Inference Studies on 1D Random Sequential Adsorption with Periodic Boundary Conditions

Chanyang Ryoo

New York University, Courant Institute of Mathematical Sciences
251 Mercer Street New York, NY 10012

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master's of Science in Mathematics
Department of Mathematics
New York University
Decemeber 2015

Miranda Holmes-Cerfon and Georg Stadler,
Faculty Advisors

Acknowledgements

To the only wise God, through Jesus Christ, be the glory forever. Amen.

I want to thank professors Holmes-Cerfon and Stadler for their gracious help. May God grant grace and peace to them and their families.

Abstract

We study inference problems of one-dimensional complete random sequential adsorption (RSA) on a finite interval I with periodic boundary conditions, also called *random parking*. The problems involve estimating the relative order of arrival of two cars and the relative order of arrival of two groups of cars. Using the likelihood of each possible arrival history of the whole set of cars, we make probabilistic predictions that infer the correct relative orders significantly better than purely random guessing (50%). This improvement from 50% increases with the size of the interval I. We also form methods of deterministic prediction using geometry of the RSA, mainly the gap size surrounding each car, and observe correct inference higher than 50% that also increases as the system size increases. Additionally, we theoretically derive the first two moments of the random variable $\Gamma(c_t)$, which is the size of the surrounding gap that a car parking at time t will have on an infinite line, and confirm the derivation of the first moment via simulations.

Contents

1 Introduction

1.1 What is Random Sequential Adsorption?

Random sequential adsorption (RSA) is a stochastic process of irreversibly placing finitely sized n-dimensional objects (e.g. line, square, circle) uniformly on a n-dimensional region (bounded or unbounded) in order such that no overlap occurs with any of the previously placed objects. Such model appears in biology and physical chemistry, for instance, when one needs to study the adsorption of bacteria on a surface, or when one desires to produce chemically active substrates by depositing particles on a surface, as studied by Feder [1] and Katira et. al. [2]. Feder showed that the iron storage protein ferritin adsorbs on carbon surface as disks would on 2D plane according to RSA, while Katira showed how proteins adsorb according to RSA to a surface functionalized with polyethylene oxide to create a chemically active surface. These are examples of 2D RSA, for which exact analytical results have not been found as in the case of 1D RSA (further discussed in next section). Studies have been done on the three dimesional case as well, modelling how spheroidal particles would adsorb onto a surface according to RSA [3]. Similar to 2D RSA, only results from simulations and theoretical approximations have been made in 3D RSA. In another more down-to-earth case, RSA has been modified to model the parking behavior of cars in streets of London by Rawal and Rodgers [4]. For this research, we focus on the one dimensional case, where an interval ('car') of length l adsorbs ('parks') itself on a circular interval of radius ρ until no more car can park (called *complete* RSA). Henceforth, we shall call the case of such random sequential adsorption 'RP', standing for 'Random Parking.' Symbolically, we have that $\mathcal{R}_{(l,\rho)}$, a realization of RP which we simulate via code, lies in $D \subset [0, 2\pi]^N$, a subset of N element vectors denoting all possible configuration of random parking, where N is the random number of cars. Although N is random, it must fall within the range $\{N_{min} = \lceil \frac{\pi\rho}{l} \rceil, \lceil \frac{\pi\rho}{l} \rceil + 1, ..., \lfloor \frac{2\pi\rho}{l} \rfloor = N_{max}\}$. The one-dimensional case was the first RSA model to be studied by a Hungarian mathematician named Alfred Renyi who started to derive analytical results, some of which we present in the next section.

1.1.1 Analytical results in 1D RSA

Some previous analytical results on infinite street RP (RP_∞) include asymptotic analysis by Alfred Renyi [5], the formulation of the density of gap size with respect to time, the density of street covered with cars with respect to time, and the density of street available for parking with respect to time [7]. Renyi proved that if unit length cars are randomly parked on a linear street of size x, then the ratio of mean number of cars parked on street and the length x approaches a limit as $x \to \infty$:

$$\lim_{x\to\infty} \frac{M(x)}{x} = \int_0^\infty \exp\big(-2\int_0^x \frac{1-e^{-y}}{y} dy\big) dx = 0.74759...$$

Also, there is an analytical expression of the population density of gap lengths as a function of time for RP_∞. This function $G(h,t)$ can be derived as a solution to an integro-differential equation that describes the dynamics of RP:

$$\frac{\partial G(h,t)}{\partial(k_a t)} = -H(h-l)(h-l)G(h,t) + 2\int_{h+l}^\infty G(s,t) ds,$$

where $H(x) = \mathbf{1}_{x>0}$ is the unit step function, k_a is the constant rate of cars entering the street, h the gap length, and l the length of car. The left hand term of the RHS describes the rate of destruction of gaps of size h via insertion of car in a gap of size h, and the right hand term describes the creation of gaps of size h via parking of cars in a gap of size greater than or equal to $h+l$. The solution to this differential equation is divided into two cases, 1) for $h \in (0, l)$ and 2) for $h > l$:

$$1)\, G(h,t) = \frac{2}{l^2} \int_0^{k_a l t} u \exp(-uh/l) \exp\Big(-2\int_0^u \frac{1-e^{-s}}{s} ds\Big) du$$

$$2)\, G(h,t) = \frac{(k_a l t)^2}{l^2} \exp\Big(-2\int_0^{k_a l t} \frac{1-e^{-u}}{u} du\Big) \exp(-k_a(h-l)t).$$

We present a plot in Figure 1 of this gap density function for $l = 1$ and $k_a = 1$ and for different times t. We observe more and more mass concentrating near $h = 0$ because as time progresses, more cars have parked near one another.

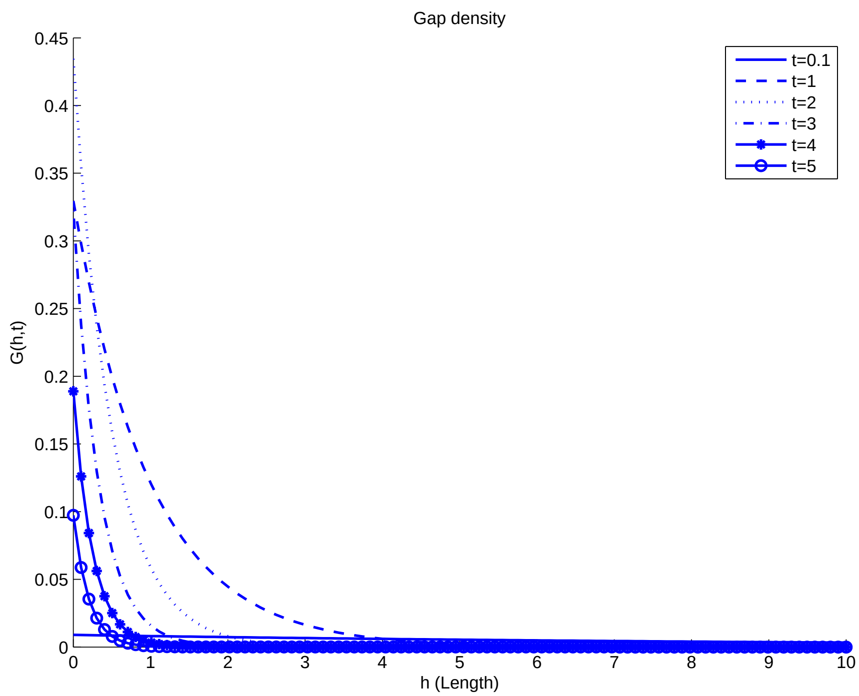

Figure 1: $G(h,t)$ is plotted for $t = 0.1, 1, 2, 3, 4, 5$, for $h \in [0, 10]$. We see that as t increases the density of gaps with size near 0 increases, while the density of gaps with size h far from 0 decreases.

Now with $G(h,t)$ one can find several other functions. First, the density of the infinite line covered by cars $\rho(t)$ is derived as

$$\rho(t) = \int_0^\infty G(h,t) dh,$$

as each gap of any size h is adjacent to a car, and the integral from 0 to ∞ gives the total density of cars on the line. We can also find the density of line available for parking $\Phi(t)$

by noting that for a given gap of size h, the available space it has is $h - l$ with which we discount the density of gap size h:

$$\begin{aligned}
\Phi(t) &= \int_l^\infty (h-l)G(h,t)dh \\
&= \frac{(k_a lt)^2}{l^2} \exp\big(-2\int_0^{k_a lt} \frac{1-e^{-u}}{u} du\big) \int_l^\infty (h-l)e^{-k_a(h-l)t} dh \\
&= \frac{(k_a lt)^2}{l^2} \exp\big(-2\int_0^{k_a lt} \frac{1-e^{-u}}{u} du\big) \int_0^\infty s e^{-k_a st} ds \\
&= \exp\big(-2\int_0^{k_a lt} \frac{1-e^{-u}}{u} du\big).
\end{aligned}$$

We know that $\Phi(t) \to 0$ as $t \to \infty$, because the density of cars on street $\rho(t)$ satisfies the differential equation

$$\frac{d\rho(t)}{dt} = k_a \Phi(t),$$

and $\rho(t) \to \lim_{x\to\infty} \frac{M(x)}{x}$ as $t \to \infty$, the Renyi's constant.

1.2 Parameters and Likelihood

Note that $\mathcal{R}_{(l,\rho)}$ (or $\mathrm{RP}_{(l,\rho)}$) is a finite sequence of some N angles, which we can describe as a pair (r, σ), where r is the set of angles of cars' positions, and $\sigma \in S_N$ is a permutation on r that describes the order in which they entered the street. Its probability density function, $p_{\mathcal{R}}$, depends on (r, σ) through the following relation:

$$p_{\mathcal{R}}(r;\sigma) = \frac{1}{2\pi\rho} \prod_{i=2}^N \frac{1}{\tilde{A}(r_{\sigma_i})}, \tag{1}$$

where $\tilde{A}(r_{\sigma_i})$ is the length of the circumference available for the ith car to park after cars $1, 2, ...(i-1)$ have parked (according to σ). For example, $\tilde{A}(r_{\sigma_2})$ will be the space available for the second car (the angle in r with assigned order 2) to park, so it will be $(2\pi\rho - 2l)$, as small intervals of length $\frac{l}{2}$ on both ends of a car are also unavailable for parking to avoid overlap.

Now if we are only given r and no σ, we can find the likelihood of each possible $\sigma \in S_N$. Likelihood of a parametergiven an outcome $\mathcal{L}(\sigma; r)$ is the probability of the outcome given the parameter i.e. Eq. 2, where we order r by σ and compute (1):

$$\mathcal{L}(\sigma;r) = \frac{1}{2\pi\rho} \prod_{i=2}^N \frac{1}{\tilde{A}(r_{\sigma_i})}. \tag{2}$$

Thus the important parameter of estimation is $\sigma \in S_N$, the arrival history.

1.3 Inference Problem

There can be different inference problems posed with the RSA model. For instance, Lieshout posed and estimated the problem of inferring car length and parking duration from a given 'core parking' (position of just the center of cars), using maximum likelihood [10]. Now the question we pose is if we can gain information given r, a realization of complete RP, about the permutation of true arrival order σ^*, such as whether one element of r arrived earlier than another element, using the knowledge of the likelihood function. For small N, it is possible to observe several things. We can calculate $\mathcal{L}$ for all possible $\sigma \in S_N$ and find $\text{argmax}_\sigma \mathcal{L}(\sigma; r)$, for instance. We can also measure informativeness of the likelihood function on σ^* by normalizing the function and treating it as a distribution to measure the Kullback-Leibler divergence (relative entropy) against other distributions. Note that the normalized likelihood (NLD), which we denote as $\mathbb{P}^{(1)}(\sigma; r) = \frac{\mathcal{L}(\sigma;r)}{\sum_\sigma \mathcal{L}(\sigma;r)}$, can be treated as a distribution because it is an update of some *prior* distribution $\mathbb{P}^{(1)}(\sigma)$ before observing sample r by Bayes' rule [6]:

$$\mathbb{P}^{(1)}(\sigma; r) = \mathbb{P}^{(1)}(\sigma) \times \frac{\mathbb{P}^{(1)}(r; \sigma)}{\mathbb{P}^{(1)}(r)}.$$

Thus, if we set $\mathbb{P}^{(1)}(\sigma)$ to be uniform since we assume that we do not know any information about the parameter σ before observation of any sample, then we retrieve the normalized likelihood function, since $\mathbb{P}^{(1)}(\sigma) = 1/|S_N|$ and $\mathbb{P}^{(1)}(r) = \sum_\sigma \mathbb{P}^{(1)}(r; \sigma)\mathbb{P}^{(1)}(\sigma) = \mathbb{P}^{(1)}(\sigma) \sum_\sigma \mathbb{P}^{(1)}(r; \sigma)$, where $\mathbb{P}^{(1)}(r; \sigma) = \mathcal{L}(\sigma; r)$ as in Eq. 2. Most importantly, we will take the partial sums of likelihoods corresponding to subsets of S_N to answer inference problems about σ^*. Finally, we would want to answer similar problems for $\frac{l}{\rho}$ small (i.e. N large).

2 Simulation and computations

2.1 RP simulation

There is a straight-forward way of simulating a RP on the computer, which replicates the exact procedure of the model. Namely, one can repeatedly generate uniform random variable over $[0, 2\pi]$, and reject if an overlap occurs or accept if not. This poses a problem of computational time for RP on large streets, since as more and more cars park, less and less space will be available on the street and the computer must generate a lot of uniform numbers to access the remaining small spaces for a complete parking. In order to circumvent this problem, we concentrate the uniform numbers to be only generated over the available gaps, with respect to their relative sizes. The algorithm is as follows and implemented in MATLAB.

(I) Optimized algorithm of simulation of $\mathcal{R}_{(l,\rho)}$:
1) Let $i = 1$. Park car $\theta_i \sim U(0, 2\pi)$, a uniform random variable from $[0, 2\pi]$
2) Let $gaps = 2\pi\rho - l$

3) While max($gaps$) $> l$
Let I_j be $[\theta_j - \frac{l}{\rho}, \theta_j + \frac{l}{\rho}]\, mod\, 2\pi$ for $j = 1, ...i$
Let $S = [0, 2\pi] - \cup_j I_j$ and decompose $S = \cup_j S_j$, where $\{S_j\}$ is open and disjoint
Let $d_j = |S_j|$
Let $K \sim U(0, \sum_j d_j)$ where the interval $(0, \sum_j d_j)$ is partitioned by $[0, d_1], [d_1, d_1 + d_2], ..., [d_1 + ... + d_{i-1}, d_1 + ... + d_i]$
Let j' denote the partition on which K fell
Let $\alpha \sim U(S_{j'})$
Let $i = i + 1$ and park car $\theta_i = \alpha$
Let $gaps$ = gap($\{\theta_j\}_{j=1}^{i}$) (see algorithm below for gap)

(II) Algorithm of finding $gaps$:
1) Sort in increasing order the existing angles θ_j, call θ'
2) Shift each index of θ' by +1, call ϕ (for the last index, we shift it to the first index)
3) Add 2π to last element of ϕ
4) Let $\Delta = \phi - \theta'$
5) Let $gaps = \rho\Delta - \mathbf{l}$, where $\mathbf{l} = [l, l, ..., l]$, whose number of elements equals that of θ'

We can run a performance comparison between the straight-forward algorithm and the optimized one, and we see that the optimized computational time is linear with respect to the radius size, whereas that for the straight-forward algorithm increases exponentially. See Figure 2. 40,000 RP samples were run and their average computing time is plotted for each radius size.

2.2 Likelihood computation

Given a sample of RP, the next computational task is to calculate its corresponding likelihood of $\sigma \in S_N$. From Eq. 2 we need to calculate $\tilde{A}(r_{\sigma_i})$ to calculate the likelihood. Note that this value is calculated as cars are still being parked according to σ. Symbolically, we let $\{\theta_1, ..., \theta_k\}$ denote the angles of the center of incompletely parked cars (i.e. $1 \leq k < N$), where θ_i is the i^{th} car to arrive according to σ. Furthermore, let I_i be as in algorithm (I), namely $[\theta_i - \frac{l}{\rho}, \theta_i + \frac{l}{\rho}]$ for $i = 1, ..., k$. Then we have $\tilde{A}(r_{\sigma_{k+1}}) = \rho(2\pi - |\cup_{i=1}^{k} I_i|\, mod\, 2\pi)$. The implementation of this calculation is simple, and to calculate the actual likelihood of σ, we re-simulate the given RP car by car according to σ, after each of which we compute $\tilde{A}$ until the $(N-1)^{th}$ car. The likelihood is then the inverse of the total product of all $\tilde{A}$. One thing to note about the likelihood computation is that once N reaches 10, the computation of all likelihoods (10! of them) becomes infeasible, as the permutations matrix in MATLAB is only practical for size less than 10.

2.3 Verifications of computation

First we verify the RP simulation code in two different ways. The first way is to confirm statistical properties of RP, namely the convergence of $M(x)/x$ to Renyi's constant. Apart from the periodic boundary condition, the RP we deal with should behave similarly to RP on a linear street. Furthermore, since as $\rho \to \infty$, $\mathrm{RP}_{(l,\rho)}$ approximate RP_∞, we simulate $10,000$

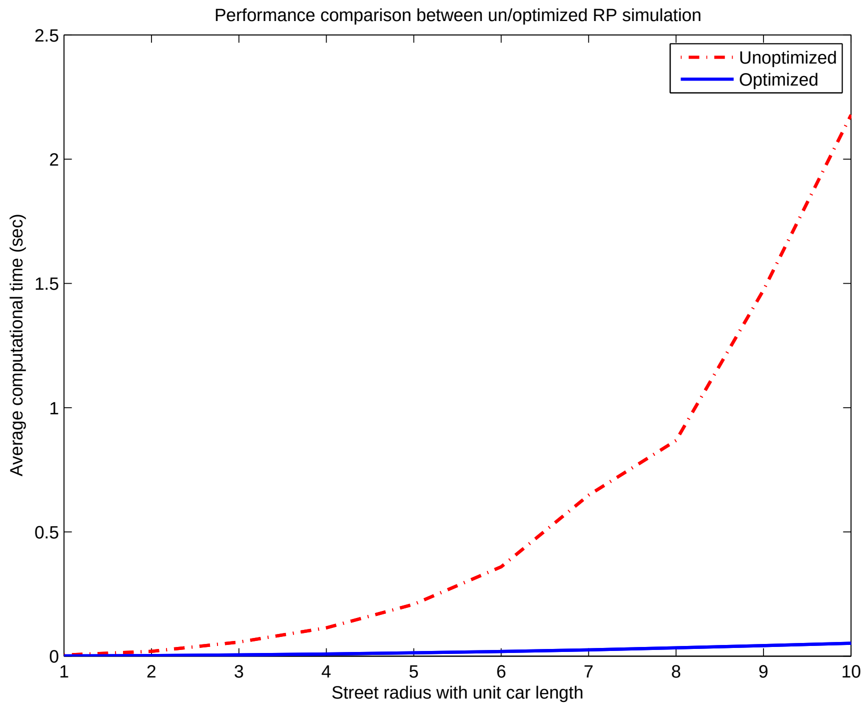

Figure 2: Average of 40,000 computational times (sec) for each parameter is plotted. We see an exponential growth for the unoptimized algorithm, compared to linear growth for the optimized one.

random parkings for $l = 1$ and $\rho = 1, 2, 3, ...20$ and observe the ratio $p_\rho = \frac{M_\rho}{2\pi\rho}$, where M_ρ is the average of number of cars parked on circle out of $10,000$ RP$_{(1,\rho)}$s. The result is shown in the plot in Figure 3. Shown with each estimate of the Renyi's constant is the 95% confidence interval $(\pm 2\sqrt{\frac{(p_\rho)(1-p_\rho)}{10000}})$. We observe that indeed the RP simulation gives convergence to Renyi's constant. As radius increases, the random parking with periodic boundary condition approximates the infinite line random parking.

The second way that we verify the RP simulation code is by graphically checking it is complete (Fig. 4). That is, we must have no available spaces for parking after simulation halts. To check this for a given RP, we boldface the intervals I_i (see algorithm I) for all $i = 1, 2, ...N$, where no car can park, and see that indeed their union cover the whole street.

Now to verify the computation of likelihood, we must verify the computation of $\tilde{A}$ as we can see from Eq. 2. To do so, we devise another independent, statistical method to calculate $\tilde{A}$ using Monte Carlo integration. After j^{th} car is parked $(j = 1, ..., N-1)$, we generate $k = 100,000$ random uniform values $\{\phi_i\}_{i=1}^{100,000}$ from $[0, 2\pi]$ and collect the ones that fall on the available parking spaces $AP_{\bar{r}} = \{\phi \in [0, 2\pi] : |\phi - \theta_i| \geq \frac{l}{\rho},\ \ \forall \theta_i \in \bar{r}\}$, where $\bar{r} = \{\theta_1, ..., \theta_j\}$. The fraction f of the number of $\phi_i \in AP_{\bar{r}}$ versus k will give us an estimate of $\tilde{A}$ by the relation $\tilde{A} \approx 2\pi\rho f$. Since there are some technical intricacies in the algorithm of this method, we present the pseudoalgorithm:

(III) Algorithm of MC $\tilde{A}$ after j^{th} car is parked:

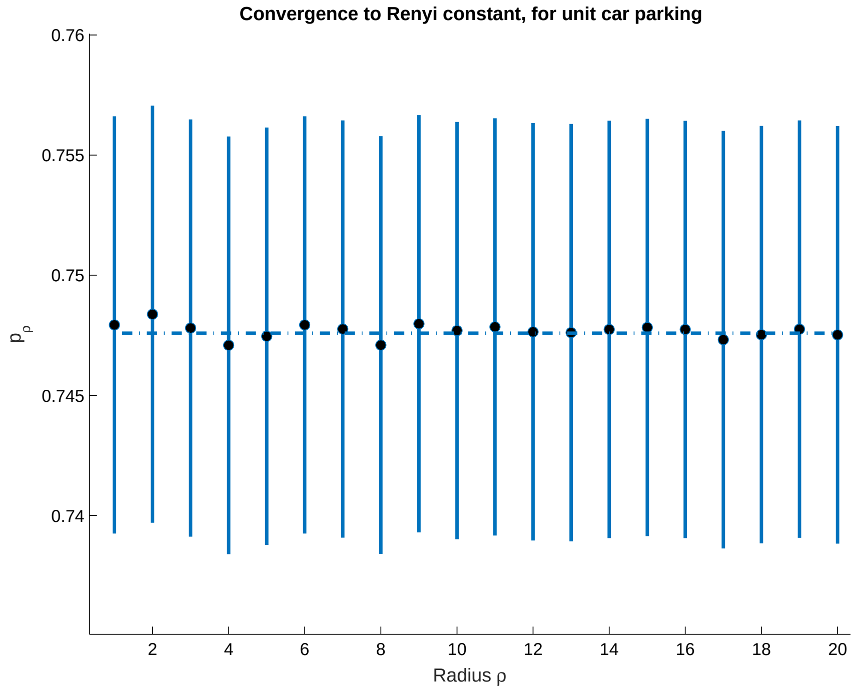

Figure 3: Even from radius 1, the average density of street covered by cars is close to the Renyi's constant, the average density of street covered by cars on an infinite line. We see that out of the 10,000 samples for each radius, the margin of error is of order 10^{-3}.

1) Let $k = 100,000$ and generate k uniform random values from $[0, 2\pi]$
2) Let M_1 be a $k \times j$ matrix whose columns are repeats of the k random values, and M_2 be a $k \times j$ matrix whose k rows are repeats of the parking angles
3) Let $\Delta = |M_1 - M_2|$, where $|\cdot|$ takes absolute value of every element
4) Count the number of rows c of Δ whose every element is larger than $\frac{l}{\rho}$
5) Let $\tilde{A} = 2\pi\rho\frac{c}{k}$

This method gives agreement to the analytical method within the 99 % confidence intervals (see a particular example in Fig. 5). To find the standard error for the MC integration, let us first write the integral we wish to estimate:

$$\tilde{A} = (2\pi\rho)(\frac{1}{2\pi}\int_0^{2\pi} \mathbf{1}_{AP_{\bar{r}}}\, d\phi) \tag{3}$$

where $\mathbf{1}_{AP_{\bar{r}}}$ is the indicator function of set $AP_{\bar{r}}$ on domain $[0, 2\pi]$. Then the integral will give the total angle value of 'open areas' and its ratio with 2π is the fraction f of such angles. Multiplying this to the whole circumference will give the area of the available space. This function has the following variance:

$$\mathrm{Var}(\mathbf{1}_{AP_{\bar{r}}}) \approx \frac{1}{k}\sum_{i=1}^{k}[\mathbf{1}_{\phi_i \in AP_{\bar{r}}} - \underline{\mathbf{1}}]^2,$$

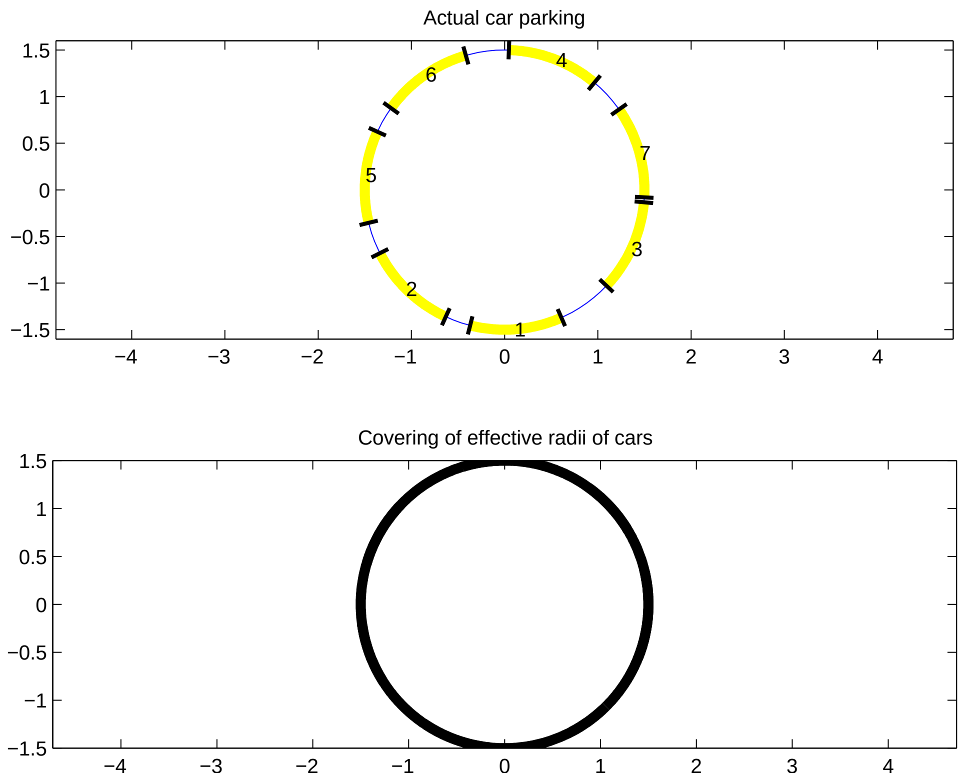

Figure 4: A 7-car RP with $l = 1, \rho = 1.5$. The whole street is covered by the *extended* cars, whose ends are extended by $\frac{1}{2}$ to indicate the region where no other cars can park. This shows the completeness of the RP.

where $\underline{\mathbf{1}} = \int_0^{2\pi} \mathbf{1}_{AP_{\bar{r}}}\, d\phi, \approx \frac{1}{k}\sum_{i=1}^{k} \mathbf{1}_{\phi_i \in AP_{\bar{r}}} = f$ and

$$\mathrm{Var}(\psi_k) = \frac{(2\pi)^2}{k}\,\mathrm{Var}(\mathbf{1}_{AP_{\bar{r}}}),$$

where $\mathrm{Var}(\psi_k)$ is the variance of the estimates ψ_k of $\tilde{A}$. So we expect the standard error (SD) to be approximately $\frac{2\pi\sqrt{\mathrm{Var}(\mathbf{1}_{AP_{\bar{r}}})}}{\sqrt{k}}$. Now $\mathrm{Var}(\mathbf{1}_{AP_{\bar{r}}}) \approx \frac{1}{k}[a(1-f)^2 + (k-a)f^2]$, where $a = |\{\phi_i \in AP_{\bar{r}}\}|$ i.e. $f = \frac{a}{k}$. This can be further simplified to $\frac{1}{k}[a(1-f)]$. So the error is

$$\mathrm{SD} \approx \frac{2\pi}{\sqrt{k}}\sqrt{f(1-f)}.$$

We calculate $\tilde{A}$ using this method and the analytical method, and see that the MC integration gives values that are within 3 SDs of the analytical value.

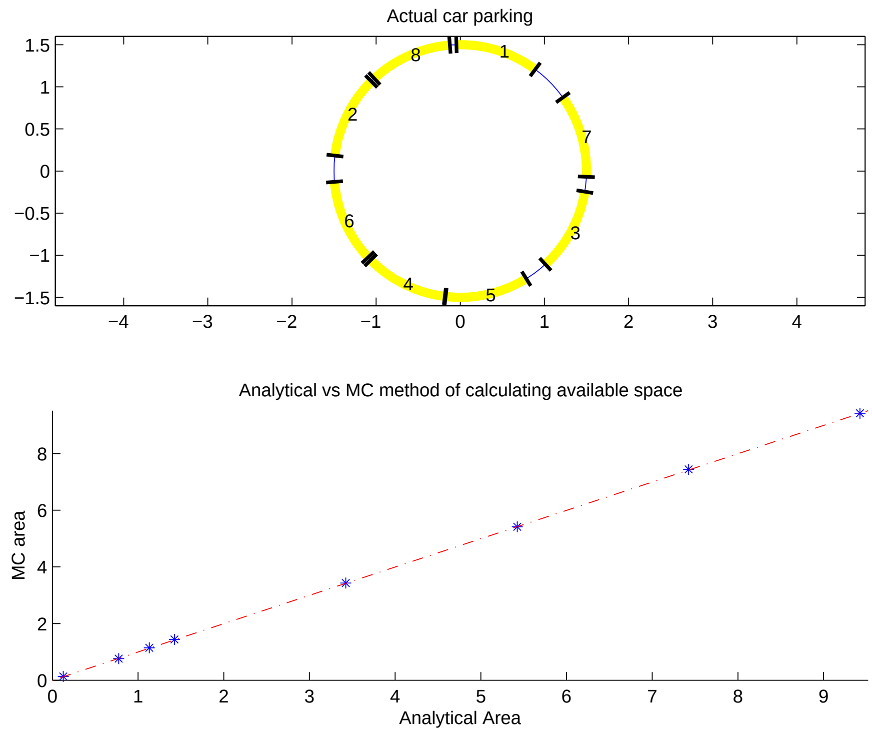

Figure 5: A 8-car RP with $l = 1, \rho = 1.5$. On the bottom, the line indicates $y = x$, and plotted on top of the line is pairs of computations of $\tilde{A}$ (MC and analytical), after each car is parked in the RP above. We observe a close alignment with $y = x$ and also see that the pair of points get close to 0 as more cars park.

3 Studies on the Likelihood Function

3.1 Interpretation as a probability distribution

Any introductory treatment of likelihoods will emphasize how a likelihood function is *not* a probability distribution. However, since likelihoods take value between 0 and 1, one can normalize a likelihood function such that its integral or sum equals 1 and treat it as a probability distribution. We seek to do so here and understand the *normalized* likelihood function (NLD) of a given RP to be the distribution of a random variable on S_N that corresponds to a guess of the arrival order of the cars. With regard to inference, a 'bad' distribution will be one that gives every permutation in S_N equal probability, without any preference to particular permutations. The best distribution will be one that gives unit mass on the true order with probability 1. Taking these into account, we establish formulas that quantify the difference between a NLD and the corresponding 'bad' and 'best' guesses.

Specifically, we use the Kullback-Leibler (KL) divergence, or also known as relative entropy, to study the differences between the NLD and uniform distribution and the NLD and unit mass distribution on σ^*. The KL divergence between P and Q measures how much information is lost when some probability distribution Q is used to approximate another probability distribution P, and is defined by:

$$D_{KL}(P||Q) = \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)},$$

where X is a discrete sample space (S_N in our case) [8]. In the case that $Q(x) = 0$ for some x, we must also have that $P(x) = 0$, else the relative entropy is not defined. In our case, there will be no such case. It is also noteworthy that relative entropy is not symmetric i.e. $D_{KL}(P||Q) \neq D_{KL}(Q||P)$ and thus not a true distance measure. We will interchangeably use the terms 'relative entropy' and 'divergence'.

In comparing between the NLD and the unit mass distribution on σ^*, we want Q to be the NLD, since the NLD should approximate reality, the unit mass on σ^*. To compare the NLD to a flat distribution, we let P be the NLD and Q be the flat distribution. This corresponds to quantifying how different our guess of parameter σ^* becomes after observing r from the uniform distribution. So to summarize, the distributions of interest are:

$$\mathbb{P}^{(1)}(\sigma; r) = \frac{\mathcal{L}(\sigma; r)}{\sum_\sigma \mathcal{L}(\sigma; r)} \quad \text{(the NLD)} \tag{4}$$

$$\mathbb{P}^{(2)}(\sigma) = \mathbf{1}_{\sigma^*}$$

and

$$\mathbb{P}^{(3)}(\sigma) = U = \frac{1}{|S_N|}.$$

Let us observe the first mentioned relative entropy between the unit mass and the normalized likelihood:

$$D_{KL}(\mathbb{P}^{(2)}||\mathbb{P}^{(1)}) = \sum_\sigma \mathbf{1}_{\sigma^*} \log \frac{\mathbf{1}_{\sigma^*}}{\frac{\mathcal{L}(\sigma;r)}{\sum_\sigma \mathcal{L}(\sigma;r)}} = \log \frac{\sum_\sigma \mathcal{L}(\sigma; r)}{\mathcal{L}(\sigma^*; r)} \tag{5}$$

First note that we desire this relative entropy to be small (small means close in distribution). In order to interpret the numerical value of the relative entropy and see how small it is, we can compare it to the maximum relative entropy that can be achieved with $\mathbb{P}^{(1)}$, i.e.,, $M = \max_f D_{KL}(f\,||\mathbb{P}^{(1)})$, where f is a distribution on S_N. We observe that a unit mass distribution on the worst permutation, $\underline{\sigma}$ (by which we mean the permutation with the lowest likelihood), i.e.,, $D_{KL}(\mathbf{1}_{\underline{\sigma}}||\mathbb{P}^{(1)}) = \log \frac{\sum_\sigma \mathcal{L}(\sigma;r)}{\mathcal{L}(\underline{\sigma};r)}$ maximizes the relative entropy with $\mathbb{P}^{(1)}$.

To see this, we write:

$$\begin{aligned} \max_f D_{KL}(f\,||\mathbb{P}^{(1)}) &= \max_f \sum_\sigma f(\sigma) \log \frac{f(\sigma)}{\frac{\mathcal{L}(\sigma;r)}{\sum \mathcal{L}(\sigma;r)}} \\ &= \max_f \sum_\sigma f(\sigma) \log \frac{cf(\sigma)}{\mathcal{L}(\sigma)}, \end{aligned}$$

where $\mathcal{L}(\sigma; r)$ has been shortened to $\mathcal{L}(\sigma)$ and $c = \sum_\sigma \mathcal{L}(\sigma)$. Further expanding, we get:

$$\max_f D_{KL}(f\,||\mathbb{P}^{(1)}) = \max_f \Big[\sum_\sigma f(\sigma) \log(cf(\sigma)) - \sum_\sigma f(\sigma) \log \mathcal{L}(\sigma)\Big]. \tag{6}$$

So we would like f to maximize

$$\sum_{\sigma} f(\sigma) \log(cf(\sigma)) = \sum_{\sigma} f(\sigma)(\log c) + \sum_{\sigma} f(\sigma) \log f(\sigma)$$

$$= \log c + \sum_{\sigma} f(\sigma) \log f(\sigma)$$

and minimize $\sum_{\sigma} f(\sigma) \log \mathcal{L}(\sigma)$ (see Eq. 6). But since $f(\sigma) \leq 1$, we have $\log f(\sigma) \leq 0$. So to maximize $\sum_{\sigma} f(\sigma) \log f(\sigma)$, we choose a unit mass distribution for f to make the sum achieve 0, its supremum. But notice that $\sum_{\sigma} f(\sigma) \log \mathcal{L}(\sigma)$ is a weighted sum of the likelihoods as $0 \leq f \leq 1$ and $\sum_{\sigma} f(\sigma) = 1$. So a correctly chosen unit mass distribution will also minimize this part. To choose the correct permutation of unit mass, we let $\underline{\sigma}$ denote the worst permutation with the least likelihood since we want to minimize $\sum_{\sigma} f(\sigma) \log \mathcal{L}(\sigma)$.

So we have $M = \log \frac{\sum_{\sigma} \mathcal{L}(\sigma)}{\mathcal{L}(\underline{\sigma})}$. Then we establish a measure on information gained with:

$$Y_1 = \frac{M - D_{KL}(\mathbb{P}^{(2)}||\mathbb{P}^{(1)})}{M}.$$

If the ratio of (5) to $D_{KL}(\mathbf{1}_{\underline{\sigma}}||\mathbb{P}^{(1)})$ is small, the likelihood analysis will be informative, as one shows how different the normalized likelihood distribution (NLD) is from the true distribution $\mathbf{1}_{\sigma^*}$ and the other how different the NLD is from the distribution corresponding to the worst guess, $\mathbf{1}_{\underline{\sigma}}$.

Next let us observe the relative entropy between $\mathbb{P}^{(3)}(\sigma) = \frac{1}{|S_N|}$ and $\mathbb{P}^{(1)}(\sigma; r)$, and see how to quantify how large it is. If it is large, then this implies that the two probability distributions are very different, meaning we gain much information by moving from the uniform distribution to the NLD. The largest value that $D_{KL}(f\,||\mathbb{P}^{(3)})$ can take is when f is some unit mass distribution, e.g. $f = \mathbb{P}^{(2)}$. Then we have $D_{KL}(\mathbb{P}^{(2)}\,||\,\mathbb{P}^{(3)}) = 1 \log \frac{1}{\frac{1}{|S_N|}} = \log |S_n| = \sum_{i=1}^{N} \log i$. We can thus compare $D_{KL}(\mathbb{P}^{(1)}|\mathbb{P}^{(3)})$ with $\sum_{i=1}^{N} \log i$. If the divergence in Eq. 7 is close to this maximum, then we will have an informative likelihood analysis:

$$D_{KL}(\mathbb{P}^{(1)}|\mathbb{P}^{(3)}) = \sum_{\sigma} \frac{\mathcal{L}(\sigma; r)}{\sum_{\sigma} \mathcal{L}(\sigma; r)} \log \frac{|S_N| \mathcal{L}(\sigma; r)}{\sum_{\sigma} \mathcal{L}(\sigma; r)}$$

$$= \sum_{\sigma} \frac{\mathcal{L}(\sigma; r)}{\sum_{\sigma} \mathcal{L}(\sigma; r)} \Big[\sum_{i=1}^{N} \log i - \log \frac{\sum_{\sigma} \mathcal{L}(\sigma; r)}{\mathcal{L}(\sigma; r)} \Big]. \tag{7}$$

Thus to quantify the information gained, we find the ratio of (7) and $\sum_{i}^{N} \log i$, which we can express as:

$$Y_2 = \frac{D_{KL}(\mathbb{P}^{(1)}|\mathbb{P}^{(3)})}{D_{KL}(\mathbb{P}^{(2)}|\mathbb{P}^{(3)})}$$

We observe that not all likelihoods of RPs give the same amount of information Y_1, Y_2. For example, we see in Fig. 6 that for a 6-car RP, $D_{KL}(\mathbb{P}^{(1)}\,||\,\mathbb{P}^{(3)}) = 0.078$ ('Divergence 1' from the uniform distribution), resulting in $Y_2 = 1.18\%$ ('Percent of max 1'), whereas for a 7 car RP (Fig. 7) with same l, ρ as the 6-car RP, has the same divergence of 1.13, which

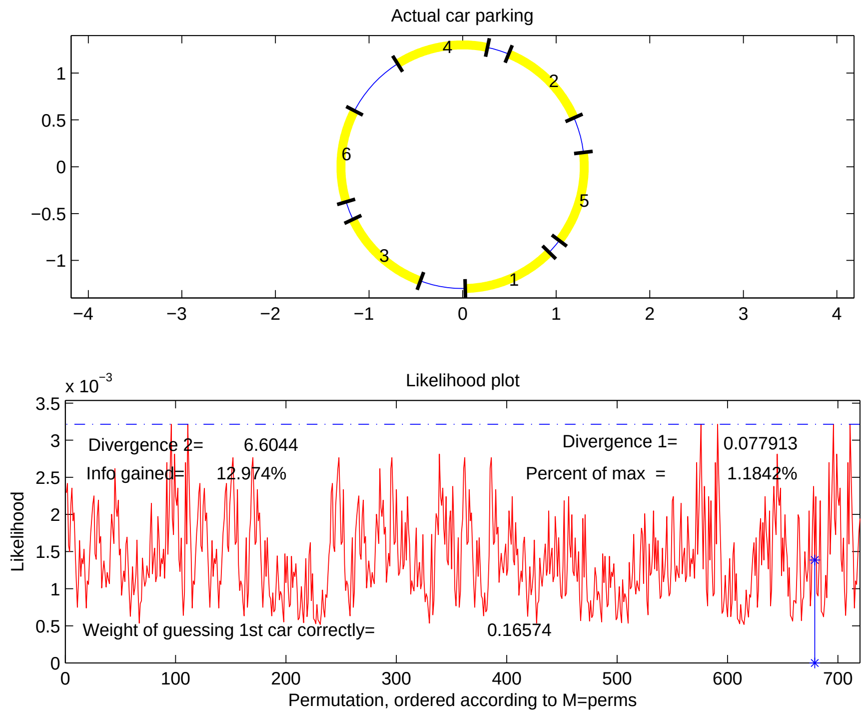

Figure 6: A 6-car RP with $l = 1, \rho = 1.5$. Below we see the likelihood function with divergences and weight of the first correct guess. Divergence 1 is $D_{KL}(\mathbb{P}^{(1)}|\mathbb{P}^{(3)})$, and divergence 2 is $D_{KL}(\mathbb{P}^{(2)}|\mathbb{P}^{(1)})$. The percentage of max is Y_2, and the information gained is Y_1. The function looks more or less flat, giving a low divergence from uniform distribution. The marker denotes the magnitude of the likelihood of σ^*.

corresponds to $Y_2 = 13.28\%$. Also for the divergence from the unit mass distribution, we see that the 6-car RP gives about $Y_1 = 13\%$, whereas the 7-car RP gives $Y_1 = 47.5\%$.

From observing these two examples, we conjecture that the geometry of the RPs affects these statistics. The 6-car RP has relatively even gaps with similar sizes compared to the 7-car RP, which has both very small to mid-sized gaps and a very large one. To study the relationship of geometry with Y_1, Y_2, we run many RP simulations with (l, ρ) fixed and plot $\{Y_i\}_{i=1}^2$ against different geometric factors.

First we look at $g_1 = \frac{N}{N_{max}}$, the *number* of cars parked normalized by the maximum number of parkable cars. We present boxplots of Y_2 for $\text{RP}_{(1,1)}$ against g_1 in Fig 8. We observe that the distribution of Y_2 'favors' higher g_1. For RPs with density $g_1^{(1)} = 4/6$, the maximum of Y_2 seems to be bounded from above at 15%, while that for RPs with density $g_1^{(3)} = 6/6$, the minimum of Y_2 seems to be bounded from below at around 10%. The median of Y_2 increases as g_1 increases, which we can interpret as resulting from the increased possibility of having more disordered RPs as more cars park. The increased disorder results in a NLD that has more curvature and thus gives more information.

So we see that the density of cars on the street influences how much average divergence the NLD will have with a flat distribution. We also look at Y_1, to see if any correlation could be made with some geometrical factor. From experiments, we see that there is correlation with

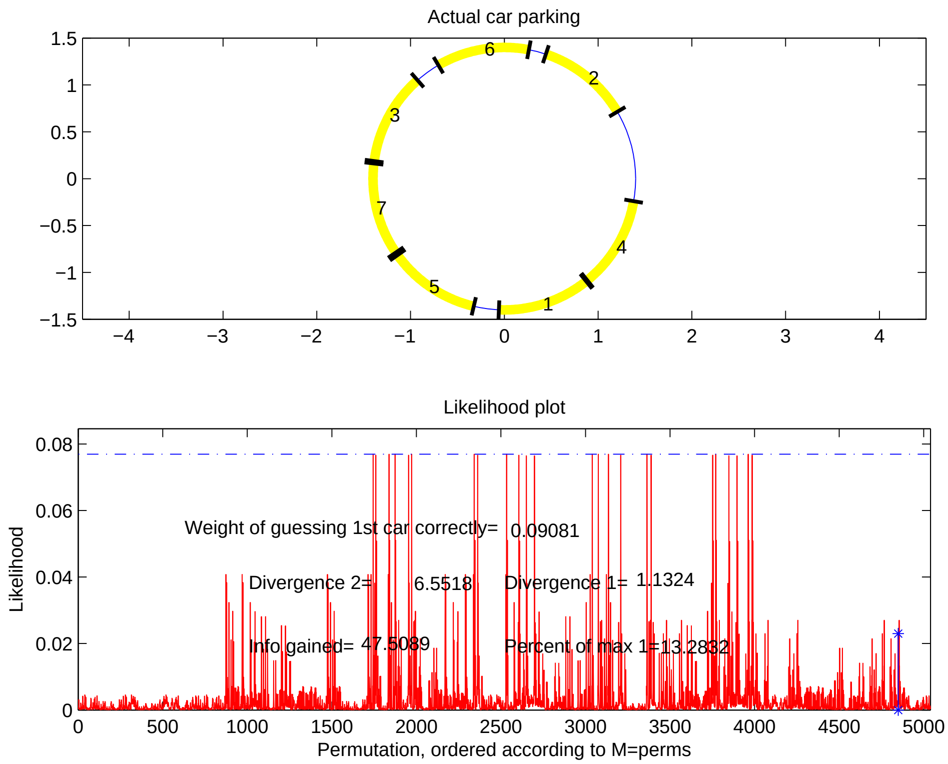

Figure 7: A 7-car RP with $l = 1, \rho = 1.4$. The likelihood function appears different from the uniform distribution, with a group of orders receiving more weight. The weight of likelihoods that guess correctly the first order is 0.09081. We see higher Y_1, Y_2 than before.

the *variance* of the gaps and Y_1. Fig. 9 shows many samples of $\mathrm{RP}_{(1,1)}$ with their variance of gaps and Y_1. While the scatter plot looks cluttered with not much correlation, one can see that it is indeed a superposition of three different scatter plots: one for each possible density $\{g_1^{(i)}\}_{i=1}^3$ (see Figs. 10-12).

Specifically, when we isolate the RPs according to their density g_1, the positive correlation between the variance of gaps and Y_1 is more pronounced, suggesting that if the gaps are evenly spaced, less information about σ^* can be gleaned from NLD compared to when the gaps are more unevenly spaced (see Figs. 10-12). To see the effect that geometry has on the likelihood, it is helpful to imagine trying to guess the order in which a set of *equally* distanced cars were sequentially parked on a circle. Even if it is known, say, that each subsequent car was parked adjacent to one another, an immediate ambiguity arises as to which car was the first to park due to symmetry.

3.2 Interpretation as information for inference

While the NLD can be studied in its entirety as a distribution compared to other distributions, parts of NLD can be analyzed to probabilistically answer inference problems. For instance, given a random parking r, we can look at the weight of likelihoods w_1^1 that guess

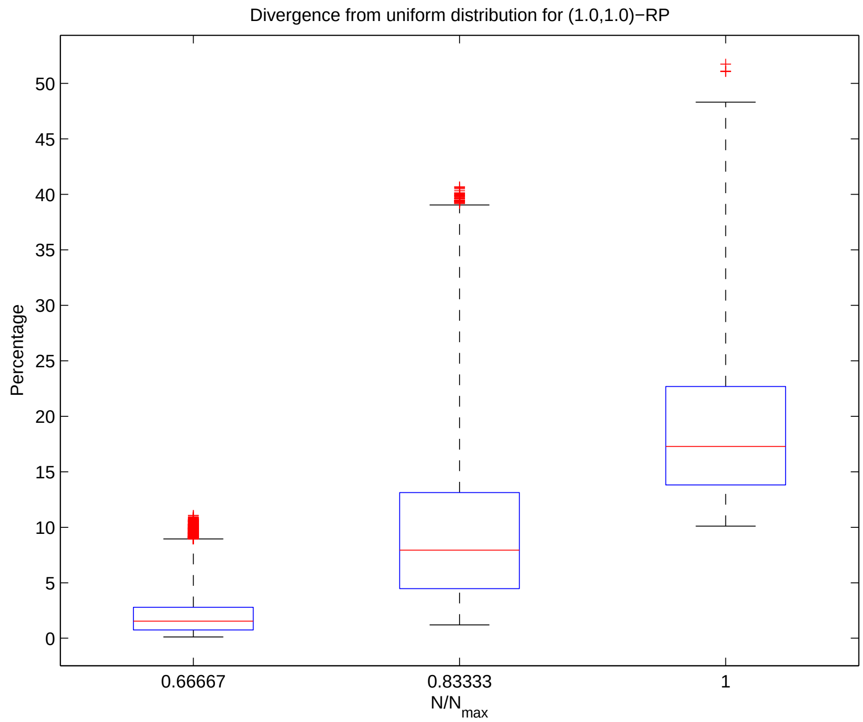

Figure 8: Boxplot of Y_2 against g_1 for $\mathrm{RP}_{(1,1)}$. 400,000 RP samples were taken, among which 127,832 had density $g_1^{(1)} = 4/6$, 265,553 had density $g_1^{(2)} = 5/6$, and $6,615$ had density $g_1^{(3)} = 6/6$. We see that the median Y_2 increases as the number of cars on the street increases, as well as the observed maximum Y_2. The crosses above are outliers.

correctly the first car:

$$w_1^1 = \frac{\sum_{\sigma:\sigma_1=1} \mathcal{L}(\sigma, r)}{\sum_{\pi} \mathcal{L}(\pi, r)},$$

where $w_i^j = \frac{\sum_{\sigma:\sigma_i=j} \mathcal{L}(\sigma,r)}{\sum_{\pi} \mathcal{L}(\pi,r)}$, with σ_i being the i^{th} element of permutation σ and j representing the true j^{th} car. To do this, we sum the likelihoods that correspond to permutations starting with 1, and divide by the sum of all likelihoods. We can extend this to look at the weight of likelihoods w_i^i that guess correctly the i^{th} car:

$$w_i^i = \frac{\sum_{\sigma:\sigma_i=i} \mathcal{L}(\sigma, r)}{\sum_{\pi} \mathcal{L}(\pi, r)}.$$

For example, we can look at Fig. 6 and 13. Note that a uniform distribution on the RP in Fig. 6 would give $w_1^1 = \frac{1}{6}$, which is slightly bigger than that from the likelihood function, 0.16574. Here is an instance of where the guessing the absolute order of car arrival (e.g. the 1^{st}, 2^{nd} car) using likelihood is more or less similar to guessing purely randomly. The next example of the 8-car RP in Fig. 13 is an instance where it is actually better: $w_1^1 = 0.14541 > \frac{1}{8}$.

This type of weight of likelihood can then be used to answer the problem of identifying the first car, for example, by finding:

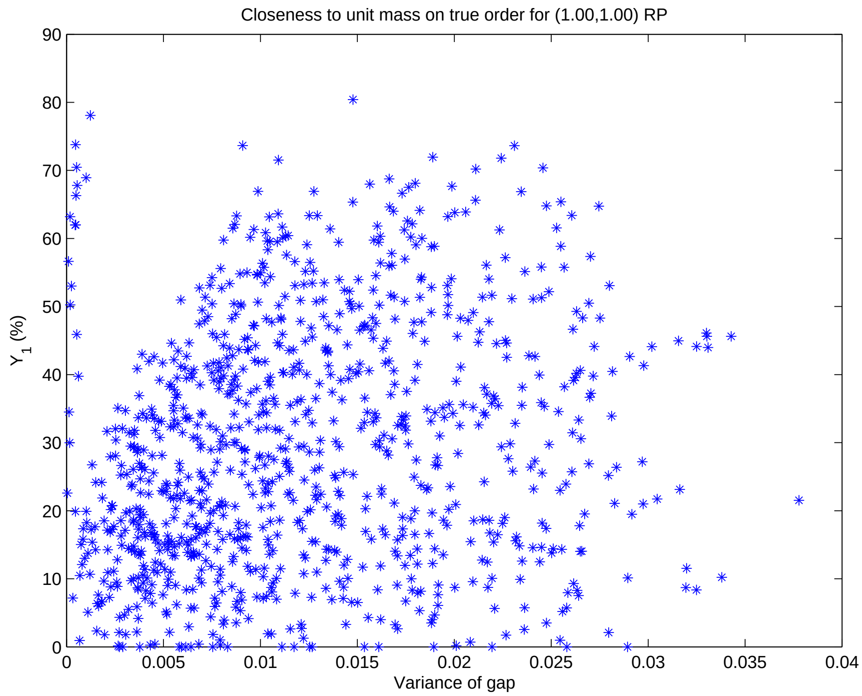

Figure 9: Scatter plot of Y_1 against variance of gaps for $\mathrm{RP}_{(1,1)}$. Although there does not seem to be a strong correlation, this is a superposition of three scatter plots shown in the next three figures, where it is clearer to see the correlation.

$$c_1 = \arg\max_{i\in[N]} \Big(\sum_{\sigma:\sigma_1=i} \mathcal{L}(\sigma, r) \Big) = \arg\max_{i\in[N]} (w_1^i),$$

where $[N] = [1, 2, 3, ...N]$. We do this by summing up the likelihoods that assign car index i to be the first, for each $i = 1$ through N, and finding the index c_1 whose sum attains the maximum. Likewise, we can look at to which car c_n the likelihood analysis gives most weight to be at the nth order,

$$c_n = \arg\max_{i\in[N]} \Big(\sum_{\sigma:\sigma_n=i} \mathcal{L}(\sigma, r) \Big) = \arg\max_{i\in[N]} (w_n^i).$$

If the likelihood analysis is informative, we expect $c_1 = 1$ and $c_n = n$ (the correct cars).

As seen on Table 1, we run a small numerical experiment using $\rho = 1, l = 1$, where we simulate 40,000 random parkings for each density $g_1^{(1)}, g_1^{(2)}, g_1^{(3)}$, find the number of times k that $c_n = n$, and calculate $p = k/40,000 \approx \mathbb{P}(c_n = n)$.

We also report in the fourth column the mean of w_n^n from the same 40,000 simulations along with the standard deviation, and remark that it is smaller than p, suggesting that the distribution of w_n^n for fixed N car parking must have a left tail that brings down the mean. Also note that the likelihood analysis gives p that is higher than what random guessing would give, since for $N = 4$, p would center around 0.25, for $N = 5$ around 0.2, and for $N = 6$ around 0.16, suggesting that using the likelihood function improves our guess of

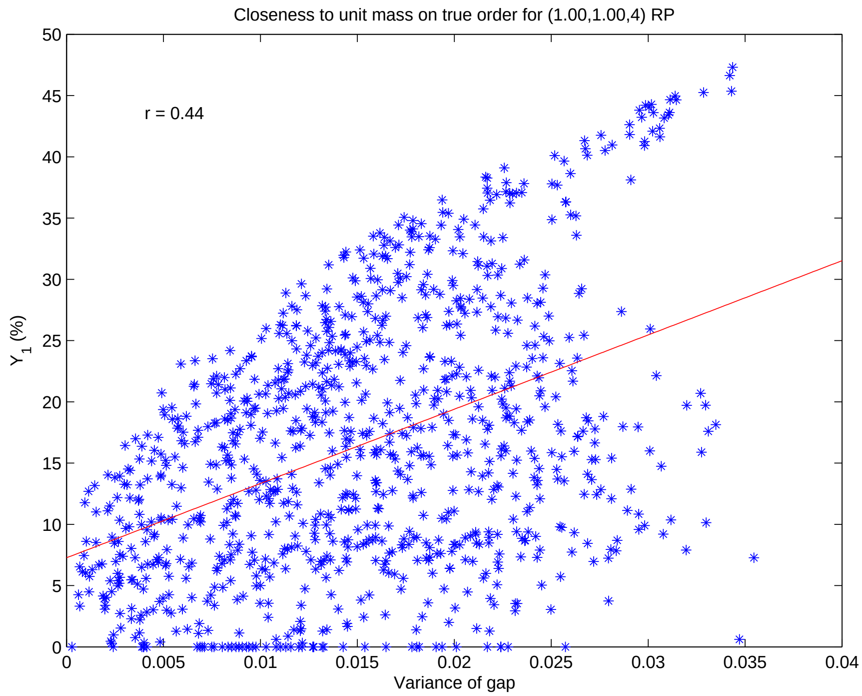

Figure 10: Scatter plot of Y_1 against variance of gaps for $\mathrm{RP}_{(1,1)}$ with density $g_1^{(1)}$. We see consistently high Y_1 after the variance of the gap exceeds about 0.027, and the linear correlation is $r = 0.44$.

absolute arrival order of single cars.

Finally, we can also look at how much weight of likelihood is given to guessing correctly the relative arrival orders. For example, we can ask how much likelihood w_{12} is given to the permutations that correctly identify the relative order of two cars, e.g., the first two to arrive:

$$w_{12} = \frac{\sum_{\sigma:\sigma(1)<\sigma(2)} \mathcal{L}(\sigma, r)}{\sum_{\pi} \mathcal{L}(\pi, r)},$$

where $\sigma(i)$ is the index of element i in the permutation. To study the statistics of these numbers, we compute the empirical distribution of w_{ij}s, for different i and j by running many RP simulations with fixed l, ρ, N. For instance, we see that for $l = 1, \rho = 1$, and

Table 1: Estimate p of $\mathbb{P}(c_n = n)$ for $\mathrm{RP}_{(1,1)}$.

N	n	p	$\widehat{w_n^n}$	SD of w_n^n
4	1	0.28208	0.25448	0.03136
4	2	0.27973	0.25456	0.03145
5	1	0.25503	0.21576	0.04534
5	2	0.25553	0.21565	0.04522
6	1	0.21005	0.1806	0.04223
6	2	0.21220	0.1805	0.04258

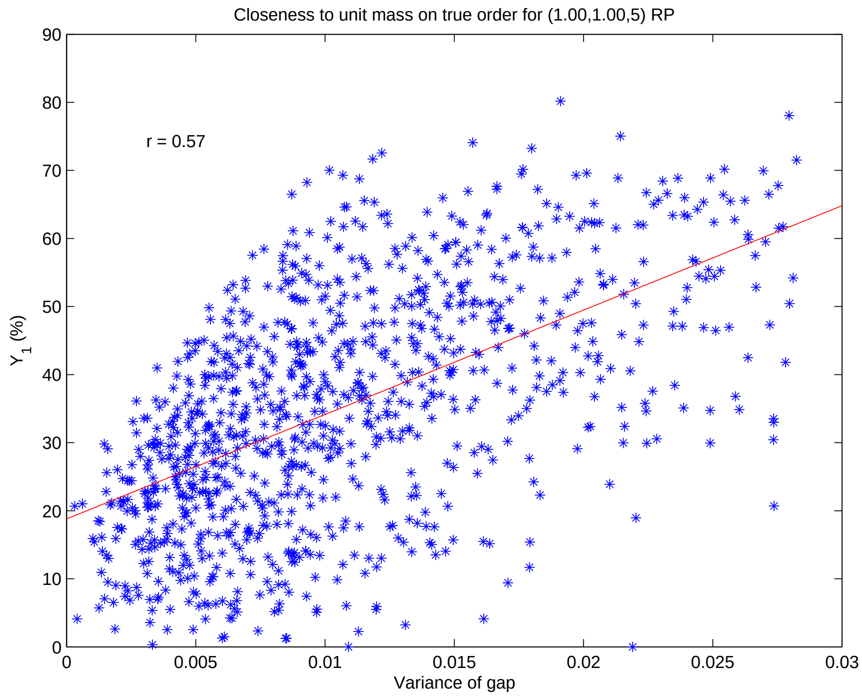

Figure 11: Scatter plot of Y_1 against variance of gaps for $\mathrm{RP}_{(1,1)}$ with density $g_1^{(2)}$. We observe a positive correlation ($r = 0.57$) between the variance of gap and Y_1.

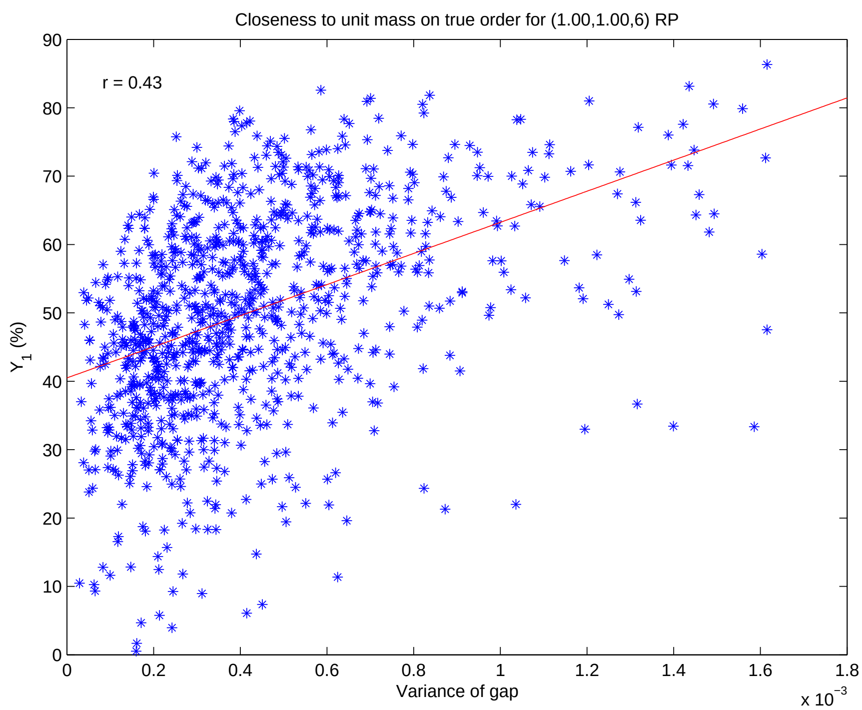

Figure 12: Scatter plot of Y_1 against variance of gaps for $\mathrm{RP}_{(1,1)}$ with density $g_1^{(3)}$. Again we see a positive correlation ($r = 0.43$) between the variance of gap and Y_1.

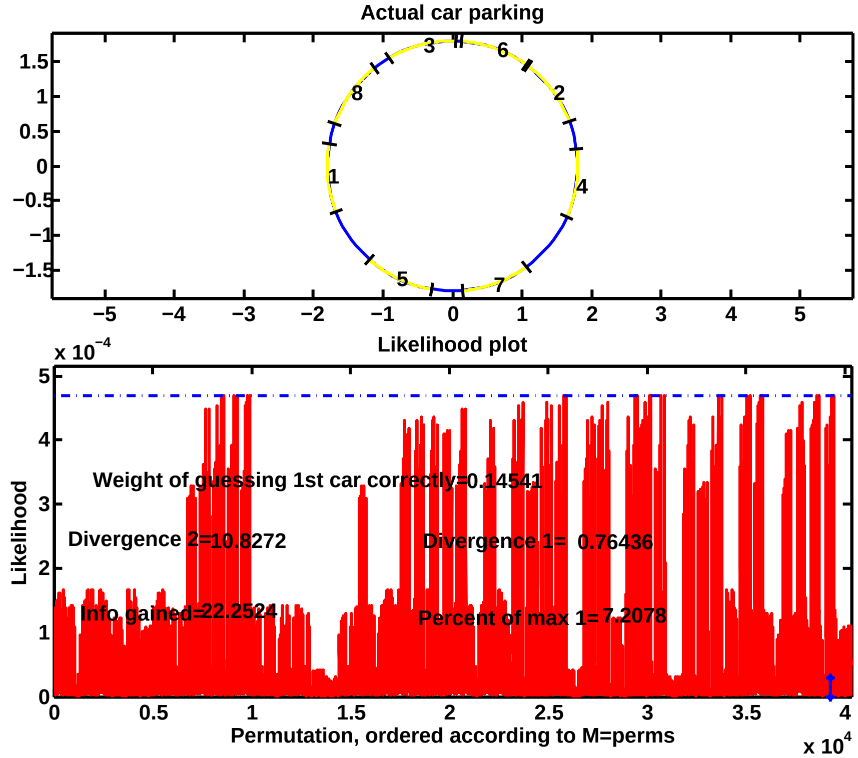

Figure 13: A 8-car RP and its likelihoods with divergences and weight of first correct guess. We see that the weight of likelihood of permutations guessing correctly the first car is more than $\frac{1}{8} = 0.125$ which uniform distribution would give.

$N = 4, 5, 6$, the distribution of w_{12} is centered around a half with a larger spread in both directions as N increases (Figs. 14, 30, 31 in Appendix). The centering around half means that NLD is similar to uniform guessing, while the increasing spread as N increases shows increasing variance of the 1^{st}-2^{nd} order guessing performance of likelihood analysis. But for certain i and j, we have more skewed distributions with positive deviation from a half, indicating better predictive analysis. For instance, with fixed $l = 1, \rho = 1$, we look at w_{1N}, for different N: $4, 5, 6$ (i.e. guessing the relative order of first and last car). We see in Figs. 15, 32, and 33 (in Appendix) that there is more mass towards the right side of half, when uniform guessing would give only a half. This seems to suggest that the likelihood analysis does better at guessing relative order of cars that came at a larger time interval.

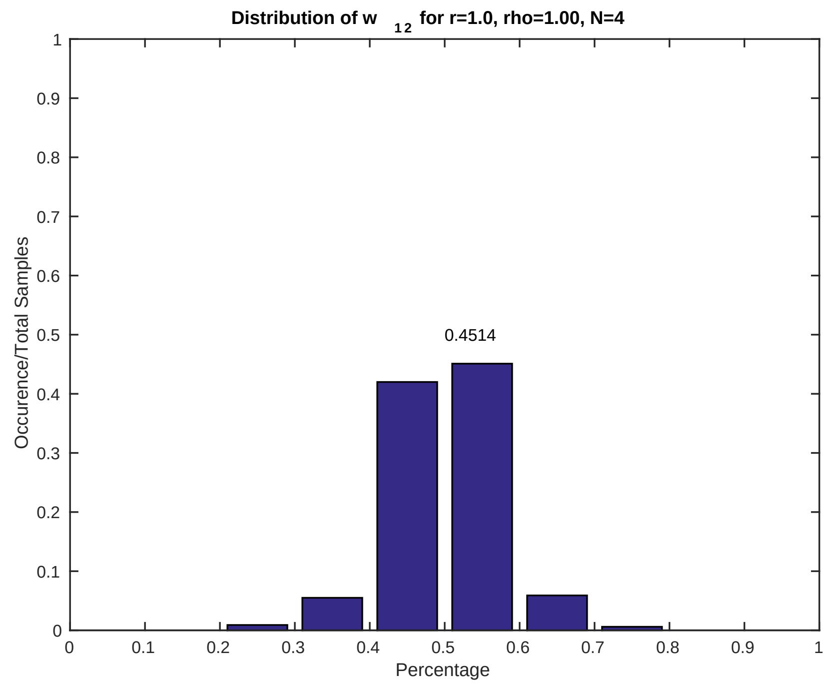

Figure 14: Distribution of w_{12} for $l = 1, \rho = 1, N = 4$. We see a centering around half, with slightly more mass towards the right of 0.5. The mean w_{12} is 0.4514.

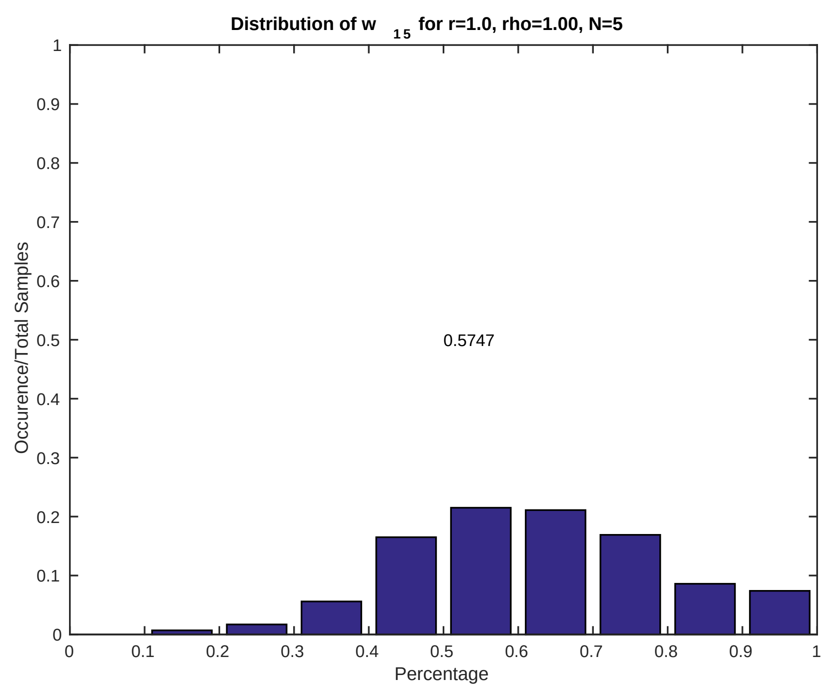

Figure 15: Distribution of w_{15} for $l = 1, \rho = 1, N = 5$. We see more mass towards the right of 0.5, indicating better prediction ability than 50%. The mean is 0.5747.

4 Inference

4.1 Problems

As mentioned above, we can pose problems about relative orderings of two cars. Although a lot of such problems can be asked, we present the ones that have a high probability of being inferred correctly.

1. Idenfity the pair of cars with the minimum gap. Which one came earler?
2. Identify the car that has clockwise and counterclockwise adjacent cars c_1, c_2 such that $w_{c_1c_2}$ is the largest out of all such pairs (that have a car in between). Which car out of c_1 and c_2 came earlier?
3. Identify the pair of adjacent cars c_1, c_2 such that $w_{c_1c_2}$ is largest out of all adjacent pairs. Which one came earlier?
4. Given that two cars are the first and the last, which is which?

We can also ask problems about relative orderings of *groups* of cars. Consider the following 'Two-color problem':

- Given that a set of cars to park are half (or $\lfloor N/2 \rfloor$ if N is odd) white and half black, and that all cars of the same color sequentially enter the street first with all the other colored cars parking second, which color was the first?

4.2 Likelihood method

4.2.1 Problems 1–4

We can use the likelihood function to probabilistically answer the above problems. For problems 1 through 4, the likelihood method involves first identifying the two cars c_1, c_2, computing $w_{c_1c_2}$ and $w_{c_2c_1} = 1 - w_{c_1c_2}$, and guessing with probability $w_c = \max\{w_{c_1c_2}, w_{c_2c_1}\}$ that c_1 came first if $w_{c_1c_2} > w_{c_2c_1}$ or c_2 if otherwise. Note: Since this method is a probabilistic method, it can be viewed as a forecast of a binary event. Forecasts are rated by skill scores, so we assess the accuracy of the likelihood method by using *Brier* skill scores (BS) [9]. In our context, BS is defined as

$$BS = \frac{1}{k} \sum_{i=1}^{k} ((w_c)_i - I_i)^2,$$

where k is the number of forecasts made and I_i is whether the i^{th} forecast was correct (1 if yes, 0 if not). Thus a small BS close to zero is an indicator of a good forecasting method. For example, the uniform distribution guesses the correct answer to problems 1)-4) with 50% accuracy, and the corresponding BS would be $0.25 = \frac{1}{k} \sum_{i=1}^{k} 0.5^2$.

4.2.2 Monte Carlo for large N

When finding $w_{c_1c_2}$ for RPs of size $N \geq 10$, we must resort to Monte Carlo methods (MC) instead of directly computing $w_{c_1c_2}$, as we run to memory problems due to the size of the permutation matrix. Specifically, given a RP of size N, we implement a Metropolis Hastings algorithm to sample from S_N according to the NLD, a random variable we will call Σ. We can use this method since

$$\begin{aligned} w_{c_1c_2} &= \mathbb{P}(\Sigma \in \Pi_{12}) \\ &= \mathbb{E}\mathbf{1}_{\Pi_{12}}, \\ &\approx \frac{1}{M}\sum_{i=1}^{M}\mathbf{1}_{\sigma_i\in\Pi_{12}} \end{aligned} \tag{8}$$

where $\{\sigma_i\}_{i=1}^{M}$ is the Markov chain with $M >> 0$ steps, and $\Pi_{12} = \{\sigma \in S_N \,|\, \sigma(c_1) < \sigma(c_2)\}$.

Our proposed move is a lazy random transposition as suggested by Diaconis [11]:

$$\begin{cases} \sigma_0 \sim U(S_N), \text{uniform over } S_N \\ \mathbb{P}(\sigma_i \to \sigma_{i+1} = \sigma_i) = \frac{1}{N}, \\ \mathbb{P}(\sigma_i \to \sigma_{i+1} = \sigma_i \cdot t) = \frac{2}{N^2}, \ \text{ where } \sigma_i \cdot t \text{ is transposition on } \sigma_i \\ \mathbb{P}(\sigma_i \to \pi) = 0, \ \text{ if } \pi \text{ is neither } \sigma_i \cdot t \text{ nor } \sigma_i. \end{cases}$$

We accept the transition move with the following acceptance criteria: Let σ_{i+1} be the proposed move, then

- Accept σ_{i+1} with probability $p = \min\{1, \frac{\mathcal{L}(\sigma_{i+1})}{\mathcal{L}(\sigma_i)}\}$
- else, $\sigma_{i+1} := \sigma_i$ and propose another move.

To illustrate that the MC method works correctly, we show several things: the convergence of the running mean of a function of σ_i, the extent of mixing across the sample space, the autocorrelation, and an approximation of a NLD for a low N-RP using the MC method, which we compare with the exact NLD. First, to observe a running mean of the Markov chain, we must decide on a function $d : S_N \to \mathbb{R}$. To do so, we assign each permutation with its distance away from the identity permutation. While there are several metrics in S_N, we choose the Kendall-Tau (KT) metric d_K [12], as it can take on more values and contains more information about the difference in relative orderings between two permutations ϕ, π:

$$d_K(\phi, \pi) = \frac{1}{N(N-1)}\sum_{i,j=1}^{N}\mathbf{1}(\phi(i) > \phi(j) \wedge \pi(i) < \pi(j)),$$

where $N(N-1)$ is a normalization factor such that $d_K(\phi, \pi) \in [0, 1]$.

Using this distance, we calculate the running mean of $\{d_K(\sigma_i, I)\}_{i=1}^{M}$, where I is the identity permutation. We see a convergence towards a mean $\mathbb{E}(d_K(\Sigma, I))$ (Fig. 16).

Similarly, we can plot $\{d_K(\sigma_i, I)\}_{i=1}^{M}$ against i and we see that the Markov chain samples both permutations whose distances from I are large and small, mixing well throughout the

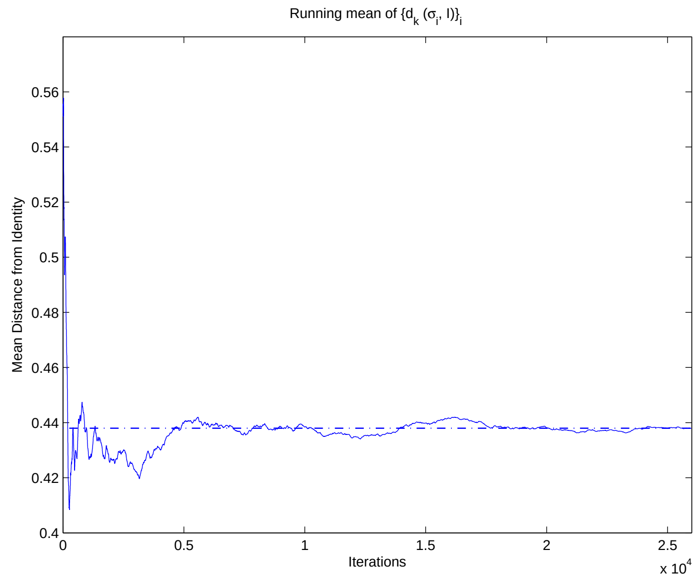

Figure 16: Running mean of distance away from I of a Metropolis MC of NLD of a 13-car RP. The first few steps of MCMC do not converge to the mean distance from I of permutations distributed by the NLD, but after 5,000 steps, the running mean distance starts to converge.

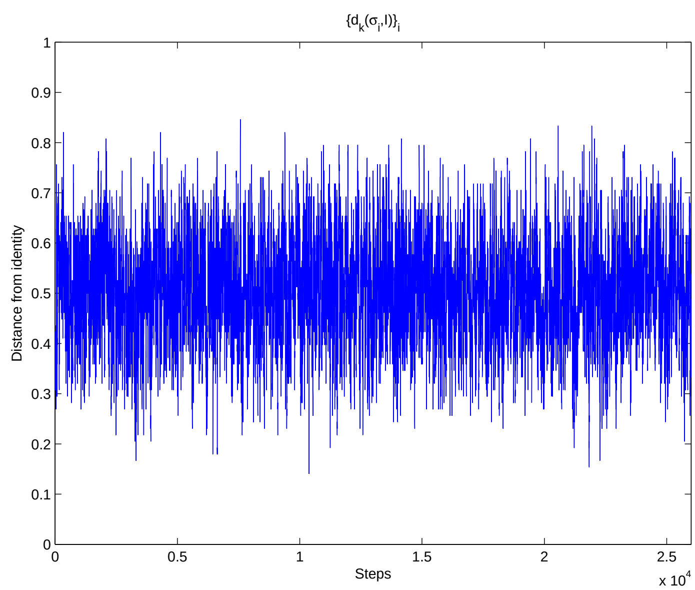

Figure 17: Sampling of S_N by Metropolis MC, $\{\sigma_i\}_{i=1}^{26000}$, for sampling NLD of 13-RP. The Kendall tau distance of MC steps from I is plotted and we see a thorough mixing across different possible distances from 0 to 1

whole space (Fig. 17). Now to see the rate of convergence, we can look at the autocorrelation function, which decays exponentially (Fig. 18). Results of autocorrelation time computed using method by Sokal [13] consistently give less than 10 steps. The autocorrelation time $\hat{\tau}$

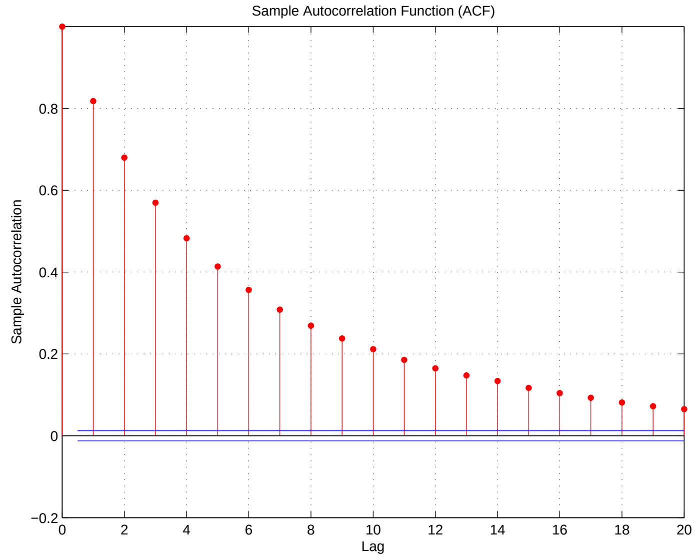

Figure 18: Autocorrelation function of Metropolis MC of $\{d_K(\sigma_i, I)\}$ for NLD of a 13 car-RP.

is computed by letting $\lambda(t) = \begin{cases} 1 & \text{if } |t| \leq N \\ 0 & \text{if } |t| > N \end{cases}$, where N is some integer, and computing

$$\hat{\tau} = \frac{1}{2} \sum_{-(M-1)}^{M-1} \lambda(t)\hat{\rho}(t),$$

where M is the total number of steps and $\rho(t)$ is the normalized autocorrelation function $C(t)/C(0)$, with $C(t)$ being the autocorrelation function. The integer N is computed as the smallest integer such that $N \geq c\hat{\tau}(N)$, where $c \approx 6$ [13]. A Metropolis MC for a 13-car RP gives $\hat{\tau} \approx 1$. Also, we calculate the acceptance ratio (acceptance of different move), and we see that simulations give acceptance rates between 58% and 60%. Finally, we run MC on a 6-car RP and draw a histogram, comparing it to the NLD which can be directly obtained. We see a good approximation in Fig. 19.

4.2.3 Two-color problem

For the two color problem the likelihood method involves summing up the likelihoods of permutations that correspond to $N/2$ (or $\lfloor N/2 \rfloor$ if N odd) single colored cars being the first half cars to arrive and comparing this sum s_1 to the sum of likelihoods s_2 corresponding to $N/2$ (or $\lfloor N/2 \rfloor + 1$) cars of the other color being the first half to arrive. Then we guess with probability $s = \max\{s_1, s_2\}/(s_1 + s_2)$ that color 1 came first if $s_1 = \max\{s_1, s_2\}$ or that color 2 came first if $s_2 = \max\{s_1, s_2\}$. This method works for systems of low $N \leq 10$. For large N we were not able to develop a suitable MC method.

4.2.4 Results for problems 1–4

Using either direct calculation of $w_{c_1c_2}$ or using Eq. 8, we answer the problems using the likelihood method of probabilistic prediction in section 4.2.1. We present the results on the

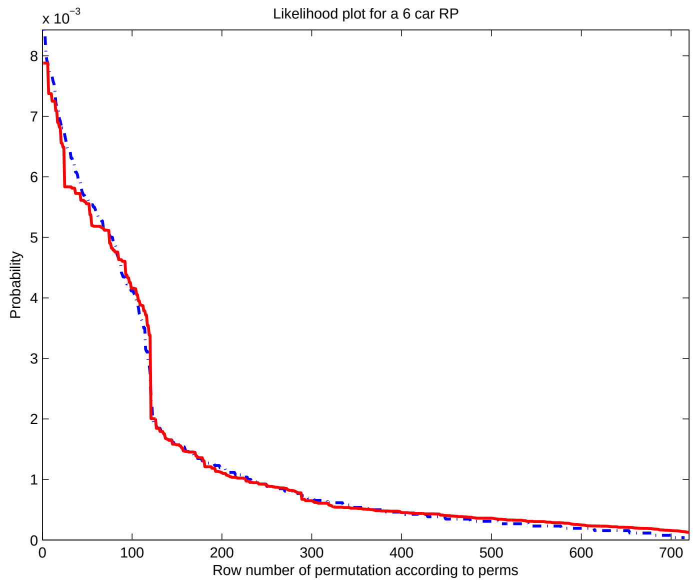

Figure 19: Approximation of NLD using MC method for 6-car RP. The x-axis is the row number of each permutation ($6! = 720$) according to MATLAB's `perms` matrix. The blue dotted line is the approximation using histogram of MCMC and the red solid line is the actual NLD.

following pages, from Figs. (20)-(23), which show that the likelihood method performs *better* than random guessing, exceeding 70% answer rate around $\rho = 6$ for problems 2-4 and 60% for problem 1. We remark that all problems show steady increase of accuracy as the street radius ρ increases. Because of the computational cost, we run 1000 samples of RP for each radius. We put all four plots (without error bars for clarity) on the same plot in Fig. 24 and see that problem 1 has the lowest accuracy, whereas the other three problems are answered with similar accuracy using the likelihood method.

4.2.5 Results for two-color problem

We sample $10,000$ unit length car RPs for small radii and apply the likelihood method for the two color problem to each RP. Since we can find exact likelihood of RP up to 10 cars, we range our radius from $\rho = 1$ to 1.7, at which $N_{max} = 10$. Similarly, we see from Fig. 25 that accuracy is consistently higher than 50% ($\geq 65\%$) and increasing as street radius increases.

4.3 Geometric estimation

4.3.1 Problems 1–4

The limitation to the likelihood method in estimating the answer to these problems is the computational cost. Even though there is way to calculate $w_{c_1c_2}$ for RPs of large N using the MC method, the Markov chain also does not have a low computational cost. Therefore, to run many samples of RP and answer inference problems using the likelihood method is impractical for N large (for $N \geq 8$ computations exceed 1 second.) However, using the fact

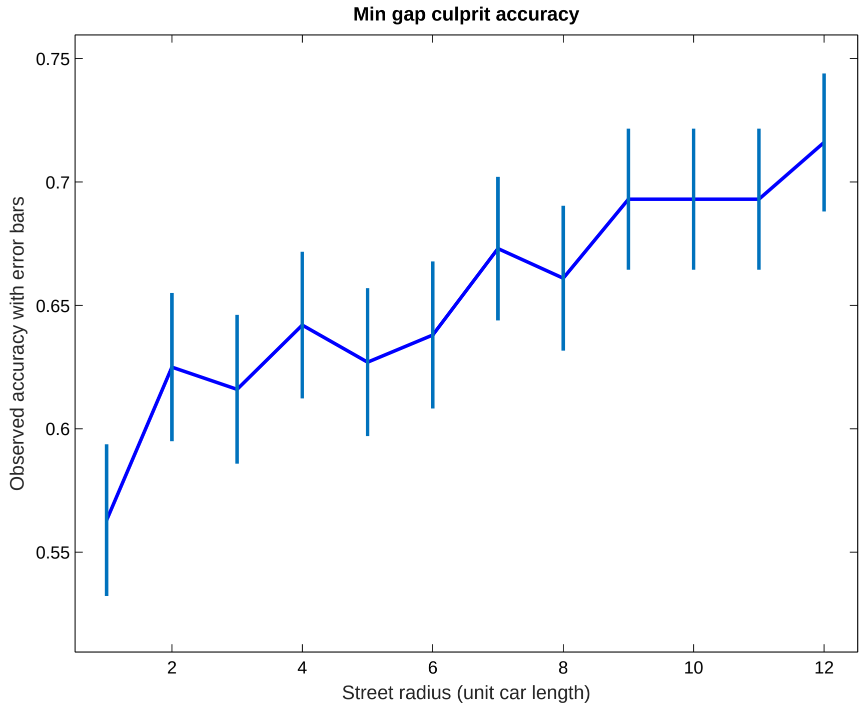

Figure 20: Accuracy of predictions to Problem 1 for RPs with $l = 1$ and $\rho = 1$ to 12. *Accuracy* is defined as the proportion of correct guesses of relative order in the 1000 simulations.

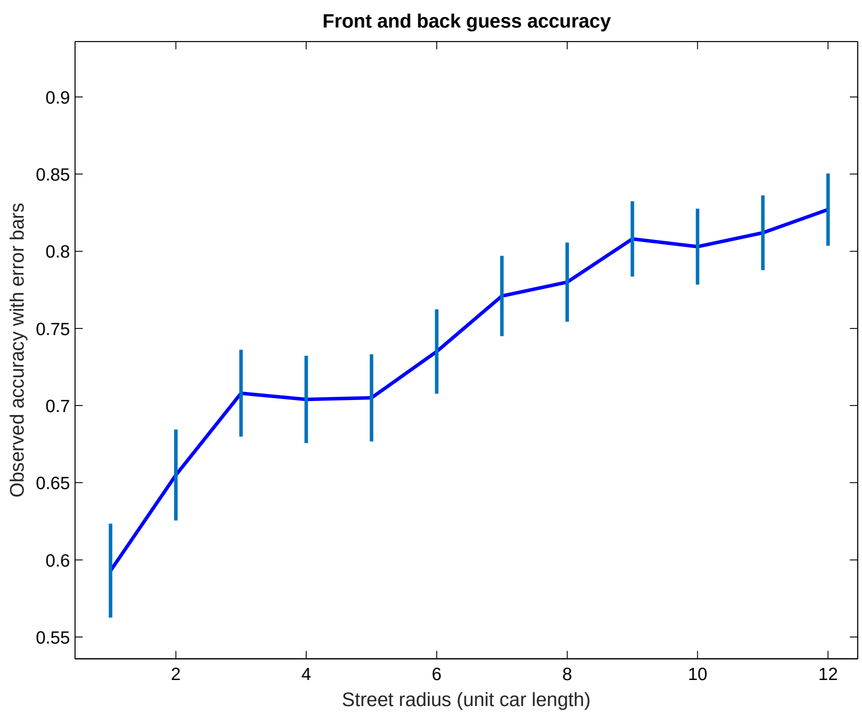

Figure 21: Accuracy of predictions to Problem 2 for RPs with $l = 1$ and $\rho = 1$ to 12.

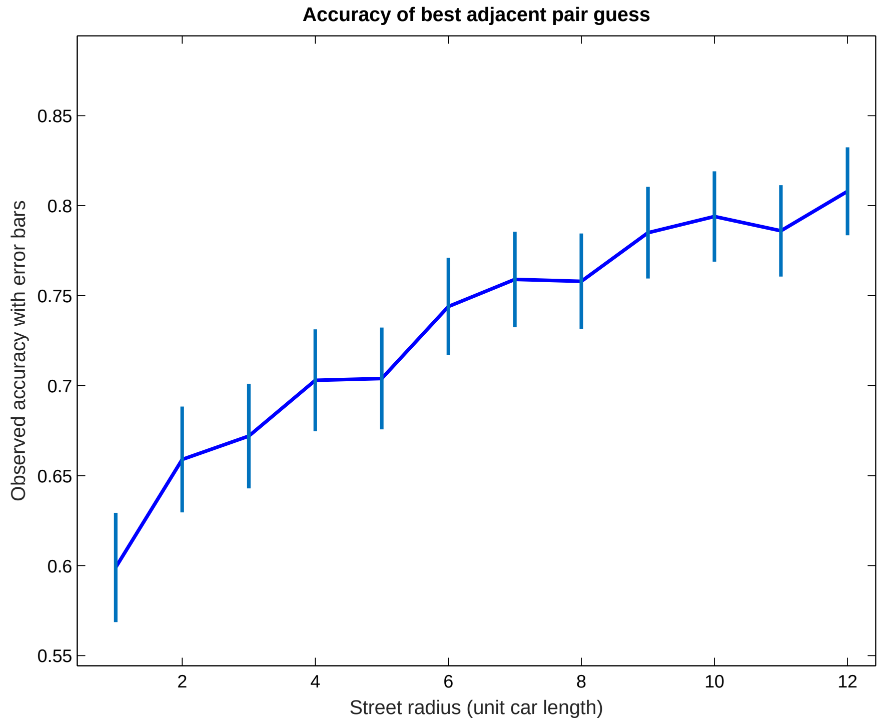

Figure 22: Accuracy of predictions to Problem 3 for RPs with $l = 1$ and $\rho = 1$ to 12.

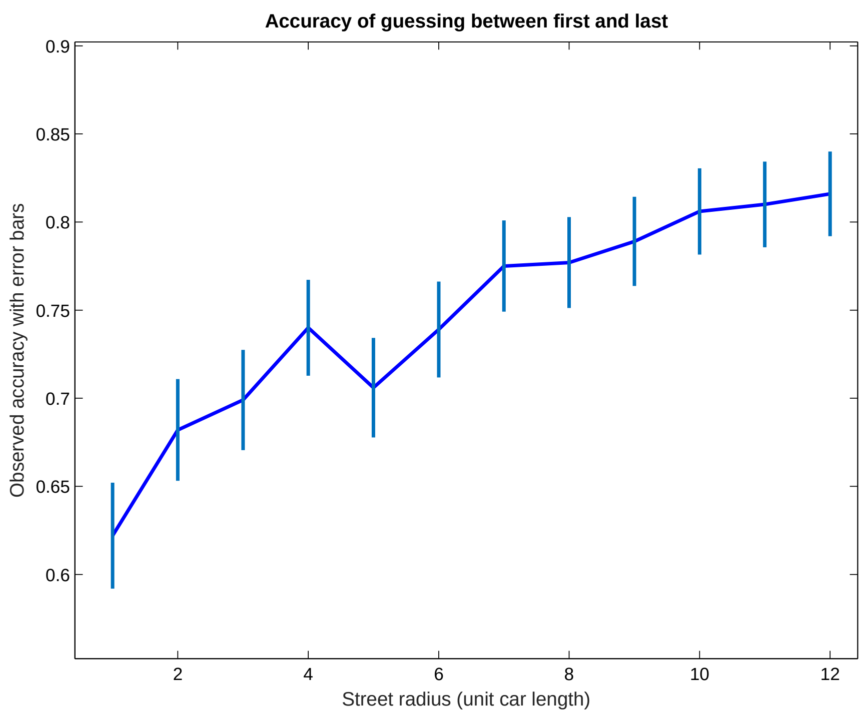

Figure 23: Accuracy of predictions to Problem 4 for RPs with $l = 1$ and $\rho = 1$ to 12.

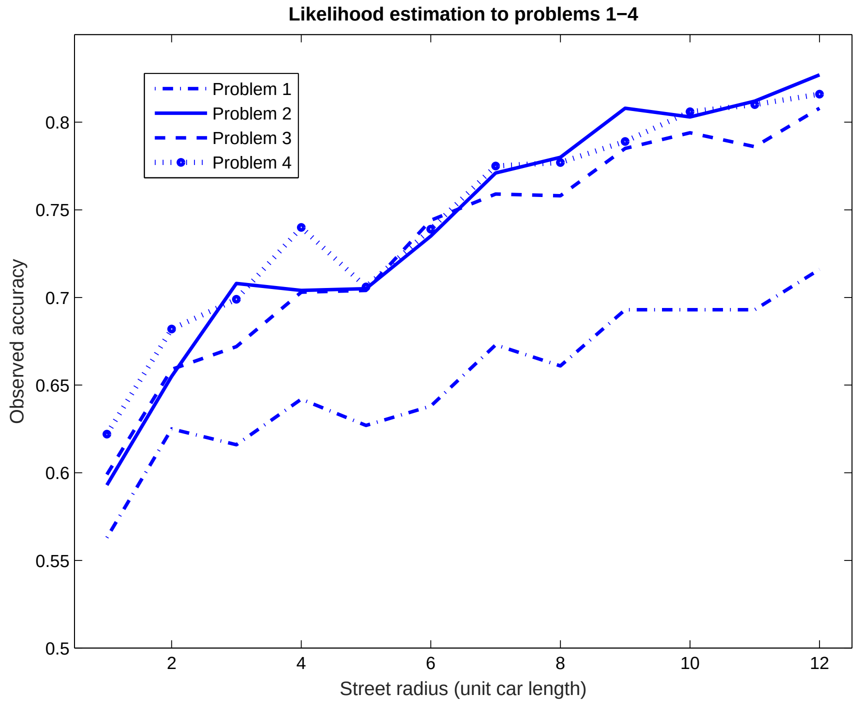

Figure 24: Accuracy of predictions to Problems 1-4 for RPs with $l = 1$ and $\rho = 1$ to 12.

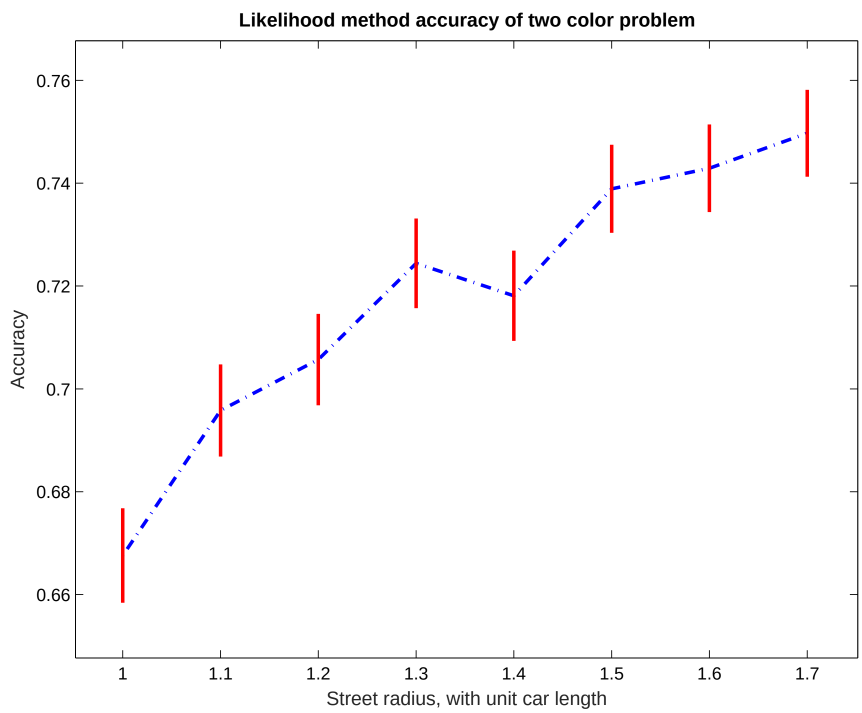

Figure 25: Accuracy of predictions to two color problem using direct likelihood computation for small $\rho = 1$ to 1.7.

that in a RP, the earlier arriving cars will have more space to park compared to the latter ones (recall that $\Phi(t) \to 0$ as $t \to \infty$ as discussed in 1.1), we can construct methods relying on the geometry of RP to estimate the answers to all the inference problems. Of course, problems 2-3 must be restated now since they involve computing $w_{c_1 c_2}$ and this geometric method will not involve computation of likelihoods. To restate the problems, let c_i^c denote the clockwise adjacent car of c_i and c_i^{cc} denote the counterclockwise adjacent car. Also, let $\gamma(c_i)^{(1)}$ and $\gamma(c_i)^{(2)}$ be the left and right gap of car c_i. Define a surrounding gap $\Gamma(c_i)$ to be $\Gamma(c_i) := \gamma(c_i)^{(1)} + \gamma(c_i)^{(2)}$.

2. Identify the car $c_{i'}$ that maximizes $|\Gamma(c_i^c) - \Gamma(c_i^{cc})|$. Which out of $c_{i'}^c$ and $c_{i'}^{cc}$ came earlier?

3. Identify the car $c_{i'}$ that maximizes $|\Gamma(c_i) - \Gamma(c_i^c)|$. Which out of $c_{i'}$ and $c_{i'}^c$ came earlier?

Then, once we are given $\{c_1, c_2\}$ whose relative order we wish to find, we estimate the order to be $[c_M, c_m]$, where c_M came earlier, where c_M and c_m are such that $\Gamma(c_M) = \max\{\Gamma(c_1), \Gamma(c_2)\}$ and $\Gamma(c_m) = \min\{\Gamma(c_1), \Gamma(c_2)\}$.

4.3.2 Two-color problem

For the two color problem, there are several geometric ways to estimate the order of the colors. First, we can use a similar method as above. This method is inspired from the fact that the available space $\Phi(t)$ goes to 0 as time progresses.

1. Identify car c such that $\Gamma(c) = \min\{\Gamma(c_i)\}_{i=1}^N$. Let the color of c be the second color. (The 'Minimum surrounding gap method.')

Another method is to observe the two-colored cars separately. This method is inspired from the observation that the first half cars can spread themselves out more than the second half cars, which are restricted to park in between the already parked first half cars:

2. Let $m_1 = \min\{\gamma(c_i)^{(1)}\}_{c_i \in G_1}$ and $m_2 = \min\{\gamma(c_i)^{(1)}\}_{c_i \in G_2}$, where G_j is the set of cars with only color j for $j = 1, 2$. Let j' be the second color where $m_{j'} = \min\{m_1, m_2\}$. ('The minmin method.')

The final method also looks at the two sets of cars separately. This method aims at exploiting the difference in the parking behavior of the first and second half of cars:

3. Find distribution Δ_1 of $\gamma(G_1)^{(1)}$ and Δ_2 of $\gamma(G_2)^{(1)}$, where $\gamma(G_i)^{(1)}$ is the set of gaps between cars of only color i (measuring gap by ignoring different colored cars that may be in between). Compare Δ_i with $G(h, t)$ at $t = t^*$, where t^* is the time that $\rho(t^*)$ is half of Renyi's constant, and choose i that matches more closely to be the first color. ('The gap density method.')

We further explain how exactly the gap density method works. We choose l and k_a to be 1 in both our simulations and in the formula for $G(h, t^*)$ (see section 1.1). The distribution Δ_i is a histogram, and we choose evenly spaced centers $\{h_j\}$ on the x-axis, find

$$\min_{i \in \{1,2\}} \sum_j |\Delta_i(h_j) - G(h_j, t^*)|,$$

and choose the i that gives the minimum as the first color.

4.3.3 Results for problems 1–4

We observe that the geometric method gives surprisingly accurate answers for problems $1-4$ (Fig. 26). The topmost line corresponds to the problem of estimating relative order between the first and last cars; the second from the top line to that of relative order between clockwise and counterclockwise neighbors of the 'best' car (difference in Γ of the pair is largest); the third line to that of relative order between cars with minimum gap; and the last line to that to relative order of the 'best' pair of adjacent cars. We see that the geometric method far outperforms the likelihood method in problem 4, as by radius 5, the geometric method has already exceeded 90% accuracy, while the likelihood method has barely exceeded 70%. For the relative order of adjacent pairs (problem 3), however, the likelihood method outperforms the geometric method, reaching close to 80% at radius 12, whereas the geometric method gives around 65% probability of correct answer at the same radius. The question of relative order between clockwise and counterclockwise cars (problem 2) seems to be answered with similar accuracy by both likelihood and geometric method.

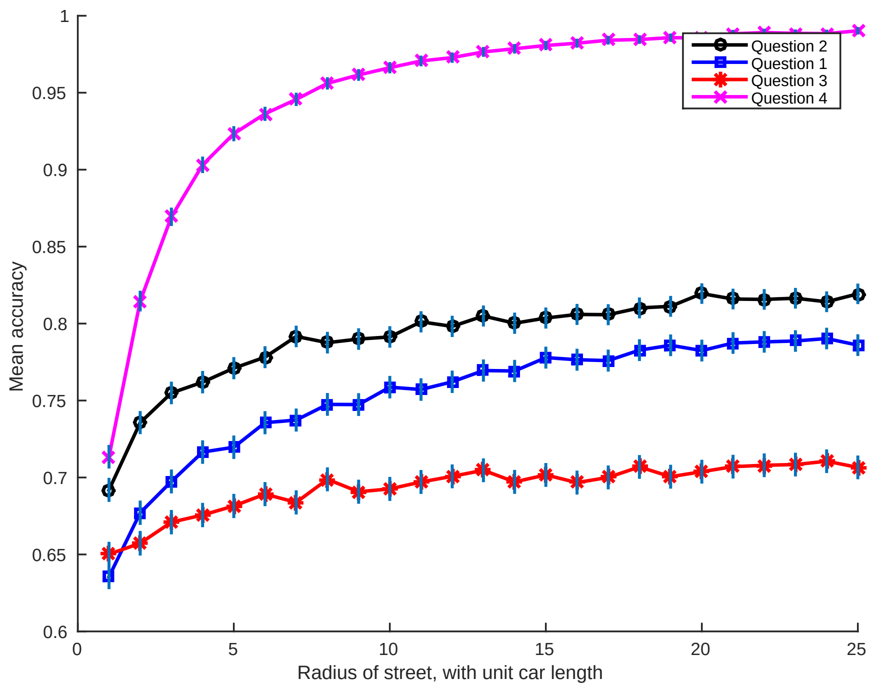

Figure 26: Accuracy of predictions to problems 1-4.

4.3.4 Results for two-color problem

We first present results of implementing method 1-2 in Figs. 27. The line that has the greater increase in accuracy from street radius 1 to 25 is the minimum surrounding gap method, and the method with slowly increasing accuracy is the minmin method. We see that these two geometric methods give consistently high accuracies, and we see that the likelihood method and the minimum surrounding gap method perform similarly for small radii around 1. The fact that the minmin method gives high accuracy tells us that the latter half cars tend to

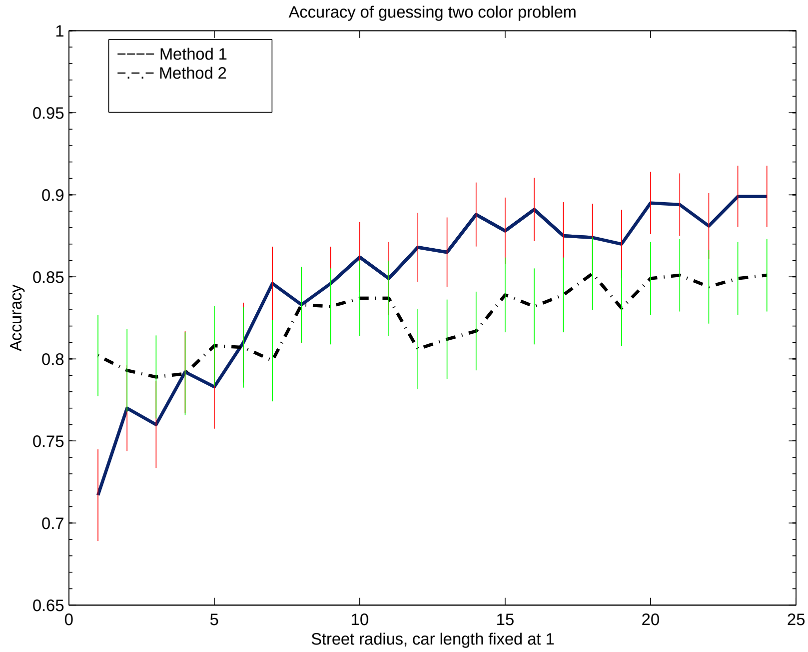

Figure 27: Accuracy of predictions to two color question using geometrical methods 1-2. The error bars are 1.95 standard deviation. The minmin method (method 2) outperforms the min-surrounding-gap method for low ρ but underperforms for larger ρ.

have cars more closely to each other. This arises from the fact that their distributions are different, which we further exploit in the gap density method.

Notice in Fig. 28 of the gap density method that there is low accuracy for small street radii as expected from approximation error, since we are approximating the gap distribution of a small street RP with that of RP_∞. As the radius becomes larger, however, the approximation gets better and we see an increasing accuracy due to the closeness of the fit of $\Delta_{i'}$ to $G(h, t^*)$, where $\Delta_{i'}$ is the distribution of gaps of the cars of first color i'.

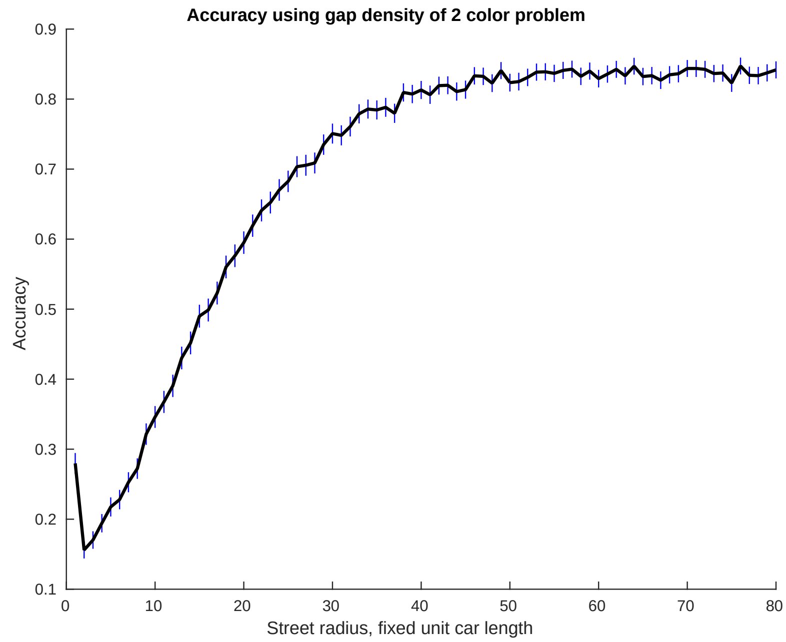

Figure 28: Accuracy of predictions to two color problem using gap density comparison. Low accuracy is observed for low ρ as these RPs do not resemble RP_∞, but higher accuracy is observed as ρ increases, as these RPs start to resemble the infinite line case.

5 Analytical explorations on $\mathbf{R}_\infty$

5.1 First two moments of $\Gamma(c_t)$

Because of the successful estimating ability of the minimum surrounding gap method in the two color problem and also of the geometrical methods in problems $1-4$, we explore to quantify the expected surrounding gap length that a car entering at time t will have on an infinite length street, which we express as $\mathbb{E}(\Gamma(c_t))$, where c_t is a car that enters the street at time t. Such study was done by Ziff [14] for random sequential adsorption of dimers on 1D lattice [15]. To find $\mathbb{E}(\Gamma(c_t))$, first we find the following integrals $E_n(t)$ that will help in our calculation of the expectation:

$$\begin{aligned} E_n(t) &= \int_l^\infty (h-l)^n \frac{G(h,t)}{\int_{\geq l} G(h',t)dh'} dh \\ &= \frac{n!}{(k_a t)^n}, \end{aligned} \tag{9}$$

See Appendix for detailed calculations. We call these integrals $E_n(t)$ as they take on the form of the moments of an exponential random variable (E for *exponential*) with parameter $k_a t$.

Now to find $\mathbb{E}(\Gamma(c_t))$, note that c_t will only park itself on one of the *available* spaces, whose conditional distribution $f(h,t)$ (conditioned on only available gaps) we can find using the available space function $\Phi(t) = \int_{\geq l}(\tilde{h}-l)G(\tilde{h},t)d\tilde{h}$. Since the probability of a gap being chosen is proportional to its length greater than the car size l, $(h-l)$ (a car cannot park on a gap of size l), we write:

$$f(h,t) = \frac{(h-l)G(h,t)}{\int_{\geq l}(\tilde{h}-l)G(\tilde{h},t)d\tilde{h}},$$

for gap $h \geq l$. Depending on which gap h that car c_t lands, the length $\Gamma(c_t)$ will be $(h-l)$. So we have

$$\begin{aligned} \mathbb{E}(\Gamma(c_t)) &= \int_l^\infty \Big(h-l\Big)\frac{(h-l)G(h,t)}{\int_{\geq l}(\tilde{h}-l)G(\tilde{h},t)d\tilde{h}}dh \\ &= \frac{2}{k_a t}. \end{aligned} \tag{10}$$

See Appendix for detailed calculations. As expected, we see that $\mathbb{E}(\Gamma(c_0)) = \infty$, as the first car will have no other cars and the whole street as its surrounding gap. A note of interest about $\mathbb{E}(\Gamma(c_t))$ is its independence of car length l.

Also note that the higher the rate k_a of cars' entering the street, which is defined as the length of time interval between two parking attempts (to a spot also chosen uniformly from the street), the faster $\mathbb{E}(\Gamma(c_t))$ decreases as time progresses. This is expected, as given a fixed period of time, more cars would likely to have parked compared to a RP with smaller k_a, resulting in smaller surrounding gaps due to the larger number of cars. We experimentally verify the inverse relationship of $\mathbb{E}(\Gamma(c_t))$ with t by running $1,000$ samples of $\text{RP}_{(1,60)}$ to approximate RP_∞ and average the surrounding gaps of cars entering at fixed times, $\{t_i\}_{i=1}^{50} = \{1,...,50\}$ (scaled by k_a). Then, without taking into account the first point (where the surrounding gap of the first car will always be $120\pi - 1$ whereas RP_∞ has infinity surrounding gap at $t \approx 0$), we fit this curve of averages to the equation $\mathbb{E} = 2/(k_a t)$ and see similar behavior between the average of the experimental results and $\mathbb{E}(\Gamma(c_t)) = \frac{2}{(2.6E-03)t}$. See Fig. 29.

The variance of $\Gamma(c_t)$, $\mathbb{V}(\Gamma(c_t))$, can also be found in a similar way:

$$\begin{aligned} \mathbb{V}(\Gamma(c_t)) &= \int_l^\infty \Big(h-l\Big)^2\frac{(h-l)G(h,t)}{\int_{\geq l}(\tilde{h}-l)G(\tilde{h},t)d\tilde{h}}dh - \frac{4}{(k_a t)^2} \\ &= \frac{\int_{\geq l}(\tilde{h}-l)^3 G(\tilde{h},t)d\tilde{h}}{\Phi(t)} - \frac{4}{(k_a t)^2} \\ &= \frac{6}{(k_a t)^2} - \frac{4}{(k_a t)^2} \\ &= \frac{2}{(k_a t)^2}. \end{aligned} \tag{11}$$

.

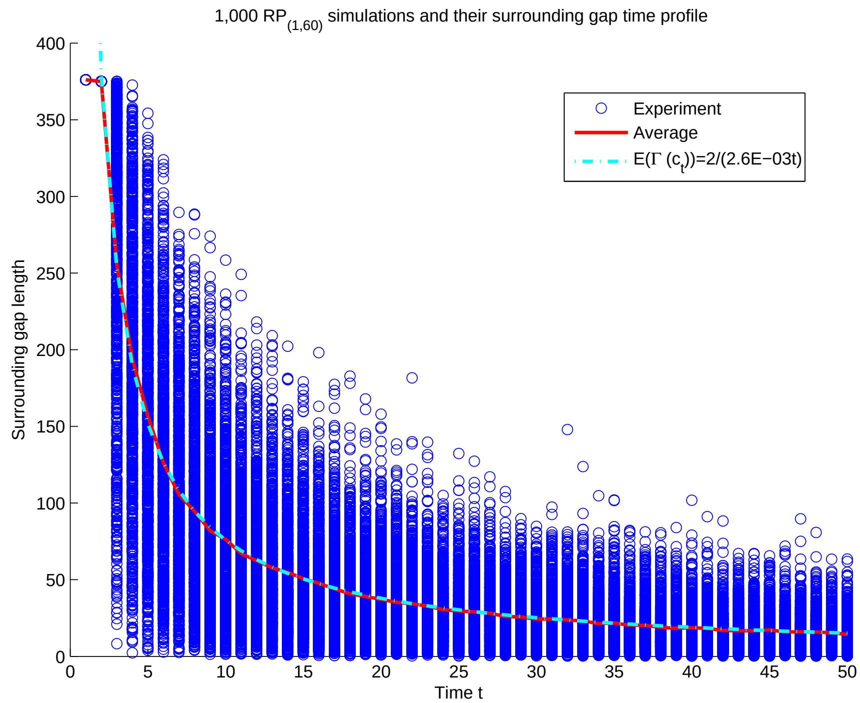

Figure 29: Experimental computation of $\mathbb{E}(\Gamma(c_t))$ for $\text{RP}_{(1,60)}$. $1,000$ samples were used to average the surrounding gaps of cars entering at the time instances $\{t_i\}$ that are scaled by k_a. For each time instance, the RP samples give different surrounding gap of the car entering at that time. Here, by the result of the fitting of the average, we see that $k_a \approx 2.6E-03$. As expected by theory, we see an inverse relationship of the average with time t, which we can model with $\mathbb{E}(\Gamma(c_t)) \approx \frac{2}{(2.6E-03)t}$.

6 Conclusion

Using MATLAB, we have implemented an optimized complete random parking simulation on a circle with linear computational time with respect to the size of the radius. After verification of the code, we have studied the average, statistical behavior of likelihood functions of random parking $\text{RP}_{(l,\rho)}$, with l, ρ fixed, viewing them as probability distributions on possible permutations of arrival history to be measured against the uniform distribution and the unit mass distribution on the true permutation. We see that geometry of each $\text{RP}_{(l,\rho)}$, such as the number of cars and the variance of the gaps, is associated with the relative entropies. For instance, the variance of gaps is negatively associated with the divergence of the normalized likelihood distribution (NLD) with the unit mass distribution on the true permutation, whereas the number of cars is somewhat positively associated with the relative entropy of NLD with the uniform distribution. We conjecture that this phenomenon is due to the general fact that the more disorderly the RP, the better we can infer information about σ^*. Instead of looking at likelihood functions as a whole, we can extract a part of the function, namely the weight of likelihoods of certain set of permutations, to make probabilistic decisions on the relative arrival orderings of pairs of cars.

We have come up with methods to estimate with high accuracy inference problems on RP.

In particular, we have looked at four problems involving relative order of arrivals of specific pairs of cars and answered them using the weights of likelihood, resulting in accuracies higher than 50%, which is what random guessing would achieve. We discovered that the accuracy increase as the street radius increases. Because of the computational time of computing likelihoods or implementing MC methods for large radius, we obtained a much faster and comparably accurate *geometrical* guessing where only the surrounding gaps of cars are taken into account. Achieving high accuracy shows that it is possible to in some sense extract information about the history of RP. Another inference question we looked at is the two color problem, for which we can obtain accuracies $\geq 70\%$ using either weights of likelihood or geometrical indicators. There are several geometrical methods to estimate the answer, all of which involve observing the gaps and increase in accuracy as the street radius increases, at the limit of which is RP_∞ where there are many established theoretical results.

A comment about the two different methods of answering the inference problems, namely the likelihood and geometrical method, is that while the former is a *probabilistic* prediction of the answer, the latter is a *deterministic* prediction making no use of probabilities. As such, the geometric method is much faster than the likelihood method because it does not require computing probabilities of each possible answer. However, the likelihood method gives information about the NLD associated to the RP in problem, as we can learn about its weights w_{ij}'s, or sample from the NLD in case of large N RPs. Through the geometrical method, no information about the NLD can be obtained.

Finally, using established theoretical results on RP_∞, we explored the random variable $\Gamma(c_t)$, which is the length of the surrounding gap that a car entering an infinite line at time t will have, and obtained its first two moments, both of which take a relatively simple form and decay to zero as $t \to \infty$.

References

[1] Feder J 1980 *Journal of Theor. Bio.* **87** 237

[2] Katira P, Agarwal A, and Hess H 2009 *Adv. Mat.* **21** 1599

[3] Weronski P and Z Adamczyk 1996 *J. Chem. Phys.* **105** 5562

[4] Rawal S and G J Rodgers 2005 *Phy. A: Stat. Mech. App.* **346** 621

[5] Renyi A 1963 *Sel. Transl. Math. Stat. Prob.* **4** 203

[6] Lee P 2004 *Bayesian Statistics, an introduction*

[7] Talbot J, Tarjus G, Tassel P R and Viot P 2000 *Colloids and Surfaces A: Phys. Eng. Asp.* **165** 287

[8] Kullback S and R A Leibler 1951 *Ann. Math. Stats.* **22** 79

[9] Brier G 1950 *Monthly Weather Review* **78** 1

[10] Lieshout M 2006 *Adv. App. Prob.* **38** 889

[11] Diaconis P and Shahshahani M 1981 *Z. Wahrs. verw. Geb.* **57** 159

[12] Fagin R, Kumar R and Sivakumar D 2003 *SIAM J. Disc. Math.* **17** 134

[13] Sokal A 1996 Lectures at the Cargese Summer School 15

[14] Ziff R 1994 *J. Phys. A: Math. Gen.* **27** L657

[15] Evans J W 1993 *Rev. Mod. Phys.* **65** 1281

7 Appendix

Details of Eq. 9: Let $F(u) = u^2 \exp\big(-2\int_0^u \frac{1-e^{-s}}{s} ds\big)$. Then,

$$\begin{aligned}
E_n(t) &= \int_l^\infty (h-l)^n \frac{G(h,t)}{\int_{\geq l} G(h',t)dh'} dh \\
&= \frac{F(k_a l t)}{l^2 \int_{\geq l} G(h',t)dh'} \int_0^\infty \tilde{h}^n e^{-k_a t \tilde{h}} d\tilde{h} \quad \text{[plugging in } G \text{ and change of variable]} \\
&= k_a t \int_0^\infty \tilde{h}^n e^{-k_a t \tilde{h}} d\tilde{h} \quad \text{[cancellation of } F \text{ with part of } \int_{\geq l} G(h',t)dh' = k_a t e^{-2\int_0^{k_a l t} \frac{(1-e^{-s})}{s} ds}] \\
&= k_a t \Big[\frac{n!}{(k_a t)^{n+1}}\Big] \\
&= \frac{n!}{(k_a t)^n}.
\end{aligned}$$

Details of Eq. 10:

$$\begin{aligned}
\mathbb{E}(\Gamma(c_t)) &= \int_l^\infty \Big(h-l\Big) \frac{(h-l)G(h,t)}{\int_{\geq l}(\tilde{h}-l)G(\tilde{h},t)d\tilde{h}} dh \\
&= \frac{\int_{\geq l}(h-l)^2 G(\tilde{h},t)d\tilde{h}}{\Phi(t)} \quad \text{[multiply } E_2(t) \text{ and } \int_{\geq l} G(s,t)ds \text{ to get numerator]} \\
&= \frac{\int_{\geq l} G(s,t)ds \frac{2}{(k_a t)^2}}{\Phi(t)} \\
&= \frac{k_a t e^{-2\int_0^{k_a l t} \frac{1-e^{-u}}{u} du} \frac{2}{(k_a t)^2}}{e^{-2\int_0^{k_a l t} \frac{1-e^{-u}}{u} du}} \\
&= \frac{2}{k_a t},
\end{aligned}$$

where the third equality comes from the rearrangement of Eq. 9.

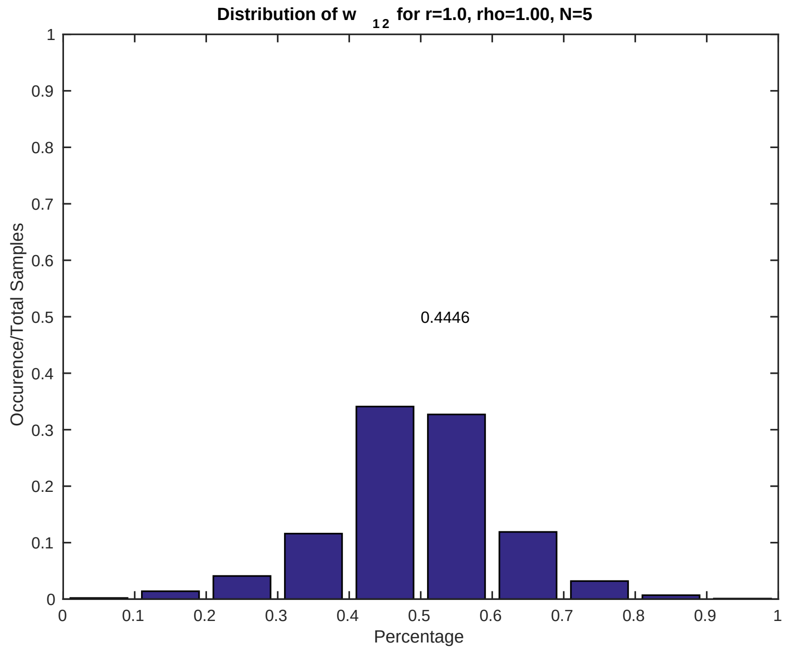

Figure 30: Distribution of w_{12} for $l = 1, \rho = 1, N = 5$. There is more spread than the case $N = 4$, and the mean is 0.4446.

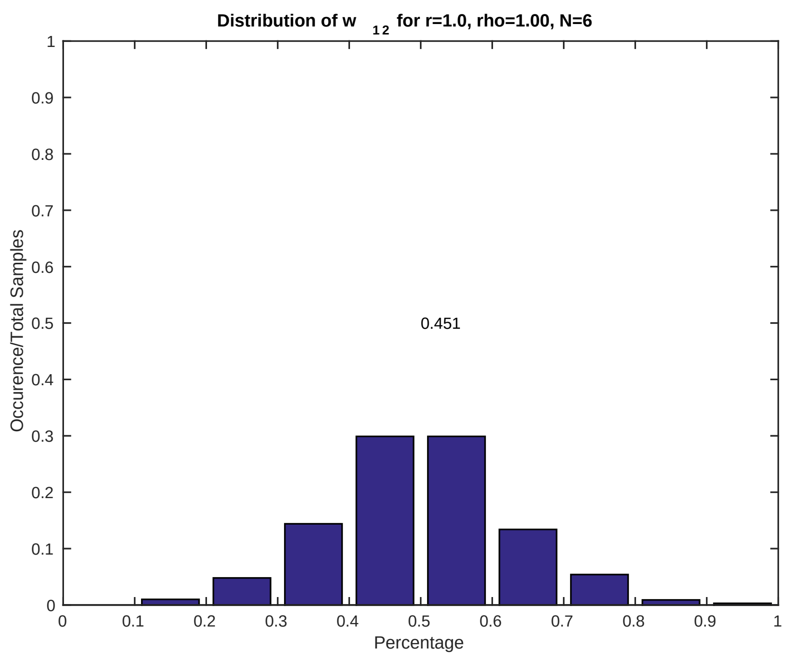

Figure 31: Distribution of w_{12} for $l = 1, \rho = 1, N = 6$. There is even more spread than the case $N = 5$, and the mean is 0.451.

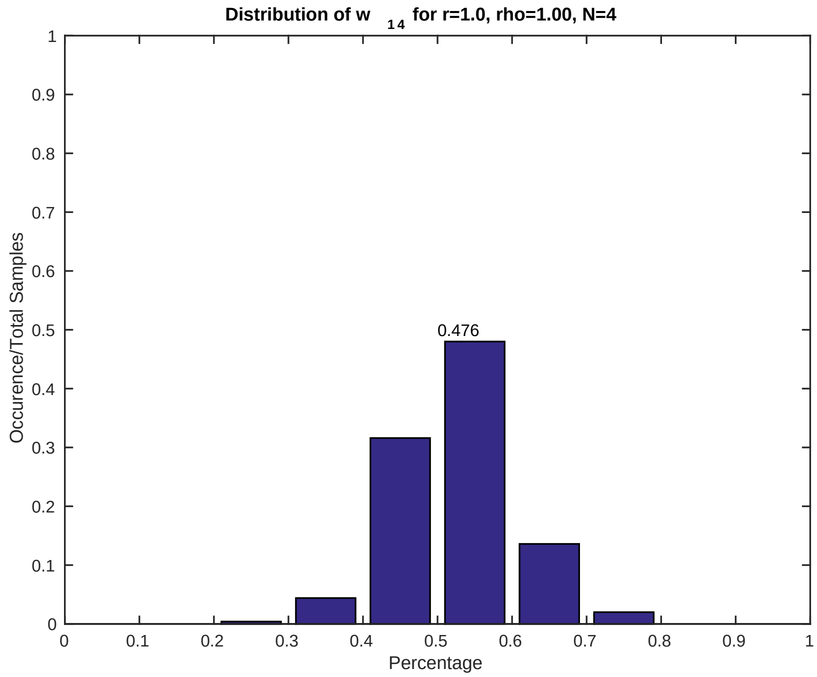

Figure 32: Distribution of w_{14} for $l = 1, \rho = 1, N = 4$. More mass is towards the right of 0.5, and the mean is 0.475.

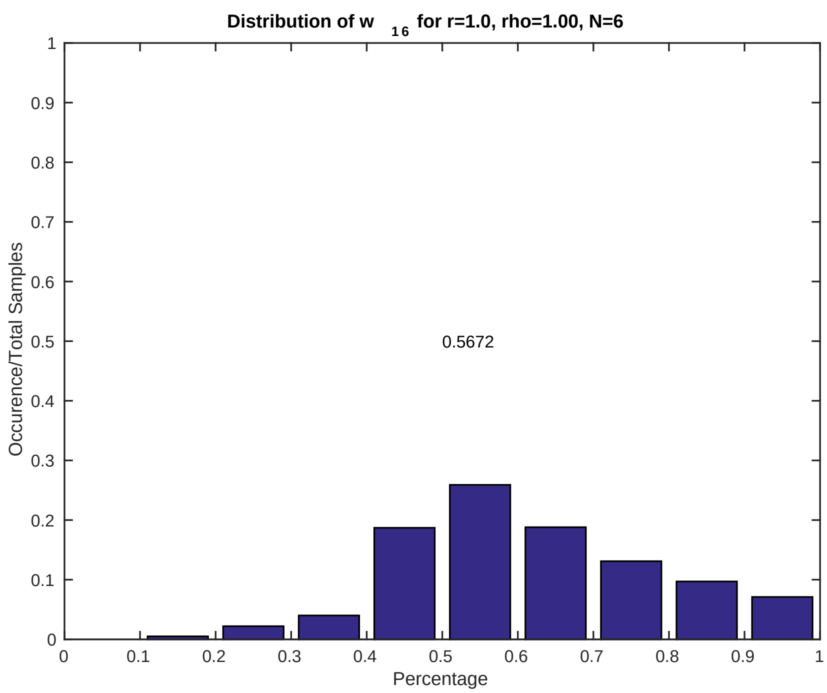

Figure 33: Distribution of w_{16} for $l = 1, \rho = 1, N = 6$. Significant more mass is toward the right of 0.5, and the mean is 0.5672.