Between 2- and 3-colorability

Alan Frieze* and Wesley Pegden†

Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA 15213, U.S.A

alan@random.math.cmu.edu, wes@math.cmu.edu

Submitted: Sep 10, 2014; Accepted: Jan 14, 2015; Published: Feb 16, 2015
Mathematics Subject Classifications: 05C80; 05C15

Abstract

We consider the question of the existence of homomorphisms between $G_{n,p}$ and odd cycles when $p = c/n$, $1 < c \leqslant 4$. We show that for any positive integer ℓ, there exists $\varepsilon = \varepsilon(\ell)$ such that if $c = 1 + \varepsilon$ then w.h.p. $G_{n,p}$ has a homomorphism from $G_{n,p}$ to $C_{2\ell+1}$ so long as its odd-girth is at least $2\ell + 1$. On the other hand, we show that if $c = 4$ then w.h.p. there is no homomorphism from $G_{n,p}$ to C_5. Note that in our range of interest, $\chi(G_{n,p}) = 3$ w.h.p., implying that there is a homomorphism from $G_{n,p}$ to C_3. These results imply the existence of random graphs with circular chromatic numbers χ_c satisfying $2 < \chi_c(G) < 2 + \delta$ for arbitrarily small δ, and also that $2.5 \leqslant \chi_c(G_{n,\frac{4}{n}}) < 3$ w.h.p.

1 Introduction

The determination of the chromatic number of $G_{n,p}$, where $p = \frac{c}{n}$ for constant c, is a central topic in the theory of random graphs. For $0 < c < 1$, such graphs contain, in expectation, a bounded number of cycles, and are almost-surely 3-colorable. The chromatic number of such a graph may be 2 or 3 with positive probability, according as to whether or not any odd cycles appear.

For $c \geqslant 1$, we find that the chromatic number $\chi(G_{n,\frac{c}{n}}) \geqslant 3$ with high probability, see for example Bollobás [6] or Janson, Łuczak and Ruciński [9]. Letting $c_k := \sup_c \chi(G_{n,\frac{c}{n}}) \leqslant k$,

*Research supported in part by NSF grant ccf1013110
†Research supported in part by NSF grant dms1363136

it is known for all k and $c \in (c_k, c_{k+1})$ that $\chi(G_{n,\frac{c}{n}}) \in \{k, k+1\}$, see Łuczak [10] and Achlioptas and Naor [3]; for $k > 2$, the chromatic number may well be concentrated on the single value k, see Friedgut [7] and Achlioptas and Friedgut [1].

In this paper, we consider finer notions of colorability for the graphs $G_{n,\frac{c}{n}}$ for $c \in (1, c_3)$, by considering homomorphisms from $G_{n,\frac{c}{n}}$ to odd cycles $C_{2\ell+1}$. Recall that a graph homomorphism from G to H is a function $f : V(G) \to V(H)$ such that $u \sim v$ implies $f(u) \sim f(v)$[1]. In particular, the existence of a homomorphism from a graph G to $C_{2\ell+1}$ implies the existence of homomorphisms to C_{2k+1} for all $k < \ell$. As the 3-colorability of a graph G corresponds to the existence of a homomorphism from G to K_3, the existence of a homomorphism to $C_{2\ell+1}$ implies 3-colorability. Thus considering homomorphisms to odd cycles $C_{2\ell+1}$ gives a hierarchy of 3-colorable graphs amenable to increasingly stronger constraint satisfaction problems. Note that a fixed graph having a homomorphism to all odd-cycles is bipartite.

Our main result is the following:

Theorem 1. *For any integer $\ell > 1$, there is an $\varepsilon > 0$ such that with high probability, $G_{n,\frac{1+\varepsilon}{n}}$ either has odd-girth $< 2\ell + 1$ or has a homomorphism to $C_{2\ell+1}$.*

Conversely, we expect the following:

Conjecture 1. *For any $c > 1$, there is an ℓ_c such that with high probability, there is no homomorphism from $G_{n,\frac{c}{n}}$ to $C_{2\ell+1}$ for $\ell \geqslant \ell_c$.*

As c_3 is known to be at least 4.03 [2], the following confirms Conjecture 1 for a significant portion of the interval $(1, c_3)$.

Theorem 2. *For any $c > 2.774$, there is an ℓ_c such that with high probability, there is no homomorphism from $G_{n,\frac{c}{n}}$ to to $C_{2\ell+1}$ for $\ell \geqslant \ell_c$.*

We also have that $\ell_4 = 2$:

Theorem 3. *With high probability, $G_{n,\frac{4}{n}}$ has no homomorphism to C_5.*

Note that as $c_3 > 4.03 > 4$, see Achlioptas and Moore [2], we see that there are triangle-free 3-colorable random graphs without homomorphisms to C_5. Our proof of Theorem 3 involves computer assisted numerical computations. The same calculations which rigorously demonstrate that $\ell_4 = 2$ suggest actually that $\ell_{3.75} = 2$ as well.

Our results can be reformulated in terms of the *circular chromatic number* of a random graph. Recall that the circular chromatic number $\chi_c(G)$ of G is the infimum r of circumferences of circles C for which there is an assignment of open unit intervals of C

[1] For a graph $G = (V, E)$ and $a, b \in V$, we write $a \sim b$ to mean that $\{a, b\} \in E$

to the vertices of G such that adjacent vertices are assigned disjoint intervals. (Note that if circles C of circumference r were replaced in this definition with line segments S of length r, then this would give the ordinary chromatic number $\chi(G)$.) It is known that $\chi(G) - 1 < \chi_c(G) \leqslant \chi(G)$, that $\chi_c(G)$ is always rational, and moreover, that $\chi_c(G) \leqslant \frac{p}{q}$ if and only if G has a homomorphism to the circulant graph $C_{p,q}$ with vertex set $\{0, 1, \ldots, q-1\}$, with $v \sim u$ whenever $\operatorname{dist}(v, u) := \min\{|v-u|, v+q-u, u+q-v\} \geqslant q$. (See [12].) Since $C_{2\ell+1,\ell}$ is the odd cycle $C_{2\ell+1}$ our results can be restated as follows:

Theorem 4. *In the following, inequalities for the circular chromatic number hold with high probability.*

1. *For any $\delta > 0$, there is an $\varepsilon > 0$ such that, $G = G_{n,\frac{1+\varepsilon}{n}}$ has $\chi_c(G) \leqslant 2 + \delta$ unless it has odd girth $\leqslant \frac{2}{\delta}$.*
2. *For any $c > 2.774$, there exists $r > 2$ such that $\chi_c(G_{n,\frac{c}{n}}) > r$.*
3. $2.5 \leqslant \chi_c(G_{n,\frac{4}{n}}) < 3$.

Note that for any c and $\ell > 1$, there is positive probability that $G_{n,\frac{c}{n}}$ has odd girth $< 2\ell+1$, and a positive probability that it does not. In particular, as the probability that $G_{n,\frac{c}{n}}$ has small odd-girth can be computed precisely, Theorem 1 gives an exact probability in $(0, 1)$ that $G_{n,\frac{1+\varepsilon}{n}}$ has a homomorphism to $C_{2\ell+1}$. Indeed, Theorem 1 implies that if $c = 1 + \varepsilon$ and ε is sufficiently small relative to ℓ, then

$$\lim_{n\to\infty} \mathbf{Pr}(\chi_c(G_{n,\frac{c}{n}}) \in (2 + \tfrac{1}{\ell+1}, 2 + \tfrac{1}{\ell}]) = e^{-\phi_\ell(c)} - e^{-\phi_{\ell+1}(c)}, \tag{1}$$

where

$$\phi_\ell(c) = \sum_{i=1}^{\ell-1} \frac{c^{2i+1}}{2(2i+1)}.$$

We close with two more conjectures. The first concerns a sort of pseudo-threshold for having a homomorphism to $C_{2\ell+1}$:

Conjecture 2. *For any ℓ, there is a $c_\ell > 1$ such that $G_{n,\frac{c}{n}}$ has no homomorphism to $C_{2\ell+1}$ for $c > c_\ell$, and has either odd-girth $< 2\ell + 1$ or has a homomorphism to $C_{2\ell+1}$ for $c < c_\ell$.*

The second asserts that the circular chromatic numbers of random graphs should be dense.

Conjecture 3. *There are no real numbers $2 \leqslant a < b$ with the property that for any value of c, $\mathbf{Pr}(\chi_c(G_{n,\frac{c}{n}}) \in (a, b)) \to 0$.*

Note that our Theorem 1 confirms this conjecture for the case $a = 2$.

2 Structure of the paper

We prove Theorem 1 in Section 3. We first prove some structural lemmas and then we show, given the properties in these lemmas, that we can algorithmically find a homomorphism. We prove Theorem 2 in Section 4 by the use of a simple first moment argument. We prove Theorem 3 in Section 5. This is again a first moment calculation, but it has required numerical assistance in its proof.

3 Finding homomorphisms

Lemma 1. *If $\alpha < 1/10$ and c is a positive constant where*

$$c < c_0 = \exp\left\{\frac{1-6\alpha}{3\alpha}\right\}$$

then w.h.p. any two cycles of length less than $\alpha \log n$ in $G_{n,p}$, $p = \frac{c}{n}$, are at distance more than $\alpha \log n$.

Proof If there are two cycles contradicting the above claim, then there exists a set S of size $s \leqslant 3\alpha \log n$ that contains at least $s+1$ edges. The expected number of such sets can be bounded as follows:

$$\begin{aligned}\sum_{s=4}^{3\alpha\log n} \binom{n}{s}\binom{\binom{s}{2}}{s+1}\left(\frac{c}{n}\right)^{s+1} &\leqslant \sum_{s=4}^{3\alpha\log n}\left(\frac{ne}{s}\right)^{s}\left(\frac{se}{2}\right)^{s+1}\left(\frac{c}{n}\right)^{s+1} \\ &\leqslant \frac{3c\alpha\log n}{n}\sum_{s=4}^{3\alpha\log n}\left(\frac{ce^2}{2}\right)^{s} \\ &< \frac{(ce^2)^{3\alpha\log n}\log n}{n},\end{aligned}$$

which tends to 0 for our choices of α, c. □

Our next lemma is concerned with cycles in K_2 which is the *2-core* of $G_{n,p}$. The 2-core of a graph is the graph induced by the edges that are in at least one cycle. When $c > 1$, the 2-core consists of a linear size sub-graph together with a few vertex disjoint cycles. By few we mean that in expectation, there are $O(1)$ vertices on these cycles.

Let $0 < x < 1$ be such that $xe^{-x} = ce^{-c}$. Then w.h.p. K_2 has

$$\nu \approx (1-x)\left(1-\frac{x}{c}\right)n \text{ vertices and } \mu \approx \left(1-\frac{x}{c}\right)^2 \frac{cn}{2} \text{ edges.}$$

(See for example Pittel [11]).

If $c = 1+\varepsilon$ for ε small and positive then $x = 1-\eta$ where $\eta = \varepsilon + a_1\varepsilon^2$, $|a_1| \leqslant 2$ for $\varepsilon < 1/10$.

The degree sequence of K_2 can be generated as follows, see for example Aronson, Frieze and Pittel [4]: Let λ be the solution to

$$\frac{\lambda(e^\lambda - 1)}{e^\lambda - 1 - \lambda} = \frac{2\mu}{\nu} \approx \frac{c-x}{1-x} = \frac{2+a_1\varepsilon}{1+a_1\varepsilon}.$$

We deduce from this that

$$\lambda \leqslant 4|a_1|\varepsilon \leqslant 8\varepsilon.$$

We will let $Z_1, \ldots, Z_n$ denote independent copies of the random variable Z where for $d \geqslant 2$,

$$\mathbf{Pr}(Z = d) = \frac{\lambda^d}{d!(e^\lambda - 1 - \lambda)}. \tag{2}$$

It is shown in [4] that, conditioned on the event that $D_1 := \sum d(i) = 2\mu$, we have that the degrees $d(1), d(2), \ldots, d(n)$ of K_2 are distributed as the $Z_1, Z_2, \ldots, Z_n$'s. Thus we will make use of the factor

$$\begin{aligned}\theta_k &= \frac{\mathbf{Pr}(d(i) = d_i, i = 1, 2, \ldots, k \mid D_1 = 2\mu)}{\mathbf{Pr}(Z_i = d_i, i = 1, 2, \ldots, k)} \\ &= \frac{\mathbf{Pr}(Z_{k+1} + \cdots + Z_n = 2\mu - (Z_1 + \cdots + Z_k))}{\mathbf{Pr}(Z_1 + \cdots + Z_n = 2\mu)}.\end{aligned}$$

It is shown in [4] that if $Z_1, Z_2, \ldots, Z_N$ are independent copies of Z then

$$\mathbf{Pr}(Z_1 + \cdots + Z_N = N\,\mathbf{E}(Z) - t) = \frac{1}{\sigma\sqrt{2\pi N}}\left(1 + O\left(\frac{t^2+1}{N\sigma^2}\right)\right) \tag{3}$$

where $\sigma^2 = \Theta(1)$ is the variance of Z.

We observe next that the maximum degree in $G_{n,p}$ and hence in K_2 is q.s.[2] at most $\log n$. It follows from this and (3) that

$$\theta_k = 1 + o(1) \text{ for } k \leqslant \log^2 n \text{ and } \theta_k = O(n^{1/2}) \text{ in general.}$$

[2] A sequence of events $\mathcal{E}_n$ is said to occur *quite surely* q.s. if $\mathbf{Pr}(\neg\mathcal{E}_n) = O(n^{-C})$ for any constant $C > 0$.

Lemma 2. *For any α, β, there exists $c_0 > 1$ such that w.h.p. any cycle C of length greater than $\alpha \log n$ in the 2-core of $G_{n,p}$, $p = \frac{c}{n}$, $1 < c < c_0$, has at most $\beta|C|$ vertices of degree $\geqslant 3$.*

Proof Suppose that

$$e^{1+8\varepsilon} \left(\frac{8\varepsilon e}{\beta} \right)^{\beta} < 1.$$

We will show then that w.h.p. K_2 does not contain a cycle C where (i) $|C| \geqslant \alpha \log n$ and (ii) C contains $\beta|C|$ vertices of degree greater than two.

We can bound the probability of the existence of a "bad" cycle C as follows: There are $\binom{\nu}{k}\frac{(k-1)!}{2}$ possible cycles in K_2. In particular, there are $\binom{\nu}{k}\frac{(k-1)!}{2}\binom{k}{\beta k}$ choices of a cycle of length k, and βk (lexicographically first, say) vertices $v_1, \ldots, v_{\beta k}$ on the cycle which have degree at least 3. We then sum over the choices $d(i), i = 1, \ldots, k$ of the degrees of the vertices on the cycle. The probability that $d(i) = d_i$ for $k = 1, \ldots, k$ is given by $\theta_k \prod_{i=1}^{k} \mathbf{Pr}(Z_i = d_i, i = 1, 2, \ldots, k)$. Given this, we switch to the configuration model of Bollobás [5] for a random graph with a fixed degree sequence. In this model, the probability that the edges of the cycle exist is $\prod_{i=1}^{k} \frac{d_i(d_i-1)}{2\mu-2k+1}$. Using the configuration model, we inflate our estimates by a constant factor C_0 to handle the problem of loops and multiple edges. Thus the probability of such a bad cycle can be bounded by

$$\begin{aligned}
\mathbf{Pr}(\exists C) &\leqslant C_0 \sum_{k=\alpha\log n}^{\nu} \binom{\nu}{k} \frac{(k-1)!}{2} \binom{k}{\beta k} \theta_k \sum_{\substack{d_1,\ldots,d_{\beta k} \geqslant 3 \\ d_{\beta k+1},\ldots,d_k \geqslant 2}} \prod_{i=1}^{k} \left(\frac{\lambda^{d_i}}{d_i!(e^\lambda - 1 - \lambda)} \cdot \frac{d_i(d_i-1)}{2\mu - 2k + 1} \right) \\
&\leqslant C_0 \sum_{k=\alpha\log n}^{\nu} \frac{1}{2k} \left(\frac{\nu}{(2\mu - 2k)(e^\lambda - 1 - \lambda)} \right)^k \lambda^{2k} \binom{k}{\beta k} \theta_k \sum_{\substack{d_1,\ldots,d_{\beta k} \geqslant 3 \\ d_{\beta k+1},\ldots,d_k \geqslant 2}} \prod_{i=1}^{k} \frac{1}{(d_i - 2)!} \\
&\leqslant C_0 \sum_{k=\alpha\log n}^{\nu} \frac{e^{k^2/\mu}}{2k} \left(\frac{\nu}{2\mu(e^\lambda - 1 - \lambda)} \right)^k \lambda^{2k} \binom{k}{\beta k} \theta_k (e^\lambda - 1)^{\beta k} e^{(1-\beta)k\lambda} \\
&= C_0 \sum_{k=\alpha\log n}^{\nu} \frac{e^{k^2/\mu}}{2k} \left(\frac{\lambda}{e^\lambda - 1} \right)^k \binom{k}{\beta k} \theta_k (e^\lambda - 1)^{\beta k} e^{(1-\beta)k\lambda} \\
&\leqslant C_0 \sum_{k=\alpha\log n}^{\nu} \frac{\theta_k}{2k} \left(e^{k/\mu} \cdot \frac{\lambda}{(e^\lambda - 1)^{1-\beta}} \cdot \left(\frac{e}{\beta} \right)^{\beta} \cdot e^{(1-\beta)\lambda} \right)^k \\
&\leqslant C_0 \sum_{k=\alpha\log n}^{\nu} \frac{\theta_k}{2k} \left(e \cdot \lambda^{\beta} \cdot \left(\frac{e}{\beta} \right)^{\beta} \cdot e^{\lambda} \right)^k,
\end{aligned}$$

which tends to 0. □

Lemma 3. *For any $k \in \mathbb{N}$, there exists $\varepsilon_0 > 0$ such that w.h.p. we can decompose the edges of the $G = G_{n,p}$, $p = \frac{1+\varepsilon}{n}$, $0 < \varepsilon < \varepsilon_0$, as $F \cup M$, where F is a forest, and where the distance in F between any two edges in M is at least k.*

Proof We fix some $\alpha < \frac{1}{10}$. By choosing $\beta < \frac{1}{2k}$ in Lemma 2 we can find, in every cycle of length $> \alpha \log n$ of the 2-core K_2 of G (which includes all cycles of G), a path of length at least $2k+1$ whose interior vertices are all of degree 2. We can thus choose in each cycle of K_2 of length $> \alpha \log n$ such a path of maximum length, and let $\mathcal{P}$ denote the set of such paths. (Note that, in general, there will be fewer paths in $\mathcal{P}$ than long cycles in K_2 due to duplicates, but that the elements of $\mathcal{P}$ are nevertheless disjoint paths in K_2.) We now choose from each path in $\mathcal{P}$ an edge from the center of the path to give a set M_1. Note that the set of cycles in $G \setminus M_1$ is the same as the set of cycles in $G \setminus \bigcup_{P \in \mathcal{P}} P$. (In particular, the only cycles which remain have length $\leqslant \alpha \log n$ and are at distance $\geqslant k$ from M.) Thus, letting M_2 consist of one edge from each cycle of $G \setminus M_1$, Lemma 1 implies that $M = M_1 \cup M_2$ is as desired. □

Proof of Theorem 1. Our goal in this section is to give a $C_{2\ell+1}$-coloring of $G = G_{n,\frac{1+\varepsilon}{n}}$ for $\varepsilon > 0$ sufficiently small. By this we will mean an assignment $c : V(G) \to \{0, 1, \ldots, 2\ell\}$ such that $x \sim y$ in G implies that $c(x) \sim c(y)$ as vertices of $C_{2\ell+1}$; that is, that $x = y \pm 1 \pmod{2\ell+1}$.

Consider a decomposition of G as $F \cup M$ as given by Lemma 3, with $k = 4\ell - 2$.

We begin by 2-coloring F. Let $c_F : V \to \{0, 1\}$ be such a coloring. Our goal will be to modify this coloring to give a good $C_{2\ell+1}$ coloring of G.

Let $\mathcal{B}$ be the set of edges $xy \in M$ for which $c_F(x) = c_F(y)$, and let B be a set of distinct representatives for $\mathcal{B}$, and for $i = 0, 1$, let $B^i = \{v \in B \mid c_F(v) = i\}$.

We now define a new $C_{2\ell+1}$ coloring $c : V \to \{0, 1, \ldots, 2\ell\}$, by

$$c(v) = \begin{cases} c_F(v) & \text{if } \mathrm{dist}_F(v, B) \geqslant 2\ell - 1 \\ c_F(x) - (-1)^j(\mathrm{dist}_F(x, v) + 1) & \text{if } \exists x \in B^j \text{ s.t. } \mathrm{dist}(x, v)_F < 2\ell - 1. \end{cases} \tag{4}$$

(Color addition and subtraction are computed modulo $2\ell + 1$.)

Since edges in M are separated by distances $\geqslant 4\ell - 2$, this coloring is well-defined (i.e., there is at most one choice for x). Moreover, c is certainly a good $C_{2\ell+1}$-coloring of F. Thus if c is a not a good $C_{2\ell+1}$-coloring of G, it is bad along some edge $xy \in M$. But if such an edge was already properly colored in the 2-coloring c_F, it is still properly colored by c, since it has distance $\geqslant 4\ell - 2 \geqslant 2\ell - 1$ from other edges in M. On the other hand, if previously we had $c_F(x) = c_F(y) = i$, and WLOG $x \in B^i$, then the definition of $c(v)$ gives that we now have that $c(x) \in \{i-1, i+1\}$ (modulo $2\ell - 1$). Thus if c is not a good $C_{2\ell+1}$-coloring of G, then there is an edge $xy \in M$ such that $x \in B^i$ and y's color also changes in the coloring c; but by the distance between edges in M, this can only happen if x and y are at F-distance $< 2\ell - 1$. Note also that $c_F(x) = c_F(y)$ implies that $\mathrm{dist}_F(x, y)$ is even. Thus in this case, $F \cup \{xy\}$ contains an odd cycle of length $\leqslant 2\ell - 1$, and so G has odd girth $< 2\ell + 1$, as desired. □

4 Avoiding homomorphisms to long odd cycles

For large ℓ, one can prove the non-existence of homomorphisms to $C_{2\ell+1}$ using the following simple observation:

Observation 4. *If G has a homomorphism to $C_{2\ell+1}$, then G has an induced bipartite subgraph with at least $\frac{2\ell}{2\ell+1}|V(G)|$ vertices.*

Proof. Delete the smallest color class. □

Proof of Theorem 2. The probability that $G_{n,\frac{c}{n}}$ has an induced bipartite subgraph on βn vertices is at most

$$\binom{n}{\beta n} 2^{\beta n} \left(1 - \frac{c}{n}\right)^{\beta^2 n^2/4} < \left(\frac{2^\beta e^{-c\beta^2/4}}{\beta^\beta (1-\beta)^{1-\beta}}\right)^n \tag{5}$$

The expression inside the parentheses is unimodal in β for fixed c, and, for $c > 2.774$, is less than 1 for $\beta > .999971$. In particular, for $c > 2.774$, $G_{n,\frac{c}{n}}$ has no homomorphism to $C_{2\ell+1}$ for $2\ell + 1 \geqslant 1,427,583$. □

5 Avoiding homomorphisms to C_5

A homomorphism of $G = G_{n,p}, p = \frac{c}{n}$ into C_5 induces a partition of $[n]$ into sets $V_i, i = 0, 1, \ldots, 4$. As explained momentarily, this partition can be chosen so that the following hold:

P1 The sets $V_i, i = 0, 1, \ldots, 4$ are all independent sets.

P2 There are no edges between V_i and $V_{i+2} \cup V_{i-2}$. Here addition and subtraction in an index are taken to be modulo 5.

P3 Every $v \in V_i, i = 1, 2, 3, 4$ has a neighbor in V_{i-1}.

P4 Every $v \in V_2$ has a neighbor in V_3.

Conditions **P1,P2** are what is required for a homomorphism to C_5. For **P3** we observe that if $v \in V_i$ has no neighbor in V_{i-1} then we can move v to V_{i+2} and still maintain **P1,P2**. As argued in Hatami [8], Lemma 2.1, applying this repeatedly eventually leads to **P1,P2,P3** holding. Given **P1,P2,P3**, if $v \in V_2$ has no neighbors in V_3 then we can move v from from V_2 to V_0 and still have a homomorphism. Furthermore, this move does not upset **P1,P2,P3**; thus we may assume **P4** as well.

We let $|V_i| = n_i$ for $i = 0, 1, \ldots, 4$. For a fixed partition we then have

$$\mathbf{Pr}(\mathbf{P1} \wedge \mathbf{P2}) = (1-p)^S \text{ where } S = \binom{n}{2} - \sum_{i=0}^{4} n_i n_{i+1}. \tag{6}$$

$$\mathbf{Pr}(\mathbf{P3} \mid \mathbf{P1} \wedge \mathbf{P2}) = \prod_{i=1}^{4} (1 - (1-p)^{n_{i-1}})^{n_i}. \tag{7}$$

$$\mathbf{Pr}(\mathbf{P4} \mid \mathbf{P1} \wedge \mathbf{P2} \wedge \mathbf{P3}) \leqslant \left(1 - \left(1 - \frac{1}{n_2}\right)^{n_3} (1-p)^{n_3}\right)^{n_2} \tag{8}$$

Equations (6) and (7) are self evident, but we need to justify (8). Consider the bipartite subgraph Γ of $G_{n,p}$ induced by $V_2 \cup V_3$. **P3** tells us that each $v \in V_3$ has a neighbor in V_2. Denote this event by $\mathcal{A}$. We will now describe the construction of a random bipartite graph Γ' on V_2, V_3 such that we can couple Γ, Γ' so that $\Gamma \subseteq \Gamma'$. The RHS of (8) is the probability that **P4** holds using Γ' and the coupling implies that this bounds the probability of **P4** in Γ. To construct Γ' we choose a random mapping ϕ from V_3 to V_2. We then create a bipartite graph Γ' with edge set $E_1 \cup E_2$. Here $E_1 = \{\{x, y\} : x \in V_3, y = \phi(x)\}$ and E_2 is obtained by independently including each of the $n_2 n_3$ possible edges between V_2 and V_3 with probability p. We now prove that we can couple Γ, Γ' so that $\Gamma \subseteq \Gamma'$.

Event $\mathcal{A}$ can be construed as follows: A vertex in $v \in V_3$ chooses B_v neighbors in V_2 where B_v is distributed as a binomial $Bin(n_2, p)$, conditioned to be at least one. The neighbors of v in V_2 will then be a random B_v subset of V_2. We only have to prove then that if v chooses B'_v random neighbors in Γ' then B'_v stochastically dominates B_v. Here B'_v is one plus $Bin(n_2 - 1, p)$ and domination is easy to confirm. We have $n_2 - 1$ instead of n_2, since we do not wish to count the edge v to $\phi(v)$ twice. Note also, included in the estimate in the RHS of (8) is the fact that the set of events $\{v \text{ is not chosen as an image of } \phi\}$ for $v \in V_2$, are negatively correlated.

We now write $n_i = \alpha_i n$ for $i = 0, \ldots, 4$. We are particularly interested in the case where $c = 4$. Now (5) implies that $G_{n,\frac{4}{n}}$ has no induced bipartite subgraph of size βn for $\beta > 0.94$. Thus we may assume that $\alpha_i \geqslant 0.06$ for $i = 0, \ldots, 4$. In which case we can write

$$\mathbf{Pr}(\mathbf{P1} \wedge \mathbf{P2} \wedge \mathbf{P3} \wedge \mathbf{P4}) \leqslant$$
$$e^{o(n)} \times \exp\left\{-c\left(\frac{1}{2} - \sum_{i=0}^{4} \alpha_i\alpha_{i+1}\right) n\right\} \times \left(\prod_{i=1}^{4}(1 - e^{-c\alpha_{i-1}})^{\alpha_i}\right)^n \times$$
$$(1 - e^{-\alpha_3/\alpha_2} e^{-c\alpha_3})^{\alpha_2 n}.$$

The number of choices for $V_0, \ldots, V_4$ with these sizes is

$$\binom{n}{n_0, n_1, n_2, n_3, n_4} = e^{o(n)} \times \left(\frac{1}{\prod_{i=0}^{4} \alpha_i^{\alpha_i}}\right)^n \leqslant 5^n.$$

Putting $\alpha_4 = 1 - \alpha_0 - \alpha_1 - \alpha_2 - \alpha_3$ and

$$b = b(c, \alpha_0, \alpha_1, \alpha_2, \alpha_3) = \frac{1}{{\alpha_0}^{\alpha_0}{\alpha_1}^{\alpha_1}{\alpha_2}^{\alpha_2}{\alpha_3}^{\alpha_3}{\alpha_4}^{\alpha_4}}$$
$$e^{c(\alpha_0\alpha_4 - \frac{1}{2})}(e^{c\alpha_0} - 1)^{\alpha_1}(e^{c\alpha_1} - 1)^{\alpha_2}(e^{c\alpha_2} - 1)^{\alpha_3}(e^{c\alpha_3} - 1)^{\alpha_4}(1 - e^{-\alpha_3/\alpha_2}e^{-c\alpha_3})^{\alpha_2},$$

we see that since there are $O(n^4)$ choices for $n_0, \ldots, n_4$ we have

$$\mathbf{Pr}(\exists \text{ a homomorphism from } G_{n,\frac{4}{n}} \text{ to } C_5) \leqslant e^{o(n)} \left(\max_{\substack{\alpha_0+\cdots+\alpha_3 \leqslant 0.94 \\ \alpha_0,\ldots,\alpha_3 \geqslant 0.06}} b(4, \alpha_0, \alpha_1, \alpha_2, \alpha_3)\right)^n. \tag{9}$$

In the next section, we describe a numerical procedure for verifying that the maximum in (9) is less than 1. This will complete the proof of Theorem 3.

6 Bounding the function.

Our aim now is to bound the partial derivatives of $b(4.0, \alpha_0, \alpha_1, \alpha_2, \alpha_3)$, to translate numerical computations of the function on a grid to a rigorous upper bound.

Before doing this we verify that w.h.p. $G_{n,p=\frac{4}{n}}$ has no independent set S of size $s = 3n/5$ or more. Indeed,

$$\mathbf{Pr}(\exists S) \leqslant 2^n(1-p)^{\binom{s}{2}} \leqslant 2^n e^{-18n/25} e^{12/5} = o(1).$$

In the calculations below we will make use of the following bounds: They assume that $0.06 \leqslant \alpha_i \leqslant 0.6$ for $i \geqslant 0$.

$$\log(\alpha_i) > -2.82; \quad -1.31 < \log(e^{4\alpha_i} - 1) < 2.31; \quad \frac{e^{4\alpha_i}}{e^{4\alpha_i} - 1} < 4.69$$

$$\frac{1}{e^{4\alpha_i}-1} < 3.69; \quad \log(e^{\alpha_3/\alpha_2+4\alpha_3}-1) > -0.91; \quad \frac{1+4\alpha_2}{e^{\alpha_3/\alpha_2}e^{4\alpha_3}-1} < 8.40.$$

We now use these estimates to bound the absolute values of the $\frac{1}{b}\cdot\frac{\partial b}{\partial \alpha_i}$. Our target value for these is 30. We will be well within these bounds except for $i=2$

Taking logarithms to differentiate with respect to α_0, we find

$$\frac{\partial b}{\partial \alpha_0} = b(c,\alpha_0,\alpha_1,\alpha_2,\alpha_3)\times \left(c\left(-\alpha_0+\alpha_1+\frac{\alpha_1}{e^{\alpha_0 c}-1}+\alpha_4\right) - \log(\alpha_0)+\log(\alpha_4)-\log(e^{\alpha_3 c}-1)\right). \quad (10)$$

In particular, for $c=4$,

$$\frac{1}{b}\cdot\frac{\partial b}{\partial \alpha_0} \geqslant -4\alpha_0+\log(\alpha_4)-\log(e^{4\alpha_3}-1) > -2.4-2.82-2.31,$$
$$\frac{1}{b}\cdot\frac{\partial b}{\partial \alpha_0} \leqslant 4\left(\alpha_1+\frac{\alpha_1}{e^{\alpha_0 c}-1}+\alpha_4\right)-\log(\alpha_0)-\log(e^{4\alpha_3}-1) < 4\times 4.69+2.82+1.31.$$

Similarly, we find

$$\frac{\partial b}{\partial \alpha_1} = b(c,\alpha_0,\alpha_1,\alpha_2,\alpha_3)\times \left(c\left(-\alpha_0+\alpha_2+\frac{\alpha_2}{e^{\alpha_1 c}-1}\right) - \log(\alpha_1)+\log(\alpha_4)+\log\left(\frac{e^{\alpha_0 c}-1}{e^{\alpha_3 c}-1}\right)\right), \quad (11)$$

and so for $c=4$,

$$\frac{1}{b}\cdot\frac{\partial b}{\partial \alpha_1} \geqslant -4\alpha_0+\log(\alpha_4)+\log(e^{4\alpha_0}-1)-\log(e^{4\alpha_3}-1) > -2.4-2.82-3.62,$$
$$\frac{1}{b}\cdot\frac{\partial b}{\partial \alpha_1} \leqslant 4\left(\alpha_2+\frac{\alpha_2}{e^{4\alpha_1}-1}\right)-\log(\alpha_1)-\log(e^{4\alpha_3}-1) < 2.4\times 4.69+2.82+1.31.$$

We next find that

$$\frac{\partial b}{\partial \alpha_2} = b(c,\alpha_0,\alpha_1,\alpha_2,\alpha_3)\times$$
$$c\left(-\alpha_0+\alpha_3+\frac{\alpha_3}{e^{\alpha_2 c}-1}\right)-\frac{\alpha_3/\alpha_2}{e^{\alpha_3/\alpha_2+c\alpha_3}-1}+$$
$$\log\alpha_4-\log\alpha_2+\log(e^{\alpha_1 c}-1)-\log(e^{\alpha_3 c}-1)-\frac{\alpha_3}{\alpha_2}-c\alpha_3-\log(e^{\alpha_3/\alpha_2+c\alpha_3}-1); \quad (12)$$

and so for $c=4$,

$$\frac{1}{b}\cdot\frac{\partial b}{\partial \alpha_2} \geqslant -4\alpha_0-\frac{\alpha_3}{\alpha_2}\frac{e^{\alpha_3/\alpha_2+c\alpha_3}}{e^{\alpha_3/\alpha_2+c\alpha_3}-1}-\log(e^{\alpha_3/\alpha_2+c\alpha_3}-1)+\log(\alpha_4)+\log\left(\frac{e^{4\alpha_1}-1}{e^{4\alpha_3}-1}\right)$$

We need to be a little careful here. Now $\alpha_3/\alpha_2 \leqslant 10$ and if $\alpha_3/\alpha_2 \geqslant 9$ then $\alpha_3 \geqslant 0.54$ and then $\alpha_i \leqslant 0.46-3\times .06 = 0.28$ for $i\neq 3$. We bound $-\frac{1}{b}\cdot\frac{\partial b}{\partial \alpha_i}$ for both possibilities. Continuing we get

$$\frac{\alpha_3}{\alpha_2} \geqslant 9 : \frac{1}{b} \cdot \frac{\partial b}{\partial \alpha_2} > -1.12 - 10.01 - 12.4 - 2.82 - 3.62 = -29.97,$$
$$\frac{\alpha_3}{\alpha_2} \leqslant 9 : \frac{1}{b} \cdot \frac{\partial b}{\partial \alpha_2} > -2.4 - 9.01 - 11.4 - 2.82 - 3.62,$$
$$\frac{1}{b} \cdot \frac{\partial b}{\partial \alpha_2} \leqslant 4\left(\alpha_3 + \frac{\alpha_3}{e^{4\alpha_2} - 1}\right) - \log(\alpha_2) + \log\left(\frac{e^{4\alpha_1} - 1}{e^{4\alpha_3} - 1}\right) - \log(e^{\alpha_3/\alpha_2 + c\alpha_3} - 1)$$
$$< 2.4 \times 3.69 + 2.82 + 3.62 + 0.91.$$

Finally, we find that

$$\frac{\partial b}{\partial \alpha_3} = b(c, \alpha_0, \alpha_1, \alpha_2, \alpha_3) \times$$
$$c\left(-\alpha_0 + \alpha_4 \frac{e^{c\alpha_3}}{e^{c\alpha_3} - 1}\right) + \frac{1 + c\alpha_2}{e^{\alpha_3/\alpha_2} e^{c\alpha_3} - 1} + \log(\alpha_4) - \log(\alpha_3) + \log\left(\frac{e^{\alpha_2 c} - 1}{e^{\alpha_3 c} - 1}\right) \quad (13)$$

and so for $c = 4$

$$\frac{1}{b} \cdot \frac{\partial b}{\partial \alpha_3} \geqslant -4\alpha_0 + \log(\alpha_4) + \log(e^{4\alpha_2} - 1) - \log(e^{4\alpha_3} - 1) > -2.4 - 2.82 - 3.62,$$
$$\frac{1}{b} \cdot \frac{\partial b}{\partial \alpha_3} \leqslant 4\alpha_4 \frac{e^{4\alpha_3}}{e^{4\alpha_3} - 1} + \frac{1 + 4\alpha_2}{e^{\alpha_3/\alpha_2} e^{4\alpha_3} - 1} - \log(\alpha_3) + \log\left(\frac{e^{4\alpha_2} - 1}{e^{4\alpha_3} - 1}\right)$$
$$< 2.4 \times 4.69 + 8.40 + 2.82 + 3.62.$$

We see that $|\frac{1}{b} \cdot \frac{\partial b}{\partial \alpha_i}| < 30$ for all $0 \leqslant i \leqslant 3$. Thus, if we know that $b(c, \alpha_0, \alpha_1, \alpha_2, \alpha_3) \leqslant B$ for some B, this means that we can bound $b(4, \alpha_0, \alpha_1, \alpha_2, \alpha_3) < \rho$ by checking that $b(4, \alpha_0, \alpha_1, \alpha_2, \alpha_3) < \rho - \varepsilon$ on a grid with step-size $\delta \leqslant \varepsilon/(2 \cdot B \cdot 30)$.

The C++ program in Appendix A checks that $b(4, \alpha_0, \alpha_1, \alpha_2, \alpha_3) < .949$ on a grid with step-size $\delta = .0008$ (it completes in around an hour or less on a standard desktop computer, and is available for download from the authors' websites). Suppose now that $B \geqslant 1$ is the supremum of $b(4, \alpha_0, \alpha_1, \alpha_2, \alpha_3)$ in the region of interest. For $\varepsilon = 60\delta B = 0.048B$, we must have at some δ-grid point that $b(4, \alpha_0, \alpha_1, \alpha_2, \alpha_3) \geqslant B - \varepsilon = .962B \geqslant .962$. This contradicts the computer-assisted bound of $< .949$ on the grid, completing the proof of Theorem 3. □

Acknowledgement

We thank the anonymous referee for a timely and useful pair of reviews.

References

[1] D. Achlioptas and E. Friedgut, A sharp threshold for k-colorability, *Random Structures and Algorithms* 14 (1999) 63-70.

[2] D. Achlioptas and C. Moore, Almost all Graphs with Average Degree 4 are 3-colorable, *Journal of Computer and System Science* 67 (2003) 441-471.

[3] D. Achlioptas and A. Naor, The two possible values of the chromatic number of a random graph, *Annals of Mathematics* 162 (2005), 1333-1349.

[4] J. Aronson, A. Frieze and B. Pittel, Maximum matchings in sparse random graphs: Karp-Sipser revisited, *Random Structures and Algorithms* **12** (1998), 111-178.

[5] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labeled graphs, *European Journal on Combinatorics* 1(1980) 311-316.

[6] B. Bollobás, Random Graphs, First Edition, Academic Press, London 1985, Second Edition, Cambridge University Press, 2001.

[7] E. Friedgut, Sharp Thresholds of Graph Properties, and the k-sat Problem, *Journal of the American Mathematical Society* 12 (1999) 1017-1054.

[8] H. Hatami, Random cubic graphs are not homomorphic to the cycle of size 7, *Journal of Combinatorial Theory B* 93 (2005) 319-325.

[9] S. Janson, M.J. Luczak and A. Ruciński, Random Graphs, John Wiley and Sons, New York, 2000.

[10] T. Łuczak, A note on the sharp concentration of the chromatic number of random graphs, *Combinatorica* 11 (1991) 295-297.

[11] B. Pittel, On Tree Census and the Giant Component in Sparse Random Graphs, *Random Structures and Algorithms* 1 (1990) 311-342.

[12] X. Zhu, Circular chromatic number: a survey, *Discrete Mathematics* 229 (2001) 371–410.

A C++ code to check function bound

```
#include <iostream>
#include <math.h>
#include <stdlib.h>
using namespace std;
int main(int argc, char* argv[]){
  double delta=.0008;      //step size
  double maxIndSet=.6;    //no independent sets larger than this fraction
  double minClass=.06;    //all color classes larger than this fraction
  double val=0;
  double maxval=0;
  double maxa0,maxa1,maxa2,maxa3; //to record the coordinates of max value
  maxa0=maxa1=maxa2=maxa3=0;
  double A23,A,B,C;         //For precomputing parts of the function
  double c=4;
  for (double a3=minClass; a3 + 4*minClass<1; a3+=delta){
    B=exp(c*a3)-1;
    for (double a2=minClass; a3 + a2 + 3*minClass<1; a2+=delta){
      A23=1/(pow(a2,a2)*pow(a3,a3)) * exp(-c/2)
                * pow(exp(c*a2)-1,a3) * pow(1-exp(-a3/a2)*exp(-c*a3),a2);
      for (double a1=minClass;
           a3+a1<maxIndSet && a3 + a2 + a1 + 2*minClass<1;
           a1+=delta){
        A=A23/pow(a1,a1)* pow(exp(c*a1)-1,a2);
        for (double a0=max(max(minClass,.4-a2-a3),.4-a1-a3);
             a2+a0<maxIndSet && a3+a0<maxIndSet && a3+a2+a1+a0+minClass<1;
             a0+=delta){
          double a4=1-a0-a1-a2-a3;
          C=exp(c*a0);
          val=1/pow(a0,a0) * A * pow(B*C/a4,a4)* pow(C-1,a1);
          if (val>maxval){
            maxval=val;
            maxa0=a0; maxa1=a1; maxa2=a2; maxa3=a3;
          }
        }
      }
    }
  }
  cout << "Max is "<<maxval<<", obtained at ("
       <<maxa0<<","<<maxa1<<","<<maxa2<<","<<maxa3<<","
       <<1-maxa0-maxa1-maxa2-maxa3<<")"<<endl;
}
```

Program output:

```
$./bound
Max is 0.948754, obtained at (0.2904,0.2568,0.1704,0.1632,0.1192)
```