Maker-Breaker games on random geometric graphs

Andrew Beveridge[*], Andrzej Dudek[†], Alan Frieze[‡], Tobias Müller[§], Miloš Stojaković[¶]

May 2, 2014

Abstract

In a Maker-Breaker game on a graph G, Breaker and Maker alternately claim edges of G. Maker wins if, after all edges have been claimed, the graph induced by his edges has some desired property. We consider four Maker-Breaker games played on random geometric graphs. For each of our four games we show that if we add edges between n points chosen uniformly at random in the unit square by order of increasing edge-length then, with probability tending to one as $n \to \infty$, the graph becomes Maker-win the very moment it satisfies a simple necessary condition. In particular, with high probability, Maker wins the connectivity game as soon as the minimum degree is at least two; Maker wins the Hamilton cycle game as soon as the minimum degree is at least four; Maker wins the perfect matching game as soon as the minimum degree is at least two and every edge has at least three neighbouring vertices; and Maker wins the H-game as soon as there is a subgraph from a finite list of "minimal graphs". These results also allow us to give precise expressions for the limiting probability that $G(n,r)$ is Maker-win in each case, where $G(n,r)$ is the graph on n points chosen uniformly at random on the unit square with an edge between two points if and only if their distance is at most r.

1 Introduction

Let $H = (X, \mathcal{F})$ be a hypergraph. That is, X is a finite set and $\mathcal{F} \subseteq 2^X$ is a collection of subsets of X. The *Maker-Breaker game* on H is played as follows. There are two players, Maker and Breaker, that take turns claiming unclaimed elements of X, with Breaker moving first. Maker wins if, after all the elements of X have been claimed, the set $M \subseteq X$ of elements claimed by him contains an element of $\mathcal{F}$ (i.e. if $F \subseteq M$ for some $F \in \mathcal{F}$). Otherwise Breaker wins.

The study Maker-Breaker games has a considerable history going back to Hales-Jewett [**?**] and Lehman [13] and the subject has become increasingly popular over the past decade or so. Often the case is considered where $X = E(G)$ is the edge set of some graph G and $\mathcal{F} \subseteq 2^{E(G)}$ is a collection of graph theoretic structures of interest, such as the set of all spanning trees, the set of all perfect matchings, the set of all Hamilton cycles, or the set of all subgraphs isomorphic to a given graph H. In these cases we speak of, respectively, the connectivity game, the perfect matching game, the Hamilton cycle game and the H-game.

Already in [6], an interesting connection between positional games on the complete graph and the corresponding properties of the random graph was pointed out, roughly noting that the course

[*]Department of Mathematics, Statistics and Computer Science, Macalester College, Saint Paul, MN. E-mail: abeverid@macalester.edu

[†]Department of Mathematics, Western Michigan University, Kalamazoo, MI. E-mail: andrzej.dudek@wmich.edu.
Research supported in part by Simons Foundation Grant #244712

[‡]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA. E-mail: alan@random.math.cmu.edu. Supported in part by NSF Grant CCF1013110

[§]Mathematical Institute, Utrecht University, Utrecht, the Netherlands. E-mail: t.muller@uu.nl. Part of this work was done while this author was supported by a VENI grant from Netherlands Organisation for Scientific Research (NWO)

[¶]Department of Mathematics and Informatics, University of Novi Sad, Serbia. Email: milos.stojakovic@dmi.uns.ac.rs. Partly supported by Ministry of Science and Technological Development, Republic of Serbia, and Provincial Secretariat for Science, Province of Vojvodina.

of a game between the two (smart) players often resembles a purely random process. This many time repeated and enriched paradigm in positional game theory later came to be known as *Erdős's probabilistic intuition*, and although still lacking the form of a precise statement, it has proved a valuable source of inspiration. See [3] for an overview of the theory of positional games.

Combining two related concepts, Maker-Breaker games on the Erdős-Rényi model of random graphs were first introduced and studied in [22]. Several works followed, including [10, 4], resulting in precise descriptions of the limiting probabilities for Maker-win in games of connectivity, perfect matching and Hamilton cycle, all played on the edges of a random graph $G \sim G(n,p)$. As for the H-game, not that much is known. The limiting probability is known precisely only for several special classes of graphs [17]. Recently, the order of the threshold probability was determined for all graphs H that are not trees, and whose maximum 2-density is not determined by a K_3 as a subgraph [18]. When it comes to other models of random graphs, some positional games on random regular graphs were studied in [5].

In the current paper we add to this line of research, by considering Maker-Breaker games played on the random geometric graph on the unit square. Given points $x_1, \dots, x_n \in \mathbb{R}^2$ and $r > 0$ we define the geometric graph $G(x_1, \dots, x_n; r)$ as follows. It has vertex set $V = \{x_1, \dots, x_n\}$, and an edge $x_i x_j$ if and only if $\|x_i - x_j\| \leq r$, where $\|.\|$ is the Euclidean norm. The random geometric graph $G(n,r) = G(X_1, \dots, X_n; r)$ is defined by taking $X_1, \dots, X_n$ i.i.d. uniform at random on the unit square $[0,1]^2$. The random geometric graph essentially goes back to Gilbert [9] who defined a very similar model in 1961. For this reason the random geometric graph is sometimes also called the *Gilbert model*. Random geometric graphs have been the subject of a considerable research effort in the last two decades or so. As a result, detailed information is now known on aspects such as (k-)connectivity [19, 20], the largest component [21], the chromatic number and clique number [16, 15] and the simple random walks on the graph [7]. A good overview of the results prior to 2003 can be found in the monograph [21]. It is of course possible to define the random geometric graph in dimensions other than two, using a probability measure other than the uniform and using a metric other than the euclidean norm, and in fact several authors do this. In the present work we have decided to stick with the two-dimensional, uniform, euclidean setting because this is the most natural choice in our view, and this setting is already challenging enough.

Recall that, formally, a graph property $\mathcal{P}$ is a collection of graphs that is closed under isomorphism. We call a graph property $\mathcal{P}$ *increasing* if it is preserved under addition of edges (i.e. $G \in \mathcal{P}$ implies $G \cup e \in \mathcal{P}$ for all $e \in \binom{V(G)}{2}$). Examples of increasing properties are being connected, being non-planar, containing a Hamilton cycle, or being Maker-win in any of the games mentioned above. We define the *hitting radius* of an increasing property $\mathcal{P}$ as:

$$\rho_n(\mathcal{P}) := \inf\{r \geq 0 : G(X_1, \dots, X_n; r) \text{ satisfies } \mathcal{P}\}.$$

Here we keep the locations of the points $X_1, \dots, X_n$ fixed as we take the infimum.

We give explicit descriptions of the hitting radius for three different games, namely the connectivity game, the Hamilton cycle game and the perfect matching game. For each game, we have a very satisfying characterization: the hitting radius for $G(n,r)$ to be Maker-win coincides exactly with a simple, necessary minimum degree condition. Each characterization engenders an extremely precise description of the behavior at the threshold value for the radius. We note that these results require some very technical lemmas that explicitly catalog the structure of the graph around its sparsest regions. Finally, we also state a general theorem for the H-game for a fixed graph H. The hitting radius obeys a similar behavior, and can be determined by finding the smallest k for which the H-game is Maker-win on the k-clique. Upper bounds for such k are available, see e.g. [2], so determining k for a given H is essentially a finite problem.

We state these four results below, as couplets consisting of a theorem recognizing the coincidence of the hitting radii, followed by a corollary that describes the behavior around that critical radius. We say that an event A_n holds *whp* (short for *with high probability*) to mean that $\mathbb{P}(A_n) = 1 - o(1)$ as $n \to \infty$.

Theorem 1.1 *The random geometric graph process satisfies*

$$\rho_n(\textit{Maker wins the connectivity game}) = \rho_n(\textit{minimum degree} \geq 2) \quad \text{whp.}$$

Theorem 1.1 allows us to derive an expression for the limiting probability that Maker wins the connectivity game on $G(n,r)$, as follows. Let $n \geq 2$ be an integer and $r > 0$ be arbitrary. Since having minimum degree at least two is a necessary condition for Maker winning the connectivity game, we have:

$$\begin{array}{rcl} \mathbb{P}(\delta(G(n,r)) \geq 2) & \geq & \mathbb{P}(\text{Maker wins on } G(n,r)) \\ & = & \mathbb{P}(\text{Maker wins on } G(n,r) \text{ and } \delta(G(n,r)) \geq 2) \\ & = & \mathbb{P}(\delta(G(n,r)) \geq 2) \\ & & -\mathbb{P}(\text{Maker loses on } G(n,r) \text{ and } \delta(G(n,r)) \geq 2) \\ & \geq & \mathbb{P}(\delta(G(n,r)) \geq 2) \\ & & -\mathbb{P}(\rho_n(\text{Maker wins}) \neq \rho_n(\text{minimum degree} \geq 2)). \end{array} \tag{1}$$

Combining this with Theorem 1.1 we see that, for any sequence $(r_n)_n$ of positive numbers:

$$\mathbb{P}(\text{Maker wins the connectivity game on } G(n,r_n)) = \mathbb{P}(\delta(G(n,r_n)) \geq 2) - o(1). \tag{2}$$

Combining this with a result of Penrose (repeated as Theorem 2.10 below), we find that:

Corollary 1.2 *Let $(r_n)_n$ be a sequence of positive numbers and let us write*

$$x_n := \pi n r_n^2 - \ln n - \ln\ln n.$$

Then the following holds in the random geometric graph $G(n,r_n)$:

$$\lim_{n\to\infty} \mathbb{P}(\textit{Maker wins the connectivity game}) = \begin{cases} 1 & \textit{if } x_n \to +\infty, \\ e^{-(e^{-x}+\sqrt{\pi e^{-x}})} & \textit{if } x_n \to x \in \mathbb{R}, \\ 0 & \textit{if } x_n \to -\infty. \end{cases}$$

Let us define $N(v)$ to be the set of neighbors of vertex v, and the edge-degree of an edge $e = uv \in E(G)$ as $d(e) = |(N(v) \cup N(u)) \setminus \{u,v\}|$.

Theorem 1.3 *The random geometric graph process satisfies, for n even:*

$$\rho_n(\textit{Maker wins the perfect matching game}) = \rho_n(\textit{min. deg.} \geq 2 \textit{ and min. edge-deg.} \geq 3) \quad \text{whp.}$$

A similar argument to the one we used for the connectivity game yields the following.

Corollary 1.4 *Let $(r_n)_n$ be a sequence of positive numbers and let us write*

$$x_n := \pi n {r_n}^2 - \ln n - \ln\ln n.$$

Then the following holds in the random geometric graph $G(n,r_n)$:

$$\lim_{\substack{n\to\infty,\\ n \text{ even}}} \mathbb{P}(\textit{Maker wins the perfect matching game}) = \begin{cases} 1 & \textit{if } x_n \to +\infty, \\ e^{-((1+\pi^2/8)e^{-x}+\sqrt{\pi}(1+\pi)e^{-x/2})} & \textit{if } x_n \to x \in \mathbb{R}, \\ 0 & \textit{if } x_n \to -\infty. \end{cases}$$

Theorem 1.5 *The random geometric graph process satisfies*

$$\rho_n(\textit{Maker wins the Hamilton cycle game}) = \rho_n(\textit{minimum degree} \geq 4) \quad \text{whp.}$$

Again, we can make use of this result to obtain the precise relation between the radius r and the probability of Maker-win in the Hamilton cycle game.

Corollary 1.6 *Let $(r_n)_n$ be a sequence of positive numbers and let us write*

$$x_n := \frac{1}{2}\left(\pi n {r_n}^2 - (\ln n + 5\ln\ln n - 2\ln(6))\right).$$

Then the following holds in the random geometric graph $G(n,r_n)$:

$$\lim_{n\to\infty} \mathbb{P}(\textit{Maker wins the Hamilton cycle game}) = \begin{cases} 1 & \textit{if } x_n \to +\infty, \\ e^{-e^{-x}} & \textit{if } x_n \to x \in \mathbb{R}, \\ 0 & \textit{if } x_n \to -\infty. \end{cases}$$

Our next result, on the H game, is a bit different from our results for the other games. In particular, the "action" now takes place long before the connectivity threshold, when the average degree $\pi n r^2$ decays polynomially in n.

Theorem 1.7 *Let H be any fixed graph and let k_H denote the least k for which the H-game is Maker's win on a k-clique, and let $\mathcal{F}_H$ denote the family of all graphs on k_H vertices for which the game is Maker-win. Then*

$$\rho_n(\textit{Maker wins the } H\textit{-game}) = \rho_n(\textit{the graph contains a subgraph} \in \mathcal{F}_H) \quad \text{whp.}$$

Again the theorem on the hitting radius allows us to determine the hitting probability. This time we make use of a theorem of Penrose [21] (stated as Theorem 2.9 below) on the appearance of small subgraphs to obtain:

Corollary 1.8 *Let H be any fixed graph, and let k_H denote the smallest k for which the H-game is Maker-win on a k-clique. If $r = c \cdot n^{-k/2(k-1)}$ with $c \in \mathbb{R}$ fixed then*

$$\mathbb{P}(\textit{Maker wins the } H\textit{-game}) \mapsto f(c),$$

where $0 < f(c) < 1$ (is an expression which can be computed explicitly in principle and) satisfies $f(c) \to 0$ as $c \to -\infty$; and $f(c) \to 1$ as $c \to \infty$.

1.1 Overview

The proofs of the main theorems leverage a characterization of the local structure of the graph $G(n, r)$. In particular, we meticulously describe the graph around its sparser regions. Section 2 contains our background results. We start with a short list of known results about Maker-Breaker games, followed by a selection of results about random geometric graphs, drawn or adapted from [21]. Next, we give some geometric preliminaries, most of which provide approximations of small areas defined by intersecting disks in $[0, 1]^2$.

Section 3 contains the technical lemmas that chart a detailed cartography of $G(n, r)$. We *dissect* the unit square $[0, 1]^2$ into *cells*, which are squares of side length roughly ηr where $\eta > 0$ is very small. We fix a large constant $T > 0$ (made explicit later). We say that a cell c is *good* if it is dense, meaning that there are at least T vertices in c. The sparse cells are *bad*. We show that bad cells come in clusters of very small diameter, which we call *obstructions*. Moreover, these obstructions are well-separated from one another. These results are collected in the *Dissection Lemma 3.3*. Next, we prove the *Obstruction Lemma 3.4*, which shows that for each obstruction, there are enough points in nearby good cells to allow Maker to overcome these bottlenecks.

Section 4 contains the straight-forward proof for the connectivity game. A classic result [13] states that the connectivity game on a graph G is Maker-win if and only if G has two disjoint spanning trees. Our characterization of obstructions quickly reveal that such a pair of trees exist.

The argument for the Hamilton cycle game, found in Section 5, is far more delicate. In order to win the game on $G(n, r_n)$, Maker plays lots of local games. The games that are played in and around the obstructions require the most judicious play: this is where Maker must directly play against Breaker to keep his desired structure viable at a local level. Ultimately, he is able to construct long paths that span the obstructions and have endpoints in nearby good cells. Meanwhile, Maker plays a different kind of game in each good cell. Therein, he creates a family of flexible *blob cycles*, which consist of cycles that also contain a fairly large clique. In addition, Maker plays to claim half the edges between the vertices in nearby local games. Once the edge claiming is over, Maker stitches together his desired structure. The soup of blob cycles in each cell gives Maker the flexibility to connect together the local games. Each local game is absorbed into a nearby blob cycle. This process is repeated by merging the current blob cycles in a good cell into one blob cycle in the cell. Along the way, we also absorb other vertices that are not yet attached, and also connect the blob cycles in nearby good cells, following a pre-determined tree structure. The final result is a Hamilton cycle.

Section 6 considers the perfect matching game. The argument and game play is similar to the Hamilton cycle game. The local games in and around the obstructions are played head-to-head with Breaker, creating a matching that saturates each obstruction (and uses some nearby vertices in good cells). Meanwhile, Maker creates a Hamilton cycle through the rest of the vertices (as in the previous game) and then takes every other edge to get the perfect matching.

In Section 7, we handle the H-game with a straight-forward argument. Once we reach the threshold for the appearance of a clique on which Maker can win the H-game, we use a Poisson argument for independent copies of such a clique appearing in well-separated regions of the graph.

Finally, we conclude in Section 8 and list some directions for future research.

2 Preliminaries

In this section, we present some preliminary results that will be useful in the sequel. We start with a modest collection of previous results on Maker-Breaker games. The Maker-Breaker connectivity game is also known as the *Shannon switching game*. A classical result by Lehman [13] states:

Theorem 2.1 ([13]) *The connectivity game played on G is Maker-win if and only if G admits two disjoint spanning trees.*

This has the immediate corollary:

Corollary 2.2 *The connectivity game is Maker-win on K_n if and only if $n \geq 4$.*

The following is a standard result.

Theorem 2.3 ([2]) *Let H be a finite graph. There is an $N = N(H)$ such that Maker can win the H-game on K_s for all $s \geq N$.*

It turns out that the Hamilton cycle game, as well as several other standard games on graphs, are easy wins for Maker when played on a sufficiently large complete graph. To make the game more balanced, Chvátal and Erdős [6] introduced *biased games*. In the $(1 : b)$ biased game, Maker claims a single edge in each move, as before, but Breaker claims b edges. The parameter b is called the bias (towards Breaker). Due to the "bias monotonicity" of Maker-Breaker games, it is straightforward to conclude that for any positional game there is some value $b(n)$ such that Maker wins the game for all $b < b(n)$, while Breaker wins for $b \geq b(n)$. We call $b(n)$ the *threshold bias* for that game. Later on, we make use of a recent break-through result of Krivelevich on the threshold bias of the Hamilton cycle game on K_n.

Theorem 2.4 (Krivelevich [12]) *The threshold bias of the Hamilton cycle game on K_n satisfies $b(n) = (1 + o(1))n/\ln n$.*

2.1 Probabilistic preliminaries

Throughout this paper, $\mathrm{Po}(\lambda)$ will denote the Poisson distribution with parameter λ, and $\mathrm{Bi}(n, p)$ will denote the binomial distribution with parameters n, p. Recall that the $\mathrm{Bi}(1, p)$-distribution is also called the *Bernoulli* distribution. We will make use of the following incarnation of the Chernoff bounds. A proof can for instance be found in Chapter 1 of [21].

Lemma 2.5 *Let Z be either Poisson or Binomially distributed, and write $\mu := \mathbb{E}Z$.*

(i) *For all $k \geq \mu$ we have*
$$\mathbb{P}(Z \geq k) \leq e^{-\mu H(k/\mu)},$$

(ii) *For all $k \leq \mu$ we have*
$$\mathbb{P}(Z \leq k) \leq e^{-\mu H(k/\mu)},$$

where $H(x) := x\ln x - x + 1$. ■

An easy inequality on the Poisson distribution that we will use below is as follows. For completeness we spell out the short proof.

Lemma 2.6 *Let Z be a Poisson random variable. Then $\mathbb{P}(Z \geq k) \leq (\mathbb{E}Z)^k$, for all $k \in \mathbb{N}$*

Proof: Using Markov's inequality we have

$$\mathbb{P}(Z \geq k) = \mathbb{P}(Z(Z-1)\cdots(Z-k+1) \geq 1) \leq \mathbb{E}\big[Z(Z-1)\dots(Z-k+1)\big] = (\mathbb{E}Z)^k,$$

where we also used the well-known, elementary fact that the k-th factorial moment of the Poisson equals the k-th power of its first moment. ■

The *total variational distance* between two integer-valued random variables X, Y is defined as:

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

Let $g : \mathbb{R}^d \to [0, \infty)$ be a bounded, measurable function. A Poisson process on $\mathbb{R}^d$ with intensity g is a random set of points $\mathcal{P} \subseteq \mathbb{R}^d$ with the properties that:

($\mathcal{PP}$-1) For every measurable $A \subseteq \mathbb{R}^d$, we have $\mathcal{P}(A) \stackrel{d}{=} \mathrm{Po}\left(\int_A g(x)\mathrm{d}x\right)$, where $\mathcal{P}(A) := |A \cap \mathcal{P}|$ denotes the number of points of $\mathcal{P}$ that fall into A;

($\mathcal{PP}$-2) If $A_1, \dots, A_n$ are disjoint and measurable, then $\mathcal{P}(A_1), \dots, \mathcal{P}(A_n)$ are independent.

An important special case is when $\int_{\mathbb{R}^d} g(x)\mathrm{d}x < \infty$. In this case we can write $g = \lambda \cdot f$, with f (the probability density of) an absolute continuous probability measure. In this case the *Poisson process* $\mathcal{P}$ can also be generated as follows. Let $X_1, X_2, \dots \in \mathbb{R}^d$ be an infinite supply of random points, i.i.d. distributed with probability density f; and let $N \stackrel{d}{=} \mathrm{Po}(\lambda)$ be independent of $X_1, X_2, \dots$. Then $\mathcal{P} \stackrel{d}{=} \{X_1, \dots, X_N\}$. A proof of this folklore result and more background on Poisson point processes can for instance be found in [11].

It is often useful to consider a *Poissonized* version of the random geometric graph. Here we mean the following. Let us take $X_1, X_2, \dots$ i.i.d. uniform at random on the unit square, and we let $N \stackrel{d}{=} \mathrm{Po}(n)$ be independent of $X_1, X_2, \dots$. Following Penrose [21], we will write

$$\mathcal{P}_n := \{X_1, \dots, X_N\}. \tag{3}$$

(Thus $\mathcal{P}_n$ is a Poisson process with intensity n on the unit square and intensity 0 elsewhere.) The *Poisson random geometric graph* is defined as

$$G_{\mathcal{P}}(n,r) := G(\mathcal{P}_n; r).$$

The properties **($\mathcal{PP}$-1)** and **($\mathcal{PP}$-2)** above make $G_{\mathcal{P}}(n,r)$ often slightly easier to deal with than the ordinary random geometric graph. For notational convenience (and again following Penrose [21]) we set:

$$\mathcal{X}_n := \{X_1, \dots, X_n\}. \tag{4}$$

The usual random geometric graph $G(n,r) = G(\mathcal{X}_n; r)$ is sometimes also called the *binomial random geometric graph*. By defining both $G(n,r)$ and $G_{\mathcal{P}}(n,r)$ on the same set of points $X_1, X_2, \dots$ we get an explicit coupling which often helps to transfer results from the Poissonized setting to the original setting.

The next theorem is especially useful for dealing with the subgraph counts and counts of other "small substructures" in the Poissonized random geometric graph. The statement and its proof are almost identical to Theorem 1.6 of [21], but for completeness we include a proof in Appendix A.

Theorem 2.7 *Let $\mathcal{P}_n$ be as in* (3)*, and let $h(a_1, \ldots, a_k; A)$ be a bounded measurable function defined on all tuples $(a_1, \ldots, a_k; A)$ with $A \subseteq \mathbb{R}^2$ finite and $a_1, \ldots, a_k \in A$. Let us write*

$$Z := \sum_{\substack{a_1,\ldots,a_k \in \mathcal{P}_n, \\ a_1,\ldots,a_k \text{ distinct}}} h(a_1, \ldots, a_k; \mathcal{P}_n).$$

Then

$$\mathbb{E}Z = n^k \cdot \mathbb{E}h(Y_1, \ldots, Y_k; \{Y_1, \ldots, Y_k\} \cup \mathcal{P}_n),$$

where $Y_1, \ldots, Y_k$ are i.i.d. uniform on the unit square, and are independent of $\mathcal{P}_n$.

The previous theorem can be used to count "small substructures" in the Poisson setting. We are primarily interested in the binomial random graph. The next lemma is useful for transferring results from the Poisson random geometric graph to the binomial random graph. The proof is relatively standard. For completeness we provide the proof in Appendix B, since it is not available in the literature as far as we are aware. Here and in the rest of the paper, $B(x, r) := \{y \in \mathbb{R}^2 : \|x - y\| < r\}$ denotes the ball of radius r around the point x.

Lemma 2.8 *Let $h_n(v_1, \ldots, v_k; V)$ be a sequence of $\{0, 1\}$-valued, measurable functions defined on all tuples $(v_1, \ldots, v_k; V)$ with $V \subseteq \mathbb{R}^2, v_1, \ldots, v_k \in V$ and set*

$$Z_n := \sum_{v_1,\ldots,v_k \in \mathcal{P}_n} h_n(v_1, \ldots, v_k; \mathcal{P}_n), \quad \tilde{Z}_n := \sum_{v_1,\ldots,v_k \in \mathcal{X}_n} h_n(v_1, \ldots, v_k; \mathcal{X}_n),$$

with $\mathcal{P}_n$ as in (3) *and $\mathcal{X}_n$ as in* (4)*. Suppose that $\mathbb{E}Z_n = O(1)$ and that there exists a sequence $(r_n)_n$ such that $\pi n r_n^2 = o(\sqrt{n})$ and the value of $h_n(v_1, \ldots, v_k; V)$ does not depend on $V \setminus (B(v_1; r_n) \cup \cdots \cup B(v_k; r_n))$. Then $Z_n = \tilde{Z}_n$ whp.*

We will need two results of Penrose [21] that were proved using Poissonization. The first of these two results is on the occurrence of small subgraphs. For H a graph, we shall denote by $N(H) = N(H; n, r)$ the number of *induced* subgraphs of $G(n, r)$ that are isomorphic to H. For H a connected geometric graph on k vertices, let us denote

$$\mu(H) := \frac{1}{k!} \int_{\mathbb{R}^2} \cdots \int_{\mathbb{R}^2} 1_{\{G(0,x_1,\ldots,x_{k-1};1) \cong H\}} \mathrm{d}x_1 \ldots \mathrm{d}x_{k-1} \tag{5}$$

(Here $G \cong H$ means that G and H are isomorphic and 1_A is the indicator function of the set A, in our case the set of all $(x_1, \ldots, x_{k-1}) \in \left(\mathbb{R}^2\right)^{k-1}$ that satisfy $G(0, x_1, \ldots, x_{k-1}; 1) \cong H$.) It can be seen that, since H is a connected geometric graph, $0 < \mu(H) < \infty$. The following is a restriction of Corollary 3.6 in [21] to the special case of the uniform distribution on the unit square and the Euclidean norm.

Theorem 2.9 (Penrose [21]) *For $k \in \mathbb{N}$, let $H_1, \ldots, H_m$ be connected, non-isomorphic geometric graphs on $k \geq 2$ vertices. Let $(r_n)_n$ be a sequence of positive numbers satisfying $r_n = \alpha \cdot n^{-\frac{k}{2(k-1)}}$ for some constant $\alpha > 0$. Then*

$$(N(H_1), \ldots, N(H_m)) \xrightarrow{d} (Z_1, \ldots, Z_m),$$

where $Z_1, \ldots, Z_m$ are independent Poisson random variables with means $\mathbb{E}Z_i = \alpha^{2(k-1)} \cdot \mu(H_i)$.

We shall also need a result on the minimum degree of the random geometric graphs. The following is a reformulation of Theorem 8.4 in [21], restricted to the case of the Euclidean metric in two dimensions.

Theorem 2.10 (Penrose [21]) *Let $(r_n)_n$ be a sequence of positive numbers. The following hold for the random geometric graph $G(n, r_n)$:*

(i) *If* $\pi n r_n^2 = \ln n + x + o(1)$ *for some fixed* $x \in \mathbb{R}$ *then*

$$\lim_{n\to\infty} \mathbb{P}[G(n, r_n) \text{ has min. deg.} \geq 1] = e^{-e^{-x}}.$$

(ii) *If* $\pi n r_n^2 = \ln n + \ln\ln n + x + o(1)$ *for some fixed* $x \in \mathbb{R}$ *then*

$$\lim_{n\to\infty} \mathbb{P}[G(n, r_n) \text{ has min. deg.} \geq 2] = e^{-(e^{-x}+\sqrt{\pi e^{-x}})}.$$

(iii) *If* $\pi n r_n^2 = \ln n + (2k-3)\ln\ln n + 2\ln((k-1)!) + 2x + o(1)$ *for some fixed* $x \in \mathbb{R}$ *and* $k > 2$ *then*

$$\lim_{n\to\infty} \mathbb{P}[G(n, r_n) \text{ has min. deg.} \geq k] = e^{-e^{-x}}.$$

Later on we will do some reverse-engineering of Theorem 2.10. For this purpose it is convenient to state also the following intermediate result that was part of the proof of Theorem 2.10.

Lemma 2.11 ([21]) *Let* $(r_n)_n$ *be such that* $\pi n r_n^2 = \ln n + \ln\ln n + x + o(1)$ *for some* $x \in \mathbb{R}$*, and let* W_n *denote the number of vertices of degree exactly one in* $G(n, r_n)$*. Then*

$$EW_n \to e^{-x} + \sqrt{\pi e^{-x}}.$$

2.2 Geometric preliminaries

This section begins with two elementary results about geometric graphs. The remainder of the section is devoted to approximating the area of intersecting regions in $[0,1]^2$. We start with a standard elementary result. Because we are not aware of a proof anywhere in the literature, we provide a proof in Appendix C.

Lemma 2.12 *Let* G *be a connected geometric graph. Then* G *has a spanning tree of maximum degree at most five.*

Note that Lemma 2.12 is best possible since $K_{1,5}$ is a connected geometric graph. We also need the following observation. We leave the straightforward proof to the reader.

Lemma 2.13 *Let* $G = (x_1, \dots, x_n; r)$ *be a geometric graph and suppose that* $x_i x_j, x_a x_b \in E(G)$ *are two edges that do not share endpoints. If the line segments* $[x_i, x_j]$ *and* $[x_a, x_b]$ *cross then at least one of the edges* $x_i x_a, x_i x_b, x_j x_a, x_j x_b$ *is also in* $E(G)$*.* ■

We now turn to approximating areas in $[0,1]^2$. First, we give an expression for the area of the difference between two disks of the same radius. We leave the proof, which is straightforward trigonometry, to the reader.

Lemma 2.14 *For* $x, y \in \mathbb{R}^2$ *we have, provided* $d := \|x - y\| \leq 2r$*:*

$$\text{area}(B(x; r) \setminus B(y; r)) = \pi r^2 - 2r^2 \arccos(d/2r) + dr\sqrt{1 - (d/2r)^2}.$$

■

Using the fact that $\frac{\pi}{2}(1-x) \leq \arccos(x) \leq \frac{\pi}{2} - x$ for $0 \leq x \leq 1$, and the Taylor approximations $\arccos(x) = \frac{\pi}{2} - x + O(x^2)$ and $\sqrt{1-x} = 1 - x/2 + O(x^2)$, we get the following straightforward consequence of Lemma 2.14 that will be useful to us later:

Corollary 2.15 *For* $x, y \in \mathbb{R}^2$ *we have, provided* $d := \|x - y\| \leq 2r$*:*

$$dr \leq \text{area}(B(x; r) \setminus B(y; r)) \leq 4dr, \tag{6}$$

and

$$\text{area}(B(x; r) \setminus B(y; r)) = 2dr - O(d^2), \tag{7}$$

as $d \downarrow 0$*.* ■

We also need expressions for the area of the intersection of a disk of radius r with the unit square. Again the proof is straightforward trigonometry that we leave to the interested reader.

Lemma 2.16 *Suppose that $r < \frac{1}{2}$ and $x \in [0, r) \times (r, 1-r)$. Writing h for the first coordinate of x, we have:*

$$\operatorname{area}(B(x;r) \cap [0,1]^2) = \pi r^2 - \arccos(h/r) \cdot r^2 + hr\sqrt{1-(h/r)^2}.$$

■

Again using $\frac{\pi}{2}(1-x) \leq \arccos(x) \leq \frac{\pi}{2} - x$ for $0 \leq x \leq 1$, and the Taylor approximations $\arccos(x) = \frac{\pi}{2} - x + O(x^2)$ and $\sqrt{1-x} = 1 - x/2 + O(x^2)$, we get:

Corollary 2.17 *Suppose that $r < \frac{1}{2}$ and $x \in [0, r) \times (r, 1-r)$. Writing h for the first coordinate of x, we have:*

$$\frac{\pi}{2}r^2 + hr \leq \operatorname{area}(B(x;r) \cap [0,1]^2) \leq \frac{\pi}{2}r^2 + 2hr, \tag{8}$$

for $0 \leq h \leq r$, and

$$\operatorname{area}(B(x;r) \cap [0,1]^2) = \frac{\pi}{2}r^2 + 2hr - O(h^2), \tag{9}$$

as $h \downarrow 0$. ■

We need one more geometric approximation, that combines the bounds from (7) and (9).

Lemma 2.18 *Suppose that $x \in [0, r) \times (2r, 1-2r), y \in B(x;r)$, with y to the right of x. Let h denote the first coordinate of x, and let α, d be defined by $v := y - x = (d\cos\alpha, d\sin\alpha)$ (see Figure 1). Then*

$$\operatorname{area}([0,1]^2 \cap (B(x;r) \cup B(y;r))) = \frac{\pi}{2}r^2 + 2hr + (1+\cos\alpha)dr + O((d+h)^2), \tag{10}$$

and

$$\operatorname{area}([0,1]^2 \cap (B(y;r) \setminus B(x;r))) = (1+\cos\alpha)dr + O(d(d+h)) \tag{11}$$

where the error terms are uniform over all $-\pi/2 \leq \alpha \leq \pi/2$.

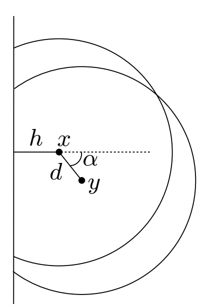

Figure 1: Computing the area of $(B(x;r) \cup B(y;r)) \cap [0,1]^2$.

Proof: Let ℓ_1 denote the vertical line through x, and let ℓ_2 denote the vertical line through the midpoint of $[x, y]$. Let A_1 denote the part of $B(x;r) \cup B(y;r)$ between the y-axis and ℓ_1; let A_2 denote the part of $B(x;r) \cup B(y;r)$ between ℓ_1 and ℓ_2; and let A_3 denote the part of $B(x;r) \cup B(y;r)$ to the right of ℓ_2.

By symmetry, and (7), we have.

$$\operatorname{area}(A_3) = \frac{1}{2}\operatorname{area}(B(x;r) \cup B(y;r)) = \frac{\pi}{2}r^2 + dr + O(d^2).$$

Observe that A_2 is contained in a rectangle of sides $d \cdot \frac{1}{2}\cos(\alpha)$ and $2r + d$. Hence, $\text{area}(A_2) \leq dr\cos(\alpha) + O(d^2)$. On the other hand, A_2 contains the portion of $B(x;r)$ between ℓ_1 and ℓ_2. Thus, using (9) we see that in fact

$$\text{area}(A_2) = dr\cos\alpha + O(d^2).$$

Similarly, A_1 is contained inside a rectangle of sides h and $2r + d$; and it contains the part of $B(x;r)$ between the y-axis and ℓ_1, giving:

$$\text{area}(A_1) = 2hr + O(h(d + h)).$$

Combining the three expressions proves (10).

Now notice that, if $B^+(x;r)$ denotes the portion of $B(x;r)$ to the right of ℓ_1:

$$\begin{array}{rcl} \text{area}([0,1]^2 \cap (B(y;r) \setminus B(x;r))) & \geq & \text{area}((A_2 \cup A_3) \setminus B^+(x;r)) \\ & = & \text{area}(A_2 \cup A_3) - \frac{\pi}{2}r^2 \\ & = & (1 + \cos\alpha)dr + O(d^2). \end{array}$$

On the other hand it is not hard to see that the portion of $B(y;r) \setminus B(x;r)$ that lies between the y-axis and ℓ_1 has area at most $h \cdot d$, so that

$$\begin{array}{rcl} \text{area}([0,1]^2 \cap (B(y;r) \setminus B(x;r))) & \leq & \text{area}((A_2 \cup A_3) \setminus B^+(x;r)) + h \cdot d \\ & = & (1 + \cos\alpha)dr + O(d(d + h)), \end{array}$$

proving (11). ∎

3 The structure of $G(n, r_n)$ near the connectivity threshold

This section contains a number of observations of varying technical difficulty that describe the structure of the random geometric graph $G(n, r_n)$ when r_n is such that the probability of Maker-win for one of the games under consideration is nontrivial. Intuitively speaking, we characterize regions of $G(n, r_n)$ as being dense or sparse. Most of the graph is dense. The sparse regions are of small diameter, and they are well-separated. Equally as important, we show that the dense region surrounding a sparse region contains enough points to enable Maker to overcome this local bottleneck. Let us begin.

Let $G = (V, r)$ be a geometric graph, where $V = \{x_1, \ldots, x_n\} \subset [0,1]^2$. We consider the structure of G with respect to a partition of $[0,1]^2$ into small squares. We introduce a vocabulary for describing the density of vertices in each square. We also categorize the vertices themselves, depending on whether they are in dense squares or sparse squares. In addition, we will pay special attention to vertices in dense squares that are also close to vertices in sparse squares.

Let $\eta > 0$ be arbitrarily small and let $m \in \mathbb{N}$ be such that $q(m) := 1/m = \eta r$. Let $\mathcal{D} = \mathcal{D}(m)$ denote the *dissection* of $[0,1]^2$ into squares of side length $q(m)$. We will call these squares *cells*. For $K \in \mathbb{N}$, we define a $K \times K$ *block of cells* in the obvious way, see Figure 2. Given $T > 0$ and $V \subseteq [0,1]^2$, we call a cell $c \in \mathcal{D}$ *good* with respect to T, V if $|c \cap V| \geq T$ and *bad* otherwise. When the choice of T and V is clear from the context we will just speak of good and bad. Let $\Gamma = \Gamma(V, m, T, r)$ denote the graph whose vertices are the good cells of $\mathcal{D}(m)$, with an edge $cc' \in E(\Gamma)$ if and only if the lower left corners of c, c' have distance at most $r - q\sqrt{2}$. (Note that this way, any $x \in c$ and $y \in c'$ have distance $\|x - y\| \leq r$.) We will usually just write Γ when the choice of V, m, T, r is clear from the context. Let us denote the components of Γ by $\Gamma_1, \Gamma_2, \ldots$ where Γ_i has at least as many cells as Γ_{i+1} (ties are broken arbitrarily). For convenience we will also write $\Gamma_{\max} = \Gamma_1$. We will often be a bit sloppy and identify Γ_i with the union of its cells, and speak of $\text{diam}(\Gamma_i)$ and the distance between Γ_i and Γ_j and so forth.

Let us call a point $v \in V$ *safe* if there is a cell $c \in \Gamma_{\max}$ such that $|B(v;r) \cap V \cap c| \geq T$. (I.e. in the geometric graph $G(V;r)$, the point v has at least T neighbours inside c.) If v is not safe and there is a good cell $c \in \Gamma_i$, $i \geq 2$, such that $|B(v;r) \cap V \cap c| \geq T$, we say that v is *risky*. Otherwise,

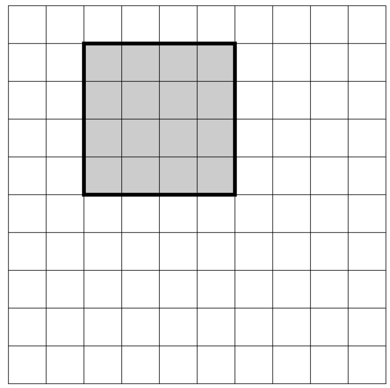

Figure 2: The dissection $\mathcal{D}(10)$, with a 4×4 block of cells highlighted.

if v is neither safe nor risky, we call v *dangerous*. Every vertex in a cell of $\Gamma_{\max}$ is safe. Every vertex in a cell of Γ_i for $i \geq 2$ is risky. Vertices in bad cells can be safe, risky or dangerous.

For $i \geq 2$ we let Γ_i^+ denote the set of all points of V in cells of Γ_i, together with all risky points v that satisfy $|B(v; r) \cap V \cap c| \geq T$ for at least one $c \in \Gamma_i$.

The following is a list of desirable properties that we would like V and $\Gamma(V, m, T, r)$ to have:

(str-1) $\Gamma_{\max}$ contains more than $0.99 \cdot |\mathcal{D}|$ cells;

(str-2) $\operatorname{diam}(\Gamma_i^+) < r/100$ for all $i \geq 2$;

(str-3) If $u, v \in V$ are dangerous then either $\|u - v\| < r/100$ or $\|u - v\| > r \cdot 10^{10}$;

(str-4) For all $i \neq j \geq 2$ the distance between Γ_i^+ and Γ_j^+ is at least $r \cdot 10^{10}$;

(str-5) If $v \in V$ is dangerous and $i \geq 2$ then the distance between v and $\Gamma_i{}^+$ is at least $r \cdot 10^{10}$;

(str-6) If $c, c' \in \Gamma_{\max}$ are two cells at Euclidean distance at most $10r$, then there is a path in $\Gamma_{\max}$ between them of (graph-) length at most 10^5.

See Figure 3 for a schematic of a geometric graph that satisfies **(str-1)**–**(str-6)**.

3.1 The Dissection Lemma

For $n \in \mathbb{N}$ and $\eta > 0$ a constant, let us define

$$m = m_n := \left\lceil \sqrt{\frac{n}{\eta^2 \ln n}} \right\rceil. \tag{12}$$

The goal of this section is to prove that if $r_n = \ln n + o(\ln n)$ then **(str-1)**–**(str-6)** hold for $\Gamma(\mathcal{X}_n, m_n, T, r_n)$ *whp*. This is stated formally in the Dissection Lemma 3.3 below. First, we prove two intermediate lemmas.

Lemma 3.1 *Let $\eta, \varepsilon, T, K > 0$ be arbitrary but fixed, where η, ε are small and T, K are large. Let m be given by* (12) *and let $\mathcal{X}_n$ be as in* (4)*. The following hold* whp *for $\mathcal{D}(m_n)$ with respect to T and $\mathcal{X}_n$:*

(i) *Out of every $K \times K$ block of cells, the area of the bad cells inside the block is at most $(1 + \varepsilon) \ln n / n$;*

(ii) *Out of every $K \times K$ block of cells touching the boundary of the unit square, the area of the bad cells inside the block is at most $(1 + \varepsilon) \ln n / 2n$.*

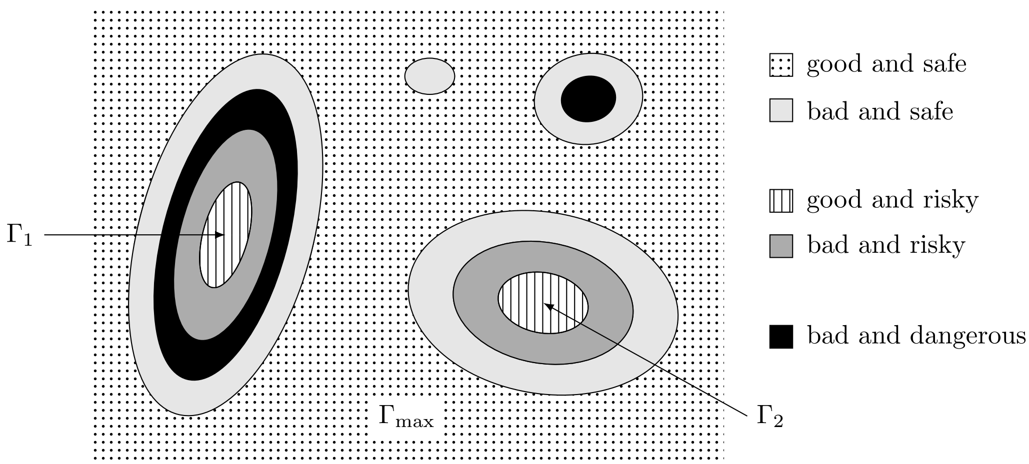

Figure 3: A schematic of part of a geometric graph that satisfies **(str-1)**–**(str-6)**. Cells are characterized as good or bad. The smaller components Γ_1, Γ_2 of Γ are surrounded by bad cells. Vertices are also characterized as safe, risky or dangerous.

(iii) *Every $K \times K$ block of cells touching a corner contains only good cells.*

Proof: The number of $K \times K$ blocks of cells is $(m - K + 1)^2 = \Theta(n/\ln n)$. The number of sets of cells of area $> (1+\varepsilon)\ln n/n$ inside a given $K \times K$ block is at most $2^{K^2} = O(1)$. If A is a union of cells inside some $K \times K$ block then the probability that all its cells are bad is at most

$$\mathbb{P}(\mathrm{Po}(n \cdot \mathrm{area}(A)) < TK^2) \le \exp\left[-n \cdot \mathrm{area}(A) \cdot H\left(\frac{TK^2}{n \cdot \mathrm{area}(A)}\right)\right],$$

using Lemma 2.5, where $H(x) = x \ln x - x + 1$. If $\mathrm{area}(A) \ge (1+\varepsilon)\ln n/n$ then $\frac{TK^2}{n \cdot \mathrm{area}(A)} \le \frac{TK^2}{(1+\varepsilon)\ln n} = o(1)$ so that $H\left(\frac{TK^2}{n\cdot\mathrm{area}(A)}\right) = 1 + o(1)$ and hence

$$\begin{aligned}\mathbb{P}(\mathrm{Po}(n \cdot \mathrm{area}(A)) < TK^2) &\le \exp[-(1+\varepsilon)\ln n \cdot (1+o(1))] \\ &= n^{-(1+\varepsilon)+o(1)}.\end{aligned}$$

Thus, the probability that there is a $K \times K$ block such that the area of the bad cells inside it is at least $(1+\varepsilon)\ln n/n$ is bounded above by:

$$m^2 \cdot 2^{K^2} \cdot n^{-(1+\varepsilon)+o(1)} = n^{-\varepsilon+o(1)} = o(1).$$

This proves part **(i)**.

The number of $K \times K$ blocks touching a side of the unit square is at most $4m = \Theta(\sqrt{n/\ln n})$, and if A is a union of cells with $\mathrm{area}(A) \ge (1+\varepsilon)\ln n/2n$, consisting of no more than K^2 cells, then the probability that all its cells are bad is at most

$$\begin{aligned}\mathbb{P}(\mathrm{Po}(n \cdot \mathrm{area}(A)) < TK^2) &\le \exp[-(1+\varepsilon)\tfrac{1}{2}\ln n \cdot (1+o(1))] \\ &= n^{-(1+\varepsilon)/2+o(1)}.\end{aligned}$$

Hence, the probability that there is a $K \times K$ block of cells touching a side of the unit square such that the union of the bad cells inside the block has area at least $(1+\varepsilon)\ln n/2n$ is at most:

$$4m \cdot 2^{K^2} \cdot n^{-(1+\varepsilon)/2+o(1)} = n^{-\varepsilon/2+o(1)} = o(1).$$

This proves part **(ii)**.

Finally, there are only four $K \times K$ blocks that touch a corner. The probability that a cell c is bad is at most

$$\mathbb{P}(\mathrm{Po}(n \cdot \mathrm{area}(c)) < T)) \leq \exp[-\Omega(\ln n) \cdot (1 + o(1))] = o(1).$$

Hence, the probability that there is a bad cell inside one of the $K \times K$ blocks of cells touching a corner of the unit square is at most:

$$4 \cdot K^2 \cdot o(1) = o(1).$$

This proves part **(iii)**. ■

Recall that a *halfplane* is one of the two connected components of $\mathbb{R}^2 \setminus \ell$ with ℓ a line. A *halfdisk* is the intersection of a disk $B(x,r)$ with a halfplane whose defining line goes through the center x. Also recall that a set $A \subseteq \mathbb{R}^2$ is a *Boolean combination* of the sets $A_1, \dots, A_n \subseteq \mathbb{R}^2$ if A can be constructed from $A_1, \dots, A_n$ by means of any number of compositions of the operations intersection, union and complement. Our next lemma shows that the area of A is well-approximated by the area of the cells that are entirely contained in A.

Lemma 3.2 *There exists a universal constant $C > 0$ such that the following holds for every constant $\eta > 0$ and all sufficiently large $n \in \mathbb{N}$. Let m_n be given by* (12)*, let $A \subseteq [0,1]^2$ be a Boolean combination of at most* 1000 *halfdisks with radii at most* $1000 \cdot \sqrt{\ln n/n}$*, and let*

$$A' := \bigcup \{c \in \mathcal{D}(m_n) : c \subseteq A\},$$

denote the union of all cells of $\mathcal{D}(m_n)$ that are contained in A. Then

$$\mathrm{area}(A') \geq \mathrm{area}(A) - C \cdot \eta \cdot \left(\frac{\ln n}{n}\right),$$

for n sufficiently large.

Proof: Let A, A' be as in the statement of the lemma. We use ∂A to denote the boundary of A. Let us define $q := 1/m$ and

$$A'' := \{z \in \mathbb{R}^2 : B(z; q\sqrt{2}) \subseteq A\} = A \setminus \left(B(0; q\sqrt{2}) + \partial A\right).$$

Then clearly $A'' \subseteq A' \subseteq A$. If A is a boolean combination of the halfdisks $A_1, \dots, A_k$ then clearly $\partial A \subseteq \partial A_1 \cup \dots \cup \partial A_k$. If A_i is a halfdisk of radius $r \leq 1000 \cdot \sqrt{\ln n/n}$, then it is easily seen that

$$\begin{aligned} \mathrm{area}\left(\partial A_i + B(0; q\sqrt{2})\right) &\leq \tfrac{1}{2}\left(\pi(r + q\sqrt{2})^2 - \pi(r - q\sqrt{2})^2\right) + 2r \cdot 2q\sqrt{2} \\ &= r \cdot q(2\pi\sqrt{2} + 4\sqrt{2}) \\ &\leq 10^5 \cdot \eta \cdot \left(\tfrac{\ln n}{n}\right). \end{aligned}$$

This gives

$$\begin{aligned} \mathrm{area}(A') &\geq \mathrm{area}(A'') \\ &\geq \mathrm{area}(A) - \textstyle\sum_{i=1}^k \mathrm{area}(\partial A_i + B(0, q\sqrt{2}) \\ &\geq \mathrm{area}(A) - 1000 \cdot 10^5 \cdot \eta \cdot \left(\tfrac{\ln n}{n}\right), \end{aligned}$$

which proves the lemma with $C := 10^8$. ■

We can now prove the Dissection Lemma: our random geometric graph satisfies **(str-1)**–**(str-6)** with high probability. It will be convenient in the proof to introduce two slightly weaker properties that will be proved before their original counterpart:

(str0-2) $\mathrm{diam}(\Gamma_i) < r/100$ for all $i \geq 2$;

(str0-4) For all $i \neq j \geq 2$ the distance between Γ_i and Γ_j is at least $r \cdot 10^{10}$;

Lemma 3.3 (Dissection Lemma) *Let $T > 0$ be arbitrary but fixed. For $\eta > 0$ sufficiently small, the following holds. Let m_n be given by* (12), *let $\mathcal{X}_n$ be as in* (4), *and let r_n be such that $\pi n r_n^2 = \ln n + o(\ln n)$. Then* **(str-1)**–**(str-6)** *hold for* $\Gamma(\mathcal{X}_n, m_n, T, r_n)$ whp.

Proof: We can assume that the conclusion of Lemma 3.1 holds with $\varepsilon := 10^{-5}$ and $K := \lceil (1+\varepsilon) \cdot 10^{100}/\eta^2 \rceil$.
Proof of (str-1): Consider a $K \times K$ block of cells $\mathcal{B}$. By Lemma 3.1, it contains at most

$$N := \frac{(1+\varepsilon)\ln n}{q^2 n} = \frac{1+\varepsilon+o(1)}{\eta^2}$$

bad cells, since $q = (1+o(1))\sqrt{\eta^2 \ln n/n}$.

Thus, at least $K - N > 0.99K$ rows of the block do not contain any bad cell. The cells of such a bad-cell-free row clearly belong to the same component of Γ (provided $q < r - q\sqrt{2}$ which is certainly true for η sufficiently small). Since there is also at least one bad-cell-free column, we see that all the bad-cell-free rows of the block belong to the same component of Γ, and this component contains at least 99 percent of the cells in the block. Let $\mathcal{C}(\mathcal{B})$ denote the component of Γ that contains more than $0.99K^2$ cells of the block $\mathcal{B}$.

Let us now consider two $K \times K$ blocks $\mathcal{B}_1, \mathcal{B}_2$, where $\mathcal{B}_2$ is obtained by shifting $\mathcal{B}_1$ to the left by one cell. Then there are at least $K - 2N > 0$ rows where both blocks don't have any bad cells. This shows that the component $\mathcal{C}(\mathcal{B}_1) = \mathcal{C}(\mathcal{B}_2)$. Clearly the same thing is true if B_2 is obtained by shifting B_1 right, down or up by one cell.

Now let $\mathcal{B}_1, \mathcal{B}_2$ be two arbitrary $K \times K$ blocks. Since we can move from $\mathcal{B}_1$ to $\mathcal{B}_2$ by repeatedly shifting left, right, down or up, we see that in fact $\mathcal{C}(\mathcal{B}_1) = \mathcal{C}(\mathcal{B}_2)$ for *any* two blocks $\mathcal{B}_1, \mathcal{B}_2$. This proves that there is indeed a component of Γ that contains more than $0.99 \cdot |\mathcal{D}|$ cells.

Proof of (str0-2): Let c be a cell that contains at least one point of Γ_i with $i \geq 2$. Let us first assume that c is at least $K/2$ cells away from the boundary of $[0,1]^2$. In this case we can center a $K \times K$ block of cells $\mathcal{B}$ on c (If K is odd we place $\mathcal{B}$ so that c is the middle cell, and if K is even we place $\mathcal{B}$ so that a corner of c is the center of $\mathcal{B}$). Reasoning as in the proof of **(str-1)**, at least one row below, one row above, at least one column to the left and at least one column to the right of c are bad-cell-free. The cells in these bad-cell-free rows and columns must belong to $\Gamma_{\max}$ (by the proof of **(str-1)**). Therefore $\Gamma_i + B(0, r - q\sqrt{2})$ is completely contained in the block $\mathcal{B}$. (Here $A + B := \{a + b : a \in A, b \in B\}$ denotes the Minkowski sum.) Let p_L be a leftmost point of Γ_i, let p_R be a rightmost point, let p_B be a lowest point and let p_T be a highest point of Γ_i. (These points need not be unique, but this does not pose a problem for the remainder of the proof.) Let D_L denote the halfdisk

$$D_L := B(p_L; r - 2q\sqrt{2}) \cap \{z \in \mathbb{R}^2 : z_x < (p_L)_x\}.$$

Then D_L cannot contain any good cell, because that would contradict that p_L is the leftmost point of Γ_i. Similarly we define the halfdisks D_R, D_B, D_T and observe that each of them cannot contain any good cell. (see Figure 4).

Now let us set $A := D_L \cup D_R \cup D_B \cup D_T$. Assuming that $\operatorname{diam}(\Gamma_i) \geq r/100$ we have either $\|p_L - p_R\| \geq r/100\sqrt{2}$ or $\|p_T - p_B\| \geq r/100\sqrt{2}$. We can assume without loss of generality $\|p_L - p_R\| \geq r/100\sqrt{2}$. Observe that, since $(p_L)_x \leq (p_T)_x \leq (p_R)_x$, we have

$$\operatorname{area}(D_T \setminus (D_L \cup D_R)) \geq \left(\frac{r/100\sqrt{2}}{2(r - 2s\sqrt{2})} \right) \cdot \operatorname{area}(D_T),$$

and similarly for $\operatorname{area}(D_B \setminus (D_L \cup D_R))$. It follows that

$$\operatorname{area}(A) \geq \pi(r - 2q\sqrt{2})^2(1 + \frac{1}{200\sqrt{2}}).$$

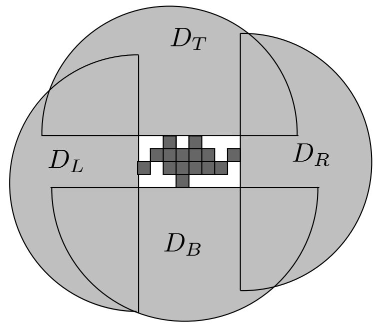

Figure 4: Dropping halfdisks onto a non-giant component of Γ.

Thus, if A' denotes the union of all cells that are contained in A then, using Lemma 3.2, we have

$$\begin{aligned} \operatorname{area}(A') &\geq \operatorname{area}(A) - C\eta\left(\tfrac{\ln n}{n}\right) \\ &\geq \pi r^2(1-6\eta)^2(1+\tfrac{1}{200\sqrt{2}}) - C\eta\left(\tfrac{\ln n}{n}\right) \\ &\geq (1+10^{-5})\left(\tfrac{\ln n}{n}\right), \end{aligned}$$

for n large enough, provided η was chosen sufficiently small. Since $A \subseteq \mathcal{B}$ this contradicts Lemma 3.1(i). Hence $\operatorname{diam}(\Gamma_i) < r/100$.

Let us now consider the case when some cell c that is within $K/2$ cells of a side of $[0,1]^2$ contains a point of Γ_i with $i \geq 2$. If c is within $K/2$ cells of two sides of $[0,1]^2$ then it is in fact in a $K \times K$ block touching a corner. Since there are no bad cells in such a block, we have from the proof of **(str-1)** that $c \in \Gamma_{\max}$ and hence $\Gamma_{\max} \cap \Gamma \neq \emptyset$, a contradiction.

Hence we can assume c is within $K/2$ cells of one of the sides, but more more than $K/2$ cells away from all other sides. By symmetry considerations, we can assume the closest side is the y-axis. Let p_L, p_R, p_T, p_B and D_L, D_R, D_T, D_B and $A := D_L \cup D_R \cup D_B \cup D_T$ be defined as before. Observe that $D_R \subseteq [0,1]^2$ and that at least half the area of D_T, D_B falls inside $[0,1]^2$. Thus

$$\operatorname{area}(A \cap [0,1]^2) \geq \frac{1}{2}\operatorname{area}(A).$$

Hence, if $A' \subseteq A$ again denotes the union of the cells contained in A:

$$\begin{aligned} \operatorname{area}(A') &\geq \operatorname{area}(A \cap [0,1]^2) - C\eta\left(\tfrac{\ln n}{n}\right) \\ &\geq \tfrac{1}{2}\operatorname{area}(A) - C\eta\left(\tfrac{\ln n}{n}\right) \\ &\geq (1+10^{-5})\left(\tfrac{\ln n}{n}\right)/2, \end{aligned}$$

for η sufficiently small (using Lemma 3.2 to get the first line, and previous computations to get the last line). But this contradicts part **(ii)** of Lemma 3.1.

It follows that $\operatorname{diam}(\Gamma_i) < r/100$, as required.

Proof of (str-3): The proof is analogous to the proof of **(str0-2)**. If u, v are two dangerous points with $\|u - v\| < r \cdot 10^{10}$ then we let p_L, p_R, p_T, p_B be the leftmost, rightmost, top and bottom points of $A = \{u, v\}$ and continue as before to find a contradiction to Lemma 3.1.

Proof of (str0-4): Again the proof is analogous to the proof of **(str0-2)**. If Γ_i, Γ_j with $2 \leq i < j$ are two components of Γ and the distance between them is at most $r \cdot 10^{10}$, then we take p_L, p_R, p_T, p_B to be the leftmost, rightmost, top and bottom points of $A = \Gamma_i \cup \Gamma_j$ and continue as before.

Proofs of (str-2),(str-4): We run through the proofs replacing Γ_i by Γ_i^+. It will still be possible to claim that D_L contains no good cells. Before, it would seem possible that such a cell was in $\Gamma_j, j \neq i$, disallowing our contradiction. Now we know that the components of Γ are too far apart for this to happen.

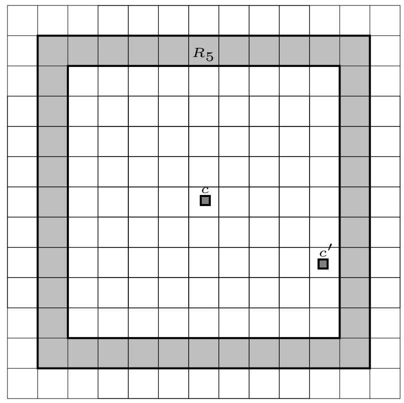

Figure 5: The squares $S_{x,y}$ and R_5.

Proof of (str-5): Once again the proof is analogous to the proof of **(str0-2)**. If v is dangerous and Γ_i with $i \geq 2$ is a component of Γ and the distance from v to Γ_i^+ is at most $r \cdot 10^{10}$, then we take p_L, p_R, p_T, p_B to be the leftmost, rightmost, top and bottom points of $A = \{v\} \cup {\Gamma_i}^+$ and continue as before.

Proof of (str-6): We assume first that c is at least $100r'$ away from the boundary, where $r' = r - q\sqrt{2}$. Let p be the lower left corner of c, and for $x, y \in \mathbb{Z}$ let us define the *square*

$$S_{x,y} := p + \left[\frac{(x-\frac{1}{2})r'}{\sqrt{5}}, \frac{(x+\frac{1}{2})r'}{\sqrt{5}}\right] \times \left[\frac{(y-\frac{1}{2})r'}{\sqrt{5}}, \frac{(y+\frac{1}{2})r'}{\sqrt{5}}\right],$$

and for $k \in \mathbb{N}$ let us set $R_k := \{S_{x,y} : \max(|x|,|y|) = k\}$, see Figure 5.

Observe that c' is contained in some $S_{x,y}$ with $|x|, |y| \leq 11$ as c, c' have distance at most $10r$. On the other hand, it cannot be that each of $R_{12}, \ldots, R_{27}$ contains a square that does not contain any good cell. This is because otherwise, if $Q_i \in R_i$ is a square that does not contain any good cell, and we set $A := \bigcup_{i=12}^{27} Q_i$ then A satisfies the conditions of Lemma 3.2. Hence setting $A' := \bigcup\{c \in \mathcal{D}(m_n) : c \subseteq A\}$, we have that

$$\begin{array}{rcl} \text{area}(A') & \geq & \text{area}(A) - C \cdot \eta \cdot \left(\frac{\ln n}{n}\right) \\ & = & \frac{16}{5}(r')^2 - C \cdot \eta \cdot \left(\frac{\ln n}{n}\right) \\ & \geq & (1 + 10^{-5})\pi r^2 - C \cdot \eta \cdot \left(\frac{\ln n}{n}\right) \\ & \geq & (1 + 10^{-10})\left(\frac{\ln n}{n}\right), \end{array}$$

for $\eta > 0$ sufficiently small (here C is the absolute constant from Lemma 3.2). But this contradicts Lemma 3.1.

Hence, there is $12 \leq k \leq 27$ such that all squares in R_k contain at least one good cell. This implies that there is a cycle C in $\Gamma_{\max}$ that is completely contained in $\bigcup_{k=12}^{27} R_k$ (note that if $a \in S_{x,y}$ and $b \in S_{x+1,y}$ then $\|a - b\| \leq r'$; and similarly if $a \in S_{x,y}$ and $b \in S_{x,y+1}$.) Let us pick a $c'' \in \Gamma_{\max}$ that has distance at least $100r'$ from both c and c' (Such a c'' exists by **(str-1)** if n is sufficiently large since $R_0, \ldots, R_{27} \subseteq B(p, 100r)$). Consider a cc''-path in Γ_{max}. It must cross C somewhere. Hence, by Lemma 2.13, we can find a path from c to a cell of C that stays inside squares of $R := R_0 \cup \cdots \cup R_{27}$. Similarly there is a path from c' to a cell of C that stays inside squares of $R_0 \cup \cdots \cup R_{27}$. Hence, there is a cc'-path that stays entirely inside squares of R.

Now let $c = c_0, c_1, \ldots, c_N = c'$ be a cc'-path that stays inside squares of R. Let p_i denote the lower left corner of the cell c_i. By leaving out vertices if necessary, we can assume that $c_i c_{i+2} \notin E(\Gamma)$ for all $i = 0, \ldots, N-2$. It follows that the balls $\{B(p_i, r'/2) : i \text{ even}\}$ are disjoint, as are the balls $\{B(p_i, r'/2) : i \text{ odd}\}$. On the other hand we have that

$$B(p_i, r') \subseteq B\left(p, r' \cdot \left(1 + \frac{(\frac{1}{2} + 27)\sqrt{2}}{\sqrt{5}}\right)\right) \subseteq B(p, 100r'),$$

for all i. Thus $\lceil \frac{N}{2} \rceil \pi (\frac{r'}{2})^2 \leq \pi(100r')^2$. This gives that $N \leq 8 \cdot 10^4 \leq 10^5$, as required.

The case when c is closer than $100r'$ to the boundary of the unit square is analogous and we leave it to the reader. ■

3.2 The Obstruction Lemma

We introduce some terminology for sets of dangerous and risky points. Suppose that $V \subseteq [0,1]^2$ and m, T, r are such that **(str-1)**–**(str-6)** above hold. Dangerous points come in groups of points of diameter $< r/100$ that are far apart. We formally define a *dangerous cluster* (with respect to V, m, T, r) to be an inclusion-wise maximal subset of V with the property that $\mathrm{diam}(A) < r \cdot 10^{10}$ and all elements of A are dangerous.

A set $A \subseteq V$ is an *obstruction* (with respect to V, m, T, r) if it is either a dangerous cluster or Γ_i^+ for some $i \geq 2$. We call A an *s-obstruction* if $|A| = s$, and we call it an $(\geq s)$-obstruction if $|A| \geq s$. By **(str-3)-(str-5)**, obstructions are pairwise separated by distance $r \cdot 10^{10}$. (One consequence: a vertex in a good cell is adjacent in G to at most one obstruction.) A point $v \in V$ is *crucial* for A if

(cruc-1) $A \subseteq B(v; r)$, and;

(cruc-2) The vertex v is safe: there is some cell $c \in \Gamma_{\max}$ such that $|B(v;r) \cap c \cap V| \geq T$.

We shall call the T vertices in **(cruc-2)** *important* for the crucial vertex v and *important* for the obstruction A. Note that this crucial vertex could be in the obstruction, or it could be a nearby safe point.

In this section, we prove an important, technical lemma concerning the neighborhood of an obstruction. In particular, we show that *whp* every obstruction has a reasonable number of crucial vertices, along with their corresponding important vertices. This condition is essential for Maker: the obstructions are bottlenecks that require judicious play. These crucial and important vertices provide Maker with the flexibility he needs to overcome the local sparsity of the obstruction. The following is the main result of this section.

Lemma 3.4 (Obstruction Lemma) *For η sufficiently small and T sufficiently large, the following holds. Let $(m_n)_n$ be given by* (12) *and let $V_n := \mathcal{X}_n$ with $\mathcal{X}_n$ as in* (4)*, let $(r_n)_n$ be a sequence of positive numbers such that*

$$\pi n r_n^2 = \ln n + (2k-3)\ln\ln n + O(1),$$

with $k \geq 2$ fixed. Then the following hold whp *with respect to V_n, m_n, T, r_n:*

- **(i)** *For every $2 \leq s < T$, every s-obstruction has at least $k + s - 2$ crucial vertices;*
- **(ii)** *Every $(\geq T)$-obstruction has at least $k + T - 2$ crucial vertices.*

We prove the Obstruction Lemma via a series of intermediate, technical results. First, we prove a general lemma about obstructions that are either near the boundary or have large diameter. The next lemma may seem like a bit of overkill at first. Let us therefore point out that it will be used in several places in the paper.

Lemma 3.5 *Let $k \geq 2$ and $0 < \delta \leq \frac{1}{100}$ and $C, D \geq 100$ be arbitrary but fixed, and let r_n be such that*

$$\pi n r_n^2 = (1 + o(1)) \ln n.$$

Let Z denote the number of k-tuples $(X_{i_1}, \ldots, X_{i_k}) \in \mathcal{P}_n^k$ such that

(i) $|(B(X_{i_j}; r_n) \setminus B(X_{i_j}; \delta r_n)) \cap \mathcal{P}_n| \leq D$ *for all* $1 \leq j \leq k$*, and;*

(ii) $\|X_{i_j} - X_{i_{j'}}\| \leq C r_n$ *for all* $1 \leq j < j' \leq k$*, and;*

(iii) *One of the following holds:*

(a) There is a $1 \leq j \leq k$ *such that* X_{i_j} *is within* Cr *of a corner of the unit square, or;*

(b) There is a $1 \leq j \leq k$ *such that* X_{i_j} *is within* Cr *but no closer than* δr *to the boundary of the unit square, or;*

(c) There are $1 \leq j < j' \leq k$ *such that* $\|X_{i_j} - X_{i_{j'}}\| \geq \delta r_n$*.*

Then $\mathbb{E}Z = O(n^{-c})$ *for some* $c = c(\delta)$*.*

Proof: Let us set $C' := 2kC$. Let Z_{cnr} denote the number k-tuples $(u_1, \ldots, u_k) \in \mathcal{P}_n$ satisfying the demands from the lemma and u_1 is within $C'r_n$ of a corner of the unit square; this includes all k-tuples that satisfy condition (a). Let Z_{sde} denote the number of such k-tuples satisfying the demands from the lemma, and u_1 is within $C'r_n$ of the boundary of the unit square, but not within $C'r_n$ of a corner; this includes all k-tuples satisfying condition (b) but not condition (a). Let Z_{mdl} denote the number of such k-tuples with u_1 more than $C'r_n$ away from the boundary of the unit square and satisfying condition (c) but not counted by Z_{cnr} or Z_{sde}.

Using Theorem 2.7 above, we find that

$$\mathbb{E}Z_{\text{cnr}} = n^k \cdot \mathbb{E}h(Y_1, \ldots, Y_k; \{Y_1, \ldots, Y_k\} \cup \mathcal{P}_n),$$

where h is the indicator function of the event that Y_1 is within $C'r_n$ of a corner and $(Y_1, \ldots, Y_k)$ satisfy the demands from the lemma; and $Y_1, \ldots, Y_k$ are chosen uniformly at random from the unit square and are independent of each other and of $\mathcal{P}_n$.

Observe that, for every $u \in [0,1]^2$ we have:

$$n \cdot \text{area}((B(u; r) \setminus B(u; \delta r)) \cap [0,1]^2) \geq \mu_{\text{cnr}} := n \cdot \frac{1}{4}\pi(1 - \delta^2)r^2$$

Let $A \subseteq [0,1]^2$ be the set of points of the unit square that are within $C'r$ of a corner, and denote $A(u) := B(u; Cr) \cap [0,1]^2$:

$$\begin{array}{rcl}
\mathbb{E}Z_{\text{cnr}} & \leq & n^k \displaystyle\int_A \int_{A(u_1)} \cdots \int_{A(u_1)} \mathbb{P}\Big(\text{Po}\left((B(u_1; r) \setminus B(u_1; \delta r) \cap [0,1]^2\right) \leq D\Big) \mathrm{d}u_k \ldots \mathrm{d}u_1 \\
& \leq & n^k \cdot 4\pi(C'r)^2 \cdot \left(\pi C^2 r^2\right)^{k-1} \cdot \mathbb{P}(\text{Po}(\mu_{\text{cnr}}) \leq D) \\
& = & O\left(\ln^k n \cdot \exp[-\mu_{\text{cnr}} \cdot H(D/\mu_{\text{cnr}})]\right) \\
& = & O\left(\ln^k n \cdot \exp\left[-(\frac{1}{4}(1-\delta^2) + o(1)) \cdot \ln n\right]\right) \\
& = & O\left(n^{-\frac{1}{4}(1-\delta^2)+o(1)}\right).
\end{array} \tag{13}$$

Here we used Lemma 2.5 and that $D/\mu_{\text{cnr}} \to 0$ so that $H(D/\mu_{\text{cnr}}) \to 1$.

Let us now consider Z_{mdl}. For $u, v \in \mathbb{R}^2$, let us write

$$A(u, v) := ((B(u; r) \setminus B(u; \delta r)) \cup (B(v; r) \setminus B(v; \delta r))).$$

If $u, v \in [0,1]^2$ are such that $\|u - v\| \geq \delta r$ then (6) gives

$$n \cdot \text{area}(A(u, v)) \geq \mu_{\text{mdl}} := (\pi + \delta - 2\pi\delta^2)nr^2.$$

Using Theorem 2.7 again, we get

$$\begin{array}{rcl}
\mathbb{E}Z_{\text{mdl}} & \leq & n^k \cdot \left(\pi C^2 r^2\right)^k \cdot \mathbb{P}(\text{Po}(\mu_{\text{mdl}}) \leq D) \\
& = & O\left(\ln^k n \cdot \exp[-\mu_{\text{mdl}} \cdot H(D/\mu_{\text{mdl}})]\right) \\
& = & O\left(n^{-1-\delta/\pi+2\delta^2+o(1)}\right).
\end{array} \tag{14}$$

We now consider $Z_{\rm sde}$. Observe that if $u, v \in [0,1]^2$ are such that v is more than δr away from one side of the unit square, and more than r from all other sides, then (8) gives that

$$\begin{array}{rcl} n \cdot \text{area}(A(u,v) \cap [0,1]^2) & \geq & n \cdot \text{area}\left([0,1]^2 \bigcap B(v;r) \setminus (B(v;\delta r) \cup B(u;\delta r))\right) \\ & \geq & (\frac{\pi}{2} + \delta - 2\pi\delta^2) n r^2. \end{array}$$

Similarly, if $u, v \in [0,1]^2$ are such that both u, v are more than r away from all but one side of the unit square and $\|u - v\| \geq \delta r$ then (6) gives

$$\begin{array}{rcl} n \cdot \text{area}([0,1]^2 \cap A(u,v)) & \geq & n \cdot \frac{1}{2}\,\text{area}(A(u,v)) \\ & \geq & \mu_{\rm sde} := \left(\frac{\pi+\delta-2\pi\delta^2}{2}\right) n r^2. \end{array}$$

We see that

$$\begin{array}{rcl} \mathbb{E}Z_{\rm sde} & \leq & n^k \cdot 4 \cdot C' r_n \cdot \left(\pi C r_n^2\right)^{k-1} \cdot \mathbb{P}(\text{Po}(\mu_{\rm sde}) \leq D) \\ & = & O\left(n^{\frac{1}{2}} \ln^{\frac{2k-1}{2}} n \cdot \exp[-\mu_{\rm sde} \cdot H\left(D/\mu_{\rm sde}\right)]\right) \\ & = & O\left(n^{-\delta/2\pi+\delta^2+o(1)}\right). \end{array} \tag{15}$$

Since $\delta \leq 1/100$, we have $\delta/2\pi > \delta^2$. Together with (13) and (14) this proves the bound on $\mathbb{E}Z$ with $c(\delta) = \delta/100\pi$. ■

Next, we prove a series of three technical lemmas about nearby pairs of points with restrictions on their common neighbours. This argument culminates in Corollary 3.7 below, which illuminates the structure of the shared neighborhood of such pairs. We start with two definitions, shown visually in Figure 6. Given $V \subseteq \mathbb{R}^2$ and $r > 0$, we will say that the pair $(u,v) \in V^2$ is an *(a,b,c)-pair* (with respect to r) if:

(pr-1) $\|u - v\| < r/100$;

(pr-2) $B(u, r - \|u - v\|) \setminus B(u, \|u - v\|)$ contains exactly a points of $V \setminus \{u, v\}$;

(pr-3) $(B(u,r) \cup B(v,r)) \setminus B(u, r - \|u - v\|)$ contains exactly b points of $V \setminus \{u, v\}$.

(pr-4) $B(u, \|u - v\|)$ contains exactly c points of $V \setminus \{u, v\}$.

We say that (u,v) is an $(a, b, \geq c)$-pair if it is an (a, b, c')-pair for some $c' \geq c$.

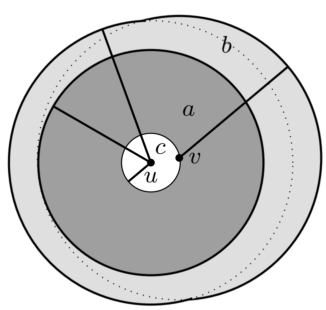

Figure 6: Vertices u, v with $\|u - v\| < r/100$ form an (a,b,c)-pair when the vertex counts in the three regions shown are a, b, and c, respectively.

Lemma 3.6 *For $k \geq 2$, let r_n be such that*

$$\pi n r_n^2 = \ln n + (2k-3)\ln\ln n + O(1),$$

and let $\mathcal{P}_n$ as in (3). Let $a, b, c \in \mathbb{N}$ be arbitrary but fixed, and let R denote the number of $(a, b, \geq c)$-pairs in $\mathcal{P}_n$ with respect to r_n. Then $\mathbb{E}R = O\left(\ln^{a-(k+c)} n\right)$.

Proof: We start by considering $\mathbb{E}R$. Let R_{cnr} denote the number of $(a,b,\geq c)$-pairs (X_i,X_j) in $\mathcal{P}_n$ for which X_i is within $100r$ of a corner of $[0,1]^2$; let R_{sde} denote the number of $(a,b,\geq c)$-pairs (X_i,X_j) in $\mathcal{P}$ for which X_i is within $100r$ of the boundary of $[0,1]^2$, but not within $100r$ of a corner, and let R_{mdl} denote the number of $(a,b,\geq c)$-pairs (X_i,X_j) for which X_i is more than $100r$ away from the boundary of $[0,1]^2$.

By Lemma 3.5 (with $k=2,\delta=1/100,D=\max(a+b,100),C=100$) we have

$$\mathbb{E}R_{\text{cnr}} = O\left(n^{-\Omega(1)}\right). \tag{16}$$

For $0<z<r/100$ let us write:

$$\begin{aligned}
\mu_1(z) &:= n\cdot\text{area}(B(u;r-z)\setminus B(u;z)),\\
\mu_2(z) &:= n\cdot\text{area}((B(u;r)\cup B(v;r))\setminus B(u;r-z)),\\
\mu_3(z) &:= n\cdot\text{area}(B(u;z)),
\end{aligned}$$

where $u,v\in\mathbb{R}^2$ are two points with $\|u-v\|=z$. Observe that

$$\mu_1(z)+\mu_2(z)+\mu_3(z) = n\cdot\text{area}(B(u,r)\cup B(v,r)) \geq n\cdot(\pi r^2+rz),$$

using Lemma 2.14. If $z\leq r/100$ then we also have

$$\mu_1(z)+\mu_2(z) \geq n\cdot(\pi r^2+rz-\pi z^2) \geq n\cdot(\pi r^2+\frac{1}{2}rz). \tag{17}$$

Let us now consider R_{mdl}. Using Theorem 2.7 and Lemma 2.6 we find

$$\begin{aligned}
\mathbb{E}R_{\text{mdl}} &= \frac{n^2}{2}\int_{[100r,1-100r]^2}\int_{B(v;r/100)}\frac{\mu_1(\|u-v\|)^a e^{-\mu_1(\|u-v\|)}}{a!}\cdot\frac{\mu_2(\|u-v\|)^b e^{-\mu_2(\|u-v\|)}}{b!}\cdot\\
&\qquad\qquad \mu_3(\|u-v\|)^c\mathrm{d}u\mathrm{d}v\\
&= \frac{n^2}{2}(1-200r)^2\int_0^{r/100}\frac{\mu_1(z)^a e^{-\mu_1(z)}}{a!}\cdot\frac{\mu_2(z)^b e^{-\mu_2(z)}}{b!}\cdot\mu_3(z)^c\cdot 2\pi z\mathrm{d}z\\
&\leq n^2\int_0^{r/100}\left(\pi nr^2\right)^a\left(4\pi nrz\right)^b\left(\pi nz^2\right)^c e^{-\pi nr^2-nrz}2\pi z\mathrm{d}z\\
&= O\left(n^{2+c}\cdot\ln^a n\cdot(nr)^b\int_0^{r/100}e^{-\ln n-(2k-3)\ln\ln n-nrz}z^{b+2c+1}\mathrm{d}z\right)\\
&= O\left(n^{1+c}\cdot(\ln n)^{a-(2k-3)}\cdot(nr)^b\int_0^{r/100}e^{-nrz}z^{b+2c+1}\mathrm{d}z\right)\\
&= O\left(n^{1+c}\cdot(\ln n)^{a-(2k-3)}\cdot(nr)^{-(2+2c)}\right)\\
&= O\left((\ln n)^{a-(2k-3)}\cdot(nr^2)^{-(1+c)}\right)\\
&= O\left((\ln n)^{a+2-2k-c}\right)\\
&= O\left((\ln n)^{a-(k+c)}\right).
\end{aligned} \tag{18}$$

Here we have used a switch to polar coordinates to get the second line; we used (17) to get the third line; the change of variables $y=nrz$ to get the sixth line; and in the last line we use $k\geq 2$.

Now we turn attention to R_{sde}. Let R'_{sde} denote the number of (a,b,c)-pairs (u,v) with u at distance at least $r/100$ and at most $100r$ from the boundary. Then, again by Lemma 3.5:

$$\mathbb{E}R'_{\text{sde}} = O\left(n^{-O(1)}\right). \tag{19}$$

On the other hand, by Theorem 2.7 and a switch to polar coordinates (and where $z = ||u - v||$ and w is the distance of the nearest of u, v to the boundary):

$$
\begin{aligned}
\mathbb{E}\left(R_{\text{sde}} - R'_{\text{sde}}\right) &\leq 8n^2 \int_0^{r/100}\int_0^{r/100} \left(\pi n r^2\right)^a \left(4\pi n r z\right)^b \left(\pi n z^2\right)^c e^{-\frac{\pi}{2}nr^2 - \frac{1}{2}nrz - nrw + \pi z^2} \pi z \mathrm{d}z\mathrm{d}w \\
&\leq 8n^2 \int_0^{r/100}\int_0^{r/100} \left(\pi n r^2\right)^a \left(4\pi n r z\right)^b \left(\pi n z^2\right)^c e^{-\frac{\pi}{2}nr^2 - \frac{1}{4}nrz - nrw} \pi z \mathrm{d}z\mathrm{d}w \\
&= O\left(n^{2+c} \cdot \ln^a n \cdot (nr)^b \int_0^{r/100}\int_0^{r/100} z^{b+2c+1} e^{-\frac{\pi}{2}nr^2 - \frac{1}{2}nrz - nrw} \mathrm{d}z\mathrm{d}w\right) \\
&= O\left(n^{\frac{3}{2}+c} \cdot (\ln n)^{a-(2k-3)/2} \cdot (nr)^b \int_0^{r/100}\int_0^{r/100} z^{b+2c+1} e^{-\frac{1}{2}nrz - nrw} \mathrm{d}z\mathrm{d}w\right) \\
&= O\left(n^{\frac{3}{2}+c} \cdot (\ln n)^{a-(2k-3)/2} \cdot (nr)^{b-(b+2c+3)}\right) \\
&= O\left((\ln n)^{a-(2k-3)/2} \cdot (nr^2)^{-(\frac{3}{2}+c)}\right) \\
&= O\left((\ln n)^{a-(2k-3)/2-\frac{3}{2}-c}\right) \\
&= O\left((\ln n)^{a-(k+c)}\right).
\end{aligned}
$$

Combining this with (16), (18) and (19) proves the lemma. ■

Clearly the definition of $(a, b, \geq c)$-pairs can reformulated in terms of a function h_n as in Lemma 2.8. As an immediate corollary of Lemmas 3.6 and 2.8 we now find:

Corollary 3.7 *For $k \geq 2$, let $(r_n)_n$ be such that*

$$\pi n r_n^2 = \ln n + (2k-3)\ln\ln n + O(1),$$

and let $\mathcal{X}_n$ be as in (4). Let $a, b, c \in \mathbb{N}$ be fixed such that $a \leq c + k - 1$. W.h.p., there are no $(a, b, \geq c)$-pairs for $\mathcal{X}_n, r_n$.

Finally, we are ready to prove the Obstruction Lemma.

Proof of Lemma 3.4: Let $\eta > 0$ be small and $T > 0$ be large, to be determined in the proof. All the probability theory needed for the current proof is essentially done. For the remainder of the proof we can assume that $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ is an arbitrary set of points in the unit square for which the conclusions of Lemma 3.1 with $\varepsilon := 1/10^{10}$ and $K = 10^{10}$ hold, **(str-1)**–**(str-5)** hold and the conclusions of Corollary 3.7 hold for all $a, b, c \leq 1000 \cdot D$, where

$$D := (T-1) \cdot \left\lfloor \pi \left(r + \frac{\sqrt{2}}{m}\right)^2 \Big/ \left(\frac{1}{m}\right)^2 \right\rfloor. \tag{20}$$

(There are at most $\lfloor \pi(r + \sqrt{2}/m)^2/(1/m)^2 \rfloor$ cells that intersect a disk of radius r, so D is an upper bound on the degree of a dangerous vertex.) That is, there are no $(a, b, \geq c)$-pairs with $b, c \leq 1000 \cdot D$ and $a \leq k + c - 1$. (Observe that D is bounded because $r = O(\sqrt{\ln n/n})$ and m is defined as in (12).)

Let A be an arbitrary obstruction with $2 \leq s := |A| < T$. Then A is necessarily a dangerous cluster. Pick $u, v \in A$ with $\operatorname{diam}(A) = \|u - v\| =: z$. Observe that $(B(u, r) \cup B(v, r)) \setminus B(u, r - z)$ cannot contain more than $2D$ points, since this would imply either $d(u) \geq D$ or $d(v) \geq D$ and this would in turn imply that one of u, v would be safe. Thus $B(u, r - z) \setminus B(u, z)$ contains at least $m := k + s - 2$ points $X_{i_1}, \ldots, X_{i_m}$. (Otherwise (u, v) would be a (a, b, c)-pair for some $s - 2 \leq c \leq D$, $a \leq c + k - 1$ and $b \leq 2D$, contradicting our assumptions.) Let us emphasize that b is big and c is small.

Observe that $X_{i_1}, \ldots, X_{i_m}$ are within r of every element of A. Also observe that $X_{i_1}, \ldots, X_{i_m}$ must be either safe or risky, because, by definition of a dangerous cluster, A is an inclusion-wise maximal subset of $\mathcal{X}_n$ with the property that $\operatorname{diam}(A) < r \cdot 10^{10}$ and all elements of A are dangerous. Thus each X_{i_j} has at least T neighbours in some good cell c_j. By **(str-5)** we must have $c_1, \dots, c_m \in \Gamma_{\max}$ so that $X_{i_1}, \dots, X_{i_m}$ are indeed crucial.

Now suppose that $|A| \geq T$. If A is a dangerous cluster then we proceed as before. Let us thus suppose that $A = \Gamma_i^+$ for some $i \geq 2$. Let

$$d_{V \setminus A}(v) := |N(v) \setminus A|,$$

and let W denote the set of those X_i that are crucial for A. If $|W| \geq k$ then we are done. Let us thus assume that $|W| < k$. Let us observe:

$$\text{Any } v \in A \setminus W \text{ has } d_{V \setminus A}(v) \leq D. \tag{21}$$

(This is because if $d_{V \setminus A}(v) > D$ then v is adjacent to at least T vertices in some cell c that does not belong to Γ_i. This cell c is therefore certainly good. By **(str-4)**, any good cell that is within $r \cdot 10^{10}$ of A, but not part of Γ_i, must belong to $\Gamma_{\max}$.)

In that case $|A \setminus W| > T - k \geq k$ (T being sufficiently large). Pick an arbitrary $u \in A \setminus W$ and set

$$A' := B(u, r/100) \cap V.$$

Let v be a point of $A' \setminus W$ of maximum distance to u, and write $z := \|u - v\|$. Observe that $(B(u,r) \cup B(v,r)) \setminus B(u, r - z)$ cannot contain more than $3D$ points, since this would imply either $d_{V \setminus A}(u) > D$ or $d_{V \setminus A}(v) > D$ and this in turn would imply that one of u, v would be crucial. Thus $B(u, r-z) \setminus B(u, z)$ contains at least k points $X_{i_1}, \dots, X_{i_k}$. (Otherwise (u, v) would be an $(a, b, \geq 0)$-pair for some $a \leq k-1, b \leq 1000 \cdot D$, contradicting our assumptions.) We claim these points must be crucial. Aiming for a contradiction, suppose $X_{i_j} \notin W$ for $1 \leq j \leq k$. What is more, by choice of A' and v, we must have $\|X_{i_j} - u\| > r/100$. Let us first suppose that u is at least $2r$ away from the boundary of the unit square. Using the bound (6) and Lemma 3.2, we see that the total area of the cells that fall inside $O := (B(X_{i_j}; r) \cup B(u; r)) \setminus B(u; r/100)$ is at least

$$\begin{array}{rcl} r^2\left(\pi + \frac{1}{100}\right) - \pi r^2 (\frac{1}{100})^2 - C\eta\left(\frac{\ln n}{n}\right) & \geq & (1 + \frac{1}{1000})\pi r^2 - C\eta\left(\frac{\ln n}{n}\right) \\ & \geq & (1 + 10^{-5})\frac{\ln n}{n}, \end{array}$$

provided we have chosen η small enough. Similarly, if u is within $2r$ of one of the sides of the unit square and at least $2r$ from all other sides then the total area of the cells inside O is at least $\frac{1}{2}\operatorname{area}(O) - C\eta\left(\frac{\ln n}{n}\right) \geq (1 + 10^{-5}) \ln n / 2n$; and if u is within $2r$ of two of the sides of the unit square then the total area of the cells inside O is at least $\frac{1}{4}\operatorname{area}(O) - C\eta\left(\frac{\ln n}{n}\right) \geq (1 + 10^{-5}) \ln n / 4n$. (Provided η was chosen small enough in both cases). Hence, by the conclusion of Lemma 3.1, at least one of these cells in O must be good. Since such a good cell is not part of the component of Γ that generates A, and is closer than $r \cdot 10^{10}$ to A, it must belong to $\Gamma_{\max}$. Hence either X_{i_j} or u is crucial. Since we chose $u \in A \setminus W$ we must have $X_{i_j} \in W$, a contradiction.

This proves that all of $X_{i_1}, \dots, X_{i_k}$ are crucial, completing the proof. ■

4 The connectivity game

To improve the presentation, we separate the proof into a deterministic part and a probabilistic part. The deterministic part is the following lemma:

Lemma 4.1 *Suppose that $V \subseteq [0,1]^2, m \in \mathbb{N}, T > 100, r > 0$ and $r \leq \rho \leq 2r$ are such that*

- **(i)** **(str-1)**–**(str-5)** *hold with respect to r, and;*
- **(ii)** *Every obstruction with respect to ρ has at least 2 crucial vertices.*

Then Maker wins the connectivity game on $G(V; \rho)$.

Proof: Observe that $\Gamma = \Gamma(V; m, T, \rho)$ also satisfies **(str-1)**–**(str-5)** if we modify them very slightly by replacing the number $r \cdot 10^{10}$ in **(str-3)**–**(str-5)** by $r \cdot 10^{10}/2$. All mentions of safe, dangerous, obstructions etc. will be with respect to ρ in the rest of the proof.

From Theorem 2.1, we know that Maker can win the game if $G(V; \rho)$ contains two edge disjoint spanning trees.

In every cell $c \in \Gamma_{\max}$ there are at least T vertices and they form a clique, so we can find two edge disjoint trees on them. Since $\Gamma_{\max}$ is connected, there is a pair of edges between every two adjacent cells, and they complete two edge disjoint trees on all vertices in cells of $\Gamma_{\max}$. Every vertex that is safe but not in any obstruction has more than T neighbours among the vertices in cells of $\Gamma_{\max}$, so we can extend our pair of trees to such vertices as well. Finally, any obstruction has at least two crucial vertices, so the vertices in the obstruction can be connected to one of the trees via one crucial vertex, and to the other one via another crucial vertex. This way, we obtain a pair of edge disjoint spanning trees in our graph. ■

Proof of Theorem 1.1: Observe that if there is a vertex of degree at most one then Breaker can isolate this vertex and win the connectivity game (recall that Breaker has the first move). This implies:

$$\mathbb{P}\left(\rho_n(\text{Maker wins the connectivity game}) \geq \rho_n(\text{min. deg.} \geq 2)\right) = 1. \tag{22}$$

It remains to see that $\rho_n(\text{Maker wins the connectivity game}) \leq \rho_n(\text{min. deg.} \geq 2)$ also holds *whp*.

Let $K > 0$ be a (large) constant, and define:

$$r_L(n) := \left(\frac{\ln n + \ln\ln n - K}{\pi n}\right)^{\frac{1}{2}}, \quad r_U(n) := \left(\frac{\ln n + \ln\ln n + K}{\pi n}\right)^{\frac{1}{2}}. \tag{23}$$

By Theorem 2.10, there is a $K = K(\varepsilon)$ such that

$$\begin{aligned}
\mathbb{P}\Big(\delta(G(n, r_L)) < 2,\, \delta(G(n, r_U)) \geq 2\Big) &\geq \mathbb{P}\Big(\delta(G(n, r_L)) < 2\Big) - \mathbb{P}\Big(\delta(G(n, r_U)) < 2\Big) \\
&= 1 - e^{-(e^K + \sqrt{\pi e^K})} - (1 - e^{-(e^{-K} + \sqrt{\pi e^{-K}})}) + o(1) \\
&= e^{-(e^{-K} + \sqrt{\pi e^{-K}})} - e^{-(e^K + \sqrt{\pi e^K})} + o(1).
\end{aligned}$$

Hence, for arbitrary $\varepsilon > 0$, we can choose K such that

$$\mathbb{P}\Big(\delta(G(n, r_L)) < 2,\, \delta(G(n, r_U)) \geq 2\Big) \geq 1 - \varepsilon + o(1).$$

This shows that

$$\mathbb{P}\Big(r_L(n) \leq \rho_n(\text{min. deg.} \geq 2) \leq r_U(n)\Big) \geq 1 - \varepsilon + o(1). \tag{24}$$

Now notice that, by Lemma 3.3 the properties **(str-1)**–**(str-5)** are satisfied with probability $1 - o(1)$ by $V = \mathcal{X}_n, m = m_n, T = 10, r = r_L(n)$ with $\mathcal{X}_n$ as given by (4), m_n as given by (12) and r_L as above. By Lemma 3.4, with probability $1 - o(1)$, $V = \mathcal{X}_n, m = m_n, T = 10, r = r_L(n)$ are such that every (≥ 2)-obstruction with respect to r_L has at least two crucial vertices. This is then clearly also true with respect to any $\rho \geq r_L$. Since a 1-obstruction is just a vertex of low degree, by definition every 1-obstruction has at least two crucial vertices for every $\rho \geq \rho_n(\text{min. deg.} \geq 2)$. Also observe that $r_L(n) \leq 2r_U(n)$. Hence, by Lemma 4.1 and (22):

$$\mathbb{P}\left(\rho_n(\text{Maker wins the connectivity game}) = \rho_n(\text{min. deg.} \geq 2)\right) \geq 1 - \varepsilon - o(1).$$

Sending $\varepsilon \downarrow 0$ gives the theorem. ■

5 The Hamilton cycle game

In this section, we prove Theorem 1.5. First, we prove that Maker can win two different path-making games. These games will be useful in constructing local paths that will be stitched together to create the Hamilton cycle. Next, we introduce *blob cycles* and prove some results about them. In particular, we give some conditions under which blob cycles can be merged into larger blob cycles. With these intermediate results in hand, we then characterize the hitting radius for the Hamilton cycle game.

5.1 Two helpful path games

We begin by considering two auxiliary path games. Maker will use these local path games to create parts of his Hamilton cycle. Both games are played on a clique in which Maker tries to make a long path. In the first game, Maker tries to make a path through all but one vertex.

Lemma 5.1 *Let $s \geq 3$ be arbitrary. Maker can make a path of length $s-2$ in K_s.*

Proof: The cases $s = 3, 4$ are straight-forward and left to the reader. Assume that $s \geq 5$ and the statement is true for $s-1$. Our strategy for Maker is as follows. He chooses an arbitrary vertex u. If Breaker picks an edge incident with u then Maker responds by taking another edge incident with u. (If there is no such free edge, he claims an arbitrary free edge and forgets about it in the remainder of the game.) If Breaker claims an edge between two vertices of $K_s \setminus u$ then Maker responds by picking another edge between two vertices of $K_s \setminus u$ according to the winning strategy on K_{s-1}.

We now argue that this is a winning strategy. Let M denote Maker's graph at the end of the game. If M contains a path spanning $K_s \setminus u$ we are done. Otherwise, M contains a path $P = v_1, \ldots, v_{s-2}$ that contains all but one vertex w of $K_s \setminus u$. Observe that u is adjacent to at least $\lfloor \frac{s-1}{2} \rfloor$ vertices of $K_s \setminus u$. If u is adjacent to v_1 or to v_{s-2} then we are done, so we can assume this is not the case.

If u is not adjacent to w, then it has at least $\lfloor \frac{s-1}{2} \rfloor > \lceil \frac{s-4}{2} \rceil$ neighbours amongst the $s-4$ interior vertices of P. But then u is adjacent to two consecutive vertices v_i, v_{i+1} of P and hence $P' := v_1, \ldots, v_i, u, v_{i+1}, \ldots, v_{s-2}$ is a path of the required type. Hence we can assume that u is adjacent to w. If u is also adjacent to v_{s-3} then $P' := v_1, \ldots, v_{s-3}, u, w$ is path of the required type. Hence we can assume this is not the case. Similarly we can assume u is not adjacent to v_2. If $s = 5$ or $s = 6$ then this in fact implies that u is not adjacent to any vertex of P, so that w is the only neighbour of u. But this contradicts the fact that u has at least $\lfloor \frac{s-1}{2} \rfloor \geq 2$ neighbours.

Considering $s > 6$, vertex u has at least $\lfloor \frac{s-1}{2} \rfloor - 1 > \lceil \frac{s-6}{2} \rceil$ neighbours amongst $v_3, \ldots, v_{s-4}$. Again it follows that u is adjacent to two consecutive vertices of P, so that we once again find a path of the required type. This concludes the proof. ■

Our second auxiliary game is the (a, b) path game. It is played on the graph $G_{a,b}$ which has a vertex set of size $a + b$, partitioned into two sets A, B with $|A| = a, |B| = b$, where B induces a stable set and all edges not between two vertices of B are present. (So in particular, A is a clique and $uv \in E(G_{a,b})$ for all $u \in A, v \in B$.)

Maker's objective is to create either a single path between two vertices of B that contains all vertices of A (and possibly some other vertices from B), or two vertex disjoint paths between vertices of B that cover all vertices of A.

Lemma 5.2 *The (a, b) path game is a win for Maker if one of the following conditions is met*

(i) *$b \geq 6$, or;*

(ii) *$a = 3$ and $b \geq 5$, or;*

(iii) *$a \in \{1, 2\}$ and $b \geq 4$.*

Proof: We prove each of these three winning conditions in turn.

Proof of part (i): We can assume that $a \geq 4$, because otherwise one of the other cases will apply. If $b \geq 6, a \geq 4$ then Maker can play as follows.

- Whenever Breaker plays an edge between two vertices of A then Maker responds by claiming another edge inside A according to the strategy from Lemma 5.1, guaranteeing a path in A that contains all but one vertex of A.
- If Breaker claims an edge between $u \in A$ and $v \in B$ then Maker claims an arbitrary unclaimed edge connecting u to a vertex in B.

If this is not possible then Maker claims an arbitrary edge.

Consider the graph at the game's end. Every vertex of A will have at least three neighbours in B. Maker has also claimed a path P in A that contains $|A|-1$ vertices. Let a_1, a_2 be the endpoints of P, and let a_3 be the sole vertex not on P. Since a_1, a_2 both have at least three neighbours in B, we can extend P to a path P' between two vertices $b_1, b_2 \in B$. If a_3 has two neighbours distinct from b_1, b_2 then we have our two vertex disjoint paths between vertices of B covering A.

Otherwise a_3 is adjacent to b_1, b_2 and a third vertex $b_3 \in B$. But then we can extend P' by going to a_3 and then to b_3 to get a single path between two B-vertices that covers A.

Proof of part (ii): We now consider the case when $a = 3, b \geq 5$. Say $A = \{a_1, a_2, a_3\}, B = \{b_1, \ldots, b_\ell\}$ with $\ell \geq 5$. The strategy for Maker is as follows. For each $i = 1, 2, 3$ he pairs the edges a_ib_1, a_ib_2 (meaning that if Breaker claims one of them, then Maker responds by claiming the other) and he pairs the edges a_ib_3, a_ib_4. This will make sure every vertex of A has at least two neighbours in B. He also pairs a_1b_5, a_2b_5. This will make sure that one of a_1, a_2 has at least three neighbours in B. Furthermore, he pairs a_1a_2, a_1a_3 and a_2a_3, a_3b_5. This ensures Maker will either claim two edges in A, or he will have one edge in A and a_3 will have three neighbours in B. This concludes the description of Maker's strategy.

To see that it is a winning strategy for Maker, let us assume first that Maker's graph has two edges in A. That is, he has claimed a path P through all three vertices of A. Since both endpoints of P have two neighbours in B, it extends to a path P' between two points of B.

Suppose then that Maker has claimed only a single edge of A, but every vertex of A is adjacent to three vertices of B. In this case we can reason as in the proof of part **(i)** to see that Maker has either a single, or two vertex disjoint paths between vertices of B that cover all vertices of A.

Proof of part (iii): If $a = 1$ and $b \geq 4$ then it is easy to see that Maker can claim at least two edges incident with the unique vertex of A, which gives a path of the required type.

Let us thus assume that $A = \{a_1, a_2\}$ and $B = \{b_1, \ldots, b_\ell\}$ with $\ell \geq 4$. The winning strategy for Maker depends on Breaker's first move.

Case 1: Breaker did not claim a_1a_2 in his first move. Without loss of generality, Breaker claimed a_1b_1. Maker responds by claiming a_1a_2, and for the remainder of the game he pairs a_1b_3, a_1b_4 and a_2b_1, a_2b_2. This way, he will clearly end up with a single path between two vertices of B that covers A.

Case 2: Breaker claimed the edge a_1a_2 in his first move. In response, Maker claims a_1b_1. For the rest of the game, Maker plays as follows. Let e_B be the first edge incident with A that Breaker claims after this point. We distinguish three subcases.

(2-a) If $e_B = a_1b_i$ for some $i \geq 1$, then Maker responds by claiming a_2b_i. Without loss of generality $i = 2$. For the rest of the game, Maker now pairs a_1b_3, a_1b_4 and a_2b_3, a_2b_4. This way, both a_1 and a_2 will have at least two neighbours in B, and a_2 has at least one neighbour not adjacent to a_1. But then there is either a single path between B-vertices covering a_1, a_2 or there are two vertex disjoint paths between B vertices, one which covers a_1 and one which covers a_2..

(2-b) If $e_B = a_2b_1$, then Maker claims a_2b_2 in response and for the remainder of the game he pairs a_1b_3, a_1b_4 and a_2b_3, a_2b_4. Again a_1 and a_2 will have at least two neighbours in B, and a_2 will have at least one neighbour not adjacent to a_1. As in case **(2-a)**, we then have the required path(s).

(2-c) If $e_B = a_2b_i$ for $i \geq 2$, then without loss of generality we can assume $i = 2$. Maker responds by claiming a_2b_3; and for the remainder of the game he pairs a_1b_2, a_1b_4 and a_2b_1, a_2b_4. Again a_1 and a_2 will have at least two neighbours in B, and a_2 will have at least one neighbour not adjacent to a_1. Again we can find the required path(s). ■

5.2 Blob Cycles

In this section, we prove some results about *blob cycles.* An s-blob cycle C is the union of a cycle with a clique on s consecutive vertices of it; this s-clique is called the *blob* of C. See Figure 7 for a depiction. When C is an s-blob cycle with $|V(C)| = m$, we typically write $V(C) = \{u_0, \ldots, u_{m-1}\}$ where u_iu_{i+1} is an edge for all i, modulo m, and $u_0, \ldots, u_{s-1}$ are the blob vertices. An *s-blob Hamilton cycle* is the union of a Hamilton cycle with a clique on s consecutive vertices of it.

Blob cycles will be the building blocks of our Hamilton cycle in $G(n, r_n)$. In the next section, we will construct our Hamilton cycle by gluing together many blob cycles in a good cell. Most of these blob cycles are used to connect to a sparser part of the graph. We then use a much larger blob cycle to conglomerate the blob cycles in a good cell into one blob cycle. The lemmas below concern adding vertices and/or edges to a blob cycle, and then finally combining many blob cycles into one.

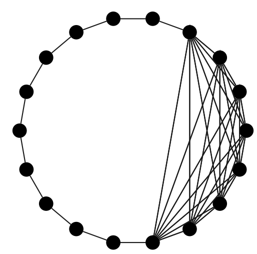

Figure 7: An 8-blob cycle on 18 vertices.

Lemma 5.3 *For $k \geq 4$, there is an $N = N(k)$ such that Maker can make a k-blob Hamilton cycle in an s-clique for $s > N$.*

Proof: By Theorem 2.3 there is an $s_0 = s_0(k)$ such that Maker can make a k-clique on K_s for all $s \geq s_0$. Now let $s_1 = s_1(k)$ be such that Maker can win the biased $(1 : b)$-Hamilton cycle game on K_s for all $s \geq s_1$, where $b := \binom{s_0}{2}$. (Such an s_1 exists by Theorem 2.4.)
We set $N := \max\left(2\binom{s_0}{2} + 1, s_1\right) + k$.

For $s \geq N$ Maker can now play as follows. First he makes a k-clique on the first s_0 vertices of K_s by playing according to the strategy given by Theorem 2.3. (If Breaker plays an edge not between two vertices among the first s_0 then Maker pretends Breaker played an arbitrary free edge between two of the first s_0 vertices and responds to that.)

Once Maker has succeeded in making a k-clique C, no more than $\binom{s_0}{2}$ edges have been played so far. He now continues playing as follows:

- If Breaker claims an edge between two vertices of $V \setminus C$ then Maker responds according to the strategy given by Theorem 2.4 that will ensure him a spanning cycle on $V \setminus C$;
- If Breaker claims an edge between $u \in V \setminus C$ and $v \in C$ then Maker claims an unclaimed edge between u and another vertex of C.

At the end of the game Maker's graph will contain a spanning cycle $\tilde{C}$ on $V \setminus C$. Note that if a vertex $u \in V \setminus C$ was not incident with any edge claimed by Breaker at the time when Maker finished claiming C then, by the end of the game, u will be incident in Maker's graph with at least $\lfloor k/2 \rfloor \geq 2$ vertices of C. By choice of N, more than half of all the vertices of $V \setminus C$ were not incident with any of Breaker's edges just after Maker finished building his k-clique C. Hence there will be two vertices u, v that are consecutive on $\tilde{C}$ and two vertices $x \neq y \in C$ such that Maker claimed the edges ux, vy. This gives the required k-blob Hamilton cycle. ■

Lemma 5.4 *For every $s \geq 4$ the following holds. Let G be a graph, $C \subseteq G$ an s-blob cycle, and $u, v \notin V(C)$ with $uv \in E(G)$ such that u, v each have at least $\lceil |V(C)|/2 \rceil$ neighbours on C. Then there is an s'-blob cycle C' with $s' \geq s-2$ and $V(C') = \{u, v\} \cup V(C)$ where $uv \in E(C')$ and the blob of C' is contained in the blob of C.*

Proof: In this proof (and the ones that follow), we continue to use the convention that $V(C) = \{u_0, \ldots, u_{m-1}\}$ where $u_i u_{i+1}$ is an edge for all i, modulo m, and $u_0, \ldots, u_{s-1}$ are the blob vertices.

Let us first assume that m is odd. In this case, since u, v each have at least $\lceil |V(C)|/2 \rceil = \lceil m/2 \rceil$ neighbours on C, there are two consecutive vertices u_i, u_{i+1} such that $uu_i, vu_{i+1} \in E(G)$. If $s-1 \leq i \leq m-1$, then $u_0, \ldots, u_i, u, v, u_{i+1}, \ldots, u_{m-1}$ is an s-blob cycle. If $i = 0$ then $u_0, u, v, u_1, \ldots, u_{m-1}$ is an $(s-1)$-blob cycle; the case $i = s-2$ is similar. If $1 \leq i \leq s-3$, we can relabel the blob vertices $u_1, u_2, \ldots, u_{s-3}$ so that $i = 1$. Then $u_0, u_1, u, v, u_2, \ldots, u_{m-1}$ is an $(s-2)$-blob cycle. In all cases, the new blob is a subset of the old blob.

Let us then assume that m is even. If there are two consecutive vertices u_i, u_{i+1} such that $uu_i, vu_{i+1} \in E(G)$, then we can proceed as before. If there no two such consecutive vertices then either u, v are both adjacent to each of $u_0, u_2, \ldots, u_{m-2}$ or both are adjacent to each of $u_1, u_3, \ldots, u_{m-1}$. In both cases we can easily relabel the vertices $u_1, \ldots, u_{s-2}$ of the blob in such a way that $uu_i, vu_{i+1} \in E(G)$ for some i, and we are again done by a previous argument. ■

Lemma 5.5 *For every ℓ there is an $s = s(\ell)$ such that the following holds. If G is a graph, $C \subseteq G$ is an s-blob cycle, and $v \notin V(C)$ has at least $|V(C)|/2 - \ell$ neighbours on C, then there exists an s'-blob cycle C' with $s' \geq s-2$ and $V(C') = \{v\} \cup V(C)$ and the blob of C' is contained in the blob of C.*

Proof: Let us set $s := 100 \cdot (\ell + 1)$, and write $V(C) = \{u_0, \ldots, u_{m-1}\}$ and $u_0, \ldots, u_{s-1}$ are the vertices of an s-clique. First, suppose that v is adjacent to at least two vertices of $u_1, \ldots, u_{s-2}$. By reordering these vertices, $vu_1, vu_2 \in E(G)$. Set $C' = u_0, u_1, v, u_2, \ldots, u_{m-1}$. Then C' is clearly an $(s-2)$-blob cycle whose blob is inside the blob of C.

Next, suppose then that v has at most one neighbour among $u_1, \ldots, u_{s-2}$. Then v must have $\lfloor m/2 \rfloor - (\ell+1)$ neighbours on the path $P = u_{s-1}, \ldots, u_{m-1}, u_0$. Observe that the biggest subset of $V(P)$ that does not contain two consecutive vertices has cardinality

$$\lceil |V(P)|/2 \rceil = \lceil (m-s+2)/2 \rceil \leq m/2 - 50(\ell+1) + 2 < \lfloor m/2 \rfloor - (\ell+1).$$

Therefore v is adjacent to two consecutive vertices on the path P, say $vu_i, vu_{i+1} \in E(G)$. This time $C' = u_0, \ldots, u_i, v, u_{i+1}, \ldots, u_{m-1}$ is clearly an s-blob cycle whose blob is identical to the blob of C. ■

Lemma 5.6 *For every ℓ there is an $s = s(\ell)$ such that the following holds. Suppose that G is a graph, $C_1, C_2 \subseteq G$ are vertex disjoint, C_1 is an s-blob cycle, C_2 is a 5-blob cycle, and every vertex of the blob of C_2 has at least $\lfloor |V(C_1)|/2 \rfloor - \ell$ neighbours on C_1. Then there exists an s'-blob cycle C where $s' \geq s-2$ with $V(C) = V(C_1) \cup V(C_2)$ and the blob of C is contained in the blob of C_1.*

Proof: Let us set $s := 100 \cdot (\ell + 1)$, and write $V(C_1) = \{u_0, \ldots, u_{m-1}\}$ where $u_0, \ldots, u_{s-1}$ are the blob vertices. Similarly $V(C_2) = \{v_0, \ldots, v_{n-1}\}$ where $v_0, \ldots, v_4$ form a clique and $v_0, \ldots, v_4$ each have at least $\lfloor m/2 \rfloor - \ell$ neighbours on C_1.

First suppose that $v_1 u_0, v_2 u_1 \in E(G)$. The graph $C' = u_0, v_1, v_0, v_{n-1}, \ldots, v_2, u_1, \ldots, u_{m-1}$ is an $(s-1)$-blob cycle of the required type. Next, when $v_1 u_1, v_2 u_2 \in E(G)$ the graph $C' = u_0, u_1,$

$v_1, v_0,\ v_{n-1}, \ldots, v_2,\ u_2, \ldots, u_{m-1}$ is an $(s-2)$-blob cycle. More generally, if any two among $v_1, \ldots, v_3$ have distinct neighbours among $u_0, \ldots, u_{s-1}$ then we can just re-order the vertices of C_1 and C_2 and apply the above argument to get an $(s-1)$-blob cycle or an $(s-2)$-blob cycle.

If any two of v_1, v_2, v_3 both have at least two neighbours among $u_1, \ldots, u_{s-2}$ then we are done by the previous argument. So assume this is not the case. Relabeling C_2 if necessary, v_1, v_2 each have at most one neighbour among $u_1, \ldots, u_{s-2}$. Then v_1, v_2 each have $\lfloor m/2 \rfloor - \ell - 1$ neighbours on $P = u_{s-1}, \ldots, u_{m-1}, u_0$. In turn, there are at least $\lfloor m/2 \rfloor - \ell - 1$ points of P *adjacent to a neighbour of* v_1. Since $(\lfloor m/2 \rfloor - \ell - 1) + (\lfloor m/2 \rfloor - \ell - 1) > m - s + 1 = |V(P)|$ there exists an $s - 1 \leq i \leq m - 1$ such that $v_1 u_i, v_2 u_{i+1} \in E(G)$. Clearly $C' = u_0, \ldots, u_i,\ v_1, v_0,\ v_{n-1}, \ldots, v_2,$ $u_{i+1}, \ldots, u_{n-1}$ is an s-blob cycle of the required type. ■

Our final conglomeration lemma is the crucial tool for merging an ensemble of blob cycles, along with some additional edges and vertices, into a single spanning cycle. Ultimately, we will use this lemma to handle every good cell in our dissection. The additional vertices and edges will be chosen so that they allow us (1) to combine spanning cycles in neighboring cells and (2) to add vertices in bad cells to our spanning cycle. See Section 5.3 for details.

Lemma 5.7 *For every ℓ_1, ℓ_2 there exists an $s = s(\ell_1, \ell_2)$ such that the following holds. Suppose G is a graph and $C, C_1, \ldots, C_{\ell_1} \subseteq G$ and $u_1, \ldots, u_{\ell_1}, v_1, \ldots, v_{\ell_1}, w_1, \ldots, w_{\ell_2} \in V(G)$ are such that:*

- *The vertices $u_1, \ldots, u_{\ell_1}, v_1, \ldots, v_{\ell_1}, w_1, \ldots, w_{\ell_2}$ are distinct;*
- *$u_i v_i \in E(G)$ for $i = 1, \ldots, \ell_1$;*
- *$C_1, \ldots, C_{\ell_1}$ are 10-blob cycles and C is an s-blob cycle;*
- *$C, C_1, \ldots, C_{\ell_1}$ are vertex disjoint and do not contain any of the u_i's, v_i's or w_i's;*
- *u_i, v_i both have at least $\lfloor |V(C_i)|/2 \rfloor$ neighbours on C_i for $i = 1, \ldots, \ell_1$;*
- *Every vertex of $\{w_1, \ldots, w_{\ell_2}\}$ has at least $\lfloor |V(C)|/2 \rfloor$ neighbours on C;*
- *For every $1 \leq i \leq \ell_1$, every vertex of the blob of C_i has at least $\lfloor |V(C)|/2 \rfloor$ neighbours on C.*

Then there is an $(s - 2\ell_1 - 2\ell_2)$-blob cycle C' such that

- *$V(C') = V(C) \cup V(C_1) \cup \cdots \cup V(C_{\ell_1}) \cup \{u_1, \ldots, u_{\ell_1}, v_1, \ldots, v_{\ell_1}, w_1, \ldots, w_{\ell_2}\}$, and;*
- *$u_1 v_1, \ldots, u_{\ell_1} v_{\ell_1} \in E(C')$, and;*
- *The blob of C' is contained in the blob of C.*

Proof: We prove the statement for $s := \max\{s_1(\ell_1), s_2(\ell_2)\} + 2\ell_1 + 2\ell_2$ where $s_1(\cdot)$ is the function provided by Lemma 5.5 and $s_2(\cdot)$ is as provided by Lemma 5.6.

First, we apply Lemma 5.4 to each triple u_i, v_i, C_i. For $1 \leq i \leq \ell_1$, this yields an 8-blob cycle C_i' containing the edge $u_i v_i$ and such that each vertex of the blob of C_i' has at least $\lfloor |V(C)|/2 \rfloor$ neighbours on C. Also, the edge $u_i v_i$ will not be part of the assumed 8-blob of C_i'.

Next, we join all of the w_i's with the s-blob cycle C. We apply Lemma 5.5 repeatedly, with $\ell = 2\ell_2$. This value of ℓ lets us pick distinct u_j, u_{j+1} for each w_i, one after another. We then add all the w_i to C in one fell swoop to get the $(s - 2\ell_1)$-blob cycle C' that contains all of the w_i's.

Finally, we merge C' with the $C_1', \ldots, C_{\ell_1}'$ in a similar manner. We apply Lemma 5.6 repeatedly, with $\ell = 2\ell_1 + 3\ell_2$. This choice of ℓ allows us to pick distinct u_j, u_{j+1} for each C_i'. We then merge $C_1', \ldots, C_{\ell_1}$ with C in one fell swoop to get the desired $(s - 2\ell_1 - 2\ell_2)$-blob cycle C''. ■

5.3 Proof of the Hamilton cycle game

We again divide the remainder of the proof of Theorem 1.5 into a deterministic and a probabilistic part. The following lemma is the deterministic part:

Lemma 5.8 *There exists a value of T such that the following holds. If $V \subseteq [0,1]^2, m \in \mathbb{N}, r > 0$ and $r \leq \rho \leq 2r$ are such that*

- **(i)** **(str-1)**–**(str-6)** *hold with respect to r, and;*
- **(ii)** *For all $2 \leq s \leq T$, every s-obstruction with respect to r has at least $s+2$ crucial vertices, and;*
- **(iii)** *Every $(\geq T)$-obstruction with respect to r has at least six crucial vertices, and;*
- **(iv)** *$G(V, \rho)$ has minimum degree at least four.*

Then Maker wins the Hamilton cycle game on $G(V; \rho)$.

Before launching into the proof, we give a high-level overview of the argument. We use the spanning tree $\mathcal{T}$ of $\Gamma_{\max}$ from Lemma 2.12 as the skeleton of the Hamilton cycle. Maker plays many local mini-games, and then stitches together the local structures to create the Hamilton cycle. Figure 8 shows a simplified version of the mini-games in a good cell. We *mark* a large, but finite, number of vertices in good cells. The majority of the marked vertices are *important vertices* in good cells that are associated to *critical vertices* for bad cells. Some marked vertices are reserved for joining the blob cycles in good cells that are adjacent in $\mathcal{T}$. Maker plays the game so that (1) in each cell, he can create a family of blob cycles spanning the unmarked vertices (Lemma 5.3, multiple times), (2) there are two independent edges between good cells that are adjacent in $\mathcal{T}$, and (3) he can construct a family of paths through the bad cells. In particular, every vertex in a bad cell will be on exactly one path between a pair of marked vertices that are in the same good cell. At this point, we place the marked vertices into two categories. Pairs of marked vertices that are endpoints of a special path are temporarily considered *marked edges $u_i v_i$*. Note that the path games create one or two paths through each obstruction. The unused marked vertices w_j are considered as *marked singletons.*

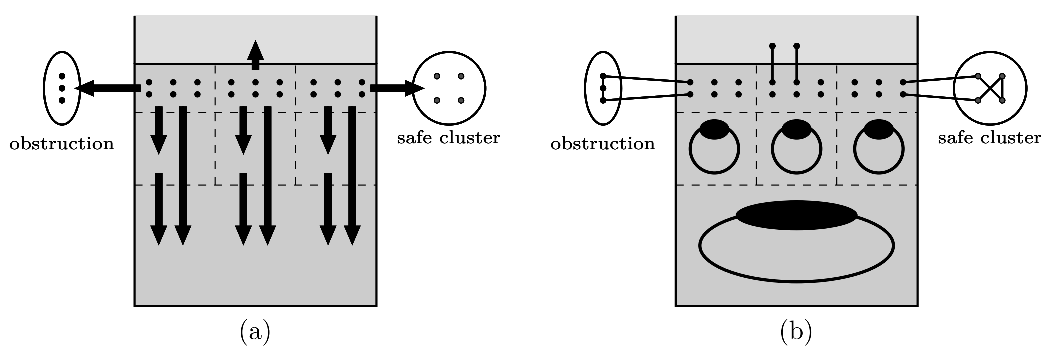

Figure 8: A schematic for the mini-games for Maker's Hamilton strategy in a good cell. (a) The marked vertices (top row) are designated to make paths through nearby obstructions and safe clusters, and to connect to good cells that are adjacent in the tree $\mathcal{T}$. The unmarked vertices (bottom two rows) are partitioned into subsets for blob cycle creations. Maker claims half the edges from each vertex to each lower level. (b) After all edges have been claimed, Maker has paths through obstructions and two independent edges to nearby good cells, and some unused marked vertices, and a soup of blob cycles. The blob cycles absorb the paths and unused vertices (using the many edges down to the lower levels in the good cell), culminating in the Hamilton cycle.

We construct our Hamilton cycle as follows. For each good cell, we conglomerate its vertices via Lemma 5.7 using the family of blob cycles and the marked edges u_iv_i and marked singletons w_j. This creates a cycle that spans all unmarked vertices, all marked edges and all marked singletons. Finally, we replace each marked edge with the graph structure it represents. This adds all of the bad vertices to our cycle, and connects every cell in $\Gamma_{\max}$. The result is the desired Hamilton cycle. With this outline in mind, we proceed with the proof.

Proof: Let T be a (large but finite) number, to be made explicit later on in the proof. For each good cell of our dissection, we will identify at most T vertices to help us with nearby vertices in bad cells. Collectively, we will refer these vertices as *marked*. The remaining vertices are *unmarked*, and these unmarked vertices will be used to create blob cycles within the good cell.

As in the proof of Lemma 4.1, observe that $\Gamma(V; m, T, \rho)$ satisfies **(str-1)**–**(str-6)** if we modify **(str-3)**–**(str-5)** very slightly by replacing the number $r \cdot 10^{10}$ by $r \cdot 10^{10}/2$. Also observe that items **(ii)** and **(iii)** clearly also hold with respect to ρ. Again all mention of safe, dangerous, obstructions and so on will be with respect to ρ from now on.

Before the game starts, Maker identifies many local games that he will play. These games fall into four categories, according to the types of vertices involved. We have games between good vertices in neighboring good cells; games for bad vertices and nearby marked good vertices; games for unmarked vertices within a good cell; and games between marked and unmarked vertices within a good cell. We now describe these games in detail.

Lemma 2.12 implies that $\Gamma_{\max}$ has a spanning tree $\mathcal{T}$ of maximum degree at most five. Before the game, we fix such a spanning tree. Maker will use this spanning tree as the skeleton for the Hamilton cycle. For each edge cc' in the spanning tree, we identify four vertices in each cell that are important for the edge cc' and are considered marked. Since $\mathcal{T}$ has maximum degree 5, we mark at most 20 vertices in each good cell. With T being large, we can and do take all these marked vertices distinct.

Maker keeps track of each obstruction. By Lemma 3.4 (with $k = 4$), every s-obstruction has at least $2 + s \geq 4$ crucial vertices for when $2 \leq s < T$, or at least 4 crucial vertices when $s \geq T$. Note that a 2-obstruction must have at least 4 crucial vertices. This for instance means we cannot have two vertices of degree four that are joined by an edge (which would have given Breaker a winning strategy).

To each obstruction, we assign either all of its crucial vertices if there are fewer than six, or six of its crucial vertices, if there are more. (No vertex can be crucial for multiple obstructions: obstructions are well-separated by **(str-3)** – **(str-5)**, so a cell contains crucial vertices for at most one obstruction.) Each of these crucial vertices has at least T important vertices inside some cell of $\Gamma_{\max}$. We assign four important vertices (all in the same cell) to each crucial vertex. (As above, we choose all important vertices to be distinct.) Every important vertex and every crucial vertex in a good cell is considered marked. In G, vertices in a good cell are adjacent to at most one obstruction, so this adds at most $6 + 24 = 30$ marked vertices to each cell.

We must also consider pairs of important vertices assigned to the obstruction. Let $c, c' \in \Gamma_{\max}$ be the cells that the two quadruples of important vertices lie in (these important vertices belong to different crucial vertices, so we might have $c \neq c'$), as shown in Figure 9. This gives an upper bound on their distance, so by **(str-6)** there is a (short) path Π between them. We orient all edges of the path towards one of the endpoints, say c. Inside each cell of the path we assign two vertices to Π. These vertices are considered marked. A given good cell c'' is near at most one obstruction and the path is short, so there are at most $\binom{6}{2}$ such paths through c''. Thus, we mark no more than $2\binom{6}{2} = 30$ additional vertices in c'' (choosing distinct vertices every time because T is large). This completes our marking of vertices for obstructions.

Next, we deal with safe vertices that are not in a cell of $\Gamma_{\max}$. Each such vertex v has at least T neighbours inside some cell $c \in \Gamma_{\max}$. We assign v arbitrarily to one such c. Next, for each cell c, we consider the set of safe points assigned to it. We partition these safe vertices into at most 36 cliques by centering a $3r \times 3r$ square on the center of c, and dividing it into $(r/2) \times (r/2)$ squares. We will refer to each of these cliques as a *safe cluster*. For each safe cluster we do the following. If it has more than six members, we declare six of them "crucial". For each such crucial vertex

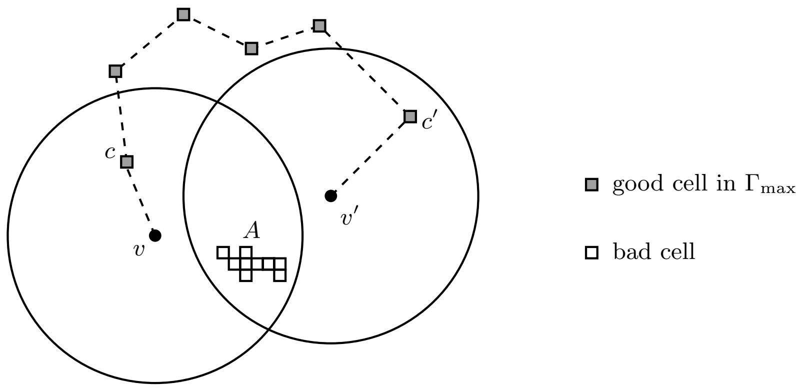

Figure 9: An obstruction A with crucial vertices v, v'. These vertices have important vertices in distinct cells c and c'. There is a short path between c and c' in $\Gamma_{\max}$, and we mark two vertices in each cell of this path.

w, we pick four of its neighbours in c (different from all marked vertices in c, since T is large). We declare these vertices to be *important for* w. If a safe cluster has at most six members, then we consider these vertices to be singleton safe clusters. We assign four *important* vertices in c to each of these vertices. These important vertices are considered marked. In the worst case (where every safe cluster size is at most six), we mark $36 \cdot 6 \cdot 4 = 864$ vertices in c.

We have accounted for all vertices outside of $\Gamma_{\max}$, and marked fewer than 1000 vertices in each good cell. The path games associate one or two paths to each obstruction or safe cluster, so we end up with at most 200 marked edges. Finally, we address all unmarked vertices in cells of $\Gamma_{\max}$. Inside each cell $c \in \Gamma_{\max}$, we partition the unmarked points into sets $C_0(c), C_1(c), \ldots, C_\ell(c)$ where $|C_i(c)| = N(10)$ for $i = 1, \ldots, \ell$, and $|C_0(c)| > N(s)$, with $N(10)$ and $N(q)$ as in Lemma 5.3, where $\ell = 1000$ and $s = s(200, 1000)$ from Lemma 5.7 are both constants. We can now specify our constant T, which must allow us to mark our distinct vertices, and to make the blob cycles. Choosing $T = 1000 + \ell N(10) + N(s)$ is sufficient.

This completes Maker's organization of the graph. During the game, Maker plays as follows:

(i) Every edge cc' of $\mathcal{T}$ has important vertices $x_1, x_2, x_3, x_4 \in c$ and $y_1, y_2, y_3, y_4 \in c'$. We pair the edges x_1y_1 and x_2y_2. When Breaker claims one of them, Maker responds by claiming the other one. Likewise, we pair edge x_3y_3 and x_4y_4. Therefore Maker claims two independent edges between c and c'.

(ii) For every crucial vertex, we pair the edges to the four important vertices assigned to it. When Breaker claims one of them, Maker responds by claiming the other one. Therefore, Maker claims at least two of these edges.

(iii) If Breaker claims an edge inside an obstruction O together with the crucial vertices C assigned to that obstruction, then Maker responds according to his winning strategy for the corresponding (a, b)-path game (Lemma 5.2), where $A = O$ and $B = C$.

(iv) If Breaker claims an edge inside a safe cluster S together with its important vertices I, Maker again responds according to his winning strategy for the (a, b)-path game, this time with $A = S$ and $B = I$.

(v) Suppose c_1c_2 is a directed edge of a path Π between two cells c, c' containing important vertices for an obstruction. Let $u_1, u_2 \in c_1$ and $v_1, v_2 \in c_2$ be the vertices assigned to Π. Maker pairs u_1v_1, u_1v_2 and pairs u_2v_1, u_2v_2.

(vi) Similarly if $u_1, u_2, u_3, u_4 \in c$ are important for an obstruction, and $v_1, v_2 \in c$ are assigned to a path Π in $\Gamma_{\max}$ then Maker pairs u_iv_1, u_iv_2 for $i = 1, \ldots, 4$.

(vii) If Breaker claims an edge between two vertices of $C_i(c)$ for some $1 \leq i \leq \ell$ and $c \in \Gamma_{\max}$ then Maker claims another edge in $C_i(c)$ according to his winning strategy for the 10-blob game (Lemma 5.3).

(viii) Likewise, if Breaker claims an edge inside $C_0(c)$ for some $c \in \Gamma_{\max}$ then Maker responds according to his winning strategy for the s-blob game (Lemma 5.3).

(ix) If Breaker claims an edge uv with $u \in c \setminus (C_0(c) \cup \cdots \cup C_\ell(c))$ for some $c \in \Gamma_{\max}$ and $v \in C_i(c)$ then Maker claims another edge between u and a vertex of $C_i(c)$.

(x) If Breaker claims an edge between $u \in C_0(c)$ and $v \in C_i(c)$ with $c \in \Gamma_{\max}$ and $1 \leq i \leq \ell$, then Maker claims another edge wv with $w \in C_0(c)$.

(xi) For any other Breaker move, Maker responds arbitrarily.

We now prove that this is a winning strategy for Maker. We must show that after the game ends, Maker's graph contains a Hamilton cycle.

P1 First, by (iv), Maker will have won the (a, b)-game for each safe cluster S. There will be either two important vertices (in the same good cell c) connected by a path that spans the safe cluster, or two pairs of such vertices connected by vertex disjoint paths that span the cluster.

P2 Similarly, by (iii), Maker will have won the (a, b)-game for every obstruction. In other words, there will be one or two paths between pairs of crucial vertices that span the obstruction. By (ii), each of these path extends to a path between important vertices in two possibly different cells $c, c' \in \Gamma_{\max}$. By (v) and (vi), these paths also extend to a path between two marked vertices in the same cell $c \in \Gamma_{\max}$. Also, every vertex that got marked crucial but is not part of such a path, will be part of a path of length two between two of its important vertices.

P3 Let u, v be a pair of important vertices in the same good cell c that are the endpoints of a path through an obstruction or safe cluster. We treat uv as a marked edge. We also have a pair of vertices marked to connect to each adjacent cell in $\mathcal{T}$. The total number of marked edges is at most 200. Meanwhile any marked vertex w that is not part of one of these paths (as either an endpoint or an interior point) is treated as a marked singleton.

P4 Next, we consider the unmarked vertices in a good cell c. Strategies (vii) and (viii) ensure that Maker will have created the blob cycles $C_0(c), C_1(c), \ldots, C_\ell(c)$. These cycles span all unmarked vertices in c. By (ix) and (x), the family of blob cycles and the marked edges and the marked singletons satisfy the conditions of Lemma 5.7. Therefore, we can construct one cycle that spans all of the vertices in c, and includes every marked edge. At this point, we have a spanning cycle in every good cell, where this cycle used marked edges.

P5 Let uv be a marked edge in c. Recall that this marked edge uv is a placeholder for another structure in Maker's graph (in fact, Breaker may have claimed the actual edge between these vertices). First, if the vertices u, v were important for the spanning tree $\mathcal{T}$ of $\Gamma_{\max}$, then there is a corresponding marked edge $u'v'$ in the spanning cycle of c'. We replace these two marked edges by edges uu' and vv', which merges the spanning cycles in c and c'. Second, if the vertices u, v were important for an obstruction O, then we replace the marked edge uv with the spanning path through O whose endpoints are u and v. Third, we replace any marked edge associated with a safe cluster with the analogous path through that cluster. Once we have replaced all the marked edges, we have our Hamilton cycle.

■

Proof of Theorem 1.5: The proof is very similar to that of Theorem 1.1. It is clear that Breaker wins if there is a vertex of degree at most three (Breaker starts). Hence

$$\mathbb{P}\left(\rho_n(\text{Maker wins the Hamilton cycle game}) \geq \rho_n(\text{min. deg. } \geq 4)\right) = 1. \tag{25}$$

We now define:

$$r_L(n) := \left(\frac{\ln n + 5\ln\ln n - K}{\pi n}\right)^{\frac{1}{2}}, \quad r_U(n) := \left(\frac{\ln n + 5\ln\ln n + K}{\pi n}\right)^{\frac{1}{2}},$$

for K a (large) constant. By Theorem 2.10, we can choose $K = K(\varepsilon)$ such that

$$\mathbb{P}\Big(r_L(n) \leq \rho_n(\text{min. deg. } \geq 4) \leq r_U(n)\Big) \geq 1 - \varepsilon + o(1). \tag{26}$$

By Lemma 3.3 the properties **(str-1)**–**(str-6)** are satisfied with probability $1 - o(1)$ by $V = \mathcal{X}_n, m = m_n, T = O(1), r = r_L(n)$ with $\mathcal{X}_n$ as given by (4), m_n as given by (12) and r_L as above. By Lemma 3.4, with probability $1 - o(1)$, the remaining conditions of Lemma 5.8 are met for any $r \leq \rho \leq 2r$ with $\delta(G(V;\rho)) \geq 4$. Hence:

$$\mathbb{P}\left(\rho_n(\text{Maker wins the Hamilton cycle game}) = \rho_n(\text{min. deg. } \geq 4)\right) \geq 1 - \varepsilon - o(1).$$

Sending $\varepsilon \downarrow 0$ gives the theorem. ■

6 The perfect matching game

We start by considering the obvious obstructions preventing Maker-win: vertices v with degree $d(v) \leq 1$, and edges uv with edge-degree $d(uv) = |(N(v) \cup N(u)) \setminus \{u, v\}| \leq 2$. Indeed, with Breaker first to move, he can isolate a vertex of degree one. For an edge with at most two neighbouring vertices, Breaker can ensure that one vertex remains unmatched by Maker.

Lemma 6.1 *Let $(r_n)_n$ be such that*

$$\pi n r_n^2 = \ln n + \ln\ln n + x + o(1),$$

for some $x \in \mathbb{R}$. Let Z_n denote the number of vertices of degree exactly one plus the number of edges of edge-degree exactly two in $G_{\mathcal{P}}(n, r_n)$. Then

$$\mathbb{E}Z_n \to (1 + \pi^2/8)e^{-x} + \sqrt{\pi}(1 + \pi)e^{-x/2}$$

as $n \to \infty$.

Proof: Let Y_n denote the number of vertices of degree exactly one and let W_n denote the number of edges of edge-degree exactly two. From Theorem 2.10 and Lemma 2.11, we have

$$\mathbb{E}Y_n = e^{-x} + \sqrt{\pi e^{-x}} + o(1). \tag{27}$$

It thus remains to compute $\mathbb{E}W_n$. Let W_{cnr} denote the number of edge-degree two edges $\{X_i, X_j\}$ with both vertices X_i, X_j within $100r$ of two of the sides of $[0,1]^2$. By Lemma 3.5 using $k = 2$ and condition (iii) (a), we have

$$\mathbb{E}W_{\text{cnr}} = O\left(n^{-O(1)}\right). \tag{28}$$

Let W_{mdl} denote the number of edge-degree two edges $\{X_i, X_j\}$ for which both X_i and X_j are at least r removed from the boundary of the unit square. For $0 \leq z \leq r$, let us write

$$\mu(z) = n \cdot \text{area}(B(u;r) \cup B(v;r)),$$

where $u, v \in \mathbb{R}^2$ are such that $\|u-v\| = z$. Let $\varepsilon > 0$ be arbitrary. By equation (7) of Corollary 2.15, there is a $\delta = \delta(\varepsilon)$ such that

$$\begin{array}{c} \pi n r^2 + (2-\varepsilon) n r z \leq \mu(z) \leq \pi n r^2 + (2+\varepsilon) n r z, \\ \text{and} \\ (1-\varepsilon)\pi n r^2 \leq \mu(z) \leq (1+\varepsilon)\pi n r^2, \end{array} \tag{29}$$

for all $0 \leq z \leq \delta r$. Let W_{mdl}^{δ} denote the number of edge-degree two edges $\{X_i, X_j\}$ for which both X_i, X_j are at least r removed from the boundary of the unit square, and $\|X_i - X_j\| \leq \delta r$. By Lemma 3.5 using $k = 2$ and condition (iii) (c), we have

$$\mathbb{E}\left(W_{\text{mdl}} - W_{\text{mdl}}^{\delta}\right) = O\left(n^{-O(1)}\right). \tag{30}$$

Let us now compute W_{mdl}^{δ}. We have

$$\begin{aligned} \mathbb{E}W_{\text{mdl}}^{\delta} &\leq \frac{1}{2} \cdot n^2 \int_{[0,1]^2} \int_{B(v;\delta r)} \frac{\mu(\|u-v\|)^2 e^{-\mu(\|u-v\|)}}{2} \mathrm{d}u\mathrm{d}v \\ &= \frac{1}{2} \cdot n^2 \int_0^{\delta r} \frac{\mu(z)^2 e^{-\mu(z)}}{2} 2\pi z \mathrm{d}z \\ &\leq \frac{1}{2} n^2 \int_0^{\delta r} \frac{\left((1+\varepsilon)\pi n r^2\right)^2 e^{-\pi n r^2 - (2-\varepsilon) n r z}}{2} 2\pi z \mathrm{d}z \\ &= (1+o(1)) \cdot \frac{(1+\varepsilon)^2 \pi}{2} \cdot n^2 \cdot \ln^2 n \cdot e^{-\pi n r^2} \int_0^{\delta r} e^{-(2-\varepsilon) n r z} z \mathrm{d}z \\ &= (1+o(1)) \cdot \frac{(1+\varepsilon)^2 \pi e^{-x}}{2} \cdot n \cdot \ln n \int_0^{\delta r} e^{-(2-\varepsilon) n r z} z \mathrm{d}z \\ &= (1+o(1)) \cdot \frac{(1+\varepsilon)^2 \pi e^{-x}}{2} \cdot n \cdot \ln n \cdot \left(\frac{1}{(2-\varepsilon) n r}\right)^2 \int_0^{(2-\varepsilon)\delta n r^2} e^{-y} y \mathrm{d}y \\ &= (1+o(1)) \cdot \frac{(1+\varepsilon)^2 \pi e^{-x}}{2(2-\varepsilon)^2} \cdot \frac{\ln n}{n r^2} \cdot \int_0^{\infty} e^{-y} y \mathrm{d}y \\ &= (1+o(1)) \cdot \frac{(1+\varepsilon)^2 \pi^2 e^{-x}}{2(2-\varepsilon)^2}. \end{aligned} \tag{31}$$

Here the factor $\frac{1}{2}$ in the first line comes from the fact that Theorem 2.7 applies to *ordered* pairs; to get the second line we switched to polar coordinates; to get the third line we applied the bounds (29); to get the fourth line we used that $\pi n r^2 = (1+o(1)) \ln n$; to get the fifth line we used that $\pi n r^2 = \ln n + \ln \ln n + x + o(1)$ so that $e^{-\pi n r^2} = (1+o(1)) e^{-x} / n \ln n$; to get the sixth line we used the change of variables $y = (2-\varepsilon) n r z$; and in the last two lines we used that $n r^2 \to \infty$ so that $\int_0^{(2-\varepsilon)\delta n r^2} e^{-y} y \mathrm{d}y \to \int_0^{\infty} e^{-y} y \mathrm{d}y = 1$.

By analogous computations we get

$$\begin{aligned} \mathbb{E}W_{\text{mdl}} &\geq \mathbb{E}W_{\text{mdl}}^{\delta} \\ &\geq \frac{1}{2} \cdot n^2 \int_{[100r, 1-100r]^2} \int_{B(v;\delta r)} \frac{\mu(\|u-v\|)^2 e^{-\mu(\|u-v\|)}}{2} \mathrm{d}u\mathrm{d}v \\ &= (1-o(1)) \cdot \frac{n^2}{2} \int_0^{\delta r} \frac{\mu(z)^2 e^{-\mu(z)}}{2} 2\pi z \mathrm{d}z \\ &\geq (1-o(1)) \cdot \frac{n^2}{2} \int_0^{\delta r} \frac{\left((1-\varepsilon)\pi n r^2\right)^2 e^{-\pi n r^2 - (2+\varepsilon) n r z}}{2} 2\pi z \mathrm{d}z \\ &= (1-o(1)) \cdot \frac{(1-\varepsilon)^2 \pi^2 e^{-x}}{2(2+\varepsilon)^2}. \end{aligned} \tag{32}$$

Combining (30), (31) and (32) and sending $\varepsilon \downarrow 0$, we find

$$\mathbb{E}W_{\text{mdl}} = (1+o(1)) \cdot \frac{\pi^2 e^{-x}}{8}. \tag{33}$$

Finally, we must consider the remaining four rectangles of width r that are adjacent to exactly one border. Let W_{sde} denote the number of edges $\{X_i, X_j\}$ with edge-degree two such that at least one of X_i, X_j is no more than r away from some side of the unit square, and both are at least $100r$ away from all other sides; and let W_{sde}^{δ} denote all such pairs for which in addition at least one of X_i, X_j is no more than δr away from a side of the unit square, and $\|X_i - X_j\| \leq \delta r$. By Lemma 3.5 using $k = 2$ with the union of conditions (iii)(b) and (iii)(c), we have

$$\mathbb{E}\left(W_{\text{sde}} - W_{\text{sde}}^{\delta}\right) = O\left(n^{-O(1)}\right). \tag{34}$$

For $0 \leq w, z \leq r$ and $-\pi/2 \leq \alpha \leq \pi/2$, let us write:

$$\mu(w, z, \alpha) := n \cdot \text{area}([0,1]^2 \cap (B(u;r) \cup B(v;r))),$$

where $u, v \in [0,1]^2$ are such that $u_x = w < v_x$, $\|u - v\| = z$ and the angle between $v - u$ and the positive x-axis is α (see Figure 1). Fix $\varepsilon > 0$. By Lemma 2.18 there is a $\delta = \delta(\varepsilon)$ such that for all $0 \leq w, z \leq \delta r$ and all $-\pi/2 \leq \alpha \leq \pi/2$:

$$\begin{array}{c} \frac{\pi}{2}nr^2 + (1 + \cos\alpha - \varepsilon)nrz + (2 - \varepsilon)nwr \\ \leq \\ \mu(w, z, \alpha) \\ \leq \\ \frac{\pi}{2}nr^2 + (1 + \cos\alpha + \varepsilon)nrz + (2 + \varepsilon)nwr, \end{array} \tag{35}$$

and

$$(1 - \varepsilon)\frac{\pi}{2}nr^2 \leq \mu(w, z, \alpha) \leq (1 + \varepsilon)\frac{\pi}{2}nr^2. \tag{36}$$

Let $B^+(v;r)$ denote the set of all $p \in B(v;r)$ with $p_x \geq v_x$. We can write

$$\begin{aligned}
\mathbb{E}W_{\text{sde}}^{\delta} &\leq 4 \cdot n^2 \int_{[0,\delta r] \times [100r, 1-100r]} \int_{B^+(v;r)} \mathbb{P}\left(\text{Po}([0,1]^2 \cap (B(u;r) \cup B(v;r))) = 2\right) \mathrm{d}u\mathrm{d}v \\
&= (1 + o(1)) \cdot 4n^2 \cdot \int_{-\pi/2}^{\pi/2} \int_0^{\delta r} \int_0^{\delta r} \frac{(\mu(w,z,\alpha))^2}{2} e^{-\mu(w,z,\alpha)} z \, \mathrm{d}w\mathrm{d}z\mathrm{d}\alpha \\
&\leq (1 + o(1)) \cdot 4n^2 \cdot \int_{-\pi/2}^{\pi/2} \int_0^{\delta r} \int_0^{\delta r} \frac{\left((1+\varepsilon)\frac{\pi}{2}nr^2\right)^2}{2} e^{-\frac{\pi}{2}nr^2 - (1+\cos\alpha-\varepsilon)nrz - (2-\varepsilon)nrw} z \, \mathrm{d}w\mathrm{d}z\mathrm{d}\alpha \\
&= (1 + o(1)) \cdot \frac{(1+\varepsilon)^2}{2} \cdot n^2 \cdot \ln^2 n \cdot e^{-\frac{\pi}{2}nr^2} \cdot \int_{-\pi/2}^{\pi/2} \int_0^{\delta r} \int_0^{\delta r} e^{-(1+\cos\alpha-\varepsilon)nrz - (2-\varepsilon)nrw} z \, \mathrm{d}w\mathrm{d}z\mathrm{d}\alpha \\
&= (1 + o(1)) \cdot \frac{(1+\varepsilon)^2}{2} \cdot n^{\frac{3}{2}} \cdot \ln^{\frac{3}{2}} n \cdot e^{-x/2} \cdot \int_{-\pi/2}^{\pi/2} \int_0^{\delta r} \int_0^{\delta r} e^{-(1+\cos\alpha-\varepsilon)nrz - (2-\varepsilon)nrw} z \, \mathrm{d}w\mathrm{d}z\mathrm{d}\alpha \\
&= (1 + o(1)) \cdot \frac{(1+\varepsilon)^2 e^{-x/2}}{2} \cdot n^{\frac{3}{2}} \cdot \ln^{\frac{3}{2}} n \cdot \left(\frac{1}{(2-\varepsilon)nr}\right) \cdot \left[1 - e^{-(2-\varepsilon)\delta nr^2}\right] \cdot \\
&\quad \int_{-\pi/2}^{\pi/2} \int_0^{\delta r} e^{-(1+\cos\alpha-\varepsilon)nrz} z \, \mathrm{d}z\mathrm{d}\alpha \\
&= (1 + o(1)) \cdot \frac{(1+\varepsilon)^2 e^{-x/2}}{2(2-\varepsilon)} \cdot n^{\frac{3}{2}} \cdot \ln^{\frac{3}{2}} n \cdot (nr)^{-1} \cdot \\
&\quad \int_{-\pi/2}^{\pi/2} \left(\frac{1}{(1+\cos\alpha-\varepsilon)nr}\right)^2 \int_0^{(1+\cos\alpha-\varepsilon)\delta nr^2} y \, e^{-y} \mathrm{d}y\mathrm{d}\alpha \\
&= (1 + o(1)) \cdot \frac{(1+\varepsilon)^2 e^{-x/2}}{2(2-\varepsilon)} \cdot n^{\frac{3}{2}} \cdot \ln^{\frac{3}{2}} n \cdot (nr)^{-3} \cdot \int_{-\pi/2}^{\pi/2} \left(\frac{1}{1+\cos\alpha-\varepsilon}\right)^2 \mathrm{d}\alpha
\end{aligned}$$

Here we have once again used Theorem 2.7 in the first line, together with symmetry considerations. The first line gives an upper bound on W_{sde}^{δ} since a vertex in $B(u;r) \cap B(v;r)$ would increase the edge degree by 2 instead of 1. In the second line we switched to polar coordinates; in the third line we used the bounds (35) and (36); in the fourth line we used $\pi nr^2 = (1 + o(1)) \ln n$;

in the fifth line we used $\pi n r^2 = \ln n + \ln\ln n + x + o(1)$; in the sixth line we integrated with respect to w; in the seventh line we applied the substitution $y = (1+\cos\alpha - \varepsilon)nrz$; in the eight line we used that $\int_0^{(1+\cos\alpha-\varepsilon)\delta n r^2} y\, e^{-y} \mathrm{d}y \to \int_0^{\infty} y\, e^{-y} y \mathrm{d}y = 1$. We get:

$$\begin{aligned} \mathbb{E}W_{\mathrm{sde}}^{\delta} &\leq (1+o(1)) \cdot \frac{(1+\varepsilon)^2 e^{-x/2}}{2(2-\varepsilon)} \cdot \ln^{\frac{3}{2}} n \cdot (nr^2)^{-\frac{3}{2}} \cdot \int_{-\pi/2}^{\pi/2} \left(\frac{1}{1+\cos\alpha-\varepsilon}\right)^2 \mathrm{d}\alpha \\ &= (1+o(1)) \cdot \frac{(1+\varepsilon)^2 e^{-x/2} \pi^{\frac{3}{2}}}{2(2-\varepsilon)} \cdot \int_{-\pi/2}^{\pi/2} \left(\frac{1}{1+\cos\alpha-\varepsilon}\right)^2 \mathrm{d}\alpha \end{aligned} \tag{37}$$

Reversing the use of the upper and lower bounds from (35) and (36), and repeating the computations giving (37) we find

$$\begin{aligned} \mathbb{E}W_{\mathrm{sde}} &\geq \mathbb{E}W_{\mathrm{sde}}^{\delta} \\ &\geq (1+o(1)) \cdot 4n^2 \cdot \int_{-\pi/2}^{\pi/2} \int_0^{\delta r} \int_0^{\delta r} \frac{\left((1-\varepsilon)\frac{\pi}{2} n r^2\right)^2}{2} e^{-\frac{\pi}{2} n r^2 - (1+\cos\alpha+\varepsilon)nrz - (2+\varepsilon)nrw} \mathrm{d}w \mathrm{d}z \mathrm{d}\alpha \\ &= (1+o(1)) \cdot \frac{(1-\varepsilon)^2 e^{-x/2} \pi^{\frac{3}{2}}}{2(2+\varepsilon)} \cdot \int_{-\pi/2}^{\pi/2} \left(\frac{1}{1+\cos\alpha+\varepsilon}\right)^2 \mathrm{d}\alpha \end{aligned}$$

Combining this with (34) and (37) and sending $\varepsilon \downarrow 0$, we find (employing the dominated convergence theorem to justify switching limit and integral):

$$\mathbb{E}W_{\mathrm{sde}} = (1+o(1)) \cdot \frac{e^{-x/2}\pi^{\frac{3}{2}}}{4} \cdot \int_{-\pi/2}^{\pi/2} \left(\frac{1}{1+\cos\alpha}\right)^2 \mathrm{d}\alpha.$$

We compute

$$\int_{-\pi/2}^{\pi/2} \left(\frac{1}{1+\cos\alpha}\right)^2 \mathrm{d}\alpha = \left[\frac{\sin(x)}{2(\cos(x)+1)} + \frac{\sin^3(x)}{6(\cos(x)+1)^3}\right]_{-\pi/2}^{\pi/2} = \frac{4}{3}.$$

Hence

$$\mathbb{E}W_{\mathrm{sde}} = (1+o(1)) \cdot \frac{e^{-x/2}\pi^{\frac{3}{2}}}{3}. \tag{38}$$

Combining (28), (33) and (38) shows

$$\mathbb{E}W = \frac{\pi^2 e^{-x}}{8} + \frac{e^{-x/2}\pi^{\frac{3}{2}}}{3} + o(1)$$

Together with (27) this proves the Lemma. ■

The following lemma can be proved via a straightforward adaptation of the proof of Theorem 6.6 in [21]. For completeness we provide a proof in the appendix.

Lemma 6.2 *Let $n \in \mathbb{N}$ and $r > 0$ be arbitrary. Let Z denote the number of vertices of degree exactly one in $G_P(n,r)$ plus the number of edges of edge-degree exactly two in $G_P(n,r)$. Then*

$$d_{TV}(Z, \mathrm{Po}(\mathbb{E}Z)) \leq 6 \cdot (I_1 + I_2 + I_3 + I_4 + I_5 + I_6)$$

where

$$
\begin{aligned}
I_1 &:= n^2 \int_{[0,1]^2} \int_{B(x;100r)} \mathbb{P}(E_x)\mathbb{P}(E_y)\mathrm{d}y\mathrm{d}x,\\
I_2 &:= n^2 \int_{[0,1]^2} \int_{B(x;100r)} \mathbb{P}(E_x^y, E_y^x)\mathrm{d}y\mathrm{d}x,\\
I_3 &:= n^4 \int_{A_3} \mathbb{P}(F_{x_1,y_1})\mathbb{P}(F_{x_2,y_2})\mathrm{d}y_2\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}x_1,\\
I_4 &:= n^4 \int_{A_3} \mathbb{P}(F_{x_1,y_1}^{x_2,y_2}, F_{x_2,y_2}^{x_1,y_1})\mathrm{d}y_2\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}x_1,\\
I_5 &:= n^3 \int_{A_5} \mathbb{P}(F_{x_1,y_1})\mathbb{P}(E_{x_2})\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}x_1,\\
I_6 &:= n^3 \int_{A_5} \mathbb{P}(F_{x_1,y_1}^{x_2}, E_{x_2}^{x_1,y_1})\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}x_1.
\end{aligned}
$$

Here

$$
\begin{aligned}
A_3 := \{(x_1,y_1,x_2,y_2) \in \left([0,1]^2\right)^4 \quad &: \quad y_1 \in B(x_1,r), y_2 \in B(x_2;r),\\
&\quad\ \ x_2,y_2 \in B(x_1;100r) \cap B(y_1;100r)\},
\end{aligned}
$$

$$
A_5 := \{(x_1,y_1,x_2) \in \left([0,1]^2\right)^3 : y_1 \in B(x_1,r), x_2 \in B(x_1;100r) \cap B(y_1;100r)\},
$$

and E_x denotes the event that one point of $\mathcal{P}$ falls inside $B(x;r)$, and E_x^y the event that one point of $\{y\} \cup \mathcal{P}$ falls inside $B(x;r)$ and $E_x^{y,z}$ denotes the event that one point of $\{y,z\} \cup \mathcal{P}$ falls inside $B(x;r)$; $F_{x,y}$ denotes the event that two points of $\mathcal{P}$ fall inside $B(x;r) \cup B(y;r)$, and $F_{x,y}^w, F_{x,y}^{w,z}$ are defined similarly as $E_x^y, E_x^{y,z}$

Lemma 6.3 *Let $(r_n)_n$ be such that*

$$
\pi n r_n^2 = \ln n + \ln\ln n + x + o(1),
$$

for some $x \in \mathbb{R}$. Let Z_n denote the number of vertices of degree **exactly** *one, plus the number of edges of edge-degree* **exactly** *two in $G_{\mathcal{P}}(n, r_n)$. Then*

$$
\mathbb{P}(Z_n = 0) \to \exp\left[-(1+\pi^2/8)e^{-x} + \sqrt{\pi}(1+\pi)e^{-x/2}\right],
$$

as $n \to \infty$.

Proof: It suffices to show that if $I_1, \ldots, I_6$ is as in Lemma 6.2, then $I_1, \ldots, I_6 \to 0$ as $n \to \infty$. Observe that I_1, I_5 is most

$$
\begin{aligned}
I_1, I_5 &\le \mathbb{E}Z_n \cdot \pi n(100 r_n)^2 \cdot \mathbb{P}(\mathrm{Po}(\pi n r_n^2/4) \le 1)\\
&= O\left(\ln n \cdot \exp[-(\tfrac{1}{4} + o(1))\ln n \cdot H(4/\pi n r_n^2)]\right)\\
&= O\left(n^{-\frac{1}{4}+o(1)}\right),
\end{aligned}
$$

where we used Lemma 2.5 and the fact that for r sufficiently small at least one quarter of $B(v;r)$ is contained in the unit square for all $v \in [0,1]^2$.

Similarly

$$
\begin{aligned}
I_3 &\le \mathbb{E}Z_n \cdot \left(\pi n(100 r_n)^2\right)^2 \cdot \mathbb{P}(\mathrm{Po}(\pi n r_n^2/4) \le 1)\\
&= O\left(n^{-\frac{1}{4}+o(1)}\right).
\end{aligned}
$$

Notice that I_2 is the expected number of (ordered) pairs of points (u,v) with $\|u-v\| \le 100 r_n$ and $d(u) = d(v) = 1$. Observe that if moreover $\|u-v\| \le r/100$, then (u,v) are in fact an $(0,0)$-pair. Thus, using Lemma 3.6 and Lemma 3.5:

$$
I_2 \le O\left(\ln^{-1} n\right) + O\left(n^{-c}\right) = o(1).
$$

Similarly, I_4 equals the expected number of 4-tuples of points $(u_1, \dots, u_4)$ with all distances $\|u_i - u_j\| \leq 100r$ and $\|u_1 - u_2\|, \|u_3 - u_4\| \leq r$, and u_1u_2, u_3u_4 each having edge-degree equal to two. Observe if in such 4-tuple that $\|u_i - u_j\| \leq r/100$ for all $1 \leq i, j \leq 4$, then one of the pairs (u_i, u_j) will in fact be a $(0,0,2)$-pair. Also observe that for each $(0,0,2)$-pair contributes at most six 4-tuples to the number of our 4-tuples with all distances $\leq r/100$. Thus, using again Lemma 3.6 and Lemma 3.5:

$$I_4 \leq 6 \cdot O\left(\ln^{-4} n\right) + O\left(n^{-c}\right) = o(1).$$

Finally, I_6 equals the expected number of 3-tuples (u_1, u_2, u_3) with all distances at most $100r$, $\|u_1 - u_2\| \leq r$, the degree of u_3 equal to two, and the edge-degree of u_1u_2 equal to two. Observe that, if all distances in such a 3-tuple are $\leq r/100$ then one of $(u_1, u_2), (u_1, u_3), (u_2, u_3)$ will be a $(0,1,1)$-pair. Also, for every $(0,1,1)$-pair, there are at most three of our 3-tuples with all distances $\leq r/100$. Hence, using Lemma 3.6 and Lemma 3.5 again:

$$I_6 \leq 3 \cdot O\left(\ln^{-3} n\right) + O\left(n^{-c}\right) = o(1).$$

This shows that

$$d_{\mathrm{TV}}(Z, \mathrm{Po}(\mathbb{E}Z)) = o(1).$$

So in particular

$$\mathbb{P}(Z = 0) = e^{-\mathbb{E}Z} + o(1) = e^{-(1+\frac{\pi^2}{8})e^{-x} + \sqrt{\pi}(1+\pi)e^{-x/2}} + o(1),$$

using Lemma 6.1. ■

Corollary 6.4 *Let $(r_n)_n$ be such that*

$$\pi n r_n^2 = \ln n + \ln \ln n + x + o(1),$$

for some $x \in \mathbb{R}$. Let Z_n denote the number of vertices of degree **at most** *one, plus the number of edges of edge-degree* **at most** *two in $G_{\mathcal{P}}(n, r_n)$. Then*

$$\mathbb{P}(Z_n = 0) \to \exp\left[-((1 + \pi^2/8)e^{-x} + \sqrt{\pi}(1+\pi)e^{-x/2})\right],$$

as $n \to \infty$.

Proof: Let Y denote the number of vertices of degree exactly one; let Y' denote the number of vertices of degree exactly zero; let W denote the number of edges of edge-degree exactly two, let W' denote the number of edges of edge-degree at most one, all in $G_{\mathcal{P}}(n, r_n)$. The probability that $Y' > 0$ can be read off from Theorem 2.10. By our choice of r_n, we have

$$\mathbb{P}(Y' > 0) = o(1). \tag{39}$$

Suppose that uv is an edge of edge-degree at most one. Then either $\|u - v\| > r/100$, or (u, v) is either a $(1,0,0)$-pair, a $(0,1,0)$-pair, a $(0,0,1)$-pair, or a $(0,0,0)$-pair. By Lemmas 3.5 and 3.6, we have

$$\mathbb{P}(W' > 0) \leq \mathbb{E}W' = O\left(n^{-c}\right) + O\left(\ln^{-1} n\right) = o(1). \tag{40}$$

Let Z_n be as in the statement of the Corollary. By (39), (40) we have

$$\mathbb{P}(Y + W = 0) - \mathbb{P}(Y' > 0) - \mathbb{P}(W' > 0) \leq \mathbb{P}(Z_n = 0) \leq \mathbb{P}(Y + W = 0),$$

and hence $\mathbb{P}(Z_n = 0) = \mathbb{P}(Y + W = 0) + o(1)$. The corollary now follows directly from Lemma 6.3. ■

Using Lemma 2.8, this last corollary immediately transfers also to the binomial case. (We can rephrase Z_n in terms of a measurable function $h_n(u, v, V)$ which equals one only if either $u \neq v$ and the pair forms an edge of edge-degree at most two or $u = v$ and the degree is at most one.)

Corollary 6.5 *Let $(r_n)_n$ be such that*

$$\pi n r_n^2 = \ln n + \ln\ln n + x + o(1),$$

for some $x \in \mathbb{R}$. Let $\tilde{Z}_n$ denote the number of vertices of degree **at most** *one, plus the number of edges of edge-degree* **at most** *two in $G(n, r_n)$. Then*

$$\mathbb{P}(\tilde{Z}_n = 0) \to \exp\left[-((1+\pi^2/8)e^{-x} + \sqrt{\pi}(1+\pi)e^{-x/2})\right],$$

as $n \to \infty$.

6.1 The proof of Theorem 1.3

We will again introduce an auxiliary game that will be helpful for the analysis of the perfect matching game on the random geometric graph. As in the (a,b) path game, The (a, b) *matching game* is played on the same graph $G_{a,b}$ as the (a, b) path game in Section 5.1. The vertices of $G_{a,b}$ are partitioned into sets A, B with $|A| = a, |B| = b$, where the only missing edges are the internal edges of B. Maker's objective is to create a matching that saturates all vertices in A (he does not care about the vertices in B). When we use this lemma for the perfect matching game, the a vertices will belong to an obstruction A and the b vertices will be important for A.

Lemma 6.6 *The (a, b) matching game is a win for Maker if one of the following conditions is met*

- **(i)** $b \geq 4$, *or;*
- **(ii)** $a \in \{2, 3\}$ *and* $b \geq 3$, *or;*
- **(iii)** $a = 1$ *and* $b \geq 2$.

Proof of part (i): If $b \geq 4$ then a winning strategy for Maker is as follows:

- Whenever Breaker plays an edge between two vertices of A then Maker responds claiming another edge inside A according to the strategy from Lemma 5.1 that will guarantee him that by the end of the game he will have a path in A that contains all but one vertex of A.
- If Breaker claims an edge between a vertex $u \in A$ and a vertex $v \in B$ then Maker claims an arbitrary unclaimed edge connecting u to a vertex in B.

If Maker cannot claim such an edge, then he claims an arbitrary edge (and we forget about it for the remainder of the game). At the end of the game, Maker's graph contains a path P through all but one vertex of A, and every vertex of A will have at least two neighbours in B. Thus, the path P contains a matching that covers all but at most two points of A, and the remaining (up to) two points can be covered by (at most two) vertex disjoint edges to vertices in B.
Proof of part (ii): If $a = 3$ then the strategy just outlined in the proof of part **(i)** also works. This time, at the end of the game, Maker's graph contains an edge between two vertices of A, and the remaining vertex of A has at least one neighbour in B.

Let us thus consider the case when $a = 2$. We can write $A = \{a_1, a_2\}$ and $B = \{b_1, \ldots, b_\ell\}$ with $\ell \geq 3$. First suppose that in Breaker's first move he does not claim a_1a_2. In that case Maker can claim a_1a_2 in his first move and win the game.

Hence we can assume that in his first move, Breaker claims a_1a_2. Maker now responds by claiming a_1b_1 in his first move and for the remainder of the game pairs the edges a_2b_2, a_2b_3, meaning that if Breaker plays one of these two edges then he claims the other (otherwise he plays arbitrarily). This way, Maker will clearly end up with a matching of the required type in the end.
Proof of part (iii): Maker wins in the obvious way. ■

Lemma 6.7 *There exists a value of T such that the following holds. If $V \subseteq [0,1]^2, m \in \mathbb{N}, r > 0$ and $r \le \rho \le 10^5 r$ are such that*

(i) **(str-1)**–**(str-5)** *hold with respect to r, and;*

(ii) *For all $3 \le s \le T$, every s-obstruction with respect to r has at least s crucial vertices, and;*

(iii) *Every $(\ge T)$-obstruction with respect to r has at least four crucial vertices, and;*

(iv) *Every edge of $G(V, \rho)$ has edge-degree at least three, and;*

(v) *$G(V, \rho)$ has minimum degree at least two.*

Then Maker wins the perfect matching game on $G(V; \rho)$.

Proof: The proof is a relatively straightforward adaptation of the proof of Lemma 5.8. In fact it is slightly simpler. This time, on each obstruction or safe cluster we play the corresponding (a, b)-matching game (which we win by Lemma 6.6). This will give a matching M_0 that saturates all vertices outside cells of $\Gamma_{\max}$ and some vertices of $\Gamma_{\max}$. A small change in **P4** will show that we will have a Hamilton cycle H through the remaining vertices R of $\Gamma_{\max}$. Indeed, we only mark the edges corresponding to vertices that are important for the spanning tree $\mathcal{T}$ of $\Gamma_{\max}$. Second, we only mark the singletons corresponding to important vertices that are not matched in the various (a, b) games. Then when we apply Lemma 5.7 to merge the blob-cycles and the marked singletons viz. the important vertices not covered by M_0.

$|R|$ is even, since the total number of points n is even. Thus the Hamilton cycle H yields a matching M_1 covering all of R. We take $M_0 \cup M_1$ as our perfect matching. ■

Proof of Theorem 1.3: This is again a straightforward adaptation of the proof of Theorem 1.1. It is clear that Breaker wins if there is a vertex of degree at most one or an edge of degree at most two. (Recall that Breaker starts the game.) Hence

$$\mathbb{P}\left[\rho_n(\text{Maker wins}) \ge \rho_n(\text{min.deg.} \ge 2 \text{ and min.edge-deg.} \ge 3)\right] = 1. \tag{41}$$

We now define:

$$r_L(n) := \left(\frac{\ln n + \ln\ln n - K}{\pi n}\right)^{\frac{1}{2}}, \quad r_U(n) := \left(\frac{\ln n + \ln\ln n + K}{\pi n}\right)^{\frac{1}{2}},$$

for K a (large) constant. By Lemma 6.1, we can choose $K = K(\varepsilon)$ such that

$$\mathbb{P}\Big[r_L(n) \le \rho_n(\text{min.deg.} \ge 2 \text{ and min.edge-deg.} \ge 3) \le r_U(n)\Big] \ge 1 - \varepsilon + o(1). \tag{42}$$

By Lemma 3.3 the properties **(str-1)**–**(str-5)** are satisfied with probability $1 - o(1)$ by $V = \mathcal{X}_n, m = m_n, T = O(1), r = r_L(n)$ with $\mathcal{X}_n$ as given by (4), m_n as given by (12) and r_L as above. By Lemma 3.4, with probability $1 - o(1)$, the remaining conditions of Lemma 6.7 are met for any $r \le \rho \le 2r$ with minimum vertex degree at least two and minimum edge degree at least three. Hence:

$$\mathbb{P}\Big[\rho_n(\text{Maker wins the perfect matching game}) = \rho_n(\text{min. deg.} \ge 2 \text{ and min. edge-deg.} \ge 3)\Big] \\ \ge \\ 1 - \varepsilon - o(1).$$

Sending $\varepsilon \downarrow 0$ gives the theorem. ■

7 The H-game

Proof of Theorem 1.7: The proof of Theorem 1.7 is a bit different from the previous proofs. Let H be a fixed connected graph and let k denote the smallest number for which the H-game is Maker's win on a k-clique; let $\mathcal{F}$ denote the family of all non-isomorphic graphs on k vertices for which the game is Maker-win. Let $H_1, \dots, H_m \in \mathcal{F}$ be those graphs in $\mathcal{F}$ that can actually be realized as a geometric graph. Observe that, since H is connected and Maker cannot win on any graph on $< k$ vertices, each H_i is connected.

Let $\varepsilon > 0$ be arbitrary and let $K > 0$ be a large constant, to be chosen later. Let us set

$$r_U := K \cdot n^{-k/2(k-1)}.$$

It follows from Theorem 2.9 that

$$\begin{array}{rcl} \mathbb{P}(G(n, r_U) \text{ contains a subgraph} \in \mathcal{F}) & = & 1 - \exp\left[-K^{2(k-1)} \cdot \sum_{i=1}^{m} \mu(H_i)\right] + o(1) \\ & \geq & 1 - \varepsilon + o(1), \end{array}$$

where $\mu(.)$ is as defined in (5) and the last inequality holds if we assume (without loss of generality) that K was chosen sufficiently large.

Observe that if $G(n, r_U)$ contains a component of order $> k$, then there exist $k+1$ points $X_{i_1}, \dots, X_{i_{k+1}}$ such that $\|X_{i_1} - X_{i_j}\| \leq (k+1)r_U$ for all $2 \leq j \leq k+1$. This gives

$$\begin{array}{rcl} \mathbb{P}(G(n, r_U) \text{ has a component of order} > k) & \leq & n^{k+1} \cdot \pi^k \big((k+1)r_U\big)^{2k} \\ & = & O\left(n^{k+1-\frac{k^2}{k-1}}\right) \\ & = & O\left(n^{-k/(k-1)}\right) \\ & = & o(1). \end{array}$$

Let E denote the event that $G(n, r_U)$ contains a subgraph $\in \mathcal{F}$ but no component on $> k$ vertices. Then it is clear that

$$\mathbb{P}\left[\rho_n(\text{Maker wins the } H\text{-game}) = \rho_n(\text{contains a subgraph} \in \mathcal{F})\right] \geq \mathbb{P}(E) \geq 1 - \varepsilon - o(1),$$

since, if all components have order $\leq k$ then Maker wins if and only if there is a component $\in \mathcal{F}$. Sending $\varepsilon \downarrow 0$ proves the theorem. ■

Proof of Corollary 1.8: This follows immediately from Theorem 2.3 and Theorem 2.9. ■

8 Conclusion and further work

In the present paper, we explicitly determined the hitting radius for the games of connectivity, perfect matching and Hamilton cycle, all played on the edges of the random geometric graph. As it turns out in all three cases, the hitting radius for $G(n, r)$ to be Maker-win coincides exactly with a simple, necessary minimum degree condition. For the connectivity game it is the minimum degree two, in the case of the perfect matching game it is again the minimum degree two accompanied by the minimum edge degree three, and for the Hamilton cycle game we have the minimum degree four. Each of these characterizations engenders an extremely precise description of the behavior at the threshold value for the radius. We also state a general result for the H-game, for a fixed graph H, where the hitting radius can be determined by finding the smallest k for which the H-game is Maker-win on the k-clique edge set and finding the list of all connected graphs on k vertices which are Maker-win.

These results are curiously similar to the hitting time results obtained for the Maker-win in the same three games played on the Erdős-Rényi random graph process. In that setting, in the connectivity game the hitting time for Maker-win is the same as for the minimum degree two [22], the same holds for the perfect matching game [4], and the condition changes to minimum degree four for the Hamilton cycle game [4]. As we can see, in the case of the connectivity game and the Hamilton cycle game the conditions are exactly the same as for the random geometric graph. The difference in the condition for the perfect matching game is not a surprise, as the existence of an induced 3-path (which clearly prevents Maker from winning) at the point when the graph becomes minimum degree two is an unlikely event in the Erdős-Rényi random graph process, but it does happen with positive probability in the random geometric graph.

With respect to the H-game much less is known on the Erdős-Rényi random graph process. The only graph for which we have a description of the hitting time is the triangle—Maker wins exactly when the first K_5 with one edge missing appears [17]. On the random geometric graph we basically have the same witness of Maker's victory, as the smallest k for which Maker can win the triangle game on edges of K_k is $k = 5$. Interestingly, for most other graphs H it is known that a hitting time result for Maker-win in the Erdős-Rényi random graph process cannot involve the appearance of a finite graph on which Maker can win—Maker-winning strategy must be of "global nature" [17, 18]. This is in contrast to the results we obtained in Theorem 1.7, showing that on the random geometric graph Maker can typically win the H-game by simply spotting a copy of one of some finite list of graphs and restricting his attention to that subgraph.

Playing a game on a random graph instead of the complete graph can be seen as a help to Breaker, as the board of the game becomes sparser, there are fewer winning sets and consequently Maker finds it harder to win. A standard alternative approach is to play the biased $(1 : b)$ game on the complete graph, where Breaker again gains momentum when b is increased. Naturally, one can combine the two approaches, playing the biased game on a random graph. This has been done in [22, 8] for the Erdős-Rényi random graph, where the threshold probability for Maker-win in the biased $(1 : b)$ game, with $b = b(n)$ fixed, was sought for several standard positional games on graphs. The same question can be asked in our random geometric graph setting.

Question 8.1 *Given a bias b, what can be said for the smallest radius r at which Maker can win the $(1 : b)$ biased game, for the games of connectivity, perfect matching, Hamilton cycle, and the H-game?*

In this paper we have only considered the random geometric graph constructed on points taken uniformly at random on the unit square, using the euclidean norm to decide on the edges. So other obvious directions for further work are:

Question 8.2 *What happens for Maker-Breaker games on random geometric graphs in dimensions $d \geq 3$? What happens if we use other probability distributions or norms in two dimensions?*

References

[1] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Probab.*, 17(1):9–25, 1989.

[2] J. Beck. Positional games and the second moment method. *Combinatorica*, 22(2):169–216, 2002.

[3] J. Beck. *Combinatorial games : Tick-Tack-Toe Theory*, volume 114 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2008.

[4] S. Ben-Shimon, A. Ferber, D. Hefetz, and M. Krivelevich. Hitting time results for Maker-Breaker games. *Random Structures Algorithms*, 41(1):23–46, 2012.

[5] S. Ben-Shimon, M. Krivelevich, and B. Sudakov. Local resilience and Hamiltonicity Maker-Breaker games in random regular graphs. *Combin. Probab. Comput.*, 20(2):173–211, 2011.

[6] V. Chvátal and P. Erdős. Biased positional games. *Ann. Discrete Math.*, 2:221–229, 1978. Algorithmic aspects of combinatorics (Conf., Vancouver Island, B.C., 1976).

[7] C. Cooper and A. Frieze. The cover time of random geometric graphs. *Random Structures Algorithms*, 38(3):324–349, 2011.

[8] A. Ferber, R. Glebov, M. Krivelevich, and A. Naor. Biased games on random graphs. manuscript.

[9] E. N. Gilbert. Random plane networks. *J. Soc. Indust. Appl. Math.*, 9:533–543, 1961.

[10] D. Hefetz, M. Krivelevich, M. Stojaković, and T. Szabó. A sharp threshold for the Hamilton cycle Maker-Breaker game. *Random Structures Algorithms*, 34(1):112–122, 2009.

[11] J. Kingman. *Poisson Processes.* Oxford University Press, Oxford, 1993.

[12] M. Krivelevich. The critical bias for the Hamiltonicity game is $(1 + o(1))n/\ln n$. *J. Amer. Math. Soc.*, 24(1):125–131, 2011.

[13] A. Lehman. A solution of the Shannon switching game. *J. Soc. Indust. Appl. Math.*, 12:687–725, 1964.

[14] C. J. H. McDiarmid. Random channel assignment in the plane. *Random Structures Algorithms*, 22(2):187–212, 2003.

[15] C. J. H. McDiarmid and T. Müller. On the chromatic number of random geometric graphs. *Combinatorica*, 31(4):423–488, 2011.

[16] T. Müller. Two-point concentration in random geometric graphs. *Combinatorica*, 28(5):529–545, 2008.

[17] T. Müller and M. Stojaković. A threshold for the Maker-Breaker clique game. *Random Structures Algorithms.* to appear.

[18] R. Nenadov, A. Steger, and M. Stojaković. On the threshold for the Maker-Breaker H-game. manuscript.

[19] M. D. Penrose. The longest edge of the random minimal spanning tree. *Ann. Appl. Probab.*, 7(2):340–361, 1997.

[20] M. D. Penrose. On k-connectivity for a geometric random graph. *Random Structures Algorithms*, 15(2):145–164, 1999.

[21] M. D. Penrose. *Random Geometric Graphs.* Oxford University Press, Oxford, 2003.

[22] M. Stojaković and T. Szabó. Positional games on random graphs. *Random Structures Algorithms*, 26(1-2):204–223, 2005.

A The proof of Theorem 2.7

Proof of Theorem 2.7: We condition on $N = m$. For convenience, let us write $\mathcal{X}_m := \{X_i : 1 \leq i \leq m\}$. We have

$$\begin{aligned}
\mathbb{E}Z &= \sum_{m=k}^{\infty} \mathbb{E}\left[Z|N=m\right] \cdot \mathbb{P}(N=m) \\
&= \sum_{m=k}^{\infty} (m)_k \cdot \mathbb{E}\left[h(X_1,\ldots,X_k;\mathcal{X}_m)\right] \cdot \mathbb{P}(N=m) \\
&= \sum_{m=k}^{\infty} (m)_k \cdot \mathbb{E}\left[h(Y_1,\ldots,Y_k;\{Y_1,\ldots,Y_k\}\cup\mathcal{X}_{m-k})\right] \cdot \frac{n^m e^{-n}}{m!} \\
&= n^k \cdot \sum_{j=0}^{\infty} \mathbb{E}\left[h(Y_1,\ldots,Y_k;\{Y_1,\ldots,Y_k\}\cup\mathcal{X}_j)\right] \cdot \frac{n^j e^{-n}}{j!} \\
&= n^k \cdot \mathbb{E}\left[h(Y_1,\ldots,Y_k;\{Y_1,\ldots,Y_k\}\cup\mathcal{P})\right],
\end{aligned}$$

as required. ■

B The proof of Lemma 2.8

Proof of Lemma 2.8: Let $\varepsilon > 0$ be arbitrary. By Markov's inequality and the fact that $\mathbb{E}Z_n = O(1)$, there exists a constant $K > 0$ such that

$$\mathbb{P}(Z_n > K) \leq \mathbb{E}Z_n/K \leq \varepsilon.$$

Let N be the Po(n)-distributed random variable used in the definition of $\mathcal{P}_n$. By Chebyschev's inequality we have

$$\mathbb{P}\left[|N-n| > K\sqrt{n}\right] \leq \mathrm{Var}(N)/(K\sqrt{n})^2 = 1/K^2 < \varepsilon.$$

(We can assume without loss of generality that $1/K^2 < \varepsilon$.)

Also observe that, since $\pi n r^2 = o(\sqrt{n})$, there exists a sequence $f(n) = o(\sqrt{n})$ such that

$$\mathbb{P}\big[\Delta(G(n+K\sqrt{n}, r_n)) > f(n)\big] = o(1),$$

where $\Delta(.)$ denotes the maximum degree. (This can for instance be seen from known results such as Theorem 2.3 in [14], or by a first moment argument using the Chernoff bound.)

For $m \in \mathbb{N}$ let us call a tuple $(X_{i_1},\ldots,X_{i_k}) \in \mathcal{X}_m^k$ a *configuration* if $h_n(X_{i_1},\ldots,X_{i_k};\mathcal{X}_m) = 1$ and let Y_m denote the number of configurations in $\mathcal{X}_m$. Let pick an arbitrary $t \leq K$ and an arbitrary pair $n - K\sqrt{n} \leq m < m' \leq n + K\sqrt{n}$.

We have

$$\begin{aligned}
\mathbb{P}(\tilde{Z}_n \neq Z_n) &= \textstyle\sum_{m=0}^{\infty}\sum_{t=0}^{\infty} \mathbb{P}(\tilde{Z} \neq t, Z = t|N=m)\mathbb{P}(N=m) \\
&\leq \textstyle\sum_{m=n-K\sqrt{n}}^{n+K\sqrt{n}}\sum_{t=0}^{K} \mathbb{P}(\tilde{Z} \neq t, Z = t|N=m)\mathbb{P}(N=m) \\
&\quad +\mathbb{P}(|N-n| > K\sqrt{n}) + \mathbb{P}(Z > K) \\
&\leq \textstyle\sum_{m=n-K\sqrt{n}}^{n+K\sqrt{n}}\sum_{t=0}^{K} \mathbb{P}(\tilde{Z} \neq t, Z = t|N=m)\mathbb{P}(N=m) + 2\varepsilon \\
&= \textstyle\sum_{m=n-K\sqrt{n}}^{n+K\sqrt{n}}\sum_{t=0}^{K} \mathbb{P}(Y_n \neq t, Y_m = t)\mathbb{P}(N=m) + 2\varepsilon
\end{aligned} \tag{43}$$

Let us now fix some $n - K\sqrt{n} \leq m \leq n$ and $t \leq K$. Note that if $Y_m = t$ and $Y_n < t$ then we can fix t configurations in $\mathcal{X}_m$, and at least one of the points $X_{m+1},\ldots,X_n$ must fall within r of one of the $k \cdot t$ points of these configurations. This gives:

$$\mathbb{P}(Y_m = t, Y_n < t) \leq \mathbb{P}(Y_n < t|Y_m = t) \leq k \cdot t \cdot r_n^2 \cdot (n-m) = o(1), \tag{44}$$

since $r_n = o(n^{-1/2})$ and $n - m = O(\sqrt{n})$.

Similarly, if $Y_m = t$ and $Y_n > t$ then we can fix $t+1$ configurations in $\mathcal{X}_n$, and at least one of the points that is either part of one of these configurations, or within distance r of a point of one of these configurations must be among $X_{m+1}, \dots, X_n$. This gives

$$\begin{aligned} \mathbb{P}(Y_m = t, Y_n > t) &\leq \mathbb{P}(\Delta(G(n + K\sqrt{n}, r_n)) > f(n)) \\ &\quad + \mathbb{P}(Y_m = t, Y_n > t, \Delta(G(n + K\sqrt{n}, r_n)) \leq f(n)) \\ &\leq o(1) + \mathbb{P}(Y_m = t | Y_n > t, \Delta(G(n + K\sqrt{n}, r_n)) \leq f(n)) \\ &\leq o(1) + k \cdot (t+1) \cdot (f(n)+1) \cdot \tfrac{n-m}{n} \\ &= o(1), \end{aligned} \tag{45}$$

using $n - m = O(\sqrt{n})$ and $f(n) = \sqrt{n}$ for the last line. Combining (44) and (45) shows that $\mathbb{P}(Y_m = t, Y_n \neq t) = o(1)$ for all $t \leq K$ and $n - K\sqrt{n} \leq m \leq n$.

Similarly, we find that $\mathbb{P}(Y_m = t, Y_n \neq t) = o(1)$ for all $t \leq K$ and $n \leq m \leq n + K\sqrt{n}$. Using (43) we now find that

$$\mathbb{P}(\tilde{Z}_n \neq Z_n) \leq \sum_{m=n-K\sqrt{n}}^{n+K\sqrt{n}} \sum_{t=0}^{K} \mathbb{P}(Y_n \neq t, Y_m = t)\mathbb{P}(N = m) + 2\varepsilon = o(1) + 2\varepsilon.$$

Sending $\varepsilon \downarrow 0$ completes the proof. ■

C The proof of Lemma 2.12

Proof of Lemma 2.12: Consider a spanning tree T of G that minimizes the sum of the edge lengths. Then T does not have any vertex of degree ≥ 7. This is because, if v were to have degree ≥ 7, then there are two neighbours u, w of v such that the angle between the segments $[v, u]$ and $[v, w]$ is strictly less than 60 degrees. We can assume without loss of generality that $[v, u]$ is shorter than $[v, w]$. Note that if we remove the edge vw and add the edge uw then we obtain another spanning tree but with strictly smaller total edge-length, a contradiction. Hence T has maximum degree at most 6.

Similarly, we see that if v has degree 6 in T then the neighbours of v have exactly the same distance to v and the angle between a pair of consecutive neighbours is exactly 60 degrees.

Let us now pick a spanning tree T' which minimizes the total edge-length (and hence has no degree ≥ 7 vertices) and, subject to this, has as few as possible degree 6 vertices and, subject to these two demands, the maximum over all degree 6 vertices of their x-coordinate is as large as possible. Let v be a degree 6 vertex with largest x-coordinate. It has two neighbours u, w with strictly larger x-coordinates. Observe that, as seen above the segments $[v, u], [v, w], [u, w]$ all have the same length. Let us remove the edge uv and add uw. This results in another spanning tree, with the same total edge length. The degree of u has not changed, the degree of v has dropped, and the degree of w has increased by one. If the degree of w has become 6 then our spanning tree is not as we assumed (there is a degree 6 vertex with x-coordinate strictly larger than that of v). Thus the degree of w after this operation is < 6. But then the number of degree 6 vertices has decreased, contradiction. ■

D The proof of Lemma 6.2

If $G = (V, E)$ is a graph and $\overline{Z} = (Z_v : v \in V)$ are random variables, then G is a *dependency graph* for $\overline{Z}$ if, whenever there is no edge between $A, B \subseteq V$, the random vectors $\overline{Z_A} := (Z_v : v \in A)$ and $\overline{Z_B} := (Z_v : v \in B)$ are independent. The proof of Lemma 6.2 relies on the following result of Arratia et al. [1]:

Theorem D.1 (Arratia et al. [1]) *Let $(Z_v : v \in V)$ be a collection of Bernouilli random variables with dependency graph $G = (V, E)$. Set $p_v := \mathbb{E}Z_v = \mathbb{P}(Z_v = 1)$ and $p_{uv} := \mathbb{E}Z_u Z_v = \mathbb{P}(Z_v = Z_u = 1)$. Set $W := \sum Z_v$, and $\lambda := \mathbb{E}W$. Then*

$$d_{TV}(W, \mathrm{Po}(\lambda)) \leq \min(3, 1/\lambda) \cdot \left(\sum_{v \in V(G)} \sum_{u \in N(v)} p_{uv} + \sum_{v \in V(G)} \sum_{u \in N(v) \cup \{v\}} p_u p_v \right).$$

Proof of Lemma 6.2: We will use Theorem D.1. We consider the dissection $\mathcal{D}(m)$ for some $m \in \mathbb{N}$. For convenience, let us order the cells of $\mathcal{D}(m)$ in some (arbitrary) way as $c_1, c_2, \dots, c_{m^2}$. For $1 \leq i \leq m^2$, let us denote by x_i the lower left-hand corner of c_i, and let ξ_i denote the indicator variable defined by;

$$\xi_i := 1_{\{\mathcal{P}(c_i)=1, \mathcal{P}(B(x_i;r)\setminus c_i)=1\}}.$$

For $1 \leq i < j \leq m^2$ with $\|x_i - x_j\| \leq r$, let us write

$$\xi_{(i,j)} := 1_{\{\mathcal{P}(c_i)=\mathcal{P}(c_j)=1, \mathcal{P}((B(x_i;r)\cup B(x_j;r))\setminus(c_i \cup c_j))=2\}}.$$

Let us set

$$Y^m := \sum_{i=1}^{m^2} \xi_i, \quad W^m := \sum_{\substack{1 \leq i < j \leq m^2, \\ \|x_i - x_j\| \leq r}} \xi_{i,j},$$

and $Z^m := Y^m + W^m$. By construction, we have

$$\lim_{m\to\infty} Y^m = Y, \quad , \lim_{m\to\infty} W^m = W, \quad \lim_{m\to\infty} Z^m = Z. \tag{46}$$

For convenience let us write $\mathcal{I}^m := [m^2], \mathcal{J}^m := \binom{[m^2]}{2}$ and $\mathcal{V}^m := \mathcal{I}^m \cup \mathcal{J}^m$. We now define a graph $\mathcal{G}^m$ with vertex set $\mathcal{V}_m$ and edges:

- $ij \in E(\mathcal{G}^m)$ if $i, j \in \mathcal{I}^m, \|x_i - x_j\| \leq 100r$;
- $uv \in E(\mathcal{G}^m)$ if $u = \{i_1, i_2\}, v = \{i_3, i_4\} \in \mathcal{J}^m$ and $\|x_{i_a} - x_{i_b}\| \leq 100r$ for all $1 \leq a, b \leq 4$;
- $uv \in E(\mathcal{G}^m)$ if $u = i_1 \in \mathcal{I}^m, v = \{i_2, i_3\} \in \mathcal{J}^m$ and $\|x_{i_a} - x_{i_b}\| \leq 100r$ for all $1 \leq a, b \leq 3$;

By the spatial independence properties of the Poisson process, this defined a dependency graph on the random indicator variables $(\xi_v : v \in \mathcal{V}^m)$. By Theorem D.1, we therefore have that

$$d_{\mathrm{TV}}(Z^m, \mathbb{E}(Z^m)) \leq 3 \cdot \left(\sum_{v \in \mathcal{V}^m} \sum_{u \in N(v)} \mathbb{E}\xi_v \xi_u + \sum_{v \in \mathcal{V}^m} \sum_{u \in N(v)} \mathbb{E}\xi_v \mathbb{E}\xi_u \right). \tag{47}$$

Let us write

$$\begin{aligned}
S_1^m &:= \sum_{v \in \mathcal{I}^m} \sum_{u \in N(v)\cap\mathcal{I}^m} \mathbb{E}\xi_v \mathbb{E}\xi_u, \quad S_2^m := \sum_{v \in \mathcal{I}^m} \sum_{u \in N(v)\cap\mathcal{I}^m} \mathbb{E}\xi_v \xi_u, \\
S_3^m &:= \sum_{v \in \mathcal{J}^m} \sum_{u \in N(v)\cap\mathcal{J}^m} \mathbb{E}\xi_v \mathbb{E}\xi_u, \quad S_4^m := \sum_{v \in \mathcal{J}^m} \sum_{u \in N(v)\cap\mathcal{J}^m} \mathbb{E}\xi_v \xi_u, \\
S_5^m &:= \sum_{v \in \mathcal{J}^m} \sum_{u \in N(v)\cap\mathcal{I}^m} \mathbb{E}\xi_v \mathbb{E}\xi_u, \quad S_6^m := \sum_{v \in \mathcal{J}^m} \sum_{u \in N(v)\cap\mathcal{I}^m} \mathbb{E}\xi_v \xi_u,
\end{aligned}$$

Then

$$\begin{aligned}
\textstyle\sum_{v \in \mathcal{V}^m} \sum_{u \in N(v)} \mathbb{E}\xi_v \xi_u = S_2^m + S_4^m + 2S_6^m, \\
\textstyle\sum_{v \in \mathcal{V}^m} \sum_{u \in N(v)} \mathbb{E}\xi_v \mathbb{E}\xi_u = S_1^m + S_3^m + 2S_5^m.
\end{aligned} \tag{48}$$

For $x \in [0,1]^2$, let us define $\varphi^m(x) := m^{-2}\mathbb{E}\xi_i$ where $i \in \mathcal{I}$ is such that $x \in c_i$. Then

$$S_1^m = \int_{[0,1]^2} \int_{B'(x;100r)} \varphi^m(x)\varphi^m(y)\mathrm{d}y\mathrm{d}x,$$

where $B'(x;100r)$ is the union of all cells c_i with $\|x - c_i\| \le 100r$. Next we claim that $\lim_{m\to\infty} S_j^m \to I_j$ for all $1 \le j \le 6$. Now notice that

$$0 \le \varphi^m(x) \le m^{-2}(nm^2)e^{-nm^2} \le n,$$

and, for every $x \in [0,1]^2$:

$$\lim_{m\to\infty} \varphi_m(x) = n \cdot \mathbb{P}(E_x).$$

We can thus apply the dominated convergence theorem to show that $\lim_{m\to\infty} S_1^m \to I_1$. Similarly, we can show that $S_j^m \to I_j$ for all $2 \le j \le 6$. Together with (46), (47) and (48) this proves the Lemma. ■