \square

For completion, we also relate this to the problem of parking n cars on a circular street with spots $[n]$.

Corollary 28. *Suppose n cars park on a circular street with n spots. The expected number of preference vectors in $[n]^n$ containing $0 \leq k < n$ unlucky cars is*

$$U_n(k) = n! \sum_{C_k \vdash \{2, \dots, n\}} \prod_{i \in C_k} \frac{i - 1}{n - (i - 1)}. \quad (3.2)$$

Proof. The proof proceeds similarly to that for Theorem 7. Note that in the denominator we subtract 1, since each car may choose among n spots only. \square

Remark 29. This inner product is identical to the generating function derived by Khidr and El-Desouky [10], which is based on the Stirling numbers of the first kind. This generating function produces the triangular array documented under OEIS [A071208](#).

Our next result provides a combinatorial interpretation for $U_n(k)$ from Theorem 7 after proper normalization.

Proposition 30. *Take $k \in \{0, 1, \dots, n\}$ arbitrary and let $U_n(k)$ be as in the statement of Theorem 7. Then $(n + 1)^{-1}U_n(k)$ is the number of classical parking functions of length n with k unlucky cars.*

Proof. The proof follows from Theorem 7 together with the following two observations. First, notice that the property of having exactly k unlucky cars is invariant under cyclic shifts. In other words, a cyclic shift of a preference vector in $[n + 1]^n$ having exactly k unlucky cars yields another preference vector in $[n + 1]^n$ which again has exactly k unlucky cars. Second, also notice that there is exactly one cyclic shift of the preference vector that keeps spot $n + 1$ unoccupied.

These two observations imply that if we group the preference vectors into equivalence classes according to the cyclic shift, then inside each equivalence class there is exactly one preference vector that leaves spot $n + 1$ unoccupied. Thus, by Theorem 7, $U_n(k)/(n + 1)$ gives us the number of preference vectors in $[n + 1]^n$ leaving spot $n + 1$ unoccupied, which we recognize is the number of classical parking functions of length n with k unlucky cars. \square

3.2 A problem on polynomials

We now consider a generating polynomial and provide a parking function interpretation for sequence OEIS [A220884](#).

Theorem 8. Let $Q_n(q)$ be the generating polynomial $\prod_{k=2}^n [(n + 1 - k)q + k]$. Then

$$Q_n(q) = \sum_{k=0}^{n-1} \frac{1}{n+1} U_n(k) q^k. \quad (3.3)$$

Remark 31. The polynomial $Q_n(q)$ appears in a paper by Gessel and Seo as $P_n(1, q, 1)$, where the coefficient of q^k counts, among other things, the number of rooted forests on n vertices with k ascents (see Theorem 9.1 in [6]).

Proof. Recall that we fix the leading preference to be spot 1 on a circular street, since the first car is always lucky. Then, rewriting $\prod_{k=2}^n [(n + 1 - k)q + k] = \prod_{k=2}^n [(k - 1)q + (n + 1) - (k - 1)]$, we see that the two coefficients in the k -th factor denote the number of preferences which make car k unlucky and lucky, respectively. Therefore, when expanding the first $0 \leq j \leq n - 1$ factors, i.e., $\prod_{k=2}^{j+1} [(k - 1)q + (n + 1) - (k - 1)]$, and grouping together same-degree terms, each coefficient of q^i , for all $i \in \{0, \dots, j\}$, represents the expected number of *partial* preference vectors in $[n + 1]^{j+1}$ containing i unlucky cars. \square

Remark 32. In the left-aligned triangular array documented under OEIS [A220884](#), the entry in row i , column j denotes the expected number of preference vectors in $[i + 1]^i$ which have j unlucky cars, with the first car parking at spot 1 always. See Table 2.

$i \backslash j$	0	1	2	3	4	5
0	1					
1	1	0				
2	2	1	0			
3	6	8	2	0		
4	24	58	37	6	0	
5	120	444	504	204	24	0

Table 2: Triangular array under OEIS [A220884](#).

3.3 Weighted Pascal's triangle

The notion of partial preference vectors in the proof of Theorem 8 alludes to a recurrence relation which bears very strong resemblance to that for Pascal's triangle.

Definition 33. Let $E_n(i, k)$, with $0 \leq k < i \leq n$, denote the expected number of preference vectors in $[n + 1]^i$ containing k unlucky cars. Then

$$E_n(i, k) = ((n + 1) - (i - 1)) \cdot E_n(i - 1, k) + (i - 1) \cdot E_n(i - 1, k - 1). \quad (3.4)$$

We provide a sample calculation to show the mechanics of the recurrence relation of (3.4). For $n = 4$, $i = 3$ and $k = 1$:

$$E_4(3, 1) = (5 - (3 - 1)) \cdot E_4(2, 1) + (3 - 1) \cdot E_4(2, 0) = 3 \cdot 1 + 2 \cdot 4 = 3 + 8 = 11.$$

We refer to the triangular array produced by (3.4) as the "Weighted Pascal's Triangle" due to the weighted coefficients. An example of this triangle, for a length $i = 4$ sequence, is presented in Table 4.

Notice how (3.4) contrasts with the recurrence relation for Pascal's triangle:

$$T(n, d) = T(n - 1, d) + T(n - 1, d - 1). \quad (3.5)$$

Note that n in this case represents the row of Pascal's triangle, and d represents the column of the triangle. An example of such a triangle for $n = 4$ is presented in Table 3.

Notice that (3.4) reduces to (3.5), when both $n - i = 1$ and $i = 2$.

$n \backslash d$	0	1	2	3	4
0	1				
1	1	1			
2	1	2	1		
3	1	3	3	1	
4	1	4	6	4	1

$i \backslash k$	0	1	2	3	4
0	0				
1	1	0			
2	4	1	0		
3	12	11	2	0	
4	24	58	37	6	0

Table 3: Pascal's triangle for $0 \leq d \leq n \leq 4$ Table 4: Weighted Pascal's triangle for $n = 4$.

It is useful to think about two limiting cases when discussing Table 4. Observe the case of $i = 0$ and $k = 0$. The entry in the table/triangle is giving us the number of expected sequences with zero unlucky cars, when there are no cars parked: such a sequence does not exist because there are no parked cars on the street! Another interesting limiting case occurs in the last row with $i = n$. Notice that each cell in this row can be computed with Theorem 7. That is, $E_n(n, k) = U_n(k)$. For example, consider the case for $i = 4$ and $k = 0$. This entry tells us the expected number of sequences which have no unlucky cars, after four cars have parked, which is $U_4(0) = 4!$.

4 Future research

The combinatorial results derived in Section 3 connect to existing works which provide promising avenues for further analysis. For example, Diaconis and Hicks [3] cite the

generating function:

$$\sum_{\forall \alpha \in \text{PF}_n} q^{L(\alpha)} = q \prod_{i=1}^{n-1} [i + (n - i + 1)q], \quad (4.1)$$

where $L(\alpha)$ counts the number of lucky cars in the one-way classical parking function $\alpha \in \text{PF}_n$, chosen uniformly. From Proposition 30, $\frac{1}{n+1}U_n(k)$ counts the number of unlucky cars in the classical case, we thus have

$$\sum_{\forall \alpha \in \text{PF}_n} q^{L(\alpha)} = \sum_{k=0}^{n-1} \frac{1}{n+1} U_n(k) q^{n-k} = q^n Q_n \left(\frac{1}{q} \right), \quad (4.2)$$

where the polynomial Q_n is defined as in Theorem 8.

In [3] it was shown that the cumulative density function of $L(\alpha)$ is asymptotically normal, which implies that the cumulative density function of $\sum U_n(k)$ as investigated in Theorem 7 is also asymptotically normal. The authors are further interested in similarly researching the distribution of the generating function for $U_n(k)$ for each valid k .

Acknowledgements

P. E. Harris was supported through a Karen Uhlenbeck EDGE Fellowship. M. Yin was supported by the University of Denver's Faculty Research Fund 84688-145601 and Professional Research Opportunities for Faculty Fund 80369-145601.

The authors thank Kevin Wang for his insights which made Section 2.6 and Section 3 possible. Pamela E. Harris and Mei Yin acknowledge helpful conversations in the Special Session on Algebraic, Geometric, and Topological Combinatorics at the 2022 AMS Central Fall Sectional Meeting in El Paso, organized by Art Duval, Caroline Klivans, and Jeremy L. Martin. Rodrigo Ribeiro would like to thank Pedro Franklin from Federal University of Uberlândia for all the helpful conversations about statistics and simulations on Friday nights.

The authors are grateful to many valuable comments from the referee, which significantly improved the quality of the paper.

References

- [1] Etienne Bellin. Asymptotic behaviour of the first positions of uniform parking functions. [arXiv:2108.08661](https://arxiv.org/abs/2108.08661), 2021.
- [2] Joshua Carlson, Alex Christensen, Pamela E. Harris, Zakiya Jones, and Andrés Ramos Rodríguez. Parking functions: Choose your own adventure. *The College Mathematics Journal*, 52(4):254–264, 2021.
- [3] Persi Diaconis and Angela Hicks. Probabilizing parking functions. *Advances in Applied Mathematics*, 89:125–155, 2017.
- [4] Carl-Gustav Esseen. Fourier analysis for distribution functions. *Acta Mathematica*, 77:1–125, 1945.
- [5] Dominique Foata and John Riordan. Mappings of acyclic and parking functions (short communication). *Aequationes Mathematicae*, 9:299–300, 1973.

- [6] Ira M. Gessel and Seunghyun Seo. A refinement of cayley’s formula for trees. *The Electronic Journal of Combinatorics*, #P11, 2006.
- [7] Kimberly P. Hadaway and Pamela E. Harris. Honk! Honk!, Part 1: An Introduction to Parking Functions. *Girls’ Angle Bulletin*, 14(2):15–20, 2021.
- [8] Kimberly P. Hadaway and Pamela E. Harris. Honk! Honk!, Part 2: An Introduction to Parking Functions. *Girls’ Angle Bulletin*, 14(3):14–20, 2021.
- [9] Richard Kenyon and Mei Yin. Parking functions: From combinatorics to probability. *Methodology and Computing in Applied Probability*, 25: 32, 2023.
- [10] A.M. Khidr and B.S. El-Desouky. A symmetric sum involving the stirling numbers of the first kind. *European Journal of Combinatorics*, 5(1):51–54, 1984.
- [11] John E. Kolassa and Peter McCullagh. Edgeworth series for lattice distributions. *The Annals of Statistics*, 18:981–985, 1990.
- [12] Alan G. Konheim and Benjamin Weiss. An occupancy discipline and applications. *SIAM Journal on Applied Mathematics*, 14:1266–1274, 1966.
- [13] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. MBK. American Mathematical Society, 2017.
- [14] Jean-Christophe Novelli and Jean-Yves Thibon. Duplicial algebras and lagrange inversion. *Algebraic Combinatorics, Resurgence, Moulds and Applications (CARMA)*, 1, 2020.
- [15] Jim Pitman. Forest volume decompositions and abel - cayley - hurwitz multinomial expansions. *Journal of Combinatorial Theory. Series A*, 98(1):175–191, 2002.
- [16] John Riordan. *Combinatorial Identities*. Wiley series in probability and mathematical statistics. Wiley, 1968.
- [17] Melanie Tian and Enrique Trevino. Generalizing parking functions with randomness. *The Electronic Journal of Combinatorics*, 29:#P3.13, 2022.