Rainbow Hamilton Cycles in Random Geometric Graphs

Alan Frieze*and Xavier Pérez-Giménez†

July 8, 2021

Abstract

Let $X_1, X_2, \ldots, X_n$ be chosen independently and uniformly at random from the unit d-dimensional cube $[0,1]^d$. Let r be given and let $\mathcal{X} = \{X_1, X_2, \ldots, X_n\}$. The random geometric graph $G = G_{\mathcal{X},r}$ has vertex set $\mathcal{X}$ and an edge X_iX_j whenever $\|X_i - X_j\| \leq r$. We show that if each edge of G is colored independently from one of $n + o(n)$ colors and r has the smallest value such that G has minimum degree at least two, then G contains a rainbow Hamilton cycle a.a.s.

1 Introduction

Given a graph $G = (V, E)$ plus an edge coloring $c : E \to [q]$, we say that $S \subseteq E$ is *rainbow colored* if no two edges of S have the same color. There has been a substantial amount of research on the question as to when does an edge colored graph contain a rainbow Hamilton cycle. The early research was done in the context of the complete graph K_n when restrictions were placed on the colorings. In this paper we deal with the case where we have a random geometric graph and the edges are colored randomly.

In the case of the Erdős-Rényi random graph $G_{n,m}$, Cooper and Frieze [5] proved that if $m \geq 21n \log n$ and each edge of $G_{n,m}$ is randomly given one of at least $q \geq 21n$ random colors then *asymptotically almost surely* (a.a.s.) there is a rainbow Hamilton cycle. Frieze and Loh [10] improved this result to show that if $m \geq \frac{1}{2}(n+o(n)) \log n$ and $q \geq (1+o(1))n$ then a.a.s. there is a rainbow Hamilton cycle. This was further improved by Ferber and Krivelevich [8] to $m = n(\log n + \log\log n + \omega)/2$ and $q \geq (1 + o(1))n$, where $\omega \to \infty$ with n. This is best possible in terms of the number of edges. The case $q = n$ was considered by Bal and Frieze [3]. They showed that $O(n \log n)$ random edges suffice.

Let $X_1, X_2, \ldots, X_n$ be chosen independently and uniformly at random from the unit d-dimensional cube $[0,1]^d$ where $d \geq 2$ is constant. Let r be given and let $\mathcal{X} = \{X_1, X_2, \ldots, X_n\}$. The random geometric graph $G_{\mathcal{X},r}$ has vertex set $[n]$ and an edge ij for each pair $i, j \in [n]$ $(i \neq j)$ satisfying $\|X_i - X_j\| \leq r$. Here $\|\cdot\|$ refers to an arbitrary ℓ_p-norm, where $1 < p \leq \infty$. We define the *length* of an edge ij to be $\|X_i - X_j\|$. Throughout the paper we tacitly assume that the points $X_1, \ldots, X_n$ are all different, which happens almost surely, and identify the vertex set with $\mathcal{X}$. (We will use the terms point and vertex interchangeably when referring to an element of $\mathcal{X}$.) Suppose now that

*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA, USA, 15213. Research supported in part by NSF grant DMS1661063.

†Department of Mathematics, University of Nebraska-Lincoln, Lincoln NE, USA, 68588. Research supported in part by Simons Foundation Grant #587019.

each edge of $G_{\mathcal{X},r}$ is given a random color from $[q]$. We call the resulting edge-colored graph $G_{\mathcal{X},r,q}$. Bal, Bennett, Pérez-Giménez and Pralat [2] considered the problem of the existence of a rainbow Hamilton cycle $G_{\mathcal{X},r,q}$. They showed that for r at the threshold for Hamiltonicity, $q = O(n)$ random colors are sufficient to have a rainbow Hamilton cycle a.a.s. The aim of this paper is to show that $q = n + o(n)$ colors suffice in this context.

Let $\theta = \theta(d,p)$ denote the volume of the unit ℓ_p-ball in d dimensions, and let

$$r^d = \frac{(2/d)\log n + (4-d-2/d)\log\log n + f}{2^{2-d}\theta n}, \tag{1}$$

for some $f = f(n)$.

Theorem 1. *Let r be as in (1) for some $f \to \infty$. Let $\eta > 0$ be an arbitrarily small constant and $q = \lceil (1+\eta)n \rceil$. Then a.a.s. $G_{\mathcal{X},r,q}$ contains a rainbow Hamilton cycle.*

We actually prove a stronger hitting-time result, for which we need some definitions. For $n \geq 3$, let

$$\hat{r} = \inf\left\{r \geq 0 : \ G_{\mathcal{X},r} \text{ has minimum degree at least } 2\right\}.$$

Clearly, $\hat{r}$ is a deterministic continuous function of the random set of points $\mathcal{X}$ and thus a random variable. The random graph $G_{\mathcal{X},\hat{r}}$ can be obtained by taking an empty graph on vertex set $\mathcal{X}$ and adding edges one by one in increasing order of lengths until the minimum degree becomes 2 or more. (If two or more edges have the same length, they should be added all at once to the graph, but this does not happen almost surely.) In particular, $G_{\mathcal{X},\hat{r}}$ has minimum degree at least 2, so the infimum in the definition of $\hat{r}$ can be safely replaced by a minimum. The asymptotic distribution of $\hat{r}$ is well known, and can be derived from Theorem 8.4 in [14]. Indeed, with r parametrized in terms of f as in (1), we have

$$\lim_{n\to\infty} \mathbb{P}(\hat{r} \leq r) = \begin{cases} 0 & f \to -\infty \\ F(\alpha) & f \to \alpha \in \mathbb{R} \\ 1 & f \to \infty, \end{cases} \tag{2}$$

where $F(\alpha)$ is a continuous distribution function. (An explicit description of $F(\alpha)$ can be found, e.g., in Corollaries 3 and 4 of [2].) We now consider the edge-colored version $G_{\mathcal{X},\hat{r},q}$ of $G_{\mathcal{X},\hat{r}}$. Our main result asserts that, if we start with the empty graph on vertex set $\mathcal{X}$ and we add randomly colored edges one by one in increasing order of lengths, then (provided that we use sufficiently many colors) a.a.s. we obtain a rainbow Hamilton cycle as soon as the minimum degree becomes at least 2.

Theorem 2. *Let $\eta > 0$ be any fixed constant and $q = \lceil (1+\eta)n \rceil$. Then $G_{\mathcal{X},\hat{r},q}$ has a rainbow Hamilton cycle a.a.s.*

Combining this and (2) immediately yields Theorem 1, so we will devote the remainder of the paper to the proof of Theorem 2.

Proof sketch. We partition $[0,1]^d$ into small cubic cells of side around εr. These cells are classified into types acording to the number of points and color repetitions they contain. Then the set of cells is endowed with a graph structure by connecting every pair of cells at distance slightly less than r. (Note that similar constructions have been fruitfully used in [2, 4, 6, 13].) In

Section 2, we derive some basic properties of this graph of cells. Then we use a variation of Pósa's rotation-extension argument to show that most cells contain a spanning family of 'not too many' rainbow paths that avoid certain forbidden colors. This type of argument has been widely applied in the study of Hamilton cycles in many other families of random graphs (e.g. the Erdős-Rényi random graph $G_{n,m}$ [12], random regular graphs [7], preferential attachment graphs [11]), but so far not before in the context of random geometric graphs. In Section 3, we introduce and analyze a greedy procedure (`Build`), which a.a.s. constructs a rainbow Hamilton cycle in $G_{\mathcal{X},\hat{r},q}$, based on the structure and properties of the graph of cells. An unusual and interesting feature of this procedure is that it sometimes introduces errors (i.e. color repetitions) which are recursively fixed by another procedure (`Problem-fix`), which may in turn trigger further errors. We show that typically these errors do not accumulate past a certain bound and the algorithm succeeds.

We finish the discussion by observing that our results are best possible in terms of the number of permitted colors q. Indeed, for dimension $d \in \{2,3\}$, if we allow only $q = n$ colors, then a standard coupon collector argument shows that a.a.s. some colors are still missing on $G_{\mathcal{X},\hat{r},q}$. More precisely, let

$$r^* = \inf\{r \geq 0 : \text{ all } n \text{ colors appear on } G_{\mathcal{X},r,n}\}.$$

Then a.a.s. $r^* \sim \sqrt[d]{2\log n/(\theta n)}$, and thus

$$\begin{cases} r^* \geq (3/2 + o(1))\hat{r} & \text{for } d \in \{2,3\} \\ r^* \sim \hat{r} & \text{for } d = 4 \\ r^* \leq (5/8 + o(1))\hat{r} & \text{for } d \geq 5. \end{cases}$$

Similarly, if we consider a slight variation of the model in which the points of $\mathcal{X}$ are placed on the torus $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$ instead of the cube $[0,1]^d$, then with the analogous definitions of $G_{\mathcal{X},\hat{r},q}$, $\hat{r}$ and r^*, we have that a.a.s. $\hat{r} \sim \sqrt[d]{\log n/(\theta n)}$ (see Theorem 8.3 in [14]) and therefore $r^* \sim \sqrt[d]{2} \cdot \hat{r}$. The difference in $\hat{r}$ between the two models is explained by the presence of vertices of degree less than 2 near the boundaries of $[0,1]^d$. In either case, it is conceivable that with exactly $q = n$ colors, as soon as the minimum degree is at least 2 and we see all the colors, we have a rainbow Hamilton cycle a.a.s. We state this as a conjecture for either the cube $[0,1]^d$ or the torus $\mathbb{T}^d$ models. Let $t_1 \vee t_2 := \max\{t_1, t_2\}$.

Conjecture 3. *$G_{\mathcal{X},\hat{r}\vee r^*,n}$ has a rainbow Hamilton cycle a.a.s.*

We also include a similar statement conditional on the event that $G_{\mathcal{X},\hat{r},n}$ has all n colors (which is a rare event for the cube model and $d \in \{2,3\}$ or for the torus model and any $d \geq 2$).

Conjecture 4. *Conditional upon $r^* \leq \hat{r}$, $G_{\mathcal{X},\hat{r},n}$ has a rainbow Hamilton cycle a.a.s.*

While this paper only discusses rainbow Hamilton cycles, analogous questions can be asked about rainbow perfect matchings with $q = n/2$ colors. Let

$$\hat{r}_1 = \inf\{r \geq 0 : \ G_{\mathcal{X},r} \text{ has minimum degree at least } 1\}.$$

and (for even n)

$$r_1^* = \inf\left\{r \geq 0 : \text{ all } n/2 \text{ colors appear on } G_{\mathcal{X},r,n/2}\right\}.$$

Conjecture 5. *For even n, $G_{\mathcal{X},\hat{r}_1\vee r_1^*,n/2}$ has a rainbow perfect matching a.a.s.*

Conjecture 6. *For even n and conditional upon $r_1^* \leq \hat{r}_1$, $G_{\mathcal{X},\hat{r}_1,n/2}$ has a rainbow perfect matching a.a.s.*

2 Notation and structural properties

Throughout the paper, $d \geq 2$ and an ℓ_p-norm $\|\cdot\|$ on $\mathbb{R}^d$ $(1 < p \leq \infty)$ are fixed. Let $\eta > 0$ be an arbitrary constant (which we will assume to be sufficiently small to satisfy all the requirements in the argument) and set

$$q = \lceil (1+\eta)n \rceil.$$

Let $Q = [q]$ denote the set of available colors. Recall that $G_{\mathcal{X},\hat{r},q}$ is obtained by assigning to each edge of $G_{\mathcal{X},\hat{r}}$ a random color in Q chosen uniformly at random and independently from all other choices.

Let $\varepsilon > 0$ be a constant which is assumed to be sufficiently small given our choices of η and d. We use the standard $o()$, $\omega()$, $O()$, $\Theta()$ 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 $a_n = \Theta(1)$, $a_n = \Omega(1)$ 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 η 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.

Henceforth, let r be defined as in (1) for some arbitrary function $f \to -\infty$, $f = o(\log\log n)$. From (2), we have

$$r \leq \hat{r} \qquad \text{and} \qquad r \sim \hat{r} \qquad \text{a.a.s.}$$

Following the construction in [2], we divide $[0,1]^d$ into a set $\mathcal{C}$ of $N = \lceil (\varepsilon r)^{-1} \rceil^d$ d-dimensional cubic cells of side $s = 1/\lceil (\varepsilon r)^{-1} \rceil \sim \varepsilon r$. We remark that

$$N \sim \frac{d\theta n}{2^{d-1}\varepsilon^d \log n}.$$

For sake of simplicity, assume that every point in $\mathcal{X}$ is contained in one single cell in $\mathcal{C}$ (which occurs almost surely, since cell boundaries have measure 0). The *graph of cells* $\mathcal{G}_\mathcal{C}$ is a graph with vertex set $\mathcal{C}$ where two cells are adjacent in $\mathcal{G}_\mathcal{C}$ if their centres are at ℓ_p-distance at most $r - 2ds$. (Here we assume that $2ds$ is much smaller than r by our choice of ε.) By the triangle inequality, any two different points $X_i, X_j \in \mathcal{X}$ which are contained in the same cell or in two cells that are adjacent in $\mathcal{G}_\mathcal{C}$ satisfy $\|X_i - X_j\| \leq r$, and therefore X_iX_j must be an edge of $G_{\mathcal{X},r}$ and a.a.s. an edge of $G_{\mathcal{X},\hat{r}}$. A cell C is *dense* if $|C \cap \mathcal{X}| \geq \varepsilon^3 \log n$. Otherwise it is *sparse*. The set of dense cells is denoted by $\mathcal{D}$, and $\mathcal{G}_\mathcal{D}$ is the subgraph of $\mathcal{G}_\mathcal{C}$ induced by the good cells. The paper [2] shows that a.a.s.

$$\mathcal{G}_\mathcal{D} \text{ contains a unique giant component } \Gamma_0 \text{ containing } N - o(N) \text{ cells.} \tag{3}$$

(The proof in [2] is adapted from an earlier article [13] that uses a less restrictive definition of dense cell.) The cells in Γ_0 are called *good*. A cell that is not good, but is adjacent to a cell in Γ_0 is called *bad*. The remaining cells are called *ugly*. Note that bad cells are sparse by definition, but ugly cells may be dense or sparse. The following two lemmas describe properties that occur a.a.s., and their proofs are in [2]:

Lemma 7. *A.a.s.*

P1 $|C \cap \mathcal{X}| \leq \log n$ *for all* $C \in \mathcal{C}$ *(cf. Lemma 5 in [2]).*

P2 *There are at most* $n^{1-\varepsilon/2}$ *bad cells (cf. Lemma 10 in [2]).*

P3 *There are at most* $n^{O(\varepsilon^{1/d})}$ *ugly cells (cf. Lemma 10 in [2]).*

P4 *The maximum degree in* $G_{\mathcal{X},\hat{r},q}$ *is at most* $O(\log n)$ *(cf. Lemma 6 in [2]).*

Lemma 8 (Lemma 13 in [2])**.** *Let* $\mathcal{X}_\mathcal{U}$ *denote the set of points in ugly cells. Then a.a.s. there is a collection of paths* $\mathcal{P}$ *such that*

Q1 $\mathcal{P}$ *covers* $\mathcal{X}_\mathcal{U}$*.*

Q2 $\mathcal{P}$ *covers at most two vertices inside any non-ugly cell.*

Q3 *Every vertex in* $\mathcal{X}$ *that is covered by* $\mathcal{P}$ *is at graph-distance at most* $2(20d)^d$ *from some vertex in* $X_\mathcal{U}$ *with respect to the graph* $G_{\mathcal{X},\hat{r}}$*.*

Q4 *For each path* $P \in \mathcal{P}$*, there is a good cell* C_P *such that the two endvertices of* P *lie in cells that are adjacent in* $\mathcal{G}_\mathcal{C}$ *to* C_P*;*

Q5 *Every pair of distinct paths in* $\mathcal{P}$ *are at* ℓ_p*-distance at least* Ar *from each other,* $A > 0$ *arbitrary.*

For a region $S \subseteq [0,1]^d$, we let $V(S) = S \cap \mathcal{X}$ and $E(S) = \binom{V(S)}{2} \cap E(G_{\mathcal{X},\hat{r}})$. (Note that this definition will be slightly modified later in (6).) A color repetition in S is a pair of edges in $E(S)$ that receive the same color in $G_{\mathcal{X},\hat{r},q}$. A cell C is *rainbow* if $E(C)$ is rainbow: that is, C has no color repetitions.

We now prove some lemmas related to the colorings of cells.

Lemma 9. *For any constant* $A > 0$*, the following hold a.a.s.*

(a) There are at most $\log^4 n$ *non-rainbow cells.*

(b) No d-dimensional cube of side at most Ar *obtained as a union of cells in* $\mathcal{C}$ *contains* 2 *color repetitions.*

(c) There are no two non-rainbow cells within ℓ_p*-distance* Ar *of each other.*

(d) There are no non-rainbow cells within ℓ_p*-distance* Ar *of the boundary of* $[0,1]^d$*.*

(e) There are no non-rainbow cells within ℓ_p*-distance* Ar *of any cell that is not good.*

Remark. In particular (with $A > 1$), a.a.s. every non-rainbow cell is good and is only adjacent in $\mathcal{G}_\mathcal{C}$ to good rainbow cells. Moreover, a.a.s. every cell contains at most one repetition.

Proof. All the statements in the lemma follow from simple first moment arguments.

(a) For a fixed cell C,

$$\mathbb{P}(C \text{ is not rainbow} \mid \mathbf{P1}) \leq (1+\eta)n\mathbb{P}\left(\mathrm{Bin}\left(\log^2 n, \frac{1}{(1+\eta)n}\right) \geq 2\right) = O\left(\frac{\log^4 n}{n}\right). \quad (4)$$

Explanation: we choose a color c. Then the number of edges of color c in cell C is dominated by the stated binomial.

We then have, by the Markov inequality that

$$\mathbb{P}(\neg\ (a) \mid \mathbf{P1}) \leq \frac{\mathbb{E}(\text{number of rainbow cells} \mid \mathbf{P1})}{\log^4 n} = O\left(\frac{N \log^4 n}{n \log^4 n}\right) = O_\varepsilon\left(\frac{1}{\log n}\right).$$

(b)–(c) Let $\mathcal{Q}$ be the set of regions $Q \subseteq [0,1]^d$ such that Q is a d-dimensional cube of side at most $(A+1)r$ obtained as a union of cells in $\mathcal{C}$. Note that $|\mathcal{Q}| = O_\varepsilon(N) = O_\varepsilon(n/\log n)$. Moreover, assuming $\mathbf{P1}$, $|V(Q)| = O_\varepsilon(\log n)$ and thus $|E(Q)| = O_\varepsilon(\log^2 n)$ for each $Q \in \mathcal{Q}$. Therefore, we have that

$$\begin{aligned}\mathbb{P}(\text{some } Q \in \mathcal{Q} \text{ has 2 color repetitions} \mid \mathbf{P1}) &\leq |\mathcal{Q}|(1+\eta)n\mathbb{P}\left(\text{Bin}\left(O_\varepsilon(\log^2 n), \frac{1}{(1+\eta)n}\right) \geq 3\right) + \\ &\quad + |\mathcal{Q}|(1+\eta)^2 n^2 \mathbb{P}\left(\text{Bin}\left(O_\varepsilon(\log^2 n), \frac{1}{(1+\eta)n}\right) \geq 2\right)^2 \\ &= O_\varepsilon\left(\frac{\log^7 n}{n}\right). \qquad (5)\end{aligned}$$

Explanation: The first term is an upper bound on the expected number of triples of edges of the same color and the second term accounts for double pairs of edges with the same color.

Clearly, (5) implies (b). It also implies (c) since any two cells within ℓ_p-distance Ar must be contained in one $Q \in \mathcal{Q}$.

(d)–(e) Assuming $\mathbf{P2}$ and $\mathbf{P3}$, there are at most $2n^{1-\varepsilon/2}$ cells that are not good. Since the graph $\mathcal{G}_\mathcal{C}$ has maximum degree $O_\varepsilon(1)$, there are at most $O_\varepsilon(n^{1-\varepsilon/2})$ cells within ℓ_p-distance Ar of some cell that is not good. Moreover, there are $O_\varepsilon(1/r^{d-1})$ cells within ℓ_p-distance Ar of the boundary of $[0,1]^d$. Arguing as in (a), we have that

$$\begin{aligned}\mathbb{P}(\neg\ (d) \text{ or } \neg(e) \mid \mathbf{P1,P2,P3}) &= O_\varepsilon\left(n^{1-\varepsilon/2} + 1/r^{d-1}\right)(1+\eta)n\mathbb{P}\left(\text{Bin}\left(\log^2 n, \frac{1}{(1+\eta)n}\right) \geq 2\right) \\ &= O_\varepsilon\left(\frac{\log^4 n}{n^{\varepsilon/2}} + \frac{\log^4 n}{nr^{d-1}}\right) = o(1). \qquad \square\end{aligned}$$

We remove all the non-rainbow cells (which must be good a.a.s.) from the giant component Γ_0 of good cells, and obtain Γ_1. We argue next that a.a.s. Γ_1 remains connected.

Lemma 10. *The graph of rainbow good cells Γ_1 is a.a.s. connected.*

Proof. Let C, C' be any two cells in Γ_1 (i.e. rainbow and good). Since Γ_0 is a.a.s. connected, there must be a C, C'-path $P = (C = C_1, C_2, \ldots, C_m = C')$ of good cells. We want to show that, after deleting all the non-rainbow good cells, there is still a C, C'-path in Γ_1. Suppose that an interior cell C_i of path P (i.e. $1 < i < m$) is non-rainbow. Let S be the union of all cells different from C_i that are within ℓ_p-distance $2r$ of C_i. Assuming that the a.a.s. statements in Lemma 9 hold, S is away from the boundary of $[0,1]^d$ and all the cells contained in S are rainbow and good. Clearly, S is topologically connected and hence the cells in S induce a connected subgraph of Γ_1. By construction, $C_{i-1}, C_{i+1} \subseteq S$, and hence we can find a C_{i-1}, C_{i+1}-path Q that uses only cells in S and thus rainbow and good. Hence, replacing the subpath C_{i-1}, C_i, C_{i+1} in P by Q, we obtain another C, C'-path that avoids C_i. Iterating this argument for all non-rainbow cells in P, we obtain a C, C'-path in Γ_1. $\square$

We now choose a collection of paths $\mathcal{P}$ in $G_{\mathcal{X},\hat{r}}$ satisfying **Q1**–**Q5**, which must exist a.a.s. in view of Lemma 8. (If there are multiple choices for $\mathcal{P}$, pick one arbitrarily.) We call paths in $\mathcal{P}$ *ugly*. Let $V(\mathcal{P})$ be the set of points in $\mathcal{X}$ covered by ugly paths, and let $E'(\mathcal{P})$ be the set of edges of $G_{\mathcal{X},\hat{r}}$ that are incident with some vertex in $V(\mathcal{P})$.

Lemma 11. *A.a.s. $E'(\mathcal{P})$ has $n^{O(\varepsilon^{1/d})}$ edges, and it is rainbow colored in $G_{\mathcal{X},\hat{r},q}$.*

Proof. Properties **P1**, **P3**, **P4** and **Q3** immediately imply that $|E'(\mathcal{P})| = n^{O(\varepsilon^{1/d})}$. Conditional upon this, the probability that $E'(\mathcal{P})$ has a color repetition can be bounded by $\binom{n^{O(\varepsilon^{1/d})}}{2} \times n^{-1} = o(1)$. □

In the sequel, we remove all vertices in $V(\mathcal{P})$ from the cells, without changing the original cell classification into good, bad and ugly. (The argument will first attempt to build a rainbow cycle H through $\mathcal{X} \setminus V(\mathcal{P})$ and then insert the paths in $\mathcal{P}$ into H.) After this operation, **Q1** and **Q2** imply that ugly cells will no longer contain any points from $\mathcal{X}$ (since they were all on ugly paths and got removed), while each good cell will contain at least $\varepsilon^3 \log n - 2$ points from $\mathcal{X}$ (since at most 2 points were removed). Note that a.a.s. non-rainbow cells are not affected by this operation, since they do not contain points in $V(\mathcal{P})$ by Lemma 9(e) (with $A > 2(20d)^d$) and **Q3**. For convenience, for each cell C, we redefine $V(C)$ and $E(C)$ to denote the sets of vertices and edges of $G_{\mathcal{X},\hat{r}} - V(\mathcal{P})$ contained in C. That is,

$$V(C) = C \cap \mathcal{X} \setminus V(\mathcal{P}) \qquad \text{and} \qquad E(C) = \binom{V(C)}{2}. \tag{6}$$

Moreover, let $E'(C)$ be the set of all edges of $G_{\mathcal{X},\hat{r}} - V(\mathcal{P})$ incident with some point in $V(C)$. Finally, we consider the set E' of all edges of $G_{\mathcal{X},\hat{r}}$ that are incident with points in $V(\mathcal{P})$ or with points in cells that are not good or not rainbow. That is,

$$E' = E'(\mathcal{P}) \cup \bigcup_{C \notin \Gamma_1} E'(C). \tag{7}$$

During the construction of the rainbow Hamilton cycle in Section 3, special care will be required to avoid repeating colors that already appear in E'. The following result will help us achieve that.

Lemma 12. *Let $k_0 = \lceil 20/\varepsilon \rceil$. A.a.s. $|E'| \le n^{1-\varepsilon/3}$, and moreover, for every bad or non-rainbow cell C, fewer than $k_0 + 2$ edges in $E'(C)$ are assigned a color in $G_{\mathcal{X},\hat{r},q}$ that is repeated on another edge in E'.*

In fact, we prove something slightly stronger by allowing C to range over all cells, not necessarily bad or non-rainbow. The reason for stating the lemma only for bad or non-rainbow cells is that when we use it in Section 3 we will only expose the colors of edges in E' and assume that the second a.a.s. conclusion of the lemma holds as stated just for these cells.

Proof. Properties **P1**–**P4** and the a.a.s. claims in Lemma 9(a) and Lemma 11 imply that (eventually, for large n)

$$|E'| \le n^{O(\varepsilon^{1/d})} + O(\log^6 n) + O(n^{1-\varepsilon/2} \log^2 n) \le n^{1-\varepsilon/3}.$$

Also, for every cell C, $|E'(C)| = O(\log^2 n)$. Conditional on all the above properties, the probability that there is a cell C with k_0 edges in $E'(C)$ whose colors in $G_{\mathcal{X},\hat{r},q}$ are also used on $E' \setminus E'(C)$ can be bounded by

$$N\binom{O(\log^2 n)}{k_0}\left(\frac{n^{1-\varepsilon/3}}{(1+\eta)n}\right)^{k_0} \leq n^{1+o(1)-k_0\varepsilon/10} = o(1).$$

As a result, a.a.s. every cell C has fewer than k_0 edges in $E'(C)$ with colors repeated on $E' \setminus E'(C)$. To finish the proof, we observe that, in view of Lemma 9(b) (with say $A = 1$), a.a.s. for every cell C the set of edges $E'(C)$ contains at most one pair of edges with repeated colors. □

Let $G_{m,p}$ denote the Erdős-Rényi-Gilbert binomial random graph on m vertices where each pair of vertices is joined by an edge with probability $p \in [0,1]$. (Here p is unrelated to the parameter associated to the ℓ_p-norm $\|\cdot\|$ in the definition of the random geometric graph $G_{\mathcal{X},\hat{r}}$.) We now prove a lemma concerning the existence of Hamilton cycles and spanning collections of paths in $G_{m,p}$.

Lemma 13. *Suppose that $0 < p < 1$ is constant. Then,*

(a) $\mathbb{P}(G_{m,p}$ is not Hamiltonian$) \leq e^{-mp/5}$ for m sufficiently large.

(b) Let $\psi(G)$ denote the minimum number of vertex disjoint paths that cover the vertices of G. Then, for fixed $k \geq 1$ and sufficiently large m,

$$\mathbb{P}(\psi(G_{m,p}) > k) \leq e^{-kmp/6}.$$

Proof. (The asymptotic notation in this proof is with respect to $m \to \infty$, and we tacitly assume that m is sufficiently large for every inequality to be true.)

(a) We consider the standard coupling $G_{m,p} \supseteq G_1 \cup G_2$, where G_1, G_2 are independent copies of $G_{m,p/2}$. Given a graph G and $S \subseteq V(G)$, let $N_G(S)$ denote the disjoint neighborhood of S, i.e. the set of vertices that are not in S but are adjacent to some vertex in S. Let A_1 be the event that, for every $S \subseteq V(G_1)$ with $1 \leq |S| \leq m/6$, $|N_{G_1}(S)| > 2|S|$. By bounding the expected number of sets S that violate this condition, we get that

$$\mathbb{P}(\neg A_1) \leq \sum_{s=1}^{\lfloor m/6 \rfloor} \binom{m}{s}\binom{m}{2s}(1-p/2)^{s(m-3s)} \leq \sum_{s=1}^{\lfloor m/6 \rfloor}\left(\frac{me}{s}\cdot\frac{m^2e^2}{4s^2}\cdot e^{-mp/4}\right)^s \leq \frac{1}{3}e^{-mp/5}.$$

Let A_2 be the event that G_1 is connected. By bounding the expected number of components of order at most $m/2$, we show that

$$\mathbb{P}(\neg A_2) \leq \sum_{s=1}^{\lfloor m/2 \rfloor}\binom{m}{s}(1-p/2)^{s(m-s)} \leq \sum_{s=1}^{\lfloor m/2 \rfloor}\left(\frac{me}{s}\cdot e^{-mp/4}\right)^s \leq \frac{1}{3}e^{-mp/5}.$$

Now let A_3 be the event that G_2 has at least $\mu := \left\lceil \binom{m}{2}p/2 - m^{7/4}\right\rceil = (1+o(1))m^2p/4$ edges. By Chernoff's bound (see e.g. Corollary 21.7 in [9]), $\mathbb{P}(\neg A_3) = e^{-\Omega(m^{3/2})}$.

We will apply Pósa's rotation-extension argument (see [9], Chapter 6 for more details). By Pósa's lemma, events A_1 and A_2 imply that, if G_1 is not Hamiltonian, then there exist a set $\mathrm{END} \subseteq V(G_1)$ and for each $x \in \mathrm{END}$ a set $\mathrm{END}_x \subseteq \mathrm{END}$ with $|\mathrm{END}| \geq |\mathrm{END}_x| \geq m/6$ with the following property. The addition of any edge $\{x,y\}$ with $x \in \mathrm{END}$ and $y \in \mathrm{END}_x$ (which we call a *booster* edge) to G_1 results in either increasing the length of the longest path or closing a Hamilton cycle. Hence, there must be at least $\binom{\lceil m/6 \rceil}{2}$ such boosters. Moreover, since A_1, A_2 are increasing properties with respect to the addition of edges, every non-Hamiltonian supergraph $G_1' \supseteq G_1$ on vertex set $V(G_1)$ must satisfy the same property. Let us condition on events A_1, A_2 and A_3, and consider an enumeration $e_1, e_2, \ldots, e_\mu, \ldots$ of the edges of G_2. Suppose that, for some $0 \le k \le \mu - 1$, the supergraph $G_1 + \{e_1, e_2, \ldots, e_k\}$ of G_1 is not Hamiltonian. Then the probability that e_{k+1} is a booster is at least $\binom{\lceil m/6 \rceil}{2} / \binom{m}{2} \ge 1/37$. (This is because we know that none of $e_1, e_2, \ldots, e_k$ are boosters of $G_1 + \{e_1, e_2, \ldots, e_k\}$.) Thus the probability that we fail to produce a Hamilton cycle after adding edges $e_1, \ldots, e_\mu$ to G_1 is at most $\mathbb{P}(\mathrm{Bin}(\mu, 1/37) \le m) \le e^{-\Omega(m^2 p)}$ (again by Chernoff's bound). Hence, we conclude that

$$\mathbb{P}(G_{m,p} \text{ is not Hamiltonian}) \le \mathbb{P}(\neg A_1) + \mathbb{P}(\neg A_2) + \mathbb{P}(\neg A_3) + e^{-\Omega(m^2 p)} \le e^{-mp/5}.$$

(b) Let V_ℓ be the set of vertices of degree at most ℓ in G_1. Then for $\ell, r = O(1)$,

$$\mathbb{P}(|V_\ell| \ge r) \le \binom{m}{r} \mathbb{P}(\mathrm{Bin}(m-r, p/2) \le \ell)^r \le m^r \left(\sum_{i=0}^{\ell} \binom{m-r}{i} (p/2)^i (1-p/2)^{m-r-i} \right)^r$$
$$\le m^{r+\ell r} e^{-r(m-r-\ell)p/2} \le \frac{1}{2} e^{-rmp/3}.$$

Suppose now that we arbitrarily add edges incident to the vertices of degree at most $3k$ in G_1 so that the new graph H has minimum degree $3k$. (We can follow any fixed deterministic rule to do that, so H is a well-defined function of G_1.) Let A_1' be the event that, for every $S \subseteq V(H)$ with $1 \le |S| \le m/6$, $|N_H(S)| > 2|S|$, and let A_2' be the event that H is connected. Then, arguing as in (a),

$$\mathbb{P}(\neg A_1') \le \sum_{s=k}^{\lfloor m/6 \rfloor} \binom{m}{s} \binom{m}{2s} (1-p/2)^{s(m-3s)} \le \sum_{s=k}^{\lfloor m/6 \rfloor} \left(\frac{me}{s} \cdot \frac{m^2 e^2}{4s^2} \cdot e^{-mp/4} \right)^s \le \frac{1}{2} e^{-kmp/5}.$$
$$\mathbb{P}(\neg A_2') \le \sum_{s=3k}^{\lfloor m/2 \rfloor} \binom{m}{s} (1-p/2)^{s(m-s)} \le \sum_{s=3k}^{\lfloor m/2 \rfloor} \left(\frac{me}{s} \cdot e^{-mp/4} \right)^s \le e^{-3kmp/5}.$$

Repeating the same Pósa rotation-extension argument from part (a), it then follows that

$$\mathbb{P}(H \cup G_2 \text{ is not Hamiltonian}) \le \mathbb{P}(\neg A_1') + \mathbb{P}(\neg A_2') + \mathbb{P}(\neg A_3) + e^{-\Omega(m^2 p)} \le e^{-kmp/5}.$$

Now suppose that $k \ge 2$. If $|V_{3k}| \le \lfloor k/2 \rfloor$ and $H \cup G_2$ is Hamiltonian, then we have $\psi(G_{m,p}) \le k$, since deleting all the edges in $E(H) \setminus E(G_1)$ from a Hamilton cycle of $H \cup G_2$ creates at most $2\lfloor k/2 \rfloor$ paths. Hence,

$$\mathbb{P}(\psi(G_{m,p}) > k) \le e^{-kmp/5} + \mathbb{P}(|V_{3k}| \ge \lfloor k/2 \rfloor + 1) \le e^{-kmp/5} + \frac{1}{2} e^{-kmp/6} \le e^{-kmp/6}.$$

The case $k = 1$ follows immediately from part (a). This finishes the proof of the lemma. □

3 Rainbow Hamilton cycle construction

We now describe how we select our rainbow Hamilton cycle. Firstly, for each point $X_i \in \mathcal{X}$, we expose the cell containing X_i (which determines which cells are good, bad and ugly), and suppose that properties **P1–P3** in Lemma 7 hold. (Note that we do not reveal the exact location of each point X_i in $[0,1]^d$ to avoid conditioning on events of measure 0.) Next, we expose the incidence structure of graph $G_{\mathcal{X},\hat{r}}$, and suppose **P4** in Lemma 7 also holds. Moreover, assume there is a collection of ugly paths $\mathcal{P}$ that satisfies **Q1**–**Q5** in Lemma 8. In the case there is more than one choice for $\mathcal{P}$, pick one arbitrarily. Recall $V(\mathcal{P})$ is the set of points in $\mathcal{X}$ covered by paths in $\mathcal{P}$. For the next part of the argument we will remove all points in $V(\mathcal{P})$ from the cells, and treat them separately (see (6) and the discussion above it, in Section 2). In view of this, for any good cell C, $|V(C)| \ge \varepsilon^3 \log n - 2$, while for every ugly cell D, $|V(D)| = 0$. Now we reveal the number of color repetitions in each cell (which determines which ones are rainbow) without exposing the actual colors of the edges yet. Assume that all the a.a.s. statements in Lemmas 9 and 10 hold. In particular, every non-rainbow cell must be good, and the graph of rainbow good cells Γ_1 is connected. Moreover, points in non-rainbow cells are not adjacent in $G_{\mathcal{X},\hat{r}}$ to points on ugly paths. Further, we expose the colors of all the edges in E' (defined in (7)). Recall that these are the edges of $G_{\mathcal{X},\hat{r}}$ that are incident with points covered by $\mathcal{P}$ or with points contained in cells that are not in Γ_1. Recall the definitions of $E'(\mathcal{P})$ and $E'(C)$ in Section 2, as well. We condition on $E'(\mathcal{P})$ being rainbow (which is a.a.s. true by Lemma 11), and suppose that the a.a.s. conclusions in Lemma 12 hold. In particular $|E'| \le n^{1-\varepsilon/3}$. We conclude this discussion with a crucial observation. Conditional on all the information about $G_{\mathcal{X},\hat{r},q}$ exposed so far, the colors on the edges X_iX_j with both endpoints in cells of Γ_1 remain uniformly random with the only restriction that, for every cell C in Γ_1, the colors on the edges in $E(C)$ must be all different.

In the remainder of this section, we will a.a.s. build a rainbow cycle H that visits all the vertices inside rainbow good cells and avoids colors assigned to edges in E'. Next, we will deterministically extend H to include all the vertices inside bad or non-rainbow cells. Finally, we will insert the ugly paths into H to create a rainbow Hamilton cycle.

3.1 Rainbow good cells

Our first goal is to build a rainbow cycle that covers all the vertices inside rainbow good cells and avoids all the colors used on E'. (Recall $|E'| \le n^{1-\varepsilon/3}$.) Pick a spanning tree T of the giant component Γ_1 consisting of all the rainbow good cells, and root it at one of its cells C_1. Note that T has maximum degree $\Delta(T) = O_\varepsilon(1)$ (since it is a subgraph of $\mathcal{G}_{\mathcal{C}}$), and contains N_1 cells with $N_1 \sim N = O_\varepsilon(n/\log n)$, in view of all our earlier a.a.s. assumptions. Suppose that $C_1, C_2, \ldots, C_{N_1}$ is an enumeration of the cells in Γ_1 that follows from a depth-first search of T from the root cell C_1. For each $1 < i \le N_1$, let $\pi(i)$ denote the index of the parent $C_{\pi(i)}$ of C_i in this search. For convenience, we write $V_i = V(C_i)$ and $E_i = E(C_i)$. Let $m_i = |V_i|$, and recall $\varepsilon^3 \log n - 2 \le m_i \le \log n$ from our previous a.a.s. assumptions. Also, for $i, j = 1, \ldots, N_1$ $(i \ne j)$, let $E_{i,j}$ denote the set of edges in $G_{\mathcal{X},\hat{r}}$ with one endpoint in V_i and one in V_j.

Below we describe procedure `Build`, in which we examine the rainbow good cells $C_1, \ldots, C_{N_1}$ in this order and, at each step $i = 1, \ldots, N_1$, attempt to construct a rainbow cycle $H_i \subseteq G_{\mathcal{X},\hat{r},q}$ through $V_1 \cup \cdots \cup V_i$ that avoids colors on E'. Roughly speaking, at each step i, we find either a rainbow cycle or a rainbow collection of paths with vertex set V_i and which does not repeat any colors used on E' or H_{i-1}. Then, we patch this cycle or each of these paths into H_{i-1} at the parent cell $C_{\pi(i)}$ by using two edges in $E_{i,\pi(i)}$. This creates the new cycle H_i, which is typically rainbow. Occasionally, though, this patching operation cannot be done without repeating some colors already used on H_{i-1}. In that case, our algorithm attempts to fix these errors by making a small number of additional modifications to H_i, recursively. In the description of procedure `Build`, it is often convenient to regard $G_{\mathcal{X},\hat{r}}$ as an oriented graph by initially assigning to each edge $\{x, y\}$ an arbitrary orientation, xy or yx, which may change over the course of the algorithm. A path or a cycle is called *directed* (with respect to an orientation) if all its vertices have in- and out-degree at most one. We do not assume paths or cycles to be directed unless explicitly stated.

As we run this procedure we will expose some additional information of $G_{\mathcal{X},\hat{r},q}$, and assume in our description that certain properties hold (see Assumptions 1–6 below). If any of these assumptions ceases to be true at any given time, then `Build` fails and immediately stops. (We will later show that a.a.s. this does not occur.) Moreover, we claim that some additional properties are satisfied (see Claims 1–5 below) at the end of each step $i = 1, \dots, N_1$ provided that procedure `Build` has been successful so far. These claims are deterministic consequences of all of our assumptions, and will be proven inductively along with the description of the procedure.

Fix $1 \le i \le N_1$, and suppose we have just completed i steps of the algorithm.

Claim 1. $H_i \subseteq G_{\mathcal{X},\hat{r},q}$ is a rainbow directed cycle on vertex set $V_1 \cup \dots \cup V_i$, and it does not use any colors assigned to E'.

Claim 2. For every $j > i$, the procedure has not yet exposed the colors on any edges in $E_j \cup E_{j,\pi(j)}$. (In particular, these colors remain uniformly distributed conditional upon E_j being rainbow.)

For convenience, we identify the cycle H_i with its edge set $E(H_i)$, so in particular $|H_i|$ denotes the number of (oriented, colored) edges in H_i. We will tacitly follow a similar abuse of notation for other subgraphs of $G_{\mathcal{X},\hat{r}}$ (and also for their corresponding edge-colored versions, given $G_{\mathcal{X},\hat{r},q}$).

Claim 3. For every $1 < j \le i$, $|H_i \cap E_{j,\pi(j)}| = O(1)$. Moreover, for every $1 \le j' < j$ with $j' \ne \pi(j)$, $|H_i \cap E_{j,j'}| = 0$.

Each of the cells $C_1, \dots, C_i$ is labelled as *safe* or *unsafe* (with cell C_1 always declared unsafe). Safe cells will be used to fix errors due to color repetitions. Note that some cells may change their status from safe to unsafe during the procedure, but never the other way around.

Claim 4. The number of unsafe cells is at most $o(N_1)$.

For technical reasons, for each $1 \le j \le i$ we select a 'reasonably large' matching M_j in $H_i \cap E_j$, and partition it into two disjoint matchings M'_j and M''_j of roughly equal size. We say that an edge e is incident with a set of edges A in a graph if e shares an endpoint with some edge in A.

Claim 5. For every $1 \le j \le i$, the following holds. $M_j, M'_j, M''_j \subseteq H_i \cap E_j$ are matchings with $M_j = M'_j \cup M''_j$ and $M'_j \cap M''_j = \emptyset$. These matchings satisfy $|M'_j|, |M''_j| \ge (\varepsilon^3/4 + o(1)) \log n$ and thus $|M_j| \ge (\varepsilon^3/2 + o(1)) \log n$. Moreover, if cell C_j is safe, then the procedure has not yet exposed the colors on any edges in $E_{j,\pi(j)}$ that are incident with $M''_{\pi(j)}$.

Procedure `Build`: We initially assign an arbitrary orientation to every edge in $G_{\mathcal{X},\hat{r}}$. First consider cell C_1. We examine the edges in E_1 one by one, reveal their color in $G_{\mathcal{X},\hat{r},q}$, and delete those edges whose color has already been used on E'. (Recall that the colors on E_1 are uniformly distributed conditional upon E_1 being rainbow.) Let G_1 denote the graph with vertex set V_1 and the edges that remain. Each edge is deleted with probability at most $|E'|/(|Q| - m_1) = o(1)$, and thus (ignoring the orientations of the edges) G_1 contains a copy of G_{m_1,p_1} with $p_1 = 1 - o(1)$.

Assumption 1. G_1 is Hamiltonian. This holds a.a.s. by Lemma 13(a).

(Recall that if any of our Assumption 1–6 fails, then `Build` stops and fails.) Then, pick a Hamilton cycle H_1 of G_1, which must be rainbow by construction, and modify the orientations of the edges of H_1 (if needed) to ensure it is a directed cycle. Next, select an arbitrary matching M_1 of size at least $(\varepsilon^3/2 + o(1)) \log n$ contained in the cycle H_1 (e.g. by taking alternating edges in H_1), and partition M_1 into two disjoint matchings M_1' and M_1'' of size at least $(\varepsilon^3/4 + o(1)) \log n$ each. We label cell C_1 as unsafe since we will require all safe cells to have a parent in T. This finalizes the first step of the procedure. Note that Claims 1–5 are trivially satisfied with $i = 1$.

Let $1 < i \leq N_1$, and suppose we have successfully run the first $i - 1$ steps of `Build`. In particular, we inductively assume that Claims 1–5 were valid at the end of step $i - 1$. We now proceed to describe step i. As in the first step, we reveal the colors of the edges in E_i one by one, and delete those edges whose color has already been used on E' or H_{i-1}. Let G_i denote the resulting graph on vertex set V_i. Each edge is deleted with probability at most

$$\frac{|E'| + |H_{i-1}|}{|Q| - m_i} \leq \frac{n^{1-\varepsilon/3} + n}{(1+\eta)n - \log n} = \frac{1 + o(1)}{1 + \eta} \leq 1 - \eta/2 + o(1),$$

for $\eta < 1$. Hence, G_i contains a copy of G_{m_i,p_i} with $p_i = \eta/2 + o(1)$. (Here we are again ignoring the current orientations of the edges.) We say that step i is a Hamiltonian step if G_i contains a Hamilton cycle (i.e. a cycle through V_i, not necessarily directed). By Lemma 13(a), step i fails to be Hamiltonian with probability at most $e^{-m_i p_i/4} \leq n^{-\varepsilon^3\eta/8+o(1)}$. (This bound is also valid if $i = 1$, although a stronger bound was used in the first step of the algorithm.)

Assumption 2. The number of non-Hamiltonian steps up to step i is at most $n^{1-\varepsilon^3\eta/9}$.

Note that the expected number of non-Hamiltonian steps at the end of the procedure is at most $N_1 n^{-\varepsilon^3\eta/8+o(1)} = o(n^{1-\varepsilon^3\eta/9})$, so Assumption 2 is a.a.s. valid by the Markov inequality.

If step i is Hamiltonian, we will perform a `cycle-patch` step (below).

Assumption 3. If step i is not Hamiltonian then G_i contains a collection of at most $\psi_0 = \lceil \frac{13}{\varepsilon^3\eta} \rceil$ vertex-disjoint paths that cover V_i.

Note that, by Lemma 13(b), the probability that Assumption 3 fails at step i is at most $e^{-\psi_0 m_i p_i/6} \leq n^{-\psi_0\varepsilon^3\eta/12+o(1)} = o(1/N_1)$. Taking a union bound over all N_1 steps in the algorithm, we conclude that a.a.s. Assumption 3 is always valid. In this case we will perform a `forest-patch` step (below).

Swaps and cycle rotations: For the description of the `cycle-patch` and `forest-patch` steps below, it is convenient to introduce the following operations in the context of a directed graphs where loops are allowed. Given two non-incident directed edges xy and uv (possibly $x = y$ or $u = v$), an xy, uv-*swap* is the operation that deletes xy and uv and replaces them by xv and uy. Note that the orientation of the edges xy and uv determines the way in which their endpoints get recombined into new edges by the xy, uv-swap. We can use swaps to merge or modify directed cycles. For instance, given two vertex-disjoint directed cycles O_1, O_2 with $xy \in O_1$ and $uv \in O_2$, the application of an xy, uv-swap to $O_1 \cup O_2$ yields one single directed cycle on the same vertex set. (Note that O_1 or O_2 could be directed cycles of lenght 1, i.e. loops, or of length 2, i.e. pairs of anti-parallel edges.) Moreover, given a directed cycle O of length at least 4 and two non-consecutive edges xy, uv in O, we can reverse the orientations of all edges along the directed path from y to v in O (so that in particular uv becomes vu), then apply an xy, vu-swap, and finally reverse the orientation of vy to yv. We call this operation an xy, uv-*rotation* of the directed cycle O. The resulting graph is a different directed cycle with the same vertex set as O.

`Cycle-patch step:` If G_i is Hamiltonian, then label cell C_i as safe and pick a Hamilton cycle D_i of G_i. We can assume that D_i is a directed cycle, by appropriately modifying the orientations of the edges if necessary. Note that $H_{i-1} \cup D_i$ is rainbow by construction and does not use any colors from edges in E'. Since D_i is a cycle with $m_i \geq (\varepsilon^3 + o(1)) \log n$ edges, we can choose a matching M_i of size at least $(\varepsilon^3/2 + o(1)) \log n$ contained in D_i. Moreover, recall that $M'_{\pi(i)} \subseteq E_{\pi(i)} \cap H_{i-1}$ is a matching of size at least $(\varepsilon^3/4 + o(1)) \log n$, by Claim 5 applied to step $i-1$. Now let us reveal the colors on all the edges in $E_{i,\pi(i)}$ that are incident with both M_i and $M'_{\pi(i)}$. These colors had not been exposed yet in view of Claim 2 at step $i-1$. Our goal is to merge H_{i-1} and D_i together into one single larger directed cycle, which we will call H_i. To do that, we will pick appropriate edges $xy \in M_i$ and $uv \in M'_{\pi(i)}$, and perform an xy, uv-swap to $H_{i-1} \cup D_i$. That is, edges xy and uv are replaced by xu and yv, by appropriately updating edge orientations in $G_{\mathcal{X},\hat{r}}$ if needed. (We could also merge H_{i-1} and D_i in a different way if we first reversed the orientation of the edges in D_i and then applied an xy, vu-swap instead, but our argument will ignore this alternative.) An xy,uv-swap is *valid* if the two added edges, xu and yv, receive different colors in $G_{\mathcal{X},\hat{r},q}$ and these colors have not already been used on $E' \cup H_{i-1} \cup E_i$. Note that, in that case, the cycle H_i resulting from the swap satisfies the properties in Claim 1.

Assumption 4. There are indeed edges $xy \in M_i$ and $uv \in M'_{\pi(i)}$ such that the xy,uv-swap is valid.

The probability that a given xy,uv-swap is valid is at least

$$\left(1 - \frac{|E'| + |H_{i-1}| + |E_i| + 1}{|Q|}\right)^2 \geq \left(1 - \frac{1 + o(1)}{1 + \eta}\right)^2 \geq \eta^2/2,$$

for $\eta < \sqrt{2} - 1$ and large enough n. Since M_i and $M'_{\pi(i)}$ are disjoint matchings, the pairs of edges added in different swaps are disjoint, and thus the events concerning the validity of different swaps are independent. Hence, the probability that Assumption 4 fails at step i is at most

$$(1 - \eta^2/2)^{|M_i||M'_{\pi(i)}|} \leq (1 - \eta^2/2)^{(\varepsilon^6/8 + o(1)) \log^2 n} = o(1/N_1).$$

Summing over all N_1 potential steps, we conclude that a.a.s. Assumption 4 holds throughout the procedure. In view of that, we pick a valid xy, uv-swap arbitrarily, apply it to cycles H_{i-1} and D_i, and call H_i the resulting cycle. After the swap, we update the matchings as follows. We delete edge uv from $M'_{\pi(i)}$ and also from $M_{\pi(i)}$. Moreover, we delete xy from M_i, and partition the resulting matching M_i into two disjoint matchings M'_i and M''_i of size at least $(\varepsilon^3/4 + o(1)) \log n$ each. This finalizes step i. We now verify that Claims 1–5 remain valid at the end of this step. By construction, H_i satisfies all the properties in Claim 1. Claim 2 is also true since we did not expose the colors on any edge incident with any vertex in V_j for $j > i$. Moreover, since the only edges in $H_i \setminus H_{i-1}$ with endpoints in different cells are xu and yv, then

$$|H_i \cap E_{j,j'}| = \begin{cases} |H_{i-1} \cap E_{j,j'}| & \text{for } 1 \leq j' < j \leq i-1 \\ 2 & \text{for } j = i \text{ and } j' = \pi(i) \\ 0 & \text{for } j = i \text{ and } j' \neq \pi(i), \end{cases}$$

which implies that Claim 3 remains valid. Claim 4 still holds since we did not label any new cell unsafe. Matchings M_i, M'_i, M''_i introduced at this step satisfy the properties in Claim 5 by construction. Note that we did not expose the colors of any edges in $E_{i,\pi(i)}$ incident with $M''_{\pi(i)}$.

Matchings M_j, M'_j, M''_j with $j < i$ satisfied Claim 5 at the previous step, but we must take into account that the sizes of matchings $M_{\pi(i)}, M'_{\pi(i)}$ were decreased by one. However, each matching can only be affected at most $\Delta(T) = O_\varepsilon(1)$ times througout the procedure as a result of a `cycle-patch` step, so Claim 5 holds.

`Forest-patch` step: Otherwise, suppose G_i is not Hamiltonian. In that case, we label cell C_i as unsafe. A *linear forest* is a graph whose connected components are paths. By Assumption 3, we can pick a spanning linear forest L_i of G_i with at most ψ_0 components. Note that $H_{i-1} \cup L_i$ is rainbow by construction and does not use any of the colors used on E'. Since L_i consists of at most ψ_0 paths with a total of at least $m_i - \psi_0 \geq (\varepsilon^3 + o(1)) \log n$ edges, we can choose a matching M_i of size at least $(\varepsilon^3/2 + o(1)) \log n$ contained in L_i, and then partition M_i into two disjoint matchings M'_i, M''_i of size at least $(\varepsilon^3/4 + o(1)) \log n$ in any arbitrary way. Moreover, recall that $M''_{\pi(i)} \subseteq E_{\pi(i)} \cap H_{i-1}$ is a matching of size at least $(\varepsilon^3/4 + o(1)) \log n$, by Claim 5 applied to step $i - 1$.

Our goal is to patch each of the path components of L_i into H_{i-1}. For each path component P of L_i, we reveal the colors on all the edges in $E_{i,\pi(i)}$ that are incident with both an endpoint of P and some edge in $M''_{\pi(i)}$. These colors had not been exposed yet in view of Claim 2 at step $i - 1$. Then, we apply a `path-patch` sub-step (below) to this path P. This extends cycle H_{i-1} to a larger rainbow cycle that contains P and does not use any colors previously used on E'. For convenience, we still call this new cycle H_{i-1}, but will rename it to H_i at the end of the step when all the paths of L_i have been inserted.

`Path-patch` sub-step: Let u, v be the endpoints of path P in cell C_i (possibly $u = v$). We can assume that path P is directed, say from v to u, by appropriately modifying the orientation of the edges if necessary. Our goal is to patch P into H_{i-1}. To do that, we will pick an appropriate edge $xy \in M''_{\pi(i)}$, delete xy from H_{i-1}, and add edges xv, uy to join directed paths $H_{i-1} - xy$ and P. (As usual, we update the orientations in $G_{\mathcal{X},\hat{r}}$ of the two added edges xv, uy, if needed.) By analogy with the `cycle-patch` step, we can regard this operation as performing an xy, uv-swap to $H_{i-1} \cup (P + uv)$, where $P + uv$ denotes the directed cycle obtained by adding edge uv to path P. (If $u = v$, then $P + uv$ is simply a loop uu; if P consists of one single edge vu, then we simply regard $P + uv$ as a pair of anti-parallel edges.) Note that the way P is inserted into H_{i-1} depends on the orientation given to P. (For simplicity, our procedure only considers one of the two possible ways of doing that.)

Given an edge $xy \in M''_{\pi(i)}$, let c_1 and c_2 denote the colors assigned in $G_{\mathcal{X},\hat{r},q}$ to the edges that would be added at the end of an xy, uv-swap (i.e. xv and uy). If color c_k is repeated on an edge $e_k \in H_{i-1}$ for some $k \in \{1, 2\}$, then we say that the xy, uv-swap causes a *problem* at edge e_k or simply that e_k is a *problem edge* (relative to that particular swap and H_{i-1}). Note that each color c_k appears on at most one edge of H_{i-1} (since H_{i-1} is rainbow), and therefore an xy, uv-swap causes at most two problems. We say that the xy, uv-swap is *ideal* if colors c_1 and c_2 are different from each other and do not appear on any edges in $E' \cup H_{i-1} \cup E_i$. (In particular, ideal swaps create no problem edges.) On the other hand, the xy, uv-swap is *acceptable* if it is not ideal but the following conditions hold: 1) colors c_1 and c_2 are different; 2) c_1 and c_2 do not appear on any edge in $E' \cup E_i$; 3) each c_k is used at most once on H_{i-1} (note that this condition is redundant since H_{i-1} is rainbow, but it will be useful later on when we consider acceptable swaps in a slightly different context that allows a few color repetitions); 4) if color c_k is used on some edge $e_k \in H_{i-1}$ for some $k \in \{1, 2\}$ (i.e. e_k is a problem edge), then $e_k \in E_{j_k}$ for some safe cell C_{j_k}; and 5) if both colors

c_1, c_2 are respectively used on edges $e_1, e_2 \in H_{i-1}$, then the safe cells C_{j_1} and C_{j_2} containing these edges (as defined in condition 4)) must be different. Later on, we will consider acceptable swaps in a context where H_{i-1} may already contain some additional edges labelled as problems, which are located at different safe cells. In view of that, it is convenient to reword condition 5) as follows: a problem edge created by the xy, uv-swap cannot be contained in the same cell as another problem edge (relative to that swap or already present in H_{i-1}). Finally, the xy, uv-swap is *forbidden* if it is neither ideal nor acceptable.

Assumption 5. Not all the xy, uv-swaps for $xy \in M''_{\pi(i)}$ are forbidden. (Hence there is at least one swap that is ideal or acceptable.)

We defer the proof that Assumption 5 is valid a.a.s., until later.

First suppose that there exists an edge $xy \in M''_{\pi(i)}$ such that the xy, uv-swap is ideal. Pick one such edge xy arbitrarily, and apply the xy, uv-swap to $H_{i-1} \cup (P + uv)$. This inserts P into H_{i-1}. The resulting cycle, which we still call H_{i-1}, is directed and rainbow by construction and does not contain any colors used on E'. After performing the swap, edge xy is removed from matching $M''_{\pi(i)}$ and thus from $M_{\pi(i)}$. Otherwise, if there is no ideal swap available, pick an arbitrary $xy \in M''_{\pi(i)}$ such that the xy, uv-swap is acceptable (there must be at least one). By definition, for at least one $k \in \{1, 2\}$ (and maybe for both), color c_k already appears on one edge $e_k \in H_{i-1}$ which is labelled as a problem edge. We then apply the xy, uv-swap to $H_{i-1} \cup (P + uv)$, and remove edge xy from the matchings $M''_{\pi(i)}$ and $M_{\pi(i)}$. In this case, the new cycle obtained after the swap, which we still denote by H_{i-1}, is directed and contains no colors used on E', but it is not rainbow since it has one or two color repetitions, one per problem edge. We now attempt to make H_{i-1} rainbow by taking a `Problem-fix` sub-step (below) for each problem edge e_k. This procedure recursively applies cycle rotations to H_{i-1}, which remove problem edges but may in turn create new problems.

Assumption 6. `Problem-fix` successfully terminates after at most $\xi = 16/(\eta\varepsilon^3)$ recursive iterations.

We defer the proof that Assumption 6 is valid a.a.s., until later.

We will show that the resulting cycle H_{i-1} is directed and rainbow, includes path P and does not share any colors with E'. This ends the `Path-patch` sub-step. If P was the last path of L_i to be inserted, then rename H_{i-1} to H_i.

`Problem-fix` sub-step: Recall that H_{i-1} is a directed cycle, but not rainbow. However, all its color repetitions are due to the presence of problem edges. Note that H_{i-1} may contain up to 2ξ problem edges. This is due to Assumption 6 and to the fact that the `Path-patch` sub-step and each additional recursive call of `Problem-fix` can create at most two new problem edges. Moreover, since problem edges originate from acceptable swaps, they must all lie in different and safe cells by construction.

Suppose that we are trying to fix a problem edge $uv \in H_{i-1} \cap E_j$ for some $j \le i - 1$. In particular, the cell C_j containing that edge must be safe, and thus by Claim 5 the colors on the edges that are incident with uv and $M''_{\pi(j)}$ have not yet been exposed. Our plan is to perform an xy, uv-rotation of H_{i-1} for some suitable $xy \in M''_{\pi(j)}$. (Note that uv and xy are not incident, since they are contained in different cells.) This operation amounts to reversing the orientation of some edges (including uv to vu) and applying an xy, vu-swap. For simplicity, we will only discuss the choice of the xy, vu-swap for $xy \in M''_{\pi(j)}$, and assume that the edge orientations are adjusted as required by the xy, uv-rotation, so the resulting cycle (which we still call H_{i-1}) is directed.

We essentially follow the same strategy as in the `Path-patch` sub-step. We reiterate Assumption 5 here, and suppose that not all the xy, vu-swaps are forbidden. (Otherwise, `Build` fails.) Then, we first attempt to perform an ideal xy, vu-swap for some $xy \in M''_{\pi(j)}$, if possible, and otherwise use an acceptable one. In the former case, the corresponding xy, uv-rotation successfully removes the problem edge uv from H_{i-1} while not creating any new problems. In the latter case, we also get rid of uv, but add one or two new problem edges (and thus color repetitions) to H_{i-1}. At the end of either case, we further delete edge xy from the matchings $M''_{\pi(j)}, M_{\pi(j)}$, and remove edge uv from any of the matchings M_j, M'_j, M''_j that may contain it (possibly none of them). Finally, we update the status of cell C_j from safe to unsafe to guarantee that it will never host other problem edges at any other step of the algorithm. If the resulting directed cycle H_{i-1} has no problem edges left, then we successfully terminate `Problem-fix`. Otherwise, we recursively apply `Problem-fix` to one of the remaining problem edges. With Assumption 6 in mind, we only allow up to ξ recursive iterations arising from one `Path-patch` sub-step. Otherwise, `Build` fails.

This ends the description of the `Forest-patch` step (and all its corresponding sub-steps), which is taken at step i if G_i is not Hamiltonian. We proceed to verify that at the end of that step Claims 1–5 remain valid. Claim 1 holds by construction, inductively assuming that it was true at step $i-1$. Indeed, after inserting all paths from L_i into H_{i-1} and recursively fixing all the problem edges, H_i is a directed rainbow cycle spanning $V_1 \cup \cdots \cup V_i$ and avoiding all colors that appear on E'. Claim 2 also remains valid since only colors on edges in $E_i \cup \bigcup_{1<j\leq i} E_{j,\pi(j)}$ were exposed at step i. To verify Claim 3, note that $|H_i \cap E_{j,j'}| = |H_{i-1} \cap E_{j,j'}|$ for each $1 \leq j' < j \leq i-1$, unless $j' = \pi(j)$ and a problem edge was created in C_j during step i, in which case $|H_i \cap E_{j,\pi(j)}| = |H_{i-1} \cap E_{j,\pi(j)}| + 2$. Moreover, $|H_i \cap E_{i,\pi(i)}| \leq 2\psi_0$ and $|H_i \cap E_{i,j}| = 0$ for $j \neq \pi(i)$. Hence, Claim 3 follows by induction and from the fact that a cell can host at most one problem edge during the whole procedure. Now recall that C_1 is always unsafe and that a cell C_j $(1 < j \leq i)$ is unsafe at the end of step i only in the following two situations: a) step j was not Hamiltonian (there are at most $n^{1-\varepsilon^3\eta/9}$ such steps, by Assumption 2); or b) step j was Hamiltonian (and thus C_j was initially declared safe), but then C_j became unsafe due to a problem edge arising from a later non-Hamiltonian step (there at most $n^{1-\varepsilon^3\eta/9}$ non-Hamiltonian steps, by Assumption 2, and each triggers at most ξ problem edges, by Assumption 6). Hence, there are at most $1 + n^{1-\varepsilon^3\eta/9} + \xi n^{1-\varepsilon^3\eta/9} = o(N_1)$ unsafe cells, and Claim 4 holds. Next we verify Claim 5 at the end of step i, inductively assuming that it was true at the previous step. Note that matchings M_i, M'_i, M''_i created in the `Forest-patch` step satisfy all the requirements by construction (recall that C_i is declared unsafe, so the last condition in the claim is trivially true). We need to check that M_j, M'_j, M''_j (for $1 \leq j \leq i-1$) still meet all the conditions at the end of step i. During that step, the sizes of matchings $M_{\pi(i)}$ and $M''_{\pi(i)}$ decreased by at most ψ_0 due to the insertion of the paths of L_i into H_{i-1}, but this can happen at most $\Delta(T) = O_\varepsilon(1)$ times throughout the entire procedure. Moreover, for each cell C_j $(1 < j \leq i-1)$ containing a problem edge at step i, we decreased the sizes of M_j, M'_j, M''_j by at most one (but this can happen only once in the procedure), and likewise the sizes of $M_{\pi(j)}, M''_{\pi(j)}$ were decreased by one (but this can happen at most $\Delta(T) = O_\varepsilon(1)$ times). We excluded $j = 1$ above since C_1 is unsafe, and thus never contains problem edges. Hence, by induction, $|M'_j|, |M''_j| \geq (\varepsilon^3/4 + o(1)) \log n$ and $|M_j| \geq (\varepsilon^3/2 + o(1)) \log n$ for all $1 \leq j \leq i$. Finally, all the cells C_j $(1 < j \leq i-1)$ for which we exposed the colors on the edges in $E_{j,\pi(j)}$ incident with $M''_{\pi(j)}$ at step i were relabelled unsafe, and therefore the last condition in Claim 5 remains valid.

We have shown that Claims 1–5 hold throughout the N_1 steps of procedure `Build` as long as it does not fail: that is, under Assumptions 1–6. In particular, at the end of the N_1-th step, by Claim 1, we obtain a rainbow cycle H_{N_1} spanning all the vertices in rainbow good cells and avoiding all the colors used on E'. It only remains to show that Assumptions 5 and 6 hold a.a.s. through all the steps (since Assumptions 1–4 have already been verified) in order to conclude that procedure `Build` succeeds a.a.s. To do so, we will bound the probability that a given swap in a `Path-path` or a `Problem-fix` sub-step is forbidden and the probability it is not ideal.

Let $u, v \in V_j$ be the endpoints of the problem edge to be fixed in a `Problem-fix` with $j \le i-1$ or the endpoints of the path to be patched in a `Path-patch` sub-step with $j = i$. Given $xy \in M''_{\pi(j)}$, the xy, uv-swap is forbidden if, for some $k \in \{1, 2\}$, color c_k is equal to c_{3-k} or one of the repeated colors on H_i (there are at most ξ of those, by Assumption 6, since each color repetition is due to a problem edge) or if c_k appears on any of the following edges: edges in $E' \cup E_i$ (where $|E' \cup E_i| \le n^{1-\varepsilon/3} + \log n$, by an earlier bound on $|E'|$ and **P1**), edges in unsafe cells (there are $o(N_1 \log n)$ of those, by Claim 4 and **P1**), edges in H_i with endpoints in different cells (there are $O(N_1)$ of those, by Claim 3) or edges in a safe cell containing another problem edge (there are at most $\xi \log n$ of these, by Assumption 6 and **P1**). Hence, the probability that a given xy, uv-swap is forbidden is at most

$$2\left(\frac{1 + \xi + |E'| + \log n + o(N_1 \log n) + O(N_1) + O(\log n)}{|Q|}\right) = o(1).$$

Since $M''_{\pi(j)}$ is a matching, events concerning different swaps are independent, and thus the probability that all the swaps are forbidden is

$$(o(1))^{|M''_{\pi(j)}|} \le (o(1))^{(\varepsilon^3/4 + o(1)) \log n} = n^{-\omega(1)} = o(1/N_1),$$

by Claim 5. Therefore, summing this bound over all N_1 potential steps times the at most $(1 + \xi)$ possible `Path-path` or `Problem-fix` sub-steps within each step, the probability that Assumption 5 fails at some point in the algorithm is $o(1)$. On the other hand, given $xy \in M''_{\pi(j)}$, the xy, uv-swap is not ideal if, for some $k \in \{1, 2\}$, color c_k is equal to c_{3-k} or it appears on some edge in $E' \cup H_{i-1} \cup E_i$. Hence, the probability that a given xy, uv-swap is not ideal is at most

$$2\left(\frac{1 + |E'| + n + m_i}{|Q|}\right) = \frac{1 + o(1)}{1 + \eta} \le 1 - \eta/2$$

(for $0 < \eta < 1$ and large enough n). Thus, the probability that we are forced to pick an acceptable swap at a given `Path-patch` or `Problem-fix` sub-step is at most

$$(1 - \eta/2)^{|M''_{\pi(j)}|} \le (1 - \eta/2)^{(\varepsilon^3/4 + o(1)) \log n} \le n^{-\eta\varepsilon^3/8 + o(1)},$$

again by Claim 5. Note that each acceptable swap introduces one or two new problem edges, which in turn require recursive iterations of `Problem-fix`. Then, the probability that from one single `Forest-patch` step we create ξ problems (which requires picking at least $\xi/2$ acceptable swaps) is at most

$$O(n^{-\xi\eta\varepsilon^3/16 + o(1)}) = o(1/N_1),$$

where we use the fact that $\xi\eta\varepsilon^3/16 > 1$. So we expect $o(1)$ violations of Assumption 6 in the N_1 steps of the algorithm, and thus Assumption 6 holds a.a.s. by Markov inequality. This completes the analysis of `Build`, and shows that a.a.s. we obtain a rainbow directed cycle $H = H_{N_1}$ through all the vertices inside rainbow good cells that avoids all the colors used on E'.

3.2 Bad or non-rainbow cells

Suppose that all the earlier a.a.s. statements in the paper hold (see the discussion at the beginning of Section 3) and also that procedure `Build` succeeds at building the rainbow directed cycle H. (Here we assume that the edges of $G_{\mathcal{X},\hat{r}}$ are oriented as in Section 3.1.) Recall that H does not use any colors on E', which is the set of edges of $G_{\mathcal{X},\hat{r}}$ that are incident with a point in a cell that is not good or not rainbow or are incident with a point in an ugly path. We will extend H by adding the points in cells that are either non-rainbow (and thus good by Lemma 9) or bad, one cell at a time. We will do that deterministically, given all our a.a.s. assumptions.

Let C be a cell that is bad or non-rainbow. By the definition of bad cell and by Lemma 9, C must be adjacent in the graph of cells $\mathcal{G}_{\mathcal{C}}$ to some rainbow good cell C_i (for some $1 \le i \le N_1$). In view of Claim 5, at the end of procedure `Build`, $H \cap E_i$ contains a matching M_i of size $|M_i| = \Omega_\varepsilon(\log n)$. An edge in $E'(C)$ (i.e. incident with some point in $V(C)$) is labelled *dangerous* if its color in $G_{\mathcal{X},\hat{r},q}$ is repeated on some other edge in E'. By Lemma 12 there can be at most $k_0+1 = O_\varepsilon(1)$ dangerous edges in $E'(C)$. Suppose first that $V(C)$ contains more than $2k_0+5$ points. Since $V(C)$ induces a clique in $G_{\mathcal{X},\hat{r}}$, we can find k_0+2 edge-disjoint spanning cycles of that clique (for instance, consider the well-known Walecki construction described in [1]). At least one of these cycles does not contain any dangerous edges. Pick one and call it H_C. We can assume that H_C is a directed cycle by adjusting the orientations of its edges as needed. Now pick an edge $uv \in H_C$ and an edge $xy \in M_i$ with the property that xy is not incident with any dangerous edge in $E(C')$. (We have at least $|M_i| - k_0 - 1 = \Omega_\varepsilon(\log n)$ choices for xy, since a dangerous edge in $E(C')$ is incident with at most one edge in M_i.) Then, by applying an xy, uv-swap to $H_C \cup H$, we merge H_C and H into one larger rainbow directed cycle that we still call H. After the swap, delete xy from M_i. Otherwise, if $V(C)$ contains $t \le 2k_0+5$ points $v_1, \ldots, v_t$, then pick t different edges $x_1y_1, \ldots, x_ty_t$ in M_i such that each x_jy_j is not incident with any dangerous edge in $E'(C)$. As before, we have plenty of freedom to do this, since we can choose from a pool of at least $|M_i| - k_0 - 1 = \Omega_\varepsilon(\log n)$ edges. Then, each vertex v_j ($1 \le j \le t$) is inserted into H by replacing x_jy_j by the directed path $x_jv_jy_j$, ajusting edge orientations if needed. The resulting cycle, which we still denote by H, is directed and rainbow by construction and includes all the points in $V(C)$. After doing that, we delete edges $x_1y_1, \ldots, x_ty_t$ from the matching M_i.

We repeat the same operation for every bad or non-rainbow cell C, one cell at a time, until H covers all vertices of $\mathcal{X}$ that are not in ugly paths. Note that, since the graph of cells has maximum degree $\Delta(\mathcal{G}_{\mathcal{C}}) = O_\varepsilon(1)$, for each good rainbow cell C_i ($1 \le i \le N_1$), the corresponding matching M_i may loose at most $(2k_0+5)\Delta(\mathcal{G}_{\mathcal{C}}) = O_\varepsilon(1)$ edges in total, so we still have $|M_i| = \Omega_e(\log n)$ throughout this procedure, as required. Moreover, the edges that were used to extend H must have all different colors and do not repeat colors from edges incident with ugly paths, thanks to the fact that we did not choose any dangerous edges. Therefore, we eventually obtain a rainbow directed cycle H that covers all vertices in $\mathcal{X} \setminus V(\mathcal{P})$ and avoids all colors on $E'(\mathcal{P})$.

3.3 Ugly paths

It only remains to patch the ugly paths into H. Let $\mathcal{P}$ be the collection of ugly paths as in Lemma 8. For each path $P \in \mathcal{P}$, let u and v be the endpoints of P, and assume that P is directed from v to u by adjusting the orientations of its edges appropriately if necessary. Let C_P be a good cell as in **Q4**, which must be also rainbow by Lemma 9. So, using the notation from Section 3.1, $C_P = C_i$ for some $1 \le i \le N_1$. Pick any edge xy from the matching M_i. (There are $|M_i| = \Omega_\varepsilon(\log n)$ choices.) Note that both x and y are adjacent with u and v in $G_{\mathcal{X},\hat{r}}$ by our choice of C_P. By applying an xy, uv-swap to $H \cup (P + uv)$, we insert P into H. The resulting cycle, which we still call H, is directed and rainbow, since the set of edges $E'(\mathcal{P})$ incident with ugly paths is rainbow by Lemma 11 and H does not use any colors appearing on $E'(\mathcal{P})$. We can repeat this operation for each $P \in \mathcal{P}$, noting that each rainbow good cell C_i will be used to patch at most one ugly path in view of **Q5**. Hence, we eventually obtain a rainbow (directed) Hamilton cycle H of $G_{\mathcal{X},\hat{r},q}$. This completes the proof of Theorem 2.

References

[1] B. Alspach, The wonderful Walecki construction, *Bulletin of the Institute of Combinatorics and its Applications* 52 (2008) 7–20.

[2] D. Bal, P. Bennett, X. Pérez-Giménez and P. Prałat, Rainbow perfect matchings and Hamilton cycles in the random geometric graph, *Random Structures and Algorithms* 51 (2017) 587–606.

[3] D. Bal and A.M. Frieze, Rainbow Matchings and Hamilton Cycles in Random Graphs *Random Structures and Algorithms* 48 (2016) 503–523.

[4] J. Balogh, B. Bollobás, M. Krivelevich, T. Müller and M. Walters, Hamilton cycles in random geometric graphs, *The Annals of Applied Probability* 21 (2011), 1053–1072.

[5] C. Cooper and A.M. Frieze, Multi-coloured Hamilton cycles in random edge-colored graphs, *Combinatorics, Probability and Computing* 11 (2002) 129–133.

[6] J. Díaz, D. Mitsche and X. Pérez-Giménez, Sharp threshold for Hamiltonicity of random geometric graphs, *SIAM Journal on Discrete Mathematics* 21 (2007) 57–65.

[7] T. Fenner and A.M. Frieze, Hamiltonian cycles in random regular graphs, *Journal of Combinatorial Theory* B (1984) 103-112.

[8] A. Ferber and M. Krivelevich, Rainbow Hamilton cycles in random graphs and hypergraphs, in *Recent trends in combinatorics, IMA Volumes in Mathematics and its applications, A. Beveridge, J. R. Griggs, L. Hogben, G. Musiker and P. Tetali, Eds.*, Springer 2016, 167–189.

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

[10] A.M. Frieze and P. Loh, Rainbow hamilton cycles in random graphs, *Random Structures and Algorithms* 44 (2014) 328–354.

[11] A. Frieze, X. Pérez-Giménez, P. Prałat and B. Reiniger, Perfect matchings and Hamiltonian cycles in the preferential attachment model, *Random Structures and Algorithms* 54 (2019), 258–288.

[12] J. Komlós and E. Szemerédi, Limit distributions for the existence of Hamilton circuits in a random graph, *Discrete Mathematics* 43 (1983) 55-63.

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

[14] M.D. Penrose. Random geometric graphs, *Oxford University Press*, 2003.