Disjoint Hamilton cycles in the random geometric graph

Tobias Müller*
Centrum Wiskunde & Informatica
P.O. Box 94079
1090 GB Amsterdam, The Netherlands
tobias@cwi.nl

Xavier Pérez-Giménez†
Department of Combinatorics and Optimization
University of Waterloo
Waterloo ON, Canada
xperez@uwaterloo.ca

Nicholas Wormald‡
Department of Combinatorics and Optimization
University of Waterloo
Waterloo ON, Canada
nwormald@uwaterloo.ca

Abstract

We consider the standard random geometric graph process in which n vertices are placed at random on the unit square and edges are sequentially added in increasing order of edge-length. For fixed $k \geq 1$, we prove that the first edge in the process that creates a k-connected graph coincides a.a.s. with the first edge that causes the graph to contain $k/2$ pairwise edge-disjoint Hamilton cycles (for even k), or $(k-1)/2$ Hamilton cycles plus one perfect matching, all of them pairwise edge-disjoint (for odd k). This proves and extends a conjecture of Krivelevich and Müller. In the special when case $k = 2$, our result says that the first edge that makes the random geometric graph Hamiltonian is a.a.s. exactly the same one that gives 2-connectivity, which answers a question of Penrose. (This result appeared in three independent preprints, one of which was a precursor to this paper.) We prove our results with lengths measured using the ℓ_p norm for any $p > 1$, and we also extend our result to higher dimensions.

1 Introduction

Many authors have studied the evolution of the random geometric graph on n labelled vertices placed independently and *uniformly at random* (u.a.r.) on the unit square $[0, 1]^2$, in which edges are added in increasing order of length (see e.g. [7]). Penrose [6] proved that the first added edge that makes the graph have minimum degree k is *asymptotically almost surely* (a.a.s., i.e. with probability tending to 1 as $n \to \infty$) the first one that makes it k-connected. He also asked whether, in the evolution of the random geometric graph, 2-connectivity occurs a.a.s. precisely when the first Hamilton cycle is created. The setting was not only for Euclidean edge lengths, but also for the ℓ_p norm for any $p > 1$. As a first step towards answering Penrose's question, Díaz, Mitsche and the second author showed in [3] that the property of

*Research partially supported by a VENI grant from Netherlands Organisation for Scientific Research (NWO).
†Partially supported by the Province of Ontario under the Post-Doctoral Fellowship (PDF) Program.
‡Supported by the Canada Research Chairs Program and NSERC.

being Hamiltonian has a sharp threshold at $r \sim \sqrt{\log n/(\theta n)}$ (here θ denotes the area of the unit ball with respect to the norm used), which coincides asymptotically with the threshold for k-connectivity for any constant k. On the other hand, a more general result is already known for the evolution of the random graph G on n labelled vertices, in which edges are added one by one. Bollobás and Frieze showed in [2] that a.a.s. as soon as G has minimum degree k, it also contains $\lfloor k/2 \rfloor$ edge-disjoint Hamilton cycles plus an additional edge-disjoint perfect matching if k is odd, where k is any constant positive integer. The main result in the present paper is that the analogue of Bollobás and Frieze's result holds for the random geometric graph. That is, we show that, in the evolution of the random geometric graph, a.a.s. as soon as the graph becomes k-connected, it immediately contains $\lfloor k/2 \rfloor$ edge-disjoint Hamilton cycles plus one additional perfect matching if k is odd. This result applies to the geometric graph in a unit hypercube of $d \geq 2$ dimensions, and for the ℓ_p norm, $1 < p \leq \infty$. This proves and extends a conjecture of Krivelevich and the first author [4], who conjectured the result for even values k. The special case $k = 2$ of our result answers Penrose's question.

Three independent but somewhat similar proofs for Penrose's question appeared in preprints by Balogh, Bollobás and Walters [1], by Krivelevich and Müller [4] and by Pérez-Giménez and Wormald [8]. The present paper presents the proof in [8] (which is only for dimension 2 but does cover arbitrary k) and additionally includes the extension to arbitrary dimension d making use of a result in [4]. Part of the original argument in [8] is considerably shortened here, by excluding probabilistically some vertex configurations which were treated in [8] by proving a stronger result on packing linear forests in graphs.

Let $\boldsymbol{X} = (X_1, \dots, X_n)$ be a random vector, where each X_i is a point in $[0,1]^2$ chosen independently with uniform distribution. Given $\boldsymbol{X}$ and a radius $r = r(n) \geq 0$, we define the *random geometric graph* $\mathscr{G}(\boldsymbol{X}; r)$ as follows: the vertex set of $\mathscr{G}(\boldsymbol{X}; r)$ is $\{1, \dots, n\}$ and there is an edge joining i and j whenever $\|X_i - X_j\|_p \leq r$. Here $\|\cdot\|_p$ denotes the standard ℓ_p norm, for some fixed $1 < p \leq \infty$. Unless otherwise stated, all distances in $[0,1]^2$ are measured according to the ℓ_p norm (i.e. $d(X,Y) = \|X - Y\|_p$). Let θ be the area of the unit ℓ_p-ball (e.g. $\theta = \pi$ for $p = 2$, and $2 \leq \theta \leq 4$ for all $1 \leq p \leq \infty$).

A continuous-time random graph process $\big(\mathscr{G}(\boldsymbol{X}; r)\big)_{0 \leq r < \infty}$ is defined in a natural way, by first choosing the random set of points $\boldsymbol{X}$ and then adding edges one by one as we increase the radius r from 0 to ∞.

Theorem 1. *Consider the random graph process* $\big(\mathscr{G}(\boldsymbol{X}; r)\big)_{0 \leq r < \infty}$ *for any* ℓ_p*-normed metric on* $[0,1]^2$, $1 < p \leq \infty$, *and let* k *be a fixed positive integer.*

(i) For even $k \geq 2$, *a.a.s. the minimum radius* r *at which the graph* $\mathscr{G}(\boldsymbol{X}; r)$ *is* k*-connected is equal to the minimum radius at which it has* $k/2$ *edge-disjoint Hamilton cycles.*

(ii) For odd $k \geq 1$, *a.a.s. the minimum radius* r *at which the graph* $\mathscr{G}(\boldsymbol{X}; r)$ *is* k*-connected is equal to the minimum radius at which it has* $(k-1)/2$ *Hamilton cycles and one perfect matching, all of them pairwise edge-disjoint. (Here asymptotics are restricted to even* n*.)*

The reason that we restrict ourselves to the ℓ_p norm with $p > 1$ to measure the edge-lengths in Theorem 1 (as opposed to a completely arbitrary norm), is that this restriction is imposed by the results of Penrose that we invoke in the proof of Theorem 1.

The next section contains the basic geometric definitions and probabilistic statements required in the argument, including proofs that several properties hold a.a.s. Then, in Section 3, we prove the main theorem, by supplying the required construction of Hamilton cycles (and perfect matching) in the random geometric graph deterministically, assuming the properties that were shown to hold a.a.s. Finally, in Section 4, we extend the argument to general dimension.

Throughout this paper, we use $N_G(v)$ to denote the set of neighbours of a vertex v in a graph G (the subscript G may be omitted when it is clear from the context).

2 Asymptotically almost sure properties

Let $k \geq 1$ be a fixed integer, and define $m = 2k-3$ if $k \geq 2$ and $m = 0$ if $k = 1$. We state a result which is a consequence of Theorem 8.4 in [7].

Proposition 2. *In the random process* $\left(\mathscr{G}(\boldsymbol{X};r)\right)_{0\leq r<\infty}$, *let* r_k *be the smallest* r *such that* $\mathscr{G}(\boldsymbol{X};r)$ *is* k*-connected. Then,*

$$\theta n {r_k}^2 - \log n - m \log\log n \tag{1}$$

is bounded in probability.

Here, we define some properties of the random geometric graph that hold a.a.s. and that will turn out to be sufficient for our construction of disjoint Hamilton cycles. In view of Proposition 2, we shall mainly focus our analysis to r satisfying $\theta n r^2 = \log n + m \log\log n + O(1)$ or sometimes just $\theta n r^2 \sim \log n$.

Henceforth we assume that the points in $\boldsymbol{X}$ are in general position—i.e. they are are all different, no three of them are collinear, and all distances between pairs of points are strictly different—since this holds with probability 1. The first lemma shows that sets of vertices with relatively few common neighbours are rare.

Lemma 3. *For any small enough constant* $\eta > 0$ *and any* r *such that* $\theta n r^2 = \log n + m \log\log n + O(1)$, *and constant integers* $j \geq 2$ *and* $K \geq j + k$, *the random geometric graph* $\mathscr{G}(\boldsymbol{X};r)$ *a.a.s. satisfies the following property. Every set* J *of* j *vertices such that* $\max_{u,v\in J}\{d(X_u, X_v)\} \leq \eta r$ *and* $\max_{u\in J} |N_{\mathscr{G}(\boldsymbol{X};r)}(u)| \leq K$ *has at least* k *common neighbours not in* J.

Proof. The vertices $v_1, \ldots, v_N$ form a *bad configuration* if there are j_1, j_2, j_3, j_4 such that $N = j_1 + j_2 + j_3 + j_4$ and the following conditions are met:

(b1) $2 \leq j_1 \leq K+1$, $j_2 \leq k-1$, $j_3 \leq K$ and $j_4 \leq K$;

(b2) X_{v_1} is closer to the boundary of $[0,1]^2$ than X_{v_2};

(b3) $\|X_{v_1} - X_{v_i}\|_2 \leq \|X_{v_1} - X_{v_2}\|_2 \leq \sqrt{2}\eta r$ for $i = 2, \ldots, j_1$;

(b4) $\boldsymbol{X} \cap B_p(X_{v_1}, r - \|X_{v_1} - X_{v_2}\|_2 \cdot \sqrt{2}) = \{X_{v_1}, \ldots, X_{v_{j_1+j_2}}\}$;

(b5) $\boldsymbol{X} \cap B_p(X_{v_1}, r) = \{X_{v_1}, \ldots, X_{v_{j_1+j_2+j_3}}\}$;

(b6) $\boldsymbol{X} \cap (B_p(X_{v_2}, r) \cup B_p(X_{v_1}, r)) = \{X_{v_1}, \ldots, X_{v_{j_1+j_2+j_3+j_4}}\}$.

Here and elsewhere in the paper, $B_p(x,s) := \{y \in \mathbb{R}^2 : \|x-y\|_p < s\}$ denotes the l_p-ball around x of radius s. See Figure 1 for a depiction of a bad configuration. Observe that if a set J violating the conclusion of the lemma exists, then there must be a bad configuration. To see this, let us think of $\{v_1, \ldots, v_{j_1}\}$ as J, where we assume w.l.o.g. that X_{v_1} and X_{v_2}

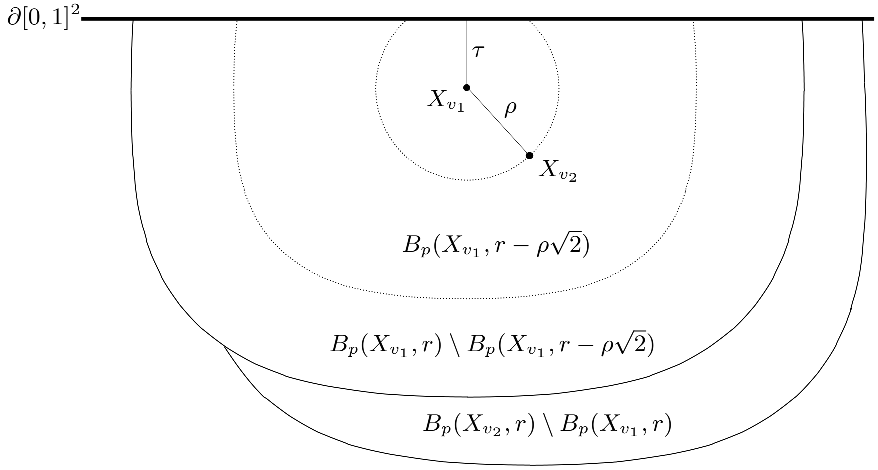

Figure 1: Visual description of a bad configuration. Points $X_{v_3}, \dots, X_{v_{j_1}}$ must lie in $B_2(X_{v_1}, \rho)$. The points $X_{v_{j_1+1}}, \dots, X_{v_{j_1+j_2}}$ must lie in $B_p(X_{v_1}, r - \rho\sqrt{2})$. The points $X_{v_{j_1+j_2+1}}, \dots, X_{v_{j_1+j_2+j_3}}$ lie in $B_p(X_{v_1}, r) \setminus B_p(X_{v_1}, r-\rho\sqrt{2})$, and $X_{v_{j_1+j_2+j_3+1}}, \dots, X_{v_{j_1+j_2+j_3+j_4}}$ lie in $B_p(X_{v_2}, r) \setminus B_p(X_{v_1}, r)$.

realise the Euclidean diameter of $\{X_v : v \in J\}$ and X_{v_1} is closer to the boundary of $[0,1]^2$ than X_{v_2} is. Also let the vertices $v_{j_1+1}, \dots, v_{j_1+j_2+j_3+j_4}$ be defined from the conditions **(b4–6)**. Note that, with this construction, **(b2–6)** are trivially satisfied, and also $2 \le j_1 = |J| \le K+1$. Moreover, $j_2 \le k-1$ because $|\bigcap_{i=1}^{j_1} N_{\mathscr{G}(\mathbf{X};r)}(v_i)| \le k-1$ (by our choice of J) and $B_p(X_{v_1}, r - \|X_{v_1} - X_{v_2}\|_2 \cdot \sqrt{2}) \subseteq \bigcap_{i=1}^{j_1} B_p(X_{v_i}, r)$ by condition **(b3)** and the fact that $\|x\|_p \le \sqrt{2}\|x\|_2$ for all $x \in \mathbb{R}^2$ and all $p \ge 1$. Finally, $j_3 \le K$ and $j_4 \le K$ because the degrees of v_1 and v_2 are at most K. Hence condition **(b1)** is also satisfied.

We distinguish three types of bad configurations according to the position of X_{v_1} in $[0,1]^2$: *corner* bad configurations are those in which X_{v_1} is at distance at most r from two of the four sides of $[0,1]^2$; *side* bad configurations are those in which X_{v_1} is at distance at most r from exactly one of the four sides of $[0,1]^2$; all other bad configurations are referred to as *interior* bad configurations.

Let T_0, T_1 and T_2 denote respectively the number of corner, side and interior bad configurations. An easy calculation shows that the expected number of vertices of degree at most K near the corners is $o(1)$, so that also $T_0 = 0$ a.a.s.

Let us now consider $\mathbf{E}T_1$. Given a bad configuration, let $\tau \le r$ denote the distance of X_{v_1} from the boundary and let $\rho := \|X_{v_1} - X_{v_2}\|_2$. It can be seen that, provided η was chosen sufficiently small, we have

$$c_1 \rho r \le \text{area}\left([0,1]^2 \cap B_p(X_{v_2}, r) \setminus B_p(X_{v_1}, r)\right) \le c_2 \rho r,$$

and

$$c_1 \rho r \le \text{area}\left([0,1]^2 \cap B_p(X_{v_1}, r) \setminus B_p(X_{v_1}, r - \rho\sqrt{2})\right) \le c_2 \rho r,$$

for suitably chosen constants $c_1, c_2 > 0$ (where $c_1 = 1/100, c_2 = 100$ will do for our purposes). We shall also use the observation that

$$\text{area}([0,1]^2 \cap B_p(X_{v_1}, r)) \ge \frac{\theta}{2}(r^2 + r\tau).$$

We see that

$$
\begin{aligned}
&\left(1-\operatorname{area}([0,1]^2\cap(B_p(X_{v_1},r)\cup B_p(X_{v_2},r)))\right)^{n-(j_1+j_2+j_3+j_4)} \\
&\qquad\qquad\qquad\qquad \leq \left(1-\theta r^2/2-\theta r\tau/2-c_1 r\rho\right)^{n-4K} \\
&\qquad\qquad\qquad\qquad \leq 2\exp\left[-\theta n r^2/2-\theta n r\tau/2-c_1 n r\rho\right],
\end{aligned}
$$

where the second inequality holds for n sufficiently large. Shortly, we will also use the bounds that the area of $[0,1]^2\cap B_2(X_{v_1},\rho)$ is at most $\pi\rho^2$ and that of $[0,1]^2\cap B_p(X_{v_1},r-\rho\sqrt{2})$ is at most θr^2. In this way, we can bound $\mathbf{E}T_1$ by summing over all possible choices of $v_1,\ldots,v_N$ the probability that they form a bad configuration. This probability can be obtained by integrating over τ and ρ the probability that the points $\{X_{v_3},\ldots,X_{v_N}\}$ lie in their respective regions according to **(b3–6)** (using also the previous bounds on the area of these regions). Integration over ρ is achieved by changing to polar coordinates, and we use the fact that the probability density function of the distance between X_{v_1} and the closest side of $[0,1]^2$ is at most 4).

$$
\begin{aligned}
\mathbf{E}T_1 &\leq \sum 8n^{j_1+j_2+j_3+j_4}\int_0^r\int_0^{\sqrt{2}\eta r}\left(\pi\rho^2\right)^{j_1-2}\left(\theta r^2\right)^{j_2}(c_2 r\rho)^{j_3+j_4} \\
&\qquad\times\exp\left[-\frac{\theta n r^2}{2}-\frac{\theta n r\tau}{2}-c_1 n r\rho\right]\pi\rho\,d\rho\,d\tau \\
&= O\left(\sum n^{j_1+j_2+j_3+j_4}r^{2j_2+j_3+j_4}\int_0^r\int_0^{\sqrt{2}\eta r}\rho^{2j_1+j_3+j_4-3}\exp\left[-\frac{\theta n r^2}{2}-\frac{\theta n r\tau}{2}-c_1 n r\rho\right]d\rho\,d\tau\right),
\end{aligned}
$$

where both sums are over all $j_1,\ldots,j_4$ that satisfy **(b1)** above. Applying the substitutions $s=c_1 n r\rho$ and $t=\theta n r\tau/2$, we get

$$
\begin{aligned}
&\int_0^r\int_0^{\sqrt{2}\eta r}\rho^{2j_1+j_3+j_4-3}e^{-\theta n r\tau/2-c_1 n r\rho}d\rho d\tau \\
&\qquad= \int_0^{\theta n r^2/2}\int_0^{c_1\sqrt{2}\eta n r^2}\left(\frac{s}{c_1 n r}\right)^{2j_1+j_3+j_4-3}e^{-s-t}\frac{ds}{c_1 n r}\frac{2dt}{\theta n r} \\
&\qquad= O\left((nr)^{-(2j_1+j_3+j_4-1)}\int_0^{\theta n r^2/2}\int_0^{c_1\sqrt{2}\eta n r^2}s^{2j_1+j_3+j_4-3}e^{-s-t}\,ds\,dt\right) \\
&\qquad= O\left((nr)^{-(2j_1+j_3+j_4-1)}\right).
\end{aligned}
$$

Since $\theta n r^2=\log n+m\log\log n+O(1)$, we have $e^{-\theta n r^2/2}=O\left(n^{-1/2}\log^{-m/2}n\right)$. Putting everything together, we find

$$
\begin{aligned}
\mathbf{E}T_1 &= O\left(\sum n^{-j_1+j_2+1}r^{-2j_1+2j_2+1}n^{-1/2}\log^{-m/2}n\right) \\
&= O\left(\sum (nr^2)^{j_2-j_1}rn^{1/2}\log^{-m/2}n\right) \\
&= O\left(\sum (\log n)^{j_2-j_1+(1-m)/2}\right) \\
&= O\left(\log^{-1}n\right)=o(1),
\end{aligned}
$$

using $nr^2=\Theta(\log n)$ and $n^{1/2}r=\Theta(\log^{1/2}n)$ to arrive at the third line, and $j_1\geq 2$, $j_2\leq k-1$, $m\geq 2k-3$ so that $j_2-j_1+(1-m)/2\leq -1$ to arrive at the last line (together with the fact that we are summing over finitely many choices of the js).

It remains to show that $T_2 = 0$ a.a.s. The computations are fairly similar to these in the analysis of T_1 (the main difference being that we do not need the parameter τ), and we omit them. We now get

$$\mathbf{E}T_2 = O\left(\sum n^{j_1+j_2+j_3+j_4} r^{2j_2+j_3+j_4} \int_0^{\sqrt{2}\eta r} \rho^{2j_1+j_3+j_4-3} \exp\left[-\theta n r^2 - c n r \rho\right] d\rho\right).$$

This is $o(1)$, by computations similar to those for $\mathbf{E}T_1$. $\square$

For the following definitions, we fix $\delta > 0$ to be a small enough constant and assume $r \to 0$. We tessellate $[0,1]^2$ into square *cells* of side $\delta' r = \lceil (\delta r)^{-1} \rceil^{-1}$. (Note that δ' is not constant, but $\delta' \leq \delta$ and $\delta' \to \delta$). Let $\mathcal{C}$ be the set of cells, and let $\mathscr{G}_{\mathcal{C}}$ be an auxiliary graph with vertex set $\mathcal{C}$ and with one edge connecting each pair of cells c_1 and c_2 iff all points in c_1 have distance at most r from all points in c_2. Note that we shall use the term *adjacent cells* to refer to cells which are adjacent vertices of the graph of cells $\mathscr{G}_{\mathcal{C}}$, while cells sharing a side boundary will be described as being *topologically adjacent*. Let Δ be the maximum degree of $\mathscr{G}_{\mathcal{C}}$. By construction, Δ is a constant only depending on δ and the chosen ℓ_p norm.

We may assume that each point in $\boldsymbol{X}$ lies strictly in the interior of a cell in the tesselation, since this happens with probability 1. Let M be a large enough but constant positive integer (its choice will only depend on Δ, thus on δ, and also on k and ℓ_p). A cell in $\mathcal{C}$ is *dense* if it contains at least M points of the random set $\boldsymbol{X}$, *sparse* if it contains at least one, but less than M, points in $\boldsymbol{X}$, and *empty* if it has no points in $\boldsymbol{X}$. Let $\mathcal{D} \subseteq \mathcal{C}$ be the set of dense cells. Note that $\mathcal{D} \neq \emptyset$, since the total number of cells is $|\mathcal{C}| = \Theta(n/\log n)$, so at least one must contain $\Omega(\log n)$ points in $\boldsymbol{X}$.

A set of cells is said to be *connected* if it induces a connected subgraph of $\mathscr{G}_{\mathcal{C}}$. (For δ small enough, this includes the situation where the union of cells is topologically connected.) The *area* of a set of cells is simply the area of the corresponding union of cells. A set of cells *touches* one side (or one corner) of $[0,1]^2$ if it contains a cell which has some boundary on that side (or corner) of the unit square.

Lemma 4. *For any constants $\delta > 0$ and $\alpha > 0$ and for any r satisfying $\theta n r^2 \sim \log n$, the following statements hold a.a.s.*

1. *All connected sets of cells of area at least $(1+\alpha)\theta r^2$ contain some dense cell.*
2. *All connected sets of cells of area at least $(1+\alpha)\theta r^2/2$ touching some side of $[0,1]^2$ contain some dense cell.*
3. *All cells contained inside a $5r \times 5r$ square on each corner of $[0,1]^2$ are dense.*

Proof. Recall that the area of each cell is $\delta'^2 r^2$. Then, in order to show the first statement in the lemma, it suffices to consider all connected sets of cells with exactly $s = \lceil (1+\alpha)\theta/\delta'^2 \rceil = \Theta(1)$ cells. Let $\mathcal{S}$ be such a set of cells. The probability that $\mathcal{S}$ has no dense cell is at most

$$\begin{aligned}\sum_{t=0}^{(M-1)s} \binom{n}{t} (s\delta'^2 r^2)^t (1 - s\delta'^2 r^2)^{n-t} &= O\left(e^{-(1+\alpha)\theta r^2 n}\right) \sum_{t=0}^{(M-1)s} (r^2 n)^t \\ &= O\left(n^{-(1+\alpha)+o(1)} \log^{(M-1)s} n\right). \qquad (2)\end{aligned}$$

To conclude the first part of the proof, multiply the probability above by the number $\Theta(1/r^2) = \Theta(n/\log n)$ of connected sets of s cells.

By a completely analogous argument, if $\mathcal{S}$ has area only $\lceil(1+\alpha)\theta/\delta'^2\rceil/2$ and touches some side of $[0,1]^2$, the probability that it has no dense cell is $O(n^{-(1+\alpha)/2)+o(1)}\log^{(M-1)s} n$. However, the number of such sets is only $\Theta(\sqrt{n/\log n})$.

Finally, there is a bounded number of cells inside any of the $5r\times 5r$ squares on the corners, and each individual cell is dense with probability $1-o(1)$. □

A set of cells is *small* if it can be embedded in a 16×16 grid of cells, and it is *large* otherwise. Consider the subgraph $\mathscr{G}_{\mathcal{C}}[\mathcal{D}]$ of $\mathscr{G}_{\mathcal{C}}$ induced by dense cells, and let $\mathcal{D}_0$ be the set of dense cells which are not in small components of $\mathscr{G}_{\mathcal{C}}[\mathcal{D}]$ (we shall see that $\mathcal{D}_0$ forms a unique large component in $\mathscr{G}_{\mathcal{C}}[\mathcal{D}]$). Most of the trouble in our argument comes from cells which are not adjacent to any dense cell in $\mathcal{D}_0$, so let $\mathcal{B}=\mathcal{C}\setminus(\mathcal{D}_0\cup N(\mathcal{D}_0))$, and call the cells in $\mathcal{B}$ *bad* cells. Also, let us denote components of $\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$ as *bad* components. Note that by construction all cells in $N(\mathcal{B})\setminus\mathcal{B}$ must be sparse (or empty) but adjacent to some cell in $\mathcal{D}_0$, while $\mathcal{B}$ itself may contain both sparse and dense cells. Figure 2 illustrates some of these definitions.

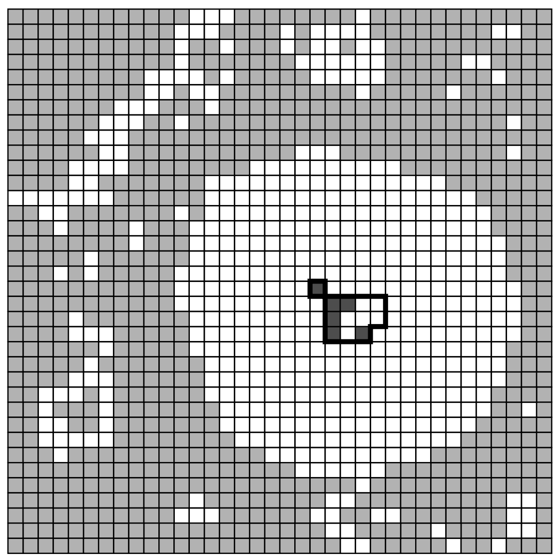

Figure 2: Example of tesselation of $[0,1]^2$ into cells. Shaded cells represent dense cells (i.e. cells in $\mathcal{D}$), while the white ones are either sparse or empty cells. The graph $\mathscr{G}_{\mathcal{C}}[\mathcal{D}]$ induced by the dense cells has two components: one large (light gray cells) and one small (dark gray cells), so $\mathcal{D}_0$ is precisely the set of light gray cells. The bad cells (i.e. cells in $\mathcal{B}$) are the ones enclosed by the thicker black countour, and $\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$ consists of one single small (bad) component. This bad component has both dense and non-dense cells.

Lemma 5. *For a small enough constant $\delta>0$ and for any r satisfying $\theta nr^2\sim\log n$, the following holds a.a.s.*

1. *All components of $\mathscr{G}_{\mathcal{C}}[\mathcal{D}]$ are small except for one large component formed by precisely the cells in $\mathcal{D}_0$.*

2. *$\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$ has only small components.*

Proof. First, we claim that the following statements are a.a.s. true. (Recall that "connected" is defined in terms of the graph $\mathscr{G}_{\mathcal{C}}$, not topological adjacency.)

1. For any large connected set of cells $\mathcal{S}$ such that $N(\mathcal{S})$ does not touch all four sides of $[0,1]^2$, $N(\mathcal{S}) \setminus \mathcal{S}$ must contain some dense cell.

2. For any pair of connected sets of cells $\mathcal{S}_1$ and $\mathcal{S}_2$ not adjacent to each other (i.e. $\mathcal{S}_2 \cap N(\mathcal{S}_1) = \emptyset$) and such that both $N(\mathcal{S}_1)$ and $N(\mathcal{S}_2)$ touch all four sides of $[0,1]^2$, $N(\mathcal{S}_1)\setminus\mathcal{S}_1$ or $N(\mathcal{S}_2) \setminus \mathcal{S}_2$ must contain some dense cell.

As an immediate consequence of this claim, by considering the maximal connected sets of dense cells, we deduce that $\mathscr{G}_\mathcal{C}[\mathcal{D}]$ must have a unique large component, consisting of all cells in $\mathcal{D}_0$ (note that $\mathcal{D}_0 \neq \emptyset$ by statement 3 in Lemma 4). Moreover, $N(\mathcal{D}_0)$ must touch all four sides of $[0,1]^2$. Now suppose that $\mathscr{G}_\mathcal{C}[\mathcal{B}]$ has some large component $\mathcal{S}$. By definition $N(\mathcal{S}) \setminus \mathcal{S}$ contains only sparse cells. Then, by the first part of the claim, $N(\mathcal{S})$ must touch the four sides of $[0,1]^2$. Hence, we apply the second part of the claim to $\mathcal{S}$ and $\mathcal{D}_0$ to deduce that such large $\mathcal{S}$ cannot exist.

It just remains to prove the initial claim. Let $\mathcal{S}$ be a connected set of cells. Observe that $\bigcup N(\mathcal{S})$ is topologically connected (and in particular $N(\mathcal{S})$ is a connected set of cells), and that the outer boundary γ of $\bigcup N(\mathcal{S})$ is a simple closed polygonal path along the grid lines in $[0,1]^2$ defined by the tessellation. If we remove from γ the segments that coincide with some side of $[0,1]^2$, each connected polygonal path that remains is called a *piece* of γ. Note that $N(\mathcal{S}) \setminus \mathcal{S}$ need not be a connected set of cells. However all cells in $N(\mathcal{S})$ along the same piece of γ must be contained in the same topological component of $\bigcup(N(\mathcal{S}) \setminus \mathcal{S})$, and thus in the same connected component of $\mathscr{G}_\mathcal{C}[N(\mathcal{S}) \setminus \mathcal{S}]$.

The argument comprises several cases. For each case, a lower bound on the area of some connected component of $\mathscr{G}_\mathcal{C}[N(\mathcal{S}) \setminus \mathcal{S}]$ is given by finding some disjoint subsets of $[0,1]^2$ of large enough area contained in the union of cells in that component. Then, Lemma 4 ensures that $N(\mathcal{S}) \setminus \mathcal{S}$ contain at least one dense cell.

Given a cell c, let $B_\nearrow(c)$ be the set of points at distance at most $(1-4\delta')r$ from the top right corner of c and above and to the right of that corner. The sets $B_\nwarrow(c)$, $B_\searrow(c)$ and $B_\swarrow(c)$ are defined analogously replacing (top, above, right) by (top, above, left), (bottom, below, right) and (bottom, below, left) respectively. Note that $B_\nearrow(c)$, $B_\nwarrow(c)$, $B_\searrow(c)$ and $B_\swarrow(c)$ are disjoint and contained in $\bigcup(N(c) \setminus \{c\})$.

Case 1. Let $\mathcal{S} \subseteq \mathcal{C}$ be a connected set of cells which is not small and such that $N(\mathcal{S})$ does not touch any side of $[0,1]^2$. Since $\mathcal{S}$ is not small, assume without loss of generality that its vertical extent is greater than $16\delta' r$. Let c_1, c_2, c_3, c_4 be respectively the topmost, bottommost, leftmost and rightmost cells in $\mathcal{S}$ (possibly not all different and not unique). Let $A_\rightarrow$ be any rectangle of height $16\delta' r$ and width $(1-20\delta')r$ glued to the right of c_4 and between the top of c_1 and the bottom of c_2. Also choose a similar rectangle $A_\leftarrow$ of the same dimensions glued to the left of c_3, and let $A_\uparrow$ and $A_\downarrow$ be rectangles of height $(1-4\delta')r$ and width $\delta' r$ placed on top of, and below, the cells c_1 and c_2 respectively. By construction, $B_\nearrow(c_1)$, $B_\nwarrow(c_1)$, $B_\searrow(c_2)$, $B_\swarrow(c_2)$, $A_\uparrow$, $A_\downarrow$, $A_\leftarrow$ and $A_\rightarrow$ are disjoint and are contained in the same topological component of $\bigcup(N(\mathcal{S}) \setminus \mathcal{S})$ (i.e. the one that touches γ), which thus has area at least

$$\theta(1-4\delta')^2 r^2 + 2\delta' r(1-4\delta')r + 32\delta' r(1-20\delta')r \geq \theta r^2(1+\delta'/3).$$

Hence, by Lemma 4, $N(\mathcal{S}) \setminus \mathcal{S}$ must contain some dense cell.

Case 2. Let $\mathcal{S} \subseteq \mathcal{C}$ be a connected set of cells which is not small and such that $N(\mathcal{S})$ touches only one side of $[0,1]^2$ (assume it is the bottom side). This is very similar to Case 1, so we just sketch the main differences in the argument.

If the vertical extent of $\mathcal{S}$ is greater than $16\delta' r$, then proceed as in Case 1 but only consider the sets $B_{\nearrow}(c_1)$, $B_{\nwarrow}(c_1)$, $A_{\uparrow}$, $A_{\leftarrow}$ and $A_{\rightarrow}$. Otherwise, the horizontal extent of $\mathcal{S}$ must be greater than $16\delta' r$, and we consider instead the sets $B_{\nearrow}(c_4)$, $B_{\nwarrow}(c_3)$, $A'_{\uparrow}$, $A'_{\leftarrow}$ and $A'_{\rightarrow}$. Here, $A'_{\leftarrow}$ and $A'_{\rightarrow}$ are rectangles of height $\delta' r$ and width $(1-4\delta')r$ placed to the left and right of cells c_3 and c_4 respectively, and $A'_{\uparrow}$ is any rectangle of height $(1-20\delta')r$ and width $16\delta' r$ glued on top of c_1 and strictly between the left side of c_3 and the right side of c_4. In both cases, we deduce that the topological component of $\bigcup(N(\mathcal{S}) \setminus \mathcal{S})$ that touches the upper piece of γ has area at least $(1+\delta'/6)\theta r^2/2$. Since some cells in this component touch one side of $[0,1]^2$, Lemma 4 implies that $N(\mathcal{S}) \setminus \mathcal{S}$ must contain some dense cell.

Case 3. Let $\mathcal{S} \subseteq \mathcal{C}$ be a connected set of cells which is not small. Suppose first that $N(\mathcal{S})$ touches exactly two sides of $[0,1]^2$ which are adjacent (say the bottom and the left sides of $[0,1]^2$). If the horizontal extent of $\mathcal{S}$ is at most $4r$, then $N(\mathcal{S}) \setminus \mathcal{S}$ has some cell inside the $5r \times 5r$ square on the bottom left corner of $[0,1]^2$. But these cells are all dense by Lemma 4 and we are done. Hence we can assume that $\mathcal{S}$ has horizontal extent greater than $4r$. In the other cases that $N(\mathcal{S})$ touches two non-adjacent sides or three sides of $[0,1]^2$, we can assume without loss of generality that $N(\mathcal{S})$ touches the left and right sides of $[0,1]^2$ but not the top side. Therefore, in all the cases considered, $\mathcal{S}$ must contain some cells intersecting each of the five first vertical stripes of width r at the left side of $[0,1]^2$. Let c_1, c_2, c_3, c_4 and c_5 be the uppermost cells in $\mathcal{S}$ intersecting each of the five vertical stripes. These cells are not necessarily all different, but for each c of these, either $B_{\nwarrow}(c)$ or $B_{\nearrow}(c)$ is completely contained in the corresponding strip. Thus, the topological component of $\bigcup(N(\mathcal{S}) \setminus \mathcal{S})$ that touches the upper piece of γ has area at least $5(1-4\delta')^2\theta r^2/4 > (1+1/8)^2\theta r^2$, and by Lemma 4, $N(\mathcal{S}) \setminus \mathcal{S}$ must contain some dense cell.

Case 4. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be connected sets of cells not adjacent to each other (i.e. $\mathcal{S}_2 \cap N(\mathcal{S}_1) = \emptyset$) and such that both $N(\mathcal{S}_1)$ and $N(\mathcal{S}_2)$ touch all four sides of $[0,1]^2$. Note that by Lemma 4 all cells inside the $5r \times 5r$ square on the top left corner of $[0,1]^2$ are dense. Assume that none of these cells belongs to $N(\mathcal{S}_1) \setminus \mathcal{S}_1$ or $N(\mathcal{S}_2) \setminus \mathcal{S}_2$ (otherwise we are done). It could happen that these cells in the top left square are either all in $\mathcal{S}_1$ or all in $\mathcal{S}_2$. Assume they are not in $\mathcal{S}_1$. Then consider, as in Case 3, the uppermost cells c_1, c_2, c_3, c_4 and c_5 in $\mathcal{S}_1$ intersecting each of the five first vertical stripes of width r at the left side of $[0,1]^2$. The same argument shows that the topological component of $\bigcup(N(\mathcal{S}_1) \setminus \mathcal{S}_1)$ that touches the upper left piece of γ has area at least $(1+1/8)^2\theta r^2$, and Lemma 4 completes the proof. □

Finally, we need to show that bad components a.a.s. have some properties to be used in the construction of the Hamilton cycles. Given a component b of $\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$, let $J = J(b) \subseteq \{1, \ldots, n\}$ be the set of indices of points in $\boldsymbol{X}$ contained in some cell of b. Moreover, for any r', consider the set $J' = J'(b, r') = N_{\mathscr{G}(\boldsymbol{X}; r')}(J) \setminus J$ (i.e. the set of strict neighbours of J in a random geometric graph of radius r').

Lemma 6. *For a small enough constant $\delta > 0$, and for any r and r' satisfying $\theta n r^2 \sim \log n$ and $r \leq r' \leq (1+1/32)r$, the following is a.a.s. true. For each small component b of $\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$, there exists a connected set of dense cells $\mathcal{R}(b) \subseteq \mathcal{D}_0$ of size $0 < |\mathcal{R}(b)| \leq 10/\delta^2$ such that*

1. *for every $i \in J'(b, r')$, the cell containing X_i is adjacent to some cell in $\mathcal{R}(b)$, and*
2. *$\mathcal{R}(b) \cap \mathcal{R}(\widetilde{b}) = \emptyset$ and $J'(b, r') \cap J'(\widetilde{b}, r') = \emptyset$, for any other small component $\widetilde{b}$ of $\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$. different from b.*

Proof. Let b be a small component of $\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$, and let g be any 16×16 grid covering b. Let O denote the geometric centre of the grid g, and let $\mathcal{S}$ be the set of cells which have some point at distance between $3r/4$ and $3r/2$ from O. Take as $\mathcal{R}(b)$ the subset $\mathcal{R} = \mathcal{S} \cap \mathcal{D}$ formed by the dense cells in $\mathcal{S}$. This set will be shown to have all the desired properties. (Note that the size of $\mathcal{R}$ is $|\mathcal{R}| \le |\mathcal{S}| < 10/\delta^2$.)

Consider a coarser tessellation of $[0,1]^2$ into larger squares of side $\lfloor 1/(16\delta') \rfloor \delta' r$ (each square containing exactly $\lfloor 1/(16\delta') \rfloor^2$ cells). We refer to each square both as a subset of $[0,1]^2$, and as the set of cells it contains. Let $\mathcal{Q}$ be the set of squares of the coarser tessellation that contain at least one point at distance exactly $5r/4$ from O. By construction, all squares in $\mathcal{Q}$ are contained inside $\mathcal{S}$. Moreover, we claim that all squares in $\mathcal{Q}$ contain some dense cell. To show this, suppose first that $N(b)$ does not touch any side of $[0,1]^2$. In fact, by choosing δ sufficiently small, we can guarantee that each square $q \in \mathcal{Q}$ has no intersection with $N(b) \setminus b$, and thus $q \cup (N(b) \setminus b)$ is a connected set of cells of area at least

$$\theta(1 - 34\delta')^2 r^2 + \lfloor 1/(16\delta') \rfloor^2 \delta'^2 r^2 \ge (\theta + 1/257) r^2.$$

Hence, assuming that statement 1 in Lemma 4 holds, $q \cup (N(b) \setminus b)$ must contain some dense cell, which must be in q since $N(b) \setminus b$ does not contain any. The cases that $N(b)$ touches one or two sides of $[0,1]^2$ are dealt with analogously by using statements 2 and 3 in Lemma 4.

Since the union of squares in $\mathcal{Q}$ is topologically connected, and each pair of cells lying in topologically adjacent squares of $\mathcal{Q}$ are also adjacent in $\mathscr{G}_{\mathcal{C}}$, the dense cells in squares of $\mathcal{Q}$ induce a connected set of cells. Moreover, for any other cell c in $\mathcal{S}$ there is some square $q \in \mathcal{Q}$ such that c is adjacent to all cells in q. Hence, $N(\mathcal{R}) \supseteq \mathcal{S}$, and also $\mathcal{R}$ induces a connected set of cells. Since $\mathcal{R}$ cannot be embedded in a 16×16 grid of cells, $\mathcal{R}$ must be contained in $\mathcal{D}_0$.

Now consider any vertex $i \in J' = J'(b, r')$. If $d(X_i, O) \le 3r/8$, then the cell c containing X_i must be in $N(b) \setminus b$. Therefore, since b is a component of $\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$, c must be sparse but adjacent to some dense cell $d \in \mathcal{D}_0$. By construction, any point in d must be at distance between $(1 - 34\delta')r$ and $(11/8 + 2\delta')r$ from O, so $d \in \mathcal{R}$. Otherwise, suppose that $d(X_i, O) > 3r/8$. We also have $d(X_i, O) \le (1 + 1/32 + 16\delta')r$, since $i \in J'$. Then the cell c containing X_i must be adjacent to all cells in some square $q \in \mathcal{Q}$, and in particular to some dense cell in $\mathcal{R}$.

To verify the other requirements, define $\mathcal{Q}'$ to be the set of squares of the coarser tessellation with some point at distance exactly $7r/4$ from O. The same argument we used for $\mathcal{Q}$ shows that all squares in $\mathcal{Q}'$ contain some dense cell. Let $\mathcal{R}'$ be the set of dense cells in squares of $\mathcal{Q}'$. Then it is immediate to verify that any point in a cell c of some other small component $\widetilde{b} \neq b$ of $\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$ must be at distance at least $41r/16$ from O since otherwise c would be adjacent to some cell in b, $\mathcal{R}$ or $\mathcal{R}'$. All remaining statements follow easily from that. □

3 Building Hamilton cycles and a perfect matching

A *factorisation* of a graph is the set of subgraphs induced by a partition of the edge set. A *hamiltonian decomposition* of a graph is a factorisation in which at most one subgraph is a perfect matching, and all the remaining ones are Hamilton cycles. The construction of a hamiltonian decomposition in the following lemma is well known (since 1892). It is attributed to Walecki by Lucas [5]. We also need the decomposition to contain a *transversal*, by which we mean a matching that contains an edge of each of the Hamilton cycles in the decomposition. (Note that the transversal does not contain an edge of the perfect matching.)

Lemma 7. *Every complete graph has a hamiltonian decomposition with a transversal.*

Note that the number of Hamilton cycles in such a decomposition of K_{k+1} will be $\lfloor k/2 \rfloor$, and thus the transversal is a not-quite-perfect matching.

Proof. First, for k even, consider the complete graph K_{k+1} on the vertices $\{0, 1, \ldots, k-1, *\}$. We shall first colour the edges of K_{k+1}. Expressions referring to vertex labels other than $*$ are interpreted mod k and expressions referring to colour labels are mod $k/2$. (In this paper, mod denotes taking the remainder on division.)

For each pair of vertices u and v in $\{0, 1, \ldots, k-1\}$, assign the colour

$$\lceil (u+v)/2 \rceil \tag{3}$$

(mod $k/2$ of course) to the edge uv. Also, assign colour i to the edges from $*$ to both vertices i and $i + k/2$. It is easy to check that, for each $i \in \{0, 1, \ldots, k/2 - 1\}$, the edges receiving colour i form a $(k+1)$-cycle $(v_0, \ldots, v_k)$ where

$$v_0 = *, \qquad v_1 = i, \qquad v_{t+1} = v_t + (-1)^t t, \quad \forall t \in \{1, \ldots, k-1\},$$

or equivalently

$$v_0 = *, \qquad v_t = i - (-1)^t \lfloor t/2 \rfloor, \quad \forall t \in \{1, \ldots, k\}.$$

Thus, the colouring induces a factorisation of K_{k+1} into $k/2$ Hamilton cycles of colours $0, 1, \ldots, k/2 - 1$, giving the required hamiltonian decomposition. See Figure 3.

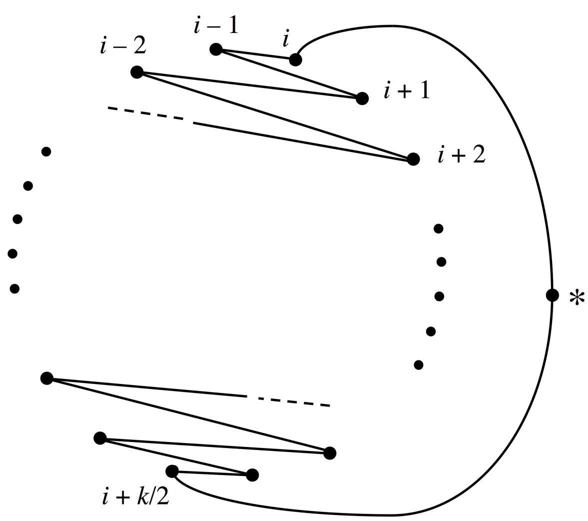

Figure 3: *One factor in a hamiltonian decomposition. Other factors are obtained by rotating the circle.*

When k is congruent to 2 mod 4, the set of edges $\{2i, 2i+1\}$ $(i = 0, \ldots, k/2 - 1)$ is a transversal. When k is divisible by 4, one transversal uses the edges $\{2i, 2i+1\}$ $(i = 0, \ldots, k/4 - 1)$, the edge from $*$ to $k/2$, and the edges $\{k/2 + 2i - 1, k/2 + 2i\}$ $(i = 1, \ldots, k/4 - 1)$.

For odd k, a perfect matching needs to be included. There is a similar colouring scheme, using the colours $0, 1, \ldots, (k-1)/2$, where colours are taken mod $(k+1)/2$. In this case, colour $\lceil (u + v \mod k)/2 \rceil$ is on the edge uv (note we assume by convention that $u + v \mod k \in \{0, \ldots, k-1\}$), each colour i $(i \in \{1, \ldots, (k-1)/2\})$ is on the edge from i to $*$, and each colour $(k+1)/2 - i$ $(i \in \{0, \ldots, (k-1)/2\})$ is on the edge from $k - i$ to $*$. The edges of colour 0 form a perfect matching, and each of the other colours gives a Hamilton cycle. Finally, the transversal for odd k is easily found, similar to the even k case, using more or less every second edge of the form $\{i, i+1\}$. □

In the remainder of this section, we use the results in the earlier lemmas to prove Theorem 1. We first give a complete proof for k even, and then provide the extra pieces of argument required for k odd.

(i) **Proof for k even.**

Let $\epsilon > 0$ be arbitrarily small. Recall the definitions of m and r_k in the beginning of Section 2. In view of Proposition 2, we can choose a large enough constant $\lambda > 0$ such that, by setting

$$r_l = \sqrt{\frac{\log n + m \log\log n - \lambda}{\theta n}}$$

we can guarantee that

$$\mathbf{Pr}\left(r_l \le r_k \le \sqrt{\frac{\log n + m \log\log n + \lambda}{\theta n}} \right) > 1 - \epsilon.$$

Hence, looking at the evolution of $\mathscr{G}(\boldsymbol{X}; r)$ for $0 \le r < \infty$, the probability that it becomes k-connected at some radius r satisfying the condition in Lemma 3 is greater than $1 - \epsilon$. Let us condition upon this event. Then, by the results in Section 2, we may assume that the properties described in Lemmas 5 and 6 hold for $r = r_l$ and some δ, and also that the property in Lemma 3 holds for $r = r_k$, $\eta = 32\delta$, $K = 2k + k2$ and all $2 \le j \le k$. So we may assume $\boldsymbol{X}$ to be an arbitrary fixed set of n points in $[0,1]^2$ in general position and satisfying these properties. The proof is completed by giving a deterministic construction of $k/2$ edge-disjoint Hamilton cycles for the geometric graph $\mathscr{G}(\boldsymbol{X}; r_k)$. Most edges will be of length at most r_l but we shall use a few of length between r_l and r_k. (The last edges creating k-connectivity arrive during this period, and they are of course necessary to construct $k/2$ edge-disjoint Hamilton cycles.) We define the edges of each Hamilton cycle by colouring some of the edges of $\mathscr{G}(\boldsymbol{X}; r_k)$, using colours $1, \dots, k/2$, such that each of these colour classes induces a Hamilton cycle.

We take $r = r_l$ (except at special points in the argument) and define $\mathscr{G}_{\mathcal{C}}$, $\mathcal{D}$, $\mathcal{B}$ and so on accordingly (see Section 2). Let T be a spanning tree of the largest component $\mathcal{D}_0$ of $\mathscr{G}_{\mathcal{C}}[\mathcal{D}]$. Next, double each edge of T to get an Eulerian multigraph F. The vertex degrees in T are bounded above by Δ, so those in F are bounded above by 2Δ. Next, pick an Eulerian circuit C of F.

Henceforth, we have no need to consider points in $[0,1]^2$ that are not members of $\boldsymbol{X}$. So, points in $\boldsymbol{X}$ contained in some cell c will simply be referred to as points in c, and they will often be identified with their corresponding vertices in $\mathscr{G}(\boldsymbol{X}; r_l)$ or $\mathscr{G}(\boldsymbol{X}; r_k)$. Also, the term *dense cell* will refer only to cells in $\mathcal{D}_0$, thus excluding these dense cells contained in bad components. For descriptive purposes, we split the rest of the argument into two parts, first treating the case that there are no bad cells, i.e. $\mathcal{B}$ is empty. For this we only need the edges of $\mathscr{G}(\boldsymbol{X}; r_l)$. Then we will show how the construction is easily modified to handle the bad components, using some edges of $\mathscr{G}(\boldsymbol{X}; r_k)$.

Part 1. $\mathcal{B}$ is empty.

In this case, the rest of the proof involves two steps, which will be used in different forms during the later arguments.

Step 1. Turning the circuits into cycles

The subgraph of $\mathscr{G}(\boldsymbol{X}; r_l)$ induced by the points contained in any dense cell is complete and has many more than k vertices. Lemma 7 provides $k/2$ edge-disjoint Hamilton cycles in this subgraph. In fact, it provides more; we just choose a subset of the Hamilton cycles that are given by that lemma. The separate cycles in all the dense cells will be 'broken' and rejoined together using C as a template. In the following discussion we assume C is oriented, so we may speak of incoming and outgoing edges of C with respect to a cell.

For any dense cell c, the deletion of c from C breaks C up into a number of paths P_i. For colour 1, do the following. Associate each path P_i with an edge z_i that joins two points in c and has already been coloured 1, using a different edge z_i for each path P_i. Uncolour the edges z_i, and associate the outgoing and incoming edges of the path P_i (with respect to the cell c) each with an endpoint of z_i. After doing the same for all dense cells, every edge cd of C, where c and d are cells, has now been associated with two points, one in c and one in d. Colour the edge joining these two points using colour 1. Doing this for all edges of C clearly joins up all the edges coloured 1 into one big cycle using all points in the dense cells.

Now do the same with colours $2, \ldots, k/2$, one after another, but each time being careful to use edges z_i in each cell that are not adjacent to such edges used with any of the previous colours. This is easily done because using an edge for one colour eliminates at most four edges of another colour (as the graph induced by edges of a given colour has maximum degree at most 2). So the process can be carried out if M is greater than $2k\Delta$. All the edges that are still coloured and were not used for joining into paths P_i we call *spare*.

Step 2. Extending the cycles into the sparse cells

There are now $k/2$ edge-disjoint coloured cycles, one of each colour, and each cycle uses precisely all the points in dense cells. Note that within each dense cell, there are still an arbitrarily large number (depending on M) of spare edges of each colour, left over from the original application of Lemma 7. To prepare for extending the cycles into the sparse cells, we will break the cycles at these spare edges.

Let c be any sparse cell. By the definition of $\mathcal{B}$ and our assumption that $\mathcal{B}$ has no cells, there is a dense cell, say c', adjacent to c in $\mathscr{G}_{\mathcal{C}}$. If c contains at most $2k$ points, for each vertex v of the geometric graph inside c do the following. Choose a spare edge z inside c' of colour 1, uncolour the spare edge z, and colour the two edges from the endpoints of z to v with the colour 1. Any edges of different colours adjacent to z should be deemed not spare after use. Then repeat for each of the other colours. After this, the edges of any given colour form a cycle containing all points in dense cells and in c.

On the other hand, if c contains more than $2k$ points, the above process could potentially require too many spare edges, so we must do something else. By Lemma 7, we can specify $k/2$ edge-disjoint Hamilton cycles around the points in c, one of each of the colours. One can then greedily choose an independent set of edges, one of each colour. (This is easily seen by noting that choosing an edge knocks out at most four adjacent edges with any particular colour. Alternatively, by a more careful argument which we give later, it can be shown that the same holds as long as c contains more than k points.) These edges can be matched up with $k/2$ spare edges that have both endpoints in c', and then each of the coloured cycles is easily extended by uncolouring each matched pair of edges and appropriately colouring the edges joining their endpoints. Again for this case, the edges of any given colour form a cycle containing all the points in dense cells and in c.

This process can be repeated for each sparse cell. Since each dense cell has at most Δ neighbours in $\mathscr{G}_{\mathcal{C}}$, the total number of spare edges required of any one colour in any dense cell can be crudely bounded above by $2k\Delta$, which is the same as the upper bound on the number of points already used up. Thus, for M sufficiently large ($4k\Delta$ should do), there will be a sufficient number of spare edges to finish with a cycle of each colour through all points in $\mathscr{G}(\boldsymbol{X}; r_l)$, using only the edges of $\mathscr{G}(\boldsymbol{X}; r_l)$. This finishes Step 2 and the proof in the case that $\mathcal{B}$ is empty.

We now turn our attention to the (much more difficult) case that $\mathcal{B}$ is nonempty, for which we use an appropriate modification of the above argument.

Part 2. The general case: $\mathcal{B}$ can be nonempty.

For this, we will need to use some edges of $\mathscr{G}(\boldsymbol{X}; r_k)$ that are not present in $\mathscr{G}(\boldsymbol{X}; r_l)$, but the definition of all structures (such as bad components) remains as determined by the graph $\mathscr{G}(\boldsymbol{X}; r_l)$. Recall the Eulerian circuit C chosen at the start of the proof. This circuit gives a directed tour of all dense cells in the graph $\mathscr{G}_{\mathcal{C}}$. (Recall that in this section the term *dense* is reserved for cells in $\mathcal{D}_0$.) We will first extend this tour to a circuit that includes routes through each bad component, and later perform modified versions of Steps 1 and 2 described above.

Pick one such bad component b, which must be small by Lemma 5, and let $\mathcal{R} = \mathcal{R}(b)$ be a set of cells as in Lemma 6. Recall that $0 < |\mathcal{R}| < 10/\delta^2$. To take care of b, we will work entirely in $\mathcal{R}$ and the bad component b. Let J denote the set of points in cells in b and set $j = |J|$ (assume that $j > 0$, since otherwise b has no role in our argument). The subgraph of $\mathscr{G}(\boldsymbol{X}; r_l)$ induced by J is a copy of K_j, since b is small and can be embedded in a 16×16 grid of cells (and assuming that $32\delta < 1$). Now consider the graph $\mathscr{G}(\boldsymbol{X}; r_k)$, which by definition is k-connected, and let $J' = N_{\mathscr{G}(\boldsymbol{X}; r_k)}(J) \setminus J$. Let H denote the induced bipartite subgraph of $\mathscr{G}(\boldsymbol{X}; r_k)$ with parts J and J', and let $G \subseteq \mathscr{G}(\boldsymbol{X}; r_k)$ be the union of H with the clique on vertex set J. Note for later reference that the set J' can possibly contain vertices in dense cells: although no cell in b is adjacent to a dense cell, points in it can be adjacent to points in a dense cell.

A *linear forest* is a forest all of whose components are paths.

Claim 1. G contains $k/2$ pairwise edge-disjoint linear forests $F_1, \ldots, F_{k/2}$, such that
(a) in each forest all vertices in J have degree 2, and
(b) the set of vertices in J' contained in some path of $F_1, \ldots, F_{k/2}$ has cardinality at most $2k^2$.

To prove the claim, we consider two cases.

Case 1: $j > k$.

Since $\mathscr{G}(\boldsymbol{X}; r_k)$ is k-connected, no vertex cut of G of size less than k can separate J from J'. Moreover, both J and J' have cardinality at least k. So (a version of) Menger's theorem implies that there is a set of k pairwise disjoint paths joining J to J'. Hence, there is a matching, T, of cardinality k, with each edge of the matching joining a point in J to a point in J'.

Consider first an arbitrary complete graph K_j, of which Lemma 7 can be used to obtain a full hamiltonian decomposition, together with a transversal containing one edge from each of the Hamilton cycles. Now choose $k/2$ of the Hamilton cycles in the decomposition, and let T' be the matching consisting of the edges of the transversal that lie in the chosen cycles.

Next, we can identify the set of vertices of K_j with the set J, such that the vertices incident with edges in T' are identified with the vertices of J that are matched by T. From each of the Hamilton cycles, delete the edge, say x, in that cycle that lies in T', and add the two edges of T adjacent to x. This gives a path P in G which starts and finishes in vertices in J'. Since T' is a matching, the end vertices of all paths comprise a set of k distinct vertices. Hence, these $k/2$ paths suffice for $F_1, \ldots, F_{k/2}$.

Case 2: $j \leq k$.

Set $K = 2k + k^2$ and define J_0 to be the set of points in J that have less than K neighbours in $\mathscr{G}(\boldsymbol{X}; r_k)$. Also, put $J_1 = J \setminus J_0$ and let $J_0' = N_{\mathscr{G}(\boldsymbol{X};r_k)}(J_0) \setminus J_0$ (note that J_0' is contained in the (disjoint) union of J_1 and J').

We first prove an analogue of the claim but with J_0 and J_0' playing the role of J and J'. Let H_0 denote the induced bipartite subgraph of $\mathscr{G}(\boldsymbol{X}; r_k)$ with parts J_0 and J_0', and let $G_0 \subseteq \mathscr{G}(\boldsymbol{X}; r_k)$ be the union of H_0 with the clique on vertex set J_0. Put $j_0 = |J_0|$. The degenerate case $j_0 = 0$ will be treated later, so we assume that $j_0 \geq 1$. Observe that we can find a set J_0'' of k points in J_0' which are common neighbours of all points in J_0 with respect to $\mathscr{G}(\boldsymbol{X}; r_k)$. This follows from Lemma 3 for $j_0 \geq 2$, and is trivially true for $j_0 = 1$ since $\mathscr{G}(\boldsymbol{X}; r_k)$ is k-connected.

Consider first an arbitrary complete graph K_{k+j_0}, of which Lemma 7 can be used to obtain a full hamiltonian decomposition, together with a transversal containing one edge from each of the Hamilton cycles. Now choose $k/2$ of the Hamilton cycles in the decomposition, and let T be the matching consisting of the edges of the transversal that lie in the chosen cycles.

Next, we can identify the set of vertices of K_{k+j_0} with the set $J_0 \cup J_0''$, so that the vertices incident with edges in T are identified with the vertices of J_0''. Each of the original Hamilton cycles in the decomposition turns into a linear forest when restricted to the edges in G_0, since at least the edge in the matching T is missing in G_0 ($J_0'' \subseteq J_0'$ has no internal edges in G_0). By construction, these $k/2$ forests have the following properties: in each forest all vertices in J_0 have degree 2; and the set of vertices in J_0' contained in some path of some forest has cardinality at most k. In addition, this statement is also trivially satisfied in the case $j_0 = 0$ by considering $k/2$ empty forests (also define $J_0'' = \emptyset$ in this case).

We complete each of these $k/2$ forests by adding all these vertices of J_1 that were not used by any path in that forest (these new vertices are interpreted as paths of length 0 in the forest, and they might belong to more than one forest). Observe that after doing that the forests stay edge-disjoint, and the set of vertices in $J_1 \cup J_0'$ contained in some path has cardinality at most $2k$.

Some of the paths in the forests have one or two end vertices in J_1. We extend these paths (edge-disjointly with paths in any of the forests) as follows. If $v \in J_1$ is an end vertex of some path, we extend the path in that direction by adding one of the neighbours of v in $\mathscr{G}(\boldsymbol{X}; r_k)$ which is not already in any path of any forest (paths of length 0 in J_1 are naturally extended in two directions). This can be done since each vertex v in J_1 has at least $K \geq |J| + |J_0''| + 2(k/2)|J_1|$ neighbours to choose. Observe that the newly added end vertices must lie in $N(J_1) \setminus J \subseteq J'$ since all vertices in J were already used in some path. The resulting forests $F_1, \ldots, F_{k/2}$ satisfy all the requirements, since all vertices in J have now degree 2 in each forest, and the set of vertices in J' contained in some path is at most $|J_0''| + (k/2)2|J_1| \leq 2k^2$. This completes the proof of the claim (see also Figure 4 for a visual description of this second case).

Let J'' be the set of vertices in J' contained in some path of $F_1, \ldots, F_{k/2}$. By the claim above, we have $|J''| \leq 2k^2$. We extend each forest F_i to a spanning forest F_i' of $J \cup J''$ by adding those vertices in J'' not used by any path of F_i (these vertices are counted as separate paths of length 0). The total number of paths of $F_1', \ldots, F_{k/2}'$ is at most $|J''|k/2 \leq k^3$, taking into account multiplicity since each 0-length path may belong to more than one forest F_i'. Moreover, by setting $r = r_l$ and $r' = r_k$ in Lemma 6, we deduce that each vertex in J'' is contained in a cell c that is either dense or adjacent to some dense cell in $\mathcal{R}$. (Note that c may be either sparse or dense.)

The next step is to associate each of the paths in F_i' with the colour i, and create circuits

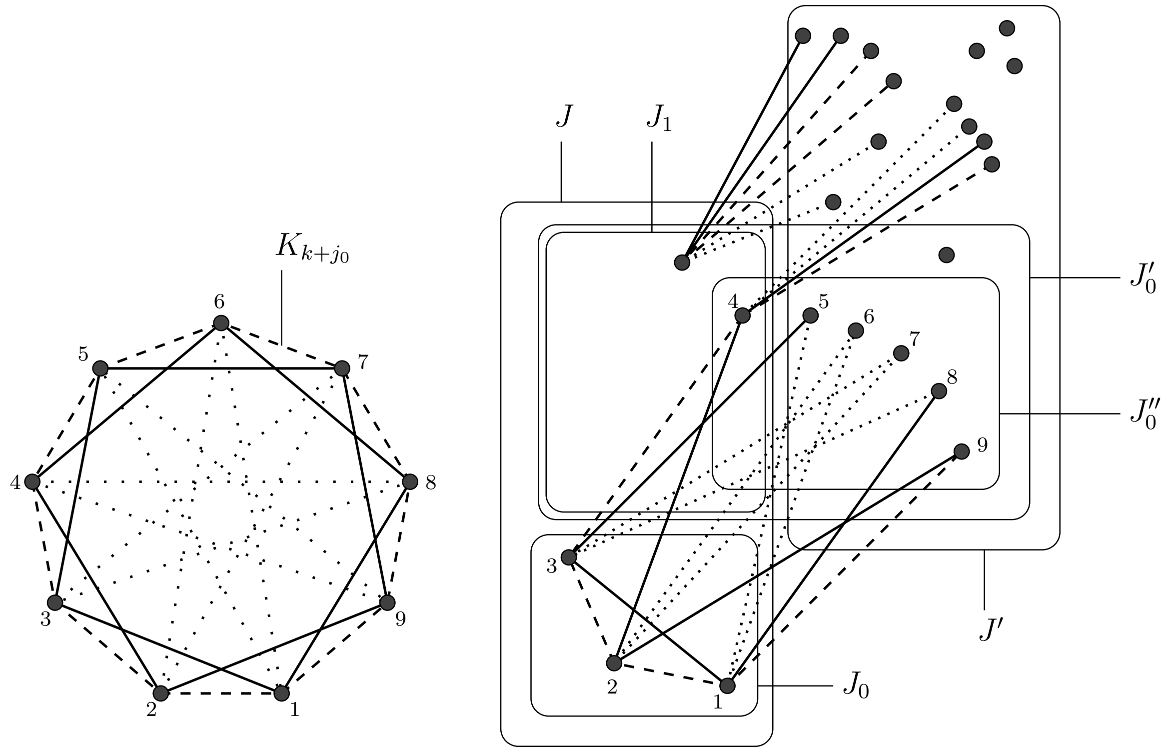

Figure 4: Illustration of Case 2 in the proof of Claim 1 for $k = 6$, $j = 5$ and $j_0 = 3$, representing the construction of the linear forests. For the sake of clarity, only a subset of the vertices in J' is shown. (Note that each vertex in J_1 should have at least $2k + k^2$ neighbours in $J \cup J'$.)

$C_1, \ldots C_{k/2}$ such that circuit C_i contains C, together with an extra cycle $C(P)$ for each path P in the forest F_i'. To construct $C(P)$, take two cells c and d in $\mathcal{R}$ (possibly $c = d$), one for each end-vertex of P, either containing the end-vertex or adjacent to a cell containing it. Then the cycle $C(P)$ consists of a special new edge (possibly a loop) joining cells c and d (we say this edge *represents* P), together with a path of cells within $\mathcal{R}$ joining those same cells c and d. Note that each cycle has length bounded above by $10/\delta^2$ (see Lemma 6).

This construction was all with respect to a particular bad component b. Now repeat the construction for all the other bad components, in each case extending the circuits $C_1, \ldots C_{k/2}$ in the same way as for b. By Lemma 6, two paths P and P' related to different bad components have no vertex in common, and also the corresponding extra cycles $C(P)$ and $C(P')$ use disjoint sets of dense cells. Hence, the number of these new cycles passing through any particular dense cell is at most k^3, so the maximum degree of dense cells in each C_i is at most $2\Delta + 2k^3$.

We next perform a version of Step 1 described in Part 1. First, let us call all the vertices lying in forests F_i with respect to bad components the *forest* vertices.

Step 1'. Turning the circuits C_i into edge-disjoint coloured cycles

The first part of this is done just as for Step 1 when $\mathcal{B}$ empty, except for two aspects. To start with, all forest vertices within dense cells are set aside and not used in the construction of the coloured cycles within the dense cells. Secondly, where an edge of the circuit between cells c and c' is one of those representing a path P in a forest, instead of a simple edge between vertices u in c and v in c', the cycle uses the path P represented by that edge, together with the edges joining P to u and v. In this way, we obtain from C_i a cycle of colour i that visits precisely all vertices that are either forest vertices (i.e. in $J(b)$ or $J''(b)$ for any small component b), or in dense cells, but *not* both. All other points in $\boldsymbol{X}$ will be called *outsiders.* They are not yet visited by any of $C_1, \ldots, C_{k/2}$, either because they are neither forest vertices nor in a dense cell, or are in both.

We next perform a version of Step 2, as follows.

Step 2'. Extending the cycles to the outsiders

This consists of extending each coloured cycle as done in Step 2, but this time extending to them only through the outsiders. The other significant difference between this and Step 2 is that the maximum degree of dense cells in each C_i is now bounded above by $2\Delta + 2k^3$ rather than 2Δ, and there are some outsiders in each dense cell, so there are fewer spare edges to work with, but the change is only a constant. So we need to adjust the lower bound on M accordingly. This completes the proof of part (i) of the theorem.

(ii) **Proof for k odd.**

We now consider odd $k \geq 1$. Recall that the total number of vertices in the geometric graph must be even. The same framework of argument is used as for k even. For k odd, colours $1, \ldots (k-1)/2$ will be used for the Hamilton cycles, and colour $(k+1)/2$ will be used for the matching. We find it convenient to refer to $(k+1)/2$ as the *match* colour. The edges that we colour using the match colour will form a spanning subgraph D of the geometric graph, whose edges form a cycle on some (possibly all) vertices and a matching of some of the other vertices; in the very last stage of the argument we will adjust this to form a perfect matching of the whole graph.

Define the multigraph F as for k even, and construct C in the same way. We next need to perform a version of Step 1. In this case, instead of creating $k/2$ edge-disjoint cycles passing through all the points in dense cells, we construct $(k+1)/2$ of them using the same construction as for k even.

In the case that there are no bad cells, the argument as in Part 1 above shows that all the cycles can be extended as in Step 2 to edge-disjoint Hamilton cycles of the graph. Then, since the graph has an even number of vertices, every second edge of the matching colour can be omitted to provide the desired colouring.

So consider the case that $\mathcal{B}$ is possibly nonempty and follow the argument in Part 2 for k even, up to the point where Claim 1 is made. This claim is replaced by the following.

Claim 2. G contains $(k+1)/2$ pairwise edge-disjoint linear forests $F_1, \ldots, F_{(k+1)/2}$, such that
(a) in each of the first $(k-1)/2$ forests all vertices in J have degree 2,
(b) the set of vertices in J' contained in some path of $F_1, \ldots, F_{(k+1)/2}$ has cardinality at most $2k^2$, and
(c) the last forest, $F_{(k+1)/2}$, is a matching that saturates each vertex in J.

To prove this claim, we again distinguish two main cases but insert an extra one.

Case 1a: $j > k + 1$.

As in the case $j > k$ for the first claim, we may find the matching T of (odd) cardinality k, and also $(k+1)/2$ edge-disjoint Hamilton cycles in $G[J]$, with a matching T' containing one edge from each Hamilton cycle. By relabelling J, we can then align the end vertices of the T' with those of T in J, but only using one end vertex of the edge in the last cycle. A set of edges in this last cycle is easily deleted so that what remains, together with the incident edge of T, is a matching either entirely contained in J or with one edge (of T) leaving J. This proves the claim in this case.

Case 1b: $j = k + 1$.

In this case, we may use the decomposition of a $K_{j+1} = K_{k+2}$ into $(k + 1)/2$ Hamilton cycles, delete any one vertex, and use the broken Hamilton cycles for the paths, plus every second edge of any one of them for the matching.

Case 2: $j \leq k$.

This is analogous to the proof of Case 2 of Claim 1, so we just sketch the main changes. Here for $j_0 \geq 1$ we use the full hamiltonian decomposition of K_{k+j_0} (see Lemma 7) to select $(k-1)/2$ edge-disjoint Hamilton cycles and one matching of all but at most one vertex (which we call the *unmatched* vertex) of K_{k+j_0}. We also obtain a matching T consisting precisely of one edge of each of the $(k - 1)/2$ cycles. Then we identify the set of vertices of K_{k+j_0} with the set $J_0 \cup J_0''$ as before, but we must additionally guarantee that the unmatched vertex (if it exists) is identified with some vertex of J_0''. Reasoning as in the proof of Claim 1, we obtain $(k + 1)/2$ edge-disjoint forests in G_0, where the first $(k - 1)/2$ of the forests have analogous properties to those in Claim 1, and the last forest is a matching that saturates each vertex in J_0. To extend the forests to all the vertices of J, we deal with the first $(k-1)/2$ forests exactly as before. The matching is extended by connecting each vertex v in J_1 that is not in the original matching to one of the neighbours of v which is not already in any path of any forest. As before, this can be done since each vertex v in J_1 has at least $K \geq |J|+|J_0''|+(2(k-1)/2+1)|J_1|$ neighbours to choose. This finishes the proof of Claim 2.

We now extend the forests to spanning forests F_i', exactly as for the even k case (this now includes the match colour). We are now prepared for the analogue of Step 1.

Step 1". Circuits into cycles and matching

For each non-match colour $1 \leq i \leq (k-1)/2$, the forests F_i' in the various bad components are treated the same way as in Step 1' in order to create circuits C_i. These are then turned into edge-disjoint cycles $\widetilde{C}_i$ passing through all vertices except for the outsiders, just as for k even. Simultaneously with this, by including an extra colour in the construction, we create a cycle in the match colour that passes through just the vertices in the dense cells apart from outsiders. For this colour, we must do something different with respect to the vertices not in dense cells. So let $i = (k + 1)/2$; so far we have a cycle $\widetilde{C}_i$, through all non-forest points in dense cells, whose edges are of colour i. The edges of a given forest F_i related to a bad component are naturally coloured with i. For each single vertex component v in the forest F_i' (i.e. each vertex v unmatched by F_i), let d be a dense cell either equal or adjacent to the cell c containing v. We may select the end-vertices v_1 and v_2 in d of a spare edge of $\widetilde{C}_i$ (note this implies that the edges vv_1 and vv_2 are currently uncoloured). Denoting these two vertices v_1 and v_2, the pair $\{v_1, v_2\}$ is called the *gate* for v. Each such vertex v is treated in this way, its gate is defined and v is added to a set W (a set of vertices which are 'waiting'). Naturally, any spare edge incident with either vertex in a gate is deemed non-spare for all subsequent choices of gates. Note that at this stage, all vertices in the graph are incident with an edge of the match colour except for the outsiders and those in W.

Next, we perform the step of extending the cycles to the outsiders.

Step 2". Extending cycles and matching to outsiders

For each non-match colour i, the cycle of colour i is extended as in Step 2'. We next show that we can also include a near-perfect matching of the match colour, which saturates all but a bounded number of outsiders, such that the matching is edge-disjoint from all the coloured cycles. This extra matching is easily obtained using the methodology of Step 2': for the match colour i, if a sparse cell c contains at most $2k$ outsiders, we may leave all outsiders unmatched.

If it contains more outsiders, simply include an extra cycle $\widehat{C}$ through all outsiders in c. This cycle should be chosen simultaneously with all the other cycles being chosen within cell c in this Step 2", using Lemma 7. Then, colour every second edge of $\widehat{C}$ with the match colour, leaving at most one outsider in this cell unmatched. Finally, in either case, for each remaining unmatched outsider v, choose a gate (v_1, v_2) exactly as described in Step 1". Note that a dense cell contains a bounded number of outsiders since these are all in the forest F_i.

After all this, every vertex is in each of the coloured cycles of colour $i \leq (k-1)/2$, but we still need to create the perfect matching of colour $i = (k+1)/2$. So far, all vertices are either matched by the match colour i, or lie in W, or lie on the cycle $\widetilde{C}_i$. To fix this, in one fell swoop, we choose simultaneously for all vertices v in W, a vertex v', in the gate for v, such that all the vertices v' have odd distance apart as measured along the cycle $\widetilde{C}_i$ of colour i. (Why this is possible will be explained shortly.) Then all such edges of the form vv' are coloured i, and finally, every second edge along C_i between these vertices is coloured i in such a way as to create a matching. The edges of colour i clearly form a perfect matching of the graph.

The only thing left to explain is why the choice of all v' as specified, creating odd distances, is feasible. Since there are two adjacent vertices on the cycle in each gate that can potentially be used as v', we may pass along the cycle $\widetilde{C}_i$ making sure that the distances between chosen vertices are odd, until returning to the starting point. The very last distance must be odd because the number of gates equals the number of vertices in W. These are precisely the vertices outside the cycle that are not already matched by colour i. Since the number of vertices in the graph is even, the parity is correct for every distance to be odd. □

4 General dimension

We can be extend our main result to general dimension, in the following sense. Modify the definition of $\boldsymbol{X} = (X_1, \ldots, X_n)$ by assuming that the X_i are chosen independently and u.a.r. from $[0,1]^d$, for some fixed integer $d \geq 2$. Redefine $\mathscr{G}(\boldsymbol{X}; r)$ analogously, using some fixed ℓ_p norm of $[0,1]^d$. Then Theorem 1 still holds for the random graph process $\big(\mathscr{G}(\boldsymbol{X}; r)\big)_{0 \leq r < \infty}$.

In fact, most of the argument in the paper is independent of d, so we shall simply sketch the main differences of those parts that need to be changed.

First of all, the definition of constant m in Section 2 is extended to

$$m = \begin{cases} 2^{d-2}(k+2-d-2/d) & \text{if } 1 \leq k < d, \\ 2^{d-1}(k+1-d-1/d) & \text{if } k \geq d \end{cases}$$

and let θ denote the volume of the unit d-dimensional ball with respect to the ℓ_p norm. Then, Proposition 2 remains valid if we change (1) to

$$\theta n {r_k}^d - \frac{2^{d-1}}{d} \log n - m \log\log n,$$

since it still follows from Theorem 8.4 in [7]. In view of that, we replace the condition on r by $\theta n r^d = \frac{2^{d-1}}{d} \log n + m \log\log n + O(1)$ in the statement of Lemma 3 and by $\theta n r^d \sim \frac{2^{d-1}}{d} \log n$ in Lemmas 4, 5 and 6.

In order to extend the proof of Lemma 3 to general dimension d, we classify bad configurations into types according to their position with respect to the boundary of $[0,1]^d$. More precisely, for each $i \in \{0, \ldots, d\}$, let T_i denote the number of bad configurations such that X_{v_1}

is at distance at most r from exactly $d-i$ *facets* (i.e. $(d-1)$-dimensional faces) of $[0,1]^d$. Setting τ_i to be the distance between X_{v_1} and the corresponding facet, we obtain by an analogous argument that, for some constant $c>0$,

$$\begin{aligned}
\mathbf{E}T_i &= O(1)\sum_{j_1,j_2,j_3,j_4} n^{j_1+j_2+j_3+j_4}\int_0^r\cdots\int_0^r\int_0^{\sqrt{d}\eta r}\rho^{d(j_1-2)+d-1}r^{dj_2}(\rho r^{d-1})^{j_3+j_4}\\
&\quad\times\exp\left[-\frac{\theta n r^d}{2^{d-i}}-cr^{d-1}n(\rho+\tau_1+\cdots+\tau_{d-i})\right]d\rho\,d\tau_1\cdots d\tau_{d-i}\\
&= O(1)\sum_{j_1,j_2,j_3,j_4} n^{j_1+j_2+j_3+j_4}r^{dj_2+(d-1)(j_3+j_4)}\\
&\quad\times\int_0^r\cdots\int_0^r\int_0^{\sqrt{d}\eta r}\rho^{dj_1+j_3+j_4-d-1}\frac{e^{-cr^{d-1}n(\rho+\tau_1+\cdots+\tau_{d-i})}}{(n^{2^{d-1}/d}\log^m n)^{1/2^{d-i}}}d\rho\,d\tau_1\cdots d\tau_{d-i}\\
&= O(1)\sum_{j_1,j_2,j_3,j_4}\frac{n^{j_1+j_2+j_3+j_4}r^{dj_2+(d-1)(j_3+j_4)}}{(n^{2^{d-1}/d}\log^m n)^{1/2^{d-i}}(r^{d-1}n)^{dj_1+j_3+j_4-i}}\\
&\quad\times\int_0^{nr^d}\cdots\int_0^{nr^d}\int_0^{\sqrt{d}\eta nr^d}x^{dj_1+j_3+j_4-d-1}e^{-c(x+y_1+\cdots+y_{d-i})}dx\,dy_1\cdots dy_{d-i}\\
&= O(1)\sum_{j_1,j_2,j_3,j_4}\frac{(nr^d)^{j_1+j_2-dj_1+i}r^{-i}}{n^{2^{i-1}/d}\log^{m/2^{d-i}} n}\\
&\quad\times\int_0^{nr^d}\cdots\int_0^{nr^d}\int_0^{\sqrt{d}\eta nr^d}x^{dj_1+j_3+j_4-d-1}e^{-c(x+y_1+\cdots+y_{d-i})}dx\,dy_1\cdots dy_{d-i}\\
&= O(1)\sum_{j_1,j_2,j_3,j_4}\frac{(\log n)^{j_1+j_2-dj_1+i-i/d-m/2^{d-i}}}{n^{(2^{i-1}-i)/d}},
\end{aligned}$$

where we used the change of variables $\{x=r^{d-1}n\rho,\ y_1=r^{d-1}n\tau_1,\ \ldots,\ y_1=r^{d-1}n\tau_1\}$ and the fact that $\theta nr^d=\frac{2^{d-1}}{d}\log n+m\log\log n+O(1)$. This last expression is trivially $o(1)$ for all $i\notin\{1,2\}$, since then we have $(2^{i-1}-i)/d>0$. To check the cases $i=1$ and $i=2$, observe that an equivalent definition of m is

$$m=\max\{2^{d-2}(k+2-d-2/d),2^{d-1}(k+1-d-1/d)\}.$$

Hence (using also that $j_1+j_2\le k+1$ and $j_1\ge 2$),

$$\begin{aligned}
\mathbf{E}T_1 &= O\left((\log n)^{k+1-2d+1-1/d-m/2^{d-1}}\right)=O(\log^{1-d}n)=o(1)\qquad\text{and}\\
\mathbf{E}T_2 &= O\left((\log n)^{k+1-2d+2-2/d-m/2^{d-2}}\right)=O(\log^{1-d}n)=o(1).
\end{aligned}$$

The cells defined in Section 2 become d-dimensional hypercubes of side $\delta' r$, and all the remaining definitions in that section are extended analogously. The a.a.s. events in Lemma 4 simply turn into: for each $i\in\{0,\ldots,d\}$, all connected sets of cells of area at least $(1+\alpha)i\theta r^d/2^{d-1}$ touching exactly $d-i$ facets of $[0,1]^d$ contain some dense cell. The proof is completely analogous, setting $s=\lceil(1+\alpha)i\theta/(2^{d-1}\delta'^d)\rceil=\Theta(1)$, changing (2) into

$$\begin{aligned}
\sum_{t=0}^{(M-1)s}\binom{n}{t}(s\delta'^d r^d)^t(1-s\delta'^d r^d)^{n-t} &= O\left(e^{-(1+\alpha)i\theta r^d n/2^{d-1}}\right)\sum_{t=0}^{(M-1)s}(r^d n)^t\\
&= O\left(n^{-(1+\alpha)i/d+o(1)}\log^{(M-1)s}n\right),
\end{aligned}$$

and observing that there are $\Theta(1/r^i) = \Theta((n/\log n)^{i/d})$ connected sets of s cells touching $d-i$ facets of $[0,1]^d$.

Redefine *small* sets of cells to be those ones that can be embedded in a $(4d^2) \times \overset{d}{\overbrace{\cdots}} \times (4d^2)$ grid of cells (i.e. a set of cells of ℓ_∞-diameter at most $4d^2\delta' r$). Call those sets of cells that are not small *large*. In these terms, we proceed to extend Lemma 5 to general dimension $d \geq 2$. From results in either [1] or [4], we deduce that a.a.s. $\mathscr{G}_{\mathcal{C}}[\mathcal{D}]$ has one very large component $\mathcal{D}_0$ and all (bad) components of $\mathscr{G}_{\mathcal{C}}[\mathcal{B}]$ have geometric diameter at most $r/10$ (see more specifically Lemma 5, Corollary 10 and Section 3 in [1], and also Proposition 5 and Section 3 in [4]). So in particular a.a.s. there is no bad component b such that $N(b)$ touches any pair of opposite facets of $[0,1]^d$. In order to show that a.a.s. all bad components are small it is sufficient to prove the following claim:

- A.a.s. for any large connected set of cells $\mathcal{S}$ such that $N(\mathcal{S})$ does not touch any two opposite facets of $[0,1]^d$, $N(\mathcal{S}) \setminus \mathcal{S}$ must contain some dense cell.

The proof of this claim is very similar to the one of the claim in the beginning of the proof of Lemma 5, and simply consists in bounding from below the area of some connected component of $N(\mathcal{S})\setminus\mathcal{S}$ by describing some disjoint suitable subsets contained in one topological component of $\cup(N(\mathcal{S}) \setminus \mathcal{S})$. We thus sketch only the main ideas, and describe the subsets of $\cup(N(\mathcal{S}) \setminus \mathcal{S})$ with the required properties.

For any $i \in \{0, \ldots, d\}$, assume that $N(\mathcal{S})$ touches exactly $d-i$ facets of $[0,1]^d$, namely $F_1, \ldots, F_{d-i}$, where w.l.o.g. $F_t = [0,1]^{t-1} \times \{0\} \times [0,1]^{d-t}$. For $i = 0$, the claim above is immediate from the extended version of Lemma 4, so we focus on the case $i > 0$.

First we need some geometric definitions that extend some of the objects we already used for $d = 2$. We call *d-ball sector* of centre $O \in \mathbb{R}^d$ to the intersection of the d-ball of centre O and radius $(1-2d\delta')r$ with one of the 2^d orthants that arise after translating the origin of coordinates to O. A d-ball sector has volume $\theta(1-2d\delta')^d r^d/2^d$. Given a cell in $[0,1]^d$, we associate to each of its 2^d corners the d-ball sector centred on the corner and in the orthant opposite to the cell. To simplify notation, let us denote the $(d-1)$th and dth coordinates in $[0,1]^d$ (or $\mathbb{R}^d$) as 'horizontal' and 'vertical' respectively, so that the usual two dimensional language can be applied when referring to these coordinates. A *vertical cylinder sector* in $\mathbb{R}^d$ is the cartesian product of a $(d-1)$-ball sector times $[a,b]$, and we say that this cylinder sector has height $b-a \geq 0$. Similarly, we can obtain a *horizontal cylinder sector* of length $b-a$ by permuting the last two coordinates. A vertical (horizontal) cylinder sector of height (length) $b-a$ has volume at least $(b-a)\theta(1-2d\delta')^{d-1}r^{d-1}/2^d$.

Now let c_1 and c_2 be respectively a topmost and a bottommost cell in $\mathcal{S}$ (possibly equal or not unique). We can find 2^{i-1} disjoint d-ball sectors above c_1 (associated to the top corners of c_1 which point towards the interior of $[0,1]^d$) which are contained in the same topological component of $\cup(N(\mathcal{S}) \setminus \mathcal{S})$. Similarly, we can find 2^{i-1} disjoint d-ball sectors below c_2 in the same component. Hence, some topological component of $\cup(N(\mathcal{S}) \setminus \mathcal{S})$ has area at least

$$2^{i-d}\theta(1-2d\delta')^d r^d.$$

Therefore, if $3 \leq i \leq d$ then $2^{i-d} > i/2^{d-1}$ and the claim follows by Lemma 4. Notice that so far we did not use the fact that $\mathcal{S}$ is large.

For the cases $i = 1, 2$, we need to achieve a better bound by finding some additional and disjoint subsets of $\cup(N(\mathcal{S}) \setminus \mathcal{S})$ with the required properties. Since $\mathcal{S}$ is large, we can assume w.l.o.g. that the vertical length of $\mathcal{S}$ is at least $4d^2\delta' r$. Consider first the case $i = 2$. Let c_3 and c_4 be respectively a leftmost and a rightmost cell in $\mathcal{S}$ (possibly equal or not unique). In addition to the previously described d-ball sectors, we consider a vertical cylinder sector of height $4d^2\delta' r$ to the right of c_4 and between the top and the bottom d-ball sectors. Similarly, consider a vertical cylinder sector of height $4d^2\delta' r$ to the left of c_3, a horizontal cylinder sector of length $\delta' r$ above c_1 and a horizontal cylinder sector of length $\delta' r$ below c_2. The total area is in this case at least

$$4\theta(1-2d\delta')^d r^d/2^d + 2(4d^2+1)\delta' r\theta(1-2d\delta')^{d-1}r^{d-1}/2^d = \\ 2\theta r^d/2^{d-1}\left[(1-2d\delta')^d + (4d^2+1)\delta'(1-2d\delta')^{d-1}/2\right],$$

and the claim follows by Lemma 4 since $(1-2d\delta')^d + (4d^2+1)\delta'(1-2d\delta')^{d-1}/2 > 1$ for δ small enough.

Finally, suppose $i = 1$. This case is similar to the previous one. We consider the initial 2^i d-ball sectors as before plus an additional vertical cylinder sector of height $4d^2\delta' r$ to the right of c_4, a horizontal cylinder sector of length $\delta' r$ above c_1 and a horizontal cylinder sector of length $\delta' r$ below c_2. The total area is at least

$$2\theta(1-2d\delta')^d r^d/2^d + (4d^2+2)\delta' r\theta(1-2d\delta')^{d-1}r^{d-1}/2^d = \\ 2\theta r^d/2^{d-1}\left[(1-2d\delta')^d + (4d^2+2)\delta'(1-2d\delta')^{d-1}/2\right],$$

and the claim follows by Lemma 4.

Lemma 6 can be extended effortlessly changing 16 to $4d^2$ and other constants appropriately. The remaining parts of the argument either are independent of d or require trivial extensions to the case of arbitrary $d \geq 3$.

References

[1] J. Balogh, B. Bollobás, and M. Walters. Hamilton cycles in random geometric graphs. Preprint, http://arxiv.org/abs/0905.4650

[2] B. Bollobás and A. Frieze, On matchings and Hamiltonian cyclesin random graphs. In *Random Graphs '83*, M. Karoński and A.Ruciński (Eds.), North-Holland, Amsterdam, pp. 23–46 (1985).

[3] J. Díaz, D. Mitsche, and X. Pérez. Sharp threshold for hamiltonicity of random geometric graphs. *SIAM Journal on Discrete Mathematics*, 21(1):57–65, 2007.

[4] M. Krivelevich, and T. Müller. Hamiltonicity of the random geometric graph. Preprint, http://arxiv.org/abs/0906.0071v2

[5] E. Lucas. Récreations Mathématiques. Vol II, Paris, 1892

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

[7] M.D. Penrose. *Random Geometric Graphs*, volume 5 of *Oxford Studies in Probability*. Oxford University Press, 2003.

[8] X. Pérez-Giménez, and N.C. Wormald. Disjoint Hamilton cycles in the random geometric graph. Preprint, http://arxiv.org/abs/0907.4459