Electron. Commun. Probab. **28** (2023), article no. 46, 1–12.
https://doi.org/10.1214/23-ECP552
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Maximum gaps in one-dimensional hard-core models

Dingding Dong* Nitya Mani†

Abstract

We study the distribution of the maximum gap size in one-dimensional hard-core models. First, we sequentially pack rods of length 2 into an interval of length L at random, subject to the hard-core constraint that rods do not overlap. We find that in a saturated packing, with high probability there is no gap of size $2 - o(L^{-1})$ between adjacent rods, but there are gaps of size at least $2 - L^{\varepsilon-1}$ for all $\varepsilon > 0$.

We subsequently study a dependent thinning-based variant of the hard-core process, the one-dimensional "ghost" hard-core model. In this model, we sequentially pack rods of length 2 into an interval of length L at random, such that placed rods neither overlap with previously placed rods *nor* previously considered candidate rods. We find that in the infinite time limit, with high probability the maximum gap between adjacent rods is smaller than $\log L$ but at least $(\log L)^{1-\varepsilon}$ for all $\varepsilon > 0$.

Keywords: random sequential adsoprtion; statistical mechanics.
MSC2020 subject classifications: Primary 60D05, Secondary 82C22.
Submitted to ECP on May 20, 2023, final version accepted on September 24, 2023.
Supersedes arXiv:2210.10088.

1 Introduction

The *Rényi parking problem* is a classical combinatorial question that gives a simple example of a *random sequential addition (RSA) process* or a one-dimensional *hard-core model*; such models are of much interest in statistical mechanics [2, 15, 18].

The setup for the parking problem proceeds as follows. Consider a closed interval $[0, L]$ for $L > 2$, into which rods of length 2 sequentially arrive at integer times. When each rod arrives, we attempt to place it uniformly at random in the interval, subject to the *hard-core* condition that rods cannot overlap with each other. The process terminates when no more rods can be added to the interval without violating the hard-core constraint, in which case we say that the rods form a *saturated configuration*. Let $N(L)$ be the random variable that denotes the number of rods in the saturated configuration. In 1958, Rényi proved the following well-known result.

Theorem 1.1 (Rényi [19]). *Let $N(L)$ be defined as above. Then,*

$$\lim_{L\to\infty} \frac{2\mathbb{E}[N(L)]}{L} = \alpha,$$

where α is the Rényi parking constant

$$\alpha := \int_0^\infty \exp\left[-2\int_0^x \frac{1-e^{-y}}{y} dy\right] dx \approx 0.7475979202.$$

*Harvard University, USA. E-mail: ddong@math.harvard.edu
†Massachusetts Institute of Technology, USA. E-mail: nmani@mit.edu

We study the distribution of *gaps* between adjacent rods in the saturated state, focusing on the upper extreme. In particular, we seek to understand the following.

Question 1.2. What can we say about the largest gap that arises in a saturated state?

Itoh [10] studied a delay integral equation that characterizes the distribution of the minimum gap sizes in a saturated configuration, following methods developed by Dvoretzky and Robbins [7]. The empirical distribution of gap sizes was also examined by Widom [26] and by Rintoul, Torquato and Tarjus [20] in the study of the nearest neighbors problem in one-dimensional RSA. In [10], Itoh observed that the expected minimum gap size in a saturated configuration on an interval of length L is smaller than any constant $\varepsilon > 0$ in the large L limit. This work was extended in [16] to approximate the upper tail of the distribution of minimum gap sizes. These works, as noted in §4 of [6], imply an analogous integral recurrence for the CDF of the maximum gap size in a saturated packing. In [6] the authors provided some preliminary observations and noted that further study of the maximum gap size was of substantial interest.

In this work, we give a threshold for the maximum gap size in a saturated configuration of the hard-core model.

Theorem 1.3. *For L sufficiently large, the following holds for the saturated configuration of the one-dimensional hard-core process:*

- *with high probability, there are no gaps of size $2 - o(1/L)$;*
- *for all $a > 0$, with positive probability, there exists a gap of size at least $2 - a/L$;*
- *for all $\varepsilon > 0$, with high probability, there exists a gap of size at least $2 - L^{\varepsilon-1}$.*

In the above one-dimensional hard-core model, rod centers are sampled uniformly from possible locations that do not violate the hard-core constraint. This *classical* process can be equivalently viewed as repeatedly randomly sampling a *candidate rod center* from $[1, L-1]$ and *rejecting* any rod whose placement would violate the hard-core constraint with respect to the previously placed rods. Note that rods that fail to park have no effect on future parking attempts.

This alternate formulation gives rise to a long-studied variant of the hard-core model, the *Matérn* or *ghost hard-core process*, studied in e.g. [13, 14, 24, 22] (related *dependent thinning* models have also been studied in a variety of other contexts [17, 24, 23]). Unlike the classical RSA process, where much is unknown even in two dimensions, Torquato and Stillinger [24] were able to analytically derive the n-point correlation functions and limiting densities of the ghost process, exactly solving the ghost hard-core model in arbitrary dimension. Specifically, they showed that the expected density of the d-dimensional ghost hard-core process in the infinite time limit was exactly $\frac{1}{2^d}$ for all $d \in \mathbb{Z}_+$.

We study the *one-dimensional ghost hard-core model*, akin to the hard-core model above, focusing on properties of this process in the infinite time limit. We give a precise definition of the ghost hard-core process below.

Definition 1.4 (One-dimensional ghost hard-core process)**.** *We attempt to place rods of length 2 on the interval $[0, L]$ as follows:*

- *Initialize:*
  - $A_0 = \emptyset$*, the collection of accepted rod centers;*
  - $Y_0 = \emptyset \subseteq [0, L]$*, the region occupied by accepted rods;*
  - $Y_0^g = \emptyset \subseteq [0, L]$*, the region occupied by all candidate rods.*
- *For $t = 1, 2, \ldots$:*
  - *If $[0, L] \setminus Y_{t-1}^g$ has no connected component of length ≥ 2, terminate and output the infinite time limit configuration $(A_\infty, Y_\infty, Y_\infty^g) := (A_{t-1}, Y_{t-1}, Y_{t-1}^g)$.*

- *Sample candidate rod center* $x \in [0, L]$ *uniformly at random.*
- *If* $x \in [0,1) \cup (L-1, L]$ *or* $(x-1, x+1) \cap Y^g_{t-1} \neq \emptyset$*, let* $A_t := A_{t-1}$*,* $Y_t := Y_{t-1}$*,* $Y^g_t := (Y^g_{t-1} \cup (x-1, x+1)) \cap [0, L]$*, and continue.*
- *Else, let* $A_t := A_{t-1} \cup \{x\}$*,* $Y_t := Y_{t-1} \cup (x-1, x+1)$*,* $Y^g_t := Y^g_{t-1} \cup (x-1, x+1)$*, and continue.*

Similarly to Theorem 1.3, we study the upper extreme of the distribution of gaps between consecutively placed rods per the configuration of rod centers given by A_∞ as defined above (i.e. in the infinite time limit of the ghost process).

Theorem 1.5. *For* L *sufficiently large, the following holds for the infinite time limit of the one-dimensional ghost hard-core process:*

- *with high probability, all gaps are smaller than* $\log L$*;*
- *for all* $\varepsilon > 0$*, with high probability, there exists a gap of size* $(\log L)^{1-\varepsilon}$*.*

Notation

Throughout this article, we consider packing rods of length 2 into an interval of length L, which we model by the closed interval $[0, L] \subseteq \mathbb{R}$. Unless stated otherwise, we study configurations in the *infinite time limit* of the two processes, the *1D classical hard-core model* and the *1D ghost hard-core model*.

We refer to the infinite time limit of the 1D classical hard-core model as *saturation*, since at this limit, no more rods can be packed without violating the hard-core constraint. Since both models considered in this article are one-dimensional and subject to the hard-core constraint, we sometimes omit the modifiers *one-dimensional* and *hard-core* when describing them.

We employ several conventions for asymptotic notation (as $L \to \infty$).

- $f = o(g)$, $g = \omega(f)$, $f \ll g$ and $g \gg f$ are all used to mean $\lim_{L\to\infty} f/g = 0$;
- $f = O(g)$ means there exists constant C such that $f \leq Cg$ for sufficiently large L;
- $f = \Omega(g)$ means there exists constant c such that $f \geq cg$ for sufficiently large L;
- $f = \Theta(g)$ means $f = O(g)$ and $f = \Omega(g)$;
- $f \sim g$ means that $\lim_{L\to\infty} f/g = 1$.

2 The maximum gap size in the hard-core model

In this section we prove Theorem 1.3. For the classical 1D hard-core model on $[0, L]$, let $G(L, r)$ denote the number of gaps of length at least r at saturation. We seek a threshold $r = r(L)$ such that as $L \to \infty$, we are likely to find gaps smaller than r and unlikely to find any gap of size much greater than r. Towards this goal, we prove Theorem 1.3.

We first consider fixing r, and show that as $L \to \infty$, $\mathbb{E}[G(L, r)]$ converges to a linear function $c_r(L+2)$, by studying a recurrence relation $\mathbb{E}[G(L, r)]$ satisfies. Since $G(L, r)$ is weakly decreasing with respect to r, so is the function $r \mapsto c_r$. By quantifying the rate of convergence, we obtain upper and lower bounds on $c_{2-\delta}$ by linear functions of δ. This implies that $\mathbb{E}[G[L, r]]$ changes from $o(1)$ to $\omega(1)$ as $(2-r)L$ does. We obtain the desired concentration around this expected value via the second moment method.

Since $G(L, r) = 0$ for $r \geq 2$, we henceforth suppose $r \in (0, 2)$. Observe that $\mathbb{E}[G(L, r)]$ satisfies the following recurrence relation:

Observation 2.1. For every $0 < r < 2$, the expectation $\mathbb{E}[G(L,r)]$ satisfies the following integral recurrence relation:

$$\mathbb{E}[G(L,r)] = \begin{cases} \frac{2}{L-2}\int_0^{L-2} \mathbb{E}[G(x,r)]\,dx & L > 2 \\ 1 & r \le L < 2 \\ 0 & L < r \end{cases} .$$

We will first show that $\mathbb{E}[G(L,r)] \sim c_r(L+2)$ for some constant $c_r > 0$. To do this, we consider some helpful auxiliary functions.

Definition 2.2. *For $r \in (0,2)$, define $f_r(L) := \mathbb{E}[G(L,r)]$ and $g_r(L) := \frac{\mathbb{E}[G(L,r)]}{L+2}$.*

Observation 2.3. For all $r \in (0,2)$ and $L > 2$, we have

$$g_r(L) = \frac{2}{L^2-4}\int_0^{L-2} g_r(x)(x+2)\,dx.$$

Proof. The relation follows from a direct calculation:

$$g_r(L) = \frac{1}{L+2}\cdot\frac{2}{L-2}\int_{x=0}^{L-2} f_r(L)\,dx = \frac{2}{L^2-4}\int_0^{L-2} g_r(x)(x+2)\,dx. \qquad \square$$

One immediate consequence of the above observation is the below result.

Observation 2.4. For all $r \in (0,2)$ and $L \ge 0$, we have $0 \le g_r(L) \le 1$.

Indeed, one can check manually that for $L \in [0,4]$, we have $g_r(L) \in [0,1]$. Recursively, we obtain the more general result by noting that for all $L \ge 4$, $\frac{2}{L^2-4}\int_0^{L-2}(x+2)dx = 1$. We now prove that $g_r(L)$ converges to a constant $c_r > 0$ as $L \to \infty$ by controlling the derivative $g_r'(L)$.

Lemma 2.5. *For every $m \in \mathbb{N}$, there exists some $N_m \in \mathbb{N}$ such that for all $r \in (0,2)$, whenever $L > 2m+1$, we have $|g_r'(L)| \le N_m/L^m$.*

Proof. We prove the lemma by applying Observation 2.3 multiple times. For all $L > 2$, applying Observation 2.3 twice, we have

$$g_r'(L) = \frac{-4L}{(L^2-4)^2}\cdot\frac{L^2-4}{2}\cdot g_r(L) + \frac{2L}{L^2-4}\cdot g_r(L-2) = \frac{2L}{L^2-4}(g_r(L-2) - g_r(L)).$$

By Observation 2.4, $0 \le g_r(L) \le 1$ for all $L \ge 0$. Therefore, for all $L \ge 3$, we have

$$|g_r'(L)| \le \frac{2L}{L^2-4}\cdot 2 \le \frac{N_1}{L}$$

for some $N_1 > 0$. Substituting this inequality again into the above expression, we see that since $g_r(L)$ is differentiable almost everywhere, for all $L \ge 5$, we have

$$\begin{aligned} |g_r'(L)| &= \frac{2L}{L^2-4}|g_r(L-2) - g_r(L)| = \frac{2L}{L^2-4}\left|\int_{L-2}^{L} g_r'(x)dx\right| \le \frac{2L}{L^2-4}\left|\int_{L-2}^{L}\frac{N_1}{L}dx\right| \\ &= \frac{2L}{L^2-4}\cdot\frac{2N_1}{L} \le \frac{N_2}{L^2} \end{aligned}$$

for some $N_2 > 0$ (e.g. we can take $N_2 \approx 4N_1$). Iterating gives the desired result. $\square$

Lemma 2.6. *For every $r \in (0,2)$, there exists $c_r \ge 0$ such that $\lim_{L\to\infty} g_r(L) = c_r$. Moreover, the convergence is uniform in r.*

Proof. Since g_r is differentiable almost everywhere, for all $L_2 > L_1 > 5$, we have

$$|g_r(L_2) - g_r(L_1)| = \left| \int_{L_1}^{L_2} g_r'(x)\,dx \right| \le N_2 \left(\frac{1}{L_1} - \frac{1}{L_2} \right).$$

Consequently, for all $\varepsilon > 0$, there exists $L_\varepsilon > 0$ such that for all $L_1, L_2 > L_\varepsilon$ we have $|g_r(L_2) - g_r(L_1)| < \varepsilon$. This implies that $\lim_{L\to\infty} g_r(L) = c_r$ for some $c_r \ge 0$. Since N_2 does not depend on r, the convergence is uniform in r. □

Lemma 2.5 actually implies a stronger result. Since

$$g_r(L) = c_r - \int_L^\infty g_r'(x)\,dx = c_r + o(L^{-m})$$

for arbitrary large $m \in \mathbb{N}$, we have the following:

Corollary 2.7. *For all $m \in \mathbb{N}$, we have $g_r(L) = c_r + o(L^{-m})$, where the convergence is uniform in r.*

Given the limiting coefficient c_r, we wish to understand its magnitude as a function of r. To this end, we define the following auxiliary functions.

Definition 2.8. *For every $r \in (0, 2]$, define function $h_r(L)$ with domain $(2, \infty)$ as follows:*

$$h_r(L) = \begin{cases} 0 & 2 < L < 2 + r \\ \frac{2}{L-2} & 2 + r \le L \le 4 \\ \frac{2}{L-2} \left(1 + \int_2^{L-2} h_r(x)\,dx \right) & L > 4. \end{cases}$$

For all $L > 4$, $f_r(L)$ is continuously differentiable with respect to r on $(0, 2)$, with $\frac{\partial f_r(L)}{\partial r} = -h_r(L)$. Moreover, the left derivative of $f_r(L)$ at $r = 2$ equals $-h_2(L)$. An analysis similar to Lemma 2.5 gives the following asymptotic expression for h_r.

Lemma 2.9. *For all $r \in (0, 2]$, there exists $\lambda_r > 0$ such that $\lim_{L\to\infty} h_r(L)/(L + 2) = \lambda_r$, where the convergence is uniform in r. Moreover, if $r_1 < r_2$, then $\lambda_{r_1} \ge \lambda_{r_2}$.*

Since $h_r(L) = -\frac{\partial f_r(L)}{\partial r}$ for all $L > 4$, one might imagine that $-\lambda_r$ is the derivative of c_r with respect to r. We make this notion precise below.

Lemma 2.10. *For all $r \in (0, 2)$, we have $\frac{\partial c_r}{\partial r} = -\lambda_r$. Moreover, the left derivative of c_r at 2 equals $-\lambda_2$.*

Proof. Consider $\{\frac{f_r(L)}{L+2} : L > 4\}$ and $\{-\frac{h_r(L)}{L+2} : L > 4\}$ as families of functions of r. Recall that for all $L > 4$, $f_r(L)$ is continuously differentiable with respect to r on $(0, 2)$, with the following properties:

- $\lim_{L\to\infty} \frac{f_r(L)}{L+2} = c_r$;
- $\frac{\partial}{\partial r}(\frac{f_r(L)}{L+2}) = -\frac{h_r(L)}{L+2}$;
- $-\frac{h_r(L)}{L+2}$ converges to $-\lambda_r$ uniformly in r.

Applying the differentiable limit theorem, we find that c_r is differentiable on $(0, 2)$, with $\frac{\partial c_r}{\partial r} = -\lambda_r$. Extending to $r = 2$ gives that the left derivative of c_r at 2 equals $-\lambda_2$. □

The above will be enough understand the threshold at which the expected number of length r gaps, $f_r(L) = \mathbb{E}[G(L, r)]$, transitions from $\Omega(1)$ to $o(1)$.

Corollary 2.11. *If a function $\gamma(L)$ has image in $(0, 2]$, then $\lambda_2(2 - \gamma(L)) \le c_{\gamma(L)}$. Moreover, if the image of γ lies in $[1, 2]$, then we also have $c_{\gamma(L)} \le \lambda_1(2 - \gamma(L))$. Therefore, $\lim_{L\to\infty} f_{\gamma(L)}(L) = c_{\gamma(L)}(L + 2) = o(1)$ if and only if $\gamma(L) = 2 - o(1/L)$.*

Proof. Since $c_2 = \lim_{L\to\infty} \frac{\mathbb{E}[G(L,2)]}{L+2} = 0$, Lemma 2.10 implies that for all $r \in (0,2)$, we have $c_r = \int_r^2 \lambda_s ds$. The desired result then follows by noting that λ_r is decreasing with respect to r. □

The last ingredient for proving Theorem 1.3 is an asymtotic second moment result on $\mathbb{E}[G(L,r)]$.

Definition 2.12. *For* $r \in (0,2)$ *and* $L > 0$*, let* $V_r(L) = \text{Var}[G(L,r)]$.

For $0 \le L \le 2$, we have $V_r(L) = 0$. For $L > 2$, we have the following recursive relation.

Observation 2.13. For all $r \in (0,2)$ and $m \in \mathbb{N}$, the following holds for $V_r(L)$ when $L > 2$:

$$V_r(L) = \frac{2}{L-2}\int_0^{L-2} V_r(x)\,dx + o(L^{-m}),$$

where the convergence is uniform in r.

Proof. For $L > 2$, let X denote the left endpoint of the first placed rod, so that X follows the uniform distribution on $[0, L-2]$. Let $G_1(L,r)$ denote the number of gaps of size at least r on the left of the first placed rod, and $G_2(L,r)$ denote the number of those on the right. Notice that $G(L,r) = G_1(L,r) + G_2(L,r)$. Furthermore, $G_1(L,r), G_2(L,r)$ are conditionally independent given X. By conditional variance, we therefore have

$$\begin{aligned}
V_r(L) &= \text{Var}[G(L,r)] = \mathbb{E}_X[\text{Var}[G(L,r)|X]] + \text{Var}[\mathbb{E}_X[G(L,r)|X]] \\
&= \mathbb{E}_X[\text{Var}[G_1(L,r)|X] + \text{Var}[G_2(L,r)|X]] + \text{Var}[\mathbb{E}_X[G(L,r)|X]] \\
&= \frac{1}{L-2}\int_0^{L-2}(V_r(x) + V_r(L-2-x))dx \\
&\quad + \frac{1}{L-2}\int_0^{L-2}(f_r(x) + f_r(L-2-x) - f_r(L))^2 dx \\
&= \frac{2}{L-2}\int_0^{L-2} V_r(x)dx + \frac{1}{L-2}\int_0^{L-2}(f_r(x) + f_r(L-2-x) - f_r(L))^2 dx.
\end{aligned}$$

We know from Corollary 2.7 that the second term is $o(L^{-m})$ for arbitrary $m \in \mathbb{N}$, which leads to the desired recurrence. □

Using the previous argument on $f_r(L)$, we obtain the following result analogous to Corollary 2.11:

Lemma 2.14. *There exists a constant* $\mu_1 > 0$ *such that for any function* $\gamma(L)$ *with image in* $[1,2]$*, we have*

$$V_{\gamma(L)}(L) \le \mu_1(2 - \gamma(L))(L+2) + o(1).$$

We are ready to prove Theorem 1.3.

Proof of Theorem 1.3. For every $\varepsilon > 0$, we wish to find $M_1, M_2 > 0$ such that for all L sufficiently large, we have

$$2 - \frac{M_1}{L} \le \max\{r : G[L,r] \ge 1\} \le 2 - \frac{M_2}{L}$$

with probability at least $1-\varepsilon$. First, choose $M_1 > 0$ such that $\frac{\lambda_2^2 M_1/4}{2\mu_1 + \lambda_2^2 M_1/4} > 1 - \frac{\varepsilon}{2}$. By Corollary 2.11, we have

$$\mathbb{E}\left[G\left(L, 2 - \frac{M_1}{L}\right)\right] = c_{2-\frac{M_1}{L}}(L+2) + o(1) \ge \lambda_2 \cdot \frac{M_1}{L}\cdot(L+2) + o(1) \ge \lambda_2 M_1/2$$

and by Lemma 2.14 we have

$$V_{2-\frac{M_1}{L}}(L) \leq \mu_1 \cdot \frac{M_1}{L} \cdot (L+2) + o(1) \leq 2\mu_1 M_1$$

for L sufficiently large. From the second moment method, we see that

$$\mathbb{P}\left[G\left(L, 2-\frac{M_1}{L}\right) > 0\right] \geq \frac{\mathbb{E}[G(L, 2-\frac{M_1}{L})]^2}{\mathbb{E}[G(L, 2-\frac{M_1}{L})^2]} \geq \frac{(\lambda_2 M_1/2)^2}{2\mu_1 M_1 + (\lambda_2 M_1/2)^2} \geq 1 - \frac{\varepsilon}{2}.$$

Next, choose $M_2 > 0$ such that $2\lambda_1 M_2 < 1$ and $\frac{2\mu_1 M_2}{(1-2\lambda_1 M_2)^2} < \frac{\varepsilon}{2}$. For L sufficiently large, we have

$$\mathbb{E}[G(L, 2-M_2/L)] = c_{2-M_2/L}(L+2) + o(1) \leq \lambda_1 \cdot M_2/L \cdot (L+2) + o(1) \leq 2\lambda_1 M_2$$

and

$$\mathrm{Var}[G(L, 2-M_2/L)] \leq 2\mu_1 M_2.$$

Therefore, by the second moment method, we have

$$\mathbb{P}[G(L, 2-M_2/L) \geq 1] \leq \frac{2\mu_1 M_2}{(1-2\lambda_1 M_2)^2} < \frac{\varepsilon}{2}.$$

Combining the above, we know that the maximum gap has size in $[2-M_2/L, 2-M_1/L]$ with probability at least $1-\varepsilon$. □

3 Maximum gaps in the one-dimensional ghost hard-core model

In this section we prove Theorem 1.5, giving a threshold for the maximum gap size in the infinite time limit of the ghost hard-core process. During the process, we use *ghosts* to denote those previously attempted rods that were eventually rejected, as they still place a constraint on locations of future rods (see Definition 1.4).

Imagine that we pause at time $t \in \mathbb{N}$, and for some choice of $\ell \ll L$, consider the collection of gaps of length $\Theta(\ell)$ between accepted rods. For each such gap of size $\Theta(\ell)$, we will compute the probability that this gap is eventually retained, using an inductive argument on the cumulative lengths of overlapping ghosts at either end. We make this idea more precise below.

Definition 3.1. *Suppose we run the ghost hard-core process in Definition 1.4. Fix $0 < \ell < L$ and $k_1, k_2 \geq 0$ such that $k_1 + k_2 \leq \ell$. Consider any time point $t \in \mathbb{N}$ and $x \in [2, L-2]$ such that $[x, x+\ell] \subseteq [0, L]$. Define $\mathbf{P}^{(\ell)}(k_1, k_2)$ to be the conditional probability*

$$\mathbb{P}\Big[[x, x+\ell] \subseteq [0, L] \setminus Y_\infty \textit{ given}$$
$$[x, x+\ell] \textit{ is a connected component of } [0, L] \setminus Y_t \textit{ and}$$
$$[x+k_1, x+\ell-k_2] \textit{ is a connected component of } [0, L] \setminus Y_t^g\Big]$$

(this quantity is independent of x and t). In other words, $\mathbf{P}^{(\ell)}(k_1, k_2)$ is the probability that a length-ℓ gap between accepted rods is eventually retained, given that its leftmost k_1-subinterval and rightmost k_2-subinterval are fully covered by overlapping ghosts.

Suppose $[x, x+\ell]$ is a gap between accepted rods (i.e., a connected component of $[0, L] \setminus Y_t$), and $[x+k_1, x+\ell-k_2]$ is a connected component of $[0, L] \setminus Y_t^g$. If $\ell - (k_1+k_2) \leq 2$, then the gap $[x, x+\ell]$ is retained with probability 1, so $\mathbf{P}^{(\ell)}(k_1, k_2) = 1$. If $\ell - (k_1+k_2) > 2$, we can deduce a recurrence relation for $\mathbf{P}^{(\ell)}(k_1, k_2)$ based on the location of the next arrived rod. We therefore have the following observation.

Observation 3.2. Suppose $0 < \ell < L$, $k_1, k_2 \geq 0$ and $k_1 + k_2 \leq \ell$.

1. If $\ell - (k_1 + k_2) \leq 2$, then $\mathbf{P}^{(\ell)}(k_1, k_2) = 1$.
2. If $\ell - (k_1 + k_2) > 2$, then

$$\begin{aligned}\mathbf{P}^{(\ell)}(k_1, k_2) &= \left(1 - \frac{\ell - k_1 - k_2 + 2}{L}\right)\mathbf{P}^{(\ell)}(k_1, k_2) + \frac{1}{L}\int_0^2 \mathbf{P}^{(\ell)}(k_1 + k, k_2)dk \\ &\quad + \frac{1}{L}\int_0^2 \mathbf{P}^{(\ell)}(k_1, k_2 + k)dk \\ &= \frac{1}{\ell - k_1 - k_2 + 2}\int_0^2 \mathbf{P}^{(\ell)}(k_1 + k, k_2)dk \\ &\quad + \frac{1}{\ell - k_1 - k_2 + 2}\int_0^2 \mathbf{P}^{(\ell)}(k_1, k_2 + k)dk.\end{aligned}$$

Since $\mathbf{P}^{(\ell)}(k_1, k_2)$ is symmetric in k_1, k_2 and only depends on the sum $k_1 + k_2$, we take $s := \ell - (k_1 + k_2)$ and define

$$\mathbf{P}^{(\ell)}(s) := \begin{cases} 1 & s \leq 2, \\ \frac{2}{s+2}\int_{s-2}^{s} \mathbf{P}^{(\ell)}(x)\,dx & s > 2. \end{cases}$$

We now prove the second half of Theorem 1.5, namely that for all $\varepsilon > 0$, with high probability a gap of size at least $(\log n)^{1-\varepsilon}$ is retained in the ghost hard-core process. To do so, we first give a lower bound on $\mathbf{P}^{(\ell)}(s)$.

Lemma 3.3. *For all $s \geq 0$, $\mathbf{P}^{(\ell)}(s) \geq s^{-s}$.*

Proof. For $0 \leq s \leq 2$ we have $\mathbf{P}^{(\ell)}(s) = 1$. When $s \geq 2$, we utilize the identity

$$\mathbf{P}^{(\ell)}(s) = \frac{2}{s+2}\int_{s-2}^{s} \mathbf{P}^{(\ell)}(x)\,dx \geq \frac{2}{\lfloor s \rfloor + 3}\int_{s-2}^{s-1} \mathbf{P}^{(\ell)}(x)\,dx$$

to show that $\mathbf{P}^{(\ell)}(s) \geq \frac{2^{\lfloor s \rfloor - 1}}{\lfloor s+3 \rfloor!/4!} \geq s^{-s}$ for all $s \geq 2$.

For $2 \leq s \leq 3$, we know from the above identity that $\mathbf{P}^{(\ell)}(s) \geq \frac{2}{5}$. Similarly, for $3 \leq s \leq 4$, we have $\mathbf{P}^{(\ell)}(s) \geq \frac{2}{6} \cdot \frac{2}{5}$. Inductively, assuming $\mathbf{P}^{(\ell)}(x) \geq \frac{2^{\lfloor x \rfloor - 1}}{\lfloor x \rfloor!/4!}$ for all $s - 2 \leq x \leq s - 1$, we have

$$\mathbf{P}^{(\ell)}(s) \geq \frac{2}{\lfloor s \rfloor + 3}\int_{s-2}^{s-1} \mathbf{P}^{(\ell)}(x)\,dx \geq \frac{2}{\lfloor s \rfloor + 3} \cdot \frac{2^{\lfloor s \rfloor - 2}}{\lfloor s+2 \rfloor!/4!} = \frac{2^{\lfloor s \rfloor - 1}}{\lfloor s+3 \rfloor!/4!}. \qquad \square$$

With $L \to \infty$, we choose parameter $\ell = \ell(L)$ with $\ell = \omega(1)$ and $\ell^{\ell+1} = o(L)$, so that $(1 - 1/\ell^\ell)^{L/\ell} = o(1)$. We will show that with positive probability, there are $\Omega(L/\ell)$ disjoint gaps of size at least ℓ at some time point in the ghost hard-core process; moreover, if ℓ is sufficiently small compared to L, then at least one of them is retained at the infinite time limit. To show the first claim, we apply the following theorem on a randomly partitioned interval:

Theorem 3.4 (Theorem 2.2 [9]). *Suppose an interval of length 1 is broken uniformly at random into n subintervals with lengths $S_1 \leq \cdots \leq S_n$. Then for every $i \in [n]$ and $r \in \mathbb{N}$, we have $\mathbb{E}[S_i^r] = \mathbb{E}[Y_i^r] \cdot \frac{\Gamma(n)}{\Gamma(n+r)}$, where*

$$Y_i = \frac{X_n}{n} + \frac{X_{n-1}}{n-1} + \cdots + \frac{X_{n-i+1}}{n-i+1},$$

and $X_1, \ldots, X_n$ are independent exponential random variables with mean 1.

To simplify the calculations below, we assume without loss of generality that all the bounds on the number of placed rods and all of the given times are integers.

Definition 3.5. *Consider the ghost hard-core process on* $[0, L]$*. For* $c > 0$ *and* $0 \le \ell \le L$*, let* $A_{c,\ell,L}$ *denote the event that at time* $L/\ell - 1$*, there are at least* cL/ℓ *pairs of adjacent accepted rods whose centers have distance at least* ℓ*.*

Lemma 3.6. *Fix arbitrary* $c \in (0, 1/e)$ *and* $\theta \in (0, 1)$*. Suppose* $\ell = \ell(L)$ *is a function of* L *such that* $\ell^{\ell+1} = o(L)$*. Then for* L *sufficiently large, we have*

$$\mathbb{P}[A_{c,\ell,L}] \ge \left(1 - \frac{1}{(1+\theta)\ln((2c)^{-1})}\right)^2 (1-\theta) + \left(1 - \frac{2}{(1+\theta)\ln((1-c)^{-1})\ell}\right)^2 (1-\theta) - 1.$$

Proof. We rescale and consider placing rods of length $2/L$ on an interval of length 1. Let $M = L/\ell$. For $0 < b < 1$, let S_{bM} denote the bM-th smallest distance between adjacent *candidate* rod centers. Without loss of generality, suppose bM is always an integer.

By Theorem 3.4, we have

$$\mathbb{E}[S_{bM}] = \frac{1}{M} \sum_{M-bM+1}^{M} \frac{1}{i} \sim \frac{\ln((1-b)^{-1})}{M},$$

$$\mathbb{E}[S_{bM}^2] = \frac{1}{M(M+1)} \left(\left(\sum_{M-bM+1}^{M} \frac{1}{i} \right)^2 + \sum_{M-bM+1}^{M} \frac{1}{i^2} \right) \sim \frac{\ln^2((1-b)^{-1})}{M^2}.$$

For M sufficiently large, by Paley-Zygmund, we have

$$\begin{aligned} \mathbb{P}(S_{M-2cM} \ge 1/M) &\ge \mathbb{P}\left(S_{M-cM} \ge \frac{\mathbb{E}[S_{M-cM}]}{(1+\theta)\ln((2c)^{-1})} \right) \\ &\ge \left(1 - \frac{1}{(1+\theta)\ln((2c)^{-1})}\right)^2 (1-\theta). \end{aligned}$$

as well as

$$\begin{aligned} \mathbb{P}[S_{cM} \ge 2/L] &\ge \mathbb{P}\left[S_{cM} \ge \frac{\mathbb{E}[S_{cM}]}{(1+\theta)\ln((1-c)^{-1})} \cdot \frac{2M}{L} \right] \\ &\ge \left(1 - \frac{2}{(1+\theta)\ln((1-c)^{-1})\ell}\right)^2 (1-\theta). \end{aligned}$$

Thus, with at least $\left(1 - \frac{1}{(1+\theta)\ln((2c)^{-1})}\right)^2 (1-\theta) + \left(1 - \frac{2}{(1+\theta)\ln((1-c)^{-1})\ell}\right)^2 (1-\theta) - 1$ probability, there are at most cM pairs of adjacent candidate rods whose centers have distance less than $2/L$, and at least $2cM$ pairs of adjacent candidate rods whose centers have distance at least $1/M$. This implies that among the latter, at least cM pairs are not affected by ghosts, so $A_{c,\ell,L}$ holds. □

We are now ready to conclude the second half of Theorem 1.5.

Lemma 3.7. *If* $\ell = (\log L)^{1-\varepsilon}$ *for* $\varepsilon \in (0, 1)$*, then for sufficiently large* L*, with high probability there exists a gap of size at least* ℓ *in the infinite time limit.*

Proof. For any $\delta \in (0, 1)$, we can choose $c \in (0, 1/e)$ and $\theta \in (0, 1)$ via Lemma 3.6 such that $\mathbb{P}[A_{c,2\ell,L}] \ge 1 - \delta/2$ for L sufficiently large. Consequently, at time $t = L/(2\ell) - 1$, there are at $\frac{cL}{2\ell}$ gaps of size at least $2\ell - 2 \ge \ell$, and thus at least $\frac{cL}{4\ell}$ gaps of size within $[\ell, 4\ell/c]$. By Lemma 3.3 and the inequality that $1 + x \le e^x$, the probability that none of these gaps is retained is at most

$$\left(1 - (4\ell/c)^{-4\ell/c}\right)^{\frac{cL}{4\ell}} \le \exp\left[-\frac{cL}{4\ell(4\ell/c)^{4\ell/c}} \right] = o(1),$$

as

$$\log\left(\frac{cL}{4\ell(4\ell/c)^{4\ell/c}}\right) = \log L - (4\ell/c+1)\log(4\ell/c) = \ell^{\frac{1}{1-\varepsilon}} - (4\ell/c+1)(\log\ell - \log c + 2)$$
$$= \omega(1).$$

Hence with probability at least $1-\delta$, there exists a gap of size at least ℓ in the infinite time limit. $\square$

To show the first half of Theorem 1.5, we give an upper bound on $\mathbf{P}^{(\ell)}(s)$.

Lemma 3.8. *There exists $C > 0$ such that for all $s \geq 0$, $\mathbf{P}^{(\ell)}(s) \leq Cs^{-s/3}$.*

Proof. Take absolute constant $M > 0$ such that for all $s > M$, we have

1. $2(s+s^{-s/3})^{-(s+s^{-s/3})/3} \geq s^{-s/3}$,
2. $$\frac{s^{-s/3}}{\int_{s-2}^{s} x^{-x/3}dx} \geq \frac{s^{-s/3}}{2(s-2)^{-s/3+2/3}} = \frac{1}{2(s-2)^{2/3}} \cdot \left(1+\frac{2}{s-2}\right)^{-s/3} \geq \frac{8}{s+2}.$$

Since the function $s \mapsto s^{-s/3}$ is strictly positive on $[0, M]$, we can take absolute constant $C \geq 10$ such that $\mathbf{P}^{(\ell)}(s) \leq Cs^{-s/3}$ for all $0 \leq s \leq M$.

For $s > M$, let $P(s)$ be the property that $\mathbf{P}^{(\ell)}(y) \leq Cy^{-y/3}$ for all $y < s$. Observe that if $P(s)$ holds, then we have

$$C(s+s^{-s/3})^{-(s+s^{-s/3})/3} \geq Cs^{-s/3}/2 \geq Cs^{-s/3}/4 + \frac{2}{s+2}\int_{s-2}^{s} Cx^{-x/3}dx$$
$$\geq s^{-s/3} + \frac{2}{s+2}\int_{s-2}^{s}\mathbf{P}^{(\ell)}(x)\,dx \geq \frac{2}{s+2}\int_{s-2}^{s+s^{-s/3}}\mathbf{P}^{(\ell)}(x)\,dx.$$

Thus, for all $s \leq y \leq s+s^{-s/3}$, we have

$$\mathbf{P}^{(\ell)}(y) \leq \frac{2}{s+2}\int_{s-2}^{s+s^{-s/3}}\mathbf{P}^{(\ell)}(x)\,dx \leq C(s+s^{-s/3})^{-(s+s^{-s/3})/3} \leq Cy^{-y/3}.$$

In other words, $P(s)$ implies $P(s+s^{-s/3})$.

Let $Y = \{s \geq M : \mathbf{P}^{(\ell)}(s) > Cs^{-s/3}\}$. By contradiction, if Y is nonempty, then $y_0 := \inf Y$ exists. Since $P(y_0)$ holds, by the above argument, $P(y_0 + y_0^{-y_0/3})$ holds as well, so $y_0 \neq \inf Y$, a contradiction. $\square$

Lemma 3.9. *For sufficiently large L, with high probability, the largest gap that remains in the infinite time limit of the ghost hard-core model has size less than $\log L$.*

Proof. Let $\ell = \log L$. At any time, there are at most L/ℓ distinct gaps of length at least ℓ between accepted rods. Applying Lemma 3.8 and a union bound, the probability that at least one of them is retained is at most

$$\frac{L}{\ell}\cdot C\ell^{-\ell/3} = \frac{CL}{\log L^{\log L/3+1}} = o(1). \qquad \square$$

Theorem 1.5 then follows by combining Lemmas 3.7 and 3.9.

4 Further directions

The higher dimensional analogues of the random sequential adsorption process are of particular importance. A primary motivating question is trying to understand the maximum density of a *sphere packing*, a maximum collection of congruent radius one spheres in $\mathbb{R}^d$ that do not overlap. Determining the densest packings in arbitrary dimensions is one of the most longstanding open problems in discrete geometry, recently resolved in $\mathbb{R}^8$ and $\mathbb{R}^{24}$ in the breakthrough works of [25, 4]. The only other dimensions in which optimal sphere packings are known are dimensions $1, 2$, and 3. One can derive lower bounds on an optimal sphere packing by studying packing procedures, such as the random sequential addition process of hard spheres in $\mathbb{R}^d$. Further, RSA in more than one dimension is in and of itself a process of much physical interest, as in [18, 21, 15, 6, 11, 8, 2, 3]. While our methods for establishing the extreme values of gaps do not generalize to more than one dimension, Theorem 1.3 offers a tantalizing glimpse into the existence of relatively large gaps in saturated hard-core packings and perhaps studying sphere packing densities.

Theorem 1.5 suggests that in the ghost hard-core process, the maximum gap size in the infinite time limit converges to a logarithmic scale, which is much larger than in the classical hard-core process. One natural question to ask, therefore, is whether the same difference occurs in higher dimensions.

Question 4.1. Given a random configuration in the infinite time limit of the d-dimensional ghost RSA process, packed with radius 1 spheres, what is the expected density of radius 1 spheres that will be added if we run the traditional RSA process beginning from there?

Two-dimensional sphere packing problems have been extensively studied; physical theories such as the Asakura-Oosawa depletion interaction seek to explain the "effective interaction" between hard-sphere particles in higher dimensions [1]. Most of the work in more than one dimension, however, remains at the heuristic level, lacking rigorous mathematical results about distributions of gaps sizes (or even asymptotic packing densities and pair correlations) [12]. For example, the following question is open.

Question 4.2. Given a random packing of radius 1 spheres into an $L \times L$ square via the ghost hard-core process, what is the expected maximum radius of a sphere that could be added to this packing, without violating the hard-core constraint?

References

[1] Sho Asakura and Fumio Oosawa. On interaction between two bodies immersed in a solution of macromolecules. *The Journal of Chemical Physics*, 22(7):1255–1256, 1954.

[2] B. Edwin Blaisdell and Herbert Solomon. On random sequential packing in the plane and a conjecture of Palasti. *J. Appl. Probab.*, 7(3):667–698, 1970. MR0282389

[3] B. Edwin Blaisdell and Herbert Solomon. Random sequential packing in euclidean spaces of dimensions three and four and a conjecture of Palasti. *J. Appl. Probab.*, 19(2):382–390, 1982. MR0649975

[4] Henry Cohn, Abhinav Kumar, Stephen D. Miller, Danylo Radchenko, and Maryna Viazovska. The sphere packing problem in dimension 24. *Ann. of Math. (2)*, 185(3):1017–1033, 2017. doi:10.4007/annals.2017.185.3.8. MR3664817

[5] François Dunlop and Thierry Huillet. Hard rods: statistics of parking configurations. *Phys. A*, 324(3-4):698–706, 2003. doi:10.1016/S0378-4371(03)00065-7. MR1982909

[6] Mathieu Dutour Sikirić and Yoshiaki Itoh. *Random Sequential Packing of Cubes*. World Scientific, 2011. MR2779667

[7] A. Dvoretzky and H. Robbins. On the "parking" problem. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 9:209–225, 1964. MR0173275

[8] James W. Evans. Random and cooperative sequential adsorption. *Reviews of Modern Physics*, 65(4):1281, 1993.

[9] Lars Holst. On the lengths of the pieces of a stick broken at random. *J. Appl. Probab.*, 17(3):623–634, 1980. URL: https://doi-org.ezp-prod1.hul.harvard.edu/10.1017/s0021900200033738, doi:10.1017/s0021900200033738.

[10] Yoshiaki Itoh. On the minimum of gaps generated by one-dimensional random packing. *J. Appl. Probab.*, 17(1):134–144, 1980. doi:10.1017/s0021900200046878.

[11] Yoshiaki Itoh and Sumie Ueda. On packing density by a discrete random sequential packing of cubes in a space of n dimension. In *Proc. Inst. Statist. Math*, volume 31, pages 65–69, 1983.

[12] Werner Krauth. *Statistical Mechanics*, volume 13 of *Oxford Master Series in Physics*. Oxford University Press, Oxford, 2006. Algorithms and computations, Oxford Master Series in Statistical Computational, and Theoretical Physics. MR2370557

[13] Bertel Matern. Meddelanden fran statens. *Skogsforskningsinstitut 49*, 5(2):1–144, 1960. MR0169346

[14] Bertil Matérn. *Spatial Variation*, volume 36. Springer Science & Business Media, 2013. MR0867886

[15] A.J. Matheson. Computation of a random packing of hard spheres. *Journal of Physics C: Solid State Physics*, 7(15):2569, 1974.

[16] J. A. Morrison. The minimum of gaps generated by random packing of unit intervals into a large interval. *SIAM J. Appl. Math.*, 47(2):398–410, 1987. doi:10.1137/0147025. MR0881357

[17] G Porod. Die röntgenkleinwinkelstreuung von dichtgepackten kolloiden systemen. II. Teil. *Kolloid-Zeitschrift*, 125(2):108–122, 1952.

[18] V. Privman, J.S. Wang, and P. Nielaba. Continuum limit in random sequential adsorption. *Physical Review B*, 43(4):3366, 1991.

[19] A Rényi. On a one-dimensional problem concerning random space-filling problem. *Publ. Math. Inst. Hungar. Acad. Sci*, 3:109–127, 1958. MR0104284

[20] M.D. Rintoul, S. Torquato, and G. Tarjus. Nearest-neighbor statistics in a one-dimensional random sequential adsorption process. *Physical Review E*, 53(1):450, 1996.

[21] T. Schreiber, Mathew D. Penrose, and J. E. Yukich. Gaussian limits for multidimensional random sequential packing at saturation. *Comm. Math. Phys.*, 272(1):167–183, 2007. doi:10.1007/s00220-007-0218-2.

[22] Dietrich Stoyan, Wilfrid S. Kendall, Sung Nok Chiu, and Joseph Mecke. *Stochastic Geometry and Its Applications*. John Wiley & Sons, 2013. MR3236788

[23] Dietrich Stoyan and Martin Schlather. Random sequential adsorption: relationship to dead leaves and characterization of variability. *Journal of Statistical Physics*, 100:969–979, 2000.

[24] S. Torquato and F. H. Stillinger. Exactly solvable disordered sphere-packing model in arbitrary-dimensional Euclidean spaces. *Phys. Rev. E (3)*, 73(3):031106, 8, 2006. doi:10.1103/PhysRevE.73.031106. MR2231345

[25] Maryna S. Viazovska. The sphere packing problem in dimension 8. *Ann. of Math. (2)*, 185(3):991–1015, 2017. doi:10.4007/annals.2017.185.3.7. MR3664816

[26] B. Widom. Random sequential addition of hard spheres to a volume. *J. Chem. Phys.*, 44(10):3888–3894, 1966. doi:10.1063/1.1726548.

Acknowledgments. We would like to thank Henry Cohn and Salvatore Torquato for helpful discussions, ideas for writing improvements, and several useful reference suggestions for background. We would also like to thank the anonymous referee for helpful comments. NM was supported by the Hertz Graduate Fellowship and by the NSF GRFP #2141064.