



HAL
open science

Clique coloring of binomial random graphs

Colin McDiarmid, Dieter Mitsche, Pawel Pralat

► **To cite this version:**

Colin McDiarmid, Dieter Mitsche, Pawel Pralat. Clique coloring of binomial random graphs. *Random Structures and Algorithms*, 2018, 10.1002/rsa.20804 . hal-02083918

HAL Id: hal-02083918

<https://hal.science/hal-02083918>

Submitted on 29 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CLIQUE COLOURING OF BINOMIAL RANDOM GRAPHS

COLIN MCDIARMID, DIETER MITSCHE, AND PAWEŁ PRAŁAT

ABSTRACT. A clique colouring of a graph is a colouring of the vertices so that no maximal clique is monochromatic (ignoring isolated vertices). The smallest number of colours in such a colouring is the clique chromatic number. In this paper, we study the asymptotic behaviour of the clique chromatic number of the random graph $\mathcal{G}(n, p)$ for a wide range of edge-probabilities $p = p(n)$. We see that the typical clique chromatic number, as a function of the average degree, forms an intriguing step function.

1. INTRODUCTION AND MAIN RESULTS

A *proper colouring* of a graph is a labeling of its vertices with colours such that no two vertices sharing the same edge have the same colour. The smallest number of colours in a proper colouring of a graph $G = (V, E)$ is called its *chromatic number*, and it is denoted by $\chi(G)$.

In this paper we are concerned with another notion of vertex colouring. A *clique* $S \subseteq V$ is a subset of the vertex set such that any pair of vertices in S is connected by an edge. Moreover, a clique S is *maximal* if there is no vertex in $V \setminus S$ connected by an edge to every vertex in S (in other words, S is not a proper subset of another clique). A *clique colouring* of a graph G is a colouring of the vertices so that no maximal clique is monochromatic, ignoring isolated vertices. The smallest number of colours in such a colouring is called the *clique chromatic number* of G , denoted by $\chi_c(G)$. Clearly, $\chi_c(G) \leq \chi(G)$ but it is possible that $\chi_c(G)$ is much smaller than $\chi(G)$. For example, for any $n \geq 2$ we have $\chi(K_n) = n$ but $\chi_c(K_n) = 2$. Note that if G is triangle-free then $\chi_c(G) = \chi(G)$.

The problem has received considerable attention for deterministic graphs: in [23] it was shown that planar graphs satisfy $\chi_c(G) \leq 3$. In [1] a necessary and sufficient condition for $\chi_c(G) \leq k$ on line graphs was given. Moreover, several graph classes are known to satisfy $\chi_c(G) \leq 2$: claw-free perfect graphs [2], co-diamond free graphs [5], claw-free planar graphs [24], powers of cycles (other than odd cycles longer than three that need three colours) [3], and also claw-free graphs with maximum degree at most 7 (except for odd cycles longer than 3) [18]. Also, circular-arc graphs are known to have $\chi_c(G) \leq 3$ (see [4]). Further results about other classes of graphs can also be found in [15], and the clique chromatic number of graphs without having long paths was studied in [9]. On the algorithmic side, it is known that testing whether $\chi_c(G) = 2$

Date: February 14, 2018.

1991 Mathematics Subject Classification. 05C80, 05C15, 05C35.

Key words and phrases. random graphs, clique chromatic number.

The research of the third author is supported in part by NSERC and Ryerson University.

for a planar graph can be performed in polynomial time [16], but deciding whether $\chi_c(G) = 2$ is *NP*-complete for 3-chromatic perfect graphs [16] and for graphs with maximum degree 3 [2]. The clique chromatic number for geometric graphs (in particular, random geometric graphs) is analysed in the accompanying paper [22].

Let us recall the classic model of random graphs that we study in this paper. The *binomial random graph* $\mathcal{G}(n, p)$ is the random graph G with vertex set $[n]$ in which every pair $\{i, j\} \in \binom{[n]}{2}$ appears independently as an edge in G with probability p . Note that $p = p(n)$ may (and usually does) tend to zero as n tends to infinity. The behaviour of many colouring problems has been investigated for $\mathcal{G}(n, p)$: the classic chromatic number has been intensively studied, see [12, 14] and the references therein; the list chromatic number (known also as the choice number) was studied among others in [17, 25], and other variants were analysed recently in [8, 6] and [10].

All asymptotics throughout are as $n \rightarrow \infty$ (we emphasize that the notations $o(\cdot)$ and $O(\cdot)$ refer to functions of n , not necessarily positive, whose growth is bounded). We use the notations $f \ll g$ for $f = o(g)$ and $f \gg g$ for $g = o(f)$. We also write $f(n) \sim g(n)$ if $f(n)/g(n) \rightarrow 1$ as $n \rightarrow \infty$ (that is, when $f(n) = (1 + o(1))g(n)$). We say that events A_n in a probability space hold *with high probability* (or whp), if the probability that A_n holds tends to 1 as n goes to infinity. Since we aim for results that hold whp, we will always assume that n is large enough. We often write $\mathcal{G}(n, p)$ when we mean a graph drawn from the distribution $\mathcal{G}(n, p)$. Finally, we use $\log n$ to denote natural logarithms.

Here is our main result. We consider the edge probability $p(n)$ ranging from the sparse case when $pn \rightarrow \infty$ arbitrarily slowly, to the dense case when $p = 1 - \varepsilon$ for an arbitrarily small $\varepsilon > 0$, and we break this range into 8 parts. After the theorem we give a corollary which is less precise but easier to read.

Theorem 1.1. *Let $\varepsilon > 0$ be a constant (arbitrarily small). Let $\omega = \omega(n)$ be a function tending to infinity with n (arbitrarily slowly), and suppose that $\omega = o(\sqrt{\log n})$. Let $G \in \mathcal{G}(n, p)$ for some $p = p(n)$. Then, the following holds whp:*

- (a) *If $pn \geq \omega$ and $pn < n^{1/2-\omega/\sqrt{\log n}}$, then $\chi_c(G) \sim \chi(G) \sim \frac{pn}{2 \log(pn)}$.*
- (b) *If $n^{1/2-\omega/\sqrt{\log n}} \leq pn < \sqrt{2n \log n}$, then $\chi_c(G) = \Omega\left(\frac{p^{3/2}n}{(\log n)^{1/2}}\right)$ and $\chi_c(G) \leq \chi(G) \sim \frac{pn}{2 \log(pn)}$.*
- (c) *If $\sqrt{2n \log n} \leq pn < n^{3/5-(6/\log n)^{1/2}}$, then $\chi_c(G) = \Theta\left(\frac{p^{3/2}n}{(\log n)^{1/2}}\right)$.*
- (d) *If $n^{3/5-(6/\log n)^{1/2}} \leq pn < n^{3/5}(\log n)^{3/5}$, then $\chi_c(G) = n^{2/5+o(1)}$.*
- (e) *If $n^{3/5}(\log n)^{3/5} \leq pn < n^{2/3-\varepsilon}$, then $\chi_c(G) = \Omega(1/p)$ and $\chi_c(G) = O(\log n/p)$.*
- (f) *If $pn = n^{2/3+o(1)}$ and $pn < n^{2/3}(\log n)^{4/9}$, then $\chi_c(G) = n^{1/3+o(1)}$.*
- (g) *If $n^{2/3}(\log n)^{4/9} \leq pn < n^{1-\varepsilon}$, then $\chi_c(G) = \Omega(1/p)$ and $\chi_c(G) = O(\log n/p)$.*
- (h) *If $p = n^{-o(1)}$ and $p \leq 1 - \varepsilon$, then $\chi_c(G) \leq (1/2 + o(1)) \log_{1/(1-p)} n$.*

Note that

$$\log_{1/(1-p)} n = \begin{cases} \Theta(\log n) & \text{if } p = \Omega(1) \text{ and } p \leq 1 - \varepsilon \text{ for some } \varepsilon > 0, \\ (1 + o(1))(\log n)/p & \text{if } p = o(1). \end{cases}$$

In places a slightly tighter result than the one given in Theorem 1.1 can be obtained. We are interested in the ratio between the upper and lower bounds on $\chi_c(G)$ for each p , but in order to keep the statement reasonably simple, we spelled out results only for large intervals of p . For instance, as cases (d) and (f) are concerned with $pn = n^{3/5+o(1)}$ and $pn = n^{2/3+o(1)}$ respectively, we treated these cases in the statement of Theorem 1.1 less fully. For the reader interested in more precise bounds on this ratio, we refer to the proofs in the sections below for more details.

In order to understand better the behaviour of the clique chromatic number, let us introduce the following function $f : (0, 1) \rightarrow \mathbb{R}$:

$$f(x) = \begin{cases} x & 0 < x < 1/2 \\ 1 + 3(x - 1)/2 & 1/2 \leq x < 3/5 \\ 1 - x & 3/5 \leq x < 1. \end{cases}$$

This function is depicted in Figure 1. We get immediately the following corollary.

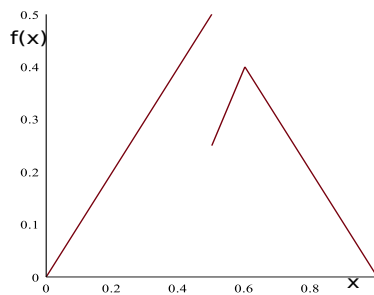


FIGURE 1. The function $f(x)$ related to the clique chromatic number of $G(n, p)$

Corollary 1.2. *Let $f : (0, 1) \rightarrow \mathbb{R}$ be defined as above. Let $G \in \mathcal{G}(n, p)$ for some $p = p(n)$. If $pn = n^{x+o(1)}$ for some $x \in (0, 1/2) \cup (1/2, 1)$, then whp $\chi_c(G) = n^{f(x)+o(1)}$.*

Figure 1 shows that there are dramatic transitions in Corollary 1.2 when p is about $n^{-1/2}$ and $n^{-2/5}$, and there is some problem at about $n^{-1/3}$ in Theorem 1.1. What is happening at these points? Here is the rough story.

For $p \ll n^{-1/2}$, whp most edges are not in triangles, and $\chi_c(G)$ is close to $\chi(G)$. For $p \gg n^{-1/2}$, whp each edge is in many triangles, and we may use triangle-free sets as colour classes; or we may find relatively small dominating sets which yield an upper bound on $\chi_c(G)$ (see Lemma 3.2). For $p \ll n^{-2/5}$, the first bound (which increases with p) is stronger, whereas from then on the second bound (which decreases with p) gives better results. For lower bounds on $\chi_c(G)$, depending on the range of p , we find a value of k , so that whp there are many k -cliques and most of them are not contained inside $(k+1)$ -cliques. (In the central range in Figure 1, for $n^{-1/2} \ll p \ll n^{-2/5}$, we use $k = 3$.) For $p \ll n^{-1/3}$ whp most triangles are not contained in 4-cliques; and for $p \gg n^{-1/3}$ whp each triangle is in many 4-cliques.

2. PRELIMINARIES

In this section, let us recall a few inequalities that are well known and can be found, for example, in [12]. We will use the following version of *Chernoff's bound*. Suppose that $X \in \text{Bin}(n, p)$ is a binomial random variable with expectation $\mu = pn$. If $0 < \delta < 1$, then

$$\mathbb{P}[X < (1 - \delta)\mu] \leq \exp\left(-\frac{\delta^2\mu}{2}\right), \quad (1)$$

and if $\delta > 0$,

$$\mathbb{P}[X > (1 + \delta)\mu] \leq \exp\left(-\frac{\delta^2\mu}{2 + \delta}\right). \quad (2)$$

We will use the following version of *Janson's inequality*, adapted to our setting.

Lemma 2.1. *Let $G = (V, E) \in \mathcal{G}(n, p)$, $S \subseteq V$, and $k \in \mathbb{N} \setminus \{1, 2\}$. For $A \subseteq S$ with $|A| = k$, let \mathcal{I}_A be the indicator variable which is 1 if the vertices in A form a clique of size k , and 0 otherwise. Let $X = \sum_{A \subseteq S, |A|=k} \mathcal{I}_A$ be the random variable counting copies of K_k in S , $\mu = \mathbb{E}[X]$, and $\overline{\Delta} = \sum_{A, B \subseteq S, |A|=|B|=k, |A \cap B| \geq 2} \mathbb{E}[\mathcal{I}_A \mathcal{I}_B]$. Then, for $0 \leq t \leq \mathbb{E}X$,*

$$\mathbb{P}(X \leq \mu - t) \leq \exp\left(-\frac{\varphi(-t/\mu)\mu^2}{\overline{\Delta}}\right),$$

where $\varphi(x) = (1 + x) \log(1 + x) - x$.

(Let us note that indicator random variables \mathcal{I}_A and \mathcal{I}_B are independent if $|A \cap B| \leq 1$; see the definition of $\overline{\Delta}$ above.)

We will use the following result of Vu to obtain bounds on the upper tail of certain subgraphs (see [26]). Denote by $e(H)$ ($v(H)$, respectively) the number of edges (vertices, respectively) of a graph H . We say that a graph H is *balanced* if for every subgraph $H' \subseteq H$ with $v(H') \geq 1$, we have $e(H')/v(H') \leq e(H)/v(H)$.

Lemma 2.2. *Let H be a fixed balanced graph on k vertices, and let $K > 0$ be a constant. Then there are constants $c = c(H, K) > 0$ and $s = s(H, K)$ such that the following holds. Let $G \in \mathcal{G}(n, p)$, let Y denote the number of appearances of H in the graph G , and let $\mu = \mathbb{E}[Y]$. If $0 < \varepsilon \leq K$ and $\varepsilon^2 \mu^{1/(k-1)} \geq s \log n$, then*

$$\mathbb{P}(Y \geq (1 + \varepsilon)\mu) \leq \exp(-c\varepsilon^2 \mu^{1/(k-1)}).$$

(Let us remark that stronger bounds on the upper tail are known; see, for example, [7]. For our purposes, however, the current one is sufficient.)

We will also make use of the following version of *Harris's inequality* (see [11]). Let $G \in \mathcal{G}(n, p)$, with vertex set $V = [n]$ and random edge set E . Let $F \subseteq \binom{V}{2}$ so F is a set of possible edges. Let Ω be the collection of all subsets of the set F . For $A \subseteq \Omega$ let $\mathbb{P}(A)$ mean $\mathbb{P}(E \cap F \in A)$. For $A \subseteq \Omega$, we say that A is a *downset* if $x \in A$ and $y \subseteq x$ implies $y \in A$, and A is an *upset* if $x \in A$ and $x \subseteq y \subseteq \Omega$ implies $y \in A$. We also may refer to the event $E \cap F \in A$ as a *down-event* or an *up-event* in these two cases.

Lemma 2.3. *With notation as above, let A and B be upsets and let C and D be downsets. Then A and B are positively correlated, i.e., $\mathbb{P}(A \cap B) \geq \mathbb{P}(A)\mathbb{P}(B)$, and similarly C and D are positively correlated; and A and C are negatively correlated, i.e., $\mathbb{P}(A \cap C) \leq \mathbb{P}(A)\mathbb{P}(C)$.*

The final lemma here is a simplified version of Lemma 2.3 of [20]. It says roughly that a random walk which is biased upwards when needed tends to move upwards as we should expect. We shall use it in the proof of Theorem 3.1 (to analyse a greedy procedure to find triangle-free sets).

Lemma 2.4. *Let $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_k$ be a filter (in a probability space). Let Y_1, Y_2, \dots, Y_k be binary random variables such that each Y_i is \mathcal{F}_i -measurable, and let $Z = \sum_{i=1}^k Y_i$. Let E_0, E_1, \dots, E_{k-1} be events where $E_i \in \mathcal{F}_i$ for $i = 0, \dots, k-1$. Let $0 \leq p \leq 1$ and let z be a positive integer. Suppose that for each $i = 1, \dots, k$*

$$\mathbb{P}(Y_i = 1 \mid \mathcal{F}_{i-1}) \geq p \quad \text{on } E_{i-1}.$$

Then

$$\mathbb{P}\left(\left(Z < z\right) \wedge \left(\bigwedge_i E_i\right)\right) \leq \mathbb{P}(\text{Bin}(k, p) < z).$$

3. UPPER BOUNDS

3.1. Upper bounds for parts (a) and (b). Part (a) and (b) of Theorem 1.1 follow immediately from the fact that $\chi_c(G) \leq \chi(G)$ for each graph. It is well known that whp

$$\chi(\mathcal{G}(n, p)) \sim \frac{pn}{2 \log(pn)},$$

provided that $pn \rightarrow \infty$ as $n \rightarrow \infty$, and $p = o(1)$ (see [19, 21]).

3.2. Upper bound for part (c). For p as here, we shall see that whp each edge is in a triangle, so we can use triangle-free sets of vertices as colour classes. We repeatedly use a natural greedy algorithm to construct maximal triangle-free sets and remove them. While many vertices remain, with high probability each set constructed has size at least about $p^{-3/2}(\log n)^{1/2}$, so we will not remove too many of them. Later phases require a negligible number of further colours. In the proof, in order to show that our greedy triangle-free sets are large, we need first to show that they do not tend to contain more edges than a typical set of the same size.

Theorem 3.1. *Suppose that $p = p(n)$ is such that $\sqrt{2 \log n/n} \leq p \leq 1/\log n$. Let $c > \sqrt{2}$ be a constant. Then whp for $G \in \mathcal{G}(n, p)$ we have*

$$\chi_c(G) \leq c p^{3/2} n (\log n)^{-1/2}.$$

Proof. First let us check that whp each edge is in a triangle, for then we can colour using triangle-free sets. The expected number of edges that are not part of any triangle is

$$\binom{n}{2} p(1-p^2)^{n-2} \leq n^2 p e^{-np^2} (1-p^2)^{-2} = o(1)$$

since $e^{-np^2} \leq n^{-2}$ and $p = o(1)$. The desired property now follows from Markov's inequality.

We will repeatedly use the following greedy algorithm to construct a triangle-free set A of vertices in a graph G with vertex set a subset of the positive integers. Consider the vertices in increasing order. Initially set $A = \emptyset$ and $W = V$. While W is non-empty: remove the first member w from W , and if $A \cup \{w\}$ is triangle-free then add w to A . Note that this procedure examines only potential edges with at least one end in A .

How large is the final triangle-free set A for $G \in \mathcal{G}(n, p)$? The first thing to consider is the number of edges of G within A when $|A| = a$. We shall see in part (1) of the proof below that this is stochastically at most $\text{Bin}(\binom{a}{2}, p)$. In part (2) of the proof we shall define appropriate integers j_0 and t such that whp we find a triangle-free set of size j_0 within t steps; and we complete the proof in part (3).

Part (1). First we consider the number of edges within the set A . Let n be a (fixed) integer, with corresponding $p = p(n)$, and embed the random graph G in an infinite random graph G' on $1, 2, 3, \dots$ with the same edge probability p . Let E' be the random edge-set of G' . For $j = 1, 2, \dots$ let A_j be the set A when $|A| = j$ (which is well defined for each j since we are working in G' and $p < 1$). Let X_j be the number of edges within A_j (and set $X_0 = 0$), and let $Y_j = X_j - X_{j-1}$.

We **claim** that $X_j \leq_s \text{Bin}(\binom{j}{2}, p)$. This is clearly true for $j = 1$ since $X_1 = 0$. Let $j \geq 2$ and suppose it is true for $j - 1$. Let $0 \leq x \leq \binom{j-1}{2}$ and let E_1 be the event that $X_{j-1} = x$. Let B be a set of $j - 1$ vertices, let F_0 be a set of x edges within B , and let w be an element of V after all of B . Let E_2 be the event that $A_{j-1} = B$ and F_0 is the set of edges of G' within B . Let E_3 be the event that w is the vertex added to A_{j-1} to form A_j .

Let F_1 be the set of $j - 1$ possible edges between w and B . Let $F'_1 = E' \cap F_1$, the random set of edges present in F_1 . Note that $|F'_1| \sim \text{Bin}(j - 1, p)$. For each real t , let $f(t) = \mathbb{P}(\text{Bin}(j - 1, p) \geq t)$: and observe that $|F'_1| \geq t$ is an up-event. Let

$$\mathcal{F}_1 = \{F \subseteq F_1 : |F \cap \{uw, vw\}| \leq 1 \text{ for each edge } uv \in F_0\},$$

and observe that $F'_1 \in \mathcal{F}_1$ is a down-event. Hence, for each t , by Harris's inequality, Lemma 2.3,

$$\begin{aligned} \mathbb{P}(Y_j \geq t \mid E_1 \wedge E_2 \wedge E_3) &= \mathbb{P}(|F'_1| \geq t \mid E_1 \wedge E_2 \wedge E_3) \\ &= \mathbb{P}(|F'_1| \geq t \mid F'_1 \in \mathcal{F}_1) \\ &\leq \mathbb{P}(|F'_1| \geq t) = f(t). \end{aligned}$$

Since this holds for each possible choice of B , F_0 and w specifying events E_2 and E_3 , we obtain

$$\mathbb{P}(Y_j \geq t \mid X_{j-1} = x) = \mathbb{P}(Y_j \geq t \mid E_1) \leq f(t).$$

Now fix $t \geq 0$, and let $g(x) = f(t-x) = \mathbb{P}(\text{Bin}(j-1, p) \geq t-x)$. By the last inequality, since $X_j = X_{j-1} + Y_j$,

$$\begin{aligned} \mathbb{P}(X_j \geq t) &= \sum_x \mathbb{P}(X_{j-1} = x) \mathbb{P}(Y_j \geq t-x \mid X_{j-1} = x) \\ &\leq \sum_x \mathbb{P}(X_{j-1} = x) g(x) = \mathbb{E}[g(X_{j-1})]. \end{aligned}$$

But $X_{j-1} \leq_s \text{Bin}(\binom{j-1}{2}, p)$ by the induction hypothesis, and $g(x)$ is non-decreasing, so

$$\begin{aligned} \mathbb{E}[g(X_{j-1})] &\leq \mathbb{E}[g(\text{Bin}(\binom{j-1}{2}, p))] \\ &= \sum_x \mathbb{P}(\text{Bin}(\binom{j-1}{2}, p) = x) \mathbb{P}(\text{Bin}(j-1, p) \geq t-x) \\ &= \mathbb{P}(\text{Bin}(\binom{j}{2}, p) \geq t). \end{aligned}$$

Thus

$$\mathbb{P}(X_j \geq t) \leq \mathbb{P}(\text{Bin}(\binom{j}{2}, p) \geq t),$$

and we have established the claim that $X_j \leq_s \text{Bin}(\binom{j}{2}, p)$ for each j . This completes part (1) of the proof.

Part (2). Still in the infinite case, let T_j be the number of vertices w tested until we find one to add to A_{j-1} . For $j = 1, 2$ the first available vertex must be added. Fix $j \geq 3$. For each positive integer i , let W^i be the i th vertex in W after the largest vertex in A_{j-1} (the i th candidate to add to A_{j-1}).

Let x, B, F, w, E_1, E_2 and \mathcal{F}_1 be as before. Observe that \mathcal{F}_1 is an intersection of x down-sets in 2^{F_1} , and for each of these downsets \mathcal{D} we have $\mathbb{P}(F'_1 \in \mathcal{D}) = (1-p^2)$. By Harris's inequality (Lemma 2.3) again,

$$\mathbb{P}(w \text{ can be added to } A_{j-1} \mid E_1 \wedge E_2) = \mathbb{P}(F'_1 \in \mathcal{F}_1) \geq (1-p^2)^x.$$

Since this holds for each possible choice of B, F and w , it follows that for each positive integer i we have

$$\mathbb{P}(W^i \text{ can be added to } A_{j-1} \mid X_{j-1} = x) \geq (1-p^2)^x.$$

Hence

$$\mathbb{P}(T_j \geq t \mid X_{j-1} = x) \leq (1 - (1-p^2)^x)^{t-1},$$

and since the upper bound is nondecreasing in x , it follows that

$$\mathbb{P}(T_j \geq t \mid X_{j-1} \leq x) \leq (1 - (1-p^2)^x)^{t-1}.$$

Let $0 < \eta < 1/\sqrt{2}$, and let $j_0 \sim \eta p^{-3/2} (\log n)^{1/2}$. Note that $j_0 = O(n^{3/4} (\log n)^{-1/4}) = o(n^{3/4})$, and $j_0 = \Omega(\log^2 n)$. Observe that $\frac{1}{2}\eta^2 < \frac{1}{4}$; and let $\frac{1}{2}\eta^2 < \alpha < \beta < \frac{1}{4}$. Let

$\mu_0 = \binom{j_0}{2}p \sim (\eta^2/2)p^{-2} \log n$; and let $x_0 \sim \alpha p^{-2} \log n$. Then for each $j = 1, \dots, j_0$, by the Chernoff bound (2)

$$\mathbb{P} \left(\text{Bin} \left(\binom{j-1}{2}, p \right) > x_0 \right) \leq \mathbb{P} \left(\text{Bin} \left(\binom{j_0}{2}, p \right) > x_0 \right) \leq e^{-\Omega(\mu_0)} = o(n^{-2}),$$

where the final step follows since $p = o(1)$. Let $t = p^{-\frac{3}{2}}n^\beta$, so $t = o(n^{\frac{3}{4}+\beta})$ and $t/j_0 = n^{\beta+o(1)}$. Then

$$\begin{aligned} \mathbb{P} \left(\sum_{j=1}^{j_0} T_j \geq t \right) &\leq \sum_{j=1}^{j_0} \mathbb{P}(T_j \geq t/j_0) \\ &\leq \sum_{j=1}^{j_0} \left(\mathbb{P}(T_j \geq t/j_0 \mid X_{j-1} \leq x_0) + \mathbb{P}(X_{j-1} > x_0) \right) \\ &\leq j_0(1 - (1 - p^2)^{x_0})^{t/j_0-1} + j_0 \mathbb{P}(X_{j_0-1} > x_0) \\ &\leq n(1 - (1 - p^2)^{x_0})^{t/j_0-1} + o(1/n) \end{aligned}$$

by the above. But since $p = o(1)$,

$$(1 - p^2)^{x_0} = e^{-(1+o(1))p^2 x_0} = e^{-(\alpha+o(1)) \log n} = n^{-\alpha+o(1)}$$

and so

$$(1 - (1 - p^2)^{x_0})^{t/j_0-1} \leq \exp(-n^{-\alpha+o(1)} \cdot n^{\beta+o(1)}) = \exp(-n^{\beta-\alpha+o(1)}).$$

Thus

$$\mathbb{P} \left(\sum_{j=1}^{j_0} T_j \geq t \right) \leq n \exp(-n^{\beta-\alpha+o(1)}) + o(1/n) = o(1/n).$$

In other words, the probability that the greedy algorithm fails to find a triangle-free set of size j_0 within the first $t - 1$ available vertices is $o(1/n)$. Now let us return to the finite case $G \in \mathcal{G}(n, p)$, and note that the last statement still applies. This completes part (2) of the proof.

Part (3). Let us repeatedly seek a maximal triangle-free set of size at least j_0 , and remove it, as long as at least t vertices remain. Let V' be the set of vertices remaining at some step (when we are about to seek a new triangle-free set). Each time, we have $|V'| \geq t$ (that is, we start with at least t vertices), and the potential edges have not been examined before. Hence, each time, the probability we fail is $o(1/n)$, and so the probability we ever fail is $o(1)$. This whole process uses at most $n/j_0 \sim \eta^{-1}p^{3/2}n(\log n)^{-1/2}$ colours. Observe for later that this number is $\Omega(n^{1/4})$.

By this stage, $|V'| < t$. While $|V'| \geq \log^2 n$, we continue to seek maximal triangle-free sets using the same greedy algorithm, but now we will be content with smaller sets. Any remaining vertices can each be given a new colour, adding at most $\log^2 n$ colours. We need new notation.

Let $s^* = \lceil p^{-3/2} \rceil$, and let $s = \min\{s^*, |V'|\}$. List the vertices in V' in their natural increasing order as v'_1, v'_2, \dots . For $i = 1, \dots, s$ let \mathcal{F}_i be the σ -field generated by the appearance or not of the edges amongst v'_1, \dots, v'_i ; let E_i be the random set of edges amongst v'_1, \dots, v'_i which appear; and let Y_i be the indicator that the i th vertex v'_i is

accepted into A . Let $X = |E_s|$. Then $X \sim \text{Bin}\left(\binom{s}{2}, p\right)$, with expected value at most about $\frac{1}{2}p^{-2} = \Omega(\log^2 n)$. Let $x_0 = \frac{2}{3}p^{-2}$. Then $\mathbb{P}(X > x_0) = o(1/n)$ by the Chernoff bound (2).

Given that $E_{i-1} = F$, the event that $Y_i = 0$ is the union over the edges $e \in F$ of the event that v'_i is adjacent to both end vertices of e . Thus, for each event $B \in \mathcal{F}_{i-1}$

$$\mathbb{P}(Y_i = 0 \mid B \wedge (|E_{i-1}| \leq x_0)) \leq x_0 p^2 \leq 2/3$$

(assuming that the conditioning event has positive probability). Hence

$$\mathbb{P}(Y_i = 1 \mid \mathcal{F}_{i-1}) \geq 1/3 \quad \text{on the event } (|E_{i-1}| \leq x_0).$$

Let Z be the value of $|A|$ after trying to add v'_1, \dots, v'_s (starting from $A = \emptyset$). Thus $Z = \sum_{i=1}^s Y_i$. Hence by Lemma 2.4 and the Chernoff bound (2)

$$\mathbb{P}((Z < s/4) \wedge (X_{s-1} \leq x_0)) \leq \mathbb{P}(\text{Bin}(s, 1/3) < s/4) = e^{-\Omega(s)} = o(1/n),$$

and so

$$\mathbb{P}(Z < s/4) \leq \mathbb{P}(X_{s-1} > x_0) + o(1/n) = o(1/n).$$

Throughout the time when $t > |V'| \geq s^*$, whp each triangle-free set A found has size at least $s^*/4$; so the number of colours used is at most $4t/s^* \sim 4n^\beta$. Throughout the time when $s^* > |V'| \geq \log^2 n$, whp each triangle-free set A found has size at least $|V'|/4$; so the number of colours used is $O(\log |V'|) = O(\log n)$. As we noted before, the final phase uses $O(\log^2 n)$ colours. Thus the total number of colours used after the first phase is $O(n^\beta)$, which is negligible compared with $n/j_0 = \Omega(n^{\frac{1}{4}})$. \square

3.3. Upper bounds for parts (d)-(g). Before we state a useful observation that holds for deterministic graphs, let us introduce a few definitions. An *independent set* (or a *stable set*) of a graph $G = (V, E)$ is a set of vertices in G , no two of which are adjacent. A *dominating set* in G is a subset D of V such that every vertex not in D is adjacent to at least one member of D . Of course, a set could be both independent and dominating; this situation happens if and only if the set is a maximal independent set. Finally, for $v \in V$ let $N(v)$ denote the set of neighbours of v .

The following lemma is part of Theorem 3 of [2]: we give a short proof here for completeness.

Lemma 3.2. *Let $G = (V, E)$ be a graph, and suppose that $A \subseteq V$ is a dominating set. Then $\chi_c(G) \leq |A| + 1$.*

Proof. Let $A = \{v_1, v_2, \dots, v_k\}$. For $i = 1, \dots, k$ in turn, assign colour i to each uncoloured neighbour of v_i . Any vertices remaining uncoloured must form an independent subset of A : give each such vertex colour 0. This gives a proper clique-colouring, since any vertices coloured 0 form an independent set; and for each $i = 1, \dots, k$ the set of vertices coloured i is dominated by v_i , which is not coloured i . \square

From Lemma 3.2, we quickly get the following bound for binomial random graphs. This proves the upper bounds in parts (d) to (g) of Theorem 1.1. We say that $p = p(n)$ is *bounded below 1* if there is some constant $\varepsilon > 0$ such that $p \leq 1 - \varepsilon$ for all n .

Lemma 3.3. *Suppose that $p = p(n)$ is bounded below 1, and $\omega = \omega(n) := pn - \log n \rightarrow \infty$ as $n \rightarrow \infty$. Let $k = k(n) = \lceil \log_{1/(1-p)} n + \log_{1/(1-p)} \omega \rceil$. Then for $G \in \mathcal{G}(n, p)$ whp*

$$\chi_c(G) \leq k + 1 \sim \log_{1/(1-p)} n.$$

Proof. We construct a maximal independent set in the usual greedy way, by fixing an arbitrary order of the vertices v_1, v_2, \dots, v_n and adding vertices (in this order, one by one) to form a maximal independent set. Observe that at the beginning of a given step t of this process, no edge emanating from v_t has been exposed so far. Note that whp the process terminates after at most k vertices are added to the independent set: indeed, the probability that a larger set is created during this greedy process is the probability that a set of size k is created at some point of the process, and this set does not dominate the remaining vertices; and (by considering just the second part) this probability is at most $n(1-p)^k \leq \omega^{-1} = o(1)$. The result now follows by Lemma 3.2. \square

3.4. Upper bound for part (h). We will now show that for dense random graphs, we can improve the bound from part (d) by a factor of about 2. This then proves part (h) of Theorem 1.1. (For sparser graphs, the improvement might also be possible but the argument would be more tedious.)

Theorem 3.4. *Suppose that $p = p(n)$ is bounded below 1, and satisfies $p = n^{-o(1)}$. Then, for $G \in \mathcal{G}(n, p)$ whp*

$$\chi_c(G) \leq (1/2 + o(1)) \log_{1/(1-p)} n.$$

Proof. First, let us start with a few simple properties of $\mathcal{G}(n, p)$.

Claim 1 Whp there is no clique of size $k = \lceil 2 \log_{1/p} n + 1 \rceil \sim 2 \log_{1/p} n$.

Indeed, noting that $k \gg 1$ we get that the expected number of cliques of size k is

$$\binom{n}{k} p^{\binom{k}{2}} \leq \left(\frac{ne}{k} p^{(k-1)/2} \right)^k \leq \left(\frac{e}{k} \right)^k = o(1).$$

The claim holds by Markov's inequality.

Claim 2 Whp there is no maximal clique of size at most $k = \lfloor \log_{1/p} n - 3 \log_{1/p} \log n \rfloor \sim \log_{1/p} n$.

Indeed, the expected number of sets of size k (not necessarily inducing cliques) for which there is no vertex adjacent to all of them is

$$\begin{aligned} \binom{n}{k} (1-p^k)^{n-k} &\leq \left(\frac{ne}{k} \right)^k \exp(-p^k(n-k)) \\ &= O(\exp(k \log n - \log^3 n)) = o(1). \end{aligned}$$

As before, the claim holds by Markov's inequality.

The rest of the proof uses ideas similar to the ones used in the proofs of Lemmas 3.2 and 3.3. Let

$$\gamma = \frac{2 \log_{1/p} \log n}{\log_{1/p} n}$$

and note that $\gamma \rightarrow 0$ as $n \rightarrow \infty$. As before, we fix an arbitrary order of the vertices and construct an independent set A greedily, but we stop once its size is equal to $k = \lceil (1/2 + \gamma) \log_{1/(1-p)} n \rceil$. As before, we use $k + 1$ colours to create colour classes that do not contain any maximal cliques. Let N be the random number of vertices not in A and not dominated by A (and so not coloured yet). Then N is stochastically at most $\text{Bin}(n, (1-p)^k)$, and so by (1) whp $N \leq N_0$, where $N_0 = \lfloor 2n(1-p)^k \rfloor \leq 2n^{1/2-\gamma}$. We will show that whp one additional colour is enough to finish the job. We may assume that $N \leq N_0$. As this part of the random graph is not exposed yet, this subgraph is distributed as a subgraph of $\mathcal{G}(N_0, p)$ and Claim 1 implies that whp the maximum size of a clique in the additional colour is at most

$$\begin{aligned} \lfloor 2 \log_{1/p} N_0 \rfloor &\leq (1 - 2\gamma) \log_{1/p} n + 2 \log_{1/p} 2 + 1 \\ &= \log_{1/p} n - 4 \log_{1/p} \log n + 2 \log_{1/p} 2 + 1 \\ &< \lfloor \log_{1/p} n - 3 \log_{1/p} \log n \rfloor. \end{aligned}$$

On the other hand, Claim 2 implies that whp no clique of such size is maximal in G . The proofs of the upper bounds of Theorem 1.1 are finished. \square

4. LOWER BOUNDS

Given a graph G , we say that a set S of vertices is *maximal-clique-free* if it contains no clique of size at least 2 which is a maximal clique in G . We let $\text{mcf}(G)$ denote the maximum size of a maximal-clique-free set of vertices. Observe that if G has n vertices then

$$\chi_c(G) \geq \frac{n}{\text{mcf}(G)}.$$

Our lower bounds on $\chi_c(G)$ are usually obtained from upper bounds on $\text{mcf}(G)$ by using this inequality.

4.1. Lower bound for part (a). Let us first consider very sparse graphs. Observe that for any graph H with n vertices and at most t vertices in triangles, we have $\chi_c(H) \geq \frac{n-t}{\alpha(H)}$. Let $G \in \mathcal{G}(n, p)$. If $pn \rightarrow \infty$ and $p = o(1)$, then $\chi(G) \sim \frac{n}{\alpha(G)} \sim \frac{pn}{2 \log(pn)}$ whp (see, for example, [19, 21]). For $pn = o(n^{1/3})$ the expected number of triangles is $\binom{n}{3} p^3 = o(n)$, so whp the number of vertices in triangles is $o(n)$. Thus for $pn \rightarrow \infty$ with $pn = o(n^{1/3})$

$$\chi_c(G) \geq (1 + o(1)) \frac{pn}{2 \log(pn)} \text{ whp,} \quad (3)$$

and for such values of pn the lower bound of part (a) is proven.

For the remaining range of pn to be considered in part (a) we may therefore clearly assume $pn > \log n + \omega$. (In fact, we may assume $p \geq n^{1/3}/\omega$ for some function $\omega = \omega(n) \rightarrow \infty$ as $n \rightarrow \infty$ but it would not simplify the argument.) Let $G \in \mathcal{G}(n, p)$, $\omega = \omega(n)$ be a function tending to infinity with n arbitrarily slowly, and let

$$c = c(n) = \frac{9 \log n}{\log(e/(np^2))}.$$

(Observe that $9 \leq c \leq 9 \log n$.)

We will start with the following elementary observation.

Lemma 4.1. *Let ω, c be defined as above, and let $\log n + \omega < pn \leq \sqrt{n}$. Let $G \in \mathcal{G}(n, p)$. Then whp no edge of G belongs to more than c triangles.*

Proof. Consider a pair u, v of distinct vertices. The number of common neighbours of u and v is a random variable X with distribution $\text{Bin}(n-2, p^2)$, which is stochastically bounded from above by $\text{Bin}(n, p^2)$. We first observe that

$$\mathbb{P}(X \geq c) \leq \binom{n}{\lceil c \rceil} (p^2)^{\lceil c \rceil} \leq \left(\frac{en}{c} \cdot p^2\right)^{\lceil c \rceil} \leq \left(\frac{np^2}{e}\right)^c,$$

since $c \geq 9 \geq e^2$ and $np^2/e \leq 1/e \leq 1$. Hence,

$$\mathbb{P}(X \geq c) \leq \exp\left(-c \log\left(\frac{e}{np^2}\right)\right) = n^{-9} = o(n^{-2}).$$

Since there are $O(n^2)$ pairs to consider, the result holds by a union bound. \square

The next lemma is less straightforward.

Lemma 4.2. *Let ω, c be defined as above. Suppose that $\log n + \omega < pn \leq n^{1/2-\omega/\sqrt{\log n}}$. Let*

$$\varepsilon = \varepsilon(n) = 3(2ecnp^2)^{1/(2c)}$$

and $\varepsilon' = \varepsilon'(n) = \max\{\varepsilon, 1/\log \log(pn)\}$. Then $\varepsilon' = o(1)$. Moreover, let $G = (V, E) \in \mathcal{G}(n, p)$. Then, whp every set $K \subseteq V$ of size

$$k = (2 + \varepsilon') \frac{\log(pn)}{p} \sim \frac{2 \log(pn)}{p}$$

has the following properties:

- (a) the graph induced by K , $G[K]$, has at least one edge,
- (b) there is no set of edge-disjoint triangles, each with one vertex in $V \setminus K$, which contains at least a $1/(2c)$ -fraction of the edges in $G[K]$.

Proof. First, recall that $c = (9 \log n)/(\log(e/(np^2)))$ and so

$$\varepsilon = \varepsilon(n) = 3(2ecnp^2)^{1/(2c)} = 3 \exp\left(-\frac{\log(e/(np^2))}{18 \log n} \log\left(\frac{n \log(e/(np^2))}{18e np^2 \log n}\right)\right)$$

is an increasing function in p (if we think of p as a variable). Hence, writing $\tilde{\varepsilon}$ as the value of ε when p attains its maximum value $n^{-\frac{1}{2}-\omega/\sqrt{\log n}}$,

$$\begin{aligned} 0 < \varepsilon \leq \tilde{\varepsilon} &= 3 \exp\left(-\frac{\omega}{9\sqrt{\log n}}(2\omega\sqrt{\log n} - O(\log \log n))\right) \\ &= 3 \exp(-(2/9 - o(1))\omega^2) = o(1), \end{aligned}$$

and so also $\varepsilon' = o(1)$, and the first assertion follows. On the other hand, it follows immediately from the definition of ε' that $\varepsilon' \geq 1/\log \log(pn) \gg 1/\log(pn)$.

For the second assertion, part (a) is standard. The probability p_0 that the desired property is not satisfied can be estimated as follows:

$$\begin{aligned}
 p_0 &\leq \binom{n}{k} (1-p)^{\binom{k}{2}} \leq \left(\frac{en}{k}\right)^k \exp\left(-p\binom{k}{2}\right) \\
 &= \exp\left(k\left(\log(en/k) - pk/2 + p/2\right)\right) \\
 &\leq \exp\left(k\left(\log(pn) + O(1) - \log\log(pn) - \frac{2+\varepsilon'}{2}\log(pn)\right)\right) \\
 &\leq \exp(-k) = o(1),
 \end{aligned}$$

where the second last step holds for n sufficiently large.

For part (b), for a given m such that $1 \leq m \leq \binom{k}{2} = O(n^2)$, we are going to estimate the probability p_m that there exists a set K of size k with precisely m edges in $G[K]$ and with at least $\lceil m/(2c) \rceil$ edges that belong to edge-disjoint triangles with the third vertex in $V \setminus K$. We have

$$p_m \leq \binom{n}{k} \binom{\binom{k}{2}}{m} p^m (1-p)^{\binom{k}{2}-m} \binom{m}{\lceil m/(2c) \rceil} (np^2)^{\lceil m/(2c) \rceil} \leq \exp(f_m),$$

where

$$\begin{aligned}
 f_m &= k \left(\log\left(\frac{en}{k}\right) - \frac{p(k-1)}{2} \right) + m \left\{ \log \left[\left(\frac{ek^2p}{2m} \right) (2ecnp^2)^{1/(2c)} \right] + p \right\} \\
 &= k \left(\log(pn) + O(1) - \log\log(pn) - \frac{2+\varepsilon'}{2}\log(pn) \right) \\
 &\quad + m \left\{ \log \left[\left(\frac{e(2+\varepsilon')^2 \log^2(pn)}{2mp} \right) \left(\frac{2ec(pn)^2}{n} \right)^{1/(2c)} \right] + p \right\}.
 \end{aligned}$$

Hence, $p_m \leq \exp(-kg_m)$, where

$$g_m = \frac{\varepsilon'}{2} \log(pn) - \frac{mp}{(2+\varepsilon') \log(pn)} \left\{ \log \left[\left(\frac{e(2+\varepsilon')^2 \log^2(pn)}{2mp} \right) \left(\frac{2ec(pn)^2}{n} \right)^{1/(2c)} \right] + p \right\}.$$

Let

$$m' = \left(\frac{e(2+\varepsilon')^2 \log^2(pn)}{2p} \right) \left(\frac{2ecd^2}{n} \right)^{1/(2c)} = o\left(\frac{\log^2(pn)}{p}\right).$$

First, note that if $m > em'$, then

$$g_m \geq \frac{\varepsilon'}{2} \log(pn) - \frac{mp}{(2+\varepsilon') \log(pn)} (-1+p) \geq 1.$$

For $m' \leq m \leq em'$, we get

$$g_m \geq \frac{\varepsilon'}{2} \log(pn) - O\left(\frac{m'p^2}{\log(pn)}\right) = \frac{\varepsilon'}{2} \log(pn) - o(1) \geq 1.$$

Finally, for $m = m'/x$ for any $x = x(n) \geq 1$, since $\max_{x \geq 1} (\log x/x) = 1/e$, we get

$$\begin{aligned} g_m &\geq \frac{\varepsilon'}{2} \log(pn) - \frac{m'p/e}{(2 + \varepsilon') \log(pn)} - o(1) \\ &= \frac{\varepsilon'}{2} \log(pn) - \left(\frac{(2 + \varepsilon') \log(pn)}{2} \right) \left(\frac{2ec(pn)^2}{n} \right)^{1/(2c)} - o(1) \\ &= \log(pn) \left[\frac{\varepsilon'}{2} - (1 + o(1)) \left(\frac{2ec(pn)^2}{n} \right)^{1/(2c)} \right] - o(1) \geq 1. \end{aligned}$$

As a result, $p_m \leq \exp(-k) = o(n^2)$ and so $\sum_{m=0}^{\binom{k}{2}} p_m = o(1)$. The result holds. \square

The last two lemmas give us easily the following lower bound on χ_c that matches (asymptotically) the upper bound. This proves part (a) of Theorem 1.1.

Theorem 4.3. *Let ω be a function tending to infinity with n arbitrarily slowly. Suppose that $p = p(n)$ is such that $\omega < pn \leq n^{1/2 - \omega/\sqrt{\log n}}$. Let $G = (V, E) \in \mathcal{G}(n, p)$. Then, whp $\text{mcf}(G) \leq (2 + o(1)) \log(pn)/p$, and so*

$$\chi_c(G) \geq (1 + o(1)) \frac{pn}{2 \log(pn)} \sim \chi(\mathcal{G}(n, p)).$$

Proof. By (3), we may assume that $pn - \log n \rightarrow \infty$ (or indeed $pn = \Omega(n^{1/3}/\omega)$ for some $\omega = \omega(n) \rightarrow \infty$). Let $k \sim 2 \log(pn)/p$ be defined as in Lemma 4.2. As we aim for a statement that holds whp, we may assume that all properties in Lemma 4.2 and Lemma 4.1 hold deterministically. Consider any colouring that yields $\chi_c(G)$. We will show that no colour class has size at least k which will finish the proof. For a contradiction, suppose that the vertices of some set K of size k are of the same colour. It follows from Lemma 4.2(a) that $G[K]$ contains at least one edge. As no maximal clique is monochromatic, each edge of $G[K]$ belongs to a triangle with the third vertex in $V \setminus K$. As, by Lemma 4.1, no edge belongs to more than c triangles, we may greedily pick a set of edge-disjoint such triangles containing at least a $1/(2c)$ -fraction of the edges in $G[K]$. But this contradicts Lemma 4.2(b). \square

4.2. Lower bounds for parts (b) and (c). We start with the following observation. We will apply the result for $k = 3$ but we state it in a more general case as the proof of this generalization is exactly the same.

Lemma 4.4. *Let $k \geq 3$ be a fixed integer, and let $\varepsilon \in (0, 2/k)$ be a fixed arbitrarily small positive constant. Suppose that $p = p(n)$ satisfies*

$$(3k)^{\frac{2}{k}} (\log n)^{\frac{2}{k(k-1)}} n^{1 - \frac{2}{k}} \leq pn \leq n^{1 - \varepsilon},$$

and let $s = s(n) := 3kp^{-k/2} (\log n)^{1/(k-1)}$. Let $G = (V, E) \in \mathcal{G}(n, p)$. Then, whp the following property holds: every set $S \subseteq V$ of $\lceil s \rceil$ vertices contains at least $\frac{1}{2} \binom{s}{k} p^{\binom{k}{2}} = \Theta\left(s^k p^{\binom{k}{2}}\right)$ copies of K_k .

(Note that the lower bound for pn is there only to make sure that $s \leq n$.)

Proof. For a fixed set $S \subseteq V$ of size $\lceil s \rceil$, let X be the random variable counting the number of copies of K_k in S . Clearly, $\mu = \mathbb{E}[X] = \binom{\lceil s \rceil}{k} p^{\binom{k}{2}}$. Using the notation of Lemma 2.1, we may deal separately with the cases $|A \cap B| = i$ with $2 \leq i \leq k$ to get

$$\begin{aligned} \overline{\Delta} &= \sum_{i=2}^k \binom{\lceil s \rceil}{2k-i} \binom{2k-i}{k} \binom{k}{i} p^{2\binom{k}{2}-\binom{i}{2}} \sim \sum_{i=2}^k \frac{s^{2k-i}}{(k-i)!2^i i!} p^{k(k-1)-\binom{i}{2}} \\ &= s^{2k} p^{k(k-1)} \sum_{i=2}^k \frac{s^{-i}}{(k-i)!2^i i!} p^{-\binom{i}{2}}. \end{aligned}$$

Let us observe that it follows from the definition of s that the i -th term of the sum is of order $(1/p)^{-\binom{i}{2}(k-i+1)} (\log n)^{-i/(k-1)}$. Since $1/p \geq n^\varepsilon$, the logarithmic factor is negligible and it is easy to see that the sum is dominated by the last term ($i = k$). Therefore,

$$\overline{\Delta} \sim \binom{\lceil s \rceil}{k} p^{\binom{k}{2}} = \mu.$$

Applying Lemma 2.1 with $t = \mathbb{E}[X]/2$ we get

$$\begin{aligned} \mathbb{P}\left(X \leq \frac{\mathbb{E}[X]}{2}\right) &\leq \exp\left(- (1 + o(1)) \varphi(-1/2) \mu\right) \leq \exp\left(-\frac{0.15s^k}{k!} p^{\binom{k}{2}}\right) \\ &= \exp\left(-\frac{0.15(3k)^{k-1} s \log n}{k!}\right) \leq \exp(-2s \log n). \end{aligned}$$

Taking a union bound over all $\binom{n}{\lceil s \rceil} \leq \exp(\lceil s \rceil \log n)$ sets of size $\lceil s \rceil$, the desired property holds. \square

We will also need the following property.

Lemma 4.5. *Let $k \geq 3$ be a fixed integer, and let $\varepsilon = \varepsilon(n)$ satisfy $(6/\log n)^{\frac{1}{k-1}} \leq \varepsilon \leq 1/k$. Suppose $p = p(n)$ satisfies $pn \leq n^{k/(k+2)-\varepsilon}$, and let $G \in \mathcal{G}(n, p)$. Then whp no edge in G belongs to as many as $r = \lceil \varepsilon^{-(k-1)} \rceil$ copies of K_{k+1} .*

Proof. Let the graph H be formed from an edge $e = uv$ together with some cliques K_{k+1} containing e (at least one such clique). We will show that we have

$$\frac{e(H) - 1}{v(H) - 2} \geq \frac{k+2}{2}.$$

To see this, we may use induction on the number of cliques in a construction of H . The inequality holds (at equality) if H is a single $(k+1)$ -clique. Suppose that there are at least two cliques in the construction, and the last one added n' new vertices and e' new edges. If $n' = 0$, then the inequality still holds after adding the last clique, so we may assume that $n' \geq 1$. Then

$$e' \geq n'(k+1-n') + \binom{n'}{2} = n' \left(k + \frac{1}{2} - \frac{n'}{2} \right) \geq n' \left(k + \frac{1}{2} - \frac{k-1}{2} \right) = n' \cdot \frac{k+2}{2},$$

where the second inequality follows since $n' \leq k-1$. Thus $\frac{e'}{n'} \geq \frac{k+2}{2}$, and the induction step follows, since from $\frac{a+b}{c+d} \geq \min\{\frac{a}{c}, \frac{b}{d}\}$ for positive a, b, c and d , we may conclude that $\frac{e(H)-1}{v(H)-2} \geq \frac{k+2}{2}$.

The number of copies of K_{k+1} on $2 + \lfloor 1/\varepsilon \rfloor$ vertices that contain the edge uv is at most

$$\binom{\lfloor 1/\varepsilon \rfloor}{k-1} \leq \left(\frac{e(1/\varepsilon)}{k-1} \right)^{k-1} < \varepsilon^{-(k-1)}$$

for $k \geq 4$ (since $e < k-1$); and $\binom{\lfloor 1/\varepsilon \rfloor}{2} < \varepsilon^{-2}$, so the result holds also for $k = 3$. Thus if H contains at least r such cliques we must have $v(H) > 2 + 1/\varepsilon$. Further H must then have a subgraph on at most $2 + r(k-1)$ vertices which contains at least r such cliques. Call the set W of vertices in G *dense* if

$$\frac{e(G[W])}{|W|-2} \geq \frac{k+2}{2}.$$

Let a_i be the expected number of dense sets of size i in G . It will suffice for us to show that

$$\sum_{i=2+\lceil 1/\varepsilon \rceil}^{2+r(k-1)} a_i = o(1). \quad (4)$$

Now

$$\begin{aligned} a_{j+2} &\leq \binom{n}{j+2} \binom{\binom{j+2}{2}}{\lceil j(k+2)/2 \rceil} p^{\lceil j(k+2)/2 \rceil} \\ &\leq \left(\frac{en}{j+2} \right)^{j+2} \left(\frac{e(j+2)(j+1)}{j(k+2)} \cdot p \right)^{\lceil j(k+2)/2 \rceil} \\ &\leq n^2 (np^{(k+2)/2})^j \left(\frac{e}{j+2} \right)^{j+2} \left(\frac{e(j+2)(j+1)}{j(k+2)} \right)^{\lceil j(k+2)/2 \rceil}. \end{aligned}$$

But $np^{(k+2)/2} \leq n^{-\varepsilon(k+2)/2}$. Thus for $j \geq \lceil 1/\varepsilon \rceil$,

$$n^2 (np^{(k+2)/2})^j \leq n^2 (n^{-\varepsilon(k+2)/2})^j \leq n^2 n^{-2\varepsilon j} n^{-(\varepsilon/2)j} \leq n^{-(\varepsilon/2)j} \leq n^{-1/2},$$

as $k \geq 3$. Since there are at most $2 + r(k-1) = O(\varepsilon^{-(k-1)}) = o(n^{1/2})$ terms in the sum in (4), the desired conclusion there follows. \square

Now, finally we are able to prove the lower bounds in parts (b) and (c) of Theorem 1.1.

Theorem 4.6. *Let $k \geq 3$ be a fixed integer, and let $\varepsilon = \varepsilon(n) = (6/\log n)^{\frac{1}{k-1}} = o(1)$. Suppose that $p = p(n)$ satisfies*

$$(3k)^{\frac{2}{k}} (\log n)^{\frac{2}{k(k-1)}} n^{1-\frac{2}{k}} \leq pn \leq n^{k/(k+2)-\varepsilon},$$

and let $s = s(n) := 3kp^{-k/2} (\log n)^{1/(k-1)}$. Let $G = (V, E) \in \mathcal{G}(n, p)$. Then, whp $\text{mcf}(G) \leq s$ and so

$$\chi_c(G) \geq \frac{n}{s} = \Omega \left(\frac{(pn)^{k/2} n^{1-k/2}}{(\log n)^{1/(k-1)}} \right) = \Omega \left(\frac{np^{k/2}}{(\log n)^{1/(k-1)}} \right).$$

Proof. For simplicity we shall ignore the fact that certain numbers should be integers: this never matters here. Let S be a fixed set of size s . We shall show that $\mathbb{P}(S \text{ maximal-clique-free})$ is very small.

If S is maximal-clique-free, then it must be possible to extend each copy J of K_k inside S to K_{k+1} by adding some vertex $v \in V \setminus S$ which is complete to J , that is, v is connected by an edge to each of the k vertices of J . Unfortunately (for our proof) these edges can be reused in extensions for different copies of K_k inside S . Our plan is to show that, if S is maximal-clique-free, then with very high probability there must be a large collection of copies of K_k inside S that are extended to $V \setminus S$ using each edge between S and $V \setminus S$ at most once; and this is very unlikely to happen.

Let $\mu = \binom{s}{k} p^{\binom{k}{2}}$, the expected number of copies of K_k inside S . Let \mathcal{E}_S be the event that S contains at least $\frac{1}{2}\mu$ copies of K_k . Let \mathcal{F} be the event that each edge appears in at most $(1/\varepsilon)^{k-1}$ copies of K_{k+1} . Let \mathcal{F}_S be the event that each edge uv with $u \in S$ and $v \in V \setminus S$ appears in at most $(1/\varepsilon)^{k-1}$ copies of K_{k+1} such that each vertex other than v is in S . Let E_0 be any set of edges within S such that \mathcal{E}_S holds, and condition on $E|_S = E_0$. Here $E|_S$ denotes the set of edges within S in the random edge set E of G . Observe that, since $E|_S$ is fixed, \mathcal{F}_S becomes a down-event in the lattice Ω of subsets of $E(S, V \setminus S)$. Now let $t \geq 1$, and consider a possible ‘edge-disjoint extension of length t ’. Suppose that J_1, \dots, J_t are distinct copies of K_k inside S , and v_1, \dots, v_t are vertices in $V \setminus S$ (not necessarily distinct), such that each possible edge appears at most once as uv_i for some vertex u in J_i . Then the event that v_i is complete to J_i for each $i \in [t]$ is the up-event in Ω that each of the kt relevant edges is present. Hence by Lemma 2.3 (Harris’s inequality)

$$\mathbb{P}(v_i \text{ is complete to } J_i \text{ for each } i \in [t] \mid (E|_S = E_0) \wedge \mathcal{F}_S) \leq p^{kt}. \quad (5)$$

Keep the conditioning on $E|_S = E_0$, condition also on \mathcal{F}_S , and suppose that S is maximal-clique-free. List the copies of K_k inside S in say lexicographic order, discarding any after the first $\frac{1}{2}\mu$; and consider the vertices in $V \setminus S$ in their natural order. Consider the first copy J of K_k on the list: test the vertices in $V \setminus S$ one by one until we find a vertex v complete to J (we must succeed since S is maximal-clique-free). Eliminate from the list the clique J and all other copies J' of K_k such that J' and J have at least one vertex in common and v is complete to J' (so the extensions by v have an edge in common). Since we conditioned on \mathcal{F}_S , at most $k(1/\varepsilon)^{k-1}$ copies of K_k are eliminated from the list in this round, and we move to the next copy of K_k that is left in the list and continue. Since we also conditioned on $E|_S = E_0$, there are at least $\mu/2$ copies of K_k in S , and so the process must last for at least t rounds, where

$$t = \frac{\mu/2}{k(1/\varepsilon)^{k-1}}.$$

Now we lower bound t : we have

$$\begin{aligned} t &= \frac{1}{2k(1/\varepsilon)^{k-1}} \binom{s}{k} p^{\binom{k}{2}} \sim \frac{\varepsilon^{k-1}}{2k \cdot k!} s^k p^{\binom{k}{2}} \\ &\geq \frac{\varepsilon^{k-1}}{2ek^2(k/e)^k} (3kp^{-k/2}(\log n)^{1/(k-1)})^k p^{\binom{k}{2}}, \end{aligned}$$

as $k! \leq ek(k/e)^k$. We get

$$t \geq (1+o(1)) \frac{(e\varepsilon)^{k-1} 3^k}{2k^2} p^{-k/2} (\log n)^{k/(k-1)} > 3\varepsilon^{k-1} k p^{-k/2} (\log n)^{k/(k-1)} = \varepsilon^{k-1} s \log n = 6s,$$

as $3^k e^{k-1} / (2k^3) > 3$ for $k \geq 3$.

The number of sets of t distinct copies J_1, \dots, J_t of K_k inside S which are on the list, and vertices v_1, \dots, v_t yielding possible edge-disjoint extensions as in (5), is at most $\binom{\frac{1}{2}\mu}{t} n^t$. Thus, by (5) and a union bound,

$$\mathbb{P}(S \text{ maximal-clique-free} \mid (E|_S = E_0) \wedge \mathcal{F}_S) \leq \binom{\frac{1}{2}\mu}{t} n^t p^{kt}.$$

Since this holds for every choice E_0 of edges within S such that \mathcal{E}_S holds, and since $\mathcal{F} \subseteq \mathcal{F}_S$, it follows that

$$\mathbb{P}((S \text{ maximal-clique-free}) \wedge \mathcal{E}_S \wedge \mathcal{F}) \leq \binom{\frac{1}{2}\mu}{t} (np^k)^t.$$

We have

$$\begin{aligned} \binom{\frac{1}{2}\mu}{t} (np^k)^t &\leq \exp\left(t \log\left(\frac{e(3e)^k (\log n)^{k/(k-1)} np^{k/2}}{2t}\right)\right) \\ &\leq \exp\left(t \log\left(\frac{e(3e)^k np^k \log n}{36k}\right)\right) \\ &\leq \exp\left(-t \left(\frac{2k}{k+2} + \varepsilon k - 1 + o(1)\right) \log n\right) \\ &\leq \exp\left(-\frac{t}{5+o(1)} \log n\right) \end{aligned}$$

which is $o(n^{-s})$ since $t \geq (6+o(1))s$. Let us rename the fixed set S as S_0 . Then, with the union below being over all s -subsets S of vertices,

$$\begin{aligned} &\mathbb{P}(\vee_{|S|=s} ((S \text{ maximal-clique-free}) \wedge \mathcal{E}_S \wedge \mathcal{F})) \\ &\leq \binom{n}{s} \mathbb{P}((S_0 \text{ maximal-clique-free}) \wedge \mathcal{E}_{S_0} \wedge \mathcal{F}) \\ &\leq \binom{n}{s} \binom{\frac{1}{2}\mu}{t} (np^k)^t = o(1). \end{aligned}$$

Finally, let $\mathcal{E} = \wedge_{|S|=s} \mathcal{E}_S$. Then

$$\mathbb{P}(\text{mcf}(G) \geq s) \leq \mathbb{P}(\vee_{|S|=s} ((S \text{ maximal-clique-free}) \wedge \mathcal{E}_S \wedge \mathcal{F})) + \mathbb{P}(\bar{\mathcal{E}}) + \mathbb{P}(\bar{\mathcal{F}}).$$

But each of the three terms in the upper bound here is $o(1)$. We have just seen this for the first term: the second and third terms are $o(1)$ by Lemmas 4.4 and 4.5 respectively. The proof of the theorem is complete. \square

4.3. Lower bounds for parts (d)-(g). Let us start with the following lemma, which is proved in a similar way to Lemma 4.4.

Lemma 4.7. *Let $k \geq 3$ be a fixed integer, and let $\varepsilon > 0$ be a fixed arbitrarily small positive constant. Let $C = C(n)$ be such that $1 \leq C = n^{o(1)}$. Suppose that $p = p(n)$ satisfies*

$$(3k)^{2/(k+2)}(\log n)^{\frac{2}{(k-1)(k+2)}} n^{\frac{k}{k+2}} C^{-2/(k+2)} \leq pn \leq n^{1-\varepsilon}.$$

Let $s = s(n) = Cpn$, and let $G = (V, E) \in \mathcal{G}(n, p)$. Then, whp the following property holds: every set $S \subseteq V$ of $\lceil s \rceil$ vertices contains at least $\frac{1}{2} \binom{\lceil s \rceil}{k} p^{\binom{k}{2}} = \Theta\left(s^k p^{\binom{k}{2}}\right)$ copies of K_k .

Proof. As before, for simplicity we shall ignore the fact that certain numbers should be integers. For a fixed set $S \subseteq V$ of size $\lceil s \rceil$, let X be the random variable counting the number of copies of K_k in S . Clearly, $\mu = \mathbb{E}[X] = \binom{\lceil s \rceil}{k} p^{\binom{k}{2}}$. Using the notation of Lemma 2.1, we may independently deal with the cases $|A \cap B| = i$ with $2 \leq i \leq k$ to get

$$\begin{aligned} \bar{\Delta} &= \sum_{i=2}^k \binom{\lceil s \rceil}{2k-i} \binom{2k-i}{k} \binom{k}{i} p^{2\binom{k}{2}-\binom{i}{2}} \sim \sum_{i=2}^k \frac{s^{2k-i}}{(k-i)!2^i i!} p^{k(k-1)-\binom{i}{2}} \\ &= s^{2k} p^{k(k-1)} \sum_{i=2}^k \frac{s^{-i}}{(k-i)!2^i i!} p^{-\binom{i}{2}}. \end{aligned} \quad (6)$$

For $i = 2, 3, \dots, k-1$, let $a_i = s^{-i} p^{-\binom{i}{2}}$ and note that the i th term in the last sum is $\Theta(a_i)$. The ratio $a_{i+1}/a_i = s^{-1} p^{-i} = \frac{1}{Cn} p^{-i-1}$, which is increasing with i . It follows that a_2, \dots, a_k is a unimodal sequence, with maximum either a_2 or a_k . Moreover, since $1/p \geq n^\varepsilon$, at most two consecutive terms can be of the same order. As a result, the sum is of order the larger of the term $i = 2$ and the term $i = k$. (Note that we did not rule out the possibility that the $(k-1)$ -st term is of order of the k -th term yet.) More precisely, comparing the two terms, the sum is of order of the term $i = k$ if $p \ll n^{-2/(k+3)} C^{-2/(k+3)}$, and of order of the term $i = 2$ otherwise. Moreover, in the case $p \ll n^{-2/(k+3)} C^{-2/(k+3)}$, by comparing the $(k-1)$ -st and the k -th term, we see that the k -th term dominates the $(k-1)$ -st term if $p \ll (Cn)^{-1/k}$, and since $n^{-2/(k+3)} C^{-2/(k+3)} \leq (Cn)^{-1/k}$ for $k \geq 3$, this condition is satisfied. Hence, if $p \ll n^{-2/(k+3)} C^{-2/(k+3)}$, the sum is in fact asymptotic to the term $i = k$.

Suppose first that $p \ll n^{-2/(k+3)} C^{-2/(k+3)}$. As we already mentioned, in this case $\bar{\Delta} \sim \mu$, and so applying Lemma 2.1 with $t = \mathbb{E}[X]/2$ we get

$$\begin{aligned} \mathbb{P}\left(X \leq \frac{\mathbb{E}[X]}{2}\right) &\leq \exp\left(- (1 + o(1)) \varphi(-1/2) \mu\right) \leq \exp\left(- \frac{0.15(Cpn)^k}{k!} p^{\binom{k}{2}}\right) \\ &\leq \exp\left(- \frac{0.15(3k)^{k-1} (Cpn) \log n}{k!}\right) \leq \exp(-2s \log n), \end{aligned}$$

by our assumption that $pn \geq (3k)^{2/(k+2)} (\log n)^{\frac{2}{(k-1)(k+2)}} n^{\frac{k}{k+2}} C^{-2/(k+2)}$. Taking a union bound over all $\binom{n}{\lceil s \rceil} \leq \exp(\lceil s \rceil \log n)$ sets of size $\lceil s \rceil$, the desired property holds.

Assume now that $p = \Omega(n^{-2/(k+3)}C^{-2/(k+3)})$. In this case we get

$$\bar{\Delta} = \Theta(\mu^2 p^{-1} s^{-2}).$$

Therefore, applying Lemma 2.1 with $t = \mathbb{E}[X]/2$ we get

$$\mathbb{P}\left(X \leq \frac{\mathbb{E}[X]}{2}\right) \leq \exp(-\Omega(\mu^2/\bar{\Delta})) = \exp\left(-\Omega\left(\frac{sC(pn)^2}{n}\right)\right) \leq \exp(-2s \log n),$$

as, by our assumption on p and C , we have $C(pn)^2/n \geq n^{2k/(k+2)-1+o(1)} \geq n^{1/5+o(1)} \gg \log n$. Taking a union bound over all $\binom{n}{\lceil s \rceil} \leq \exp(\lceil s \rceil \log n)$ sets of size $\lceil s \rceil$, the desired property holds. \square

We can now prove the following theorem that implies the lower bounds in parts (e) and (g) of Theorem 1.1.

Theorem 4.8. *Let $k \geq 3$ be a fixed integer. Suppose that $p = p(n)$ satisfies*

$$(\log n)^{\frac{2k}{(k-1)(k+2)}} n^{\frac{k}{k+2}} \leq pn \leq (4e^2 k \log^2 n)^{-1/k} n^{(k-1)/k}.$$

Let $s = s(n) = Cp$, where

$$C = C(n) = \left(\frac{32k^2 \log n}{-\log(p^k n 4ek \log^2 n)} \right)^{1/(k-1)}. \quad (7)$$

Let $G = (V, E) \in \mathcal{G}(n, p)$. Then, whp $\text{mcf}(G) \leq s$, and so

$$\chi_c(G) \geq \frac{n}{s} = \frac{1}{Cp} = \frac{n^{o(1)}}{p}.$$

Before we move to the proof of this result, let us make a few comments. First, obviously, if there are several values of k that satisfy the assumptions on pn , one should consider the one that gives the best lower bound for $\chi_c(G)$. In particular, the case $k = 3$ covers values of pn between $n^{3/5}(\log n)^{3/5}$ and $n^{2/3}(12e^2 \log^2 n)^{-1/3}$, and the case $k = 4$ covers the range of pn between $n^{2/3}(\log n)^{4/9}$ and $n^{3/4}(16e^2 \log^2 n)^{-1/4}$. The case $k = 5$ can be applied already if pn is at least $n^{5/7}(\log n)^{5/14}$, and so for values of pn in the interval $[n^{5/7}(\log n)^{5/14}, n^{3/4}(16e^2 \log^2 n)^{-1/4}]$ both $k = 4$ and $k = 5$ satisfy the conditions of the theorem. Similarly, if we fix $\varepsilon > 0$ (arbitrarily small), all values of pn that belong to the interval $[n^{5/7}(\log n)^{5/14}, n^{1-\varepsilon}]$ are covered by at least two values of k at most the constant $2/\varepsilon$ (since $p \geq n^{-\frac{2}{k+2}}$, for $p \leq n^{-\varepsilon}$ we are concerned only with k such that $n^{-\frac{2}{k+2}} \leq n^{-\varepsilon}$, that is $k \leq 2/\varepsilon - 2 \leq 2/\varepsilon$). If p is in the interval for $k - 1$ then $p \leq n^{-1/(k-1)} \leq n^{-1/k-\eta}$, for some $\eta \geq 1/(k-1) - 1/k = 1/k(k-1) \geq \varepsilon^2/4$. Thus for at least one of the relevant values k (any one except the smallest) we have $pn \leq n^{(k-1)/k-\eta}$, and so $C \leq (32k/(\eta + o(1)))^{1/(k-1)} = O(1)$. This in turn implies that $\text{mcf}(G) = O(pn)$ and so $\chi_c(G) = \Omega(1/p)$. The only range not covered by this theorem is when $pn = n^{3/5+o(1)}$ or $pn = n^{2/3+o(1)}$ which will be done separately later on.

Proof. The proof is similar to the proof of Theorem 4.6. However, one additional idea is needed: informally speaking, we now have too many copies of K_k , but still only $\Theta(s)$ of them are going to be extended to K_{k+1} using disjoint edge sets. The union bound

over all choices of these would be too large to be successfully applied, and so we focus on a randomly chosen subset of the K_k 's. To this end, define

$$L := \left\lceil \frac{n^{k-1} p^{(k+2)(k-1)/2}}{(k-1)! (\log n)^2} \right\rceil \geq \frac{(\log n)^{k-2}}{(k-1)!}.$$

Now, in order to obtain the desired smaller subset of cliques, we do the following, independently of the random graph G : for each set of k vertices, we independently colour it *red* with probability $1/L$. Our first goal is to show that no edge uv belongs to too many cliques of size $k+1$ which after removal of v form a red k -set.

Let $\eta = \eta(n) = o(1)$, say $\eta = n^{-\frac{1}{11}}$. For each pair of distinct vertices u and v , let $\mathcal{E}_{u,v}$ be the event that

$$||N(u) \cap N(v)| - np^2| \leq \eta np^2. \quad (8)$$

Let \mathcal{E} be the event that $\mathcal{E}_{u,v}$ holds for each pair $u \neq v$. Since $\eta^2 np^2 \geq n^{1/55}$, by the Chernoff bounds (1) and (2), $\mathbb{P}(\bar{\mathcal{E}}_{u,v}) = o(n^{-2})$; and so by a union bound, whp \mathcal{E} holds.

Fix any ordered pair of distinct vertices u, v , and any set $I \subseteq V \setminus \{u, v\}$ with $|I| = k-1$. Let $X_{u,v,I} = 1$ if I induces a clique, u and v are both complete to I (we do not require u and v to be adjacent). Finally, let $X_{u,v} = \sum_{I \subseteq V \setminus \{u,v\}} X_{u,v,I}$. Let $m = m(n) = \lfloor (1+\eta)np^2 \rfloor$. Let Z be the number of $(k-1)$ -cliques in the random graph with vertex set $[m]$ and edge-probability $p(n)$. Clearly, conditional on $\mathcal{E}_{u,v}$, $X_{u,v}$ is stochastically at most Z . We have

$$\mathbb{E}Z = \binom{m}{k-1} p^{\binom{k-1}{2}} \sim L \log^2 n \geq \frac{\log^k n}{(k-1)!} \gg (\log n)^{k-2}. \quad (9)$$

Hence, by Lemma 2.2, once n is sufficiently large that $\mathbb{E}Z \leq (4/3)L \log^2 n$,

$$\begin{aligned} \mathbb{P}(X_{u,v} \geq 2L \log^2 n \mid \mathcal{E}_{u,v}) &\leq \mathbb{P}(Z \geq (3/2)\mathbb{E}Z) \\ &\leq \exp\left(-\Omega((\mathbb{E}Z)^{1/(k-2)})\right) \\ &= \exp\left(-\omega(\log n)\right) = o(n^{-2}). \end{aligned}$$

Thus

$$\mathbb{P}(X_{u,v} \geq 2L \log^2 n) \leq \mathbb{P}(X_{u,v} \geq 2L \log^2 n \mid \mathcal{E}_{u,v}) + \mathbb{P}(\bar{\mathcal{E}}_{u,v}) = o(n^{-2}).$$

By taking a union bound over all ordered pairs of distinct vertices, we see that whp for each $u \neq v$ we have $X_{u,v} \leq 2L \log^2 n$. Now let $Y_{u,v,I}$ be defined as $X_{u,v,I}$ with the additional condition that $I \cup \{u\}$ is red, and let $Y_{u,v} = \sum_I Y_{u,v,I} \leq X_{u,v}$. Clearly, conditional on $X_{u,v} \leq 2L \log^2 n$, $Y_{u,v}$ is stochastically bounded above by the binomial random variable $\text{Bin}(2L \log^2 n, 1/L)$ with expectation $2 \log^2 n$. It is a straightforward application of (2) (together with a union bound over all ordered pairs u, v) to see that whp for each pair $u \neq v$ we have $Y_{u,v} \leq 4 \log^2 n$. Let \mathcal{D} be the event that this property holds, so that \mathcal{D} holds whp. This completes the first part of the proof.

For the rest of the proof, we argue very much as in the proof of Theorem 4.6. As before, we ignore rounding issues for s , since this does not matter. Let $s = Cpn$ and let S be a set of size s . Let \mathcal{D}_S be the event that each edge uv with $u \in S$ and $v \in V \setminus S$ appears in at most $4 \log^2 n$ copies of K_{k+1} which are such that the k vertices other than

v form a red k -set in S . Note that $\mathcal{D} \subseteq \mathcal{D}_S$. Let $\mu = \binom{s}{k} p^{\binom{k}{2}}$, the expected number of k -cliques in S ; so μ/L is the expected number of red k -cliques in S . Let \mathcal{H}_S be the event that S contains at least $\mu/(4L)$ red cliques of size k ; and let $\mathcal{H} = \bigwedge_{|S'|=s} \mathcal{H}_{S'}$. (We shall check later that \mathcal{H} holds whp)

Denote the random colouring of the k -sets of vertices by R ; and let $R|_S$ denote its restriction to the k -sets in S . Let E_0 be any set of edges within S and let R_0 be any colouring of the k -sets within S such that \mathcal{H}_S holds when $E|_S = E_0$ and $R|_S = R_0$. Condition on $E|_S = E_0$ and $R|_S = R_0$. Let $t \geq 1$, and consider a possible ‘edge-disjoint extension of length t ’. Suppose that J_1, \dots, J_t are distinct red cliques of size k inside S , and v_1, \dots, v_t are vertices in $V \setminus S$ (not necessarily distinct), such that each possible edge appears at most once as uv_i for some vertex u in J_i . Then, as in (5),

$$\mathbb{P}(v_i \text{ is complete to } J_i \text{ for each } i \in [t] \mid (E|_S = E_0) \wedge (R|_S = R_0) \wedge \mathcal{D}_S) \leq p^{kt}. \quad (10)$$

Also, as before, if S is maximal-clique-free, then by considering a list of red k -cliques of length $\mu/4L$, there must be a disjoint extension of length $t := (\mu/4L)/(4k \log^2 n)$. Hence, by a union bound,

$$\mathbb{P}(S \text{ maximal-clique-free} \mid (E|_S = E_0) \wedge (R|_S = R_0) \wedge \mathcal{D}_S) \leq \binom{\mu/4L}{t} n^t p^{kt}.$$

Since this holds for each choice of E_0 and R_0 such that \mathcal{H}_S holds,

$$\mathbb{P}(S \text{ maximal-clique-free} \mid \mathcal{H}_S \wedge \mathcal{D}_S) \leq \binom{\mu/4L}{t} (np^k)^t.$$

Note that

$$\begin{aligned} \binom{\mu/4L}{t} (np^k)^t &\leq (4eknp^k \log^2 n)^t \\ &= \exp\left(\frac{\mu}{16kL \log^2 n} \log(4eknp^k \log^2 n)\right) \\ &= \exp\left((1+o(1)) \frac{sC^{k-1}}{16k^2} \log(4eknp^k \log^2 n)\right). \end{aligned}$$

Plugging in the value of C , we see that this expectation is at most $\exp(-(2+o(1))s \log n)$. Hence

$$\binom{n}{s} \mathbb{P}(S \text{ maximal-clique-free} \wedge \mathcal{H}_S \wedge \mathcal{D}_S) = o(1).$$

But, letting $\mathcal{H} = \bigwedge_{|S|=s} \mathcal{H}_S$,

$$\mathbb{P}(\text{mcf}(G) \geq s) \leq \binom{n}{s} \mathbb{P}(S \text{ maximal-clique-free} \wedge \mathcal{H}_S \wedge \mathcal{D}_S) + \mathbb{P}(\bar{\mathcal{H}}) + \mathbb{P}(\bar{\mathcal{D}}).$$

We have just seen that the first term in the upper bound is $o(1)$, and we noted earlier that \mathcal{D} holds whp; so it remains to show that $\mathbb{P}(\bar{\mathcal{H}}) = o(1)$.

Let S_0 be a fixed set of s vertices, and let S now be an arbitrary such set. Let \mathcal{G}_S be the event that S contains at least $\frac{1}{2}\mu = \frac{1}{2} \binom{s}{k} p^{\binom{k}{2}}$ copies of K_k ; and let $\mathcal{G} = \bigwedge_{|S|=s} \mathcal{G}_S$.

By Lemma 4.7 we have $\mathbb{P}(\bar{\mathcal{G}}) = o(1)$. Also, by the Chernoff bound (1),

$$\mathbb{P}(\bar{\mathcal{H}}_S \mid \mathcal{G}_S) = \exp(-\Omega(\mu/L)) = \exp(-\Omega(s \log^2 n));$$

and so, by a union bound,

$$\mathbb{P}(\bigvee_{|S|=s}(\bar{\mathcal{H}}_S \wedge \mathcal{G}_S)) \leq \binom{n}{s} \mathbb{P}(\bar{\mathcal{H}}_{S_0} \mid \mathcal{G}_{S_0}) = o(1).$$

But

$$\mathbb{P}(\bar{\mathcal{H}}) = \mathbb{P}(\bigvee_{|S|=s} \bar{\mathcal{H}}_S) \leq \mathbb{P}(\bigvee_{|S|=s}(\bar{\mathcal{H}}_S \wedge \mathcal{G}_S)) + \mathbb{P}(\bar{\mathcal{G}}).$$

and so $\mathbb{P}(\bar{\mathcal{H}}) = o(1)$, as required. This completes the proof of the theorem. \square

Finally, we deal with the missing gaps when $pn = n^{3/5+o(1)}$ or $pn = n^{2/3+o(1)}$ which will finish the lower bounds in parts (d) and (f) of Theorem 1.1.

Theorem 4.9. *Suppose that $p = p(n)$ satisfies*

- (a) $n^{3/5-6/(\log n)^{1/2}} \leq pn \leq n^{3/5}(\log n)^{3/5}$, or
- (b) $pn = n^{2/3+o(1)}$ and $pn \leq n^{2/3}(\log n)^{4/9}$.

Let $G = (V, E) \in \mathcal{G}(n, p)$. Then, whp,

$$\chi_c(G) \geq \frac{n^{o(1)}}{p}.$$

Sketch of the proof. Let us focus on part (a) first. Since the proof of this theorem is almost identical to the one of Theorem 4.8 (with $k = 3$), we only mention two technical issues that are relatively easy to deal with. First, note that in order to apply Lemma 4.7, we cannot keep C as defined in (7); this time, we set

$$C = \max \left(\left(\frac{32k^2 \log n}{-\log(p^k n 4ek \log^2 n)} \right)^{1/(k-1)}, 3k(\log n)^{1/(k-1)}(pn)^{-(k+2)/2} n^{k/2} \right).$$

We do not necessarily have $C = O((\log n)^{1/2})$ anymore but still it is the case that $1 \leq C = n^{o(1)}$. (Recall that $k = 3$ in the proof of part (a).) This time there is no need to reduce the number of cliques for the union bound to work, so we may keep $L = 1$, which simplifies the argument slightly. The last difference is with the application of Lemma 2.2. This time the expected value of $X_{u,v} = X_{u,v}(p)$ (see (9)) is not large enough for the lemma to be directly applied. However, as standard in such situations, one can increase the probability p to some value $p' \geq p$ for which the expected value of $X_{u,v}(p')$ is of order greater than $(\log n)^{k-2}$ (which equals $\log n$ in the case $k = 3$ we deal with in part(a)) as required. For such a value p' , it follows that whp for any pair u, v we have $X_{u,v}(p') \leq 2 \log^2 n$, and by standard coupling arguments the same holds for p . The rest of the proof is not affected.

Exactly the same adjustments are required for part (b), this time with $k = 4$. \square

5. CONCLUDING REMARKS

Let us pick up two points for further thought.

- We investigated the clique colouring number $\chi_c(G)$ for random graphs $G \in \mathcal{G}(n, p)$, and in Theorem 1.1 we obtained fairly good estimates for values of $p = p(n)$, other than $p = n^{-\frac{1}{2}+o(1)}$ where χ_c drops dramatically as p increases. By parts (a) and (c), for suitable $\varepsilon(n) = o(1)$ (going to 0 slowly), if $p = n^{-\frac{1}{2}-\varepsilon}$ then $\chi_c(G) = n^{\frac{1}{2}-o(1)}$ whp; whereas if $p = n^{-\frac{1}{2}+\varepsilon}$ then $\chi_c(G) = n^{\frac{1}{4}+o(1)}$ whp. For intermediate values of p , all we say in part (b) is that $\chi_c(G)$ lies whp between the values $n^{\frac{1}{4}-o(1)}$ and $n^{\frac{1}{2}+o(1)}$. It would be interesting to learn more about this jump.
- A second natural question for random graphs $G \in \mathcal{G}(n, p)$ concerns the dense case, when p is a constant with $0 < p < 1$. We have seen that $\chi_c(G)$ is $O(\log n)$ whp but what about a lower bound?

REFERENCES

- [1] T. Andreae, M. Schughart, Zs. Tuza, Clique-transversal sets of line graphs and complements of line graphs, *Discrete Math.* **88**, 1991, 11–20.
- [2] G. Bacsó, S. Gravier, A. Gyárfás, M. Preissmann, A. Sebő, Coloring the maximal cliques of graphs, *SIAM Journal on Discrete Mathematics* **17**, 2004, 361–376.
- [3] C.N. Campos, S. Dantas, C.P. de Mello, Colouring clique-hypergraphs of circulant graphs. *Electron. Notes Discret. Math.* **30**, 2008, 189–194.
- [4] M.R. Cerioli, A.L. Korenchender, Clique-coloring Circular-Arc graphs. *Electron. Notes Discret. Math.* **35**, 2009, 287–292.
- [5] D. Défossez, Clique-coloring some classes of odd-hole-free graphs, *J. Graph Theory* **53**, 2006, 233–249.
- [6] R.M. del Rio-Chanona, C. MacRury, J. Nicolaidis, X. Perez-Gimenez, P. Prałat, and K. Ternovsky, Injective colouring of binomial random graphs, preprint.
- [7] R. DeMarco, J. Kahn, Tight upper-tail bounds for cliques, *Random Structures & Algorithms* **41**(4), 2012, 469–487.
- [8] A. Dudek, D. Mitsche, and P. Prałat, The set chromatic number of random graphs, *Discrete Applied Mathematics* **215**, 2016, 61–70.
- [9] S. Gravier, C. Hoàng, F. Maffray, Coloring the hypergraph of maximal cliques of a graph with no long path, *Discrete Math.*, **272**, 2003, 285–290.
- [10] A. Frieze, D. Mitsche, X. Pérez-Giménez, P. Prałat, On-line list colouring of random graphs, *Electronic Journal of Combinatorics* **22**(2), 2015, #P2.41.
- [11] T.E. Harris, A lower bound for the critical probability in a certain percolation, *Proceedings of the Cambridge Philosophical Society*, **56**, 1960, 13–20.
- [12] S. Janson, T. Łuczak, A. Ruciński, *Random Graphs*, Wiley, New York, 2000.
- [13] S. Janson, A. Ruciński, The deletion method for upper tail estimates, *Combinatorica*, **24** (4) (2004), 615–640.
- [14] R.J. Kang, C. McDiarmid, Colouring random graphs, in L. Beineke and R.J. Wilson, eds, *Topics in Chromatic Graph Theory*, 2015.
- [15] S. Klein, A. Morgana, On clique-colouring with few P_4 's, *J. Braz. Comput. Soc.* **18**, 2012, 113–119.
- [16] J. Kratochvíl, Zs. Tuza, On the complexity of bicoloring clique hypergraphs of graphs, *J. Algorithms*, **45**, 2002, 40–54.
- [17] M. Krivelevich, B. Sudakov, V.H. Vu, N.C. Wormald, On the probability of independent sets in random graphs, *Random Structures & Algorithms* **22** (1), 2003, 1–14.

- [18] Z. Liang, E. Shan, L. Kang, Clique-coloring claw-free graphs, *Graphs and Combinatorics*, DOI10.1007/s00373-015-1657-8, 2015, 1–16.
- [19] T. Luczak, The chromatic number of random graphs, *Combinatorica* **11**(1), 1991, 45–54.
- [20] M. Luczak, C. McDiarmid, E. Upfal, On-line routing of random calls in networks, *Probability Theory and Related Fields* **125**, 2003, 457–482.
- [21] C.J.H. McDiarmid, On the chromatic number of random graphs, *Random Structures & Algorithms* **1**(4), 1990, 435–442.
- [22] C.J.H. McDiarmid, D. Mitsche, P. Pralat, Clique colouring of geometric graphs, in preparation.
- [23] B. Mohar, R. Skrekovski, The Grötzsch Theorem for the hypergraph of maximal cliques. *Electronic J. Comb.* **6**, 1999, R26.
- [24] E. Shan, Z. Liang, L. Kang, Clique-transversal sets and clique-coloring in planar graphs, *European J. Combin.* **36**, 2014, 367–376.
- [25] V.H. Vu, On some degree conditions which guarantee the upper bound of chromatic (choice) number of random graphs, *Journal of Graph Theory*, **31**, 1999, 201–226.
- [26] V. Vu, A large deviation result on the number of small subgraphs of a random graph, *Combinatorics, Probability and Computing* **10**(1), 2001, 79–94.

DEPARTMENT OF STATISTICS, OXFORD UNIVERSITY, 24 - 29 ST GILES, OXFORD OX1 3LB, UK
E-mail address: cmcd@stats.ox.ac.uk

UNIVERSITÉ DE NICE SOPHIA-ANTIPOLIS, LABORATOIRE J-A DIEUDONNÉ, PARC VALROSE,
06108 NICE CEDEX 02
E-mail address: dmitsche@unice.fr

DEPARTMENT OF MATHEMATICS, RYERSON UNIVERSITY, TORONTO, ON, CANADA AND THE
FIELDS INSTITUTE FOR RESEARCH IN MATHEMATICAL SCIENCES, TORONTO, ON, CANADA
E-mail address: pralat@ryerson.ca