1 Introduction

We consider the $t$-improper chromatic number of the Erdős-Rényi random graph $G_{n,p}$. As usual, $G_{n,p}$ denotes a random graph with vertex set $[n] = \{1, \ldots, n\}$ in which the edges are included independently at random with probability $p$. A set of vertices is $t$-dependent if it induces a subgraph of maximum degree at most $t$. The $t$-improper chromatic number $\chi^t(G)$ is the smallest number of colours needed in a $t$-improper colouring — a colouring of the vertices in which colour classes are $t$-dependent sets. When $t = 0$, we are simply considering the usual notion of proper colouring.

The $t$-improper chromatic number was introduced about twenty years ago independently by Andrews and Jacobson [1], Harary and Fraughnaugh (née Jones) [7,8], and Cowen et al [4]. In the first paper, the authors considered
various general lower bounds for the $t$-improper chromatic number; in the second, the authors studied $\chi^t$ as part of the larger setting of generalised chromatic numbers; in the third, the authors established best upper bounds on $\chi^t$ for planar graphs to generalise the Four Colour Theorem.

The chromatic number of random graphs has been well studied. Fix $0 < p < 1$ and let $\gamma = 2 / \log \frac{1}{1-p}$. In 1975, Grimmett and McDiarmid [6] showed that, for any $\varepsilon > 0$, the expected number $C_{n,j}$ of $j$-colourings of $G_{n,p}$ satisfies the following asymptotically almost surely (a.a.s.):

$$C_{n,j} \to \begin{cases} 0 & \text{if } j \leq (1 - \varepsilon) \frac{n}{\gamma \log n} \\ \infty & \text{if } j \geq (1 + \varepsilon) \frac{n}{\gamma \log n} \end{cases}$$  (1)

thus showing in particular that $\chi(G_{n,p}) \geq (1 - \varepsilon) \frac{n}{\gamma \log n}$ a.a.s. Furthermore, they conjectured that $\chi(G_{n,p}) \sim \frac{n}{\gamma \log n}$ a.a.s. This conjecture remained a tantalising open problem for over a decade, until Bollobás [2] and Matula and Kučera [12] used martingale techniques to establish the conjecture. Luczak [11] extended the result to sparse random graphs. For further background into the colouring of random graphs, consult [3,9].

For any graph $G$, clearly $\chi^t(G) \leq \chi(G)$. Also $\chi(G) \leq (t + 1)\chi^t(G)$, since any $t$-dependent set of vertices can be properly coloured with $(t + 1)$ colours. Furthermore, we make use of a result of Lovász [10] to conclude the following. ($\Delta(G)$ denotes the maximum degree of $G$.)

Proposition 1.1 For any graph $G$ and integer $t$,

$$\frac{\chi(G)}{t+1} \leq \chi^t(G) \leq \min \left\{ \left\lceil \frac{\Delta(G)+1}{t+1} \right\rceil, \chi(G) \right\}.$$

In the dense case, i.e. when the edge probability $p$ is constant, we show that $\chi^t(G_{n,p})$ is likely to be close to the upper end of this range, as long as $t(n) = o(\log n)$ or $t(n) = \omega(\log n)$. Recall that $\Delta(G_{n,p}) \sim np$ a.a.s. in this case. More fully, we have

Theorem 1.2 For fixed edge probability $0 < p < 1$, the following hold:

(a) if $t(n) = o(\log n)$, then $\chi^t(G_{n,p}) \sim \left( \frac{1}{2} \log \frac{1}{1-p} \right) \frac{n}{\log n}$ a.a.s.;

(b) if $t(n) = \Theta(\log n)$, then $\chi^t(G_{n,p}) = \Theta \left( \frac{n}{\log n} \right) = \Theta \left( \frac{np}{t} \right)$ a.a.s.;

(c) if $t(n) = \omega(\log n)$ and $t(n) = o(n)$, then $\chi^t(G_{n,p}) \sim \frac{np}{t}$ a.a.s.;

(d) if $t(n)$ satisfies $\frac{np}{t} \to x$, where $0 < x < \infty$ and $x$ is not integral, then $\chi^t(G_{n,p}) = \lceil x \rceil$ a.a.s.
Let us more closely examine part (b), where the above theorem is imprecise. Suppose that \( t(n) \sim \tau \log n \) for some fixed \( \tau > 0 \). We identify a constant \( \kappa > 0 \) (depending on \( p \) and \( \tau \)) such that the expected number of \( t \)-improper \( j \)-colourings of \( G_{n,p} \) tends to 0 if \( j \leq (1 - \varepsilon) \frac{n}{\kappa \log n} \) and to \( \infty \) if \( j \geq (1 + \varepsilon) \frac{n}{\kappa \log n} \). Note that this result is analogous to (and extends) (1). It is natural now to conjecture that \( \chi^t(G_{n,p}) \sim \frac{n}{\gamma \log n} \) a.a.s. We later describe this conjecture in more detail.

Throughout the paper, unless stated otherwise, we assume that \( 0 < p < 1 \) is fixed.

2 Proof of Theorem 1.2

In this section, we give almost the entire proof of the main theorem; the proof of part (b) is deferred to the full version of the paper. It follows from Proposition 1.1 that we need only establish lower bounds on \( \chi^t(G_{n,p}) \). We will do this by estimating the \( t \)-dependence number \( \alpha^t \) — the size of a largest \( t \)-dependent set — and then applying the straightforward observation that \( \chi^t(G) \geq |V(G)|/\alpha^t(G) \) for any graph \( G \).

Lemma 2.1 If \( t(n) = o(\log n) \), then \( \chi^t(G_{n,p}) \sim \frac{n}{\gamma \log n} \) a.a.s., where \( \gamma = \frac{2}{\log \frac{1}{1-p}} \).

Proof. Since the result holds for \( t = 0 \), we assume \( t \geq 1 \). Fix \( \varepsilon > 0 \) and set \( k = k(n) = \lceil \frac{1}{1-\varepsilon} \gamma \log n \rceil \). We shall show that the expected number of \( t \)-dependent \( k \)-sets approaches zero.

Let \( g(k, t) \) denote the number of graphs on \( [k] = \{1, \ldots, k\} \) with average degree at most \( t \). Clearly, the number of graphs on \( [k] \) with maximum degree at most \( t \) is at most \( g(k, t) \). Now, since a graph on \( k \) vertices with average degree \( d' \) has \( d'k/2 \) edges, it is clear that

\[
g(k, t) \leq \sum_{s=0}^{s_0} \binom{k}{s},
\]

putting \( s_0 = \lceil \frac{tk}{2} \rceil \). If \( t \leq \frac{k-1}{3} \), then \( \frac{tk}{2} \leq \frac{1}{3} \left( \binom{k}{2} + 1 \right) \); furthermore, if \( s \leq \frac{1}{3} (x + 1) \), then \( \binom{x}{s} = \frac{s}{x-s+1} \binom{x}{s} \leq \frac{1}{2} \binom{x}{s} \); therefore, since \( t(n) = o(k) \), it follows that for large enough \( n \)

\[
g(k, t) \leq \sum_{s=0}^{s_0} 2^{-s} \binom{k}{s_0} \leq 2 \binom{k}{s_0} \leq 2 \left( \frac{e(k-1)}{t} \right)^{\frac{tk}{2}}.
\]

The last inequality is due to a well-known consequence of Stirling’s formula: \( \binom{x}{s} \leq \left( \frac{ex}{s} \right)^s \).
Lemma 2.2
Fix \((b)\) is the subject of the next section.

Proof. Again, we want to estimate the maximum size of edges). Now let \(k\) is a graph on \([k]\) with maximum degree at most \(t\), the expected number of \(t\)-improper \(k\)-sets in \(G_{n,p}\) is at most

\[
(n)_k q^\left(\frac{k-1}{2}\right) - \frac{tk}{q} g(k, t) \leq \left(\frac{en}{k} q^\left(\frac{k-1}{2}\right) 2^\left(\frac{k}{t}\right) \left(\frac{ek}{t}\right)^\frac{1}{2}\right)^k.
\]

where \(q = 1-p\). Let us examine the expression \(B = \frac{en}{k} q^\left(\frac{k-1}{2}\right) 2^\left(\frac{k}{t}\right) \left(\frac{ek}{t}\right)^\frac{1}{2}\). Taking its logarithm, we obtain:

\[
\log B = \log n - \log k - \frac{k-t}{2} + \frac{t}{2} \log k + t + O(1)
\]

Now \(t \log \frac{\log n}{t} = \log n \cdot \frac{\log n}{\log n} = o(\log n)\) since \(\frac{\log n}{t} \to \infty\). Thus, \(\log B \to -\infty\) and the expected number of \(t\)-dependent \(k\)-sets in \(G_{n,p}\) approaches zero as \(n \to \infty\). So, with probability going to one, \(\alpha^t(G_{n,p}) \leq \frac{1}{1-\varepsilon} \gamma \log n\) and \(\chi^t(G_{n,p}) \geq (1-\varepsilon) \gamma \frac{n}{\log n} \).

Next, we consider the case \(t(n) = \omega(\log n)\). By Proposition 1.1 and the fact that \(\Delta(G_{n,p}) \sim np\) a.a.s., it follows that \(\chi^t(G_{n,p}) \leq \lceil (1+\varepsilon) \frac{np}{t} \rceil\) for any fixed \(\varepsilon > 0\). Thus, Lemma 2.2 implies parts (c) and (d) of Theorem 1.2. Part (b) is the subject of the next section.

Lemma 2.2 Fix \(\varepsilon > 0\). If \(t/\log n \to \infty\), then \((1-\varepsilon) \frac{np}{t} \leq \chi^t(G_{n,p})\) a.a.s.

Proof. Again, we want to estimate the maximum size \(\alpha^t\) of a \(t\)-dependent set in \(G_{n,p}\) and we will pass to average degree (or, equivalently, the number of edges). Now let \(k = k(n) = \lceil \frac{t}{(1-\varepsilon)p} \rceil + 1\). Clearly, \(p(k-1) \geq \frac{t}{(1-\varepsilon)^{1/2}}\) for large enough \(n\). Let \(E = E([k])\) be the set of edges induced on \([k]\). Then \(|E| \sim \text{Bin} \left(\binom{k}{2}, p\right)\) and, using a Chernoff bound (and the fact that \(t \leq p(k-1)\)),

\[
\Pr(\Delta(G_{k,p}) \leq t) \leq \Pr(\text{deg}_{\text{avg}}(G_{k,p}) \leq t) = \Pr(|E| \leq kt/2) = \Pr(|E| - p\binom{k}{2} \leq \frac{k}{2} (p(k-1) - t)) \leq \exp \left(-2 \left(\frac{k}{2} (p(k-1) - t)\right)^2 / \binom{k}{2} \right) \leq \exp \left(- (p(k-1) - t)^2 \right)
\]

We have that \(p(k-1) - t \geq \frac{t}{1-\varepsilon^{1/2}} - t = \frac{\varepsilon}{2-\varepsilon} t = \Omega(k)\) and hence the expected number of \(t\)-dependent \(k\)-sets in \(G_{n,p}\) is at most

\[
\binom{n}{k} e^{-\Omega(k^2)} \leq \left(\frac{en}{k} e^{-\Omega(k)}\right)^k = \exp(1 + \log n - \log k - \Omega(k))^k \to 0
\]
as \( n \to \infty \) since \( k/\log n \to \infty \). So, with probability going to one, \( \alpha^t(G_{n,p}) \leq \frac{t}{(1-\epsilon)p} \) and \( \chi^t(G_{n,p}) \geq (1 - \epsilon)^{np/t} \).

3 The intermediate case: an open problem

For this section, we assume that \( \tau > 0 \) is some fixed constant and that \( t(n) \sim \tau \log n \). As mentioned before, \( \chi^t(G_{n,p}) \leq \lceil (1 + \epsilon)^{np/t} \rceil = O(\frac{n}{\log n}) \) a.a.s. for any fixed \( \epsilon > 0 \). We believe that this upper bound can be improved and conjecture a value for the asymptotic constant suggested by Theorem 1.2(b). This value is obtained by using large deviation methods to give more precise estimates for \( \alpha^t(G_{n,p}) \). For further background into large deviations, consult [5]. Denote

\[
\Lambda^*(x) = \begin{cases} 
  x \log \frac{x}{q} + (1 - x) \log \frac{1-x}{p} & \text{for } x \in [0, 1] \\
  \infty & \text{otherwise}
\end{cases}
\]

(where \( \Lambda^*(0) = \log \frac{1}{q} \) and \( \Lambda^*(1) = \log \frac{1}{p} \)). This is the Fenchel-Legendre transform of the logarithmic moment generating function associated with the Bernoulli distribution with probability \( p \) (cf. Exercise 2.2.23(b) of [5]). The following theorem is stated without proof.

**Theorem 3.1** Fix \( \kappa > \tau/p \) and suppose \( k(n) \sim \kappa \log n \) and \( j(n) \sim \frac{n}{\kappa \log n} \). Let \( E_{n,k} \) be the expected number of \( t \)-dependent \( k \)-sets in \( G_{n,p} \) and \( C_{n,j} \) be the expected number of \( t \)-improper \( j \)-colourings of \( G_{n,p} \). Then

(a) \( E_{n,k} = \exp \left( k \log n \left( 1 - \frac{\kappa}{2} \Lambda^* \left( \frac{\tau}{\kappa} \right) + o(1) \right) \right) \)

(b) \( C_{n,j} = \exp \left( n \log n \left( 1 - \frac{\kappa}{2} \Lambda^* \left( \frac{\tau}{\kappa} \right) + o(1) \right) \right) \)

**Lemma 3.2** There exists a unique \( \kappa > \tau/p \) such that \( \frac{\kappa}{2} \Lambda^* \left( \frac{\tau}{\kappa} \right) = 1 \).

**Proof.** Consider \( f : (0, \infty) \to \mathbb{R} \) defined as \( f(\kappa) = \frac{1}{2} \Lambda^* \left( \frac{\tau}{\kappa} \right) - \frac{1}{\kappa} \). Then \( f \) is continuous, strictly increasing on \( \kappa \in (\tau/p, \infty) \), \( f(\tau/p) = -p/\tau < 0 \) and \( f(\kappa) \to \frac{1}{2} \log \frac{1}{q} > 0 \) as \( \kappa \to \infty \). \( \Box \)

Theorem 3.1 suggests the following:

**Conjecture 3.3** Let \( \kappa \) be the unique value satisfying \( 1 - \frac{\kappa}{2} \Lambda^* \left( \frac{\tau}{\kappa} \right) = 0 \) and \( \kappa > \tau/p \). Then \( \chi^t(G_{n,p}) \sim \frac{n}{\kappa \log n} \) a.a.s.
4 Sparse graphs

By utilising the large deviations techniques mentioned in the previous section, the investigations can be extended to the case when $p = p(n) = o(1)$. As alluded to in the introduction, Łuczak showed the following result for the colouring of sparse random graphs:

**Theorem 4.1 (Łuczak [11])** Suppose $0 < p(n) < 1$, $p(n) = o(1)$ and $\varepsilon > 0$. Set $d(n) = np(n)$. There exists constant $d_0$ such that, if $d(n) \geq d_0$, then $(1 - \varepsilon) \frac{d}{2 \log d} \leq \chi(G_{n,p}) \leq (1 + \varepsilon) \frac{d}{2 \log d}$ a.a.s.

The next theorem, stated without proof, is an example of how Theorem 4.1 may be extended to $t$-improper colouring (in the specific case when $t$ is small enough with respect to $d$).

**Theorem 4.2** Suppose $0 < p(n) < 1$ and $p(n) = o(1)$. Set $d(n) = np(n)$.

(a) Fix $\varepsilon > 0$ and suppose $t(n) = t_0$ for some fixed $t_0 \geq 0$. There exists constant $d_0$ such that, if $d(n) \geq d_0$, then $(1 - \varepsilon) \frac{d}{2 \log d} \leq \chi^t(G_{n,p}) \leq (1 + \varepsilon) \frac{d}{2 \log d}$ a.a.s.

(b) If $d(n) = \omega(1)$ and $t(n) = o(\log d)$, then $\chi^t(G_{n,p}) \sim \frac{d}{2 \log d}$ a.a.s.

**References**


