Barrier coverage

Paul Balister*, Béla Bollobás†‡§, Amites Sarkar¶

August 31, 2011

Abstract

Suppose sensors are deployed randomly in a long thin strip, and suppose each sensor can detect objects within a fixed distance. We say that the sensors achieve barrier coverage if there is no path across the strip that a small object can follow that avoids detection by the sensors. We give fairly precise results on the probability that barrier coverage is achieved as a function of the range of the sensors, the height and length of the strip, and the number of sensors deployed. In particular, we show that the most likely obstruction — a rectangular region crossing the strip which is devoid of sensors — does not in general dominate the probability of failure of barrier coverage.

1 Introduction

Imagine a collection of randomly placed sensors deployed in some region in the plane, each capable of detecting events or objects within a given fixed distance. A natural and well studied question is whether or not they cover the region of interest, in the sense that every point in the region is in the sensing range of some sensor (see for example [11, 6, 7, 9, 3]). In this paper we consider a somewhat different question. Instead of fully covering the

*Department of Mathematical Sciences, University of Memphis, USA. E-mail: pbalistr@memphis.edu
†Research supported in part by NSF grant CCF-0728928.
‡DPMMS, Cambridge University, UK, and Department of Mathematical Sciences, University of Memphis, USA. E-mail: B.Bollobas@dpmms.cam.ac.uk
§Research supported in part by NSF grants DMS-0906634, CNS-0721983 and CCF-0728928, ARO grant W911NF-06-1-0076, and TAMOP-4.2.2/08/1/2008-0008 program of the Hungarian National Development Agency.
¶Department of Mathematics, Western Washington University, Bellingham, USA. E-mail: amites.sarkar@wwu.edu
region, one simply wishes to block the movement of an intruder say, across a long thin strip. In this case, fully covering the strip is not necessary. It is simply enough that there be no path across the strip avoiding the sensing regions. We call such coverage barrier coverage and give fairly precise results on the probability of barrier coverage as a function of the range of the sensors, the height and length of the strip, and the number of sensors deployed.

Consider therefore a horizontal strip \([a, b] \times [0, h]\) of height \(h\) and length \(b - a \gg h\). We shall assume each sensor has sensing range \(\frac{r}{2}\), so that two sensors will have overlapping sensing regions when they are within distance \(r\) of each other. (The reason for this choice of normalization will become clearer when we discuss the graphs \(G_{h,r}\) and \(G_{h,r}(a, b)\) defined below.) As scaling all distances by a constant factor does not change the model, we will assume without loss of generality that the scale is chosen so that there is on average one sensor per unit area. If the sensors are required to lie inside the strip, then a vertical crossing path at \(x = a\) or \(x = b\) is more likely to avoid the sensors than one near the middle of the strip. Thus the regions of the strip close to the ends are in some sense atypical. To avoid these “boundary problems”, we shall extend the strip horizontally and allow sensors at points \((x, y)\) with \(x < a\) or \(x > b\). We shall however always insist that \(y \in [0, h]\). To do this, rather than fixing the total number of sensors, we shall instead place the sensors according to a Poisson point process with intensity 1 in the infinite strip \(\mathbb{R} \times [0, h]\). This now ensures that any horizontal translate of a crossing path has the same chances of avoiding the sensors. Of course, if \(b - a \to \infty\), then there will almost surely be a crossing path that avoids the sensors. Thus we shall mainly be concerned with the frequency of breaks along the strip where a crossing path can avoid sensors. Results for a finite strip will then be deduced from this infinite model. This therefore motivates the following formalized version of our problem.

We consider an infinite strip \(S_h = \mathbb{R} \times [0, h]\) of width (or height) \(h\) and place sensors inside this strip randomly according to a Poisson point process \(\mathcal{P}\) of intensity 1. We shall assume that each sensor has the ability to detect intruders strictly within a (Euclidean) distance \(\frac{r}{2}\) and write \(D_v = \{x \in \mathbb{R}^2 : \|x - v\|_2 < \frac{r}{2}\}\) for the sensing region of a sensor located at a point \(v \in S_h\). Construct an infinite random geometric graph \(G_{h,r}\) with vertex set given by the set of sensors \(\mathcal{P}\), by joining every sensor to every other sensor that is strictly within distance \(r\), that is, whenever their sensing regions intersect. For \(a < b\), define \(G_{h,r}(a, b)\) to be the subgraph of \(G_{h,r}\) consisting of sensors with \(x\)-coordinate lying in the interval \((a - \frac{r}{2}, b + \frac{r}{2})\), and with two extra “virtual” sensors \(s\) and \(t\), where \(s\) is joined to all sensors of \(\mathcal{P}\) strictly within distance \(\frac{r}{2}\) of the line \(x = a\) and \(t\) is joined to all sensors of \(\mathcal{P}\) strictly within distance \(\frac{r}{2}\) of the line \(x = b\) (see Figure 1). The sensors \(s\) and \(t\) are never joined directly to each other.

Define a separating path to be a continuous simple path in \(S_h\) starting at some point...
Figure 1: Model with sensor range $\frac{r}{2}$ (sensing regions indicated by shading). Dotted line indicates a separating path — either a possible path of an undetected intruder, or a path disconnecting $s$ from $t$ in the network.

on the line $y = 0$, ending at some point on the line $y = h$ and not passing strictly within distance $\frac{r}{2}$ of any sensor. The following is immediate from the above definitions and standard topological properties of the plane.

**Lemma 1.** The graph $G_{h,r}(a,b)$ contains a path from $s$ to $t$ if and only if there is no separating path of the infinite graph $G_{h,r}$ that lies entirely within the rectangle $[a,b] \times [0,h]$.

Thus the questions of avoiding separating paths, and obtaining long paths along $G_{h,r}$ are effectively equivalent, and barrier coverage can also be thought of as a question of “percolation” in the Gilbert disk model when confined to a thin strip. Of course, strictly speaking there can be no percolation in a strip of bounded width, as an infinite strip almost surely has a large gap somewhere with no sensors, and hence no infinite component can exist in $G_{h,r}$. However, for large $r$ and $h$, components can exist that are exponentially large, and it is these components, and the gaps between them, that we shall be interested in.

The presence of a separating path indicates a “break” in the coverage. The aim will be to determine the frequency or *intensity* $I_{h,r}$ of these breaks, defined so that the number of breaks in $[a,b] \times [0,h]$ is on average $(b-a)(I_{h,r} + o(1))$ as $b-a \to \infty$ (the $o(1)$ error including, for example, end effects near $x = a$ and $x = b$). We also wish to estimate the probability that a break occurs in a finite interval, or more generally the probability distribution of the number of breaks. We aim to define the notion of a break so as to make breaks “almost independent”. This will imply that the number of breaks in $G_{h,r}(a,b)$ is given approximately by a Poisson variable of mean $(b-a)I_{h,r}$, and in particular, the probability that no break occurs is about $\exp(-(b-a)I_{h,r})$.

The definition of a break is more delicate than it might appear at first. For example, in Figure 2, should the two separating paths be considered as defining the same break, or two different breaks? Since there may be several small connected “islands” in the break, one may be able to construct very many homotopically distinct separating paths. The two
paths in Figure 2 are not really independent since the existence of one makes the existence of the other much more likely. If we were to count these paths as distinct breaks, then the probability distribution of the number of breaks in $G_{h,r}(a,b)$ could be very far from Poisson, and the probability of a break could be much less than the expected number of breaks, even when the expected number of breaks is much less than 1. For this reason, we wish to consider the situation in Figure 2 as a single break. There are several suitable definitions of a break, which although different, give the same asymptotic frequencies when $h$ and $r$ are large. The following definition is chosen since it is easy to compute in simulations and is fairly convenient theoretically.

Define a good component as a (graph) component $C$ of the infinite graph $G_{h,r}$ which contains a sensor strictly within distance $\sqrt{3}/2 r$ of the top boundary $\partial S_h^+$ of $S_h$, and also contains a sensor strictly within distance $\sqrt{3}/2 r$ of the bottom boundary $\partial S_h^-$ of $S_h$. We shall show (see the comments before Lemma 7 below) that there is a natural left-to-right ordering of good components. Indeed, the $\sqrt{3}/2 r$ in the definition was chosen precisely so that good components cannot “jump” past one another\footnote{One could use any distance that is at most $\sqrt{3}/2 r$ here, however the choice $\sqrt{3}/2 r$ is also important in the proof of Lemma 23.}. Now define a break to be the gap between two consecutive good components. In other words, a break is a partition of the set of good components into two classes, those on the left of the break, and those on the right, which is compatible with the left-right ordering of the good components.

Our main results are the following.

Theorem 2. Assume $r \geq 7$. Then the intensity of breaks almost surely exists and satisfies

$$I_{h,r} = r^{1/3} \varepsilon(h r^{-1/3}) e^{-hr+O(hr^{-5/3})}$$

where $\varepsilon(z)$ is the function defined in Section 6 below, and the constant implicit in the $O(\cdot)$ notation is independent of both $r$ and $h$. 

\footnote{One could use any distance that is at most $\sqrt{3}/2 r$ here, however the choice $\sqrt{3}/2 r$ is also important in the proof of Lemma 23.}
Note that \( h \) is arbitrary in Theorem 2, so the intensity is determined up to a factor of \( 1 + o(1) \) when \( h = o(r^{5/3}) \). Even for larger \( h \), the logarithm of the intensity is determined up to a factor of \( 1 + o(1) \) when \( r \to \infty \). More details are given on the function \( \varepsilon(z) \) in Section 6, but we just note here the following, which allows us to give a simple explicit expression for the intensity when \( h = \omega(r^{1/3}) \).

**Theorem 3.** There exist constants \( \alpha > 0 \), \( \beta \), and \( \eta > 0 \), such that

\[
\varepsilon(z) = \exp\left\{ \alpha z + \beta + O(e^{-\eta z^{1/9}}) \right\}
\]
as \( z \to \infty \).

**Remark.** The most "obvious" obstruction to barrier coverage is an \( r \times h \) rectangle containing no sensor (see Section 2). Based on this one would expect a break intensity of order \( h e^{-hr} \). However, as \( h \to \infty \) the actual break intensity is exponentially larger, so the most obvious obstructions do not dominate the break intensity. Indeed, the contribution given by the obvious obstructions corresponds to taking \( \varepsilon(z) = z \). However, \( \varepsilon(z) \gg z \) when \( z \gg 1 \) (see also (26)). Equivalently, non-obvious obstructions dominate when \( h \gg r^{1/3} \).

Having estimated the intensity of breaks, we have the following result on the distribution of breaks. Here we say that \( G_{h,r}(a,b) \) contains a break, if there is a separating path that lies entirely within \([a, b] \times [0, h]\) that induces this break (see Lemma 7). Note that by Lemma 1, \( G_{h,r}(a,b) \) contains an s-t path if and only if it contains no break.

**Theorem 4.** Fix \( r \geq 7 \), \( x > 0 \), and \( k \geq 0 \). Then the probability that \( G_{h,r}(0, x/I_{h,r}) \) contains exactly \( k \) breaks tends to \( e^{-x} x^k / k! \) as \( h \to \infty \). In particular, the probability of barrier coverage in \([0, x/I_{h,r}] \times [0, h]\) tends to \( e^{-x} \) as \( h \to \infty \).

## 2 Heuristics

Before giving proofs and further results, we introduce some non-rigorous heuristics about the types of component of \( G_{h,r} \) and their frequency. The purpose is to help provide some intuitive explanation for our results. We generally assume \( r \) is large, so that in most areas the graph \( G_{h,r} \) is highly connected. The main idea is the concept of "excluded area", that is that rare configurations can be described by the absence of sensors within some nice region \( A \). Outside \( A \), the density of sensors will be assumed to be close to the expected density, which will result in a high degree of connectivity in the graph \( G_{h,r} \).

Given a region \( A \subseteq S_h \), the probability that it contains no sensors is \( e^{-|A|} \), where \( |A| \) is the area of \( A \). We consider the minimal regions \( A \) that can force a component, or a break, under the assumptions that there is a reasonable density of sensors outside of \( A \).
For example, an excluded $r \times h$ rectangle across the strip is very likely to cause a break (see Figure 3). On either side of this rectangle, the components are very likely to be good (for large $r$), but the region disconnects $G_{h,r}$. This region is the smallest such empty region that results in a break, so we might expect most breaks to be approximately rectangular, and the frequency of breaks $I_{h,r}$ to be about $e^{-hr}$. (More precisely $he^{-hr}$, since such a rectangle can be placed after any sensor, and there are about $h$ sensors per unit distance along the strip.) Small components can form near the boundary of $S_h$ with an excluded area of $\frac{1}{2}\pi r^2$, or in the interior of $S_h$ with excluded area $\pi r^2$. Hence these are likely to occur with frequencies about $e^{-\pi r^2/2}$ and $e^{-\pi r^2}$ respectively. Note that if $h < \frac{\pi}{2}r$, then breaks should be more common than these small components, and so most components are likely to be good. On the other hand, if $h > \frac{\pi}{2}r$, then most components are likely to be small (bad) components that do not form breaks.

Compound breaks (breaks containing homotopically inequivalent separating paths) need an excluded area of at least $rh + (\frac{\pi}{3} - \frac{\sqrt{3}}{4})r^2 > rh$ (see fourth example in Figure 3), so at first sight these seem far rarer than simple breaks. However, we need to take care of combinatorial issues — how many ways such breaks can occur. It is possible for compound breaks to be more common than simple breaks if $h$ is extremely large, in particular if $1 \ll r^2 \ll \log h$. This is because although we lose a factor of $e^{-O(r^2)}$ in the frequency of compound breaks due to the extra $O(r^2)$ excluded area, we gain a combinatorial factor of order $h$ due to the choice of the vertical position of the small component inside the break. (The excluded area is still $hr + O(r^2)$ even if the small component is in the center of the break.)

Although vertical excluded rectangles are the most obvious breaks, diagonal breaks are also possible. If the top of the break is displaced by a distance $d$, the excluded area becomes $r\sqrt{h^2 + d^2} \approx rh + \frac{rd^2}{\sqrt{h}}$. If $rd^2 \sim h$ this does not impose a large penalty, so we may expect breaks to often deviate by about $O(\sqrt{h/r})$ from vertical.

These heuristics tend to be good for large $r$, however, corrections are needed for the effects of finite $r$. These corrections tend to make breaks more likely. Indeed, the main difference between the $e^{-hr}$ heuristic and Theorem 2 is given by the $\varepsilon(hr^{-1/3})$ factor, which by Theorem 3 gives an extra positive $\Theta(hr^{-1/3})$ term in the exponent. Determining this extra term is rather difficult; in fact this is the main contribution of this paper. This extra term arises as a trade off between combinatorial factors counting the many different types of break one can have, versus the penalty one pays in extra excluded area for non-rectangular breaks. Roughly speaking, one gains an exponential factor per unit length ($h$) of the break due to the extra “entropy” allowed by making the walls of the excluded area fluctuate. Larger $r$ reduces this effect as these walls cannot curve so much without increasing the excluded area significantly.
Figure 3: Minimal excluded areas for breaks, small components near the boundary and in the interior, compound breaks, and diagonal breaks.

3 Good components and crossing sensor-paths

Recall that a good component is a component of $G_{h,r}$ that has sensors within distance $\frac{\sqrt{3}}{2}r$ of both the top and bottom boundaries of the strip $S_h = \mathbb{R} \times [0, h]$.

Define a crossing sensor-path to be a path from the bottom of $S_h$ to the top of $S_h$ consisting of line segments, with the first line segment being a vertical line segment $bv_1$ of length at most $\frac{\sqrt{3}}{2}r$ from some point $b \in \partial S_h$ to some sensor $v_1$ of $G_{h,r}$. Then line segments corresponding to a graph theoretic walk $^2v_1, \ldots, v_n$ through $G_{h,r}$, and finally a vertical line segment $v_nt$ of length at most $\frac{\sqrt{3}}{2}r$ to a point $t \in \partial S_h^+$ (see Figure 4). Note that a crossing sensor-path need not be a simple curve and there exists a crossing sensor-path going though some sensor if and only if that sensor is contained in a good component.

More generally, define a sensor-path to be a walk through $G_{h,r}$ which may optionally also include an initial and/or final vertical segment of length at most $\frac{\sqrt{3}}{2}r$ to the boundary $\partial S_h$ of $S_h$. Sensor-paths can of course exist in any component of $G_{h,r}$.

Notice that the concepts of good components and sensor-paths are defined in terms of the infinite graph $G_{h,r}$, and are not related to the particular interval $[a, b]$ along the $x$-axis that we are concerned with.

We should also note that from now on, when we refer to, for example, a point to the right of a separating path or crossing sensor-path $\gamma$, we mean right in the topological sense of belonging to the infinite component of $S_h \setminus \gamma$ that is unbounded to the right. Hence points to the left of such a path may in fact have a larger $x$-coordinate than another point that is to the right of the same path. We now give some simple consequences of these definitions.

Lemma 5. If $C$ is a component of $G_{h,r}$ and $v$ is a sensor that lies within distance $\frac{r}{2}$ of a sensor-path $\gamma_C$ through $C$, then $v \in C$.

$^2$A walk through a graph is a sequence of vertices $v_1, \ldots, v_n$ such that each $v_iv_{i+1}$ is an edge. We allow both edges and vertices to be repeated.
Proof. Let $z$ be the closest point on $\gamma_C$ to $v$, so $\|z-v\| < \frac{r}{2}$. If $z$ is on some edge $v_iv_{i+1}$ of $\gamma_C$ then $\|v_i-v_{i+1}\| < r$, so either $\|z-v_i\| < \frac{r}{2}$ or $\|z-v_{i+1}\| < \frac{r}{2}$. Thus either $\|v-v_i\| < r$ or $\|v-v_{i+1}\| < r$ implying that $v$ is adjacent to either $v_i$ or $v_{i+1}$ in $G_{h,r}$, and so $v \in C$. Now assume $v$ lies on a vertical segment $bv_1$, say, meeting $\partial S_h^{-}$. If $z = v_1$ we are done, since then $v$ is within distance $\frac{r}{2}$ of the sensor $v_1 \in C$. Otherwise, the segment $vz$ must be horizontal and $\|v-v_1\|^2 = \|v-z\|^2 + \|z-v_1\|^2 < (\frac{r}{2})^2 + (\frac{\sqrt{3}}{2}r)^2 = r^2$. Thus once again, $v$ is within distance $r$ of a vertex of $C$, so $v \in C$.

Lemma 6. If $C$ and $C'$ are distinct components, then the minimum distance between any two sensor-paths $\gamma_C$ and $\gamma_{C'}$, through $C$ and $C'$ respectively, is at least $\frac{r}{2}$.

Proof. Let $z$ and $z'$ be the closest pair of points with $z$ on $\gamma_C$ and $z'$ on $\gamma_{C'}$, and assume $\|z-z'\| < \frac{r}{2}$. Since the paths are piecewise linear, we may assume one of $z$ and $z'$ (say $z$) is a corner or an endpoint of its path. If $z$ is a corner, then it is one of the vertices of $C$. But then $z \in C'$ by Lemma 5, a contradiction since $C$ and $C'$ are distinct components. Thus we may assume $z$ is either a top $t$ or bottom $b$ boundary point of $S_h$ on $\gamma_C$, say $z = b \in \partial S_h^−$. But since $z$ is the closest point on a vertical segment of $\gamma_C$ to $z'$, and $z' \in S_h$, $z'$ must be a bottom boundary point $b'$ of $S_h$ on $\gamma_{C'}$. But then either $v_1$ or $v'_1$ must be within distance $\frac{r}{2}$ of $\gamma_{C'}$ or $\gamma_C$ respectively, where $bv_1$ and $b'v'_1$ are vertical segments of $\gamma_C$ and $\gamma_{C'}$. Thus either $v_1 \in C'$ or $v'_1 \in C$, a contradiction.

By Lemma 6 we see that crossing sensor-paths of distinct components cannot intersect. On the other hand, we can find a crossing sensor-path of a good component $C$ that intersects all other crossing sensor-paths of $C$ (by, for example, having it pass through every vertex of $C$). Thus we can order the good components from left to right according to the order of their crossing sensor-paths. Recall that a break is a gap between two consecutive good components.

Lemma 7. Any separating path $\gamma$ partitions the good components into those that lie to the left of $\gamma$ and those that lie to the right of $\gamma$, so in particular defines a break. Conversely, for any break there exists such a separating path $\gamma$. Indeed, if the break occurs between good components $C$ and $C'$ then we can choose $\gamma$ to lie between every pair of crossing sensor-paths $\gamma_C$ and $\gamma_{C'}$ through $C$ and $C'$ respectively.

Proof. If some sensor $x$ lies to the left of $\gamma$ and some sensor $y$ lies to the right of $\gamma$, then the line segment $xy$ meets $\gamma$ at some point $z$, say. Since $\|x-z\|, \|y-z\| \geq \frac{r}{2}$ we have $\|x-y\| \geq r$, and so $x$ and $y$ are not adjacent in $G_{h,r}$. Thus any separating path disconnects $G_{h,r}$, and so defines a break. For the converse, suppose $C$ is a good component. Define $S_C := (\bigcup_{\gamma_C} \gamma_C) \cup (\bigcup_{v \in C} \overline{D_v})$ to be the closed region consisting of the union of all crossing sensor-paths $\gamma_C$ through $C$ and the closures of the sensing regions of $C$. Note that $S_C$
forms a connected subset of the plane. Indeed, \( \bigcup_{v \in C} \overline{D_v} \) is connected, and all crossing sensor-paths \( \gamma_C \) meet this set. No sensor outside of \( C \) is within distance \( \frac{r}{2} \) of any \( \gamma_C \) (by Lemma 5) or within distance \( r \) of any sensor in \( C \). Thus no sensor outside of \( C \) is within distance \( \frac{r}{2} \) of \( S_C \). Since each \( \gamma_C \) intersects both the top and bottom boundaries of \( S_h \), so does \( S_C \). Let \( \gamma \) be the rightmost boundary of \( S_C \) in \( S_h \), i.e., the intersection of \( S_C \) with the closure of the component of \( S_h \setminus S_C \) that is unbounded to the right. Now \( \gamma \) is a separating path (made up from line segments and arcs of circles), no part of which is to the left of any \( \gamma_C \). If \( C' \) is a good component that lies to the right of \( C \) and \( \gamma_{C'} \) is a crossing sensor-path through \( C' \), then no point of \( \gamma_{C'} \) can lie to the left of \( \gamma \), otherwise some point of \( \gamma_{C'} \) would be within distance \( \frac{r}{2} \) of some sensor in \( C \), contradicting Lemma 5, or would meet some \( \gamma_C \), contradicting Lemma 6. Thus \( \gamma \) lies between any pair of crossing sensor-paths \( \gamma_C \) and \( \gamma_{C'} \).

\[
\gamma_C \quad \gamma_{C'}
\]

Figure 4: Crossing sensor-path \( \gamma_C \) (solid path) and \( x^\pm_C \). Crosses indicate the rightmost point to the left of every \( \gamma_C \), or leftmost point to the right of every \( \gamma_C \) that a good separating path can pass through. Note that good separating paths cannot pass beneath \( v_1 \).

4 Widths of good components and breaks

Call a separating path good if it is as in Lemma 7, i.e., it does not cross any crossing sensor-path.

For any good component \( C \) define \( x^-_C \) to be the smallest \( x \)-coordinate of any point on any good separating path to the right of \( C \). Similarly define \( x^+_C \) to be the largest \( x \)-coordinate of any point on any good separating path to the left of \( C \) (see Figure 4). Define the width of a good component \( C \) as \( x^+_C - x^-_C \) and define the width of a break between consecutive good components \( C \) and \( C' \) as \( x^-_{C'} - x^+_C \).

Note that the width of a good component may be negative, as a good separating path to the right of \( C \) may extend further to the left than the rightmost point of a good separating path on the left of \( C \). The width of a break, between \( C \) and \( C' \) say, is however always
positive. In fact this width is (almost surely) the maximum horizontal extent of a good
separating path $\gamma$ between $C$ and $C'$. Indeed, if there were good separating paths meeting
$(x^C, y^C)$ and $(x^-C, y^-C)$, say, but no good separating path meeting both, then there would
be a collection $C''$ of sensors whose combined sensing region $\bigcup_{v \in C'} D_v$ separates these two
points. But then there would be a good component between $C$ and $C'$.

**Lemma 8.** If $\gamma$ is a separating path between good components $C$ and $C'$ that lies entirely
in $[x, x'] \times [0, h]$, then there is a good separating path between $C$ and $C'$ that also lies in
$[x, x'] \times [0, h]$. Also, some point of $\gamma$ lies in $[x^C, x^-C] \times [0, h]$.

**Proof.** Let $\gamma^L$ be the good separating path constructed in Lemma 7 as the rightmost
boundary path of $S_C = (\bigcup_{v \in C} \gamma^C) \cup (\bigcup_{v \in C} D_v)$. The rightmost point $(x_L, y_L)$ of $\gamma^L$ lies
on the boundary of some disk $D_v$, $v \in C$. As $\gamma$ must pass to the right of $v$ and cannot
approach within distance $\frac{r}{2}$ of $v$, we must have $x' \geq x_L$. Thus replacing $\gamma$ by the rightmost
boundary of $\gamma \cup \gamma^L$ results in a separating path that does not cross any crossing sensor-path
$\gamma_C$ of $C$ and still lies in $[x, x'] \times [0, h]$. Repeating this process using the good separating
path $\gamma^R$ which is the leftmost boundary of $S_{C'} = (\bigcup_{v \in C'} \gamma^{C'}) \cup (\bigcup_{v \in C'} D_v)$ results in a good
separating path in $[x, x'] \times [0, h]$.

Now take $[x, x']$ minimal, so that $[x, x']$ is the projection onto the $x$-axis of $\gamma$. Then as
any good separating path lies in $[x^C, x^-C] \times [0, h]$, $[x, x']$ must intersect $[x^C, x^-C]$, and so $\gamma$
contains a point in $[x^C, x^-C] \times [0, h]$.

As a consequence of Lemma 8, we see that Lemma 1 holds even if we insist on the
separating path being good.

Define the *intensity* $I_{h, r}$ of breaks as the limit $\lim_{\ell \to \infty} \frac{N_\ell}{\ell}$, where $N_\ell$ is the number
of breaks of $G_{h, r}$ containing separating paths in $[0, \ell] \times [0, h]$. It is easy to see that this limit
exists. Indeed, translational invariance and long range independence imply that horizontal
translation is an ergodic transformation on the probability space of this model. Thus almost
surely, breaks (or any other event that can be defined in a translational invariant manner)
occur with a well defined and deterministic frequency along the strip. In particular, $I_{h, r}$ is
almost surely well defined and constant. It is clear that $I_{h, r}$ is also equal to the intensity
of good components (defined in any reasonable way). Moreover, it makes sense to talk
about widths of random good components and breaks. For example, if $W$ is the width of a random break then $\mathbb{P}(W > x)$ is almost surely equal to the asymptotic proportion of
breaks with width $W > x$, and is well defined for any $x \in \mathbb{R}$. We now prove (fairly crude)
bounds on the average widths of breaks and good components.

**Lemma 9.** Assume $r \geq 7$. Then the average width of a break is at most
$$\max\{5h, 1/h + 2h\}.$$
Proof. We first deal with the case when $h < \frac{\sqrt{3}}{2}r$. In this case every vertex is within distance $\frac{\sqrt{3}}{2}r$ of both the top and the bottom of $S_h$, so in particular, every component is good. Consider two points of the Poisson process, $(x_1, y_1)$ and $(x_2, y_2)$, that are consecutive in the horizontal ordering of the points given by their $x$-coordinates. Counting breaks is equivalent to counting such pairs that lie in distinct components of $G_{h,r}$. Let $d = \max\{y_1, h - y_1\}$ and assume $(x_i, y_i)$ lies in the component $C_i$, $i = 1, 2$. If $d \leq \frac{r}{2}$ then the sensing region of $(x_1, y_1)$ covers the strip $S_h$ up to $x$-coordinate $x_1 + (\frac{r^2}{4} - d^2)^{1/2} \geq x_1 + \frac{r}{2} - d \geq x_1 + \frac{r}{2} - h$. If $d > \frac{r}{2}$ then $h > \frac{r}{2}$, but there is still a vertical crossing sensor-path at $x = x_1$. Thus in all cases $x_{C_1} \geq x_1 + \frac{r}{2} - h$. Similarly $x_{C_2} \leq x_2 - \frac{r}{2} + h$, so the width of the break is at most $(x_2 - \frac{r}{2} + h) - (x_1 + \frac{r}{2} - h) = (x_2 - x_1 - r) + 2h$. Let $Z = x_2 - x_1 - r$. If $Z < 0$ a break may or may not be formed, but if it is then its width is at most $2h$. On the other hand, if $Z \geq 0$ then $x_2 - x_1 \geq r$, so there is always a break and its width is at most $Z + 2h$. Conditioned on $Z \geq 0$, $Z$ is an exponential random variable with mean $1/h$. Thus the average width of a break is at most $1/h + 2h$ conditioned on $Z \geq 0$ and at most $2h$ otherwise. Thus the unconditioned average width of a break is also at most $1/h + 2h$.

We may now assume that $h \geq \frac{\sqrt{3}}{2}r$. We shall now bound the intensity of breaks with large width. Tile $S_h$ with $a \times b$ rectangles, where $a$ and $b$ are chosen so that

(i) $h/b$ is an integer;

(ii) the diameter $\sqrt{a^2 + b^2}$ of these rectangles is $\frac{r}{2}$;

(iii) the area $ab$ of these rectangles is maximized subject to (i) and (ii).

For large $h/r$, the rectangles are approximately square with side length $r/\sqrt{8}$, however to tile $S_h$ we need $h/b$ to be integral, and this decreases the area $ab$ slightly. It is easy to see that for any $h \geq \frac{\sqrt{3}}{2}r$,

$$a, b \in \left[\frac{r}{\sqrt{12}}, \frac{r}{\sqrt{3}}\right], \quad \text{and} \quad ab \geq \frac{\sqrt{3}}{12}r^2;$$

the worst case being when $h = \frac{\sqrt{3}}{2}r$, $h/b = 3$, and as one increases $h$, extreme values of $a$ and $b$ occur when the aspect ratio of the rectangles is $k : k + 1$ for some $k \geq 3$.

Fix one of these tiling rectangles $R$ adjacent to $\partial S_h^-$. We estimate the probability $p$ that there exists a good separating path $\gamma$ of width at least $w$ meeting $R$, with $R$ being the leftmost rectangle meeting both $\gamma$ and $\partial S_h^-$. Take the set of rectangles that intersect $\gamma$ and regard this set as the vertices of a graph $G$, rectangles being joined if they share a common edge. Thus $G$ is a subgraph of the square lattice. The rectangles of $G$ cannot contain any point of the Poisson process, since all the points of these rectangles lie within $\frac{r}{2}$ of $\gamma$ and $\gamma$ is a separating path. Ignoring probability zero events, it is clear that we
can, without loss of generality, assume that $\gamma$ goes through no corner of any rectangle, so that $G$ is a connected graph that joins $R$ to $\partial S_h^+$, and joins the leftmost rectangle to the rightmost rectangle meeting $\gamma$. It is easy to see that one can take a connected subset $G'$ of these rectangles containing exactly $n = h/b + \lfloor w/a \rfloor$ rectangles, still meeting our original rectangle $R$, although $G'$ may now fail to meet the top of $S_h$ or the leftmost/rightmost rectangles. For example, any spanning tree of $G$ rooted at $R$ contains at least $\frac{h}{b} - 1$ vertical and at least $\lfloor w/a \rfloor$ horizontal edges, and so $G$ contains at least $n$ vertices. Now repeatedly prune leaves until $G'$ contains exactly $n$ vertices. Is is well known that the number of connected subgraphs of order $n$ of a graph of maximum degree $\Delta$ containing a specified vertex is at most $(e\Delta)^n$ (see, for example, [5]). Thus the number of such choices of the subgraph $G'$ of the square lattice rooted at $R$ is at most $(4e)^n$. However, we can do a bit better. Using the results of Klarner and Rivest [10] (equation (7) and the preceding discussion\(^3\)) the number of connected subsets of size $n$ of the square lattice up to translation is at most
\[
\sum_k \binom{n}{k, k+1, n-2k-1} 2^{n-k-1} = \frac{1}{2} \left( \sum_k \binom{n}{k, k+1, n-2k-1} 2^{n-k-1} + \sum_k \binom{n}{k+1, k, n-2k-1} 2^{n-k-1} \right) \\
\leq \frac{1}{2\sqrt{2}} 2^n \left( 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \right)^n \\
< \frac{1}{2} (2 + 2\sqrt{2})^n.
\]
As we shall use this repeatedly, we shall state it as a lemma.

**Lemma 10.** For all $n \geq 1$, the number of equivalence classes of connected subsets of the square lattice of size $n$ up to translation is at most $\frac{1}{2} \mu^n$, where $\mu = 2 + 2\sqrt{2}$.

Since we assumed that $R$ is the leftmost rectangle on the lowest level of this subset, fixing $R$ determines which translate of this set we have. Thus there are at most $\frac{1}{2} \mu^n$ possible choices for $G'$, where $\mu = 2 + 2\sqrt{2}$. These rectangles form an empty region of area $nab$, and so the probability $p$ that there exists such a separating path $\gamma$ satisfies
\[
p \leq \frac{1}{2} \mu^n e^{-nab} \leq \frac{1}{2} (\mu e^{-ab})^n \leq \frac{1}{2} (\mu e^{-ab})^h/b + w/a - 1,
\]
where we are using the fact that $\lfloor w/a \rfloor \geq w/a - 1$ and $\mu e^{-ab} < 1$. It is conceivable (if rather implausible) that two breaks may give rise to the same graph $G'$. However, it is clear that three breaks cannot all give rise to the same graph. Indeed, if $\gamma$, $\gamma'$, $\gamma''$ are good separating paths giving rise to three successive breaks, then there are crossing sensor-paths

---

\(^3\)Although they prove a slightly sharper asymptotic bound on the number of these “lattice animals”, it is not so obvious what their stronger upper bound is for fixed $n$. Hence we use this slightly weaker result.
\(\gamma_C\) and \(\gamma_{C'}\) corresponding to good components \(C\) and \(C'\) such that \(\gamma < \gamma_C < \gamma' < \gamma_{C'} < \gamma''\) in the left-right ordering of these paths (none of these paths cross). Since no point of \(\gamma_C\) is within \(\frac{r}{2}\) of any point of \(\gamma_{C'}\), it follows that no point of \(\gamma\) is within distance \(\frac{r}{2}\) of any point of \(\gamma''\), and thus the sets of rectangles these paths pass through are disjoint. Consequently, the intensity \(I_{h,r}(w)\) of breaks of width at least \(w\) can be bounded above by \(2p/a\), that is
\[
I_{h,r}(w) \leq \frac{1}{a}(\mu e^{-ab})^{h/b+w/a-1}.
\]
Now we need a lower bound on the total intensity of breaks. Consider a vertical strip of rectangles of the tiling, each containing at least one point of \(P\). The vertices in these rectangles will form (part of) a good component. If the \(r \times h\) rectangle immediately to the right of this strip is empty, it will certainly force a break. The probability of this occurrence is
\[
(1 - e^{-ab})^{h/b} e^{-hr},
\]
so the intensity \(I_{h,r}\) of breaks can be bounded below by
\[
I_{h,r} \geq \frac{1}{a}(1 - e^{-ab})^{h/b} e^{-hr}.
\]
Hence, if \(W\) is the (positive-valued) random variable expressing the width of a break and \(w \geq a\), then
\[
\mathbb{P}(W \geq w) = \frac{I_{h,r}(w)}{I_{h,r}} \leq (\mu e^{-ab})^{h/b+w/a-1}(1 - e^{-ab})^{-h/b} e^{hr} \leq e^{hr}\left(\frac{\mu e^{-ab}}{1 - e^{-ab}}\right)^{h/b+(w-a)/a} = e^{(r-\kappa a)h - ab(w-a)}
\]
where
\[
\kappa := -\frac{1}{ab} \log \left(\frac{\mu e^{-ab}}{1 - e^{-ab}}\right) = 1 - \frac{1}{ab} \log \left(\frac{\mu}{1 - e^{-ab}}\right) \geq 0.72.
\]
The last inequality here follows from the fact that \(\mu = 2 + 2\sqrt{2}\) and, by (1), \(ab \geq \frac{\sqrt{2}}{12} r^2 \geq 5.77\) as \(r \geq 7\). Now, for all \(x \geq a\) we have by (3)
\[
\mathbb{E}(W) \leq x + \int_x^\infty \mathbb{P}(W \geq w) \, dw \leq x + \frac{1}{kb} e^{(r-\kappa a)h - \kappa b(x-a)}.
\]
To minimize the right hand side, we set \(x = a + \frac{(r-\kappa a)h}{\kappa b} \geq a\), at which point the exponent in the second term is 0. Hence
\[
\mathbb{E}(W) \leq a + \frac{(r-\kappa a)h}{\kappa b} + \frac{1}{kb} = \frac{r}{\kappa b} h + \frac{\kappa ab - \kappa ah + 1}{\kappa b}.
\]
Now \( h/b \geq 3 \) and \( \kappa_{ab} \geq (0.72)(5.77) > 1 \), so \( \kappa_{ab} - \kappa_{ah} + 1 \leq -2\kappa_{ab} + 1 < 0 \). Also, by (1), \( b \geq r/\sqrt{12}, \) so

\[
\mathbb{E}(W) \leq (\sqrt{12}/\kappa)h \leq 5h.
\]

\[\square\]

**Lemma 11.** Assume \( r \geq 7 \). Then the proportion of good components with width less than \( w \geq 0 \) is at most

\[
(w + 3)e^{-h/3}.
\]

**Proof.** As in the proof of Lemma 9, we consider the case \( h < \sqrt{3} r/2 \) first. In this case all components are good. The width of a good component is at least as large as the horizontal distance between its leftmost and rightmost points as we can take vertical crossing sensor-paths through these points and no good separating path can cross these. Any two points within horizontal distance \( 3 < \frac{r}{2} \) are within distance \( r \) and so are in the same component. Thus we can stochastically bound from below the width of a good component by a random variable

\[
X = \sum_{i=1}^{T-1} X_i
\]

where \( X_i \) are i.i.d. exponential random variables of mean \( 1/h \) giving the horizontal distance between consecutive sensors, and \( T = \min\{i : X_i \geq 3\} \). Consider \( \mathbb{E}e^{-\alpha X} \) for \( \alpha > 0 \). Conditioning on the value of \( X_1 \) we obtain

\[
\mathbb{E}(e^{-\alpha X}) = \mathbb{P}(X_1 \geq 3) + \int_0^3 \mathbb{E}(e^{-\alpha(X' + z)})he^{-hz} dz
\]

\[
= e^{-3h} + \mathbb{E}(e^{-\alpha X'}) \int_0^3 he^{-(h+\alpha)z} dz
\]

\[
\leq e^{-3h} + \mathbb{E}(e^{-\alpha X'}) \frac{h}{h+\alpha},
\]

where \( X' = \sum_{i=2}^{T'} X_i, T' = \min\{i > 1 : X_i \geq 3\} \), has the same distribution as \( X \). Since \( e^{-\alpha X} \) is bounded,

\[
\mathbb{E}(e^{-\alpha X}) \leq \frac{h+\alpha}{\alpha} e^{-3h},
\]

so

\[
\mathbb{P}(X < w) \leq \mathbb{E}(e^{\alpha w - \alpha X}) \leq \frac{h+\alpha}{\alpha} e^{\alpha w} e^{-3h}.
\]

Setting \( \alpha = 1/w \) gives \( \mathbb{P}(X < w) \leq (wh + 1)e^{1-3h} \). But \( he^{1-h} \leq 1 \) for all \( h \), so

\[
\mathbb{P}(X < w) \leq we^{-2h} + ee^{-3h} \leq (w + 3)e^{-h/3}
\]

(with room to spare).

Turning to the heart of the proof, we now assume \( h \geq \sqrt{3}/2 r \). For each good component \( C \) we shall define a “leftmost” crossing sensor-path \( \gamma_C \) through \( C \). Given any good separating path \( \gamma \) to the left of \( C \), we “explore” the Poisson process \( \mathcal{P} \) to the right of \( \gamma \) in order to
find $\gamma^L_C$. Let $U$ be the set of points $u \in S_h$ that are at distance at least $(\sqrt{3} - 1)\frac{r}{2}$ from the boundary of $S_h$, are to the right of $\gamma$, and such that $D_u = \{x : \|x - u\| < \frac{r}{2}\}$ does not contain any vertex of $P$. Let $U_0$ be the component of $U$ (as a subset of $\mathbb{R}^2$) that is connected to $\gamma$. Note that one can determine $U_0$ by examining just the union of the regions $\overline{D}_u$ with $u \in U_0$. Indeed, one can imagine the points of $P$ as pegs, and $\overline{D}_u$ a disk that can be moved around as long as it is not obstructed by these pegs, and does not get too close to $\partial S_h$. If we let $S = \bigcup_{u \in U_0} \overline{D}_u$ then the rightmost boundary of $S$ consists of a sequence of arcs joining points of $P$. The points of $P$ encountered along this boundary form the path $\gamma^C_C$ (see Figure 5). Indeed, the endpoints of the arcs forming the boundary of $S$ must be strictly within $r$ of the next endpoint, or within $\frac{\sqrt{3}}{2}r$ of the boundary of $S_h$. Thus they form a crossing sensor-path. Conversely, given any crossing sensor-path $\gamma_C$, $U_0$ must lie to the left of $\gamma_C$ as the center $u$ of the disk $\overline{D}_u$ cannot cross an edge between two vertices of $\gamma_C$ and cannot cross the vertical segments of $\gamma_C$ joining it to the boundary of $S_h$. This justifies our description of $\gamma^L_C$ as the leftmost crossing sensor-path. Since there are good separating paths to the left of $C$ that go through any given point of $U_0$, the region $S$ (and hence $\gamma^L_C$) lies entirely to the left of the vertical line $x = x_C^+ + \frac{r}{2}$.

![Figure 5: Leftmost crossing sensor-path $\gamma^L_C$ of good component $C$. Region to the right is “unexplored”. Dotted line is a separating path to the left of $C$.](image)

We now aim to bound the width of the good component conditioned on a particular value of $S$. It is important in what follows that we have not conditioned on anything to the right of $S$. Tile the strip $S_h$ with $a \times b$ rectangles as in the proof of Lemma 9. For definiteness, align the rectangles horizontally so that the line $x = x_C^- + \frac{r}{2}$ lies on a vertical boundary between two rectangles. Recall that $S$ lies entirely to the left of this line.

We now bound the probability that the width $W$ of $C$ is less than $w$ in two stages. First we bound the probability that $\frac{r}{2} \leq W < w$. Consider a good separating path $\gamma$ to the right of $C$ that meets the line $x = x_C^+$ and mark all the rectangles that intersect $\gamma$. Then no marked rectangle contains a sensor. If $W \geq \frac{r}{2}$ then $x_C^+ \geq x_C^- + \frac{r}{2}$, so $\gamma$ must stay to the right of the line $x = x_C^- + \frac{r}{2}$, and thus none of the marked rectangles intersect the region $S$. Also, at least one of the marked rectangles lies in the column $L$ of rectangles.
intersecting the line \( x = x^+_C \).

As in the proof of Lemma 9, we take the connected subgraph \( G \) of marked rectangles, so that \( G \) joins the top to the bottom of \( S_h \), consists entirely of empty rectangles, and also intersects \( L \). We bound the probability that any such graph \( G \) exists meeting \( L \) (ignoring the requirement that \( L \) is in fact the leftmost column intersecting \( G \)). Let \( R \) be the leftmost rectangle of \( G \) that meets the bottom boundary of \( S_h \). Suppose \( R \) is horizontally \( k \) rectangles to the right of \( L \) (or \(-k \) rectangles to the left of \( L \)). Then there are at least \( h/b + |k| \) rectangles in \( G \). Moreover, \( R \) is uniquely determined by \( k \) and \( L \). All these rectangles are empty, and are in regions not yet examined when constructing \( \gamma^L_C \). Thus, by Lemma 10, we can bound the probability that such a \( \gamma \) exists with a given choice of \( R \) by

\[
\sum_{n \geq h/b + |k|} \frac{1}{2} \mu^n e^{-nab} = \frac{1}{2} (\mu e^{-ab})^{h/b+|k|}/(1 - \mu e^{-ab}),
\]

where \( \mu = 2 + 2\sqrt{2} \). Summing over \( k \in \mathbb{Z} \), we have a bound of

\[
p_0 = \frac{1}{2} (\mu e^{-ab})^{h/b}/(1 - \mu e^{-ab})^3
\]

on the probability that any \( \gamma \) exists intersecting \( L \). (Here we have used \( \sum x^{[k]} = 1 + 2x + 2x^2 + \cdots \leq 1/(1 - x)^2 \). As in the proof of Lemma 9, we can bound \((\mu e^{-ab})^{h/b}\) by \( e^{-\kappa h} \), where \( \kappa \geq 0.72 \) by (4). But \( \kappa a \geq \frac{a}{\sqrt{12}} \geq 1 \) and \( 1 - \mu e^{-ab} \geq 0.99 \). Thus

\[
p_0 \leq e^{-h}.
\]

As \( L \) contains the line \( x = x^+_C \) and \( x^+_C + \frac{a}{2} \leq x^+_C < x^-_C + w \), there are at most \( \lceil (w - \frac{a}{2})/a \rceil \) choices for the column \( L \) and so

\[
P(\frac{a}{2} \leq W < w) \leq \lfloor (w - \frac{a}{2})/a \rfloor p_0 \leq we^{-h/3},
\]

where we have used the facts (from (1)) that \( \frac{a}{2} \geq a \) and \( a \geq \frac{a}{\sqrt{12}} \geq 1 \).

Now we consider the probability that \( W < \frac{a}{2} \). Unfortunately in this case it is possible for some of the rectangles of \( G \) to intersect the set \( S \), which we have already conditioned on. Indeed, if we are not careful, almost all of \( G \) may be covered by \( S \). Dealing with this case is therefore significantly more complicated.

If \( W < \frac{a}{2} \), the separating path \( \gamma \) must pass to the left of \( x = x^-_C + \frac{a}{2} \) (as it meets \( x = x^+_C \) and \( x^+_C - x^-_C = W < \frac{a}{2} \)) and also to the right of \( x = x^-_C \) (as it is to the right of \( \gamma^L_C \)). Inductively add to \( G \) any rectangles adjacent to \( G \) and (topologically) to the right of \( G \) that contain no point of \( P \), making \( G \) a maximal connected set of empty rectangles each of which either intersects \( \gamma \) or is to the right of \( \gamma \). If one can then remove rectangles from \( G \) that intersect \( S \) and still connect the top and bottom of \( S_h \), then do this. We
may have removed all rectangles intersecting \( x = x_C^+ \) in this process, but at least one of the remaining rectangles must be adjacent to a rectangle that is to the left of \( x = x_C^- + \frac{r}{2} \), either because it is adjacent to a rectangle meeting \( S \) that was removed, or because no rectangles were removed and \( \gamma \) passes to the left of \( x = x_C^- + \frac{r}{2} \). In particular, the remaining set of rectangles meets one of three columns \( L \), the one meeting \( x = x_C^- \), \( x = x_C^- + a \), or \( x = x_C^- + 2a > x_C^- + \frac{r}{2} \). Thus we can bound the probability that such a \( G \) exists by \( 3p_0 \) where \( p_0 \) is as above.

If every crossing path in \( G \) meets \( S \), then by removing rectangles if necessary, we may assume \( G \) is a path from the bottom to the top of \( S_h \). By taking \( G \) to be the shortest such path, we may assume that only one of our rectangles is adjacent to \( \partial S^- \), and only one of our rectangles is adjacent to \( \partial S^+ \). By assumption, \( G \) still meets \( S \).

If a large portion of \( G \) is covered by \( S \) then we will have difficulty proving a sufficiently strong bound on the probability of existence of \( \gamma \). Thus we need to add some extra area that we know to be devoid of sensors. Let \( B \) be the union of the rectangles of \( G \) and let \( T \) be the region topologically to the right of \( G \) (and hence to the right of \( \gamma \)) that lies within distance \( r \) of some vertex of \( \gamma \). Then we know that \( T \) contains no sensors, as any such sensor would be to the right of \( \gamma \) but joined to the component \( C \). Thus \( B \cup T \) contains no sensors and \((B \cup T) \setminus S \) has not been examined. There are at most \( 3^n - 3 \) possible paths \( G \) of length \( n \) starting at a fixed rectangle \( R \) as there are at most 3 choices at each of \( n - 1 \) steps, but the first and last steps go in a known direction. Thus we can bound the probability that \( C \) is of width \( \leq \frac{r}{2} \) by

\[
p_1 = 3p_0 + \sum_{n \geq h/b} (2 + 2(n - h/b))3^{n-3} \sup_{|G| = n} e^{-|B \cup T \setminus S|}.
\]

Here we have included \( 3p_0 \) to take into account the case when we could have removed rectangles to obtain a configuration that avoids \( S \), together with \( 2 + 2(n - h/b) \) times the probability of a path from a fixed \( R \) crossing \( S_h \). The factor of \( 2 + 2(n - h/b) \) counts the possible choices of \( R \), given that the path can make at most \( n - h/b \) horizontal steps, and must intersect one of the two columns intersecting \( x = x_C^- \) or \( x = x_C^- + a \) if it is to meet \( S \) (which is to the left of \( x = x_C^- + \frac{r}{2} < x_C^- + 2a \)) and also pass to the right of \( \gamma \) (and hence to the right of \( x = x_C^- \)).

Using (5), the bound (7) proved below in Lemma 12, and summing the series in \( n \), we have

\[
p_1 \leq 3p_0 + e^{-(rh/(4\sqrt{3}))}(h/b-1) \sum_{n \geq h/b} 2(1 + (n - h/b))3^{n-3} e^{-(\pi r^2/96)(n-h/b)}
\]

\[
\leq 3e^{-h} + 2 \frac{3^{h/b-3} e^{-(rh/(4\sqrt{3}))}(h/b-1)}}{(1 - 3e^{-\pi r^2/96})^2}.
\]
Now $1 - 3e^{-\pi r^2/96} > 1/3$ as $r \geq 7$, and $r \geq 7 > 4\sqrt{3}$. Thus
\[ p_1 \leq 3e^{-h} + 2.3^{h/5} e^{-b(h/b-1)} = 3e^{-h} + 2e^{-(2b/3 - \log(3))(h/b-1)+b/3}e^{-h/3}. \]

Now $b \geq r/\sqrt{12} \geq 2$, so $2b/3 - \log(3) \geq 4/3 - \log(3) > 0$. Thus we can replace $h/b$ by its smallest value $h/b = 3$ to get
\[ p_1 \leq 3e^{-h} + 2e^{-(2b/3 - \log(3))2 + b/3} \leq 3e^{-h} + 18e^{-b}e^{-h/3}. \]

But $b \geq 2$ and $h \geq \sqrt{a}\sqrt{3}/2 \geq 5$, so
\[ \mathbb{P}(W < \frac{r}{2}) \leq p_1 \leq (3e^{-10/3} + 18e^{-2})e^{-h/3} \leq 3e^{-h/3}. \]

The result then follows on adding (6).

**Lemma 12.** Let $G$, $B$, and $T$ be as in the proof of Lemma 11. Then
\[ |(B \cup T) \setminus S| \geq \frac{\pi r^2}{96} (n - \frac{h}{b}) + \frac{rb}{4\sqrt{3}} (\frac{h}{b} - 1), \]

where $n$ is the number of rectangles in $G$.

**Proof.** We first show that any point to the right of $B \cup T$ that is not in the bottom or top row of rectangles must be at least distance $c := r/\sqrt{12}$ from $S$. Note that $c \leq a$, $b$ by (1).

Any such point $P$ must be to the right of $\gamma$ and at distance at least $r$ from the sensors in $\gamma_C$. Any point on the right boundary of $S$ is either on an arc of a circle between two sensors of $\gamma_C$ or on an arc of a circle joining a sensor to the boundary of $S$ in $\gamma_C$. In both cases one can check using simple geometry that $P$ is at distance at least $c$ from $S$ (see Figure 6).

In the case when the closest point on $\partial S$ to $P$ is on an arc joining two vertices $v$ and $v_{1+i}$ of $\gamma_C$, the minimum distance of $P$ from $S$ is at least
\[ x := \frac{\sqrt{3}r}{2} - \frac{r}{2} \geq \frac{r}{\sqrt{12}} = c. \]

The worst case however is near the boundary when the closest point to $P$ on $\partial S$ is on an arc meeting $\partial S_h$. In this case $P$ could be within distance
\[ y := (r^2 - (\sqrt{3}r - b)^2 + (\frac{\sqrt{3}r}{2} - \frac{r}{2} - b)^2)^{1/2} - \frac{r}{2} \geq \frac{r}{2\sqrt{12}} = c \]

of $S$. If $P$ is in the top or bottom rows of rectangles it can be closer to $S$, however in all cases it must be at least a horizontal distance of
\[ z := \frac{r}{2} - \left( \frac{r^2}{4} - (\frac{\sqrt{3}r}{2} - \frac{r}{2})^2 \right)^{1/2} \geq \frac{r}{2\sqrt{12}} = c. \]
Figure 6: Left: Points $P$ at distance at least $r$ from any vertex of $\gamma_L^C$ are at horizontal distance $z \geq \frac{c}{2}$ from $S$, and if $P$ is at least $b$ from $\partial S_h$ then it is also at distance at least $\min\{x, y\} \geq c$ from $S$. Right: Horizontal line through $Z$ that approaches within distance $t$ of $Q$ intersects $(B \cup T) \setminus S$ in an interval of length at least $2(c^2 - t^2)^{1/2}$.

We now estimate the area $|(B \cup T) \setminus S|$ in two ways. First we estimate the area in a row of rectangles, and add up the contributions from the $h/b$ rows, then we estimate by columns. Averaging these two estimates will then give the result.

First fix a horizontal row of rectangles. We call a rectangle $R$ in this row right-facing if $R$ lies in $B$, the rectangle to the right of $R$ does not lie in $B$, and the rectangle to the right of $R$ is in the component (topologically) to the right of the path $G$ of rectangles. We call a rectangle $L$ in this row left-facing if $L$ lies in $B$, the rectangle to the left of $L$ does not lie in $B$, but the rectangle to the left of $L$ is in the component to the right of the $B$. (This can occur if the path of rectangles $B$ loops round and travels downwards across this row, see Figure 6.) Note that some rectangles of $B$ may be neither left nor right-facing, but no rectangle can be both as the path $G$ is of minimal length. Since the path begins at the bottom of $S_h$ and ends at the top, there will be one more right-facing than left-facing rectangle in this row. Consider the rightmost rectangle $R$ of $B$ in this row. Clearly it is right-facing. The right boundary of the excluded area $T$ is at horizontal distance at least $c$ from $S$ ($c/2$ if the row is adjacent to $\partial S_h$). Thus any horizontal line $y = y_0$ intersects $(B \cup T) \setminus S$ in an interval of length at least $\min\{a, c\} = c$ (or $\min\{a, c/2\} = c/2$) to the right of the left boundary of $R$. Integrating over $y$ gives a contribution to the area of $bc$ (or $bc/2$) for this rectangle. The remaining left and right-facing rectangles can be paired up, with a right-facing rectangle $R$ followed by some rectangles not in $B$, followed by a
left-facing rectangle \( L \). Note that rows adjacent to the boundary of \( S_h \) contain only a single right-facing rectangle, so we may assume that the row is not adjacent to \( \partial S_h \). Let \( Z \) be the rectangular region consisting of \( R, L \), and the rectangles between \( R \) and \( L \) on this row (see Figure 6). As the path \( B \) is minimal, not all of \( Z \) lies in \( B \cup T \). (Otherwise adding \( Z \) to \( B \) and cutting off the “loop” between \( R \) and \( L \) would shorten the path of rectangles.) Let \( Q \) be the region in \( Z \) that does not lie in \( B \cup T \). (Otherwise adding \( Z \) to \( B \) and cutting off the “loop” between \( R \) and \( L \) would shorten the path of rectangles.) Let \( Q \) be the region in \( Z \) that does not lie in \( B \cup T \). Then \( Q \) is at distance at least \( c \) from \( S \). Consider the intersection \( I \) of a horizontal line \( y = y_0 \) with \( (B \cup T) \setminus S \cap Z \). If the line \( y = y_0 \) meets \( Q \), then there is an interval of length at least \( \min\{a, c\} = c \) in \( I \) to the left of \( Q \), and an interval of length at least \( \min\{a, c\} = c \) in \( I \) to the right of \( Q \) that does not meet \( S \). If the line \( y = y_0 \) passes within distance \( t \) of \( Q \), then there is an interval of length \( 2(c^2 - t^2)^{1/2} \) in \( I \) all of whose points are within distance \( c \) of \( Q \) and so does not meet \( S \). Integrating over \( y \) thus gives an area in \( Z \) in \( (B \cup T) \setminus S \) at least \( \pi c^2 / 2 \), i.e., the minimum area of intersection of the disk of radius \( c \) about a point in \( Z \setminus (R \cup L) \). Thus, summing over rows, we have

\[
| (B \cup T) \setminus S | \geq \left( \frac{h}{b} - 2 \right) bc + 2 \frac{bc}{2} + \frac{1}{2} (n_h - \frac{h}{b}) \frac{\pi c^2}{2},
\]

where \( n_h \) is the total number of left- or right-facing rectangles. A similar calculation can be performed for columns of rectangles, giving

\[
| (B \cup T) \setminus S | \geq \frac{1}{2} n_v \frac{\pi c^2}{2},
\]

where \( n_v \) is the total number of up- or down-facing rectangles. (The contribution from “outward”-facing rectangles is \( bc \geq \frac{\pi c^2}{2} \) and rectangles adjacent to \( \partial S_h \) are never outward-facing.) As we travel along the path \( G \), \( n_v + n_h \) counts the number of segments of the right boundary of \( B \). Thus each counterclockwise turn adds 2 to \( n_v + n_h \), and each clockwise turn adds 0 to \( n_v + n_h \). Each straight edge adds 1, so since there are as many clockwise as counterclockwise turns we have \( n_v + n_h = n \), the total number of rectangles. Thus by averaging (8) and (9) we get

\[
| (B \cup T) \setminus S | \geq \frac{\pi c^2}{8} (n - \frac{h}{b}) + \frac{bc}{2} \left( \frac{h}{b} - 1 \right).
\]

The result follows as \( c = r / \sqrt{12} \). \[\square\]

5 The distribution of breaks

To show that breaks have an approximately Poisson distribution, we use the Stein-Chen method, a particularly convenient form of which can be found in [1]. The following theorem is immediate consequence of Theorem 1 of [1].
Theorem 13. Let \( \{X_I : I \in \mathcal{I} \} \) be a collection of Bernoulli random variables indexed by a countable collection of subsets \( I \) of some ground set. Suppose that \( \sum_I \mathbb{E}(X_I) = \lambda \), and for each \( I \), \( X_I \) is independent of \( \{X_J : I \cap J = \emptyset \} \). Let 
\[
b_1 = \sum_{I,J : I \cap J \neq \emptyset, I \neq J} \mathbb{E}(X_I) \mathbb{E}(X_J), \quad b_2 = \sum_{I,J : I \cap J \neq \emptyset, I \neq J} \mathbb{E}(X_I X_J), \quad \text{and write} \quad X = \sum_I X_I. \]
Then for all \( k \),
\[
| \mathbb{P}(X = k) - e^{-\lambda} \lambda^k / k! | \leq b_1 + b_2. 
\]

\( \square \)

Theorem 1 in [1] includes another term \( b_3 \) that bounds dependency when \( I \cap J = \emptyset \), but in our case \( b_3 = 0 \).

Proof of Theorem 4. Let \( W_g \) be a random variable giving the width of a good component, and \( W_b \) be a random variable giving the width of the subsequent break. Then \( W_b \geq 0 \), \( W_g + W_b \geq 0 \), and \( \mathbb{E}(W_g + W_b) I_{h,r} = 1 \). By Lemma 11, \( \mathbb{E}(W_g + W_b) \) grows exponentially with \( h \), while by Lemma 9, \( \mathbb{E}(W_b) \) grows at most linearly with \( h \). Thus \( \mathbb{E}(W_b)/\mathbb{E}(W_g + W_b) \to 0 \), and so \( \mathbb{E}(W_b) I_{h,r} \to 0 \) as \( h \to \infty \). Set \( \ell = x/I_{h,r} = x \mathbb{E}(W_g + W_b) \), so that \( \ell \to \infty \) as \( h \to \infty \).

Let \( I \subseteq [0, \ell] \) be a half-open interval with integer endpoints, say \( I = [p, q) \), with \( q - p > r \), and let \( I' = [p + \frac{r}{2}, q - \frac{r}{2}] \). Define \( B_I \) to be the event that there are two crossing sensor-paths, \( \gamma_1 \) and \( \gamma_2 \), that lie in \( I' \times [0, h] \), and a separating path lies in \( I' \times [0, h] \) between \( \gamma_1 \) and \( \gamma_2 \) (so \( \gamma_1 \) and \( \gamma_2 \) go through distinct good components). Clearly the event \( B_I \) depends only on the Poisson process in \( I \times [0, h] \). Now let \( X_I \) be the indicator function of the event that \( B_I \) holds, but that \( B_J \) does not hold for any proper integer subinterval \( J \) of \( I \). Clearly \( X_I \) also depends only on the Poisson process in \( I \times [0, h] \) and if \( X_I = 1 \) then \( B_I \) holds, so there is a break. Moreover this break has break interval \( [x_C^+, x_C^-] \subseteq I' \subseteq I \) as no good separating path corresponding to this break can cross the crossing sensor-paths \( \gamma_1 \) and \( \gamma_2 \) in \( I' \times [0, h] \).

From the construction of the leftmost crossing sensor-path in Lemma 11, there is always a crossing sensor-path through a good component \( C \) that lies entirely to the left of \( x_C^+ + \frac{r}{2} \). Similarly, there is a crossing sensor-path entirely to the right of \( x_C^- - \frac{r}{2} \). Thus, every break interval \( [x_C^+, x_C^-] \) that is a subset of \( [r, \ell - r] \) results in an interval \( I \) with \( X_I = 1 \). Indeed, we have
\[
I \subseteq ([x_C^+ - r], [x_C^- + r]),
\]
so the width of \( I \) is at most \( 2r + 1 \) more than the width of the break it contains.

If there are two breaks resulting in the event \( B_I \) holding, then there must be distinct proper integer subintervals \( J_1, J_2 \subset I \) such that \( B_{J_1} \) and \( B_{J_2} \) hold. This is because there would be three crossing sensor-paths, \( \gamma_1, \gamma_2, \gamma_3 \), in \( I' \times [0, h] \) corresponding to three distinct good components. But these would have to be separated by a distance of at least \( \frac{r}{2} > 3 \) from each other. Hence \( B_J \) would hold with \( J = [p, q - 3) \) and with \( J = [p + 3, q) \). Thus
each \( I \) with \( X_I = 1 \) corresponds to a single break. Conversely, each break can give rise to only one minimal interval \( I \) with \( X_I = 1 \). Hence if no break interval intersects either \([0, r]\) or \([\ell - r, \ell]\), then \( X = \sum X_I \) exactly counts the number of breaks in \( G_{h,r}(0, x/I_{h,r}) \), and is also the number of coordinates \( x_{C}^{+} \) that lie in \([0, \ell]\). The probability that a break intersects either \([0, r]\) or \([\ell - r, \ell]\) is at most \( 2(\mathbb{E}(W_b) + r)I_{h,r} \), which is at most \((10h + 2r)I_{h,r}\) by Lemma 9 when \( h \geq 1 \). But \( I_{h,r} \) tends 0 exponentially with \( h \). Thus with high probability \( X \) is the number of breaks in \( G_{h,r}(0, x/I_{h,r}) \), and so \( \mathbb{E}(X) \to x \) as \( h \to \infty \).

From the definition of \( X_I \) it is clear that \( X_I \) is independent all \( X_J \) with \( I \cap J = \emptyset \). If \( I \cap J \neq \emptyset, I \neq J \), and \( X_I X_J = 1 \), then by (10) we have \([x_{C}^{+} - r] < [x_{C}^{−} + r]\) for any good component \( C \) between the breaks corresponding to \( I \) and \( J \), and hence all these good components must have width at most \( 2r + 1 \). In particular, one of the two good components adjacent to the break given by \( I \) must have width at most \( 2r + 1 \). Thus, by Lemma 11, the expected number of \( I \)'s such that \( X_I X_J = 1 \) for some \( J \neq I \), \( I \cap J \neq \emptyset \), is at most \( 2(2r + 4)e^{-h/3}2I_{h,r} = 2(2r + 4)e^{-h/3}x \). Let \( I = [p, q] \) and suppose there is such a \( J \). There must be a crossing sensor-path \( \gamma_C \) in \( I \times [0, h] \) intersecting \([q - \frac{x}{2} - 1, q] \times [0, h] \) or else \( B_{[p,q-1]} \) would hold. Each \( J = [p', q'] \) with \( X_J = 1 \) corresponding to a break on the right of \( I \) must contain a crossing sensor-path meeting \([p' + \frac{x}{2}, p' + 1 + \frac{x}{2}] \times [0, h] \), otherwise \( B_{[p',q'+1]} \) would hold. As distinct \( J \)'s give rise to distinct breaks, all but at most one of these crossing sensor-paths must lie to the right of \( \gamma_C \). Thus if there are \( k \) intervals \( J \) intersecting \( I \) and belonging to a break on the right of \( I \), then there are \( k - 1 \) components with crossing sensor-paths intersecting \([q - \frac{x}{2} - 1, \frac{p'}{2} + \frac{1}{2}] \times [0, h] \subseteq [q - \frac{x}{2} - 1, q + \frac{x}{2}] \times [0, h] \). However, we cannot have more than \( O(h/r) \) crossing sensor-paths corresponding to distinct components intersecting this region as each is at distance at least \( \frac{x}{2} \) from the others. Thus for each \( I \) with \( X_I = 1 \) there can be at most \( O(h/r + 1) \) \( J \)'s such that \( I \cap J \neq \emptyset \) and \( X_I X_J = 1 \). Thus \( b_2 = O(h/r + 1)(2r + 4)e^{-h/3}x \), and so \( b_2 \to 0 \) as \( h \to \infty \).

Fixing \( I \), \( \sum_{J \cap I \neq \emptyset} \mathbb{E}(X_J) \leq \sum_{k} (|I| + k - 1)\mathbb{P}(X_{[0,k]} = 1) \) as there are \(|I| + k - 1 \) translates of an interval of length \( k \) intersecting \( I \) and \( \mathbb{P}(X_{I} = 1) \) depends only on \(|I| \). But if \( X_{[0,k]} = 1 \) then \( k \leq W_b + 2r + 1 \) where \( W_b \) is the width of the break corresponding to this interval. Thus \( \sum_{J \cap I \neq \emptyset} \mathbb{E}(X_J) \leq (|I| + \mathbb{E}(W_b) + 2r)I_{h,r} \). Now \(|I| \leq W'_b + 2r + 1 \), where \( W'_b \) is the width of the break corresponding to \( I \). Hence

\[
b_1 \leq (\mathbb{E}(W'_b) + \mathbb{E}(W_b) + 4r + 1)I_{h,r}^2 \leq (10h + 4r + 1)xI_{h,r},
\]

where the second inequality follows from Lemma 9 and the fact that we may assume \( h \geq 1 \). Since \( I_{h,r} \) decreases exponentially with \( h \), \( b_1 \to 0 \) as \( h \to \infty \). The result follows. \( \square \)

For practical purposes one would aim to show that the existence of any break occurring in a short strip is unlikely. In this case one can use the following simpler result.
Theorem 14. If \( r \geq 7 \) and \( h \geq 1 \) then
\[
P(G_{h,r}(0,\ell) \text{ has no } s-t \text{ path}) \leq (\ell + 5h)I_{h,r}.
\]

Proof. If \( G_{h,r}(0,\ell) \) has no \( s-t \) path then, by Lemma 1 and Lemma 8, a good separating path exists in \([0,\ell] \times [0,h]\). Thus \([0,\ell]\) must intersect one of the intervals \([x_{C_i}^+, x_{C_{i+1}}^-]\) corresponding to the breaks in \( G_{h,r} \). Equivalently, either \( x_{C_i}^+ \in [0,\ell] \), or \( 0 \in [x_{C_i}^+, x_{C_{i+1}}^-]\) for some \( i \), where \( C_i, i \in \mathbb{Z} \), are the good components ordered horizontally. The expected number of \( i \)'s with the first property is \( \ell I_{h,r} \) since the asymptotic density of points \( x_{C_i}^+ \) equals the density of breaks \( I_{h,r} \). The expected number of \( i \)'s with the second property is at most \( 5h I_{h,r} \) since, by Lemma 9, the expected width of a break is at most \( 5h \) (for \( h \geq 1 \)) and their density is \( I_{h,r} \). Thus the probability that \( G_{h,r}(0,\ell) \) is not \( s-t \) connected is at most \( (\ell + 5h)I_{h,r} \).

\[\square\]

6 Small \( h \) and the function \( \varepsilon(z) \)

Having shown in Theorem 4 that breaks occur with an approximately Poisson distribution, it remains to derive the break intensity. Once we have the break intensity, it is a simple matter to either estimate or bound the probability that no breaks exist in a strip of given length (which by Lemma 1 is equivalent to the strip being barrier covered) using either Theorem 4 or Theorem 14.

The most interesting case is when \( h \) is larger than \( r \), however, to estimate \( I_{h,r} \) for large \( h \), we shall reduce to the case of small \( h \). Thus we shall need to study the small \( h \) case in some detail first. We count the number of vertices that are the rightmost point of some good component, since this is equivalent to counting good components, and hence breaks.

For \( h < \frac{\sqrt{3}}{2}r \), all components are good, so the probability that a fixed sensor \( v \) is the rightmost sensor of a good component is given by the probability that there is no sensor \( w \) to the right of \( v \) that is adjacent in \( G_{h,r} \) to \( v \), or to any sensor \( u \) to the left of \( v \) (as in this case either \( u \) or \( w \) would be joined to \( v \) by Lemma 5 applied to the vertical crossing sensor-path through \( v \)). To calculate this probability, fix \( v \), and place sensors to the left of \( v \) in \( S_h \) according to a Poisson point process. Let \( A \) be the region in \( S_h \) to the right of \( v \) that is within distance \( r \) of \( v \) or any sensor to the left of \( v \) (see Figure 7). Then \( v \) is the rightmost sensor of a component if and only if the region \( A \) is empty. Thus conditioned on the process to the left of \( v \), the probability that \( v \) is the rightmost sensor of a component is \( e^{-|A|} \). The overall probability that we are interested in is just \( E(e^{-|A|}) \), where the expectation is over the position of \( v \) and the state of the Poisson process to the left of \( v \). The intensity of such sensors, and hence of breaks, is then given by

\[
I_{h,r} = h E(e^{-|A|}),
\]

(11)
since the intensity of sensors $v$ per unit length of the strip is just $h$. It remains to calculate $\mathbb{E}(e^{-|A|})$.

The excluded area is the union of a number of disks of radius $r$. We first approximate these areas by parabolic regions, replacing the disk $(x - x_0)^2 + (y - y_0)^2 < r^2$ about a sensor $(x_0, y_0)$ by the parabolic region $x - x_0 < r - (y - y_0)^2/2r$. We then estimate the excluded area as $|A| \approx rh - |B|$, where $B$ is the shaded region in Figure 7 which lies to the right of all these approximating parabolas and to the left of the vertical line which is at distance $r$ to the right of $v$. The advantage of this approximation is that the two parameters $r$ and $h$ can be reduced to a single parameter $z = hr^{-1/3}$ by rescaling the strip by a factor $r^{1/3}$ along the $x$-axis and by a factor $r^{-1/3}$ along the $y$-axis. This rescaling leaves the density of the Poisson process and all areas unchanged, but puts the parabolas in the form $x = z_0 - (y - y_0)^2/2$ where the points $(z_0, y_0)$ are the locations of the sensors, shifted a constant amount to the right so that they lie on the vertices of the parabolas. For convenience we now swap the $x$ and $y$-coordinates. Then $B$ is defined by placing a Poisson point process with intensity 1 in the half infinite strip $[0, z] \times [0, \infty)$ plus one more point chosen uniformly at random on $[0, z] \times \{0\}$ and then taking the area below all the parabolas $y = y_0 + (x - x_0)^2/2$, where $(x_0, y_0)$ ranges over all of these points. Define $\varepsilon(z)$ by

$$
\varepsilon(z) = z \mathbb{E}(e^{1/|B|}), \tag{12}
$$

so that

$$
I_{h,r} \approx r^{1/3} \varepsilon(zr^{-1/3}) e^{-hr}.
$$

Note that taking a Poisson point process in $[0, z] \times [0, \infty)$ and then adding one random point in $[0, z] \times \{0\}$ is equivalent to taking a Poisson process in $[0, z] \times [0, \infty)$ and then shifting all these points down until the lowest point is on the $x$-axis. Let $\mathcal{P}$ be defined as the set of these points.

It is clear that $|B| \leq z^3/6$, the worst case being when there is just one point in $\mathcal{P}$ that is located at the origin. Thus $\varepsilon(z) \leq z e^{27/6}$. We wish to improve this bound to $e^{O(z)}$ first, before showing that $\varepsilon(z) = e^{ax \beta \gamma + o(1)}$. Moreover, we wish to show that in (12), one can, without too much error, restrict $\mathcal{P}$ so that the lowest parabola above any point $(x_0, 0)$ has vertex not too far from the line $x = x_0$. Note that the horizontal distance to the vertex of one the parabolas is just the slope of the parabola at that point, so this is equivalent to restricting the maximum slope of the upper boundary of the region $B$. In doing so, we shall also bound the error introduced by approximating the excluded disks by parabolas.

We now leave the task of estimating $I_{h,r}$ to Section 8. The rest of this section will be devoted to studying the function $\varepsilon(z)$.

**Lemma 15.** Let $s(B)$ be the maximum absolute value of the slope of the upper boundary of $B$. Define $\varepsilon_s(z) = z \mathbb{E}(e^{1/|B|} 1_{s(B) \geq s})$. Then there are absolute constants $c_1, c_2 > 0$
Figure 7: Approximating excluded disks by parabolas. The vertex $v$ is the rightmost vertex of the good component, and $u$ is another vertex to the left of $v$. The excluded area $A$ consists of all points within the circles that lie to the right of $v$, so $|A| \approx hr - |B|$. For clarity, the picture is drawn with $h > r$. The diagram on the right is used in the proof of Lemma 17.

independent of $z$ and $s$ such that $\varepsilon_s(z) \leq e^{c_1 z - c_2 s^3}$.

Note that $\varepsilon(z) = \varepsilon_0(z)$, so Lemma 15 implies that $\varepsilon(z) \leq e^{c_1 z}$.

Proof. First we note that $\varepsilon_s(z) \leq \varepsilon(z) \leq ze^{z^3/6}$ and $s(B) \leq z$, so by choosing $c_1$ sufficiently large, and $c_2$ sufficiently small, we may assume the result holds for $z \leq 1$.

Now assume $z > 1$ and tile $[0, z] \times [0, \infty]$ with rectangles of width $\alpha$ and height $\alpha^2/2$. Here we choose $\alpha = z/\lceil z \rceil$ so that $n = z/\alpha = \lceil z \rceil$ is an integer and $\frac{1}{2} \leq \alpha \leq 1$. Let $a_i$, $i = 0, \ldots, n - 1$, be the number of rectangles above the interval $[i\alpha, (i + 1)\alpha]$ that intersect $B$. In other words, we bound $B$ by a step function which has height $a_i \alpha^2/2$ on the interval $[i\alpha, (i + 1)\alpha]$. Thus $|B| \leq \frac{\alpha^3}{2} \sum_i a_i$. Let $b_i$, $i = 0, \ldots, n - 1$, be defined as the maximum of

$$a_i - 2, \quad a_i + 1 - 5, \quad a_i + 2 - 10, \quad \ldots \quad a_i + r - 1 - (r + 1)^2, \quad \ldots$$

Then there can be no point of $\mathcal{P}$ in the lowest $b_i$ rectangles above $[i\alpha, (i + 1)\alpha]$. Indeed, if say $b_i = a_i + r - 1 - (r + 1)^2$ and there were a point $(x_0, y_0) \in \mathcal{P}$ with $y_0 \leq b_i \frac{\alpha^2}{2}$ and $x_0 \in [i\alpha, (i + 1)\alpha]$, then any point $(x, y) \in B$ with $x \in [(i + r)\alpha, (i + 1 + r)\alpha]$ would satisfy

$$y \leq y_0 + \frac{1}{2}(x - x_0)^2 \leq b_i \alpha^2 + (r + 1)^2 \frac{\alpha^2}{2} \leq (a_i + r - 1) \frac{\alpha^2}{2},$$

contradicting the definition of $a_i + r$. The probability of a configuration occurring with a particular sequence $(a_i)_i$ is thus at most $\exp(-\frac{\alpha^3}{2} \sum_i b_i)$. Since $|B| \leq \frac{\alpha^3}{2} \sum_i a_i$, the contribution to $\varepsilon(z)$ from such configurations is at most $z \exp(\frac{\alpha^3}{2} \sum_i (a_i - b_i))$. 25
Let $\delta_i = a_i - a_{i-1}$, $i = 1, \ldots, n - 1$. Then the sequence $(\delta_i)_i$ determines $(a_i)_i$ up to the addition of a constant. However, $a_i = 1$ when the point of $P$ on the $x$-axis lies in $[i\alpha, (i+1)\alpha]$. Thus the contribution to $\varepsilon(z)$ from configurations with a fixed sequence $(\delta_i)_i$ is at most

$$nz \exp \left( \frac{\alpha^3}{2} \sum_i (a_i - b_i) \right)$$

and the differences $a_i - b_i$ depend only the sequence $(\delta_i)_i$. Now $b_i \geq a_i - 2, a_{i+1} - 5$, so $a_i - b_i \leq \min\{5, 2, 5 - \delta_i, 5 - \delta_{i+1}\}$, where we define $\delta_0 = \delta_n = 0$ when $i = 0$ or $n - 1$. Thus

$$a_i - b_i \leq \frac{1}{2}(5 - \max\{-\delta_i, 0\}) + \frac{1}{2}(5 - \max\{\delta_i, 0\}) = 5 - \frac{1}{2}\delta_i^+ - \frac{1}{2}\delta_i^-,$$  \hspace{1cm} (13)

where $\delta_i^+ = \max\{\delta_i, 0\}$ and $\delta_i^- = \max\{-\delta_i, 0\}$. Now $|\delta_i| = \delta_i^+ + \delta_i^-$, so

$$\sum_{i=0}^{n-1} (a_i - b_i) \leq 5n - \frac{1}{2}(\delta_0 + |\delta_1| + \cdots + |\delta_{n-1}| + \delta_n^+) = 5n - \frac{1}{2}\sum_{i=1}^{n-1} |\delta_i|.$$  \hspace{1cm} (14)

Thus

$$nz \exp \left( \frac{\alpha^3}{2} \sum_{i=0}^{n-1} (a_i - b_i) \right) \leq nz e^{5na^3/2} \prod_{i=1}^{n-1} e^{-|\delta_i|a^3/4}.$$  \hspace{1cm} (15)

Summing over all possible values of $\delta_1, \delta_2, \ldots, \delta_{n-1} \in \mathbb{Z}$ in turn gives

$$\sum_{(\delta_i)} nz \exp \left( \frac{\alpha^3}{2} \sum_i (a_i - b_i) \right) \leq nz e^{5na^3/2} \left( \frac{2}{1 - e^{-a^3/4}} - 1 \right)^{n-1} \leq nz e^{c_3n},$$

for some absolute constant $c_3 > 0$. Substituting $n = z/\alpha$ and recalling that $\alpha \geq \frac{1}{2}$ gives $\varepsilon(z) \leq 2z^2e^{2c_3z}$. Thus the result for $s = 0$ (or even bounded $s$) follows.

Now suppose the upper boundary of $B$ has slope $s = s(B) > 0$ at the point $(x_0 + s, y_0 + \frac{s^2}{2})$, so that the slope is determined by the vertex $(x_0, y_0) \in P$. Let $i, j \in \mathbb{Z}$ be such that

$$x_0 \in [i\alpha, (i+1)\alpha] \quad \text{and} \quad y_0 \in [(j - 1)\alpha^2, j\alpha^2].$$

Then $a_{i+r} \leq (r+1)^2 + j$ for all $r \geq 0$. Now choose $t \in \mathbb{Z}$ so that $x_0 + s \in [(i+t)\alpha, (i+t+1)\alpha]$. Note that $(t-1)\alpha \leq s \leq (t+1)\alpha$. Then $a_{i+t} \alpha^2 \geq y_0 + \frac{s^2}{2} \geq (j - 1 + (t - 1)^2)\frac{\alpha^2}{2}$. Hence for $0 \leq r \leq t$

$$b_{i+r} \geq (j - 1 + (t - 1)^2) - 1 - (t - r + 1)^2.$$  \hspace{1cm} (15)

Thus

$$a_{i+r} - b_{i+r} \leq (r + 1)^2 + j - (j - 1) - (t - 1)^2 + 1 + (t - r + 1)^2 = 4t + 3 - 2r(t - r).$$
Since the maximum slope of the upper boundary of $B$ is $s \leq (t + 1)\alpha$, we have $|\delta_k| \leq t + 1$ for all $k$, and so

$$a_{i+r} - b_{i+r} \leq (5 - \frac{1}{2}\delta_{i+r} - \frac{1}{2}\delta_{i+r+1}) + 5t - 2r(t - r).$$

Summing, and using (13) when $j < i$ or $j > i + t$, gives

$$\sum_{j=0}^{n-1} (a_j - b_j) \leq 5n - \frac{1}{2}\sum_{j=1}^{n-1} |\delta_j| - \sum_{r=0}^{t} (2r(t - r) - 5t).$$

This last sum is $\Theta(t^2) = \Theta(s^3)$, so summing over the $(\delta_i)_i$ as in (15) now gives the result for $\varepsilon_s(z)$, so summing over the $(\delta_i)_i$ as in (15) now gives the result for $\varepsilon_s(z)$. \qed

We observe that we only used the parabolic bound $y_0 + (x - x_0)^2/2$ on the upper boundary of $B$ when $|x - x_0| < 2\alpha \leq 2$ for $s = 0$, or $|x - x_0| < s + 2\alpha \leq s + 2$ for $s > 0$. Clearly modifying the parabola slightly does not affect the proof of Lemma 15, so we can generalize it to the following.

**Lemma 16.** Assume $f : \mathbb{R} \times \mathbb{R} \to [0, \infty]$ is such that $f(x_0, x_0) = 0$ and

$$\frac{\partial}{\partial x} f(x, x_0) = \eta(x, x_0)(x - x_0), \quad 1 \leq \eta(x, x_0) \leq K,$$

for all $|x - x_0| < s + 2$. Then the conclusion of Lemma 15 also holds for $\varepsilon_{f,s}(z)$, where we define $\varepsilon_{f,s}(z)$ as for $\varepsilon_s(z)$, but with the curves $y = y_0 + f(x, x_0)$ used instead of the parabolas $y = y_0 + (x - x_0)^2/2$ in the definition of the region $B$, and constants $c_1$ and $c_2$ depending on $K$.

**Proof.** By the monotonicity of the expression for $\varepsilon_{f,s}(z)$ in terms of $e^{\|B\|_1(B)}s$, it is enough to prove the result when $f(x, x_0) = K(x - x_0)^2/2$ and the slope bound is replaced with $s(B) \geq Ks$. In this case the result follows from the proof of Lemma 15 by using $K\frac{r^2}{2}$ for the height of the rectangles instead of $\frac{r^2}{2}$. \qed

To show that $\varepsilon(z)$ is of the form $e^{\alpha z + \beta + o(1)}$ we now show that $\varepsilon(z_1 + z_2) \approx K\varepsilon(z_1)\varepsilon(z_2)$ for some constant $K$. Let $z = z_1 + z_2$ and decompose the area $B$ into two pieces, the part above $[0, z_1]$ and the part above $[z_1, z_1 + z_2]$. If $y_1$ (respectively $y_2$) is the height of the lowest vertex on the left (respectively right), then the set of vertices can be written as a union of $\mathcal{P}_1 + (0, y_1)$ and $\mathcal{P}_2 + (z_1, y_2)$, where $\mathcal{P}_1$ and $\mathcal{P}_2$ are corresponding configurations used in the definition of $\varepsilon(z_1)$ and $\varepsilon(z_2)$. The area $|B|$ can be calculated as $|B_1| + |B_2| + |R| - |C|$, where $B_1$ and $B_2$ are the corresponding areas for $\mathcal{P}_1$ and $\mathcal{P}_2$, $R$ is either $[0, z_1] \times [0, y_1]$ or $[z_1, z_1 + z_2] \times [0, y_2]$ depending on whether $y_2$ or $y_1$ is zero, and $C$ is the set of points.
that are below all the parabolas on their own side, but above a parabola on the opposite side (see Figure 7). Fix \( \mathcal{P}_1 \) and \( \mathcal{P}_2 \) and imagine sliding one of the regions \([0, z_1] \times \mathbb{R} \) or \([z_1, z_1 + z_2] \times \mathbb{R} \) either up or down. Let \( t \) be the \( y \)-coordinate of the top left point of \( B_2 \) minus the \( y \)-coordinate of the top right point of \( B_1 \). We consider the region \( C = C(t) \) to be a function of \( t \). Note in particular that \( C(0) = \emptyset \) and that \( C \) is to the right (respectively left) of the line \( x = z_1 \) when \( t > 0 \) (respectively \( t < 0 \)), and \( C \) intersects the line \( x = z_1 \) in an interval of length exactly \( |t| \).

**Lemma 17.**

\[
\varepsilon(z_1 + z_2) = z_1 z_2 E_{\mathcal{P}_1} E_{\mathcal{P}_2} e^{\left| B_1 \right| e^{\left| B_2 \right|}} \int_{-\infty}^{\infty} e^{-\left| C(t) \right|} dt,
\]

where \( E_{\mathcal{P}_1} \) and \( E_{\mathcal{P}_2} \) are expectations over the corresponding configurations and, as above, \( t \) is the distance of the top left point of \( B_2 \) over the top right point of \( B_1 \).

**Proof.** The set \( \mathcal{P} \) can be constructed with the correct distribution by setting \( \mathcal{P} = (\mathcal{P}_1 + (0, y_1)) \cup (\mathcal{P}_2 + (z_1, y_2)) \), where with probability \( z_1/z, y_1 = 0 \) and \( y_2 \) has an exponential distribution with mean \( 1/z_2 \), while with probability \( z_2/z, y_2 = 0 \) and \( y_1 \) has an exponential distribution with mean \( 1/z_1 \). Let \( t_0 \) be the value of \( t \) when \( y_1 = y_2 = 0 \). Then recalling that \( |B| = |B_1| + |B_2| + |R| - |C| \),

\[
\varepsilon(z_1 + z_2) = z E_{\mathcal{P}_1} E_{\mathcal{P}_2} e^{\left| B_1 \right| e^{\left| B_2 \right|}} \left( \frac{z_1}{z} \int_0^\infty e^{z_{22} - |C(t)|} z_2 e^{-\tau z_2} d\tau + \frac{z_2}{z} \int_0^\infty e^{z_{21} - |C(t)|} z_1 e^{-\tau z_1} d\tau \right) = z_1 z_2 E_{\mathcal{P}_1} E_{\mathcal{P}_2} e^{\left| B_1 \right| e^{\left| B_2 \right|}} \int_{-\infty}^{\infty} e^{-\left| C(t) \right|} dt.
\]

The area \( |C(t)| \) can perhaps be more easily visualized by translating the points \((x, y)\) down vertically to \((x, y - (x - z_1)^2/2)\). In this way, the area \( |C(t)| \) remains constant, but the parabolas used to define the regions \( B, B_1, B_2 \), and \( C \) become straight lines. The upper boundary of \( B \), \( B_1 \), \( B_2 \) become polygonal paths that are concave down, and \( C \) becomes a polygonal region. Lines bounding \( B_1 \), \( B_2 \), or \( B \) that slope upwards arise from points to the left of \( x = z_1 \), and lines that slope downwards arise from points to the right of \( x = z_1 \). Note that the slopes of the upper boundaries of \( B \), \( B_1 \), and \( B_2 \) at \( x = z_1 \) are unaffected by this transformation (see Figure 8).

Define

\[
Q_{\mathcal{P}_1, \mathcal{P}_2} = \int_{-\infty}^{\infty} e^{-\left| C(t) \right|} dt.
\]

The next result shows that \( Q_{\mathcal{P}_1, \mathcal{P}_2} \) can “almost” be factored as a function of \( \mathcal{P}_1 \) times a function of \( \mathcal{P}_2 \). This will be the key step in the proof of Theorem 3.
Lemma 18. For any choice of \( P_1, P'_1, P_2, P'_2 \), we have

\[
\frac{Q_{P_1,P_2} Q_{P'_1,P'_2}}{Q_{P_1,P'_2} Q_{P'_1,P_2}} \leq 8\sqrt{2\pi} \left( \min\{z_1,z_2\} \right)^{3/2} + 16.
\]

Proof. Write \( Q^L_{P_1,P_2} \) (respectively \( Q^R_{P_1,P_2} \)) for the integral of \( e^{-|C(t)|} \) over \( t < 0 \) (respectively \( t > 0 \)), i.e., over \( t \) such that the region \( C(t) = C_{P_1,P_2}(t) \) is to the left (respectively right) of the line \( x = z_1 \). Write \( P^0_1 \) for the single point \( (z_1,0) \) and \( P^0_2 \) for the single point \( (0,0) \), so that both represent a single point on the \( x \)-axis at the intersection of the two strips. For \( t < 0 \) and any \( P_2 \) we have \( |C_{P_1,P_2}(t)| \leq |C_{P_1,P_2}(0)| \). Indeed, the lowest boundary on \( B \) to the left of \( x = z_1 \) we can force by placing points to the right of \( x = z_1 \) and with a given value of \( t \) is obtained when we place points on \( x = z_1 \). Hence \( Q^L_{P_1,P_2} \geq Q^L_{P_1,P_2} \). On the other hand, if \( C_{P_1,P_2}(t/2) \not\subseteq C_{P_1,P_2}(t) \) (keeping the vertical alignment of \( P_1 \) fixed), then \( C_{P_1,P_2}(t) \setminus C_{P_1,P_2}(t/2) \) is not affected by the points in \( P_1 \) and \( |C_{P_1,P_2}(t) \setminus C_{P_1,P_2}(t/2)| \geq |C_{P_1,P_2}(-t/2)| \) (see Figure 8). Thus

\[
e^{-|C_{P_1,P_2}(t)|} \leq e^{-|C_{P_1,P_2}(t/2)|} + e^{-|C_{P_1,P_2}(-t/2)|},
\]

so, by integrating from \( t = -\infty \) to \( t = 0 \), we have

\[
Q^L_{P_1,P_2} \leq 2(Q^L_{P_1,P_2} + Q^R_{P_1,P_2}).
\]

Hence, writing \( c_L = Q^L_{P_1,P_2} \) and \( c_R = Q^R_{P_1,P_2} \), we have

\[
c_L \leq Q^L_{P_1,P_2} \leq 2(c_L + c_R),
\]

and by symmetry

\[
c_R \leq Q^R_{P_1,P_2} \leq 2(c_L + c_R).
\]

Now

\[
Q_{P_1,P_2} = Q^L_{P_1,P_2} + Q^R_{P_1,P_2},
\]

so

\[
c_L + c_R \leq Q_{P_1,P_2} \leq 4(c_L + c_R).
\]

Thus the quotient in the statement of the lemma is at most

\[
\frac{4(c_L + c_R)4(c'_L + c'_R)}{(c_L + c'_R)(c'_L + c_R)},
\]

where \( c'_L \) and \( c'_R \) are defined using \( P'_1 \) instead of \( P_1 \). Without loss of generality we may assume \( z_1 \leq z_2 \) and \( c'_L \geq c_L \). Then

\[
\frac{4(c_L + c_R)4(c'_L + c'_R)}{(c_L + c'_R)(c'_L + c_R)} \leq 16 \frac{c'_L + c'_R}{c_L + c'_R} \leq 16 \frac{c_L}{c_L}
\]

29
The smallest value of $c_L$ occurs when $P_1 = P_1^0$ as in this case $|C(t)|$ is as large as possible. Indeed, $|C(t)| = |t| z_1$ as $C(t)$ is of constant height $|t|$ and width $z_1$. Thus $c_L = \int_0^\infty e^{-t z_1} \, dt = 1/z_1$. The largest value of $c_L'$ occurs when $P_1' = \{(0,0)\}$ as this causes the height of $C(t)$ to decrease as rapidly as possible away from the line $x = z_1$ and hence minimizes the area $|C(t)|$. In this case the height of $C(t)$ decreases linearly away from $x = z_1$ with slope $z_1$ and so

$$|C(t)| = \begin{cases} |t|^2/(2 z_1) & \text{if } t \leq z_1^2; \\ (|t| - z_1^2/2) z_1 & \text{otherwise.} \end{cases}$$

(The first expression occurring when the shifted $C(t)$ is a triangular region and the second when it forms a trapezoid that extends all the way to the line $x = 0$.) Hence in this case we have

$$c_L' = \int_0^{z_1^2} e^{-t^2/(2 z_1)} \, dt + \int_{z_1^2}^\infty e^{-(t-z_1^2/2) z_1} \, dt \leq \sqrt{\pi z_1/2} + 1/z_1$$

Thus

$$\frac{Q_{P_1, P_2} Q_{P_1', P_2'}}{Q_{P_1, P_2'}} \leq 16 \sqrt{\pi z_1/2} + 1/z_1 \leq 8 \sqrt{2 \pi z_1^{3/2}} + 16.$$
Let $A = (a_{ij})$ be a square matrix with non-negative entries. For any $K \geq 1$, we say $A$ is $K$-nearly rank 1 if for all $i, j, k, l$,

$$a_{ij}a_{kl} \leq Ka_{il}a_{kj}.$$  \hfill (16)

Note that a matrix that is 1-nearly rank 1 is in fact of rank 1.

**Lemma 19.** If $A$ is $K$-nearly rank 1 and $B$ has non-negative entries, then the matrix product $AB$ (or $BA$) is $K$-nearly rank 1 (with the same value of $K$).

**Proof.** Writing $A = (a_{ij}), B = (b_{ij}),$ and $AB = (c_{ij})$, we have

$$a_{ij}a_{kl}b_{jp}b_{lq} \leq Ka_{il}a_{kj}b_{jp}b_{lq}.$$ 

Summing over $j$ and $l$ gives

$$c_{ip}c_{kp} \leq Kc_{iq}c_{kp}$$

as required. \hfill \Box

The following lemma is a quantitative version of Perron’s Theorem on eigenvalues of strictly positive matrices (see [8]).

**Lemma 20.** Suppose $A$ is a square matrix with strictly positive entries which is $K$-nearly rank 1. Then for any $\varepsilon$, $0 < \varepsilon < \frac{1}{2}$, and any $N \geq 1 + K\log(3K/\varepsilon)$, the matrix $A^N$ is $(1+\varepsilon)$-nearly rank 1.

Note that the bound on $N$ is independent of the dimensions of the matrix $A$.

**Proof.** By Perron’s Theorem, there exists a strictly positive eigenvector $v$, $Av = \lambda v$, with $\lambda$ the (unique) maximal eigenvalue. From (16), and writing $A = (a_{ij})$,

$$\lambda a_{ij}v_k = \sum_l a_{ij}a_{kl}v_l \leq \sum_l Ka_{il}a_{kj}v_l = K\lambda v_i v_{kj}.$$ 

Define $u_j$ for each $j > 0$ to be the maximal $u_j$ such that $a_{ij} \geq v_i u_j$ for all $i$. Then for each $j$ there exists an $k$ such that $a_{kj} = v_k u_j$. Thus $\lambda a_{ij}v_k \leq K\lambda v_i v_k u_j$, and so $a_{ij} \leq K v_i u_j$ for all $i$ and $j$. Let $B = (v_i u_j)$. Then $B$ is a rank 1 matrix that approximates $A$ in the sense that $B \leq A \leq KB$ (inequality holding entry-wise). Now $v$ is an eigenvector for $B$ with eigenvalue $\lambda' = u^T v$. But $Bv \leq Av \leq KBv$, so $\lambda' \leq \lambda \leq K\lambda'$. Now fix a basis vector $\delta_k = (\ldots,0,1,0,\ldots)$ and consider $A^N \delta_k$. Write

$$A^N \delta_k = (A - B + B)^{N-1} A \delta_k = (A - B)^{N-1} A \delta_k + c_k v = w + c_k v,$$
where $c_k$ is some scalar. (For any vector $x$, $M_1M_2M_3\ldots M_{N-1}x$ is proportional to $v$ when all the $M_i \in \{A - B, B\}$ and at least one $M_i = B$.) We aim to show that $w$ is small compared with $c_kv$. Now $A\delta_k \leq KB\delta_k = Ku_kv$, so
\[
w = (A - B)^{N-1}A\delta_k \leq K(\lambda - \lambda')^{N-1}u_kv.
\]
But $\lambda' \geq \lambda/K$, so
\[
K(\lambda - \lambda')^{N-1} \leq K(1 - 1/K)^{N-1}\lambda^{N-1} \leq Ke^{-(N-1)/K}\lambda^{N-1} \leq \varepsilon'\lambda^{N-1},
\]
where $\varepsilon' = \varepsilon/3$. Hence $w \leq \varepsilon'\lambda^{N-1}u_kv$. Now $A^N\delta_k \geq A^{N-1}B\delta_k = \lambda^{N-1}u_kv$, so $w \leq \varepsilon'A^N\delta_k = \varepsilon'(w + c_kv)$. Hence $w \leq \varepsilon'(1 - \varepsilon')^{-1}c_kv$, and so
\[
c_kv \leq w + c_kv = A^N\delta_k \leq (\varepsilon'(1 - \varepsilon')^{-1} + 1)c_kv = (1 - \varepsilon')^{-1}c_kv.
\]
Hence $v_i c_j \leq (A^N)_{ij} \leq (1 - \varepsilon')^{-1}v_i c_j$. The result now follows since $(A^N)_{ij}(A^N)_{kl}$ and $(A^N)_{id}(A^N)_{kj}$ both lie between $v_i v_k c_j c_l$ and $(1 - \varepsilon')^{-2}v_i v_k c_j c_l$, and $(1 - \varepsilon')^{-2} \leq 1 + \varepsilon$. \hfill \Box

**Proof of Theorem 3.** By Lemma 15 we note that in estimating $\varepsilon(z)$ for large $z$ we need only consider configurations where the slope of the upper boundary of $B$ is $O(z^{1/3})$. That is, configurations where the upper boundary of $B$ is determined by points of $P$ within horizontal distance $O(z^{1/3})$. Fix $z_0 \geq 4$ and define $f(x, x_0)$ by
\[
f(x, x_0) = \begin{cases} (x - x_0)^2/2, & \text{if } 2\left[\frac{x_0}{z_0}\right] - 1 \leq x < (2\left[\frac{x_0}{z_0}\right] + 3)z_0; \\ \infty, & \text{otherwise.} \end{cases}
\]
Now define $\varepsilon_f(z) = \varepsilon_{f,0}(z)$ as in Lemma 16. The rather bizarre definition of $f$ is chosen so that (20) below is an exact equality, rather than just an approximation. Indeed, $f$ is defined so that if $x_0 \in [(2k - 1)z_0, (2k + 3)z_0]$ then $f$ is a quadratic precisely when $x \in [(2k - 1)z_0, (2k + 3)z_0]$. In particular $f(x, x_0) = (x - x_0)^2/2$ for $|x - x_0| \leq z_0$, so the regions $B$ used in the definition of $\varepsilon_f(z)$ and $\varepsilon(z)$ are the same unless there is a point on the upper boundary of (either version of) $B$ with slope at least $z_0$. Moreover, $f(x, x_0) \geq (x - x_0)^2/2$ so $\varepsilon_f(z) \geq \varepsilon(z)$. Thus by Lemma 16
\[
0 \leq \varepsilon_f(z) - \varepsilon(z) \leq \varepsilon_{f,z=0}(z) \leq e^{c_1 z^2 - c_2(z_0 - 2)^3} \leq e^{c_1 z^2 - c_2 z_0^3/8}. \tag{17}
\]
Now assume that $z \leq c_2 z_0^3$ where $c > 0$ is chosen so that $c < c_2/(8(c_1 + 1))$ and hence $c_1 z - c_2 z_0^3/8 < c_1 z - (c_1 + 1)z = -z$. Since $\varepsilon(z) \geq z$ for all $z \geq 0$, we have
\[
\varepsilon_f(z)/\varepsilon(z) = 1 + O(e^{-z}). \tag{18}
\]
32
Divide $[0, z]$ into three intervals, $[0, 2z_0]$, $[2z_0, z - 2z_0]$, and $[z - 2z_0, z]$. Let $y_1, y_2, y_3$ be the height of the lowest point of $P$ above each interval, so that $\min\{y_1, y_2, y_3\} = 0$, and conditioned on $y_i = 0$, the other $y_j$ are independent exponential random variables. Write $P = (P_1 + (0, y_1)) \cup (P_2 + (2z_0, y_2)) \cup (P_3 + (z - 2z_0, y_3))$ as in the proof of Lemma 17. Define for $z \geq 4z_0$ a function $P^{(z)}$ of a pair of distributions of points by

$$P^{(z)}_{P_1, P_3} = \mathbb{E}_{P_2, y_1, y_2, y_3} e^{\varepsilon |\mathcal{R}|}.$$ 

In other words, $P^{(z)}_{P_1, P_3}$ is given by the same formula as for $\varepsilon_f(z)$ except conditioned on $P_1$ and $P_3$. In particular

$$\varepsilon_f(z) = \mathbb{E}_{P_1} \mathbb{E}_{P_3} P^{(z)}_{P_1, P_3}. \tag{19}$$

Now consider the interval $[0, z]$ with $z = 4z_0$. Split $P = (P_1 + (0, y_1)) \cup (P_2 + (2z_0, y_2))$ as in the proof of Lemma 17 and define a function $Q$ of a pair of distributions of points above by

$$Q_{P_1, P_2} = \int_{-\infty}^{\infty} e^{-|C(t)|} dt$$

where $C(t)$ is defined as in Lemma 17, except using $f$ instead of parabolas. In this case $C(t)$ is restricted to lie above $[z_0, 3z_0]$, and would be unaffected by points of $P$ lying to the left of $x = 0$ or to the right of $x = 4z_0$.

Using the same argument as in Lemma 17 we have that

$$P^{(4z_0+z)}_{P_1, P_3} = \mathbb{E}_{P_2, P_4} P^{(4z_0)}_{P_1, P_2} P_{P_2, P_3} P^{(z)}_{P_3, P_4}, \quad z \geq 4z_0, \tag{20}$$

where we have split the strip into two strips of widths $4z_0$ and $z$ respectively, $P_1, P_2$ represent the points in the leftmost and rightmost $2z_0$ of the first strip and $P_3, P_4$ represent the points in the leftmost and rightmost $2z_0$ of the second strip (up to vertical translations). Here we use the fact that only points of $P_2$ and $P_3$ can affect the area of the region $C(t)$ used in the definition of $Q_{P_3, P_2}$. We wish to use Lemma 20 to show that $P^{(z)}_{P_1, P_2}$ increases exponentially with $z$ to a high degree of accuracy. First, since Lemma 20 considers only finite matrices, we discretize the probability space of possible values of the $P_i$. Since the bounds in Lemma 20 are independent of the size of the matrix, this can be done to arbitrary accuracy. We now can regard $P = (P_{P, P'})_{P, P'}$ and $Q = (Q_{P, P'})_{P, P'}$ as finite matrices whose rows and columns are indexed by the possible configurations of $P_i$. Then (20) becomes

$$P^{(4z_0+z)} = P^{(4z_0)}EQEP^{(z)}, \quad z \geq 4z_0,$$

where $E$ is a diagonal matrix with entry $E_{P, P}$ being the probability of $P$. Thus by induction

$$P^{(4kz_0+z)} = (P^{(4z_0)}EQE)^k P^{(z)} \quad \text{for } k \geq 1, \ z \geq 4z_0. \tag{21}$$
By Lemma 18 we know that $Q$ is $K$-nearly rank 1 with $K = O(z_0^{3/2})$. (The proof of Lemma 18 applies to the slightly modified parabolas with only a slight change in the constants.) Thus by Lemma 19 we know that $P^{(4z_0)}EQE$ is also $K$-nearly rank 1. Thus if $k$ is larger than $Cz_0^{11/6}$ then $A = (P^{(4z_0)}EQE)^k$ is $(1 + \varepsilon)$-nearly rank 1 with $\varepsilon = e^{-\Omega(z_0^{1/3})}$. Now by (19) and (21), $\varepsilon f(4kz_0 + z)$ is of the form $u^T Av$ and $\varepsilon f(8kz_0 + z)$ is of the form $u^T A^2 v$ for some positive vectors $u$ and $v$. Using $B \leq A \leq (1 + \varepsilon)B$ with $B$ of rank 1, $B^2 = \lambda B$, as in the proof of Lemma 20, we obtain

$$\lambda'(1 + \varepsilon)^{-1}u^T Av \leq \lambda u^T Bv = u^T B^2 v \leq u^T A^2 v$$

$$\leq (1 + \varepsilon)^2 u^T B^2 v = \lambda'(1 + \varepsilon)^2 u^T Bv \leq \lambda'(1 + \varepsilon)^2 u^T Av,$$

so

$$|\log \varepsilon f(8kz_0 + z) - \log \varepsilon f(4kz_0 + z) - \log \lambda'| < 2\varepsilon$$

for $z \geq 4z_0$, $k \geq Cz_0^{11/6}$. Now $\lambda'$ depends on $z_0$ and $k$, but not on $z$, and $\varepsilon = e^{-\Omega(z_0^{1/3})}$, so setting $z_1 = 4kz_0$ and replacing $z + 4kz_0$ by $z$, we have

$$\log \varepsilon f(z + z_1) - \log \varepsilon f(z) = h(z_1) + O(e^{-\eta' z_0^{1/3}})$$

for $z \geq 2z_1$ and some fixed $\eta' > 0$ independent of $z_0$ or $z_1$. Using (18) we deduce that provided $2z_1 \leq z \leq c_0^3 - z_1$,

$$\log \varepsilon(z + z_1) - \log \varepsilon(z) = h(z_1) + O(e^{-\eta'' z_1^{1/9}}), \quad (22)$$

for some $\eta'' > 0$ independent of $z$, $z_0$, and $z_1$.

Now for sufficiently large $z_1$ we can find a $z_0 \geq 4$ and $k \geq Cz_0^{11/6}$ with $z_1 = 4kz_0$ and $10z_1 \leq c_0^3 - z_1$, say. Indeed, it is enough to find $z_0$ with $z_0^{17/6} \ll z_1 \ll z_0^3$. We fix such a choice and now regard $z_0$, $k$, and hence $h(z_1)$, as functions of $z_1$ only. Thus (22) holds for all sufficiently large $z_1$ and $2z_1 \leq z \leq 10z_1$.

Assume $z_1 \leq z_2 \leq 2z_1$. Then applying (22) twice with $(z, z_1) = (z_1 + 2z_2, z_1), (3z_1 + 2z_2, z_2)$, adding, and comparing with the result for $(2z_1 + 2z_2, z_1 + z_2)$, gives

$$h(z_1 + z_2) = h(z_1) + h(z_2) + O(e^{-\eta'' z_1^{1/9}}). \quad (23)$$

Using $\{(\lfloor n^{-1} \rfloor + 1)z_1, (\lfloor n^{-1} \rfloor + \zeta)z_1\}$ for $\{z_1, z_2\}$ in (23), we can show by induction that $h((n + \zeta)z_1) = nh(z_1) + h(\zeta z_1) + nO(e^{-\eta'' z_1^{1/9}})$ for all $n \geq 0$ and $\zeta \in [1, 2]$. Letting $n \to \infty$ we see that both $\limsup h(z)/z$ and $\liminf h(z)/z$ are of the form $(h(z_1) + O(e^{-\eta'' z_1^{1/9}}))/z_1$ for every sufficiently large $z_1$. Therefore the limit $\alpha = \lim_{z \to \infty} h(z)/z$ exists, and $h(z_1) = \alpha z_1 + O(e^{-\eta'' z_1^{1/9}})$. 

34
Set $g(z) = \log \varepsilon(z) - \alpha z$. Then by (22) we have

$$g(z + z_1) = g(z) + O(e^{-\eta z_1 \varepsilon})$$

(24)

for all $z$ with $2z_1 \leq z \leq 10z_1$. From this one can deduce that the limit $\beta = \lim_{z \to \infty} g(z)$ exists. Indeed, setting $z = 2z_1$ and using induction gives

$$g(1.5^k(2z_1)) = g(2z_1) + \sum_{i=0}^{k} O(e^{-\eta'(1.5^k z_1)^{1/9}}) = g(2z_1) + O(e^{-\eta z_1^{1/9}}).$$

One more application of (24) gives $g(z) = g(2z_1) + O(e^{-\eta z_1^{1/9}})$ for all $z$ with $3(1.5)^k 2z_1 \leq z \leq 11(1.5)^k 2z_1$, and hence for all sufficiently large $z$. As this holds for all sufficiently large $z_1$, $\beta = \lim_{z \to \infty} g(z)$ exists and

$$\log \varepsilon(z) = \alpha z + g(z) = \alpha z + \beta + O(e^{-\eta'(z/2)^{1/9}}) = \alpha z + \beta + O(e^{-\eta z/9}).$$

Finally we note that $\varepsilon(z) = z \mathbb{E}(e^{||B||}) \geq z \to \infty$ as $z \to \infty$, so $\alpha > 0$.

Monte Carlo computer simulations were performed to estimate $\varepsilon(z)$ for various values of $z$ up to $z = 8$ (see Figure 9 for a plot up to $z = 5$). Using these results, the constants in Theorem 3 were estimated as

$$\alpha = 1.12794 \pm 0.00001$$

$$\beta = -1.05116 \pm 0.00005$$

(25)

(errors are $\pm 1$ standard deviation). The $O(e^{-\eta z^{1/9}})$ error term in Theorem 3 appears to be conservative as it actually seems to tend to zero extremely rapidly as $z \to \infty$. Indeed, the approximation $\varepsilon(z) \approx e^{\alpha z + \beta}$ is within $2\%$ of the correct value when $z > 0.85$, and for $z > 3$ the error is insignificant.

For small $z$, one can expand $\varepsilon(z)$ as a power series in $z$. One can show that the only non-zero terms are of the form $c z^{1+3k}$. The first few terms are

$$\varepsilon(z) = z + \frac{1}{12} z^4 + 0z^7 + \frac{1}{64800} z^{10} - \frac{1}{2721600} z^{13} + \cdots$$

(26)

We obtained these coefficients by expanding $\varepsilon(z) = z \mathbb{E}(e^{||B||}) = z + z \mathbb{E}|B| + z\mathbb{E}(|B|^2)/2 + \cdots$, and evaluating terms via symbolic integration using Mathematica.
7 Extending to large $h$

We now consider the issues that occur when $h$ is large. The main problem is that the break may be topologically complicated, and may loop around on itself. We wish to show an approximately exponential dependence of $I_{h,r}$ as a function of $h$, and so we aim to compare $I_{2h,r}$ with $I_{h,r}^2$. This will be done by considering $S_{2h}$ as two copies of $S_h$, one on top of the other, and matching breaks up on these two independent strips. However, to combine breaks of two thin strips to form a break of a thicker strip needs these breaks to look “nice” at the point at which they meet.

We use a technique involving “surgery”, that is, we take breaks that are “bad” and map them in a measure preserving way into breaks that are “good” by cutting and pasting certain regions in $S_h$. As long as this map is a “not too many”-to-1 map, we can then lower-bound the proportion of breaks that are good. We shall use the following general lemma, leaving the precise definition of a “good” break to later.

Lemma 21. Suppose we have, for each bad break $b$, an at most $k$-to-1 map $\psi_b$ on the probability space that converts this break into a good break without destroying any break or converting any other good break into a bad break. Suppose further that it can only change bad breaks to good breaks if they are originally adjacent to $b$, and can generate at most one new break which is then immediately to the right of $b$. If $\psi_b$ is measure preserving when restricted to any subset where it is injective, then the intensity of good breaks is at least $I_{h,r}/(12k + 1)$.

Proof. Fix a large interval $[0, x]$ and let $T = \{b_1, \ldots, b_n\}$ be a linearly ordered set of symbols representing the breaks in $[0, x] \times [0, h]$ in their left-to-right order. Let $p_{T,S}$ be the probability that there are $|T|$ breaks and the subset $S \subseteq T$ corresponds to those breaks
that are good. The existence of \( \psi_b \) implies that for \( b \in T \setminus S \),
\[
p_{T,S} \leq k \left( \sum_{A \subseteq \{b^-,b^+\}} p_{T,S \cup \{b\}} + \sum_{A \subseteq \{b^-,b^+,b'\}} p_{T \cup \{b'\}, S \cup A \cup \{b\}} \right)
\]
where \( b^\pm \) are the breaks adjacent to \( b \) in \( T \), and \( b' \) is a new break inserted immediately to the right of \( b \) in \( T \) (between \( b \) and \( b^+ \)). Indeed, the right hand side gives at most \( k \) times the probability of the image of \( \psi_b \), which includes the cases when \( b \) is made good, \( b^\pm \) are possibly made good, and either zero or one new break \( b' \) is generated, which may be either good or bad.

Summing over all pairs \( (S,b) \) with \( |S| = r \) and \( b \notin S \) and letting \( p_{n,r} \) denote the probability that there are \( n \) breaks of which \( r \) are good, we have
\[
(n-r)p_{n,r} \leq k \left( \sum_{i=0}^{2} \binom{n}{i} (r+1+i)p_{n,r+1+i} + \sum_{i=0}^{3} \binom{n}{i} (r+1+i)p_{n+1,r+1+i} \right).
\]
The factor of \( n-r \) gives the number of choices for \( b \) for each choice of \( S \), and the factors of \( r+1+i \) on the right hand side count the number of choices of \( b \) for each fixed choice of the set \( S \cup A \cup \{b\} \) when \( |A| = i \). Summing over \( r \) and \( n \) gives
\[
E_B - E_G \leq k(4E_G + 8E_G),
\]
where \( E_B = \sum_{n,r} np_{n,r} \) is the expected number of breaks and \( E_G = \sum_{n,r} rp_{n,r} \) the expected number of good breaks in \([0,x] \times [0,h]\). Thus \( E_G \geq E_B/(12k+1) \), and so the result follows on letting \( x \to \infty \).

For any point \( v \in \mathbb{R}^2 \), write \( B_r(v) = \{x \in \mathbb{R}^2 : \|x-v\| < r\} \) for the open disk of radius \( r \) about \( v \). For \( z \in \mathbb{R} \times [0,r] \), write \( z \) for the rightmost point on the \( x \)-axis that lies within distance \( r \) of \( z \). We shall make use of the following simple lemma.

**Lemma 22.** If \( v_i = (x_i, y_i) \in \mathcal{P} \), \( i = 1,2,3 \), lie in distinct components of \( G_{h,r} \), \( y_i \in \[0,r\) \), and \( x_1 < x_2 < x_3 \), then \( r_{v_3} > r_{v_1} \).

**Proof.** If \( y_1 > y_2 \) then we must have \( r_{v_1} \leq x_2 \) as \((x_2,0)\) is at least as far from \( v_1 \) as \( v_2 = (x_2,y_2) \) and \( \|v_1 - v_2\| \geq r \). Thus \( r_{v_1} < x_3 < r_{v_3} \). Similarly, if \( y_1 > y_3 \) then \( r_{v_1} \leq x_3 < r_{v_3} \). Thus we may assume \( y_1 \leq \min\{y_2, y_3\} \). Set \( v_4 = (x_4, y_4) := (v_{v_1}, 0) \), \( \tilde{v}_1 := (x_1, r) \), and assume \( r_{v_3} \leq r_{v_1} \). Then \( x_3 < r_{v_3} \leq r_{v_1} \) and so \( v_2, v_3 \in (x_1, x_4) \times \{y_1, r\} \). But \( r_{v_3} \leq r_{v_1} \) implies that \( v_3 \notin B_r(v_4) \). Also \( |x_4 - x_1| \leq r \), so \((x_1, x_4) \times \{y_1, r\} \subseteq B_r(\tilde{v}_1) \cup B_r(v_4) \). Hence \( v_3 \in B_r(\tilde{v}_1) \) and so \( \tilde{v}_1 \in B_r(v_3) \). But then \((x_1, x_3) \times \{y_3, r\} \subseteq B_r(v_3) \). Also \((x_1, x_3) \times \{y_1, y_3\} \subseteq B_r(v_1) \cup B_r(v_3) \), so there is no possible location for the vertex \( v_2 \in ((x_1, x_3) \times \{y_1, r\}) \setminus (B_r(v_1) \cup B_r(v_3)) = \emptyset \), a contradiction. \( \square \)
A break between good components $C$ and $C'$ is called a *bottom left clean* break if the following holds. Assume $C$ is to the left of $C'$ and $v \in C$ is the rightmost vertex of $C$ that is within distance $r$ of $\partial S_h^-$ Then there do not exist vertices $u, u'$ in some component $C'' \neq C$ (possibly $C'' = C'$) and within distance $r$ of $\partial S_h^-$, with $u$ on the left of $v$, $u'$ on the right of $v$, and $r_{u'} > r_v$. We define bottom right clean and top left/right clean similarly.

**Lemma 23.** There exists $h_0 > 0$ such that for all $h \geq h_0$ and $r \geq 7$, the intensity of breaks that are bottom left clean and lie between two good components, each of width at least $e^{h/4}$, is at least $I_{h,r}/38$.

**Proof.** If $h \leq \frac{\sqrt{3}}{2}r$ then all breaks are bottom left clean as no component $C''$ can cross the vertical crossing sensor-path through $v$. The result then follows from Lemma 11 for sufficiently large $h$. Thus we may assume $h > \frac{\sqrt{3}}{2}r$. Fix a large constant $K$ and choose $h_0$ sufficiently large so that for $h \geq h_0$, 

$$e^{h/4} \geq 2Kh > Kr.$$ 

We define a break $b$ between two good components $C$ and $C'$ to be *bad* if $C$ and $C'$ are of width at least $e^{h/4}$, but the break is not bottom left clean. Hence a break is *good* if either it is bottom left clean, or it is adjacent to a good component of width less than $e^{h/4}$. Suppose $b$ is bad and let $v \in C$ and $u, u' \in C''$ be as in the above definition, with $u$ the leftmost vertex for which such a pair $u, u'$ exists. Let $\gamma$ be a sensor-path in $C''$ joining $u$ to $u'$. Then this path passes either above or below the vertex $v$ (see Figure 10). In both cases we shall construct a map $\psi_b$ as in Lemma 21 which converts the break into a bottom left clean (and hence good) break. In each case, the map $\psi_b$ will only affect vertices within distance $O(r)$ of the break, and hence (for large enough $K$) will not affect components to the left of $C$ or to the right of $C'$. In particular it will not affect any break that is not adjacent to $b$. Moreover, the only effect on $C$ and $C'$ relevant to adjacent breaks will be to possibly change their widths (by $O(r)$). This may cause an adjacent break to become good by reducing the width of $C$ or $C''$ below $e^{h/4}$.

**Case 1.** The sensor-path $\gamma$ passes below $v$.

Let $\gamma_C$ be a crossing sensor-path that passes through $v$. As the width of $C$ is at least $Kr \geq 3r$, we may assume that $\gamma_C$ meets $\partial S_h^-$ at a point at least $2r$ to the left of $v$. (We can choose $\gamma_C$ so that $x_C^-$ is at most $\frac{r}{2}$ to the left of this point, and $x_C^+$ is at most $\frac{r}{2}$ to the right of $v$.) We first show that in this case there cannot be another pair $\tilde{u}, \tilde{u}'$ in a component $\tilde{C}''$ demonstrating the uncleaness of the break unless $\tilde{C}'' = C''$ and any sensor-path $\tilde{\gamma}$ joining $\tilde{u}, \tilde{u}'$ also passes below $v$. Indeed, if the path $\tilde{\gamma}$ between them went below $v$ and $\tilde{C}'' \neq C''$, there would be two sensor-paths from different components passing below $w := v$ (see Figure 10, left), while if $\tilde{\gamma}$ went above $v$, then there would be two sensor-paths
\( \gamma \) joining \( u \) to \( u' \) passes below \( v \) (left) or above \( v \) (right). Other paths \( \tilde{\gamma} \) demonstrating the uncleanness of break cannot exist before (left) or after (right) surgery as sensor-paths in distinct components would be then be too close.

(\( \gamma_C \) and \( \gamma \) with \( \gamma_C \) above \( \gamma \)) from different components passing below \( w := \tilde{u} \) (\( u \) is to the left of \( \tilde{u} \) by choice of \( u \)). In either case this would contradict either Lemma 5 or Lemma 6 as either the two sensor-paths would approach within distance \( \frac{r}{2} \) of each other, or the higher one would approach within distance \( \frac{r}{2} \) of \( w \).

Define the point \( s \) to be vertically below \( v \) at distance \( r \) from \( v \) (so \( s \) will actually lie outside of the strip \( S_h \)). Let \( p \) be the point at distance \( r \) from \( v \), but \( \sqrt{3} \) \( r \) above \( s \) and to the right of \( s \) (see Figure 11). The point \( q \) is defined to be the nearest point to \( s \) that is at \( r \) from both \( v \) and \( p \), so that \( vpq \) forms an equilateral triangle. The point \( s' \) is \( r \) to the left of \( s \), \( ws's \) is equilateral with \( w \) above \( s's' \), and \( g \) is at distance \( r \) from \( w \), at the same height as \( v \), and to the left of \( w \).

Define the region \( D^+ = (B_r(s') \cap B_r(v)) \setminus (B_r(g) \cup B_r(q)) \). Define the region \( D^- \) to be those points that are outside and below \( B_r(v) \) but above \( s \), to the left of \( q \), and either within horizontal distance \( r \) of \( q \) (if below \( q \)) or within distance \( r \) of \( q \) (if above \( q \)). Let \( L \) be the set of points in \( \mathbb{R}^2 \) that are above and to the left of \( s \).

Let \( \phi: D^- \to D^+ \) be the injective area-preserving map given by Lemma 24 below. Since \( \gamma \) joins \( u \) and \( u' \) and crosses \( vs \), the last vertex \( u_L \) of \( \gamma \) lying in \( L \) must lie in \( (B_r(s) \setminus B_r(v)) \cap L \). This vertex \( u_L \) lies in \( C'' \) and is within distance \( r \) of any point of \( D^+ \). But any point of \( D^+ \) is within distance \( r \) of \( v \) \( \in C \). Thus \( \mathcal{P} \cap D^+ \) must be empty, otherwise the components \( C \) and \( C'' \) would be joined. Move all the vertices \( x \in \mathcal{P} \cap D^- \) to their corresponding positions \( \phi(x) \in D^+ \). Some points of \( D^- \) may lie outside \( S_h \), but if \( D^- \cap S_h \neq \emptyset \) then \( D^+ \) must lie entirely within \( S_h \). We define \( \psi_b \) in this case to be the induced map on (this subset of) the probability space of \( G_{h,r} \).

Claim 1.1. This map neither creates nor destroys any break.

Equivalently, \( \psi_b \) does not create or destroy any good component. All vertices originally in \( D^- \) were in \( C'' \) as consecutive vertices on \( \gamma \) to the left and right of \( s \) are, between them,
adjacent to every vertex in $D^-$. Thus $C''$ is the only component that can lose vertices. On the other hand, $C$ gains the vertices that were moved since they are all now adjacent to $v$. Since $C$ does not lose any vertices, $C$ remains good.

The vertices that have been moved are now not adjacent to anything outside $B_r(v)$ to the right of $q$, and there are no longer any vertices in $D^-$, so they are no longer adjacent to anything outside of $C$ to the right of $s$. Consider a subpath of $\gamma_C$ joining $v$ to $\partial S_h^-$ (to the left of $vs$). Let $E \subseteq S_h$ be the region cordoned off by this subpath and $vs$. Suppose a vertex $\phi(x) \in D^+$ is adjacent to a vertex $z$ outside of $E$. Let $x' = x$ if $x \in L$, and $x = u_L$, the last vertex of $\gamma$ in $L$, if $x \notin L$. Then $x' \in E$ as a path in $C''$ joining $x'$ to $u'$, say, cannot cross $\gamma_C$. By assumption, the vertical segment of $\gamma_C$ meeting $\partial S_h^-$ is far from $\phi(x)$, so either $\phi(x)z$ or $x'\phi(x)$ must cross an edge $v_iv_{i+1}$ of $\gamma_C$. But $\|\phi(x) - x'\| \leq (\sqrt{3} - 1)r$ by Lemma 24. Thus there is a curve joining $x' \in C''$ to $z$ crossing $v_iv_{i+1}$ which is of length at most $\|z - \phi(x)\| + \|\phi(x) - x'\| \leq \sqrt{3}r$. Hence either $x'$ or $z$ is within distance $\frac{\sqrt{3}}{2}r$ of the line segment $v_iv_{i+1}$. But any such vertex must be adjacent to either $v_i$ or $v_{i+1}$. As $x' \notin C$, we must have $z \in C$. Thus if $\phi(x)$ is adjacent to a vertex $z \notin C$, then $z$ must be in a component contained within $E$. This component cannot be good (and most importantly, it cannot be $C'$), as a crossing sensor-path for this component would cross $\gamma_C$. In particular, we have not merged two good components.

Now consider the component $C''$. Suppose $C''$ is good (so that $C'' = C'$). Consider the segment of a crossing sensor-path $\gamma_{C'}$ from $u$ to $\partial S_h^+$. This must go through some vertex in $D^-$. Let $u_D$ be chosen so that it is the last such vertex in $D^-$ on this path. The next vertex on $\gamma_{C'}$ must lie to the right of $q$, outside of $B_r(v)$, and within $r$ of $D^-$. Thus it lies below $p$ and hence within $\frac{\sqrt{3}}{2}r$ of $\partial S_h^-$. Thus the remaining component (those vertices that are joined to the vertices of $\gamma_{C'}$ after $u_D$) is still good. If on the other hand $C''$ is not
good, removing vertices from $C''$ will not change this fact, so $C''$ remains bad.

Hence the resulting configuration is still a break between good components, and no other breaks have been created or destroyed.

**Claim 1.2.** The break $b$ is transformed into a bottom left clean break.

Indeed, $v$ is still the rightmost vertex of $C$ within $r$ of $\partial S_h^-$ as all vertices added to $C$ lie in $E$. Also there is now no sensor-path outside of $C$ passing below $v$ as such a path would contain a vertex in $D^-$. If some pair $\tilde{u}, \tilde{u}'$ demonstrated the uncleanness of this break, then the path joining $\tilde{u}$ and $\tilde{u}'$ would pass above $v$ and not contain any vertices in $D^+$ (as these vertices have been absorbed into $C$). But then $\tilde{u}$ and $\tilde{u}'$ would have demonstrated the uncleanness of the original break, which we have already ruled out.

**Claim 1.3.** The map $\psi_b$ is injective and measure preserving on Case 1 configurations.

As $v$ is still the rightmost vertex of $C$ within $r$ of $\partial S_h^-$, the sets $D^+$ and $D^-$ are determined by the resulting configuration. Hence this transformation can be reversed, and $\psi_b$ is injective on this subset of the probability space. As we can regard $\psi_b$ as just swapping regions in $S_h$, it is clearly measure-preserving. Thus the claim follows.

![Figure 12: The case when $\gamma$ passes above $v$.](image)

**Case 2.** The path $\gamma$ passes above $v$.

The path $\gamma_C$ which joins $v$ to the rest of the good component $C$ must pass under the vertex $u$ on $\gamma$. We now repeat the above argument with $u$ in place of $v$ and with left and right reversed (see Figure 12). However, there are a few differences in the proof. Note that the surgery is now being applied to the good component $C$. Let $C_r$ be the set of vertices of $C$ that are within distance $r$ of $\partial S_h^-$ and are to the right of the region $D^-$. Note that if $v \in D^-$ then $C_r = \emptyset$, otherwise $v$ is the rightmost vertex of $C_r$. Let $C_l$ be the set of vertices of $C$ that are within distance $r$ of $\partial S_h^-$ and are to the left of the region $D^-$. Note that $C_l \neq \emptyset$ as, for example, it contains the first vertex of the crossing sensor-path $\gamma_C$ which is at least $(K - \frac{r}{2})r$ to the left of $u$. (There must be a good separating path to the right of
$C$ passing to the left of $u$, so $u$ is to the right of the line $x = x_C^r$.) Finally, let $C^+$ be the set of vertices $\phi(x), x \in \mathcal{P} \cap D^-$, that have been moved by $\psi_b$.

**Claim 2.1.** If a sensor-path $\gamma_y$ avoiding $C'' \cup D^-$ joins a vertex $y \in C^+ \cup C_r$ to a vertex which is to the left of $u$, then it must pass through a vertex $z \notin C^+$ which is to the right of $u'$, within distance $r$ of $\partial S^r_h$, and satisfies $r_z > r_v$.

Note that by replacing $y$ by the last vertex of $\gamma_y$ in $C^+ \cup C_r$, we can assume without loss of generality that $y$ is the only vertex of $\gamma_y$ in $C^+ \cup C_r$. In particular, all vertices of $\gamma_y$ except possibly $y$ are vertices in the original graph $G_{h,r}$.

As there is no sensor-path below $u$ avoiding $C'' \cup D^-$, $\gamma_y$ must pass to the right of $u'$ and then above $u$. Indeed, $\gamma_y$ cannot cross the sensor-path $\gamma$ joining $u$ and $u'$ by the same argument as in Claim 1.1. Hence $\gamma_y$ must cross the horizontal ray $u' + ([0, \infty) \times \{0\})$ from $u'$.

Suppose $y$ is to the left of $u'$. Since the vertices of $\gamma_y$ other than $y$ cannot approach within distance $r$ of $u' \in C''$, there must be some vertex $z$ on $\gamma_y$ which is below and to the right of $u'$. However, any such vertex $z$ would satisfy $r_z > r_u > r_v$ as required. Thus we can assume that $y$ is already to the right of $u'$, and as $u'$ is to the right of $v$ we must then have $y \notin C_r$ by choice of $v$. Hence $y = \phi(x) \in C^+$.

Now let $z$ be the vertex adjacent to $y$ on $\gamma_y$. As $y$ is to the right of $u'$, $y \in D^+$ must be below $u'$ as otherwise $u' \in B_r(s)$ and hence $B_r(u')$ would contain $v \in C$. (Recall $v$ is to the left of $u'$ but to the right of $s$.) As $z \notin C_r$, $z \notin B_r(s)$. Thus as $z \notin B_r(u)$, $z$ must be to the right of $y \in D^+$ and hence to the right of $u'$. If $z$ is within distance $r$ of $\partial S^r_h$, then $z \notin C$ as $z \notin C_r$. Applying Lemma 22 to $\{v, u', z\}$ then implies $r_z > r_v$ as required. Hence we can assume that $z$ is not within distance $r$ of $\partial S^r_h$. But that means that $z$ must be closer to $u$ and higher than the point $g$ in Figure 12. But as $z \in B_r(y), z \in B_r(w)$. However, $u' \notin B_r(z)$, so $u'$ must be to the left of $w$. In particular, $u'$ is at most $\frac{r}{2}$ to the right of $v$ and $\|u' - v\| \geq r$. But this implies $r_u' \leq r_v$, contradicting the choice of $u'$.

**Claim 2.2.** This map does not destroy any break, and can generate at most one new break immediately to the right of $b$.

After surgery, the component containing $C_l$ in the transformed graph is good (by the argument in Claim 1.1 showing that $C''$ remains good if it was originally so). As before, it is possible that the moved vertices $\phi(x) \in D^+$ may join components. However, to show that $\psi_b$ does not destroy breaks, it is enough to show that there is no path from any $\phi(x) \in D^+$ to $C_l$. Taking the shortest such path, we can assume that such a path does not meet $C''$, as no path could join $C''$ and $C_l$ without first going through a vertex in $C^+$. Claim 2.1 then implies $C_l$ is joined in the original graph to a vertex $z$ within distance $r$ of $\partial S^r_h$ and to the right of $u'$ and hence to the right of $v$. But then $z \in C$ contradicts the choice of $v$.  

42
It is possible that a new good component may be generated. This may happen if $C''$ becomes good by the addition of the vertices in $C^+$ and possibly by amalgamation with some other bad components. However, it is clear that at most one new good component is formed as all the new vertices are joined to a single component. Hence at most one new break is formed, and that this break is immediately to the right of $b$, which we now identify with the break immediately after the good component containing $C_l$.

**Claim 2.3.** The break $b$ is transformed into a bottom left clean break.

Let $C_\psi$ be the (good) component containing $C_l$ in the transformed graph. Let $u'$ be the rightmost vertex of $C_\psi$ that is within distance $r$ of $\partial S_h^-$. Then Claim 2.1 implies that $u'$ is to the left of $u$, as otherwise (as in Claim 2.2) we would have a vertex $z \in C_r$ to the right of $v$. Suppose there is a pair $u, u'$ demonstrating the uncleanness of this break (see Figure 10, right) and let $\tilde{\gamma}$ be the path between them. As $u'$ is to the left of $u$, $\tilde{u}$ must also be to the left of $u$.

Now let $u \notin C''$, as otherwise we could have originally used the pair $(\tilde{u}, u')$ in place of $(u, u')$, contradicting the choice of $u$ as the leftmost vertex for which such a pair $(u, u')$ exists. Suppose then that $\tilde{\gamma}$ contains one of the moved vertices $\phi(x) \in C^+$ and consider the shortest subpath of $\tilde{\gamma}$ from $\tilde{u}$ to $C^+$. This path avoids $C''$ as otherwise $\tilde{u} \in C''$. Thus by Claim 2.1 $\tilde{\gamma}$ contains a vertex $z$ with $r_z > r_v$. But the segment of this path from $\tilde{u}$ to $z$ does not meet $C^+$. Thus this segment exists in the original graph. But then $(\tilde{u}, z)$ demonstrates the uncleanness of the original break and $\tilde{u}$ is to the left of $u$, contradicting the choice of $u$. Hence we may assume $\tilde{\gamma}$ avoids $C^+$ and so is a path in the original graph $G_{h,r}$.

Now consider $\tilde{u}'$. As $\tilde{\gamma}$ is a sensor-path in the original graph and $\tilde{u} \notin C''$, we must have $\tilde{u}' \notin C''$. If $u'$ were to the right of $u'$, then applying Lemma 22 to the vertices $\{v, u', \tilde{u}'\}$ gives $r_{u'} > r_v$, so the pair $(\tilde{u}, \tilde{u}')$ would demonstrate the uncleanness of the original break, again contradicting the choice of $u$. Thus $u'$ is to the left of $u'$. But then $\tilde{u}'$ cannot be to the right of $u$ as either $u$ or $\tilde{u}'$ would have two sensor-paths from distinct components passing beneath it, contradicting the fact that these vertices are within distance $r$ of $\partial S_h^-$ (see Figure 10, right). Thus $\tilde{u}'$ is to the left of $u$. But then $\tilde{u}'$ must be at least $\sqrt{3} r$ from $\partial S_h^-$ as the path from $v'$ to $v$ in the original component $C$ passes below $\tilde{u}'$. As $\tilde{u}'$ is to the left of $u$, within vertical distance $r - \frac{\sqrt{3}}{2} r$ of $u$, and $\|\tilde{u}' - u\| \geq r$, $u'$ must be at least $\left(1 - \left(1 - \frac{\sqrt{3}}{2}\right)^2\right)^{1/2} r$ to the left of $u$, and hence is to the left of $p$ in Figure 12. By assumption $r_{u'} > r_v$, so $v' \notin B_r(\tilde{u}') \cup B_r((r_{u'}, 0))$. Since $v'$ is within distance $r$ of $\partial S_h^-$, $v'$ must be at least $\frac{r}{2}$ to the left of $\tilde{u}'$. But then it is at horizontal distance more than $r$ from $D^-$. However, there is a path in $C$ joining $v'$ to some point in $D^-$ and the last vertex on this path before entering $D^-$ is within distance $r$ of $\partial S_h^-$. This contradicts the fact that $v'$ is the rightmost vertex of $C_\psi$ within distance $r$ of $\partial S_h^-$. 

**Claim 2.4** The map $\psi_b$ is at most 2-to-1 on Case 2 configurations, and is measure-
preserving when restricted to subsets where it is injective.

Consider the rightmost vertex $v'$ of $C_\phi$ that is within distance $r$ of $\partial S_h^-$. The $x$-coordinate of $v'$ must be within $r$ of that of $q$, so determines the $x$-coordinate position of $u$ up to an interval of length $r$. The $y$-coordinate of $u$ must lie in $[\sqrt{3}/2, r]$. Thus, given $v'$, $u$ is restricted to lie in an $r$ by $(\sqrt{3} - 1) \frac{r}{2}$ rectangle $R$. As all points of $D^+$ are within distance $\sqrt{3}r$ of $\partial S_h^-$, all vertices in $R$ lie in the original graph $G_{h,r}$. The subgraph of $G_{h,r}$ consisting of vertices in $R$ can have at most two components. Once a component is selected, $u$ is uniquely determined as the leftmost vertex in this component in $R$. Thus there are at most two choices for $u$. Thus the mapping is at most $2$-to-$1$. It is clear that it is also measure-preserving on subsets where it is injective.

Putting both cases together, we have an (at most) $3$-to-$1$ locally measure preserving mapping on an unclean break converting it into a clean break which can, at worst, generate a new break immediately to the right of $b$ and convert neighboring breaks into good breaks. By Lemma 21, the intensity of good breaks is then at least $I_{h,r}/37$. The proportion of breaks that are not surrounded by good components of widths at least $e^{h/4}$ is at most $2(e^{h/4} + 3)e^{-h/3}$ by Lemma 11. This is at most $10^{-4}$, say, if $h \geq h_0$ and $h_0$ is chosen sufficiently large. Thus the intensity of bottom left clean breaks surrounded by wide good components is at least $I_{h,r}/37 - I_{h,r}/10^4 \geq I_{h,r}/38$. \hfill \square

**Lemma 24.** There exists an injective area-preserving map $\phi: D^- \to D^+$ such that

(a) if $x \in D^- \cap L$ then $\|\phi(x) - x\| \leq (\sqrt{3} - 1)r$,

(b) if $x \in D^- \setminus L$ then $\|\phi(x) - x'\| \leq (\sqrt{3} - 1)r$ for all $x' \in (B_r(s) \setminus B_r(v)) \cap L$,

where $D^\pm$ are as in Figure 11 and $L$ is the set of points above and to the left of $s$.

**Proof.** Define $D_i^+$ to be the subset of $D^-$ which is to the right of the point $t$, where $t$ is $r/4$ to the left of $s$ (see Figure 11). Let $D_i^+$ be the subset of points of $D^+$ that lie within $(\sqrt{3} - 1)r$ of $s$. It can be shown that $D_i^+$ lies to the left of $w$, and hence all points of $D_i^+$ are also within distance $(\sqrt{3} - 1)r$ of $s'$. It is then easy to check that all points of $D_i^+$ are within distance $(\sqrt{3} - 1)r$ of any point $x' \in (B_r(s) \setminus B_r(v)) \cap L$.

Define $D_3^-$ to be the set of points of $D^-$ that are within distance $(\sqrt{3} - 1)r$ of $w$, and $D_3^+$ to be the set of points of $D^+$ that are not within distance $(\sqrt{3} - 1)r$ of either $t$ or $t'$, where $t'$ is the bottom left corner of $D^-$. Let $D_3^+ = D^+ \setminus (D_2^+ \cup D_3^-)$. Then all points of $D_2^+$ are within distance $(\sqrt{3} - 1)r$ of any point in $D_2^-$, and all points of $D_3^-$ are within distance $(\sqrt{3} - 1)r$ of any point of $D_3^+$.

A rather tedious calculation shows that $|D_i^-| < |D_i^+|$ for $i = 1, 2, 3$. Hence there exists an area-preserving injective function $\phi: D^- \to D^+$ so that $\phi(D_i^-) \subseteq D_i^+$. Any such map satisfies conditions (a) and (b), so the result follows. \hfill \square
Lemma 25. There exists $c > 0$ such that for all $h \geq \sqrt{3} r$, $r \geq 7$,
\[
    cr^{-1}I_{h,r}^2 \leq I_{2h,r} \leq 100hI_{h,r}^2.
\]

Proof. Cut the strip $S_{2h}$ horizontally into two strips $S_{2h}^{(t)}$ and $S_{2h}^{(b)}$ with $S_{2h}^{(t)}$ lying above $S_{2h}^{(b)}$. If a (good) separating path $\gamma$ exists for $S_{2h}$ then its restrictions to $S_{2h}^{(t)}$ and $S_{2h}^{(b)}$ must contain separating paths $\gamma^{(t)}$ and $\gamma^{(b)}$ of $S_{h}^{(t)}$ and $S_{h}^{(b)}$ respectively. The path $\gamma$ may cross the dividing line $y = h$ in more than one point, and hence may include points that are not on either $\gamma^{(t)}$ or $\gamma^{(b)}$. Thus we cannot assume that $\gamma^{(t)}$ and $\gamma^{(b)}$ join up. Moreover, we cannot assume that $\gamma^{(t)}$ and $\gamma^{(b)}$ are good separating paths, even though $\gamma$ is. However, the interval $J = [x_C^+, x_C^-]$ corresponding to the break in $G_{2h,r}$ containing $\gamma$ must intersect the intervals $J^{(t)} = [x_{C(t)}^+, x_{C(t)}^-]$ and $J^{(b)} = [x_{C(b)}^+, x_{C(b)}^-]$ corresponding to the breaks containing $\gamma^{(t)}$ and $\gamma^{(b)}$ in $S_{h}^{(t)}$ and $S_{h}^{(b)}$ respectively. Indeed, by Lemma 8, $\gamma^{(t)}$ and $\gamma^{(b)}$ intersect $J^{(t)} \times [h, 2h]$ and $J^{(b)} \times [0, h]$, but both paths are subpaths of $\gamma$ which lies entirely within $J \times [0, 2h]$.

Write $W$, $W^{(t)}$, and $W^{(b)}$ for the random variables representing the widths of the breaks in $S_{2h}$, $S_{h}^{(t)}$, and $S_{h}^{(b)}$ respectively. Since $h \geq \sqrt{3} r > 1$ by hypothesis, we know from Lemma 9 that $\mathbb{E}(W) \leq 10h$ and that $\mathbb{E}(W^{(t)}) = \mathbb{E}(W^{(b)}) \leq 5h$. Therefore by Markov’s inequality $\mathbb{P}(W \geq 20h) \leq \frac{1}{2}$. Fix a break in $S_{h}^{(t)}$ of width $w$. The probability that a break occurs within distance $20h$ of this break in $S_{h}^{(b)}$ is at most $\mathbb{E}(40h + w + W^{(b)})I_{h,r} \leq (45h + w)I_{h,r}$, as the break must start within an interval of length $40h + w + W^{(b)}$. Here the expectation is over the Poisson process in $S_{h}^{(b)}$, which is independent of the process in $S_{h}^{(t)}$. Thus the frequency of pairs of breaks lying within $20h$ of each other is at most $\mathbb{E}(45h + W^{(t)})I_{h,r}^2 \leq 50hI_{h,r}^2$. Since at least half of all breaks in $S_{2h}$ give rise to such a pair, $I_{2h,r} \leq 100hI_{h,r}^2$.

For the lower bound, note that $I_{h,r} > 0$ and the region $\{(h, r) : h_0 \geq h \geq \sqrt{3} r, r \geq 7\} \subseteq \mathbb{R}^2$ is compact, so by making $c$ smaller if necessary we may assume the result holds for all $h \leq h_0$. Thus in what follows we shall assume $h \geq h_0$ is sufficiently large.

Now fix the Poisson process in $S_{h}^{(t)}$ and $S_{h}^{(b)}$ and suppose that $C_{L}^{(t)}$ and $C_{R}^{(t)}$ are consecutive good components of widths at least $e^{h/4}$ separated by a bottom left clean break in $S_{h}^{(t)}$, while $C_{L}^{(b)}$ and $C_{R}^{(b)}$ are consecutive good components of widths at least $e^{h/4}$ separated by a top right clean break in $S_{h}^{(b)}$. Let $v^{(t)}$ be the rightmost vertex of $C_{L}^{(t)}$ that is within distance $r$ of the bottom of $S_{h}^{(t)}$. Define $\tilde{C}_{L}^{(t)}$ as the union of all components $C$ of $S_{h}^{(t)}$ which meet a vertex $u$ which is to the left of $v^{(t)}$ and within distance $r$ of the bottom of $S_{h}^{(t)}$. Define $\tilde{C}_{R}^{(t)}$ as the complement of $\tilde{C}_{L}^{(t)}$ in $G_{h,r}^{(t)}$. As the width of $C_{R}^{(t)}$ is $\gg r$, there is a vertex $u' \in C_{R}^{(t)}$ within distance $\sqrt{3} r < r$ of $\partial S_{h}^{(t)}$ with $r_{u'} > r_{v^{(t)}}$. By the bottom left cleanness of the break, $C_{R}^{(t)}$ cannot then contain a vertex $u$ within $r$ of $\partial S_{h}^{(t)}$ which is to the left of $v^{(t)}$. Thus $C_{R}^{(t)} \subseteq \tilde{C}_{R}^{(t)}$. Similarly define $\tilde{C}_{R}^{(b)}$ and its complement $\tilde{C}_{L}^{(b)}$, with $v^{(b)}$
the leftmost vertex of \( C_R^{(b)} \) that is within distance \( r \) of the top of \( S_h^{(b)} \). We now try to shift \( S_h^{(t)} \) horizontally so that the breaks in \( S_h^{(t)} \) and \( S_h^{(b)} \) line up to form a break in \( S_{2h} \).

Define \( C_L(t) \) as the component of \( G_{2h,r} \) containing the vertices of \( C_L^{(t)} \) after \( S_h^{(t)} \) has been shifted a distance \( t \) to the right relative to \( S_h^{(b)} \). We claim that there exists \( t_m \in \mathbb{R} \) and \( \varepsilon > 0 \) such that, for all \( t \in (t_m - \varepsilon, t_m) \), \( C_L(t) \) is a good component and there exists a break in \( G_{2h,r} \) just to the right of this component and before \( C_R^{(t)} \). We choose \( t_m \) to be the minimum \( t \) such that after the shift we have \( d(\tilde{C}_L^{(t)}), \tilde{C}_R^{(b)} = r \). Note that \( t_m \) almost surely exists, as \( \tilde{C}_L^{(t)} \) contains all good components to the right of \( C_L^{(t)} \) and \( \tilde{C}_R^{(b)} \) contains all good components to the right of \( C_R^{(b)} \), and both sets almost surely contain vertices within \( \frac{1}{2} \) of the boundary between \( S_h^{(t)} \) and \( S_h^{(b)} \).

Let \( x \in \tilde{C}_L^{(t)} \) and \( y \in \tilde{C}_R^{(b)} \) be such that after a shift of \( t_m \), \( \|x - y\| = r \). Almost surely there is an \( \varepsilon > 0 \) such that no vertex of \( \tilde{C}_R^{(b)} \) is within distance \( r + \varepsilon \) of \( x \) and no vertex of \( \tilde{C}_L^{(t)} \) is within distance \( r + \varepsilon \) of \( y \). Now fix \( t \in (t_m - \varepsilon, t_m) \). Then with a shift of \( t \) we have \( r \leq d(\tilde{C}_L^{(t)}, \tilde{C}_R^{(b)}) \leq \|x - y\| < r + \varepsilon \). Also, \( y \) is to the right of \( x \).

We now claim that if \( u \in \tilde{C}_L^{(t)} \) and \( v \in \tilde{C}_R^{(t)} \), then \( \|u - v\| > r \). Suppose otherwise. Then both \( u \) and \( v \) are within distance \( r \) of the line \( y = h \). If \( x \in \tilde{C}_L^{(t)} \) then \( x \) is to the left of \( u \) (by definition of \( \tilde{C}_L^{(t)} \)) while \( v \) is to the right of \( v \) (by definition of \( \tilde{C}_R^{(t)} \)). On the other hand, if \( x \in \tilde{C}_L^{(t)} \setminus \tilde{C}_L^{(t)} \) then \( v \) lies to the right of \( x \), and \( v \) lies in distinct components of \( G_{h,R} \). Moreover, \( r_x \leq r_v \) by the bottom left cleanness of the break. Applying Lemma 22 to \( \{v, v, x\} \), we see that \( v \) cannot be to the left of \( x \). Thus in all cases \( v \) lies to the right of \( x \). As \( \|x - v\| \geq r + \varepsilon \geq \|x - y\| \), \( v \in S_h^{(t)} \) cannot lie below the line segment \( xy \). Also, if \( v \) lies above \( x \), then as \( \|x - v\| \geq r \), the line \( uv \) cannot pass to the left of \( x \). Thus the ray from \( v \) through \( u \) crosses the ray from \( x \) through \( y \). Similarly, the ray from \( u \) through \( v \) crosses the ray from \( y \) through \( x \). Thus the finite line segments \( xy \) and \( uv \) intersect at some point \( z \), say. Now \( 2r + 2\varepsilon \leq \|u - y\| + \|v - x\| \leq \|u - z\| + \|z - y\| + \|v - z\| + \|z - x\| = \|x - y\| + \|u - v\| \leq 2r + \varepsilon \), a contradiction. Hence \( \|u - v\| > r \).

The component in \( G_{2h,r} \) containing \( C_L^{(t)} \) must be contained within \( \tilde{C}_L^{(t)} \cup \tilde{C}_R^{(b)} \) as there is no edge from this set to its complement \( C_R^{(t)} \cup C_L^{(b)} \). It remains to estimate the frequency of these breaks for which \( \varepsilon \) is not too small and the resulting component \( C_L^{(t)} \) is good. The sets \( \tilde{C}_L^{(t)} \) and \( \tilde{C}_R^{(b)} \) can be found by exploring the Poisson process in \( S_h^{(t)} \) and \( S_h^{(b)} \) without encountering any vertex of \( \tilde{C}_R^{(t)} \) or \( \tilde{C}_L^{(b)} \). As the vertices \( x \) and \( y \) depend only on \( \tilde{C}_L^{(t)} \) and \( \tilde{C}_R^{(b)} \) we can condition on \( \tilde{C}_L^{(t)} \) and \( \tilde{C}_R^{(b)} \) and ask for the probability that no vertex of \( \tilde{C}_L^{(t)} \) is within \( r + \varepsilon \) of \( x \) and no vertex of \( \tilde{C}_R^{(b)} \) is within \( r + \varepsilon \) of \( y \). The extra excluded area is at most \( \pi (r + \varepsilon)^2 - \pi r^2 = 2\pi r \varepsilon + \pi \varepsilon^2 \) in each case. If \( \varepsilon = \varepsilon_0/r \) this area is of order \( \varepsilon_0 \). Taking \( \varepsilon_0 \) to be sufficiently small, this occurs with probability at least \( 1 - 10^{-4} \), say. Thus the
(2-dimensional) intensity of pairs of clean breaks where this occurs is at least \((I_{h,r}/39)^2\), say, by Lemma 23.

Now since \(h\) is sufficiently large and the widths of the good component \(C^{(t)}_L\) is at least \(e^{h/4}\), we can show that with extremely high probability, \(C^{(t)}_L\) joins with part of \(\tilde{C}^{(b)}_L\) to form a good component \(C_L(t)\) in \(S_{2h}\). Indeed, all that is necessary is a column of non-empty \(a \times b\) rectangles (\(a\) and \(b\) as in Lemma 9) spanning one of \(a^{-1}e^{h/4}\) columns of \(S_{2h}\). The probability that one column consists entirely of non-empty rectangles is \((1 - e^{-ab})^{2h/b} \geq e^{-h/300}\) (as \(ab \geq 5.77\) and \(b \geq 2\)). The probability that none of the \(a^{-1}e^{h/4}\) columns have this property is \((1 - e^{-h/300})e^{-h/4}\). But this is at most \(\exp(-e^{h/5})\) for large enough \(h\). Since this is far smaller than the intensity of breaks, we now have a lower bound on \(I_{2h,r}\) of the form \((\varepsilon_0/r)(1 - o(1))(I_{h,r}/39)^2 \geq cr^{-1}I_{h,r}^2\).

\[\Box\]

The following lemma shows that breaks are usually “almost rectangular” when \(h \ll r^2\). We shall need this so that we can approximate the excluded areas by parabolic regions even when \(h\) is reasonably large, extending the argument in Section 6, which only applied when \(h \leq \sqrt{3}r\), to all \(h \leq \delta r^2\).

**Lemma 26.** There is a constant \(c > 0\) such that, for all \(r \geq 7\), the proportion of good components \(C\) for which the rightmost boundary of the excluded region to the right of \(C\) deviates more than an angle \(\theta \in [0, \pi/2]\) from vertical at any point is at most \(e^{-\langle \theta - \sin \theta \rangle r^2 + ch}\).

**Proof.** Let \(\gamma_b\) denote the boundary of the excluded region, that is, the rightmost boundary of \(\bigcup_{v \in C} B_r(v)\) crossing \(S_h\). We first note that for \(h \ll r\), \(\gamma_b\) cannot deviate by more than \(O(h/r)\) from vertical and so \(-\langle \theta - \sin \theta \rangle r^2 + ch \geq ch - O(h^3/r)\). Thus by increasing \(c\) if necessary, the result is automatic for \(h = O(1)\). Hence from now on we shall assume \(h\) is bounded away from zero.

Let \(A\) denote the set of points at distance at most \(r\) from \(\gamma_b\) and to the left of \(\gamma_b\). The interior of \(A\) cannot contain any points of the Poisson process \(\mathcal{P}\), although the leftmost boundary \(\gamma_l\) of \(A\) must contain the points of \(\mathcal{P}\) giving rise to \(\gamma_b\). The curves \(\gamma_b\) and \(\gamma_l\) are composed of sequences of arcs of circles of radii \(r\) centered on points on \(\gamma_l\) and \(\gamma_b\) respectively (see Figure 13). We can calculate the area of \(A\) by dividing \(A\) into sectors. Each sector has area \(\frac{\pi}{2}\) times the length of the arc that gives rise to it (area \(\phi r^2\) is \(\frac{\pi}{2}\) times arc length \(\phi r\) when the angle of the sector is \(\phi\)). Since the radii of the sectors start and end horizontally, the sum of the angles of the sectors giving rise to \(\gamma_l\) is the same as the sum of the angles of the sectors giving rise to \(\gamma_b\). Hence \(|\gamma_l| = |\gamma_b|\), where \(|\cdot|\) denotes the lengths of these paths. Thus, provided \(\gamma_l\) does not self-intersect (so that areas are not double counted), we obtain

\[|A| = \frac{\pi}{2}(|\gamma_l| + |\gamma_b|) = r|\gamma_b|\].

47
It is clear from the definition of $\gamma_b$ that $\gamma_b$ does not self-intersect, however $\gamma_l$ may self-intersect (see Figure 13). In this case, exclude sectors (equivalently arcs of $\gamma_l$) that intersect previously encountered sectors. The remaining arcs of $\gamma_l$ still have length at least $h$ since the remaining arcs connect the top and bottom of $S_h$. Thus we still have

$$|A| \geq \frac{r}{2}(h + |\gamma_b|),$$

even when $\gamma_l$ does self-intersect.

We now estimate the length $|\gamma_b|$ of $\gamma_b$. Let $\theta(s)$ be the (counterclockwise) angle of $\gamma_b$ from vertical at the point on $\gamma_b$ that is distance $sr$ along $\gamma_b$ from the point where $\gamma_b$ meets $\partial S_h^-$. The function $\theta(s)$ is piecewise linear with derivative 1 except at the corners of $\gamma_b$ where it discontinuously drops. Since $\gamma_b$ goes a total vertical distance of $h$, we have

$$h = r \int_0^{|\gamma_b|/r} \cos \theta(s) \, ds.$$

If we let $\theta_0 < \theta_1$ be any two values of $\theta(s)$ then there are subintervals of values of $s$ where $\theta(s)$ increases (with derivative 1) which between them cover all values from $\theta_0$ to $\theta_1$. Thus

$$h \leq |\gamma_b| + r \int_{\theta_0}^{\theta_1} (\cos \phi - 1) \, d\phi = |\gamma_b| - r(\theta_1 - \sin \theta_1) + r(\theta_0 - \sin \theta_0).$$

In the case when $\gamma_l$ does not self-intersect, we can take $\theta_0 = 0$, $\theta_1 = \theta$ (or $\theta_0 = -\theta$, $\theta_1 = 0$ if $\theta$ measured clockwise). Note that $\theta(0) < 0 < \theta(|\gamma_b|/r)$ so 0 is always a value of $\theta(s)$. Then $|\gamma_b| \geq h + r(\theta - \sin \theta)$, so for any fixed constant $c > 0$,

$$|A| - c|\gamma_b| = (r - c)|\gamma_b| \geq rh - ch + r^2(\theta - \sin \theta) - O(\min\{r, |\gamma_b|\}).$$

In the case when $\gamma_l$ does self-intersect, we can choose $\theta_1 = \theta_0 + \pi$ for some $\theta_0$ as the curve $\gamma_b$ must change direction by a total of more than $\pi$ in order that $\gamma_l$ self-intersects. Then $|\gamma_b| \geq h + \pi r$, so

$$|A| - c|\gamma_b| \geq \frac{rh}{2} + (\frac{\pi}{2} - c)|\gamma_b| \geq rh - ch + r^2\frac{\pi}{2} - O(r) \geq rh - ch + r^2(\theta - \sin \theta) - O(r).$$

If $h \leq \frac{\pi}{2}$ then $|\gamma_b| = O(h)$ and $\gamma_l$ does not self-intersect, otherwise $r = O(h)$. Thus for any $h$ we have

$$|A| - c|\gamma_b| \geq rh + r^2(\theta - \sin \theta) - O(h)$$

regardless of whether or not $\gamma_l$ self-intersects.

Now tile $S_h$ with squares of side length $a \approx 1$ (which we will assume divides $h$). Let $T$ be the set of squares of our tiling intersecting $\gamma_b$ and assume $T$ contains $n$ squares. By
Lemma 10, there are at most $\frac{1}{2} \mu^n$ choices for the set $T$ starting at some fixed square $R$, where $R$ is the leftmost square on the bottom row of $T$ adjacent to $\partial S_k^{-}$.

As $\gamma_b$ passes through every square of $T$ and there is no sensor within distance $r$ of $\gamma_b$ on the left of $\gamma_b$, there can be no sensor within distance $r - \sqrt{2}a$ of $T$ on the left of $T$. This excludes an area of at least $|A| - \sqrt{2}a(|\gamma_l| + |\gamma_b|) \geq |A| - 2\sqrt{2}a|\gamma_b|$. Indeed, the excluded area contains all of $A$ except those points within $\sqrt{2}a$ and to the left of $\gamma_b$ (as these may be inside $T$), and those points within $\sqrt{2}a$ and to the right of $\gamma_l$ (as from $T$ we cannot guarantee that these are sufficiently close to $\gamma_b$). As all of $T$ lies within distance $\sqrt{2}a$ of $\gamma_b$, the area of $T$ is at most $2\sqrt{2}a|\gamma_b|$ and hence $n \leq 2\sqrt{2}a^{-1}|\gamma_b|$. Thus the intensity of such breaks is bounded by

$$a^{-1} \sum_{n \leq 2\sqrt{2}a^{-1}|\gamma_b|} \frac{1}{2} \mu^n e^{-|A|+2\sqrt{2}a|\gamma_b|} \leq e^{-|A|+c|\gamma_b|}$$

for some constant $c > 0$. Thus by (28), the intensity of such breaks is

$$e^{-hr - r^2(\theta - \sin \theta) + O(h)}.$$

The total intensity of breaks is at least $e^{-hr + O(h)}$ by (2) for $h \geq \frac{\sqrt{3}}{2} r$, and at least $he^{-hr} = e^{-hr + O(h)}$ for $h < \frac{\sqrt{3}}{2} r$, $h = \Omega(1)$, by (11). Thus the proportion of breaks with angle $\theta$ is at most $e^{-r^2(\theta - \sin \theta) + O(h)}$, as claimed.

\[\Box\]

Figure 13: Estimating excluded area in Lemma 26.

## 8 Proof of Main Theorem

Proof of Theorem 2. The intensity of breaks is equal to the intensity of good components, or equivalently, the frequency of points of $P$ along the strip that are the rightmost vertices of good components. A vertex $v$ is the rightmost vertex of a good component if and only if
1. it lies on some crossing sensor-path, and

2. there are no vertices to the right of \( v \) that are within distance \( r \) of a vertex \( u \), say, to the left of \( v \) which is in the same component as \( v \).

Assume first that \( h \leq \max\{\delta r^2, \sqrt[3]{2} r\} \) with \( \delta > 0 \) a small constant to be determined. Using Lemma 26, there exists a \( c > 0 \) so that at most a proportion \( e^{-cr^2} \) of breaks are such that \( \gamma_b \) makes an angle of more that \( \frac{\pi}{3} \), say, with the vertical when \( h \leq \delta r^2 \) and \( \delta \) is sufficiently small. Moreover, this proportion is exactly zero if \( h \leq \frac{\sqrt[3]{2}}{2} r \). Since \( e^{-cr^2} = O(hr^{-5/3}) \) for \( h \geq \frac{\sqrt[3]{2}}{2} r \), we only need to consider the intensity of these “almost rectangular” breaks, as the inclusion of the non-rectangular breaks will only change the intensity by a factor of \( e^{O(hr^{-5/3})} \).

We now show that the probability of a vertex \( v \) being the rightmost vertex of a good component followed by an almost rectangular break is at most \( z^{-1}\varepsilon_{f_r}(z) e^{-hr} \), where \( z = hr^{-1/3} \) and \( f_r(x, x_0) \) is defined by

\[
f_r(x, x_0) = \begin{cases} 
    r^{4/3} - r^{2/3}(r^{4/3} - (x - x_0)^2)^{1/2} & \text{if } |x - x_0| \leq r^{2/3}; \\
    \infty & \text{otherwise.}
\end{cases}
\]

Condition first on the presence of the vertices of a crossing sensor-path \( \gamma^R_C \), and consider the probability that \( \gamma^R_C \) is the rightmost crossing sensor-path (as constructed in the proof of Lemma 11) of a good component whose rightmost vertex is \( v \). This probability is given by \( \mathbb{E}e^{-|A|-|S|} \) where \( S \) is the region to the left of \( \gamma^R_C \) examined in the construction of \( \gamma^R_C \) in Lemma 11 (which must be empty), the expectation is over a Poisson process \( \mathcal{P}' \) to the left of \( S \), and \( A \) is the area to the right of \( \gamma^R_C \) that is within distance \( r \) of the sensors of \( \gamma^R_C \cup \mathcal{P}' \). (No sensor-path outside of \( C \) can cross \( \gamma^R_C \), so excluding points within distance \( r \) of \( \gamma^R_C \cup \mathcal{P}' \) is the same as excluding points within distance \( r \) of \( \gamma^R_C \cup (\mathcal{P}' \cap C) \).

Now rotate \( \gamma^R_C \) through 90°, scale horizontally by \( r^{-1/3} \) and vertically by \( r^{1/3} \). Conditioned on the presence of vertices corresponding to those of the transformed \( \gamma^R_C \) in \([0, z] \times [0, \infty]\), the probability that the region corresponding to \( S \) is empty and there are no vertices below \( \gamma^R_C \) is \( \mathbb{E}e^{-|D|-|S|} \), where \( D \) is region of \([0, z] \times [0, \infty]\) below the rotated and scaled \( \gamma^R_C \). Now in the formula defining \( \varepsilon_{f_r}(x) \), this event contributes \( ze^{1|B|} \). But \( f_r(x, x_0) \) is defined so that \( |A| + |B| = |D| + hr \) provided the slope of the boundary of the original set \( A \) never becomes horizontal. Thus the contribution to \( z^{-1}\varepsilon_{f_r}(z) e^{-hr} \) is the same as to the probability that \( v \) is the leftmost vertex of a good component in this case. Integrating over all choices of \( \gamma^R_C \) for which the original excluded area \( A \) could have right boundary within \( \frac{\pi}{3} \) of vertical gives that the probability that \( v \) is the rightmost vertex of a good component which is followed by an almost rectangular break is at most \( z^{-1}\varepsilon_{f_r}(z) e^{-hr} \). (In fact, whether or not the right boundary of \( A \) stays within \( \frac{\pi}{3} \) of vertical is determined by
satisfies I where

Now in general for bounded random variables $r \to \infty$ the probability that a vertex is the rightmost vertex of a good component is automatic when $r$ other (or $S$ most (Lemma 15). Thus by Lemma 15, the remaining configurations contribute a fraction at upper boundary of $\partial S$ $\gamma_x$ as a function of $r$. Indeed, the only points of $C_R$ the vertices of $\eta$ has slope at most $B$ for the region $\gamma_x \leq \frac{1}{2}$ and $\eta(x, x_0) = (1 - \frac{t}{r})^4|B \leq \frac{1}{2}|B|$. Indeed, the second inequality follows from the convexity of $\eta_r(x, x_0)$ (defined in (31) below) as a function of $r^{-4/3}$, followed by integration of (30).

Write $B_r$ for the region $B$ used in the definition of $\varepsilon_f_r(z)$, and $B_{\infty}$ for the limit as $r \to \infty$, which gives the region $B$ used in the definition of $\varepsilon(z)$. We have, using $r_0 = 7$,

$$|B_\infty| \leq |B_r| \leq (1 - \frac{t}{r})^4|B_\infty| + (\frac{t}{r})^4|B_r|.$$ 

Now in general for bounded random variables $X$ and $Y$, Hölder’s inequality gives that

$$\mathbb{E}(e^{(1-\varepsilon)X+\varepsilon Y}) \leq (\mathbb{E}e^X)^{1-\varepsilon}(\mathbb{E}e^Y)^{\varepsilon}.$$ 

Thus

$$\varepsilon(z) \leq \varepsilon_f_r(z) \leq \varepsilon(z)(\varepsilon_f_r(z)/\varepsilon(z))^{(7/r)^{4/3}}.$$ 

Now

$$\frac{\partial}{\partial x} f_r(x, x_0) = (x - x_0)\eta_r(x, x_0),$$ 

where

$$\eta_r(x, x_0) = (1 - r^{-4/3}(x - x_0)^2)^{-1/2}$$ 

satisfies $1 \leq \eta_r(x, x_0) \leq 2$ for all $|x - x_0| \leq 2$ and $r \geq 7$. Thus $\varepsilon_f_r(z)/\varepsilon(z) = e^{O(z)} = e^{O(hr^{-1/3})}$ by Lemma 16. Hence we have

$$\varepsilon_f_r(z) = \varepsilon(z)e^{O(hz^{-5/3})}.$$ 

51
The theorem follows for $h \leq \max\{\delta r^2, \sqrt{3}/2 \cdot r\}$, since there are $h$ vertices per unit length along the strip and so $I_{h,r} = h z^{-1} \varepsilon_r(z) e^{-hr + O(hr^{-5/3})} = r^{1/3} \varepsilon(z) e^{-hr + O(hr^{-5/3})}$.

For $h > \max\{\delta r^2, \sqrt{3}/2 \cdot r\}$, we use Lemma 25 to reduce inductively to the case $h \leq \max\{\delta r^2, \sqrt{3}/2 \cdot r\}$. Choose $k$ minimal so that $h_0 = h/2^k \leq \max\{\delta r^2, \sqrt{3}/2 \cdot r\}$. By induction on $k$, Lemma 25 implies that

$$(c/r)^{2k-1} I_{h_0,r} \leq I_{h,r} \leq (200h_0)^{2k-1} I_{h_0,r}.$$  

As $h_0 = \Theta(r^2)$ we have

$$I_{h,r} = I_{h_0,r} e^{O(2^k \log r)} = r^{2k/3} \varepsilon(z/2^k)^{2k} e^{-2^k h r + O(2^k h r^{-5/3}) + O(2^k \log r)}.$$  

But $z/2^k = h_0 r^{-1/3} = \Theta(r^{5/3})$ is bounded away from zero. Thus by Theorem 3 we have $\varepsilon(z/2^k)^{2k} / \varepsilon(z) = \exp(O(2^k))$. Hence

$$I_{h,r} = r^{1/3} \varepsilon(z) e^{-hr + O(hr^{-5/3}) + O(2^k \log r)}.$$  

However, the $O(hr^{-5/3})$ absorbs the $O(2^k \log r) = O(h \log r / r^2)$ term, giving the result.

![Figure 14: Plots of log(I_h,r) against h for various values of r. Dotted lines indicate the approximation r^{1/3} \varepsilon(hr^{-1/3}) e^{-hr} from Theorem 2.](image)
Figure 15: Plots of \( \log(I_{h,r}) + \alpha_r h + \beta_r \) against \( z = hr^{-1/3} \) for \( r = 3, 4, 6, 8, 16 \) (outer to inner solid lines respectively). Dotted line indicates the estimate \( \log \varepsilon(z) - \alpha z - \beta \). Here \( \alpha_r \) and \( \beta_r \) are as given in Table 1.

9 Simulation Results

In this section, we provide results of our simulations and compare them with our estimates. The main result here is that our estimates are almost indistinguishable from that observed in simulations for large \( r \), and there is good agreement even for \( r \) significantly less than 7. All our simulations are sufficiently extensive that statistical error bounds are not visible on the graphs.

Our theoretical results strongly suggest that Theorem 2 can be strengthened to

\[
I_{h,r} = e^{-\alpha_r h - \beta_r} + o(1),
\]

as \( h \to \infty \), where

\[
\alpha_r = r - \alpha r^{-1/3} + O(r^{-5/3}), \\
\beta_r = -\frac{1}{3} \log r - \beta + O(r^{-4/3})
\]

as \( r \to \infty \). This does not quite follow from the above, since we may still have a \( O(\log h) \) error term in the exponent resulting from the inductive use of Lemma 25. We believe however that this is just an artifact of the proof. Indeed, a \( O(\log h) \) term would arise from the factor of \( h \) on the right hand side of (27), which in turn comes from the \( O(h) \) bound on the expected width of a break (Lemma 9) used in Lemma 25. In practice, one should use the amount by which the breaks in \( S_h^{(b)} \) and \( S_h^{(b)} \) can be moved relative to one another and still form a break in \( S_{2h} \). This should be \( O(1) \), and in particular should not depend noticeably on \( h \).
Table 1: Estimates of $\alpha_r$ and $\beta_r$ from simulations, together with the approximations given by (34). Numbers in square brackets indicate approximate 1 standard deviation errors in the last decimal place.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$\alpha_r$</th>
<th>$\beta_r$</th>
<th>$\alpha_r^{\text{est}}$</th>
<th>$\beta_r^{\text{est}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.25</td>
<td>0.0428[1]</td>
<td>3.6200[15]</td>
<td>0.0650</td>
<td>1.1773</td>
</tr>
<tr>
<td>1.5</td>
<td>0.3668[1]</td>
<td>1.6790[15]</td>
<td>0.4129</td>
<td>1.0732</td>
</tr>
<tr>
<td>1.75</td>
<td>0.7018[1]</td>
<td>1.2198[10]</td>
<td>0.7353</td>
<td>0.9927</td>
</tr>
<tr>
<td>2.5</td>
<td>1.6152[2]</td>
<td>0.8490[10]</td>
<td>1.6255</td>
<td>0.8253</td>
</tr>
<tr>
<td>3</td>
<td>2.1809[2]</td>
<td>0.7548[10]</td>
<td>2.1859</td>
<td>0.7474</td>
</tr>
<tr>
<td>4</td>
<td>3.2678[2]</td>
<td>0.6342[10]</td>
<td>3.2696</td>
<td>0.6316</td>
</tr>
<tr>
<td>6</td>
<td>5.3686[2]</td>
<td>0.4801[10]</td>
<td>5.3692</td>
<td>0.4787</td>
</tr>
<tr>
<td>8</td>
<td>7.4296[1]</td>
<td>0.3751[2]</td>
<td>7.4298</td>
<td>0.3749</td>
</tr>
<tr>
<td>16</td>
<td>15.5504[1]</td>
<td>0.1337[2]</td>
<td>15.5504</td>
<td>0.1337</td>
</tr>
</tbody>
</table>

Table 1: Estimates of $\alpha_r$ and $\beta_r$ from simulations, together with the approximations given by (34). Numbers in square brackets indicate approximate 1 standard deviation errors in the last decimal place.

Figure 14 plots the logarithm of the frequency of breaks observed in simulation against $h$ for various values of $r$ together with the estimate from Theorem 2. We observe that this approximation provide an extremely good fit to the simulation data except when $r$ is close to the critical radius for the Gilbert model $r_c \approx 1.1984$ (see [12]). The network becomes highly disconnected in this case, and so the analysis breaks down.

The simulated values of $\alpha_r$ do indeed appear consistent with (33) and (25). Using simulations, one can estimate the error terms for $\alpha_r$ and $\beta_r$ giving

$$\alpha_r \approx \alpha_r^{\text{est}} := r - 1.12794r^{-1/3} - 0.20r^{-5/3}$$

$$\beta_r \approx \beta_r^{\text{est}} := -\frac{1}{3}\log r + 1.05116 + 0.27r^{-4/3}$$

(34)

Note that the constants 1.12794 and 1.05116 are the constants from Theorem 3, and only the last coefficients (0.20 and 0.27) were estimated from the simulation estimates of $\alpha_r$ and $\beta_r$. From Table 1 one sees that the approximations in (34) are extremely good for $r \geq 3$, but get progressively less accurate for smaller values of $r$.

By comparison with Theorem 3, the $o(1)$ term in (32) should be approximately equal to $\log \varepsilon(z) - \alpha z - \beta$ where $z = hr^{-1/3}$. Figure 15 shows the values of this error term obtained from simulations for $r \geq 3$ (for $r < 3$ the error is much larger). Once again, the theoretical result is very close to the results from simulations.
References


