Anti-Ramsey Properties of Random Graphs

Tom Bohman†, Alan Frieze‡, Oleg Pikhurko†
Department of Mathematical Sciences,
Carnegie Mellon University,
Pittsburgh, PA 15213

Cliff Smyth
Department of Mathematics,
Massachusetts Institute of Technology,
Cambridge, MA 02139

August 25, 2006

Abstract
We call a coloring of the edge set of a graph $G$ a $b$-bounded coloring if no color is used more than $b$ times. We say that a subset of the edges of $G$ is rainbow if each edge is of a different color. A graph has property $A(b,H)$ if every $b$-bounded coloring of its edges has a rainbow copy of $H$. We estimate the threshold for the random graph $G_{n,p}$ to have property $A(b,H)$.

1 Introduction

We call a coloring of the edge set of a graph $G$ a $b$-bounded coloring if no color is used more than $b$ times. We say that a subset of the edges of $G$ is rainbow (or polychromatic) if each edge is of a different color. We consider the following question: What relationship between $b, G$ and $H$ implies that every $b$-bounded coloring of the graph $G$ contains a rainbow copy of the graph $H$ (i.e. a copy of $H$ in which $E(H)$ is rainbow colored)? Note that this can be viewed as a variation on classical Ramsey theory, but here instead of a homogeneous (i.e. monochromatic) copy of $H$ we are interested in a heterogeneous (i.e. rainbow) copy of $H$. Questions of this form have been studied in a number of contexts. Erdős, Simonovits and Sós considered the minimum number of colors needed to ensure a rainbow copy of $H$ in every coloring of the edge set of $K_n$ where we require that every color is used at least once [6]. Leffmann, Rödl and Wysocka considered some variations on this question where the restriction that each color is used at least once is replaced by other natural restrictions, including $b$-bounded coloring [15]. The existence of rainbow Hamilton cycles in edge colored copies of complete graphs was studied in [1], [5], [9], [12]. The existence of rainbow stars was studied in Hahn [10], [11] and Fraisse, Hahn and Sotteau [8]. The complexity of finding rainbow sub-graphs was studied by Fenner and Frieze.

†Research supported in part by NSF grant DMS-0502793
‡Research supported in part by NSF grant DMS-0457512
Cooper and Frieze [4] studied the existence of polychromatic Hamilton cycles in random graphs. In this paper we study the existence of rainbow copies of a fixed graph \( H \) in \( b \)-bounded colorings of the random graph \( G_{n,p} \).

Let \( H \) be a fixed graph. Let \( v_H \) and \( e_H \) denote the number of vertices and edges of \( H \) respectively. For a positive integer \( b \) let \( A(b, H) \) denote the following graph property: \( G \in A(b, H) \) iff every \( b \)-bounded coloring of \( E(G) \) has a rainbow copy of \( H \). Define

\[
m_H = \frac{e_H - 1}{v_H - 2},
\]
\[
m^*_H = \max\{m_{H'} : H' \subseteq H, v_{H'} \geq 3\},
\]
\[
p^* = \frac{1}{n^{1/m_{H^*}}},
\]

One can show that, unless maximum degree \( \Delta_H = 1 \), it enough to consider only connected sub-graphs \( H' \).

Note that, when \( p \) is not too small, \textbf{w.h.p} the number of copies of \( H \) in \( G_{n,p} \) is \( \Theta(n^{v_H} p^{e_H}) \) while the number of edges in \( G_{n,p} \) is \( \Theta(np^2) \). (\textbf{W.h.p} stands for \textbf{with high probability}, that is, with probability \( 1 - o(1) \) as \( n \to \infty \).) Thus if \( p \ll p^* \) then the number of copies of \( H \) in \( G_{n,p} \) is much fewer than the number of edges in \( G_{n,p} \) and so it should be the case that \textbf{w.h.p} it is easy to color the edges so that there is no rainbow copy of \( H \). On the other hand, when \( p \gg p^* \) there are so many copies of \( H \) relative to the number of edges that \textbf{w.h.p} a rainbow copy of \( H \) should be unavoidable. So, at first glance, it is natural to expect \( p^* \) to be the threshold for the anti-Ramsey property \( A(b, H) \). Of course, this reasoning can also be applied to the classical Ramsey property, and \( p^* \) is (with a few exceptions) indeed the threshold for the Ramsey property that every coloring of \( G_{n,p} \) with a set of \( r \) colors has a monochromatic copy of \( H \) (as shown by Rödl and Ruciński [16]).

There is one immediate exception to this general framework for the anti-Ramsey property \( A(b, H) \). Note that if \( H \) is a forest then \( m_{H^*}^* = 1 \) (assuming that \( \Delta_H \geq 2 \)) but it turns out that there are trees that have the property \( A(b, H) \). Since \( p = n^{-(k+1)/k} \) is the threshold probability for having a copy of every tree with \( k \) edges, it follows that \( p = 1/n^{m_{H^*}^*} = 1/n \) is not the threshold for the anti-Ramsey property \( A(b, H) \).

So we begin with a general result for arbitrary graphs that are not acyclic.

**Theorem 1** Suppose that \( H \) is a graph containing at least one cycle and that \( b \) is sufficiently large. Then there exist \( c_1 = c_1(b, H) \) and \( c_2 = c_2(b, H) \) such that if \( p = cn^{-1/m_{H^*}^*} \) then

\[
\lim_{n \to \infty} \Pr(G_{n,p} \in A(b, H)) = \begin{cases} 
0 & \text{if } c \leq c_1 \\
1 & \text{if } c \geq c_2.
\end{cases}
\]

(1)

We study the threshold for \( A(b, K_3) \) in more detail. For \( b = 2 \) and \( H = K_3 \), the situation is completely resolved.

**Theorem 2** Let \( p = \frac{c_n}{n^{2/3}} \). Then

\[
\lim_{n \to \infty} \Pr(G_{n,p} \in A(2, K_3)) = \begin{cases} 
0 & \text{if } c_n \to 0 \\
1 - e^{-c^6/24} & \text{if } c_n \to c \\
1 & \text{if } c_n \to \infty.
\end{cases}
\]
Note that Theorem 2 shows that some condition on $b$ is necessary in Theorem 1 (since $m_{K_3}^* = 1/2$). When $b = 3$ and $H = K_3$ there is an intriguing gap in our result.

**Theorem 3** Let $p = \frac{c}{n^{1/2}}$. Then,

$$\lim_{n \to \infty} \Pr(G_n \in A(3, \mathcal{A})) = \begin{cases} 1 - e^{-c^{10}/120} & \text{if } c < \sqrt{2} \\ 1 & \text{if } c > \sqrt{2}. \end{cases}$$

Theorem 3 leaves open the possibility of a ‘one-sided-sharp’ phase transition; to be precise, there could be a critical value $c \in [1/\sqrt{2}, \sqrt{2}]$ at which the probability that $G_n/c/\sqrt{n}$ has property $A(3, K_3)$ quickly jumps from $1 - e^{-c^{10}/120}$ to 1. Finally, we note that (1) holds for $H = K_3$ and $b \geq 4$ (this follows immediately from the proof of Theorem 3).

We now turn to the anti-Ramsey thresholds for forests. For a tree $T$, let $s(b, T)$ be the minimum value $s$ such that there exists a tree with $s$ edges having property $A(3, T)$. For a fixed forest $F$, the threshold for $A(b, F)$ will then be $p = n^{-(s+1)/s}$ where $s$ is the maximum of $s(b, T)$ over all connected components $T$ of $F$. So the study of thresholds for $A(b, F)$ amounts to the study of $s(b, T)$.

**Theorem 4** Let $T$ be a fixed tree with diameter $l$, and set $m = \lfloor l/2 \rfloor$. Then (letting $b \to \infty$) we have

$$s(b, T) = \Theta(b^m).$$

The upper bound in Theorem 4 is given by a certain class of trees which we conjecture always determines $s(b, T)$. Let $T$ be a tree, $e$ be an edge in $T$ and $b$ be a positive integer. In Section 5 we define the tree $B_{T,e,b}$ (which we dub the $b$-blow-up of $T$ centered at $e$) and show that $B_{T,e,b} \in A(b, T)$.

**Conjecture 1** For any $b \geq 2$ and tree $T$,

$$s(b, T) = \min_{e \in T} \{|E(B_{T,e,b})|\}.$$

In support of this conjecture, we verify it for paths and rooted trees with a constant branching factor. Using similar proof techniques we have verified the conjecture for a few other special classes of trees (e.g. the $m$-fork which consists of $m$ leaves added to an endpoint of a path of length 3). The details for these other classes of trees are omitted for the sake of brevity.

**Theorem 5**

(a) Let $P_l$ be the path with $l$ edges. We have

$$s(b, P_l) = \begin{cases} (b + 1) \sum_{i=0}^{k-1} b^i & \text{if } l = 2k \\ 1 + 2 \sum_{i=1}^{k} b^i & \text{if } l = 2k + 1. \end{cases}$$

(b) Let $T_{d,l}$ be a rooted tree, with all leaves at distance $l$ from the root such that every non-leaf has the same degree $d$. Then

$$s(b, T_{d,l}) = 1 + 2 \sum_{i=1}^{l-1} (b(d - 1))^i + (b(d - 1))^l.$$
We prove our theorems in the following order. Theorem 2 is proved first in Section 2. Theorem 3 is proved in Section 3. The general theorem, Theorem 1, is proved in Section 4, and we discuss trees in Section 5.

A few words on our notation. We will use ‘⊆’ to denote inclusion. The relation ‘⊂’ excludes the case of equality. The expression $a_n \sim b_n$ means that $\lim_{n \to \infty} a_n / b_n = 1$. The $O()$-notation is standard.

2 Proof of Theorem 2

We begin by noting that $K_4$ has the anti-Ramsey property $A(2, K_3)$ (by proving the following, more general statement).

Lemma 6

$$K_{r+2} \in A(r, K_3) \text{ for } r \geq 1.$$  

**Proof** Assume for the sake of contradiction that a given $r$-bounded coloring of $K_{r+2}$ does not have a rainbow triangle. Let $C$ be a largest connected component, in terms of number of vertices, induced by edges of the same color, red say. The number of vertices in $C$ is at most $r+1$ and so there is a vertex $v \not\in C$. Consider the edges from $v$ to $C$. They cannot be colored red and as there are no rainbow triangles they must all be the same color, blue say. But then the connected component induced by the blue edges that contains $v$ has more vertices than $C$, contradiction. \qed

Now assume that $p = \frac{c}{n^{2/3}}$ and let $Z_4$ denote the number of copies of $K_4$ in $G_{n,p}$. Thus

$$E(Z_4) = \binom{n}{4} p^6 \to \frac{c^6}{24}.$$  

It is well known ([2],[13]) that in this case $Z_4$ is asymptotically Poisson and so

$$\Pr(Z_4 = 0) \to e^{-c^6/24}.$$  

Since $K_4 \in A(2, K_3)$ and the property $A(b, H)$ is monotone, we can prove Theorem 2 by showing that if $p = \frac{c}{n^{2/3}}$ then

$$\lim_{n \to \infty} \Pr(G_{n,p} \in A(2, K_3) \mid G_{n,p} \text{ is } K_4\text{-free}) = 0. \tag{2}$$

We now define a **triangle graph** $\Gamma = (W, X)$ where $W$ is the set of triangles of $G_{n,p}$ and $(T_1, T_2) \in X$ iff the triangles $T_1, T_2$ share an edge. If $C = \{T_1, T_2, \ldots, T_\ell\}$ is a connected component of $\Gamma$ we define the **base graph** of $C$ to be the sub-graph $G(C)$ of $G_{n,p}$ with vertex set $V(C) = \bigcup_{i=1}^{\ell} V(T_i)$ and edge set $E(C) = \bigcup_{i=1}^{\ell} E(T_i)$.

We say that a graph $K$ is **$d$-degenerate** if there is an ordering $v_1, v_2, \ldots, v_k$ of the vertices of $K$ such that each vertex $v$ has at most $d$ neighbors that appear before $v$ in this ordering; to be precise,  

$$|\{j : j < i \text{ and } \{v_i, v_j\} \in E(K)\}| \leq d$$

for every $i = 1, \ldots, k$. Note that for any component of $\Gamma$ we have

$$|E(C)| \geq 2|V(C)| - 3$$

with equality iff $G(C)$ is 2-degenerate.
Lemma 7 Let $\Gamma$ be the triangle graph of $G_{n,p}$ with $p = c/n^{2/3}$. Whp every component $C$ of $\Gamma$ satisfies one of the following two conditions

(a) $G(C)$ is isomorphic to $K_4$, or
(b) $G(C)$ is 2-degenerate.

Proof We first show that whp $|V(C)| \leq 6$ for all components $C$ of $\Gamma$. Indeed, if there exists a component $C$ of $\Gamma$ such that $|V(C)| \geq 7$ then there is a set of 7 vertices in $G_{n,p}$ that spans at least 11 edges. A simple first moment calculation shows whp that no such sub-graph of $G_{n,p}$ exists.

It remains to show that whp there are no components $C$ of $\Gamma$ such that $G(C)$ is not 2-degenerate and $V(C) = 5$ or 6. However, these correspond to sub-graphs of $G_{n,p}$ with 5 vertices and 8 edges and sub-graphs of $G_{n,p}$ with 6 vertices and 10 edges, respectively. By the first moment method no such sub-graphs of $G_{n,p}$ exist. \qed

We are now ready to prove (2). Suppose $G_{n,p}$ is $K_4$-free and that every component $C$ of $\Gamma$ has $G(C)$ 2-degenerate. We color the edge set of $G_{n,p}$ by considering each component of $\Gamma$ in turn. Consider a 2-degenerate ordering $v_1, \ldots, v_k$ of the vertices of $G(C)$. We introduce one color for each vertex and color the edge $\{v_i, v_j\}$ with the color corresponding to the maximum of $i$ and $j$. If $\{v_a, v_b, v_c\}$ is a triangle in $C$ then the color corresponding to the maximum of $a$, $b$ and $c$ appears on 2 of the edges in triangle. Thus, this gives a 2-bounded coloring of the edges of $G_{n,p}$ with no rainbow $K_3$.

3 Proof of Theorem 3

Suppose first that $p = \frac{c}{n^{1/2}}$ and $c < 1/\sqrt{2}$.

Let $Z_5$ denote the number of copies of $K_5$ in $G_{n,p}$. We have

$$E(Z_5) = \binom{n}{5} p^{10} \to \frac{c^{10}}{120}$$

and

$$\Pr(Z_5 = 0) \to e^{-c^{10}/120}.$$ 

Since $K_5 \in \mathcal{A}(3, K_3)$ by Lemma 6, we can prove the first part of Theorem 3 by showing that if $p = \frac{c}{n^{1/2}}$ and $c < 1/\sqrt{2}$ then

$$\lim_{n \to \infty} \Pr(G_{n,p} \in \mathcal{A}(3, K_3) \mid G_{n,p} \text{ is } K_5\text{-free}) = 0. \quad (3)$$

Let the triangle graph $\Gamma$ be as defined in Section 2. A component $C$ of $\Gamma$ is safe if

$$|E(C)| \leq 2|V(C)|.$$

Lemma 8 Whp every connected component $C$ of $\Gamma$ is safe.

Proof Consider the following process that generates all connected components of $\Gamma$. Choose 3 vertices $u,v,w$ and let $V_0 = \{u,v,w\}$ and let $E_0 = \{\{u,v\}, \{u,w\}, \{v,w\}\}$. If $u,v,w$ generate a triangle in $G$ continue as follows: Suppose that we have generated a disjoint sequence of vertex sets $V_0, V_1, \ldots, V_k$ and edge sets $E_1, E_2, \ldots, E_k$. Initialize $V_{k+1} = E_{k+1} = \emptyset$ and then perform the following steps:
A. For each \( z \notin V^{(k)} = \bigcup_{i=0}^{k} V_k \) and \( e = \{x, y\} \in E_k \), see if both edges \( \{x, z\}, \{y, z\} \) exist in \( G_{n,p} \). If so, add these edges to \( E_{k+1} \) and \( z \) to \( V_{k+1} \). This is done one vertex at a time and for each vertex it is done one edge at a time. We place \( z \) in \( V_{k+1} \) on the first success and then move onto the next vertex.

B. For each pair of vertices consisting of a vertex \( z \) in \( V_{k+1} \) and a vertex \( a \) in \( V^{(k+1)} \) see if the edge \( \{z, a\} \) is in \( G_{n,p} \). If so add this edge to \( E_{k+1} \).

Of course, we terminate when \( V_{k+1} = \emptyset \) after step A. Let \( V_{\text{final}} \) and \( E_{\text{final}} \) be the vertex and edge sets, respectively, that are formed at the end of this process and let \( C_{u,v,w} \) be the triangle component containing the triangle \( u, v, w \) (if this triangle appears). Note that \( E_{\text{final}} \) is not necessarily equal to \( E(C_{u,v,w}) \) as we add edges in step B that are not necessarily involved in triangles. However, we do have \( E(C_{u,v,w}) \subseteq E_{\text{final}} \). Also,

\[
|E_{\text{final}} \setminus E(C_{u,v,w})| \geq 2|V_{\text{final}} \setminus V(C_{u,v,w})|.
\]

Thus, if \( |E_{\text{final}}| \) is at most \( 2|V_{\text{final}}| \) then \( C_{u,v,w} \) is safe. Since edges and vertices join at a ratio of 2 edges to each vertex during step A, we have \( |E_{\text{final}}| \leq 2|V_{\text{final}}| \) iff the number of edges that join during a step B is at most 3.

Note that throughout our process the conditioning we impose on \( G_{n,p} \) is of a very special form. At any given point we have fully queried certain edges (i.e., we are conditioning on the event that some set of edges appears and some other set of edges does not appear). Furthermore, since we have checked to see if certain pairs of edges appear in \( G_{n,p} \) in step A we also condition on the event that a certain collection of pairs of edges do not appear. Since the latter is a downwardly closed event, it follows from the FKG inequality that when we condition on this event the probability that any set of \( k \) edges (that have not been fully queried) lie in \( G_{n,p} \) is at most \( p^k \).

We view our process as a sort of branching process in which the edges are the individuals and each edge that joins has a ‘parent’ edge that it attaches to. Let \( A_i \) be the number of step A children of the \( i^{th} \) edge to join. Let \( B_i \) be the number of step B children of the \( i^{th} \) edge to join. Note that there is ambiguity in the parent of a type B edge. We assign paternity to an arbitrarily chosen incident - and previously appearing - edge. Note also that each of these \( B_i \) edges shares a common vertex with its parent.

Define the constant \( c' = 6/(\delta c)^2 \) where \( \delta > 0 \) is defined by \( c + \delta = 1/\sqrt{2} \). Let \( K = c' \log n \). Let \( X_1, X_2, \ldots \) be a sequence of i.i.d. \( Bi(n, p^2) \) random variables. Let \( Y_1, Y_2, \ldots \) be a sequence of i.i.d. \( Bi(2K, p) \) random variables. Note that \( 2X_i \) dominates \( A_i \) for all \( i \) while \( Y_i \) dominates \( B_i \) for \( i = 1, \ldots, K \).

Let \( S \) be the event that the edges \( \{u, v\}, \{u, w\}, \{v, w\} \) appear in \( G_{n,p} \). If \( S \) occurs and \( |E_{\text{final}}| > K \) then \( \sum_{i=1}^{K} A_i > K - \sum_{i=1}^{K} B_i - 3 \). If \( S \) occurs and \( K \geq |E_{\text{final}}| > 2|V_{\text{final}}| \) then \( \sum_{i=1}^{K} B_i \geq 4 \).

Thus we have

\[
Pr \left( |E_{\text{final}}| > 2|V_{\text{final}}| \mid S \right) \leq Pr \left( |E_{\text{final}}| > K \mid S \right) + Pr \left( |E_{\text{final}}| > 2|V_{\text{final}}| \mid S \cap |E_{\text{final}}| \leq K \right) \leq \left[ Pr \left( \sum_{i=1}^{K} 2X_i > K - 6 \right) + Pr \left( \sum_{i=1}^{K} Y_i \geq 4 \right) \right] + Pr \left( \sum_{i=1}^{K} Y_i \geq 4 \right).
\]
Now we apply the Chernoff bounds. Since the sum \( \sum_{i=1}^{K} X_i \) is distributed as \( Bi(Kn, p^2) \) we have

\[
Pr \left( \sum_{i=1}^{K} X_i \geq \frac{K - 6}{2} \right) \leq Pr \left( \sum_{i=1}^{K} X_i \geq Knp^2(1 + \delta) \right) \leq \exp \left\{ -\delta^2 Knp^2 / 3 \right\} = \frac{1}{n^2}.
\]

(Note that we use the fact \( (1/\sqrt{2} - x)^2 (1 + x) < 1/2 \) for \( x \) in the interval \( (0, 1/\sqrt{2}) \) and that we assume that \( n \) is sufficiently large.) For the sum of the \( Y_i \)’s we simply have

\[
Pr \left( \sum_{i=1}^{K} Y_i \geq 4 \right) \leq \left( \frac{2K^2}{4} \right) \frac{p^4}{4} = O \left( \frac{(\log n)^8}{n^2} \right).
\]

Therefore, by the union bound, the probability that there is a triangle component \( C \) that is not safe is

\[
O \left( n^3 \left( \frac{1}{\sqrt{n}} \right)^3 \frac{(\log n)^8}{n^2} \right) = o(1).
\]

Assume that all triangle components \( C \) are safe. We give an algorithm for coloring each triangle component in such a way that no triangle is rainbow. Consider a fixed component \( C \) of \( \Gamma \). We define the graph \( D \) to be \( K_6 \) minus a perfect matching. Let \( v_1, v_2, v_3, \ldots, v_\ell \) be the vertices of \( C \) listed so that

(i) If \( C \) contains a copy of \( D \) then this graph comes at the beginning of the sequence. If there is no copy of \( D \), but there is a copy of \( K_5 - e \) then this graph comes at the beginning of the sequence. Finally, if there is no copy of \( D \) or \( K_5 - e \) but there is a copy of \( K_4 \) then this graph comes at the beginning of the sequence.

If \( C \) does not contain any of these graphs then the first three vertices in the sequence form a triangle.

(ii) Let \( v_k \) be the last vertex in our initial graph as defined in (i). Each subsequent vertex \( v_i, i > k \) has at least 2 neighbors (called back-neighbors below) among \( v_1, \ldots, v_{i-1} \) and the set of neighbors of \( v_i \) among \( v_1, \ldots, v_{i-1} \) span at least one edge.

Property (ii) follows from the facts that the ordering of vertices by their addition to the \( C \) satisfies it and that we can start growing \( C \) from any triangle, in particular, from one belonging to the targeted initial graph.

For \( i = k + 1, \ldots, v \) let \( d_i \) be the number of neighbors \( v_i \) has among \( v_1, \ldots, v_{i-1} \). By assumption \( d_i \geq 2 \) for all \( i > k \). Let \( I_t = \{ i > k : d_i = t \} \), \( t \geq 3 \) and \( I = I_3 \cup I_4 \cup I_5 \). Note that our assumption implies that \( I_t = \emptyset \) for \( t \geq 6 \) and \( |I_3| + 2|I_4| + 3|I_5| \leq 3 \). Furthermore,

\[
\begin{align*}
C \text{ contains } D & \implies I = \emptyset \\
C \text{ contains } K_5 - e & \implies I_4 \cup I_5 = \emptyset \\
C \text{ contains } K_4 & \implies I_5 = \emptyset \text{ and } |I_3| + 2|I_4| \leq 2.
\end{align*}
\]

We first check that \( K_5 - e \) and \( D \) can be colored without creating a rainbow triangle.

\( K_5 - e \): Suppose that \( e = \{4, 5\} \). The following table shows a coloring without a rainbow triangle:
D: Suppose that the deleted matching is \{1, 4\}, \{2, 5\}, \{3, 6\}. The following table shows a coloring without a rainbow triangle:

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

We then use the following basic coloring algorithm to color the remainder of \(E(C)\): color the edges between \(v_i\) and \(v_1, \ldots, v_{i-1}\) with the same color \(i\). This always gives a coloring with no rainbow \(K_3\) (the color of the last vertex in each triangle appears on 2 of the edges in the triangle). However, the coloring 3-bounded only if \(d_i \leq 3\) for all \(i\). For example, the algorithm succeeds if \(C\) contains a copy of \(D\) or \(K_5 - e\), because here \(I_4 = I_5 = \emptyset\). We henceforth assume that \(C\) does not contain either of these graphs. We now describe how to modify this algorithm for the remaining cases. The availability of free colors (that is, colors used less than three times in this basic coloring) will help us in this task. For the sake of brevity, we will mention only the changes needed to fix this coloring.

**Case 1:** There exists an \(i\) such that \(d_i = 5\).

In this case we have \(d_j = 2\) for \(j > k, j \neq i\). If the back-neighbors of \(v_i\) are \(v_{i_1}, \ldots, v_{i_k}\), then we recolor each \(\{v_i, v_{i_s}\}\) with color \(i_s\). Any triangle formed by \(i\) and 2 of its back-neighbors can be expressed as \(v_i, v_{i_s}, v_{i_t}\) where \(i_s > i_t\), say. This triangle will then have two edges of color \(i_s\).

**Case 2:** There exists an index \(i\) such that \(d_i = 4\).

Let the back-neighbors of \(v_i\) be \(v_{i_1}, \ldots, v_{i_4}\) where \(i_1 < i_2 < i_3 < i_4\).

**Case 2a:** \(d_{i_4} = 2\).

Here, we use the color \(i_4\) for the edge \(\{v_{i_4}, v_i\}\).

Note that if \(C\) contains a copy of \(K_4\) then, assuming that \(d_i = 4\), we are in Case 2a (otherwise \(i_4 \leq k = 4\) and we have a copy of \(K_5\), a contradiction). Assume for the remaining sub-cases that \(C\) does not contain a copy of \(K_4\).

**Case 2b:** \(d_{i_4} = 3\).

We have

\[d_j = 2\] for \(j < i, j \neq i_4\).

Now we consider the graph \(X\) induced by \(\{v_{i_1}, \ldots, v_{i_4}\}\). By assumption \(X\) has at most 4 edges (otherwise we have a \(K_5\) or \(K_5 - e\)). Since \(C\) does not contain a copy of \(K_4\), \(X\) does not contain a triangle. We may assume that \(v_{i_4}\) is adjacent to \(v_{i_3}\); Otherwise we can just recolor the edge \(\{v_{i_3}, v_i\}\) with color \(i_3\). Also, we may assume that \(X\) has no isolated vertex \(v_p\): Otherwise \(p = i_1\) or \(i_2\), and we can recolor the edge \(\{v_p, v_i\}\) with color \(p\). Therefore, we can now restrict our attention to one of the following cases listed below:

**Case 2bi:** \(X\) is 2 disjoint edges.
One of the edges in $X$ is $\{v_{i_1}, v_{i_2}\}$. The color $i_2$ is a free color, so we can use it to recolor the edge $\{v_{i_1}, v_{i_2}\}$.

**Case 2bii:** $X$ is a path of length 3.

If $v_{i_2}$ is an endpoint connected to $v_{i_1}$, then we are done by recoloring $\{v_{i_2}, v_{i_1}\}$ with color $i_2$. Thus we can assume that our path is the union of two sub-paths going monotonely up and ending in $v_{i_4}$. (One of the sub-paths can be empty.) Take the longer sub-path, let it begin with edge $\{v_{i_6}, v_{i_5}\}$, $b < a \leq 3$. We recolor the edge $\{v_{i_b}, v_{i_1}\}$ with color $i_2$ (thus color $i_a$ forms a path of length 3 after the recoloring).

**Case 2biii:** $X$ is a 3-star.

Note that the center of the star cannot be $v_{i_4}$ (since the back-neighbors of $v_{i_4}$ must span an edge). Therefore, one of the edges in the star has a free color. Use this color on the edge from the corresponding leaf to $i$.

**Case 2biv:** $X$ is 4-cycle.

Let $v_p$ be the vertex not in $\{v_{i_1}, v_{i_2}, v_{i_3}\}$ that is a back-neighbor of $v_{i_4}$. Let $v_q$ be the neighbor of $v_{i_4}$ in $\{v_{i_1}, v_{i_2}\}$. Let $v_s$ be the other vertex in $\{v_{i_1}, v_{i_2}\}$. We have the following sub-cases.

- $v_p$ is not adjacent to $v_q$.
  The edge $\{v_{i_4}, v_q\}$ has no conditions on its color (relative to $v_{i_4}$) i.e. $v_{i_4}, v_p, v_q$ and $v_{i_4}, v_{i_3}, v_q$ do not form triangles. Using this observation we can proceed as follows. We recolor edge $\{v_{i_4}, v_q\}$ with color $i$. We then color edge $\{v_{i_4}, v_{i_3}\}$ with color $i_4$. We color edge $\{v_{i_4}, v_{i_3}\}$ with color $i_3$ and the edges $\{v_{i_1}, v_{i_4}\}$ keep color $i$.

- $v_p$ is adjacent to $v_q$ but not to $v_{i_3}$.
  We replace the color on $\{v_{i_4}, v_{i_4}\}$ with the color on $\{v_{i_4}, v_{i_3}\}$. Then use color $i_4$ on the other edges incident to $v_{i_4}$, including the edge to $v_{i_1}$.

- $v_p$ is adjacent to $v_q$ and to $v_{i_3}$.
  Note first that $p < i_3$ as otherwise the back neighbors of $v_p$ do not span an edge. Since $v_p$ and $v_s$ are the (only) back-neighbors of $v_{i_3}$ there must be an edge between them. Now we have a copy of $D$ with the deleted matching being $\{v_{i_3}, v_{i_3}, v_{i_4}\}$, $\{v_{i_4}, v_{i_4}\}$, $\{v_{i_1}, v_{i_4}\}$, contradiction.

This completes the proof of (3) and the first part of the proof of Theorem 3.

Suppose now that $c > \sqrt{2}$. Whp $G_{n,p}$ has $(1 + o(1))cn^{3/2}/2$ edges, $(1 + o(1))c^3n^{3/2}/6$ triangles and $o(n^{3/2})$ copies of $K_4$. Suppose that we have a 3-bounded coloring and $A_i$ is the set of colors that are used $i$ times and $a_i = |A_i|$ for $i = 1, 2, 3$. Thus,

$$a_1 + 2a_2 + 3a_3 = (1 + o(1))cn^{3/2}/2.$$  \hspace{1cm} (4)

Suppose that there are no rainbow triangles. Then each triangle $T$ contains a pair of edges of the same color $c(T)$. For color $x$ let $t(x)$ be the number of triangles $T$ such that $c(T) = x$. So
\( t(x) = 0 \) for \( x \in A_1 \), \( t(x) \leq 1 \) for \( x \in A_2 \) and \( t(x) \leq 2 \) for \( x \in A_3 \), unless \( x \) is used to color three edges of a copy of \( K_4 \). These latter colors are relatively rare (since the total number of \( K_4 \)-sub-graphs is \( o(n^{3/2}) \)) and so we have

\[
a_2 + 2a_3 \geq (1 + o(1))c^3 n^{3/2}/6. \tag{5}
\]

It follows from (4) and (5) that

\[
\frac{c^3}{4} \leq \frac{c}{2} \text{ or } c \leq \sqrt{2}.
\]

This contradiction completes the proof of Theorem 3.

**Remark.** The upper bound in Theorem 3 can be improved. For example we could remove from our accounting those edges that are not in triangles. Or we could note that isolated triangles (which *whp* form a non-negligible proportion of the triangles) must be accounted for by colors \( x \) such that \( t(x) = 1 \). While these arguments improve the upper bound, they do not completely close the gap between the upper and lower bounds in Theorem 3.

### 4 Proof of Theorem 1

#### 4.1 Small \( c \)

Let \( p = cn^{-1/m_H} \). We first consider the case where \( c \) is sufficiently small. We can assume in fact that

\[
m_H > m_{H'} \text{ for all } H' \subset H \text{ with } v_{H'} \geq 3.
\]

For if not, and \( m^*_H = m_{H'} \) for a sub-graph \( H' \) of \( H \) then we can show that it is possible to color \( G_{n,p} \) without creating a rainbow copy of \( H' \), which of course shows there is no rainbow copy of \( H \). It follows that if \( H' \subset H \) and \( v_{H'} \geq 3 \) then

\[
\delta_H' = \frac{e_H - e_{H'}}{m_H} - v_H + v_{H'} = (v_{H'} - 2) \left( 1 - \frac{m_{H'}}{m_H} \right) > 0.
\]

Define

\[
\delta_H = \min \{ \delta_H' : H' \subset H, v_{H'} > 2 \}.
\]

We follow a similar strategy to that in the previous section. In place of the triangle graph \( \Gamma \) we will have the \( H \)-graph \( \Gamma_H \) whose vertices are the copies of \( H \) in \( G_{n,p} \) and in which two vertices \( H_1, H_2 \) are joined by an edge in \( \Gamma_H \) if \( H_1, H_2 \) share at least one edge in \( G_{n,p} \).

A component \( C \) of \( \Gamma_H \) is **safe** if \( G(C) \) is \( b(H) \)-degenerate where we set

\[
b(H) = \Delta_H + m_H v_H - e_H + 1,
\]

and \( \Delta_H \) is the maximum degree in \( H \). Recall that \( G(C) \) is \( b(H) \)-degenerate if we can order \( V(C) = \{v_1, v_2, \ldots, v_\ell\}, \ell = |V(C)| \) such that each \( v_i \) has at most \( b(H) \) neighbors among \( v_1, v_2, \ldots, v_{i-1} \).

**Lemma 9** *Whp* every connected component \( C \) of \( \Gamma_H \) is safe.
Proof In analogy to the proof of Lemma 8, we consider a process where we choose a set of vertices \( V_0 = \{v_1, v_2, \ldots, v_H\} \), let \( E_0 \) consist of all edges spanned by \( V_0 \), and if \( E_0 \) contains a copy of \( H \), we do a search that generates an edge set \( E_{\text{final}} \) that contains \( E(C) \) where \( C \) is the corresponding component of \( \Gamma_H \). We generate sets \( V_i, E_i, i = 1, 2, \ldots, k \) via an iterative application of the following 2 steps until \( V_{k+1} = \emptyset \) after step A:

A. For each set of \( v_H \) vertices that contains some \( z \notin V^{(k)} \) \( \overset{\text{def}}{=} \bigcup_{i=0}^k V_i \) and some \( e \in E_k \) determine if this set of vertices gives a copy of \( H \). When we find such a copy of \( H \) we add \( V(H) \setminus V^{(k)} \) to \( V_{k+1} \), add \( E(H) \setminus E^{(k)} \) to \( E_{k+1} \) and move on. It is important to stress that once a vertex \( z \) is added to \( V_{k+1} \) we do not query any other vertex set that contains \( z \).

B. For each pair of vertices consisting of a vertex \( z \) in \( V_{k+1} \) and a vertex \( a \) in \( V^{(k+1)} \) see if the edge \( \{z, a\} \) is in \( G_{n, p} \). If so add this edge to \( E_{k+1} \).

As in the proof of Lemma 8, the conditioning on \( G_{n, p} \) imposed by this search is of a very special form. At any stage, certain edges are fully queried and we further condition on the event that certain sets of edges do not appear. Under any conditioning of this form, the probability that any set of \( k \) (not fully queried) edges appears in \( G_{n, p} \) is at most \( p^k \). Note further that after step B we have fully queried all edges within \( V^{(k+1)} \). Let \( E_{\text{final}} \) be the edge set generated when this process terminates.

Again, we view this as a branching process where the edges are individuals. Here we have three ways in which an edge \( e \in E_k \) can have offspring:

1. Copies of \( H \) found in step A such that \( V(H) \cap V^{(k)} = e \),
2. Copies of \( H \) found in step A such that \( e \in E(H) \) but \( |V(H) \cap V^{(k)}| \geq 3 \), and
3. Edges added during step B.

Of course, there is some ambiguity in assigning the paternity of edges of types 2 and 3. This is done arbitrarily. Let \( A_i, B_i \) and \( C_i \) be the number of type 1, 2 and 3 offspring, respectively, of the \( i \)th edge to join \( E_{\text{final}} \). For simplicity of our formulas, the edges that are in \( E_0 \) but not in the initial copy of \( H \), are accounted for by increasing appropriate \( C_i \)'s.

Let \( K = C \log n \), where \( C = C(c, H) \) is a sufficiently large constant. Let \( X_1, X_2, \ldots \) be i.i.d. \( Bi(n^{v_H - 2}, p^{e_H - 1}) \) random variables, let \( Y_1, Y_2, \ldots \) be i.i.d.

\[
\sum_{H' \subseteq H, v_{H'} \geq 3} Bi(K^{e_{H'} - 2} n^{v_H - v_{H'}}, p^{e_{H'} - e_{H'}})
\]

random variables, and let \( Z_1, Z_2, \ldots \) be i.i.d. \( Bi(2K, p) \) random variables. We see that \( A_i, B_i \) and \( C_i \) are dominated by \( (e_H - 1)X_i, (e_H - 1)Y_i \) and \( Z_i \), respectively, for \( i \leq K \). We have

\[
E \left[ \sum_{i=1}^{K} (e_H - 1)X_i \right] = K(e_H - 1)n^{v_H - 2}p^{e_H - 1} = K(e_H - 1)e^{e_H - 1}.
\]

So if \( c \) is sufficiently small the Chernoff bound implies

\[
Pr \left( \sum_{i=1}^{K} (e_H - 1)X_i \geq K - e_H - (m_H v_H - e_H + 1) - \frac{e_H - 1}{\delta_H} \left[ v_H - \frac{e_H}{m_H} + 1 \right] \right) = O \left( n^{e_H/m_H - v_H - 1} \right). \quad (6)
\]
The sum $\sum_{i=1}^{K} Y_i$ is distributed as
$$\sum_{H' \subset H, v_H \geq 3} Bi \left( K \cdot K^{v_H - 2} n^{v_H - v_{H'}}, p^{v_H - e_{H'}} \right).$$

Let $I$ denote the set of sequences of non-negative integers $(i_{H'} : H' \subset H, v_{H'} \geq 3)$ such that the sum of the $i_{H'}$'s is $\left\lceil \frac{1}{\delta_H} \left[ v_H - \frac{e_H}{m_H} + 1 \right] \right\rceil$. The probability that $\sum_{i=1}^{K} (e_H - 1)Y_i$ is at least $\frac{e_H - 1}{\delta_H} \left[ v_H - \frac{e_H}{m_H} + 1 \right]$ is bounded by

$$\sum_{(i_{H'}) \in I} \prod_{H' \subset H, v_{H'} \geq 3} \left( K \cdot K^{v_H - 2} n^{v_H - v_{H'}} \right) (p^{v_H - e_{H'}})^{i_{H'}} \leq K^{O(1)} \sum_{(i_{H'}) \in I} \prod_{H' \subset H, v_{H'} \geq 3} n^{-\delta_H i_{H'}} \leq K^{O(1)} \sum_{(i_{H'}) \in I} n^{e_H/m_H - v_H - 1}. $$

(Note that $\delta_H \leq \delta_{H'}$ for any $H' \subset H$ by definition.) Since there are $|I| = K^{O(1)}$ sequences we have

$$\Pr \left( \sum_{i=1}^{K} (e_H - 1)Y_i \geq \frac{e_H - 1}{\delta_H} \left[ v_H - \frac{e_H}{m_H} + 1 \right] \right) = K^{O(1)} n^{e_H/m_H - v_H - 1}. \quad (7)$$

Finally, we have

$$\Pr \left( \sum_{i=1}^{K} Z_i \geq m_H v_H - e_H + 1 \right) \leq \left( \frac{2K^2}{m_H v_H - e_H + 1} \right)^{m_H v_H - e_H + 1} \leq K^{O(1)} n^{e_H/m_H - v_H - 1/m_H}. \quad (8)$$

Since the expected number of the initial graphs $H$ is at most $n^{v_H - e_H/m_H}$, the union bound applied to (6), (7) and (8) shows that $\text{whp}$ every component of $\Gamma_H$ has at most $K$ edges. The desired $b(H)$-degenerate ordering then follows from (8).

\[\square\]

Of course, if every component of $\Gamma_H$ is safe and $b \geq b(H)$ then one can color the edges of $G$ so that there are no rainbow copies of $H$. To color $E(C)$ for a component $C$, we simply use the same new color for every edge from $v_i$ to $\{v_1, v_2, \ldots, v_{i-1}\}$ for $1 \leq i \leq |V(C)|$. Then every copy of $H$ in $C$ has a last vertex in the order and our coloring prevents this copy being rainbow. (Note that we use the fact that $H$ has minimum degree at least 2, which follows from the assumption $m_H > m_{H'}$ for all sub-graphs $H'$ and the inequality $m_H > 1$.)
4.2 Large $c$

We now consider the case where $c$ is sufficiently large and show that whp any $b$-bounded coloring contains a rainbow copy of $H$. As before, for any graph $K$ let $v_K$ be the number of vertices in $K$ and let $e_K$ be the number of edges in $K$. Let $X_K$ be the number of copies of $K$ in $G_{n,p}$. We present here some of the facts about the distribution of $X_K$ that we will need.

Lemma 10 If $p = cn^{-\alpha}$ and $v_L > \alpha e_L$ for all sub-graphs $L$ of $K$ then whp we have

$$X_K = (1 + o(1))E(X_K).$$

See, for example, Bollobás [3] for a proof and discussion of Lemma 10. This result will apply to our situation in view of the following easy observation.

Lemma 11 Let $H$ be an arbitrary graph with maximum degree $\Delta_H \geq 2$. Then for any non-empty sub-graph $L$ of $H$ we have $v_L > e_L/m_H^*.$

Proof If $\Delta_L = 1$ then $v_L \geq 2e_L > e_L/m_H^*$, so assume otherwise. (Note that $m_H^* \geq 1$ in view of $\Delta_H \geq 2$.) Also, it is enough to consider sub-graphs $L$ without isolated vertices. Then we have

$$m_H^* \geq m_L = \frac{e_L - 1}{v_L - 2} > \frac{e_L}{v_L},$$

as required. \hfill $\square$

We also use a theorem of Kim and Vu regarding extensions of a fixed graph. Let $H$ be a fixed graph and let $R \subset V(H)$. The extension from $R$ to $H$ is the graph $J$ with vertex set $V(H)$ and edge set $E(H) \setminus \binom{R}{2}$. Theorem 4.2.4 of [14] (which strengthens a theorem of Spencer [17]) gives a criterion for the number of copies of $J$ that extend from vertex set $I$ to be approximately equal to the expected number for every $I \in \binom{[n]}{|R|}$. For a fixed set of $I$ of $|R|$ vertices let $X_{I,J}$ be the number of copies of $J$ in $G_{n,p}$ for which the vertices in $I$ correspond to $R$.

Lemma 12 If $p = cn^{-\alpha}$ and $v_J - |R| > \alpha e_J$ for all sub-graphs $J'$ of $J$ that contain $R$ then

$$X_{I,J} = (1 + o(1))E[X_{I,J}]$$

for all $I \in \binom{[n]}{|R|}$.

Let $\mathcal{F}$ be the collection of induced sub-graphs $K$ of $H$ with $v_K \geq 3$ and no isolated vertices.

We will define a collection of parameters $\{\rho_B(c) : B \in \mathcal{F}\}$ such that $0 < \rho_B(c) < 1$ and

$$\lim_{c \to \infty} \rho_B(c) = 0 \quad (9)$$

for each $B \in \mathcal{F}$. We will prove by induction on the partially ordered set on $\mathcal{F}$ given by inclusion that for $p = cn^{-\alpha}, 0 < \alpha \leq 1/m_B^*$, whp at most $\rho_B K_B n^{v_B - \alpha e_B}$ copies of $B$ in $G_{n, cn^{-\alpha}}$ are not rainbow for all $B \in \mathcal{F}$.

One base case is a path $P_2$ of length 2. The number of non-rainbow copies of $P_2$ is at most \frac{|E_{n,p}|}{b} \binom{2}{2}$ where $E_{n,p}$ is the edge set of $G_{n,p}$. Note that whp

$$|E_{n,p}| \sim \frac{n^2}{2} p = \frac{c}{2} n^{2-\alpha} \leq \frac{c}{2} n^{3-2\alpha},$$

since $\alpha \leq 1$. Thus we can take $\rho_{P_2}(c) = 1/(c K_{P_2})$. The other base case, namely $K_3$ (if it is in $\mathcal{F}$), is done similarly, using the inequality $\alpha \leq 1/2$. 

13
4.2.1 Graphs $H$ with $m_H^* = m_H$

In this section we establish the existence of $\rho_H$ in the case when $m_H^* = m_H$. Let $p = cn^{-\alpha}$ where $0 < \alpha \leq 1/m_H^*$.

**Lemma 13** If $B \in \mathcal{F}$ then whp

\[ X_B \sim E[X_B] \sim K_Bc^B n^{v_B - \alpha e_B} \]

where $K_B$ is a constant that depends on $B$. Furthermore, if $D$ is a graph that consists of two copies of $B \in \mathcal{F}$ that intersect in some graph $C \in \mathcal{F}$ then whp we have

\[ X_D \sim E[X_D] \sim K_Dc^{2e_B - e_C} n^{2v_B - v_C - \alpha (2e_B - e_C)} \]

where $K_D$ is a constant that depends on $D$.

**Proof** The first part follows from Lemmas 10 and 11.

Let $D$ be a graph that consists of two copies of $B \in \mathcal{F}$ that intersect in the graph $C \in \mathcal{F}$. Let $L$ be some sub-graph of $D$.

We can view $E(L)$ as the union of $L_1 \cup L_2 \cup L_{1,2}$, where $L_i$ is the extension from $L_{1,2} \overset{\text{def}}{=} C \cap L$ to the intersection of $L$ with the $i$-th copy of $B$, $i = 1, 2$. Let $v_{1,2} = v(L_{1,2})$, $e_{1,2} = |E(L_{1,2})|$, $v_i = v(L_i) - v_{1,2}$, and $e_i = |E(L_i)|$, $i = 1, 2$. If $v_{1,2} > 0$, then we have

\[ \frac{e_L}{v_L} = \frac{e_{1,2} + e_1 + e_2}{v_{1,2} + v_1 + v_2}. \]  

(11)

By Lemma 11, we have

\[ \frac{e_{1,2}}{v_{1,2}} < m_H. \]  

(12)

The graph $F_i$ spanned by $V(C) \cup V(L_i)$ is obtained from $C$ by adding $v_i$ vertices and at least $e_i$ edges. We obtain

\[ \frac{e_C + e_i - 1}{v_C + v_i - 2} \leq m_{F_i} \leq m_H = \frac{e_C - 1}{v_C - 2}, \]

which implies that $e_i/v_i \leq m_H$. Applying (11) we obtain $e_L/v_L < m_H$. (The inequality is strict because (12) is strict.)

If $v_{1,2} = 0$, then we apply Lemma 11 to each $L_i \subset B$:

\[ \frac{e_L}{v_L} = \frac{e_1 + e_2}{v_1 + v_2} \leq \max \left( \frac{e_1}{v_1}, \frac{e_2}{v_2} \right) < m_H, \]

In any case, we always have $e_L/v_L < m_H \leq 1/\alpha$. The claim follows from Lemma 10.

Let $B$ be an arbitrary fixed graph in $\mathcal{F}$ and let $G$ be the collection of graphs in $\mathcal{F}$ that are proper sub-graphs of $B$.

Let a $b$-bounded coloring of the edges of $G_{n,p}$ be given. We say that a non-rainbow copy of $B$ is isolated if there is a pair of edges in $B$ that have the same color but are not in any other copy of $B$. We count the number of non-rainbow copies of $B$ as follows. First we bound the
number of copies of $B$ that are not rainbow by virtue of containing a non-rainbow copy of some graph $C \in \mathcal{G}$. It then remains to bound the number of isolated, non-rainbow copies of $B$. We do this by appealing to the bound we have on the number of edges in $G_{n,p}$.

Let $C \in \mathcal{G}$ and let $C_1, \ldots, C_m$ be the copies of $C$ that are not rainbow. Let $x_i$ be the number of copies of $B$ that contain $C_i$. Note that $C_i$ is the center of at most $\binom{x_i}{2}$ copies of the graph $D$ that consists of 2 copies of $B$ that intersect at $C$. Applying Jensen’s inequality and Lemma 13 we have

$$m \left( \frac{1}{m} \sum_{i=1}^{m} x_i \right) \leq \sum_{i=1}^{m} \left( \frac{x_i}{2} \right) < 2K_D e^{2e_B - \varepsilon_C} n^{2v_B - v_C - \alpha(2e_B - \varepsilon_C)}.$$  

Therefore (applying our inductive assumption $m \leq \rho_C K_C e^{e_C} n^{v_C - \alpha_C}$)

$$\sum_{i=1}^{m} x_i < 3\sqrt{\rho_C K_D K_C e^{e_B} n^{v_B - \alpha_B}}.$$  

Therefore, the number of non-isolated, non-rainbow copies of $B$ is at most

$$\left[ \sum_{C \in \mathcal{G}} \frac{3\sqrt{\rho_C K_D K_C}}{K_B} \right] K_B e^{e_B} n^{v_B - \alpha_B}.$$  

As there are at most

$$\frac{|E_{n,p}|}{b} \left( \frac{b}{2} \right) \sim c \frac{b - 1}{4} n^{v_B - \alpha_B - (1/m_B - \alpha)e_B}$$

isolated, non-rainbow copies of $B$, the number of non-rainbow copies of $B$ is at most

$$\left[ \sum_{C \in \mathcal{G}} \frac{3\sqrt{\rho_C K_D K_C}}{K_B} + \frac{b - 1}{4e^{e_B - 1} n^{1/m_B - \alpha}} \right] K_B e^{e_B} n^{v_B - \alpha_B} < \left[ \sum_{C \in \mathcal{G}} \frac{4\sqrt{\rho_C K_D K_C}}{K_B} \right] K_B e^{e_B} n^{v_B - \alpha_B}.$$  

This expression defines $\rho_B(c)$. Noting that $\rho_C(c)$ satisfies (9) for each $C \in \mathcal{G}$, we see that $\rho_B(c)$ also satisfies (9).

4.2.2 General $H$

Now we assume that $H$ is a graph with at least one cycle such that $m_H^* > m_H$. We still have $p = cn^{-\alpha}$ where $\alpha \leq 1/m_H^*$.

For each subset $X$ of $V(H)$ let $v_X = |X|$ and $e_X = |E(H[X])|$ and $m_X = m_H[X]$ where $H[X]$ denotes the sub-graph of $H$ induced by vertex set $X$. We say that a proper subset $X$ of $V(H)$ is thick if

$$m_X > m_Y$$

for all $X \subset Y \subset V(H)$. Note, for example, that if $X$ is a maximal set satisfying $m_H^* = m_X$ then $X$ is thick. For each $X \subset V(H)$ we let $J_X$ be the extension from $X$ to $H$.

Lemma 14 If $X$ is thick then whp

$$X_{I,J_X} \sim E[X_{I,J_X}] \sim K_{X,H} n^{v_H - v_X} p^{e_H - e_X}$$

for all $I \in \binom{[n]}{|X|}$. 

15
Proof Let $Y \subseteq V(H) \setminus X$. Let $e' = e_{X \cup Y} - e_X$. Then
\[
\frac{e_{X \cup Y} - 1}{v_X + v_Y - 2} < \frac{e_X - 1}{v_X - 2} \implies \frac{e'}{v_Y} = \frac{(e_{X \cup Y} - 1) - (e_X - 1)}{(v_X + v_Y - 2) - (v_X - 2)} < \frac{e_X - 1}{v_X - 2} = m_X \leq m_H^* \leq \frac{1}{\alpha}.
\]
Since this strict inequality is true for every $Y$, Lemma 12 implies the result. \hfill \square

We bound the number of non-rainbow copies of $H$ as follows. For each pair of edges $e, f \in E(H)$, we bound the number of copies of $H$ such that $e$ and $f$ have the same color by carefully choosing a vertex set $Y$ such that $e \cup f \subseteq Y \subset V(H)$, and bounding both the number of non-rainbow copies of $Y$ and the number of extensions from $Y$ to $H$. Note that $m^*_H[Y] \leq m^*_H \leq 1/\alpha$, so we can apply the inductive assumption to $H[Y]$.

If there is a thick set $Y$ such that $e, f \subseteq Y$ then, by induction and Lemma 14, the number of copies of $H$ that have $e$ and $f$ the same color is at most
\[
\rho_{H[Y]}(e) [K_H[Y]^{n^{v_Y} p^{e_Y}}] \cdot [K_Y^{-H} n^{v_Y} p^{e_Y} - e_Y] \rho_{H[Y]}(e) [K_H[Y]^* K_Y^{-H} n^{v_H} p^{e_H}]. \quad (13)
\]
Suppose now that $e, f$ are edges in $H$ such that $X = e \cup f$ has the property that for every $X \subseteq Z \subset V(H)$ the set $Z$ is not thick. Note that in this situation we have
\[
m_H^* > m_H \geq m_Y \quad \text{for all } X \subseteq Y \subset V(H), \quad (14)
\]
for otherwise a maximal $Z$ such that $Y \subseteq Z \subset V(H)$ and $m_Z > m_H$ would be thick.

We consider the additional condition
\[
\exists Y \supset X \text{ such that } \frac{v_Y - v_X}{e_Y - e_X} \leq \alpha. \quad (15)
\]
If (15) does not hold then we can apply Lemma 12 to the number of extensions from $X$ to $H$ and we can count as in (13).

Suppose (15) does hold. Let initially $Y_0 = X$ and $i = 0$. Repeat the following as long as possible: if there is a set $Y_{i+1} \supset Y_i$ such that
\[
\frac{v_{Y_{i+1}} - v_{Y_i}}{e_{Y_{i+1}} - e_{Y_i}} \leq \alpha, \quad (16)
\]
choose a minimal such set and increase $i$ by 1. Let $Y = Y_i$ be the final set. By our assumption $i \geq 1$.

Theorem 1.2 in Vu [18], see also [14], implies that \textbf{whp} for every $j \in [i]$ and every $|Y_{j-1}|$-subset $I$ of $G_{n,p}$ we have, say, $X_{I; H[Y_j]} \leq (\log n)^{|Y_j|}$, that is, the number of extensions from $H[Y_{j-1}]$ to $H[Y_j]$ is bounded by a polylogarithmic in $n$ function. We do not give the precise statement of the theorem, but remark that for this application it is enough to verify that for every $Z$ with $Y_{j-1} \subseteq Z \subset Y_j$, we have
\[
\frac{v_{Y_j} - v_Z}{e_{Y_j} - e_Z} \leq \alpha. \quad (16)
\]
The last inequality is true because otherwise, by (16) with $i = j - 1$, we would have $Z \neq Y_{j-1}$,
and $Z$ would contradict the minimality of $Y_j$.

Thus, every $|X|$-subset of $G_{n,p}$ admits at most polylog many extensions to $H[Y]$. Therefore, the number of non-rainbow copies of $H$ arising here is at most

$$O\left((n^2p) \cdot \text{polylog} \cdot n^{v_H-e_Y} p^{e_H-e_Y}\right).$$

(Note that the first term here is simply the bound on the number of pairs of edges with the same color). In order to demonstrate that only a vanishing fraction of the copies of $H$ are non-rainbow in this way, it is enough to show that

$$2 - \alpha + v_H - v_Y - \alpha(e_H - e_Y) < v_H - \alpha e_H .$$

The last inequality is equivalent to $\alpha < 1/m_Y$, which directly follows from $\alpha \leq 1/m_H^*$ and (14).

## 5 Trees

We first define the tree $B_{T,e,b}$ and prove that $B_{T,e,b} \in \mathcal{A}(b,T)$.

Let $e = \{x, y\}$ be an edge of the tree $T$. For each vertex $v$ in $T$ let $\ell_v$ be the distance from $v$ to $e$ (so $\ell_x = \ell_y = 0$). For each vertex $v$ in $T$ let $S_v$ be the set of all strings of the form $(v,i_1,i_2,\ldots,i_{\ell_v})$ where $i_1, i_2, \ldots, i_{\ell_v}$ are integers in the set $\{1,2,\ldots,b\}$. Note that we have $S_x = \{(x)\}$ and $S_y = \{(y)\}$. The vertex set of the $B_{T,e,b}$ is $\bigcup_{v \in V(T)} S_v$. In addition to the edge $\{(x),(y)\}$, we place an edge between vertex $(v,i_1,\ldots,i_{\ell_v})$ and $(w,j_1,\ldots,j_{\ell_w})$ if and only if

(a) $w$ and $v$ are adjacent in $T$, and

(b) $i_k = j_k$ for $k = 1,\ldots,\ell_v$ (where we assume $\ell_w = \ell_v + 1$).

We call the set of edges in $B_{T,e,b}$ between a vertex in $S_v$ and a set of vertices in $S_w$, where $\ell_w = \ell_v + 1$, a **bundle** of edges. We also let the singleton edge $\{(x),(y)\}$ form a bundle. Note that the edge set of $B_{T,e,b}$ is the disjoint union of the set $\mathcal{B}$ of bundles.

Let $\Omega$ be the set of colors in an arbitrary $b$-bounded coloring of $B_{T,e,b}$. For each bundle $B \in \mathcal{B}$ let $C_B$ be the set of colors used on the edges in $B$. Let $X \subseteq \mathcal{B}$. Since the coloring is $b$-bounded we have

$$\left| \bigcup_{B \in X} C_B \right| \geq \frac{1}{b} \sum_{B \in X} |B| \geq \frac{(|X|-1)b+1}{b} .$$

Since the cardinality of this union is an integer, it is at least $|X|$. So, by Hall’s Theorem, there is a system of distinct representatives of the sets $\{C_X : X \in \mathcal{B}\}$.

This system of distinct representatives corresponds to a set $Y$ of edges in $B_{T,e,b}$ such that there is exactly one edge from each bundle in $Y$ and the colors on the edges in $Y$ are all different. This set of edges defines a rainbow copy of $T$ (as well as some extra components) and shows that $B_{T,e,b} \in \mathcal{A}(b,T)$.

### 5.1 Special Cases: Proof of Theorem 5

We begin by showing that for the path $P_l$ with $l$ edges we have

$$s(b, P_l) = \begin{cases} 
(b + 1) \sum_{i=0}^{m-1} b^i & \text{if } l = 2m \\
1 + 2 \sum_{i=1}^{m} b^i & \text{if } l = 2m + 1.
\end{cases} \quad (17)$$
Observe first that since the \( b \)-blow up of \( P_l \) centered on the edge \( e \) at the middle of the path is in \( A(b, P_l) \), the above expression is an upper bound on \( s(b, P_l) \).

For the lower bound we use induction on \( l \) with cases \( l = 1, 2 \) being trivial. Let a tree \( U \) give a rainbow \( P_l \) for every \( b \)-bounded coloring of \( U \). Partition \( E(U) \) as \( X \cup F_1 \cup \cdots \cup F_k \) so that

(i) for each \( 1 \leq i \leq k \), \( |F_i| \leq b \),

(ii) for each \( 1 \leq i \leq k \) and every path \( (x_1, x_2, x_3, x_4) \) in \( H \), if the edge \( \{x_2, x_3\} \) belongs to \( F_i \), then \( F_i \) contains the edge \( \{x_1, x_2\} \) or the edge \( \{x_3, x_4\} \) (in other words each \( F_i \) consists of all edges in \( U \) that intersect some set of vertices), and

(iii) \( |X| \) is as small as possible (given (i) and (ii)).

Note that every \( b \)-bounded coloring of the forest \( X \) yields a rainbow \( P_{l-2} \); otherwise, we color the forest \( X \) with no rainbow \( P_{l-2} \) and each \( F_i \) with its own color to give a coloring of \( U \) with no rainbow \( P_l \). Thus for some component \( Y \) of \( X \) we have \( |E(Y)| \geq s(b, P_{l-2}) \). By induction \( |E(Y)| \) is bounded below by the expression in (17).

In order to count the edges in \( U \), we assign the other components of \( X \) and the parts \( F_i \) to vertices of \( Y \) according to their vertex of attachment. (I.e. the vertex \( z \) and its incident edges are assigned to \( y \in Y \) if the path from \( z \) to \( y \) is edge disjoint from \( E(Y) \).) We claim that for each vertex \( y \in Y \) of \( Y \)-degree \( d \) there are at least \( b-d+1 \) edges attached to \( Y \) in this way. Indeed, if this is not true, then form a new \( F_i \)-set by putting together all edges of \( Y \) incident to \( y \), plus all parts attached to \( y \). The new \( F_i \) has at most \( b \) edges and \( |X| \) has strictly decreased.

Take any path \( (x_1, x_2, x_3, x_4) \) with the edge \( \{x_2, x_3\} \in F_i \). If \( y \notin \{x_2, x_3\} \), then both edges \( \{x_1, x_2\} \) and \( \{x_3, x_4\} \) are in \( F_i \). If, say, \( y = x_2 \) then the edge \( \{x_1, x_2\} \in F_i \). The claim has been proved.

For \( x \in V(Y) \) let \( d_Y(x) \) be the \( Y \)-degree and \( f_Y(x) \) the aggregate number of edges of \( E(U \setminus Y) \) assigned to \( x \). We have

\[
\sum_{x \in V(Y)} (d_Y(x) + f_Y(x)) \geq (b + 1)|V(Y)|.
\]

But this sum equals \( 2|E(Y)| + |E(U) \setminus E(Y)| = |E(U)| + |E(Y)| \). So,

\[
|E(U)| \geq (b + 1)|V(Y)| - |E(Y)| = b|E(Y)| + b + 1 \geq b \cdot s(b, P_{l-2}) + b + 1,
\]

as required to complete the proof of (17).

Now we turn to part (b) of Theorem 5. We will show that

\[
s(b, T_{d,l}) = 1 + 2 \sum_{i=1}^{l-1} (b(d-1))^i + (b(d-1))^l \tag{18}
\]

where \( T_{d,l} \) is a rooted tree, with all leaves at distance \( l \) from the root such that every non-leaf has the same degree \( d \).

Observe first that the tree \( B_{T_{d,l}, e} \) with \( e \) being any edge incident with the root shows that our expression is an upper bound for \( s(b, T_{d,l}) \).

For the lower bound we again proceed by induction on \( l \). The case \( l = 1 \) is simple: a tree with at most \( b(d-1) \) edges can be colored using only \( d-1 \) colors. Let \( l \geq 2 \) and let \( U \) be a tree with
the coloring property (i.e. \( U \in A(b, T_{d,l}) \)). We again grow classes \( F_i \) as in the proof of (17) but this time the restriction on their size is \((d - 1)b\) (to be precise, we partition \( U \) into \( X, F_1, \ldots, F_k \) such that we have \(|F_i| \leq (d - 1)b\), (ii) and (iii)). Note that every \( b \)-bounded coloring of the forest \( X \) yields a rainbow \( T_{d,l-1} \); otherwise, we color the forest \( X \) with no rainbow \( T_{d,l-1} \) and each \( F_i \) with its own set of \( d - 1 \) colors to give a coloring of \( U \) with no rainbow \( T_{d,l} \). Therefore, \( X \) has a component \( Y \) such that \(|E(Y)| \geq s(b, T_{d,l-1})\), which is equal to the expression in (18) by induction. We again assign the other components of \( X \) and the parts \( F_i \) to vertices of \( Y \) according to their vertex of attachment and let \( d_Y(x) \) be the \( Y \)-degree and \( f_Y(x) \) the aggregate number of edges of \( E(U \setminus Y) \) assigned to \( x \). We have

\[ |E(U)| + |E(Y)| = \sum_{x \in V(Y)} (d_Y(x) + f_Y(x)) \geq ((d - 1)b + 1)|V(Y)|. \]

But then

\[ |E(U)| \geq ((d - 1)b + 1)|V(Y)| - |E(Y)| = (d - 1)b|E(Y)| + (d - 1)b + 1, \]

giving the required lower bound.

### 5.2 Proof of Theorem 4

For the upper bound, consider a \( b \)-blow-up of \( T \) centered on an edge \( e \) that is in the middle of a longest path in \( T \). The upper bound follows from the fact that this blow-up is in \( A(b, T) \) and has \( O(b^m) \) vertices.

For the lower bound it is enough to note that if a tree \( H \) is a sub-graph of \( T \) then \( s(b, H) \leq s(b, T) \). Since \( T \) contains the path \( P_l \) and \( s(b, P_l) = \Omega(b^m) \), we have the desired lower bound.

### References


