Rainbow perfect matchings and Hamilton cycles in the random geometric graph

Deepak Bal[*] Patrick Bennett[†] Xavier Pérez-Giménez[‡] Paweł Prałat[§]

Abstract

Given a graph on n vertices and an assignment of colours to the edges, a rainbow Hamilton cycle is a cycle of length n visiting each vertex once and with pairwise different colours on the edges. Similarly (for even n) a rainbow perfect matching is a collection of $n/2$ independent edges with pairwise different colours. In this note we show that if we randomly colour the edges of a random geometric graph with sufficiently many colours, then a.a.s. the graph contains a rainbow perfect matching (rainbow Hamilton cycle) if and only if the minimum degree is at least 1 (respectively, at least 2). More precisely, consider n points (i.e. vertices) chosen independently and uniformly at random from the unit d-dimensional cube for any fixed $d \geq 2$. Form a sequence of graphs on these n vertices by adding edges one by one between each possible pair of vertices. Edges are added in increasing order of lengths (measured with respect to the ℓ_p norm, for any fixed $1 < p \leq \infty$). Each time a new edge is added, it receives a random colour chosen uniformly at random and with repetition from a set of $\lceil Kn \rceil$ colours, where $K = K(d)$ is a sufficiently large fixed constant. Then, a.a.s. the first graph in the sequence with minimum degree at least 1 must contain a rainbow perfect matching (for even n), and the first graph with minimum degree at least 2 must contain a rainbow Hamilton cycle.

1 Introduction

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ be n i.i.d. points in $[0,1]^d$ chosen with the uniform distribution, where $d \geq 2$ is fixed. Fix $1 < p \leq \infty$. Unless otherwise stated, distances and lengths in $[0,1]^d$ are measured with respect to the ℓ_p norm. We construct the *random geometric graph* $\mathscr{G}(\boldsymbol{X}; r)$ of radius r as follows. The vertices of $\mathscr{G}(\boldsymbol{X}; r)$ are indexed by $[n]$, and each pair of different vertices $i, j \in [n]$ are joined by an edge if and only if X_i and X_j are within (ℓ_p-normed) distance r. The *length* of an edge ij is defined to be $\|X_i - X_j\|_p$, and is always at most r by construction. With probability 1 all points in $\boldsymbol{X}$ are different, and fall in general position. Therefore, we will often identify vertex i and point X_i (i.e. we regard X as the vertex set), and assume all edges have different lengths. Random geometric graphs (or more precisely a slight variation of the model defined above) were first introduced by Gilbert [12], and have been widely investigated ever since. They provide a theoretical model for wireless ad-hoc networks, and have relevant applications in statistics. We refer the reader to Penrose's monograph [16] and a more recent survey by Walters [17] for further details and references on the subject.

---

[*]Department of Mathematical Sciences, Montclair State University, Montclair, NJ, 07043 U.S.A. `deepak.bal@montclair.edu`

[†]Department of Mathematics, Western Michigan University, Kalamazoo, MI, 49008 U.S.A. `patrick.bennett@wmich.edu`

[‡]Department of Mathematics, Ryerson University, Toronto, ON, Canada, M5B 2K3. `xperez@ryerson.ca`

[§]Department of Mathematics, Ryerson University, Toronto, ON, Canada, M5B 2K3. Research is supported in part by NSERC and Ryerson University. `pralat@ryerson.ca`

We consider the natural coupling in which all $\mathscr{G}(\boldsymbol{X};r)$ with $r \in [0,\infty)$ share one common vertex set $\boldsymbol{X}$. We call this coupling the *random geometric graph process*, and denote it by $\left(\mathscr{G}(\boldsymbol{X};r)\right)_{r \geq 0}$. Intuitively, the process starts at time $r = 0$ with an empty graph on vertex set $\boldsymbol{X}$ (almost surely, assuming that all vertices are at different positions). Then, as we increase r from 0 to ∞, we add edges one by one in increasing order of length. By construction, each snapshot of the process at a given time r is distributed precisely as a copy of $\mathscr{G}(\boldsymbol{X};r)$. Finally, $\mathscr{G}(\boldsymbol{X};r)$ is deterministically the complete graph for all $r \geq D$, where $D = \|(1,1,.\overset{d}{.}.,1)\|_p = d^{1/p}$ is the distance between two opposite corners of $[0,1]^d$.

A lot of work has been done to describe the connectivity properties of random geometric graphs in this process and the emergence of spanning subgraphs (such as perfect matchings and Hamilton cycles). A celebrated result of Penrose [15] asserts that a.a.s.[1] the first edge added during the process that gives minimum degree at least k also makes the graph k-connected. More precisely, let

$$\widehat{r}_{\delta \geq k} = \widehat{r}_{\delta \geq k}(\boldsymbol{X}) = \min\left\{r \geq 0 : \mathscr{G}(\boldsymbol{X};r) \text{ has minimum degree at least } k\right\} \quad \text{and}$$
$$\widehat{r}_{k\text{-conn}} = \widehat{r}_{k\text{-conn}}(\boldsymbol{X}) = \min\left\{r \geq 0 : \mathscr{G}(\boldsymbol{X};r) \text{ is } k\text{-connected}\right\}.$$

Then, for every constant $k \in \mathbb{N}$,

$$\lim_{n \to \infty} \mathbf{Pr}\left(\widehat{r}_{\delta \geq k} = \widehat{r}_{k\text{-conn}}\right) = 1. \tag{1}$$

In view of this, Penrose (cf. [16]) asked whether a.a.s. that first edge in the process that gives minimum degree at least 2 (and ensures 2-connectivity) is also responsible for the emergence of a Hamilton cycle. A first step in this direction was achieved by Díaz, Mitsche and Pérez [6], who showed (for dimension $d = 2$) that, given any constant $\varepsilon > 0$, a.a.s. $\mathscr{G}(\boldsymbol{X};(1+\varepsilon)\widehat{r}_{\delta \geq 2})$ contains a Hamilton cycle. Some of their ideas were recently extended by three research teams (Balogh, Bollobás and Walters; Krivelevich and Müller; and Pérez-Giménez and Wormald), who independently settled Penrose's question in the affirmative (but only two papers [4] and [14] were finally published). In particular, a more general packing result in [14] implies that

$$\text{a.a.s. } \begin{cases} \mathscr{G}(\boldsymbol{X};\widehat{r}_{\delta \geq 2}) \text{ contains a Hamilton cycle, and} \\ \mathscr{G}(\boldsymbol{X};\widehat{r}_{\delta \geq 1}) \text{ contains a perfect matching (for even } n). \end{cases} \tag{2}$$

Clearly, this claim is the best possible, since any graph with minimum degree less than 1 (less than 2) cannot have a perfect matching (respectively, Hamilton cycle).

In this paper we consider an edge-coloured version of the random geometric graph. (Throughout the manuscript, we will use the term *edge colouring* (and other terms alike) to denote an assignment of colours to the edges, not necessarily proper in a graph-theoretical sense.) Let $\boldsymbol{Z} = (Z_{ij})_{1 \leq i < j \leq n}$ be a random vector of colours, chosen independently and with replacement from a set of colours of size c. We use this vector $\boldsymbol{Z}$ to colour the edges of the random geometric graph: each edge ij of $\mathscr{G}(\boldsymbol{X};r)$ $(1 \leq i < j \leq n)$ is assigned colour Z_{ij}. We denote this model by $\mathscr{G}(\boldsymbol{X};\boldsymbol{Z};r)$. Similarly, we consider the coupled process $\left(\mathscr{G}(\boldsymbol{X};\boldsymbol{Z};r)\right)_{r \geq 0}$ in which, as we increase r from 0 to ∞, we add new edges in increasing order of length, and each new edge ij is coloured according to Z_{ij}.

Given a graph with colours assigned to the edges, we say the graph (or its edge set) is *rainbow* if all edges receive different colours. Recently, there have been many papers written on the subject of rainbow spanning structures in randomly edge coloured random graphs and digraphs (see *e.g.* [1, 2, 3, 5, 8, 9, 10, 11, 13]). Let $G_c(n,p)$ denote the binomial random graph $G(n,p)$ where each edge

---

[1]We say that a sequence of events H_n holds *asymptotically almost surely* (a.a.s.) if $\lim_{n \to \infty} \mathbf{Pr}(H_n) = 1$.

has independently been assigned a uniformly random colour from a set of size c. For graphs H with maximum degree $\Delta = \Delta(H)$ and n vertices, Ferber, Nenadov and Peter [10] showed that a.a.s. $G_c(n,p)$ contains a rainbow copy of H, provided that $p = n^{1/\Delta} \operatorname{polylog}(n)$ and $c = (1 + o(1))e(H)$. Here, the number of colours is asymptotically optimal, whereas the bound on p most likely is not. For Hamilton cycles tighter results are known. In [5], Cooper and Frieze determined that a.a.s. $G_c(n,p)$ contains a rainbow Hamilton cycle if $p \geq 42 \log n/n$ and $c \geq 21n$. This was later improved by Frieze and Loh [11] and recently even further improved by Ferber and Krivelevich [8] who showed that it holds when $c = (1 + o(1))n$ and $p = \frac{\log n + \log \log n + \omega(1)}{n}$. Here, the number of colours is asymptotically optimal and the bound on p is optimal. Bal and Frieze [3] examined the case when the number of colours is exactly optimal, showing that $G_n(n,p)$ a.a.s. contains a rainbow Hamilton cycle as long as $p = \Omega\left(\frac{\log n}{n}\right)$.

In this manuscript, we investigate the emergence of rainbow spanning structures in the random geometric graph. Our main contribution is to extend (2) to a rainbow context. We show that a.a.s. the first edge in the edge-coloured random geometric graph process $\left(\mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; r)\right)_{r \geq 0}$ that gives minimum degree at least 2 also creates a rainbow Hamilton cycle, provided that the number of colours is at least $c = \lceil Kn \rceil$, where $K = K(d) > 0$ is a sufficiently large constant. Similarly, under the same assumptions (with n even), the first edge in the process that ensures that the minimum degree is at least 1 creates a rainbow perfect matching. To state the result more precisely, let

$$\widehat{r}_{\mathrm{RPM}} = \widehat{r}_{\mathrm{RPM}}(\boldsymbol{X}, \boldsymbol{Z}) = \inf\left\{r \geq 0 : \mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; r) \text{ contains a rainbow perfect matching}\right\} \quad \text{and}$$
$$\widehat{r}_{\mathrm{RHC}} = \widehat{r}_{\mathrm{RHC}}(\boldsymbol{X}, \boldsymbol{Z}) = \inf\left\{r \geq 0 : \mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; r) \text{ contains a rainbow Hamilton cycle}\right\},$$

where we use the convention that $\inf \emptyset = \infty$. Note that whenever $\widehat{r}_{\mathrm{RPM}} < \infty$ ($\widehat{r}_{\mathrm{RHC}} < \infty$) the infimum in the above definition is actually a minimum, and it is precisely the length of the first edge in the process $\left(\mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; r)\right)_{r \geq 0}$ that creates a rainbow perfect matching (respectively, rainbow Hamilton cycle).

Theorem 1. *Given a fixed integer $d \geq 2$, there exists a sufficiently large constant $K = K(d) > 0$ satisfying the following. Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ be n i.i.d. points in $[0,1]^d$ chosen uniformly at random, and let $\boldsymbol{Z} = (Z_{ij})_{1 \leq i < j \leq n}$ be a random vector of colours, choosen independently and with replacement from a set of colours of size $\lceil Kn \rceil$. Consider the random geometric graph process $\left(\mathscr{G}(\boldsymbol{X}; r)\right)_{r \geq 0}$ (for any fixed ℓ_p-normed distance, $1 < p \leq \infty$) with a random colouring of the edges given by $\boldsymbol{Z}$. Then,*

$$\lim_{n \to \infty} \mathbf{Pr}\big(\widehat{r}_{\mathrm{RHC}}(\boldsymbol{X}, \boldsymbol{Z}) = \widehat{r}_{\delta \geq 2}(\boldsymbol{X})\big) = 1,$$

and for even n

$$\lim_{\substack{n \to \infty \\ (n \text{ even})}} \mathbf{Pr}\big(\widehat{r}_{\mathrm{RPM}}(\boldsymbol{X}, \boldsymbol{Z}) = \widehat{r}_{\delta \geq 1}(\boldsymbol{X})\big) = 1.$$

Remark 2. *For $p = 1$, we can only claim that*

$$\lim_{n \to \infty} \mathbf{Pr}\big(\widehat{r}_{\mathrm{RHC}}(\boldsymbol{X}, \boldsymbol{Z}) = \widehat{r}_{\text{2-conn}}(\boldsymbol{X})\big) = 1 \quad and \quad \lim_{\substack{n \to \infty \\ (n \text{ even})}} \mathbf{Pr}\big(\widehat{r}_{\mathrm{RPM}}(\boldsymbol{X}, \boldsymbol{Z}) = \widehat{r}_{\text{1-conn}}(\boldsymbol{X})\big) = 1, \quad (3)$$

and in fact it is not known whether (1) holds (see [15] and [16]). We include a justification of (3) in Section 5 for completeness.

Combining Theorem 1 and Theorem 8.4 in [16], we immediately obtain the limiting probabilities of having a rainbow perfect matching and having a rainbow Hamilton cycle, assuming that the number of colours is sufficiently large.

Corollary 3. *Under the same assumptions of Theorem 1, put*

$$r = \sqrt[d]{\frac{(2/d)\log n + (3 - d - 2/d)\log\log n + x}{2^{2-d}\theta n}},$$

and let $f = \log\left(2^{1-2/d}(\theta d)^{3-2/d}\theta'^{d-2}/\binom{d}{2}\right)$, where θ and θ' are the volumes of the d-dimensional and $(d-1)$-dimensional unit ℓ_p-balls, respectively. Then

$$\lim_{\substack{n\to\infty \\ (n \text{ even})}} \mathbf{Pr}\big(\mathscr{G}(\mathbf{X};\mathbf{Z};r) \text{ has a rainbow perfect matching}\big) = \begin{cases} 0 & x \to -\infty; \\ \exp\left(-e^{-\alpha-f}\right) & x \to \alpha; \\ 1 & x \to \infty. \end{cases}$$

Corollary 4. *Under the same assumptions of Theorem 1, put*

$$r = \sqrt[d]{\frac{(2/d)\log n + (4 - d - 2/d)\log\log n + y}{2^{2-d}\theta n}},$$

and let f, θ and θ' be as in Corollary 3. Then if $d \geq 3$,

$$\lim_{n\to\infty} \mathbf{Pr}\big(\mathscr{G}(\mathbf{X};\mathbf{Z};r) \text{ has a rainbow Hamilton cycle}\big) = \begin{cases} 0 & y \to -\infty; \\ \exp\left(-2e^{-\alpha-f}/d\right) & y \to \alpha; \\ 1 & y \to \infty. \end{cases}$$

Otherwise if $d = 2$,

$$\lim_{n\to\infty} \mathbf{Pr}\left(\begin{matrix} \mathscr{G}(\mathbf{X};\mathbf{Z};r) \text{ has a} \\ \text{rainbow Hamilton cycle} \end{matrix}\right) = \begin{cases} 0 & y \to -\infty; \\ \exp\left(-e^{-\alpha/2}\left(e^{-\alpha/2} + \frac{2\sqrt{\theta}}{\theta'}\right)\right) & y \to \alpha; \\ 1 & y \to \infty. \end{cases}$$

In Section 2, we will discuss how we partition $[0,1]^d$ into a grid of small d-dimensional cubic cells. Having this partition will simplify our task by allowing us to locally search for short rainbow paths or cycles within each cell or small cluster of cells. In Section 3, we will show how to find these paths or cycles. In Section 4, we will show how to connect the pieces together into one rainbow Hamilton cycle. A simple adaptation of the argument can be used to build a rainbow perfect matching. Finally, in Section 5 we will discuss the case $p = 1$, and pose some open questions in Section 6.

2 Tessellation and graph of cells

The main goal in this section is to prove Lemmas 13 and 14, which will be crucial in the construction of the rainbow Hamilton cycle and perfect matching. We will adapt and extend many ideas from [14]. Throughout the paper, $d \geq 2$ and the ℓ_p-norm ($1 < p \leq \infty$) in $\mathbb{R}^d$ are fixed. Note that the volume θ of the unit d-dimensional ℓ_p-ball satisfies

$$2^d/d! \leq \theta \leq 2^d \tag{4}$$

since the ℓ_p-ball contains the ℓ_1-ball and is contained in the ℓ_∞-ball, which have volume $2^d/d!$ and 2^d respectively. We will also make frequent use of the following inequality throughout the argument, often without explicitly mentioning it. For any $X \in \mathbb{R}^d$,

$$\|X\|_\infty \leq \|X\|_p \leq d\|X\|_\infty. \tag{5}$$

4

Moreover, we will pick a sufficiently small constant $\varepsilon = \varepsilon(d) > 0$ so that several requirements in the argument are met. Later we will choose K sufficiently large with respect to this ε (recall $\lceil Kn \rceil$ is the number of colours). We remark that our choice of ε does not depend on p or other parameters that may be introduced later in the statements.

We use the standard $o()$, $O()$, $\Theta()$ and $\Omega()$ asymptotic notation as $n \to \infty$ with the following extra considerations. We do not assume any sign on a sequence a_n satisfying $a_n = o(1)$ or $a_n = O(1)$, but on the other hand a sequence satisfying $a_n = \Theta(1)$ or $a_n = \Omega(1)$ is assumed to be positive for all but finitely many n. Furthermore, the constants involved in the bounds of the definitions of $O()$, $\Theta()$ and $\Omega()$ may depend on d, but not on p or ε. Whenever these constants depend on our choice of ε (in addition to d), we use the alternative notation $O_\varepsilon()$, $\Theta_\varepsilon()$ and $\Omega_\varepsilon()$ instead. Our asymptotic statements are always uniform for all $1 < p \leq \infty$ as a consequence of bounds (4) and (5).

Let $\omega = \omega(n) \to \infty$ be some function tending to infinity sufficiently slowly (in particular, $\omega = o(\log \log n)$). We define r_0 and r_1 by

$$\theta n r_0{}^d = (2^{d-1}/d) \log n + 2^{d-2}(3 - d - 2/d) \log \log n - \omega \qquad \text{and}$$
$$\theta n r_1{}^d = (2^{d-1}/d) \log n + 2^{d-2}(4 - d - 2/d) \log \log n + \omega.$$

Then, by Theorem 8.4 in [16], the respective lengths $\widehat{r}_{\delta \geq 1}$ and $\widehat{r}_{\delta \geq 2}$ of the critical edges of the process $\big(\mathscr{G}(\boldsymbol{X}; r) \big)_{r \geq 0}$ that give minimum degree 1 and 2 satisfy

$$r_0 \leq \widehat{r}_{\delta \geq 1} \leq \widehat{r}_{\delta \geq 2} \leq r_1 \sim r_0 \qquad \text{a.a.s.} \tag{6}$$

Our argument will use edges of length at most r_0 to construct most of the rainbow perfect matching or Hamilton cycle, and only use a few longer edges of length up to $\widehat{r}_{\delta \geq 1}$ or $\widehat{r}_{\delta \geq 2}$ at some exceptional places.

Let $s' = (2\varepsilon d\theta)^{1/d} r_0/2$. We tessellate $[0, 1]^d$ into d-dimensional cubic *cells* of side length

$$s = \big\lceil (s')^{-1} \big\rceil^{-1} \sim s' = \Theta(\varepsilon^{1/d} r_0),$$

that is, of volume

$$s^d \sim \varepsilon d 2^{1-d} \theta r_0{}^d \sim \varepsilon \log n / n,$$

arranged in a grid fashion. Let $\mathcal{C}$ denote the set of cells. There are $\Theta_\varepsilon(n/\log n) = o(n)$ cells in $\mathcal{C}$, where we recall that the constant hidden in the $\Theta_\varepsilon(\cdot)$ notation depends on ε (and d). Clearly (assuming that ε is sufficiently small given d and by (4)), the vertices inside each cell induce a clique in $\mathscr{G}(\boldsymbol{X}; r_0)$, and in fact a stronger property holds in view of the following definition.

Definition 1. The *graph of cells* $\mathscr{G}_\mathcal{C}$ is a graph with vertex set $\mathcal{C}$ (i.e. the set of cells of the tessellation), and two cells are adjacent in $\mathscr{G}_\mathcal{C}$ if they are at (ℓ_p-normed) distance at most $r_0 - 2ds = (1 - \Theta(\varepsilon^{1/d})) r_0$.

This implies that for any pair of adjacent cells and any pair of points $X_i, X_j \in \boldsymbol{X}$ that belong to these cells, X_i and X_j must be adjacent in the graph $\mathscr{G}(\boldsymbol{X}; r_0)$. (This is true regardless of the ℓ_p-norm being used, in view of (5).) The degree of a cell C in the graph of cells $\mathscr{G}_\mathcal{C}$ is at most the number of cells contained in a ball of radius r_0 centered at the center of C. As each cell has volume $\Theta(\varepsilon r_0{}^d)$ and the ball of radius r_0 has volume $\theta r_0{}^d$, we deduce that

$$\text{the maximum degree of the graph of cells is } \Delta(\mathscr{G}_\mathcal{C}) = O(1/\varepsilon). \tag{7}$$

Note that the number of points of $\boldsymbol{X}$ that fall into each cell is distributed as $\text{Bin}(n, s^d)$ with expectation $s^d n \sim \varepsilon \log n$. Then, we can easily bound the maximum number of points in a cell.

Lemma 5. *A.a.s. no cell in $\mathcal{C}$ contains more than $\log n$ vertices of $\boldsymbol{X}$.*

Proof.

$$\mathbf{Pr}\left[\mathrm{Bin}(n, s^d) \geq \log n\right] \leq \binom{n}{\lceil\log n\rceil} s^{d\lceil\log n\rceil} \leq \left(e\varepsilon + o(1)\right)^{\lceil\log n\rceil} = o(1/n),$$

since $e\varepsilon < e^{-1}$ (assuming ε is sufficiently small). Thus, a union bound over all $O_\varepsilon(n/\log n)$ cells shows that a.a.s. none has more than $\log n$ many points. $\square$

Similarly, we can perform analogous calculations to bound the number of vertices of $\boldsymbol{X}$ that fall inside of a ball of radius ℓr_1 centered around a vertex $X \in \boldsymbol{X}$, take a union bound over all n choices of X, and conclude the following.

Lemma 6. *Given any constant $\ell \in \mathbb{N}$, the maximum degree of the power graph $\mathscr{G}(\boldsymbol{X}; r_1)^\ell$ is a.a.s. $O(\log n)$.*

Definition 2. We say a cell is *dense* if it has at least $\varepsilon^3 \log n$ points of $\boldsymbol{X}$ in it, and is otherwise *sparse*.

Note that this definition is different from the corresponding notions in [14] and [4], which only require dense cells to contain a large but constant number of points of $\boldsymbol{X}$. We will show we cannot have too many sparse cells or too large connected sets of sparse cells in the graph of cells $\mathscr{G}_\mathcal{C}$ (or in its ℓ-th power $\mathscr{G}_\mathcal{C}^\ell$, for a fixed $\ell \in \mathbb{N}$). We shall also take into account whether these cells are "close" to the boundary of the cube $[0,1]^d$. To make this precise, define $F_j^\beta = [0,1]^{j-1} \times \{\beta\} \times [0,1]^{d-j}$ for $\beta \in \{0,1\}$ and $j \in \{1,2,\ldots,d\}$. These are the $2d$ facets (i.e. $(d-1)$-dimensional faces) of the boundary of $[0,1]^d$.

Lemma 7.

1. *A.a.s. the number of sparse cells is at most $n^{1-\varepsilon/2}$.*

2. *Moreover, for any arbitrary constants $A > 0$ and $\ell \in \mathbb{N}$, a.a.s. the power graph $\mathscr{G}_\mathcal{C}^\ell$*

   (a) *has no connected set $\mathcal{S}$ of at least $(1+\varepsilon)/\varepsilon$ cells which are all sparse;*

   (b) *has no connected set $\mathcal{S}$ of at least $\frac{d-i}{d}(1+\varepsilon)/\varepsilon$ cells which are all sparse and such that some cell in $\mathcal{S}$ lies within distance Ar_0 from at least i facets of $[0,1]^d$ (for $i \in \{0,1,\ldots,d-1\}$);*

   (c) *has no sparse cell within distance Ar_0 from d facets of $[0,1]^d$.*

Remark 8.

1. *The property "of at least $\frac{d-i}{d}(1+\varepsilon)/\varepsilon$ cells" in the statement can be replaced by "of total volume at least $(1+\varepsilon)\frac{d-i}{2^{d-1}}\theta r_0{}^d$", and the claim is still valid.*

2. *Note that this lemma provides an analogue of Lemma 4 in [14]. However, the latter gives a $\frac{d-i}{d}(1+\alpha)/\varepsilon$ bound on the size of $\mathcal{S}$, for α arbitrarily small (possibly much smaller than any fixed function of ε). Here, we cannot achieve that, given our more restrictive definition of dense cell (i.e. sparse cells are more abundant). However, the current statement will suffice for our purposes. In particular, Lemma 4 in [14] is used in the proof of Lemma 5 in [14] with $\alpha = \Theta(\varepsilon^{1/d})$, which is greater than ε (if ε is sufficiently small), and thus this situation is covered by our present statement. Hence, Lemma 5 in [14] is still valid with our definition of dense cells, since we can replace all uses of Lemma 4 [14] in the proof by its counterpart in this manuscript.*

Proof of Lemma 7. Recalling that the number of points in any fixed cell is distributed as $\text{Bin}(n, s^d)$, the probability that a cell is sparse is (with the convention that $(a/0)^0 = 1$ for all $a \in \mathbb{R}$)

$$\sum_{k=0}^{\lceil \varepsilon^3 \log n \rceil - 1} \binom{n}{k} s^{dk} \left(1 - s^d\right)^{n-k} \leq \sum_{k=0}^{\lceil \varepsilon^3 \log n \rceil - 1} \left(\frac{ens^d}{k(1 - s^d)}\right)^k e^{-s^d n}$$

$$\leq \lceil \varepsilon^3 \log n \rceil \left(\frac{e + o(1)}{\varepsilon^2}\right)^{\varepsilon^3 \log n} e^{-(\varepsilon + o(1)) \log n}$$

$$= n^{-\varepsilon + \varepsilon^3 \log(e/\varepsilon^2) + o(1)}$$

$$\leq n^{-\varepsilon(1 - \varepsilon/2)} \quad \text{(for large n)},$$

provided that ε is chosen so that $\varepsilon \log(e/\varepsilon^2) < 1/2$. Note that this is possible since $\varepsilon \log(e/\varepsilon^2) \to 0$ as $\varepsilon \to 0$. In particular, this condition implies that $\varepsilon < 1$ and so the number of sparse cells is a.a.s. at most $n^{1-\varepsilon/2}$ by Markov's inequality. This proves part 1.

Fix $A > 0$, $\ell \in \mathbb{N}$ and $i \in \{0, 1, \ldots, d-1\}$. In order to prove 2(b), it is enough to show that $\mathscr{G}_{\mathcal{C}}^\ell$ contains no connected set of exactly $\lceil \frac{d-i}{d}(1 + \varepsilon)/\varepsilon \rceil$ sparse cells within distance Ar_0 from i facets of $[0, 1]^d$. Observe that the events that two or more cells are sparse are negatively correlated. Therefore,

the probability that k given cells are sparse is at most $n^{-k\varepsilon(1-\varepsilon/2)}$. (8)

If we also choose ε small enough so that $(1 - \varepsilon/2)(1 + \varepsilon) > 1$, the probability that a given set of $\lceil ((d - i)/d)(1 + \varepsilon)/\varepsilon \rceil$ cells are all sparse is $o(n^{-(d-i)/d})$. Since there are only $O_\varepsilon\left((n/\log n)^{(d-i)/d}\right)$ possible connected sets of $\lceil ((d - i)/d)(1 + \varepsilon)/\varepsilon \rceil$ cells in the power graph $\mathscr{G}_{\mathcal{C}}^\ell$ lying within distance Ar_0 from i facets of $[0, 1]^d$, we can take the union bound and complete the proof of part 2(b). Part 2(a) follows as a particular case of part 2(b) taking $i = 0$. Finally, the expected number of sparse cells within distance Ar_0 from d facets of $[0, 1]^d$ is at most $O_\varepsilon(1)$ times $n^{-\varepsilon(1-\varepsilon/2)}$, which is $o(1)$. This immediately yields part 2(c) and completes the proof. $\qquad \square$

Given a set of cells $\mathcal{S} \subseteq \mathcal{C}$, we denote by $\mathscr{G}_{\mathcal{C}}[\mathcal{S}]$ the subgraph of the graph of cells induced by $\mathcal{S}$.

Definition 3. Let $\mathcal{D}$ be the set of dense cells, and let $\mathcal{G}$ be the set of cells in the largest component of $\mathscr{G}_{\mathcal{C}}[\mathcal{D}]$ (i.e. the subgraph of $\mathscr{G}_{\mathcal{C}}$ induced by dense cells). If there are two or more such largest components, pick one according to any arbitrary deterministic rule (we will see that a.a.s. $\mathcal{G}$ is very large, so the choice of $\mathcal{G}$ is unique). Call cells in $\mathcal{G}$ *good*. Cells that are not good, but are adjacent in the graph of cells to some good cell are called *bad*. Bad cells must be sparse by construction. Cells that are not adjacent to good cells are called *ugly*. Note that ugly cells may be dense or sparse. Let $\mathcal{B}$ and $\mathcal{U}$ denote the set of bad and ugly cells, respectively.

As a crucial ingredient in our argument, we will use Lemma 5 in [14], which shows that a.a.s. ugly cells (which are called "bad" in that paper) appear in small clusters far enough from each other. (Note that this result is stated for dimension $d = 2$, and then extended to general $d \geq 2$ in Section 4 of [14].) Unfortunately, the definition of dense cell we use in the present manuscript is more restrictive than the one in [14] (they only require a dense cell to contain at least M points, for a large constant $M > 0$; whilst here we require at least $\varepsilon^3 \log n$ points). In order to overcome this minor obstacle, we simply observe that our Lemma 7 extends Lemma 4 in [14] to a less restrictive notion of sparse cell (although with a slightly weaker bound). In view of Remark 8(2), the proof of Lemma 5 in [14] is also valid in our setting by trivially replacing Lemma 4 in [14] by Lemma 7 of the present paper. Hence, adapting Lemma 5 in [14] to our current notation, we obtain the following statement.

Lemma 9 ([14])**.** *A.a.s. all connected components of $\mathscr{G}_C[\mathcal{U}]$ have ℓ_∞-diameter at most $4d^2s$.*

Next, we obtain useful bounds on the number of bad and ugly cells.

Lemma 10. *A.a.s. there are at most $n^{1-\varepsilon/2}$ bad cells and at most $n^{O\left(\varepsilon^{1/d}\right)}$ ugly cells.*

In particular, this implies that a.a.s. our choice of $\mathcal{G}$ in Definition 3 was unique.

Proof. Since bad cells are sparse by definition, the first part of the statement follows trivially from Lemma 7(1).

To prove the second part, consider a cell $C = [a_1, a_1 + s] \times \cdots \times [a_d, a_d + s]$ which is at distance at most r_0 from exactly i facets of the cube $[0, 1]^d$, for some $0 \leq i \leq d$. Without loss of generality, assume these facets are precisely $F_1^0, \ldots, F_i^0$. Let $P = (a_1 + s, \ldots, a_d + s)$ (i.e. P is the point in C with largest coordinates), and let $B = B(0; r_0 - 4ds)$ denote the ℓ_p ball with centre $0 \in \mathbb{R}^d$ and radius $r_0 - 4ds$. By construction, the set

$$S = P + \left(B \cap \left([0, \infty)^i \times \mathbb{R}^{d-i} \right) \right)$$

contains precisely those points within distance $r_0 - 4ds$ of P and "above" P with respect to the first i coordinates. Moreover, S is fully contained in the cube $[0, 1]^d$, and every cell $C' \neq C$ intersecting S must belong to the set $\mathcal{N}$ of cells that are adjacent to C in the graph of cells $\mathscr{G}_C$. Therefore, the number of cells in $\mathcal{N}$ satisfies

$$
\begin{aligned}
|\mathcal{N}| + 1 &\geq \mathrm{vol}(S)/s^d \\
&= 2^{-i}\theta r_0{}^d (1 - 4ds/r_0)^d/s^d \\
&\geq \left(2^{d-i-1}/d + o(1) \right) \varepsilon^{-1}(1 - 4d^2s/r_0) \\
&\geq \left(2^{d-i-1}/d \right) \varepsilon^{-1}(1 - \alpha\varepsilon^{1/d}) \quad \text{(eventually)}
\end{aligned}
\tag{9}
$$

for some constant $\alpha = \alpha(d) > 0$, where we used $r_0/s \sim 2(\varepsilon d2\theta)^{-1/d}$ and the fact that $(1 - x)^d \geq 1 - dx$ for $0 \leq x \leq 1$. Furthermore, let E_C be the event that C is ugly and is adjacent in $\mathscr{G}_C$ to at most $(4d^2)^d - 1$ other ugly cells. This event implies that there must be a set $\mathcal{N}'$ of at least

$$|\mathcal{N}| - (4d^2)^d \geq \left(2^{d-i-1}/d \right) \varepsilon^{-1}(1 - 2\alpha\varepsilon^{1/d}) \tag{10}$$

sparse cells adjacent to C. Note that the last line follows since we can choose ϵ sufficiently small, given d (and α). Thus, by (8) and summing over all $O_\varepsilon(1)$ possible choices of such $\mathcal{N}'$, we obtain that the probability of E_C is at most

$$O_\varepsilon(1) \cdot n^{-\left(2^{d-i-1}/d \right)(1-2\alpha\varepsilon^{1/d})(1-\varepsilon/2)} \leq n^{-\left(2^{d-i-1}/d \right)\left(1-\Theta\left(\varepsilon^{1/d} \right) \right)} \quad \text{(eventually)},$$

assuming that ε is sufficiently small. Hence, summing over i and over the $O_\varepsilon\left((n/\log n)^{(d-i)/d} \right)$ possible choices of C, we show that the expected number of cells that are ugly and are adjacent in $\mathscr{G}_C$ to at most $(4d^2)^d$ other ugly cells is

$$\sum_{i=0}^d O_\varepsilon \left((n/\log n)^{(d-i)/d} n^{-\left(2^{d-i-1}/d \right)\left(1-\Theta\left(\varepsilon^{1/d} \right) \right)} \right) = n^{O\left(\varepsilon^{1/d} \right)},$$

since $d - i \leq 2^{d-i-1}$. By Markov's inequality and in view of Lemma 9 (which implies that a.a.s. there are no ugly cells that are adjacent to more than $(4d^2)^d$ other ugly cells), we conclude that a.a.s. there are at most $n^{O\left(\varepsilon^{1/d} \right)}$ ugly cells. This finishes the proof of the second statement. $\square$

Lemma 11. *Given any constant $A > 0$, a.a.s. every two ugly cells lying in different components of $\mathscr{G}_C[\mathcal{U}]$ are at (ℓ_p-normed) distance at least Ar_0 apart.*

Proof. We will assume that the a.a.s. conclusions of Lemmas 7 and 9 hold, and deterministically prove that any pair of nonadjacent cells in $\mathscr{G}_C$ at distance less than Ar_0 cannot both be ugly. Thus, consider any two different cells $C = [a_1, a_1+s] \times \cdots \times [a_d, a_d+s]$ and $C' = [a'_1, a'_1+s] \times \cdots \times [a'_d, a'_d+s]$ that are not adjacent in the graph of cells $\mathscr{G}_C$, but are at distance less than Ar_0 from each other. Suppose that there are exactly i facets of the d-cube $[0,1]^d$ (for some $0 \leq i \leq d$) at distance less than r_0 from C or C'. Assume, without loss of generality, that these facets are precisely $F_1^0, \ldots, F_i^0$. In particular, both C and C' must be at distance at most $(A+1)r_0 + 2ds \leq (A+2)r_0$ from these i facets (for ε sufficiently small given d), and at distance at least r_0 from any other facet.

We will proceed in a similar fashion as in the proof of Lemma 10 in order to describe a large set of cells that are adjacent to C or C'. Let $P = (a_1 + s, \ldots, a_d + s)$ and $P' = (a'_1 + s, \ldots, a'_d + s)$ be respectively the points in C and C' with largest coordinates. The hyperplane in $\mathbb{R}^d$ orthogonal to vector $P' - P$ (with respect to the Euclidean inner product) and passing through the origin splits $\mathbb{R}^d$ into two halfspaces

$$H = \{Q \in \mathbb{R}^d : \langle Q, P' - P \rangle \leq 0\} \qquad \text{and} \qquad H' = \{Q \in \mathbb{R}^d : \langle Q, P' - P \rangle \geq 0\}.$$

Consider the set

$$B_i = B \cap \left([0, \infty)^i \times \mathbb{R}^{d-i}\right)$$

where $B = B(0; r_0 - 4ds)$ denotes the ℓ_p ball with centre $0 \in \mathbb{R}^d$ and radius $r_0 - 4ds$. Since B_i has volume $2^{-i}\theta r_0{}^d(1 - 4ds/r_0)^d$, at least one of the sets $B_i \cap H$ or $B_i \cap H'$ has volume at least half of this. Assume it is $B_i \cap H'$ (by otherwise reversing the roles of P and P'). Define the sets

$$S = P + B_i \qquad \text{and} \qquad S' = P' + B_i \cap H'.$$

Since C and C' are not adjacent in $\mathscr{G}_C$, P and P' must be at distance greater than $r_0 - 2ds$. In particular, $P' \notin S$. Moreover, by our choice of H', any other point in S' is further away from P than P' is, so S and S' must be disjoint sets. Let $\mathcal{N}$ be the set of cells different than C and C' intersecting $S \cup S'$. By construction, every cell in $\mathcal{N}$ must be adjacent in the graph of cells $\mathscr{G}_C$ to C or C'. Moreover, since $\mathrm{vol}(S \cup S') \geq (3/2)\,\mathrm{vol}(S)$, we can use (9) to infer

$$|\mathcal{N}| + 2 \geq \mathrm{vol}(S \cup S')/s^d \geq (3/2)\left(2^{d-i-1}/d\right)\varepsilon^{-1}(1 - \alpha\varepsilon^{1/d}) \geq (4/3)\frac{d-i}{d}\varepsilon^{-1},$$

for ε sufficiently small given d, where we used that $d - i \leq 2^{d-i-1}$. In view of Lemma 9 (assuming its a.a.s. conclusion is true), if both C and C' are ugly, all but at most $2(4d^2)^d$ cells in $\mathcal{N}$ must be sparse. Then, we obtained a set of at least $(4/3)\frac{d-i}{d}\varepsilon^{-1} - 2(4d^2)^d \geq (5/4)\frac{d-i}{d}\varepsilon^{-1}$ sparse cells within distance at most $(A+2)r_0$ from each other (i.e. pairwise adjacent in the power graph $\mathscr{G}_C^\ell$ for $\ell = \lceil A+2 \rceil$). This contradicts the a.a.s. conclusion of Lemma 7. $\quad\square$

Lemma 12. *A.a.s. the following holds. Given any ugly cell C, let $\mathcal{G}' \subseteq \mathcal{G}$ be the set of good cells at ℓ_∞-distance at most $3r_0$ from C. Then, the subgraph of $\mathscr{G}_C$ induced by $\mathcal{G}'$ is connected and has diameter (as a graph) at most $2(20d)^d$.*

Proof. Let $t' = r_0/(3d)$. Tesselate $[0,1]^d$ by d-dimensional cubes of side $t = \lceil (t')^{-1} \rceil^{-1} \sim r_0/(3d)$. We call these cubes *boxes* to distinguish them from the cells. Note that a cell may intersect a box and not be fully contained in it. Regardless of that, each box contains at least

$$(t/s - 2)^d \geq \varepsilon^{-1}/\left(\theta(2d)^{d+1}\right) \tag{11}$$

9

cells. We say that two different boxes B and B' are adjacent if the set $B \cup B'$ is topologically connected (i.e. they share at least one point). (Recall that the term adjacent has a different meaning for cells.) If two different cells C and C' intersect the same box or two adjacent boxes, then C and C' must be at ℓ_p-distance at most $2dt \sim 2r_0/3$, and thus must be adjacent in the graph of cells $\mathscr{G}_\mathcal{C}$.

Fix an ugly cell C. We will assume that all the a.a.s. properties leading to the conclusions of Lemmas 7, 9, 10 and 11 hold, and deterministically prove the statement for cell C. Let $R_1 \subseteq [0,1]^d$ be the union of all boxes at ℓ_∞-distance between $1.1r_0$ and $3r_0 - t$ from cell C. Since the side of a box is $t < r_0/5$, then R_1 is topologically connected. Moreover, $[0,1]^d \setminus R_1$ has two connected components R_0 and R_2 with, say, $C \subseteq R_0$. (It is worth noting that these properties of R_1 hold regardless of how close cell C is from some facets of $[0,1]^d$, but this may cease to be true if we replaced ℓ_∞ by some other ℓ_p in the definition of R_1.) Let $\mathcal{G}_0$ and $\mathcal{G}_2$ be the set of good cells contained in R_0 and R_2 (respectively), and let $\mathcal{G}_1$ be the set of good cells that intersect R_1. Every good cell in $\mathcal{G}$ must belong to exactly one of the sets $\mathcal{G}_0$, $\mathcal{G}_1$ and $\mathcal{G}_2$. Observe that every cell in $\mathcal{G}_0$ is at ℓ_∞-distance at most $1.1r_0 + t \le 1.3r_0$ from C, and every cell in $\mathcal{G}_2$ is at ℓ_∞-distance at least $3r_0 - t \ge 2.8r_0$ from C. Therefore, if $C_0 \in \mathcal{G}_0$ and $C_2 \in \mathcal{G}_2$, then C_0 and C_2 must be at ℓ_∞-distance at least $1.5r_0 - ds \ge r_0$. In particular, no cell in $\mathcal{G}_0$ is adjacent to any cell in $\mathcal{G}_2$ with respect to the graph of cells $\mathscr{G}_\mathcal{C}$.

Claim 1. Every box $B \subseteq R_1$ contains a good cell.

Since R_1 is a topologically connected union of boxes and pairs of cells contained in adjacent boxes are adjacent in the graph of cells, Claim 1 implies that $\mathcal{G}_1$ must induce a connected subgraph of $\mathscr{G}_\mathcal{C}$. Recall that the graph induced by the set $\mathcal{G} = \mathcal{G}_0 \cup \mathcal{G}_1 \cup \mathcal{G}_2$ is also connected by the definition of good cells. Hence, since there are no edges between $\mathcal{G}_0$ and $\mathcal{G}_2$ in the graph of cells, we deduce that $\mathcal{G}_0 \cup \mathcal{G}_1$ also induces a connected graph. Let $\mathcal{G}'$ be the set of good cells at ℓ_∞-distance at most $3r_0$ from cell C. Observe that $\mathcal{G}_0 \cup \mathcal{G}_1 \subseteq \mathcal{G}'$. Moreover, and every cell $C' \in \mathcal{G}' \setminus (\mathcal{G}_0 \cup \mathcal{G}_1)$ must intersect a box B' which is adjacent to a box $B \subseteq R_1$, so C' is adjacent in the graph of cells to some cell in $\mathcal{G}_0 \cup \mathcal{G}_1$. Hence, $\mathcal{G}'$ induces a connected graph as well.

We proceed to bound the diameter (as a graph) of $\mathscr{G}_\mathcal{C}[\mathcal{G}']$. Consider any two cells $C', C'' \in \mathcal{G}'$ and a path of cells $C' = C_1, C_2, \ldots, C_j = C''$ in $\mathcal{G}'$ of minimal length. If two cells in the path intersect the same box, then they are adjacent in the graph of cells, so they must be consecutive in the path by our minimal length assumption. Similarly, we cannot have more than two consecutive cells in the path intersecting one same box. Therefore, we deduce that the number j of cells in the path is at most twice the number of boxes that may be potentially intersected by cells in $\mathcal{G}'$, which is at most

$$2((6r_0 + s)/t + 2)^d \le 2(20d)^d.$$

This gives the desired upper bound on the diameter of the graph.

It only remains to prove Claim 1. In order to do so, suppose our ugly cell C is at distance at most r_0 from exactly i facets of the cube $[0,1]^d$, for some $0 \le i \le d$. By (10) in the proof of Lemma 10, we can find a set $\mathcal{N}'$ of at least $\left(2^{d-i-1}/d\right) \varepsilon^{-1}(1 - 2\alpha\varepsilon^{1/d})$ sparse cells within ℓ_p-distance at most r_0 from C, where $\alpha = \alpha(d)$ is a positive constant.

Pick any box $B \subseteq R_1$. Every cell contained in B is at ℓ_∞-distance (and thus at ℓ_p-distance) at least $1.1r_0$ from C, so it cannot belong to $\mathcal{N}'$. Suppose that all cells contained in B are sparse. Then, by (11), we have at least

$$\left(2^{d-i-1}/d\right) \varepsilon^{-1}(1 - 2\alpha\varepsilon^{1/d}) + \varepsilon^{-1}/(\theta(2d)^{d+1}) \ge \left(2^{d-i-1}/d\right) \varepsilon^{-1}(1 + \varepsilon) \ge \frac{d-i}{d}(1 + \varepsilon)\varepsilon^{-1}$$

sparse cells at ℓ_∞-distance at most $3r_0$ from C. Thus, these sparse cells are within ℓ_p-distance $d(6r_0 + s) \le 7d(r_0 - 2ds)$ from each other and at distance at most $d(3r_0 + s) + r_0 \le (4d+1)r_0$ from

10

i facets of $[0,1]^d$. This contradicts the a.a.s. conclusion of Lemma 7 (with parameters $A = (4d+1)$ and $\ell = 7d$). Therefore, every box B in $\mathcal{R}$ must contain at least one dense cell, which must be also good in view of Lemmas 9 and 11. This proves Claim 1, and completes the proof of the lemma. $\square$

We say that a collection $\mathfrak{P}$ of paths in a graph covers a vertex if the vertex belongs to some path in $\mathfrak{P}$. The next lemma provides us with an appropriate collection of paths $\mathfrak{P}$ that covers the ugly vertices and will be a crucial ingredient in the construction of a rainbow Hamilton cycle.

Lemma 13. *Let $\boldsymbol{X}_{\mathcal{U}}$ denote the set of vertices in $\boldsymbol{X}$ that belong to ugly cells. Given any constant $A > 0$, a.a.s. there is a collection $\mathfrak{P}$ of vertex-disjoint paths in $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$ such that:*

1. *$\mathfrak{P}$ covers all vertices of $\boldsymbol{X}_{\mathcal{U}}$;*

2. *$\mathfrak{P}$ covers at most two vertices inside of each non-ugly cell;*

3. *every vertex in $\boldsymbol{X}$ that is covered by $\mathfrak{P}$ is at graph-distance at most $2(20d)^d$ from some vertex in $\boldsymbol{X}_{\mathcal{U}}$ with respect to the graph $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$;*

4. *for each path $P \in \mathfrak{P}$, there is a good cell C_P such that the two endvertices of P lie in cells that are adjacent (in the graph of cells) to C_P;*

5. *every two different paths in $\mathfrak{P}$ are at ℓ_p-distance at least Ar_0 from each other.*

Proof. Given any component $\mathcal{K}$ of the graph $\mathscr{G}_{\mathcal{C}}[\mathcal{U}]$ induced by the ugly cells, let $\boldsymbol{X}_{\mathcal{K}}$ denote the set of all vertices in $\boldsymbol{X}$ contained in cells of $\mathcal{K}$. Assume that the a.a.s. conclusions of (1), (6) and Lemmas 9, 11 and 12 hold. Then, for each component $\mathcal{K}$ of $\mathscr{G}_{\mathcal{C}}[\mathcal{U}]$, we will deterministically find a path $P = P_{\mathcal{K}}$ in $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$ such that:

1'. *P covers all vertices in $\boldsymbol{X}_{\mathcal{K}}$;*

2'. *P covers at most two vertices inside of each non-ugly cell;*

3'. *P only covers vertices within graph-distance $2(20d)^d$ from $\boldsymbol{X}_{\mathcal{K}}$ with respect to $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$;*

4'. *the two endvertices of P are contained in cells that are both adjacent (in $\mathscr{G}_{\mathcal{C}}$) to the same good cell C_P.*

Set $A' = A + 8(20d)^d$. In view of Lemma 11, for any two components $\mathcal{K}$ and $\mathcal{K}'$ of $\mathscr{G}_{\mathcal{C}}[\mathcal{U}]$, the corresponding sets of vertices $\boldsymbol{X}_{\mathcal{K}}$ and $\boldsymbol{X}_{\mathcal{K}'}$ must be at distance at least $A'r_0$ from each other. Therefore, the paths $P_{\mathcal{K}}$ and $P_{\mathcal{K}'}$ must be at distance at least $A'r_0 - 4(20d)^d \widehat{r}_{\delta \geq 2} \geq Ar_0$ from each other (since $\widehat{r}_{\delta \geq 2} \leq 2r_0$, by (6)). Combining this and properties (1'–4') above, the collection $\mathfrak{P}$ of all such paths $P_{\mathcal{K}}$ will trivially satisfy all the conditions of the lemma. It only remains to prove the existence of these paths.

Pick any connected component $\mathcal{K}$ of $\mathscr{G}_{\mathcal{C}}[\mathcal{U}]$. If $\boldsymbol{X}_{\mathcal{K}}$ is empty, there is nothing to be done. Otherwise, the vertices in $\boldsymbol{X}_{\mathcal{K}}$ induce a clique in $\mathscr{G}(\boldsymbol{X}; r_0) \subseteq \mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$ (by Lemma 9). Moreover, since $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$ is 2-connected (by (1)), we can find four different vertices $u, u' \in \boldsymbol{X}_{\mathcal{K}}$ and $v, v' \in \boldsymbol{X} \setminus \boldsymbol{X}_{\mathcal{K}}$ (unless $|\boldsymbol{X}_{\mathcal{K}}| = 1$, in which case we set $u = u'$) such that uv and $u'v'$ are edges of $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$. Therefore, we can connect all vertices in $\boldsymbol{X}_{\mathcal{K}}$ by a path P_0 with endvertices u and u', and extend this path to a longer path vP_0v' in $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$. Note that all edges in that path have length at most r_0 except for possibly uv and $u'v'$, which have length at most $\widehat{r}_{\delta \geq 2}$. Let D and D' be respectively the cells containing v and v' (possibly $D = D'$). Note that D and D' may be good or bad cells but not ugly since $v, v' \in \boldsymbol{X} \setminus \boldsymbol{X}_{\mathcal{K}}$, so in particular they must be adjacent in the graph

of cells to some good cell. If $D = D'$, then the path $P = vP_0v'$ already satisfies properties (1′–4′), with C_P being any good cell adjacent to $D = D'$. Similarly, if D and D' are adjacent in $\mathscr{G}_{\mathcal{C}}$ and D' is a good cell, then we pick any vertex $v'' \neq v'$ in cell D' (it must exist, since D' is dense), and set $P = v''vP_0v'$. The obtained path also meets our requirements, with C_P being any good cell adjacent to D'. The symmetric case can be dealt with analogously. Thus, we can restrict ourselves to the case in which $D \neq D'$ and moreover D and D' are both bad or non-adjacent. In particular, each one of them is adjacent to some good cell not in $\{D, D'\}$. Under these assumptions, we will extend vP_0v' to a longer path with the desired properties, by adding at most $2(20d)^d - 1$ extra edges of length at most r_0.

Fix any arbitrary ugly cell $C \in \mathcal{K}$, and let $\mathcal{G}'$ be the set of good cells at ℓ_∞-distance at most $3r_0$ from C. By Lemma 9, the union of all cells in $\mathcal{K}$ has ℓ_∞-diameter at most $4d^2s$. Therefore, D and D' must be at ℓ_∞-distance at most $\widehat{r}_{\delta\geq 2} + 4d^2s \leq 3r_0/2$ from cell C (by (6) and our choice of ϵ sufficiently small). From our assumptions, each of D and D' must be adjacent in $\mathscr{G}_{\mathcal{C}}$ to some good cell (not in $\{D, D'\}$) at ℓ_∞-distance at most $3r_0/2 + s + (r_0 - 2ds) \leq 3r_0$ from C. That is, D and D' must be adjacent to some cell in $\mathcal{G}' \setminus \{D, D'\}$. In view of Lemma 12, there exists a path of cells $D', D_1, D_2, \ldots, D_jD$ in the graph $\mathscr{G}_{\mathcal{C}}$ such that cells $D_1, D_2, \ldots, D_j$ belong to $\mathcal{G}' \setminus \{D, D'\}$ and $1 \leq j \leq 2(20d)^d$. Choose one vertex $v_i \in \boldsymbol{X}$ inside each cell D_i ($1 \leq i \leq j$). Then, the path $P = vP_0v'v_1v_2\cdots v_{j-1}$ satisfies our desired properties (with cell $C_P = D_j$). $\qquad\square$

Finally, we provide a (simpler) analogue of Lemma 13 that will be used for the construction of the rainbow perfect matching.

Lemma 14. *Let $\boldsymbol{X}_{\mathcal{U}}$ denote the set of vertices in $\boldsymbol{X}$ that belong to ugly cells. Given any constant $A > 0$, a.a.s. there is a collection $\mathfrak{P}$ of vertex-disjoint paths in $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta\geq 1})$ such that:*

1″. $\mathfrak{P}$ covers all vertices of $\boldsymbol{X}_{\mathcal{U}}$;

2″. every path in $\mathfrak{P}$ has an even number of vertices, and contains at most one vertex not in $\boldsymbol{X}_{\mathcal{U}}$;

3″. every two different paths in $\mathfrak{P}$ are at ℓ_p-distance at least Ar_0 from each other.

Proof (sketch). We proceed similarly as in the proof of Lemma 13, so we just sketch the main differences. For each component $\mathcal{K}$ of $\mathscr{G}_{\mathcal{C}}[\mathcal{U}]$, the vertices in $\boldsymbol{X}_{\mathcal{K}}$ induce a clique in $\mathscr{G}(\boldsymbol{X}; r_0) \subseteq \mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta\geq 1})$ (by Lemma 9). If the number of vertices in $\boldsymbol{X}_{\mathcal{K}}$ is even, we simply connect them all by a path $P_{\mathcal{K}}$. If it is odd, we use the fact that $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta\geq 1})$ is a.a.s. 1-connected (by (1)) to find vertices $u \in \boldsymbol{X}_{\mathcal{K}}$ and $v \in \boldsymbol{X} \setminus \boldsymbol{X}_{\mathcal{K}}$ at distance at most $\widehat{r}_{\delta\geq 1}$ from each other (i.e. uv is an edge of $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta\geq 1})$). Then we pick a path that connects all vertices in $\boldsymbol{X}_{\mathcal{K}}$ and has u as an endpoint, and extend it to $P_{\mathcal{K}}$ by adding edge uv. $\qquad\square$

3 Finding rainbow paths and cycles

In this section, we will build a rainbow spanning graph of $\mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; \widehat{r}_{\delta\geq k})$, for $k \in \{1, 2\}$, consisting of small (path and cycle) components, which will be later used in the construction of a rainbow perfect matching ($k = 1$) or a rainbow Hamilton cycle ($k = 2$). We will proceed in a greedy fashion, and build the rainbow small pieces in a specific order, since the calculations will only work if we reveal the colours of certain edges before others.

At this stage, we expose all the points of $\boldsymbol{X}$ (which determine which cells are ugly, bad and good), and assume henceforth that the a.a.s. conclusions of all statements in Section 2 hold. In particular, all probabilistic statements in the sequel will refer only to the random assignment $\boldsymbol{Z} = (Z_{ij})_{1\leq i<j\leq n}$ of colours to edges. In the case that our final goal is to build a rainbow Hamilton

12

cycle in $\mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; \widehat{r}_{\delta \geq 2})$, we pick a collection $\mathfrak{P}$ of vertex-disjoint paths satisfying properties 1–5 in Lemma 13 with $A = 3$ (or $A = 1000$ for that matter, as we only need to guarantee that paths in $\mathfrak{P}$ are far enough from each other). Otherwise, in order to obtain a rainbow perfect matching in $\mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; \widehat{r}_{\delta \geq 1})$, we pick a collection $\mathfrak{P}$ of vertex-disjoint paths satisfying properties $1'$–$3'$ in Lemma 14. Note that in the latter case, since $\widehat{r}_{\delta \geq 1} \leq \widehat{r}_{\delta \geq 2}$, all the edges in the paths of $\mathfrak{P}$ also belong to $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$ and moreover $\mathfrak{P}$ satisfies conditions 1, 2, 3 and 5 (but not necessarily 4) in Lemma 13. These are in fact all the assumptions on $\mathfrak{P}$ that we will need in this section.

Definition 4. We will refer to the paths in $\mathfrak{P}$ as *ugly paths* (since they cover all the vertices in ugly cells). Let $\boldsymbol{X}'$ be the set of all vertices in ugly paths, and let $\boldsymbol{X}'' = \boldsymbol{X} \setminus \boldsymbol{X}'$.

Note that non-ugly (i.e. good or bad) cells may contain both vertices of $\boldsymbol{X}'$ and vertices of $\boldsymbol{X}''$. However, by property 2 of Lemma 13, each of these cells contains at most two vertices of $\boldsymbol{X}'$. In view of this and since bad cells are sparse and good cells are dense, we conclude that each bad cell contains at most $\varepsilon^3 \log n$ vertices of $\boldsymbol{X}''$, and each good cell contains at least $\varepsilon^3 \log n - 2$ vertices of $\boldsymbol{X}''$. We will first analyze the colours of the edges in the ugly paths, and then "delete" the corresponding vertices (i.e. $\boldsymbol{X}'$) from all bad and good cells in $\mathcal{C}$. A high-level description of our construction can be summarized in the following steps.

1. First we reveal the colours of the edges in the paths of $\mathfrak{P}$ (i.e. the ugly paths). We will show that a.a.s. we do not get any repeated colours, so the ugly paths form a rainbow forest.

2. Next we consider bad cells one by one. In each bad cell, we fix some arbitrary Hamilton path on the vertices of $\boldsymbol{X}''$ contained in that cell (ignore those in $\boldsymbol{X}'$), reveal the colours of all the edges on this path, and then greedily discard edges with previously used colours in the process (i.e. in that same or previous bad cells or in ugly paths). We will see that a.a.s., only at most a few edges are removed from the path, leaving behind a small number of paths. These paths together with the ugly ones are all rainbow by construction.

3. Next we run through good cells one by one (we again restrict our attention to vertices of $\boldsymbol{X}''$ in those cells and ignore those in $\boldsymbol{X}'$). Within a particular good cell, we run through the vertices revealing the colours of the edges to other vertices in the same cell. We will construct a rainbow Hamilton cycle within each good cell, and a.a.s. these cycles will have no colour collisions with each other or any of the previously constructed parts.

4. Finally we hook up all the parts to build either a rainbow perfect matching or a rainbow Hamilton cycle. This requires adding some extra edges that connect vertices in different cells (one of which is always good). Fortunately, we have plenty of these edges available and many unused colours, so we can a.a.s. find the required edges with no colour collisions.

In the remaining of the section, we will focus on steps 1–3, which produce a rainbow collection of vertex-disjoint paths and cycles covering all vertices in $\boldsymbol{X}$. The final step is described in Section 4.

3.1 Ugly paths

We will show that the number of edges in ugly paths is so small that we do not expect colour collisions among them, so in particular the collection of ugly paths is a rainbow forest. Given a graph G, let $E(G)$ denote the edge set of G.

Lemma 15. *Assume that $\boldsymbol{X}$ satisfies all the a.a.s. statements in Section 2, and pick a collection $\mathfrak{P}$ of (ugly) paths satisfying conditions 1,2,3,5 of Lemma 13. Then a.a.s. $\bigcup_{P \in \mathfrak{P}} E(P)$ is rainbow with respect to the random edge colouring $\boldsymbol{Z}$.*

Proof. Recall that $\boldsymbol{X}_{\mathcal{U}}$ is the set of vertices in ugly cells, and $\boldsymbol{X}'$ is the set of vertices in ugly paths. Let $\widehat{\boldsymbol{X}}_{\mathcal{U}}$ be the set of vertices in $\boldsymbol{X}$ within graph-distance $2(20d)^d$ from some vertex in $\boldsymbol{X}_{\mathcal{U}}$ with respect to the graph $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$. By condition 3 of Lemma 13, we have $\boldsymbol{X}' \subseteq \widehat{\boldsymbol{X}}_{\mathcal{U}}$. We wish to obtain a bound on $|\boldsymbol{X}'|$. Note that $|\boldsymbol{X}_{\mathcal{U}}| \leq n^{0.4}$ by Lemmas 5 and 10 (assuming ε is sufficiently small). Moreover, by Lemma 6, every vertex of $\boldsymbol{X}_{\mathcal{U}}$ has at most $O(\log n)$ vertices within graph-distance $\ell = 2(20d)^d$ in $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2}) \subseteq \mathscr{G}(\boldsymbol{X}, r_1)$ (see also (6)). Therefore,

$$|\boldsymbol{X}'| \leq |\widehat{\boldsymbol{X}}_{\mathcal{U}}| = O(n^{0.4} \log n), \tag{12}$$

and the total number of edges in ugly paths is at most $|\boldsymbol{X}'| - 1$. Then, the probability that no colour is repeated across these edges is at least

$$\left(1 - \frac{|\boldsymbol{X}'|}{Kn}\right)^{|\boldsymbol{X}'|} = \exp\left(O\left(\frac{(n^{0.4} \log n)^2}{Kn}\right)\right) = 1 - o(1).$$

So in particular, a.a.s. all the edges in ugly paths receive distinct colours. $\qquad\square$

Therefore, we can expose all the colours of the edges in the ugly paths and save them for future use. In other words, we are entitled to use any of these edges for our rainbow perfect matching or Hamilton cycle, but cannot use any of their colours anywhere else.

3.2 Bad cells

In this section, we restrict our atention to vertices in $\boldsymbol{X}''$ contained in bad cells (and ignore $\boldsymbol{X}'$). Recall that the set of vertices in any cell induces a clique in $\mathscr{G}(\boldsymbol{X}''; r_0)$, so we may use any possible edge between two vertices in a cell. Our goal is to show that a.a.s. every bad cell contains a rainbow spanning linear forest with at most $4/\varepsilon$ many path components. Further, the colours used across all of these forests are distinct and also separate from the colours used on the paths in $\mathfrak{P}$.

Lemma 16. *Under the same assumptions as in Lemma 15, a.a.s. for every bad cell $C \in \mathcal{B}$ the complete graph on the vertices of $\boldsymbol{X}''$ inside of C contains a spanning forest F_C consisting of at most $4/\varepsilon$ many paths (possibly isolated vertices) such that $\bigcup_{C \in \mathcal{B}} E(F_C) \cup \bigcup_{P \in \mathfrak{P}} E(P)$ is rainbow.*

Proof. By Lemma 10, we may assume that there are at most $n^{1-\varepsilon/2}$ many bad cells. We now proceed greedily, moving through all the bad cells one by one. In a given bad cell C, we ignore any vertices which belong to a path of $\mathfrak{P}$ (there are at most 2 such vertices), and connect the remaining vertices by any arbitrary Hamilton path P_C. Then, delete any edge from this path which receives a colour used already in P_C or in a previous bad cell or in a path of $\mathfrak{P}$. Call the resulting graph F_C, which is a spanning linear forest by construction and satisfies the required rainbow conditions. It only remains to prove that we did not delete too many edges from path P_C, so that F_C has at most $4/\varepsilon$ components. Thus far, we have used at most $O(n^{1-\varepsilon/2} \log n)$ colours: at most $O(n^{0.4} \log n)$ on paths of $\mathfrak{P}$ (by (12)) and at most $\varepsilon^3 \log n \leq \log n$ other colours in each of the previous bad cells. So the probability that an edge is deleted from P_C is $O(n^{1-\varepsilon/2} \log n/(Kn)) = O(\log n/n^{\varepsilon/2})$. Since these events are independent, the probability of deleting more than $3/\varepsilon$ edges in a cell is at most

$$\binom{\lceil \log n \rceil}{\lceil 3/\varepsilon \rceil} \cdot \left(O\left(\frac{\log n}{n^{\varepsilon/2}}\right)\right)^{3/\varepsilon} \leq n^{-3/2+o(1)} = o(1/n).$$

Hence, a.a.s. it will not happen for any cell in $\mathcal{B}$ in the greedy process. $\qquad\square$

As we already mentioned, we keep these paths and forbid their colours for future use.

3.3 Good cells

In this section we will prove the following lemma which shows that each good cell contains a rainbow Hamilton cycle which does not use any previously used colours. As in Section 3.2, we restrict our attention to the vertices of $\boldsymbol{X}''$ contained in each cell.

Lemma 17. *Under the same assumptions as in Lemma 15, a.a.s. for every good cell $C \in \mathcal{G}$ the complete graph on the vertices of $\boldsymbol{X}''$ inside of C contains a Hamilton cycle H_C such that $\bigcup_{C \in \mathcal{G}} E(H_C) \cup \bigcup_{C \in \mathcal{B}} E(F_C) \cup \bigcup_{P \in \mathfrak{P}} E(P)$ is rainbow (where $\bigcup_{C \in \mathcal{B}} E(F_C)$ is the linear forest obtained in Lemma 16).*

Proof. We visit all good cells one at a time. Inside of each good cell, we build a rainbow Hamilton cycle restricted to vertices of $\boldsymbol{X}''$ only (and where each pair of vertices is regarded as a potential edge). Good cells are dense, so each cell contains at least $\varepsilon^3 \log n - 2$ vertices of $\boldsymbol{X}''$ (by condition 2 of Lemma 13). More specifically, we do the following within each good cell: we examine each edge one at a time, revealing its colour. We keep an edge if its colour: (i) has not been used on $\bigcup_{C \in \mathcal{B}} E(F_C)$ or $\bigcup_{P \in \mathfrak{P}} E(P)$, (ii) has not been used in a rainbow Hamilton cycle from a previous good cell, and (iii) has not been seen previously in this cell. Otherwise, we delete the edge. At each step, the number of unusable colours for an edge is at most $n + o(n)$ (at most n colours are referred to in (ii) and we may add $o(n)$ to account for (i) and (iii)). Hence, an edge is not present within a good cell with probability at most $(1 + o(1))/K$. Suppose a good cell has $x \geq \varepsilon^3 \log n - 2$ vertices of $\boldsymbol{X}''$ in it. The probability that, when we reveal the edges incident to some fixed vertex, we see that $\lceil x/2 \rceil$ of them are unusable is at most

$$\binom{x}{\lceil x/2 \rceil} \left(\frac{1 + o(1)}{K} \right)^{x/2} \leq 2^x \left(\frac{1 + o(1)}{K} \right)^{x/2} \leq (5/K)^{(\varepsilon^3/2) \log n} = o(1/n)$$

so long as we choose our constants such that $(\varepsilon^3/2) \log(K/5) > 1$. (Note that earlier arguments required ε to be sufficiently small, but K can be chosen large enough with respect to this ε.) Thus, by the union bound over all vertices of $\boldsymbol{X}''$ in dense cells, we conclude that a.a.s. each good cell contains a rainbow Dirac graph (a graph on s vertices with minimum degree at least $s/2$). Such graphs are Hamiltonian [7] and so a.a.s. each good cell contains a rainbow Hamilton cycle. Moreover, by construction, any colour used in such a cycle is not used in any other such cycle, nor in any path or forest constructed in ugly and bad cells. $\qquad\square$

4 Connecting the good, the bad and the ugly

In this section, we complete the proof of Theorem 1. We will first show that a.a.s. $\mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; \widehat{r}_{\delta \geq 2})$ contains a rainbow Hamilton cycle, and then adapt the argument for the corresponding statement about a rainbow perfect matching.

We assume hereafter that $\boldsymbol{X}$ is fixed and satisfies all the a.a.s. statements in Section 2. Moreover, we pick a collection $\mathfrak{P}$ of ugly paths that meets all the requirements in Lemma 13 (with $A = 3$), and assume that the a.a.s. conclusions of Lemmas 15, 16 and 17 hold. This implies that we have a rainbow graph with edge set $\bigcup_{C \in \mathcal{G}} E(H_C) \cup \bigcup_{C \in \mathcal{B}} E(F_C) \cup \bigcup_{P \in \mathfrak{P}} E(P)$ that covers all vertices of $\boldsymbol{X}$ and is made of path and cycle components. Furthermore, all edges in the ugly paths are of length at most $\widehat{r}_{\delta \geq 2}$, whilst the remaining ones are of length at most $ds \leq r_0$ (since they join pairs of vertices inside the same cell), so in particular the above graph is a rainbow subgraph of $\mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; \widehat{r}_{\delta \geq 2})$. We will obtain a rainbow Hamilton cycle by adding a few extra edges (and deleting some others accordingly) that join together the cycle and path components and preserve the rainbow condition.

These new edges will be chosen so that their endpoints lie in different but adjacent cells in the graph of cells $\mathscr{G}_\mathcal{C}$ (so they have length at most r_0), and at least one of their endpoints is in $\boldsymbol{X''}$ (i.e. not in an ugly path). In particular, the colours of these new edges have never been revealed during the greedy process that lead us to Lemmas 15, 16 and 17, and thus remain random.

Pick a spanning tree T of $\mathscr{G}_\mathcal{C}[\mathcal{G}]$ (the large component induced by the good cells in the graph of cells). Then its maximum degree satisfies $\Delta(T) \leq \Delta(\mathscr{G}_\mathcal{C}) = O(1/\varepsilon)$. Now by definition, each bad cell is adjacent (in $\mathscr{G}_\mathcal{C}$) to some cell in $\mathcal{G}$. Thus we may define a tree T' on vertex set $\mathcal{G} \cup \mathcal{B}$ by connecting each bad cell to one of the good cells it is adjacent to in $\mathscr{G}_\mathcal{C}$. Then the cells of $\mathcal{B}$ appear as leaves in T'. Again we have the bound $\Delta(T') = O(1/\varepsilon)$ since T' is a subgraph of $\mathscr{G}_\mathcal{C}$. Finally, consider the collection of ugly paths $\mathfrak{P}$ we picked from Lemma 13. Recall that each path $P \in \mathfrak{P}$ has a corresponding good cell C_P. We define tree T'' by adding each $P \in \mathfrak{P}$ as new a leaf of T' attached to good cell C_P. Since paths in $\mathfrak{P}$ are far apart (by property 5 of Lemma 13), no good cell has more than one such pendant edge attached and so we again have $\Delta(T'') = O(1/\varepsilon)$.

We now use T'' as a template to create the rainbow Hamilton cycle. For any ugly path P, we need to find an edge in the cycle H_{C_P} of the good cell C_P adjacent to P (in T''), remove that edge and attach path P to the endpoints of that edge. This extends H_{C_P} to a larger cycle that covers P. Similarly, for each bad cell $C \in \mathcal{B}$, we will attach each of the paths of F_C to the cycle $H_{C'}$ of the good cell C' that is adjacent to C in T''. We refer to these paths as *bad paths*. Note that in this operation each bad path of F_C is attached to the endpoints of a different edge of $H_{C'}$, which is immediately deleted. This uses at most $4/\varepsilon$ edges of $H_{C'}$ (by Lemma 16). Finally, we will attach the cycles within adjacent (in T'') good cells to each other. This operation requires deleting one edge in each cycle and joining the endpoints of the resulting paths together so that we obtain one longer cycle. Iterating this procedure for all edges of T'' produces a Hamilton cycle in $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta \geq 2})$ provided that we have enough edges in the cycles of the good cells.

It remains to show that we can a.a.s. connect all the pieces together in the manner described in the above paragraph and at the same time satisfy the rainbow condition. The potentially most delicate steps in our construction are when we hook up ugly/bad paths to cycles inside good cells. Consider such an ugly/bad path P with endpoints u, v contained in cell C, and let C' be the good cell adjacent to C in T''. We will show that we can find some edge $u'v'$ in the cycle $H_{C'}$ of the good cell such that the new edges uu' and vv' connecting path P to cycle $H_{C'}$ are assigned previously unused colours. We call $u'v'$ the *hook edge* for the ugly/bad path P. Then, the hook edge $u'v'$ can be removed and replaced by the ugly/bad path together with the new edges uu' and vv'. Recall that, since cell C' is good, it contains $x \geq \varepsilon^3 \log n - 2$ vertices of $\boldsymbol{X''}$, and therefore cycle $H_{C'}$ has x potential hook edges. To ensure independence, we only consider every second edge in the cycle. Moreover, at most $O(1/\varepsilon^2)$ edges are perhaps already used as hook edges to attach some other ugly/bad paths to cycle $H_{C'}$, (since the degree of T'' is $O(1/\varepsilon)$ and each bad cell contains at most $4/\varepsilon$ bad paths). In any case, there are at least $\varepsilon^3 \log n/2 - O(1/\varepsilon^2) \geq \varepsilon^3 \log n/3$ edges in the cycle we can still use. The probability that no such edge has the property we seek (i.e. that the edges connecting the endpoints of the edge to the endpoints of the ugly/bad path are not both of unused colours) is at most

$$\left(1 - \left(\frac{K - 1 + o(1)}{K} \right)^2 \right)^{\varepsilon^3 \log n/3} \leq \left(\frac{2}{K} \right)^{\varepsilon^3 \log n/3} = \exp\left(-\frac{\varepsilon^3 \log(K/2)}{3} \log n \right) = o\left(\frac{1}{n} \right)$$

so long as we ensure that K is large enough so that $\log(K/2)\varepsilon^3/3 > 1$. Recall that we have $o(n)$ ugly/bad paths in total. Hence, a union bound over all the ugly/bad paths implies that we will a.a.s succeed at finding a hook edge for each such path. The argument to merge cycles H_C and $H_{C'}$ of two good cells C and C' together is similar. This time we need to find two hook edges uv in H_C and $u'v'$

in $H_{C'}$ such that the new edges uu' and vv' receive previously unused colours. We have $\Omega_\varepsilon(\log^2 n)$ choices of pairs of hook edges and so the failure probability is at most $\left(\frac{2}{K}\right)^{\Omega_\varepsilon(\log^2 n)} = o\left(\frac{1}{n}\right)$. This completes the proof of the first part of Theorem 1.

Finally, we will show that a.a.s. $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta\geq 1})$ contains a perfect matching (for even n). The argument will reuse most of the ideas in the construction of a rainbow Hamilton cycle earlier in this section, so we will only sketch the main differences. This time, we will assume that $\mathfrak{P}$ satisfies conditions $(1''\text{--}3'')$ from Lemma 14, and deterministically build a perfect matching. Recall that our new assumptions on $\mathfrak{P}$ also imply that conditions 1,2,3,5 (but not necessarily 4) in Lemma 13 are true, so we are entitled to assume all the claims in Section 3. Additionally, the ugly paths in $\mathfrak{P}$ have an even number of vertices and (by construction) only use edges of length at most $\widehat{r}_{\delta\geq 1}$.

We proceed as before but using T' instead of T'' as a template and thus ignoring ugly paths. That is, we hook up each bad path to the cycles in the corresponding good cell and also the cycles within any two good cells that are adjacent in T'. However, we do not attach ugly paths to anything (in fact, we may not be able to do so, since $\mathfrak{P}$ may not satisfy property 4 in Lemma 13). This procedure a.a.s. creates a big rainbow cycle H in $\mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; r_0) \subseteq \mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; \widehat{r}_{\delta\geq 1})$ that covers all vertices except for those in ugly paths, and moreover $H \cup \bigcup_{P\in\mathfrak{P}} P$ is rainbow. If we restrict asymptotics to even n, then cycle H has even length since all paths of $\mathfrak{P}$ have an even number of vertices. By removing alternating edges adequately from H and the ugly paths, we obtain a rainbow perfect matching in $\mathscr{G}(\boldsymbol{X}; \widehat{r}_{\delta\geq 1})$ as desired. This implies the second statement of Theorem 1, and finishes the proof.

5   Case $p = 1$

In this section, we consider the case $p = 1$, and sketch how to adapt the argument of Theorem 1 in order to obtain (3). Recall that for $p = 1$, it is not known whether or not (1) holds (this is due to some technical parts of the argument in [15] that break down for $p = 1$). However, we can still claim that, for $p = 1$ and any constant $\eta > 0$,

$$(1 - \eta)r_0 \leq \widehat{r}_{1\text{-conn}} \leq \widehat{r}_{2\text{-conn}} \leq (1 + \eta)r_1 \qquad \text{a.a.s.}$$

This follows from the fact that (1) and (6) are valid for every $p > 1$ and by continuity of $\theta = \theta(p)$ at $p \geq 1$. Then we can pick η sufficiently small, and replace r_0, $\widehat{r}_{\delta\geq 1}$, $\widehat{r}_{\delta\geq 2}$ and r_1 by $(1 - \eta)r_0$, $\widehat{r}_{1\text{-conn}}$, $\widehat{r}_{2\text{-conn}}$ and $(1 + \eta)r_1$, respectively, in the proof of Theorem 1. The argument is still valid with virtually no adaptation, and yields (3).

6   Open questions

In this paper, we showed that a.a.s. the first edge in the edge-coloured random geometric graph process $\left(\mathscr{G}(\boldsymbol{X}; \boldsymbol{Z}; r)\right)_{r\geq 0}$ that gives minimum degree at least 2 (or 1) and such that at least n (or $n/2$) colours have appeared also creates a rainbow Hamilton cycle (or perfect matching), provided that the number of colours is at least $c = \lceil Kn \rceil$, where $K = K(d) > 0$ is a sufficiently large constant. This condition on c ensures that at least n (or $n/2$) colours have appeared long before the minimum degree becomes 2 (or 1).

Thus the most intriguing open question is to prove that these statements hold for any number of colours, c. Of course for Hamilton cycles, we must have $c \geq n$ and for perfect matchings, we must have $c \geq n/2$. This problem may be particularly interesting in the case of perfect matchings when $d = 2$ and $c = n/2$. In this case, the first appearance of a perfect matching and of $n/2$

distinct colours occurs once $(1+o(1))\frac{n}{2}\log n$ many edges have arrived. The case of Hamilton cycles when $d = 4$ and $c = n$ is interesting for the analogous reason. The first results on packing rainbow Hamilton cycles (that is, finding a collection of edge disjoint Hamilton cycles) in $G_c(n,p)$ were recently obtained in [9]. We believe that using some of the ideas in [14], it should be relatively easy to extend our argument to find a constant number of edge-disjoint rainbow Hamilton cycles and perfect matchings in $\mathscr{G}(\boldsymbol{X};\boldsymbol{Z};r)$ as well. It would be interesting to consider further extensions in which the number of rainbow Hamilton cycles or perfect matchings in the packing grows to infinity as a function of n.

References

[1] D. Bal, P. Bennett, C. Cooper, A. Frieze, and P. Prałat. Rainbow arborescence in random digraphs. to appear in *Journal of Graph Theory*, 2016.

[2] D. Bal, P. Bennett, A. Frieze, and P. Prałat. Power of κ choices and rainbow spanning trees in random graphs. *Electron. J. Combin.*, 22(1):Paper 1.29, 22, 2015.

[3] D. Bal and A. Frieze. Rainbow matchings and hamilton cycles in random graphs. `arXiv:1311.6423`, 2013.

[4] J. Balogh, B. Bollobás, M. Krivelevich, T. Müller, and M. Walters. Hamilton cycles in random geometric graphs. *Ann. Appl. Probab.*, 21(3):1053–1072, 2011.

[5] C. Cooper and A. Frieze. Multi-coloured Hamilton cycles in random edge-coloured graphs. *Combin. Probab. Comput.*, 11(2):129–133, 2002.

[6] J. Díaz, D. Mitsche, and X. Pérez. Sharp threshold for Hamiltonicity of random geometric graphs. *SIAM J. Discrete Math.*, 21(1):57–65 (electronic), 2007.

[7] G. A. Dirac. Some theorems on abstract graphs. *Proc. London Math. Soc. (3)*, 2:69–81, 1952.

[8] A. Ferber and M. Krivelevich. Rainbow hamilton cycles in random graphs and hypergraphs. `arXiv:1506.02929`, 2015.

[9] A. Ferber, G. Kronenberg, F. Mousset, and C. Shikhelman. Packing a randomly edge-colored random graph with rainbow k-outs. `arXiv:1410.1803`, 2014.

[10] A. Ferber, R. Nenadov, and U. Peter. Universality of random graphs and rainbow embedding. `arXiv:1311.7063`, 2013.

[11] A. Frieze and P.-S. Loh. Rainbow Hamilton cycles in random graphs. *Random Structures Algorithms*, 44(3):328–354, 2014.

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

[13] S. Janson and N. Wormald. Rainbow Hamilton cycles in random regular graphs. *Random Structures Algorithms*, 30(1-2):35–49, 2007.

[14] T. Müller, X. Pérez-Giménez, and N. Wormald. Disjoint Hamilton cycles in the random geometric graph. *J. Graph Theory*, 68(4):299–322, 2011.

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

[16] M. D. Penrose. *Random geometric graphs*, volume 5 of *Oxford Studies in Probability*. Oxford University Press, Oxford, 2003.

[17] M. Walters. Random geometric graphs. In *Surveys in combinatorics 2011*, volume 392 of *London Math. Soc. Lecture Note Ser.*, pages 365–401. Cambridge Univ. Press, Cambridge, 2011.