Multicoloured Hamilton cycles in random graphs; an anti-Ramsey threshold.

Colin Cooper
School of Mathematical Sciences,
University of North London,
London N7 8DB, U.K. *

Alan Frieze
Department of Mathematics,
Carnegie Mellon University,
Pittsburgh PA15213, U.S.A.†

Submitted: August 25, 1995; Accepted October 1, 1995

Abstract

Let the edges of a graph G be coloured so that no colour is used more than k times. We refer to this as a *k-bounded colouring*. We say that a subset of the edges of G is *multicoloured* if each edge is of a different colour. We say that the colouring is *$\mathcal{H}$-good*, if a multicoloured Hamilton cycle exists i.e., one with a multicoloured edge-set.
Let $\mathcal{AR}_k = \{G :$ every k-bounded colouring of G is $\mathcal{H}$-good$\}$. We establish the threshold for the random graph $G_{n,m}$ to be in $\mathcal{AR}_k$.

*Research carried out whilst visiting Carnegie Mellon University
†Supported by NSF grant CCR-9225008

1 Introduction

As usual, let $G_{n,m}$ be the random graph with vertex set $V = [n]$ and m random edges. Let $m = n(\log n + \log\log n + c_n)/2$. Komlós and Szemerédi [14] proved that if $\lambda = e^{-c}$ then

$$\lim_{n\to\infty} \mathbf{Pr}(G_{n,m}\text{is Hamiltonian}) = \begin{cases} 0 & c_n \to -\infty \\ e^{-\lambda} & c_n \to c \\ 1 & c_n \to \infty \end{cases},$$

which is $\lim_{n\to\infty} \mathbf{Pr}(\delta(G_{n,m}) \geq 2)$, where δ refers to minimum degree.

This result has been generalised in a number of directions. Bollobás [3] proved a hitting time version (see also Ajtai, Komlós and Szemerédi [1]); Bollobás, Fenner and Frieze [6] proved an algorithmic version; Bollobás and Frieze [5] found the threshold for $k/2$ edge disjoint Hamilton cycles; Bollobás, Fenner and Frieze [7] found a threshold when there is a minimum degree condition; Cooper and Frieze [9], Łuczak [15] and Cooper [8] discussed pancyclic versions; Cooper and Frieze [10] estimated the number of distinct Hamilton cycles at the threshold.

In quite unrelated work various researchers have considered the following problem: Let the edges of a graph G be coloured so that no colour is used more than k times. We refer to this as a *k-bounded colouring*. We say that a subset of the edges of G is *multicoloured* if each edge is of a different colour. We say that the colouring is *$\mathcal{H}$-good*, if a multicoloured Hamilton cycle exists i.e., one with a multicoloured edge-set. A sequence of papers considered the case where $G = K_n$ and asked for the maximum growth rate of k so that every k-bounded colouring is $\mathcal{H}$-good. Hahn and Thomassen [13] showed that k could grow as fast as $n^{1/3}$ and conjectured that the growth rate of k could in fact be linear. In unpublished work Rödl and Winkler [18] in 1984 improved this to $n^{1/2}$. Frieze and Reed [12] showed that there is an absolute constant A such that if n is sufficiently large and k is at most $\lceil n/(A \ln n)\rceil$ then any k-bounded colouring is $\mathcal{H}$-good. Finally, Albert, Frieze and Reed [2] show that k can grow as fast as cn, $c < 1/32$.

The aim of this paper is to address a problem related to both areas of activity. Let $\mathcal{AR}_k = \{G :$

every k-bounded colouring of G is $\mathcal{H}$-good$\}$. We establish the threshold for the random graph $G_{n,m}$ to be in $\mathcal{AR}_k$.

Theorem 1 *If $m = n(\log n + (2k-1)\log\log n + c_n)/2$ and $\lambda = e^{-c}$, then*

$$\begin{aligned}\lim_{n\to\infty} \mathbf{Pr}\left(G_{n,m} \in \mathcal{AR}_k\right) &= \begin{cases} 0 & c_n \to -\infty \\ \sum_{i=0}^{k-1} \frac{e^{-\lambda}\lambda^i}{i!} & c_n \to c \\ 1 & c_n \to \infty \end{cases} \\ &= \lim_{n\to\infty} \mathbf{Pr}(G_{n,m} \in \mathcal{B}_k),\end{aligned} \tag{1}$$

where $\mathcal{B}_k = \{G : G$ has at most $k-1$ vertices of degree less than $2k\}$.

Note that the case $k = 1$ generalises the original theorem of Komlòs and Szemerèdi. We use $\mathcal{AR}_k$ to denote the *anti-Ramsey* nature of the result and remark that there is now a growing literature on the subject of the Ramsey properties of random graphs, see for example the paper of Rödl and Ruciński [17].

2 Outline of the proof of Theorem 1

We will prove the result for the independent model $G_{n,p}$ where $p = 2m/n$ and rely on the monotonicity of property $\mathcal{AR}_k$ to give the theorem as stated, see Bollobás [4] and Łuczak [16]. With a little more work, one could obtain the result that the hitting times for properties $\mathcal{AR}_k$ and $\mathcal{B}_k$ in the graph process are coincidental **whp**[1].

We will follow the basic idea of [12] that, given a k-bounded colouring we will choose a multicoloured set of edges $E_1 \cup E_2$ and show that **whp** $H = (V = [n], E_1 \cup E_2)$ contains a Hamilton cycle. E_1 is chosen randomly, pruned of multiple colours and colours that occur on edges incident with vertices of low degree. E_2 is chosen carefully so as to ensure that vertices of low degree get at least 2 incident

[1]with high probability i.e. probability 1-o(1) as $n \to \infty$

edges and vertices of large degree get a substantial number of incident edges. H is multicoloured by construction. We then use the approach of Ajtai, Komlós and Szemerédi [1] to show that H is Hamiltonian **whp**.

3 Required graph properties

We say a vertex v of $G = G_{n,p}$ is *small* if its degree $d(v)$ satisfies $d(v) < \log n/10$ and *large* otherwise. Denote the set of small vertices by SMALL and the remaining vertices by LARGE. For $S \subseteq V$ we let

$$N_G(S) = N(S) = \{w \notin S : \ \exists v \in S \text{ such that } \{v, w\} \text{ is an edge of } G\}.$$

We now give a rather long list of properties. We claim

Lemma 1 *If $p = (\log n + (2k-1)\log\log n + c)/n$ then $G_{n,p}$ has properties P1 – P9 below* **whp** *and property P10 with probability equal to the RHS of (1).*

P1 $|\text{SMALL}| \le n^{1/3}$.

P2 SMALL contains no edges.

P3 No $v \in V$ is within distance 2 of more than one small vertex.

P4 $S \subseteq \text{LARGE}$, $|S| \le n/\log n$ implies that $|N(S)| \ge |S|\log n/20$.

P5 $T \subseteq V$, $|T| \le n/(\log n)^2$ implies T contains at most $3|T|$ edges.

P6 $A, B \subseteq V$, $A \cap B = \emptyset$, $|A|, |B| \ge 15n\log\log n/\log n$ implies G contains at least $|A||B|\log n/2n$ edges joining A and B.

P7 $A, B \subseteq V$, $A \cap B = \emptyset$, $|A| \le |B| \le 2|A|$ and $|B| \le Dn\log\log n/\log n$ ($D \ge 1$) implies that there are at most $10D|A|\log\log n$ edges joining A and B.

P8 If $|A| \leq Dn \log\log n / \log n$ $(D \geq 1)$ then A contains at most $10D|A| \log\log n$ edges.

P9 G has minimum degree at least $2k-1$.

P10 G has at most $k-1$ vertices of degree $2k-1$.

The proof that $G_{n,p}$ has properties P1–P4 **whp** can be carried out as in [6]. Erdős and Rényi [11] contains our claim about P9, P10. The remaining claims are simple first moment calculations and are placed in the appendix.

4 A simple necessary condition

We now show the relevance of P9, P10. Suppose a graph G has k vertices $v_1, v_2, \ldots, v_k$ of degree $2k-1$ or less and these vertices form an independent set. (The latter condition is not really necessary.) We can use colour $2i-1$ at most k times and colour $2i$ at most $k-1$ times to colour the edges incident with v_i, $1 \leq i \leq k-1$. Now use colours $2, 4, 6, \ldots, 2k-2$ at most once and colour $2k-1$ at most k times to colour the edges inicident with v_k. No matter how we colour the other edges of G there is no multicoloured Hamilton cycle. Any such cycle would have to use colours $1, 2, \ldots, 2k-2$ for its edges incident with $v_1, v_2, \ldots, v_{k-1}$ and then there is only one colour left for the edges incident with vertex v_k.

Let $\mathcal{N}_k$ denote the set of graphs satisfying P1–P10. It follows from Lemma 1 and the above that we can complete the proof of Theorem 1 by proving

$$\mathcal{N}_k \subseteq \mathcal{AR}_k. \tag{2}$$

5 Binomial tails

We make use of the following estimates of tails of the Binomial distribution several times in subsequent proofs.

Let X be a random variable having a Binomial distribution $Bin(n, p)$ resulting from n independent trials with probability p. If $\mu = np$ then

$$\mathbf{Pr}(X \le \alpha\mu) \le \left(\frac{e}{\alpha}\right)^{\alpha\mu} e^{-\mu} \qquad 0 < \alpha \le 1 \tag{3}$$

$$\mathbf{Pr}(X \ge \alpha\mu) \le \left(\frac{e}{\alpha}\right)^{\alpha\mu} e^{-\mu} \qquad 1 \le \alpha. \tag{4}$$

6 Main Proof

Assume from now on that we have a *fixed* graph $G = (V, E) \in \mathcal{N}_k$. We randomly select a multicoloured subgraph H of G, $H = (V, E_1 \cup E_2)$ and prove that it is Hamiltonian **whp**. From now on all probabilistic statements are with respect to the selection of the random set $E_1 \cup E_2$ and *not* the choice of $G = G_{n,p}$.

6.1 Construction of the multicoloured subgraph H

The sets E_1 and E_2 are obtained as follows.

6.1.1 Selection of E_1

(i) Choose edges of the subgraph of G induced by LARGE independently with probability ϵ/k, $\epsilon = e^{-200k}$, to obtain $\widetilde{E}_1$.

(ii) Remove from $\widetilde{E}_1$ all edges whose colour occurs more than once in $\widetilde{E}_1$ and also edges whose colour is the same as that of any edge incident with a small vertex.

Denote the edge set chosen by E_1, and denote by $E_1^\star$ the subset of edges of E which have the same colour as that of an edge in E_1.

Lemma 2 *For $v \in LARGE$ let $d'(v)$ denote the degree of v in $(V, E\backslash E_1^\star)$. Then* **whp**

$$d'(v) > \frac{9}{100k}\log n,$$

for all $v \in LARGE$.

Proof Suppose that large vertex v has edges of $r = r(v)$ different colours $c_1, c_2, \ldots, c_r$ incident with it in G, where $d(v)/k \le r \le d(v)$. Let X_i, $1 \le i \le r$ be an indicator for the event that E_1 contains an edge of colour c_i which is incident with v. Let k_i denote the number of times colour c_i is used in G and let ℓ_i denote the number of edges of colour c_i which are incident with v. Then

$$\begin{aligned}\mathbf{Pr}(X_i = 1) &\le \ell_i \frac{\epsilon}{k}\left(1 - \frac{\epsilon}{k}\right)^{k_i - 1} \\ &\le \epsilon.\end{aligned}$$

The random variables $X_1, X_2, \ldots, X_r$ are independent and so $X = X_1 + X_2 + \cdots + X_r$ is dominated by $Bin(r, \epsilon)$. Thus, by (4),

$$\begin{aligned}\mathbf{Pr}\left(X \ge \frac{r}{10}\right) &\le (10e\epsilon)^{\frac{r}{10}} \\ &\le (10e\epsilon)^{\frac{\log n}{100k}} \\ &\le n^{-3/2},\end{aligned}$$

when $\epsilon = e^{-200k}$. Hence **whp**,

$$d'(v) > \frac{9}{10} r \ge \frac{9}{100k}\log n$$

for every $v \in$ LARGE. □

Assume then that

$$d'(v) > \frac{9}{100k} \log n$$

for $v \in$ LARGE.

6.1.2 Selection of E_2

We show we can choose a monochromatic subset E_2 of $E \setminus E_1^*$ in which

D1 The vertices of SMALL have degree at least 2,

D2 The vertices of LARGE have degree at least $\lfloor \frac{9}{200k^2} \log n \rfloor$.

In order to select E_2, we first describe how to choose for each vertex $v \in V$, a subset A_v of the edges of $E \backslash E_1^\star$ incident with v. These sets A_v, $v \in V$ are pairwise disjoint.

The vertices v of SMALL are independent (P2) and we take A_v to be the set of edges incident with v if $d(v) = 2k-1$, and A_v to be an mk subset otherwise, where $m = \lfloor d(v)/k \rfloor$.

The subgraph F of $E \backslash E_1^\star$ induced by LARGE, is of minimum degree greater than $(9 \log n)/100k$. We orient F so that $|d^-(v) - d^+(v)| \leq 1$ for all $v \in$ LARGE. We now choose a subset A_v of edges directed outward from v by this orientation, of size $\lfloor (9 \log n)/200k^2 \rfloor \, k$.

The following lemma, applied to the sets A_v defined above, gives the required monochromatic set E_2.

Lemma 3 *Let $A_1, A_2, \ldots, A_n$ be disjoint sets with $|A_i| = 2k-1$, $1 \leq i \leq r \leq k-1$ and $|A_i| = m_i k$, $r+1 \leq i \leq n$, where the m_i's are positive integers. Let $A = A_1 \cup A_2 \cup \cdots \cup A_n$. Suppose that the elements of A are coloured so that no colour is used more than k times. Then there exists a multicoloured subset B of A such that $|A_i \cap B| = 2$, $1 \leq i \leq r$ and $|A_i \cap B| = m_i$, $r+1 \leq i \leq n$.*

Proof For $i = 1, ..., r$ partition A_i into $B_{i,1}, \ B_{i,2}$ where $|B_{i,1}| = k-1$ and $|B_{i,2}| = k$, and let $m_i = 2$. For $i = r+1, ..., n$ partition A_i into subsets $B_{i,j}$ $(j = 1, ..., m_i)$ of size k.

Let $X = \{B_{i,j} : i = 1, ..., n, \ j = 1, ..., m_i\}$ and let Y be the set of colours used in the k-bounded colouring of A. We consider a bipartite graph Γ with bipartition (X, Y), where (x, y) is an edge of Γ if colour $y \in Y$ was used on the elements of $x \in X$.

We claim that Γ contains an X-saturated matching. Let $S \subseteq X, \ |S| = s$, and suppose t elements of S are sets of size $k-1$ and $s-t$ are of size k. We have

$$\begin{aligned} |\bigcup_{B_{i,j} \in S} B_{i,j}| &= (s-t)k + t(k-1) \\ &= sk - t. \end{aligned}$$

Thus the set of neighbours $N_\Gamma(S)$ of S in Γ satisfies

$$|N_\Gamma(S)| \geq \lceil s - \tfrac{t}{k} \rceil \geq \lceil s - (\tfrac{k-1}{k}) \rceil = |S|,$$

and Γ satisfies Hall's condition for the existence of an X-saturated matching $M = \{(B_{i,j}, y_{i,j})\}$. Now construct B by taking an element of colour $y_{i,j}$ in $B_{i,j}$ for each (i, j).

□

6.2 Properties of $H = (V, E_1 \cup E_2)$

We first state or prove some basic properties of H.

Lemma 4 *H is multicoloured, and $\delta(H) \geq 2$.*

Lemma 5 *With high probability*

D3 $S \subseteq LARGE, \ |S| \leq \frac{n}{100 \log n} \Longrightarrow |N_H(S)| \geq \frac{\epsilon \log n}{300 k^2} |S|.$

Proof

Case of $|S| \leq n/(\log n)^3$

If $S \subseteq$ LARGE, then $T = N_H(S) \cup S$ contains at least $\lfloor \frac{9}{200k^2} \log n \rfloor |S|/2$ edges in E_2. No subset T of size at most $n/(\log n)^2$ contains more than $3|T|$ edges (by P5). Thus $|T| \geq \lfloor \frac{9}{200k^2} \log n \rfloor |S|/6$ and so

$$|N_H(S)| \geq \frac{3}{500k^2} \log n |S|.$$

Case of $n/(\log n)^3 < |S| \leq n/100 \log n$

By P4, G satisfies $|N(S)| \geq (|S| \log n)/20$ and we can choose a set M of

$$\lfloor (|S| \log n)/20 - (k|\mathrm{SMALL}| \log n)/10 \rfloor$$

edges which have one endpoint in S, the other a distinct endpoint not in S and of a colour different to that of any edge incident with a vertex of SMALL. This set of edges contains at least $|M|/k$ colours. If M contains t edges of colour i and G contains r edges of colour i in total, then the probability ρ that an edge of M of colour i is included in E_1 satisfies

$$\rho \geq \frac{t\epsilon}{k} \left(1 - \frac{\epsilon}{k}\right)^{r-1} \geq \frac{t\epsilon}{k}(1-\epsilon) > \frac{\epsilon}{2k}. \tag{5}$$

Thus $|N_H(S)|$ dominates $Bin(\frac{|M|}{k}, \frac{\epsilon}{2k})$, and by (3)

$$\mathbf{Pr}\left(|N_H(S)| \leq \frac{|M|\epsilon}{4k^2}\right) \leq \left(\frac{2}{e}\right)^{|M|\epsilon/4k^2}.$$

Hence the probability that some set has less than the required number of neighbours to its neighbour set is

$$\begin{aligned} \sum_{s=n/(\log n)^3}^{n/(100 \log n)} \binom{n}{s} \left(\frac{2}{e}\right)^{(\epsilon s \log n)/100k^2} &\leq \sum_s \left[\exp - \left\{\frac{\epsilon \log(e/2)}{100k^2} \log n - 4 \log \log n\right\}\right]^s \\ &= o(1). \end{aligned}$$

□

Lemma 6 *Let* $D \geq \frac{32k^2}{\epsilon}$*; if* $|A|, |B| \geq Dn\frac{\log\log n}{\log n}$ *then* **whp**

D4 H *contains more than* $\lfloor \frac{2}{D}|A|\ |B|\frac{\log n}{n} \rfloor$ *edges between* A *and* B*.*

Proof The proof follows that of Lemma 5. By P6, the number of edges between A and B in G of a colour different to that of any edge incident with a vertex of SMALL is at least $M = \lfloor (|A||B| \log n/2n) - (k|\mathrm{SMALL}| \log n)/10 \rfloor$. Thus the number of E_1-edges between these sets dominates $Bin(M/k, \epsilon/2k)$. Let $K = (1-o(1))\frac{8(1-(\log 4e)/4)}{D}\frac{\log n}{n}$. The probability that there exist sets A, B with at most $\lfloor \frac{2}{D}|A|\ |B|\frac{\log n}{n} \rfloor$ E_1-edges between them is (by (3)) at most

$$
\begin{aligned}
\sum_{a,b} \binom{n}{a}\binom{n}{b}\left(\frac{(4e)^{\frac{1}{4}}}{e}\right)^{(1-o(1))\frac{M\epsilon}{2k^2}} &\leq \sum_{a,b}\left(\frac{ne}{a}\right)^a\left(\frac{ne}{b}\right)^b e^{-Kab} \qquad (6)\\
&\leq \sum_{a,b} \exp\{(a+b)\log\log n - Kab\}\\
&\leq \sum_{a,b} \exp\left\{ab\left(\left(\frac{1}{a}+\frac{1}{b}\right)\log\log n - K\right)\right\}\\
&\leq \sum_{a,b} \exp\left\{ab\left(\frac{2\log n}{Dn} - \frac{3\log n}{Dn}\right)\right\}\\
&\leq n^2 \exp\left\{-Dn\frac{(\log\log n)^2}{\log n}\right\}\\
&= o(1).
\end{aligned}
$$

□

Assume from now on that H satisfies D1–D4. We note the following immediate Corollary.

Corollary 1 **whp** H *is connected.*

Proof If H is not connected then from D4 its has a component C of size at most $Dn\frac{\log\log n}{\log n}$. But then D3 and P3 imply $C \cap \mathrm{LARGE} = \emptyset$. Now apply D1 and P2 to get a contradiction. □

7 Proof that H is Hamiltonian

Let us suppose we have selected a G satisfying properties P1–P10, and sampled a suitable H which satisfies D1–D4. We now show that it must follow that H contains a multicoloured Hamilton cycle.

7.1 Construction of an initial long path

We use rotations and extensions in H to find a maximal path with large rotation endpoint sets, see for example [6], [14]. Let $P_0 = (v_1, v_2, \dots, v_l)$ be a path of maximum length in H. If $1 \le i < l$ and $\{v_l, v_i\}$ is an edge of H then $P' = (v_1 v_2 \dots v_i v_l v_{l-1} \dots v_{i+1})$ is also of maximum length. It is called a *rotation* of P_0 with *fixed endpoint* v_1 and *pivot* v_i. Edge (v_i, v_{i+1}) is called the *broken* edge of the rotation. We can then, in general, rotate P' to get more maximum length paths.

Let $S_t = \{v \in \mathrm{LARGE} :\ v \ne v_1$, is the endpoint of a path obtainable from P_0 by t rotations with fixed endpoint v_1 and all broken edges in $P_0\}$.

It follows from P3 and D3 that $S_1 \ne \emptyset$. It then follows that if $|S_t| \le n/(100 \log n)$ then $|S_{t+1}| \ge \epsilon \log n |S_t|/(1000k^2)$, for making this inductive assumption which is true for $|S_1|$ by D2,

$$
\begin{aligned}
|S_{t+1}| &\ge |N_H(S_t)|/2 - (1 + |S_1| + |S_2| + \cdots |S_t|) \\
&\ge \epsilon \log n |S_t|/(600k^2) - (1 + |S_1| + |S_2| + \cdots |S_t|) \\
&\ge \epsilon \log n |S_t|/(1000k^2).
\end{aligned}
$$

Thus there exists $t_0 \le (1 + o(1)) \log n / \log\log n$ such that $|S_{t_0}| \ge cn$, $c = \epsilon/(10^6 k^2)$. Let $B(v_1) = S_{t_0}$ and $A_0 = B(v_1) \cup \{v_1\}$. Similarly, for each $v \in B(v_1)$ we can construct a set of endpoints

$B(v)$, $|B(v)| \geq cn$ of endpoints of maximum length paths with endpoint v. Note that $l \geq cn$ as every vertex of B_0 lies on P_0.

In summary, for each $a \in A_0$, $b \in B(a)$ there is a maximum length path $P(a,b)$ joining a and b and this path is obtainable from P_0 by at most $(2+o(1))\log n/\log\log n$ rotations.

7.2 Closure of the maximal path

This section follows closely both the notation and the proof methodology used in [1].

Given path P_0 and a set of vertices S of P_0, we say $s \in S$ is an *interior* point of S if both neighbours of s on P_0 are also in S. The set of all interior points of S will be denoted by $int(S)$.

Lemma 7 *Given a set S of vertices with $|int(S)| \geq 7Dn\frac{\log\log n}{\log n}$, $D \geq 32k^2/\epsilon$ there is a subset $S' \subseteq S$ such that, for all $s' \in S'$ there are at least $m = \frac{1}{D}\frac{\log n}{n}|int(S)|$ edges between s' and $int(S')$. Moreover, $|int(S')| \geq |int(S)|/2$.*

Proof We use the proof given in [1]. If there is a $s_1 \in S$ such that the number of edges from s_1 to $int(S)$ is less than m we delete s_1, and define $S_1 = S\backslash\{s_1\}$. If possible we repeat this procedure for S_1, to define $S_2 = S_1\backslash\{s_2\}$ (etc). If this continued for $r = \lfloor\frac{1}{6}|int(S)|\rfloor$ steps, we would have a set S_r and a set $R = \{s_1, s_2, \ldots, s_r\}$, with

$$|int(S_r)| \geq |int(S)| - 3|R| \geq |int(S)| - 3r \geq \frac{|int(S)|}{2}.$$

This step follows because deleting a vertex of S removes at most 3 vertices of $int(S)$. However, there are fewer than

$$\begin{aligned} m|S| &\leq \frac{1}{D}\frac{\log n}{n}|int(S)|\ |R| \\ &\leq \frac{2}{D}\frac{\log n}{n}|int(S_r)|\ |R|, \end{aligned}$$

edges from R to $int(S_r)$, which contradicts our assumption D4. □

In Section 7.1 we proved the existence of maximum length paths $P(a,b),\ b \in B(a),\ a \in A_0$ where $|A_0|,\ |B(a)| \geq cn$. Thus there are at least c^2n^2 distinct endpoint pairs (a,b) and for each such pair there is a path $P(a,b)$ derived from at most $\rho = (2+o(1))\log n/\log\log n$ rotations starting with some fixed maximal path P_0.

We consider P_0 to be directed and divided into 2ρ segments $I_1, I_2, \ldots, I_{2\rho}$ of length at least $\lfloor |P_0|/2\rho \rfloor$, where $|P_0| \geq cn$. As each $P(a,b)$ is obtained from P_0 by at most ρ rotations, the number of segments of P_0 which occur on this path, although perhaps reversed, is at least ρ. We say that such a segment is *unbroken*. These segments have an absolute orientation given by P_0, and another, relative to this by $P(a,b)$, which we regard as directed from a to b. Let t be a fixed natural number. We consider sequences $\sigma = I_{i_1}, ..., I_{i_t}$ of unbroken segments of P_0, which occur in this order on $P(a,b)$, where we consider that σ also specifies the relative orientation of each segment. We call such a sequence σ a *t-sequence*, and say $P(a,b)$ *contains* σ.

For given σ, we consider the set $L = L(\sigma)$ of ordered pairs (a,b), $a \in A_0,\ b \in B(a)$ which contain the sequence σ.

The total number of such sequences of length t is $(2\rho)_t 2^t$. Any path $P(a,b)$ contains at least $\rho \geq \log n/\log\log n$ unbroken segments, and thus at least $\binom{\rho}{t}$ t-sequences. The average, over t-sequences, of the number of pairs containing a given t-sequence is therefore at least

$$\frac{c^2n^2\binom{\rho}{t}}{(2\rho)_t 2^t} \geq \alpha n^2,$$

where $\alpha = c^2/(4t)^t$. Thus there is a t-sequence σ_0 and a set $L = L(\sigma_0)$, $|L| \geq \alpha n^2$ of pairs (a,b) such that for each $(a,b) \in L$ the path $P(a,b)$ contains σ_0. Let $\hat{A} = \{a :\ L$ contains at least $\alpha n/2$ pairs with a as first element$\}$. Then $|\hat{A}| \geq \alpha n/2$. For each $a \in \hat{A}$ let $\hat{B}(a) = \{b :\ (a,b) \in L\}$.

Let $t = 1700D^2/c$, $D = 32k^2/\epsilon$ and let C_1 denote the union of the first $t/2$ segments of σ_0, in the fixed order and with the fixed relative orientation in which they occur along *any* of the paths

$P(a,b)$, $(a,b) \in L$. Let C_2 denote the union of the second $t/2$ segments of σ_0. C_1 and C_2 contain at least $\frac{t}{2}\, cn\, \frac{\log\log n}{4\log n}(1-o(1))$ interior points which from Lemma 7 gives sets C_1', C_2' with at least

$$\frac{tc(1-o(1))}{16} n^{\frac{\log\log n}{\log n}} \geq 100D^2 n^{\frac{\log\log n}{\log n}}$$

interior points.

It follows from D4 that there exists $\hat{a} \in \hat{A}$ such that H contains an edge from $\hat{a}$ to C_1'. Similarly, H contains an edge joining some $\hat{b} \in \hat{B}(\hat{a})$ to C_2'. Let x be some vertex separating C_1' and C_2' along $\hat{P} = P(\hat{a}, \hat{b})$. We now consider the two half paths P_1, P_2 obtained by splitting $\hat{P}$ at x. We consider rotations of P_i, $i = 1, 2$ with x as a fixed endpoint. We show that in both cases the finally constructed endpoint sets V_1, V_2 are large enough so that D4 guarantees an edge from V_1 to V_2. We deduce that H is Hamiltonian as the path it closes is of maximum length and H is connected.

Consider P_1. Let $T_i = \{v \in C_1' :\ v \neq x$ is the endpoint of a path obtainable from P_1 by t rotations with fixed endpoint x, pivot in $int(C_1')$ and all broken edges in $P_1\}$. We claim we can choose sets $U_i \subseteq T_i$, $i = 1, 2, \ldots$ such that $|U_1| = 1$ and $|U_{i+1}| = 2|U_i|$, as long as $|U_i| \leq Dn^{\frac{\log\log n}{\log n}}$. Thus there is an i^* such that $|U_{i^*}| \geq Dn^{\frac{\log\log n}{\log n}}$ and we are done. Note that $T_1 \neq \emptyset$ because $\hat{a}$ has an H-neighbour in $int(C_1')$. Note also that if we make a rotation with pivot in $int(C_1')$ and broken edge in P_1 then the new endpoint created is C_1'.

Let y be a vertex of U_i. Then by Lemma 7 there are at least $100D\log\log n$ edges between y and $int(C_1')$. Thus the number of edges from U_i to $int(C_1')$ is at least $50|U_i|D\log\log n$. As $|\bigcup_{j=1}^{i} U_i| < 2|U_i|$ at most $20D|U_i|\log\log n$ of these edges are contained in $\bigcup_{j=1}^{i} U_j$ (from P8), and so by P7 we have $|T_{i+1}| > 2|U_i|$ and we select a subset of size exactly $2|U_i|$.

8 Similar Problems

We note that it is straightforward to extend the above analysis to find the corresponding thresholds when Hamilton cycle is replaced by perfect matching or spanning tree. Now **whp** one needs enough edges so that the following replacements for conditions P9 and P10 hold true.

P9a G has minimum degree $k-1$.

P10a G has at most $k-1$ vertices of degree $k-1$

That these conditions are necessary can be argued as in Section 4, since connectivity and the existence of a perfect matching require minimum degree one. Lemma 3 is replaced by

Lemma 8 *Let $A_1, A_2, \ldots, A_n$ be disjoint sets with $|A_i| = k-1$, $1 \leq i \leq r \leq k-1$ and $|A_i| = m_i k$, $r+1 \leq i \leq n$, where the m_i's are positive integers. Let $A = A_1 \cup A_2 \cup \cdots \cup A_n$. Suppose that the elements of A are coloured so that no colour is used more than k times. Then there exists a multicoloured subset B of A such that $|A_i \cap B| = 1$, $1 \leq i \leq r$ and $|A_i \cap B| = m_i$, $r+1 \leq i \leq n$.*

The proof is the same.

We choose E_1 and E_2 in the same way as before. The fact that H is connected proves the existence of a multicoloured tree. For a perfect matching one can remove from H all vertices of degree one together with their neighbours and argue that the graph that remains is Hamiltonian (assuming n is even). The proof is essentially that of Section 7.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi. *The first occurrence of Hamilton cycles in random graphs.* Annals of Discrete Mathematics 27 (1985) 173-178.

[2] M.J. Albert, A.M. Frieze and B. Reed, *Multicoloured Hamilton Cycles.* Electronic Journal of Combinatorics 2 (1995) R10.

[3] B. Bollobás. *The evolution of sparse graphs.* Graph Theory and Combinatorics. (Proc. Cambridge Combinatorics Conference in Honour of Paul Erdős (B. Bollobás; Ed)) Academic Press (1984) 35-57

[4] B. Bollobás. *Random Graphs.* Academic Press (1985)

[5] B. Bollobás and A.M. Frieze, *On matchings and Hamiltonian cycles in random graphs.* Annals of Discrete Mathematics 28 (1985) 23-46.

[6] B. Bollobás, T.I. Fenner and A.M. Frieze. *An algorithm for finding Hamilton cycles in random graphs.* Combinatorica 7 (1987) 327-341.

[7] B. Bollobás, T. Fenner and A.M. Frieze. *Hamilton cycles in random graphs with minimal degree at least k.* (A Tribute to Paul Erdős (A.Baker, B.Bollobas and A.Hajnal; Ed)) (1990) 59-96.

[8] C. Cooper. *1-pancyclic Hamilton cycles in random graphs.* Random Structures and Algorithms 3.3 (1992) 277-287

[9] C. Cooper and A. Frieze. *Pancyclic random graphs.* Proc. 3rd Annual Conference on Random Graphs, Poznan 1987. Wiley (1990) 29-39

[10] C. Cooper and A. Frieze. *On the lower bound for the number of Hamilton cycles in a random graph.* Journal of Graph Theory 13.6 (1989) 719-735

[11] P. Erdős and A. Rényi. *On the strength of connectedness of a random graph.* Acta. Math. Acad. Sci. Hungar. 12 (1961) 261-267.

[12] A. Frieze and B. Reed. *Polychromatic Hamilton cycles.* Discrete Maths. 118 (1993) 69-74.

[13] G. Hahn and C. Thomassen. *Path and cycle sub-Ramsey numbers, and an edge colouring conjecture.* Discrete Maths. 62 (1986) 29-33

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

[15] T. Łuczak. *Cycles in random graphs.* Discrete Maths. (1987)

[16] T. Łuczak, *On the equivalence of two basic models of random graph*, Proceedings of Random graphs 87, Wiley, Chichester (1990), 151-159.

[17] Rödl and A. Ruciński, *Threshold functions for Ramsey properties.* (to appear)

[18] Rödl and Winkler. Private communication (1984)

Appendix: Proofs of P6–P8

P5 $T \subseteq V$, $|T| \leq n/(\log n)^2$ implies T contains at most $3|T|$ edges.

The number of edges in T is $Bin\left(\binom{|T|}{2}, p\right)$. By (4) the probability that there exists T with $3|T|$ edges is at most

$$\begin{aligned}\sum_{t=7}^{n/(\log n)^2} \binom{n}{t} \left(\frac{e\binom{t}{2}p}{3t}\right)^{3t} e^{-\binom{t}{2}p} &\leq \sum_t \left(\frac{ne}{t}\right)^t \left(\frac{(1+o(1))et\log n}{6n}\right)^{3t} \\ &\leq \sum_t \left(\frac{(\log n)^3 t^2}{n^2}\right)^t \\ &= o(1).\end{aligned}$$

P6 $A, B \subseteq V$, $A \cap B = \emptyset$, $|A|, |B| \geq 15n \log\log n/\log n$ implies G contains at least $|A||B|\log n/2n$ edges joining A and B.

The number of edges between A and B is $Bin(|A|\ |B|, p)$. By (3), the probability there exist sets A, B with less than half the expected number of edges between them, is at most

$$\begin{aligned}\sum_{a,b} \binom{n}{a}\binom{n}{b}\left(\frac{2}{e}\right)^{abp} &\leq \exp\left\{a\log(ne/a) + b\log(ne/b) - abp\log(e/2)\right\} \\ &\leq n^2 \exp\left\{-\frac{n(\log\log n)^2}{\log n}(15)^2\left(\log(e/2) - 2/15\right)\right\} \\ &= o(1),\end{aligned}$$

by the same arguments as those following (6) in Lemma 6.

P7 $A, B \subseteq V$, $A \cap B = \emptyset, |A| \leq |B| \leq 2|A|$ and $|B| \leq Dn\log\log n/\log n$ $(D \geq 1)$ implies that there are at most $10D|A|\log\log n$ edges joining A and B.

Let $|B| = 2|A| = 2a$. We have then that the probability that there exist A, B such that there are at least $10D|A|\log\log n$ edges between the sets is at most

$$\begin{aligned}\sum_a \binom{n}{a}\binom{n}{2a}\binom{2a^2}{10Da\log\log n}p^{10Da\log\log n} &\le \sum_a \left[\left(\frac{ne}{a}\right)\left(\frac{ne}{2a}\right)^2\left(\frac{ae\log n}{5Dn\log\log n}\right)^{10D\log\log n}\right]^a \\ &\le \sum_a \left[\frac{e^3}{4}\left(\frac{\log n}{2D\log\log n}\right)^3\left(\frac{e}{5}\right)^{10D\log\log n}\right]^a \\ &\le \sum_a \left(\frac{1}{\log n}\right)^a \\ &= o(1).\end{aligned}$$

P8 If $|A| \le Dn\log\log n/\log n$ $(D \ge 1)$ then A contains at most $10D|A|\log\log n$ edges.

We may assume $|A| > n/(\log n)^2$ by P5. The number of induced edges in A is $Bin\left(\binom{|A|}{2}, p\right)$. By (4) the probability there exists a set A with at least $20(1-o(1))$ times the expected number of edges is at most,

$$\begin{aligned}\sum_a \binom{n}{a}\left(\frac{e}{19}\right)^{10Da\log\log n} &\le \sum_a n\exp\{a\log(ne/a) - 10Da\log\log n\log(19/e)\} \\ &\le \sum_a n\exp\left\{-\frac{Dn\log\log n}{(\log n)^2}\left(10\log(19/e)-2\right)\right\} \\ &= o(1).\end{aligned}$$