

Minimum degree conditions for monochromatic cycle partitioning

Dániel Korándi* Richard Lang† Shoham Letzter‡ Alexey Pokrovskiy§

Abstract

A classical result of Erdős, Gyárfás and Pyber states that any r -edge-coloured complete graph has a partition into $O(r^2 \log r)$ monochromatic cycles. Here we determine the minimum degree threshold for this property. More precisely, we show that there exists a constant c such that any r -edge-coloured graph on n vertices with minimum degree at least $n/2 + c \cdot r \log n$ has a partition into $O(r^2)$ monochromatic cycles. We also provide constructions showing that the minimum degree condition and the number of cycles are essentially tight.

1 Introduction

Monochromatic cycle partitioning is a combination of Ramsey-type and covering problems. Given an edge-coloured host graph G , one seeks to partition the vertex set of G into as few monochromatic cycles as possible.¹ The case where the number of used cycles f can be upper-bounded by a function of the number of colours r is of particular interest. A classical result in this area is due to Erdős, Gyárfás and Pyber [13], who showed that any r -edge-coloured complete graph $G = K_n$ admits a partition into $\lceil 25r^2 \log r \rceil$ monochromatic cycles. The same authors conjectured that their result could in fact be improved to r cycles. For $r = 2$, this had been suggested about 20 years earlier by Lehel in a stronger sense, i.e. with the cycles having distinct colours. Lehel's conjecture was first proved for large n by Łuczak, Rödl and Szemerédi [33] and then for all n by Bessy and Thomassé [3], after preliminary work by Gyárfás [16]. For $r \geq 3$, the conjecture of Erdős, Gyárfás and Pyber turned out to be false. Pokrovskiy [34] provided colourings of the complete graph that require r cycles and a single additional vertex for a partition. He conjectured, however, that a partition into r cycles and a constant number of vertices $c(r)$ should nevertheless be sufficient. In support of his conjecture, Pokrovskiy showed the case of $r = 3$ with $c(r) = 43000$. This was independently confirmed by Letzter [29] with $c(r) = 60$.

*Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, United Kingdom. Research supported in part by SNSF grants 200020-162884 and 200021-175977, and an SNSF Postdoc.Mobility fellowship. Email: daniel.korandi@maths.ox.ac.uk.

†Institute for Computer Science, University of Heidelberg, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany. The research leading to these results was partially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 42821240 (R. Lang). Email: lang@informatik.uni-heidelberg.de.

‡Department of Mathematics, University College London, Gower Street, London, WC1E 6BT, UK. Email: s.letzter@ucl.ac.uk. The research was supported by Dr. Max Rössler, the Walter Haefner Foundation, the ETH Zurich Foundation and the Royal Society.

§Department of Mathematics, University College London, Gower Street, London, WC1E 6BT, UK. Email: Dr.Alexey.Pokrovskiy@gmail.com.

¹To avoid some trivial cases, we consider the empty set, vertices and edges to be cycles.

The best known general upper bound for the number of monochromatic cycles required to partition any r -coloured complete graph is $100r \log r$, established by Gyárfás, Ruszinkó, Sárközy and Szemerédi [18].

In the past decades monochromatic partitions of the complete graph have been researched in many ways, such as partitioning into graphs other than cycles [15, 38], more general colourings [9] and partitions of hypergraphs [7, 8, 14, 19]. For a broader overview, we refer the reader to the recent survey of Gyárfás [17]. Another natural problem arises when we consider host graphs that need not be complete. In particular, for which families of graphs can we still partition the vertex set into few monochromatic cycles? This question has been investigated for complete bipartite graphs [21], graphs with fixed independence number [37], infinite graphs [12, 36] and random graphs [26, 28] among others. Here we are interested in families of graphs characterized by a large minimum degree.

The study of minimum degree conditions for spanning substructures has a long tradition in extremal graph theory, Dirac's theorem being a classical example. Recent milestones of this area include the resolution of the Pósa-Seymour conjecture by Komlós, Sárközy and Szemerédi [24], the Bandwidth theorem by Böttcher, Schacht and Taraz [5], and the Hamilton decomposition theorem by Csaba, Kühn, Lo, Osthus and Treglown [10]. Many other results in this line of research are covered in the survey of Kühn and Osthus [27].

For monochromatic cycle partitions, the research of minimum degree conditions was initiated by Balogh, Barát, Gerbner, Gyárfás and Sárközy [2] with a strengthening of Lehel's conjecture. They showed that every 2-edge-coloured graph G on n vertices of minimum degree $(3/4 + \varepsilon)n$ admits a partition of all but $o(n)$ vertices into two monochromatic cycles of distinct colours. They also conjectured that this can be improved to a proper partition even without the term of εn . (An easy construction shows that this is best possible.) The extension to a proper partition was verified by DeBiasio and Nelsen [11] and the full conjecture was subsequently proved by Letzter [30]. Given these advances, Pokrovskiy [35] conjectured that for a 2-edge-coloured graph G with $\delta(G) \geq 2n/3$ and $\delta(G) \geq n/2$ a partition into 3 and 4, respectively, cycles is possible. (Again, constructions show that this is essentially best possible.) Allen, Böttcher, Lang, Skokan and Stein [1] confirmed the first part of this conjecture approximately, i.e. for $\delta(G) \geq (2/3 + \varepsilon)n$. Thus the problem for two colours is increasingly well understood.

The goal of this research was to determine the minimum degree threshold for partitioning an r -edge-coloured graph into $f(r)$ monochromatic cycles for general r , for any function $f(r)$ that depends only on r . A lower bound of $n/2$ for this threshold is shown by the simple example of a slightly unbalanced complete bipartite graph. However, a more involved construction shows that a minimum degree below $n/2 + O(\log n / \log \log n)$ already requires $\Omega(\log n / \log \log n)$ monochromatic cycles for a partition.

Proposition 1.1. *Let n be sufficiently large. Then there is a 2-edge-coloured graph G on n vertices with $\delta(G) \geq n/2 + \log n / (16 \log \log n)$ whose vertices cannot be partitioned into fewer than $\log n / (32 \log \log n)$ monochromatic cycles.*

Our main contribution states that a minimum degree slightly larger than this is in turn sufficient for a partition into $O(r^2)$ cycles.

Theorem 1.2. *For $r \geq 2$, let n be sufficiently large. Then any r -edge-coloured graph G on n vertices with $\delta(G) \geq n/2 + 1200r \log n$ admits a partition into $10^7 r^2$ monochromatic cycles.*

We also provide a construction that shows that the number of cycles of Theorem 1.2 is essentially best possible.

Proposition 1.3. *Let $\varepsilon > 0$ and r be sufficiently large. Then there is an r -edge-coloured graph G on $n = n(\varepsilon, r)$ vertices with $\delta(G) \geq (1 - 4\varepsilon)n$, whose vertices cannot be covered by fewer than $\varepsilon^2(r - 1)^2/4$ monochromatic trees.*

It is worth mentioning here a recent paper of Bucić, Korándi and Sudakov [6], who were interested in covering r -coloured random graphs $G(n, p)$ by monochromatic trees. Similarly to our results, they proved that the minimum number of monochromatic trees needed is $\Theta(r^2)$ when p is just above the threshold for the existence of a covering with a bounded number of trees.

Our results imply in particular, that we can determine the smallest number of cycles necessary for a partition of bounded degree graphs up to a constant factor. This stands in contrast to the situation for complete graphs, where the gap between upper and lower bound remains a factor of $\log r$.

2 Overview

A brief overview of the proof of Theorem 1.2 is as follows. Using Szemerédi’s regularity lemma, we obtain a regular partition $\{V_0, V_1, \dots, V_m\}$ of the vertices of G , and define the corresponding reduced graph \mathcal{G} . We then select $O(r^2)$ monochromatic components of \mathcal{G} in such a way that their union, denoted by \mathcal{H} , robustly contains a perfect matching. The robustness roughly translates into \mathcal{H} having a perfect matching even after removing any small set of vertices. So, in particular, \mathcal{H} has a perfect matching \mathcal{M} . We can turn \mathcal{M} into $O(r^2)$ disjoint monochromatic cycles $\mathcal{C}^{\mathcal{H}}$ covering almost all of G using a method of Łuczak [32].

The plan is now to add the remaining vertices $V(G) \setminus \mathcal{C}^{\mathcal{H}}$ into the cycles of $\mathcal{C}^{\mathcal{H}}$. More precisely, we intend to use the blow-up lemma to find monochromatic spanning paths in the regular pairs corresponding to \mathcal{M} . There are two obstacles to this. First, there might be a small number of “bad” vertices blocking the use of the blow-up lemma. Second, the clusters $V_i \setminus \mathcal{C}^b$ might be slightly different in size, which prevents us from even allocating spanning paths in the pairs.

We deal with the irregular vertices by covering them with $O(r^2)$ additional cycles \mathcal{C}^b , exploiting their large degrees. We then balance the clusters by carefully extending the cycles of $\mathcal{C}^{\mathcal{H}}$. At this point, the robustness under which \mathcal{H} has a perfect matching is crucial. Having overcome these two issues, we can finish by applying the blow-up lemma to add the remaining vertices into $\mathcal{C}^{\mathcal{H}}$. Thus $\mathcal{C}^{\mathcal{H}} \cup \mathcal{C}^b$ presents the desired cycle partition.

This method works as long as \mathcal{G} admits a spanning subgraph \mathcal{H} , which robustly contains a perfect matching, but unfortunately, we cannot always guarantee this. However, if such subgraph \mathcal{H} does not exist, then we can show that G must be balanced bipartite after the removal of $O(r)$ monochromatic cycles \mathcal{C} . (At this point, we use the additional $1200r \log n$ in the minimum

degree.) Thus we can apply a bipartite analogue of the above detailed approach to cover the rest with cycles $\mathcal{C}^{\mathcal{H}} \cup \mathcal{C}^b$. In this case the cycle partition consists of $\mathcal{C}^{\mathcal{H}} \cup \mathcal{C}^b \cup \mathcal{C}$.

3 Preliminaries

In this section we introduce some notations and tools needed for the proof of Theorem 1.2.

3.1 Notation

Let $G = (V, E)$ be a graph. The *order* of G is $|V(G)|$ and the *size* of G is $|E(G)|$. We denote the neighbourhood of a vertex v by $N_G(v)$ and write $N_G(v, W) = N_G(v) \cap W$ for a set of vertices $W \subseteq V(G)$. We denote the degree of v by $\deg_G(v) = |N_G(v)|$ and $\deg_G(v, W) = |N_G(v, W)|$. For a set of vertices $S \subseteq G$ we write $N_G(S) = (\bigcup_{s \in S} N(s)) \setminus S$. When the underlying graph G is clear from the context, we often omit the subscript G . For another graph H , the union $G \cup H$ is the graph on vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H)$. The independence number of G is denoted by $\alpha(G)$. For disjoint sets $X, Y \subseteq V(G)$, we denote by $G[X, Y]$ the induced bipartite subgraph of G with bipartition $\{X, Y\}$.

An r -edge-colouring of G assigns one *colour* from the set $[r] = \{1, 2, \dots, r\}$ to each edge of G . For $i \in [r]$, we use G_i to denote the subgraph of G whose edges are those that have colour i . A (connectivity) component of colour i of G is a (connectivity) component of G_i . When the context is clear, we will simplify our notation by using i instead of G_i in the subscript of N and \deg . For example, a vertex v with $\deg_i(v) = 0$ is in an i -coloured component of order 1.

A v - w -path is a path that starts at v and ends at w . As said above, we allow the empty set, single vertices and edges in our cycle partitions. We occasionally use the term ‘‘proper cycle’’ to emphasize that a cycle is not an empty set, a vertex or an edge.

In some of our statements, we will make assumptions of the form $x \ll y$ for certain parameters x and y . This should be understood as equivalent to the condition $x \leq f(y)$ for some unspecified increasing function f . In our usage, this is always a strengthening of $x \leq y$.

3.2 Regularity

Given a graph G and disjoint vertex sets $V, W \subseteq V(G)$ we denote the number of edges between V and W by $e(V, W)$ and the density of (V, W) by $d(V, W) = e(V, W)/(|V||W|)$. The pair (V, W) is called ε -regular, if all subsets $X \subseteq V$, $Y \subseteq W$ with $|X| \geq \varepsilon|V|$ and $|Y| \geq \varepsilon|W|$ satisfy $|d(V, W) - d(X, Y)| \leq \varepsilon$.

We say that a vertex $v \in V$ has *typical degree* in (V, W) , if $\deg(v, W) \geq (d(V, W) - \varepsilon)|W|$. It follows directly from the definition of ε -regularity that

$$\text{all but at most } \varepsilon|V| \text{ vertices in } V \text{ have typical degree in } (V, W). \quad (3.1)$$

The next lemma allows us to find (spanning) paths in regular pairs. A similar statement can be deduced from the well-known blow-up lemma [23], but it can also be proved independently, see Appendix A.

Lemma 3.1 (Long paths in regular pairs). *Let n be an integer and let ε, d be numbers with $0 < 1/n \ll \varepsilon \ll d < 1$. Suppose that (V_1, V_2) is an ε -regular pair of density $d = d(V_1, V_2)$ and with $|V_1| = |V_2| = n$ in a graph G . For $i \in \{1, 2\}$, let $v_i \in V_i$ and let $U_i \subseteq V_i$ be a set of size at least $n/6$ which contains at least $2\varepsilon n$ neighbours of v_{3-i} .*

Then for every $2 \leq k \leq (1 - 24\varepsilon) \cdot \min\{|U_1|, |U_2|\}$, there is a v_1 - v_2 -path of order $2k$ in $G[U_1 \cup \{v_1\}, U_2 \cup \{v_2\}]$.

If, additionally, $\delta(G[U_1, U_2]) \geq 5\varepsilon n$, then $G[U_1 \cup \{v_1\}, U_2 \cup \{v_2\}]$ contains a v_1 - v_2 -path of order $2k$ for every $2 \leq k \leq \min\{|U_1 \cup \{v_1\}|, |U_2 \cup \{v_2\}|\}$.

Szemerédi's Regularity Lemma [40] allows one to partition the vertex set of a graph into clusters of vertices, in a way that most pairs of clusters are regular. We will use the regularity lemma in its degree form (see [25]), with r colours and a prepartition.

Lemma 3.2 (Regularity Lemma). *For every $\varepsilon > 0$ and integers r, ℓ there is an $M = M(\varepsilon, r, \ell)$ such that the following holds. Let G be a graph on $n \geq 1/\varepsilon$ vertices whose edges are coloured with r colours, let $\{W_1, \dots, W_{\ell'}\}$ be an equipartition of $V(G)$ for some $1 \leq \ell' \leq \ell$, and let $d > 0$. Then there is a partition $\{V_0, \dots, V_m\}$ of $V(G)$ and a subgraph G' of G with vertex set $V(G) \setminus V_0$ such that the following conditions hold.*

- (a) $1/\varepsilon \leq m \leq M$,
- (b) $|V_0| \leq \varepsilon n$ and $|V_1| = \dots = |V_m| \leq \varepsilon n$,
- (c) for every $i \in [m]$, there is $j \in [\ell']$ with $V_i \subseteq W_j$,
- (d) for every $j \in [\ell']$, there are equally many $i \in [m]$ with $V_i \subseteq W_j$,
- (e) $\deg_{G'}(v) \geq \deg_G(v) - (rd + \varepsilon)n$ for each $v \in V(G) \setminus V_0$,
- (f) $G'[V_i]$ contains no edges for $i \in [m]$, and
- (g) all pairs (V_i, V_j) are ε -regular in G' and have in each colour either density 0 or at least d .

Let G be an r -edge-coloured graph with a partition $\{V_0, \dots, V_m\}$ obtained from Lemma 3.2 with parameters ε and d . We define the (ε, d) -reduced graph \mathcal{G} to be a graph with vertex set $V(\mathcal{G}) = \{x_1, \dots, x_m\}$ where two vertices x_i and x_j are connected by an edge of colour c , if (V_i, V_j) is an ε -regular pair of density at least d in colour c (if this holds for multiple colours, we choose one of them arbitrarily). Note that if G was balanced ℓ -partite with partition $\{W_1, \dots, W_{\ell}\}$, then \mathcal{G} is a balanced ℓ -partite graph, as well. It is often convenient to refer to a cluster V_i via its corresponding vertex in the reduced graph, i.e. $V_i = V(x_i)$.

The following properties of the reduced graph are easy to check. We provide a proof in Appendix A.

Proposition 3.3.

- (a) If $\deg_G(v) \geq cn$ for some $v \in V_i$, $i \in [m]$, then $\deg_{\mathcal{G}}(x_i) \geq (c - rd - \varepsilon)m$.
- (b) If $\deg_G(v) \geq cn$ for all but ηn vertices $v \in V(G)$, then $\deg_{\mathcal{G}}(x) \geq (c - rd - \varepsilon)m$ for all but $(\eta + \varepsilon)m$ vertices $x \in V(\mathcal{G})$.
- (c) If $\bigcup_{x_i \in X} V_i$ induces at least cn^2 edges in G for some $X \subseteq V(\mathcal{G})$, then X induces at least $(c - rd - \varepsilon)m^2$ edges.

The next lemma allows us to connect clusters by short paths if the corresponding vertices in the reduced graph lie in the same component. This is the basis of our application of the connected matchings method, a technique that goes back to Łuczak [32] and is by now a standard way of constructing long paths and cycles in dense graphs. See Appendix A for a proof.

Lemma 3.4 (Connecting Paths). *Let n be an integer and let ε, d be numbers with $0 < 1/n \ll \varepsilon \ll d \leq 1$. Let $G = (V, E)$ be an r -edge-coloured graph on n vertices with a partition $\{V_0, \dots, V_m\}$ and an (ε, d) -reduced graph \mathcal{G} obtained from Lemma 3.2. Suppose that $W \subseteq V$ is a vertex set such that $|W \cap V_i| \leq (d/2) \cdot |V_i|$ for every $i \in [m]$. Let $x_i x_j, x_{i'} x_{j'} \in E(\mathcal{G})$ be two edges in a component of colour c .*

Then for any two vertices $v \in V_i$ and $w \in V_{j'}$ of typical degree in colour c in (V_i, V_j) and $(V_{i'}, V_{j'})$, G contains a c -coloured v - w -path P of order at most $2m$ that avoids all vertices of W .

3.3 b -matchings

Adaptations of Łuczak's connected matchings method usually proceed by applying the regularity lemma, and finding large matchings in the components of the reduced graph (a matching whose edges belong to the same component is called a *connected matching*). The point is that a connected matching in the reduced graph can easily be converted into a cycle in the original graph.

In our case, it will be more convenient to work with 2-matchings, i.e. subgraphs, where each vertex can be incident to at most two edges. These convert to cycles the same way as matchings.

Definition 3.5 (Perfect b -matching). Let $b : V(G) \rightarrow \mathbb{Z}^{\geq 0}$ be a function on the vertices of a graph G . A *perfect b -matching* of G is a non-negative function $\omega : E(G) \rightarrow \mathbb{Z}^{\geq 0}$ on the edges, such that $\sum_{w \in N(v)} \omega(vw) = b(v)$ for every vertex v . When b is the constant 2 function, we call ω a *perfect 2-matching*.

It is easy to see that perfect 2-matchings correspond to vertex-disjoint cycles and edges that cover all the vertices. For example, a perfect matching with weight 2 on each edge is a perfect 2-matching. The following analogue of Tutte's theorem is a convenient characterization of graphs that admit a perfect 2-matching (see [39, Corollary 30.1a]).

Theorem 3.6 (Tutte). *A graph G has a perfect 2-matching if and only if every independent set $S \subseteq V(G)$ satisfies $|N(S)| \geq |S|$.*

However, we will need stronger conditions so that our graph is guaranteed to have a perfect 2-matching even after slight modifications.

3.4 Robustly matchable graphs

Definition 3.7 ((μ, ν) -robustly 2-matchable graphs). A graph H on n vertices is (μ, ν) -robustly 2-matchable if any of the following two conditions holds.

- (1) $\delta(H) \geq (1/2 - \mu)n$ and every set of $(1/2 - \nu)n$ vertices spans at least νn^2 edges.
- (2) H is a balanced bipartite graph with parts A, B (of size $n/2$) such that
 - $\delta(H) \geq (1/32 - \mu)n$, and
 - all but at most $(1/64 + \mu)n$ vertices in H have degree at least $(1/3 - \mu)n$.

We will distinguish robustly 2-matchable graphs of type 1 and type 2 accordingly.

Note that every (μ', ν') -robustly 2-matchable graph with $\mu' \leq \mu$ and $\nu' \geq \nu$ is automatically (μ, ν) -robustly 2-matchable, as well.

The following claim explains why we call these graphs “2-matchable”.

Lemma 3.8. *Every (μ, ν) -robustly 2-matchable graph H with $\mu \leq \nu < 1/1000$ contains a perfect 2-matching.²*

Proof. If H is a type 1 robustly 2-matchable graph, then for every non-empty independent set S , we have

$$|S| \leq (1/2 - \nu)n \leq (1/2 - \mu)n \leq \delta(H) \leq |N(S)|$$

(using the independence of S in the first and last step), so H satisfies the conditions of Theorem 3.6.

Now suppose H is of the second type with bipartition $V(H) = A \cup B$. By König’s theorem, it is enough to check that every independent set has size at most $n/2$. Indeed, this would guarantee the existence of a perfect matching, and hence a perfect 2-matching. So let S be an independent set in H , and let $S_A = S \cap A$ and $S_B = S \cap B$. We may assume that $|S_A| \leq |S_B|$, and note that $N(S_A) \subseteq B \setminus S_B$.

If $1 \leq |S_A| \leq (1/64 + \mu)n$, then $|N(S_A)| \geq \delta(H) \geq (1/32 - \mu)n \geq (1/64 + \mu)n \geq |S_A|$. (If S_A is empty, then trivially $|N(S_A)| \geq |S_A|$.) So $|S| = |S_A| + |S_B| \leq |N(S_A)| + |S_B| \leq |B| = n/2$. On the other hand, if $|S_A| > (1/64 + \mu)n$, then there is a vertex $v \in S_A$ of degree at least $(1/3 - \mu)n \geq n/4$, so $|N(S_A)| \geq n/4$. As $|S_B| \geq |S_A| > (1/64 + \mu)n$, we similarly get $|N(S_B)| \geq n/4$. But then $|S| = |S_A| + |S_B| \leq n - |N(S_A)| - |N(S_B)| \leq n/2$, as needed. \square

The next two statements illustrate the robustness of the above definition.

Lemma 3.9. *Suppose H is a (μ, ν) -robustly 2-matchable graph on n vertices and let $\varepsilon > 0$. Suppose H' is a spanning subgraph of H such that $\deg_{H'}(v) \geq \deg_H(v) - \varepsilon n$ for every vertex v . Then H' is $(\mu + \varepsilon, \nu - \varepsilon)$ -robustly 2-matchable whose type coincides with that of H .*

²Strictly speaking, it would be enough to bound μ from above by $1/1000$: monotonicity implies that H is also (μ, μ) -robustly 2-matchable. We impose the upper bound on ν purely to keep our parameter hierarchy simpler.

Proof. If H is a type 1 robustly 2-matchable graph, then $\delta(H') \geq \delta(H) - \varepsilon n \geq (1/2 - (\mu + \varepsilon))n$, as needed. Also, H' loses at most εn^2 edges compared to H , so every set of $(1/2 - \nu)n$ vertices spans at least $(\nu - \varepsilon)n^2$ edges. In particular, the same holds for every set of $(1/2 - (\nu - \varepsilon))n$ vertices.

On the other hand, if H is of the second type, then we similarly get $\delta(H') \geq (1/32 - (\mu + \varepsilon))n$, as well as $\deg_{H'}(v) \geq (1/3 - (\mu + \varepsilon))n$ for all but $(1/64 + \mu)n$ vertices v . \square

Lemma 3.10. *Suppose H is an r -edge-coloured (μ, ν) -robustly 2-matchable graph on n vertices. Let \mathcal{H} be the (ε, d) -reduced graph of H obtained from Lemma 3.2 with some parameters $\varepsilon, d > 0$ and $\ell = 2$ (and the corresponding bipartition if H is of type 2).*

Then \mathcal{H} is $(\mu + rd + 2\varepsilon, \nu - rd - 2\varepsilon)$ -robustly 2-matchable. Moreover, the type of \mathcal{H} coincides with the type of H .

Proof. Suppose that \mathcal{H} has m vertices. If H is a robustly 2-matchable graph of the first type, then Proposition 3.3(a) guarantees that $\delta(\mathcal{H}) \geq (1/2 - \mu - rd - \varepsilon)m$, and Proposition 3.3(c) implies that every set of $(1/2 - \nu)m$ vertices spans at least $(\nu - rd - \varepsilon)m^2$ edges.

Now suppose that H is of type 2 with bipartition $\{A, B\}$. Then \mathcal{H} is balanced bipartite, as well. By Proposition 3.3(a), we have $\delta(\mathcal{H}) \geq (1/32 - \mu - rd - \varepsilon)m$. By Proposition 3.3(b), we have $\deg_{\mathcal{H}}(x) \geq (1/3 - \mu - rd - \varepsilon)m$ for all but at most $(1/64 + \mu + \varepsilon)m$ vertices $x \in V(\mathcal{H})$. \square

We will also need the following lemma, which provides sufficient conditions for the existence of b -matchings in a graph.

Lemma 3.11. *Let t, γ be constants, and let H be a (μ, ν) -robustly 2-matchable graph on m vertices such that $m/t \leq \gamma \leq \mu \leq \nu/4 < 1/4000$. Then H has a perfect b -matching for every function $b : V(H) \rightarrow \mathbb{Z}^{\geq 0}$ such that*

- (a) $(1 - \gamma)t \leq b(x) \leq t$ for every $x \in V(H)$,
- (b) $\sum_{x \in \Psi} b(x)$ is even for every component Ψ of H , and
- (c) if H is of type 2 with bipartition $\{X, Y\}$, then $\sum_{x \in X} b(x) = \sum_{y \in Y} b(y)$.

Proof. As $\sum_{x \in \Psi} b(x)$ is even in every component Ψ , we can pair up the vertices with odd $b(x)$ within each component. Consider a family \mathcal{P} that contains one path in H between each such pair and let $\omega_0 : E(H) \rightarrow \mathbb{Z}^{\geq 0}$ be the function for which $\omega_0(e)$ is the number of paths in \mathcal{P} containing e . Then it is easy to see that $b_0(x) = \sum_{y \in N(x)} \omega_0(xy)$ is odd if and only if $b(x)$ is odd, so $b_1(x) = b(x) - b_0(x)$ is even for every x . Then for every vertex x ,

$$(1 - 2\gamma)t \leq (1 - \gamma)t - m \leq b_1(x) \leq t,$$

and if H is of type 2, then we also have

$$\sum_{x \in X} b_1(x) = \sum_{x \in X} b(x) - \sum_{e \in E(H)} \omega_0(e) = \sum_{y \in Y} b(y) - \sum_{e \in E(H)} \omega_0(e) = \sum_{y \in Y} b_1(y).$$

Let H' denote the graph obtained by replacing each vertex x by a set $W(x)$ of size $b_1(x)/2$ and replacing each edge xy by a complete bipartite graph with bipartition $W(x) \cup W(y)$. Then H' has $n = \sum_{x \in V(H)} b_1(x)/2$ vertices, and $(1 - 2\gamma)tm/2 \leq n \leq tm/2$. We will show that H' has a perfect 2-matching ω' . Then $\omega_1(xy) = \sum_{x' \in W(x), y' \in W(y)} \omega'(x'y')$ is a perfect b_1 -matching in H , and hence $\omega(xy) = \omega_0(xy) + \omega_1(xy)$ is a perfect b -matching in H .

Let us first consider the case when H is a robustly 2-matchable graph of type 1. As $\delta(H) \geq (1/2 - \mu)m$, we readily get $\delta(H') \geq (1/2 - \mu)(1 - 2\gamma)tm/2 \geq (1/2 - \mu - \gamma)n$. As in the proof of Lemma 3.8, it is enough to show that every independent set in H' has size at most $(1/2 - \mu - \gamma)n$, because then $|N_{H'}(S)| \geq |S|$ holds for every independent S , and we can apply Theorem 3.6 to get a perfect 2-matching.

So take any independent set S in H' , and observe that if $u \in S \cap W(x)$ and $v \in S \cap W(y)$ for some x, y in H , then xy is not an edge of H (otherwise v and w are adjacent in H'). So $S \subseteq \bigcup_{x \in U} W(x)$ for some independent set U in H . Since H is of type 1, we have $|U| \leq (1/2 - \nu)m$. Thus

$$|S| \leq (1/2 - \nu)tm/2 \leq \frac{1/2 - \nu}{1 - 2\gamma}n \leq (1/2 - \nu + 2\gamma)n \leq (1/2 - \mu - \gamma)n,$$

as needed.

Now suppose that H is of type 2. In this case, $\sum_{x \in X} b_1(x) = \sum_{y \in Y} b_1(y)$ guarantees that H' is also balanced bipartite. Also, $\delta(H) \geq (1/32 - \mu)m$ implies $\delta(H') \geq (1/32 - \mu - \gamma)n$ as before. Moreover, for every vertex x of degree at least $(1/3 - \mu)m$ in H , we get that every vertex in $W(x)$ has degree at least $(1/3 - \mu - \gamma)n$ in H' . As there are at most $(1/64 + \mu)tm/2 < (1/64 + \mu + \gamma)n$ exceptions, we see that H' is $(\mu + \gamma, \nu)$ -robustly 2-matchable. In particular, by Lemma 3.8, it has a perfect 2-matching. \square

3.5 Cycle covers in unbalanced bipartite graphs

Another tool we need is the following variant of a lemma of Erdős, Gyárfás and Pyber [13]. It finds a monochromatic cycle cover of the smaller part of an unbalanced bipartite graph if this part has large minimum degree.

Lemma 3.12 (Erdős–Gyárfás–Pyber [13]). *Let H be an r -coloured bipartite graph with bipartition $\{A, B\}$. Suppose that $|A| \geq 100^3 r^3 |B|$ and that every vertex in B has at least $|A|/100$ neighbours in A . Then there are $100r^2$ monochromatic pairwise vertex-disjoint proper cycles and edges that together cover all vertices of B .*

We give a short proof of this lemma for completeness in Appendix A, although our argument is nearly identical to the original proof.

3.6 Random sampling

The following lemma is a well-known Chernoff-type bound on the tail of the binomial distribution (see e.g. [22, Theorem 2.1]).

Lemma 3.13 (Chernoff bound). *Let $X \sim \text{Bin}(n, p)$ be a binomial random variable. Then the following bounds hold for every $0 \leq a \leq 1$.*

- $\Pr[X < (1 - a)np] \leq e^{-a^2np/2}$, and
- $\Pr[X > (1 + a)np] \leq e^{-a^2np/3}$.

In our proof, we will need a small set of vertices that contains many neighbours of every large-degree vertex. As shown by the next result, a randomly chosen set satisfies these properties. The proof is a routine application of Chernoff-type bounds; see Appendix A.

Proposition 3.14. *Let G be a graph on n vertices with (ε, d) -regular partition $\{V_0, \dots, V_m\}$ as provided by Lemma 3.2. Also, let p be a positive parameter, and let $B \subseteq V = V(G)$ be a vertex set satisfying $V_0 \subseteq B$ and $|B \cap V_i| \leq 10p|V_i|$ for every $i \in [m]$. If $m \log n / \sqrt{n} < p < 1/100$ and $\varepsilon < 1/10$, then there is a set $A \subseteq V \setminus B$ with the following properties.*

- (a) $|A| \geq (p/2)n$,
- (b) $|A \cap V_i| \leq 2p|V_i|$ for every $i \in [m]$,
- (c) $\deg(v, A \cap V_i) \geq (p/2) \deg(v, V_i)$ for every $v \in V$ and $i \in [m]$ with $\deg(v, V_i) > 30p|V_i|$,
- (d) $\deg(v, A) \geq |A|/100$ for every vertex $v \in V$ with $\deg(v, V \setminus B) > n/40$.

4 Main proof

The proof of Theorem 1.2 comes as a combination of the following two results.

Theorem 4.1. *Let $1/n \ll \mu \ll 1$, and let G be an r -edge-coloured graph on n vertices with minimum degree $\delta(G) \geq n/2 + 1200r \log n$. Then the vertices of G can be partitioned into at most $400r + 2$ monochromatic cycles and a $(\mu, 20\mu)$ -robustly 2-matchable graph H on at least $n/2$ vertices.*

Theorem 4.2. *Let $1/n \ll \mu \leq \nu/20 \ll 1$. Every r -edge-coloured (μ, ν) -robustly 2-matchable graph on n vertices can be partitioned into $(1/\mu + 200)r^2$ monochromatic cycles.*

Let us first elaborate on the conditions that are implicit in our \ll notation. We will need to select the five parameters ν, μ, d, ε and n , in this order. As a point of reference, we describe here the exact constraints that come from the proofs of Theorems 4.1 and 4.2:

$$\begin{aligned}
 \nu &< \frac{1}{1000}, \\
 \mu &< \min \left\{ \frac{1}{700000}, \frac{\nu}{20} \right\}, \\
 d &\leq \frac{\mu}{r}, \\
 \varepsilon &< \min \left\{ \frac{1}{10^{13}r^6}, \frac{\mu^4}{20^4}, \frac{d^2}{4000}, \varepsilon_{3.1}(d), \varepsilon_{3.4}(d) \right\}, \\
 n &> \max \left\{ \frac{4}{\varepsilon} (M_{3.2}(\varepsilon, r, 2))^4, n_{3.1}(\varepsilon), n_{3.4}(\varepsilon) \right\},
 \end{aligned} \tag{\ll}$$

where $M_{3.2}, \varepsilon_{3.1}, \varepsilon_{3.4}, n_{3.1}$ and $n_{3.4}$ are the appropriate constants coming from Lemmas 3.1, 3.2 and 3.4. Let us emphasize that $1/n < \varepsilon < d < \mu < \nu < 1$.

The proof of Theorem 4.1 is a somewhat technical argument that shows that either G is already robustly 2-matchable of type 1, or it can be turned into a type 2 robustly 2-matchable graph by deleting few monochromatic cycles. We defer its proof to Section 5, and proceed with the proof of Theorem 4.2.

Proof of Theorem 4.2. Let $G = (V, E)$ be an r -edge-coloured (μ, ν) -robustly 2-matchable graph with $20\mu \leq \nu$. If G is of type 2, then it is a balanced bipartite graph, and we denote its bipartition by $\{A, B\}$. By (\ll) , we are guaranteed a partition V_0, V_1, \dots, V_m of $V(G)$ as detailed in Lemma 3.2. Let \mathcal{G} be the corresponding (ε, d) -reduced graph. If G is of type 2, then \mathcal{G} is also balanced bipartite, and we denote its bipartition by $\{\mathcal{A}, \mathcal{B}\}$. Note that \mathcal{G} has $m \leq M_{3.2}(\varepsilon, r, 2)$ vertices.

By Lemma 3.10, and using $\varepsilon \leq \mu/2$ and $d \leq \mu/r$, we know that \mathcal{G} is $(3\mu, \nu - 2\mu)$ -robustly 2-matchable. Let \mathcal{H} denote the subgraph of \mathcal{G} that consists of all edges contained in monochromatic components of order at least $(\mu/r)m$. Then \mathcal{H} is the union of at most $(1/\mu)r^2$ monochromatic components, and $\deg_{\mathcal{H}}(x) \geq \deg_{\mathcal{G}}(x) - \mu m$ for every vertex x in \mathcal{G} . By Lemma 3.9, \mathcal{H} is $(4\mu, \nu - 3\mu)$ -robustly 2-matchable. Moreover, the type of \mathcal{H} coincides with the type of G . As $20\mu \leq \nu \leq 1/1000$ (by (\ll)), Lemma 3.8 implies that \mathcal{H} contains a perfect 2-matching ω . We denote by \mathcal{M} the edges of \mathcal{H} that have non-zero weight under ω .

Let us now call a vertex $v \in V_i$ ($i \in [m]$) *good* if v has typical degree in each regular pair (V_i, V_j) that corresponds to an edge of \mathcal{M} . In other words, v is good if $\deg_c(v, V_j) \geq (d - \varepsilon)|V_j|$ for each edge $x_i x_j \in \mathcal{M}$ of colour c . We call all other vertices of G *bad*.

Claim 4.3. *There is a collection \mathcal{C}^b of at most $100r^2$ vertex-disjoint monochromatic proper cycles and edges in G covering all bad vertices such that*

$$|V_i \cap V(\mathcal{C}^b)| \leq 5\sqrt{\varepsilon}|V_i| \quad \text{for every } i \in [m]. \quad (4.1)$$

Proof. Let B be the set of bad vertices (note that $V_0 \subseteq B$). By (3.1), and because \mathcal{M} is a 2-matching, we know that $|B \cap V_i| \leq 2\varepsilon|V_i|$ for every $i \in [m]$. In particular, $|B| \leq 2\varepsilon|V_i| \cdot m + |V_0| \leq 2\varepsilon|V_i| \cdot m + \varepsilon n \leq 3\varepsilon n$.

This together with (\ll) means that we can apply Proposition 3.14 with $p = 2\sqrt{\varepsilon}$ to obtain a set A of size $|A| \geq \sqrt{\varepsilon}n \stackrel{(\ll)}{\geq} 3 \cdot 100^3 r^3 \varepsilon n \geq 100^3 r^3 |B|$ such that $|A \cap V_i| \leq 4\sqrt{\varepsilon}|V_i|$ for every $i \in [m]$, and each vertex $v \in G$ with $\deg_G(v, V \setminus B) > n/40$ has at least $|A|/100$ neighbours in A . As $\delta(G) \geq (1/32 - \mu)n$ and $|B| < 3\varepsilon n$, this actually holds for every vertex of G , and in particular for every vertex in B . But then Lemma 3.12 provides a set \mathcal{C}^b of at most $100r^2$ disjoint monochromatic proper cycles and edges covering B . Note that the vertices of \mathcal{C}^b are contained in $A \cup B$, so (4.1) clearly holds. \square

Claim 4.4. *There is a collection $\mathcal{C}^{\mathcal{H}}$ of at most $(1/\mu)r^2$ vertex-disjoint monochromatic proper cycles and edges in G , all disjoint from \mathcal{C}^b , such that*

- (a) for every edge $e = x_i x_j$ of \mathcal{H} , there is an edge $u_e v_e$ of colour $c(e)$ in $\mathcal{C}^{\mathcal{H}}$ between vertices $u_e \in V_i$ and $v_e \in V_j$ that have typical degree in the regular pair (V_i, V_j) , and
- (b) $|V_i \cap V(\mathcal{C}^{\mathcal{H}})| \leq \varepsilon |V_i|$ for every $i \in [m]$.

Proof. We will apply a simple algorithm to find one cycle for each monochromatic component of \mathcal{H} . For this, take a component Φ of colour c , and let $e_1, \dots, e_s \in E(\mathcal{H})$ be its edges. We perform the following two steps:

- (1) For $i = 1, \dots, s$, let $e_i = y_i z_i$, and pick $u_i \in V(y_i)$ and $v_i \in V(z_i)$ that are not yet used, but have typical degree in the regular pair $(V(y_i), V(z_i))$, and $u_i v_i$ is a c -coloured edge in G .
- (2) Use Lemma 3.4 to find a c -coloured v_i - u_{i+1} path P_i in G of order at most $2m$ that avoids all previously used vertices (except v_i and u_{i+1}), for $i = 1, \dots, s$.

If these steps work, then $C_\Phi = u_1 v_1 P_1 u_2 v_2 P_2 \dots u_s v_s P_s u_1$ is a c -coloured cycle that takes care of all edges in Φ . Repeating this for every component gives us $(1/\mu)r^2$ disjoint monochromatic cycles satisfying Condition (a). Condition (b) is also satisfied because the edges and paths produced by these steps use at most $|E(\mathcal{H})| \cdot 2m \leq m^3 \stackrel{(\ll)}{\leq} \varepsilon |V_i|$ vertices in G (using $n > 2m^4/\varepsilon$ and $|V_i| > n/(2m)$ in the second inequality). We just need to check that these steps can indeed be applied.

For Step (1), note that by (3.1), $V(y_i)$ and $V(z_i)$ each have at least $(1-\varepsilon)|V(y_i)|$ typical vertices, of which at most $2\varepsilon|V(y_i)|$ have been used in former steps and at most $5\sqrt{\varepsilon}|V_i|$ are in \mathcal{C}^b by (4.1), as noted above. But then there is an edge between unused typical vertices in colour c because $\varepsilon < 1/100$ and $(V(y_i), V(z_i))$ is ε -regular. For Step (2), we just need to apply Lemma 3.4 with the set W consisting of the vertices of \mathcal{C}^b in $V(y_i) \cup V(z_i)$, as well as all previously used vertices except v_i and u_{i+1} . This is possible because $|W| < 12\sqrt{\varepsilon}|V_i|$. \square

Note that \mathcal{C}^b and $\mathcal{C}^{\mathcal{H}}$ together contain at most $(1/\mu + 100)r^2$ cycles. For parity reasons, we need another small collection \mathcal{C}^s of single vertices. For every component Ψ of (the uncoloured graph) \mathcal{H} , add a single vertex of $\bigcup_{x \in \Psi} V(x) \setminus V(\mathcal{C}^b \cup \mathcal{C}^{\mathcal{H}})$ to \mathcal{C}^s if $|\bigcup_{x \in \Psi} V(x) \setminus V(\mathcal{C}^b \cup \mathcal{C}^{\mathcal{H}})|$ is odd. Since \mathcal{H} is $(4\mu, \nu - 3\mu)$ -robustly 2-matchable, \mathcal{H} has at most two components. Thus we have $|\mathcal{C}^s| \leq 2$. Moreover, if \mathcal{H} is of type 2, then it has only one component with an even number of vertices. So in this case \mathcal{C}^s is just the empty set. Write $\mathcal{C}^0 = \mathcal{C}^b \cup \mathcal{C}^{\mathcal{H}} \cup \mathcal{C}^s$, and note that

$$|V_i \cap V(\mathcal{C}^0)| \leq |V_i \cap V(\mathcal{C}^b \cup \mathcal{C}^{\mathcal{H}} \cup \mathcal{C}^s)| \leq (5\sqrt{\varepsilon} + \varepsilon)|V_i| + 2 \leq 6\sqrt{\varepsilon}|V_i| \quad (4.2)$$

for every $i \in [m]$. The rest of the proof will extend the cycles in $\mathcal{C}^{\mathcal{H}}$ so that they cover all the remaining vertices.

More precisely, $\mathcal{C}^{\mathcal{H}}$ will serve as the ‘‘skeleton’’ of our cycle cover in the sense that we will use Lemma 3.1 to replace each edge $u_e v_e$ (corresponding to some $e = x_i x_j$ in \mathcal{H}) with a u_e - v_e path P_e in (V_i, V_j) . But first we need to decide how long these paths should be. So fix an ℓ such that

$$(1 - \varepsilon^{1/4})|V_i| \leq \ell \leq (1 + \varepsilon^{1/4})|V_i| + 2 \text{ and } \ell \text{ is divisible by } 2. \quad (4.3)$$

Our plan is to cover at least ℓ vertices in each cluster by the paths corresponding to the edges of the 2-matching \mathcal{M} . By (4.2), this leaves $b(x_i)$ vertices in V_i , where

$$\varepsilon^{1/4}|V_i| \geq b(x_i) = |V_i \setminus V(\mathcal{C}^0)| - \ell \geq \varepsilon^{1/4}|V_i| - |V_i \cap \mathcal{C}^0| - 2 \geq 0 \quad (4.4)$$

vertices in each V_i .

Claim 4.5. \mathcal{H} has a perfect b -matching $\omega_0 : E(\mathcal{H}) \rightarrow \mathbb{Z}^{\geq 0}$.

Proof. By (4.4) and since $|V_i \cap \mathcal{C}^0| + 2 \leq 12\sqrt{\varepsilon}|V_i| + 2 \stackrel{(\ll)}{\leq} \mu\varepsilon^{1/4}|V_i|$, we have

$$\varepsilon^{1/4}|V_i| \geq b(x_i) \geq \varepsilon^{1/4}|V_i| - |V_i \cap \mathcal{C}^0| - 2 \geq (1 - \mu)\varepsilon^{1/4}|V_i|.$$

Moreover, the definition of \mathcal{C}^s implies that $\sum_{x \in \Psi} b(x)$ is even for every component Ψ of \mathcal{H} . Recall that \mathcal{H} is $(4\mu, \nu - 3\mu)$ -robustly 2-matchable. If G is of type 1, then \mathcal{H} is of type 1 and we can finish by Lemma 3.11.

Suppose that G is of type 2 and thus \mathcal{H} is of type 2 as well. Recall that G has bipartition $\{A, B\}$ and \mathcal{H} has bipartition $\{\mathcal{A}, \mathcal{B}\}$. Moreover, $\bigcup_{x \in \mathcal{A}} V(x) \subseteq A$ and $\bigcup_{y \in \mathcal{B}} V(y) \subseteq B$. Since G is balanced bipartite and $\mathcal{C}^b \cup \mathcal{C}^{\mathcal{H}}$ consists of proper cycles and edges, it follows that $\sum_{x \in \mathcal{A}} b(x) = \sum_{y \in \mathcal{B}} b(y)$. Therefore, we can finish by Lemma 3.11. Note that the set \mathcal{C}^s did not play a role in this case (and is anyways empty). \square

Let ω_0 be the perfect b -matching guaranteed by Claim 4.5. Define $\omega : E(\mathcal{H}) \rightarrow \mathbb{Z}^{\geq 0}$ as

$$\omega(x_i x_j) = \begin{cases} \omega_0(x_i x_j) & \text{for } x_i x_j \notin \mathcal{M}, \\ \omega_0(x_i x_j) + \frac{\ell}{\deg_{\mathcal{M}}(x_i)} & \text{for } x_i x_j \in \mathcal{M}. \end{cases}$$

Note that this is well-defined because $\deg_{\mathcal{M}}(x_i) = \deg_{\mathcal{M}}(x_j)$, and integral because ℓ is even and $\deg_{\mathcal{M}}(x_i) \in [2]$. Then for every vertex x_i in \mathcal{H} , we have $\sum_{x_i e \in \mathcal{H}} \omega(e) = |V_i \setminus V(\mathcal{C}^0)|$ and $\sum_{x_i e \in \mathcal{H} \setminus \mathcal{M}} \omega(e) \leq b(x_i) \leq \varepsilon^{1/4}|V_i|$.

Claim 4.6. For every edge $e = x_i x_j$ in $E(\mathcal{H})$, there is a u_e - v_e path P_e of colour $c(e)$ in $G[V_i, V_j]$ that contains exactly $\omega(e) + 1$ vertices in each of V_i and V_j . Moreover, these paths can be chosen so that they are internally vertex-disjoint from each other and from \mathcal{C}^0 .

Proof. Let us first apply Proposition 3.14 with $p = 2\sqrt{\varepsilon}$ and $B = V(\mathcal{C}^0)$ to get a set S^1 with the properties given in the statement of the proposition, and then apply it again with the same p and $B = V(\mathcal{C}^0) \cup S^1$ to get another such set S^2 . This is possible because $V_0 \subseteq V(\mathcal{C}^0)$ holds, and we also have $|V(\mathcal{C}^0) \cap V_i| \leq 6\sqrt{\varepsilon}|V_i|$, and thus $|S^1 \cap V_i| \leq 4\sqrt{\varepsilon}|V_i|$ for every $i \in [m]$. Let $S_i^b = S^b \cap V_i$ for every $i \in [m]$ and $b \in [2]$. Then

- (a) $|S_i^b| \leq 4\sqrt{\varepsilon}|V_i|$ for every $i \in [m]$ and $b \in [2]$, and
- (b) for every edge $x_i x_j$ in \mathcal{H} of colour c and every vertex $v \in V_j$ with typical degree in the regular pair (V_i, V_j) , we have $\deg_c(v, S_i^b) \geq 6\varepsilon|V_i|$ for $b \in [2]$.

To see (b), note that every such vertex v of typical degree satisfies $\deg_c(v, V_i) \geq (d - \varepsilon)|V_i| > 60\sqrt{\varepsilon}|V_i|$ (using $d \stackrel{(\ll)}{>} 61\sqrt{\varepsilon}$), so by Proposition 3.14, $\deg_c(v, S_i^b) \geq \sqrt{\varepsilon}(d - \varepsilon)|V_i| \stackrel{(\ll)}{>} 60\varepsilon|V_i|$.

Let us first consider the edges e_1, \dots, e_s of $\mathcal{H} \setminus \mathcal{M}$. We will find the u_k - v_k paths P_k (where $u_k v_k$ is the edge in $\mathcal{C}^{\mathcal{H}}$ corresponding to e_k , as obtained in Claim 4.4) one by one so that for every k , the vertex set of $\mathcal{P}_k = \bigcup_{j=1}^{k-1} P_j$ is disjoint from each S_i^2 , and intersects each S_i^1 in at most $k - 1$ vertices. Suppose we have already found P_1, \dots, P_{k-1} . Let us also assume that $u_k \in V_i$ and $v_k \in V_j$ (so $e_k = x_i x_j$), and let c be the colour of e_k .

If $\omega(e_k) = 0$, then there is nothing to do: we can take $P_k = u_k v_k$. If $\omega(e_k) = 1$, then $\deg_c(u_k, S_i^1 \setminus V(\mathcal{P}_k)) \geq 4\varepsilon|V_j| - k \stackrel{(\ll)}{\geq} \varepsilon|V_j|$ (using $\varepsilon|V_i| > \varepsilon n / (2m) > m^2$) and similarly, $\deg_c(v_k, S_j^1 \setminus V(\mathcal{P}_k)) \geq \varepsilon|V_i|$. Hence, by regularity, we can find adjacent vertices $u \in S_i^1 \setminus V(\mathcal{P}_k)$ and $v \in S_j^1 \setminus V(\mathcal{P}_k)$ such that $P_k = u_k v u v_k$ is a c -coloured path, as needed.

So suppose $\omega(e_k) > 1$. Let $W = V(\mathcal{C}^0) \cup S^1 \cup S^2 \cup V(\mathcal{P}_k)$ be the set of “forbidden” vertices. We will again need neighbours $u \in S_i^1 \setminus V(\mathcal{P}_k)$ and $v \in S_j^1 \setminus V(\mathcal{P}_k)$ of v_k and u_k respectively, but this time we want to apply Lemma 3.1 to connect them with a u - v path of the right length that avoids W .

We have seen above that $\deg_c(u_k, S_j^1 \setminus V(\mathcal{P}_k)) \geq \varepsilon|V_j|$. Also,

$$|V_i \setminus W| \geq |V_i| - |V_i \cap V(\mathcal{C}^0)| - |S_i^1| - |S_i^2| - b(x_i) \geq (1 - 14\sqrt{\varepsilon} - \varepsilon^{1/4})|V_i| \stackrel{(\ll)}{\geq} |V_i|/2,$$

so by regularity, there is a neighbour $v \in S_j^1 \setminus V(\mathcal{P}_k)$ of u_k such that $\deg_c(v, V_i \setminus W) \geq (d - \varepsilon)|V_i|/2 \stackrel{(\ll)}{\geq} \varepsilon|V_i|$. Similarly, there is a neighbour $u \in S_i^1 \setminus V(\mathcal{P}_k)$ of v_k such that $\deg_c(u, V_j \setminus W) \geq \varepsilon|V_i|$. As $\omega(e_k) \leq \varepsilon^{1/4}|V_i| \stackrel{(\ll)}{\leq} (1 - \sqrt{\varepsilon})|V_i \setminus W|$, we can apply Lemma 3.1 (with $U_1 = V_i \setminus W$ and $U_2 = V_j \setminus W$) to find a c -coloured v - u path P' of order $2\omega(e_k)$ that is internally vertex-disjoint from W . But then $P_k = u_k v P' u v_k$ is a path satisfying our requirements.

Finally, let e_{s+1}, \dots, e_{s+t} be the edges of \mathcal{M} . Note that each vertex of \mathcal{H} is incident to exactly one or two of these edges. Using the same notation as before, we will find the u_k - v_k paths P_k so that \mathcal{P}_k is disjoint from S_i^2 unless $e_k = x_i x_j$ is the last edge at x_i (according to the ordering), and similarly for x_j .

Fix k , and let $U_i = V_i \setminus (V(\mathcal{C}^0) \cup V(\mathcal{P}_k))$ if e_k is the last edge at x_i , and let $U_i = V_i \setminus (V(\mathcal{C}^0) \cup V(\mathcal{P}_k) \cup S_i^2)$ otherwise. Using $|S_i^2| \leq \ell/2$ and the assumption that ε is small (see (\ll)), it is easy to check from the definitions that we have $|U_i| \geq \omega(e_k) \geq \ell/2 \geq |V_i|/3$. We similarly get $|U_j| \geq \omega(e_k) \geq |V_i|/3$ for the analogously defined U_j .

We want to use Lemma 3.1 to find the required u_k - v_k path P_k of order $2(\omega(e_k) + 1)$. As $\min\{|U_i \cup \{u_k\}|, |U_j \cup \{v_k\}|\} \geq \omega(e_k) + 1$, we just need to check that $\delta(G[U_i, U_j]) \geq 5\varepsilon|V_i|$. This follows from the properties of S_i^1 and S_i^2 : If e_k is the last edge at x_i , then $S_i^2 \subseteq U_i$, and otherwise all but k vertices of S_i^1 are in U_i . Either way, by (b) we obtain $\deg_c(v, U_i) \geq 6\varepsilon|V_i| - k \geq 5\varepsilon|V_i|$ for every $v \in U_j$, and similarly, $\deg_c(u, U_j) \geq 5\varepsilon|V_j|$ for every $u \in U_i$, as needed. (Here we also used that the vertices of non-typical degree are all in $V(\mathcal{C}^0)$ by Claim 4.3.) \square

Now for every $e \in E(\mathcal{H})$, we replace the edge $u_e v_e$ with the path P_e in the appropriate cycle of $\mathcal{C}^{\mathcal{H}}$. This gives us $(1/\mu + 100)r^2 + 2 \leq (1/\mu + 200)r^2$ monochromatic cycles that cover all vertices in V_0 , and (by the definition of the function ω) $|V_i|$ vertices in each V_i . In other words, we find a monochromatic cycle partition of G , as needed. \square

5 The structural lemma

Let us now prove the main structural theorem from Section 4.

Our proof makes use of a classical result of Bondy and Simonovits on the extremal number of even cycles.

Theorem 5.1 (Bondy–Simonovits, [4]). *Let G be a graph on n vertices with at least e edges, and let ℓ be such that $\ell \leq e/(100n)$ and $\ell n^{1/\ell} \leq e/(10n)$. Then G contains a cycle of length exactly 2ℓ .*

More precisely, we need the following immediate corollary:

Corollary 5.2. *Let $\ell \geq (2/\log 10) \log n$ be even. Then every graph on n vertices with average degree at least 100ℓ contains a cycle of length ℓ .*

The main idea of the proof of Theorem 4.1 is to show that if G is not robustly 2-matchable of type 1, then it has a bipartition $\{X, Y\}$ such that $G[X, Y]$ is essentially robustly 2-matchable of type 2, except it might be unbalanced. We use Corollary 5.2 to balance out this bipartite subgraph by covering some vertices of G with cycles induced by X and Y .

Proof of Theorem 4.1. Let us assume that G is not $(\mu, 20\mu)$ -robustly 2-matchable of type 1. The following claim provides us with useful information regarding the structure of G . Its proof, while somewhat technical, is routine.

Claim 5.3. *There is a partition $\{X, Y\}$ of the vertices with the following properties.*

- (a) $|X| \geq |Y| \geq \frac{n}{2} - 5400\mu n$,
- (b) $\delta(G[X, Y]) \geq \frac{n}{16} - 10800\mu n$,
- (c) all but $10800\mu n$ vertices have degree at least $\frac{2n}{5} - 10800\mu n$ in $G[X, Y]$, and
- (d) if $|X| > \frac{n}{2}$ we have $\Delta(G[X]) \leq \frac{n}{16}$.

Proof. As $\delta(G) \geq n/2$, the assumption that G is not type 1 $(\mu, 20\mu)$ -robustly 2-matchable implies the existence of a set S_0 of size at least $(1/2 - 20\mu)n$ that spans fewer than $20\mu n^2$ edges. This set S_0 cannot contain more than $480\mu n$ vertices v satisfying $\deg_G(v, S_0) \geq n/12$, so there is a subset $S_1 \subseteq S_0$ of size exactly $(1/2 - 500\mu)n$ such that $\Delta(G[S_1]) \leq n/12$.

Now let T be the set of vertices not in S_1 that send at least $2n/5$ edges into S_1 , and let S_2 be the set of vertices not in $S_1 \cup T$. If q denotes the size of S_2 , then we have $|S_1| = n/2 - 500\mu n$ and

$|T| = n/2 + 500\mu n - q$. We can bound q by double-counting the edges of G between S_1 and $T \cup S_2$. Indeed, counting from S_1 , the number of such edges is at least $|S_1| \cdot n/2 - 40\mu n^2 = (n/2)^2 - 290\mu n^2$. On the other hand, counting from $T \cup S_2$, there are at most

$$|T||S_1| + |S_2| \cdot \frac{2n}{5} \leq \left(\frac{n}{2} + 500\mu n - q\right) \frac{n}{2} + q \cdot \frac{2n}{5} = \left(\frac{n}{2}\right)^2 + 250\mu n^2 - q \cdot \frac{n}{10}$$

such edges. Putting these together, we get that $q \cdot n/10 \leq 540\mu n^2$, so $q \leq 5400\mu n$.

Setting $S = S_1 \cup S_2$, we obtain the following bounds on the degrees in $G[S, T]$.

$$\begin{aligned} \text{For every } v \in T, \quad \deg(v, S) &\geq \frac{2n}{5}; \\ \text{for every } v \in S_1, \quad \deg(v, T) &\geq \frac{n}{2} - \Delta(G[S_1]) - |S_2| \geq \frac{2n}{5} - 5400\mu n; \\ \text{for every } v \in S_2, \quad \deg(v, T) &\geq \frac{n}{2} - \frac{2n}{5} - |S_2| > \frac{n}{16} - 5400\mu n. \end{aligned} \tag{5.1}$$

Now let X_0 be the larger of the two sets S and T , and let Y_0 the smaller one. We then have $|X_0| = n/2 + k$ and $|Y_0| = n/2 - k$, where $k = |500\mu n - q| \leq 5400\mu n$. Let $Z \subseteq X_0$ be the set of vertices in X_0 with at least $n/16$ neighbours in X_0 .

If $|Z| \geq k$, then let $Z_0 \subseteq Z$ be a subset of size k . We claim that $X = X_0 \setminus Z_0$ and $Y = Y_0 \cup Z_0$ satisfy the conditions. Indeed, as $|Z_0| \leq 5400\mu n$, we get $\deg(v, X) \geq \deg(v, X_0) - 5400\mu n$ for every vertex v in Y , and $\deg(v, Y) \geq \deg(v, Y_0)$ for every $v \in X$. Combining this with (5.1), we see that $\delta(G[X, Y]) \geq n/16 - 10800\mu n$, and every vertex not in $S_2 \cup Z_0$ has degree at least $2n/5 - 10800\mu n$, establishing (b) and (c). As $|X| = |Y| = n/2$, (a) and (d) are also satisfied.

If $|Z| < k$, then we take $Z_0 = Z$ instead. The same argument shows that (b) and (c) hold. The definition of Z implies $\Delta(G[X]) < n/16$, establishing (d), while (a) holds because $k \leq 5400\mu n$. \square

Let X and Y be as in the claim. If $|X| = |Y|$, then set $H = G[X, Y]$. As we will see, this satisfies the conditions. In the meantime, we may assume that $|X| = \lceil n/2 \rceil + k$ for some $0 \leq k \leq 5400\mu n$.

Let us first consider the case when $k \geq 400r \log n$. We can write $2k = \ell_1 + \dots + \ell_t$ as the sum of $t \leq 400r + 1$ even numbers such that $k/(400r) \leq \ell_i \leq k/(200r)$ for every i . We will find pairwise disjoint monochromatic cycles C_1, \dots, C_t in X , where each C_i has length ℓ_i .

Suppose we have already found C_1, \dots, C_{i-1} with these properties. We want to apply Corollary 5.2 to find C_i . As $\delta(G) \geq n/2$, the minimum degree of $G[X]$ is at least k , therefore X induces at least $k|X|/2 \geq kn/4$ edges. On the other hand, (d) implies that the vertices of C_1, \dots, C_{i-1} are incident to at most $2k \cdot n/16 = kn/8$ edges. That is, at least half of the edges in $G[X]$ are not incident to any of the cycles C_1, \dots, C_{i-1} , and hence the average degree induced by $X' = X \setminus V(C_1 \cup \dots \cup C_{i-1})$ is at least $k/2$. The average degree of $G[X']$ in the most common colour (say blue) is then at least $k/(2r) \geq 100\ell_i$, so Corollary 5.2 provides a blue cycle of length ℓ_i , as needed (using $\ell_i \geq \log n \geq (2/\log 10) \log n$).

Let \mathcal{C} be the set of cycles C_1, \dots, C_t . If n is odd, we add another vertex in X as a singleton cycle to \mathcal{C} . Then \mathcal{C} contains at most $400r + 2$ cycles, and $A = X \setminus V(\mathcal{C})$ and $B = Y \setminus V(\mathcal{C})$ satisfy $|A| = |B| = n/2 - k \geq (1/2 - 5400\mu)n$. In this case, we choose $H = G[A, B]$.

Finally, suppose $0 \leq k < 400r \log n$, and let $\log n \leq \ell < \log n + 2$ be even. We can write $\ell + 2k = \ell_1 + \dots + \ell_t$ as the sum of $t \leq 200r + 1$ even numbers such that $\log n \leq \ell_i \leq 2 \log n$ for every i . We will again find a monochromatic cycle C_i of length ℓ_i in $G[X]$ for every i , but this time we will also need an ℓ -cycle C induced by Y to balance out the graph.

The minimum degree of $G[Y]$ is at least $n/2 + 1200r \log n - (n/2 + k) \geq 100r\ell$, so the most common colour (say blue) has average degree at least 100ℓ . By Corollary 5.2, there is a blue cycle C of length ℓ in Y .

To find C_1, \dots, C_t , we use the same argument as before. Suppose we already have C_1, \dots, C_{i-1} . The minimum degree of $G[X]$ is at least $1200r \log n$, so X induces at least $1200r|X| \log n/2 \geq 300rn \log n$ edges. Out of these, at most $2k \cdot n/16 = kn/8 \leq 50rn \log n$ are incident to some of the cycles C_1, \dots, C_{i-1} . Hence, the average degree in $G[X']$, where $X' = X \setminus V(C_1 \cup \dots \cup C_{i-1})$, is at least $200r \log n$, and the average degree in the majority colour (say blue) is at least $200 \log n \geq 100\ell_i$. As $\ell_i \geq \log n$, we can use Corollary 5.2 to find a blue cycle C_i of length ℓ_i in X' , as needed.

Again, let \mathcal{C} be the set of cycles C, C_1, \dots, C_t and possibly a singleton in X (if n is odd). Then \mathcal{C} contains at most $200r + 3$ cycles, and $A = X \setminus V(\mathcal{C})$ and $B = Y \setminus V(\mathcal{C})$ satisfy $|A| = |B| = n/2 - k - \ell$. We set $H = G[A, B]$.

In either of the cases, H is obtained from $G[X, Y]$ by deleting at most $2k + 2 \log n + 4$ vertices, so each of the degrees can decrease by at most this value compared to (b) and (c). Assuming $\mu < 1/700000$ and $n > 100000$, we have $2k + 2 \log n + 4 \leq 10800\mu n + 2 \log n + 4 \leq n/64$. Then it is easy to check that H is a balanced bipartite graph on at least $n/2$ vertices, such that $\delta(H) \geq n/16 - 10800\mu n - n/64 \geq n/32$, and all but $n/64$ vertices have degree at least $n/3$. This H is indeed a $(\mu, 20\mu)$ -robustly 2-matchable graph of type 2. \square

6 Sharpness for the minimum degree

In this section we prove Proposition 1.1, which shows that the minimum degree condition in Theorem 1.2 is almost best possible.

Proof of Proposition 1.1. Our construction is based on the following claim.

Claim 6.1. *For every sufficiently large n , there is a graph on n vertices with minimum degree at least $\log n$ that does not contain any proper cycle of length shorter than $\log n / (4 \log \log n)$.*

Proof. During this proof, all cycles will be proper. Our construction is probabilistic. We start with the Erdős–Rényi random graph $G(n, p)$ on n vertices, where any two vertices are connected by an edge independently with probability $p = 8 \log n / n$. Next, we take a maximal collection \mathcal{C} of edge-disjoint cycles of length less than $k = \log n / (4 \log \log n)$. By maximality, $G' = G(n, p) \setminus E(\mathcal{C})$ has no cycles whose length is shorter than k . So it is enough to show that $\delta(G') \geq \ell = \log n$ with positive probability.

To see this, first note that $\deg_{G(n, p)}(v) \sim \text{Bin}(n-1, p)$ for every vertex v , so we can use the Chernoff bounds to bound the probability that the degree is small. Let \mathcal{A} be the event that

some vertex of $G(n, p)$ has degree less than $3 \log n$. By Lemma 3.13 (with $a = 5/8$) and a union bound over the vertices, we get $\Pr[\mathcal{A}] \leq n \cdot e^{-(25/16) \log n} < 1/3$ for large enough n .

Now for every vertex v of $G(n, p)$, let \mathcal{B}_v be the event that v is incident to at least ℓ pairwise edge-disjoint cycles of length shorter than k . Note that there are at most $n^{(k_1-1)+\dots+(k_\ell-1)}$ potential sets of ℓ edge-disjoint cycles of lengths k_1, \dots, k_ℓ incident to v , and each of them is a subgraph of $G(n, p)$ with probability $p^{k_1+\dots+k_\ell}$. Hence

$$\Pr[\mathcal{B}_v] \leq \sum_{3 \leq k_1, \dots, k_\ell < k} n^{k_1+\dots+k_\ell-\ell} p^{k_1+\dots+k_\ell} \leq \left(\frac{1}{n} \cdot \sum_{k' < k} (np)^{k'} \right)^\ell \leq \left(\frac{(np)^k}{n} \right)^\ell \leq \left(\frac{1}{\sqrt{n}} \right)^{\log n}$$

using $(np)^k = (8 \log n)^{\log n / (4 \log \log n)} \leq \sqrt{n}$. This means that the probability that $\mathcal{B} = \bigcup \mathcal{B}_v$ holds is at most $n \cdot n^{-(\log n)/2} < 1/3$ for large enough n .

So with positive probability, neither \mathcal{A} nor \mathcal{B} occur. But then every vertex v is incident to at least $3 \log n$ edges in $G(n, p)$, and at most $2 \log n$ of those can appear in \mathcal{C} . This implies $\delta(G') \geq \log n$, as needed. \square

Let A and B be disjoint sets, such that $|B| = n/2 + \log n / (16 \log \log n)$, and $|A| = n - |B|$. Let G be a graph with vertex set $A \cup B$, where all A - B edges are present and are red, and $G[B]$ is a blue graph provided by Claim 6.1. So $G[B]$ has minimum degree at least $\log |B| \geq \log n - 2 \geq \log n / (8 \log \log n)$, and it induces no proper cycle shorter than $\log |B| / (4 \log \log |B|) \geq \log n / (8 \log \log n)$.

Note that G has minimum degree at least $n/2 + \log n / (16 \log \log n)$. Also, every red cycle in G is either a singleton, or it covers an equal number of vertices in A and B . Moreover, every blue cycle is either a singleton, an edge, or has length at least $\log n / (8 \log \log n)$.

Let \mathcal{C} be a collection of vertex-disjoint monochromatic cycles covering all vertices of G . If \mathcal{C} contains a proper blue cycle (of length at least $\log n / (8 \log \log n)$), then the remaining cycles of \mathcal{C} must cover at least $\log n / (16 \log \log n)$ more vertices in A than in B . But A is independent, so this is only possible if \mathcal{C} contains at least $\log n / (16 \log \log n)$ singletons. So \mathcal{C} cannot contain any proper blue cycle. But then, as \mathcal{C} covers $\log n / (16 \log \log n)$ more vertices in B than in A , it must contain at least $\log n / (32 \log \log n)$ singletons or blue edges. Hence, in any case, \mathcal{C} consists of at least $\log n / (32 \log \log n)$ cycles, as desired. \square

7 Sharpness for the number of cycles

The goal of this section is to prove Proposition 1.3, which shows that the number of cycles we use to partition the vertices is best possible, up to a constant factor.

We start with some preliminary lemmas.

Lemma 7.1. *Every n -vertex graph G with $e(G) \geq (1 - \varepsilon^2)n^2/2$ has a subgraph H with $\delta(H) \geq (1 - 2\varepsilon)n$.*

Proof. Let S be the set of vertices v in G with $\deg(v) \leq (1 - \varepsilon)n$. We have $2e(G) \leq (n - |S|)n + |S|(1 - \varepsilon)n = n^2 - \varepsilon|S|n$ which combined with $e(G) \geq (1 - \varepsilon^2)n^2/2$ gives $|S| \leq \varepsilon n$. Then $H = G \setminus S$ is a graph with $\delta(H) \geq (1 - \varepsilon)n - |S| \geq (1 - 2\varepsilon)n$. \square

Our argument uses the following theorem.

Theorem 7.2 (Gyárfás–Sárközy [20]). *Every properly edge-coloured graph G on n vertices with $\delta(G) \leq n/2$ has a rainbow matching of size $\delta(G) - 2\delta(G)^{2/3}$.*

Corollary 7.3. *For $\varepsilon > 0$ and large enough n , every properly edge-coloured n -vertex graph G with $\delta(G) \geq n/2$ has a rainbow matching of size $(1 - \varepsilon)n/2$.*

Proof. Delete edges from G to get a spanning subgraph H with $\delta(H) = n/2$, and apply Theorem 7.2 to H . We get a rainbow matching of size $n/2 - 2(n/2)^{2/3} \geq (1 - \varepsilon)n/2$ (for sufficiently large n). \square

Lemma 7.4. *For $\varepsilon > 0$ and large enough n , every properly edge-coloured n -vertex graph G with $e(G) \geq (1 - \varepsilon^2)n^2/2$ has a rainbow matching of size $(1 - 3\varepsilon)n/2$.*

Proof. Apply Lemma 7.1 to get a subgraph H on m vertices with $\delta(H) \geq (1 - 2\varepsilon)n$. Apply Corollary 7.3 to H (using that $\delta(H) \geq m/2$) in order to get a rainbow matching that has size $(1 - \varepsilon)m/2 \geq (1 - \varepsilon)\delta(H)/2 \geq (1 - \varepsilon)(1 - 2\varepsilon)n/2 \geq (1 - 3\varepsilon)n/2$. \square

The following lemma is a bipartite version of the theorem we are aiming for.

Lemma 7.5. *For any $\varepsilon > 0$ and sufficiently large r , there is an r -edge-coloured bipartite graph G with parts X and Y such that*

- (a) $\deg(x) \geq (1 - 3\varepsilon)|Y|$ for all $x \in X$, and
- (b) X cannot be covered by fewer than $\varepsilon^2 r^2/4$ monochromatic components in G .

Proof. Let Y be a set of size r . Let K_Y be an auxiliary properly r -edge-coloured complete graph on vertex set Y . Let X be the set of rainbow matchings in K_Y of size $(1 - 3\varepsilon)r/2$.

The graph G is defined as follows: for any $x \in X$ and $y \in Y$, we add the edge xy to G in colour i if the rainbow matching x of K_Y contains a colour- i edge through y . If the rainbow matching x does not contain any edge through y , then xy is not present in G .

To see that (a) holds, notice that every $x \in X$ is connected to all the vertices of Y that appear in the rainbow matching x of K_Y . Since the rainbow matching x has size $(1 - 3\varepsilon)r/2$, we get $\deg(x) \geq (1 - 3\varepsilon)|Y|$.

Let uv be an i -coloured edge of K_Y , and let $X_{uv} \subseteq X$ be the set of rainbow matchings containing uv . We claim that $T_{uv} = \{u, v\} \cup X_{uv}$ is an i -coloured component of G . Indeed, u and v are only adjacent to X_{uv} in colour i because K_Y is properly coloured. Also, the matchings in X_{uv} contain no i -coloured edges other than uv because they are rainbow. This shows that every monochromatic component of G is either of the form T_{uv} or is a singleton.

Let T_1, \dots, T_k be any family of $k \leq \varepsilon^2 r^2/4$ monochromatic components in G . We will find a vertex in X that is not covered by these monochromatic components. Using the previous paragraph, we may assume that the component T_i has form $T_{u_i v_i}$ for some edge $u_i v_i \in K_Y$. Consider $H = K_Y \setminus \{u_1 v_1, \dots, u_k v_k\}$. Then $e(H) \geq (1 - \varepsilon^2)r^2/2$, so by Lemma 7.4, H has a rainbow matching M of size $(1 - 3\varepsilon)r/2$. This M thus corresponds to a vertex $x_M \in X$. However, as M does not contain the edge $u_i v_i$ for any i , the vertex x_M does not belong to any T_i . In other words, T_1, \dots, T_k do not cover the vertex x_M , establishing (b). \square

We are now ready to prove the main result of this section.

Proof of Proposition 1.3. Let us apply Lemma 7.5 with $r - 1$ colours to obtain an $(r - 1)$ -edge-coloured bipartite graph with parts X and Y satisfying (a) and (b). To construct G from H , we blow up all vertices in Y by $n' = n'(\varepsilon, r)$ vertices, and add all the edges inside Y in some previously unused colour. Formally, we introduce a set of vertices V_y of size $n' = |X|/\varepsilon$ for every $y \in Y$, and set $Y' = \bigcup_{y \in Y} V_y$. The vertices of G are $V(G) = X \cup Y'$. For any $x \in X$ and $v \in V_y$, the edge xv is present in G in colour i precisely when the edge xy is present in H in colour i . For $u, v \in Y'$, the edge uv is present in G in colour r .

Let n denote the number of vertices in G . Then $n = |X|(1 + (r - 1)/\varepsilon)$, so in particular, $|X| \leq \varepsilon n - 1$. This together with the definition of G implies that for every vertex $y \in Y'$, we have $\deg_G(y) \geq |Y'| - 1 \geq (1 - \varepsilon)n$. Using (a) we also get that for every $x \in X$,

$$\deg_G(x) = \deg_H(x)n' \geq (1 - 3\varepsilon)|Y|n' = (1 - 3\varepsilon)|Y'| \geq (1 - 3\varepsilon)(1 - \varepsilon)n \geq (1 - 4\varepsilon)n.$$

As G was constructed from H by blowing up Y and then adding some edges in it using a new colour, we see that every monochromatic component of G touching X is of the form $\{x \in T \cap X\} \cup \{v \in V_y : y \in T \cap Y\}$ for some monochromatic component T of H . But then any covering of G by fewer than $\varepsilon^2(r - 1)^2/4$ monochromatic components would give a covering of X in H by fewer than $\varepsilon^2(r - 1)^2/4$ monochromatic components, contradicting (b). \square

Acknowledgements

We would like to thank the anonymous referees for their helpful comments and Louis DeBiasio for stimulating discussions.

References

- [1] P. Allen, J. Böttcher, R. Lang, J. Skokan, and M. Stein, *Partitioning a 2-edge-coloured graph with minimum degree $2n/3 + o(n)$ into three monochromatic cycles*, in preparation (2020).
- [2] J. Balogh, J. Barát, D. Gerbner, A. Gyárfás, and G. N. Sárközy, *Partitioning 2-edge-colored graphs by monochromatic paths and cycles*, *Combinatorica* **34** (2014), no. 5, 507–526.
- [3] S. Bessy and S. Thomassé, *Partitioning a graph into a cycle and an anticycle: a proof of Lehel’s conjecture*, *J. Combin. Theory Ser. B* **100** (2010), 176–180.

- [4] J. A. Bondy and M. Simonovits, *Cycles of even length in graphs*, J. Combin. Theory, Ser. B **16** (1974), no. 2, 97–105.
- [5] J. Böttcher, M. Schacht, and A. Taraz, *Proof of the bandwidth conjecture of Bollobás and Komlós*, Mathematische Annalen **343** (2009), no. 1, 175–205.
- [6] M. Bucić, D. Korándi, and B. Sudakov, *Covering graphs by monochromatic trees and Helly-type results for hypergraphs*, Combinatorica, to appear.
- [7] S. Bustamante, J. Corsten, N. Frankl, A. Pokrovskiy, and J. Skokan, *Partitioning edge-colored hypergraphs into few monochromatic tight cycles*, SIAM J. Disc. Math. **34** (2020), no. 2, 1460–1471.
- [8] S. Bustamante, H. Hàn, and M. Stein, *Almost partitioning 2-colored complete 3-uniform hypergraphs into two monochromatic tight or loose cycles*, J. of Graph Theory **91** (2019), no. 1, 5–15.
- [9] D. Conlon and M. Stein, *Monochromatic cycle partitions in local edge colorings*, J. Graph Theory **81** (2016), no. 2, 134–145.
- [10] B. Csaba, D. Kühn, A. Lo, D. Osthus, and A. Treglown, *Proof of the 1-factorization and Hamilton decomposition conjectures*, vol. 244, American Mathematical Society, 2016.
- [11] L. DeBiasio and L. Nelsen, *Monochromatic cycle partitions of graphs with large minimum degree*, J. Combin. Theory Ser. B **122** (2017), 634–667.
- [12] M. Elekes, D. T. Soukup, L. Soukup, and Z. Szentmiklóssy, *Decompositions of edge-colored infinite complete graphs into monochromatic paths*, Discrete Math. **340** (2017), no. 8, 2053–2069.
- [13] P. Erdős, A. Gyárfás, and L. Pyber, *Vertex coverings by monochromatic cycles and trees*, J. Combin. Theory Ser. B **51** (1991), 90–95.
- [14] F. Garbe, R. Lang, A. Lo, R. Mycroft, and N. Sanhueza-Matamala, *Partitioning a 3-uniform complete hypergraph into a red and a blue tight cycle*, in preparation (2020).
- [15] A. Grinshpun and G. N. Sárközy, *Monochromatic bounded degree subgraph partitions*, Discrete Math. **339** (2016), no. 1, 46–53.
- [16] A. Gyárfás, *Vertex coverings by monochromatic paths and cycles*, J. Graph Theory **7** (1983), 131–135.
- [17] ———, *Vertex covers by monochromatic pieces – a survey of results and problems*, Discrete Math. **339** (2016), no. 7, 1970–1977.
- [18] A. Gyárfás, M. Ruszinkó, G. N. Sárközy, and E. Szemerédi, *An improved bound for the monochromatic cycle partition number*, J. Combin. Theory Ser. B **96** (2006), 855–873.
- [19] A. Gyárfás and G. N. Sárközy, *Monochromatic path and cycle partitions in hypergraphs*, Electron. J. Combin. **21** (2012), no. 2, P2.36.
- [20] A. Gyárfás and G. N. Sárközy, *Rainbow matchings and cycle-free partial transversals of Latin squares*, Discrete Math. **327** (2014), 96–102.
- [21] P. E. Haxell, *Partitioning complete bipartite graphs by monochromatic cycles*, J. Combin. Theory Ser. B **69** (1997), 210–218.
- [22] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*, John Wiley & Sons, 2000.
- [23] J. Komlós, G. Sárközy, and E. Szemerédi, *Blow-Up Lemma*, Combinatorica **17** (1997), no. 1, 109–123.

- [24] J. Komlós, G. N. Sárközy, and E. Szemerédi, *On the Pósa-Seymour conjecture*, J. Graph Theory **29** (1998), no. 3, 167–176.
- [25] J. Komlós and M. Simonovits, *Szemerédi’s Regularity Lemma and its applications in graph theory*, Combinatorics, Paul Erdős is Eighty (D. Miklós, V. T. Sós, and T. Szőnyi, eds.), vol. 2, Bolyai Society Mathematical Studies, 1996, pp. 295–352.
- [26] D. Korándi, F. Mousset, R. Nenadov, N. Škorić, and B. Sudakov, *Monochromatic cycle covers in random graphs*, Random Structures & Algorithms **53** (2018), no. 4, 667–691.
- [27] D. Kühn and D. Osthus, *Hamilton cycles in graphs and hypergraphs: an extremal perspective*, Proc. ICM 2014 (Seoul, Korea), vol. 4, 2014, pp. 381–406.
- [28] R. Lang and A. Lo, *Monochromatic cycle partitions in random graphs*, arXiv:1807.06607 (2018).
- [29] S. Letzter, *Monochromatic cycle partitions of 3-coloured complete graphs*, in preparation (2018).
- [30] ———, *Monochromatic cycle partitions of 2-coloured graphs with minimum degree $3n/4$* , Electron. J. Combin. **26** (2019), P1.19.
- [31] L. Lovász, *Combinatorial problems and exercises*, AMS/Chelsea publication, North-Holland Publishing Company, 1979.
- [32] T. Łuczak, $R(C_n, C_n, C_n) \leq (4 + o(1))n$, J. Combin. Theory Ser. B **75** (1999), 174–187.
- [33] T. Łuczak, V. Rödl, and E. Szemerédi, *Partitioning two-colored complete graphs into two monochromatic cycles*, Combin. Probab. Comput. **7** (1998), 423–436.
- [34] A. Pokrovskiy, *Partitioning edge-coloured complete graphs into monochromatic cycles and paths*, J. Combin. Theory Ser. B **106** (2014), 70–97.
- [35] ———, *Partitioning a graph into a cycle and a sparse graph*, Submitted (2016).
- [36] R. Rado, *Monochromatic paths in graphs*, Ann. Discrete Math **3** (1978), 191–194.
- [37] G. N. Sárközy, *Monochromatic cycle partitions of edge-colored graphs*, J. Graph Theory **66** (2011), 57–64.
- [38] G. N. Sárközy and S. Selkow, *Vertex partitions by connected monochromatic k -regular graphs*, J. Combin. Theory, Ser. B **78** (2000), 115–122.
- [39] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, vol. 24, Springer Science & Business Media, 2003.
- [40] E. Szemerédi, *Regular partitions of graphs*, Colloq. Internat. CNRS **260** (1976), 399–401.

A Proofs of preliminary results

Proof of Lemma 3.1

Proof. The following claim is the main ingredient of the proof.

Claim A.1. *Let $W_1 \subseteq V_1$ and $W_2 \subseteq V_2$ satisfy $|W_1|, |W_2| \geq n/10$ and $\delta(G[W_1, W_2]) \geq 3\epsilon n$. Then $G[W_1, W_2]$ has a cycle of length $2 \min\{|W_1|, |W_2|\}$.*

Proof. Let $P = u_1 \dots u_{2\ell}$ be a longest path of even order in $H := G[W_1, W_2]$, where $u_1 \in W_1$ and $u_{2\ell} \in W_2$. Note that $\ell \geq 2\epsilon m$ by the minimum degree assumption.

We claim that H has a cycle of length 2ℓ . To see this, we will show that one of the following holds:

- u_1u_i is an edge and $u_{2\ell}, u_{i-2}$ have a common neighbour $u \notin V(P)$, for some $i \in [2\ell]$,
- $u_iu_{2\ell}$ is an edge and u_1, u_{i+2} have a common neighbour $u \notin V(P)$, for some $i \in [2\ell]$,
- $u_1u_i, u_ju_{2\ell}$ and $u_{i-1}u_{j+1}$ are edges, for some $1 \leq i < j \leq 2\ell$,
- $u_1u_j, u_iu_{2\ell}$ and $u_{i+1}u_{j+1}$ are edges, for some $1 \leq i < j \leq 2\ell$.

Note that in each of these cases, a cycle of length 2ℓ exists. Indeed, here are the appropriate cycles: $(u_1 \dots u_{i-2}uu_{2\ell} \dots u_i)$, $(u_{2\ell} \dots u_{i+2}uu_1 \dots u_i)$, $(u_1 \dots u_{i-1}u_{j+1} \dots u_{2\ell}u_j \dots u_i)$, and $(u_1 \dots u_iu_{2\ell} \dots u_{j+1}u_{i+1} \dots u_j)$.

For a set $S \subseteq V(P)$, define $S^- = \{u_{i-1} : u_i \in S\}$ and $S^+ = \{u_{i+1} : u_i \in S\}$. We now prove that one of the above four conditions holds.

Note that either u_1 does not have neighbours in $W_2 \setminus V(P)$ or $u_{2\ell}$ does not have neighbours in $W_1 \setminus V(P)$; without loss of generality, the latter holds.

Let X be the set of neighbours of u_1 that are not in P , let Y be the set of neighbours of u_1 in P , and let Z be the set of neighbours of $u_{2\ell}$. If $|X| \geq \varepsilon m$, then by regularity there is an edge between X and Z^{++} , as required for the second item above. We now assume that $|X| \leq \varepsilon n$, so $|Y| \geq 2\varepsilon n$. Let $\alpha \in [2\ell]$ be such that the sets $Y_1 = Y \cap \{u_1, \dots, u_\alpha\}$ and $Y_2 = Y \cap \{u_\alpha, \dots, u_{2\ell}\}$ each have size at least εn . Write $Z_1 = Z \cap \{u_1, \dots, u_{\alpha-1}\}$ and $Z_2 = Z \cap \{u_{\alpha+1}, \dots, u_{2\ell}\}$. If $|Z_1| \geq \varepsilon n$, then by regularity there is an edge between Z_1^- and Y_2^+ , as required for the third item. Otherwise, $|Z_2| \geq \varepsilon n$, and, again by regularity, there is an edge between Z_2^+ and Y_1^+ , as required for the fourth item.

We have established that there is a cycle C of length 2ℓ ; abusing notation slightly, write $C = (u_1 \dots u_{2\ell})$. We now show that $\ell = \min\{|W_1|, |W_2|\}$. Suppose otherwise.

First, suppose that there is a vertex $v \notin V(C)$ with at least εn neighbours outside of C ; let A be the set of these neighbours. By regularity, there is an edge between A and C . It follows that there is a path of length $2\ell + 2$, a contradiction to the choice of ℓ . Next, we may assume that every vertex not in C has at least $2\varepsilon n$ neighbours in C . Let $v \in W_1 \setminus V(C)$ and $w \in W_2 \setminus V(C)$. Let A and B be the neighbourhoods of v and w in C , respectively. By regularity, there is an edge between A^- and B^- ; let i, j be such that $u_i \in A, u_j \in B$, and $u_{i-1}u_{j-1}$ is an edge. Then $vu_i \dots u_{j-1}u_{i-1} \dots u_jw$ is a path of length $2\ell + 2$, a contradiction to the choice of ℓ . \square

Let us now see how to deduce Lemma 3.1 from Claim A.1. Assume, without loss of generality, that $|U_1| \leq |U_2|$, and write $m = |U_1|$.

Consider the first part. For $i \in \{1, 2\}$, let $A_i \subseteq U_i$ be a set of $\lceil \varepsilon n \rceil$ neighbours of v_{3-i} , and let W_i be the set of vertices $u \in U_i \setminus A_i$ that have a neighbour in A_{3-i} and have degree at least $7\varepsilon n$ in $G[U_1, U_2]$. By regularity, $|W_i| \geq |U_i| - |A_i| - 2\varepsilon n \geq |U_i| - 4\varepsilon n \geq (1 - 4\varepsilon)m$. It follows that $G[W_1, W_2]$ has minimum degree at least $3\varepsilon n$. Thus, by Claim A.1, there is a path $u_1 \dots u_{2\ell}$ in $G[W_1, W_2]$, for any ℓ with $1 \leq \ell \leq \min\{|W_1|, |W_2|\}$; without loss of generality

$u_1 \in W_1$ and $u_{2\ell} \in W_2$. By choice of W_1, W_2 , there are vertices $x_1 \in A_1$ and $x_2 \in A_2$ such that $v_1 x_2 u_1 \dots u_{2\ell} x_1 v_2$ is a path. It follows that there is a v_1 - v_2 -path of order $2k$ for every $3 \leq k \leq (1 - 24\varepsilon)m$. To complete the proof of the first part, observe that, by regularity, there is an edge between A_1 and A_2 , implying that there is a v_1 - v_2 -path of order 4.

Now we prove the second part. Note that it suffices to prove the existence of a v_1 - v_2 -path of order $2(k+1)$ with $(1 - 24\varepsilon)m \leq k \leq m$. Let U'_1 be a subset of U_1 of size k , chosen uniformly at random among such sets. Then, with high probability, $G[U'_1, U_2]$ has minimum degree at least $3\varepsilon n$, and U'_1 contains at least εn neighbours of v_2 . It follows from Claim A.1 that there is a cycle $C = (u_1 \dots u_k)$ in $G[U'_1, U_2]$. Let X be the set of neighbours of v_1 in C , let Y be the set of neighbours of v_1 outside of C , and let Z be the set of neighbours of v_2 in C . By regularity, there is an edge between Z^+ and X^+ or between Z^{++} and Y (where X^+ , Z^+ and Z^{++} are defined as in the proof of Claim A.1). One can easily check that in either case, a v_1 - v_2 -path of order $2(k+1)$ exists. \square

Proof of Proposition 3.3

Proof.

- (a) Note that $\deg_{G'}(v) \geq (c - rd - \varepsilon)n$, so v is adjacent to vertices from at least $(c - rd - \varepsilon)n/|V_i| \geq (c - rd - \varepsilon)m$ clusters in G' . By the definition of G' , this means that x_i is adjacent to the corresponding vertices in \mathcal{G} .
- (b) We may assume $\eta + \varepsilon < 1$, as otherwise there is nothing to prove. Let X be the set of vertices x in \mathcal{G} with $\deg_G(x) < (c - rd - \varepsilon)m$. By (a), we know that every vertex v in the clusters corresponding to X must have $\deg_G(v) < cn$. So if $|X| > (\eta + \varepsilon)m$, then at least $|X||V_i| > (\eta + \varepsilon)(1 - \varepsilon)n = (\eta + \varepsilon(1 - \eta - \varepsilon))n > \eta n$ vertices $v \in V(G)$ have $\deg_G(v) < cn$, contradicting our assumption.
- (c) As G' is obtained by deleting at most $(rd + \varepsilon)n^2$ edges from $G \setminus V_0$, we know that $\bigcup_{x_i \in X} V_i$ induces at least $(c - rd - \varepsilon)n^2$ edges in G' . But then there are at least $(c - rd - \varepsilon)n^2/|V_i|^2 \geq (c - rd - \varepsilon)m^2$ pairs $\{x_i, x_j\}$ in X with an edge between the corresponding clusters V_i and V_j in G' . By the definition of G' , this means that X induces at least $(c - rd - \varepsilon)m^2$ edges in \mathcal{G} . \square

Proof of Lemma 3.4

Proof. Let $i_1, \dots, i_\ell \in [m]$ be a sequence such that $i_1 = i$, $i_2 = j$, $i_{\ell-1} = i'$ and $i_\ell = j'$; $x_{i_s} x_{i_{s+1}}$ is an edge of colour c in \mathcal{G} for $s \in [\ell - 1]$; and $5 \leq \ell \leq m + 2$ (it is easy to see that such a sequence exists).

We claim that there exists a path $v_1 \dots v_\ell$ such that $v_s \in V_{i_s} \setminus W$ for $s \in [\ell]$; $v_1 = v$ and $v_\ell = w$; and v_s has typical degree in $V_{i_{s+1}}$ for $s \in [\ell - 3]$. Indeed, suppose that $v_1 \dots v_{s-1}$ is a path with the required properties, for $2 \leq s \leq \ell - 2$. If $s \leq \ell - 3$, as v_{s-1} has typical degree in V_{i_s} , there is a neighbour $v_s \in V_{i_s} \setminus (W \cup \{v_1, \dots, v_{s-1}, v_\ell\})$ of v_{s-1} with typical degree

in $V_{i_{s+1}}$, using (3.1). If $s = \ell - 2$, as $v_{\ell-3}$ and v_ℓ have typical degrees in $V_{i_{\ell-2}}$ and $V_{i_{\ell-1}}$, respectively, and by regularity, there are vertices $v_{\ell-2}$ and $v_{\ell-1}$ in $V_{i_{\ell-2}} \setminus (W \cup \{v_1, \dots, v_{\ell-3}, v_\ell\})$ and $V_{i_{\ell-1}} \setminus (W \cup \{v_1, \dots, v_{\ell-3}, v_\ell\})$, respectively, such that $v_{\ell-3}v_{\ell-2}v_{\ell-1}v_\ell$ is a path. The existence of the required path follows, proving Lemma 3.4. \square

Proof of Lemma 3.12

Proof. An important ingredient is the following classic result of Pósa (see [31]).

Theorem A.2 (Pósa). *The vertices of any graph G can be covered with at most $\alpha(G)$ vertex-disjoint cycles.*

For $x \in B$ let $N_i(x)$ denote the set of vertices in A adjacent to x in colour i . Let $\{B_1, B_2, \dots, B_r\}$ be a partition of B such that for any $1 \leq i \leq r$ and $x \in B_i$, we have $|N_i(x)| \geq |A|/(100r)$ (clearly, such a partition exists). Define a graph G_i on vertex set B_i for every $i \in [r]$ as follows. For $x, y \in B_i$, let xy be an edge of G_i if and only if $|N_i(x) \cap N_i(y)| \geq |A|/(100^3 r^3)$. We claim that the independence number of G_i is at most $100r$ for $i \in [r]$.

Indeed, suppose otherwise, and let $x_1, \dots, x_{100r+1} \in B_i$ be pairwise non-adjacent. Then

$$\begin{aligned} |A| &\geq \left| \bigcup_{1 \leq j \leq 100r+1} N_i(x_j) \right| \\ &\geq \sum_{1 \leq j \leq 100r+1} |N_i(x_j)| - \sum_{1 \leq j < k \leq 100r+1} |N_i(x_j) \cap N_i(x_k)| \\ &\geq |A| \left(\frac{100r+1}{100r} - \frac{\binom{100r+1}{2}}{100^3 r^3} \right) > |A|, \end{aligned}$$

a contradiction.

So $\alpha(G_i) \leq 100r$, and thus by Theorem A.2 G_i can be partitioned into a family \mathcal{C}_i of at most $100r$ vertex-disjoint cycles. Using the definition of G_i , we can then greedily replace the edges xy in each \mathcal{C}_i with i -coloured paths xwy (where $w \in N_i(x) \cap N_i(y) \subseteq A$), where each edge uses a distinct vertex w , to find at most $100r^2$ monochromatic vertex-disjoint proper cycles and edges in G that cover B . Additionally, for every singleton $\{x\}$ in \mathcal{C}_i we replace it by an edge xw (where $w \in N_i(x)$), so that each singleton uses a distinct vertex w , which is also distinct from the vertices chosen previously. \square

Proof of Proposition 3.14

Proof. Let A be a random subset of $V \setminus B$ where every vertex is included in A independently with probability p . We will show that the event that A satisfies all of the properties has positive probability.

- (a) Note that $|A| \sim \text{Bin}(|V \setminus B|, p)$, and by assumption, $|B| < 10pn + \varepsilon n < n/3$. Hence, by

Lemma 3.13,

$$\Pr \left[|A| < \frac{p}{2}n \right] \leq \Pr \left[|A| < \frac{3}{4}|V \setminus B|p \right] < e^{-|V \setminus B|p/32} \leq e^{-np/48}.$$

- (b) Again, for any $i \in [m]$ fixed, $|A \cap V_i| \sim \text{Bin}(V_i \setminus B, p)$ so by Lemma 3.13 and $|V_i \setminus B| \geq (9/10)|V_i| > n/(2m)$,

$$\Pr [|A \cap V_i| > 2p|V_i|] < \Pr [|A \cap V_i| > 2p|V_i \setminus B|] < e^{-|V_i \setminus B|p/3} < e^{-np/(10m)}.$$

- (c) Let $v \in V$ and $i \in [m]$ be such that $\deg(v, V_i) > 30p|V_i|$. Here $\deg(v, A \cap V_i) \sim \text{Bin}(\deg(v, V_i \setminus B), p)$, and $\deg(v, V_i \setminus B) \geq \deg(v, V_i) - |B \cap V_i| > (2/3)\deg(v, V_i)$. We can then apply Lemma 3.13 to get

$$\Pr \left[\deg(v, A \cap V_i) < \frac{p}{2} \deg(v, V_i) \right] < e^{-\deg(v, V_i)p/48} < e^{-|V_i|p^2/2} < e^{-np^2/(4m)}.$$

- (d) Let $v \in V$ with $\deg(v, V \setminus B) > n/40$. Like in the previous argument, $\deg(v, A) \sim \text{Bin}(\deg(v, V \setminus B), p)$, so by Lemma 3.13 and as $\deg(v, V \setminus B) > n/40$,

$$\Pr \left[\deg(v, A) < \frac{pn}{50} \right] \leq \Pr \left[\deg(v, A) < \frac{4p}{5} \deg(v, V \setminus B) \right] < e^{-\deg(v, V \setminus B)p/50} \leq e^{-np/2000}.$$

But the argument for Property (b) shows that $\Pr [|A| > 2pn] < me^{-np/(10m)}$, so

$$\Pr \left[\deg(v, A) < \frac{|A|}{100} \right] < e^{-np/2000} + me^{-np/(10m)}.$$

Now taking a union bound over all choices of $i \in [m]$ in (b) and (c) and all choices of $v \in V$ in (c) and (d), we get that A satisfies all properties with probability at least $1 - n^2e^{-np^2/(4m)} > 0$. (Here we use that $p > m \log n / \sqrt{n}$ and in particular $m \ll n$). In particular, there is such a set A . \square