Cycles in Edge-Coloured Graphs and Subgraphs of Random Graphs

Matthew White
The Queen's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2011

This thesis is dedicated to my parents for all their love and support.

Acknowledgements

I would like to express my gratitude to my supervisor Professor Alex Scott for his encouragement, help and support throughout the past four years. My thanks also go to Professor Colin McDiarmid and Professor Oliver Riordan for their help during the process of Transfer and Confirmation of Status.

This thesis would not have been possible without the financial support of the Engineering and Physical Sciences Research Council, and the emotional support of Emily and my parents. For this I thank them and the many others who have supported me in any respect during the completion of this thesis.

Statement of Authorship

The material in this thesis has not previously been submitted for a degree at this University, or elsewhere. The work of Chapter 2 and Chapter 5 is joint work with Professor Alex Scott. All other work in this thesis to the best of my knowledge contains no material previously published or written by another person except where due acknowledgement is made.

Cycles in Edge-Coloured Graphs and Subgraphs of Random Graphs

Matthew White

The Queen's College
University of Oxford

A thesis submitted for the degree of D.Phil

Trinity 2011

Abstract

This thesis will study a variety of problems in graph theory. Initially, the focus will be on finding minimal degree conditions which guarantee the existence of various subgraphs. These subgraphs will all be formed of cycles, and this area of work will fall broadly into two main categories.

First to be considered are cycles in edge-coloured graphs and, in particular, two questions of Li, Nikiforov and Schelp. It will be shown that a 2-edge-coloured graph with minimal degree at least $\frac{3}{4}n$ either is isomorphic to the complete 4-partite graph with classes of order $\frac{1}{4}n$, or contains monochromatic cycles of all lengths $\ell \in \left[4, \left\lceil \frac{1}{2}n \right\rceil\right]$. This answers a conjecture of Li, Nikiforov and Schelp. Attention will then turn to the length of the longest monochromatic cycle in a 2-edge-coloured graph with minimal degree at least cn. In particular, a lower bound for this quantity will be proved which is asymptotically best possible.

The next chapter of the thesis then shows that a hamiltonian graph with minimal degree at least $\frac{5-\sqrt{7}}{6}n$ contains a 2-factor with two components.

The thesis then concludes with a chapter about X_H, which is the number of copies of a graph H in the random graph $\mathcal{G}(n,p)$. In particular, it will be shown that, for a connected graph H, the value of X_H modulo k is approximately uniformly distributed, provided that k is not too large a function of n.

Contents

Chapter 1

Introduction

1.1 Subgraphs of graphs.

This thesis will be concerned with graphs with given properties, and so begins by making some elementary definitions concerning graphs, before introducing the problems to be considered.

Throughout, some notation describing the asymptotic behaviour of functions is required. Note that, unless stated otherwise, all integers in this thesis are assumed to be positive.

Definition 1.1. *Let $f(n)$ and $g(n)$ be two positive real-valued functions defined on the natural numbers.*

- *$f(n)$ lies in the class $O(g(n))$ if there are constants C and N such that $f(n) \leq Cg(n)$ for all natural numbers $n > N$.*

- *$f(n)$ is in the class $\Omega(g(n))$ if there exists c and N such that $f(n) \geq cg(n)$ for all $n > N$.*

- *$f(n)$ lies in the class $\Theta(g(n))$ if $f(n) \in O(g(n))$ and $f(n) \in \Omega(g(n))$.*

- *$f(n)$ is in the class $o(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.*

If f is a function of several variables $n_1, \ldots, n_k$, then f lies in the class $O_{n_2,\ldots,n_k}(g(n_1))$ if there are constants C and N (which will generally depend on $n_2, \ldots n_k$) such that $f(n_1, \ldots, n_k) \leq Cg(n_1)$ for all natural numbers $n_1 > N$. This notation may be extended to all of the other classes defined above.

Note the definition can be extended to functions defined on the real numbers. For simplicity, $f(n) = o(g(n))$ will often be used to mean $f(n) \in o(g(n))$, with the other notation also used in a similar way. Indeed $o(g(n))$ will frequently be treated as any other function for algebraic manipulation, so that $f(n) = m(n)(h(n) + o(g(n)))$ means that $\frac{f(n)}{m(n)} - h(n) \in o(g(n))$. In practice, such expressions can be justified using the original definitions.

The notation for graphs will generally follow that of [5]. A *simple graph* (or just *graph*) is defined to be a pair of sets (V, E) where $E \subseteq V^{(2)} = \{e \subseteq V : |e| = 2\}$. The set V is the *vertex set* of G and E is the *edge set*, and elements of these sets are referred to as *vertices* and *edges* respectively. This terminology is used as graphs are often drawn as a picture, with points representing the vertices, and lines between the points representing edges, as in Figure 1.1.

To avoid confusion when dealing with more than one graph, $V(G)$ and $E(G)$ refer to the vertex and edge sets of G respectively. The *order* of G is the number of vertices of G, which is often written as $|G|$, and the *size* of G is the number of edges of G, which is often written as $e(G)$. Two graphs G and G' are said to be *isomorphic* if there is a bijection θ from $V(G)$ to $V(G')$

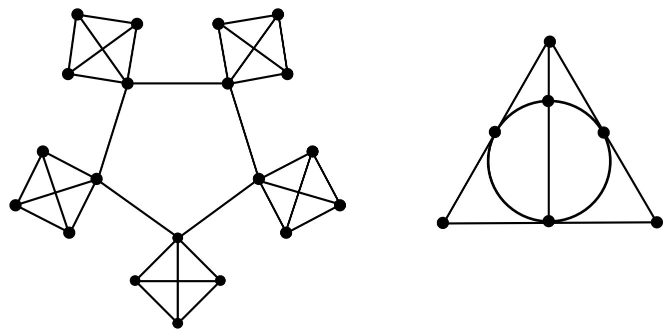


Figure 1.1: Two graphs

such that $uv \in E(G)$ if, and only if, $\theta(u)\theta(v) \in E(G')$. In this case θ is an *isomorphism* from G to G'.

A graph H is a *subgraph* of a graph G if $H = (W, F)$ where $W \subseteq V$ and $F \subseteq E \cap W^{(2)}$. That H is a subgraph of G is denoted by $H \subseteq G$. The graph G *contains a copy of* H when G has a subgraph which is isomorphic to H. A subgraph H of G is *spanning* if $V(H) = V(G)$, while H is *induced* if $E(H) = E(G) \cap V(H)^{(2)}$. The unique induced subgraph of G with vertex set W is denoted by $G[W]$.

The *complete graph on a given vertex set* V is the graph whose edge set is precisely $V^{(2)}$. All complete graphs of the same order are isomorphic, and the complete graph with n vertices is denoted by K_n.

A *path* is a graph with vertex set $\{v_1, \ldots, v_{\ell+1}\}$ whose edges are $v_i v_{i+1}$ for $1 \leq i \leq \ell$, and is denoted by $v_1 \ldots v_{\ell+1}$. Similarly, a *cycle* has vertex set $\{v_1, \ldots, v_\ell\}$ and edge set $\{v_1v_2, \ldots, v_\ell v_1\}$ and is denoted by $v_1 \ldots v_\ell v_1$. For both a path and a cycle, the *length* is the number of edges. Note that, in a cycle, the number of edges is the same as the number of vertices, while in a path, there is one less edge then there are vertices. An arbitrary path of length ℓ is denoted P_ℓ, while an arbitrary cycle of length ℓ is denoted by C_ℓ. The graphs C_4 and P_4 are illustrated in Figure 1.2. Note that, by the definition of a graph, all cycles have length at least three. For a graph G, a *Hamilton* cycle or path of G is a spanning subgraph of G which is respectively a cycle or path. A graph G is *hamiltonian* if it contains a Hamilton cycle.

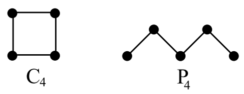


Figure 1.2: C_4 and P_4

Let u and v be vertices of a graph G. Then G *contains a path from u to v* if there is some subgraph of G which is a path $v_1 \dots v_r$ with $u = v_1$ and $v = v_r$. Define an equivalence relation on G by setting $v \sim u$ either when $v = u$ or when there is some path in G from v to u. The equivalence classes thus defined are referred to as *connected components* or just *components* and a graph G containing only one component is said to be *connected.*

For a vertex v of a graph G, the *neighbourhood* of v is the set of vertices u for which uv is an edge of G. This is denoted by $\Gamma(v)$. The *degree* of v, denoted $\deg(v)$, is the size of the neighbourhood of v. When it may be unclear which graph is referred to, $\Gamma_G(v)$ and $\deg_G(v)$ respectively indicate the neighbourhood and degree of v in graph G. If X is a subset of $V(G)$, then $\deg_X(v)$ is the number of neighbours of v in X.

Let G be a graph, with vertices $v_1, \dots, v_n$. The *degree sequence* of G is the sequence $d_1, \dots, d_n$, where $d_i = \deg(v_i)$ for all $1 \leq i \leq n$. It is usual to arrange the vertices so that $d_1 \leq d_2 \leq \cdots \leq d_n$. In this case, the *maximal* *and minimal degree of* G are defined to be the values d_n and d_1 respectively, and are denoted $\Delta(G)$ and $\delta(G)$.

This thesis will consider problems concerning the subgraphs of a graph. Such problems often ask for some condition which forces a graph to contain a particular subgraph. Often this condition will be a condition on the number of edges or the minimal degree of G. The subgraph considered may be some fixed 'small' graph H, as in the following result of Bondy and Simonovits [9], in which H is a cycle of fixed even length.

Theorem 1.2 (Bondy and Simonovits [9])**.** *Let* G *be a graph of order* n *and let* k *be an integer. If* $e(G) > 100kn^{1+1/k}$*, then* G *contains a cycle of length* $2k$*.*

Alternatively, the subgraph under consideration may be 'large' compared to the base graph. For example, a theorem of Dirac [16] states that a graph with order $n \geq 3$ and minimal degree at least $\frac{1}{2}n$ contains a Hamilton cycle.

Theorem 1.3 (Dirac [16])**.** *Let* G *be a graph of order* $n \geq 3$*. If* $\delta(G) \geq \frac{1}{2}n$*, then* G *is hamiltonian.*

The first half of this thesis will be concerned mainly with finding 'large' subgraphs. In particular, some generalisations of Dirac's Theorem will be considered. These problems will be introduced in Sections 1.2 and 1.3. The second half of the thesis will be concerned with small subgraphs. In fact, in this case, subgraphs of a random graph will be considered, and this concept will be introduced in Section 1.4.

1.2 Introducing Ramsey-Turán theory

Given a graph G with edge set $E(G)$, an *r-edge-colouring* of G is a partition $E(G) = \bigcup_{i=1}^{r} E(G_i)$ of the edges of G, where each G_i is a spanning subgraph of G. If $r = 2$, the partition is $E(G) = E(R) \cup E(B)$, where R is regarded as the red graph and B as the blue graph. A subgraph of G is said to be *monochromatic* if it is contained in one of the subgraphs G_i. For a vertex v in a 2-edge-coloured graph, the red and blue neighbours of v are defined in the obvious way, as are the red and blue degree of v.

The study of monochromatic subgraphs of an edge-coloured graph was begun by Ramsey [42], after whom the entire subject is now named. Given an integer t, define the *Ramsey number* $R(t)$ to be the smallest integer such that, if $n \geq R(t)$, and K_n is 2-edge-coloured, there is a monochromatic copy of K_t. The following result is the starting point of Ramsey theory.

Theorem 1.4 (Ramsey [42])**.** *For all $t \geq 2$, the number $R(t)$ is finite.*

Only for very small values of t is the value of the Ramsey number known. There are many generalisations of this problem. For example, for any set of graphs $H_1, \ldots, H_r$, the *Ramsey number* $R_r(H_1, \ldots, H_r)$ is defined to be the smallest integer such that, if $n \geq R_r(H_1, \ldots, H_r)$, and K_n is r-edge-coloured as $E(K_n) = \bigcup_{i=1}^{r} E(G_i)$, then, for some $1 \leq i \leq r$, there is a copy of H_i in G_i.

In extremal graph theory, study is made of how 'large' a graph may be before it is forced to have some property. In these instances, large normally refers to the number of edges or the minimal degree of the graph. A classical problem of this type is to consider how many edges a graph can have while not containing a copy of K_t. It is clear that the following graph contains no copy of K_t.

Definition 1.5. *Let $n \geq t \geq 2$ be integers. The graph $T_{n,t-1}$ is defined by partitioning n vertices as equally as possible into $t-1$ classes, and taking uv to be an edge if, and only if, u and v are in different classes.*

The following theorem of Turán shows that in fact, this graph has the maximum number of edges without containing K_t.

Theorem 1.6 (Turán [48]). *Let $n \geq t \geq 2$ be integers. If G is a graph with n vertices such that $e(G) \geq e(T_{n,t-1})$, then either $G \cong T_{n,t-1}$, or G contains a copy of K_t.*

For a given property and n, a graph on n vertices is *extremal* if it contains the maximum number of edges without having the property among all graphs on n vertices. Note that Turán's theorem shows that $T_{n,t-1}$ is the only extremal graph for the property of containing K_t.

A development of both Ramsey theory and extremal graph theory is Ramsey–Turán theory (see, for example, the survey paper of Simonovits and Sós [46]). Here a typical question would be the following.

Question 1.7. *Let $L_1, \ldots, L_k$ be fixed graphs. What is the maximum number of edges a k-edge-colored graph can have under the condition that it does not contain an L_i in the ith color, for any $1 \leq i \leq r$?*

The first half of this thesis will focus on a new type of Ramsey–Turán problem, introduced by Li, Nikiforov and Schelp [38], in which the existence of monochromatic cycles in r-edge-coloured graphs is considered. Independently, both Rosta [43] and Faudree and Schelp [24] found the value of the Ramsey number $R_2(C_r, C_s)$ for each r and s.

Theorem 1.8. *The Ramsey number for two cycles has the following values.*

- $R_2(C_3, C_3) = R_2(C_4, C_4) = 6$.
- *If* $3 \leq s \leq t$, s *is odd, and* $(t, s) \neq (3, 3)$, *then*

$$R_2(C_t, C_s) = 2t - 1.$$

- *If* $4 \leq s \leq t$, t *and* s *are both even and* $(t, s) \neq (4, 4)$, *then*

$$R_2(C_t, C_s) = t + \frac{1}{2}s - 1.$$

- *If* $4 \leq s < t$, t *is odd and* s *is even, then*

$$R_2(C_t, C_s) = \max\left\{t + \frac{1}{2}s - 1, 2s - 1\right\}.$$

These values imply that, for $n > 3$, any 2-edge-colouring of K_{2n-1} contains a monochromatic cycle of length ℓ for all $\ell \in [3, n]$. In [41], Nikiforov and Schelp show that, for sufficiently large n, this remains true if K_{2n-1} is replaced by any graph of order $2n - 1$ with large minimal degree. This lead to the following conjecture of Li, Nikiforov and Schelp [38].

Conjecture 1.9. *Let* $n \geq 4$ *and let* G *be a graph of order* n *with* $\delta(G) > \frac{3}{4}n$. *If* $E(G) = E(R) \cup E(B)$ *is a 2-edge-colouring, then* $C_\ell \subseteq R$ *or* $C_\ell \subseteq B$ *for* *all* $\ell \in \left[4, \left\lceil \frac{1}{2}n \right\rceil\right]$.

The conjectured bound on the degree is tight, as there is a 2-edge colouring of $K_{n/4,n/4,n/4,n/4}$ with no monochromatic odd cycles, as illustrated in Figure 1.3.

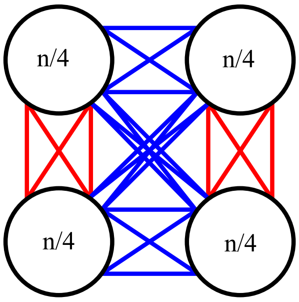


Figure 1.3: A 2-edge-coloured graph with minimal degree $\frac{3}{4}n$ and no monochromatic odd cycles.

Also, note that Conjecture 1.9 is not true for ℓ outside of the stated range. For example, let G be the complete 5-partite graph, with each part having the same number of vertices, so that $\delta(G) = \frac{4}{5}|G|$. Colouring as in Figure 1.4, there is no monochromatic triangle in G.

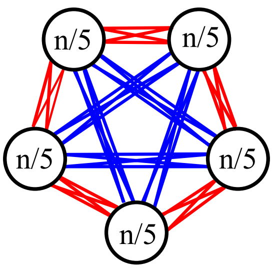


Figure 1.4: A 2-edge-coloured graph with minimal degree $\frac{4}{5}n$ and no monochromatic triangle.

Similarly, let G be the complete graph on n vertices, R be the complete bipartite graph with vertex classes of order $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$, and B have all remaining edges of G. This is a 2-colouring of the complete graph with no monochromatic odd cycle of length $\ell > \lceil \frac{n}{2} \rceil$. In [38], the following partial result was proven.

Theorem 1.10. *Let $\epsilon > 0$, and let G be a graph of sufficiently large order n, with $\delta(G) > \frac{3}{4}n$. If $E(G) = E(R) \cup E(B)$ is a 2-edge-colouring, then $C_\ell \subseteq R$ or $C_\ell \subseteq B$ for all $\ell \in \left[4, \left\lfloor \left(\frac{1}{8} - \epsilon\right) n \right\rfloor\right]$.*

Note that a monochromatic cycle in a 1-edge-coloured graph may be regarded as a cycle in a graph. A graph G is *pancyclic* if it contains a cycle of length ℓ, for all $\ell \in [3, |G|]$. The following result of Bondy shows that graphs with a large number of edges are pancyclic.

Theorem 1.11 (Bondy [8])**.** *If G is a hamiltonian graph of order n such that $e(G) \geq \frac{n^2}{4}$, then either G is pancyclic or n is even and $G \cong K_{n/2,n/2}$.*

As noted by Bondy [8], a corollary of Theorem 1.11 is that a graph of order n which has minimal degree at least $\frac{1}{2}n$ is pancyclic or isomorphic to $K_{n/2,n/2}$.

Corollary 1.12. *Let G be a graph of order $n \geq 3$. If $\delta(G) \geq \frac{1}{2}n$, then either G is pancyclic or n is even and $G \cong K_{n/2,n/2}$.*

Proof. Let G be a graph of order $n \geq 3$ with $\delta(G) \geq \frac{1}{2}n$. By Dirac's theorem (Theorem 1.3), G is hamiltonian. Hence by Theorem 1.11, either G is pancyclic or n is even and $G \cong K_{n/2,n/2}$ as claimed. $\square$

Note that Corollary 1.12 states that there is a unique extremal graph (with respect to minimal degree) for the property of not being pancyclic. As cycles in graphs can be regarded as monochromatic cycles in 1-edge-coloured graphs, Corollary 1.12 may be viewed as a result about 1-edge-coloured graphs. It can be seen that Conjecture 1.9 is a generalisation to the case of 2-edge-coloured graphs, in that the conclusion is that there are monochromatic cycles of all lengths in a given range. However Conjecture 1.9 does not claim that there is a unique extremal graph with minimal degree $\frac{3}{4}n$. It will be shown in Chapter 2 that Conjecture 1.9 is true for sufficiently large n and, further, that $K_{n/4,n/4,n/4,n/4}$ is the unique extremal graph.

Having looked at a 2-edge-coloured version of pancyclicity, this thesis will turn, in Chapter 3, to consideration of a 2-edge-coloured version of the circumference of a graph. This study also begins in a question posed by Li Nikiforov and Schelp [38]. However, before introducing this, there will be a discussion of results about the circumference of a graph.

Definition 1.13. *Let G be a graph. The* circumference *of G is the length of the longest cycle and is denoted $c(G)$.*

Note that a hamiltonian graph G has circumference $|G|$, and this is the largest value the circumference may take. Erdős and Gallai [19] proved the following result, which states that a graph with small circumference must have relatively few edges. This result is best possible, as can be seen by considering a graph consisting of multiple copies of K_L which mutually intersect in a single vertex.

Theorem 1.14 (Erdős and Gallai [19])**.** *If G is a graph with order n and* *circumference at most L, then $e(G) \leq \frac{1}{2}(n-1)L$. If G is a graph with order n with no paths of length at least $L+1$, then $e(G) \leq \frac{1}{2}nL$.*

The previous theorem concerns a relationship between the circumference of a graph and the number of edges it contains. However, as with many problems in extremal graph theory, the same type of question may be posed with a minimal degree condition replacing the condition on the number of edges. Independently, Bollobás and Häggkvist [7] and Egawa and Miyamoto [17] proved the following result on the circumference of a graph G with a given minimal degree. Note that this gives an exact answer to the question of how small the circumference of a graph with order n and minimal degree d can be.

Theorem 1.15. *Let d and n be positive integers such that $2 \leq d \leq n-1$, and set $m = \left\lceil \frac{1}{d}n \right\rceil$. Then*

$$\min\{c(G) : |G| = n \text{ and } \delta(G) \geq d\} = \left\lceil \frac{1}{m-1}n \right\rceil.$$

The work in Chapter 3 concerns the monochromatic circumference of a 2-edge-coloured graph, as defined below.

Definition 1.16 (Monochromatic circumference)**.** *Let G be an r-edge-coloured graph with colouring $E(G) = \bigcup_{\nu=1}^{r} E(G_\nu)$. The* monochromatic circumference *of G is the length of the longest monochromatic cycle of G.*

The following question about the monochromatic circumference of a 2-edge-coloured graph was posed by Li, Nikiforov and Schelp [38].

Question 1.17. *Let $0 < c < 1$ and G be a graph of sufficiently large order n, with a 2-edge-colouring. If $\delta(G) > cn$, how large must the monochromatic circumference of G be?*

In Chapter 3, an asymptotic version of this question will be posed and answered. Using Theorem 1.8, Faudree, Lesniak and Schiermeyer [23] were able to deduce the following result about the monochromatic circumference of a 2-edge colouring of the complete graph.

Theorem 1.18. *Let $n \geq 6$, and suppose that K_n is 2-edge-coloured with colouring $E(K_n) = E(R) \cup E(B)$. Then either R or B contains a cycle of length at least $\lceil \frac{2}{3}n \rceil$.*

Further, there is a 2-edge-colouring of K_n such that both R and B have circumference at most $\lceil \frac{2}{3}n \rceil$.

As stated previously, Chapter 2 will contain a proof that Conjecture 1.9 is true for sufficiently large n, and Chapter 3 will answer an asymptotic version of Question 1.17. Both Conjecture 1.9 and Question 1.17 concern finding monochromatic cycles in 2-edge-coloured graphs, and so these two chapters are intimately linked. In particular, the same method will be used in both chapters to find monochromatic cycles, and many of the results of Chapter 2 will be of use in Chapter 3.

1.3 2-factors

Let G be a graph. For a given function $f : V(G) \to \mathbb{N}$, a subgraph H of G is an *f-factor of G* if $V(H) = V(G)$ and $\deg_H(v) = f(v)$ for all $v \in V(G)$. For a positive integer r, an *r-factor* is an f-factor where $f(v) = r$ for all $v \in V(G)$. Note that G need not contain an f-factor. For example if $f(v) > \deg_G(v)$ for some vertex v, then G does not contain an f-factor.

In Chapter 4 the existence of 2-factors in graphs will be considered. Note that a 2-factor of a graph G is a vertex-disjoint union of cycles, which contain all vertices of G. If a 2-factor is connected, it must be a Hamilton cycle. Hence Theorem 1.3 implies that a graph with order $n \geq 3$ and minimal degree at least $\frac{1}{2}n$ contains a 2-factor with one component. The following result of Brandt, Chen, Faudree, Gould and Lesniak [12] extends Theorem 1.3.

Theorem 1.19. *Let k and n be positive integers, with $n \geq 4k$. Every graph G on n vertices with minimal degree at least $\frac{1}{2}n$ has a 2-factor with k components.*

Note that, if $k \geq 2$ and $n \geq 3k+2$, not all 2-factors with k components are isomorphic. For example, in a graph of order n, a 2-factor with two components may be isomorphic to any of the graphs $\left\{C_t \cup C_{n-t} : 3 \leq t \leq \left\lfloor \frac{1}{2}n \right\rfloor \right\}$.

If n is odd, the graph $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$ has n vertices and minimal degree $\frac{1}{2}(n-1)$. However, it does not contain any 2-factors, as all cycles must contain the same number of vertices in each of the two classes. Hence, both Theorem 1.3 and Theorem 1.19 are best possible with respect to the bound on the minimal degree. However, as $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$ does not contain any 2-factors, a natural question to ask is whether the minimal degree condition can be weakened for a graph which is already known to contain a 2-factor. The simplest case is to find 2-factors with k components in a hamiltonian graph. This question was introduced by Faudree, Gould, Jacobsen, Lesniak and Saito [22], who proved the following result.

Theorem 1.20. *For $n \geq 6$, any hamiltonian graph with order n and minimal degree at least $\frac{5}{12}n + 2$ contains a 2-factor with two components.*

Provided that $n \geq 4k$, Theorem 1.19 implies that any hamiltonian graph with order n and minimal degree at least $\frac{1}{2}n$ contains a 2-factor with k components. It was shown by G.Sárközy [44] that, for sufficiently large n, the minimal degree condition can be improved.

Theorem 1.21. *There exists a real number $\epsilon > 0$ such that for every integer $k \geq 2$ there exists an integer $n_0 = n_0(k)$ such that any hamiltonian graph with order $n \geq n_0$ and minimal degree at least $(\frac{1}{2} - \epsilon)n$ contains a 2-factor with k components.*

In Chapter 4, an improvement of the minimal degree condition in Theorem 1.20 will be demonstrated.

1.4 Random graphs

There are many different models for random graphs. Two of the most common and important models are the *binomial random graph* and *uniform random graph*, both of which have their roots in work of Erdős [18]. The model for a uniform random graph was formalised by Erdős and Rényi [20], who wrote a series of papers about this model, while the binomial random graph was introduced by Gilbert [25]. In this thesis, the focus will be on the binomial random graph, as defined below.

Definition 1.22. *Let n be an integer and $p \in [0,1]$. The* binomial random graph *$\mathcal{G}(n,p)$ has vertex set $[n]$, and each possible edge is included independently, with probability p. Wehn p is understood from context, a graph with this distribution may also be denoted G_n.*

Note that, if G is a labelled graph with M edges, then the probability that $\mathcal{G}(n,p) = G$ is $p^M(1-p)^{N-M}$, where $N = \binom{n}{2}$. The case when $p = \frac{1}{2}$ is a particularly special case as $\mathcal{G}(n, 1/2)$ is equally likely to be any of the graphs on n vertices.

Definition 1.23. *A* graph property *is a property of graphs which is preserved under all isomorphisms of a graph.*

In the study of $\mathcal{G}(n,p)$, it is often the case that n is considered to be tending to infinity, while p is a (possibly constant) function of n. For some graph property P, say that *almost every graph in $\mathcal{G}(n,p)$ has property P* if, as n tends to infinity, the probability that $\mathcal{G}(n,p)$ has P tends to one. For example, it was shown by Komlós and Szemerédi [35] and Korsunov [36] that if $p = \frac{1}{n}(\log n + \log\log n + \omega(n))$ and $\omega(n) \to \infty$, then almost every graph in $\mathcal{G}(n,p)$ contains a Hamilton cycle. An overview of the study of random graphs is provided by the books of Bollobás [6] and Janson, Łuczak and Ruciński [30].

In this thesis, a slightly different type of problem will be considered, which uses the concept of modular arithmetic.

Definition 1.24. *Let k be a positive integer. Two integers a and b are said to be* equivalent modulo k *if $a - b$ is divisible by k. This will be denoted $a \underset{k}{\equiv} b$.*

For any fixed graph H, the number of copies of H in $\mathcal{G}(n,p)$ is denoted by X_H. Various properties of X_H have been studied. For example, Bollobás [4] showed the following theorem.

Theorem 1.25. *Let H be a graph with at least one edge, and let*

$$m(H) = \max\left\{\frac{e(H')}{|H'|} : H' \subseteq H \text{ and } |H'| > 0\right\}.$$

Then, if $pn^{\frac{1}{m(H)}} \to 0$, almost every graph does not contain a copy of H, while if $pn^{\frac{1}{m(H)}} \to \infty$, almost every graph contains a copy of H, where the underlying distribution is that of $\mathcal{G}(n,p)$.

Independently, Bollobás [4] and Karoński and Ruciński [31] also considered the distribution of X_H when $pn^{\frac{1}{m(H)}} \to c$ for some constant $c > 0$.

In this thesis, p will always be assumed to be a fixed constant and the probability that $X_H \underset{k}{\equiv} a$ for some integer a will be considered. The main aim will be to show that X_H is approximately uniformly distributed modulo k, in the following sense.

Definition 1.26. *Let $\epsilon > 0$ and $k \in \mathbb{N}$. An integer-valued random variable Y is* ϵ-uniform modulo k *if*

$$\max_{1 \le a \le k} \left| \mathbb{P}\left(Y \underset{k}{\equiv} a\right) - \frac{1}{k} \right| < \frac{\epsilon}{k}.$$

Note that $X_{K_1} = n$ and so, for small values of ϵ, the variable X_{K_1} can not be ϵ-uniformly distributed modulo k. For the purposes of this thesis, the graph K_1 will be referred to as the trivial graph. However, X_{K_2} has the binomial distribution, with parameters $\binom{n}{2}$ and p. It will be seen in

Chapter 5 that, when k is not too large with respect to n, a variable with the binomial distribution, with parameters n and p, is $\exp\{-\Omega_p(n/k^2)\}$-uniformly distributed modulo k, provided that k is not too large. Hence X_{K_2} is approximately uniformly distributed modulo k, provided that k is not too large.

Loebl, Matoušek and Pangrác [39] proved that, for $p = \frac{1}{2}$, X_{K_3} is ϵ-uniform modulo k for any $\epsilon > 0$ and prime k with $k = O(\log n)$. This work was extended by Scott and Tateno [45] to show that, for any fixed t and p, $X_{K_t}(G_n)$ is ϵ-uniform modulo k for any $\epsilon > 0$, provided that $k = o(\sqrt{n}/\log n)$.

In Chapter 5, it will be shown that, for constant p and $k = o\left(\sqrt{n/\log n}\right)$, and a non-trivial connected graph H, the variable X_H is $\exp\{-\Omega_p(n/k^2)\}$-uniform modulo k. Further, for any finite set of connected non-trivial graphs $\mathcal{C}$, the vector variable $(X_H)_{H\in\mathcal{C}}$ is $\exp\{-\Omega_p(n/k^2)\}$-uniform modulo k. Independently, Kolaitis and Kopparty [32] showed this for constant k. Note that only connected graphs are considered because, for example, two distinct edges from a graph either form a $K_{1,2}$ or a $2K_2$. Hence

$$\binom{X_{K_2}}{2} = X_{2K_2} + X_{K_{1,2}},$$

and so X_{2K_2} is determined by $X_{K_{1,2}}$ and X_{K_2}.

Chapter 2

Cycles in 2-edge-coloured graphs

2.1 Cycles of fixed lengths in 2-edge-coloured graphs

In this chapter, cycles in 2-edge-coloured graphs will be considered. Recall from Chapter 1 the following conjecture of Li, Nikiforov and Schelp [38].

Conjecture 2.1. *Let $n \geq 4$ and let G be a graph of order n with $\delta(G) > \frac{3}{4}n$. If $E(G) = E(R) \cup E(B)$ is a 2-edge-colouring, then $C_\ell \subseteq R$ or $C_\ell \subseteq B$ for all $\ell \in \left[4, \left\lceil \frac{1}{2}n \right\rceil\right]$.*

This chapter will contain a proof that this conjecture is true for sufficiently large n. Let $K_m(t)$ be the complete m-partite graph with t vertices in each class. In fact it will be shown that $K_4(n/4)$ is the unique extremal graph for Conjecture 2.1. There are many colourings showing that $K_4(n/4)$ is extremal but they are all 2-bipartite as defined below.

Definition 2.2. *A* 2-bipartite 2-edge-colouring *of a graph G is a 2-edge-colouring $E(G) = E(R) \cup E(B)$ such that both R and B are bipartite.*

Note that such a colouring contains no monochromatic cycles of odd length. The main result of this chapter is Theorem 2.3, which proves Conjecture 1.9 for sufficiently large n.

Theorem 2.3. *There exists an integer N such that the following holds. Let G be a 2-edge-coloured graph of order $n \geq N$ with $\delta(G) \geq \frac{3}{4}n$ with 2-edge-colouring $E(G) = E(R) \cup E(B)$. Then either $G \cong K_4(n/4)$ and the colouring is a 2-bipartite 2-edge-colouring or there is a monochromatic copy of C_ℓ for all $\ell \in \left[4, \left\lceil \frac{1}{2}n \right\rceil\right]$.*

Given a 2-bipartite 2-edge-colouring of $K_4(n/4)$, let $V_1 \cup V_2$ be the bipartition of R and $W_1 \cup W_2$ be the bipartition of B. Let $U_{i,j} = V_i \cap W_j$ for all $i, j \in \{1, 2\}$. Then $U_{i,j}$ is an independent set of $K_4(n/4)$ for all $i, j \in \{1, 2\}$. As the sets $U_{i,j}$ partition the vertices of $K_4(n/4)$, they must be the four independent sets of order $n/4$. Hence, a 2-bipartite 2-edge-colouring of $K_4(n/4)$ forces a labelling of the independent sets $\{U_{i,j} : i, j \in \{1, 2\}\}$ such that:

- all edges between $U_{1,1}$ and $U_{1,2}$ and between $U_{2,1}$ and $U_{2,2}$ are blue;
- all edges between $U_{1,1}$ and $U_{2,1}$ and between $U_{1,2}$ and $U_{2,2}$ are red;
- edges between $U_{1,1}$ and $U_{2,2}$ and between $U_{2,1}$ and $U_{1,2}$ can be either colour.

Note that in this situation there are $2^{\frac{1}{8}n^2}$ ways to colour the edges between $U_{1,1}$ and $U_{2,2}$ and between $U_{2,1}$ and $U_{1,2}$. However, $K_4\left(n/4\right)$ has $24\left(\left(\frac{1}{4}n\right)!\right)^4 = 2^{O(n\log n)}$ automorphisms. Hence the number of distinct 2-bipartite 2-edge-colourings of $K_4\left(n/4\right)$ is $2^{\frac{1}{8}n^2-O(n\log n)}$.

It is clear that Theorem 2.3 is an immediate corollary of the following result. The reason for proving the stronger result is that it also implies a bound on the monochromatic circumference of G, which will be used in Chapter 3

Theorem 2.4. *Let $0 < \delta < \frac{1}{294}$. There exists an $n_0 = n_0\left(\delta\right)$ such that the following holds. Let G be a 2-edge-coloured graph of order $n \geq n_0$ with $\delta\left(G\right) \geq \frac{3}{4}n$. Suppose that $E(G) = E(R_G) \cup E(B_G)$ is the 2-edge-colouring of G. Then one of the following occurs:*

- *n is a multiple of four, $G \cong K_4\left(n/4\right)$ and the colouring is a 2-bipartite 2-edge-colouring;*
- *G contains a monochromatic cycle of length at least $\left\lceil\left(\frac{2}{3}+\frac{\delta}{2}\right)n\right\rceil$ and, for all $\ell \in \left[4, \left\lceil\frac{1}{2}n\right\rceil\right]$, either R_G or B_G contains a cycle of length ℓ;*
- *either R_G or B_G contains cycles of all lengths $\ell \in \left[3, \left\lceil\left(\frac{2}{3}-\delta\right)n\right\rceil\right]$.*

The proof of Theorem 2.4 is divided into two parts. In Section 2.3, short (up to constant length) cycles will be considered, while long cycles are considered in Section 2.4. The proof of Theorem 2.4 will require some results about matchings in 2-edge-coloured graphs, which are proved in Section 2.5. The proof of Theorem 2.4 is then given in Section 2.6. The following section will introduce the regularity method, which is used to find long cycles.

2.2 Regularity and other useful results

In proving results about cycles in r-edge-coloured graphs, the Regularity Lemma and Blow-up Lemmas will be used to find long cycles. Before introducing these, some preliminary definitions are required.

Definition 2.5. *Let G be a graph and X and Y be disjoint subsets of $V(G)$. The* density *of the graph $G[X, Y]$ is the value*

$$d_G(X, Y) := \frac{e(X, Y)}{|X||Y|}.$$

Where it is clear which graph is being considered, the density will be denoted $d(X, Y)$. Also, if G is a graph with edge-colouring $E(G) = \bigcup_{\nu=1}^{r} E(G_\nu)$, then $d_\nu(X, Y) := d_{G_\nu}(X, Y)$ and $e_\nu(X, Y) := e_{G_\nu}(X, Y)$.

A regular pair is one in which the density between not-too-small subgraphs of X and Y is close to the density between X and Y.

Definition 2.6 (Regularity)**.** *Let $\epsilon > 0$. Let G be a graph and X and Y be disjoint subsets of $V(G)$. The graph $G[X, Y]$ is* ϵ-regular *if, for all $X' \subseteq X$ and $Y' \subseteq Y$ satisfying $|X'| \geq \epsilon|X|$ and $|Y'| \geq \epsilon|Y|$,*

$$|d(X, Y) - d(X', Y')| \leq \epsilon.$$

In this case, (X, Y) may also be referred to as an ϵ-regular pair (for G)*.*

It is often useful to also have a bound on the degree of vertices in X and Y, and for this the notion of super-regularity is introduced.

Definition 2.7 (Super-regularity)**.** *Let ϵ and δ be positive real numbers. Let G be a graph and X and Y be disjoint subsets of $V(G)$. The graph $G[X,Y]$ is* (ϵ,δ)-super-regular *if, for all $X' \subseteq X$ and $Y' \subseteq Y$ satisfying $|X'| \geq \epsilon|X|$ and $|Y'| \geq \epsilon|Y|$,*

$$e(X',Y') > \delta|X'||Y'|,$$

and furthermore, $\deg_Y(v) > \delta|Y|$ *for all $v \in X$ and* $\deg_X(v) > \delta|X|$ *for all $v \in Y$.*

Note that a super-regular pair need not be regular, as the number of edges between subsets is only bounded below. The following useful lemma shows that large subgraphs of a regular graph are still regular, and similarly large subgraphs of a super-regular graph are still super-regular.

Lemma 2.8. *Let $0 < \eta, \epsilon < \frac{1}{2}$ and let δ be a positive constant. Let G be a graph and X and Y be disjoint subsets of $V(G)$. Let X' be a subset of X and Y' be a subset of Y.*

If $|X'| \geq \frac{1}{2}|X|$, $|Y'| \geq \frac{1}{2}|Y|$ and (X,Y) is an ϵ-regular pair for G with density at least δ, then $G[X',Y']$ is 2ϵ-regular, and has density at least $\delta-\epsilon$.

If $|X'| \geq (1-\eta)|X|$, $|Y'| \geq (1-\eta)|Y|$ and (X,Y) is an (ϵ,δ)-super-regular pair for G, then $G[X',Y']$ is $(2\epsilon,\delta-\eta)$-super-regular.

Proof. Suppose first that (X,Y) is an ϵ-regular pair for G and that $|X'| \geq \frac{1}{2}|X|$ and $|Y'| \geq \frac{1}{2}|Y|$. Let X^* be a subset of X' and Y^* be a subset of Y' such that $|X^*| \geq 2\epsilon|X'| \geq \epsilon|X|$ and $|Y^*| \geq 2\epsilon|Y'| \geq \epsilon|Y|$. The regularity of $G[X,Y]$ implies that

$$|d(X,Y) - d(X',Y')| < \epsilon \tag{2.1}$$

and

$$|d(X,Y) - d(X^*,Y^*)| < \epsilon.$$

Hence, by the triangle inequality,

$$|d(X',Y') - d(X^*,Y^*)| < 2\epsilon,$$

and so $G[X',Y']$ is 2ϵ-regular. Moreover, (2.1) implies that $G[X',Y']$ has density at least $\delta - \epsilon$.

Now suppose that (X,Y) is an (ϵ,δ)-super-regular pair for G and that $|X'| \geq (1-\eta)|X|$ and $|Y'| \geq (1-\eta)|Y|$. Let X^* be a subset of X' and Y^* be a subset of Y' such that $|X^*| \geq 2\epsilon|X'| > \epsilon|X|$ and $|Y^*| \geq 2\epsilon|Y'| > \epsilon|Y|$. As $G[X,Y]$ is super-regular,

$$e(X^*,Y^*) > \delta|X^*||Y^*|.$$

Let $v \in X'$. Then

$$\begin{aligned}\deg_{Y'}(v) &\geq \deg_Y(v) - (|Y| - |Y'|) \\ &> \delta|Y| - |Y| + (1-\eta)|Y| = (\delta-\eta)|Y| \geq (\delta-\eta)|Y'|.\end{aligned}$$

Similarly the degree in X' of any vertex in Y' is greater than $(\delta-\eta)|X'|$. Hence $G[X',Y']$ is $(2\epsilon,\delta-\eta)$-super-regular. □

The (uncoloured) Regularity Lemma, due to Szemerédi [47] states that, for sufficiently large n, the vertex set of a graph of order n can be partitioned into clusters, such that, for most pairs of clusters, the bipartite graph between them is regular. Many different forms of the Regularity Lemma can be found (see, for example, the survey paper of Komlós and Simonovits [34]) and of particular use in this thesis is the degree form of the lemma, which here is stated for r-edge coloured graphs. Although statements of an edge-coloured version of the Regularity Lemma are common, it is difficult to locate a proof and so, for completeness, a proof is given in Appendix A.

Theorem 2.9 (Degree form of r-coloured Regularity Lemma)**.** *For every $\epsilon > 0$ and integers r and m there is an $M = M(\epsilon, r, m)$ such that, if G is an r-edge-coloured graph of sufficiently large order n with edge-colouring $E(G) = \bigcup_{\nu=1}^{r} E(G_\nu)$, and $\eta \in [0,1]$ is any real number, then there is a partition $\mathcal{P} = (V_i)_0^k$ with $m \leq k \leq M$ and a subgraph $G' \subseteq G$ with the following properties.*

- $|V_0| \leq \epsilon |V|$*;*
- *all clusters V_i, $i \geq 1$, are of the same size;*
- $\deg_{G'}(v) > \deg_G(v) - (r\eta + \epsilon)|V|$ *for all $v \in V$;*
- $e(G'(V_i)) = 0$ *for all $i \geq 1$;*
- *for all $1 \leq i < j \leq k$ and $1 \leq \nu \leq r$, the graph $G'_\nu[V_i, V_j]$ is ϵ-regular with density either 0 or greater than η (where, for all $1 \leq \nu \leq r$, $E(G'_\nu) = E(G') \cap E(G_\nu)$).*

The following notion of a *reduced graph* is particularly useful. It will be useful to introduce a new type of edge-colouring prior to defining the reduced graph.

Definition 2.10 (edge-multi-colouring). *Given a graph G with edge set $E(G)$, an* r-edge-multi-colouring *of G is a set $(G_i)_{i=1}^r$ of spanning subgraphs of G such that $E(G) = \bigcup_{i=1}^r E(G_i)$. Note that if this is a partition of $E(G)$, the edge-multi-colouring is an edge-colouring as defined earlier.*

Definition 2.11 (Reduced graph). *For a given $\epsilon > 0$ and $\eta \in [0, 1]$, let G' be a 2-edge-coloured graph and $\mathcal{P} = (V_0)_0^k$ be a partition of $V(G')$ with the properties listed in the Regularity Lemma (Theorem 2.9). The* reduced graph *H has vertex set $\{v_1, \ldots v_k\}$ and a 2-edge-multi-colouring $E(H) = E(R_H) \cup E(B_H)$, where B_H and R_H are defined as follows. Let $\{v_i, v_j\}$ be an edge of B_H when $B_{G'}[V_i, V_j]$ has density at least η and let $\{v_i, v_j\}$ be an edge of R_H when $R_{G'}[V_i, V_j]$ has density at least η.*

Note that this definition is symmetric with respect to the two colours, a fact that will be used in the proof of Theorem 2.4.

The main tool for finding long cycles in a 2-edge-coloured graph G will be to apply the Regularity Lemma to G and obtain a reduced graph H. The structure of H will then be used to find long cycles in G. There are two main tools for doing so. The first is the Embedding Lemma. There are many forms of this result. Proofs are given by, for example, Diestel [15, p.184] and Komlós and Simonovits [34].

Theorem 2.12 (Embedding Lemma). *Given $d > \epsilon > 0$, a graph H, and a positive integer t, let G be a graph obtained by replacing each vertex of H with a set of order t, and replacing the edges of H with ϵ-regular pairs of density at least d. For a fixed integer a, let $H(a)$ be the graph defined by replacing each vertex of H with a set of order a, and replacing the edges of*

H with complete bipartite graphs.

Let F be a subgraph of $H(a)$ with and maximum degree $\Delta > 0$, and let $\eta = d - \epsilon$ and $\epsilon_0 = \eta^{\Delta} / (2 + \Delta)$. If $\epsilon \leq \epsilon_0$ and $a - 1 \leq \epsilon_0 t$, then $F \subseteq G$.

Note that in the Embedding Lemma, the subgraphs F under consideration have order at most $a|H| = \frac{a}{t}|G| \leq \left(\epsilon_0 + \frac{1}{t}\right)|G|$, and so are small compared to G. To prove Theorem 2.4, however, cycles of length $\frac{2}{3}|G|$ must be shown to lie in G. To embed larger graphs into G the Blow-up Lemma of Komlós, Sárközy and Szemerédi [33] will be used. Note that, in this case, edges of H are replaced by *super*-regular pairs.

Theorem 2.13 (Blow-up Lemma)**.** *Given a graph H of order s and positive parameters δ, Δ and c, there exist positive numbers $\epsilon_0 = \epsilon_0(\delta, \Delta, s, c)$ and $\alpha = \alpha(\delta, \Delta, s, c)$ such that the following holds. Let t be an arbitrary positive integer, and replace the vertices $v_1, \ldots, v_s$ of H with pairwise disjoint sets $V_1, \ldots, V_s$ of order t. Two graphs are constructed on the same vertex set $V = \bigcup V_i$. The first graph G_1 is obtained by replacing each edge $\{v_i, v_j\}$ of H with the complete bipartite graph between the corresponding vertex sets V_i and V_j. A sparser graph G_2 is constructed by replacing each edge $\{v_i, v_j\}$ arbitrarily with an (ϵ_0, δ)-super-regular pair between V_i and V_j.*

If a graph F with $\Delta(F) \leq \Delta$ is embeddable into G_1 then it is also embeddable into G_2. This remains true even if for every i there are certain vertices x to be embedded into V_i whose images are restricted to certain sets $C_x \subseteq V_i$, provided that:

(i) each C_x is of order at least ct;

(ii) the number of such restrictions within a set V_i is not more than αt.

This section concludes with some other results used in the proof of Theorem 2.4. In Section 2.5, some results about matchings in 2-edge-multi-coloured graphs will be proved. For bipartite graphs, Hall's Theorem [28] gives an exact condition for the existence of a large matching.

Theorem 2.14 (Hall's Theorem)**.** *Let G be a bipartite graph between sets X and Y. If*

$$|\Gamma(S)| \geq |S| - d$$

for all $S \subseteq X$, then G contains a matching from X to Y with at least $|X| - d$ edges.

For general graphs, Tutte's 1-Factor Theorem [49] gives an exact condition for the existence of a perfect matching, and this theorem will be used throughout Section 2.5 and again in Chapter 3. Before stating the theorem, the following definition is required.

Definition 2.15. *For any graph G, let $q(G)$ be the number of components of G of odd order.*

The following defect version of Tutte's 1-Factor Theorem was first noted by Berge [2].

Theorem 2.16 (Tutte's 1-Factor Theorem)**.** *A graph G contains a set of independent edges covering all but at most d of the vertices if, and only if,*

$$q(G - S) \leq |S| + d$$

for all $S \subseteq V$.

The following result of Menger [40] is about connectivity, and will be used to show that there are paths between subsets of $V(G)$.

Theorem 2.17. *Let k be a positive integer. Let G be a graph and let S and T be disjoint subsets of $V(G)$, both of order at least k. Then one of the following occurs.*

- *There are k vertex-disjoint paths from S to T contained in G.*
- *There is a set $U \subseteq V(G)$ of order at most $k-1$ such that, in $G-U$, there are no paths between $S \setminus U$ and $T \setminus U$.*

Results about cycles in uncoloured graphs will be used to prove results about cycles in 2-edge coloured graphs. For a graph G with $E(G) = E(R) \cup E(B)$, this will often take the form of showing that there is some set $S \subseteq V(G)$ with $|S| \geq \frac{1}{2}|G|$ such that $R[S]$ is hamiltonian and contains lots of edges (the choice of R over B being arbitrary). Bondy's Theorem (Theorem 1.11) then implies that either $R[S]$ is a complete bipartite graph or $R[S]$ is pancyclic and hence R contains a cycle of length ℓ for all $\ell \in \left[3, \left\lceil \frac{1}{2}n \right\rceil\right]$. In most cases, Dirac's Theorem (Theorem 1.3) will suffice to show that $R[S]$ is hamiltonian. However, there are a couple of cases in which the stronger result below is required. Recall that the degree sequence of a graph G with vertices $\{v_1, \ldots, v_n\}$ is the sequence $d_1, \ldots, d_n$, where $d_i = \deg(v_i)$ for all $1 \leq i \leq n$.

Theorem 2.18 (Chvátal [14])**.** *Let G be a graph of order $n \geq 3$ with degree sequence $d_1 \leq d_2 \leq \cdots \leq d_n$ such that*

$$d_k \leq k < \frac{1}{2}n \Rightarrow d_{n-k} \geq n-k.$$

Then G contains a Hamilton cycle.

The following result of Bollobás [3, p.150] is also concerned with cycles in uncoloured graphs, and will be used in Section 2.3 to find short monochromatic cycles, as will the result of Györi, Nikiforov and Schelp below.

Theorem 2.19 (Bollobás [3]). *If G is a graph of order n, with $e(G) > \frac{1}{4}n^2$, then G contains C_k for all $k \in \left[3, \lceil \frac{n}{2} \rceil\right]$.*

Theorem 2.20 (Györi, Nikiforov and Schelp [27]). *Let k, m be positive integers. There exist $n_0 = n_0(k, m)$ and $c = c(k, m) > 0$ such that for every nonbipartite graph G on $n > n_0$ vertices with minimum degree*

$$\delta > \frac{n}{2(2k+1)} + c,$$

if $C_{2s+1} \subseteq G$, for some $k \leq s \leq 4k+1$, then $C_{2s+2j+1} \subseteq G$ for every $j \in [m]$.

2.3 Existence of short cycles

The aim of this section is to show that a graph satisfying the conditions of Theorem 2.4 either contains monochromatic cycles of all short lengths or is the extremal graph identified in Theorem 2.4. The proof of the following lemma follows exactly the method used by Li, Nikiforov and Schelp [38] in their proof of Theorem 1.10 to show the existence of short monochromatic cycles.

Lemma 2.21. *Let L be an integer. Let n be sufficiently large, let G be a 2-edge-coloured graph of order n with $\delta(G) \geq \frac{3}{4}n$, and let $E(G) = E(R_G) \cup$*

$E(B_G)$ be the 2-edge-colouring of G. If G contains a monochromatic C_3 or C_5, then it contains a monochromatic C_ℓ for all odd $\ell \in [5, 2L+1]$.

Proof. Suppose first that $\Delta(B) > \frac{1}{2}n + 4L$. Let v be a vertex with blue degree $\Delta(B)$, and U be the neighbourhood of v in B. Suppose that $B[U]$ contains a path $v_1 \dots v_{2L+1}$ of length $2L$. Then, as illustrated in Figure 2.1, $vv_1 \dots v_i v$ is a blue cycle of length $i+1$ for all $2 \le i \le 2L$.

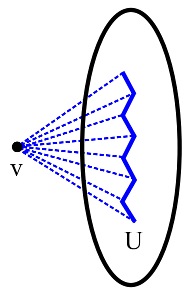


Figure 2.1: If the blue neighbourhood U of a vertex v contains a long blue path, then B contains cycles of many lengths.

Suppose now that $B[U]$ does not contain a path of length $2L$. Theorem 1.14 implies that $e(B[U]) \le L|U|$. However, any vertex $v \in U$ has at most $\frac{1}{4}n$ non-neighbours in U and so at least $|U| - \frac{1}{4}n$ neighbours. Hence

$$\begin{aligned} e(G[U]) &= \frac{1}{2}\sum_{u \in U} \deg_U(u) \\ &\ge \frac{1}{2}|U|\left(|U| - \frac{1}{4}n\right) \\ &> \frac{1}{2}|U|\left(\frac{1}{2}|U| + 2L\right). \end{aligned}$$

Hence $e(R[U]) = e(G[U]) - e(B[U]) > \frac{1}{4}|U|^2$, and so Theorem 2.19 implies that $R[U]$ has cycles of all lengths from 3 to $\frac{1}{2}|U|$.

Suppose now that $\Delta(B) \leq \frac{1}{2}n + 4L$. This implies that, for n sufficiently large

$$\delta(R) \geq \frac{1}{4}n - 4L > \frac{1}{6}n + c(1, L),$$

where $c(1, L)$ is the constant from Theorem 2.20. Hence, if $C_{2s+1} \subseteq R$ for some $1 \leq s \leq 5$, then Theorem 2.20 implies that $C_{2s+2j+1} \subseteq R$ for all $1 \leq j \leq L$. Similarly, it may be assumed that $\delta(B) > \frac{1}{6}n + c(1, L)$, and so, if $C_{2s+1} \subseteq B$ for some $1 \leq s \leq 5$, then $C_{2s+2j+1} \subseteq B$ for all $1 \leq j \leq L$. Hence, if G contains a monochromatic C_3 or C_5 then it contains a monochromatic C_ℓ (of the same colour) for all odd $\ell \in [5, 2L+1]$ as required. □

Lemma 2.21 may now be used to prove the following lemma, the main result of this section. Note that this result can not be proven by directly applying Theorem 1.10, as the assumption in that theorem is that $\delta(G) > \frac{3}{4}n$, which is stronger than the assumption below.

Lemma 2.22. *Let K be an integer. Let G be a 2-edge-coloured graph of sufficiently large order n with $\delta(G) \geq \frac{3}{4}n$, and let $E(G) = E(R_G) \cup E(B_G)$ be the 2-edge-colouring of G. Then either $G \cong K_4(n/4)$ and the colouring is a 2-bipartite 2-edge-colouring, or $C_\ell \subseteq R$ or $C_\ell \subseteq B$ for all $\ell \in [4, K]$.*

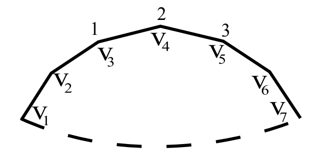
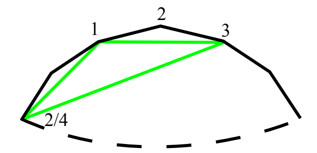
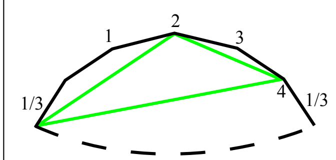
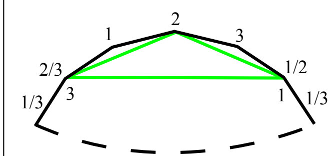
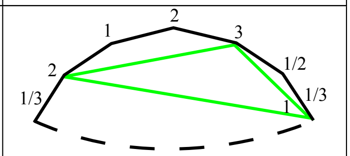
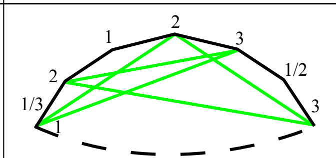
Proof. The minimal degree condition implies that G contains at least $\frac{3}{8}n^2$ edges. Hence it may be assumed that R_G has at least $\frac{3}{16}n^2$ edges, and so Theorem 1.2 implies that R_G contains C_ℓ for all even $\ell \in [4, K]$. Hence, by Lemma 2.21, it is sufficient to prove that either $G \cong K_4(n/4)$ and the colouring is a 2-bipartite 2-edge-multi-colouring or there is a monochromatic C_3 or C_5.

Suppose that G does not contain a monochromatic C_3 or C_5. Any 2-edge-multi-colouring of K_5 contains a monochromatic C_3 or C_5. Hence $K_5 \nsubseteq G$. However $e(G) \geq \frac{3}{8}n \geq e(T_4(n))$ and hence Turán's Theorem (Theorem 1.6) implies that $G \cong T_4(n)$. However, $\delta(G) \geq \frac{3}{4}n$ implies that in fact n is a multiple of four and hence $G \cong K_4\,(n/4)$. Let U_i $(1 \leq i \leq 4)$ be the independent sets of G of order $n/4$. It only remains to show that both R and B are bipartite.

Suppose that R is not bipartite. Let $C = v_1 v_2 \ldots v_r$ be a shortest odd cycle of R; by assumption $r \geq 7$. Define a vertex colouring c of R by setting $c(v) = j$ when $v \in U_j$. Note that c is a proper colouring as each U_j is an independent set. As C is an odd cycle, there must be three consecutive vertices of C with different colours under c. Without loss of generality, assume that $c(v_3) = 1$, $c(v_4) = 2$ and $c(v_5) = 3$. Suppose now that $G[V(C)]$ does not contain a triangle or 5-cycle which is edge-disjoint from C. Table 2.1 gives a case-by-case analysis, eventually yielding a contradiction to this assertion.

Thus $G[V(C)]$ contains a triangle or 5-cycle which is edge-disjoint from C. An edge of that triangle or 5-cycle must be red, else it is monochromatic. But this red edge, together with some segment of C, forms a shorter red odd cycle than C. This contradicts the assumption that C was minimal. Hence R is bipartite, and similarly so is B, completing the proof. □

At each stage in the following table, the numbers on the outer edge of the cycle are the only potential colours for that vertex after previous cases, and the numbers on the inner edge are those used in the particular case under consideration. In each case, the green edges are a triangle or 5-cycle of G, which is edge-disjoint from C.

By assumption, $c(v_3) = 1$, $c(v_4) = 2$ and $c(v_5) = 3$. Note that $\|C\| \geq 7$ and so v_1v_7 may or may not be an edge of C.	
If $c(v_1)$ is 2 or 4, then G contains the triangle $v_1v_3v_5$, as these vertices lie in different U_j. Hence $c(v_1) \in \{1,3\}$, and similarly $c(v_7) \in \{1,3\}$.	
If $c(v_6) = 4$, then G contains the triangle $v_1v_4v_6$. Hence $c(v_6) \neq 4$ and similarly $c(v_2) \neq 4$. As $c(v_3) = 1$ and $c(v_5) = 3$, this implies that $c(v_2) \in \{2,3\}$ and $c(v_6) \in \{1,2\}$.	
If $c(v_2) = 3$ and $c(v_6) = 1$, then G contains the triangle $v_2v_4v_6$. Hence, at least one of $c(v_2)$ and $c(v_6)$ must be 2. By symmetry, assume that $c(v_2) = 2$ and $c(v_6) \in \{1,2\}$.	
If $c(v_7) = 1$, then G contains the triangle $v_2v_5v_7$. Hence $c(v_7) = 3$.	
If $c(v_1) = 1$, then $v_1v_5v_2v_7v_4$ is a 5-cycle in G, not containing any edges of C. Hence $c(v_1) = 3$.	

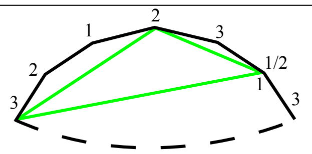
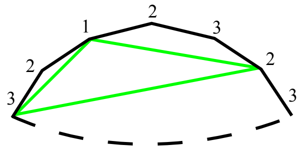
If $c(v_6) = 1$, then G contains the triangle $v_1v_4v_6$. Hence $c(v_6) = 2$.	
But then G contains the triangle $v_1v_3v_6$, contradicting the assumption that $G[V(C)]$ does not contain a triangle or 5-cycle which is edge-disjoint from C.	

Table 2.1: Pictorial representation showing that $G\,[V(C)]$ contains a C_3 or C_5 which is edge-disjoint from C.

2.4 Tools for finding long cycles

As noted already, the main method of finding long monochromatic cycles will be to apply the Regularity Lemma to a 2-edge-coloured graph and consider the reduced graph H, as defined in Section 2.2. Recall that H has a 2-edge-multi-colouring $E(H) = E(R_H) \cup E(B_H)$. Suppose that a component of R_H contains a large matching. Applying the Embedding Lemma in the red paths of H between two edges of the matching, red paths of G linking the pairs of clusters corresponding to the matching can be found. The Blow-Up Lemma can then be used within each pair of clusters in G corresponding to a matching edge of H to join these red paths together, creating long red cycles. The following Lemma formalises this argument.

Lemma 2.23. *Given $0 < \eta < 1$, there exists $\epsilon = \epsilon\,(\eta)$ such that the following* *holds. Let H be a connected graph of order s' containing a matching M on $2s \geq 4$ vertices. Let t be a positive integer satisfying*

$$t \geq \max\left\{\frac{8s}{\eta}, 12s', \frac{4s}{\alpha\left(\frac{1}{8}\eta, 2, 2, \frac{1}{8}\eta\right)}\right\}$$

where α is as defined in the Blow-Up Lemma (Theorem 2.13).

Replace the vertices $\{v_i : 1 \leq i \leq s'\}$ of H with pairwise disjoint sets $V_1, \ldots, V_{s'}$ of order t. Let G be a graph defined by replacing each edge v_iv_j of H with a 2ϵ-regular pair between V_i and V_j of density at least $\frac{1}{2}\eta$, with the added condition that, for each edge v_iv_j of M, the graph $G[V_i, V_j]$ is also $\left(2\epsilon, \frac{1}{2}\eta\right)$-super-regular. For all $1 \leq i \leq s'$, let $S_i \subseteq V_i$, with $|S_i| \leq \epsilon t - 2$.

Suppose that $v_i \in V(M)$ and that x_1 and x_1' are not necessarily distinct vertices lying in $V_i \setminus S_i$. Then, for all $2 \leq \ell \leq \left\lceil\left(1 - \epsilon - \frac{1}{4}\eta\right) st\right\rceil$, the graph G contains a path of length 2ℓ from x_1 to x_1', which does not contain any vertex in $\bigcup_{i=1}^{s'} S_i$.

Note that, setting $x_1 = x_1'$, this lemma implies that, for all $2 \leq \ell \leq \left\lceil\left(1 - \epsilon - \frac{1}{4}\eta\right) st\right\rceil$, the graph G contains a cycle of length 2ℓ.

Proof. Let $\epsilon = \min\left\{\frac{1}{4}\epsilon_0\left(\frac{1}{8}\eta, 2, 2, \frac{1}{8}\eta\right), \frac{1}{18}\eta, \frac{1}{144}\eta^2\right\}$, where ϵ_0 is as defined in the Blow-Up Lemma. It may be assumed that $E(M) = \{v_{2i-1,2i} : 1 \leq i \leq s\}$ and that x_1 and x_1' lie in V_1. For all $1 \leq i \leq s'$, let V_i' be a subset of $V_i \setminus (S_i \cup \{x_1, x_1'\})$ with order $\lceil (1 - \epsilon) t \rceil$. Let G' be the restriction of G to the set $\bigcup_{i=1}^{s'} V_i'$.

As H is connected, it contains a path P_j joining $\{v_{2j-1}, v_{2j}\}$ to $\{v_1, v_2\}$, for all $2 \leq j \leq s$. Note that it may be assumed that v_{2j-1} is one endpoint of

P_j. If v_2 is the other endpoint, then extend P_j using the edge v_1v_2, so that it may always be assumed that P_j runs from v_1 to v_{2j-1}.

For a path $P = v_{i_0}v_{i_1}\cdots v_{i_p}$ in H, a path $Q = x_{i_0}x_{i_1}\cdots x_{i_q}$ in G' is said to *agree* with P if $p = q$ and, for all $0 \leq l \leq p$, the vertex x_{i_l} is contained in V'_{i_l}. The proof continues by constructing vertex-disjoint paths $Q_2, Q'_2, \ldots, Q_s, Q'_s$ in G' such that, for all $2 \leq j \leq s$, the paths Q_j and Q'_j both agree with P_j.

Indeed, suppose that such paths $Q_2, Q'_2, \ldots, Q_{j-1}, Q'_{j-1}$ have been constructed for some $2 \leq j \leq s$. Each path uses each set V'_i at most once. Hence each set V'_i corresponding to a vertex v_i of P_j contains at most $2s \leq \frac{1}{4}\eta t$ vertices in $\bigcup_{2\leq l\leq j-1}(V(Q_l) \cup V(Q'_l))$.

Recall that, for each edge of P_j, the corresponding bipartite graph in G is 2ϵ-regular with density at least $\frac{1}{2}\eta$. Hence Lemma 2.8 implies that, after removing the sets $\bigcup_{2\leq l\leq j-1}(V(Q_l) \cup V(Q'_l))$ and $V_i \setminus V'_i$ from the sets V_i, this bipartite graph remains 4ϵ-regular with density at least $\frac{1}{2}\eta - 2\epsilon$. Hence the path Q_j may be obtained by applying Theorem 2.12 to the graph P_j. Note that this requires that $\epsilon \leq \frac{\left(\frac{1}{2}\eta-6\epsilon\right)^2}{4}$ and $\epsilon \leq \frac{\frac{1}{2}\eta-6\epsilon}{3}$, which are both true by the choice of ϵ. Repeating the argument, the path Q'_j may be obtained in the same way. For $2 \leq j \leq s$, let $x_j \in V'_1$ and $y_j \in V'_{2j-1}$ be the endpoints of Q_j and $x'_j \in V'_1$ and $y'_j \in V'_{2j-1}$ be the endpoints of Q'_j. Hence G has the structure illustrated in Figure 2.2.

For $1 \leq i \leq 2s$, let W_i be the vertices of V'_i, which are not vertices of Q_j or Q'_j for any $2 \leq j \leq s$. Note that $|W_i| \geq |V'_i| - 2s \geq \left(1 - \epsilon - \frac{1}{4}\eta\right)t$. Deleting vertices where appropriate, it may be assumed that each W_i has order $t' = \left\lceil\left(1 - \epsilon - \frac{1}{4}\eta\right)t\right\rceil$. Note that, for all $1 \leq i \leq s$, the graph $G'[W_{2i-1}, W_{2i}]$ is $\left(4\epsilon, \frac{1}{4}\eta - \epsilon\right)$-super-regular by Lemma 2.8, and hence is $\left(\epsilon_0, \frac{1}{8}\eta\right)$-super-regular,

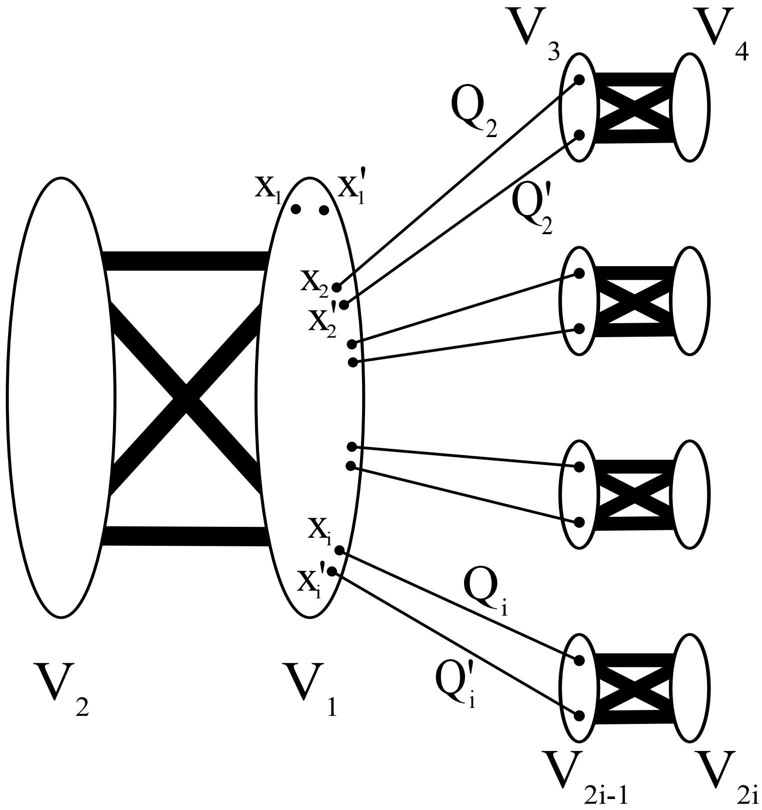


Figure 2.2: Super-regular pairs are shown with thick lines and paths Q_j and Q'_j with thin lines.

where $\epsilon_0 = \epsilon_0\left(\frac{1}{8}\eta, 2, 2, \frac{1}{8}\eta\right)$, as defined in the Blow-Up Lemma.

For $1 \leq i \leq s$, and any vertex $z \in V_{2i-1}$, let C_z be the neighbourhood of z in W_{2i}. Fix $1 \leq \ell \leq t' - 1$. In the complete bipartite graph from W_{2i-1} to W_{2i}, there is a path of length 2ℓ between C_{y_i} and $C_{y'_i}$. As $G\left[V_{2i-1}, V_{2i}\right]$ is $\left(2\epsilon, \frac{1}{2}\eta\right)$-super-regular, both y_i and y'_i have at least $\frac{1}{2}\eta t$ neighbours in V_{2i}. Hence, as $|V_{2i} \setminus W_{2i}| \leq \left(\epsilon + \frac{1}{4}\eta\right)t$, both C_{y_i} and $C_{y'_i}$ have order at least $\frac{1}{8}\eta t'$.

The Blow-Up Lemma thus implies that $G'\left[W_{2i-1}, W_{2i}\right]$ contains a path of length 2ℓ between C_{y_i} and $C_{y'_i}$. Extend these paths to paths between x_i and $x_{i'}$ using Q_i and Q'_i. Hence, for all $2 \leq i \leq r$ and $1 \leq \ell \leq t' - 1$, there is a path of length $2(\ell + |V(P_i)|)$ between x_i and $x_{i'}$ in G', which is contained wholly in $W_{2i-1} \cup W_{2i} \cup V(Q_i) \cup V(Q_i')$. This is illustrated in Figure 2.3.

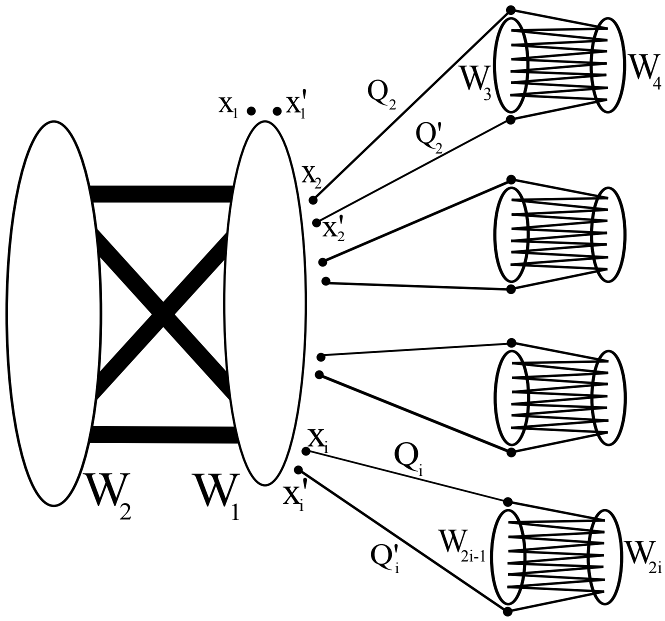


Figure 2.3

Fix $1 \leq i \leq s$. Let $i \leq \ell_1 \leq t' - i$. In the complete bipartite graph from W_1 to W_2, there are paths $Q_1'', Q_2'', \ldots, Q_i''$ such that Q_i'' is a path of length $2(\ell_1 - (i-1))$ between C_{x_i} and $C_{x_1'}$ and, for $1 \leq j \leq i-1$, the path Q_j'' is a path of length two between C_{x_j} and $C_{x_{j+1}'}$. Note that there are at most $2s$ sets of the form C_{x_j} or $C_{x_{j'}}$. However, $t' \geq \frac{1}{2}t \geq \frac{2s}{\alpha}$, where $\alpha = \alpha\left(\frac{1}{8}\eta, 2, 2, \frac{1}{8}\eta\right)$ is as defined in the Blow-Up Lemma. Hence the Blow-Up Lemma implies that such paths $Q_1'', Q_2'', \ldots, Q_i''$ exist in $G'[W_1, W_2]$.

By extending both ends of each of the paths $Q_1'', Q_2'', \ldots, Q_i''$ by a single edge, paths $Q_1^*, Q_2^*, \ldots, Q_i^*$ can be obtained such that:

- for all $1 \leq j \leq i-1$, Q_j^* is a path of length four from x_j to x'_{j+1} and is otherwise contained in $W_1 \cup W_2$;

- Q_1^* is a path of length $2(\ell_1 + 2 - i)$ from x_i to x'_1 and is otherwise contained in $W_1 \cup W_2$.

For all $2 \leq j \leq i$, let $1 \leq \ell_j \leq t' - 1$. Then, for all $2 \leq j \leq i$, there is a path of length $2(\ell_j + |V(P_j)|)$ between x_j and $x_{j'}$ in G', which is contained wholly in $W_{2j-1} \cup W_{2j} \cup V(Q_j) \cup V(Q'_j)$. Concatenating these paths with the paths $Q_1^*, Q_2^*, \ldots, Q_i^*$ gives a path between x_1 and x'_1. This has length 2ℓ, where

$$\ell = \ell_1 + 2 - i + 2(i-1) + \sum_{j=2}^{i} (\ell_j + |V(P_j)|) = i + \sum_{j=2}^{i} |V(P_j)| + \sum_{j=1}^{i} \ell_j.$$

This holds for any $2i - 1 \leq \sum_{j=1}^{i} \ell_j \leq i(t'-2)+1$. Hence, for all $1 \leq i \leq s$, the graph G contains a path of length 2ℓ from x_1 to x'_1, for all $\ell \in [L_i, L'_i]$, where

$$L_i = 3i - 1 + \sum_{j=2}^{i} |V(P_j)|$$

and

$$L'_i = i(t'-1) + \sum_{j=2}^{i} |V(P_j)| + 1.$$

However $t' \geq \frac{1}{2}t \geq 6s'$ implies that, for any $1 \leq i \leq s-1$,

$$4i + 2 + |V(P_{i+1})| \leq 6st',$$

and so $L'_{i+1} \leq L_i$. Hence G contains a path of length 2ℓ from x_1 to x'_1, for all $L_1 \leq \ell \leq L'_s$. However, $L_1 = 2$ and

$$L'_s = s\left(t' - 1\right) + \sum_{j=2}^{s} \left|V\left(P_j\right)\right| + 1 \geq st' \geq \left(1 - \epsilon - \frac{1}{4}\eta\right) st.$$

Hence, as L'_s is an integer, the proof is complete. $\square$

The previous lemma requires that the edges of the matching in H correspond to super-regular pairs in G. However, if H is the reduced graph obtained having applied the Regularity Lemma to a graph G, all edges of H will correspond to regular pairs in G. In this case, a partition with the properties required by Lemma 2.23 can be obtained by removing a small proportion of the vertices in each cluster. This process is used frequently throughout both this and the following chapter, and so is formalised in the following lemma.

Lemma 2.24. *Let $0 < \epsilon < \frac{1}{2}\eta < \frac{1}{2}$. Let H be a connected graph with vertex set $\{v_1, \ldots, v_{s'}\}$ and suppose that M is a maximal matching of H with edges $\{v_{2i-1}v_{2i} : 1 \leq i \leq s\}$. Let t be a positive integer.*

Replace the vertices of H with pairwise disjoint sets $V_1, \ldots, V_{s'}$ of order t. Let G be a graph defined by replacing each edge v_iv_j of H with an ϵ-regular pair between V_i and V_j of density at least η. Then there are sets $(W_i)_{i=1}^{s'}$ such that the following holds.

- *For all $1 \leq i \leq s'$, W_i is a subset of V_i of order $t' = \lceil (1 - \epsilon)\, t \rceil$.*

- *For all edges v_iv_j of H, the graph $G\left[W_i, W_j\right]$ is 2ϵ-regular with density at least $\eta - \epsilon$.*

- *For all edges $v_i v_j$ of M, the graph $G[W_i, W_j]$ is $(2\epsilon, \eta - 2\epsilon)$-super-regular.*

- *For all $1 \le i \le s'$, there is some $1 \le j \le 2s$ such that $v_j \in V(M)$, and any vertex in W_i has at least $(\eta - 2\epsilon)\, t'$ neighbours in W_j.*

Proof. For $1 \le i \le s$, let W_{2i-1} be the set of vertices of V_{2i-1} with at least $(\eta - \epsilon)\, t$ neighbours in V_{2i} and W_{2i} be the set of vertices of V_{2i-1} with at least $(\eta - \epsilon)\, t$ neighbours in V_{2i}. Note that, for $1 \le i \le r$

$$\begin{aligned} d_G\left(V_{2i} \setminus W_{2i}, V_{2i-1}\right) &= \frac{e_G\left(V_{2i} \setminus W_{2i}, V_{2i-1}\right)}{|V_{2i-1}||V_{2i} \setminus W_{2i}|} \\ &= \frac{\sum_{v \in V_{2i} \setminus W_{2i}} \deg_{V_{2i-1}}(v)}{t|V_{2i} \setminus W_{2i}|} \\ &< \eta - \epsilon. \end{aligned}$$

Hence, as $G[V_{2i}, V_{2i-1}]$ is ϵ-regular with density at least η, it follows that $|W_{2i}| \ge (1-\epsilon)|V_{2i}|$. Similarly $|W_{2i-1}| \ge (1-\epsilon)|V_{2i-1}|$, for all $1 \le i \le r$.

For any vertex $v_i \in V(H) \setminus V(M)$, there is a path in H from v_i to $V(M)$, which does not contain any edges of M. As M is maximal, this path cannot contain an edge which has no endpoint in $V(M)$ and hence v_i has a neighbour $v_j \in V(M)$. Let W_i be the set of vertices of V_i with at least $(\eta - \epsilon)\, t$ neighbours in W_j. By the previous argument, the density between W_j and $V_i \setminus W_i$ is less than $\eta - \epsilon$ and hence $|W_i| \ge (1-\epsilon)\, t$. By deleting vertices as necessary, it may be assumed that $|W_i| = t' = \lceil (1-\epsilon)t \rceil$ for all $1 \le i \le s'$.

Fix $1 \le i \le s$. Let W'_{2i-1} and W'_{2i} be subsets of W_{2i-1} and W_{2i} respectively, with $|W'_{2i-1}| \ge 2\epsilon|W_{2i-1}| \ge \epsilon|V_{2i-1}|$ and $|W'_{2i}| \ge 2\epsilon|W_{2i}| \ge \epsilon|V_{2i}|$.

Then, as $G[V_{2i}, V_{2i-1}]$ is ϵ-regular with density at least η,

$$e\left(W'_{2i-1}, W'_{2i}\right) > (\eta - \epsilon)\,|W'_{2i-1}||W'_{2i}|.$$

Also, any vertex in W_{2i-1} has at least $(\eta - \epsilon)\,t$ neighbours in V_{2i} and so at least $(\eta - 2\epsilon)\,t$ neighbours in W_{2i} (and vice versa). Hence $G\left[W_{2i-1}, W_{2i}\right]$ is $(2\epsilon, \eta - 2\epsilon)$-super-regular.

Note that Lemma 2.8 implies that, for any edge $v_i v_j$ of H, the graph $G[W_i, W_j]$ is 2ϵ-regular with density at least $\eta - \epsilon$. The final condition is immediate from the construction of w_i. □

These two lemmas have the following immediate corollary.

Corollary 2.25. *Given $\eta > 0$, there exists $\epsilon = \epsilon(\eta)$ such that the following holds. Let H be a connected graph of order s' containing a matching M on at least $2s$ vertices. Let t be a positive integer satisfying*

$$(1-\epsilon)\,t \geq \max\left\{\frac{8s}{\eta}, 12s', \frac{4s}{\alpha\left(\frac{1}{8}\eta, 2, 2, \frac{1}{8}\eta\right)}\right\}$$

where α is as defined in the Blow-Up Lemma.

Replace the vertices $\{v_i : 1 \leq i \leq s\}$ of H with pairwise disjoint sets $V_1, \ldots, V_s$ of order t. Let G be a graph defined by replacing each edge $v_i v_j$ of H with an ϵ-regular pair between V_i and V_j of density at least η. Then, for all $2 \leq \ell \leq \left\lceil (1-\epsilon)\left(1-\epsilon-\frac{1}{4}\eta\right) st \right\rceil$, G contains a cycle of length 2ℓ.

2.5 Matchings in 2-edge-coloured graphs

For an r-edge-multi-coloured graph G with r-edge-multi-colouring $E(G) = \bigcup_{i=1}^{r} E(G_i)$, a *monochromatic component* of G is a component of one of the subgraphs G_i. The lemmas of the previous section show that, when attempting to find cycles in G, it is useful to look for a monochromatic component of the reduced graph which contains a large matching. The following corollary of Hall's Theorem (Theorem 2.14) proves that a graph with large minimal degree contains large matchings between the parts of any partition of the vertex set into two parts.

Lemma 2.26. *Let $\delta > 0$ and let G be a graph of order n with $\delta(G) \geq \left(\frac{3}{4} - \delta\right) n$. Suppose that $V(G) = V_1 \cup V_2$ is a partition of the vertices of G, with $|V_1| \leq |V_2|$. Then G contains a matching with at least* $\min\left\{|V_1|, \left(\frac{1}{2} - 2\delta\right) n\right\}$ *edges between V_1 and V_2.*

Proof. For a subset S of V_1, a lower bound on $|\Gamma_{V_2}(S)|$ is obtained by splitting into the cases that $|S| > \left(\frac{1}{4} + \delta\right) n$ and $|S| \leq \left(\frac{1}{4} + \delta\right) n$.

If $S \subseteq V_1$ and $|S| > \left(\frac{1}{4} + \delta\right) n$, consider a vertex $v \in V_2$. As v has at most $\left(\frac{1}{4} + \delta\right) n$ non-neighbours in G, it has a neighbour in S. Hence $\Gamma_{V_2}(S) = V_2$, and so $|\Gamma_{V_2}(S)| = |V_2| \geq |V_1| \geq |S|$.

If $S \subseteq V_1$ and $|S| \leq \left(\frac{1}{4} + \delta\right) n$, then any vertex in S has at least $|V_2| -$

$\left(\frac{1}{4}+\delta\right)n$ neighbours in V_2. Hence

$$\begin{aligned} |\Gamma_{V_2}(S)| &\geq |V_2| - \left(\frac{1}{4}+\delta\right)n \\ &= |S| - \left(\left(\frac{1}{4}+\delta\right)n + |S| - |V_2|\right) \\ &\geq |S| - \left(\left(\frac{1}{2}+2\delta\right)n - |V_2|\right). \end{aligned}$$

Thus Theorem 2.14 implies that G contains a matching with at least

$$|V_1| - \max\left\{0, \left(\left(\frac{1}{2}+2\delta\right)n - |V_2|\right)\right\}$$

edges between V_1 and V_2. As $|V_1|+|V_2| = n$, this is the required matching. $\square$

The following definition uses Tutte's Theorem to force a graph without a matching to have a particular structure. Recall that, for any graph G, the number of odd components of G is denoted by $q\,(G)$.

Definition 2.27. *Let a and b be integers and G be a graph, which does not contain a matching on at least b vertices.*

- *Choose $S_G \subseteq V(G)$ to be a set such that*

$$q\,(G\,[V(G) - S_G]) > |S_G| + |V(G)| - b.$$

Note that as G does not contain a matching on at least b vertices, such a set exists by Theorem 2.16. If $|V\,(G)\,| < b$, then it will be assumed that S_G is empty.

- *For all $i \geq 1$, let $T_{G,i}$ be the set of vertices of G which lie in components of $G\left[V(G) - S_G\right]$ of order i.*

- *Let $T_G = \bigcup_{1 \leq i \leq a} T_{G,i}$.*

- *Let $U_G = V(G) \setminus (S_G \cup T_G)$.*

Note that, although the definitions of S_G, T_G and U_G do depend on the values a and b, in practice a and b will be fixed throughout any use of this notation.

The following result about the sets just defined is used not only in this chapter, but also in Chapter 3.

Lemma 2.28. *Let a and b be integers and G be a graph, which does not contain a matching on at least b vertices. Then*

$$\begin{aligned}
|S_G| &< \frac{1}{2}b, \\
|T_G| &> \left(1 - \frac{1}{a}\right)|V(G)| + |S_G| - b,
\end{aligned}$$

and

$$2|S_G| + |U_G| < b + \frac{1}{a}|V(G)|.$$

Furthermore, $G\left[V(G) \setminus S_G\right]$ has maximum degree at most $3\left(\frac{1}{2}b - |S_G|\right)$.

Proof. It is immediate that

$$q\left(G\left[V(G) - S_G\right]\right) \leq |V(G)| - |S_G|$$

and hence the definition of S_G implies that

$$|V(G)| - |S_G| > |S_G| + |V(G)| - b.$$

This implies the first equation.

There are at most $|T_G|$ components of $G\left[V(G) - S_G\right]$ of order at most a. However, there are at most $\frac{1}{a}|V(G)|$ components of $G\left[V(G) - S_G\right]$ of order at least a. Hence

$$q\left(G\left[V(G) - S_G\right]\right) \leq |T_G| + \frac{1}{a}|V(G)|.$$

This implies that

$$\begin{aligned} |T_G| &\geq q\left(G\left[V(G) - S_G\right]\right) - \frac{1}{a}|V(G)| \\ &> \left(1 - \frac{1}{a}\right)|V(G)| + |S_G| - b \end{aligned}$$

and so the second equation holds.

However, $|V(G)| = |S_G| + |T_G| + |U_G|$, which implies that

$$\begin{aligned} 2|S_G| + |U_G| &= |S_G| + |V(G)| - |T_G| \\ &< b + \frac{1}{a}|V(G)|. \end{aligned}$$

For all $1 \leq i \leq n$, there are exactly $\frac{1}{i}|T_{G,i}|$ components of order i in

$G[V(G) \setminus S_G]$. Hence, by the definition of S_G,

$$\begin{aligned}\sum_{i\geq 1} \frac{1}{2i-1}|T_{G,2i-1}| &= q\left(G\left[V(G) - S_G\right]\right) \\ &> |S_G| + |V(G)| - b.\end{aligned}$$

However,

$$\begin{aligned}\sum_{i\geq 1} \frac{1}{2i-1}|T_{G,2i-1}| &\leq |T_{G,1}| + \frac{1}{3}\sum_{i\geq 2}|T_{G,2i-1}| \\ &\leq |T_{G,1}| + \frac{1}{3}\left(|V(G)| - |S_G| - |T_{G,1}|\right).\end{aligned}$$

Combining these inequalities gives

$$\frac{2}{3}|T_{G,1}| > \frac{4}{3}|S_G| + \frac{2}{3}|V(G)| - b,$$

and so

$$|T_{G,1}| > 2|S_G| + |V(G)| - \frac{3}{2}b. \tag{2.2}$$

However, $T_{G,1}$ is a set of isolated vertices in $G[V(G) \setminus S_G]$. Hence, any vertex in $V(G) \setminus S_G$ has degree in $G[V(G) \setminus S_G]$ at most $|V(G) \setminus (S_G \cup T_{G,1})|$. Hence (2.2) implies that

$$\Delta\left(G\left[V(G) \setminus S_G\right]\right) \leq |V(G)| - |T_{G,1}| - |S_G| < 3\left(\frac{1}{2}b - |S_G|\right)$$

completing the proof of Lemma 2.28. $\square$

The following lemma about the component structure of a 2-edge-multi-coloured graph with large minimal degree is also useful.

Lemma 2.29. *Let $0 < \delta < \frac{1}{12}$ and let G be a graph of sufficiently large order n with $\delta(G) \geq \left(\frac{3}{4} - \delta\right) n$. Suppose that G has a 2-edge-multi-colouring $E(G) = E(R) \cup E(B)$. Let R' be the largest component of R and B' be the largest component of B. Also let $U_1 = V(R') \cap V(B')$, $U_2 = V(R') \setminus V(B')$, $U_3 = V(B') \setminus V(R')$ and $U_4 = V(G) \setminus (V(R') \cup V(B'))$. Then one of the following holds.*

- *One of R or B is connected.*
- *Both R' and B' have order at least $\left(\frac{3}{4} - \delta\right) n$ and $V(G) = V(R') \cup V(B')$.*
- *Each set U_i has order at most $\left(\frac{1}{4} + \delta\right) n$.*

Proof. Suppose that none of the above holds. Then, in particular, both R and B are disconnected.

Suppose first that $V(G) = V(R') \cup V(B')$. As neither R nor B is connected, it must be the case that $V(R') \setminus V(B')$ is non-empty. Let $v \in V(R') \setminus V(B')$. Note that v has no neighbours of either colour in $V(B') \setminus V(R')$. Hence $|V(B') \setminus V(R')| \leq \left(\frac{1}{4} + \delta\right) n$ which implies that $|R'| \geq \left(\frac{3}{4} - \delta\right) n$. A similar argument shows that $|B'| \geq \left(\frac{3}{4} - \delta\right) n$, contradicting the assumption that none of the three cases hold. Hence it may be assumed that $V(G) \setminus (V(R') \cup V(B'))$ is non-empty.

Suppose now that $|R'| \leq \left(\frac{5}{12} - \delta\right) n$. Then $\Delta(R) < \left(\frac{5}{12} - \delta\right) n$ and hence $\delta(B) > \frac{1}{3} n$. Since B is disconnected, it has exactly two components $B_1 = B'$

and B_2 with $\frac{1}{3}n < |B_2| \leq |B_1| < \frac{2}{3}n$. Note that $U_i = V(B_i) \cap V(R')$, for $i \in \{1,2\}$.

Suppose that $U_i \neq \emptyset$, for some $i \in \{1,2\}$, and let $v \in U_i$. Then v has no neighbours outside $R' \cup B_i$. Hence, as $U_{3-i} = R' \setminus B_i$,

$$|U_{3-i}| \geq |\Gamma_G(v)| - |B_i| > \left(\frac{1}{12} - \delta\right) n. \tag{2.3}$$

In particular $U_{3-i} \neq \emptyset$, and so U_1 is non-empty if and only if U_2 is non-empty. As $V(R') = U_1 \cup U_2$, this implies that both U_1 and U_2 are non-empty.

However, (2.3) implies that U_i has order at least $\left(\frac{3}{4} - \delta\right) n - |B_i|$, for $i \in \{1,2\}$. Hence R' has order at least $\left(\frac{3}{2} - 2\delta\right) n - |B_1| - |B_2| = \left(\frac{1}{2} - 2\delta\right) n$. This contradicts the assumption that $|R'| \leq \left(\frac{5}{12} - \delta\right) n$. It may therefore be assumed that $|R'| > \left(\frac{5}{12} - \delta\right) n$, and similarly $|B'| > \left(\frac{5}{12} - \delta\right) n$.

If some U_i is non-empty, let $x \in U_i$. The graph G contains no edges between x and U_{5-i} and hence $|U_{5-i}| \leq \left(\frac{1}{4} + \delta\right) n$. Recall that it may be assumed that U_4 is non-empty and hence $|U_1| \leq \left(\frac{1}{4} + \delta\right) n$. However both R' and B' have order at least $\left(\frac{5}{12} - \delta\right) n$ and hence both U_2 and U_3 are non-empty. This implies that both U_2 and U_3 have order at most $\left(\frac{1}{4} + \delta\right) n$ and so U_1 is non-empty. Finally, this in turn implies that U_4 has order at most $\left(\frac{1}{4} + \delta\right) n$, contradicting the assumption that the third case of the lemma does not hold. This completes the proof. $\square$

Note that the Regularity Lemma implies that the minimal degree of the reduced graph H is not too much smaller than $\frac{|H|}{|G|}$ times the minimal degree of G. In view of this, and the results of Section 2.5, the following lemma can be used to prove Theorem 2.4.

Lemma 2.30. *Let $0 < \delta < \frac{1}{36}$. There exists $n_0 = n_0(\delta)$ such that the following holds.*

Let G be a graph of order $n \geq n_0$ with $\delta(G) \geq \left(\frac{3}{4} - \delta\right) n$. Suppose that G has a 2-edge-multi-colouring $E(G) = E(R) \cup E(B)$. Then one of the following holds.

(i) There is a set S of order at least $\left\lfloor \left(\frac{2}{3} - \frac{1}{2}\delta\right) n \right\rfloor$ such that either $R[S]$ or $B[S]$ has maximal degree at most $11\delta n$.

(ii) There is a component of R or B which contains a matching on at least $\left(\frac{2}{3} + \delta\right) n$ vertices. Further, there is a monochromatic component of G containing a matching on at least $\left(\frac{1}{2} + \delta\right) n$ vertices such that either this component contains an odd cycle, or this component contains all but at most $5\delta n$ vertices of $|V(G)|$.

(iii) There is a partition $V(G) = U_1 \cup \cdots \cup U_4$ with $\max_i |U_i| \leq \left(\frac{1}{4} + \delta\right) n$ such that there are no red edges from $U_1 \cup U_2$ to $U_3 \cup U_4$ and no blue edges from $U_1 \cup U_3$ to $U_2 \cup U_4$. Further, the bipartite graphs $R[U_1, U_2]$, $R[U_3, U_4]$, $B[U_1, U_3]$ and $B[U_2, U_4]$ all contain a matching on at least $\left(\frac{1}{2} - 6\delta\right) n$ vertices.

Proof. Let $0 < \delta < \frac{1}{36}$. Assume throughout that n is sufficiently large. Suppose that G is a graph of order n with $\delta(G) \geq \left(\frac{3}{4} - \delta\right) n$ and that G has a 2-edge-multi-colouring $E(G) = E(R) \cup E(B)$. Let R' be the largest component of R and B' be the largest component of B. Also let $U_1 = V(R') \cap V(B')$, $U_2 = V(R') \setminus V(B')$, $U_3 = V(B') \setminus V(R')$ and $U_4 = V(G) \setminus (V(R') \cup V(B'))$, as in Lemma 2.29.

Suppose first that $\max_i |U_i| \leq \left(\frac{1}{4}+\delta\right)n$. Note that there are no red edges from $U_1 \cup U_2$ to $U_3 \cup U_4$ and no blue edges from $U_1 \cup U_3$ to $U_2 \cup U_4$. By Lemma 2.26, G contains a matching from $U_1 \cup U_2$ to $U_3 \cup U_4$ with at least $\min\left\{|U_1|+|U_2|, |U_3|+|U_4|, \left(\frac{1}{2}-2\delta\right)n\right\}$ edges. However, $|U_i|+|U_j| \geq \left(\frac{1}{2}-2\delta\right)n$ for all $1 \leq i < j \leq 4$.

All edges from $U_1 \cup U_2$ to $U_3 \cup U_4$ are blue and hence B contains a matching M from $U_1 \cup U_2$ to $U_3 \cup U_4$ with at least $\left(\frac{1}{2}-2\delta\right)n$ edges. All edges of M lie between U_1 and U_3 or between U_2 and U_4. Both $B[U_1, U_3]$ and $B[U_2, U_4]$ contain at most $\left(\frac{1}{4}+\delta\right)n$ edges of M, and hence $B[U_1, U_3]$ and $B[U_2, U_4]$ both contain a matching on at least $\left(\frac{1}{2}-6\delta\right)n$ vertices. A similar argument shows that $R[U_1, U_2]$ and $R[U_3, U_4]$ both contain a matching on at least $\left(\frac{1}{2}-6\delta\right)n$ vertices.

Hence, if $\max_i |U_i| \leq \left(\frac{1}{4}+\delta\right)n$, then the third case of Lemma 2.30 is satisfied. Hence, for the remainder of the proof it is assumed that

$$\max_i |U_i| > \left(\frac{1}{4}+\delta\right)n. \tag{2.4}$$

The rest of the proof is split into two cases, depending on whether one of R' or B' contain a matching on at least $\left(\frac{2}{3}+\delta\right)n$ vertices.

Case 1 - neither R' nor B' contains a matching on at least $\left(\frac{2}{3}+\delta\right)n$ vertices.

The aim is to show that the first case of Lemma 2.30 holds and so there is a large set on which one of R or B has a low maximum degree.

Let $S_{R'}$, $T_{R'}$, $U_{R'}$, $S_{B'}$, $T_{B'}$ and $U_{B'}$ be defined using Definition 2.27, with

$a = \lceil \delta^{-1} \rceil$ and $b = \left\lceil \left(\frac{2}{3} + \delta \right) n \right\rceil$. Then Lemma 2.28 implies that

$$|S_{R'}| < \left\lceil \left(\frac{1}{3} + \frac{1}{2}\delta \right) n \right\rceil, \tag{2.5}$$
$$|T_{R'}| > |V(R')| + |S_{R'}| - \left\lceil \left(\frac{2}{3} + 2\delta \right) n \right\rceil - 1, \tag{2.6}$$

and

$$2|S_{R'}| + |U_{R'}| < \left\lceil \left(\frac{2}{3} + 2\delta \right) n \right\rceil + 1. \tag{2.7}$$

Similar equations hold for $S_{B'}$, $T_{B'}$ and $U_{B'}$. The following fact will be sufficient to complete the proof in this case.

Fact 2.31. *Suppose that C is a component of B with $|V(C)| \leq \left(\frac{5}{12} - 2\delta \right) n$. Then $V(C) \subseteq S_{R'}$. Similarly, if C is a component of R with $|V(C)| \leq \left(\frac{5}{12} - 2\delta \right) n$, then $V(C) \subseteq S_{B'}$*

Proof. The proof of both statements is identical, so assume that C is a component of B with $|V(C)| \leq \left(\frac{5}{12} - 2\delta \right) n$. This implies that B is not connected. Hence Lemma 2.29 and the assumption that (2.4) holds implies that R' contains at least $\left(\frac{3}{4} - \delta \right) n$ vertices.

A vertex in C has blue degree at most $|V(C)| - 1$. Hence any vertex in C must have red degree at least

$$\begin{aligned} \delta(G) - |V(C)| + 1 &\geq \left(\frac{3}{4} - \delta \right) n - \left(\frac{5}{12} - 2\delta \right) n + 1 \\ &= \left(\frac{1}{3} + \delta \right) n + 1. \end{aligned} \tag{2.8}$$

A vertex in $V(G) \setminus V(R')$ has red degree at most

$$n - |V(R')| - 1 \leq \left(\frac{1}{4} + \delta\right) n - 1.$$

However, a vertex in $T_{R'}$ has at most a red neighbours outside of the set $S_{R'}$ and hence (2.5) implies that, for all $v \in T_{R'}$,

$$\begin{aligned} \deg_R(v) &\leq a + |S_{R'}| \\ &< \left\lceil \left(\frac{1}{3} + \frac{1}{2}\delta\right) n \right\rceil + a. \end{aligned}$$

Hence (2.8) implies that $V(C)$ does not intersect $T_{R'}$ or $V(G) \setminus V(R')$.

Suppose that there exists a vertex v in $V(C) \cap U_{R'}$. All blue neighbours of v lie in $V(C) \subseteq S_{R'} \cup U_{R'}$. However, by definition, a vertex in $U_{R'}$ has all of its red neighbours in $S_{R'} \cup U_{R'}$. Hence, in G, all of the neighbours of v are in $S_{R'} \cup U_{R'}$, and so (2.7) implies that v has degree in G less than $\left\lceil \left(\frac{2}{3} + 2\delta\right) n \right\rceil + 1$. This contradicts $\delta(G) \geq \left(\frac{3}{4} - \delta\right) n$, and so $V(C) \subseteq S_{R'}$. □

Using this fact, the proof may now be completed in Case 1. Suppose first that $|V(B')| < \left(\frac{2}{3} + \delta\right) n$. Then Lemma 2.29 and the assumption that (2.4) holds implies that R is connected.

As B' is the largest component of B, if $|V(B')| \leq \left(\frac{5}{12} - 2\delta\right) n$, then all components of B have order at most $\left(\frac{5}{12} - 2\delta\right) n$. Fact 2.31 implies that $S_{R'}$ contains all components of B, and hence has order n, contradicting (2.5).

If, however, $\left(\frac{5}{12} - 2\delta\right) n < |V(B')| \leq \left(\frac{2}{3} - \frac{1}{2}\delta\right) n$, then Lemma 2.26 implies that G contains a matching between $V(B')$ and $V(G) \setminus V(B')$ with at least $\left(\frac{1}{3} + \frac{1}{2}\delta\right) n$ edges. As all edges between $V(B')$ and $V(G) \setminus V(B')$ are red, this implies that R' contains a matching on at least $\left(\frac{2}{3}+\delta\right)n$ vertices, which was assumed not to occur.

Finally, if $\left(\frac{2}{3}-\frac{\delta}{2}\right)n < |B'| < \left(\frac{2}{3}+\delta\right)n$, then all components of B other than B' have order at most $\left(\frac{1}{3}+\frac{\delta}{2}\right)n$. Hence, Fact 2.31 implies that $S_{R'}$ contains $V(G)\setminus V(B')$ and so $|S_{R'}| > \left(\frac{1}{3}-\delta\right)n$. Let $S = V(G)\setminus S'_R$. Then (2.5) implies that $|S| \geq \left\lfloor\left(\frac{2}{3}-\frac{1}{2}\delta\right)\right\rfloor$. Hence, as $R' = R$, Lemma 2.28 implies that the first case of Lemma 2.30 holds.

Hence it may be assumed that both $V(B')$ and $V(R')$ contain at least $\left(\frac{2}{3}+\delta\right)n$ vertices. Suppose that $T_{R'}\cap T_{B'}$ is non-empty and let $v \in T_{R'}\cap T_{B'}$. Then v has at most $|S_{R'}|+a$ red neighbours and at most $|S_{B'}|+a$ blue neighbours. However, (2.5) implies that both $S_{R'}$ and $S_{B'}$ have order at most $\left\lceil\left(\frac{1}{3}+\frac{\delta}{2}\right)n\right\rceil$, and so v has at most $\left(\frac{2}{3}+\delta\right)n+2a+2$ neighbours in G. This contradicts the minimal degree of G, and hence $T_{R'}\cap T_{B'}$ must be empty. Fact 2.31 implies that $T_{B'}$ does not intersect $V(G)\setminus V(R')$ and vice versa. Hence $T_{R'}\subseteq S_{B'}\cup U_{B'}$ and $T_{B'}\subseteq S_{R'}\cup U_{R'}$.

Suppose that $T_{B'}\cap U_{R'}$ is non-empty and let $v\in T_{B'}\cap U_{R'}$. Then all of the red neighbours of v are contained in $S_{R'}\cup U_{R'}$. However, v has at most $\left(\frac{1}{4}+\delta\right)n$ non-neighbours, and hence v has at least $n-|S_{R'}\cup U_{R'}|-\left(\frac{1}{4}+\delta\right)n$ neighbours outside of $S_{R'}\cup U_{R'}$. Thus v has at least $n-|S_{R'}\cup U_{R'}|-\left(\frac{1}{4}+\delta\right)n$ blue neighbours. However $v\in T_{B'}$ and so all but a of its blue neighbours are in $S_{B'}$. Hence

$$|S_{B'}| \geq n-(|S_{R'}|+|U_{R'}|)-\left(\frac{1}{4}+\delta\right)n-a > |S_{R'}|+\left(\frac{1}{12}-3\delta\right)n-a-2,$$

where the second inequality follows from (2.7).

Similarly, if $T_{R'} \cap U_{B'}$ is non-empty, then

$$|S_{R'}| > |S_{B'}| + \left(\frac{1}{12} - 3\delta\right) n - a - 2.$$

As these can not both occur, one of $T_{B'} \cap U_{R'}$ or $T_{R'} \cap U_{B'}$ is empty.

Assume without loss of generality that $T_{B'} \subseteq S_{R'}$. Then $S_{R'}$ contains the disjoint sets $T_{B'}$ and $V(G) \setminus V(B')$. Hence, (2.6) implies that

$$|S_{R'}| \geq |T_{B'}| + n - |V(B')| > |S_{B'}| + \left\lfloor \left(\frac{1}{3} - 2\delta\right) n \right\rfloor - 1.$$

However (2.5) implies that $|S_{R'}| < \left\lceil \left(\frac{1}{3} + \frac{1}{2}\delta\right) n \right\rceil$, and hence $|S_{B'}| \leq 3\delta n$. Note also that $|S_{R'}| > \left\lfloor \left(\frac{1}{3} - 2\delta\right) n \right\rfloor$. Lemma 2.28 implies that $R'[V(R') \setminus S_{R'}]$ has maximum degree at most $11\delta n$. As $V(G) \setminus V(R') \subseteq S_{B'}$, it must be the case that $|V(R')| \geq (1 - 3\delta) n$, and so $R[V(G) \setminus S_{R'}]$ has maximum degree at most $11\delta n$, implying that the first case of Lemma 2.30 holds.

Case 2 - one of R' or B' contains a matching on at least $\left(\frac{2}{3} + \delta\right) n$ vertices.

In this case, if there is a monochromatic component C which contains a matching on at least $\left(\frac{1}{2} + \delta\right) n$ vertices such that either C contains an odd cycle or $|V(C)| \geq (1 - 5\delta) n$, then the second case of Lemma 2.30 holds, completing the proof. Without loss of generality, suppose that R' contains a matching on at least $\left(\frac{2}{3} + \delta\right) n$ vertices. It may be assumed that $|V(R')| < (1 - 5\delta) n$ and R' is bipartite with classes $Y_{R'}$ and $Z_{R'}$, otherwise taking $C = R'$ will complete the proof. As R is not connected, and assumption (2.4) still holds, Lemma 2.29 implies that $|V(B')| \geq \left(\frac{3}{4} - \delta\right) n$.

It may be assumed that $|Z_{R'}| \geq |Y_{R'}|$ and so

$$|Y_{R'}| \leq \frac{1}{2}|V(R')| < \left(\frac{1}{2} - \frac{5}{2}\delta\right) n.$$

As each edge of the matching in R' contains one vertex from $Z_{R'}$ and one from $Y_{R'}$,

$$\left(\frac{1}{3} + \frac{\delta}{2}\right) n \leq |Y_{R'}| \leq |Z_{R'}| < \left(\frac{2}{3} - \frac{11}{2}\delta\right) n. \tag{2.9}$$

Note that any vertex in $V(G)\backslash V(R')$ has red degree at most $n - |V(R')| \leq \frac{1}{3}n$, and any vertex in $Z_{R'}$ has red degree at most $|Y_{R'}| < \left(\frac{1}{2} - \frac{5\delta}{2}\right) n$. Hence $Y_{R'}$ contains all vertices with red degree at least $\left(\frac{1}{2} - 2\delta\right) n$.

However a vertex v in $V(G) \setminus V(B')$ has no blue neighbours in $V(B')$, and hence has blue degree at most $\left(\frac{1}{4} + \delta\right) n$. Hence v has red degree at least $\left(\frac{1}{2} - 2\delta\right) n$, and so $v \in Y_{R'}$. Thus $V(G) \setminus V(B') \subseteq Y_{R'}$.

Suppose that B' does not contain a matching on at least $\left(\frac{1}{2} + \delta\right) n$ vertices. Let $S_{B'}$, $T_{B'}$ and $U_{B'}$ be defined using Definition 2.27, with $a = \lceil \delta^{-1} \rceil$ and $b = \left\lceil \left(\frac{1}{2} + \delta\right) n \right\rceil$. Then Lemma 2.28 implies that

$$|S_{B'}| < \left(\frac{1}{4} + \frac{1}{2}\delta\right) n + \frac{1}{2}, \tag{2.10}$$

and

$$|T_{B'}| > |V(B')| + |S_{B'}| - \left(\frac{1}{2} + 2\delta\right) n - 1. \tag{2.11}$$

Let v be a vertex in $T_{B'}$. Then v has at most a blue neighbours outside of

$S_{B'}$ and hence (2.10) implies that v has blue degree at most $\left(\frac{1}{4}+\frac{1}{2}\delta\right)n+a+1$. Thus v has red degree at least $\left(\frac{1}{2}-\frac{3}{2}\delta\right)n-a-1$, and so $v \in Y_{R'}$. Hence $Y_{R'}$ contains the disjoint sets $T_{B'}$ and $V(G)\setminus V(B')$, and so (2.11) implies that

$$\begin{aligned}
|Y_{R'}| &\geq |T_{B'}| + |V(G)\setminus V(B')| \\
&> |V(B')| + |S_{B'}| - \left(\frac{1}{2}+2\delta\right)n - 1 + n - |V(B')| \\
&\geq \left(\frac{1}{2}-2\delta\right)n - 1.
\end{aligned}$$

However this contradicts $|Y_{R'}| < \left(\frac{1}{2}-\frac{5\delta}{2}\right)n$. Hence B' contains a matching on at least $\left(\frac{1}{2}+\delta\right)n$ vertices.

Let u be a vertex of $Z_{R'}$, and so necessarily $u \in V(B')$. Then u has at most $\left(\frac{1}{4}+\delta\right)n$ non-neighbours in $Z_{R'}$, and so at least one neighbour $v \in Z_{R'}$. As $Z_{R'}$ is an independent set in R, the edge uv must be blue. However, in G, both u and v have at most $\left(\frac{1}{4}+\delta\right)n$ non-neighbours, and so u and v have at least $\left(\frac{1}{2}-2\delta\right)n$ common neighbours. Hence, in particular, u and v have a common neighbour $w \in V(G)\setminus Y_{R'}$. By definition, all red neighbours of both u and v lie in $Y_{R'}$, and so uw and vw are both blue edges. Hence B' contains a cycle of length three, and so taking $C = B'$, the second case of Lemma 2.30 holds, completing the proof. $\square$

2.6 Proof of Theorem 2.4

In this section, Theorem 2.4 will be proved. The following preliminary lemmas will be of use in the proof.

Lemma 2.32. *Let n be sufficiently large and let G be a graph of order n with $\delta(G) \geq \frac{3}{4}n$. Suppose that $E(G) = E(R_G) \cup E(B_G)$ is a 2-edge-multi-colouring.*

Suppose that there is a set $S \subseteq V(G)$ with $|S| > \frac{1}{2}n$ and $\Delta(B_G[S]) \leq \frac{1}{4}|S| - \frac{1}{8}n$. Then $C_\ell \subseteq R_G$ for all $\ell \in [3, |S|]$.

Suppose that B_G is bipartite. Either $C_\ell \subseteq R_G$ for all $\ell \in [4, \lceil \frac{n}{2} \rceil]$, or $G \cong K_4(n/4)$ and the colouring is a 2-bipartite 2-edge-multi-colouring.

Note that this result also holds with the colours exchanged.

Proof. Suppose that there is a set S with $|S| \geq \frac{1}{2}n$ such that $\Delta(B_G[S]) \leq \frac{1}{4}|S| - \frac{1}{8}n$. All vertices in S have at most $\frac{1}{4}n$ non-neighbours in G, and so

$$\delta(R_G[S]) \geq |S| - \frac{1}{4}n - \left(\frac{1}{4}|S| - \frac{1}{8}n\right) = \frac{3}{4}|S| - \frac{1}{8}n \geq \frac{1}{2}|S|. \qquad (2.12)$$

Corollary 1.12 implies that either $R_G[S]$ is pancyclic, or $R_G[S] \cong K_{|S|/2,|S|/2}$. In the latter case, $\delta(R_G[S]) = \frac{1}{2}|S|$, and so equality holds in (2.12), implying that $|S| = \frac{1}{2}n$. Hence, if $|S| > \frac{1}{2}n$, then $C_\ell \subseteq R_G$ for all $\ell \in [3, |S|]$.

Suppose that B_G is bipartite with classes S_1 and S_2, labelled so that $|S_1| \geq |S_2|$. If $|S_1| > \frac{1}{2}n$, then $C_\ell \subseteq R_G$ for all $\ell \in [3, |S_1|]$ and the proof is complete. So it may be assumed that n is even and $|S_1| = |S_2| = \frac{1}{2}n$. By the above, either $C_\ell \subseteq R_G$ for all $\ell \in [3, \frac{1}{2}n]$, or both $R_G[S_1]$ and $R_G[S_2]$ are isomorphic to $K_{n/4,n/4}$. In the first case, the proof is complete so assume that both $R_G[S_1]$ and $R_G[S_2]$ are isomorphic to $K_{n/4,n/4}$. This implies that n is divisible by four. Both $B_G[S_1]$ and $B_G[S_2]$ are isomorphic to the empty graph and so $G \cong K_4(n/4)$.

For $i \in \{1, 2\}$, let $S_{i,1}$ and $S_{i,2}$ be the independent sets of G partitioning S_i. If R_G is not bipartite, then, without loss of generality, there are red edges between $S_{1,1}$ and both $S_{2,1}$ and $S_{2,2}$. Hence there is a red path of length either two or four between a vertex of $S_{2,1}$ and a vertex of $S_{2,2}$, with all internal vertices in S_1. However R_G is complete between $S_{2,1}$ and $S_{2,2}$ and so R_G contains C_ℓ for all $\ell \in [4, \lceil \frac{1}{2}n \rceil]$. If, however, R_G is bipartite then the colouring is a 2-bipartite 2-edge-multi-colouring. $\square$

Lemma 2.33. *Let n be sufficiently large and let G be a graph of order n with $\delta(G) \geq \frac{3}{4}n$. There is no set S of order at most three such that $V(G) \setminus S$ can be partitioned into non-empty sets $X_1, \ldots, X_4$ such that G has no edges between X_1 and X_4 and no edges between X_2 and X_3.*

Proof. Suppose that there is such a set S. Then $\sum_{i=1}^{4} |X_i| \geq n - 3$, and so $|X_i| \geq \frac{1}{4}(n-3)$, for some $1 \leq i \leq 4$. Let $v \in X_{5-i}$, which is assumed to be non-empty. Then v has no neighbours in X_i and is also not adjacent to itself. Hence $d_G(v) \leq n - (|X_i| + 1) < \frac{1}{4}n$, contradicting the minimal degree of G. $\square$

Using the lemmas above, it is now possible to prove Theorem 2.4.

Proof of Theorem 2.4. Let G be a 2-edge-coloured graph of order n with $\delta(G) \geq \frac{3}{4}n$. Let $E(G) = E(R) \cup E(B)$ be the 2-edge-colouring of G. Fix $0 < \delta < \frac{1}{294}$, and let $\eta = \frac{1}{10}\delta$ and $\epsilon = \min\{\frac{1}{4}\epsilon(\eta), \frac{1}{6}\eta\}$, where $\epsilon(\eta)$ is as defined in Lemma 2.23. Let $m = \max\{n_0, \epsilon^{-1}\}$, where $n_0(\delta)$ is defined as in Lemma 2.30. Finally, let $K = 2M(\epsilon, 2, m)$, where $M(\epsilon, 2, m)$ is as defined in the Regularity Lemma (Theorem 2.9).

Lemma 2.22 implies that either $G \cong K_4(n/4)$ and the colouring is a 2-bipartite 2-edge-colouring or G contains a monochromatic C_ℓ for all $\ell \in [4, K]$. Hence, to complete the proof, it is sufficient to show that one of the following holds:

1. $G \cong K_4(n/4)$ and the colouring is a 2-bipartite 2-edge-multi-colouring;

2. both

   (a) G contains a monochromatic cycle of length at least $\left\lceil \left(\frac{2}{3} + \frac{\delta}{2}\right) n \right\rceil$;

   (b) for all $\ell \in \left[K, \left\lceil \frac{1}{2} n \right\rceil\right]$, either R or B contains a cycle of length ℓ;

3. one of R or B contains cycles of all lengths $\ell \in \left[3, \left\lceil \left(\frac{2}{3} - \delta\right) n \right\rceil\right]$.

Suppose that G does not satisfy any of the above; this will be referred to as the *main assumption* of the proof.

Apply Theorem 2.9 to G, with parameters m, $r = 2$, η and ϵ. Hence there is a partition $\mathcal{P} = (V_i)_{i=0}^k$ of $V(G)$ and a subgraph G' of G with the properties given in that Theorem. Let H be the reduced graph defined as in Definition 2.11 by G' and $\mathcal{P}$, with 2-edge-multi-colouring $E(H) = E(R_H) \cup E(B_H)$. Recall that $|V_i| = t \leq \frac{n}{k} \leq \lceil \epsilon n \rceil$ for all $1 \leq i \leq k$, and note that

$$(1 - \epsilon) n \leq \sum_{i=1}^{k} |V_i| = kt \leq n. \tag{2.13}$$

Theorem 2.9 implies that $\delta(G') \geq \left(\frac{3}{4} - 2\eta - \epsilon\right) n$. Suppose that $\delta(H) < \left(\frac{3}{4} - \delta\right) k$ and let v_i be a vertex of H such that $\deg_H(v_i) < \left(\frac{3}{4} - \delta\right) k$. In G', a vertex $v \in V_i$ has neighbours only in V_0, or in V_j for those j such that $v_i v_j$

is an edge of H. Hence

$$\deg_{G'}(v) < \left(\frac{3}{4} - \delta\right) kt + |V_0| \leq \left(\frac{3}{4} - \delta + \epsilon\right) n.$$

This is a contradiction, as $\delta > 2\eta + 2\epsilon$. Hence $\delta(H) \geq \left(\frac{3}{4} - \delta\right) k$.

Lemma 2.30 implies that one of the following holds.

(i) There is a set $S \subseteq V(H)$ of order at least $\left\lfloor \left(\frac{2}{3} - \frac{\delta}{2}\right) k \right\rfloor$ such that either $\Delta(R_H[S]) \leq 11\delta k$ or $\Delta(B_H[S]) \leq 11\delta k$.

(ii) There is a component of R_H or B_H which contains a matching on at least $\left(\frac{2}{3} + \delta\right) k$ vertices. Further, there is a monochromatic component of H containing a matching on at least $\left(\frac{1}{2} + \delta\right) k$ vertices such that either this component contains an odd cycle, or this component contains all but at most $5\delta k$ vertices of $|V(H)|$.

(iii) There is a partition $V(H) = U_1 \cup \cdots \cup U_4$ with $\max_i |U_i| \leq \left(\frac{1}{4} + \delta\right) k$ such that R_H contains no edges from $U_1 \cup U_2$ to $U_3 \cup U_4$ and B_H contains no edges from $U_1 \cup U_3$ to $U_2 \cup U_4$. Further, the bipartite graphs $R_H[U_1, U_2]$, $R_H[U_3, U_4]$, $B_H[U_1, U_3]$ and $B_H[U_2, U_4]$ all contain a matching on at least $\left(\frac{1}{2} - 6\delta\right) k$ vertices.

In each of the three cases, a contradiction to the main assumption will be obtained.

Case (i)

As the definition of the reduced graph is symmetric, it may be assumed that there is a set $S \subseteq V(H)$ of order at least $\left\lfloor \left(\frac{2}{3} - \frac{\delta}{2}\right) k \right\rfloor$ such that $\Delta(R_H[S]) \leq 11\delta k$. Let $S' \subseteq V(G)$ be the union of all clusters V_i such that $v_i \in S$. Note that (2.13) implies that

$$|S'| \geq \left\lfloor \left(\frac{2}{3} - \frac{\delta}{2}\right) k \right\rfloor t \geq \left(\frac{2}{3} - \delta\right) n.$$

Hence $|S'| \geq \left\lceil \left(\frac{2}{3} - \delta\right) n \right\rceil$.

Let $v \in V_i \subseteq S'$ and suppose that vu is a red edge of G for some $u \in S'$. Then, either vu is not an edge of G', or $u \in V_j$ for some j such that $v_i v_j \in E\left(R_H[S]\right)$. Hence

$$\deg_{R_G[S']}(v) \leq (\deg_G(v) - \deg_{G'}(v)) + 11\delta kt \leq 12\delta n.$$

This implies that $\Delta(R_G[S']) \leq 12\delta n$. However $|S'| \geq \left(\frac{2}{3} - \delta\right) n$ implies that $\frac{1}{4}|S'| \geq \left(\frac{1}{8} + 12\delta\right) n$ and thus $\Delta(R_G[S']) \leq \frac{1}{4}|S'| - \frac{1}{8}n$. Hence Lemma 2.32 implies that B_G contains a cycle of length ℓ for all $\ell \in \left[3, \left\lceil \left(\frac{2}{3} - \delta\right) n \right\rceil\right]$. This contradicts the main assumption (part 3).

Case (ii)

In this case, there is a component C_1 of R_H or B_H which contains a matching M_1 on at least $\left(\frac{2}{3} + \delta\right) k$ vertices. Further, there is a monochromatic component C_2 of H containing a maximal (within C_2) matching M_2 on at least $\left(\frac{1}{2} + \delta\right) k$ vertices such that either C_2 contains an odd cycle, or C_2 contains all but at most $5\delta k$ vertices of $|V(H)|$.

Assume without loss of generality that C_1 is a component of R_H. Corollary 2.25 implies that $R_{G'}$ contains a cycle of length at least

$$\left\lceil (1 - \epsilon) \left(1 - \epsilon - \frac{1}{4}\eta\right) \left(\frac{2}{3} + \delta\right) kt \right\rceil \geq \left(\frac{2}{3} + \frac{1}{2}\delta\right) n.$$

Hence R_G contains a cycle of length at least $\left\lceil \left(\frac{2}{3}+\frac{1}{2}\delta\right) n \right\rceil$, and so part 2.($a$) of the main assumption holds. Hence if part 2.(b) of the main assumption also holds, there is a contradiction.

Similarly, assume without loss of generality that C_2 is a component of R_H. Assume also that $V(C_2) = \{v_1, \ldots, v_{s'}\}$ where $s' = |V(C_2)|$, and that $E(M_2) = \{v_{2i-1}, v_{2i} : 1 \leq i \leq s\}$. Lemma 2.24 applied to C_2 implies that there are sets $(W_i)_{i=1}^{s'}$ such that the following holds.

- For all $1 \leq i \leq s'$, W_i is a subset of V_i of order $t' = \lceil (1-\epsilon)\, t \rceil$.
- For all edges $v_i v_j$ of H, the graph $R_{G'}\,[W_i, W_j]$ is 2ϵ-regular with density at least $\eta - \epsilon$.
- For all edges $v_i v_j$ of M, the graph $R_{G'}\,[W_i, W_j]$ is $(2\epsilon, \eta - 2\epsilon)$-super-regular.
- For all $1 \leq i \leq s'$, there is some $1 \leq j \leq 2s$ such that $v_j \in V(M)$, and any vertex in W_i has at least $(\eta - 2\epsilon)\, t'$ neighbours in W_j.

Note that

$$\left(1 - \epsilon - \frac{1}{4}\eta\right) t' \left(\frac{1}{2} + \delta\right) k \geq \left\lceil \frac{1}{2} n \right\rceil . \tag{2.14}$$

Hence Lemma 2.23 implies that R_G contains a cycle of length ℓ, for all even $4 \leq \ell \leq \left\lceil \frac{1}{2} n \right\rceil$.

Suppose that, for some $1 \leq i \leq 2s$, R_G contains a path P of odd length at most $K - 4$ between two vertices x_1 and x_1' of W_i. It may be assumed that $i = 1$. Note that P contains at most $K - 4 \leq \epsilon t'$ vertices in each cluster.

Hence Lemma 2.23 and (2.14) imply that, for all even $4 \leq \ell \leq \lceil \frac{1}{2} n \rceil$, the graph R_G contains a path of length ℓ between x_1 and x_1' which is internally vertex disjoint from P. Hence, for all odd $K \leq \ell \leq \lceil \frac{1}{2} n \rceil$, the graph R_G contains a cycle of length ℓ. In particular, R_G contains a cycle of length ℓ for all $K \leq \ell \leq \lceil \frac{1}{2} n \rceil$. This implies that part 2.(b) of the main assumption holds, and so the following fact holds.

Fact 2.34. *For all $1 \leq i \leq 2s$, R_G does not contain a path of odd length at most $K - 4$ between two vertices x_1 and x_1' of W_i.*

Suppose that C_2 contains an odd cycle D. As M_2 is maximal, it may be assumed that $V(D)$ intersects $V(M_2)$ and that $v_1 \in V(D) \cap V(M_2)$. The Embedding Lemma implies that there is a path of odd length at most $k \leq M(\epsilon) = \frac{1}{2} K$ between two vertices x_1 and x_1' of W_1. This contradicts Fact 2.34 and hence C_2 does not contain an odd cycle.

It may be assumed that $|V(C_2)| \geq (1 - 5\delta)\, n$ and that C_2 is bipartite. Let C_X and C_Y be the two classes of the bipartite graph $R_H[C_2]$. As each edge of M_2 must have one end in C_X and the other in C_Y, it may be assumed that $v_{2i-1} \in C_X$ and $v_{2i} \in C_Y$, for all $1 \leq i \leq s$. Let X be the union of all W_i such that $v_i \in C_X$ and Y be the union of all W_i such that $v_i \in C_Y$. By the definition of W_j the following fact holds.

Fact 2.35. *Any vertex of X has at least $(\eta - \epsilon)\, t'$ neighbours in $\bigcup_{i=1}^{s} W_{2i}$, and similarly, any vertex of Y has at least $(\eta - \epsilon)\, t'$ neighbours in $\bigcup_{i=1}^{s} W_{2i-1}$.*

Suppose that R_G contains a path P of odd length at most $K - 9 - k$ between two vertices v and u in X. Then u and v have distinct neighbours in $\bigcup_{i=1}^{s} W_{2i}$. Hence R_G contains a path of odd length at most $K - 7 - k$

between two distinct vertices $v' \in W_{2i}$ and $u' \in W_{2j}$, for some $1 \le i \le j \le r$. There is a path in R_H of length at most $k-1$ from v_{2i} to v_{2j}, which must have even length as $R_H[C_2]$ is bipartite. By the Embedding Lemma, R_G contains a path P' of even length at most $k-1$ between two vertices $v'' \in W_{2i}$ and $u'' \in W_{2j}$. However, Lemma 2.23 implies that R_G contains a path of length four from u'' to u', with no internal vertices in $V(P) \cup V(P')$, as illustrated in Figure 2.4. Hence there is a path of odd length at most $K-4$ between vertices v' and v'' of W_{2i}, contradicting Fact 2.34. Hence R_G does not contain a path of odd length at most $K-9-k$ between two vertices of X. Similarly R_G does not contain a path of odd length at most $K-9-k$ between two vertices of Y.

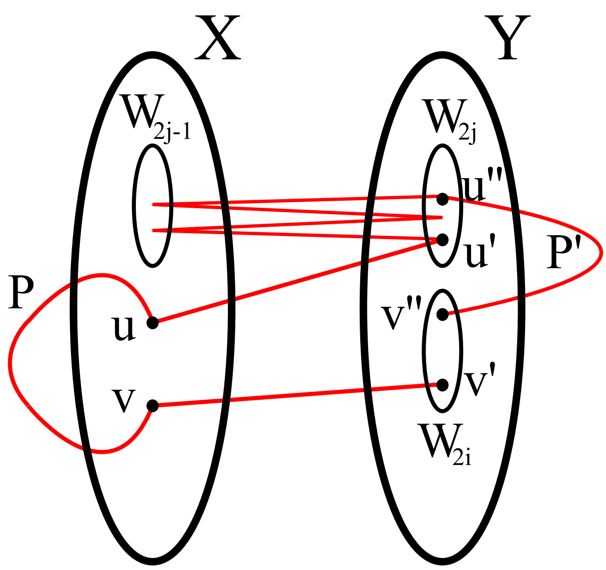


Figure 2.4: The straight red lines are edges and the curved red lines are paths.

In particular, $R_G[X \cup Y]$ is bipartite. Note that

$$|X \cup Y| \geq (1 - 5\delta)\, kt(1 - \epsilon) \geq (1 - 6\delta)n.$$

If $|X| > \frac{1}{2}n$, then Lemma 2.32 implies that B_G contains a cycle of length ℓ for all $\ell \in [3, |X|]$, contradicting the main assumption. Hence $|X| \leq \frac{1}{2}n$ and $|Y| \leq \frac{1}{2}n$, implying that $\min\{|X|, |Y|\} \geq \left(\frac{1}{2} - 6\delta\right) n$.

If any vertex v of $V \setminus (X \cup Y)$ has at least one red neighbour in both X and Y, then Fact 2.35 implies that there is a path of length three between two vertices of X. Hence all vertices of $V \setminus (X \cup Y)$ have no red neighbours in at least one of X or Y. Let X' be the set of vertices of $V \setminus (X \cup Y)$ with at least two red neighbours in Y, and Y' be the set of vertices of $V \setminus (X \cup Y)$ with at least two red neighbours in X. There are no red edges between X' and X or between Y' and Y. If there is a red edge uv within X', let u' and v' be distinct vertices of Y such that uu' and vv' are both red edges. Then $u'uvv'$ is a red path of length three between vertices of Y, which can not occur. Hence both $X \cup X'$ and $Y \cup Y'$ contain no red edges. In view of Lemma 2.32, it may be assumed that both $X \cup X'$ and $Y \cup Y'$ have order at most $\frac{1}{2}n$.

Let $Z = V(G) \setminus (X \cup X' \cup Y \cup Y')$. Note that vertices in Z have at most one red neighbour in $X \cup Y$. A vertex in Z has at most $6\delta n$ neighbours in $V \setminus (X \cup Y)$ and hence at least $\left(\frac{3}{4} - 6\delta\right) n$ neighbours in $X \cup Y$. Let X'' be the set of vertices in Z with at least $\left(\frac{3}{8} - 3\delta\right) n$ neighbours in X, and Y'' be the set of vertices in Z with at least $\left(\frac{3}{8} - 3\delta\right) n$ neighbours in Y. Let $X_0 = X \cup X' \cup X''$ and $Y_0 = Y \cup Y' \cup Y''$ Then V is the (not necessarily disjoint) union of X_0 and Y_0.

Without loss of generality, suppose that $|X_0| \geq \frac{1}{2}n$. By definition, all vertices in X'' have at least $\left(\frac{3}{8} - 3\delta\right) n$ neighbours in X. Recall that all vertices in X'' have at most one red neighbour in X. Let X''' be the set of vertices in X with a red neighbour in X''. Note that $|X'''| \leq |X''|$. All vertices in $X \cup X'$ have at most $\frac{1}{4}n$ non-neighbours in G and so at least $|X_0| - \frac{1}{4}n$ neighbours in X_0. Vertices in $X \setminus X'''$ have no red neighbours in X_0, while vertices in $X' \cup X'''$ have at most $|X''|$ red neighbours in X_0. In summary, the degree of a vertex in $B_G[X_0]$ satisfies

$$\deg_{B_G[X_0]}(v) \geq \begin{cases} |X_0| - \frac{1}{4}n & v \in X \setminus X''' \\ |X_0| - \frac{1}{4}n - |X''| & v \in X''' \cup X' \\ \left(\frac{3}{8} - 3\delta\right) n - 1 & v \in X''. \end{cases} \tag{2.15}$$

However, $\frac{1}{2}n \leq |X_0| \leq \left(\frac{1}{2} + 6\delta\right) n$. Hence vertices of X'' and $X \setminus X'''$ have degree at least $\frac{1}{2}|X_0|$ in $B_G[X_0]$. But

$$\begin{aligned} \max\{|X''|, |X'| + |X'''|\} &\leq |X' \cup X''| \\ &\leq |V(G) \setminus (X \cup Y)| \leq 6\delta n \leq 12\delta|X_0|. \end{aligned} \tag{2.16}$$

Hence there are at most $12\delta|X_0|$ vertices of X_0 with degree less than $\frac{1}{2}|X_0|$ in $B_G[X_0]$, and each of these vertices has degree at least $\left(\frac{1}{2} - 12\delta\right)|X_0|$ in $B_G[X_0]$. Letting $d_1 \leq \cdots \leq d_{|X_0|}$ be the degree sequence of $B_G[X_0]$, this implies that there is no $a < \frac{1}{2}|X_0|$ such that $d_a \leq a$. Hence Theorem 2.18 implies that $B_G[X_0]$ is Hamiltonian.

Note that (2.15) implies that

$$
\begin{aligned}
2e\left(B_G[X_0]\right) &\geq \left(|X_0| - \frac{1}{4}n\right)|X \setminus X'''| + \left(\left(\frac{3}{8} - 3\delta\right)n - 1\right)|X''| \\
&\quad + \left(|X_0| - \frac{1}{4}n - |X''|\right)\left(|X'''| + |X'|\right) \\
&= \left(|X_0| - \frac{1}{4}n\right)\left(|X| + |X'|\right) \\
&\quad + \left(\left(\frac{3}{8} - 3\delta\right)n - 1 - \left(|X'''| + |X'|\right)\right)|X''|.
\end{aligned}
$$

However, (2.16) and $|X_0| \leq \left(\frac{1}{2} + 6\delta\right)n$ imply that

$$
\left(\frac{3}{8} - 3\delta\right)n - 1 - \left(|X'''| + |X'|\right) \geq \left(\frac{3}{8} - 9\delta\right)n - 1 \geq |X_0| - \frac{1}{4}n + \delta n.
$$

Hence

$$
2e\left(B_G[X_0]\right) \geq \left(|X_0| - \frac{1}{4}n\right)\left(|X| + |X'| + |X''|\right) + \delta n|X''|,
$$

and, noting that $|X_0| \geq \frac{1}{2}n$,

$$
e\left(B_G[X_0]\right) \geq \frac{1}{4}|X_0|^2 + \frac{1}{2}\delta n|X''|.
$$

Thus Theorem 1.11 implies that either $B_G[X_0]$ is pancyclic, or $B_G[X_0] \cong K_{|X_0|/2,|X_0|/2}$ and $e(B_G[X_0]) = \frac{1}{4}|X_0|^2$. However, if $B_G[X_0]$ is pancyclic, then $C_\ell \subseteq B_G$ for all $\ell \in [3, |X_0|]$ contradicting the main assumption. Hence $B_G[X_0] \cong K_{|X_0|/2,|X_0|/2}$ and $e(B_G[X_0]) = \frac{1}{4}|X_0|^2$, which implies that X'' is empty. Thus $X_0 = X \cup X'$ is an independent set in R_G. Hence Lemma 2.32 and the main assumption imply that $|X_0| \leq \frac{1}{2}n$. But then $|Y_0| \geq \frac{1}{2}n$, and a similar argument shows that Y'' is empty. Hence Y_0 is an independent set in R_G and so R_G is bipartite. Thus Lemma 2.32 implies a contradiction to part 1. of the main assumption.

Case (iii)

In this case, there is a partition $V(H) = U_1 \cup \cdots \cup U_4$ with $\max_l |U_l| \leq \left(\frac{1}{4} + \delta\right) k$ such that R_H contains no edges from $U_1 \cup U_2$ to $U_3 \cup U_4$ and B_H contains no edges from $U_1 \cup U_3$ to $U_2 \cup U_4$. Further, the bipartite graph $R_H[U_l, U_{l+1}]$ contains a matching $M_{l,l+1}$ on at least $\left(\frac{1}{2} - 6\delta\right) k$ vertices for $l \in \{1, 3\}$, and the bipartite graph $B_H[U_l, U_{l+2}]$ contains a matching $M_{l,l+2}$ on at least $\left(\frac{1}{2} - 6\delta\right) k$ vertices for $l \in \{1, 2\}$. Note that this implies that each U_i has order at least $\left(\frac{1}{4} - 3\delta\right) k$. Also, the minimal degree of H implies that both R_H and B_H have exactly two components.

Arguing as in Lemma 2.24, there are sets $(W_i)_{i=1}^{k}$ such that the following hold.

- For all $1 \leq i \leq k$, W_i is a subset of V_i of order $t' = \lceil (1 - 2\epsilon)\, t \rceil$.
- For all edges $v_i v_j$ of R_H, the graph $R_{G'}[W_i, W_j]$ is 2ϵ-regular with density at least $\eta - \epsilon$.
- For all edges $v_i v_j$ of B_H, the graph $B_{G'}[W_i, W_j]$ is 2ϵ-regular with density at least $\eta - \epsilon$.
- For $l \in \{1, 3\}$ and all edges $v_i v_j$ of $M_{l,l+1}$, the graph $R_{G'}[W_i, W_j]$ is $(2\epsilon, \eta - 2\epsilon)$-super-regular.
- For $l \in \{1, 2\}$ and all edges $v_i v_j$ of $M_{l,l+2}$, the graph $B_{G'}[W_i, W_j]$ is $(2\epsilon, \eta - 2\epsilon)$-super-regular.

For $1 \leq l \leq 4$, let $X_l \subseteq V$ be the union of W_i such that $v_i \in U_l$. Also let X_0 be the union of the exceptional set V_0 and $\bigcup_{i=1}^k V_i \setminus W_i$. Then $|X_0| \leq \epsilon n + 2\epsilon tk \leq \delta n$, and, for all $1 \leq l \leq 4$,

$$\left(\frac{1}{4} - 4\delta\right) n \leq \left(\frac{1}{4} - 3\delta\right) kt' \leq |X_l| \leq \left(\frac{1}{4} + \delta\right) kt' \leq \left(\frac{1}{4} + \delta\right) n.$$

As R_H contains no edges from $U_1 \cup U_2$ to $U_3 \cup U_4$ and B_H contains no edges from $U_1 \cup U_3$ to $U_2 \cup U_4$, the definition of the reduced graph implies that in G', there are no red edges from $X_1 \cup X_2$ to $X_3 \cup X_4$ and no blue edges from $X_1 \cup X_3$ to $X_2 \cup X_4$. In G', a vertex in X_1 has no neighbours in X_4, and so, in G, a vertex in X_1 has at most $(2\eta + \epsilon)\, n$ neighbours in X_4. Hence, in G, a vertex in X_1 has at least $\left(\frac{1}{4} - 5\delta\right) n$ non-neighbours in X_4, and so at most $5\delta n$ non-neighbours in $X_1 \cup X_2 \cup X_3$. A similar result holds for each X_l. The following fact, a corollary of Lemma 2.23, will be of use throughout analysis of this case.

Fact 2.36. *Let x and y be any two not necessarily distinct vertices of X_1, and, for $1 \leq i \leq 2$, let S_i be a subset of X_i containing at most ten vertices. Let $3 \leq \ell \leq \left\lceil \left(\frac{1}{4} - 4\delta\right) n \right\rceil$. Then $R_G[X_1 \cup X_2]$ contains a path of length 2ℓ from x to y with no internal vertices in $S_1 \cup S_2$.*

Note that the same result holds if the roles of X_1 and X_2 are reversed, or if the graph $R_G[X_1 \cup X_2]$ is replaced by any of the graphs $R_G[X_3 \cup X_4]$, $B_G[X_1 \cup X_3]$ or $B_G[X_2 \cup X_4]$.

Proof. First note that there are no blue edges between X_1 and X_2 in G'. Hence, in G, both x and y have at most δn blue neighbours in X_2 and at most $5\delta n$ non-neighbours in X_2. Hence, both x and y have at least $|X_2|-6\delta n$ red neighbours in X_2. In particular, as $|X_2| \geq \left(\frac{1}{4}-3\delta\right)n$ both x and y have at least $\frac{3}{4}|X_2|$ red neighbours in X_2. Hence, there is some $W_i \subseteq X_2$ such that both x and y have at least two red neighbours in W_i. Let x' and y' be vertices of W_i such that xx' and yy' are edges of R_G.

Let $2 \leq \ell \leq \left\lceil\left(\frac{1}{4}-4\delta\right)n\right\rceil \leq \left\lceil\left(1-\epsilon-\frac{1}{4}\eta\right)t'\left(\frac{1}{4}-3\delta\right)k\right\rceil$. Applying Lemma 2.23 to the graph $R_G[X_1, X_2]$ implies that R_G contains a path of length 2ℓ from x' to y' with no internal vertices in $S_1 \cup S_2 \cup \{x, y\}$. Extending this path by the edges xx' and yy' gives the required paths from x to y. □

Recall that each vertex of X_1 has at most $5\delta n$ non-neighbours in $X_1 \cup X_2 \cup X_3$. In particular, there is an edge of G contained within X_1. Suppose, without loss of generality, that there is a red edge xy of G contained within X_1. Fact 2.36 implies that R_G contains a path of length 2ℓ from x to y for all $3 \leq \ell \leq \left\lceil\left(\frac{1}{4}-3\delta\right)n\right\rceil$ and a cycle of length 2ℓ for all $3 \leq \ell \leq \left\lceil\left(\frac{1}{4}-3\delta\right)n\right\rceil$. Hence R_G contains a cycle of length ℓ for all $\ell \in \left[7, \left\lceil\left(\frac{1}{2}-6\delta\right)n\right\rceil\right]$.

To complete the proof, it is sufficient to show the following statement, which shows that part 2. of the main assumption holds.

(A) G contains a monochromatic cycle of length at least $\left\lceil\left(\frac{2}{3}+\frac{1}{2}\delta\right)n\right\rceil$, and a monochromatic cycle of length ℓ for all $\ell \in \left[\left\lceil\left(\frac{1}{2}-6\delta\right)n\right\rceil+1, \left\lceil\frac{1}{2}n\right\rceil\right]$.

Suppose that R_G contains two vertex-disjoint paths P and P' from $X_1 \cup X_2$ to $X_3 \cup X_4$, with all internal vertices in X_0, and that R_G also contains two edges in X_1. Adding edges at both ends if necessary, it may be assumed that both P and P' have endpoints in X_1 and X_3 and contain at most one vertex in each X_l. Let P have endpoints x in X_1 and y in X_3 and let P' have endpoints x' in X_1 and y' in X_3. Let $3 \leq \ell_1 \leq \left\lceil \left(\frac{1}{4} - 4\delta\right) n \right\rceil$. Fact 2.36 implies that $R_G[X_3 \cup X_4]$ contains a path of length $2\ell_1$ from y to y' which is internally vertex-disjoint from P and P'. Let P_{ℓ_1} be the path formed by concatenation of this path with P and P'.

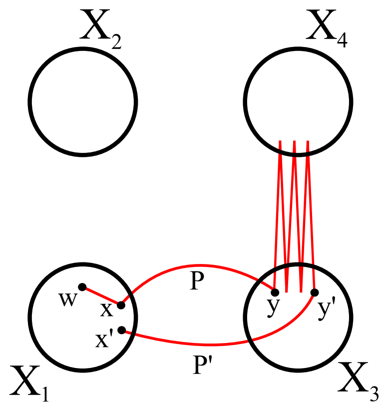


Figure 2.5: The straight red lines are edges and the curved red lines are paths.

By assumption R_G contains two edges in X_1, and hence contains some edge in X_1 which is not xx', as in Figure 2.5. Let this edge be ww', where it may be assumed that $w \notin \{x, x'\}$ and $w' \neq x'$. Note that, either $w' = x$, or w' and x have a common red neighbour in X_2. It will be assumed from now on that $w' = x$, the case when $w' \neq x$ being very similar. Let $3 \leq \ell_2 \leq \left\lceil \left(\frac{1}{4} - 4\delta\right) n \right\rceil$. Fact 2.36 implies that $R_G[X_1 \cup X_2]$ contains a path from x to x' of length $2\ell_2$ which is internally vertex-disjoint from P and P'. However, Fact 2.36 also implies that $R_G[X_1 \cup X_2]$ contains a path from w to x' of length $2\ell_2$ which is internally vertex-disjoint from P and P'. Adding the edge wx, this implies that R_G contains a path of length $2\ell_2 + 1$ from x to x' which is internally vertex-disjoint from P and P'.

In particular, $R_G\left[X_1 \cup X_2\right]$ contains a path from x to x' of length ℓ_3 which is internally vertex-disjoint from P and P', for all $6 \leq \ell_3 \leq \left\lceil \left(\frac{1}{2} - 8\delta\right) n \right\rceil$. Concatenating these paths with the path P_{ℓ_1} shows that R_G contains a cycle of length ℓ for all

$$|V(P)| + |V(P')| + 12 \leq \ell \leq \lceil (1 - 16\delta)\, n \rceil + |V(P)| + |V(P')| - 2.$$

Hence, if R_G contains two vertex-disjoint paths P and P' from $X_1 \cup X_2$ to $X_3 \cup X_4$ and at least two edges in some X_l then (A) holds. However this contradicts the main assumption, and so the following fact holds.

Fact 2.37. *If R_G contains two vertex-disjoint paths P and P' from $X_1 \cup X_2$ to $X_3 \cup X_4$, then R_G contains at most one edge in each X_l.*

Note that an identical result holds for B_G. Recall that each vertex of X_l has at least $|X_l| - 4\delta n$ neighbours in X_l, and so X_l contains at least three edges. Hence, it may be assumed that R_G does not contain two vertex-disjoint paths P and P' from $X_1 \cup X_2$ to $X_3 \cup X_4$. By Theorem 2.17, there is a vertex v_R such that there are no red paths from $X_1 \cup X_2$ to $X_3 \cup X_4$ in $G - \{v_R\}$.

Suppose that there is also a vertex v_B such that there are no blue paths from $X_1 \cup X_3$ to $X_2 \cup X_4$ in $G - \{v_B\}$. Let $S = \{v_R, v_B\}$. Consider a vertex $v \in X_0 \setminus S$. By the definition of v_R, it cannot have red neighbours in both $X_1 \cup X_2$ and $X_3 \cup X_4$, and similarly it cannot have blue neighbours in both $X_1 \cup X_3$ and $X_2 \cup X_4$. Hence, there is some X_i such that v has no neighbours in X_i, and so v may be added to X_{5-i} without adding edges between X_i and X_{5-i}. Repeating this argument for each vertex of $X_0 \setminus S$ shows that S is a contradiction to Lemma 2.33, and so there is no such vertex v_B.

Hence Theorem 2.17 implies that there are two vertex disjoint blue paths between $X_1 \cup X_3$ and $X_2 \cup X_4$. Thus Fact 2.37 implies that B_G contains at most one edge in each X_l. Note that, using the above argument, B_G certainly contains a cycle of length at least $\left\lceil \left(\frac{2}{3} + \frac{1}{2}\delta\right) n \right\rceil$, and so, to prove (A), it just remains to show that G contains a monochromatic cycle of length ℓ for all $\ell \in \left[\left\lceil \left(\frac{1}{2} - 6\delta\right) n \right\rceil + 1, \left\lceil \frac{1}{2} n \right\rceil\right]$. The proof continues by bounding below the red degree of all vertices of G.

In G, a vertex of X_1 has at most $5\delta n$ non-neighbours in $X_1 \cup X_2$. However, in G', a vertex of X_1 has no blue neighbours in X_2, and hence, in G a vertex of X_1 has at most $(2\eta + \epsilon)\, n$ blue neighbours in X_2. A vertex of X_1 has at most one blue neighbour in X_1. This implies the following fact.

Fact 2.38. *In R_G, all vertices in X_1 have at least $|X_1| + |X_2| - 6\delta n$ neighbours in $X_1 \cup X_2$.*

Similar bounds hold for all vertices in $\bigcup_{i=1}^{4} X_i$. The proof now turns to bounding below the red degree of all vertices of X_0.

Suppose that some vertex $v \in X_0$ has at least $\left(\frac{1}{2} + 4\delta\right) n$ blue neighbours. At most δn of these neighbours lie in X_0, and at most $\left(\frac{1}{4} + \delta\right) n$ of these neighbours lie in each X_l. Hence v must have at least δn blue neighbours in at least three of the sets X_i. Suppose that v has at least two blue neighbours in each of X_1, X_2 and X_3, the other cases being similar. If B_G contains no path P from $X_1 \cup X_3$ to $X_2 \cup X_4$ in $G - \{v\}$, then $S = \{v, v_R\}$ contradicts Lemma 2.33, after adding each vertex of X_0 to one of $X_1, \ldots, X_4$ as above.

Hence B_G contains a blue path P from $X_1 \cup X_3$ to $X_2 \cup X_4$ in $G - \{v\}$.

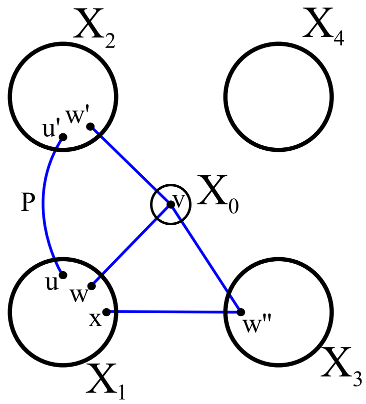


Figure 2.6: The straight blue lines are edges and the curved blue lines are paths.

By adding an edge at each end if necessary, it may be assumed that P has endpoints $u \in X_1$ and $u' \in X_2$ and contains at most one vertex in each X_l. Let $w \in X_1$, $w' \in X_2$ and $w'' \in X_3$ be blue neighbours of v which are not vertices of P. Note that w'' has at most $5\delta n$ non-neighbours in $X_1 \cup X_3 \cup X_4$, and has no red neighbour in $X_1 \setminus \{v_R\}$. Hence w'' has a blue neighbour x in $X_1 \setminus \{u, w\}$, as in Figure 2.6.

Let $\ell_1, \ell_2 \in \left[3, \left\lceil \left(\frac{1}{4} - 4\delta\right) n \right\rceil\right]$. Fact 2.36 implies that $B_G[X_1, X_3]$ contains a path P^{ℓ_1} from x to u of length $2\ell_1$ and a path P^{ℓ_2} from w to u of length $2\ell_2$, both of which contain no internal vertices in $V(P) \cup \{w''\}$.

Then the blue path $u'PuP^{\ell_1}xw''vw'$ has length $|V(P)| - 1 + 2\ell_1 + 3$, and the blue path $u'PuP^{\ell_2}wvw'$ has length $|V(P)| - 1 + 2\ell_2 + 2$. Hence B_G contains a path of length ℓ from u' to w' for all $|V(P)| + 7 \leq \ell \leq \left\lceil \left(\frac{1}{2} - 8\delta\right) n \right\rceil$. Note that the only internal vertices of these paths in $X_2 \cup X_4$ are internal vertices of P. Let $\ell_3 \in \left[3, \left\lceil \left(\frac{1}{4} - 4\delta\right) n \right\rceil\right]$. Fact 2.36 implies that $B_G[X_2, X_4]$ contains a path P^{ℓ_3} from w' to u' of length $2\ell_3$, which is internally disjoint from P. Hence B_G contains a cycle of length ℓ for all $|V(P)| + 13 \leq \ell \leq \lceil (1 - 16\delta) n \rceil$.

This contradicts the main assumption.

Thus each vertex in X_0 has at most $\left(\frac{1}{2}+4\delta\right)n$ blue neighbours, and so at least $\left(\frac{1}{4}-4\delta\right)n$ red neighbours. In particular, this implies that each vertex of X_0 has a red neighbour in $X_1 \cup X_2 \cup X_3 \cup X_4$. Hence there are exactly two components in $R_G[V-\{v_R\}]$, one of which contains $X_1 \cup X_2$ and the other of which contains $X_3 \cup X_4$.

Let D be the vertex set of a component of $R_G[V-\{v_R\}]$ which contains at least $\frac{1}{2}n$ vertices, if such a component exists. Otherwise, n is odd and both components of $R_G[V-\{v_R\}]$ contain $\frac{1}{2}(n-1)$ vertices. As v_R has red degree at least $\left(\frac{1}{4}-4\delta\right)n$ it has at least $\left(\frac{1}{8}-2\delta\right)n$ neighbours in one of the components of $R_G\left[V-\{v_R\}\right]$. In this case, let D be the union of v_R and the vertex set of a component of $R_G\left[V-\{v_R\}\right]$ in which v_R has at least $\left(\frac{1}{8}-2\delta\right)n$ neighbours.

Without loss of generality, assume that $X_1 \cup X_2 \subseteq D$. Fact 2.38 implies that all vertices in $X_1 \cup X_2$ have at least $|X_1|+|X_2|-6\delta n$ red neighbours in $X_1 \cup X_2$. Let $d_1 \leq \cdots \leq d_{|D|}$ be the degree sequence of $R_G[D]$. Note that $|D| \leq \left(\frac{1}{2}+3\delta\right)n$ and $d_1 \geq \left(\frac{1}{8}-2\delta\right)n$. Also, all but at most δn vertices of D lie in $X_1 \cup X_2$. However, a vertex in $X_1 \cup X_2$ has at least $|X_1|+|X_2|-5\delta n \geq |D|-6\delta n$ red neighbours in $X_1 \cup X_2$. This implies that $d_{\lfloor \delta n \rfloor+1} \geq \frac{1}{2}|D|$ and hence $d_i \geq i$ for all $i \leq \frac{1}{2}|D|$. Hence Theorem 2.18 implies that $R_G[D]$ is hamiltonian.

Finally $R_G[D]$ has at least as many edges as $R_G[X_1 \cup X_2]$. However, $R_G[X_1 \cup X_2]$ has at least $\frac{1}{2}\left(|D|-6\delta n\right)\left(|D|-\delta n\right) > \frac{1}{4}|D|^2$ edges. Hence, by Theorem 1.11, $R[D]$ is pancyclic, contradicting the main assumption. $\square$

Chapter 3

Monochromatic circumference

3.1 The circumference of a graph

Like Chapter 2, this chapter will consider monochromatic cycles in edge-coloured graphs. Recall that the *monochromatic circumference* of an r-edge coloured graph G is the length of the longest monochromatic cycle. Li, Nikiforov and Schelp [38] posed the following question about the monochromatic circumference of a 2-edge-coloured graph.

Question 3.1. *Let $0 < c < 1$ and G be a graph of sufficiently large order n. If $\delta(G) > cn$ and G is 2-edge-coloured, what is the minimum value the monochromatic circumference may take?*

The monochromatic circumference of a 2-edge colouring of the complete graph was studied by Faudree, Lesniak and Schiermeyer [23], who proved the following result.

Theorem 3.2. *Let $n \geq 6$, and suppose that $E(K_n) = E(R) \cup E(B)$ is an* *edge-colouring of K_n. Then either R or B contains a cycle of length at least $\left\lceil\frac{2}{3}n\right\rceil$. Further, there is a 2-edge-colouring of K_n such that both R and B have circumference at most $\left\lceil\frac{2}{3}n\right\rceil$.*

The following theorem extends this by showing that any 2-edge-coloured graph G with sufficiently large order n and minimal degree at least $\frac{3}{4}n$ has monochromatic circumference at least $(1 - o_n(1))\,\frac{2}{3}n$.

Theorem 3.3. *Let $0 < \delta < \frac{1}{294}$. Let G be a graph of sufficiently large order n with $\delta(G) \geq \frac{3}{4}n$. Suppose that $E(G) = E(R_G) \cup E(B_G)$ is a 2-edge-colouring. Then either G has monochromatic circumference at least $\left\lceil\left(\frac{2}{3} + \frac{\delta}{2}\right)n\right\rceil$, or one of R_G and B_G contains C_ℓ for all $\ell \in \left[3, \left\lceil\left(\frac{2}{3} - \delta\right)n\right\rceil\right]$.*

For integers $t \leq s$, define the following 2-edge-coloured graph. Note that for $s = 2t$, this is the 2-edge-colouring of K_n described by Faudree, Lesniak and Schiermeyer [23] in their proof of Theorem 3.2.

Definition 3.4 (The graph $F_{s,t}$)**.** *Let $t \leq s$ be integers. Let $F_{s,t}$ be the complete graph on $t + s$ vertices, with $V = V(F_{s,t})$. Let $A \subseteq V$ be a set of order s. The edges of $F_{s,t}$ are 2-coloured by letting all edges contained in A be red, and all other edges be blue, as illustrated in Figure 3.1.*

In $F_{s,t}$, the red graph has circumference s. All blue edges have at least one end in $V \setminus A$. Hence any blue cycle has at least half of its vertices in $V \setminus A$ and so the blue graph has circumference $2|V \setminus A| = 2t$. Thus the monochromatic circumference of $F_{s,t}$ is $\max\{s, 2t\}$.

Let $n = 3t$. Then $|F_{2t,t}| = n$, $\delta(F_{2t,t}) = n-1$ and $F_{2t,t}$ has monochromatic circumference $\frac{2}{3}n$. Hence Theorem 3.3 is asymptotically sharp. Results in

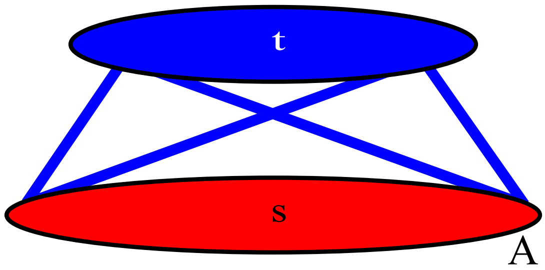


Figure 3.1: The graph $F_{s,t}$.

this chapter will be proved using the same methods as in the previous chapter. In particular, Theorem 3.3 is a consequence of Theorem 2.4, as demonstrated in the following proof.

Proof of Theorem 3.3. Let G be a graph of sufficiently large order n with $\delta(G) \geq \frac{3}{4}n$. Suppose that $E(G) = E(R_G) \cup E(B_G)$ is a 2-edge-colouring. Theorem 2.4 implies that, unless $G \cong K_4(n/4)$ and the colouring is 2-bipartite, one of the two conclusions of Theorem 3.3 holds.

If $G \cong K_4(n/4)$ and the colouring is 2-bipartite, let $V_1 \cup V_2$ be the bipartition of R_G and $W_1 \cup W_2$ be the bipartition of B_G. Let $U_{i,j} = V_i \cap W_j$ for all $i, j \in \{1, 2\}$. Then the $U_{i,j}$ are four independent sets of G covering the vertices, and so must be the four independent sets of order $n/4$. Hence:

- all edges between $U_{1,1}$ and $U_{1,2}$ and between $U_{2,1}$ and $U_{2,2}$ are blue;
- all edges between $U_{1,1}$ and $U_{2,1}$ and between $U_{1,2}$ and $U_{2,2}$ are red;
- edges between $U_{1,1}$ and $U_{2,2}$ and between $U_{2,1}$ and $U_{1,2}$ can be either colour.

For $n \geq 12$, there are three independent edges of G which lie between $U_{1,1}$ and $U_{2,2}$. In particular, there are two independent edges uv and $u'v'$ of

G with the same colour such that $u, u' \in U_{1,1}$ and $v, v' \in U_{2,2}$. Without loss of generality, assume that uv and $u'v'$ are blue edges. However, $B_G[U_{1,1}, U_{1,2}]$ contains a path of length $\frac{1}{2}n - 2$ from u to u', and $B_G[U_{2,1}, U_{2,2}]$ contains a path of length $\frac{1}{2}n - 2$ from v to v'. Hence B_G contains a cycle of length at least $n - 2$. $\square$

Note that Theorem 3.3 does not give an exact answer to the question posed by Li, Nikiforov and Schelp. Hence, the following definition is made, allowing an asymptotic version of Question 3.1 to be considered.

Definition 3.5. *For any positive integer r, let Φ_r be the following function with domain $(0, 1)$. For $0 < c < 1$, let $\Phi_r(c)$ be the supremum of the set of ϕ such that any r-edge-coloured graph G of sufficiently large order n with minimal degree at least cn has monochromatic circumference at least ϕn.*

Recall that Theorem 1.15 gives a tight bound on how large the circumference of a graph with a given order and minimal degree may be. This Theorem may be use to determine the value of $\Phi_1(c)$ for all $c \in (0, 1)$.

Corollary 3.6. *For any positive integer m, $\Phi_1(c) = \frac{1}{m}$ for all $c \in \left[\frac{1}{m+1}, \frac{1}{m}\right)$.*

Proof. Let m be a positive integer and suppose that $c \in \left[\frac{1}{m+1}, \frac{1}{m}\right)$. Let G be a graph with order n and minimal degree at least cn. Note that G has minimal degree at least $d := \lceil cn \rceil$. However $\frac{1}{d}n \leq \frac{1}{c} \leq m + 1$, and so $\left\lceil \frac{1}{d}n \right\rceil \leq m + 1$. Hence Theorem 1.15 implies that G has circumference at least $\left\lceil \frac{1}{m}n \right\rceil \geq \frac{1}{m}n$. Thus $\Phi_1(c) \geq \frac{1}{m}$.

Let t be an integer greater than $\frac{1}{1-cm}$. Let G be the graph consisting of m cliques, each of order t. Then G is a graph of order $n := mt$, and minimal degree $t-1 \geq cn$. However G has circumference $t = \frac{1}{m}n$. This implies that $\Phi_1(c) \leq \frac{1}{m}$. $\square$

The remainder of this chapter will focus mainly on the value of Φ_2. For $c \geq \frac{3}{4}$, Theorem 3.3 implies that $\Phi_2(c) \geq \frac{2}{3}$. However, the graph $F_{2t,t}$ shows that $\Phi_2(c) \leq \frac{2}{3}$ for all c. In fact, the following theorem will give the value of $\Phi_2(c)$ for all $c \in (0,1)$.

Theorem 3.7. *The function Φ_2 takes the following values.*

$$\Phi_2(c) = \begin{cases} \frac{2}{3} & \text{for } c \in [\frac{3}{4}, 1) \\ \min\{\frac{1}{2}, \frac{2}{3}c\} & \text{for } c \in \left[\frac{5}{9}, \frac{3}{4}\right) \\ \min\{\frac{1}{3}, \frac{2}{3}c\} & \text{for } c \in \left(\frac{7}{16}, \frac{5}{9}\right) \\ \min\{\frac{1}{m}, \frac{2}{3}c\} & \text{for } c \in \left(\frac{2m+1}{(m+1)^2}, \frac{2m-1}{m^2}\right] \text{ for any integer } m \geq 4. \end{cases}$$

The function Φ_2 is illustrated in Figure 3.2.

First note that, for any $c \in (0, \frac{7}{16}]$, there is a unique integer $m_c \geq 4$ such that $c \in \left(\frac{2m_c+1}{(m_c+1)^2}, \frac{2m_c-1}{m_c^2}\right]$, and hence this theorem does give the value of Φ_2 for all $c \in (0,1)$. Also note that, for any $m \geq 2$, Theorem 3.7 implies that

$$\Phi_2(c) = \min\{\frac{1}{m}, \frac{2}{3}c\} \quad \text{for } c \in \left(\frac{2m+1}{(m+1)^2}, \frac{2m-1}{m^2}\right),$$

and that the difference in the last three cases is only that Φ_2 is right-continuous at $\frac{3}{4}$ and $\frac{5}{9}$. By considering the minimum of $\frac{2}{3}c$ and $\frac{1}{m_c}$, it can be seen that $\Phi_2(c) = \frac{2}{3}c$ for $\frac{5}{9} \leq c < \frac{3}{4}$, for $\frac{7}{16} < c \leq \frac{1}{2}$ and for $\frac{9}{25} < c \leq \frac{3}{8}$, and no other values of c.

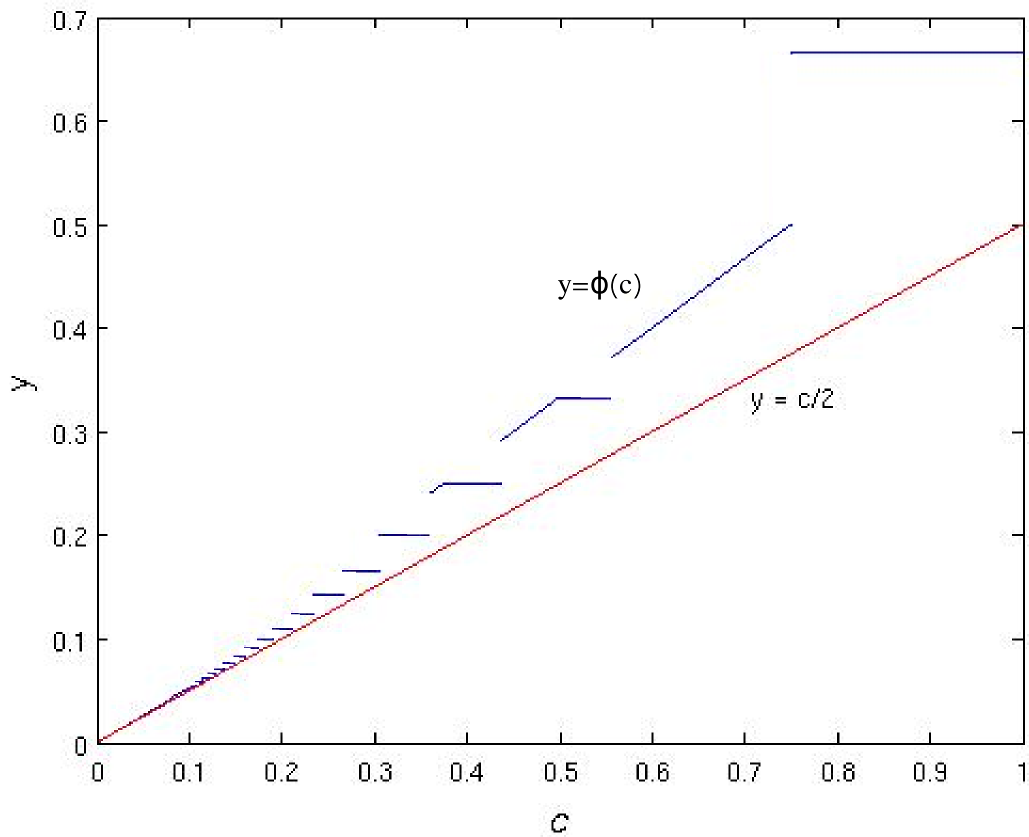


Figure 3.2: The function Φ_2.

For any integer $m \geq 2$, Theorem 3.7 implies that Φ_2 has a discontinuity at $\frac{2m-1}{m^2}$. Compare this to Corollary 3.6, where it can be seen that Φ_1 has a discontinuity at $\frac{1}{m}$ for any integer m. These discontinuities arise for similar reasons. A good way to obtain a graph with high minimal degree and short circumference is to take a graph which is a union of m cliques. If the cliques are taken to have roughly the same order, then this gives a graph with circumference approximately $\frac{1}{m}n$. However, in the 1-edge-coloured case a graph with m components has minimal degree less than $\frac{1}{m}n$, which explains why Φ_1 has a discontinuity at $\frac{1}{m}$.

In the 2-edge-coloured case, suppose that the red and blue graphs consist of m cliques of order approximately $\frac{1}{m}n$ each. It can be checked that the minimal degree is maximised in this case by letting each red clique intersect each blue clique in approximately $\frac{1}{m^2}n$ vertices, giving a graph with minimal degree approximately $\frac{2m-1}{m^2}$. Hence for $c > \frac{2m-1}{m^2}$, either the red or blue graph has at most $m-1$ components, explaining the discontinuity of Φ_2 at $\frac{2m-1}{m^2}$.

Note that a graph G with minimal degree at least cn has at least $\frac{c}{2}n^2$ edges. If G has a 2-edge-colouring, either the red or the blue graph has at least $\frac{c}{4}n^2$ edges. Hence Theorem 1.14 implies that G has monochromatic circumference at least $\frac{c}{2}n$. This implies that $\Phi_2(c) \geq \frac{1}{2}c$ for all $c \in (0,1)$. However, for any $c \in (0, \frac{7}{16}]$,

$$\frac{1}{m_c} - \frac{1}{2}c \leq \frac{2\left(m_c+1\right)^2 - \left(2m_c+1\right)m_c}{2m_c\left(m_c+1\right)^2} = \frac{3m_c+2}{2m_c\left(m_c+1\right)^2} = O\left(\frac{1}{m_c^2}\right).$$

As $c \to 0$, the integer m_c becomes large, and $\Phi_c = \frac{1}{m_c} = \frac{1}{2}c\left(1 + O\left(\frac{1}{m_c}\right)\right)$. Hence $\frac{\Phi_c}{\frac{1}{2}c} \to 1$ as $c \to 0$, showing that, as $c \to 0$, the bound implied by Theorem 1.14 becomes tight.

As in the proof of Theorem 3.3, the regularity method will be used to prove Theorem 3.7. In Section 3.2, results about matchings in 2-edge-multi-coloured graphs will be proved. This will lead to the proof of Theorem 3.7, which is given in Section 3.3. Section 3.4 contains some discussion about the monochromatic circumference of an r-edge-coloured graph.

3.2 Matchings in 2-edge-coloured graphs revisited

In this section, the methods of Section 2.5 will be applied to graphs with minimal degree less than $\frac{3}{4}n$. In particular, for a graph G with no large matching, recall Definition 2.27 which defined S_G, T_G and U_G as follows.

Definition 3.8. *Let a and b be integers and G be a graph, which does not contain a matching on at least b vertices.*

- *Choose $S_G \subseteq V(G)$ to be a set such that*

$$q\left(G\left[V(G) - S_G\right]\right) > |S_G| + |V(G)| - b.$$

  If $|V\left(G\right)| < b$, then it is assumed that S_G is empty.

- *For all $i \geq 1$, let $T_{G,i}$ be the set of vertices of G which lie in components of $G\left[V(G) - S_G\right]$ of order i.*

- *Let $T_G = \bigcup_{1 \leq i \leq a} T_{G,i}$.*

- *Let $U_G = V(G) \setminus (S_G \cup T_G)$.*

The first result of the section is the following result, which shows that when a 2-edge-multi-coloured graph does not have a monochromatic component containing a large matching, all of its monochromatic components are small.

Lemma 3.9. *Let $m \geq 2$ be an integer, and let $0 < \delta < \frac{1}{5(m+1)^2}$. Let $c \geq \frac{2m+1}{(m+1)^2} - \delta$, and $\phi = \min\{\frac{2}{3}c, \frac{1}{m}\} - \delta$. Suppose that G is a 2–edge-multi-*

coloured graph of sufficiently large order n, with minimal degree at least cn, which does not contain a monochromatic component containing a matching on at least ϕn vertices. Then both the largest red component and the largest blue component have order less than $\frac{1}{m}n$.

Note that, in particular, under the assumptions of this lemma both the red and the blue graph have at least $m+1$ components.

Proof. Let $m \geq 2$ be an integer, and let $0 < \delta < \frac{1}{5(m+1)^2}$. Let $c \geq \frac{2m+1}{(m+1)^2} - \delta$, and $\phi = \min\{\frac{2}{3}c, \frac{1}{m}\} - \delta$. Let G be a graph of sufficiently large order n, such that $\delta(G) \geq cn$ and let $E(G) = E(R) \cup E(B)$ be a 2-edge colouring. Suppose that there is no component of R or B containing a matching on at least ϕn vertices.

Suppose that R has r components $R_1, \ldots, R_r$ and B has b components $B_1, \ldots B_b$, where it is assumed that $|V(R_i)| \geq |V(R_{i+1})|$ for all $1 \leq i \leq r-1$ and that $|V(B_i)| \geq |V(B_{i+1})|$ for all $1 \leq i \leq b-1$. Define $0 \leq r' \leq r$ to be the unique integer such that $|V(R_i)| \geq \phi n$ for all $1 \leq i \leq r'$ and $|V(R_i)| < \phi n$ for all $r'+1 \leq i \leq r$. Similarly define $0 \leq b' \leq b$.

Note that, if $r' \geq m+1$, then $n \geq \sum_{i=1}^{r'} |V(R_i)| \geq (m+1)\,\phi n$. Hence $\phi \leq \frac{1}{m+1}$, and, as $\delta < \frac{1}{m(m+1)}$, this implies that $\phi < \frac{1}{m} - \delta$. Hence $\min\{\frac{2}{3}c, \frac{1}{m}\} = \frac{2}{3}c$ and so $\phi = \frac{2}{3}c - \delta$. However $\phi \leq \frac{1}{m+1}$ implies that $c \leq \frac{3}{2(m+1)} + \frac{3}{2}\delta$. Recall that $c \geq \frac{2m+1}{(m+1)^2} - \delta$, and so $\frac{2m+1}{(m+1)^2} \leq \frac{3}{2(m+1)} + \frac{5}{2}\delta$. This implies that

$$4m + 2 \leq 3(m+1) + 5\delta\,(m+1)^2 < 3m + 4.$$

This contradicts the assumption that $m \geq 2$. Hence $r' \leq m$, and similarly $b' \leq m$.

By assumption no monochromatic component of G contains a matching on $b = \lceil \phi n \rceil$ vertices. For any red component R_i, recall that S_{R_i} is a set such that

$$q\left(R\left[V(R_i) - S_{R_i}\right]\right) > |S_{R_i}| + |V(R_i)| - \lceil \phi n \rceil ,$$

where S_{R_i} is assumed to be empty for $i > r'$. Also T_{R_i} is the set of vertices in components of $R\left[V(R_i) - S_{R_i}\right]$ of order at most $a := \lceil 3m\delta^{-1} \rceil$, and U_{R_i} is the set $V(R_i) \setminus (S_{R_i} \cup T_{R_i})$. Lemma 2.28 implies that

$$|S_{R_i}| < \frac{1}{2} \lceil \phi n \rceil , \tag{3.1}$$

$$|T_{R_i}| > |V(R_i)| + |S_{R_i}| - \lceil \phi n \rceil - \frac{\delta}{3m} n, \tag{3.2}$$

and

$$2|S_{R_i}| + |U_{R_i}| < \lceil \phi n \rceil + \frac{\delta}{3m} n. \tag{3.3}$$

Note that for any B_j, similar sets S_{B_j}, T_{B_j} and U_{B_j} can be defined, and that analogues of the above equations hold for these sets.

Suppose, for some $1 \leq i \leq r$, that $v \in T_{R_i}$. All but a vertices of the red neighbourhood of v lie in S_{R_i}, and so (3.1) implies that v has red degree less than $\frac{1}{2}\phi n + a + 1$. Similarly, for $1 \leq j \leq b$, any vertex in T_{B_j} has blue degree less than $\frac{1}{2}\phi n + a + 1$.

However, for any $1 \leq i \leq r$, a vertex in U_{R_i} only has red neighbours in $U_{R_i} \cup S_{R_i}$. Hence (3.3) implies that a vertex in U_{R_i} has red degree less than $\lceil \phi n \rceil + \frac{\delta}{3m} n$. Similarly, for $1 \leq j \leq b$, any vertex in U_{B_j} has blue degree less than $\lceil \phi n \rceil + \frac{\delta}{3m} n$.

Suppose that $T_{R_i} \cap \left(T_{B_j} \cup U_{B_j}\right)$ is non-empty for some $1 \leq i \leq r$ and $1 \leq j \leq b$, and let v be a vertex of $T_{R_i} \cap \left(T_{B_j} \cup U_{B_j}\right)$. Then, for n sufficiently large, v has red degree less than $\frac{1}{2}\phi n + a + 1$ and blue degree less than $\phi n + \frac{\delta}{3m} n + 1$. Hence, in G, the degree of v is less than $\left(\frac{3}{2}\phi + \frac{\delta}{3m}\right) n + a + 2$. However $\phi \leq \frac{2}{3}c - \delta$, and so the degree of v is less than $\left(c - \frac{9m-2}{6m}\delta\right) n + a + 2$. For sufficiently large n, this contradicts $\delta(G) \geq cn$. Hence, for all $1 \leq i \leq r$, the set T_{R_i} is contained in $\bigcup_{1 \leq j \leq b'} S_{B_j}$ and similarly, for all $1 \leq j \leq b$, the set T_{B_j} is contained in $\bigcup_{1 \leq i \leq r'} S_{R_i}$.

Note that this implies that $\sum_{i=1}^{r'} |T_{R_i}| \leq \sum_{j=1}^{b'} |S_{B_j}|$ and similarly that $\sum_{i=1}^{b'} |T_{B_j}| \leq \sum_{j=1}^{r'} |S_{R_i}|$. However (3.2) implies that

$$\sum_{j=1}^{b'} \left(|T_{B_j}| - |S_{B_j}|\right) > \sum_{j=1}^{b'} \left(|V(B_j)| - \lceil \phi n \rceil - \frac{\delta}{3m} n\right)$$

and

$$\sum_{i=1}^{r'} \left(|T_{R_i}| - |S_{R_i}|\right) > \sum_{i=1}^{r'} \left(|V(R_i)| - \lceil \phi n \rceil - \frac{\delta}{3m} n\right).$$

Together these equations imply that

$$\begin{aligned} 0 &\geq \sum_{i=1}^{r'} \left(|T_{R_i}| - |S_{R_i}|\right) + \sum_{j=1}^{b'} \left(|T_{B_j}| - |S_{B_j}|\right) \\ &> \sum_{j=1}^{b'} \left(|V(B_j)| - \left(\lceil \phi n \rceil + \frac{\delta}{3m} n\right)\right) + \sum_{i=1}^{r'} \left(|V(R_i)| - \left(\lceil \phi n \rceil + \frac{\delta}{3m} n\right)\right). \end{aligned}$$

Hence, as $|V(R_{r'})| \geq \lceil \phi n \rceil$ and $|V(B_{b'})| \geq \lceil \phi n \rceil$

$$\begin{aligned}|V(R_1)| &< (r'+b')\left(\lceil \phi n \rceil + \frac{\delta}{3m}n\right) - \sum_{j=1}^{b'} |V(B_j)| - \sum_{i=2}^{r'} |V(R_i)| \\ &\leq \lceil \phi n \rceil + (b'+r')\,\frac{\delta}{3m}n.\end{aligned}$$

However $\phi \leq \frac{1}{m} - \delta$ and $b' + r' \leq 2m$, and so

$$|V(R_1)| \leq \frac{1}{m}n + 1 - \frac{1}{3}\delta n < \frac{1}{m}n.$$

Similarly $|V(B_1)| < \frac{1}{m}n$. □

The previous lemma can now be used to prove the following result, which is one of the main results of this section.

Lemma 3.10. *Let $m \geq 2$ be an integer, and let $0 < \delta < \frac{1}{5(m+1)^2}$. There exists $n_1 = n_1(m, \delta)$ such that the following holds.*

Let $c \geq \frac{2m+1}{(m+1)^2}$, and $\phi = \min\{\frac{2}{3}c, \frac{1}{m}\} - \delta$. If G is a 2-edge-multi-coloured graph of order $n \geq n_1$, with minimal degree at least cn, then there is some monochromatic component of G containing a matching on at least ϕn vertices.

Proof. Let $m \geq 2$ be an integer, and let $0 < \delta < \frac{1}{5(m+1)^2}$. Let $c \geq \frac{2m+1}{(m+1)^2}$, and $\phi = \min\{\frac{2}{3}c, \frac{1}{m}\} - \delta$. Let G be a graph of sufficiently large order n, such that $\delta(G) \geq cn$ and let $E(G) = E(R) \cup E(B)$ be a 2-edge colouring. Suppose that there is no component of R or B containing a matching on at least ϕn vertices, this will be referred to as the *main assumption.*

Suppose that R has r components $R_1, \ldots, R_r$ and B has b components

$B_1, \dots B_b$, where it is assumed that $|V(R_i)| \geq |V(R_{i+1})|$ for all $1 \leq i \leq r-1$ and that $|V(B_i)| \geq |V(B_{i+1})|$ for all $1 \leq i \leq b-1$. By the main assumption Lemma 3.9 implies that $|V(R_1)| < \frac{1}{m}n$ and $|V(B_1)| < \frac{1}{m}n$. Hence, both R and B have at least $m+1$ components.

This implies that $|V(R_r)| \leq \frac{1}{m+1}n$ and $|V(B_b)| \leq \frac{1}{m+1}n$. Let J be the set of $1 \leq j \leq b$ such that $V(B_j)$ intersects $V(R_r)$ and I be the set of $1 \leq i \leq r$ such that $V(R_i)$ intersects $V(B_b)$. Note that, if $V(R_r)$ is contained in $V(B_j)$ for some j, the neighbourhood of any vertex in R_r is contained in $V(B_j)$. This implies that $|V(B_j)| \geq cn$, contradicting the fact that $|V(B_1)| < \frac{1}{m}n$. Hence it may be assumed that $|J| > 1$, and similarly that $|I| > 1$.

Let $j \in J$, and consider $v \in V(B_j) \cap V(R_r)$. The only neighbours of v are vertices in $V(B_j) \cup V(R_r)$, which implies that

$$|V(B_j)| + |V(R_r) \setminus V(B_j)| \geq cn + 1.$$

Summing this equation for all $j \in J$ implies that

$$\sum_{j \in J} |V(B_j)| + (|J| - 1)\,|V(R_r)| \geq c|J|n + |J|$$

and hence

$$c|J|n < \sum_{j=1}^{|J|} |V(B_j)| + (|J| - 1)\,|V(R_r)|. \tag{3.4}$$

Note that $\sum_{j=1}^{|J|} |V(B_j)| \leq n$ and recall that $|V(R_r)| \leq \frac{1}{m+1}n$ and $c \geq$

$\frac{2m+1}{(m+1)^2}$. Hence (3.4) implies that

$$\frac{(2m+1)\,|J|}{(m+1)^2}n < n + \frac{|J|-1}{m+1}n$$

and thus

$$|J|\,(2m+1) < (m+1)^2 + (|J|-1)\,(m+1)\,.$$

This implies that $|J| < m+1$. Thus $|J| \leq m < b$. Similarly, $|I| \leq m$.

Note that

$$(m+1-|J|)\,|V(B_b)| \leq \sum_{j=|J|+1}^{b} |V(B_j)| = n - \sum_{j=1}^{|J|} |V(B_j)|.$$

Also (3.4) implies that

$$n - \sum_{j=1}^{|J|} |V(B_j)| < (1-c|J|)\,n + (|J|-1)\,|V(R_r)|,$$

and hence

$$(m+1-|J|)\,|V(B_b)| < \left(1 - \frac{(2m+1)\,|J|}{(m+1)^2}\right)n + (|J|-1)\,|V(R_r)|.$$

As $|V(R_r)| \leq \frac{1}{m+1}n$, this implies that

$$\begin{aligned}(m+1-|J|)\,|V(B_b)| &< \left(1 - \frac{(2m+1)\,|J|}{(m+1)^2}\right)n + \frac{(|J|-1)}{m+1}n \\ &= \frac{(m+1)^2 - m|J| - (m+1)}{(m+1)^2}n \\ &= \frac{m\,(m+1-|J|)}{(m+1)^2}n.\end{aligned}$$

Hence, as $|J| \leq m$, this implies that $|V(B_b)| \leq \frac{m}{(m+1)^2}n$. Similarly, $|V(R_r)| \leq \frac{m}{(m+1)^2}n$.

Recall that $|V(B_1)| < \frac{1}{m}n$, which implies that $\sum_{j=1}^{|J|} |V(B_j)| < \frac{|J|}{m}n$. Hence (3.4) implies that

$$\frac{(2m+1)\,|J|}{(m+1)^2}n < \frac{|J|}{m}n + \frac{(|J|-1)\,m}{(m+1)^2}n.$$

Thus

$$m\,(2m+1)\,|J| < |J|\,(m+1)^2 + (|J|-1)\,m^2$$

which implies that $|J| > \frac{m^2}{m+1}$. Hence $|J| > m-1$, and so $|J| = m$. Similarly, $|I| = m$.

Equation (3.4) implies that

$$\sum_{j=1}^{m} |V(B_j)| > \frac{(2m+1)\,m}{(m+1)^2}n - (m-1)\,|V(R_r)| \tag{3.5}$$

As $|V(R_r)| \leq \frac{m}{(m+1)^2}n$, this implies that $\sum_{j=1}^{m} |V(B_j)| > \frac{m(m+2)}{(m+1)^2}n$ and so $|B_b| < \frac{1}{(m+1)^2}n$. Similarly $|R_r| < \frac{1}{(m+1)^2}n$. Hence (3.5) implies that

$$\sum_{j=1}^{m} |V(B_j)| > \frac{(2m+1)\,m - (m-1)}{(m+1)^2}n \geq n,$$

a contradiction. □

The final result of this section is similar to Lemma 3.10, but considers the case when $m = 2$ and c is a little below $\frac{5}{9}$. Note that this is necessary for finding the value of $\Phi_2(\frac{5}{9})$ which behaves differently to $\Phi_2(c)$ for all other $c < \frac{3}{4}$.

Lemma 3.11. *Let $0 < \delta < \frac{1}{45}$. There exists $n_2 = n_2(\delta)$ such that, if G is a 2-edge-multi-coloured graph of order $n \geq n_2$, with minimal degree at least $\left(\frac{5}{9} - \delta\right) n$, then one of the following occurs:*

- *there is some monochromatic component of G containing a matching on at least $\left(\frac{10}{27} - \frac{5}{3}\delta\right) n$ vertices;*

- *G has exactly three red components and exactly three blue components; furthermore the intersection of the vertex set of any red component and the vertex set of any blue component has order at most $\left(\frac{1}{9} + 7\delta\right) n$*

Proof. Let $0 < \delta < \frac{1}{45}$. Let G be a graph of sufficiently large order n, such that $\delta(G) \geq cn$ and let $E(G) = E(R) \cup E(B)$ be a 2-edge colouring. Suppose that there is no component of R or B containing a matching on at least $\left(\frac{10}{27} - \frac{5}{3}\delta\right) n$ vertices, this will be referred to as the *main assumption.*

Suppose that R has r components $R_1, \ldots, R_r$ and B has b components $B_1, \ldots B_b$, where it is assumed that $|V(R_i)| \geq |V(R_{i+1})|$ for all $1 \leq i \leq r-1$ and that $|V(B_i)| \geq |V(B_{i+1})|$ for all $1 \leq i \leq b-1$. Lemma 3.9 implies that both R_1 and B_1 have order less than $\frac{1}{2}n$, and so r and b are both at least 3.

Let J be the set of $1 \leq j \leq b$ such that $V(B_j)$ intersects $V(R_r)$ and I be the set of $1 \leq i \leq r$ such that $V(R_i)$ intersects $V(B_b)$. If $V(R_r)$ is contained in $V(B_j)$ for some j, the neighbourhood of any vertex in R_r is contained in $V(B_j)$. This implies that $|V(B_j)| \geq cn$, contradicting the fact that $|V(B_1)| < \frac{1}{2}n$. Hence it may be assumed that $|J| > 1$, and similarly that $|I| > 1$.

Using the same method as that used to deduce (3.4) in the previous lemma, it can be shown that

$$\left(\frac{5}{9}-\delta\right)|J|n < \sum_{j=1}^{|J|}|V(B_j)| + (|J|-1)\,|V(R_r)|. \tag{3.6}$$

Suppose first that $r \geq 4$ and so $|V(R_r)| \leq \frac{1}{4}n$. As $\sum_{j=1}^{|J|}|V(B_j)| \leq n$, this implies that

$$\left(\frac{5}{9}-\delta\right)|J|n < n + (|J|-1)\,\frac{1}{4}n.$$

Hence $|J|\left(\frac{11}{36}-\delta\right) < \frac{3}{4}$. As $\delta < \frac{1}{45}$, this implies that $|J| < \frac{45}{17}$ and hence $|J| = 2$. But then (3.6) implies that

$$\begin{aligned} |V(B_1)| + |V(B_2)| &> \left(\frac{10}{9}-2\delta\right)n - |V(R_r)| \\ &\geq \left(\frac{31}{36}-2\delta\right)n. \end{aligned} \tag{3.7}$$

Hence $|V(B_b)| < \left(\frac{5}{36}+2\delta\right)n \leq \frac{1}{4}n$. This implies that $|I| = 2$ and so

$$|V(R_1)| + |V(R_2)| + |V(B_b)| > \left(\frac{10}{9}-2\delta\right)n.$$

Combined with (3.7), this contradicts $\sum_j |V(B_j)| + \sum_i |V(R_i)| = 2n$.

Hence $r = 3$, and similarly $b = 3$. Let $W_{i,j} = V(R_i) \cap V(B_j)$. A vertex in $W_{i,j}$ has neighbours only in $V(R_i) \cup V(B_j)$. This implies that, for any non-empty $W_{i,j}$,

$$|V(R_i)| + |V(B_j)| - |W_{i,j}| \geq \left(\frac{5}{9}-\delta\right)n. \tag{3.8}$$

Suppose that some $W_{i,j}$ is empty. As no red component is contained in a blue component (or vice versa), each of the sets $W_{i,j'}$ for $j' \neq j$ and $W_{i',j}$ for $i' \neq i$ are non-empty. Hence (3.8) implies that, for $j' \neq j$,

$$|V(R_i)| + |V(B_{j'})| - |W_{i,j'}| \geq \left(\frac{5}{9} - \delta\right) n.$$

Hence, as $V(R_i) = \cup_{j' \neq j} W_{i,j'}$,

$$\begin{aligned} |V(R_i)| + \sum_{j' \neq j} |V(B_{j'})| &= 2|V(R_i)| + \sum_{j' \neq j} (|V(B_{j'})| - |W_{i,j'}|) \\ &\geq \left(\frac{10}{9} - 2\delta\right) n. \end{aligned}$$

Similarly

$$|V(B_j)| + \sum_{i' \neq i} |V(R_{i'})| \geq \left(\frac{10}{9} - 2\delta\right) n.$$

This contradicts $\sum_j |V(B_j)| + \sum_i |V(R_i)| = 2n$.

Hence each $W_{i,j}$ is non-empty. For any i, summing (3.8) for all j implies that

$$2|V(R_i)| + n \geq \left(\frac{5}{3} - 3\delta\right) n.$$

Hence $|V(R_i)| \geq \left(\frac{1}{3} - \frac{3}{2}\delta\right) n$ for any $1 \leq i \leq 3$. Similarly $|V(B_j)| \geq \left(\frac{1}{3} - \frac{3}{2}\delta\right) n$ for any $1 \leq j \leq 3$. This implies that $|V(R_i)| \leq \left(\frac{1}{3} + 3\delta\right) n$ for any $1 \leq i \leq 3$ and $|V(B_j)| \leq \left(\frac{1}{3} + 3\delta\right) n$ for any $1 \leq j \leq 3$. Hence (3.8) implies that

$$|W_{i,j}| \leq |V(R_i)| + |V(B_j)| - \left(\frac{5}{9} - \delta\right) n \leq \left(\frac{1}{9} + 7\delta\right) n$$

for all i and j, as claimed. $\square$

3.3 The monochromatic circumference of a 2-edge-coloured graph with minimal degree at least cn

In this section, a proof of Theorem 3.7 will be given. Firstly, Lemma 3.10 and Lemma 3.11 will be used to obtain the following result about the existence of long monochromatic cycles.

Lemma 3.12. *Let $m \geq 2$ be an integer, and let $c > \frac{2m+1}{(m+1)^2}$. Let δ be a positive constant with*

$$\delta < \min \left\{ c - \frac{2m+1}{(m+1)^2}, \frac{1}{5(m+1)^2} \right\}.$$

Let G be a 2-edge-coloured graph with sufficiently large order n and minimal degree at least cn. Then G contains a monochromatic cycle of length at least $\left(\min\{\frac{2}{3}c, \frac{1}{m}\} - 2\delta\right) n$.

Now let $0 < \delta < \frac{1}{540}$. If G is a graph with sufficiently large order n and minimal degree at least $\frac{5}{9}n$, it contains a monochromatic cycle of length at least $\left(\frac{10}{27} - 2\delta\right) n$.

Proof. Let $m \geq 2$ be an integer, and let $c > \frac{2m+1}{(m+1)^2}$. Let δ be a positive constant with

$$\delta < \min\left\{c - \frac{2m+1}{(m+1)^2}, \frac{1}{5(m+1)^2}\right\}.$$

Let $\phi = \min\{\frac{2}{3}(c-\delta), \frac{1}{m}\} - \delta$. Let G be a 2-edge-coloured graph with order n and minimal degree at least cn.

Let $\eta = \min\{\frac{2}{27\phi}, \frac{1}{5}\}\delta$ and $\epsilon = \epsilon(\eta)$, where $\epsilon(\eta)$ is as defined in Corollary 2.25 (note that $\epsilon(\eta) < \eta$). Let $m = n_1$, where n_1 is as defined in Lemma 3.10. Apply the Regularity Lemma (Theorem 2.9) to G, with parameters $r = 2$, m, η and ϵ. Hence there is a partition $\mathcal{P} = (V_i)_{i=0}^k$ of $V(G)$ and a subgraph G' of G with the properties given in that theorem. Let H be the reduced graph defined as in Definition 2.11 by G' and $\mathcal{P}$, with 2-edge-multi-colouring $E(H) = E(R_H) \cup E(B_H)$.

The Regularity Lemma implies that $\delta(G') \geq (c - 2\eta - \epsilon)n$. Suppose that $\delta(H) < (c-\delta)k$ and let v_i be a vertex of H such that $\deg_H(v_i) < (c-\delta)k$. In G', a vertex $v \in V_i$ has neighbours only in V_0, or in V_j for those j such that v_iv_j is an edge of H. Hence, recalling that for $i \neq 0$ all sets V_i have the same order,

$$\deg_{G'}(v) < (c-\delta)k|V_1| + |V_0| \leq (c - \delta + \epsilon)n.$$

This is a contradiction, as $\delta > 2\eta + 2\epsilon$. Hence $\delta(H) \geq (c-\delta)k$.

Note that $c - \delta > \frac{2m+1}{(m+1)^2}$. Hence Lemma 3.10 implies that H has a monochromatic component C which contains a matching M on at least ϕk vertices. Suppose that C is a component of R_H. Corollary 2.25 implies that R_G contains a cycle of length at least $(1-\epsilon)\left(1-\epsilon-\frac{1}{4}\eta\right)|V_1|\phi k$. However

$k|V_1| = n - |V_0| \geq (1-\epsilon)\, n$, and

$$\phi\,(1-\epsilon)^2\left(1-\epsilon-\frac{1}{4}\eta\right) \geq \phi\left(1-3\epsilon-\frac{1}{4}\eta-\frac{1}{4}\epsilon^2\eta-\epsilon^3\right) \geq \phi-\frac{1}{3}\delta.$$

Hence R_G contains a cycle of length at least $\left(\min\{\frac{2}{3}c, \frac{1}{m}\} - 2\delta\right) n$. Similarly, if C is a component of B_H, then B_G contains a cycle of length at least $\left(\min\{\frac{2}{3}c, \frac{1}{m}\} - 2\delta\right) n$.

Now let $0 < \delta < \frac{1}{540}$. Suppose that G is a graph with sufficiently large order n and minimal degree at least $\frac{5}{9}n$. Assume that G does not contain a monochromatic cycle of length at least $\left(\frac{10}{27} - 2\delta\right) n$.

Apply the Regularity Lemma to G as above to obtain a 2-edge-multi-coloured reduced graph H with minimal degree at least $\left(\frac{5}{9} - \delta\right) n$. Hence Lemma 3.11 implies that one of the following occurs:

- H has a monochromatic component containing a matching on at least $\left(\frac{10}{27} - \frac{5}{3}\delta\right) n$ vertices

- H has exactly three red components and exactly three blue components, such that the intersection of the vertex set of any red component and the vertex set of any blue component has order at most $\left(\frac{1}{9} + 7\delta\right) k$.

If H has a monochromatic component containing a matching on at least $\left(\frac{10}{27} - \frac{5}{3}\delta\right) k$ vertices, then Corollary 2.25 implies that G contains a monochromatic cycle of length at least $\left(\frac{10}{27} - 2\delta\right) n$. This contradicts the earlier assumption.

Hence it may be assumed that H has exactly three red components and three blue components and that the intersection of the vertex set of any red component and the vertex set of any blue component has order at most $\left(\frac{1}{9}+3\delta\right)k$. Partition $V(G)\setminus V_0$ into three sets X_1, X_2 and X_3, where each X_i contains the clusters corresponding to vertices of exactly one of the red components of H. Similarly partition $V(G)\setminus V_0$ into three sets Y_1, Y_2 and Y_3, where each Y_j contains the clusters corresponding to vertices of exactly one of the blue components of H. Let $U_{i,j}=X_i\cap Y_j$. Note that for all i and j, the set $U_{i,j}$ has order at most $\left(\frac{1}{9}+7\delta\right)n$, and that the sets $U_{i,j}$ partition the set $V(G)\setminus V_0$. The proof proceeds with the following fact.

Fact 3.13. *Let $1\le i\le 3$. For any two vertices u and v in X_i, the graph $R_G[X_i]$ contains a Hamilton path from u to v. Similarly, for all $1\le j\le 3$, and for any two vertices u and v in Y_j, the graph $B_G[Y_i]$ contains a Hamilton path from u to v.*

Proof. The result will be proved for X_1, all other cases being similar. First, a bound on the minimal degree of $R_G[X_1]$ will be calculated. Recall that G' is the subgraph of G defined by the Regularity Lemma. In G', a vertex in $U_{1,i}$ has no neighbours outside $X_1\cup Y_i\cup V_0$, and has at least $\left(\frac{5}{9}-\delta\right)n$ neighbours in total. However,

$$|X_1\cup Y_i|+|V_0|\le\left(\frac{5}{9}+35\delta\right)n+\epsilon n\le\left(\frac{5}{9}+36\delta\right)n.$$

Hence, in G', a vertex in $U_{1,i}$ has at most $37\delta n$ non-neighbours in X_1.

However, the blue neighbours in G' of a vertex in $U_{1,i}$ are contained in Y_i. Hence, in G', a vertex in $U_{1,i}$ has at most $|U_{1,i}|$ blue neighbours in X_1. This implies that, in G', a vertex in $U_{1,i}$ has at least $|X_1|-|U_{1,i}|-37\delta n$ red neighbours in X_1. As each $U_{i,j}$ has order at most $\left(\frac{1}{9}+7\delta\right)n$, this implies that $R_{G'}[X_1]$ has minimal degree at least $|X_1|-\left(\frac{1}{9}+44\delta\right)n$. As $R_{G'}$ is a subgraph of R_G, this implies that $R_G[X_1]$ has minimal degree at least $|X_1|-\left(\frac{1}{9}+44\delta\right)n$.

Let u and v be vertices of X_1, and let $X_1' = X_1 \setminus \{u,v\}$. As V_0 contains at most ϵn vertices, and each $U_{i,j}$ has order at most $\left(\frac{1}{9}+7\delta\right)n$, the set X_1' has order at least $\left(\frac{1}{3}-43\delta\right)n$. Hence $R_G[X_1']$ has minimal degree at least $|X_1'|-\left(\frac{1}{9}+44\delta\right)n \geq \frac{1}{2}|X_1'|$.

Theorem 1.3 implies that $R_G[X_1']$ contains a Hamilton cycle C. However both u and v have at least $|X_1'|-\left(\frac{1}{9}+20\delta\right)n \geq \frac{1}{2}|X_1'|+1$ red neighbours in X_1' and so there are vertices u' and v' of X_1' such that uu' and vv' are both red edges and $u'v' \in E(C)$. Hence the remainder of C forms a Hamilton path from u' to v' in $R_G[X_1']$. Adding the two edges uu' and vv' to this path gives a Hamilton path from u to v in $R_G[X_i]$. $\square$

Suppose that R_G contains two disjoint paths P and P' from X_1 to X_2. By taking subpaths if necessary, it may be assumed that both paths contain no internal vertices in $X_1 \cup X_2$. Fact 3.13 implies that $R_G[X_1]$ contains a Hamilton path P_1 between the endpoint of P in X_1 and the endpoint of P' in X_1. Similarly, $R_G[X_2]$ contains a Hamilton path P_2 between the endpoint of P in X_2 and the endpoint of P' in X_2. Concatenating the paths P, P_1, P' and P_2 gives a cycle in R_G with length at least $|X_1|+|X_2| \geq \left(\frac{2}{3}-10\delta\right)n$.

Hence it may be assumed that R_G does not contain two disjoint paths from X_1 to X_2. Theorem 2.17 implies that there is a vertex x_1 such that $R_G[V(G)\setminus\{x_1\}]$ contains no paths from X_1 to X_2. Thus, in $R_G\left[V(G)\setminus\{x_1\}\right]$, there are no paths from X_3 to at least one of X_1 and X_2. Without loss of generality, there are no paths in $R_G[V(G)\setminus\{x_1\}]$ from X_1 to X_3. Considering the original graph R_G, the above argument shows that there is a vertex x_2 such that $R_G[V(G)\setminus\{x_2\}]$ contains no paths from X_2 to X_3. Similarly, there are vertices y_1 and y_2 such that $B_G[V(G)\setminus\{y_1\}]$ contains no paths from Y_1 to $Y_2\cup Y_3$, and $B_G[V(G)\setminus\{y_2\}]$ contains no paths from Y_2 to Y_3.

Let $S=\{x_1,x_2,y_1,y_2\}$ (note that these vertices need not be distinct). Note that $G[V(G)\setminus S]$ contains no blue paths between any two of Y_1, Y_2 and Y_3 and no red paths between any two of X_1, X_2 and X_3. Considering the minimal degree of G, and by adding vertices of $V_0\setminus S$ to an appropriate $U_{i,j}$ if necessary, it may be assumed that X_1, X_2 and X_3 are the vertex sets of components of $R_G[V(G)\setminus S]$, and that Y_1, Y_2 and Y_3 are the vertex sets of components of $B_G[V(G)\setminus S]$.

For $1\le i,j\le 3$, let $a_{i,j}$ be the number of vertices in S whose neighbourhood in G intersects $U_{i,j}$. Then a vertex in $U_{i,j}$ has neighbours only in $X_i\cup Y_j\cup S$, and has at most $a_{i,j}$ neighbours in S. Noting that a vertex may not be adjacent to itself, this implies that

$$\frac{5}{9}n\le |X_i|+|Y_j|-1-|U_{i,j}|+a_{i,j}.$$

Summing the above equation over all i and j yields

$$\begin{aligned}5n&\le 3\left(\sum_i|X_i|+\sum_j|Y_j|\right)-9-\sum_{i,j}|U_{i,j}|+\sum_{i,j}a_{i,j}\\&\le 5\,(n-|S|)-9+\sum_{i,j}a_{i,j}.\end{aligned}$$

Hence $\sum_{i,j}a_{i,j}\ge 5|S|+9$. As each $a_{i,j}$ is at most $|S|$, this implies that S is a set of three or four elements.

For a vertex $v \in S$, let b_v be the number of sets $U_{i,j}$ such that the neighbourhood of v in G intersects $U_{i,j}$. Suppose that x_1 is distinct from each of x_2, y_1 and y_2. Then $x_1 \in V(G) \setminus \{x_2\}$ and hence, without loss of generality, has no red neighbours in X_3. However, $x_1 \in V(G) \setminus \{y_1, y_2\}$, and so has no blue neighbours in at least two of the sets Y_1, Y_2 and Y_3. Assuming that x_1 has no blue neighbours in $Y_1 \cup Y_2$, it has no neighbours of either colour in $U_{3,1} \cup U_{3,2}$. Hence $b_{x_1} \leq 7$. Similar results hold if any of x_2, y_1 and y_2 are distinct from each of the other vertices x_1, x_2, y_1 and y_2.

If S contains three distinct elements, two of x_1, x_2, y_1 and y_2 are the same vertex and the other two are distinct from that one and each other. Hence $\sum_{v\in S} b_v \leq 9+7+7 = 5|S|+8$. Similarly, if S contains four distinct elements then $\sum_{v\in S} b_v \leq 28 = 5|S| + 8$. However, by a double counting argument, $\sum_{v\in S} b_v = \sum_{i,j} a_{i,j}$, and so this contradicts $\sum_{i,j} a_{i,j} \geq 5|S| + 9$. $\square$

Theorem 3.7 can now be proved.

Proof of Theorem 3.7. Firstly, suppose that $c \geq \frac{3}{4}$, and that $0 < \delta' < \frac{1}{294}$. Let $\phi = \frac{2}{3} - \delta'$. Let G be a 2-edge-coloured graph with minimal degree at least $cn \geq \frac{3}{4}n$. Theorem 3.3 implies that G has monochromatic circumference at least $\left\lceil \left(\frac{2}{3} - \delta'\right) n \right\rceil \geq \phi n$. Hence $\Phi_c \geq \phi$. In particular, this implies that $\Phi_c \geq \frac{2}{3}$, for all $c \geq \frac{3}{4}$.

Let $m \geq 2$ be an integer, and let $c > \frac{2m+1}{(m+1)^2}$. Let δ be a positive constant with

$$\delta < \min\left\{c - \frac{2m+1}{(m+1)^2}, \frac{1}{5\,(m+1)^2}\right\}.$$

Let G be a 2-edge-coloured graph with sufficiently large order n and minimal degree at least cn. Then Lemma 3.12 implies that G contains a monochromatic cycle of length at least $\left(\min\{\frac{2}{3}c, \frac{1}{m}\} - 2\delta\right) n$. Hence $\Phi(c) \geq \min\{\frac{2}{3}c, \frac{1}{m}\}$ as required.

Finally, suppose that G is a graph with sufficiently large order n and minimal degree at least $\frac{5}{9}n$. Lemma 3.12 implies that G contains a monochromatic cycle of length at least $\left(\frac{10}{27} - 2\delta\right) n$, for any $0 < \delta < \frac{1}{540}$. Hence $\Phi\left(\frac{5}{9}\right) \geq \frac{10}{27}$.

To complete the proof, edge-colourings of graphs will be defined which imply an upper bound on Φ_c. Recall that $F_{2t,t}$ is a graph with $3t$ vertices, minimal degree $3t-1$ and monochromatic circumference $2t$. Let $c \in \left[\frac{3}{4}, 1\right)$ and choose t large enough so that $3t \geq \frac{1}{1-c}$. Then $F_{2t,t}$ has minimal degree at least $c|F_{2t,t}|$ and monochromatic circumference $\frac{2}{3}|F_{2t,t}|$. This implies that $\Phi_c \leq \frac{2}{3}$ for all $c \in \left[\frac{3}{4}, 1\right)$.

For the remaining graph constructions, the following notation will be used. For a set A, the complete graph with vertex set A is denoted by K_A, while for a family of disjoint sets $A_1, \ldots, A_m$, the complete m-partite graph whose parts are the A_i will be denoted $K_{A_1,\ldots,A_m}$. For a graph G, the complement $\overline{G}$ of G is the graph with the same vertex set as G whose edges are precisely the pairs of distinct vertices which are not edges of G. Also, for graphs G and H, the graph $G \cup H$ is obtained by taking the disjoint union of a copy of G and a copy of H, while the graph $G + H$ is obtained from $G \cup H$ by adding all edges between the copy of G and the copy of H.

Let $m \geq 2$ be an integer. Let $\{A_{i,j} : 1 \leq i \leq m, 1 \leq j \leq m\}$ be a family of sets of order t. Let $G_{m,t}$ be the graph with vertex set $\bigcup_{i,j} A_{i,j}$ whose edge set is the union of the edge sets of the graphs R and B defined below.

- $B = \bigcup_{j=1}^{m} K_{A_{j,1},A_{j,2},\ldots,A_{j,m}}$.

- $R = \bigcup_{i=1}^{m} K_{A_{1,i} \cup A_{2,i} \cup \cdots \cup A_{m,i}}$.

The graph $G_{3,t}$ is illustrated in Figure 3.3. Define $G^*_{m,t}$ by adding two vertices u and v to $G_{m,t}$, and letting R contain all edges from u to $\bigcup_{i,j} A_{i,j}$ and B contain all edges from v to $\bigcup_{i,j} A_{i,j}$.

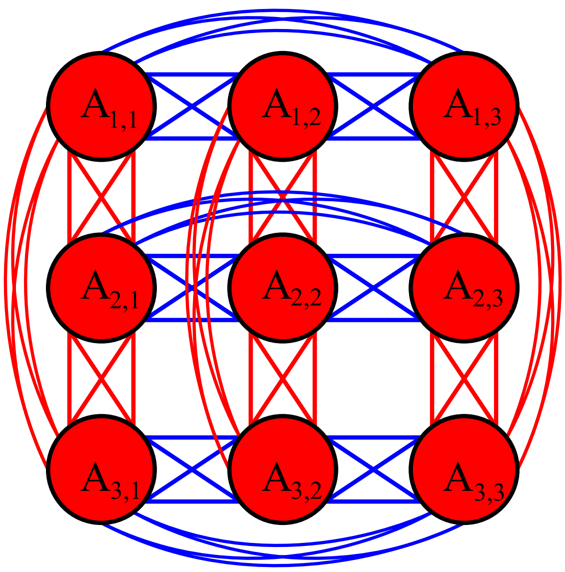


Figure 3.3: The graph $G_t^{(3)}$.

Note that $|G^*_{m,t}| = m^2t + 2$ and $G^*_{m,t}$ has minimal degree $(2m-1)t + 1$. However, the graph R consists of m cliques of order $mt + 1$, which mutually intersect in the vertex u. Hence R has circumference $mt+1$, as does B. Thus $G^*_{m,t}$ has monochromatic circumference $mt + 1$.

Let $m \geq 4$ and suppose that $c \leq \frac{2m-1}{m^2}$. Let $\phi > \frac{1}{m}$, and let t be an integer such that $\frac{m-2}{m} < \left(\phi - \frac{1}{m}\right)(m^2t + 2)$. Then $G^*_{m,t}$ has minimal degree at least

$$(2m-1)\,t + 1 = \frac{2m-1}{m^2}\left(m^2t + 2\right) + \frac{m^2 - 4m + 2}{m^2} \geq c\left|G^*_{m,t}\right|,$$

where the second inequality used that $m \geq 4$. Also, $G^*_{m,t}$ has monochromatic circumference $\frac{1}{m}|G^*_{m,t}| + \frac{m-2}{m} < \phi|G^*_{m,t}|$. Hence $\Phi(c) \leq \phi$ for all $\phi > \frac{1}{m}$, and so $\Phi(c) \leq \frac{1}{m}$.

Suppose now that $c < \frac{3}{4}$. Let t be an integer such that $t > \frac{1}{3-4c}$. Then $G_{2,t}$ has order $4t$, minimal degree $3t-1 \geq c|G_{2,t}|$ and monochromatic circumference $2t = \frac{1}{2}|G_{2,t}|$. Hence $\Phi(c) \leq \frac{1}{2}$. Similarly, for $c < \frac{5}{9}$, the graph $G_{3,t}$ shows that $\Phi(c) \leq \frac{1}{3}$.

It remains to show the bound $\Phi(c) \leq \frac{2}{3}c$ for all $c < \frac{3}{4}$. Note that $\frac{1}{m_c} \leq \frac{2}{3}c$ for all $c \in (0, \frac{9}{25}]$, and so $G^*_{m_c,t}$ gives this bound for sufficiently large t. For $\frac{9}{25} \leq c < \frac{3}{4}$ and n sufficiently large, a graph $H_{n,c}$ will be constructed with order n, minimal degree at least $cn-6$ and monochromatic circumference at most $\frac{2}{3}cn + 16$.

Suppose that $\frac{3}{5} \leq c < \frac{3}{4}$ and let n be a positive integer. Let A_1 be a set of order $\left\lfloor \frac{1}{3}cn \right\rfloor$, A_2 and A_3 be sets of order $\lfloor (1-c)\, n \rfloor$, and A_4 be a set of order $n - |A_1| - |A_2| - |A_3|$. Hence

$$\begin{aligned} \frac{1}{3}cn - 1 \leq & |A_1| \leq \frac{1}{3}cn, \\ (1-c)\, n - 1 \leq & |A_2| = |A_3| \leq (1-c)\, n \end{aligned}$$

and

$$\left(\frac{5}{3}c - 1\right) n \leq |A_4| \leq \left(\frac{5}{3}c - 1\right) n + 3.$$

Define two graphs on vertex set $\bigcup_i A_i$ as follows.

- $R = K_{A_3 \cup A_4} \cup \left(K_{A_1} + \overline{K_{A_2}}\right)$.

- $B = K_{A_1,A_3} \cup \left(K_{A_2} + \overline{K_{A_4}}\right)$.

Let $H_{n,c}$ be the graph with vertex set $\bigcup_i A_i$ and $E(H_{n,c}) = E(R) \cup E(B)$, as illustrated in Figure 3.4.

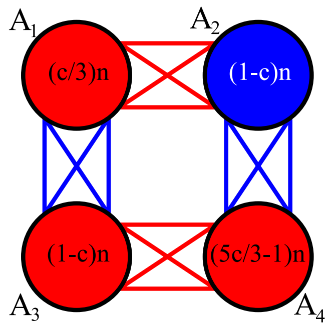


Figure 3.4: The graph $H_{n,c}$, for $\frac{3}{5} \leq c < \frac{3}{4}$. Note the approximate order of each set is given.

Note that $|H_{n,c}| = n$, and, in $H_{n,c}$, a vertex in A_i is adjacent to all vertices but itself and vertices in A_{5-i}. Hence $H_{n,c}$ has minimal degree $n-1-\max_i |A_i|$. Recall that $c < \frac{3}{4}$. Hence for n sufficiently large, $\max_i |A_i| = |A_2| \leq (1-c)\,n$, and so $\delta(H_{n,c}) \geq cn - 1$.

The blue graph B contains two components. One has vertex set $A_2 \cup A_4$, and hence has circumference at most $|A_2| + |A_4| \leq \frac{2}{3}cn + 3$. The other blue component is the complete bipartite graph between A_1 and A_3, and hence has circumference $2\min\{|A_1|, |A_3|\} \leq \frac{2}{3}cn$. Hence the circumference of B is at most $\frac{2}{3}cn + 3$. Similarly the circumference of R is at most $\frac{2}{3}cn + 3$.

Let $\frac{1}{2} \leq c < \frac{3}{5}$, and let n be a positive integer. Let A_1 and A_2 be sets of order $\left\lfloor \frac{2}{3}cn \right\rfloor$, A_3 and A_4 be sets of order $\left\lfloor \left(1 - \frac{5}{3}c\right) n \right\rfloor$, and A_5 be a set of order $n-2|A_1|-2|A_2|$. Hence

$$\begin{aligned}\frac{2}{3}cn-1 \leq & |A_1| = |A_2| \leq \frac{2}{3}cn, \\ \left(1-\frac{5}{3}c\right)n-1 \leq & |A_3| = |A_4| \leq \left(1-\frac{5}{3}c\right)n\end{aligned}$$

and

$$(2c-1)\,n \leq |A_5| \leq (2c-1)\,n+4.$$

Define two graphs on vertex set $\bigcup_i A_i$ as follows.

- $R = K_{A_1} \cup \left(K_{A_4 \cup A_5} + \overline{K_{A_2}}\right) \cup \overline{K_{A_3}}$.
- $B = K_{A_2} \cup \left(K_{A_3} + K_{A_1,A_5}\right) \cup \overline{K_{A_4}}$.

Let $H_{n,c}$ be the graph with vertex set $\bigcup_i A_i$ and $E(H_{n,c}) = E(R) \cup E(B)$, as illustrated in Figure 3.5.

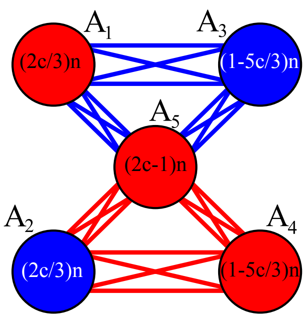


Figure 3.5: The graph $H_{n,c}$, for $\frac{1}{2} \leq c < \frac{3}{5}$. Note that the approximate order of each set is given.

Note that $|H_{n,c}| = n$. In $H_{n,c}$, a vertex in A_5 is adjacent to all vertices but itself, a vertex in $A_1 \cup A_3$ is adjacent to all vertices in $A_1 \cup A_3 \cup A_5$ but itself, and a vertex in $A_2 \cup A_4$ is adjacent to all vertices in $A_2 \cup A_4 \cup A_5$ but itself. Hence $H_{n,c}$ has minimal degree $|A_1| + |A_3| + |A_5| - 1 \geq cn - 3$.

The red graph R contains two non-trivial components. One is the complete graph with vertex set A_1, and hence has circumference $|A_1| \leq \frac{2}{3}cn$. The other red component is $K_{A_4 \cup A_5} + \overline{K_{A_2}}$. Hence in any cycle of this component, at least half of the vertices must lie in $A_4 \cup A_5$. Thus this component has circumference at most $2\left(|A_4| + |A_5|\right) \leq \frac{2}{3}cn + 8$. Hence the circumference of R is at most $\frac{2}{3}cn + 8$. Similarly the circumference of B is at most $\frac{2}{3}cn + 8$.

Let $\frac{1}{3} \leq c < \frac{1}{2}$, and let n be a positive integer. Let $A_{1,2}$, $A_{1,3}$, $A_{2,1}$ and $A_{3,1}$ be sets of order $t_1 := \left\lfloor \left(\frac{1}{2} - \frac{5}{6}c\right) n \right\rfloor$. Let $A_{2,2}$, $A_{2,3}$, $A_{3,2}$ and $A_{3,3}$ be sets of order $t_2 := \left\lfloor \left(\frac{3}{4}c - \frac{1}{4}\right) n \right\rfloor$, and let $A_{1,1}$ be a set of order $t_3 = n - 4t_1 - 4t_2$. Hence

$$\begin{aligned} \left(\frac{1}{2} - \frac{5}{6}c\right) n - 1 \leq & t_1 \leq \left(\frac{1}{2} - \frac{5}{6}c\right) n, \\ \left(\frac{3}{4}c - \frac{1}{4}\right) n - 1 \leq & t_2 \leq \left(\frac{3}{4}c - \frac{1}{4}\right) n \end{aligned}$$

and

$$\frac{1}{3}cn \leq t_3 \leq \frac{1}{3}cn + 8.$$

Define two graphs on vertex set $\bigcup_{i,j} A_{i,j}$ as follows.

- $R = \left(K_{A_{1,1}} + \overline{K_{A_{2,1} \cup A_{3,1}}}\right) \cup K_{A_{1,2} \cup A_{2,2} \cup A_{3,2}} \cup K_{A_{1,3} \cup A_{2,3} \cup A_{3,3}}$.
- $B = K_{A_{1,1}, A_{1,2} \cup A_{1,3}} \cup \left(K_{A_{2,1}} + K_{A_{2,2}, A_{2,3}}\right) \cup \left(K_{A_{3,1}} + K_{A_{3,2}, A_{3,3}}\right)$.

Let $H_{n,c}$ be the graph with vertex set $\bigcup_i A_i$ and $E(H_{n,c}) = E(R) \cup E(B)$, as illustrated in Figure 3.6.

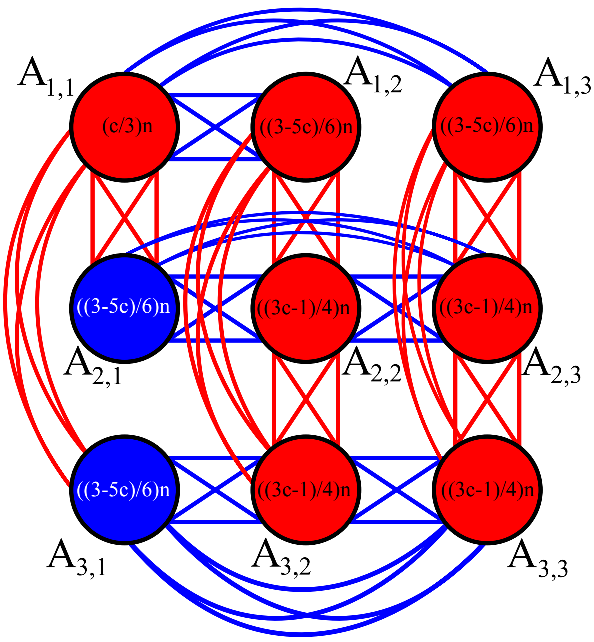


Figure 3.6: The graph $H_{n,c}$, for $\frac{1}{3} \le c < \frac{1}{4}$. Note that the approximate order of each set is given.

Note that $|H_{n,c}| = n$. In $H_{n,c}$, a vertex in $A_{i,j}$ only has neighbours in $\bigcup_{j'} A_{i,j'} \cup \bigcup_{i'} A_{i',j}$. Also, no vertex is adjacent to itself, and there are no edges between $A_{1,2}$ and $A_{1,3}$, nor between $A_{2,1}$ and $A_{3,1}$. Hence the degree of a vertex in $A_{1,1}$ is $t_3 + 4t_1 - 1 \ge (2-3c)\,n - 5$, the degree of a vertex in $A_{1,2} \cup A_{1,3} \cup A_{2,1} \cup A_{3,1}$ is $t_3 + t_1 + 2t_2 - 1 \ge cn - 4$, and the degree of a vertex in $A_{2,2} \cup A_{2,3} \cup A_{3,2} \cup A_{3,3}$ is $2t_1 + 3t_2 - 1 \ge \left(\frac{7}{12}c + \frac{1}{4}\right) n - 6$. Hence, for $\frac{1}{3} \le c < \frac{1}{2}$ and n sufficiently large, $H_{n,c}$ has minimal degree at least $cn - 4$.

The red graph R contains three components. Two are the complete graphs with vertex sets $A_{2,1} \cup A_{2,2} \cup A_{2,3}$ and $A_{3,1} \cup A_{3,2} \cup A_{3,3}$ respectively. These both have circumference $t_1 + 2t_2 \le \frac{2}{3}cn$. The third component of R is $K_{A_{1,1}} + \overline{K_{A_{2,1} \cup A_{3,1}}}$. Hence in any cycle of this component, at least half of the vertices must lie in $A_{1,1}$. Thus this component has circumference at most $2|A_{1,1}| \leq \frac{2}{3}cn + 16$. Hence the circumference of R is at most $\frac{2}{3}cn + 16$. Similarly the circumference of B is at most $\frac{2}{3}cn + 16$.

Hence, for all $\frac{1}{3} \leq c < \frac{3}{4}$, a graph $H_{n,c}$ with minimal degree at least $cn - 6$ and monochromatic circumference at most $\frac{2}{3}cn + 16$ has been defined. Let $\phi = \frac{2}{3}c + \epsilon$ for some $0 < \epsilon < \frac{3}{4} - c$. Then, for n sufficiently large, $H_{n,c+\epsilon}$ has minimal degree at least $(c+\epsilon)\,n - 6 \geq cn$ and monochromatic circumference at most $\frac{2}{3}\,(c+\epsilon)\,n + 16 \leq \phi n$. Hence $\Phi(c) \leq \frac{2}{3}c + \epsilon$, for any $\epsilon > 0$, which implies that $\Phi(c) \leq \frac{2}{3}c$. Thus the proof of the upper bound on $\Phi(c)$ is complete. $\square$

3.4 The monochromatic circumference of an r-edge-coloured graph

In this section, the monochromatic circumference of r-edge-coloured graphs is considered. Note that if G has order n and minimal degree at least cn, then G has at least $\frac{1}{2}cn^2$ edges. Thus, if $E(G) = \bigcup_{\nu=1}^{r} G_\nu$ is an r-edge-colouring of G, there is some $1 \leq \nu \leq r$ such that G_ν has at least $\frac{1}{2r}cn^2 > \frac{1}{2}\,(n-1)\left(\frac{1}{r}cn\right)$ edges. Hence Theorem 1.14 implies that G_ν has circumference at least $\frac{1}{r}cn$. This shows that $\Phi_r(c) \geq \frac{1}{r}c$ for any positive integer r and $c \in (0,1)$.

Using combinatorial designs, graphs can be constructed showing that this bound is close to being tight for some values of c. Orthogonal Latin squares as defined below were first introduced by Euler [21]. For a general overview of combinatorial designs, see Ionin and Shrikhande [29] and for an in-depth treatment of Latin squares, see Laywine and Mullen [37].

Definition 3.14 (Mutually orthogonal Latin squares)**.** *A* Latin square *of order m is an $m \times m$ array with entries $1, \ldots, m$, having the property that each entry occurs exactly once in each row and each column. For a Latin square M, denote by $M(i,j)$ the entry in the ith and jth column. Two Latin squares M_1 and M_2 are* orthogonal *if, for every $k, l \in \{1, \ldots, m\}$, there is a unique i and j such that $M_1(i,j) = k$ and $M_2(i,j) = l$. A set $\{M_1, \ldots, M_r\}$ of Latin squares of order n is a set of* mutually orthogonal Latin squares *if any two distinct squares in the set are orthogonal.*

It is easy to check that there are at most $m-1$ mutually orthogonal Latin squares of order m. Assuming that they exist (see [29]), mutually orthogonal Latin squares can be used to construct r-edge-coloured graphs with short monochromatic circumference.

Lemma 3.15. *Let t be an integer. If there are r mutually orthogonal Latin squares of order m, then there is an $(r+2)$-edge-coloured graph $G^{(t)}$ with order $m^2 t$, minimal degree $((r+2)(m-1)+1)\,t-1$ and monochromatic circumference mt.*

Proof. Suppose that $\{M_1, \ldots, M_r\}$ is a set of mutually orthogonal Latin squares. Let $V = \{(i,j) : 1 \leq i, j \leq m\}$. For $1 \leq \nu \leq r$, let G_ν be the graph with vertex set V whose edges join any two pairs (i,j) and (i',j') such that $M_\nu(i,j) = M_\nu(i',j')$. Also, let G_{r+1} be the graph with vertex set V whose edges join any two pairs (i,j) and (i',j') such that $i = i'$ and G_{r+2} be the graph with vertex set V whose edges join any two pairs (i,j) and (i',j') such that $j = j'$.

Note that the graphs G_ν for $1 \le \nu \le r+2$ are mutually edge-disjoint, as $\{M_1, \ldots, M_r\}$ is a set of mutually orthogonal Latin squares. Let G be the $(r+2)$-edge-coloured graph with vertex set $\{(i,j) : 1 \le i, j \le m\}$ defined by $E(G) = \bigcup_{\nu=1}^{r+2} E(G_\nu)$.

For any integer t, let $H^{(t)}$ be the graph obtained by replacing each vertex (i,j) of G with a set $A_{i,j}$ of order t, and each edge with colour ν with a complete bipartite graph in the same colour. The graph $G^{(t)}$ is obtained from $H^{(t)}$ by joining all pairs of vertices which lie in a common set $A_{i,j}$ with an edge (of any colour). It is clear that $G^{(t)}$ has order $m^2 t$, minimal degree $((r+2)(m-1)+1)\,t-1$ and monochromatic circumference mt as claimed. $\square$

The bound implied by Theorem 1.14 implies that $\Phi_{r+2}\left(\frac{r+2}{m}\right) \ge \frac{1}{m}$. However, Lemma 3.15 implies that, when there are r mutually orthogonal Latin squares of order m, $\Phi_{r+2}(c) \le \frac{1}{m}$ for any $c < \frac{((r+2)(m-1)+1)}{m^2}$. For fixed r, as m becomes large, this upper bound on Φ_{r+2} is valid for values of c approaching $\frac{r+2}{m}$ and so is close to best possible.

Chapter 4

The existence of 2-factors

4.1 2-factors in hamiltonian graphs

Recall that a 2-factor of a graph G is a 2-regular spanning subgraph of G. In this chapter, a minimal degree condition for a hamiltonian graph to contain a 2-factor with k components will be considered.

Definition 4.1. *For positive integers k and n, let $c(n,k)$ be the smallest integer such that every hamiltonian graph of order n with minimal degree at least $c(n,k)$ has a 2-factor with k components.*

Theorem 1.20 shows that $c\,(n,k) \leq \frac{5}{12}n + 2$. The following result is an improvement of Theorem 1.20. Note, for the purpose of comparison, that $\frac{5-\sqrt{7}}{6} \approx 0.392$ while $\frac{5}{12} \approx 0.417$.

Theorem 4.2. *For $n \geq 6$, $c(n,2) \leq \frac{5-\sqrt{7}}{6}n + 2$.*

Theorem 4.2 will be proved in Section 4.2. This section concludes with a brief discussion of lower bounds on $c\,(n,k)$. Let $t \geq 3$. Faudree, Gould, Jacobsen, Lesniak and Saito [22] constructed the following graph H_t, which shows that $c(5t, 2) > 4$. For $1 \le i \le t$, let $V_1, \ldots, V_t$ be disjoint sets of order five, and let u_i and v_i be particular distinct vertices of V_i. The graph H_t has vertex set $\bigcup_{i=1}^{t} V_i$. The edge set of H_m consists of:

- for $1 \le i \le m$, all edges within V_i except $u_i v_i$;
- the edges $u_i v_{i+1}$ for $1 \le i \le t-1$ and the edge $u_t v_1$.

The latter set of edges will be referred to as *joining edges*. See Figure 4.1 for an illustration of H_4.

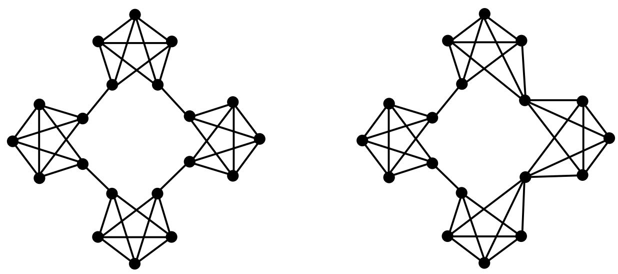


Figure 4.1: The graphs H_4 and $H_{4,2}$

The graph H_t is 4-regular and hamiltonian. Suppose that H_t has a 2-factor with two components C_1 and C_2. As $t \ge 3$, one of the components, C_1 say, must intersect at least two of the sets V_i and hence contains a joining edge. However, C_1 is a cycle, and a cycle of H_t containing one joining edge must contain them all. Hence C_2 contains no joining edges and is contained within some V_j. As both $u_j v_{j+1}$ and $u_{j-1} v_j$ are contained in C_1, C_2 must contain exactly the three remaining vertices of V_j. However, there is no edge

u_jv_j, and so C_1 can not be a cycle. Hence H_m does not have a 2-factor with two components. This shows that $c(5t, 2) > 4$.

Note that H_t has a 2-factor with t components. However, a similar argument to the one above shows that H_t does not have a 2-factor with k components, for any $k \notin \{1, t\}$. For $t \geq 3$ and $1 \leq i \leq t$, the following graphs also have this property, and their orders are not restricted to multiples of five. Let $H'_{t,i}$ be the graph obtained by contracting the joining edges u_jv_{j+1} of H_m for all $1 \leq j \leq i$. See Figure 4.1 for an illustration of $H_{4,2}$. These graphs imply a lower bound on $c(n, k)$.

Corollary 4.3. *$c(n, k) > 4$ for $n \geq 4(k + 1)$.*

Proof. For all $n \geq 4(k + 1)$, there are values $t \geq k + 1$ and $1 \leq i \leq t$ such that $n = 5t - i$. Then $H'_{t,i}$ is a hamiltonian graph with minimal degree four, but no 2-factor with k components. □

4.2 2-factors with two components in hamiltonian graphs.

In this Section, a proof will be given of Theorem 4.2. The method of proof is based on that used by Faudree, Gould, Jacobsen, Lesniak and Saito [22] in their proof of Theorem 1.20. The following notation will be used in the proof of Theorem 4.2. Suppose that C is a cycle with orientation $\overrightarrow{C}$. For $v \in V(C)$, let v^+ be the outneighbour of v in $\overrightarrow{C}$, and v^- be the in-neighbour of v in $\overrightarrow{C}$. For integers i, iteratively define v^{+i} to be $\left(v^{+(i-1)}\right)^+$ and v^{-i} to be $\left(v^{-(i-1)}\right)^-$. Throughout, indices will be assumed to be considered modulo

$n = |G|$.

Proof of Theorem 4.2. Let G be a hamiltonian graph with n vertices and minimal degree at least $d = \frac{5-\sqrt{7}}{6}n + 2$. To obtain a contradiction, assume that G does not contain a 2-factor with two components. Let C be a Hamilton cycle of G and label the vertices of G so that $C = x_1 \ldots x_n$. Let $\overrightarrow{C}$ be the orientation of C with $\overrightarrow{C} = x_1 x_2 \ldots x_n$ and $\overleftarrow{C}$ be the opposite orientation of C.

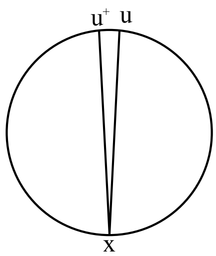


Figure 4.2: An insertion

For $x \in V(G)$ and $e = uu^+ \in E(C)$ call (x, e) an *insertion* if both u and u^+ are neighbours of x, as in Figure 4.2. Let I denote the set of all insertions in G. For each edge $e \in E(C)$, define

$$j(e) = |\{(x, e) \in I : x \in V(G)\}|.$$

Similarly, for each $x \in V(G)$, define

$$i(v) = |\{(v, f) \in I : f \in E(C)\}|.$$

Counting the number of insertions in two ways implies that

$$|I| = \sum_{e \in E(C)} j(e) = \sum_{v \in V(G)} i(v).$$

Say that $x \in V(G)$ *has an insertion* when there is an insertion (x, e) for some $e \in E(C)$. By considering the value of $j(xx^+)$ in the two cases when x does or does not have an insertion, a lower bound on $|I|$ will be obtained.

If x does not have an insertion, let y be a vertex such that $y^+ \in \Gamma(x)$, other than one of $\{x^{+2}, x, x^{-2}\}$. If y is adjacent to x in G, there is an insertion (x, yy^+), contradicting the assumption that x does not have an insertion. Hence y is not adjacent to x and so in particular, y is neither x^+ or x^-.

Suppose that y is adjacent to x^+. Then both $x^+\overrightarrow{C}yx^+$ and $xy^+\overrightarrow{C}x$ are cycles in G as $y \notin \{x^{-2}, x^-, x, x^+, x^{-2}\}$. These two cycles form a 2-factor with two components in G, as illustrated in Figure 4.3. Hence y is adjacent to neither x or x^+. This implies that, when x does not have an insertion,

$$\left|\left\{y : y \notin \Gamma(x), y \notin \Gamma\left(x^+\right)\right\}\right| \geq |\Gamma(x)| - 3 \geq d - 5.$$

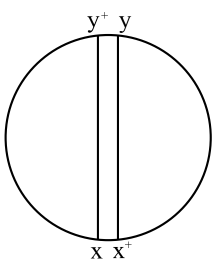


Figure 4.3

Suppose now that x does have insertions. Let y be a vertex such that $y^+x^- \in E(G)$, other than any of $\{x^{+2}, x^+, x, x^-, x^{-3}\}$. Note that $y \neq x^{-2}$, as

$y^+x^- \in E(G)$. If $y \in \Gamma(x)$, then $x\overrightarrow{C}yx$ and $x^-y^+\overrightarrow{C}x^-$ form a 2-factor with two components, as in Figure 4.3. If, however, $y \notin \Gamma(x)$ and $y \in \Gamma(x^+)$, there are two disjoint cycles $x^+\overrightarrow{C}yx^+$ and $x^-y^+\overrightarrow{C}x^-$ which cover all vertices of G except x.

By assumption x has an insertion, which is not (x, yy^+) as $y \notin \Gamma(x)$. Without loss of generality, assume that x has an insertion (x, vv^+), where vv^+ is an edge of $y^+\overrightarrow{C}x^-$. Then $x^+\overrightarrow{C}yx^+$ and $x^-y^+\overrightarrow{C}vxv^+\overrightarrow{C}x^-$ form a 2-factor with two components, as illustrated in Figure 4.4. Hence y is adjacent to neither x or x^+. Thus, if x has an insertion,

$$\left|\left\{y : y \notin \Gamma(x), y \notin \Gamma\left(x^+\right)\right\}\right| \geq |\Gamma(x)| - 5 \geq d - 5.$$

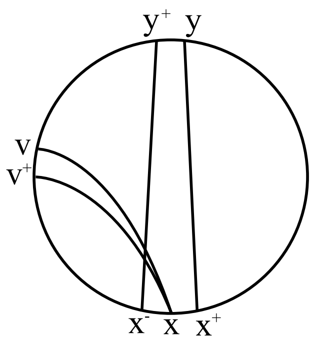


Figure 4.4

Hence, whether x has an insertion or not,

$$\left|\Gamma(x) \cup \Gamma\left(x^+\right)\right| \leq n - d + 5.$$

This implies that, for all $x \in V(G)$,

$$j\left(xx^{+}\right) = \left|\Gamma\left(x\right) \cap \Gamma\left(x^{+}\right)\right| \geq d + d - (n - d + 5) = 3d - n - 5,$$

and hence

$$|I| \geq n\left(3d - n - 5\right) > n. \tag{4.1}$$

The remainder of the proof follows exactly the method used by Faudree, Gould, Jacobsen, Lesniak and Saito [22] in their proof of Theorem 1.20. However, in that proof, $|I|$ had been bounded below by $\frac{n^2}{8}$, and it is the improvement given by (4.1) that allows a better bound to be obtained in Theorem 4.2.

Let x_i be the vertex with the most insertions. The bound in (4.1) implies that $i\left(x_i\right) > 1$. Let

$$\begin{aligned} X &= \Gamma\left(x_i\right) - \left\{x_{i-1}, x_{i+1}\right\}, \\ Y &= \left\{y : y^{-}x_{i+1} \in E(G)\right\} - \left\{x_{i+3}, x_i\right\}, \\ Z &= \left\{y : y^{+}x_{i-1} \in E(G)\right\} - \left\{x_{i-3}, x_i\right\}. \end{aligned}$$

Note that it is immediate that $x_{i+2} \notin Y$ and $x_{i-2} \notin Z$.

If $X \cap Y$ is non-empty, let $y \in X \cap Y$. Then $y \notin \{x_{i-1}, x_i, x_{i+1}, x_{i+2}, x_{i+3}\}$, and both yx_i and $y^{-}x_{i+1}$ are edges. Hence there is a 2-factor with two components of the form illustrated in Figure 4.3. Thus $X \cap Y$ is empty and similarly $X \cap Z$ is empty. Hence

$$|Y \cap Z| = |X| + |Y| + |Z| - |X \cup Y \cup Z| \geq 3\,(d-2) - n > 0, \qquad (4.2)$$

and so, in particular, $Y \cap Z$ is non-empty.

Let $x_j \in Y \cap Z - \{x_{i-1}, x_{i+1}\}$. Suppose that x_j has an insertion (x_j, uu^+). As x_i has more than one insertion, there is an insertion (x_i, vv^+) with $v \neq u$. Note that $x_j \notin X$ and so $x_i x_j$ is not an edge. Hence both edges uu^+ and vv^+ lie either within $x_{i+1}\overrightarrow{C}x_{j-1}$ or within $x_{j+1}\overrightarrow{C}x_{i-1}$. Thus there is a 2-factor with two components with cycles of the form $x_{i+1}\overrightarrow{C}ux_ju^+\overrightarrow{C}x_{j-1}x_{i+1}$ and $x_{j+1}\overrightarrow{C}vx_iv^+\overrightarrow{C}x_{i-1}x_{j+1}$, as shown in Figure 4.5. Note that the exact form of the cycles depends on the locations of u and v. Hence no vertex $x_j \in Y \cap Z - \{x_{i-1}, x_{i+1}\}$ has an insertion.

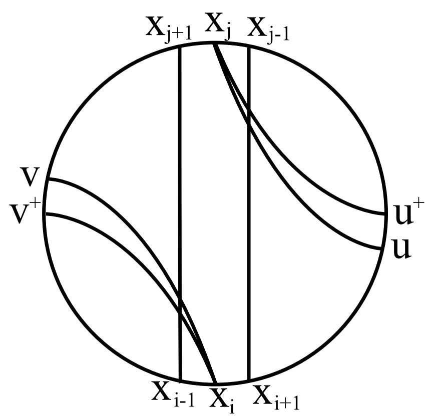


Figure 4.5

Hence (4.2) implies that there are at most $n - (3\,(d-2) - n - 2) = 2n - 3d + 8$ vertices which have insertions. Recalling that x_i is a vertex with the most insertions, (4.1) implies that x_i has at least $\frac{n(3d-n-5)}{2n-3d+8}$ insertions.

Let $x_j \in Y \cap Z$, where Y and Z are as defined earlier. Note that $x_{j-1} \in \Gamma(x_{i+1})$ and $x_{j+1} \in \Gamma(x_{i-1})$, and also recall that x_j has no insertions, and is not adjacent to x_i. Let T be the neighbourhood of x_j. Let S be the vertices adjacent to x_i and not x_j. Finally, let R be vertices which are adjacent to neither x_i nor x_j. Note that $R \cup S \cup T$ is a partition of $V(G)$, and that both x_i and x_j lie in R.

Suppose that $i+1 \leq l \leq j-2$, and that $x_i x_{l+1}$ and $x_j x_l$ are both edges of G. Then, as illustrated in Figure 4.6, $x_i \overrightarrow{C} x_\ell x_j \overleftarrow{C} x_{\ell+1} x_i$ and $x_{j+1} \overrightarrow{C} x_{i-1} x_{j+1}$ form a 2-factor with two components. Similarly, if $j+2 \leq \ell \leq i-1$, and both $x_i x_{\ell-1}$ and $x_j x_\ell$ are edges of G, then G contains a 2-factor with two components. As G is assumed not to contain a 2-factor with two components, neither of these two events occurs.

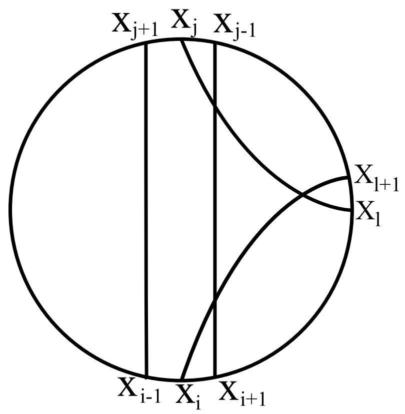


Figure 4.6

Suppose that $(x_i, x_l x_{l+1})$ is an insertion for some $i+1 \leq l \leq j-2$. Then x_l is a neighbour of x_i, and is not a neighbour of x_j by the above. Hence

$x_l \in S$. Similarly, if $(x_i, x_{l-1}x_l)$ is an insertion for some $j+2 \leq l \leq i-1$, then $x_l \in S$. As x_i is not adjacent to itself or x_j, all of its insertions have the above form and hence

$$|S| \geq i\,(x_i)\,. \tag{4.3}$$

Let x_l be a neighbour of x_j for some $i+1 \leq l \leq j-2$. Then x_{l+1} is not a neighbour of x_i. However, x_j has no insertions and so x_{l+1} is not a neighbour of x_j. Hence $x_{l+1} \in R$. Similarly, if x_l is a neighbour of x_j with $j+2 \leq \ell \leq i-1$, then $x_{l-1} \in R$. In particular, for each neighbour of x_j except x_{j+1} and x_{j-1}, a unique element of $R \setminus \{x_i, x_j\}$ can be obtained. However both x_i and x_j lie in R and so

$$|R| \geq |T|\,.$$

Hence, (4.3) implies that

$$n = |R| + |S| + |T| \geq 2\,|T| + i\,(x_i)\,.$$

However T is the neighbourhood of x_j and thus contains at least d vertices. This implies that

$$n \geq 2d + \frac{n\,(3d-n-5)}{2n-3d+8}.$$

Hence

$$(n-2d)\,(2n-3d+8) \geq n\,(3d-n-5)$$

and so

$$3n^2 + 13n \geq 10dn + 16d - 6d^2. \tag{4.4}$$

However $d = \frac{5-\sqrt{7}}{6}n + 2$ and so

$$\begin{aligned} d\,(10n + 16 - 6d) &= \left(\frac{5-\sqrt{7}}{6}n + 2\right)\left(\left(5+\sqrt{7}\right)n + 4\right) \\ &> 3n^2 + \frac{40 + 4\sqrt{7}}{3}n. \end{aligned}$$

This is a contradiction to (4.4) and completes the proof. □

Chapter 5

The distribution of X_H modulo k, for a connected graph H

5.1 Partitions of a random graph

Recall, for a graph H, that X_H is the number of copies of H in $\mathcal{G}(n,p)$. This chapter will study, for a connected graph H, the distribution of X_H, when taken modulo an integer function $k(n)$. The following notation concerning modular arithmetic is used throughout this chapter.

Definition 5.1. *For a set S, an S-tuple is a sequence $(x_i)_{i \in S}$.*

- *For two sets S and T, the set S^T is the set of all T-tuples formed by elements of S.*
- *As standard, $[s]$ shall denote the set $\{1, \ldots, s\}$ and here it will also denote the set of equivalence classes modulo k. Note that the set $\{1, \ldots, k\}$* is *a set of representatives of the equivalence classes modulo k.*

- *For two S-tuples of integers $\mathbf{x}$ and $\mathbf{y}$, let $\mathbf{x} \underset{k}{\equiv} \mathbf{y}$ when $x_i \underset{k}{\equiv} y_i$ for all $i \in S$. For an S-tuple of integers $\mathbf{x}$ and a set A of S-tuples of integers let $\mathbf{x} \underset{k}{\in} A$ when $\mathbf{x} \underset{k}{\equiv} \mathbf{y}$ for some $\mathbf{y} \in A$.*

In Chapter 1, a notion of approximate uniformity modulo k was defined for integer-valued random variables. Here this definition is extended to vectors of integer-valued random variables.

Definition 5.2. *Let $\epsilon > 0$ and $k \in \mathbb{N}$. An integer-valued random variable $\mathbf{Y} = (Y_i)_{i \in S}$ is* ϵ-uniform modulo k *if*

$$\max_{\mathbf{a} \in [k]^S} \left| \mathbb{P}\left(\mathbf{Y} \underset{k}{\equiv} \mathbf{a}\right) - \frac{1}{k^{|S|}} \right| < \frac{\epsilon}{k^{|S|}}.$$

The main result of this chapter is that $(X_H(G_n))_{H \in \mathcal{C}}$ is $e^{-\Omega\left(n/k^2\right)}$-uniform modulo k, provided that $k = o\left(\sqrt{n/\log n}\right)$.

Theorem 5.3. *Let $\mathcal{C}$ be any finite set of connected graphs, $p \in (0,1)$ be a fixed real number and let $G_n \in \mathcal{G}(n,p)$. Then, for every $k \in \mathbb{N}$ with $k = o\left(\sqrt{n/\log n}\right)$, the distribution of $(X_H(G_n))_{H \in \mathcal{C}}$ is $e^{-\Omega\left(n/k^2\right)}$-uniform modulo k.*

As noted in Chapter 1, Kolaitis and Kopparty [32] independently showed the following result, for constant k.

Theorem 5.4. *Let k be a fixed integer and $p \in (0,1)$. Let $\mathcal{C}$ be a finite set of connected non-trivial graphs. Then the distribution of $(X_H)_{H \in \mathcal{C}}$ is $\left(2^{-\Omega(n)}\right)$-uniform modulo k.*

Theorem 5.3 is an immediate corollary of the following theorem.

Definition 5.5. *Let $\mathcal{C}_t$ be the set of all non-trivial connected graphs with t or fewer vertices (recall that the trivial graph is K_1).*

Theorem 5.6. *Let T be a fixed integer, $p \in (0,1)$ be a fixed real number and let $G_n \in \mathcal{G}(n,p)$. Then, for every $k \in \mathbb{N}$ with $k = o\left(\sqrt{n/\log n}\right)$, the distribution of $(X_H(G_n))_{H \in \mathcal{C}_T}$ is $e^{-\Omega\left(n/k^2\right)}$-uniform modulo k.*

Suppose that H is a graph on T vertices. The value of X_H modulo k will be calculated in the following way. Fix a vertex v of the vertex set V of $\mathcal{G}(n,p)$. This allows decomposition of the random graph into a random graph $\mathcal{G}(n-1,p)$ on vertex set $V - \{v\}$ and a random partition of $V - \{v\}$ into neighbours and non-neighbours of v. Conditional on the graph $\mathcal{G}(n-1,p)$, the calculation becomes that of counting the number of copies of H which contain the vertex v. This can be done by counting the number of copies in $\mathcal{G}(n-1,p)$ of the set of graphs $H \setminus \{x\}$ where $x \in V(G)$. Note that all such graphs have order $T-1$. However, a copy of $H \setminus \{x\}$ in $\mathcal{G}(n-1,p)$ will not always give a copy of H in $\mathcal{G}(n,p)$. For example, a copy of $H \setminus \{x\}$ amongst non-neighbours of v is not a copy of H in $\mathcal{G}(n,p)$. In order to count copies of H containing v it is therefore necessary to colour the graphs $H \setminus \{x\}$ in a certain way and count the number of occurrences of the coloured graphs. Thus partitions of the vertex set of the random graph and colourings of the small graph are now introduced.

Throughout the remainder of this section S and T will represent a general indexing set, although in practice they will always be Cartesian products of finite sets of integers. The random graph $\mathcal{G}(n,p)$ will always be assumed to have vertex set $[n]$. The following define partitions of an integer and of a set.

Definition 5.7. *An* (n, S)-partition *is an* S*-tuple of non-negative integers, summing to* n*. For an* (n, S)*-partition* $\mathbf{x}$*, let* $\min(\mathbf{x}) := \min_{i \in S} x_i$*.*

Definition 5.8. *For a set* V*, an* S-indexed partition *of* V *is a family* $\mathcal{P}$ *of subsets* $(V_i)_{i \in S}$ *of* V *such that* $V = \bigcup_{i \in S} V_i$ *and* $V_i \cap V_j = \emptyset$ *for all* $i \neq j$*. The* profile *of* $\mathcal{P}$ *is the* (n, S)*-partition* $\mathbf{x}$ *defined by* $x_i = |P_i|$ *for all* $i \in S$*. Let* $\min(\mathcal{P}) = \min(\mathbf{x})$*, where* $\mathbf{x}$ *is the profile of* $\mathcal{P}$*.*

The following definition introduces coloured graphs, and extends the definition of $\mathcal{C}_t$.

Definition 5.9. *For a set* S *and graph* H*, an* S-colouring of H *is a function* c *from* $V(H)$ *to* S*. Throughout* (H, c)*, will denote a graph* H *with colouring* c*.*

- *Let* $\mathrm{col}(H, S)$ *be the set of* S*-colourings of* H*.*
- *Two coloured graphs* (H, c) *and* (H', c') *are* isomorphic *if there is an isomorphism* θ *from* H *to* H' *such that* $c(v) = c'(\theta(v))$ *for all* $v \in V(H)$*.*
- *An* automorphism *of* (H, c) *is an isomorphism from* (H, c) *to itself.*
- *For* $c \in \mathrm{col}(H, S)$ *and* $c' \in \mathrm{col}(H, T)$*, let* $c \times c' \in \mathrm{col}(H, S \times T)$ *be the colouring where each vertex* $v \in V(H)$ *is coloured by* $(c(v), c'(v))$*.*
- *For* $t \in \mathbb{N}$*, the set* $\mathcal{C}_{t,S}$ *consists of one representative of each isomorphism class of* S*-coloured non-trivial connected graphs with at most* t *vertices. Thus* $\mathcal{C}_{t,S}$ *is the set of* (H, c) *with* $H \in \mathcal{C}_t$ *and* $c \in \mathrm{col}(H, S)$*.*
- *Let* $\mathcal{C}_{t,s} = \mathcal{C}_{t,[s]}$ *and* $C(t, s) = |\mathcal{C}_{t,s}|$*. Note that* $|\mathcal{C}_{t,S}| = C(t, s)$ *for any* S *with* $|S| = s$*.*

- *For $1 \leq i \leq j \leq s$, let $(K_2, (i, j))$ be the coloured graph consisting of a single edge with one vertex coloured i and the other coloured j. Hence*

$$\mathcal{C}_{2,s} = \{(K_2, (i, j)) : 1 \leq i \leq j \leq s\}.$$

The proof of Theorem 5.6 requires counting of the number of times a small graph with a colouring is contained in a large graph with a partition of its vertex set. Such a containment is now formally defined.

Definition 5.10. *Let G be a graph on vertex set V with a partition $\mathcal{P}$ indexed by S, and let $\mathcal{C}$ be a finite set of S-coloured graphs. For $(H, c) \in \mathcal{C}$, a* copy of (H, c) in $(G, \mathcal{P})$ *is a subgraph of G, isomorphic to H, such that for all $v \in H$ the vertex corresponding to v is in $P_{c(v)}$. Let $X^{\mathcal{P}}_{(H,c)}(G)$ be the number of copies of (H, c) in G and let $X^{\mathcal{P}}_{\mathcal{C}}(G)$ be the vector $\left(X^{\mathcal{P}}_{(H,c)}(G)\right)_{(H,c)\in\mathcal{C}}$.*

The following simple lemma means that when considering a partition of the vertex set of $\mathcal{G}(n, p)$, it is the profile of the partition which is important when counting coloured subgraphs.

Lemma 5.11. *Let $G_n \in \mathcal{G}(n, p)$, let $\mathcal{C}$ be a finite set of S-coloured graphs and let $\mathcal{P}$ be an S-indexed partition of $[n]$.*

(i) Let $|S'| = |S|$ and let ϕ be a bijection from S to S'. Define $\mathcal{C}'$ from $\mathcal{C}$ by applying ϕ to the colouring of each vertex of each graph. Similarly, define $\mathcal{P}'$ by applying ϕ to the indices of the partition $\mathcal{P}$. Then $X^{\mathcal{P}}_{\mathcal{C}}(G_n)$ is ϵ-uniform modulo k if, and only if, $X^{\mathcal{P}'}_{\mathcal{C}'}(G_n)$ is ϵ-uniform modulo k.

(ii) Let $\mathcal{D}$ be a subset of $\mathcal{C}$. If $X^{\mathcal{P}}_{\mathcal{C}}(G_n)$ is ϵ-uniform modulo k, then $X^{\mathcal{P}}_{\mathcal{D}}(G_n)$ is also ϵ-uniform modulo k. □

The method used to prove Theorem 5.6 requires use of random partitions, which are introduced below.

Definition 5.12. *Suppose that $\mathcal{P}$ is an S-partition of $[n]$ with profile $\mathbf{x}$.*

- *An $(n, S \times T)$-partition $\mathbf{y}$* refines $\mathbf{x}$ *if $\sum_{j\in T} y_{i,j} = x_i$ for all $i \in S$.*
- *A T*-refinement *of $\mathcal{P}$ is an $S \times T$-partition $\mathcal{Q}$ of $[n]$ with $\bigcup_{j\in T} Q_{i,j} = P_i$ for all $i \in S$.*
- *For an $(n, S \times T)$-partition $\mathbf{y}$ refining $\mathbf{x}$, let $\mathcal{R}(\mathcal{P}, \mathbf{y})$ be the set of all T-refinements of $\mathcal{P}$ with profile $\mathbf{y}$.*

As noted earlier, when proving Theorem 5.6, the random graph $\mathcal{G}(n, p)$ is decomposed into a random graph $\mathcal{G}(n-1, p)$ on vertex set $V - \{v\}$ and a random partition of $V - \{v\}$, where v is some vertex of the random graph. The proof proceeds by conditioning on the graph $\mathcal{G}(n-1, p)$, and counting the number of subgraphs, depending on the random partition. Hence some notion of a fixed graph being 'nice' under a random partition is required. This is defined below.

Definition 5.13. *Suppose that $\mathcal{P}$ is an S-partition of $[n]$ with profile $\mathbf{x}$ and that $\mathbf{y}$ is an $(n, S \times T)$-partition refining $\mathbf{x}$. Let G be a graph with vertex set $[n]$ and let $\mathcal{C}$ be a set of $S \times T$-coloured graphs. For $\epsilon > 0$ and $k \in \mathbb{N}$, G is* ϵ-refinable modulo k for $\mathcal{C}$, $\mathcal{P}$ and $\mathbf{y}$ *if, when $\mathcal{Q}$ is chosen uniformly at random from $\mathcal{R}(\mathcal{P}, \mathbf{y})$, $X_{\mathcal{C}}^{\mathcal{Q}}(G)$ is ϵ-uniform modulo k.*

Thus, if G is ϵ-refinable modulo k for $\mathcal{C}$, then the values of $X_{(H,c\times c')}^{\mathcal{Q}}(G)$ for $(H, c \times c') \in \mathcal{C}$ are approximately independent modulo k. However, for any $c \in \mathrm{col}(H,S)$, the value of $X^{\mathcal{P}}_{(H,c)}(G)$ is determined before the random partition. Also, for any T-refinement $\mathcal{Q}$ of $\mathcal{P}$,

$$X^{\mathcal{P}}_{(H,c)}(G) = \sum_{c' \in \mathrm{col}(H,T)} X^{\mathcal{Q}}_{(H,c\times c')}(G).$$

Hence, given the value of $X^{\mathcal{Q}}_{(H,c\times c')}(G)$ modulo k for all but one $c' \in \mathrm{col}(H,T)$, the remaining value is determined.

This implies that, for small ϵ, no graph is ϵ-refinable for $\mathcal{C}_{t,S\times T}$ modulo k. Thus the following set of graphs is defined, excluding one of the possible colourings for each small graph.

Definition 5.14. *Suppose that T has a total ordering, and let 1_T be the minimal element of T. Let*

$$\mathcal{D}_{t,S,T} = \mathcal{C}_{t,S\times T} \setminus \mathcal{C}_{t,S\times\{1_T\}}.$$

Hence $\mathcal{D}_{t,S,T}$ is the set of $(H, c_1 \times c_2)$ with $(H,c_1) \in \mathcal{C}_{t,S}$ and c_2 a T-colouring of H, other than that for which $c_2(v) = 1_T$ for all $v \in V(H)$. Let $\mathcal{D}_{t,s,d} = \mathcal{D}_{t,[s],[d]}$ and let $D(t,s,d) = |\mathcal{D}_{t,s,d}|$.

The formal proof of Theorem 5.6 will follow from two lemmas, proved in Sections 5.3 and 5.4. These essentially prove the following statements:

- $X_{\mathcal{C}_{t,sd^2}}$ is uniform $\implies$ almost all graphs are refinable for $\mathcal{D}_{t,s,d}$;
- almost all graphs are refinable for $\mathcal{D}_{t,s,2^s}$ $\implies$ $X_{\mathcal{C}_{t+1,s}}$ is uniform.

The formal statements corresponding to these implications are Lemma 5.23 and Lemma 5.28. Section 5.2 will introduce some probabilistic results needed in the proof, and will prove the base case of Theorem 5.6. Sections 5.3 and 5.4 will contain the proofs of the two statements above. Finally, Section 5.5 will contain the proof of Theorem 5.6.

5.2 Useful probabilistic results

This section contains some probabilistic facts, which will be used in the proof of Theorem 5.6. The first lemma will be used to show that a given event has probability close to some constant. It is often easier to prove this conditional on some other event, and the following elementary fact can be used to do this.

Lemma 5.15. *Let C, D, ϵ, δ and η lie in $[0,1]$. Let $I = J \cup J'$ be some indexing set. Let E and F be events in the same probability space, such that F is the disjoint union of events F_i for $i \in I$. Suppose that $|\mathbb{P}(F) - D| \leq \delta$. Suppose further that $|\mathbb{P}(E|F_i) - C| < \epsilon$ for all $i \in J$ and $\mathbb{P}\left(\bigcup_{i\in J'} F_i\right) < \eta$. Then $|\mathbb{P}(E \cap F) - CD| < \eta + \epsilon(D + \delta) + C\delta$.*

Proof. Note that

$$\mathbb{P}(E \cap F) = \sum_{i\in I} \mathbb{P}(E \mid F_i)\mathbb{P}(F_i).$$

Hence, using the triangle inequality,

$$|\mathbb{P}(E \cap F) - CD| \leq \sum_{i\in I} \mathbb{P}(F_i)\, |\mathbb{P}(E \mid F_i) - C| + C\left|\sum_{i\in I} \mathbb{P}(F_i) - D\right|$$

However, the first term on the right hand side can be divided according to whether i is in J or J'. Thus the assumed bounds imply that

$$\begin{aligned}|\mathbb{P}(E \cap F) - CD| &< \sum_{i \in J'} \mathbb{P}(F_i) + \sum_{i \in J} \mathbb{P}(F_i)\epsilon + C\delta \\ &< \eta + \epsilon(D + \delta) + C\delta,\end{aligned}$$

as required. □

Most of the results of this section show that some random variable is likely to be close to its mean, and the following result is a general bound of this form.

Lemma 5.16 (Chebyshev's inequality)**.** *Let X be a random variable with mean μ and finite variance σ^2. Then for any real number $\alpha > 0$*

$$\mathbb{P}(|X - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}.$$

For some distributions, the previous bound can be improved, as in the following two lemmas.

Definition 5.17. *The random variable X has the* hypergeometric distribution with parameters N, M and n *if, for all positive integers a,*

$$\mathbb{P}(X = a) = \frac{\binom{M}{a}\binom{N-M}{n-a}}{\binom{N}{n}}.$$

Note that if choosing n items from a set of N, of which M are 'successes', the number of 'successes' chosen has the hypergeometric distribution with parameters N, M and n. The following Lemma is due to Chvátal [13].

Lemma 5.18. *Suppose that X has the hypergeometric distribution with parameters N, M and n. Then, for any constant $t > 0$,*

$$\mathbb{P}\left(X \leq \frac{nM}{N} - tn\right) \leq e^{-2t^2 n}.$$

For a proof of the following Chernoff-type bound, see, for example, the book of Alon and Spencer [1, Appendix A].

Lemma 5.19. *Let $X = \sum_{i=1}^{n} X_i$, where independently X_i takes value one with probability p_i and zero otherwise. Set $p = \frac{1}{n}\sum_{i=1}^{n} p_i$. For any real $a \geq 0$*

$$\mathbb{P}(X - pn < -a) < e^{-\frac{a^2}{2pn}}$$
$$\mathbb{P}(X - pn \geq a) < e^{-\frac{a^2}{2pn + \frac{2}{3}a}}$$

and so immediately

$$\mathbb{P}(|X - pn| \geq a) < 2e^{-\frac{a^2}{2pn + \frac{2}{3}a}}.$$

When $p_i = p$ for all $1 \leq i \leq n$, note that X has *binomial distribution with parameters n and p*. This distribution is of particular importance, and is denoted by $\mathrm{Bin}(n, p)$. The next two results are about $\mathrm{Bin}(n, p)$. The first is from a book of Bollobás [6, p.6].

Lemma 5.20. *For $p \in (0, 1)$, as $n \to \infty$*

$$\max_{0 \leq r \leq n} \mathbb{P}(\mathrm{Bin}(n, p) = r) \sim \frac{1}{\sqrt{2\pi np(1-p)}},$$

with the maximum occurring at any integer r satisfying $p(n+1)-1 \leq r \leq p(n+1)$.

Note that when $p(n+1)$ is an integer,

$$\mathbb{P}(\mathrm{Bin}(n,p) = p(n+1)-1) = \mathbb{P}(\mathrm{Bin}(n,p) = p(n+1)).$$

In the random graph $\mathcal{G}(n,p)$, the number of edges has the binomial distribution with parameters $\binom{n}{2}$ and p. The simplest case of Theorem 5.6 thus follows from the following result, showing that the binomial distribution is very close to uniformly distributed modulo k, provided that k is not too large. When $p = \frac{1}{2}$, the following result is immediate from a result of Graves and Hamilton [26] concerning sums of binomial coefficients.

Theorem 5.21. *Let n and k be integers and $p \in (0,1)$. For fixed $a \in [k]$, let $\Pi_{n,a} = \mathbb{P}(\mathrm{Bin}(n,p) \underset{k}{\equiv} a)$. Then*

$$\Pi_{n,a} = \frac{1}{k}\sum_{j=0}^{k-1}(1-p+p\omega^j)^n\omega^{-aj}$$

where $\omega = e^{2\pi i/k}$.

In particular, for an integer n which is sufficiently large, $\mathrm{Bin}(n,p)$ *is* $\exp\left\{-\frac{4p(1-p)n}{3k^2}\right\}$*-uniform modulo k, provided that $k = o\left(\sqrt{\frac{n}{\log n}}\right)$.*

Proof. Let $\omega = e^{2\pi i/k}$. That

$$\sum_{j=0}^{k-1}\omega^{j(b-a)} = k\mathbb{1}(b \underset{k}{\equiv} a)$$

is well-known. Hence

$$\begin{aligned}
\Pi_{n,a} &= \sum_{b=0}^{n} \mathbb{1}(b \underset{k}{\equiv} a)(1-p)^{n-b} p^b \binom{n}{b} \\
&= \frac{1}{k} \sum_{b=0}^{n} (1-p)^{n-b} p^b \binom{n}{b} \sum_{j=0}^{k-1} \omega^{j(b-a)} \\
&= \frac{1}{k} \sum_{j=0}^{k-1} \omega^{-ja} \sum_{b=0}^{n} (1-p)^{n-b} p^b \binom{n}{b} \omega^{jb} \\
&= \frac{1}{k} \sum_{j=0}^{k-1} (1-p+p\omega^j)^n \omega^{-aj}
\end{aligned}$$

as claimed.

It remains to show that

$$\max_{a \in [k]} \left| \Pi_{n,a} - \frac{1}{k} \right| \le \frac{1}{k} \exp\left\{ -\frac{4p(1-p)n}{3k^2} \right\}.$$

When $k = 1$, this is trivial. Suppose that $k > 1$ and fix $a \in [k]$. The triangle inequality implies that

$$\left| \Pi_{n,a} - \frac{1}{k} \right| = \left| \frac{1}{k} \sum_{j=1}^{k-1} (1-p+p\omega^j)^n \omega^{-aj} \right| \le \frac{1}{k} \sum_{j=1}^{k-1} \left| 1-p+p\omega^j \right|^n . \qquad (5.1)$$

Let $\theta_j = 2\pi j/k$ and $x = 1 - \cos\theta_1$. Then $\omega^j = \cos\theta_j + i\sin\theta_j$ and so

$$\begin{aligned}
\left| 1-p+p\omega^j \right|^2 &= (1-p+p\cos\theta_j)^2 + (p\sin\theta_j)^2 \\
&= (1-p)^2 + p^2 + 2(1-p)p\cos\theta_j \\
&\le 1 - 2p(1-p)x.
\end{aligned}$$

Hence (5.1) implies that

$$\left|\Pi_{n,a} - \frac{1}{k}\right| \leq (1 - 2p(1-p)x)^{n/2}.$$

However, some simple calculus shows that $1 - y \leq e^{-y}$ for positive y and hence

$$\left|\Pi_{n,a} - \frac{1}{k}\right| \leq e^{-p(1-p)xn}.$$

The Taylor expansion of cosine is $\cos\theta = \sum_{l=0}^{\infty} \frac{(-\theta^2)^l}{(2l)!}$. Recall that $k \geq 2$, which implies that $\left(\pi^2 - \frac{\pi^4}{3k^2}\right) \geq \frac{2}{3}$. Hence

$$\cos\theta_1 < 1 - \frac{\theta_1^2}{2!} + \frac{\theta_1^4}{4!} = 1 - \frac{2}{k^2}\left(\pi^2 - \frac{\pi^4}{3k^2}\right) \leq 1 - \frac{4}{3k^2} \tag{5.2}$$

and so $x > \frac{4}{3k^2}$. Thus, provided that $k = o\left(\sqrt{\frac{n}{\log n}}\right)$,

$$\left|\Pi_{n,a} - \frac{1}{k}\right| \leq \frac{1}{k}\exp\left\{-\frac{4p(1-p)n}{3k^2}\right\}.$$

□

The following result is the base case of Theorem 5.6.

Lemma 5.22. *Let $p \in (0,1)$ and $s \in \mathbb{N}$. Suppose that $\mathbf{x}$ is an $(n,[s])$-partition and that $k \in \mathbb{N}$ satisfies $k = o\left(\frac{\min(\mathbf{x})}{\sqrt{\log(\min(\mathbf{x}))}}\right)$. Let $G_n \in \mathcal{G}(n,p)$ and $\mathcal{P}$ be a partition of $[n]$ with profile $\mathbf{x}$. Then, for sufficiently large n, the random variable $X_{\mathcal{C}_{2,s}}^{\mathcal{P}}(G_n)$ is $2C(2,s)\exp\left\{-\frac{5p(1-p)\min(\mathbf{x})^2}{8k^2}\right\}$-uniform modulo k.*

Proof. Recall that $\mathcal{C}_{2,s} = \{(K_2,(i,j)) : 1 \leq i \leq j \leq s\}$. Then

$$X^{\mathcal{P}}_{(K_2,(i,j))}(G_n) \sim \begin{cases} \operatorname{Bin}\left(\binom{x_i}{2}, p\right) & i = j \\ \operatorname{Bin}\left(x_i x_j, p\right) & i \neq j \end{cases}$$

These variables are jointly independent as they depend on disjoint sets of edges in $\mathcal{G}(n,p)$.

Theorem 5.21 and the assumption that $k = o\left(\frac{\min(\mathbf{x})}{\sqrt{\log(\min(\mathbf{x}))}}\right)$ imply that

$$\begin{aligned}\mathbb{P}\left(X^{\mathcal{P}}_{(K_2,(i,i))}(G_n) \underset{k}{\equiv} a_{i,i}\right) &\leq \frac{1}{k}\left(1 + \exp\left\{\frac{-4p(1-p)\binom{x_i}{2}}{3k^2}\right\}\right) \\ &\leq \frac{1}{k}\left(1 + \exp\left\{\frac{-5p(1-p)\min\left(\mathbf{x}\right)^2}{8k^2}\right\}\right),\end{aligned}$$

provided that $\min\left(\mathbf{x}\right) \geq 16$. However, this is a consequence of the assumption that $k = o\left(\frac{\min(\mathbf{x})}{\sqrt{\log(\min(\mathbf{x}))}}\right)$.

As a similar argument holds for $X^{\mathcal{P}}_{(K_2,(i,j))}(G_n)$, the independence of the variables $X^{\mathcal{P}}_{(K_2,(i,j))}(G_n)$ implies that

$$\mathbb{P}(X^{\mathcal{P}}_{\mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a}) \leq \frac{1}{k^{C(2,s)}}\left(1 + \exp\left\{\frac{-5p(1-p)\min\left(\mathbf{x}\right)^2}{8k^2}\right\}\right)^{C(2,s)}.$$

As $k = o\left(\frac{\min(\mathbf{x})}{\sqrt{\log(\min(\mathbf{x}))}}\right)$, this implies that

$$\mathbb{P}(X^{\mathcal{P}}_{\mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a}) \leq \frac{1}{k^{C(2,s)}}\left(1 + 2C\left(2,s\right)\exp\left\{\frac{-5p(1-p)\min\left(\mathbf{x}\right)^2}{8k^2}\right\}\right).$$

Similarly

$$\mathbb{P}(X^{\mathcal{P}}_{\mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a}) \geq \frac{1}{k^{C(2,s)}} \left(1 - 2C\left(2,s\right) \exp\left\{\frac{-5p(1-p)\min\left(\mathbf{x}\right)^2}{8k^2}\right\}\right),$$

and so

$$\left|\mathbb{P}(X^{\mathcal{P}}_{\mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a}) - \frac{1}{k^{C(2,s)}}\right| \leq \frac{2C\left(2,s\right)}{k^{C(2,s)}} \exp\left\{\frac{-5p(1-p)\min\left(\mathbf{x}\right)^2}{8k^2}\right\},$$

completing the proof. □

5.3 The uniformity of $\mathcal{C}_{t,sd^2}$ implies that almost all graphs are refinable for $\mathcal{D}_{t,s,d}$

In this section, the first of the two results mentioned at the end of Section 5.1 is proved. The formal statement of this result is as follows.

Lemma 5.23. *Let $d \geq 2$, s and t be fixed integers, and let $p \in (0,1)$. Let $\lambda > 0$ and $0 < \mu < 1$ be real constants, and let $\nu = \min\{\lambda, \mu^4/2\}$. Let $k = k(n)$ be any integer valued function of n such that $k^{2D(t,s,d)} = o\left(e^{n/153}\right)$, and let $G_n \in \mathcal{G}(n,p)$.*

Suppose that the following two statements hold.

- *For every $[sd^2]$-partition $\mathcal{A}$ of $[n]$ with $\min\left(\mathcal{A}\right) \geq \mu^2 n/2$, the random variable $X^{\mathcal{A}}_{\mathcal{C}_{t,sd^2}}(G_n)$ is $e^{-\lambda n/k^2}$-uniform modulo k.*
- *For every $[sd]$-partition $\mathcal{A}$ of $[n]$ with $\min\left(\mathcal{A}\right) \geq \mu n$, the random variable $X^{\mathcal{A}}_{\mathcal{C}_{t,sd}}(G_n)$ is $e^{-\lambda n/k^2}$-uniform modulo k.*

Then, for any $[s]$-labelled partition $\mathcal{P}$ of $[n]$ and any $(n, [s] \times [d])$-partition $\mathbf{y}$ refining the profile of $\mathcal{P}$ such that $\min(\mathbf{y}) \geq \mu n$*, G_n is $e^{-\nu n/3k^2}$-refinable modulo k for $\mathcal{D}_{t,s,d}$, $\mathcal{P}$ and $\mathbf{y}$ with probability at least $1 - 16k^{D(t,s,d)}e^{-\nu n/3k^2}$.*

Recall that refinability is a property of a fixed graph G, which is assumed to have vertex set $[n]$. Let $\mathcal{P}$ be an $[s]$-labelled partition of $[n]$ and let $\mathbf{y}$ be an $(n, [s] \times [d])$-partition refining the profile of $\mathcal{P}$. G is refinable when, for a refinement $\mathcal{Q}$ chosen uniformly at random from $\mathcal{R}(\mathcal{P}, \mathbf{y})$, the variable $X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G)$ is ϵ-uniform modulo k. The proof that almost all graphs $G_n \in \mathcal{G}(n,p)$ are refinable will use a counting argument and the second moment method on the number of refinements $\mathcal{Q} \in \mathcal{R}(\mathcal{P}, \mathbf{y})$ such that $X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G) \underset{k}{\equiv} \mathbf{a}$.

A simple preliminary result is needed. The second moment method will require consideration of two refinements $\mathcal{Q}, \mathcal{Q}' \in \mathcal{R}(\mathcal{P}, \mathbf{x})$. Define $\mathcal{Q} \diamond \mathcal{Q}'$ to be the $T \times T$-refinement of $\mathcal{P}$ where $(Q \diamond Q')_{i,j,j'} = Q_{i,j} \cap Q'_{i,j'}$. In this case, the colourings associated with $\mathcal{Q} \diamond \mathcal{Q}'$ are from $[s] \times [d] \times [d]$, and hence, in light of Lemma 5.11, the assumption of Lemma 5.23 may be used as long as $\min(\mathcal{Q} \diamond \mathcal{Q}')$ is not too small. Lemma 5.24 will show that, for almost all $\mathcal{Q}, \mathcal{Q}' \in \mathcal{R}(\mathcal{P}, \mathbf{x})$, it is the case that $\min(\mathcal{Q} \diamond \mathcal{Q}')$ is not too small.

Lemma 5.24. *Let $0 < \mu < 1$, and let s and d be positive integers. Let n be a sufficiently large integer. Let $\mathcal{A}$ be an $[s]$-partition of $[n]$. Let $\mathbf{y}$ be an $(n, [s] \times [d])$-partition with* $\min(\mathbf{y}) \geq \mu n$*, such that $\mathbf{y}$ refines the profile of $\mathcal{A}$. Suppose that $\mathcal{B}$ and $\mathcal{C}$ are chosen uniformly at random from $\mathcal{R}(\mathcal{A}, \mathbf{y})$. Then* $\min(\mathcal{B} \diamond \mathcal{C}) \geq \frac{1}{2}\mu^2 n$ *with probability at least $1 - sd^2e^{-\frac{1}{2}\mu^4 n}$.*

Proof. It is sufficient to show that, for all $1 \le i \le s$ and $1 \le j, j' \le d$,

$$\mathbb{P}\left(|(B \diamond C)_{i,j,j'}| < \frac{1}{2}\mu^2 n\right) \le e^{-\frac{1}{2}\mu^4 n}.$$

Then by the union bound, as there are s choices for i and d^2 choices for j and j',

$$\mathbb{P}\left(\min\left(\mathcal{B} \diamond \mathcal{C}\right) < \frac{1}{2}\mu^2 n\right) \le sd^2 e^{-\frac{1}{2}\mu^4 n}$$

as required.

Fix i, j and j'. The sets $B_{i,j}$ and $C_{i,j'}$ are independently chosen uniformly at random from subsets of A_i of order $y_{i,j}$ and $y_{i,j'}$ respectively. Hence, it can be seen that $|(B \diamond C)_{i,j,j'}|$ has the hypergeometric distribution with parameters $|A_i|$, $y_{i,j}$ and $y_{i,j'}$. Note that

$$\frac{1}{2}\mu^2 n = \mu^2 n - \frac{1}{2}\mu^2 n \le \frac{y_{i,j} y_{i,j'}}{|A_i|} - \frac{1}{2}\mu^2 n.$$

Hence Lemma 5.18 implies that

$$\begin{aligned}\mathbb{P}\left(|(B \diamond C)_{i,j,j'}| < \frac{1}{2}\mu^2 n\right) &\le \mathbb{P}\left(|(B \diamond C)_{i,j,j'}| < \frac{y_{i,j} y_{i,j'}}{|A_i|} - \frac{1}{2}\mu^2 n\right) \\ &\le e^{-\frac{1}{2}\mu^4 n^2/|A'|} \le e^{-\frac{1}{2}\mu^4 n}\end{aligned}$$

as claimed. □

It is now possible to prove Lemma 5.23.

Proof of Lemma 5.23. Let $\mathcal{P}$ be an S-labelled partition of $[n]$, and let $\mathbf{y}$ be an $(n, [s] \times [d])$-partition refining the profile of $\mathcal{P}$ with $\min(\mathbf{y}) \ge \mu n$. The aim is to show that $G_n \in \mathcal{G}(n,p)$ is $e^{-\nu n/3k^2}$-refinable modulo k for $\mathcal{D}_{t,s,d}$, $\mathcal{P}$ and $\mathbf{y}$ with probability at least $1 - 16k^{D(t,s,d)}e^{-\nu n/3k^2}$.

Let G be a fixed graph with vertex set $[n]$. From Section 5.1, recall that G is $e^{-\nu n/3k^2}$-refinable modulo k for $\mathcal{D}_{t,s,d}$, $\mathcal{P}$ and $\mathbf{y}$ if, when $\mathcal{Q}$ is chosen uniformly at random from $\mathcal{R}(\mathcal{P},\mathbf{y})$,

$$\max_{\mathbf{a}\in[k]^{\mathcal{D}_{t,s,d}}} \left| \mathbb{P}\left(X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G) \underset{k}{\equiv} \mathbf{a} \right) - \frac{1}{k^{D(t,s,d)}} \right| < \frac{e^{-\nu n/3k^2}}{k^{D(t,s,d)}}.$$

Let $N := |\mathcal{R}(\mathcal{P},\mathbf{y})|$ and for $\mathbf{a} \in [k]^{\mathcal{D}_{t,s,d}}$, let $N_{\mathbf{a}} = N_{\mathbf{a}}(G)$ be the number of $\mathcal{Q} \in \mathcal{R}(\mathcal{P},\mathbf{y})$ such that $X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G) \underset{k}{\equiv} \mathbf{a}$. Then, when choosing $\mathcal{Q}$ uniformly at random from $\mathcal{R}(\mathcal{P},\mathbf{y})$, the value $\frac{N_{\mathbf{a}}(G)}{N}$ is precisely the probability that $X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G) \underset{k}{\equiv} \mathbf{a}$. Hence G is $e^{-\nu n/3k^2}$-refinable modulo k for $\mathcal{D}_{t,s,d}$, $\mathcal{P}$ and $\mathbf{y}$ if, and only if,

$$\max_{\mathbf{a}\in[k]^{\mathcal{D}_{t,s,d}}} \left| N_{\mathbf{a}}(G) - \frac{N}{k^{D(t,s,d)}} \right| < \frac{e^{-\nu n/3k^2}N}{k^{D(t,s,d)}}.$$

It is sufficient to prove that, for all $\mathbf{a} \in [k]^{\mathcal{D}_{t,s,d}}$,

$$\mathbb{P}\left(\left| N_{\mathbf{a}}(G_n) - \frac{N}{k^{D(t,s,d)}} \right| \geq \frac{e^{-\nu n/3k^2}N}{k^{D(t,s,d)}} \right) \leq 16e^{-\nu n/3k^2}. \tag{5.3}$$

Then, as there are $k^{D(t,s,d)}$ choices for $\mathbf{a}$, the union bound implies that

$$\mathbb{P}\left(\max_{\mathbf{a}\in[k]^{\mathcal{D}_{t,s,d}}} \left| N_{\mathbf{a}}(G_n) - \frac{N}{k^{D(t,s,d)}} \right| \geq \frac{e^{-\nu n/3k^2}N}{k^{D(t,s,d)}} \right) \leq 16k^{D(t,s,d)}e^{-\nu n/3k^2},$$

completing the proof.

Fix $\mathbf{a} \in [k]^{\mathcal{D}_{t,s,d}}$. Note that $N_{\mathbf{a}}(G_n)$ is a random variable, given by

$$N_{\mathbf{a}}(G_n) = \sum_{\mathcal{Q} \in \mathcal{R}(\mathcal{P},\mathbf{y})} \mathbb{1}\left(X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G_n) \underset{k}{\equiv} \mathbf{a}\right).$$

Hence, taking expectations

$$\mathbb{E}(N_{\mathbf{a}}(G_n)) = \sum_{\mathcal{Q} \in \mathcal{R}(\mathcal{P},\mathbf{y})} \mathbb{P}\left(X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G_n) \underset{k}{\equiv} \mathbf{a}\right). \tag{5.4}$$

Recall that $\nu = \min\{\lambda, \mu^2/9\}$. Fix $\mathcal{Q} \in \mathcal{R}(\mathcal{P}, \mathbf{y})$. By assumption, $X^{\mathcal{A}}_{\mathcal{C}_{t,sd}}(G_n)$ is $e^{-\lambda n/k^2}$-uniform modulo k, for every $[sd]$-partition $\mathcal{A}$ of $[n]$ with $\min(\mathcal{A}) \geq \mu n$. Hence, as $\mathcal{Q}$ has profile $\mathbf{y}$ and $\min(\mathbf{y}) \geq \mu n$, the variable $X^{\mathcal{Q}}_{\mathcal{C}_{t,sd}}(G_n)$ is $e^{-\lambda n/k^2}$-uniform modulo k. By Lemma 5.11, this implies that $X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G_n)$ is $e^{-\lambda n/k^2}$-uniform modulo k. Hence

$$\left|\mathbb{P}\left(X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G_n) \underset{k}{\equiv} \mathbf{a}\right) - \frac{1}{k^{D(t,s,d)}}\right| < \frac{e^{-\lambda n/k^2}}{k^{D(t,s,d)}},$$

and, summing over $\mathcal{Q}$, it follows from (5.4) that

$$\left|\mathbb{E}(N_{\mathbf{a}}(G_n)) - \frac{N}{k^{D(t,s,d)}}\right| < \frac{e^{-\lambda n/k^2} N}{k^{D(t,s,d)}}. \tag{5.5}$$

The second moment of $N_{\mathbf{a}}$ will now be bounded. Note first that

$$N_{\mathbf{a}}(G_n)^2 = \sum_{(\mathcal{Q},\mathcal{Q}') \in \mathcal{R}(\mathcal{P},\mathbf{y})^2} \mathbb{1}\left(X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G_n) \underset{k}{\equiv} \mathbf{a}\right) \mathbb{1}\left(X^{\mathcal{Q}'}_{\mathcal{D}_{t,s,d}}(G_n) \underset{k}{\equiv} \mathbf{a}\right)$$

and so

$$\mathbb{E}N_{\mathbf{a}}(G_n)^2 = \sum_{(\mathcal{Q},\mathcal{Q}')\in\mathcal{R}(\mathcal{P},\mathbf{y})^2} \mathbb{P}\left(X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G_n) \underset{k}{\equiv} \mathbf{a}, X^{\mathcal{Q}'}_{\mathcal{D}_{t,s,d}}(G_n) \underset{k}{\equiv} \mathbf{a} \right). \tag{5.6}$$

For $(\mathcal{Q},\mathcal{Q}') \in \mathcal{R}(\mathcal{P},\mathbf{y})^2$, recall that $\mathcal{Q}\diamond\mathcal{Q}'$ is the refinement of $\mathcal{P}$ where $(Q\diamond Q')_{i,j,j'} = Q_{i,j}\cap Q'_{i,j'}$ for all $i \in S$ and $j, j' \in [d]$. For $(H, c_1 \times c_2) \in \mathcal{D}_{t,s,d}$, a copy of $(H, c_1 \times c_2)$ in $(G_n, \mathcal{Q})$ must necessarily be a copy of $(H, c_1 \times c_2 \times c')$ in $(G_n, \mathcal{Q}\diamond\mathcal{Q}')$ for some $c' \in \mathrm{col}(H, [d])$, and vice versa. Hence

$$X^{\mathcal{Q}}_{(H,c_1\times c_2)}(G_n) = \sum_{c'\in\mathrm{col}(H,[d])} X^{\mathcal{Q}\diamond\mathcal{Q}'}_{(H,c_1\times c_2\times c')}(G_n) \tag{5.7}$$

and similarly

$$X^{\mathcal{Q}'}_{(H,c_1\times c_2)}(G_n) = \sum_{c'\in\mathrm{col}(H,[d])} X^{\mathcal{Q}\diamond\mathcal{Q}'}_{(H,c_1\times c'\times c_2)}(G_n). \tag{5.8}$$

Let $\Lambda = \Lambda(\mathbf{a})$ be the set of $\mathbf{b} \in [k]^{\mathcal{C}_{t,[s]\times[d]\times[d]}}$ such that, for all coloured graphs $(H, c_1 \times c_2) \in \mathcal{D}_{t,s,d}$,

$$b_{(H,c_1\times c_2\times\underline{1})} \underset{k}{\equiv} a_{(H,c_1\times c_2)} - \sum_{c'\in\mathrm{col}(H,[d])\setminus\{\underline{1}\}} b_{(H,c_1\times c_2\times c')}$$

and

$$b_{(H,c_1\times\underline{1}\times c_2)} \underset{k}{\equiv} a_{(H,c_1\times c_2)} - \sum_{c'\in\mathrm{col}(H,[d])\setminus\{\underline{1}\}} b_{(H,c_1\times c'\times c_2)},$$

where $\underline{1}$ is the colouring where each vertex of H has colour 1.

It is clear from (5.7) and (5.8) that $X^{\mathcal{Q}\diamond\mathcal{Q}'}_{\mathcal{C}_{t,[s]\times[d]\times[d]}}(G_n) \underset{k}{\in} \Lambda$ if, and only if, both $X^{\mathcal{Q}}_{\mathcal{D}_{t,s,d}}(G_n) \underset{k}{\equiv} \mathbf{a}$ and $X^{\mathcal{Q}'}_{\mathcal{D}_{t,s,d}}(G_n) \underset{k}{\equiv} \mathbf{a}$.

By definition, for all $(H, c_1 \times c_2) \in \mathcal{D}_{t,s,d}$ the colouring $c_2 \neq \underline{1}$. Hence, for each choice of $(H, c_1 \times c_2) \in \mathcal{D}_{t,s,d}$, both $b_{(H,c_1\times c_2\times\underline{1})}$ and $b_{(H,c_1\times\underline{1}\times c_2)}$ appear only on the left hand side of the above equations, and so are uniquely determined, given the other values $b_{(H,c_1\times c_2\times c_3)}$. As there are $D(t, s, d)$ choices for H, c_1 and c_2, and each gives us two dependent values, the set Λ has order $k^{C(sd^2,t)-2D(s,d,t)}$. Recall from Section 5.1 that G_n cannot be refinable for $\mathcal{C}_{t,sd}$; this is the point in the proof at which this argument fails for that set.

By assumption, if $\min(\mathcal{Q}\diamond\mathcal{Q}') \geq \mu^2 n/2$, the variable $X^{\mathcal{Q}\diamond\mathcal{Q}'}_{\mathcal{C}_{t,[s]\times[d]\times[d]}}(G_n)$ is $e^{-\lambda n/k^2}$-uniform modulo k and so, for all $\mathbf{c} \in [k]^{\mathcal{C}_{t,[s]\times[d]\times[d]}}$,

$$\mathbb{P}\left(X^{\mathcal{Q}\diamond\mathcal{Q}'}_{\mathcal{C}_{t,[s]\times[d]\times[d]}}(G_n) \underset{k}{\equiv} \mathbf{c}\right) \leq \frac{1+e^{-\lambda n/k^2}}{k^{C(t,sd^2)}}. \tag{5.9}$$

However, by Lemma 5.24, as $\min(\mathbf{y}) \geq \mu n$,

$$\left|\left\{(\mathcal{Q},\mathcal{Q}') \in \mathcal{R}(\mathcal{P},\mathbf{y})^2 : \min(\mathcal{Q}\diamond\mathcal{Q}') < \mu^2 n/2\right\}\right| < N^2 s d^2 e^{-\frac{1}{2}\mu^4 n}.$$

Dividing the sum in (5.6) into those $\mathcal{Q},\mathcal{Q}' \in \mathcal{R}(\mathcal{P},\mathbf{y})$ with $\min(\mathcal{Q}\diamond\mathcal{Q}') < \mu^2 n/2$ and those with $\min(\mathcal{Q}\diamond\mathcal{Q}') \geq \mu^2 n/2$ gives

$$\begin{aligned}\mathbb{E}N_{\mathbf{a}}(G_n)^2 &= \sum_{(\mathcal{Q},\mathcal{Q}')\in\mathcal{R}(\mathcal{P},\mathbf{y})^2} \mathbb{P}\left(X^{\mathcal{Q}\diamond\mathcal{Q}'}_{\mathcal{C}_{t,[s]\times[d]\times[d]}}(G_n) \underset{k}{\in} \Lambda\right) \\ &\leq N^2 s d^2 e^{-\frac{1}{2}\mu^4 n} + \sum_{\substack{\mathcal{Q},\mathcal{Q}'\in\mathcal{R}(\mathcal{P},\mathbf{y}) \\ \min(\mathcal{Q}\diamond\mathcal{Q}')\geq\frac{1}{2}\mu^2 n}} \sum_{\mathbf{b}\in\Lambda} \mathbb{P}\left(X^{\mathcal{Q}\diamond\mathcal{Q}'}_{\mathcal{C}_{t,[s]\times[d]\times[d]}}(G_n) \underset{k}{\equiv} \mathbf{b}\right).\end{aligned}$$

But then, as $\nu = \min\{\mu^4/2, \lambda\}$ and $k^{2D(t,s,d)} = o\left(e^{n/153}\right)$, it follows from (5.9) that

$$\begin{aligned}\mathbb{E}N_{\mathbf{a}}(G_n)^2 &\leq N^2 sd^2 e^{-\frac{1}{2}\mu^4 n} + N^2 k^{C\left(t,sd^2\right)-2D(t,s,d)} \frac{1+e^{-\lambda n/k^2}}{k^{C(t,sd^2)}} \\ &\leq \frac{N^2}{k^{2D(t,s,d)}}\left(1+2e^{-\nu n/k^2}\right). \end{aligned} \tag{5.10}$$

Hence, by (5.5) and (5.10),

$$\begin{aligned}\mathrm{Var}\left(N_{\mathbf{a}}(G_n)\right) &\leq \frac{N^2}{k^{2D(t,s,d)}}\left(1+2e^{-\nu n/k^2}\right) - \left(\frac{N}{k^{D(t,s,d)}} - \frac{Ne^{-\nu n/k^2}}{k^{D(t,s,d)}}\right)^2 \\ &\leq \frac{4N^2 e^{-\nu n/k^2}}{k^{2D(t,s,d)}}\end{aligned}$$

and so, by Lemma 5.16, with $a = \frac{e^{-\nu n/3k^2}N}{2k^{D(t,s,d)}}$, it follows from (5.5) that

$$\begin{aligned}\mathbb{P}\left(\left|N_{\mathbf{a}}(G_n) - \frac{N}{k^{D(t,s,d)}}\right| \geq \frac{e^{-\nu n/3k^2}N}{k^{D(t,s,d)}}\right) &\leq \mathbb{P}(|N_{\mathbf{a}}(G_n) - \mathbb{E}(N_{\mathbf{a}}(G_n))| \geq a) \\ &\leq \frac{4e^{-\nu n/k^2}}{\left(\frac{1}{2}e^{-\nu n/3k^2}\right)^2} \\ &= 16e^{-\nu n/3k^2}\end{aligned}$$

as required in (5.3). This completes the proof. □

5.4 Almost all graphs are refinable for $\mathcal{D}_{t,s,2^s}$ implies that $\mathcal{C}_{t+1,s}$ is uniform

In this section, the second of the two results mentioned at the end of Section 5.1 is proved. This states that, if almost all graphs are refinable for $\mathcal{D}_{t,s,2^s}$, then $X_{\mathcal{C}_{t+1,s}}$ is uniform. The formal statement and proof of this result are given in Lemma 5.28, which appears later in this section.

Throughout this section, k will be a function depending on n. Where it is not specified, it will be assumed that $k = k(n)$. Also, throughout this section, S will be the set $\{0,1\}^s$, where $\mathbf{0}$ is the minimal element of S. Recall that $\mathcal{D}_{t,[s],S}$ is the set of $(H, c_1 \times c_2)$ with $H \in \mathcal{C}_t$, where c_1 is some $[s]$-colouring of H and c_2 is some S-colouring of H other than the monochromatic colouring of H with colour $\mathbf{0}$.

Before beginning the proof of Lemma 5.28, the following discussion will consider the case $s = 1$. In this case, there is only one $[s]$-labelled partition $\mathcal{P}$ of $[n]$ and so proving that $X_{\mathcal{C}_{t+1}}(G_n)$ is approximately uniform modulo k is sufficient. The initial assumption is that, when $\mathbf{y}$ is a refinement of the profile of $\mathcal{P}$ and $\min(\mathbf{y})$ is not too small, with high probability a graph $G_n \in \mathcal{G}(n,p)$ is approximately refinable modulo k for $\mathcal{D}_{t,1,2}$, $\mathcal{P}$ and $\mathbf{y}$.

As mentioned in Section 5.1, the main idea of the proof is to fix a vertex $u \in [n]$. This allows decomposition of G_n into a graph $G'_{n-1} = G_n \setminus \{u\} \in \mathcal{G}(n-1,p)$ and a random partition $\mathcal{B}$ of $[n] - \{u\}$ into neighbours and non-neighbours of u. Note that the properties of $\mathcal{G}(n,p)$ imply that G'_{n-1} and $\mathcal{B}$ are independent.

Given the value of $d(u)$, the partition $\mathcal{B}$ is chosen uniformly at random from $\mathcal{R}([n]-\{u\},(d(u),n-1-d(u)))$, where (with high probability) neither $d(u)$ or $n-1-d(u)$ is too small. By assumption, it follows that with high probability G'_{n-1} is approximately refinable modulo k for $\mathcal{D}_{t,1,2}$, $\mathcal{P}$ and $(d(u),n-1-d(u))$.

Now condition on any fixed G'_{n-1} and $\mathbf{y}=(d(u),n-1-d(u))$ such that G'_{n-1} is approximately refinable modulo k for $\mathcal{D}_{t,1,2}$, $\mathcal{P}$ and $\mathbf{y}$. For $H\in\mathcal{C}_{t+1}\setminus\mathcal{C}_2$, a copy of H in G_n is either a copy of H in G'_{n-1} or contains the vertex u.

Suppose that there is some copy of H in G_n which contains u. Let $v\in V(H)$ be the vertex appearing at u so that there is a copy of $H\setminus\{v\}$ in G'_{n-1}. In fact, there is a copy of $(H\setminus\{v\},c)$ in $\left(G'_{n-1},\mathcal{B}\right)$ for some S-colouring c of H. Any neighbour of v in H must, in this copy, be a neighbour of u, and so, in particular, c is not monochromatic with colour $\mathbf{0}$.

Hence $X_H(G_n)-X_H(G'_{n-1})$ is the sum of $X^{\mathcal{B}}_{(F,c)}(G'_{n-1})$ for some graphs (F,c) where F is of smaller order than H and c is not monochromatic with colour $\mathbf{0}$. In particular, any such graph (F,c), if connected, is a member of $\mathcal{D}_{t,1,2}$. By assumption, this means that $X_H(G_n)$ is approximately uniform modulo k.

There are a number of things to notice about this approach which complicate matters. Firstly, the set $\mathcal{D}_{t,1,2}$ consists of connected graphs, and if H contains a cutvertex v, then the graph $H\setminus\{v\}$ is not connected. However, Lemma 5.25 will show that, for a disconnected graph F, the value of $X^{\mathcal{B}}_{(F,c)}(G'_{n-1})$ is determined by $X^{\mathcal{B}}_{\mathcal{D}_{|F|-1,1,2}}(G'_{n-1})$. This is sufficient to allow the above method to proceed.

Also, the above analysis shows that for each graph $H\in\mathcal{C}_{t+1}\setminus\mathcal{C}_2$, the random variable $X_H(G_n)$ is approximately uniform modulo k. However, the aim is to show that $X_{\mathcal{C}_{t+1}\setminus\mathcal{C}_2}(G_n)$ is approximately uniform modulo k and it is possible that some of the variables $X_H(G_n)$ are dependent modulo k. Hence it is necessary to show that the random variables summing to $X_H(G_n) - X_H(G'_{n-1})$ take particular forms implying this is not the case. This is done in Lemma 5.27.

Note also that graphs in $\mathcal{C}_2$ (i.e. the graph K_2) are not considered. The reason for this is simple, as the number of edges in G_n is simply the sum of the number of edges in G'_{n-1} and the degree of the vertex u, and hence is fixed by the conditioning. However, the degree of the vertex u is binomially distributed, and so, with care, Theorem 5.21 may be used to show that this too is approximately uniformly distributed modulo k.

Finally, the case with $s > 1$ is similar. However, instead of just considering one vertex, a set of vertices, one in each set of the partition $\mathcal{P}$ is considered.

The following lemma is that mentioned above concerning disconnected graphs. Note that the first part of the lemma appears in [32] as Lemma 6.2.

Lemma 5.25. *Let G be a graph and $\mathcal{A}$ be a $T \times T'$-indexed partition of $V(G)$. Suppose that $(H, c_1 \times c_2)$ is a $T \times T'$-coloured disconnected graph. Then $X^{\mathcal{A}}_{(H,c_1\times c_2)}(G)$ is a function of $X^{\mathcal{A}}_{\mathcal{D}_{(H,c_1\times c_2)}}(G)$ for some set*

$$\mathcal{D}_{(H,c_1\times c_2)} \subseteq \mathcal{C}_{|H|-1,T\times T'} \cup \{(K_1,(i,j)) : (i,j) \in T \times T'\}.$$

Further, suppose that, under the colouring c_2, no component of H is *monochromatic with colour $1_{T'}$. Then*

$$\mathcal{D}_{(H,c_1\times c_2)} \subseteq \mathcal{D}_{|H|-1,T,T'} \cup \{(K_1,(i,j)) : (i,j) \in T \times T'\}.$$

Proof. Suppressing unnecessary notation let $X_{(H',c_1\times c_2)} = X^{\mathcal{A}}_{(H',c_1\times c_2)}(G)$ for any graph H'. The proof of both claims is by induction on $|H| \geq 2$. If H is a disconnected graph of order two, then it contains two isolated vertices u and v. If both $c_1(u) = c_1(v)$ and $c_2(u) = c_2(v)$, then $X_{(H,c_1\times c_2)} = \binom{X_{(K_1,(c_1(u),c_2(u)))}}{2}$. Otherwise, $X_{(H,c_1\times c_2)} = X_{(K_1,(c_1(u),c_2(u)))}X_{(K_1,(c_1(v),c_2(v)))}$.

Suppose now that H has order at least three and let

$$(H_1, c_{1,1} \times c_{1,2}), \cdots, (H_r, c_{r,1} \times c_{r,2})$$

be the distinct non-colour-isomorphic components of $(H, c_1 \times c_2)$ with each $(H_i, c_{i,1} \times c_{i,2})$ having multiplicity m_i. Note that each H_i is connected and of smaller order than H.

For $1 \leq i \leq r$, let $\left(H_i^{(j)}\right)_{j=1}^{m_i}$ be m_i distinct copies of $(H_i, c_{i,1} \times c_{i,2})$ in $(G, \mathcal{A})$. If the graphs $H_i^{(j)}$ are vertex-disjoint, then their union is a copy of $(H, c_1 \times c_2)$ in $(G, \mathcal{A})$. If not, their union is, in $(G, \mathcal{A})$, a copy of a graph obtained from $(H, c_1 \times c_2)$ by identifying some sets of vertices which have the same colour under both c_1 and c_2, subject to the condition that no two vertices in the same component are identified.

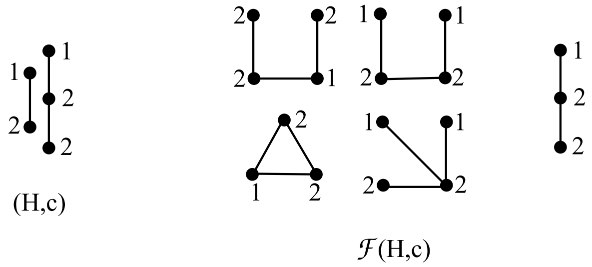


Figure 5.1: A [2]-coloured graph (H, c) and $\mathcal{F}(H, c)$. Note that (H, c) is a graph with two components, while $\mathcal{F}(H, c)$ is a set of five connected graphs.

As illustrated in Figure 5.1, let $\mathcal{F}(H, c_1 \times c_2)$ be the set of graphs obtained from $(H, c_1 \times c_2)$ by identifying some sets of vertices as described above. For each $(F, \ell) \in \mathcal{F}(H, c_1 \times c_2)$, let $\lambda((H, c_1 \times c_2); (F, \ell))$ be the number of ways a copy of (F, ℓ) can be obtained in this way from m_i copies of each $(H_i, c_{i,1} \times c_{i,2})$. Then

$$X_{(H,c_1\times c_2)} = \prod_{i=1}^{r} \binom{X_{(H_i,c_{i,1}\times c_{i,2})}}{m_i} - \sum_{(F,\ell)\in\mathcal{F}(H,c_1\times c_2)} \lambda((H, c_1 \times c_2); (F, \ell)) X_{(F,\ell)}.$$

All graphs in $\mathcal{F}(H, c_1 \times c_2)$ are of smaller order than H. If $(F, \ell) \in \mathcal{F}(H, c_1 \times c_2)$ is disconnected, then by the inductive hypothesis, $X_{(F,\ell)}$ is a function of $X_{\mathcal{D}_{(F,\ell)}}$ for some set $\mathcal{D}_{(F,\ell)} \subseteq \mathcal{C}_{|F|-1,T\times T'} \cup \{(K_1, (i, j)) : (i, j) \in T \times T'\}$. Hence the first part of the result holds, with

$$\mathcal{D}_{(H,c_1\times c_2)} = \{(H_i, c_{i,1} \times c_{i,2}) : 1 \le i \le r\} \cup \bigcup_{(F,\ell)\in\mathcal{F}(H,c_1\times c_2)} \mathcal{D}_{(F,\ell)},$$

where $\mathcal{D}_{(F,\ell)} = \{(F, \ell)\}$ when F is connected.

Suppose that, under the colouring c_2, no component of H is monochromatic with colour $1_{T'}$. By definition, this is also true for all graphs in $\mathcal{F}(H, c_1 \times c_2)$. Hence, the inductive hypothesis implies that

$$\bigcup_{(F,\ell)\in\mathcal{F}(H,c_1\times c_2)} \mathcal{D}_{(F,\ell)} \subseteq \mathcal{D}_{|H|-1,T,T'} \cup \{(K_1,(i,j)) : (i,j) \in T \times T'\}.$$

However, each component of H is either an isolated vertex or a connected graph of order at least two and hence

$$\{(H_i, c_{i,1} \times c_{i,2}) : 1 \leq i \leq r\} \subseteq \mathcal{D}_{|H|-1,T,T'} \cup \{(K_1,(i,j)) : (i,j) \in T \times T'\}.$$

This completes the proof. □

Prior to stating the second preliminary lemma, the following definition formalises the situation described earlier, in which one vertex from each set of a partition is considered. Recall that $S = \{0,1\}^s$.

Definition 5.26. *Let $\mathcal{P}$ be an $[s]$-labelled partition of $[n]$, and let G be any graph with vertex set $[n]$. The* standard situation *defines sets U and V' and partitions $\mathcal{P}'$ and $\mathcal{B}$ in the following way. Firstly choose $U = \{u_1, \ldots, u_s\}$ to be a set consisting of one vertex u_i from each set P_i. The set V' is defined to be $[n] \setminus U$ and the partition $\mathcal{P}'$ is the restriction of the partition $\mathcal{P}$ to the set V'. Finally, let $\mathcal{B}$ be the S-refinement of $\mathcal{P}'$ where, for all $1 \leq i \leq s$ and $\alpha \in S$,*

$$B_{i,\alpha} = \{w \in P_i : wu_j \in E(G) \textit{ if, and only if, } \alpha_j = 1\}.$$

The second of the two preliminary lemmas follows. Recall that $S =$

$\{0,1\}^s$.

Lemma 5.27. *Let $t \geq 2$ and s be integers. There exists an injection ρ from $\mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$ to $\mathcal{D}_{t,[s],S}$ and a partial order $\prec$ on the set $\mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$ such that the following holds, for any graph G with vertex set $[n]$.*

Let $Q((H,c)) = \mathcal{D}_{t,s,S} \setminus \{\rho(H',c') : (H,c) \prec (H',c')\}$. Let $\mathcal{P}$ be an $[s]$-labelled partition of $[n]$, and suppose that U, V', $\mathcal{P}'$ and $\mathcal{B}$ are defined according to the standard situation. Then, for all $(H,c) \in \mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$,

$$X^{\mathcal{P}}_{(H,c)}(G) - X^{\mathcal{P}'}_{(H,c)}(G[V']) = X^{\mathcal{B}}_{\rho((H,c))}(G[V']) + f_{(H,c)},$$

where $f_{(H,c)}$ depends only on the graph $G[U]$, the size of the sets $B_{i,\alpha} \in \mathcal{B}$ and the value of $X^{\mathcal{B}}_{(F,\ell)}(G[V'])$ for graphs $(F,\ell) \in Q((H,c))$.

Proof. For $(H,c) \in \mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$, let $w_{(H,c)}$ be a non-cutvertex of H with minimal degree (if there is more than one such vertex, take any of them). Let $\rho(H,c) = \left(H - w_{(H,c)}, c \times \ell^{(H,c)}\right)$ where the colouring $\ell^{(H,c)} \in \operatorname{col}(H,S)$ is given by

$$\left(\ell^{(H,c)}(u)\right)_i = \begin{cases} 1 & uw_{(H,c)} \in E(H) \text{ and } i = c\left(w_{(H,c)}\right) \\ 0 & \text{otherwise} \end{cases}.$$

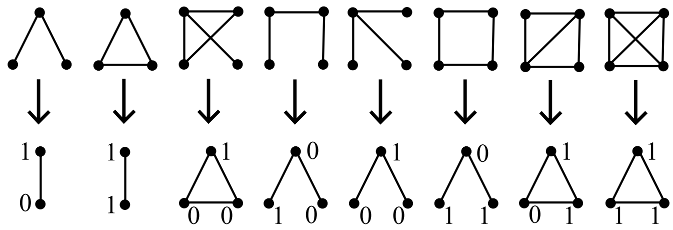


Figure 5.2: The function ρ from $\mathcal{C}_{4,1} \setminus \mathcal{C}_{2,1}$ (top) to $\mathcal{D}_{3,1,\{0,1\}}$ (bottom); the trivial colouring with colour 1 is supressed.

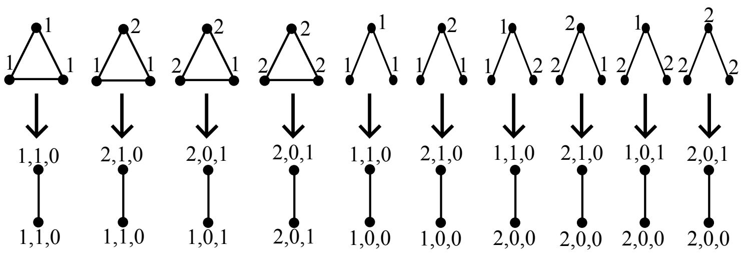


Figure 5.3: The function ρ from $\mathcal{C}_{3,2} \setminus \mathcal{C}_{2,2}$ (top) to $\mathcal{D}_{2,2,\{0,1\}^2}$ (bottom).

Also, let $(H_0, c_0) \prec (H_1, c_1)$ if $|H_0| < |H_1|$ or $|H_0| = |H_1|$ and $d(w_{(H_0,c_0)}) > d(w_{(H_1,c_1)})$. For the case $s = 1$ and $t = 3$, ρ and $\prec$ are illustrated in Figures 5.2 and 5.4 respectively, while Figure 5.3 illustrates ρ the case $t = 2$ and $s = 2$. It is simple to check that ρ is an injection from $\mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$ to $\mathcal{D}_{t,s,S}$ and that $\prec$ is a partial order on the set $\mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$.

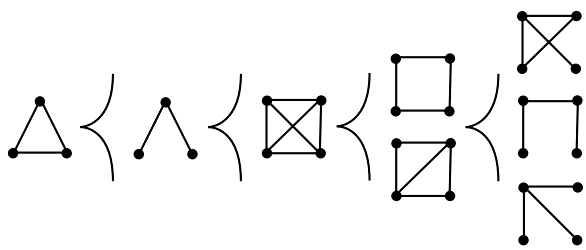


Figure 5.4: The partial order $\prec$ on $\mathcal{C}_{4,1} \setminus \mathcal{C}_{2,1}$, graphs on the left precede graphs on the right.

Fix $(H, c) \in \mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$ and suppose that the *standard situation* occurs. For $I \subseteq [s]$, let $U_I = \{u_i : i \in I\}$. Let $(H, c)^I_*$ be the set of all I-tuples $(w_i)_{i \in I}$ for which each w_i is a vertex of H coloured i. Recall that an *automorphism* of (H, c) is an automorphism of H which fixes the colouring c. Then, for all $I \subseteq S$, the automorphism group of (H, c) acts on $(H, c)^I_*$ in the obvious way. Let $(H, c)^I$ be a set of representatives of each orbit of this action.

A copy of (H, c) in $(G, \mathcal{P})$ intersects U in U_I for some $I \subseteq [s]$. Let $W = \{w_i : i \in I\}$ be the set of vertices of (H, c) intersecting U in this copy, with w_i the vertex of (H, c) at u_i. For all $i \in I$, the vertex u_i lies in P_i, and so the colour of w_i under c must equal i for all $i \in I$. Hence, after relabelling, W lies in $(H, c)^I$.

For $I \subseteq [s]$ and $W \in (H, c)^I$, let $X^{\mathcal{P}}_{((H,c);W)}(G; U_I)$ be the number of copies of (H, c) in $(G, \mathcal{P})$ which intersect U in U_I, with u_i being the image of w_i. Then, as it is unlabelled copies being counted,

$$X^{\mathcal{P}}_{(H,c)}(G) = \sum_{I \subseteq [s]} \sum_{W \in (H,c)^I} X^{\mathcal{P}}_{((H,c);W)}(G; U_I). \tag{5.11}$$

Fix $I \subseteq [s]$ and $W \in (H, c)^I$. Let $L_{H,W}$ be the set of S-colourings ℓ of $H - W$ such that, for all $v \in H - W$ and $1 \leq i \leq s$, the value of $(\ell(v))_i$ is one when $i \in I$ and vw_i is an edge (and may be one or zero otherwise). Suppose there is a copy of (H, c) in $(G, \mathcal{P})$ which intersects U in U_I, with u_i being the image of w_i. Then a vertex corresponding to $v \in H - W$ must lie in $P'_{c(v)}$ and hence in $B_{c(v),\alpha(v)}$ for some $\alpha(v) \in S$. However, the vertex corresponding to $v \in H - W$ must lie in the neighbourhood of u_i for all i such that vw_i is an edge of H. Hence $(\alpha(v))_i = 1$ for any $i \in I$ for which $vw_i \in E(H)$ and so the colouring α is in $L_{H,W}$. Thus, a copy of (H,c) in $(G,\mathcal{P})$ which intersects U in U_I, with u_i being the image of w_i consists of:

- a copy of $(H[W],c)$ in $(G[U],\mathcal{P})$;
- a copy of $(H-W,c\times\ell)$ in $(G[V'],\mathcal{B})$ for some $\ell\in L_{H,W}$.

A copy of $(H[W],c)$ in $(G[U],\mathcal{P})$ and a copy of $(H-W,c\times\ell)$ in $(G[V'],\mathcal{B})$ for some $\ell\in L_{H,W}$ corresponds to a fixed number, $\tau_{(H,c);W;\ell}$ of copies of (H,c) in $(G,\mathcal{P})$ which intersect U in U_I, with u_i being the image of w_i. Note that, in particular,$\tau_{(H,c);\{w_{(H,c)}\};\ell^{(H,c)}}=1$.

Let $E_{W,U}$ be the event that there is a copy of $(H[W],c)$ in $(G[U],\mathcal{P})$. If $E_{W,U}$ does not hold then there are no copies of (H,c) in $(G,\mathcal{P})$ which intersect U in U_I, with u_i being the image of w_i. If, however, $E_{W,U}$ does hold then there is a one-to-$\tau_{(H,c);W;\ell}$ correspondence between copies of $(H-W,c\times\ell)$ in $(G[V'],\mathcal{B})$ for some $\ell\in L_{H,W}$ and copies of (H,c) in $(G,\mathcal{P})$ which intersect U in U_I, with u_i being the image of w_i. Hence

$$X^{\mathcal{P}}_{((H,c);W)}(G;U_I)=\mathbb{1}\left(E_{W,U}\right)\sum_{\ell\in L_{H,W}}\tau_{(H,c);W;\ell}X^{\mathcal{B}}_{(H-W,c\times\ell)}(G[V']).$$

Together with (5.11), this implies that

$$X^{\mathcal{P}}_{(H,c)}(G)=\sum_{I\subseteq S}\sum_{W\in(H,c)^I}\mathbb{1}\left(E_{W,U}\right)\sum_{\ell\in L_{H,W}}\tau_{(H,c);W;\ell}X^{\mathcal{B}}_{(H-W,c\times\ell)}(G[V']).$$

Note that $E_{W,U}$ depends only on $G[U]$. When $I=\emptyset$, W must also be empty, $E_{W,U}$ must always occur, and $L_{H,W}=\mathrm{col}\,(H,S)$. Also $\tau_{(H,c);W;\ell}=1$ for all $\ell\in L_{H,W}$. Hence, in this case, the inner sum is equal to $X^{\mathcal{P}'}_{(H,c)}(G[V'])=$

$\sum_{\ell\in\mathrm{col}(H,S)} X^{\mathcal{B}}_{(H,c\times\ell)}(G[V'])$.

However, with $I = \{c(w_{(H,c)})\}$, $W = \{w_{(H,c)}\}$ and $\ell = \ell^{(H,c)}$, the right-hand side of the above sum has a term equal to $X^{\mathcal{B}}_{\rho((H,c))}(G[V'])$. Hence

$$X^{\mathcal{P}}_{(H,c)}(G) - X^{\mathcal{P}'}_{(H,c)}(G[V']) = X^{\mathcal{B}}_{\rho((H,c))}(G[V']) + f_{(H,c)}$$

where

$$\begin{aligned} &f_{(H,c)} + X^{\mathcal{B}}_{\rho((H,c))}(G[V']) \\ &\qquad = \sum_{\emptyset\neq I\subseteq[s]} \sum_{W\in(H,c)^I} \mathbb{1}\left(E_{W,U}\right) \sum_{\ell\in L_{H,W}} \tau_{(H,c);W;\ell} X^{\mathcal{B}}_{(H-W,c\times\ell)}(G[V']). \end{aligned} \tag{5.12}$$

It remains to show that the right hand side of (5.12) depends only on the graph $G[U]$, the size of the sets $B_{i,\alpha}\in\mathcal{B}$ and the value of $X^{\mathcal{B}}_{(F,\ell)}(G[V'])$ for graphs $(F,\ell)\in Q\left((H,c)\right)$. It is clear that $\mathbb{1}\left(E_{W,U}\right)$ depends only on the graph $G[U]$. It remains to show, for all $\emptyset\neq I\subseteq[s]$, $W\in(H,c)^I$ and $\ell\in L_{H,W}$, that $X^{\mathcal{B}}_{(H-W,c\times\ell)}(G[V'])$ depends only on the graph $G[U]$, the size of the sets $B_{i,\alpha}\in\mathcal{B}$ and the value of $X^{\mathcal{B}}_{(F,\ell)}(G[V'])$ for graphs $(F,\ell)\in Q\left((H,c)\right)$.

If $|H-W| = 0$, then, for any colouring $\ell\in L_{H,W}$, it is clear that $X^{\mathcal{B}}_{(H-W,c\times\ell)}(G[V']) = 1$. However, if $H-W = \{v\}$, then the number of copies of $(H-W, c\times\ell)$ in $(G[V'],\mathcal{B})$ is simply the order of the set $B_{c(v),\ell(v)}$. Hence, if $|H-W|\le 1$, then $X^{\mathcal{B}}_{(H-W,c\times\ell)}(G[V'])$ depends only on the size of the sets $B_{i,\alpha}\in\mathcal{B}$.

As $I\neq\emptyset$, all sets $W\in(H,c)^I$ are also non-empty. For any component of $H-W$, there is an edge from that component to W, as H is connected. Hence all colourings $\ell\in L_{H,W}$ colour at least one vertex of each component of $H - W$ with a colour other than $\mathbf{0}$. Hence Lemma 5.25 implies that, if $H - W$ is disconnected, $X^{\mathcal{B}}_{(H-W,c\times\ell)}(G[V'])$ is a function of $X^{\mathcal{B}}_{\mathcal{D}_{(H-W,c\times\ell)}}(G[V'])$ for some set $\mathcal{D}_{(H-W,c\times\ell)} \subseteq \mathcal{D}_{|H|-|W|-1,[s],S} \cup \{(K_1,(i,\alpha)) : (i,\alpha) \in [s] \times S\}$.

Note that all graphs in $\mathcal{D}_{|H|-|W|-1,[s],S}$ have order at most $|H| - |W| - 1 \leq |H| - 2$. Hence no such graph is the inverse image under ρ of a graph preceded by (H,c) under $\prec$. However $X^{\mathcal{B}}_{(K_1,(i,\alpha))}(G[V']) = |B_{i,\alpha}|$, and so, if $H - W$ is disconnected, the random variable $X^{\mathcal{B}}_{(H-W,c\times\ell)}(G[V'])$ depends only on the size of the sets $B_{i,\alpha} \in \mathcal{B}$ and the value of $X^{\mathcal{B}}_{(F,\ell)}(G[V'])$ for graphs $(F,\ell) \in Q\left((H,c)\right)$.

Note that, if $H - W$ is connected and $|H - W| \geq 2$, then $(H - W, c \times \ell)$ is an element of $\mathcal{D}_{t,[s],S}$ for any $\ell \in L_{H,W}$. Hence, either the proof is complete or there exists some $\emptyset \neq I \subseteq [s]$, $W \in (H,c)^I$ and $\ell \in L_{H,W}$ such that $(H - W, c \times \ell) \in \mathcal{D}_{t,[s],S} \setminus Q\left((H,c)\right)$. Suppose that this is so. Then $(H - W, c \times \ell) = \rho\left(H', c'\right)$ for some $(H',c') \in \mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$ with $(H,c) \prec (H',c')$. However, $\rho\left(H',c'\right) = \left(H' - w_{(H',c')}, c' \times \ell^{(H',c')}\right)$. As $|W| \geq 1$, this implies that $|H'| - 1 = |H| - |W| \leq |H| - 1$. However, $(H,c) \prec (H',c')$ and so $|H'| \geq |H|$, which implies that $|W| = 1$ and $|H| = |H'|$.

Let $W = \{w\}$, and $i = c(w)$. Then

$$(H - \{w\}, c \times \ell) = \left(H' - w_{(H',c')}, c' \times \ell^{(H',c')}\right).$$

The definition of $L_{H,\{w\}}$ implies that

$$d_H(w) \leq \left|\{u \in H - \{w\} : \left(\ell(u)\right)_i = 1\}\right|.$$

However, by definition

$$\left(\ell^{(H',c')}(u)\right)_i = \begin{cases} 1 & uw_{(H',c')} \in E(H') \text{ and } i = c\left(w_{(H',c')}\right) \\ 0 & \text{otherwise} \end{cases}$$

and hence

$$|\{u \in H - \{w\} : (\ell(u))_i = 1\}| = \begin{cases} d_{H'}(w_{(H',c')}) & i = c\left(w_{(H',c')}\right) \\ 0 & i \neq c\left(w_{(H',c')}\right) \end{cases}$$

and so $d_H(w) \leq d_{H'}(w_{(H',c')})$. But $(H,c) \prec (H',c')$ and $|H| = |H'|$ implies that $d_{H'}(w_{(H',c')}) < d_H(w_{(H,c)})$. This gives $d_H(w) < d_H(w_{(H,c)})$, contradicting the choice of $w_{(H,c)}$. Hence, for all $\emptyset \neq I \subseteq S$, $W \in (H,c)^I$ and $\ell \in L_{H,W}$ with $H - W$ connected, $(H - W, c \times \ell) \in Q\left((H,c)\right)$, completing the proof. $\square$

Lemma 5.28 may now be formally stated and proved.

Lemma 5.28. *Let s and t be fixed integers and $p \in (0,1)$. Let $\lambda \in (0,1)$ and $\mu \in (0,1)$ be real constants and let $\nu = \min\left\{\frac{7\lambda}{8}, \mu\frac{(1-p^s)^2}{4p^s}, \mu\frac{(1-(1-p)^s)^2}{4(1-p)^s}\right\}$. Let $k = k(n)$ be an integer-valued function with $k = o\left(\sqrt{n/\log n}\right)$ and let $G_n \in \mathcal{G}(n,p)$.*

Suppose that, for every $[s]$-labelled partition $\mathcal{A}$ of $[n-s]$ and every $(n-s, S)$-partition $\mathbf{y}$ refining the profile of $\mathcal{A}$ such that $\min(\mathbf{y}) \geq \mu n$, the random graph G_{n-s} is $e^{-\lambda n/k^2}$-refinable modulo k for $\mathcal{D}_{t,[s],S}$, $\mathcal{A}$ and $\mathbf{y}$ with failure probability at most $e^{-\lambda n/k^2}$.

Then, for any $[s]$-labelled partition $\mathcal{P}$ of $[n]$ with $\min(\mathcal{P}) \geq \frac{\mu n}{p^s(1-p)^s} + 1$, the variable $X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}}(G_n)$ is $6e^{-\nu n/k^2}$-uniform modulo k.

Proof of Lemma 5.28. Let $\mathcal{P}$ be an $[s]$-labelled partition of $[n]$ satisfying $\min(\mathcal{P}) \geq \frac{\mu n}{p^s(1-p)^s} + 1$. Recall that G_n is assumed to have vertex set $[n]$. Let the sets U and V' and the partitions $\mathcal{P}'$ and $\mathcal{B}$ be defined according to the standard situation described in Definition 5.26.

For an $(n-s, [s] \times S)$-partition $\mathbf{y}$, let $D_{\mathbf{y}}$ be the event that $\mathcal{B}$ has profile $\mathbf{y}$. Note that this can occur only if $\mathbf{y}$ is a refinement of the profile of $\mathcal{P}'$. If $\mathbf{y}$ is a refinement of the profile of $\mathcal{P}'$, then, given the event $D_{\mathbf{y}}$, the partition $\mathcal{B}$ is chosen uniformly at random from $\mathcal{R}(\mathcal{P}', \mathbf{y})$. For a graph Γ on vertex set V', let Φ_Γ be the event that $G_n[V'] = \Gamma$ and for a graph γ on vertex set U, let Ψ_γ be the event that $G_n[U] = \gamma$.

The lemma's conclusion is that $X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}}(G_n)$ is $6e^{-\nu n/k^2}$-uniform modulo k. As outlined above, the first step in the proof is to prove that, for most choices of $\mathbf{y}$ and the graphs Γ and γ, $X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}}(G_n)$ is approximately uniform modulo k, conditional on $D_{\mathbf{y}}$, Φ_Γ and Ψ_γ.

The properties of $\mathcal{G}(n,p)$ imply that the graphs $G_n[V']$ and $G_n[U]$ and the partition $\mathcal{B}$ are jointly independent. Recall that, given the event $D_{\mathbf{y}}$, the partition $\mathcal{B}$ is chosen uniformly at random from $\mathcal{R}(\mathcal{P}', \mathbf{y})$. Suppose that the graph Γ is $e^{-\lambda n/k^2}$-refinable modulo k for $\mathcal{D}_{t,[s],S}$, $\mathcal{P}'$ and $\mathbf{y}$. Then

$$\max_{\mathbf{b} \in [k]^{\mathcal{D}_{t,[s],S}}} \left| \mathbb{P}\left(X^{\mathcal{B}}_{\mathcal{D}_{t,[s],S}}(G_n[V']) \underset{k}{\equiv} \mathbf{b} \mid \Psi_\gamma, \Phi_\Gamma, D_{\mathbf{y}} \right) - \frac{1}{k^{D(t,[s],S)}} \right| < \frac{e^{-\lambda n/k^2}}{k^{D(t,[s],S)}}. \tag{5.13}$$

Lemma 5.27 implies that there exists an injection ρ from $\mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$ to $\mathcal{D}_{t,s,S}$ and a partial order $\prec$ on the set $\mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$ such that, for all

$(H, c) \in \mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}$,

$$X^{\mathcal{P}}_{(H,c)}(G_n) = X^{\mathcal{P}'}_{(H,c)}(G_n[V']) + X^{\mathcal{B}}_{\rho(H,c)}(G_n[V']) + f_{(H,c)} \tag{5.14}$$

where $f_{(H,c)}$ depends only on the graph $G_n[U]$, the size of the sets $B_{i,\alpha} \in \mathcal{B}$ and the value of $X^{\mathcal{B}}_{(F,\ell)}(G_n[V'])$ for graphs

$$(F, \ell) \in Q\left((H, c)\right) = \mathcal{D}_{t,s,\{0,1\}^s} \setminus \{\rho\left(H', c'\right) : (H, c) \prec (H', c')\}.$$

However, given Ψ_γ, Φ_Γ and $D_{\mathbf{y}}$, the graph $G_n[U]$ and the size of the sets $B_{i,\alpha} \in \mathcal{B}$ are fixed. Hence, in this case $f_{(H,c)}$ depends only on $X^{\mathcal{B}}_{(F,\ell)}(\Gamma)$ for $(F, \ell) \in Q\left((H, c)\right)$, where the random choice is of $\mathcal{B} \in \mathcal{R}\left(\mathcal{P}', \mathbf{y}\right)$.

Given any $\mathbf{c} \in [k]^{\mathcal{D}_{t,[s],S} \setminus \mathrm{Im}(\rho)}$, and a graph Γ on vertex set V', extend to $\mathbf{c} \in [k]^{\mathcal{D}_{t,[s],S}}$ by setting

$$c_{\rho(H,c)} = a_{(H,c)} - X^{\mathcal{P}'}_{(H,c)}(\Gamma) - f_{(H,c)}\left(c_{(F,\ell)} : (F, \ell) \in Q\left((H, c)\right)\right).$$

Note this is well-defined, by solving in increasing order under $\prec$. Let $\Lambda = \Lambda\left(\mathbf{a}\right)$ be the set of $\mathbf{c} \in [k]^{\mathcal{D}_{t,[s],S}}$ obtained in this way. Note that

$$|\Lambda| = k^{\left|\mathcal{D}_{t,[s],S} \setminus \mathrm{Im}(\rho)\right|} = k^{D(t,s,2^s) - C(t+1,s) + C(2,s)}.$$

The definition of Λ and (5.14) imply that, conditional on Ψ_γ, Φ_Γ and $D_{\mathbf{y}}$,

$$X^{\mathcal{P}}_{\mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a} \text{ if, and only if, } X^{\mathcal{B}}_{\mathcal{D}_{t,[s],S}}\left(G_n[V']\right) \underset{k}{\in} \Lambda.$$

Hence

$$\mathbb{P}\left(X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}\setminus\mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a} \mid \Psi_\gamma, \Phi_\Gamma, D_{\mathbf{y}}\right)$$
$$= \mathbb{P}\left(X^{\mathcal{B}}_{\mathcal{D}_{t,[s],S}}(G_n[V']) \underset{k}{\in} \Lambda \mid \Psi_\gamma, \Phi_\Gamma, D_{\mathbf{y}}\right)$$

Thus (5.13) implies that, when Γ is $e^{-\lambda n/k^2}$-refinable modulo k for $\mathcal{D}_{t,[s],S}$, $\mathcal{P}'$ and $\mathbf{y}$,

$$\left|\mathbb{P}\left(X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}\setminus\mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a} \mid \Psi_\gamma, \Phi_\Gamma, D_{\mathbf{y}}\right) - \frac{1}{k^{C(t+1,s)-C(2,s)}}\right| < \frac{e^{-\lambda n/k^2}}{k^{C(t+1,s)-C(2,s)}}. \tag{5.15}$$

The remainder of the proof is technical and involves applying Lemma 5.15 twice to remove the dependence in (5.15) on Φ_Γ and $D_{\mathbf{y}}$. If $\min(\mathbf{y})$ is not too small, then, by assumption, almost all graphs are refinable for $\mathcal{D}_{t,[s],S}$, $\mathcal{P}$ and $\mathbf{y}$. This enables the dependence on Φ_Γ to be removed with a small error term. Using Lemma 5.19 to show that, almost surely, $\min(\mathbf{y})$ is not too small, the dependence on $D_{\mathbf{y}}$ can be removed. However, in (5.15), the variable is $X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}\setminus\mathcal{C}_{2,s}}(G_n)$, while the result is about $X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}}(G_n)$. This can be dealt with when removing the dependence on $D_{\mathbf{y}}$, but this requires that the number of coloured edges in $G_n[V']$ is kept track of when removing the dependence on Φ_Γ.

Fix $\mathbf{r} \in [k]^{\mathcal{C}_{2,s}}$ and let $I_{\mathbf{r}}$ be the set of graphs Γ on vertex set V' such that $X^{\mathcal{P}'}_{\mathcal{C}_{2,s}}(\Gamma) \underset{k}{\equiv} \mathbf{r}$. For a fixed $\mathbf{y}$ and γ, apply Lemma 5.15 with:

- E the event that $X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}\setminus\mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a}$;
- F the event that $X^{\mathcal{P}'}_{\mathcal{C}_{2,s}}(G_n[V']) \underset{k}{\equiv} \mathbf{r}$, which is the union of the events

Φ_Γ for $\Gamma \in I_{\mathbf{r}}$;

- J the set of those $\Gamma \in I_{\mathbf{r}}$ which are $e^{-\lambda n/k^2}$-refinable modulo k for $\mathcal{D}_{t,[s],S}$, $\mathcal{P}$ and $\mathbf{y}$.

By assumption, if $\mathbf{y}$ refines the profile of $\mathcal{P}'$ and $\min(\mathbf{y}) \geq \mu n$, the random graph $G_n[V']$ is $e^{-\lambda n/k^2}$-refinable modulo k for $\mathcal{D}_{t,[s],S}$, $\mathcal{P}'$ and $\mathbf{y}$ with failure probability at most $e^{-\lambda n/k^2}$. Hence, as the graphs $G_n[V']$ and $G_n[U]$ and the partition $\mathcal{B}$ are jointly independent,

$$\mathbb{P}\left(\bigcup_{\Gamma \in I_{\mathbf{r}} \setminus J} \Phi_\Gamma \mid \Psi_\gamma, D_{\mathbf{y}}\right) < e^{-\lambda n/k^2}. \tag{5.16}$$

It is immediate from Lemma 5.22 and $\min(\mathcal{P}') \geq \frac{\mu n}{p^s(1-p)^s}$ that

$$\begin{aligned}
&\left|\mathbb{P}\left(X^{\mathcal{P}'}_{\mathcal{C}_{2,s}}(G[V']) \underset{k}{\equiv} \mathbf{r} \mid \Psi_\gamma, D_{\mathbf{y}}\right) - \frac{1}{k^{C(2,s)}}\right| \\
&\qquad \leq \frac{2C(2,s)}{k^{C(2,s)}} \exp\left\{\frac{-5\mu^2 n^2}{8\left(p(1-p)\right)^{2s-1} k^2}\right\} < \frac{e^{-\nu n/k^2}}{2k^{C(2,s)}}.
\end{aligned} \tag{5.17}$$

Lemma 5.15 and equations (5.15), (5.16) and (5.17) imply that, for all $\mathbf{y}$ refining the profile of $\mathcal{P}'$ with $\min(\mathbf{y}) \geq \mu n$,

$$\begin{aligned}
&\left|\mathbb{P}\left(X^{\mathcal{P}}_{\mathcal{C}_{t+1,s} \setminus \mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a}, X^{\mathcal{P}'}_{\mathcal{C}_{2,s}}(G_n[V']) \underset{k}{\equiv} \mathbf{r} \mid \Psi_\gamma, D_{\mathbf{y}}\right) - \frac{1}{k^{C(t+1,s)}}\right| \\
&\qquad < e^{-\lambda n/k^2} + \frac{e^{-\lambda n/k^2}}{k^{C(t+1,s)-C(2,s)}} \frac{\left(1 + \frac{1}{2}e^{-\nu n/k^2}\right)}{k^{C(2,s)}} + \frac{e^{-\nu n/k^2}}{2k^{C(t+1,s)}} \\
&\qquad < \frac{e^{-\nu n/k^2}}{k^{C(t+1,s)}},
\end{aligned} \tag{5.18}$$

using that $k = o\left(\sqrt{n/\log n}\right)$.

The proof continues by removing the dependence on $\mathbf{y}$, with γ and $\mathbf{r}$ fixed. It needs to be ensured that the variable $X^{\mathcal{P}}_{\mathcal{C}_{s,2}}(G_n)$ takes the right value. Recall that $\mathcal{C}_{s,2} = \{(K_2,(i,j)) : 1 \le i \le j \le s\}$. Also

$$X^{\mathcal{P}}_{(K_2,(i,i))}(G_n) = X^{\mathcal{P}'}_{(K_2,(i,i))}(G_n[V']) + \deg_{P'_i}(v_i),$$

and for $i \neq j$,

$$\begin{aligned} &X^{\mathcal{P}}_{(K_2,(i,j))}(G_n) - X^{\mathcal{P}'}_{(K_2,(i,j))}(G_n[V']) \\ &\qquad = \mathbb{1}\left(v_iv_j \in E(G_n[U])\right) + \deg_{P'_j}(v_i) + \deg_{P'_i}(v_j)\,. \end{aligned}$$

Hence, if $X^{\mathcal{P}'}_{\mathcal{C}_{2,s}}(G_n[V']) \underset{k}{\equiv} \mathbf{r}$ and $G_n[U] = \gamma$, then $X^{\mathcal{P}}_{\mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a}$ if and only if

$$\begin{aligned} \deg_{P'_i}(v_i) &\underset{k}{\equiv} a_{(K_2,(i,i))} - r_{(K_2,(i,i))} \\ \deg_{P'_j}(v_i) + \deg_{P'_i}(v_j) &\underset{k}{\equiv} a_{(K_2,(i,j))} - r_{(K_2,(i,j))} - \mathbb{1}\left(v_iv_j \in E(\gamma)\right). \end{aligned}$$

Note that $\deg_{P'_j}(v_i) = \sum_{\alpha:\alpha_i=1} |B_{j,\alpha}|$. Let $\Upsilon_{\mathbf{r},\gamma}$ be the set of $(n,[s]\times\{0,1\}^s)$-partitions $\mathbf{y}$ refining the profile of $\mathcal{P}'$ such that

$$\begin{aligned} \sum_{\alpha:\alpha_i=1} y_{i,\alpha} &\underset{k}{\equiv} a_{(K_2,(i,i))} - r_{(K_2,(i,i))} \\ \sum_{\alpha:\alpha_i=1} y_{j,\alpha} + \sum_{\alpha:\alpha_j=1} y_{i,\alpha} &\underset{k}{\equiv} a_{(K_2,(i,j))} - r_{(K_2,(i,j))} - \mathbb{1}\left(v_iv_j \in E(\gamma)\right). \end{aligned}$$

Hence, if $X^{\mathcal{P}'}_{\mathcal{C}_{2,s}}(G_n[V']) \underset{k}{\equiv} \mathbf{r}$ and $G_n[U] = \gamma$, then $X^{\mathcal{P}}_{\mathcal{C}_{2,s}}(G_n) \underset{k}{\equiv} \mathbf{a}$ if, and only if, $D_{\mathbf{y}}$ holds for some $\mathbf{y} \in \Upsilon_{\mathbf{r},\gamma}$.

For a fixed γ and $\mathbf{r}$, apply Lemma 5.15 with:

- E the event that both $X^{\mathcal{P}}_{\mathcal{C}_{s,t+1}\setminus\mathcal{C}_{s,2}}(G_n) \underset{k}{\equiv} \mathbf{a}$ and $X^{\mathcal{P}'}_{\mathcal{C}_{s,2}}(G_n[V']) \underset{k}{\equiv} \mathbf{r}$;
- F the union of the events $D_{\mathbf{y}}$ for $\mathbf{y} \in \Upsilon_{\mathbf{r},\gamma}$;
- J the set of those $\mathbf{y} \in \Upsilon_{\mathbf{r},\gamma}$ such that $\min(\mathbf{y}) \geq \mu n$, that is those $\mathbf{y}$ for which (5.18) applies.

As $d_{P'_j}(v_i) \sim \mathrm{Bin}(|P_j|-1, p)$, $k = o\left(\sqrt{\frac{n}{\log n}}\right)$ and $\min(\mathcal{P}) \geq \frac{\mu n}{p^s(1-p)^s} + 1$, Lemma 5.21 implies that

$$\left|\mathbb{P}\left(\bigcup_{\mathbf{y}\in\Upsilon_{\mathbf{r},\gamma}} D_{\mathbf{y}} \mid \Psi_\gamma\right) - \frac{1}{k^{C(2,s)}}\right| < \frac{1}{k^{C(2,s)}} \exp\left\{-\frac{4\mu n}{3k^2\left(p(1-p)\right)^{s-1}}\right\}.$$

However $\nu < \frac{4\mu}{3(p(1-p))^{s-1}}$ and so

$$\left|\mathbb{P}\left(\bigcup_{\mathbf{y}\in\Upsilon_{\mathbf{r},\gamma}} D_{\mathbf{y}} \mid \Psi_\gamma\right) - \frac{1}{k^{C(2,s)}}\right| < \frac{e^{-\nu n/k^2}}{k^{C(2,s)}}. \tag{5.19}$$

Lemma 5.19 will be used to bound $\mathbb{P}\left(\bigcup_{\mathbf{y}\in\Upsilon_{\mathbf{r},\gamma}\setminus J} D_{\mathbf{y}}\right)$. Suppose that $\mathbf{y} \in \Upsilon_{\mathbf{r},\gamma} \setminus J$. Then $y_{j,\alpha} < \mu n$ for some $1 \leq j \leq s$ and $\alpha \in S$. Hence

$$\mathbb{P}\left(\bigcup_{\mathbf{y}\in\Upsilon_{\mathbf{r},\gamma}\setminus J} D_{\mathbf{y}}\right) \leq \sum_{j\in[s]}\sum_{\alpha\in S} \mathbb{P}\left(|B_{j,\alpha}| < \mu n\right). \tag{5.20}$$

For $\alpha \in S$, let $|\alpha| = |\{i : \alpha_i = 1\}|$. As $|B_{j,\alpha}| \sim \mathrm{Bin}(|P_j|-1, p^{|\alpha|}(1-p)^{s-|\alpha|})$, Lemma 5.19 implies that

$$\mathbb{P}\left(|B_{j,\alpha}| < \mu n\right) \leq \exp\left\{\frac{-\left(p^{|\alpha|}(1-p)^{s-|\alpha|}(|P_j|-1) - \mu n\right)^2}{2p^{|\alpha|}(1-p)^{s-|\alpha|}(|P_j|-1)}\right\}.$$

However, $|P_j| \geq \frac{\mu n}{p^s(1-p)^s} + 1$ and so

$$\begin{aligned}\mathbb{P}\left(|B_{j,\alpha}| < \mu n\right) &\leq \exp\left\{\frac{-(|P_j|-1)\left(p^{|\alpha|}(1-p)^{s-|\alpha|} - p^s(1-p)^s\right)^2}{2p^{|\alpha|}(1-p)^{s-|\alpha|}}\right\} \\ &\leq \exp\left\{\frac{-\mu n\left(1-p^{s-|\alpha|}(1-p)^{|\alpha|}\right)^2}{2p^{s-|\alpha|}(1-p)^{|\alpha|}}\right\}.\end{aligned}$$

Hence (5.20) implies that

$$\mathbb{P}\left(\bigcup_{\mathbf{y}\in\Upsilon_{\mathbf{r},\gamma}\setminus J} D_{\mathbf{y}}\right) \leq s2^s \exp\left\{\frac{-\mu n\left(1-p^{s-|\alpha|}(1-p)^{|\alpha|}\right)^2}{2p^{s-|\alpha|}(1-p)^{|\alpha|}}\right\}.$$

However, $\nu \leq \frac{1}{2}\min\left\{\mu\frac{(1-p^s)^2}{2p^s}, \mu\frac{(1-(1-p)^s)^2}{2(1-p)^s}\right\}$, and some simple calculus implies that $\nu \leq \frac{1}{2}\left(\frac{\mu(1-p^{s-|\alpha|}(1-p)^{|\alpha|})^2}{2p^{s-|\alpha|}(1-p)^{|\alpha|}}\right)$ for all $\alpha \in S$. Hence

$$\mathbb{P}\left(\bigcup_{\mathbf{y}\in\Upsilon_{\mathbf{d},\gamma}\setminus J} D_{\mathbf{y}}\right) < \frac{e^{-\nu n/k^2}}{k^{C(2,s)+C(t+1,s)}}. \tag{5.21}$$

Recall that when $X^{\mathcal{P}'}_{\mathcal{C}_{2,s}}(G_n[V']) \underset{k}{\equiv} \mathbf{r}$ and $G_n[U] = \gamma$, the event $X^{\mathcal{P}}_{\mathcal{C}_{2,s}}(G) \underset{k}{\equiv} \mathbf{a}$ is equivalent to the event $\bigcup_{\mathbf{y}\in\Upsilon_{\mathbf{d},\gamma}} D_{\mathbf{y}}$. Applying Lemma 5.15 to (5.18), (5.19) and (5.21) gives

$$\begin{aligned}\zeta_{\gamma,\mathbf{r}} &:= \left|\mathbb{P}\left(X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}}(G_n) \underset{k}{\equiv} \mathbf{a}, X^{\mathcal{P}'}_{\mathcal{C}_{2,s}}(G_n[V']) \underset{k}{\equiv} \mathbf{r} \mid \Psi_\gamma\right) - \frac{1}{k^{C(t+1,s)+C(2,s)}}\right| \\ &\leq \frac{e^{-\nu n/k^2}}{k^{C(t+1,s)+C(2,s)}} + \frac{e^{-\nu n/k^2}}{k^{C(t+1,s)}}\frac{1+e^{-\nu n/k^2}}{k^{C(2,s)}} + \frac{1}{k^{C(t+1,s)}}\frac{e^{-\nu n/k^2}}{k^{C(2,s)}} \\ &< \frac{4e^{-\nu n/k^2}}{k^{C(t+1,s)+C(2,s)}}.\end{aligned}$$

Finally

$$
\begin{aligned}
&\mathbb{P}\left(X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}}(G_n) \underset{k}{\equiv} \mathbf{a}\right) \\
&\quad = \sum_{\mathbf{r}} \sum_{\gamma} \mathbb{P}\left(X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}}(G_n) \underset{k}{\equiv} \mathbf{a}, X^{\mathcal{P}'}_{\mathcal{C}_{2,s}}(G_n[V']) \underset{k}{\equiv} \mathbf{r} \mid \Psi_\gamma\right) \mathbb{P}\left(\Psi_\gamma\right)
\end{aligned}
$$

and so, as there are $k^{C(2,s)}$ choices for $\mathbf{r}$,

$$
\begin{aligned}
\left| \mathbb{P}\left(X^{\mathcal{P}}_{\mathcal{C}_{t+1,s}}(G) \underset{k}{\equiv} \mathbf{a}\right) - \frac{1}{k^{C(t+1,s)}} \right| &\leq \sum_{\mathbf{r}} \sum_{\gamma} \mathbb{P}\left(\Psi_\gamma\right) \zeta_{\gamma,\mathbf{r}} \\
&< \frac{4e^{-\nu n/k^2}}{k^{C(t+1,s)}}
\end{aligned}
$$

as claimed. □

5.5 X_H is uniformly distributed modulo k

Using the results of the previous sections, Theorem 5.6 can now be proved. It is immediate from the following lemma.

Lemma 5.29. *Let $G_n \in \mathcal{G}(n,p)$. For fixed integers T and s, and fixed $0 < \mu \leq 1$, the variable $X^{\mathcal{A}}_{\mathcal{C}_{T,s}}(G_n)$ is $\exp\left\{-\Omega\left(\frac{n}{k^2}\right)\right\}$-uniform modulo k for every $[s]$-labelled partition $\mathcal{A}$ with $\min\left(\mathcal{A}\right) \geq \mu n$.*

Proof. The proof is by induction on $T \geq 2$, for all fixed s and μ. The base case is immediate from Lemma 5.22, with the assumption that $k = o\left(\sqrt{n/\log n}\right)$.

The inductive step is proved by applying Lemmas 5.23 and 5.28. Let $s \in \mathbb{N}$ and $0 < \mu \leq 1$ be fixed. Let $T \geq 3$ and suppose that the inductive hypothesis holds. Then both of the following hold.

- For any $[4^s s]$-labelled partition $\mathcal{A}$ of $[n-s]$ with $\min(\mathcal{A}) \geq \frac{(p(1-p))^{2s}\mu^2}{8}n$, the random variable $X^{\mathcal{A}}_{\mathcal{C}_{T-1,4^s s}}(G_{n-s})$ is $\exp\left\{-\Omega\left(\frac{n}{k^2}\right)\right\}$-uniform modulo k.

- For any $[2^s s]$-labelled partition $\mathcal{A}$ of $[n-s]$ with $\min(\mathcal{A}) \geq \frac{(p(1-p))^{s}\mu}{2}n$, the random variable $X^{\mathcal{A}}_{\mathcal{C}_{T-1,2^s s}}(G_{n-s})$ is $\exp\left\{-\Omega\left(\frac{n}{k^2}\right)\right\}$-uniform modulo k.

Let $\mathcal{P}$ be any $[s]$-labelled partition of $[n-s]$, and let $\mathbf{y}$ satisfying $\min(\mathbf{y}) \geq \frac{(p(1-p))^{s}\mu}{2}n$ be any $(n-s, [s] \times \{0,1\}^s)$-partition refining the profile of $\mathcal{P}$. Lemma 5.23 and Lemma 5.11 imply that, with failure probability at most $\exp\left\{-\Omega\left(\frac{n}{k^2}\right)\right\}$, the random graph G_{n-s} is $\exp\left\{-\Omega\left(\frac{n}{k^2}\right)\right\}$-refinable modulo k for $\mathcal{D}_{T-1,s,\{0,1\}^s}$, $\mathcal{P}$ and $\mathbf{y}$.

Hence Lemma 5.28 implies that for any $[s]$-labelled partition $\mathcal{A}$ of $[n]$ with $\min(\mathcal{A}) \geq \mu n \geq \frac{(p(1-p))^{s}\mu}{2(p(1-p))^s}n + 1$, the random variable $X^{\mathcal{A}}_{\mathcal{C}_{T,s}}(G_n)$ is $\exp\left\{-\Omega\left(\frac{n}{k^2}\right)\right\}$-uniform modulo k. This completes the proof. $\square$

Chapter 6

Conclusion

6.1 Cycles in r-edge-coloured graphs

In Chapter 2 and Chapter 3, results were proved about the existence of cycles of certain lengths in 2-edge-coloured graphs. In both cases, these results could be viewed as generalisations of previous results concerning 1-edge-coloured graphs. Hence it is natural to consider the more general setting of r-edge-coloured graphs.

First consider Theorem 2.3. This states that a 2-edge-coloured graph with sufficiently large order n and minimal degree at least $\frac{3}{4}n$ is either isomorphic to $K_4\left(\frac{1}{4}n\right)$, or contains a monochromatic cycle of length ℓ for all $\ell \in \left[4, \left\lceil \frac{n}{2} \right\rceil\right]$. This is a 2-colour version of Corollary 1.12, which states that graphs with order n and minimal degree at least $\frac{1}{2}n$ are either pancyclic or isomorphic to $K_2\left(\frac{1}{2}n\right)$.

Note that $K_2\left(\frac{1}{2}n\right)$ is the bipartite graph with largest minimal degree, while $K_4\left(\frac{1}{4}n\right)$ has the largest minimal degree among all graphs with a 2-

bipartite 2-edge-colouring. If an r-edge-coloured graph has no odd cycles, then it certainly can not have a cycle of length ℓ for all ℓ in some large range of consecutive integers. Hence, in trying to generalise the above results to r-edge-coloured graphs, the following definition is of use.

Definition 6.1. *An* r-bipartite r-edge-colouring *of a graph G is an r-edge-colouring $E(G) = \bigcup_{\nu=1}^{r} E(G_\nu)$ such that each G_ν is bipartite.*

Suppose that G is a graph which has an r-bipartite r-edge-colouring. This colouring defines a partition of the vertex set of G into at most 2^r independent sets. Hence G has minimal degree at most $\left(1 - \frac{1}{2^r}\right)|G|$. This leads to the following conjecture, which is a generalisation of Theorem 2.3.

Conjecture 6.2. *Let $n \geq 3$, and r be an integer. There exists an integer $R = R(r)$ such that the following holds. Let G be a graph of order n with $\delta(G) \geq \left(1 - \frac{1}{2^r}\right)n$. If $E(G) = \cup_{\nu=1}^{r} E(G_\nu)$ is an r-edge-colouring, then either:*

- *for all $\ell \in \left[R, \left\lceil \frac{1}{2^{r-1}}n \right\rceil\right]$ there is some $1 \leq \nu \leq r$ such that $C_\ell \subseteq G_i$;*
- *n is divisible by 2^r, G is the complete 2^r-partite graph with classes of order $2^{-r}n$, and the colouring is a r-bipartite r-edge-colouring.*

As noted before the conjecture, there are no graphs with minimal degree greater than $\left(1 - \frac{1}{2^r}\right)n$ with an r-bipartite r-edge-colouring. However, the following example shows that there is a graph with order n and minimal degree $\left(1 - \frac{1}{2^r}\right)n$ which has an r-bipartite r-edge-colouring, namely the conjectured extremal graph $K_{2^r}\left(\frac{1}{2^r}n\right)$.

Let $(U_\alpha)_{\alpha \in \{0,1\}^r}$ be 2^r disjoint sets of order t. For all $\alpha \neq \alpha'$ in $\{0,1\}^r$, let $H_{\alpha,\alpha'}$ be the complete bipartite graph with vertex classes U_α and $U_{\alpha'}$. For all $1 \le \nu \le r$, let Z_ν be the set of (α, α') such that, for all $1 \le i \le \nu - 1$, $\alpha_i = \alpha_i$ while $\alpha_\nu = 1$ and $\alpha'_\nu = 0$. Let G_ν be the union of the graphs $H_{\alpha,\alpha'}$ where $(\alpha, \alpha') \in Z_\nu$. Let G'_r be the graph G_r with the addition of all edges contained in U_α, for all $\alpha \in \{0,1\}^r$. For all $1 \le \nu \le r$, the graph G_ν is a subgraph of the complete bipartite graph with vertex classes $\bigcup_{\alpha:\alpha_\nu=0} U_\alpha$ and $\bigcup_{\alpha:\alpha_\nu=1} U_\alpha$. Hence $\bigcup_{\nu=1}^{r} E(G_\nu)$ is an r-bipartite r-edge-colouring of $K_{2^r}(t)$. Similarly $\bigcup_{\nu=1}^{r-1} E(G_\nu) \cup E(G'_r)$ is an r-edge-colouring of $K_{2^r t}$ with no odd cycles of length greater than $2t$.

When n is a multiple of 2^r, setting $t = \frac{1}{2^r}n$ gives an r-bipartite r-edge-colouring of $K_{2^r}\left(\frac{1}{2^r}n\right)$, and an r-edge-colouring of K_n with no monochromatic odd cycles of length greater than $\frac{1}{2^{r-1}}n$. This is why the conjectured interval for which a graph with minimal degree at least $\left(1 - \frac{1}{2^r}\right)n$ other than $K_{2^r}\left(\frac{1}{2^r}n\right)$ has cycles of length ℓ only goes up to $\left\lceil \frac{1}{2^{r-1}}n \right\rceil$.

Note that the case when $r = 1$ in Conjecture 6.2 corresponds to Corollary 1.12 and so is true with $R(1) = 3$. Theorem 2.3 implies that the case $r = 2$ is true for sufficiently large n, with $R(2) = 4$.

In Chapter 3, the monochromatic circumference of an edge-coloured graph was considered. Let r be a positive integer and $0 < c < 1$. Recall that $\Phi_r(c)$ is the supremum of the set of real-valued ϕ such that any r-edge-coloured graph G of sufficiently large order n with minimal degree at least cn has monochromatic circumference at least ϕn.

In Chapter 3, the value of $\Phi_1(c)$ and $\Phi_2(c)$ was given for all $c \in (0,1)$. As noted in that chapter, Theorem 1.14 implies that $\Phi_r(c) \ge \frac{1}{r}c$ for all positive integers r and $c \in (0,1)$. This gives a lower bound on Φ_r for all r, and Lemma 3.15 gave an upper bound using graph constructions. In fact it was shown that, if there are r mutually orthogonal Latin squares of order m, then $\Phi_{r+2}(c) \leq \frac{1}{m}$ for any $c < \frac{((r+2)(m-1)+1)}{m^2}$.

Suppose that, for $r \geq 1$ fixed, there is some constant M such that, for any $m \geq M$, there are r mutually orthogonal Latin squares of order m. Then, for any $m \geq M$ and $\frac{(r+2)m+1}{(m+1)^2} \leq c < \frac{(r+2)(m-1)+1}{m^2}$,

$$\frac{\Phi_{r+2}(c)}{\frac{1}{r+2}c} \leq \frac{r+2}{cm} \leq \frac{(r+2)(m+1)^2}{((r+2)m+1)\,m} = 1 + O\left(\frac{1}{m}\right).$$

Note that, for all $c < \frac{(r+2)(M-1)+1}{M^2}$, there is some $m_c \geq M$ for which

$$\frac{(r+2)m_c+1}{(m_c+1)^2} \leq c < \frac{(r+2)(m_c-1)+1}{m_c^2}$$

and that $m_c \to \infty$ as $c \to 0$. Hence $\Phi_{r+2}(c) \sim \frac{1}{r+2}c$ as $c \to 0$.

That, for any $m \geq 1$ there is a Latin square of order m is trivial. It was proved by Bose, Shrikhande and Parker [11] that, for any $m \geq 7$, there are two mutually orthogonal Latin squares of order m. This implies that $\Phi_3(c) \sim \frac{1}{3}c$ as $c \to 0$ and that $\Phi_4(c) \sim \frac{1}{4}c$ as $c \to 0$.

It is not in general known, for a fixed r, for which values of m there are r mutually orthogonal Latin squares of order m. However, it was shown by Bose [10] that, if m is a prime power and $r-1 \leq m$, then there are r mutually orthogonal Latin squares of order m. By considering prime numbers, this implies that $\limsup_{c\to 0} \frac{\Phi_r(c)}{\frac{1}{r}c} = 1$. This leads to the following conjecture.

Conjecture 6.3. *For any $r \geq 5$, $\Phi_r(c) \sim \frac{1}{r}c$ as $c \to 0$.*

More generally, it may be asked what the value of $\Phi_r(c)$ is for any $r \geq 3$ and $c \in (0,1)$.

6.2 2-factors with three or more components

Chapter 4 considered 2-factors in hamiltonian graphs. In particular, Theorem 4.2 showed that a hamiltonian graph with order n and minimal degree at least $\frac{5-\sqrt{7}}{6}n+2$ contains a 2-factor with two components. Recall that $c(n,k)$ is the smallest integer such that every hamiltonian graph of order n with minimal degree at least $c(n,k)$ has a 2-factor with k components. The method used to prove Theorem 4.2 compares three neighbourhoods and so is likely to be limited to showing that $c(n,2) \leq dn$ for $d \geq \frac{1}{3}$.

For any fixed $k \geq 3$, Theorem 1.21 implies that $c\,(n,k) < \frac{1}{2}$ for sufficiently large n, which is only a slight improvement on the trivial bound implied by Theorem 1.19. One possible way to improve the bound given by Theorem 1.21 is outlined below. This relies on Conjecture 6.4, which states that a hamiltonian graph with large enough minimal degree contains a 2-factor with two components in which one of the components is small. Note that a small adaptation in the proof of Theorem 4.2 implies that, for $\epsilon > 0$, a hamiltonian graph with minimal degree at least $\left(\frac{5-\sqrt{7}}{6}+\epsilon\right)n$ has a 2-factor with two components, in which both components have order at least ϵn.

Conjecture 6.4. *There exists $c < \frac{1}{2}$ such that, for all $\epsilon > 0$, there exists $n_0 = n_0(\epsilon)$ such that the following holds. Any hamiltonian graph with order $n > n_0$ and minimal degree at least cn has a 2-factor with two components in which one of the components has order at most ϵn.*

If true for some c significantly less than $\frac{1}{2}$, this conjecture would imply the following theorem.

Theorem 6.5. *Assume Conjecture 6.4 holds with $c = c_0$. For all $\epsilon > 0$ and for all integers $k \geq 2$, there exists $n_1(\epsilon, k)$ such that the following holds. Any hamiltonian graph with order $n > n_1$ and minimal degree at least $(c_0 + (k-2)\epsilon)n$ has a 2-factor with k components in which all but one of the components have order at most ϵn.*

Proof. Fix $\epsilon > 0$, and let $n_1(\epsilon, k) = \frac{n_0(\epsilon)}{(1-(k-2)\epsilon)}$ where $n_0(\epsilon)$ comes from Conjecture 6.4. Proceed by induction on k. The case $k = 2$ follows immediately from Conjecture 6.4.

Let G be a hamiltonian graph of order $n > n_1$ with minimal degree at least $(c_0 + (k-2)\epsilon)n$. As $n_1(\epsilon, k) \geq n_1(\epsilon, k-1)$, the inductive hypothesis implies that G contains a 2-factor with $k-1$ components in which $k-2$ of the components have order at most ϵn. Let $C_1, \ldots, C_{k-1}$ be the cycles of this 2-factor, with $|C_i| \leq \epsilon n$ for all $1 \leq i \leq k-2$.

Let $G_1 = G[V(C_{k-1})]$ and $G_2 = G[V(G) - V(C_{k-1})]$. Then

$$\delta(G_1) \geq (c_0 + (k-2)\epsilon)n - |G_2| \geq c_0 n \geq c_0 |G_1|.$$

Hence, if Conjecture 6.4 is true, G_1 has a 2-factor with two components C'_{k-1} and C'_k, where $|C'_{k-1}| \leq \epsilon n$. Thus $C_1 \cup \cdots C_{k-2} \cup C'_{k-1} \cup C'_k$ is the required 2-factor of G with k components. $\square$

If Conjecture 6.4 holds, then Theorem 6.5 immediately implies the following upper bound on $c(n, k)$.

Corollary 6.6. *Assume Conjecture 6.4 holds with $c = c_0$. For all $\epsilon > 0$ and* *all integers $k \geq 2$, there exists $n_1(\epsilon, k)$ such that, for all $n \geq n_1(\epsilon, k)$,*

$$c(n,k) \leq (c_0 + (k-2)\epsilon)n.$$

6.3 The distribution of X_H modulo k for a disconnected graph H

In Chapter 5, it was shown that, for constant p and $k = o\left(\sqrt{n/\log n}\right)$, $X_{\mathcal{C}}(G_n)$ is $\exp\{-\Omega(n/k^2)\}$-uniform modulo k, for any finite set $\mathcal{C}$ of connected graphs. There are several questions which may be posed, based on this result.

Firstly, what happens for larger values of k? Note that, for any finite set $\mathcal{C}$ of graphs, $X_{\mathcal{C}}(G_n)$ is $O\left(k^{|\mathcal{C}|}\right)$-uniform modulo k for any k. Hence, if $k \geq n^2$, then $X_{\mathcal{C}}(G_n)$ is trivially $\exp\{-\Omega(n/k^2)\}$-uniform modulo k. Lemma 5.22 implies that, if $k = o\left(\frac{n}{\sqrt{\log n}}\right)$, then $X_{K_2}(G_n)$ is $\exp\{-\Omega(n^2/k^2)\}$-uniform modulo k. This leads to the following conjecture.

Conjecture 6.7. *Let $p \in (0,1)$ be a fixed real number and let $G_n \in \mathcal{G}(n,p)$. Let $\mathcal{C}$ be a finite set of connected graphs. Let k be an integer-valued function of n with $k = o\left(\sqrt{n/\log n}\right)$. Then $X_{\mathcal{C}}(G_n)$ is $\exp\{-\Omega(n^2/k^2)\}$-uniform modulo k.*

Another possible extension would be to consider what happens if $p \to 0$ as $n \to \infty$. Note that, in this case, there may be graphs H such that X_H is almost surely equal to zero.

Note that Lemma 5.25 implies that, for any disconnected graph H, the value of X_H modulo k can be determined from the value of $X_{\mathcal{C}_{|H|}}$ modulo k.

Suppose that H is the disjoint union of a single edge and a path of length three. The following equation is obtained by considering how an edge and a path of length three may intersect. Let H_0 and H_1 be the graphs pictured in Figure 6.1. Then

$$\begin{aligned} &X_{P_3}(G)X_{P_1}(G) \\ &= X_{P_3 \cup P_1}(G) + 2X_{P_4}(G) + 2X_{H_1}(G) + 4X_{C_4}(G) + 2X_{H_0}(G) + 3X_{P_3}(G). \end{aligned} \tag{6.1}$$

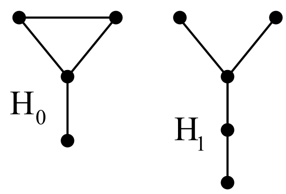


Figure 6.1: The graphs H_0 and H_1.

In particular,

$$X_{P_3 \cup P_1} \underset{2}{\equiv} X_{P_3}\left(1 + X_{P_1}\right),$$

and hence Theorem 5.3 implies that

$$\begin{aligned} \left|\mathbb{P}\left(X_{P_3 \cup P_1} \underset{2}{\equiv} 1\right) - \frac{1}{4}\right| &= \left|\mathbb{P}\left(X_{P_3} \underset{2}{\equiv} 1 \text{ and } X_{P_1} \underset{2}{\equiv} 0\right) - \frac{1}{4}\right| \\ &< \exp\left\{-\Omega\left(n\right)\right\}. \end{aligned}$$

Hence $X_{P_3 \cup P_1}$ is not approximately uniform modulo 2.

This phenomenon appears to be rare, and it is conjectured that, in fact, it is rare.

Conjecture 6.8. *There is a finite set $\mathcal{H}$ of graphs such that the following holds. Let k be a constant, and let ϵ be a constant, which is significantly less than $\frac{1}{k}$. Then, for any graph $H \notin \mathcal{H}$ with no isolated vertices, the variable X_H is ϵ-uniform modulo k.*

Appendix A

Proof of the r-coloured Regularity Lemma

In this appendix, a proof is given of the degree form of the r-edge-coloured version of the Regularity Lemma (Theorem 2.9). The first stage will be to prove Theorem A.3, which is a general form of the r-edge-coloured version of the Regularity Lemma. The following two definitions are properties of partitions, which allow Theorem A.3 to be stated simply.

Definition A.1 (Equitable partition)**.** *Let $\mathcal{P} = (V_i)_{i=0}^{k}$ be a partition of a set V. Then $\mathcal{P}$ is an* equitable partition *with* exceptional class *V_0 if, for all $1 \leq i \leq k$, the sets V_i have the same order.*

Definition A.2 (Regular partition)**.** *Let G be a graph of order n with an r-edge-colouring $E(G) = \bigcup_{\nu=1}^{r} E(G_\nu)$. For $\epsilon > 0$, an equitable partition $\mathcal{P}$ is* ϵ-regular *if $|V_0| \leq \epsilon n$, and, with the exception of at most ϵk^2 pairs $1 \leq i < j \leq k$, the graph $G_\nu [V_i, V_j]$ is ϵ-regular for all $1 \leq \nu \leq r$.*

Theorem A.3 (The r-coloured Regularity Lemma)**.** *For any* $0 < \epsilon < \frac{1}{2}$ *and integers* m *and* r*, there exists an* M *such that, if* G *is a graph of order* $n \geq m$ *with an* r*-edge-colouring* $E(G) = \bigcup_{\nu=1}^{r} E(G_\nu)$*, then there is an* ϵ*-regular partition* $(V_i)_{i=0}^{k}$ *of* G *with* $m \leq k \leq M$*.*

The proof of Theorem A.3 will follow that of the (non-coloured) regularity lemma given by Bollobás [5, p.124-129]. The following two lemmas stated in [5] are required.

Lemma A.4. *Let* $0 < \gamma, \delta < 1$*. Suppose that* X *and* Y *are disjoint sets of vertices of a graph* G*, and* $X^* \subseteq X$ *and* $Y^* \subseteq Y$ *are such that* $|X^*| \geq (1-\gamma)|X| > 0$ *and* $|Y^*| \geq (1-\delta)|Y| > 0$*. Then*

$$|d(X^*, Y^*) - d(X, Y)| \leq \gamma + \delta.$$

Lemma A.5. *Let* $1 \leq t < s$ *be integers and, for* $1 \leq i \leq s$*, let* d_i *be a real number. Set* $D = \frac{1}{s}\sum_{i=1}^{s} d_i$ *and* $d = \frac{1}{t}\sum_{i=1}^{t} d_i$*. If* $t \geq \gamma s$ *and* $|D - d| \geq \delta > 0$*, then*

$$\frac{1}{s}\sum_{i=1}^{s} d_i^2 \geq D^2 + \gamma\delta^2.$$

Of crucial use in the proof will be the index of an equitable partition of the vertex set of a graph.

Definition A.6. *Let* G *be a graph of order* n *with an* r*-edge-colouring* $E(G) = \bigcup_{\nu=1}^{r} E(G_\nu)$*. For an equitable partition* $\mathcal{P} = (V_i)_{i=0}^{k}$ *of* $V(G)$*, let*

$$\operatorname{ind}(\mathcal{P}) = \frac{1}{k^2}\sum_{\nu=1}^{r}\sum_{i=1}^{k}\sum_{j=i+1}^{k} d_\nu^2(V_i, V_j).$$

The following lemma shows that for an equitable partition $\mathcal{P}$ which is not ϵ-regular, there is an equitable partition $\mathcal{Q}$ with a significantly higher index than $\mathcal{P}$, such that the exceptional set of $\mathcal{Q}$ is not too much bigger than that of $\mathcal{P}$. This, together with the fact that $\operatorname{ind}(\mathcal{P}) < \frac{1}{2}$ for any equitable partition $\mathcal{P}$, will prove Theorem A.3.

Lemma A.7. *Let G be a graph of order n with an r-edge-colouring $E(G) = \bigcup_{\nu=1}^{r} E(G_\nu)$. Let $\mathcal{P} = (V_i)_{i=0}^{k}$ be an equitable partition of $V(G)$, with $|V_i| = s \geq 2^{3k+1}$ for all $1 \leq i \leq k$. Suppose that $\mathcal{P}$ is not ϵ-regular, for some $0 < \epsilon < \frac{1}{2}$ with $2^{-k} \leq \frac{\epsilon^5}{8}$. Then, setting $\ell = k\left(4^k - 2^{k-1}\right)$, there is an equitable partition $\mathcal{Q} = (U_i)_{i=0}^{\ell}$ such that*

- *$V_0 \subseteq U_0$ and $|U_0| \leq |V_0| + \frac{1}{2^k} n$;*
- *$\operatorname{ind}(Q) \geq \operatorname{ind}(P) + \frac{\epsilon^5}{4}$.*

Proof. Suppose that (V_i, V_j) is not ϵ-regular for G_ν for some $1 \leq \nu \leq r$, and let ν_{ij} be the minimal such ν. Let $V_{ij} \subseteq V_i$ and $V_{ji} \subseteq V_j$ be sets such that $|V_{i,j}| \geq \epsilon |V_i|$, $|V_{ji}| \geq \epsilon |V_j|$ and

$$\left| d_{\nu_{ij}}(V_{ij}, V_{ji}) - d_{\nu_{ij}}(V_i, V_j) \right| \geq \epsilon. \tag{A.1}$$

If a pair (V_i, V_j) is ϵ-regular for G_ν for all $1 \leq \nu \leq r$, let $V_{ij} = V_{ji} = \emptyset$.

For $1 \leq i \leq k$, and any $I \subseteq [k] \setminus \{i\}$, let

$$W_{i,I} = \bigcap_{j \in I} V_{ij} \cap \bigcap_{j \notin I} V_i \setminus V_{ij},$$

so that the $W_{i,I}$ are the *atoms* of V_i with respect to the sets V_{ij}. Let $H =$

$4^k - 2^{k-1}$, and $d = \lfloor s/4^k \rfloor \geq 2^{k+1}$, so that $4^k d \leq s \leq 4^k(d+1)$.

For all $1 \leq i \leq k$, partition each atom of V_i into as many as possible subsets of order d and a small set of leftover vertices. In doing so, each atom has at most $d-1$ leftover vertices, and there are at most 2^{k-1} non-empty atoms. As

$$Hd + 2^{k-1}(d-1) = d4^k - 2^{k-1} < s = |V_i|,$$

there are at least H such subsets of V_i, which are now labelled as D_{ih}.

Let $\overline{V_i} = \bigcup_h D_{ih}$ and $\overline{V_{ij}} = \bigcup_{D_{ih} \subseteq V_{ij}} D_{ih}$. It will be shown that, for all i and j, the sets $V_i \setminus \overline{V_i}$ and $V_{ij} \setminus \overline{V_{ij}}$ are small. Firstly,

$$\frac{\left|V_i \setminus \overline{V}_i\right|}{|V_i|} \leq \frac{s - Hd}{4^k d} \leq \frac{4^k + 2^{k-1}d}{4^k d} = \frac{1}{d} + \frac{1}{2^{k+1}} \leq 2^{-k} \leq \frac{\epsilon^5}{8}. \tag{A.2}$$

Clearly, if (V_i, V_j) is ϵ-regular for G_ν for all $1 \leq \nu \leq r$, then $V_{ij} = \overline{V_{ij}} = \emptyset$. Suppose that (V_i, V_j) is not ϵ-regular for $G_{\nu_{ij}}$. Then, by (A.2),

$$\frac{|V_{ij} \setminus \overline{V_{ij}}|}{|V_{ij}|} \leq \frac{\left|V_i \setminus \overline{V}_i\right|}{\epsilon |V_i|} \leq \frac{\epsilon^4}{8} \tag{A.3}$$

and

$$|\overline{V_{ij}}| \geq |V_{ij}| - |V_i \setminus \overline{V_i}| \geq \left(\epsilon - \epsilon^5/8\right)|V_i| \geq \left(1 - 2^{-7}\right)\epsilon|\overline{V_i}|. \tag{A.4}$$

Using the bounds (A.2) and (A.3), Lemma A.4 now implies that, for all $1 \leq i < j \leq k$ and $1 \leq \nu \leq r$,

$$\left|d_\nu\left(\overline{V_i}, \overline{V_j}\right) - d_\nu\left(V_i, V_j\right)\right| \leq \frac{\epsilon^5}{4}, \tag{A.5}$$

and

$$\left|d_\nu\left(\overline{V_{ij}},\overline{V_{ji}}\right)-d_\nu\left(V_{ij},V_{ji}\right)\right|\leq\frac{\epsilon^4}{4}. \tag{A.6}$$

Also, applying Lemma A.4 to the uncoloured graph G gives

$$\left|d\left(\overline{V_i},\overline{V_j}\right)-d\left(V_i,V_j\right)\right|\leq\frac{\epsilon^5}{4}.$$

This implies that

$$\begin{aligned}
&\sum_{\nu=1}^{r} d_\nu^2\left(V_i,V_j\right)-\sum_{\nu=1}^{r} d_\nu^2\left(\overline{V_i},\overline{V_j}\right)\\
&\quad=\sum_{\nu=1}^{r}\left(d_\nu\left(V_i,V_j\right)+d_\nu\left(\overline{V_i},\overline{V_j}\right)\right)\left(d_\nu\left(V_i,V_j\right)-d_\nu\left(\overline{V_i},\overline{V_j}\right)\right)\\
&\quad\leq 2\left(d\left(V_i,V_j\right)-d\left(\overline{V_i},\overline{V_j}\right)\right)\\
&\quad\leq\frac{\epsilon^5}{2}. && \text{(A.7)}
\end{aligned}$$

For any pair (V_i,V_j) and colour ν the Cauchy-Schwarz inequality implies that

$$\begin{aligned}
\frac{1}{H^2}\sum_{u=1}^{H}\sum_{v=1}^{H} d_\nu^2(D_{iu},D_{jv}) &\geq\left(\frac{1}{H^2}\sum_{u=1}^{H}\sum_{v=1}^{H} d_\nu(D_{iu},D_{jv})\right)^2\\
&=d_\nu^2(\overline{V_i},\overline{V_j}). && \text{(A.8)}
\end{aligned}$$

Using Lemma A.5, this bound can be improved for pairs (V_i,V_j) which are not ϵ-regular for $G_{\nu_{i,j}}$. For such a pair (V_i,V_j) the equations (A.1), (A.5) and

(A.6) imply that

$$\left|d_\nu\left(\overline{V_i},\overline{V_j}\right)-d_\nu\left(\overline{V_{ij}},\overline{V_{ji}}\right)\right|\geq\epsilon-\frac{\epsilon^4}{4}-\frac{\epsilon^5}{4}>\frac{15}{16}\epsilon. \tag{A.9}$$

Hence, Lemma A.5 and the bound (A.4) imply that

$$\begin{aligned}\frac{1}{H^2}\sum_{u=1}^{H}\sum_{v=1}^{H}d_\nu^2(D_{iu},D_{jv})&\geq d_\nu^2(\overline{V_i},\overline{V_j})+\frac{|\overline{V_{ij}}||\overline{V_{ji}}|}{|\overline{V_i}||\overline{V_j}|}\left(\frac{15}{16}\epsilon\right)^2\\&\geq d_\nu^2(\overline{V_i},\overline{V_j})+\frac{3}{4}\epsilon^4.\end{aligned} \tag{A.10}$$

Relabel the sets D_{ih} as $(U_i)_{i=1}^{\ell}$ and let $U_0=V_0\cup\bigcup_{i=1}^{k}V_i\setminus\overline{V_i}$. It remains to show that the partition $\mathcal{Q}=(U_i)_{i=0}^{\ell}$ has the desired properties. Firstly

$$|U_0|-|V_0|=\sum_{i=1}^{k}\left|V_i\setminus\overline{V_i}\right|\leq\sum_{i=1}^{k}\frac{1}{2^k}|V_i|\leq\frac{n}{2^k},$$

where the middle inequality follows from (A.2).

The proof is complete if the index of $\mathcal{Q}$ can be shown to be much larger than the index of $\mathcal{P}$. Firstly,

$$\begin{aligned}\operatorname{ind}(\mathcal{Q})&=\frac{1}{\ell^2}\sum_{\nu=1}^{r}\sum_{1\leq i<j\leq\ell}d_\nu^2(U_i,U_j)\\&\geq\frac{1}{k^2}\sum_{\nu=1}^{r}\sum_{1\leq i<j\leq k}\left(\frac{1}{H^2}\sum_{u=1}^{H}\sum_{v=1}^{H}d_\nu^2(D_{iu},D_{jv})\right)\end{aligned}$$

Hence, as at least ϵk^2 pairs are not regular for some colour, (A.10) and (A.8) imply that

$$\begin{aligned}
\operatorname{ind}(\mathcal{Q}) &\geq \frac{1}{k^2}\left(\sum_{\nu=1}^{r}\sum_{1\leq i<j\leq k} d_\nu^2(\overline{V_i},\overline{V_j}) + \epsilon k^2 \frac{3}{4}\epsilon^4\right)\\
&= \frac{1}{k^2}\left(\sum_{1\leq i<j\leq k}\sum_{\nu=1}^{r} d_\nu^2(\overline{V_i},\overline{V_j})\right) + \frac{3}{4}\epsilon^5.
\end{aligned}$$

Finally, (A.7) implies that

$$\begin{aligned}
\operatorname{ind}(\mathcal{Q}) &\geq \frac{1}{k^2}\left(\sum_{1\leq i<j\leq k}\left(\sum_{\nu=1}^{r} d_\nu^2(V_i,V_j) - \frac{\epsilon^5}{2}\right)\right) + \frac{3}{4}\epsilon^5\\
&= \operatorname{ind}(\mathcal{P}) + \frac{1}{4}\epsilon^5
\end{aligned}$$

as claimed. □

Lemma A.7 can now be used to deduce Theorem A.3.

Proof of Theorem A.3. Set $t = \lfloor 2\epsilon^{-5}\rfloor$. Define $k_0, k_1, \ldots, k_t$ by setting k_0 to be the minimal integer satisfying $k_0 \geq m$ and $2^{-k_0} \leq \epsilon^5/8$, and, for all $1 \leq i \leq t$, setting $k_{i+1} = k_i\left(4^{k_i} - 2^{k_i - 1}\right)$. The proof will show that $M = k^t$ will do.

Let G be an r-edge-coloured graph of order $n \geq m$. If $n \leq M$, partitioning G into n singletons gives a 0-regular partition. Hence it can be assumed that $n > M$.

Let $\mathcal{P}_0 = \left(V_i^{(0)}\right)_{i=0}^{k_0}$ be a partition of $V(G)$ chosen so that $\left|V_i^{(0)}\right| = \lfloor n/k_0\rfloor$ for all $1 \leq i \leq k_0$ and hence $\left|V_0^{(0)}\right| \leq k_0 < \frac{\epsilon}{2}n$. If $\mathcal{P}_0$ is ϵ-regular, then the proof is complete. If not, let $\mathcal{P}_1 = \left(V_0^{(1)}\right)_{i=0}^{k_1}$ be the partition obtained from Lemma A.7, with $\left|V_0^{(1)}\right| \leq \left|V_0^{(0)}\right| + \frac{1}{2^{k_0}}n$ and $\operatorname{ind}(\mathcal{P}_1) \geq \operatorname{ind}(\mathcal{P}_0) + \frac{\epsilon^5}{4}$. Repeat this process until an ϵ-regular partition is obtained.

Note that there is some ϵ-regular partition $\mathcal{P}_j$ for $0 \leq j \leq t$, else

$$\operatorname{ind}(\mathcal{P}_{t+1}) \geq \operatorname{ind}(\mathcal{P}_0) + (t+1)\frac{\epsilon^5}{4} \geq \frac{1}{2},$$

a contradiction as $\operatorname{ind}(\mathcal{P}) \leq \frac{1}{2}$ for any partition $\mathcal{P}$. Finally

$$\left|V_0^{(j)}\right| \leq \left|V_0^{(0)}\right| + n\sum_{\ell=1}^{j} 2^{-k_i} < n\left(\frac{\epsilon}{2} + 2^{1-k_0}\right) < \epsilon n$$

as required. $\square$

The degree form of the 2-edge-coloured Regularity Lemma can now be deduced.

Proof of Theorem 2.9. Suppose that $\epsilon > 0$, and integers r and m are given. Let $\epsilon' = \min\left\{\frac{1}{18}\epsilon^2, \frac{1}{6r}\epsilon\right\}$ and $m' = \left\lceil \frac{m}{1-\frac{1}{3}\epsilon} \right\rceil$. Let G be an r-edge-coloured graph with order $n \geq m'$. Apply Theorem A.3 to G to obtain an ϵ'-regular partition $\mathcal{Q} = (W_i)_0^l$, with $m' \leq l \leq M$. Let $t' := |W_1|$, so that $|W_i| = t'$ for all $1 \leq i \leq \ell$. Also $t'l \leq n$.

For $1 \leq i \leq l$, let J_i be the set of $j \neq i$ such that, for some $1 \leq \nu \leq r$, the graph $G_\nu(W_i, W_j)$ is not ϵ'-regular. Let I be the set of $1 \leq i \leq l$ such that $|J_i| \geq \frac{1}{3}\epsilon l$. By the definition of a regular partition, $\sum_{i=1}^{l} |J_i| \leq 2\epsilon' l^2$, and so $|I| \leq \frac{1}{3}\epsilon l$. Relabelling if necessary, it can be assumed that $I = [k+1, l]$, for some integer $k \geq \left(1 - \frac{1}{3}\epsilon\right) l \geq m$.

For $1 \leq i \leq k$ and $1 \leq \nu \leq r$, let $K_{i,\nu}$ be the set of $j \notin J_i$ such that $d_\nu(W_i, W_j) \leq \eta + \epsilon'$ and let $\mathcal{W}_{i,\nu} = \bigcup_{j \in K_{i,\nu}} W_j$. Let $U_{i,\nu}$ be the set of $v \in W_i$ such that there at least $(\eta + 2\epsilon')|\mathcal{W}_{i,\nu}|$ edges from v to $\mathcal{W}_{i,\nu}$ in

G_ν. Then $d_\nu(U_{i,\nu}, \mathcal{W}_{i,\nu}) \geq \eta + 2\epsilon'$, and so there is some $j \in K_{i,\nu}$ such that $d_\nu(U_{i,\nu}, W_j) \geq \eta + 2\epsilon'$. As $G_\nu(W_i, W_j)$ is ϵ'-regular and $d_\nu(W_i, W_j) \leq \eta + \epsilon'$, this implies that $|U_{i,\nu}| \leq \epsilon'|W_i| = \epsilon' t'$.

For $1 \leq i \leq k$, let $U_i = \bigcup_{1 \leq \nu \leq r} U_{i,\nu}$ and $V_i = W_i - U_i$. By adding vertices to U_i as necessary, ensure that all sets U_i have the same order $t'' \leq r\epsilon' t'$. Also let

$$V_0 = W_0 \cup \bigcup_{i \in I} W_i \cup \bigcup_{i \notin I} U_i.$$

Let G' be the subgraph of G obtained by removing the following sets of edges:

- edges within V_i for any $1 \leq i \leq k$;
- edges between V_i and V_j for $1 \leq i < j \leq k$ such that, for some $1 \leq \nu \leq r$, the graph $G_\nu(W_i, W_j)$ is not ϵ'-regular.
- edges coloured ν between V_i and V_j for $1 \leq i < j \leq k$ such that $j \in K_{i,\nu}$.

The proof is completed by showing that the graph G' and the partition $V_0 \cup \bigcup_{i \notin I} V_i$ satisfy the conditions of the Theorem.

Note that there are $k = l - |I|$ non-exceptional clusters, and that $m \leq k \leq M$. Also,

$$\begin{aligned} |V_0| &= |V_0'| + \sum_{i \in I} |W_i| + \sum_{i=1}^{k} |U_i| \\ &< \epsilon' n + \frac{1}{3}\epsilon l t' + l r \epsilon' t' \\ &< \epsilon n \end{aligned}$$

as required. By construction, all clusters V_i have the same order $t := t' - t''$, and there are no edges within any V_i.

Suppose that $1 \leq i \leq k$, and $v \in V_i$. The edges incident to v which are deleted to form G' are those from v to $\bigcup_{j \in \{i\} \cup J_i} V_j$, and edges coloured ν from v to $\bigcup_{j \in K_{i,\nu}} V_j \subseteq \mathcal{W}_{i,\nu}$. However, $\deg_{V_j}(v) \leq t'$, for all $1 \leq j \leq k$. Also $i \notin I$ implies that $|J_i| < \frac{1}{3}\epsilon l$. Finally, as $v \notin U_{i,\nu}$, there at most $(\eta + 2\epsilon') |\mathcal{W}_{i,\nu}|$ edges from v to $\mathcal{W}_{i,\nu}$ in G_ν. Hence

$$\begin{aligned}
\deg_{G'}(v) &\geq \deg_G(v) - \left(1 + \frac{1}{3}\epsilon l\right) t' - \sum_{\nu=1}^{r} (\eta + 2\epsilon') |\mathcal{W}_{i,\nu}| \\
&\geq \deg_G(v) - \frac{2}{3}\epsilon n - r(\eta + 2\epsilon')n \\
&\geq \deg_G(v) - (r\eta + \epsilon)n,
\end{aligned}$$

as required.

The proof is complete if, for all $1 \leq i < j \leq k$ and $1 \leq \nu \leq r$, the graph $G'_\nu(V_i, V_j)$ is ϵ-regular with a density either zero or at least η. If there are no edges in G'_ν between V_i and V_j then $G'_\nu(V_i, V_j)$ is ϵ-regular with density zero. Suppose now that there are edges in G'_ν between V_i and V_j. Then $G_\nu(W_i, W_j)$ is ϵ'-regular, with density at least $\eta + \epsilon'$ and $G_\nu(W_i, W_j) = G'_\nu(W_i, W_j)$. As $|V_i| \geq (1 - \epsilon) |W_i|$ and $|V_j| \geq (1 - \epsilon) |W_j|$, Lemma 2.8 implies that $G'_\nu(V_i, V_j)$ is ϵ-regular with density at least η, completing the proof. $\square$

Bibliography

[1] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley-Interscience, 2nd ed., 2000.

[2] C. Berge, *Sur le couplage maximum d'un graphe*, C.R. Acad. Sci. Paris, 247 (1958), pp. 258–259. (in French).

[3] B. Bollobás, *Extremal Graph Theory*, Academic Press Inc., 1978.

[4] B. Bollobás, *Combinatorics (Swansea, 1981) London Math. Soc. Lecture Note Ser. 52*, CUP, 1981, ch. Random graphs, pp. 80–102.

[5] B. Bollobás, *Modern Graph Theory*, Springer, 1998.

[6] ——, *Random Graphs*, CUP, 2nd ed., 2001.

[7] B. Bollobás and R. Häggkvist, *A tribute to Paul Erdős*, CUP, 1990, ch. The circumference of a graph with a given minimal degree., pp. 97–140.

[8] J. Bondy, *Pancyclic graphs I*, J. Comb. Theory, Ser. B, 11 (1971), pp. 80–84.

[9] J. Bondy and M. Simonovits, *Cycles of even lengths in graphs*, J. Comb. Theory, Ser. B, 16 (1974), pp. 97–105.

[10] R. Bose, *On the application of the properties of Galois fields to the problem of construction of hyper-Græco-Latin squares*, Sankhyā, 3 (1938), pp. 323–338.

[11] R. Bose, S. Shrikhande, and E. Parker, *Further results on the construction of sets of mutually orthogonal Latin squares and the falsity of Euler's conjecture*, Canada. J. Math, 12 (1960), pp. 189–203.

[12] S. Brandt, G. Chen, R. Faudree, R. Gould, and L. Lesniak, *Degree conditions for 2-factors*, J. Graph Theory, 24 (1997), pp. 165–173.

[13] Chvátal, *The tail of the hypergeometric distribution*, Disc. Math., 25 (1979), pp. 285–287.

[14] V. Chvátal, *On Hamilton's ideals*, J. Comb. Theory, Ser. B, 12 (1972), pp. 163–168.

[15] R. Diestel, *Graph Theory*, Springer, 2006.

[16] G. Dirac, *Some theorems on abstract graphs*, Proc. London Math. Soc., 2 (1952), pp. 69–81.

[17] Y. Egawa and T. Miyamoto, *The longest cycles in a graph G with minimum degree at least $|G|/k$*, J. Comb. Theory, Ser. B, 46 (1989), pp. 356–362.

[18] P. Erdős, *Some remarks on the theory of graphs*, Bull. Am. Math. Soc., 53 (1947), pp. 292–294.

[19] P. Erdős and T. Gallai, *On maximal paths and circuits of graphs*, Acta Math. Acad. Sci. Hungar., 10 (1959), pp. 337–356.

[20] P. Erdős and A. Rényi, *On random graphs I*, Publ. Math. Debrecen, 6 (1959), pp. 290–297.

[21] L. Euler, *Recherches sur une nouvelie espéce de quarres magiques*, Verh. Zeeuwsch Genootsch. Wetensch. Vlissingen, 9 (1782), pp. 85–239. (in French).

[22] R. Faudree, R. Gould, M. Jacobson, L. Lesniak, and A. Saito, *A note on 2-factors with two components*, Disc. Math., 300 (2005), pp. 218–224.

[23] R. Faudree, L. Lesniak, and I. Schiermeyer, *On the circumference of a graph and its complement*, Disc. Math., 309 (2009), pp. 5891–5893.

[24] R. Faudree and R. Schelp, *All ramsey numbers for cycles in graphs*, Disc. Math., 8 (1974), pp. 313–329.

[25] E. Gilbert, *Random graphs*, Ann. Math. Stat, 30 (1959), pp. 1141–1144.

[26] C. Graves and W. Hamilton, *On a theorem related to the binomial coefficient*, Proc. of the Royal Irish Acad., 9 (1865), pp. 297–302.

[27] E. Györi, V. Nikiforov, and R. Schelp, *Nearly bipartite graphs*, Disc. Math., 272 (2003), pp. 187–196.

[28] P. Hall, *On representatives of subsets*, J. London Math. Soc., 10 (1935), pp. 26–30.

[29] Y. Ionin and M. Shrikhande, *Combinatorics of Symmetric Designs*, Cambridge, 2006.

[30] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley-Interscience, 2000.

[31] M. Karoński and A. Ruciński, *On the number of strictly balanced subgraphs of a random graph*, in Graph Theory, Proceedings (Lagów, 1981) Lecture Notes in Math. 1018, M. Borowiecki, J. Kennedy, and M. Sysło, eds., Springer, 1983, pp. 79–83.

[32] P. Kolaitis and S. Kopparty, *Random graphs and the parity quantifier.* arXiv:0904.2436v1 [math.CO].

[33] J. Komlós, G. Sárközy, and E. Szemerédi, *Blow-up Lemma*, Combinatorica, 17 (1997), pp. 109–123.

[34] J. Komlós and M. Simonovits, *Szemerdi's regularity lemma and its applications in graph theory*, in Combinatorics, Paul Erdős is Eighty, Vol. 2 Bolyai Society Mathematical Studies 2, D. Miklós, V. Sós, and M. Sysło, eds., János Bolyai Math. Soc., Budapest, 1996, pp. 295–352.

[35] J. Komlós and E. Szemerédi, *Limit distributions for the existence of Hamilton cycles in a random graph*, Disc. Math., 43 (1983), pp. 55–63.

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

[37] C. Laywine and G. Mullen, *Discrete mathematics using Latin squares.*, John Wiley & Sons, 1998.

[38] H. Li, V. Nikiforov, and R. Schelp, *A new type of Ramsey-Turán problems*, Disc. Math., 310 (2010), pp. 3579–3583.

[39] M. Loebl, J. Matoušek, and O. Pangrác, *Triangles in random graphs*, Disc. Math., 289 (2004), pp. 181–185.

[40] K. Menger, *Zur allgemeinen Kurventheorie*, Fund. Mathematicae, 10 (1927), pp. 96–115. (in German).

[41] V. Nikiforov and R. Schelp, *Cycles and stability*, J. Comb. Theory, Ser. B, 98 (2008), pp. 69–84.

[42] F. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc., 30 (1930), pp. 264–286.

[43] V. Rosta, *On a Ramsey type problem of J.A.Bondy and P.Erdős, I & II*, J. Comb. Theory, Ser. B, 15 (1973), pp. 94–120.

[44] G. Sárközy, *On* 2*-factors with k components*, Disc. Math., 308 (2008), pp. 1962–1972.

[45] A. Scott and A. Tateno, *Complete subgraphs of random graphs*. Manuscript from personal correspondence.

[46] M. Simonovits and V. Sós, *Ramsey-Turán theory*, Disc. Math., 229 (2001), pp. 239–340.

[47] E. Szemerédi, *Regular partitions of graphs*, Colloques Internationaux C.N.R.S - Problèmes Combinatoires et Théorie des Graphes, 260 (1976), pp. 399–401.

[48] P. Turán, *On an extremal problem in graph theory*, Mat. Lapok, 48 (1941), pp. 436–452. (in Hungarian).

[49] W. Tutte, *A factorization of linear graphs*, J. London Math. Soc., 22 (1947), pp. 107–111.