Cycle Structure of Random Parking Functions

J. E. Paguyo

University of Southern California

Permutation Patterns 2022

Introduction to Parking Functions

- Consider n parking spaces on a one-way street arranged in increasing order.
- There are n cars, each with a preferred spot. Let π_i, with $1 \leq \pi_i \leq n$, be the preference of car i, for $1 \leq i \leq n$.
- Car i parks at spot π_i if it is available; otherwise it takes the next available spot to the right if it exists.

Definition

A parking function is a sequence $\pi = (\pi_1, \ldots, \pi_n)$ with $1 \leq \pi_i \leq n$ so that all cars can park. Let PF_n be the set of parking functions of size n.

Parking functions are combinatorial objects with applications to combinatorics, probability, and computer science. We would like to explore them by asking: "What does a typical parking function look like?"

Introduction to Parking Functions (cont.)

Example

$\pi = 13531$ is a parking function since this sequence of preferences results in the parking

$$\underline{1} \quad \underline{1} \quad \underline{3} \quad \underline{3} \quad \underline{5}$$

- Note that $S_n \subset \mathrm{PF}_n$ and that parking functions are invariant under permutations.
- Equivalently, π is a parking function if and only if $\pi_{(i)} \leq i$, $i \in [n]$, where $(\pi_{(1)}, \ldots, \pi_{(n)})$ is π sorted in weakly increasing order.

History and Previous Results

- Introduced by Konheim and Weiss ('66) in their study of the hash storage structure. They showed that

$$|\mathrm{PF}_n| = (n+1)^{n-1}$$

- Connections to various combinatorial objects:
  1. (Stanley, '97-'98). set partitions and hyperplane arrangements.
  2. (Pitman-Stanley, '02). volume polynomials of polytopes.
  3. (Haiman, '94). symmetric functions.
  4. (Cori-Rossin, '00). abelian sandpiles.
- Some variants and generalizations:
  1. (Kung-Yan, '03). $\boldsymbol{u}$-parking functions ($\pi_{(i)} \leq u_i$).
  2. (Postnikov-Shapiro, '04). G-parking functions, where G is a digraph.
  3. (Gorsky-Mazin-Vazirani, '16). rational parking functions
  4. (Ehrenborg-Happ, '16). parking with variable car sizes.

History and Previous Results (cont.)

- Probabilistic questions have also been considered, but tend to be harder.
  - (Chassaing-Marckert, '01). Found bijection with rooted labeled trees on $n+1$ vertices. Showed that the queue length in a BFS on a uniformly random tree converges to Brownian excursion.
  - (Flajolet-Poblete-Viola, '98; Janson, '01). Considered generalized parking functions (m cars, n spots, $m \leq n$) and showed that the distribution of the displacement statistic

$$A(\pi) = \binom{n+1}{2} - (\pi_1 + \cdots + \pi_n)$$

    converges to normal, Poisson, or Airy distributions, depending on the ratio between m and n.
  - (Diaconis-Hicks, '17). Studied the asymptotic distribution of coordinates (among other statistics).
  - (Kenyon-Yin, '21). Extended results to generalized and $\boldsymbol{u}$-parking functions.

Equivalence of Ensembles

- Let $\mathcal{F}_n = \{f : [n] \to [n]\}$ be the set of mappings from $[n]$ to $[n]$.
- Note that $\mathrm{PF}_n \subseteq \mathcal{F}_n$ and $\frac{|\mathrm{PF}_n|}{|\mathcal{F}_n|} = \frac{(n+1)^{n-1}}{n^n} \asymp \frac{1}{n}$.
- It is natural to expect that for some statistics, the distribution of statistics in the "micro-canonical ensemble", PF_n, should be close to those in the "canonical ensemble", $\mathcal{F}_n$.
- (Diaconis-Hicks, '17). The ensembles PF_n and $\mathcal{F}_n$ have the same limiting distributions for:
  - Number of repeats: $R(\pi) = |\{i : \pi_i = \pi_{i+1}\}|$
  - Number of "lucky" cars, $L(\pi)$, able to park in their preferred spot
  - Number of 1's: $N_1(\pi) = |\{i : \pi_i = 1\}|$
- (Diaconis-Hicks, '17). The ensembles have exactly the same distributions for: descents, ascents, inversions (and more)

Equivalence of Ensembles (cont.)

- Let $N_i(\pi)$ and $N_i(f)$ be the number of i's in $\pi \in \mathrm{PF}_n$ and $f \in \mathcal{F}_n$, respectively.
  - For random parking functions, $P(N_n(\pi) = 1) \sim \frac{1}{e}$ and $P(N_n(\pi) = 0) \sim 1 - \frac{1}{e}$.
  - For random mappings, $N_i(f)$ is approximately $\mathrm{Poisson}(1)$ distributed for all i. In particular, $P(N_n(f) = 0) \sim \frac{1}{e}$.
- This provides an example of a distinct difference between PF_n and $\mathcal{F}_n$.
- Diaconis and Hicks asked for other statistics on random parking functions for which the "equivalence of ensembles" heuristic holds.
- One suggestion they gave was to look at the cycle structure of random parking functions.

Parking Function Multi-Shuffle

Parking completions provide the link between enumeration and probability. First we define the parking function multi-shuffle.

- Fix $\pi_{\ell+1}, \ldots, \pi_n$. Define

$$A_{\pi_{\ell+1},\ldots,\pi_n} := \{\boldsymbol{u} = (u_1, \ldots, u_\ell) : \pi = (u_1, \ldots, u_\ell, \pi_{\ell+1}, \ldots, \pi_n) \in \mathrm{PF}_n\}.$$

- By a switch of coordinates, $\pi \in \mathrm{PF}_n$ if and only if its rearrangement is in PF_n, so we may assume that $\boldsymbol{u} = (u_1, \ldots, u_\ell)$ is in increasing order.
- If $A_{\pi_{\ell+1},\ldots,\pi_n}$ is nonempty, then there exists a unique maximal element (in component-wise partial order) $\boldsymbol{u} \in [n]^\ell$ with $u_i \geq i$ for all $1 \leq i \leq \ell$, and so $A_{\pi_{\ell+1},\ldots,\pi_n} = [\boldsymbol{u}]$.
- Thus given the last $n - \ell$ parking preferences, it suffices to find the maximal possible first ℓ preferences.

Parking Function Multi-Shuffle (cont.)

Definition

Let $1 \leq \ell \leq n$ and let $\boldsymbol{u} = (u_1, \ldots, u_\ell) \in [n]^\ell$ be in increasing order. We say that $(\pi_{\ell+1}, \ldots, \pi_n)$ is a parking function multi-shuffle of $(\alpha_1, \ldots, \alpha_{\ell+1})$, where $\alpha_1 \in \mathrm{PF}_{u_1-1}$, $\alpha_2 \in \mathrm{PF}_{u_2-u_1-1}$,..., $\alpha_\ell \in \mathrm{PF}_{u_\ell-u_{\ell-1}-1}$, $\alpha_{\ell+1} \in \mathrm{PF}_{n-u_\ell}$ if $\pi_{\ell+1}, \ldots, \pi_n$ is any permutation of the union of the $\ell+1$ words α_1, $\alpha_2 + (u_1, \ldots, u_1)$,..., $\alpha_{\ell+1} + (u_\ell, \ldots, u_\ell)$. We denote this as
$(\pi_{\ell+1}, \ldots, \pi_n) \in \mathrm{MS}(u_1 - 1, u_2 - u_1 - 1, \ldots, u_\ell - u_{\ell-1} - 1, n - u_\ell)$.

Example

For $n = 8$ and $u_1 = 4$, we have that $(2, 5, 1, 6, 8, 2, 7)$ is a parking function multi-shuffle of $\alpha_1 = (2, 1, 2) \in \mathrm{PF}_3$ and $\alpha_2 = (1, 2, 4, 3) \in \mathrm{PF}_4$.

Parking Completions

Theorem (Diaconis-Hicks, '17; Yin, '21)

Let $1 \leq \ell \leq n$ *and let* $\boldsymbol{u} = (u_1, \ldots, u_\ell) \in [n]^\ell$ *be in increasing order. Then* $A_{\pi_{\ell+1}, \ldots, \pi_n} = [\boldsymbol{u}]$ *if and only if* $(\pi_{\ell+1}, \ldots, \pi_n) \in \mathrm{MS}(u_1 - 1, u_2 - u_1 - 1, \ldots, u_\ell - u_{\ell-1} - 1, n - u_\ell)$.

- Consider the original parking problem. Suppose ℓ of the n spots are already occupied, with preferences given by $\boldsymbol{v} = (v_1, \ldots, v_\ell)$, where we can assume that the entries are in increasing order by permutation invariance.
- We want to find parking preferences for the remaining $n - \ell$ cars so that they can all successfully park.
- The set of successful preference sequences are the parking completions for $\boldsymbol{v} = (v_1, \ldots, v_\ell)$, denoted by $\mathrm{PC}_n(\boldsymbol{v})$.

Parking Completions (cont.)

Theorem (Adeniran et al, '20; Yin, '21)

Let $1 \leq \ell \leq n$ and let $\boldsymbol{v} = (v_1, \ldots, v_\ell) \in [n]^\ell$ be in increasing order. The number of parking completions for $\boldsymbol{v} = (v_1, \ldots, v_\ell)$ is

$$|\mathrm{PC}_n(\boldsymbol{v})| = \sum_{\boldsymbol{s} \in L_n(\boldsymbol{v})} \binom{n-\ell}{\boldsymbol{s}} \prod_{i=1}^{\ell+1} (s_i + 1)^{s_i - 1},$$

where

$$L_n(\boldsymbol{v}) = \left\{ \boldsymbol{s} = (s_1, \ldots, s_{\ell+1}) \in \mathbb{N}^{\ell+1} \,\middle|\, s_1 + \cdots + s_i \geq v_i - i \ \ \forall i \in [\ell], \sum_{i=1}^{\ell+1} s_i = n - \ell \right\}$$

- Note that this quantity stays constant if all $v_i \leq i$ and decreases as each v_i increases past i since there are fewer resulting summands.

Expected Number of k-Cycles

- Let $C_k(\pi)$ be the number of k-cycles of $\pi \in \mathrm{PF}_n$.
- For example, $\pi = 6\,1\,2\,4\,1\,9\,1\,6\,8\,4\,2\,10 \in \mathrm{PF}_{12}$ has one 3-cycle $(6\,9\,8)$ and one fixed point at 4.
- We can decompose $C_k(\pi)$ into a sum of indicator random variables:

$$C_k(\pi) = \sum_{\alpha \in A_k} \mathbb{1}_{\{\alpha \text{ is a } k\text{-cycle in } \pi\}}$$

where $A_k = \{(i_1, \ldots, i_k) : 1 \leq i_1 < \cdots < i_k \leq n\}$.

Proposition (P., '22+)

Let $\pi \in \mathrm{PF}_n$ be a uniformly random parking function. Then $E(C_k(\pi)) \sim \frac{1}{k}$ for all $k \in [n]$.

Expected Number of k-Cycles (cont.)

Proof Sketch:

- By linearity of expectation,

$$\begin{aligned} E(C_k(\pi)) &= \sum_{1 \le i_1 < \cdots < i_k \le n} P(\pi_1 = i_1, \ldots, \pi_k = i_k) \\ &= \frac{1}{|\mathrm{PF}_n|} \sum_{1 \le i_1 < \cdots < i_k \le n} (k-1)! |\mathrm{PC}_n((i_1, \ldots, i_k))|. \end{aligned}$$

- Applying the parking completion theorem yields

$$\begin{aligned} &\sum_{1 \le i_1 < \cdots < i_k \le n} |\mathrm{PC}_n((i_1, \ldots, i_k))| \\ &= \sum_{1 \le i_1 < \cdots < i_k \le n} \sum_{\mathbf{s} \in L_n((i_1, \ldots, i_k))} \binom{n-k}{\mathbf{s}} \prod_{i=1}^{k+1} (s_i + 1)^{s_i - 1} \end{aligned}$$

Expected Number of k-Cycles (cont.)

- A change of variables and some technical computations give:

$$\sum_{1\le i_1<\cdots<i_k\le n} |\mathrm{PC}_n((i_1,\ldots,i_k))|$$

$$= \left[\sum_{s_1=0}^{n-k}\sum_{s_2=0}^{n-k-s_1}\cdots\sum_{s_k=0}^{n-k-\sum_{i=1}^{k-1}s_i}\binom{n-k}{s_1,\ldots,s_k,n-k-\sum_{i=1}^{k}s_i}\right.$$

$$\left.\times\left(n-k-\sum_{i=1}^{k}s_i+1\right)^{n-k-\sum_{i=1}^{k}s_i}\prod_{i=1}^{k}(s_i+1)^{s_i-1}\right]\frac{n^{k-1}}{k!}(1+o(1))$$

- The expression in the bracket can be handled using Abel's multinomial theorem (Pitman, '02). So finally we get:

$$\frac{1}{|\mathrm{PF}_n|}\sum_{1\le i_1<\cdots<i_k\le n}(k-1)!|\mathrm{PC}_n((i_1,\ldots,i_k))| = \frac{(n+1)^{n-k}n^{k-1}}{(n+1)^{n-1}k}(1+o(1))$$

Asymptotic Distribution of Cycles

Our main result is:

Theorem (P., '22+)

Let $\pi \in \mathrm{PF}_n$ be a uniformly random parking function. The process of cycle counts converges in distribution to a process of independent Poisson random variables

$$(C_1, C_2, \ldots) \xrightarrow{D} (Y_1, Y_2, \ldots)$$

as $n \to \infty$, where $Y_k \sim Poisson(1/k)$.

- We prove something stronger. Recall that the total variation distance between two random variables X, Y is

$$d_{TV}(X, Y) = \sup_{A \subseteq \Omega} |P(X \in A) - P(Y \in A)|.$$

Let $W = (C_1, \ldots, C_d)$ and $Y = (Y_1, \ldots, Y_d)$. Suppose $d = o(n^{1/5})$. Then $d_{TV}(W, Y) = O\left(\frac{d^5}{n-d}\right) \to 0$ as $n \to \infty$.

A Brief Introduction to Stein's Method

Our proof uses Stein's method. Main advantage: provides explicit error bounds for the distributional approximation.

- Introduced by Charles Stein in 1972 as a novel approach to proving CLTs for sums of dependent random variables.
- Developed for many target distributions: e.g. Poisson (Chen, '75), Gamma (Luk, '94), Dickman (Bhattacharjee-Goldstein, '19).
- The starting point is Stein's lemma; e.g. for Poisson approximation:
  - Let $\mathcal{A}f(k) = \lambda f(k+1) - kf(k)$ be the characterizing operator of the Poisson distribution.
  - Then $Y \sim \mathrm{Poisson}(\lambda)$ if and only if $E(\mathcal{A}f(Y)) = 0$ for all bounded functions f.

A Brief Introduction to Stein's Method (cont.)

- Let $Y \sim \mathrm{Poisson}(\lambda)$ and $A \subseteq \mathbb{N} \cup \{0\}$. The Stein equation for the Poisson distribution

$$\lambda f_A(k+1) - k f_A(k) = \mathbb{1}_{\{k \in A\}} - P(Y \in A)$$

  has a unique bounded solution f_A.
- As a corollary: If $W \geq 0$ is an integer-valued random variable with mean λ, then

$$d_{TV}(W, Y) = \sup_{A \subseteq \Omega} |E[\lambda f_A(W+1) - W f_A(W)]|.$$

- Bounding the distance between the distributions of W and Y is reduced to bounding $E[\lambda f_A(W+1) - W f_A(W)]$. If W is approximately Poisson, then this should be close to 0.
- Several approaches:
  - dependency graphs (Chen, '75; Arratia-Goldstein-Gordon, '90)
  - exchangeable pairs (Stein, '86; Chatterjee-Diaconis-Meckes, '05)
  - size-bias couplings (Barbour-Holst-Janson, '92)

Stein's Method and Exchangeable Pairs

We use the following multivariate version:

Theorem (Chatterjee-Diaconis-Meckes, '05)

Let $W = (W_1, \dots, W_d)$ be a random vector with $EW_i = \lambda_i$. Let $Z = (Z_1, \dots, Z_d)$ have independent coordinates with $Z_i \sim Poisson(\lambda_i)$. Let $W' = (W'_1, \dots, W'_d)$ be defined on the same probability space as W with (W, W') an exchangeable pair. Then

$$d_{TV}(W, Z) \leq \sum_{k=1}^{d} \alpha_k [E|\lambda_k - c_k P(A_k)| + E|W_k - c_k P(B_k)|],$$

with $\alpha_k = \min\{1, 1.4\lambda_k^{-1/2}\}$, any choice of the $\{c_k\}$, and

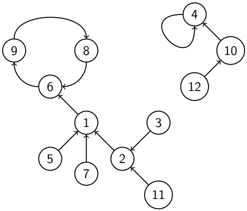
$$A_k = \{W'_k = W_k + 1, W_j = W'_j \text{ for } k+1 \leq j \leq d\},$$
$$B_k = \{W'_k = W_k - 1, W_j = W'_j \text{ for } k+1 \leq j \leq d\}.$$

Proof Sketch

- The pair (W, W') of random variables is an exchangeable pair if $(W, W') =_d (W', W)$.
- To construct an exchangeable pair, one typically applies a "small perturbation" to the original variable W. Let π' be the parking function obtained by applying a random transposition to our initial parking function π. Set $C_k' = C_k(\pi')$. Then (C_k, C_k') is an exchangeable pair.
- For fixed k, we upper bound the summands. This involves finding expressions for $P(A_k)$ and $P(B_k)$.
- We do this geometrically by considering the digraph representation of π. There are n labeled vertices with a directed edge from i to π_i so that each vertex has outdegree one.

Digraph Representation (cont.)

- For example, $\pi = 6\,1\,2\,4\,1\,9\,1\,6\,8\,4\,2\,10 \in \mathrm{PF}_{12}$ has the following digraph representation:



- Digraphs of parking functions consist of connected components, where each connected component is comprised of rooted trees arranged in a cycle.
- The idea is to determine the transpositions which increase or decrease the number of k-cycles by one, but neither create nor destroy any j-cycles for $k+1 \leq j \leq d$. This involves a lot of case analysis.

Final Remarks

- In a work in progress, we are carrying out the Diaconis-Hicks probabilistic program for rational parking functions. Diaconis and Hicks state "These are at the forefront of current research with applications to things like the cohomology of affine Springer fibers".
- One can also conduct a probabilistic study on $\boldsymbol{u}$-parking functions and G-parking functions.
- Harris et al considered k-Naples parking functions, where the cars can move backwards to check vacant spots before moving forward. There are no probabilistic developments in this direction.
- Adeniran and Pudwell introduced the notion of pattern-avoiding parking functions. A parallel probabilistic study on these objects is fully warranted.

Thank you! Questions?