CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Rainbow Hamiltonicity in uniformly coloured perturbed digraphs

Kyriakos Katsamaktsis*, Shoham Letzter[†] (iD) and Amedeo Sgueglia[†] (iD)

Department of Mathematics, University College London, London, UK
Corresponding author: Shoham Letzter; Email: s.letzter@ucl.ac.uk

Abstract

We investigate the existence of a rainbow Hamilton cycle in a uniformly edge-coloured randomly perturbed digraph. We show that for every $\delta \in (0, 1)$ there exists $C = C(\delta) > 0$ such that the following holds. Let D_0 be an n-vertex digraph with minimum semidegree at least δn and suppose that each edge of the union of D_0 with a copy of the random digraph $\mathbf{D}(n, C/n)$ on the same vertex set gets a colour in $[n]$ independently and uniformly at random. Then, with high probability, $D_0 \cup \mathbf{D}(n, C/n)$ has a rainbow directed Hamilton cycle.

This improves a result of Aigner-Horev and Hefetz ((2021) *SIAM J. Discrete Math.* **35**(3) 1569–1577), who proved the same in the undirected setting when the edges are coloured uniformly in a set of $(1 + \varepsilon)n$ colours.

1. Introduction

Determining which minimum degree forces the containment of a given spanning subgraph is a central theme in extremal combinatorics. The prototypical example is Dirac's theorem [10], which says that for an n-vertex graph G the condition $\delta(G) \geq n/2$ is sufficient to guarantee that G is Hamiltonian, and that this minimum degree condition is best possible. The question has been investigated in the setting of digraphs as well. Here a natural notion of minimum degree is the minimum *semidegree*, denoted by $\delta^0(D)$, which is the minimum over all in- and out-degrees of the vertices of D. Ghouila-Houri [15] proved a directed analogue of Dirac's theorem, showing that, if D is an n-vertex digraph with $\delta^0(D) \geq n/2$, then D is Hamiltonian. Here by a digraph D being *Hamiltonian* we mean that D contains a *directed Hamilton cycle*, i.e. a Hamilton cycle with all edges oriented consistently.

On the other hand, one of the main pursuits in probabilistic combinatorics is understanding the minimum p such that $\mathbf{G}(n, p)$, the binomial random graph on $[n]$ with edge probability p, contains a given subgraph *with high probability*.[1] Following the breakthrough of Pósa [24], it was proven [19, 20] that, with high probability, $\mathbf{G}(n, p)$ is Hamiltonian if it has minimum degree at least 2, implying that $\log n/n$ is a *sharp threshold* for Hamiltonicity. Here $\hat{p} = \hat{p}(n)$ is said to be

---

[1]We say that a sequence of events $(A_n)_{n\in\mathbb{N}}$ holds *with high probability* if $\mathbb{P}[A_n] \to 1$ as $n \to \infty$.

a *sharp threshold* for a graph property $\mathcal{P}$ if, for every $\varepsilon > 0$, with high probability the following holds.

$$G(n, p) \notin \mathcal{P} \text{ if } p \leq (1 - \varepsilon)\hat{p},$$

$$G(n, p) \in \mathcal{P} \text{ if } p \geq (1 + \varepsilon)\hat{p}.$$

The directed analogue of $\mathbf{G}(n, p)$, which we call the *random directed graph* and denote by $\mathbf{D}(n, p)$, is the random digraph on $[n]$ where each ordered pair of distinct vertices forms a directed edge independently with probability p. McDiarmid showed [21] that the threshold for Hamiltonicity in $\mathbf{D}(n, p)$ is at most that of $\mathbf{G}(n, p)$, and Frieze [14] proved a directed analogue of the above result about Hamiltonicity of random graphs with minimum degree at least 2, from which it follows that $\log n/n$ is a sharp threshold for Hamiltonicity for directed random graphs as well. Recently, Montgomery [22] determined sharp thresholds for all possible orientations of a Hamilton cycle in $\mathbf{D}(n, p)$.

As an interpolation between the deterministic and probabilistic models, Bohman et al. [8] introduced the *perturbed* model for graphs and digraphs. Given a fixed $\delta > 0$, let D_0 be a digraph on vertex set $[n]$ with minimum semidegree at least δn. The perturbed digraph is $D_0 \cup \mathbf{D}(n, p)$, i.e. it is the union of a digraph on vertex set $[n]$ with minimum semidegree at least δn, and the random digraph $\mathbf{D}(n, p)$ on the same vertex set. The perturbed graph model is defined similarly as $G_0 \cup \mathbf{G}(n, p)$, where G_0 is an n-vertex graph with minimum degree at least δn. In [8] they proved that there exists $C > 0$, depending only on δ, such that for each n-vertex digraph D_0 with minimum semidegree at least δn, the perturbed digraph $D_0 \cup \mathbf{D}(n, C/n)$ has with high probability a directed Hamilton cycle. That is, for every digraph with linear minimum semidegree, adding linearly many random edges results in a graph that with high probability contains a directed Hamilton cycle. Up to the dependence on δ and C, this is best possible for all $\delta \in (0, 1/2)$, since the complete bipartite digraph with parts of size δn and $(1 - \delta)n$, with each edge appearing with both orientations, requires $\Omega(n)$ edges to be Hamiltonian. (When $\delta \geq 1/2$ no random edges are needed, due to Ghouila-Houri's theorem.) This was generalised very recently by Araujo et al. [5], who showed that with high probability the same hypotheses ensure that the perturbed digraph contains every orientation of a cycle of every possible length, simultaneously.

In this paper we consider a rainbow variant of the theorem of Bohman, Frieze and Martin. A subset of the edges of an edge-coloured graph or digraph is called *rainbow* if no two edges share a colour, and a subgraph is called rainbow if its edge set is rainbow. For a finite set of colours $\mathcal{C}$, a graph or digraph is *uniformly coloured* in $\mathcal{C}$ if each edge gets a colour in $\mathcal{C}$ independently and uniformly at random. The problem of finding rainbow subgraphs of uniformly coloured graphs is well studied, in particular for $\mathbf{G}(n, p)$ [6, 9, 11–13]. The analogous problem in perturbed graphs was considered only more recently [1, 2, 4]. In particular, the problem of containing a rainbow Hamilton cycle was first addressed by Anastos and Frieze [4], who showed that if the number of colours is at least about $120n$, then $G \sim G_0 \cup \mathbf{G}(n, C/n)$ has a rainbow Hamilton cycle with high probability, for suitable C depending only on δ. Later, Aigner-Horev and Hefetz [1] improved this result by showing that, at the same edge probability in the random graph, $(1 + \varepsilon)n$ colours suffice for every $\varepsilon > 0$ (where C now depends also on ε). We prove that the optimal number of colours suffices.

Theorem 1.1. *For any $\delta \in (0, 1/2)$ there exists $C > 0$ such that the following holds. Let G_0 be an n-vertex graph with minimum degree at least δn and let $G \sim G_0 \cup \mathbf{G}(n, C/n)$ be uniformly coloured in $[n]$. Then with high probability G contains a rainbow Hamilton cycle.*

In fact, we can prove the corresponding result for uniformly coloured perturbed digraphs.

Theorem 1.2. *For any $\delta \in (0, 1/2)$ there exists $C > 0$ such that the following holds. Let D_0 be an n-vertex digraph with minimum semidegree at least δn and let $D \sim D_0 \cup \mathbf{D}(n, C/n)$ be uniformly coloured in $[n]$. Then with high probability D contains a rainbow directed Hamilton cycle.*

As explained above, for $\delta \in (0, 1/2)$, both results have the optimal edge probability, up to the dependence of C on δ. We remark that the first two authors proved Theorem 1.1 in unpublished work,[2] using a somewhat simpler version of Lemma 3.1. Here, we will deduce Theorem 1.1 from Theorem 1.2 by a variation of McDiarmid's celebrated coupling argument [21] (see Section 3).

The paper is structured as follows. In Section 2 we sketch the proof of Theorem 1.2. In Section 3 we prove Theorem 1.2 assuming Lemma 3.1, the key lemma of the paper, and deduce Theorem 1.1 from Theorem 1.2. In Section 4 we state and prove some preliminary results that we will need later. Section 5 is the most technical, where we prove the existence of the 'gadgets' which underpin Lemma 3.1. In Section 6 we prove Lemma 3.1. We finish with some concluding remarks in Section 7.

Notation. Given a digraph D and $x, y \in V(D)$, we write xy for the edge directed from x to y. Given $X, Y \subseteq V(D)$, we write $E_D(X, Y)$ for the set of edges xy with $x \in X$ and $y \in Y$, and $e_D^+(X, Y) = |E_D(X, Y)|$ for its size, and similarly $e_D^-(X, Y) = |E_D(Y, X)|$. The out-neighbourhood of a vertex v is denoted by $N_D^+(v)$ and its size by $\deg_D^+(v)$. For a vertex set Y, we write $N_D^+(v, Y) = N_D^+(v) \cap Y$ and denote its size by $\deg_D^+(v, Y) = |N_D^+(v, Y)|$. Similarly we will use $N_D^-(v)$, $\deg_D^-(v)$, $N_D^-(v, Y)$ and $\deg_D^-(v, Y)$. We will suppress D when the digraph in question is clear. We recall that the *length* of a (directed) path is the number of its edges. Given an edge-coloured digraph D, we denote the colour of an edge e by $\mathcal{C}(e)$ and the set of colours on the edges of a subdigraph D' by $\mathcal{C}(D')$. Moreover we say that D' is *spanning* in a colour set $\mathcal{C}'$ if $\mathcal{C}(D') = \mathcal{C}'$.

Throughout the paper, we will assume that n is sufficiently large. Asymptotic notation hides absolute constants: if for some $x, \varepsilon, n > 0$ we write $x = O(\varepsilon n)$, then there is an absolute constant $C > 0$, which does not depend on x, ε, n or any other parameters, such that $x \leq C\varepsilon n$. We write $x = \Omega(y)$ if $y = O(x)$, and we write $x = \Theta(y)$ if both $x = \Omega(y)$ and $x = O(y)$. For $x, y \in (0, 1)$, we write $x \ll y$ if $x < f(y)$ for an implicit positive increasing function f.

2. Proof sketch

Our proof uses the *absorption method*. This method is typically applicable when one searches for a spanning subgraph, and involves two stages: finding an almost spanning subgraph; and dealing with the remainder, by having a 'special' set of vertices, put aside at the beginning, that can cover any sufficiently small set of vertices. For finding the special set of vertices, we prove Lemma 3.1. It states that with high probability the perturbed digraph has a subgraph H_{abs} such that, for any small sets of vertices V' and colours $\mathcal{C}'$ with $|V'| = |\mathcal{C}'|$, disjoint from the vertices and colours of H_{abs} the following holds: there exists a rainbow directed path Q with vertex set $V' \cup V(H_{\text{abs}})$ and colours $\mathcal{C}' \cup \mathcal{C}(H_{\text{abs}})$, whose ends can be chosen arbitrarily within V'.

Let us briefly sketch the proof of Lemma 3.1. We first put aside a subset of the vertices and a subset of the colours, which are typically called the 'reservoir' or the 'flexible set', that have the following property: for any sets of vertices and colours V' and $\mathcal{C}'$ of the same small size (much smaller than the reservoir) which are disjoint from the reservoir, we can find a rainbow directed path Q_0 that uses $V', \mathcal{C}'$, and a set of $O(|V'|)$ vertices and colours from the reservoir. The question is then how to cover the rest of the reservoir; to this end, we build an 'absorbing structure' (H_{abs} above) which has the following property: it can 'absorb' any subset of vertices and colours of the same size of the reservoir in a rainbow directed path Q_{abs}. Then combining Q_{abs} and Q_0 gives Q.

The 'absorbing structure' H_{abs} is built by putting together several 'gadgets' or 'absorbers', graphs on $\Theta(1)$ vertices with the following property: each gadget has two directed paths with the same endpoints such that one avoids a designated vertex v and colour c in the reservoir, and the other one 'absorbs' v and c (see Fig. 1 in Section 5). This absorbing structure is based on one that was introduced by Gould et al. [16] for constructing rainbow Hamilton paths in

random optimal colourings of the (undirected) complete graph, which in turn is based on ideas of Montgomery [23].

3. Proof of Theorems 1.2 and 1.1

In this section we prove the main theorem, Theorem 1.2, and use it to deduce Theorem 1.1. We will use Lemma 3.1 below, which we prove in Section 6.

Lemma 3.1. *Let $1/n \ll 1/C \ll \eta \ll \gamma \ll \delta$. Let D_0 be a digraph on n vertices with minimum semidegree at least δn, and suppose that $D \sim D_0 \cup \mathbf{D}(n, C/n)$ is uniformly coloured in $\mathcal{C} = [n]$. Then, with high probability, D contains a digraph H_{abs} on at most γn vertices with the following property. For any $V' \subseteq V \setminus V(H_{abs})$, $\mathcal{C}' \subseteq \mathcal{C} \setminus \mathcal{C}(H_{abs})$ with $2 \leq |V'| = |\mathcal{C}'| \leq \eta n$ and distinct $x, y \in V'$, there exists a rainbow directed path Q such that*

- *Q is a path from x to y,*
- *$V(Q) = V(H_{abs}) \cup V'$,*
- *$\mathcal{C}(Q) = \mathcal{C}(H_{abs}) \cup \mathcal{C}'$.*

The next lemma is a rainbow directed version of a commonly used consequence of the depth first search algorithm [7], which we will use to find an almost spanning rainbow directed path. The undirected version of the lemma was used in [12, Lemma 2.17] in the binomial random graph and in [1, Proposition 2.1] for finding a Hamilton cycle in the undirected perturbed graph. Lemma 3.2 can be proven by following almost verbatim the proof of [1, Proposition 2.1].

Lemma 3.2 (cf. Proposition 2.1 [1] and Lemma 2.17 [12]). *Let D be an n-vertex edge-coloured digraph. If every two disjoint sets of vertices X, Y of size k satisfy $|\mathcal{C}(E(X, Y))| \geq n$, then D has a rainbow directed path of length at least $n - 2k$.*

The next lemma can easily be proved using Chernoff's bound (cf. Theorem 4.1).

Lemma 3.3. *Let $1/n \ll 1/C \ll \alpha \leq 1/2$ and let $D \sim \mathbf{D}(n, C/n)$ be uniformly coloured in $\mathcal{C} = [n]$. Then, with high probability, every two disjoint sets of vertices X, Y of size αn satisfy $|\mathcal{C}(E(X, Y))| \geq (1 - \alpha)n$.*

Our main theorem now follows easily from the previous three lemmas.

Proof of Theorem 1.2. Let η and γ satisfy

$$1/n \ll 1/C \ll \eta \ll \gamma \ll \delta,$$

and write $V = V(D)$ and $\mathcal{C} = [n]$. By Lemma 3.1 we may assume that there exists a subdigraph $H_{abs} \subseteq D$ with $|V(H_{abs})| \leq \gamma n$ and the properties stated in Lemma 3.1, so in particular $|\mathcal{C}(H_{abs})| = |V(H_{abs})| - 1$; and by Lemma 3.3 (on input $\alpha = \eta/4$) that any two disjoint subsets $X, Y \subseteq V$ of size $k = \eta n/4$ satisfy $|\mathcal{C}(E(X, Y))| \geq (1 - \eta/4)n = n - k$. Let $\mathcal{C}_2 = \mathcal{C} \setminus \mathcal{C}(H_{abs})$ and let V_2 be an arbitrary subset of $V \setminus V(H_{abs})$ of size $n - |V(H_{abs})| - k$. Then $|\mathcal{C}_2| \geq n - \gamma n$, $|V_2| = |\mathcal{C}_2| - k - 1$, and, for any two disjoint subsets $X, Y \subseteq V_2$ of size k,

$$|\mathcal{C}(E(X, Y)) \cap \mathcal{C}_2| \geq |\mathcal{C}_2| - k \geq |V_2|.$$

Then, by Lemma 3.2, applied to the subgraph of D spanned by edges in $D[V_2]$ coloured in $\mathcal{C}_2$, there is a rainbow directed path P_2 of length at least $|V_2| - 2k$ with vertices in V_2 and colours in $\mathcal{C}_2$. Let $V_2' = V \setminus (V(H_{abs}) \cup V(P_2))$ and $\mathcal{C}_2' = \mathcal{C} \setminus (\mathcal{C}(H_{abs}) \cup \mathcal{C}(P_2))$ be the set of vertices and colours used by neither H_{abs} nor P_2. Observe that $|V_2'| \leq 3k \leq \eta n - 2$ and $|\mathcal{C}_2'| = |V_2'| + 2$. Denote the first and last vertices of P_2 by x and y. Then by the property of H_{abs} there exists a rainbow directed path Q from y to x, with $V(Q) = V(H_{abs}) \cup V_2' \cup \{x, y\}$, and $\mathcal{C}(Q) = \mathcal{C}(H_{abs}) \cup \mathcal{C}_2'$. The concatenation of P_2 and Q gives a rainbow directed Hamilton cycle, as desired. □

Finally we prove Theorem 1.1.

Proof of Theorem 1.1. Let C be given by Theorem 1.2 such that for any digraph D_0 on $[n]$ with minimum semidegree at least $\delta n/4$, the perturbed digraph $D_0 \cup \mathbf{D}(n, \frac{C}{2n})$ uniformly coloured in $[n]$ has with high probability a rainbow directed Hamilton cycle. Let $N = \binom{n}{2}$ and $e_1, \ldots, e_N$ be an enumeration of the edges of the (undirected) complete graph on $[n]$. For each $0 \leq i \leq N$, define a randomly edge-coloured digraph Γ_i on $[n]$ as follows, recalling that G_0 is an n-vertex graph with minimum degree at least δn, and writing $e_j = \{x, y\}$.

a)   For $j > i$:

   1.   If $e_j \in E(G_0)$, then add both xy, yx to $E(\Gamma_i)$, and colour both edges with the same colour, chosen uniformly in $[n]$.
   2.   If $e_j \notin E(G_0)$, then add both xy, yx to $E(\Gamma_i)$ together with probability $\frac{C}{n}$, and colour both edges with the same colour, chosen uniformly in $[n]$.

b)   For $j \leq i$:

   1.   if $e_j \in E(G_0)$, toss a coin that comes up heads with probability $1/3$. If it comes up heads, add both xy, yx to $E(\Gamma_i)$ and colour each independently and uniformly at random in $[n]$. If it comes up tails, add each xy, yx independently with probability $\frac{C}{2n}$ to $E(\Gamma_i)$, and colour each independently and uniformly at random in $[n]$.
   2.   If $e_j \notin E(G_0)$, add each of xy, yx independently with probability $\frac{C}{2n}$ to $E(\Gamma_i)$, and colour each independently and uniformly at random in $[n]$.

The sequence is coupled, i.e. Γ_{i-1} and Γ_i agree as probability spaces (but not as digraphs) on all edges apart from e_i, for $i \in [N]$. Clearly $\Gamma_0 \sim G_0 \cup \mathbf{G}(n, C/n)$ and is uniformly coloured in $[n]$, where we view the undirected edges as two parallel directed edges. It is also easy to see that $\Gamma_N \sim D_0 \cup \mathbf{D}(n, \frac{C}{2n})$ and it is uniformly coloured in $[n]$, where D_0 is a random subgraph of G_0, with each edge selected independently (as two parallel edges) with probability $1/3$. Indeed, for $e = \{x, y\} \notin E(G_0)$, clearly each of xy, yx is in $E(\Gamma_N)$ independently with probability $\frac{C}{2n}$. For $e \in E(G_0)$, with probability $1/3$, both orientations of e are added to $E(D_0)$, and each is coloured independently and uniformly in $[n]$; otherwise each orientation is added independently to $E(\Gamma_0)$ with probability $\frac{C}{2n}$ and coloured independently and uniformly in $[n]$. A straightforward application of the union bound and Chernoff's bound implies that with high probability the minimum semidegree of D_0 is at least $\delta n/4$. Hence, by the choice of C, with high probability Γ_N has a directed rainbow Hamilton cycle. We will show that for each $i \in [N]$

$$\mathbb{P}\big[\Gamma_{i-1} \text{ has a rainbow Hamilton cycle}\big] \geq \mathbb{P}\big[\Gamma_i \text{ has a rainbow Hamilton cycle}\big], \qquad (1)$$

which then implies the theorem.

To this end, reveal the randomness on all edges apart from $e_i = xy$, i.e. on all those edges that Γ_{i-1}, Γ_i automatically agree on. There are three possibilities:

1. $\Gamma_{i-1} \cup \{xy, yx\}$ has no rainbow Hamilton cycle, irrespective of the colouring of xy, yx. Then, regardless of the outcome of e_i, both Γ_{i-1} and Γ_i have no rainbow Hamilton cycle.
2. Γ_{i-1} has a rainbow Hamilton cycle without using $\{xy, yx\}$. Then, regardless of the outcome of e_i, both Γ_{i-1} and Γ_i have a rainbow Hamilton cycle.
3.   Both events above do not hold. In other words, Γ_{i-1} has a rainbow Hamilton cycle using one of xy, yx, for some choice of colours, but it has no rainbow Hamilton cycle that avoids both of xy, yx. Let $\mathcal{C}_1, \mathcal{C}_2 \subseteq \mathcal{C}$, satisfy the following: $\Gamma_{i-1} \cup xy$ has a rainbow Hamilton cycle

through xy if and only if $\mathcal{C}(xy) \in \mathcal{C}_1$, and $\Gamma_{i-1} \cup yx$ has a rainbow Hamilton cycle through yx if and only if $\mathcal{C}(yx) \in \mathcal{C}_2$. (So $\mathcal{C}_1 \cup \mathcal{C}_2 \neq \emptyset$.)

Conditioning on either of the first two cases, both Γ_{i-1}, Γ_i have the same probability to have a rainbow Hamilton cycle. Conditioning on the third case, the probability that Γ_{i-1} has a rainbow Hamilton cycle is

$$
\begin{cases}
\frac{|\mathcal{C}_1 \cup \mathcal{C}_2|}{n} & \text{if } e_i \in E(G_0), \\
\frac{C}{n} \cdot \frac{|\mathcal{C}_1 \cup \mathcal{C}_2|}{n} & \text{if } e_i \notin E(G_0).
\end{cases}
$$

and the probability that Γ_i has a rainbow Hamilton cycle is

$$
\begin{cases}
\frac{1}{3} \cdot \left(1 - \left(1 - \frac{|\mathcal{C}_1|}{n}\right)\left(1 - \frac{|\mathcal{C}_2|}{n}\right)\right) + \frac{2}{3} \cdot \left(\frac{C}{2n} \cdot \frac{|\mathcal{C}_1| + |\mathcal{C}_2|}{n} - \frac{C^2}{4n^2} \cdot \frac{|\mathcal{C}_1||\mathcal{C}_2|}{n^2}\right) \le \frac{|\mathcal{C}_1| + |\mathcal{C}_2|}{2n} \le \frac{|\mathcal{C}_1 \cup \mathcal{C}_2|}{n} & \text{if } e_i \in E(G_0), \\
\frac{C}{2n} \cdot \frac{|\mathcal{C}_1| + |\mathcal{C}_2|}{n} - \frac{C^2}{4n^2} \cdot \frac{|\mathcal{C}_1||\mathcal{C}_2|}{n^2} \le \frac{C}{n} \cdot \frac{|\mathcal{C}_1| + |\mathcal{C}_2|}{2n} \le \frac{C}{n} \cdot \frac{|\mathcal{C}_1 \cup \mathcal{C}_2|}{n} & \text{if } e_i \notin E(G_0).
\end{cases}
$$

This shows that in either case the probability that Γ_{i-1} has a rainbow Hamilton cycle is at least as large as the probability that Γ_i has one, thereby proving (1) and thus the theorem. □

4. Preliminaries

Next we collect three preliminary results that we need: the Chernoff bound, Theorem 4.1; that random sparse subgraph of dense hypergraphs have large matchings, Lemma 4.2; and that in the perturbed digraph, between any two vertices, there is a large rainbow collection of directed paths of length three, Lemma 4.3.

Theorem 4.1 (Chernoff Bound, [17, eq. (2.8) and Theorem 2.8]). *For every $\varepsilon > 0$ there exists $c_\varepsilon > 0$ such that the following holds. Let X be the sum of mutually independent indicator random variables and write $\mu = \mathbb{E}[X]$. Then*

$$
\mathbb{P}[|X - \mu| \ge \varepsilon\mu] \le 2\exp(-c_\varepsilon\,\mu).
$$

The next lemma, despite its technical appearance, proves the following straightforward statement: quite sparse random subgraphs of dense hypergraphs contain, with very high probability, a matching of linear size. A *matching* in a hypergraph is a collection of pairwise vertex-disjoint edges. The *degree* of a vertex v is the number of edges incident to v.

Lemma 4.2. *Let $1/n \ll \rho \ll \alpha, c, 1/r$ where $r \ge 2$ is an integer. Let $\mathcal{H}$ be an r-uniform hypergraph on n vertices with at least αn^r edges.*

Writing $m = cn$, let $\mathcal{H}_m$ be the random subgraph of $\mathcal{H}$ that consists of m edges of $\mathcal{H}$, chosen with replacement and uniformly at random. Then, with probability at least $1 - \exp\left(-\frac{c\alpha^2 n}{3}\right)$, the hypergraph $\mathcal{H}_m$ has a matching of size at least ρn.

Writing $p = cn^{-r+1}$, let $\mathcal{H}_p$ be the random subgraph of $\mathcal{H}$ where we keep each edge independently with probability p. Then, with probability at least $1 - \exp\left(-\frac{c\alpha^2 n}{3r}\right)$, the hypergraph $\mathcal{H}_p$ has a matching of size at least ρn.

Proof. Write $\beta(\mathcal{G})$ for the size of a largest matching of a hypergraph $\mathcal{G}$.

It is not hard to see that $\mathcal{H}$ contains an induced subgraph $\mathcal{H}'$ of minimum degree at least αn^{r-1} on at least αn vertices.

We first prove the result for $\mathcal{H}_m$. Let $\mathcal{H}'_m$ be the hypergraph with vertices $V(\mathcal{H}')$ and edges $E(\mathcal{H}_m) \cap E(\mathcal{H}')$. We will show that, with high probability, $\mathcal{H}'_m$ has a matching of size at least ρn. Clearly $\beta(\mathcal{H}_m) \ge \beta(\mathcal{H}'_m)$, so the lemma readily follows.

Suppose $\beta(\mathcal{H}'_m) < \rho n$, and let M be a maximal matching in $\mathcal{H}'_m$. Then $S = V(\mathcal{H}') \setminus V(M)$ is an independent set in $\mathcal{H}'_m$ and $|S| \geq (\alpha - r\rho)n$. By the minimum degree condition of $\mathcal{H}'$, the number of edges with all vertices in S is at least $\frac{1}{r}|S|(\alpha n^{r-1} - r\rho n^{r-1}) \geq \frac{1}{r}(\alpha - r\rho)(\alpha - r\rho)n^r \geq \frac{\alpha^2}{2r}n^r$. This gives

$$\mathbb{P}[S \text{ is independent in } \mathcal{H}'_m] = \left(1 - \frac{e(\mathcal{H}[S])}{e(\mathcal{H})}\right)^m \leq \exp\left(-cn \cdot \frac{\frac{\alpha^2}{2r}n^r}{\binom{n}{r}}\right)$$

$$\leq \exp\left(-\frac{\alpha^2}{2}(r-1)! \, cn\right).$$

Then, since $\alpha, r\rho < 1/2$, the number of $S \subseteq V(\mathcal{H})$ with $|S| \geq (\alpha - r\rho)n$ is at most

$$n\binom{n}{(\alpha - r\rho)n} \leq n\binom{n}{r\rho n} \leq n \cdot \left(\frac{e}{r\rho}\right)^{r\rho n} \leq e^{2r\rho n}(r\rho)^{-r\rho n} = \exp\left((2r\rho - r\rho \log(r\rho))n\right).$$

Thus, by the union bound, $\mathbb{P}[\beta(\mathcal{H}'_m) < \rho n] \leq \exp(-f_{c,r,\alpha}(\rho)n)$, where

$$f_{c,r,\alpha}(\rho) = -2r\rho + r\rho \log(r\rho) + \frac{(r-1)!}{2}c\alpha^2.$$

Since $f_{c,r,\alpha}(\rho)$ is continuous near 0 and $f_{c,r,\alpha}(\rho) \to \frac{(r-1)!}{2}c\alpha^2$ as $\rho \to 0$, for ρ sufficiently small $f_{c,r,\alpha}(\rho) \geq \frac{(r-1)!}{3}c\alpha^2 \geq \frac{c\alpha^2}{3}$, which gives the first part of the lemma.

For the second part of the lemma observe that the same argument works: with S as above, in $\mathcal{H}_p$ we have

$$\mathbb{P}[S \text{ is independent in } \mathcal{H}_p] = (1-p)^{e(\mathcal{H}[S])} \leq e^{-pe(\mathcal{H}[S])} \leq \exp\left(-\frac{c\alpha^2 n}{2r}\right)$$

and a similar calculation as above shows that, for ρ sufficiently small, the probability there is such an S is at most $\exp\left(-\frac{c\alpha^2 n}{3r}\right)$. $\qquad\square$

Lemma 4.3 (Triangles and short paths). *Let $1/n \ll 1/C \ll \lambda \ll \rho \ll \delta, q$. Let $\mathcal{C}$ be a set of colours of size qn, let D_0 be a digraph on n vertices with minimum semidegree at least δn, and suppose that $D \sim D_0 \cup \mathbf{D}(n, C/n)$ is uniformly coloured in $\mathcal{C}$. Then, with probability at least $1 - \exp(-\lambda n)$, the following holds. For any two vertices $u, v \in V(D)$ there is a matching M of size at least ρn such that $\bigcup_{xy \in M}\{ux, xy, yv\}$ is rainbow.*

Moreover, with probability at least $1 - \exp(-\lambda n)$, for any $u \in V(D)$ there is a matching M of size at least ρn such that $\bigcup_{xy \in M}\{ux, xy, yu\}$ is rainbow.

Proof. Let ρ_1 satisfy $1/C \ll \lambda \ll \rho \ll \rho_1 \ll \delta, q$. Fix $u, v \in V$. By the minimum semidegree assumption on D_0, there exist disjoint subsets $N_u \subseteq N_{D_0}^+(u)$, $N_v \subseteq N_{D_0}^-(v)$, each of size $\delta n/2$. Consider the oriented bipartite graph with bipartition (N_u, N_v) and edges

$$\{zw \in E(\mathbf{D}(n, C/n)) : z \in N_u, \ w \in N_v\}$$

Then, by Lemma 4.2 (applied with $(\alpha, c, r, n)_{4.2} = (1/4, \delta C, 2, \delta n)$), with probability $1 - e^{-\Omega(\delta Cn)}$, it has a matching M of size $\rho_1 n$. For each $zw \in M$, reveal whether the directed path $uzwv$ is rainbow, without exposing the colours. Then each $uzwv$ is rainbow independently with probability $1 - o(1)$. Hence by Chernoff's bound (Theorem 4.1), with probability $1 - e^{-\Omega(\rho_1 n)}$, there is $M' \subseteq M$ with $|M'| \geq \rho_1 n/2$ such that for each $zw \in M'$, $uzwv$ is rainbow. Let $\mathcal{P}' = \{uzwv : zw \in M'\}$, and observe that the same colour may still repeat on different paths in $\mathcal{P}'$.

We now show that we can find a large subset of $\mathcal{P}'$ where the paths are pairwise colour-disjoint. Reveal the colours on the edges of the paths in $\mathcal{P}'$. By symmetry, each triple of distinct colours in $\mathcal{C}$ is equally likely to appear in $\mathcal{P}'$. Hence $\mathcal{P}'$ corresponds to selecting uniformly at random with replacement $|\mathcal{P}'| \geq \rho_1 n/2$ edges from the complete 3-graph with vertex set $\mathcal{C}$. Thus, by Lemma 4.2

(applied with $(\alpha, c, r, n)_{4.2} = (1/7, \rho_1/(2q), 3, qn))$, with probability $1 - e^{-\Omega(\rho_1 n)}$, this 3-graph has a matching of size ρn. This corresponds to an $M'' \subseteq M'$ of size ρn so that $\bigcup_{xy \in M''}\{ux, xy, yv\}$ is rainbow.

The probability this fails for some pair u, v is, by the union bound, at most

$$n^2 \cdot \left(e^{-\Omega(\delta C n)} + e^{-\Omega(\rho_1 n)} + e^{-\Omega(\rho_1 n)} \right) \leq e^{-\lambda n},$$

proving the first statement of the lemma.

The second statement of the lemma follows similarly by finding a large matching between disjoint subsets of $N_{D_0}^-(u)$ and $N_{D_0}^+(u)$. □

5. Finding absorbers

In this section we show how to find 'absorbers', which are the building blocks for the digraph H_{abs} in Lemma 3.1.

Definition 5.1 (*Absorber*). *Let v be a vertex and c a colour. A (v, c)-absorber is an edge-coloured digraph $A_{v,c}$ with $v \in V(A_{v,c})$ and $c \in \mathcal{C}(A_{v,c})$ that has two directed paths P, P' with the following properties:*

- *they are rainbow;*
- *they have the same first and last vertex;*
- *P is spanning in $V(A_{v,c})$ and $V(P') = V(P) \setminus \{v\} = V(A_{v,c}) \setminus \{v\}$;*
- *P is spanning in $\mathcal{C}(A_{v,c})$ and $\mathcal{C}(P') = \mathcal{C}(P) \setminus \{c\} = \mathcal{C}(A_{v,c}) \setminus \{c\}$.*

We call P the (v, c)-absorbing path and P' the (v, c)-avoiding path. The internal vertices of $A_{v,c}$ are $V(A_{v,c}) \setminus \{v\}$ and the internal colours are $\mathcal{C}(A_{v,c}) \setminus \{c\}$. The first and last vertices of the absorber are the first and last vertices of P and P'.

The aim of this section is to prove the following lemma, which says that for any vertex v, colour c and any small (but linear in size) sets of forbidden vertices and colours, we can find a (v, c)-absorber.

Lemma 5.2. *Let $1/n \ll 1/C \ll v \ll \delta < 1/2$. Let D_0 be a digraph on n vertices with minimum semidegree at least δn, and suppose that $D \sim D_0 \cup \mathbf{D}(n, C/n)$ is uniformly coloured in $\mathcal{C} = [n]$. Then, with high probability, the following holds. For any $v \in V(D)$ and $c \in \mathcal{C}$, and for all $V' \subseteq V(D)$ and $\mathcal{C}' \subseteq \mathcal{C}$ each of size at least $(1 - v)n$, there exists a (v, c)-absorber on 13 vertices with internal vertices in V' and internal colours in $\mathcal{C}'$.*

Our absorbers are depicted in Fig. 1. We describe their structure here, though it might be easier to read off the structure from the figure. They consist of vertices $\{v, v_1, v_2\}$ that induce an oriented triangle with edges $\{v_1 v, vv_2, v_1 v_2\}$, and vertices $\{x, y, z, u, w_1, w_2\}$ inducing an oriented $K_{2,4}$ with edges $\{xy, xz, yu, zu, w_2 y, w_2 z, yw_1, zw_1\}$, and two directed paths of length 3 – P_1 from v_2 to x and P_2 from w_1 to w_2 – whose interiors are vertex-disjoint and disjoint of previously mentioned vertices. The absorber is equipped with an edge colouring satisfying the following: $\mathcal{C}(yu) = c$, $\mathcal{C}(v_1 v) = \mathcal{C}(xy)$, $\mathcal{C}(vv_2) = \mathcal{C}(zu)$, $\mathcal{C}(v_1 v_2) = \mathcal{C}(xz)$, $\mathcal{C}(w_2 y) = \mathcal{C}(w_2 z)$, $\mathcal{C}(yw_1) = \mathcal{C}(zw_1)$, and the colours of the edges of P_1 and P_2 are different from one another and from those on the rest of the absorber. The (v, c)-absorbing path P and (v, c)-avoiding path P' are

$$P = v_1 v v_2 P_1 x z w_1 P_2 w_2 y u, \qquad P' = v_1 v_2 P_1 x y w_1 P_2 w_2 z u.$$

For the existence of the oriented triangle (on vertices $\{v, v_1, v_2\}$) and the directed paths of length three P_1 and P_2 we use Lemma 4.3. The existence of the oriented $K_{2,4}$ (with the required colouring)
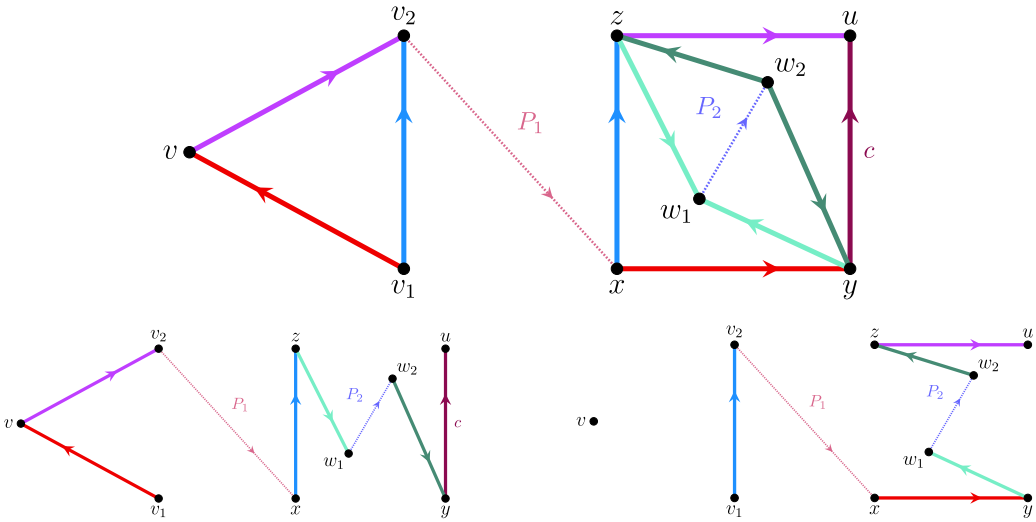
Figure 1. At the top is the (v, c)-absorber used in Lemma 5.2. At the bottom the first figure shows the (v, c)-absorbing path and the second figure the (v, c)-avoiding path. The directed paths P_1, P_2 have length 3 and are rainbow with colours disjoint of one another and of the other colours in the figure.

is more involved and uses the regularity lemma for digraphs. This is the most technical part of our proof and is accomplished in Lemma 5.11.

5.1 Finding oriented squares with diagonal directed paths

The aim of this section is to prove Lemma 5.11, which allows us to find a $K_{2,4}$ as required in Fig. 1.

5.1.1 Regularity preliminaries

The (directed) *densities* of a pair of non-empty disjoint vertex sets V, W in a digraph are $d^+(V, W) = \frac{e^+(V,W)}{|V||W|}$ and $d^-(V, W) = \frac{e^-(V,W)}{|V||W|}$. For simplicity $d(V, W)$ will always stand for $d^+(V, W)$.

Definition 5.3. *An ordered pair of disjoint vertex sets* (V, W) *is ε-regular if for any $V' \subseteq V$, $W' \subseteq W$ with $|V'| \geq \varepsilon |V|$, $|W'| \geq \varepsilon |W|$ we have $|d(V', W') - d(V, W)| < \varepsilon$.*

Since we need regular pairs to be sufficiently dense, we will find the following definition useful.

Definition 5.4. *An ordered pair of disjoint vertex sets (V, W) is (ε, ρ)-super-regular if it is ε-regular and for every $v \in V$, $\deg^+(v, W) \geq \rho |W|$ and for every $w \in W$, $\deg^-(w, V) \geq \rho |V|$.*

We state without proof the following straightforward consequences of Definition 5.4.

Lemma 5.5. *If (V, W) is ε-regular, then for all but at most $\varepsilon |V|$ vertices $v \in V$ we have $\deg^+(v, W) \geq (d^+(V, W) - \varepsilon) |W|$, and for all but at most $\varepsilon |W|$ vertices $w \in W$ we have $\deg^-(w, V) \geq (d^-(V, W) - \varepsilon) |V|$.*

Lemma 5.6. *Let $1/n \ll \varepsilon, \rho, p$. Let (V, W) be (ε, ρ)-super-regular with $|V| = |W| = n$. Sample each edge independently with probability p. Then, with probability at least $1 - \exp(-\Omega(p\rho n))$, the resulting pair is $(\varepsilon, p\rho/2)$-super-regular.*

We use the following degree form of the regularity lemma for digraphs, from [18]. The original version of Szemerédi's regularity lemma [25] for digraphs was first proved by Alon and Shapira [3]. The version we use is Lemma 7 in [18].

Theorem 5.7 (Degree form of the regularity lemma for digraphs). *For every $\varepsilon \in (0, 1)$ and every integer M', there are integers M and n_0 such that the following holds. If D is a digraph on $n \geq n_0$ vertices and $d \in [0, 1]$ is any real number, then there is a partition $(V_0, \ldots, V_k)$ of $V(D)$, a spanning subgraph D' of D and a set U of ordered pairs (i, j), $1 \leq i, j \leq k$ and $i \neq j$, such that the following hold:*

- *$M' \leq k \leq M$;*
- *$|V_0| \leq \varepsilon n$;*
- *$|V_1| = \cdots = |V_k|$;*
- *for all $x \in V(D)$, $\deg_{D'}^+(x) > \deg_D^+(x) - (d + \varepsilon)n$ and $\deg_{D'}^-(x) > \deg_D^-(x) - (d + \varepsilon)n$;*
- *$|U| \leq \varepsilon k^2$;*
- *for every ordered pair $(i, j) \notin U$ with $i \neq j$ the bipartite digraph $D[V_i, V_j]$ is ε-regular;*
- *D' is obtained from D by deleting the following edges of D: all edges with both endpoints in V_i, for all $i \in [k]$; all edges $E_D(V_i, V_j)$ for all $(i, j) \in U$; and all edges $E_D(V_i, V_j)$ for all $(i, j) \notin U$, $i \neq j$, with $d_D(V_i, V_j) < d$.*

We call $V_1, \ldots, V_k$ the *clusters* of the partition and V_0 the *exceptional set*. Note that the last two conditions of the lemma imply that for all $1 \leq i, j \leq k$ with $i \neq j$, the bipartite graph $D'[V_i, V_j]$ is ε-regular and has density either 0 or at least d.

5.1.2 Finding gadgets

In the following lemma we show that, if (A, B) and (C, B) are two dense regular pairs in a digraph whose edges are uniformly coloured, and given a dense bipartite graph H with bipartition (A, C), with very high probability there is a dense subgraph $H' \subseteq H$, such that for every edge ac in H' there is an oriented path of length 2 from a to c through B which is monochromatic (and of course both edges are directed outwards from a and c).

By applying this lemma twice (reversing edges for the second application, see the proof of Lemma 5.11), given four dense regular pairs $(W_2, Y), (W_2, Z), (Y, W_1), (Z, W_1)$ in a digraph whose edges are uniformly coloured, with very high probability we can find $(y, z, w_1, w_2) \in Y \times Z \times W_1 \times W_2$ such that $\mathcal{C}(w_2 y) = \mathcal{C}(w_2 z) \neq \mathcal{C}(yw_1) = \mathcal{C}(zw_1)$, [3] in line with the edges spanned by $\{y, z, w_1, w_2\}$ in our gadgets.

Lemma 5.8. *Let $1/n \ll \gamma, \varepsilon, \lambda \ll \alpha, \rho, 1/q$. Let D be a digraph, and let A, B, C be disjoint sets of vertices in D of size n, such that $(A, B), (C, B)$ are ε-regular and have density at least ρ. Let H be a digraph with $E(H) \subseteq A \times C$ and $e(H) \geq \alpha n^2$. Assign each edge in D a colour in $\mathcal{C} = [qn]$ uniformly and independently. Then, with probability at least $1 - \exp(-\lambda n)$, the number of edges ac in H for which there exists $b \in B$ such that $ab, cb \in E(D)$ and $\mathcal{C}(ab) = \mathcal{C}(cb)$, is at least γn^2.*

Proof. Let β be a constant satisfying $\gamma, \lambda \ll \beta \ll \alpha, \rho, 1/q$.

Let H_1 be the subgraph of H spanned by edges ac such that $|N^+(a) \cap N^+(c) \cap B| \geq \rho^2 n/2$. We claim that $e(H_1) \geq \alpha n^2/2$. Indeed, let $A' = \{a \in A : |N^+(a) \cap B| \geq (\rho - \varepsilon)n\}$. Then $|A'| \geq (1 - \varepsilon)n$ by ε-regularity. Given $a \in A'$, we have that $|N^+(a) \cap N^+(c) \cap B| \geq (\rho - \varepsilon)^2 n \geq \rho^2 n/2$ for all but at most εn vertices $c \in C$, using $\varepsilon \ll \rho$ for the last inequality. Thus $e(H_1) \geq e(H) - |A \setminus A'| \cdot n - |A'| \cdot \varepsilon n \geq e(H) - 2\varepsilon n^2 \geq \alpha n^2/2$, using $\varepsilon \ll \alpha$.

---

[3] In fact, we can find quadratically many such quadruples, with distinct pairs (y, z).

Let $M_1, \ldots, M_k$ be a maximal collection of pairwise edge-disjoint matchings in H_1 of size at least $\alpha n/4$. We claim that $k \geq \alpha n/4$. Indeed, otherwise, since $|M_i| \leq n$ for every i, we have $e(H_1 \setminus (M_1 \cup \ldots \cup M_k)) \geq \alpha n^2/2 - \alpha n^2/4 = \alpha n^2/4$. But the graph $H_1 \setminus (M_1 \cup \ldots \cup M_k)$ can be decomposed into at most n matchings, showing that there is a matching M_{k+1} in $H \setminus (M_1 \cup \ldots \cup M_k)$ of size at least $\alpha n/4$, a contradiction.

Let M_i' be the (random) submatching $\{ac \in M_i : \text{there is } b \in B \text{ such that } \mathcal{C}(ab) = \mathcal{C}(cb)\}$. For $ac \in M_i$, we have

$$\mathbb{P}\big[ac \in M_i'\big] = 1 - \left(1 - \frac{1}{qn}\right)^{|N^+(a) \cap N^+(c) \cap B|}$$

$$\geq 1 - \exp\left(-\frac{|N^+(a) \cap N^+(c) \cap B|}{qn}\right)$$

$$\geq 1 - \exp\left(-\frac{\rho^2}{2q}\right) = 8\alpha^{-1}\beta.$$

Thus, $\mathbb{E}\big[|M_i'|\big] \geq |M_i| \cdot 8\alpha^{-1}\beta \geq 2\beta n$. Because, for each $i \in [k]$, the events $\{ac \in M_i'\}$ with $ac \in M_i$ are independent, by Chernoff's bound we have

$$\mathbb{P}\big[|M_i'| \leq \beta n\big] \leq \mathbb{P}\left[|M_i'| \leq \frac{\mathbb{E}\big[|M_i'|\big)}{2}\right] \leq \exp\big(-\Omega\left(\mathbb{E}\big[|M_i'|\big]\right)\big) \leq \exp(-\Omega(\beta n)).$$

By a union bound, with probability at least $1 - 2n\exp(-\Omega(\beta n)) \geq 1 - \exp(-\lambda n))$, we have $|M_i'| \geq \beta n$ for every $i \in [k]$, showing $|M_1' \cup \ldots \cup M_k'| \geq (\alpha n/4) \cdot \beta n = \gamma n^2$. This proves the lemma. $\square$

In the next lemma we construct the part spanned by $\{x, y, z, u\}$. Recall that we would like the colours of zu, xz, xy to match the colours of the edges vv_2, v_1v_2, v_1v in the oriented triangle in the absorber. To achieve this, we let $\mathcal{C}_0$ be a collection of $\Omega(n)$ triples of distinct colours, which are pairwise disjoint, and ensure that the colours of zu, xz, xy match the colours of one of the triples in $\mathcal{C}_0$.

Lemma 5.9. *Let $1/n \ll \lambda \ll \varepsilon \ll \alpha, \rho, q_0 \leq 1$ and let $1/n \ll \mu \leq 1$. Let X, Y, Z, U be disjoint sets of μn vertices in a digraph D such that $(X, Y), (X, Z), (Y, U), (Z, U)$ are (ε, ρ)-super-regular. Let H be a digraph with $E(H) \subseteq Y \times Z$ and $e(H) \geq \alpha \mu^2 n^2$. Let $\mathcal{C} = [n]$, let $c \in \mathcal{C}$, and let $\mathcal{C}_0$ be a set of pairwise disjoint triples of colours in $\mathcal{C} \setminus \{c\}$ of size at least $q_0 n$.*

Assign each edge in D a colour from $\mathcal{C}$, uniformly and independently. Then, with probability at least $1 - \exp(-\lambda n)$, there is a quadruple $(x, y, z, u) \in X \times Y \times Z \times U$ with $xy, xz, yu, zu \in E(D)$ and $yz \in E(H)$, such that $(\mathcal{C}(zu), \mathcal{C}(xz), \mathcal{C}(xy)) \in \mathcal{C}_0$ and $\mathcal{C}(yu) = c$.

Proof. Let $\gamma, \gamma_1, \gamma_3, \lambda'$ satisfy

$$1/n \ll \lambda \ll \lambda' \ll \gamma_3 \ll \gamma_1 \ll \gamma \ll \varepsilon, \mu.$$

Let D_0 be the subgraph of D consisting of edges in D contained in $X \times (Y \cup Z)$ or in $(Y \cup Z) \times U$. We find a quadruple (x, y, z, u) satisfying the requirements of the lemma, by revealing the colours of $E(D_0)$ in the order $E(Y, U), E(Z, U), E(X, Y), E(X, Z)$. At each step we find a linear size monochromatic matching between the corresponding vertex sets. To ensure that each new matching extends nicely previously chosen matchings, we will 'clean up' the graph before finding it.

First, let D_1 be the subgraph of D_0 obtained by removing edges $yu \in Y \times U$ with $|N_{D_0}^-(u) \cap N_{D_0}^+(y)| \leq (\alpha \rho \mu/4)n$. We claim that $e(D_1[Y, U]) \geq (\alpha \rho \mu^2/4)n^2$. Indeed, writing $Y' = \{y \in Y : d_H^+(y) \geq (\alpha/2)|Z|\}$, observe that $|Y'| \geq (\alpha/2)|Y|$ (otherwise $e(H) \leq (|Y \setminus Y'| \cdot (\alpha/2)|Z| + |Y'| \cdot |Z| < |Y| \cdot (\alpha/2)|Z| + (\alpha/2)|Y| \cdot |Z| = \alpha|Y||Z| = \alpha \mu^2 n^2$, a contradiction). For every $y \in Y'$,

all but at most $\varepsilon|U|$ vertices u in U have at least $(\rho - \varepsilon)|N_H^+(y)| \geq (\alpha\rho\mu/4)n$ in-neighbours in $N_H^+(y)$. This implies that $e(D_1[Y, U]) \geq e(D_0[Y', U]) - |Y'| \cdot \varepsilon|U| \geq (\rho - 2\varepsilon)|Y'||U| \geq (\rho/2) \cdot (\alpha/2) \cdot \mu^2 n^2 = (\alpha\rho\mu^2/4)n^2$.

Reveal the colours of $E_{D_1}(Y, U)$. Since $e(D_1[Y, U]) \geq (\alpha\rho\mu^2/4)n^2$, and each edge is coloured c independently with probability $1/n$, Lemma 4.2 yields that, with probability at least $1 - \exp(-\Omega(\lambda'n))$, there exists a matching $M \subseteq D_1[Y, U]$ of size at least γn with all edges coloured c.

For $(c_1, c_2, c_3) \in \mathcal{C}_0$, say an edge xz in $E_D(X, Z)$ is *good* for $(c_1, c_2, c_3) \in \mathcal{C}_0$, if there exist $u \in U$ and $y \in Y$ such that $yz \in E(H)$, $xy, yu, zu \in E(D)$, $\mathcal{C}(xy) = c_3$, $\mathcal{C}(yu) = c$, and $\mathcal{C}(zu) = c_1$. Notice that this definition does not depend on the colours of $E_D(X, Z)$. Observe also that if xz is good for $(c_1, c_2, c_3) \in \mathcal{C}_0$ and $\mathcal{C}(xz) = c_2$, then (x, y, z, u) satisfies the requirements of the lemma. Therefore, to complete the proof it suffices to show that, with high probability, for some $(c_1, c_2, c_3) \in \mathcal{C}_0$ there exists at least one good edge $xz \in E_D(X, Z)$ with $\mathcal{C}(xz) = c_2$. For that, the crucial argument is the following claim, where we show that, with very high probability, there are many good edges for (c_1, c_2, c_3). We do so by 'extending' M with a monochromatic matching $M_1 \subseteq D_1[Z, U]$ in colour c_1, and a monochromatic matching $M_3 \subseteq D_1[X, Y]$ in colour c_3.

For $(c_1, c_2, c_3) \in \mathcal{C}_0$, write

$$G(c_1, c_2, c_3) := \{e \in E_D(X, Z) : e \text{ is good for } (c_1, c_2, c_3)\}.$$

Claim 5.10. *Fix* $(c_1, c_2, c_3) \in \mathcal{C}_0$. *With probability at least* $1 - \exp(-\Omega(\lambda'n))$, *we have* $|G(c_1, c_2, c_3)| \geq \gamma_3 n$.

Proof. For $y \in Y \cap V(M)$, denote by $M(y)$ the neighbour of y in M. Let D_2 be the subdigraph of D_1 obtained by removing all edges in $Z \times U$ except edges of the form $zM(y)$ satisfying $yz \in E(H)$ and $|N^-(y) \cap N^-(z) \cap X| \geq (\rho^2\mu/2)n$, for some $y \in Y$.

We claim that $e(D_2[Z, U]) \geq (\alpha\rho\mu\gamma/8)n^2$. Indeed, given $y \in Y \cap V(M)$, by (ε, ρ)-super-regularity, we have $|N_{D_1}^-(y) \cap X| \geq \rho|X|$, and all but at most $\varepsilon|Z|$ vertices $z \in Z$ satisfy

$$|N_{D_1}^-(z) \cap N_{D_1}^-(y) \cap X| \geq (\rho - \varepsilon)|N_{D_1}^-(y) \cap X| \geq (\rho^2/2)|X| = (\rho^2\mu/2)n.$$

Thus, using that $|N_{D_1}^-(M(y)) \cap N_H^+(y) \cap Z| \geq (\alpha\rho\mu/4)n = (\alpha\rho/4)|Z|$, by choice of D_1,

$$\left|N_{D_2}^-(M(y)) \cap Z\right| \geq \left|N_{D_1}^-(M(y)) \cap N_H^+(y) \cap Z\right| - \varepsilon|Z|$$

$$\geq (\alpha\rho/4 - \varepsilon)|Z| \geq (\alpha\rho/8)|Z| = (\alpha\rho\mu/8)n.$$

Hence, $e(D_2[Z, U]) \geq |M| \cdot (\alpha\rho\mu/8)n \geq (\alpha\rho\mu\gamma/8)n^2$, as claimed. Reveal the colours of $E_{D_2}[Z, U]$. Since each edge is coloured c_1 independently with probability $1/n$, using $e(D_2[Z, U]) \geq (\alpha\rho\mu\gamma/8)n^2$, by Lemma 4.2, with probability at least $1 - \exp(-\Omega(\lambda'n))$, there is a matching M_1 in $D_2[Z, U]$ coloured c_1 that has size at least $\gamma_1 n$.

Denote by $M_1(y)$ the neighbour of $M(y)$ in M_1 (if it exists). Notice that, by the definition of D_2, the matching M_1 'extends' the matching M in the sense that $V(M_1) \cap U \subseteq V(M)$, and, moreover, for every $y \in Y$ such that $M_1(y)$ exists, so does $M(y)$, and $yM_1(y) \in E(H)$.

Next, we will find a large matching M_3 coloured c_3 which, along with M_1 and M, will give us a large number of good edges for (c_1, c_2, c_3). To this end, let D_3 be the spanning subgraph of D_2 obtained by removing all edges $xy \in X \times Y$ except those for which $M_1(y)$ exists and $xM_1(y)$ is an edge in D_2.

We claim that $e(D_3[X, Y]) \geq (\rho^2\mu\gamma_1/2)n^2$. Indeed, let $y \in Y$ be such that $M(y)$ and $M_1(y)$ exist. Then $|N_{D_2}^-(y) \cap N_{D_2}^-(M_1(y)) \cap X| \geq (\rho^2\mu/2)n$, by choice of D_2, showing $|N_{D_3}^-(y)| \geq (\rho^2\mu/2)n$. Hence, $e(D_3[X, Y]) \geq |M_1| \cdot (\rho^2\mu/2)n \geq (\rho^2\mu\gamma_1/2)n^2$.

Reveal the random colouring of $D_3[X, Y]$. Since each edge of $D_3[X, Y]$ is coloured c_3 independently with probability $1/n$, using Lemma 4.2, with probability at least $1 - \exp(-\Omega(\lambda'n))$, there exists a matching M_3 in $D_3[X, Y]$ coloured c_3 that has size at least $\gamma_3 n$.

For $y \in Y$, denote by $M_3(y)$ the neighbour of y in M_3 (if it exists). Notice that, by the definition of D_3, the matching M_3 'extends' the matching M_1 in the sense that if $M_3(y)$ exists then so does $M_1(y)$ and $M_1(y)M_3(y) \in E(D_3)$.

Let

$$G_0(c_1, c_2, c_3) := \{M_3(y)M_1(y) : y \in Y \text{ and } M_1(y), M_3(y) \text{exist}\}.$$

Notice that $|G_0(c_1, c_2, c_3)| = |M_3| \geq \gamma_3 n$. Moreover, $M_3(y)M_1(y)$ is good for (c_1, c_2, c_3), as can be seen by taking $u = M(y)$, since we have: $yM_1(y) \in E(H)$; $M_3(y)y, yM(y), M_1(y)M(y) \in E(D)$; $\mathcal{C}(M_3(y)y) = c_3$, $\mathcal{C}(yM(y)) = c$, $\mathcal{C}(M_1(y)M(y)) = c_1$. Hence $|G(c_1, c_2, c_3)| \geq |G_0(c_1, c_2, c_3)| \geq \gamma_3 n$, with probability at least $1 - \exp(-\Omega(\lambda n))$, as required for the claim. □

By the union bound over the $q_0 n$ triples $(c_1, c_2, c_3) \in \mathcal{C}_0$, Claim 5.10 implies that, with probability at least $1 - \exp(-\Omega(\lambda' n))$, for each $(c_1, c_2, c_3) \in \mathcal{C}_0$, we have $|G(c_1, c_2, c_3)| \geq \gamma_3 n$. We finally show that, with probability at least $1 - \exp(-\Omega(\lambda' n))$, for some $(c_1, c_2, c_3) \in \mathcal{C}_0$ there exists $e \in G(c_1, c_2, c_3)$ satisfying $\mathcal{C}(e) = c_2$.

For $e \in E_{D_3}(X, Z)$, let

$$G'(e) := \{c_2 \in \mathcal{C} : \text{there exist } c_1, c_3 \text{ such that } (c_1, c_2, c_3) \in \mathcal{C}_0 \text{ and } e \in G(c_1, c_2, c_3)\}.$$

Then, using that no two triples in $\mathcal{C}_0$ share a colour, we have

$$\sum_{e \in E_D(X,Z)} |G'(e)| = \sum_{(c_1, c_2, c_3) \in \mathcal{C}_0} |G(c_1, c_2, c_3)| \geq |\mathcal{C}_0| \cdot \gamma_3 n = q_0 \gamma_3 n^2.$$

Now we reveal the colours of $E_{D_3}(X, Z)$. For $e \in E_{D_3}(X, Z)$, let A_e be the event that e gets a good colour, i.e. $\mathcal{C}(e) \in G'(e)$. Each edge is coloured independently, so the events A_e are mutually independent, and it holds that $\mathbb{P}[A_e] = |G'(e)|/n$. Hence, the probability that no event A_e holds is

$$\prod_{e \in E_{D_3}(X,Z)} (1 - \mathbb{P}[A_e]) \leq \exp\left(-\sum_{e \in E_{D_3}(X,Z)} \mathbb{P}[A_e]\right)$$

$$= \exp\left(-\sum_{e \in E_{D_3}(X,Z)} \frac{|G'(e)|}{n}\right) \leq \exp(-q_0 \gamma_3 n).$$

Hence, with probability at least $1 - \exp(-q_0 \gamma_3 n)$, at least one edge gets a good colour, i.e. there exists $e \in E_{D_3}(X, Z)$ such that $\mathcal{C}(e) \in G'(e)$, as required for the lemma. Altogether, all the required events hold simultaneously with probability at least $1 - \exp(-\Omega(\lambda' n)) - \exp(-q_0 \gamma_3 n) \geq 1 - \exp(-\lambda n)$, as required. □

Next, we show how to find a copy of $K_{2,4}$ with the orientation in Fig. 1. As above, we additionally make sure that the edges zu, xz, xy are coloured according to a triple of colours from a linear set $\mathcal{C}_0$ of pairwise disjoint triples of distinct colours. When applying Lemma 5.11 to find a (v, c)-absorber (cf. Lemma 5.2 and Section 6), $\mathcal{C}_0$ will be precisely the set of colour triples seen on a collection of rainbow triangles intersecting only on v.

Lemma 5.11 (Finding a $K_{2,4}$). *Let $1/n \ll \lambda \ll \delta, q_0$. Let $\mathcal{C} = [n]$, $c \in \mathcal{C}$, and $\mathcal{C}_0 \subseteq \mathcal{C}^3$ be a collection of colour triples that are pairwise disjoint and avoid c, with $|\mathcal{C}_0| = q_0 n$. Let D be a digraph on n vertices with minimum semidegree at least δn, which is coloured uniformly in $\mathcal{C}$.*

Then, with probability at least $1 - \exp(-\lambda n)$, the following holds. There exists an oriented copy K of $K_{2,4}$, with $V(K) = \{x, y, z, u, w_1, w_2\}$ and $E(K) = \{xz, xy, yu, zu, zw_1, yw_1, w_2z, w_2y\}$. Moreover, the edges of K are coloured as follows: $(\mathcal{C}(zu), \mathcal{C}(xz), \mathcal{C}(xy)) \in \mathcal{C}_0$, $c = \mathcal{C}(yu)$, and $\mathcal{C}(yw_1) = \mathcal{C}(zw_1) \neq \mathcal{C}(w_2y) = \mathcal{C}(w_2z)$, with $\mathcal{C}(yw_1), \mathcal{C}(w_2y)$ disjoint from $\{c, \mathcal{C}(zu), \mathcal{C}(xz), \mathcal{C}(xy)\}$.

Proof. Let $M', M, \alpha', \alpha'', \lambda'$ satisfy

$$\lambda \ll \lambda' \ll 1/M \ll 1/M' \ll \varepsilon \ll \alpha'' \ll \alpha' \ll \rho \ll \delta, q_0.$$

Consider a partition $(V_0, \ldots, V_k)$, with $M' \leq k \leq M$, of $V(D)$ and a spanning subgraph D', both given by Theorem 5.7 with parameters $\varepsilon_{5.7} = \varepsilon$, $d_{5.7} = \rho$ and M'. Let R be the reduced digraph associated with this partition, that is the digraph on $[k]$ where $ij \in E(R)$ if and only if $e_{D'}(V_i, V_j) > 0$. Observe that if $ij \in E(R)$ then $D'[V_i, V_j]$ is ε-regular and has density at least ρ.

Claim 5.12. *The digraph R contains a copy of K.*

Proof. We observe that the minimum semidegree of R satisfies $\delta^+(R) \geq \delta k/2$ because, otherwise, there would be vertices with out-degree at most $(\delta k/2) \cdot (n/k) < \delta n - (\rho + \varepsilon)n$ in D', contradicting that $\deg_{D'}^+(x) > \deg_D^+(x) - (\rho + \varepsilon)n$ for all $x \in V(D)$.

We will use, twice, the following assertion.

($\circledast$) If U is a set of at least $4/\delta$ vertices in R, then there are two distinct vertices $u_1, u_2 \in U$ such that $|N^+(u_1) \cap N^+(u_2)| \geq \frac{\delta^2 k}{10}$.

First, let us prove ($\circledast$). Suppose that U is a set contradicting ($\circledast$). Write $\ell = |U|$, and note that we may assume $\ell = \lceil 4/\delta \rceil$, so $4/\delta \leq \ell \leq 5/\delta$. Then,

$$k \geq \left| \bigcup_{u \in U} N^+(u) \right| \geq \sum_{u \in U} |N^+(u)| - \sum_{u,v \in U \text{ distinct}} |N^+(u) \cap N^+(v)|$$

$$\geq \ell \cdot \frac{\delta k}{2} - \binom{\ell}{2} \frac{\delta^2 k}{10}$$

$$= \frac{\ell \delta k}{2} \left(1 - \frac{(\ell-1)\delta}{10} \right) > k,$$

a contradiction.

Now, using ($\circledast$) (with $U = V(R)$), we can find two distinct vertices x, w_2 with $|N^+(x) \cap N^+(w_2)| \geq \delta^2 k/10 \geq 4/\delta$. Apply ($\circledast$) again, with $U = N^+(x) \cap N^+(w_2)$ to find two distinct out-neighbours of both x and w_2, denoted y, z, such that $|N^+(y) \cap N^+(z)| \geq \delta^2 k/10 \geq 4$. Now pick distinct $u, w_1 \in (N^+(y) \cap N^+(z)) \setminus \{x, w_2\}$. Then $\{x, y, z, u, w_1, w_2\}$ spans a copy of K, as required. $\qquad \square$

Let $X, Y, U, Z, W_1, W_2 \subseteq V(D)$ be the vertex clusters corresponding to the vertices of the copy of K in R given by Claim 5.12. Using Lemma 5.5, by removing up to $O(\varepsilon n)$ vertices from each cluster, we can ensure that $(X, Y), (X, Z), (W_1, Y), (W_1, Z), (Y, U), (Z, U), (Y, W_1), (Z, W_1)$ are $(\varepsilon, \rho/2)$-super-regular and that X, Y, Z, U, W_1, W_2 all have the same size, which we denote by n'. Thus, without loss of generality, we assume this is the case for the remainder of the proof.

Delete some triples from $\mathcal{C}_0$ if necessary so that $\mathcal{C}_0$ covers at most $n/2 - 1$ colours, and let $\mathcal{C}_1$ and $\mathcal{C}_2$ be disjoint sets of $n/4$ colours that do not appear in triples in $\mathcal{C}_0$ and are not c. For each edge of $E(Y \cup Z, W_1)$, and for these edges only, reveal if its colour belongs to $\mathcal{C}_1$, and delete if it does not. Similarly, for each edge of $E(W_2, Y \cup Z)$, and for these edges only, reveal if its colour belongs to $\mathcal{C}_2$, and delete it if not. This amounts to sampling each edge in $E(Y \cup Z, W_1) \cup E(W_2, Y \cup Z)$ with probability $1/4$. Hence, by Lemma 5.6, with probability at least $1 - \exp(-\Omega(\rho n))$, the remaining edges in $E(Y \cup Z, W_1) \cup E(W_2, Y \cup Z)$ span $(\varepsilon, \rho/16)$-super-regular pairs.

Now reveal the colours of the remaining edges in $E(Y \cup Z, W_1)$, which are chosen uniformly from $\mathcal{C}_1$. By Lemma 5.8, with $(n, \varepsilon, \alpha, \rho, q, A, B, C)_{5.8} = (n', \varepsilon, 1, \rho/16, n/(4n'), Y, W_1, Z)$ and H with $E(H) = Y \times Z$, with probability at least $1 - \exp(-\lambda' n)$, the following holds: there is a subgraph $H' \subseteq H$ with $e(H') \geq \alpha'(n')^2$, such that for every edge $yz \in E(H')$, there is a vertex $w_1 \in$

W_1 such that $\mathcal{C}(yw_1) = \mathcal{C}(zw_1)$. Next, reveal the colours of the remaining edges in $E(W_2, Y \cup Z)$; they are coloured uniformly in $\mathcal{C}_1$. By the same lemma, with $(n, \varepsilon, \alpha, \rho, q, A, B, C, H)_{5.8}$ $= (n', \varepsilon, \alpha', \rho/16, n/(4n'), Y, W_2, Z, H')$, with probability at least $1 - \exp(-\lambda' n)$, the following holds: there is a subgraph $H'' \subseteq H'$ with $e(H'') \geq \alpha''(n')^2$, such that for every edge $yz \in E(H'')$ there exists $w_2 \in W_2$ such that $\mathcal{C}(w_2 y) = \mathcal{C}(w_2 z)$.[4]

Finally, by Lemma 5.9, applied with $(\varepsilon, \alpha, \rho, \mu, H)_{5.9} = (\varepsilon, \alpha'', \rho/16, n'/n, H'')$, with probability at least $1 - \exp(-\lambda' n)$, there is a quadruple $(x, y, z, u) \in X \times Y \times Z \times U$ such that: xy, xz, yu, zu are edges in D; yz is an edge in H''; $(\mathcal{C}(zu), \mathcal{C}(xz), \mathcal{C}(xy)) \in \mathcal{C}_0$; and $\mathcal{C}(yu) = c$. By choice of H'', there are vertices $w_1 \in W_1$ and $w_2 \in W_2$ such that $yw_1, zw_1, w_2 u, w_2 z$ are edges and $\mathcal{C}(zw_1) = \mathcal{C}(yw_1) \in \mathcal{C}_1$ and $\mathcal{C}(w_2 z) = \mathcal{C}(w_2 y) \in \mathcal{C}_2$. Altogether, all the required events hold with probability at least $1 - \exp(-\Omega(\rho n)) - \exp(-\Omega(\lambda' n)) \geq 1 - \exp(-\lambda n)$. $\qquad\square$

5.2 Proof of Lemma 5.2

Proof of Lemma 5.2. Let λ, ρ, ρ' satisfy $\nu \ll \lambda, \rho, \rho' \ll \delta$. Fix $v \in V(D)$, $c \in \mathcal{C}$ and $V' \subseteq V(D)$, $\mathcal{C}' \subseteq \mathcal{C}$ of size at least $(1 - \nu)n$.

For the next claim, it is useful to refer to Fig. 1.

Claim 5.13. *With probability $1 - \exp(-\Omega(\lambda n))$ the following holds. In $D[V']$ there exist an oriented triangle T with $V(T) = \{v, v_1, v_2\}$ and $E(T) = \{vv_2, v_1 v, v_1 v_2\}$, and an oriented copy K of $K_{2,4}$, with $V(K) = \{x, y, z, u, w_1, w_2\}$ and $E(K) = \{xz, xy, yu, zu, yw_1, zw_1, w_2 y, w_2 z\}$. Moreover, the edges of K and T are coloured as follows. $\mathcal{C}(yu) = c$, $\mathcal{C}(xz) = \mathcal{C}(v_1 v_2)$, $\mathcal{C}(xy) = \mathcal{C}(v_1 v)$, $\mathcal{C}(zu) = \mathcal{C}(vv_2)$, $\mathcal{C}(zw_1) = \mathcal{C}(yw_1)$ and $\mathcal{C}(w_2 z) = \mathcal{C}(w_2 y)$.*

Proof. Let $(V_\triangle, V_\square)$ be a random partition of V'. Then, from Chernoff's bound, a union bound over the vertices of V', and $\nu \ll 1$, with probability $1 - \exp(-\Omega(\delta n))$, the digraphs $D_{\delta, n}[V_\triangle]$, $D_{\delta, n}[V_\square]$ have minimum semidegree at least $\delta n/3$, and $|V_\triangle|, |V_\square| \geq n/3$.

First reveal the random edges and colours of $D[V_\triangle]$.

Then, by Lemma 4.3, with probability $1 - \exp(-\Omega(\lambda n))$, there is a collection Δ_v of ρn rainbow triangles $\{v, v_1, v_2\}$, oriented as in the claim, that pairwise intersect only on v, are pairwise colour-disjoint, and with $V(\Delta_v) \subseteq V_\triangle \cup \{v\}$. For $T \in \Delta_v$, denote its vertices by v, $v_1(T), v_2(T)$ such that $E(T) = \{vv_2(T), v_1(T)v, v_1(T)v_2(T)\}$. Let $\mathcal{C}_v$ be the collection of triples $(\mathcal{C}(vv_2(T)), \mathcal{C}(v_1(T)v_2(T)), \mathcal{C}(v_1(T)v))$ with $T \in \Delta_v$, whose three colours are in $\mathcal{C}'$. Then $|\mathcal{C}_v| \geq \rho n - 3\nu n \geq (\rho/2)n$.

Next, reveal the colours of edges in $D[V_\square]$. By setting $\mathcal{C}_0 = \mathcal{C}_v$ in Lemma 5.11, it follows that, with probability $1 - \exp(-\Omega(\lambda n))$, $D[V_\square]$ contains a subgraph K isomorphic to $K_{2,4}$ with the orientation as in the claim, whose colours match those of a triangle $T \in \Delta_v$ as in the statement of the claim.

We fail to find a T or K as required with probability at most $\exp(-\Omega(\lambda n))$. $\qquad\square$

By Lemma 4.3, with probability $1 - \exp(-\Omega(\lambda n))$, for any two vertices $a, b \in V'$ there are $\rho' n$ rainbow directed paths of length three from a to b which are pairwise colour disjoint and internally vertex disjoint. Hence, with probability $1 - \exp(-\Omega(\lambda n))$, this and the conclusion of Claim 5.13 hold simultaneously.

Then, using $\nu \ll \rho'$, there exist two colour- and vertex-disjoint rainbow directed paths P_1, P_2 of length 3 such that: P_1 is directed from v_2 to x; P_2 is directed from w_1 to w_2; the interiors of P_1 and P_2 are in $V' \setminus (V(K) \cup V(T))$; and the colours of P_1 and P_2 are in $\mathcal{C}' \setminus (\mathcal{C}(K) \cup \mathcal{C}(T))$. Then the graph $A_{v,c}$, defined as

$$A_{v,c} = K \cup T \cup P_1 \cup P_2,$$

---

[4]Technically, here we apply the lemma to the digraph obtained by reversing the direction of edges in $E(W_2, Y \cup Z)$.

is a (v, c)-absorber: the (v, c)-absorbing path and the (v, c)-avoiding path are

$$v_1 v v_2 P_1 xzw_1 P_2 w_2 yu \quad \text{and} \quad v_1 v_2 P_1 xyw_1 P_2 w_2 zu$$

and it is straightforward to check they satisfy Definition 5.1 and have 13 vertices.

The number of $V' \subseteq V$ of size at least $(1 - v)n$ is at most $n\binom{n}{vn} = \exp(O(v \log v) n)$, and the same bound holds for the number of $C' \subseteq C$ of the same size. Using $v \ll \lambda$, the probability we fail to find an absorber for some v, c, V', C', is by the union bound, at most

$$n^2 e^{O(v \log v)n} \cdot e^{-\Omega(\lambda n)} = o(1). \qquad \square$$

6. Proof of Lemma 3.1

In this section we prove Lemma 3.1. The following lemma enables us to cover any small set of vertices and colours.

Lemma 6.1 (Flexible sets). *Let $1/n \ll 1/C \ll \zeta \ll \mu \ll \delta \leq 1$. Let D_0 be a digraph on n vertices with minimum semidegree at least δn, and suppose that $D \sim D_0 \cup \mathbf{D}(n, C/n)$ is uniformly coloured in $C = [n]$. Then there exist $V_{flex} \subseteq V$, $C_{flex} \subseteq C$ of size $2\mu n$ such that with high probability the following holds. For all distinct $u, v \in V$, $c \in C$, and $V'_{flex} \subseteq V_{flex}$, $C'_{flex} \subseteq C_{flex}$ of size at least $(2\mu - \zeta)n$, there exists a rainbow directed path of length seven from u to v, with internal vertices in V'_{flex} and colours in $C'_{flex} \cup \{c\}$, that contains the colour c.*

Proof. Let γ satisfy $\zeta \ll \gamma \ll \mu$.

For a colour c, let M_c be a largest oriented matching of colour c in D, and for distinct vertices u, v, let $\mathcal{P}_{u,v}$ be a largest collection of pairwise vertex- and colour-disjoint rainbow directed paths of length three with endpoints u, v. By Lemmas 4.2, 4.3, with probability at least $1 - \exp(-\gamma n)$, we have $|M_c| \geq \gamma n$ and $|\mathcal{P}_{u,v}| \geq \gamma n$ for every colour c and distinct vertices u, v.

Let V' be a random subset of V, obtained by including each vertex independently with probability μ, and let C' be a random subset of C, obtained by including each colour independently with probability μ.

Then, by Chernoff and union bounds, with high probability, the following properties hold.

- $|V'|, |C'| \leq 2\mu n$,
- at least $\mu^2 \gamma n/2$ edges in M_c have both endpoints in V', for every $c \in C$,
- at least $\mu^5 \gamma n/2$ paths in $\mathcal{P}_{u,v}$ have their interior vertices in V' and all colours in C', for all distinct $u, v \in V$.

Suppose that all three properties hold, and let V_{flex} be a subset of V that contains V' and has size $2\mu n$ and let C_{flex} be a subset of C that contains C' and has size $2\mu n$.

We show that these sets satisfy the requirements of the lemma. Indeed, fix u, v, c and $V'_{\text{flex}}, C'_{\text{flex}}$ as in the lemma. Then, as $\zeta \ll \mu, \gamma$, there is an edge $e = xy \in M_c$ with both ends in V'_{flex}. Similarly, there are paths $P_1 \in \mathcal{P}_{u,x}$ and $P_2 \in \mathcal{P}_{y,v}$ that are vertex- and colour-disjoint, their interiors are in V'_{flex}, and their colours are in $C'_{\text{flex}} \setminus \{c\}$. Then $P_1 \cup e \cup P_2$ is a path that satisfies the requirements of the lemma. $\qquad \square$

We will put together several (v, c)-absorbers to construct the digraph H_{abs} in Lemma 3.1, by having a (v, c)-absorber for each edge of a bipartite graph which has the following property. This follows an idea introduced by Montgomery [23], which was adapted to the rainbow setting by Gould et al. [16].

Definition 6.2 (Definition 3.3, [16]). *Let H be a balanced bipartite graph with bipartition (A, B). We say H is robustly matchable with respect to A', B', for some $A' \subseteq A$ and $B' \subseteq B$ of equal size, if*

for every pair of sets $X \subseteq A'$, $Y \subseteq B'$ with $|X| = |Y| \leq |A'|/2$, there is a perfect matching in $H[A \setminus X, B \setminus Y]$. We call A', B' the flexible sets of H.

Proposition 6.3 (Lemma 4.5, [16]). *For every large enough $m \in \mathbb{N}$, there exists a 256-regular bipartite graph with bipartition (A, B) and $|A| = |B| = 7m$, which is robustly matchable with respect to some $A' \subseteq A$, $B' \subseteq B$ with $|A'| = |B'| = 2m$.*

Proof of Lemma 3.1. Let ζ, μ, ν satisfy $\eta \ll \zeta \ll \mu \ll \nu \ll \gamma$.

By the union bound, the conclusions of Lemmas 4.3, 6.1, 5.2 hold simultaneously with high probability. Let V_{flex}, $\mathcal{C}_{\text{flex}}$ be sets as given by Lemma 6.1; so $|V_{\text{flex}}| = |\mathcal{C}_{\text{flex}}| = 2\mu n$.

Let V_{buf}, $\mathcal{C}_{\text{buf}}$ be arbitrary subsets of $V \setminus V_{\text{flex}}$, $\mathcal{C} \setminus \mathcal{C}_{\text{flex}}$ of size $5\mu n$. Let H be a bipartite graph on $(V_{\text{flex}} \cup V_{\text{buf}}, \mathcal{C}_{\text{flex}} \cup \mathcal{C}_{\text{buf}})$ that is isomorphic to the graph in Proposition 6.3 such that V_{flex}, $\mathcal{C}_{\text{flex}}$ are the flexible sets.

Claim 6.4. *There is collection of absorbers $A_{v,c}$ on 13 vertices each and rainbow directed paths $P_{v,c}$ of length three each, for every edge vc in H, with the following properties: the internal vertices of $A_{v,c}$ and of $P_{v,c}$ are pairwise disjoint and disjoint of $V_{\text{flex}} \cup V_{\text{buf}}$; the internal colours of $A_{v,c}$ and the colours of $P_{v,c}$ are pairwise disjoint and disjoint of $\mathcal{C}_{\text{flex}} \cup \mathcal{C}_{\text{buf}}$; and for some ordering of the edges of H, the path $P_{v,c}$ starts with the last vertex of $A_{v',c'}$ and ends with the first vertex of $A_{v,c}$, where $v'c'$ is the predecessor of vc in the ordering (with $P_{vc} = \emptyset$ the trivial path for the first edge vc).*

Proof. Let H_0 be a maximal subgraph of H with some ordering of its edges, for which we can find a collection of absorbers and paths as in the claim. Suppose for contradiction $H_0 \neq H$ and let $v_1 c_1 \in E(H \setminus H_0)$ and $v_0 c_0$ be the last edge of H_0 in the ordering that has absorber A_{v_0,c_0}.

Let $V_0, \mathcal{C}_0$ be the union of the vertices and colours spanned by the absorbers for $E(H_0)$, the paths connecting them, and $V_{\text{flex}} \cup V_{\text{buf}}$, $\mathcal{C}_{\text{flex}} \cup \mathcal{C}_{\text{buf}}$. Then, since each absorber has 13 vertices and each path connecting consecutive absorbers has four vertices, we have

$$|V_0|, |\mathcal{C}_0| = O(e(H_0)) = O(\mu n) < \nu n / 2.$$

Hence, by the assumption that Lemma 5.2 holds, there exists a (v_1, c_1)-absorber A_{v_1,c_1} on 13 vertices with internal vertices and internal colours disjoint from V_0 and $\mathcal{C}_0$. Moreover, from Lemma 4.3 there exists a rainbow directed path P_{v_1,c_1} of length three between the last vertex of A_{v_0,c_0} and the first vertex of A_{v_1,c_1}, with internal vertices disjoint of $V_0 \cup V(A_{v_1,c_1})$ and colours disjoint of $\mathcal{C}_0 \cup \mathcal{C}(A_{v_1,c_1})$. Then the subgraph of H with edges $E(H_0) \cup \{v_1 c_1\}$ satisfies the conditions of the claim and properly contains H_0, contradicting the maximality of H_0. $\square$

Set

$$H_{\text{abs}} := \bigcup_{vc \in E(H)} (A_{v,c} \cup P_{v,c}),$$

where $A_{v,c}$ and $P_{v,c}$ are as in Claim 6.4, with $P_{v,c} = \emptyset$ for the first vc,

Let w be the first vertex in the first absorber in H_{abs} with the ordering given by Claim 6.4, and let w' be the last vertex of the last absorber. We will now show how to construct a directed path Q from x to y as in the statement of the lemma, given $V' \subseteq V \setminus V(H_{\text{abs}})$ and $\mathcal{C}' \subseteq \mathcal{C} \setminus \mathcal{C}(H_{\text{abs}})$ of size between 2 and ηn. Let $c_0 \in \mathcal{C}'$. From Lemma 6.1, there exists a rainbow directed path Q_1 from x to w, with internal vertices in V_{flex} and colours in $\mathcal{C}_{\text{flex}} \cup \{c_0\}$, which includes the colour c_0 and has length 7.

Claim 6.5. *There exists a rainbow directed path Q_2 from w' to y, with internal vertices $V''_{\mathrm{flex}} \cup (V' \setminus \{x\})$, and colours $\mathcal{C}''_{\mathrm{flex}} \cup (\mathcal{C}' \setminus \{c_0\})$, for some $V''_{\mathrm{flex}} \subseteq V_{\mathrm{flex}} \setminus V(Q_1)$, $\mathcal{C}''_{\mathrm{flex}} \subseteq \mathcal{C}_{\mathrm{flex}} \setminus \mathcal{C}(Q_1)$ with $\left|V''_{\mathrm{flex}}\right| = \left|\mathcal{C}''_{\mathrm{flex}}\right| \leq \mu n - 7$.*

Proof. The Claim will follow by applying greedily Lemma 6.1.

Let $v_0, \ldots, v_k$ be an enumeration of the vertices in $(V' \cup \{w'\}) \setminus \{x\}$, with $v_0 = w'$ and $v_k = y$, and let $c_1, \ldots, c_k$ be an enumeration of $\mathcal{C}' \setminus \{c_0\}$. We claim that there exist directed paths $P_1, \ldots, P_k$ as follows: P_i is a rainbow directed path from v_{i-1} to v_i of length 7; the interiors of the paths P_i are pairwise vertex-disjoint, contained in V_{flex}, and disjoint of $V' \cup \{w'\}$; P_i contains an edge coloured c_i, and all other colours are in $\mathcal{C}_{\mathrm{flex}}$; and no colour appears on more than one path P_i.

To see this, suppose that $P_1, \ldots, P_{i-1}$ are defined. Let $V'_{\mathrm{flex}} = V_{\mathrm{flex}} \setminus (V(P_1) \cup \ldots \cup V(P_{i-1}) \cup V(Q_1))$, $\mathcal{C}'_{\mathrm{flex}} = \mathcal{C}_{\mathrm{flex}} \setminus (\mathcal{C}(P_1) \cup \ldots \cup \mathcal{C}(P_{i-1}) \cup \mathcal{C}(Q_1))$. Then $|V'_{\mathrm{flex}}| \geq 2\mu n - 7k \geq (2\mu - \zeta)n$, using that $k \leq \eta n$, $|V(Q_1)|, |V(P_j)| = 8$, and $\eta \ll \zeta$, and similarly $|\mathcal{C}'_{\mathrm{flex}}| \geq (2\mu - \zeta)n$. Hence, by Lemma 6.1 there is a rainbow directed path P_i from v_{i-1} to v_i of length 7, that contains an edge with colour c_i, and whose internal vertices and other colours are in V'_{flex} and $\mathcal{C}'_{\mathrm{flex}}$, as required.

Write $Q_2 = P_1 \ldots P_k$. Let $V''_{\mathrm{flex}} = V(Q_2) \cap V_{\mathrm{flex}}$, $\mathcal{C}''_{\mathrm{flex}} = \mathcal{C}(Q_2) \cap \mathcal{C}_{\mathrm{flex}}$. Then $\left|V''_{\mathrm{flex}}\right| = \left|\mathcal{C}''_{\mathrm{flex}}\right| < |Q_2| \leq \zeta n < \mu n - 7$, as required. □

Let $V'''_{\mathrm{flex}} = (V(Q_1) \cup V(Q_2)) \cap V_{\mathrm{flex}}$ and $\mathcal{C}'''_{\mathrm{flex}} = (\mathcal{C}(Q_1) \cup \mathcal{C}(Q_2)) \cap \mathcal{C}_{\mathrm{flex}}$. Then we have $\left|V'''_{\mathrm{flex}}\right| = \left|\mathcal{C}'''_{\mathrm{flex}}\right| \leq \mu n$. Hence by choice of H there is a matching M' between $V_{\mathrm{flex}} \setminus V'''_{\mathrm{flex}}$ and $\mathcal{C}_{\mathrm{flex}} \setminus \mathcal{C}'''_{\mathrm{flex}}$.

For $vc \in E(H)$ let $Q_{M'}(vc)$ be the (v, c)-absorbing path of $A_{v,c}$ if $vc \in E(M')$ and the avoiding path otherwise. Let

$$Q_{\mathrm{abs}} = \bigcup_{vc \in E(H)} (Q_{M'}(vc) \cup P_{v,c}).$$

Then Q_{abs} is a rainbow directed path that is spanning in $V(H_{\mathrm{abs}}) \setminus V'''_{\mathrm{flex}}$ and $\mathcal{C}(H_{\mathrm{abs}}) \setminus \mathcal{C}'''_{\mathrm{flex}}$ with first vertex w and last vertex w', with every colour used once. Therefore $Q = Q_1 \cup Q_{\mathrm{abs}} \cup Q_2$ is a rainbow path, spanning in $V(H_{\mathrm{abs}}) \cup V'$ and $\mathcal{C}(H_{\mathrm{abs}}) \cup \mathcal{C}'$ with first vertex x and last vertex y. □

7. Conclusion

In this paper we considered the problem of containing a rainbow directed Hamilton cycle in uniformly coloured perturbed digraphs, when the number of colours is n (optimal) and the probability is $\Theta(n^{-1})$ (optimal up to a constant factor). A natural open problem to consider next is whether, under the same assumptions, it is possible to guarantee with high probability a rainbow copy of every fixed orientation of a Hamilton cycle, or even to prove a *universality* result, that is showing that we can find a rainbow copy of every orientation simultaneously. In the uncoloured setting such a universality result was recently obtained by Araujo et al. [5], for cycles of arbitrary lengths (not just Hamilton cycles). They proved that for every $\delta \in (0, 1)$ there exists $C > 0$, such that for every fixed orientation D of a cycle of length between 3 and n, with probability at least $1 - e^{-n}$, the perturbed digraph $D_0 \cup \mathbf{D}(n, C/n)$ (where as usual δn is a lower bound on the semidegree of D_0) contains D. Taking the union bound over the at most $n2^n$ choices for D gives the universality result. This greatly generalises the original result of Bohman et al. [8] where, under the same assumptions, they showed that with high probability the perturbed digraph has a directed Hamilton cycle. It is plausible that our methods in combination with those of [5] could extend the result of [5] to the rainbow setting, i.e. that we can find with probability $1 - e^{-n}$ a rainbow copy of a fixed oriented cycle of length $\ell \in [2, n]$ when colouring uniformly with n colours.

References

[1] Aigner-Horev, E. and Hefetz, D. (2021) Rainbow Hamilton cycles in randomly colored randomly perturbed dense graphs. *SIAM J. Discrete Math.* **35**(3) 1569–1577.

[2] Aigner-Horev, E., Hefetz, D. and Lahiri, A. (2023) Rainbow trees in uniformly edge-coloured graphs. *Random Struct. Algor.* **62**(2) 287–303.

[3] Alon, N. and Shapira, A. (2003) Testing subgraphs in directed graphs. In Proceedings of the 35th Annual ACM Symposium on Theory of Computing, pp. 700–709.

[4] Anastos, M. and Frieze, A. (2019) How many randomly colored edges make a randomly colored dense graph rainbow Hamiltonian or rainbow connected? *J. Graph Theory* **92**(4) 405–414.

[5] Araujo, I., Balogh, J., Krueger, R. A., Piga, S. and Treglown, A. (2024) On oriented cycles in randomly perturbed digraphs. *Combin. Probab. Comput.* **33**(2) 157–178.

[6] Bal, D. and Frieze, A. (2016) Rainbow matchings and Hamilton cycles in random graphs. *Random Struct. Algor.* **48**(3) 503–523.

[7] Ben-Eliezer, I., Krivelevich, M. and Sudakov, B. (2012) The size Ramsey number of a directed path. *J. Comb. Theory Ser. B* **102**(3) 743–755.

[8] Bohman, T., Frieze, A. and Martin, R. (2003) How many random edges make a dense graph Hamiltonian? *Random Struct. Algor.* **22**(1) 33–42.

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

[10] Dirac, G. A. (1952) Some theorems on abstract graphs. *Proc. London Math. Soc.* **s3-2**(1) 69–81.

[11] Ferber, A. (2015) Closing gaps in problems related to Hamilton cycles in random graphs and hypergraphs. *Electron. J. Comb.* **22**(1). Paper 1.61, 7 pp.

[12] Ferber, A. and Krivelevich, M. (2016) Rainbow Hamilton cycles in random graphs and hypergraphs. In *Recent Trends in Combinatorics*, Vol. **159**, Springer, pp. 167–189.

[13] Frieze, A. and Loh, P. (2014) Rainbow Hamilton cycles in random graphs. *Random Struct. Algor.* **44**(3) 328–354.

[14] Frieze, A. M. (1988) An algorithm for finding Hamilton cycles in random directed graphs. *J. Algor.* **9**(2) 181–204.

[15] Ghouila-Houri, A. (1960) Une condition suffisante d'existence d'un circuit Hamiltonien. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* **251** 495–497(in French).

[16] Gould, S., Kelly, T., Kühn, D. and Osthus, D. (2022) Almost all optimally coloured complete graphs contain a rainbow Hamilton path. *J. Comb. Theory Ser. B* **156** 57–100.

[17] Janson, S., Ruciński, A. and Łuczak, T. (2011) *Random Graphs*. John Wiley & Sons.

[18] Kelly, L., Kühn, D. and Osthus, D. (2008) A Dirac-type result on Hamilton cycles in oriented graphs. *Comb. Prob. Comput.* **17**(5) 689–709.

[19] Komlós, J. and Szemerédi, E. (1983) Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete Math.* **43**(1) 55–63.

[20] Korshunov, A. D. (1977) A solution of a problem of P. Erdős and A. Rényi about Hamilton cycles in non-oriented graphs. *Metody Diskr. Anal. Teoriy Upr. Syst.,Sb. Trudov Novosibirsk* **31** 17–56 (in Russian).

[21] McDiarmid, C. (1980) Clutter percolation and random graphs. In *Mathematical Programming Studies*, pp. 17–25.

[22] Montgomery, R. (to appear) Spanning cycles in random directed graphs. *Random Struct. Algorithms*.

[23] Montgomery, R. (2019) Spanning trees in random graphs. *Adv. Math.* **356** 106793.

[24] Pósa, L. (1976) Hamiltonian circuits in random graphs. *Discrete Math.* **14**(4) 359–364.

[25] Szemerédi, E. (1976) Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay)*, pp. 399–401.