The intersection of a random geometric graph with an Erdős-Rényi graph

Patrick Bennett[*] Alan Frieze[†] Wesley Pegden[‡]

Abstract

We study the intersection of a random geometric graph with an Erdős-Rényi graph. Specifically, we generate the random geometric graph $G(n,r)$ by choosing n points uniformly at random from $D=[0,1]^2$ and joining any two points whose Euclidean distance is at most r. We let $G(n,p)$ be the classical Erdős-Rényi graph, i.e. it has n vertices and every pair of vertices is adjacent with probability p independently. In this note we study $G(n,r,p) := G(n,r) \cap G(n,p)$. One way to think of this graph is that we take $G(n,r)$ and then randomly delete edges with probability $1-p$ independently. We consider the clique number, independence number, connectivity, Hamiltonicity, chromatic number, and diameter of this graph where both $p(n) \to 0$ and $r(n) \to 0$; the same model was studied by Kahle, Tian and Wang (2023) for $r(n) \to 0$ but p fixed.

1 Introduction

In this note we study the intersection of a random geometric graph with an Erdős-Rényi graph. Specifically, we generate the random geometric graph $G(n,r)$ by choosing n points uniformly at random from $D=[0,1]^2$ and joining any two points whose Euclidean distance is at most r. See Penrose [3] for research monograph on this model. For convenience we work on the torus, so for example we consider $(0,0),(1,0),(0,1)$ and $(1,1)$ to the the same point. We let $G(n,p)$ be the classical Erdős-Rényi graph, i.e. it has n vertices and every pair of vertices is adjacent with probability p independently. In this note we study the graph we will call $G(n,r,p) := G(n,r) \cap G(n,p)$. One way to think of this graph is that we take $G(n,r)$ and then randomly delete edges with probability $1-p$ independently.

Kahle, Tian and Wang [2] studied the clique number of a random graph which is more general than our $G(n,r,p)$ (but for a more limited range of the parameter p). In [2] they study the *noisy random geometric graph*, which is the result of taking a random geometric graph, deleting each edge independently with some probability, and also adding random edges that were not present. Thus, our model is the *deletion-only case* of the model in [2]. However, in [2] they consider the case where only $r=r(n)$ can be a function of n, but p is a fixed constant. We allow $p=p(n) \to 0$ as well. For convenience we define the parameter

$$q := \pi r^2 p,$$

[*]Department of Mathematics, Western Michigan University, Kalamazoo MI 49008, Research supported in part by Simons Foundation Grant #426894.

[†]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213, Research supported in part by NSF grant DMS1952285.

[‡]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213, Research supported in part by NSF grant DMS1700365.

which is the probability that that two given vertices are adjacent.

We begin with a discussion of the clique number of $G = G(n, r, p)$.

Theorem 1. *For each fixed $\varepsilon > 0$ there exists $C = C_\varepsilon$ such that we have the following. If $nr^2 \geq \log n$ and $C/(nr^2) \leq 1 - p \leq \log^{-3} n$, then*

$$\frac{(2-\varepsilon)\log\left[(1-p)nr^2\right]}{1-p} \leq \omega(G) \leq \frac{2\log\left[(1-p)nr^2\right]}{1-p} \tag{1}$$

We next consider connectivity.

Theorem 2. *Let $\varepsilon > 0$. There exists some $r_0 > 0$ such that we have the following for all $r \leq r_0$. If $q \leq (1-\varepsilon)\log n/n$, then w.h.p. $G(n, r, p)$ is disconnected and in particular it has isolated vertices. On the other hand if we have $q \geq (1+\varepsilon)\log n/n$ then w.h.p. $G(n, r, p)$ is connected.*

We next consider Hamiltonicity. Our result here is not as precise as that for connectivity.

Theorem 3. *There exists some $r_0 > 0$ and $K > 0$ such that we have the following for all $r \leq r_0$, if $q \geq K \log n/n$ then w.h.p. $G(n, r, p)$ is Hamiltonian.*

After this we consider the independence number.

Theorem 4. *For each fixed $\varepsilon > 0$ and fixed p with $0 < p < 1$ we have the following. If $nr^2 \geq n^\varepsilon$, then*

$$\alpha(G) = \Theta\left(r^{-2}\log_{1/(1-p)}(nr^2)\right) \tag{2}$$

We then consider the chromatic number.

Theorem 5. *For each fixed $\varepsilon > 0$ and fixed p with $0 < p < 1$ we have the following. If $nr^2 \geq n^\varepsilon$, then*

$$\chi(G) = \Theta\left(\frac{nr^2}{\log_{1/(1-p)}(nr^2)}\right) \tag{3}$$

We finally consider the diameter.

Theorem 6. *If $r < 1/2$ and $nr^2 p \geq \log^2 n$ then w.h.p.*

$$diam(G) = \Theta\left(r^{-1} + \frac{\log n}{\log(nr^2 p)}\right).$$

2 Degree sequence

The degree of a vertex is distributed as $\mathrm{Bin}(n-1, q)$. We will use the following standard tool to show concentration.

Theorem 7 (Chernoff–Hoeffding bound)**.** *Let X be distributed as $Bin(m, s)$ and $0 < \delta < 1$. Then*

$$\mathbb{P}(|X - ms| > \delta ms) \leq 2\exp(-\delta^2 ms/3)$$

Suppose $q = \omega(\log n/n)$. Letting $\delta = [4\log n/nq]^{1/2} = o(1)$, we see that

$$\mathbb{P}\left(\left|deg(v) - (n-1)q\right| > \delta(n-1)q\right) \leq 2\exp(-\delta^2(n-1)q/3) = o(1/n).$$

Thus by the union bound, w.h.p. every vertex in $G(n, r, p)$ has degree $(1 + o(1))(n-1)q = (1 + o(1))nq$.

3 Cliques: proof of Theorem 1

Proof. We first prove the upper bound. We partition D into $1/r^2$ cells with side length r, and we say that a *block* consists of a 10 by 10 set of cells. Each block is anchored by a single cell in the top left hand corner. Thus there are $1/r^2$ blocks. We first reveal the location of the points. The probability that the ith point belongs to a given block is $100r^2$, and thus the expected number of points in each block is $100nr^2$, and the Chernoff bounds tell us that the probability that a given block has more than $200nr^2$ points is at most $e^{-100nr^2/3}$, which is $o(r^2)$ for $r \geq \left(\frac{\log n}{n}\right)^{1/2}$, and in particular this bound on r implies that, by a union bound over the $1/(100r^2)$ blocks, no block has more than $200nr^2$ points.

Any clique must have all of its vertices lying in one of our blocks. The graph induced on the vertices in a block is distributed as $G(n', p)$ where $n' \leq 200nr^2$. For our upper bound we can take $n' = 200nr^2$ since adding more points to a block cannot decrease the clique size. The expected number of cliques of size $k := \frac{2\log\left[(1-p)nr^2\right]}{1-p}$ in $G(n', p)$ is then at most

$$\begin{aligned}
\binom{n'}{k} p^{\binom{k}{2}} \leq \left(\frac{en'p^{\frac{k-1}{2}}}{k}\right)^k &\leq \left(\frac{O(1)nr^2 p^{\frac{2\log\left[(1-p)nr^2\right]}{2(1-p)}}}{k}\right)^k \\
&= \left(\frac{O(1)nr^2\left[(1-p)nr^2\right]^{\frac{2\log p}{2(1-p)}}}{k}\right)^k \\
&= \left(\frac{O(1)nr^2\left[(1-p)nr^2\right]^{-1}}{\frac{\log[(1-p)nr^2]}{1-p}}\right)^k \quad \text{since } \log p/(1-p) = -(1+o(1)) \\
&= \left(\frac{O(1)}{\log\left[(1-p)nr^2\right]}\right)^k \\
&= \left(\frac{O(1)}{\log C}\right)^{\Omega(\log^2 n)}
\end{aligned}$$

which (for large C) is small enough to beat a union bound over the blocks. Thus w.h.p. G has no clique of size k. This completes the proof of the upper bound.

Now we prove the lower bound. Here we use cells with side length $r/\sqrt{2}$. By the pigeonhole principle one of the cells contains at least $n'' := nr^2/2$ points, which are all within distance r of each other. W.h.p. (see, for example, [1]) since $C/(nr^2) \leq 1-p \leq \log^{-3} n$, $G(n'', p)$ has a clique of size at least

$$\frac{(2-\varepsilon/2)\log[(1-p)n'']}{1-p} \geq \frac{(2-\varepsilon)\log\left[(1-p)nr^2\right]}{1-p}.$$

Explanation: to justify the first expression for the clique number, observe that for small ρ we have $\alpha(G(n,\rho)) = (2\pm\varepsilon)\log(n\rho)/\rho$ and so for $p = 1-\rho$ close to 1 we have $\omega(G(n,p) = (2\pm\varepsilon)\log(n(1-p))/(1-p)$. This completes the proof of the upper bound. □

4 Connectivity: proof of Theorem 2

We will use $r \leq r_0$, for example, to conclude that a ball of radius r has volume πr^2 (since we have wrap around, this ball would intersect itself if r were too big).

Proof. Suppose first that $q \le (1-\varepsilon)\log n/n$. Let X be the number of isolated vertices, so

$$\mathbb{E}[X] = n(1-q)^{n-1} = n\exp\Big(-(n-1)q + O(nq^2)\Big) \ge (1+o(1))n^{\varepsilon}.$$

We will now bound the expected number of pairs of isolated vertices. Fix two points $u, v \in [0,1]^2$. Say the volume of $B_r(u) \cap B_r(v)$ is x. The probability that a third vertex w is nonadjacent to both u and v is

$$1 - \Big(2\pi r^2 - x\Big)p - x\Big(1-(1-p)^2\Big) = 1 - 2\pi r^2 p - xp(1-p) \le 1 - 2q.$$

Thus the expected number of pairs u, v that are both isolated is at most

$$n^2(1-2q)^{n-2} \le n^2\exp\Big(-2(n-2)q\Big) = (1+o(1))(\mathbb{E}[X])^2.$$

Thus, by the second moment method, w.h.p. $X > 0$ and we have isolated vertices.

Now suppose that $q = (1+\varepsilon)\log n/n$. Note that this case suffices for all $q \ge (1+\varepsilon)\log n/n$ since the property of being connected is monotone in p (and also in r). We use cells of side length $\lambda := \eta r$ for some $\eta \ll \varepsilon$. We say a cell is *good* if it has at least $g := \eta^2 n\lambda^2$ points in it, and *bad* otherwise. We say that a cell is *great* if it is has say $x \ge g$ points (i.e. it is good) and it has a connected component on at least $x - g/2$ vertices. We say that such a component is a *great component*

Claim 1. *W.h.p. every vertex has a neighbor in a great component.*

Proof. Consider a fixed vertex v and the ball $B_r(v)$. Now since the diameter of a cell is $\sqrt{2}\eta r < 2\eta r$, we have that for any point v' with $||v - v'|| < (1-2\eta)r$, the cell containing v' lies entirely in $B_r(v)$. Thus, the number of cells contained entirely in $B_r(v)$ is at least

$$k := \left\lceil \frac{Vol(B_{(1-2\eta)r}(v))}{\lambda^2} \right\rceil = \left\lceil \pi(1-2\eta)^2\eta^{-2} \right\rceil.$$

Note that

$$\pi\eta^{-2} - 4\eta^{-1} \;\le k \le \pi\eta^{-2} + 1. \tag{4}$$

So we have some cells $C_1, \dots, C_k$ each contained in $B_r(v)$. Suppose that each cell C_i has j_i points in it (not counting v itself which is of course in one of them). Let $\rho(j_i)$ be the probability that the graph induced on C_i either fails to have a component on at least $j_i - g/2$ vertices, or else there is such a component but v is not adjacent to any of vertices in it. Then the probability that our claim fails for the vertex v is at most

$$\sum_{\substack{0 \le j \le n-1 \\ j_1, \dots, j_k \ge 0 \\ j_1 + \dots + j_k = j}} \binom{n-1}{j_1, \dots, j_k, n-1-j} \lambda^{2j}\left(1 - k\lambda^2\right)^{n-1-j} \prod_{i: j_i \ge g} \rho(j_i). \tag{5}$$

Explanation: we set $j = j_1 + \dots j_k$ and sum over all possibilities for these values. We must choose which points go where which explains the multinomial coefficient. Then for each of the j points going into one of the cells $C_1, \dots, C_k$, it has a probability of λ^2 of being in the correct cell. The rest of our $n - j$ points each has a probability of $1 - k\lambda^2$ of not being in any cell $C_1, \dots, C_k$. Finally, in order for our claim to fail at v, within each cell C_i having $j_i \ge g$ points, the event whose probability is defined as $\rho(j_i)$ must happen (and these events are independent since the C_i induce disjoint Erdős-Rényi graphs).

We first handle the terms of (5) with small j. Using the fact that $m! \ge (m/e)^m$ for all $m \ge 0$ (if we agree that $0^0 = 1$ and $0 \log 0 = 0$), we have

$$(j_1)! \dots (j_k)! \ge \left(\frac{j_1}{e}\right)^{j_1} \dots \left(\frac{j_k}{e}\right)^{j_k} = \exp\left\{ j_1 \log\left(\frac{j_1}{e}\right) + \dots j_k \log\left(\frac{j_k}{e}\right)\right\} \ge \exp\left\{ j \log\left(\frac{j}{ke}\right)\right\} = \left(\frac{j}{ke}\right)^j,$$

and so

$$\binom{n}{j_1,\ldots,j_k,n-j} \le \frac{n^j}{(j_1)!\ldots(j_k)!} \le \left(\frac{ekn}{j}\right)^j.$$

For each j there are at most k^j choices for the $j_1,\ldots j_k$ summing to j. Thus, the sum of all terms in (5) for $j \le \eta k\lambda^2 n$ is at most

$$\sum_{0\le j\le \eta k\lambda^2 n} k^j \left(\frac{ekn}{j}\right)^j \lambda^{2j}\left(1-k\lambda^2\right)^{n-j} \le \exp\{-k\lambda^2 n\} \sum_{0\le j\le \eta k\lambda^2 n} \left(\frac{ek^2\lambda^2 n}{j(1-k\lambda^2)}\right)^j. \tag{6}$$

Using (4) and $\lambda = \eta r$ gives

$$(\pi - 4\eta)r^2 \le k\lambda^2 \le (\pi+\eta^2)r^2, \tag{7}$$

so by choosing r_0 not too large (and η small) we can guarantee say $k\lambda^2 < 1/2$. Now the ratio of consecutive terms in (6) is

$$\left(\frac{ek^2\lambda^2 n}{(j+1)(1-k\lambda^2)}\right)^{j+1}\left(\frac{ek^2\lambda^2 n}{j(1-k\lambda^2)}\right)^{-j} = \frac{1}{j+1}\left(\frac{j}{j+1}\right)^j \frac{ek^2\lambda^2 n}{1-k\lambda^2} \ge 2,$$

since $nr^2 = \Omega(\log n)$. So the sum is on the order of its last term. Thus (6) is at most (explanation follows)

$$\begin{aligned}
\exp\{-k\lambda^2 n\}\cdot O\left(\left(\frac{ek}{\eta(1-k\lambda^2)}\right)^{\eta k\lambda^2 n}\right) &= O\left(\exp\left\{-\left(1-\eta\log\left(\frac{2e(\pi\eta^{-2}+1)}{\eta}\right)\right)k\lambda^2 n\right\}\right)\\
&= O\left(\exp\left\{-\left(1-\eta\log\left(\frac{2e(\pi\eta^{-2}+1)}{\eta}\right)\right)(\pi-4\eta)r^2 n\right\}\right)\\
&= O\left(\exp\left\{-\frac{(1+\varepsilon/2)\log n}{p}\right\}\right) = o(1/n)
\end{aligned}$$

On the first line above we used $k\lambda^2 < 1/2$ and (4), and on the second line we used (7). The last line follows since for $\eta \ll \varepsilon$ the coefficient $\left(1-\eta\log\left(\frac{2e(\pi\eta^{-2}+1)}{\eta}\right)\right)(\pi-4\eta)$ from the line above is at least say $(1-\varepsilon/10)\pi$, and $\pi r^2 n = nq/p = (1+\varepsilon)\log n/p$.

Now we take care of the terms of (5) with larger j. Suppose $j > \eta k\lambda^2 n$. For any $j_i \ge g$ we have

$$\begin{aligned}
\rho(j_i) &\le (1-p)^{j_i - g/2} + \sum_{g/2\le m\le j_i/2}\binom{j_i}{m}(1-p)^{m(j_i-m)}\\
&\le (1-p)^{j_i-g/2} + \sum_{g/2\le m\le j_i/2}\left(\frac{ej_i(1-p)^{j_i}}{m(1-p)^m}\right)^m\\
&\le (1-p)^{j_i-g/2} + \sum_{g/2\le m\le j_i/2}\left(\frac{2ej_i(1-p)^{j_i/2}}{g}\right)^m\\
&\le (1-p)^{j_i-g/2} + \sum_{g/2\le m\le j_i/2}\left(\frac{4}{gp}\right)^m = (1+o(1))(1-p)^{j_i-g/2},
\end{aligned} \tag{8}$$

where the last line follows from $xe^{-ax} \le 1/(ea)$ and $gp = \Omega(\log n)$.

The total sum of all values j_i with $j_i < g$ is at most kg, and so the sum of the rest of the j_i is at least $j - kg$. Thus the product in (5) is at most

$$\prod_{i:j_i\ge g}\rho(j_i) \le \prod_{i:j_i\ge g}(1+o(1))(1-p)^{j_i-g/2} \le (1+o(1))(1-p)^{j-3kg/2}$$

The sum of the terms of (5) with $j > \eta k\lambda^2 n$ is therefore bounded above by

$$
\begin{aligned}
&(1+o(1)) \sum_{\substack{\eta k\lambda^2 n < j \le n \\ j_1,\ldots,j_k \ge 0 \\ j_1+\ldots+j_k = j}} \binom{n}{j_1,\ldots,j_k,n-j} \lambda^{2j} \left(1-k\lambda^2\right)^{n-j} (1-p)^{j-3kg/2} \\
&\le (1+o(1)) \exp\{3kgp/2\} \sum_{\substack{0 \le j \le n \\ j_1,\ldots,j_k \ge 0 \\ j_1+\ldots+j_k = j}} \binom{n}{j_1,\ldots,j_k,n-j} \lambda^{2j} \left(1-k\lambda^2\right)^{n-j} (1-p)^{j} \\
&= (1+o(1)) \exp\{3kgp/2\} \Big(k\lambda^2(1-p) + (1-k\lambda^2)\Big)^n \qquad (9)
\end{aligned}
$$

where the last line follows from the multinomial theorem. Simplifying and using $1+x \le e^x$, the above is at most

$$
\begin{aligned}
(1+o(1)) \exp\{3kgp/2 - k\lambda^2 pn\} &\le (1+o(1)) \exp\left\{-\left(1-\frac{3}{2}\eta^2\right) k\lambda^2 pn\right\} \\
&\le (1+o(1)) \exp\left\{-\left(1-\frac{3}{2}\eta^2\right) (\pi - 4\eta) r^2 pn\right\} \\
&\le \exp\{-(1+\varepsilon/2)\log n\} = o(1/n)
\end{aligned}
$$

where on the first line we replaced $g = \eta^2 n\lambda^2$, on the second line we used (7), and on the last line we used $\eta \ll \varepsilon$ and the fact that $q = \pi r^2 p = (1+\varepsilon)\log n/n$. Thus (5), our bound on the probability that the claim fails for a single vertex, is $o(1/n)$. The claim now follows from the union bound over n vertices. $\square$

We now define an auxiliary graph Γ whose vertex set is our set of great cells in $G(n,r,p)$, and where two cells are adjacent in Γ if the diameter of their union is at most r (i.e. every point in one cell is distance at most r from every point in the other cell). Note that if we have two cells adjacent in Γ the probability there is no edge between their great components is at most $(1-p)^{g^2/4}$, small enough to union bound over pairs of cells and conclude that w.h.p. every such pair of cells has an edge between the great components.

Recalling that we have Claim 1, to finish the proof of Theorem 2 it suffices to prove the following:

Claim 2. *W.h.p.* Γ *is connected.*

Proof. The probability that any given cell is bad is at most

$$
\begin{aligned}
\sum_{j=0}^{g} \binom{n}{j} \lambda^{2j} \left(1-\lambda^2\right)^{n-j} &\le 2g \binom{n}{g} \lambda^{2g} \left(1-\lambda^2\right)^{n-g} \\
&\le 2g \exp\left\{g \log\left(\frac{ne}{g}\right) + g\log\left(\lambda^2\right) - (n-g)\lambda^2\right\} \\
&\le 2g \exp\left\{\eta^2 n\lambda^2 \left(\log\left(\frac{e}{\eta^2}\right) + \lambda^2\right) - n\lambda^2\right\} \\
&\le \exp\left\{-(1-\varepsilon/10)n\lambda^2\right\}. \qquad (10)
\end{aligned}
$$

Note that if we condition on one cell being bad, the event that another cell is bad only becomes less likely.

We consider blocks composed of $100\eta^{-4}$ cells as defined in Section 3. The probability that a block contains $k-10\eta^{-1}$ bad cells is at most

$$\binom{100\eta^{-4}}{k-10\eta^{-1}}\exp\left\{-(k-10\eta^{-1})(1-\varepsilon/10)n\lambda^2\right\} = O\left(\exp\left\{-(\pi\eta^{-2}-14\eta^{-1})(1-\varepsilon/10)n(\eta r)^2\right\}\right)$$
$$= O\left(\exp\left\{-(1-\varepsilon/10)(\pi-14\eta)nr^2\right\}\right)$$
$$= O\left(\exp\left\{-(1+\varepsilon/2)\log n/p\right\}\right) = o(1/n)$$

where we have used (4) and $\eta \ll \varepsilon$. Since the number of blocks is $O((\eta^2 r)^{-2}) = o(n)$, by the union bound we have that no block contains $k-10\eta^{-1}$ bad cells.

Following part of the same calculation in (8), the probability that a good cell having $j \geq g$ points fails to be great is at most

$$\sum_{g/2\leq m\leq j/2}\binom{j}{m}(1-p)^{m(j-m)} \leq \sum_{g/2\leq m\leq j/2}\left(\frac{2ej(1-p)^{j/2}}{g}\right)^m$$
$$\leq \frac{j}{2}\left(\frac{4}{gp}\right)^m \leq e^{-\Omega(\log n\log\log n)},$$

and so by the union bound w.h.p. every good cell is great.

Now within a block there are $10\eta^{-2} > 2k$ rows and columns (i.e. vertical or horizontal strips of cells). Thus, within a block we have that more than half of the rows (and more than half of the columns) contain only great cells. Call such a row or column great. Thus, for any two blocks that are disjoint except for sharing a side, there is some great row or column in one of the blocks touching a great row or column in the other block. Thus, the cells that are in great rows or columns are all in the same component of Γ. Any other component of Γ must be contained entirely in one block and consist of cells bounded by great rows and columns. Suppose there is such a component Γ_0 of Γ. Consider a "highest" cell C in Γ_0, i.e. a cell whose y-coordinate of its center (x_C, y_C) is largest. There are at least k cells which contain only points of distance at most r from (x_C, y_C). At most $2\eta^{-1}$ of these cells are at the same height as C, and so at least $(k-2\eta^{-1})/2$ of them are above C. These $(k-2\eta^{-1})/2$ cells must not be great. We can find an additional $(k-2\eta^{-1})/2$ cells which are not great by considering the "lowest" cell in Γ_0. Thus, Γ_0 is contained in a block which has $k-2\eta^{-1}$ cells which are not great, which is a contradiction. This completes the proof of our claim, and the proof of Theorem 2.

□

□

5 Hamilton Cycles: proof of Theorem 3

Proof. We begin by decomposing $G_{n,p}$ into $G_{n,p_1} \cup G_{n,p_2}$ where $p_1 = p_2$ and $1-p = (1-p_1)(1-p_2)$. Then we partition D into cells $C_1, C_2, \ldots, C_m$, $m = 5/r^2$ of side $r/\sqrt{5}$. We assume that the cell order is such that it defines a sequence where C_i, C_{i+1} share an edge for $1 \leq i < m$. Now the expected number of points in a cell is $nr^2/5 \geq K\log n/(5p)$. So we can assume that K is large enough so that w.h.p. every cell has at least $10\log n/p$ points.

Now if the are n_i points in a cell i then these points induce a random graph $\Gamma_i = G_{n_i,p_1}$ where we have assumed that $n_i \geq 10\log n/p$ for $i \in [m]$. It is not difficult to show that $\mathbb{P}(G_{N,\xi}$ is not Hamiltonian$) \leq 2N^2qe^{-Nq}$ when $\xi = O(\log N/N)$ (see for example Chapter 6 of Frieze and Karoński [1] where we see that this probability is dominated by the probability of the existence of vertices of degree less than two). It follows from this that w.h.p., Γ_i is Hamilton for $i \in [m]$. Indeed,

$$\mathbb{P}(\exists i : \Gamma_i \text{ is not Hamiltonian}) \leq 2\sum_{i=1}^{m} n_i^2 p_1 e^{-n_i p_1} \leq 2n^3 p_1 n^{-10p_1/p} = o(1),$$

since $p_1, p_2 \geq p/2$.

Assume then that Γ_i contains a Hamilton cycle for $i \in [m]$. Divide C_i into two paths $P_{i,-}, P_{i,+}$ of length $\lfloor n/2 \rfloor$ or $\lceil n/2 \rceil$. We will show next that w.h.p. for all $i \in [m-1]$ there is are edges $\{a_i, b_i\} \in P_{i,+}, \{c_i, d_i\} \in P_{i+1,-}$ such that G_{n,p_2} contains both edges $\{a_i, c_i\}, \{b_i, d_i\}$. If for $i \in [m-1]$, we delete $\{a_i, b_i\}, \{c_i, d_i\}$ from $\bigcup_{i=1}^m C_i$ and replace them by $\{a_i, c_i\}, \{b_i, d_i\}$ then we obtain a Hamilton cycle. The probability that we fail to make this transformation can be bounded by

$$\sum_{i=1}^{m-1}(1-p_2^2)^{n_i n_j/16} \leq me^{-100p_2^2\log^2 n/16p^2} = o(1).$$

Here we have used every other edge in the paths $P_{i,+}, P_{i,-}$ to avoid dependencies. □

6 Independence number: proof of Theorem 4

Proof. We start with the lower bound. We use cells of side length r. By removing every other row and every other column of cells, we obtain a set of $1/(4r^2)$ cells so that no two cells in our set intersect share a boundary (not even a corner). The distance between any two distinct cells is at least r, so no point in one would ever be adjacent to any point in another. By Chernoff, w.h.p. every cell in C has at least $n' = nr^2/2$ points. W.h.p. (see, for example, [1]) $G(n', p)$ has an independent set of size at least $\log_{1/(1-p)}(n')$. Thus, for constant p w.h.p. we have

$$\alpha(G) \geq |C|\log_{1/(1-p)}(n') = \frac{1}{4r^2}\log_{1/(1-p)}(nr^2/2) = \Omega\left(r^{-2}\log_{1/(1-p)}(nr^2)\right).$$

We move on to the upper bound. Use cells of side length $r/\sqrt{2}$, so there are $2r^{-2}$ cells. The Chernoff bounds imply that w.h.p. no cell has more than $n'' := 2n\left(r/\sqrt{2}\right)^2 = nr^2$ points. The expected number of independent sets of size $k := 3\log_{1/(1-p)} n''$ in $G(n'', p)$ is at most

$$\binom{n''}{k}(1-p)^{\binom{k}{2}} \leq \left(n''(1-p)^{\frac{k-1}{2}}\right)^k \leq \left(O(1)nr^2(1-p)^{\frac{3}{2}\log_{1/(1-p)}(nr^2)}\right)^k = \left(O((nr^2)^{-1/2})\right)^k,$$

small enough for the union bound to conclude that w.h.p. no cell has an independent set larger than k. Summing over the cells we have

$$\alpha(G) \leq n''k = 2r^{-2} \cdot 3\log_{1/(1-p)}\left(nr^2\right) = O\left(r^{-2}\log_{1/(1-p)}(nr^2)\right).$$

This completes the proof of the upper bound. □

7 Chromatic number: proof of Theorem 5

Proof. First we prove the upper bound. Use cells of side length r. We will use 4 disjoint palettes of colors. Each palette will be used on a set of $(2r)^{-2}$ cells in every other row and every other column, so that two cells using the same palette are always distance at least r apart. The Chernoff bound implies that w.h.p. every cell has at most $n' := 2nr^2$ points. We will use the following result:

Theorem 8 (Theorems 7.7, 7.8 in [1]). *For constant $0 < p < 1$ we have*

$$\mathbb{E}(\chi(G(n,p))) = (1+o(1))\frac{n}{2\log_{1/(1-p)} n}$$

and

$$\mathbb{P}(|\chi(G(n,p) - \mathbb{E}(\chi(G(n,p)))| \geq t) \leq \exp(-t^2/2n).$$

Thus we have

$$\mathbb{P}\left(G(n',p) \geq \frac{n'}{\log_{1/(1-p)} n'}\right) \leq \exp\left(-\frac{(1+o(1))\left(\frac{n'}{2\log_{1/(1-p)} n'}\right)^2}{2n'}\right) = \exp\left(-\Omega\left(\frac{n'}{\log^2 n'}\right)\right),$$

which is small enough to beat the union bound over all cells. Thus, w.h.p. within each cell we have a graph which requires at most $n'/\log_{1/(1-p)} n'$ colors. Thus each palette can have that many colors, and we have

$$\chi(G) \leq \frac{4n'}{\log_{1/(1-p)}(n')} = O\left(\frac{nr^2}{\log_{1/(1-p)}(nr^2)}\right).$$

Now we prove the lower bound. Using Theorem 4, we have

$$\chi(G) \geq \frac{n}{\alpha(G)} = \Omega\left(\frac{nr^2}{\log_{1/(1-p)}(nr^2)}\right).$$

□

8 Diameter: proof of Theorem 6

Proof. We first prove the lower bound. We use cells of side length r, and the Chernoff bounds imply that w.h.p. every cell has a point in it. Taking two points in farthest-apart cells, they have Euclidean distance $\Omega(1)$ and so their graph distance is $\Omega(r^{-1})$. Thus

$$\operatorname{diam}(G) = \Omega(r^{-1}). \tag{11}$$

Now note that by the Chernoff bounds, w.h.p. every vertex has degree at most $2nq = 2\pi nr^2 p$, i.e. about twice the expected degree. Assuming this, the number of vertices within a distance i of a point v is at most

$$1 + 2\pi nr^2 p + \ldots + (2\pi nr^2 p)^i = (1+o(1))(2\pi nr^2 p)^i.$$

Therefore if $(2\pi n r^2 p)^i < n/2$ then $\mathrm{diam}(G) > i$. In particular,

$$\mathrm{diam}(G) > \frac{\log(n/2)}{\log(2\pi n r^2 p)} = \Omega\left(\frac{\log n}{\log(nr^2p)}\right). \tag{12}$$

Combining (11) and (12) completes our lower bound:

$$\mathrm{diam}(G) = \Omega\left(r^{-1} + \frac{\log n}{\log(nr^2p)}\right).$$

Now we prove the upper bound. Here we will use cells of side length $r/2\sqrt{2}$. If two cells intersect at all (even at a single point) then every point in one cell is within distance r of every point in the other cell. The Chernoff bounds imply that w.h.p. every cell has at least $n' := nr^2/16$ points (half the expected number).

Consider any two points u, v. Let

$$\ell := \max\left\{\frac{\sqrt{2}}{r}, \frac{\log(n'/10)}{\log(n'p/2)}\right\}$$

Since $\ell \geq \sqrt{2}/r$, there is a sequence of (not necessarily distinct) cells $C_0, \ldots, C_\ell$ such that $u \in C_0$, $v \in C_\ell$, and any pair of consecutive cells in our sequence intersects. Let $S_0 = \{u\}$ and for $i = 1, \ldots \ell$, let S_i be some set of points in C_i that have neighbors in S_{i-1}. We will show that w.h.p. we can take

$$|S_i| = \min\left\{\left(\frac{n'p}{2}\right)^i, \frac{n'}{10}\right\}, \qquad i = 0, \ldots, \ell. \tag{13}$$

We will prove this by induction on i. The base case $i = 0$ holds since $|S_0| = 1$. We have already revealed the location of all the points in the plane, so our inductive proof can proceed by revealing adjacencies between points in S_i and points in C_{i+1} (so each edge is present with probability p). For the induction step, assume it holds for some $i < \ell$. The number X of vertices in C_{i+1} and not in S_i that are adjacent to some vertex in S_i is distributed as $\mathrm{Bin}\left(n'', 1 - (1-p)^{|S_i|}\right)$, where $n'' \geq 9n'/10$ is the number of points in C_{i+1} and not S_i.

Case 1: Suppose we have $\left(\frac{n'p}{2}\right)^{i+1} \leq \frac{n'}{10}$. Then $|S_i| = \left(\frac{n'p}{2}\right)^i \leq \frac{1}{5p}$, so

$$\begin{aligned}
\mathbb{E}[X] = n''\left[1 - (1-p)^{|S_i|}\right] &\geq \frac{9n'}{10}\left[1 - \exp\{-|S_i|p\}\right] \\
&\geq \frac{9n'}{10}\left[|S_i|p - \frac{1}{2}(|S_i|p)^2\right] \\
&\geq \frac{9}{10}\left[1 - \frac{1}{50}\right]|S_i|n'p > \frac{3}{4}\left(\frac{n'p}{2}\right)^i n'p.
\end{aligned} \tag{14}$$

Since the above expectation is $\Omega(n'p) = \Omega(\log^2 n)$, the Chernoff bounds imply that with probability at least $1 - \exp(-\Omega(\log^2 n))$, the random variable X is at least $\left(\frac{n'p}{2}\right)^{i+1}$ (2/3 of (14)). The failure probability is small enough to beat the union bound over ℓ steps. This completes the proof for Case 1.

Case 2: Suppose we have $\left(\frac{n'p}{2}\right)^{i+1} > \frac{n'}{10}$. Then $|S_i| \geq \left(\frac{n'p}{2}\right)^i > \frac{1}{5p}$, so

$$\mathbb{E}[X] = n''\left[1 - (1-p)^{|S_i|}\right] \geq \frac{9n'}{10}\left[1 - \exp\{-|S_i|p\}\right] > \frac{9n'}{10}\left[1 - \exp\{-1/5\}\right] > 0.15n'. \tag{15}$$

Since the above is $\Omega(\log^2 n)$, again Chernoff gives us that with probability at least $1 - \exp(-\Omega(\log^2 n))$, the random variable X is at least $\frac{n'}{10}$ (2/3 of (15)). This completes the proof for Case 2. Thus w.h.p. (13) holds.

Because $\ell \geq \frac{\log(n'/10)}{\log(n'p/2)}$, we have

$$\left(\frac{n'p}{2}\right)^{\ell} \geq \frac{n'}{10}$$

and so $|S_\ell| = n'/10$. The probability there is no edge from v to S_ℓ is then $(1-p)^{n'/10} = \exp\{-\Omega(n'p)\} = \exp\{-\Omega(\log^2 n)\}$, so w.h.p. there is such an edge. Thus the distance from u to v is at most

$$\ell + 1 = 1 + \max\left\{\frac{\sqrt{2}}{r}, \frac{\log(n'/10)}{\log(n'p/2)}\right\} = O\left(r^{-1} + \frac{\log n}{\log(nr^2p)}\right)$$

□

References

[1] A.M. Frieze and M. Karoński, Introduction to Random Graphs, Cambridge University Press, 2015.

[2] M. Kahle, M. Tian and Y. Wang, On the clique number of noisy random geometric graphs, Random Structures and Algorithms 63 (2023) 242-279.

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