The t-stability number of a random graph

Nikolaos Fountoulakis

Max-Planck-Institut für Informatik
Campus E1 4
Saarbrücken 66123
Germany

Ross J. Kang[*]

School of Engineering and Computing Sciences
Durham University
South Road, Durham DH1 3LE
United Kingdom

Colin McDiarmid

Department of Statistics
University of Oxford
1 South Parks Road
Oxford OX1 3TG
United Kingdom

Abstract

Given a graph $G = (V, E)$, a vertex subset $S \subseteq V$ is called t-stable (or t-dependent) if the subgraph $G[S]$ induced on S has maximum degree at most t. The t-stability number $\alpha_t(G)$ of G is the maximum order of a t-stable set in G. The theme of this paper is the typical values that this parameter takes on a random graph on n vertices and edge probability equal to p. For any fixed $0 < p < 1$ and fixed non-negative integer t, we show that, with probability tending to 1 as $n \to \infty$, the t-stability number takes on at most two values which we identify as functions of t, p and n. The main tool we use is an asymptotic expression for the expected number of t-stable sets of order k. We derive this expression by performing a precise count of the number of graphs on k vertices that have maximum degree at most t.

1 Introduction

Given a graph $G = (V, E)$, a vertex subset $S \subseteq V$ is called t-stable (or t-dependent) if the subgraph $G[S]$ induced on S has maximum degree at most t. The t-stability number

$\alpha_t(G)$ of G is the maximum order of a t-stable set in G. The main topic of this paper is to give a precise formula for the t-stability number of a dense random graph.

The notion of a t-stable set is a generalisation of the notion of a stable set. Recall that a set of vertices S of a graph G is *stable* if no two of its vertices are adjacent. In other words, the maximum degree of $G[S]$ is 0, and therefore a stable set is a 0-stable set.

The study of the order of the largest t-stable set is motivated by the study of the *t-improper chromatic number* of a graph. A *t-improper colouring* of a graph G is a vertex colouring with the property that every colour class is a t-stable set, and the *t-improper chromatic number* $\chi_t(G)$ of G is the least number of colours necessary for a t-improper colouring of G. Obviously, a 0-improper colouring is a proper colouring of a graph, and the 0-improper chromatic number is the chromatic number of a graph.

The t-improper chromatic number is a parameter that was introduced and studied independently by Andrews and Jacobson [1], Harary and Fraughnaugh (née Jones) [11, 12], and by Cowen *et al.* [7]. The importance of the t-stability number in relation to the t-improper chromatic number comes from the following obvious inequality: if G is a graph that has n vertices, then

$$\chi_t(G) \geqslant \frac{n}{\alpha_t(G)}.$$

The t-improper chromatic number also arises in a specific type of radio-frequency assignment problem. Let us assume that the vertices of a given graph represent transmitters and an edge between two vertices indicates that the corresponding transmitters interfere. Each interference creates some amount of noise which we denote by N. Overall, a transmitter can tolerate up to a specific amount of noise which we denote by T. The problem now is to assign frequencies to the transmitters and, more specifically, to assign as few frequencies as possible, so that we minimise the use of the electromagnetic spectrum. Therefore, any given transmitter cannot be assigned the same frequency as more than T/N nearby transmitters — that is, neighbours in the transmitter graph — as otherwise the excessive interference would distort the transmission of the signal. In other words, the vertices/transmitters that are assigned a certain frequency must form a T/N-stable set, and the minimum number of frequencies we can assign is the T/N-improper chromatic number.

Given a graph $G = (V, E)$, we let $S_t = S_t(G)$ be the collection of all subsets of V that are t-stable. We shall determine the order of the largest member of S_t in a random graph $G_{n,p}$. Recall that $G_{n,p}$ is a random graph on a set of n vertices, which we assume to be $V_n := \{1, \ldots, n\}$, and each pair of distinct vertices is present as an edge with probability p independently of every other pair of vertices. Our interest is in dense random graphs, which means that we take $0 < p < 1$ to be a fixed constant.

We say that an event occurs *asymptotically almost surely (a.a.s.)* if it occurs with probability that tends to 1 as $n \to \infty$.

1.1 Related background

The t-stability number of $G_{n,p}$ for the case $t = 0$ has been studied thoroughly for both fixed p and $p(n) = o(1)$. Matula [20, 21, 22] and, independently, Grimmett and McDiarmid [10] were the first to notice and then prove asymptotic concentration of the stability number using the first and second moment methods. For $0 < p < 1$, define $b := 1/(1-p)$ and

$$\alpha_{0,p}(n) := 2\log_b n - 2\log_b \log_b n + 2\log_b(e/2) + 1.$$

For fixed $0 < p < 1$, it was shown that for any $\varepsilon > 0$ a.a.s.

$$\lfloor \alpha_{0,p}(n) - \varepsilon \rfloor \leqslant \alpha_0(G_{n,p}) \leqslant \lfloor \alpha_{0,p}(n) + \varepsilon \rfloor, \tag{1}$$

showing in particular that $\chi(G_{n,p}) \geqslant (1-\varepsilon)n/\alpha_{0,p}(n)$. Assume now that $p = p(n)$ is bounded away from 1. Bollobás and Erdős [4] extended (1) to hold with $p(n) > n^{-\delta}$ for any $\delta > 0$. Much later, with the use of martingale techniques, Frieze [9] showed that for any $\varepsilon > 0$ there exists some constant C_ε such that if $p(n) \geqslant C_\varepsilon/n$ then (1) holds a.a.s.

Efforts to determine the chromatic number of $G_{n,p}$ took place in parallel with the study of the stability number. For fixed p, Grimmett and McDiarmid conjectured that $\chi(G_{n,p}) \sim n/\alpha_{0,p}(n)$ a.a.s. This conjecture was a major open problem in random graph theory for over a decade, until Bollobás [2] and Matula and Kučera [19] used martingales to establish the conjecture. It was crucial for this work to obtain strong upper bounds on the probability of nonexistence in $G_{n,p}$ of a stable set with just slightly fewer than $\alpha_{0,p}(n)$ vertices. Łuczak [18] fully extended the result to hold for sparse random graphs; that is, for the case $p(n) = o(1)$ and $p(n) \geqslant C/n$ for some large enough constant C. Consult Bollobás [3] or Janson, Łuczak and Ruciński [15] for a detailed survey of these as well as related results.

For the case $t \geqslant 1$, the first results on the t-stability number were developed indirectly as a consequence of broader work on hereditary properties of random graphs. A graph property — that is, an infinite class of graphs closed under isomorphism — is said to be *hereditary* if every induced subgraph of every member of the class is also in the class. For any given t, the class of graphs that are t-stable is an hereditary property. As a result of study in this more general context, it was shown by Scheinerman [25] that, for fixed p, there exist constants $c_{p,1}$ and $c_{p,2}$ such that $c_{p,1} \ln n \leqslant \alpha_t(G_{n,p}) \leqslant c_{p,2} \ln n$ a.a.s. This was further improved by Bollobás and Thomason [5] who characterised, for any fixed p, an explicit constant c_p such that $(1-\varepsilon)c_p \ln n \leqslant \alpha_t(G_{n,p}) \leqslant (1+\varepsilon)c_p \ln n$ a.a.s. For any fixed hereditary property, not just t-stability, the constant c_p depends upon the property but essentially the same result holds. Recently, Kang and McDiarmid [16, 17] considered t-stability separately, but also treated the situation in which $t = t(n)$ varies (i.e. grows) in the order of the random graph. They showed that, if $t = o(\ln n)$, then a.a.s.

$$(1-\varepsilon)2\log_b n \leqslant \alpha_t(G_{n,p}) \leqslant (1+\varepsilon)2\log_b n \tag{2}$$

(where $b = 1/(1-p)$, as above). In particular, observe that the estimation (2) for $\alpha_t(G_{n,p})$ and the estimation (1) for $\alpha_0(G_{n,p})$ agree in their first-order terms. This implies that as long as $t = o(\ln n)$ the t-improper and the ordinary chromatic numbers of $G_{n,p}$ have roughly the same asymptotic value a.a.s.

1.2 The results of the present work

In this paper, we restrict our attention to the case in which the edge probability p and the non-negative integer parameter t are fixed constants. Restricted to this setting, our main theorem is an extension of (1) and a strengthening of (2).

Theorem 1 *Fix $0 < p < 1$ and $t \geqslant 0$. Set $b := 1/(1-p)$ and*

$$\alpha_{t,p}(n) := 2\log_b n + (t-2)\log_b \log_b n + \log_b(t^t/t!^2) + t\log_b(2bp/e) + 2\log_b(e/2) + 1.$$

Then for every $\varepsilon > 0$ a.a.s.

$$\lfloor \alpha_{t,p}(n) - \varepsilon \rfloor \leqslant \alpha_t(G_{n,p}) \leqslant \lfloor \alpha_{t,p}(n) + \varepsilon \rfloor.$$

We shall see that this theorem in fact holds if $\varepsilon = \varepsilon(n)$ as long as $\varepsilon \gg \ln \ln n / \sqrt{\ln n}$.

We derive the upper bound with a first moment argument, which is presented in Section 3. To apply the first moment method, we estimate the expected number of t-stable sets that have order k. In particular, we show the following.

Theorem 2 *Fix $0 < p < 1$ and $t \geqslant 0$. Let $\alpha_t^{(k)}(G)$ denote the number of t-stable sets of order k that are contained in a graph G. If $k = O(\ln n)$ and $k \to \infty$ as $n \to \infty$, then*

$$\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) = \left(e^2 n^2 b^{-k+1} k^{t-2} \left(\frac{tbp}{e} \right)^t \frac{1}{t!^2} \right)^{k/2} (1 + o(1))^k.$$

(Note that by (2) the condition on k is not very restrictive.) Using this formula, we will see in Section 3 that the expected number of t-stable sets with $\lfloor \alpha_{t,p}(n) + \varepsilon \rfloor + 1$ vertices tends to zero as $n \to \infty$.

The key to the calculation of this expected value is a precise formula for the number of degree sequences on k vertices with a given number of edges and maximum degree at most t. In Section 2, we obtain this formula by the inversion formula of generating functions — applied in our case to the generating function of degree sequences on k vertices and maximum degree at most t. This formula is an integral of a complex function that is approximated with the use of an analytic technique called saddle-point approximation. Our proof is inspired by the application of this method by Chvátal [6] to a similar generating function. For further examples of the use of the saddle-point method, consult Chapter VIII of Flajolet and Sedgewick [8].

The lower bound in Theorem 1 is derived with a second moment argument in Section 4.

We remark that Theorems 1 and 2 are both stated to hold for the case $t = 0$ (if we assume that $0^0 = 1$) in order to stress that these results generalise the previous results of Matula [20, 21, 22] and Grimmett and McDiarmid [10]. Our methods apply for this special case, however in our proofs our main concern will be to establish the results for $t \geqslant 1$.

In Section 5 we give a quite precise formula for the t-improper chromatic number of $G_{n,p}$. For $t = 0$, that is, for the chromatic number, McDiarmid [23] gave a fairly tight estimate on $\chi(G_{n,p})(= \chi_0(G_{n,p}))$ proving that for any fixed $0 < p < 1$ a.a.s.

$$\frac{n}{\alpha_{0,p}(n) - 1 - o(1)} \leqslant \chi_0(G_{n,p}) \leqslant \frac{n}{\alpha_{0,p}(n) - 1 - \frac{1}{2} - \frac{1}{1-(1-p)^{1/2}} + o(1)}.$$

Panagiotou and Steger [24] recently improved the lower bound showing that a.a.s.

$$\chi_0(G_{n,p}) \geqslant \frac{n}{\alpha_{0,p}(n) - \frac{2}{\ln b} - 1 + o(1)},$$

and asked if better upper or lower bounds could be developed. In Section 5, we improve upon McDiarmid's upper bound and we generalise (for $t \geqslant 1$) both this new bound and the lower bound of Panagiotou and Steger.

Theorem 3 *Fix $0 < p < 1$ and $t \geqslant 0$. Then a.a.s.*

$$\frac{n}{\alpha_{t,p}(n) - \frac{2}{\ln b} - 1 + o(1)} \leqslant \chi_t(G_{n,p}) \leqslant \frac{n}{\alpha_{t,p}(n) - \frac{2}{\ln b} - 2 - o(1)}.$$

Given a graph G, let the *colouring rate* $\overline{\alpha_0}(G)$ of G be $|V(G)|/\chi_0(G)$, which is the maximum average size of a colour class in a proper colouring of G. Then the case $t = 0$ of Theorem 3 implies for any fixed $0 < p < 1$ that a.a.s.

$$\alpha_{0,p}(n) - \frac{2}{\ln b} - 2 - o(1) \leqslant \overline{\alpha_0}(G_{n,p}) \leqslant \alpha_{0,p}(n) - \frac{2}{\ln b} - 1 + o(1).$$

Thus the colouring rate of $G_{n,p}$ is a.a.s. contained in an explicit interval of length $1 + o(1)$. We remark that Shamir and Spencer [27] showed a.a.s. $\tilde{O}(\sqrt{n})$-concentration of $\chi_0(G_{n,p})$ — see also a recent improvement by Scott [26]. (The $\tilde{O}$ notation ignores logarithmic factors.) It therefore follows that $\overline{\alpha_0}(G_{n,p})$ is a.a.s. $\tilde{O}(n^{-1/2})$-concentrated.

The above discussion extends easily to t-improper colourings.

2 Counting degree sequences of maximum degree t

Given non-negative integers k, t with $t < k$, we let

$$C_{2m}(t, k) := \sum_{(d_1,\ldots,d_k), \sum_i d_i = 2m, d_i \leqslant t} \frac{1}{\prod_i d_i!}.$$

(Here, the d_i are non-negative integers.) Given a fixed degree sequence $(d_1, \ldots, d_k)$ with $\sum_i d_i = 2m$, the number of graphs on k vertices $(v_1, \ldots, v_k)$ where v_i has degree d_i is at most

$$\frac{1}{\prod_i d_i!} \frac{(2m)!}{m! 2^m}.$$

See for example [3] in the proof of Theorem 2.16 or Section 9.1 in [15] for the definition of the configuration model, from which the above claim follows easily. Therefore, $C_{2m}(t, k)(2m)!/(m!2^m)$ is an upper bound on the number of graphs with k vertices and m edges such that each vertex has degree at most t. Note also that $(2m)!C_{2m}(t, k)$ is the number of allocations of $2m$ balls into k bins with the property that no bin contains more than t balls.

In the proof of Theorem 2, we need good estimates for $C_{2m}(t, k)$, when $2m$ is close to tk. In particular, as we will see in the next section (Lemma 7) we will need a tight estimate for $C_{2m}(t, k)$ when $t - \ln k / \sqrt{k} < 2m/k < t - 1/(\sqrt{k} \ln k)$, since in this range the expected number of t-stable sets having m edges is maximised. We require a careful and specific treatment of this estimation due to the fact that $2m/k$ is not bounded below t.

For $t \geqslant 1$, note that $C_{2m}(t, k)$ is the coefficient of z^{2m} in the following generating function:

$$G(z) = R_t(z)^k = \left(\sum_{i=0}^{t} \frac{z^i}{i!} \right)^k.$$

Cauchy's integral formula gives

$$C_{2m}(t, k) = \frac{1}{2\pi i} \int_C \frac{R_t(z)^k}{z^{2m+1}} dz,$$

where the integration is taken over a closed contour containing the origin.

Before we state the main theorem of this section, we need the following lemma, which follows from Note IV.46 in [8].

Lemma 4 *Fix $t \geqslant 1$. The function $rR_t'(r)/R_t(r)$ is strictly increasing in r for $r > 0$. For each $y \in (0, t)$, there exists a unique positive solution $r_0 = r_0(y)$ to the equation $rR_t'(r)/R_t(r) = y$ and furthermore the function $r_0(y)$ is a continuous bijection between $(0, t)$ and $(0, \infty)$. Thus, if we set*

$$s(y) = r_0(y) \frac{d}{dx} \left. \frac{xR_t'(x)}{R_t(x)} \right|_{x=r_0(y)},$$

then $s(y) > 0$.

We will prove a "large powers" theorem to obtain a very tight estimate on $C_{2m}(t, k)$ when $2m/k$ is quite close to t. A version of this theorem holds if we instead assume that $2m/k$ is bounded away from t; indeed, this immediately follows from Theorem VIII.8 of [8]. However, our version, where $2m/k$ approaches t, is necessary in light of Lemma 7 below.

Theorem 5 *Assume that $t \geqslant 1$ is fixed and $k \to \infty$. Suppose that m and k are such that $t - \ln k / \sqrt{k} \leqslant 2m/k \leqslant t - 1/(\sqrt{k} \ln k)$ for any $\varepsilon > 0$, and r_0 and s are defined as in Lemma 4. Then uniformly*

$$C_{2m}(t, k) = \frac{1}{\sqrt{2\pi k s(2m/k)}} \frac{R_t(r_0(2m/k))^k}{r_0(2m/k)^{2m}} (1 + o(1)).$$

In the proof of the theorem (as well as in later sections), we make frequent use of the following lemma, whose proof is postponed until the end of the section.

Lemma 6 *If $y = y(k) \to t$ as $k \to \infty$ (and $y < t$) and r_0 and s are defined as in Lemma 4, then*

$$r_0 = \frac{t}{t-y} + O(1), \tag{3}$$

$$\frac{dr_0}{dy} = \frac{r_0{}^2}{t}\left(1 + O\left(\frac{1}{r_0}\right)\right), \quad and \tag{4}$$

$$s = \frac{t}{r_0}\left(1 + O\left(\frac{1}{r_0}\right)\right). \tag{5}$$

Proof of Theorem 5 The proof is inspired by [6]. Throughout, we for convenience drop the subscript and write $R(z)$ in the place of $R_t(z)$. Recall that $r_0 = r_0(2m/k)$ is the unique positive solution of the equation $rR'(r)/R(r) = 2m/k$, where $t - \ln k/\sqrt{k} \leqslant 2m/k \leqslant t - 1/(\sqrt{k}\ln k)$, and let C be the circle of radius r_0 centred at the origin. Using polar coordinates, we obtain

$$C_{2m}(t,k) = \frac{1}{2\pi i}\int_C \frac{R(r_0 e^{i\varphi})^k}{r_0{}^{2m+1} e^{i2m\varphi} e^{i\varphi}}d(r_0 e^{i\varphi}) = \frac{1}{2\pi r_0{}^{2m}}\int_{-\pi}^{\pi}\frac{R(r_0 e^{i\varphi})^k}{e^{i2m\varphi}}d\varphi.$$

We let $\delta = \delta(k) := \ln k\sqrt{r_0/k}$ and write

$$C_{2m}(t,k) = \frac{1}{2\pi r_0{}^{2m}}\left(\int_\delta^{2\pi-\delta}\frac{R(r_0 e^{i\varphi})^k}{e^{i2m\varphi}}d\varphi + \int_{-\delta}^{\delta}\frac{R(r_0 e^{i\varphi})^k}{e^{i2m\varphi}}d\varphi\right). \tag{6}$$

Note that, since $2m/k < t - 1/(\ln k\sqrt{k})$, it follows from (3) that $\delta \to 0$ as $k \to \infty$. We shall analyse the two integrals of (6) separately.

To begin, we consider the first integral of (6) and we wish to show that it makes a negligible contribution to the value of $C_{2m}(t,k)$. Note that

$$\left|R(r_0 e^{i\varphi})\right|^2 = \left(\sum_{j=0}^{t}\frac{r_0{}^j}{j!}\cos(j\varphi)\right)^2 + \left(\sum_{j=0}^{t}\frac{r_0{}^j}{j!}\sin(j\varphi)\right)^2$$

$$= \sum_{0\leqslant j_1,j_2\leqslant t}\frac{r_0{}^{j_1+j_2}}{j_1!j_2!}\left(\cos(j_1\varphi)\cos(j_2\varphi) + \sin(j_1\varphi)\sin(j_2\varphi)\right)$$

$$= \sum_{0\leqslant j_1,j_2\leqslant t}\frac{r_0{}^{j_1+j_2}}{j_1!j_2!}\cos\left((j_1 - j_2)\varphi\right)$$

$$= R(r_0)^2 - \sum_{0\leqslant j_1<j_2\leqslant t}\frac{2r_0{}^{j_1+j_2}}{j_1!j_2!}\left(1 - \cos\left((j_1 - j_2)\varphi\right)\right). \tag{7}$$

Note that $r_0 \to \infty$ as $k \to \infty$. Hence, from (7),

$$\left|R(r_0 e^{i\varphi})\right|^2 \leqslant R(r_0)^2\left(1 - \frac{\frac{2r_0{}^{2t-1}}{t!(t-1)!}(1-\cos\varphi)}{\frac{r_0{}^{2t}}{t!^2} + \Theta(r_0{}^{2t-1})}\right) = R(r_0)^2\left(1 - (1+o(1))\frac{2t}{r_0}(1-\cos\varphi)\right).$$

It follows that for k large enough

$$\left| \int_\delta^{2\pi-\delta} \frac{R(r_0 e^{i\varphi})^k}{e^{i2m\varphi}} d\varphi \right| \leqslant 2\pi R(r_0)^k \left(1 - (1+o(1))\frac{2t}{r_0}(1-\cos\delta) \right)^{k/2}$$

$$\leqslant 2\pi R(r_0)^k \exp\left(-\frac{tk}{2r_0}(1-\cos\delta) \right)$$

$$= 2\pi R(r_0)^k \exp\left(-\frac{t}{2} \cdot \frac{k\delta^2}{r_0 \ln k} \cdot \frac{1-\cos\delta}{\delta^2} \cdot \ln k \right). \qquad (8)$$

Since $\delta \to 0$, we have that $(1-\cos\delta)/\delta^2 \to 1/2$. By the choice of δ, we also have that $k\delta^2/(r_0 \ln k) \to \infty$ as $k \to \infty$, and it follows from Inequality (8) that

$$\left| \int_\delta^{2\pi-\delta} \frac{R(r_0 e^{i\varphi})^k}{e^{i2m\varphi}} d\varphi \right| < R(r_0)^k/k, \qquad (9)$$

for large enough k.

In order to precisely estimate the second integral of (6), we consider the function $f : \mathbb{R} \to \mathbb{C}$ given by

$$f(\varphi) := R(r_0 e^{i\varphi}) \exp\left(-i\frac{2m}{k}\varphi \right) = \exp\left(-i\frac{2m}{k}\varphi \right) \left(\sum_{j=0}^t \frac{r_0{}^j}{j!}(\cos(j\varphi) + i\sin(j\varphi)) \right).$$

The importance of the function f is that

$$\int_{-\delta}^\delta \frac{R(r_0 e^{i\varphi})^k}{e^{i2m\varphi}} d\varphi = \int_{-\delta}^\delta f(\varphi)^k d\varphi.$$

We will show that the real part of $f(\varphi)^k$ is well approximated by $R(r_0)^k \exp(-sk\varphi^2/2)$ when $|\varphi|$ is small — see (12) below. The imaginary part can be ignored as the integral approximates a real quantity.

To this end we will apply Taylor's Theorem, and in order to do this we shall need the first, second and third derivatives of f with respect to φ. First,

$$f'(\varphi) = \exp\left(-i\frac{2m}{k}\varphi \right) \left(\sum_{j=0}^t \frac{r_0{}^j}{j!}\left(\frac{2m}{k} - j \right)(\sin(j\varphi) - i\cos(j\varphi)) \right).$$

Note that

$$f'(0) = -i\left(\frac{2m}{k}\sum_{j=0}^t \frac{r_0{}^j}{j!} - \sum_{j=0}^t \frac{r_0{}^j}{j!}j \right) = -i\left(\frac{2m}{k}R(r_0) - r_0 R'(r_0) \right) = 0$$

by the choice of r_0. Next,

$$f''(\varphi) = -i\frac{2m}{k}f'(\varphi) + \exp\left(-i\frac{2m}{k}\varphi \right) \left(\sum_{j=0}^t \frac{r_0{}^j}{j!}\left(\frac{2m}{k} - j \right)j(\cos(j\varphi) + i\sin(j\varphi)) \right).$$

Therefore,

$$f''(0) = -i\frac{2m}{k}f'(0) + \sum_{j=0}^{t}\frac{r_0{}^j}{j!}\left(\frac{2m}{k} - j\right)j$$

$$= \frac{2m}{k}\sum_{j=1}^{t}\frac{r_0{}^j}{j!}j - \sum_{j=1}^{t}\frac{r_0{}^j}{j!}j(j-1) - \sum_{j=1}^{t}\frac{r_0{}^j}{j!}j$$

$$= \left(\frac{r_0 R'(r_0)}{R(r_0)}\right)r_0 R'(r_0) - r_0{}^2 R''(r_0) - r_0 R'(r_0)$$

$$= -r_0\left(\frac{-r_0 R'(r_0)^2}{R(r_0)} + r_0 R''(r_0) + R'(r_0)\right)$$

$$= -R(r_0)r_0\left(\frac{(r_0 R''(r_0) + R'(r_0))R(r_0) - r_0 R'(r_0)^2}{R(r_0)^2}\right)$$

$$= -R(r_0)r_0\frac{d}{dx}\left.\frac{xR'(x)}{R(x)}\right|_{x=r_0} = -R(r_0)s(2m/k). \tag{10}$$

Thus, $f''(0) < 0$ by Lemma 4. Last, we have

$$f'''(\varphi) = -i\frac{2m}{k}f''(\varphi) - i\frac{2m}{k}\exp\left(-i\frac{2m}{k}\varphi\right)\left(\sum_{j=0}^{t}\frac{r_0{}^j}{j!}\left(\frac{2m}{k} - j\right)j(\cos(j\varphi) + i\sin(j\varphi))\right)$$

$$+ \exp\left(-i\frac{2m}{k}\varphi\right)\left(\sum_{j=0}^{t}\frac{r_0{}^j}{j!}\left(\frac{2m}{k} - j\right)j^2(-\sin(j\varphi) + i\cos(j\varphi))\right).$$

Since $r_0 \to \infty$ as $k \to \infty$, there is a positive constant a such that $a \leqslant r_0$, for k sufficiently large. Clearly, $f(0) = R(r_0) > a^t/t! > 0$. The continuity of f on the compact set $-\pi \leqslant \varphi \leqslant \pi$ implies that there is a positive constant δ_0 such that whenever $|\varphi| \leqslant \delta_0$ we have $Re(f(\varphi)) > 0$. Since the first two derivatives of $Im(f(\varphi))$ with respect to φ vanish when $\varphi = 0$, and also $Im(f(0)) = 0$, Taylor's Theorem implies that

$$|Im(f(\varphi))| \leqslant \sup_{|\varphi| \leqslant \delta_0} |Im(f'''(\varphi))|\frac{\varphi^3}{6}$$

if $|\varphi| \leqslant \delta_0$. Now, note that $Re(f(\varphi))$ and $Im(f'''(\varphi))$ can be considered as polynomials of degree t with respect to r_0. The leading term of $Re(f(\varphi))$ is

$$Re\left(\exp\left(-i\frac{2m}{k}\varphi\right)(\cos(t\varphi) + i\sin(t\varphi))\right)\frac{r_0{}^t}{t!};$$

thus, $Re(f(\varphi)) = \Omega(r_0{}^t)$. On the other hand, using the derivative computations above and simplifying, it follows that the leading term of $Im(f'''(\varphi))$ is

$$Im\left(\exp\left(-i\frac{2m}{k}\varphi\right)(\sin(t\varphi) + i\cos(t\varphi))\right)\left(t - \frac{2m}{k}\right)^3\frac{r_0{}^t}{t!}.$$

By (3), $t - 2m/k = (1 + o(1))t/r_0$ and thus $Im(f'''(\varphi)) = O({r_0}^{t-1})$. So, there exists $c_1 > 0$ such that for every φ with $|\varphi| \leqslant \delta_0$

$$\frac{\sup_{|\varphi| \leqslant \delta_0} |Im(f'''(\varphi))|}{|Re(f(\varphi))|} < \frac{c_1}{r_0},$$

and therefore

$$\left| \frac{Im(f(\varphi))}{Re(f(\varphi))} \right| \leqslant \frac{c_1 \varphi^3}{6 r_0},$$

for any φ with $|\varphi| \leqslant \delta_0$. On the other hand, we have (see pages 15–16 of [6] for the details)

$$\left| \frac{Re(z^k)}{Re(z)^k} - 1 \right| \leqslant \epsilon \left(k, \left| \frac{Im(z)}{Re(z)} \right| \right),$$

with

$$\epsilon(k, x) = (1 + x)^k - 1 - xk \leqslant e^{xk} - 1$$

(for $x \geqslant 0$). Since $\epsilon(k, x)$ increases in x for $x \geqslant 0$, we have

$$1 - \epsilon \left(k, \frac{c_1 \delta^3}{6 r_0} \right) \leqslant \frac{Re(f(\varphi)^k)}{Re(f(\varphi))^k} \leqslant 1 + \epsilon \left(k, \frac{c_1 \delta^3}{6 r_0} \right), \tag{11}$$

whenever $|\varphi| \leqslant \delta \leqslant \delta_0$.

Next, we approximate the function $\ln Re(f(\varphi))$. First,

$$\frac{d}{d\varphi}(\ln Re(f(\varphi))) \bigg|_{\varphi=0} = \frac{Re(f'(\varphi))}{Re(f(\varphi))} \bigg|_{\varphi=0} = 0.$$

Second, we have

$$\frac{d^2}{d\varphi^2}(\ln Re(f(\varphi))) = \frac{d}{d\varphi} \left(\frac{Re(f'(\varphi))}{Re(f(\varphi))} \right) = \frac{Re(f''(\varphi))Re(f(\varphi)) - Re(f'(\varphi))^2}{Re(f(\varphi))^2};$$

therefore, by Equation (10),

$$\frac{d^2}{d\varphi^2}(\ln Re(f(\varphi))) \bigg|_{\varphi=0} = \frac{Re(f''(0))Re(f(0)) - Re(f'(0))^2}{Re(f(0))^2} = \frac{-R(r_0)s}{R(r_0)} = -s$$

Now, the numerator of the third derivative with respect to φ is

$$(Re(f''(\varphi))Re(f(\varphi)) - Re(f'(\varphi))^2)' Re(f(\varphi))^2$$
$$- 2Re(f(\varphi))(Re(f''(\varphi))Re(f(\varphi)) - Re(f'(\varphi))^2)$$
$$= Re(f(\varphi)) \Big((Re(f''(\varphi))Re(f(\varphi)) - Re(f'(\varphi))^2)' Re(f(\varphi))$$
$$- 2(Re(f''(\varphi))Re(f(\varphi)) - Re(f'(\varphi))^2) \Big).$$

Thus an elementary calculation gives that (for $|\varphi| \leqslant \delta_0$)

$$\frac{d^3}{d\varphi^3}(\ln Re(f(\varphi)))$$
$$= \frac{Re(f'''(\varphi))Re(f(\varphi))^2 - 3Re(f''(\varphi))Re(f'(\varphi))Re(f(\varphi)) + 2Re(f'(\varphi))^3}{Re(f(\varphi))^3}.$$

If, as we did earlier for $Re(f(\varphi))$ and $Im(f'''(\varphi))$, we consider $Re(f'(\varphi))$, $Re(f''(\varphi))$ and $Re(f'''(\varphi))$ as polynomials with respect to r_0, we can show that $Re(f'(\varphi)) = O(r_0^{t-1})$, $Re(f''(\varphi)) = O(r_0^{t-1})$ and $Re(f'''(\varphi)) = O(r_0^{t-1})$. It then follows that there exists $c_2 > 0$ such that for every φ with $|\varphi| \leqslant \delta_0$

$$\left| \frac{d^3}{d\varphi^3}(\ln Re(f(\varphi))) \right| \leqslant \frac{c_2}{r_0}.$$

Therefore, Taylor's Theorem implies that for every φ with $|\varphi| \leqslant \delta_0$ we have

$$\left| \ln Re(f(\varphi)) - \left(\ln R(r_0) - \frac{s\varphi^2}{2} \right) \right| \leqslant \frac{c_2\varphi^3}{6r_0}.$$

It follows that

$$\exp\left(-\frac{c_2 k\delta^3}{6r_0} \right) \leqslant \frac{Re(f(\varphi))^k}{R(r_0)^k \exp(-sk\varphi^2/2)} \leqslant \exp\left(\frac{c_2 k\delta^3}{6r_0} \right).$$

The condition that $2m/k < t - 1/(\ln k\sqrt{k})$ and (3) together imply that $r_0 < t \ln k\sqrt{k} + O(1)$. Therefore, $k\delta^3/r_0 = \sqrt{r_0/k}\ln^3 k \to 0$ as $k \to \infty$, and we have

$$\exp\left(\frac{c_2 k\delta^3}{6r_0} \right) = 1 + o(1) \quad \text{and} \quad \epsilon\left(k, \frac{c_1\delta^3}{6r_0} \right) \leqslant \exp\left(\frac{c_1 k\delta^3}{6r_0} \right) - 1 = o(1),$$

proving that

$$Re(f(\varphi)^k) = R(r_0)^k \exp(-sk\varphi^2/2)(1 + o(1)) \tag{12}$$

uniformly for $|\varphi| \leqslant \delta$. From (6), (9) and (12), we obtain

$$2\pi r_0{}^{2m} C_{2m}(t,k) = R(r_0)^k \left(\int_{-\delta}^{\delta} \exp(-sk\varphi^2/2)d\varphi + o(1) \right). \tag{13}$$

Using a change of variables $\psi = \sqrt{sk}\varphi$, observe that

$$\int_{-\delta}^{\delta} \exp\left(-\frac{sk\varphi^2}{2} \right) d\varphi = \frac{1}{\sqrt{sk}} \int_{-\delta\sqrt{sk}}^{\delta\sqrt{sk}} \exp\left(-\frac{\psi^2}{2} \right) d\psi = \sqrt{\frac{2\pi}{sk}}(1 + o(1)),$$

as $k \to \infty$ since $\delta\sqrt{sk} \sim \sqrt{t}\ln k \to \infty$. Thus, Equation (13) becomes

$$2\pi r_0{}^{2m} C_{2m}(t,k) = \sqrt{\frac{2\pi}{ks}} R(r_0)^k(1 + o(1))$$

and the result follows.

2.1 Proof of Lemma 6

Proof of Equation (3) First, note that $r_0 = r_0(y) \to \infty$ as $k \to \infty$ by Lemma 4. So

$$r_0 R'(r_0) = \frac{r_0{}^t}{(t-1)!} \left(1 + \frac{t-1}{r_0} + O\left(\frac{1}{r_0{}^2}\right) \right),$$

$$R(r_0) = \frac{r_0{}^t}{t!} \left(1 + \frac{t}{r_0} + O\left(\frac{1}{r_0{}^2}\right) \right).$$

Thus,

$$\frac{r_0 R'(r_0)}{R(r_0)} = t\frac{1 + \frac{t-1}{r_0} + O\left(\frac{1}{r_0{}^2}\right)}{1 + \frac{t}{r_0} + O\left(\frac{1}{r_0{}^2}\right)} = t\left(1 + \frac{t-1}{r_0} + O\left(\frac{1}{r_0{}^2}\right)\right)\left(1 - \frac{t}{r_0} + O\left(\frac{1}{r_0{}^2}\right)\right)$$

$$= t\left(1 - \frac{t}{r_0} + \frac{t-1}{r_0} + O\left(\frac{1}{r_0{}^2}\right)\right) = t\left(1 - \frac{1}{r_0} + O\left(\frac{1}{r_0{}^2}\right)\right). \tag{14}$$

Since $r_0 R'(r_0)/R(r_0) = y = t(1 - (t-y)/t)$, we obtain

$$1 - \frac{t-y}{t} = 1 - \frac{1}{r_0} + O\left(\frac{1}{r_0{}^2}\right) \tag{15}$$

which can be rewritten as

$$r_0 = \frac{t}{t-y} \left(1 + O\left(\frac{1}{r_0}\right)\right),$$

and this implies the desired expression.

Proof of Equation (4) A more careful treatment of the computations for the proof of (3) shows that the $O(1/r_0{}^2)$ error term in (15) may instead be written $\eta(1/r_0)/r_0{}^2$ where η is a power series with positive radius of convergence. In particular, as $r_0 R'(r_0)$ and $R(r_0)$ are polynomial functions of r_0, (14) yields, for some power series η_1, η_2 and $\hat{\eta}_2$ with positive radius of convergence, that

$$\frac{y}{t} = \frac{r_0 R'(r_0)}{t R(r_0)} = \frac{1 + \frac{t-1}{r_0} + \frac{\eta_1(1/r_0)}{r_0{}^2}}{1 + \frac{t}{r_0} + \frac{\eta_2(1/r_0)}{r_0{}^2}} = \left(1 + \frac{t-1}{r_0} + \frac{\eta_1(1/r_0)}{r_0{}^2}\right)\left(1 - \frac{t}{r_0} + \frac{\hat{\eta}_2(1/r_0)}{r_0{}^2}\right)$$

$$= 1 - \frac{1}{r_0} + \eta\left(\frac{1}{r_0}\right)\frac{1}{r_0{}^2}.$$

Then, by differentiating both sides of this expression with respect to y, we obtain

$$\frac{1}{t} = \frac{d}{dr_0}\left(1 - \frac{1}{r_0} + \eta\left(\frac{1}{r_0}\right)\frac{1}{r_0{}^2}\right)\frac{dr_0}{dy}.$$

We have that

$$\frac{d}{dr_0}\left(1 - \frac{1}{r_0} + \eta\left(\frac{1}{r_0}\right)\frac{1}{r_0{}^2}\right) = \frac{1}{r_0{}^2} - \eta\left(\frac{1}{r_0}\right)\frac{2}{r_0{}^3} - \eta'\left(\frac{1}{r_0}\right)\frac{1}{r_0{}^4} = \frac{1}{r_0{}^2} + O\left(\frac{1}{r_0{}^3}\right)$$

and (4) immediately follows.

Proof of Equation (5) By the definition of r_0, it follows from the chain rule that

$$1 = \frac{d}{dy}\frac{r_0 R'(r_0)}{R(r_0)} = \frac{d}{dr_0}\frac{r_0 R'(r_0)}{R(r_0)}\frac{dr_0}{dy}.$$

Thus,

$$\frac{d}{dx}\frac{xR'(x)}{R(x)}\Bigg|_{x=r_0(y)} = \left(\frac{dr_0(y')}{dy'}\Bigg|_{y'=y}\right)^{-1},$$

implying that

$$s(y) = r_0(y)\left(\frac{dr_0(y')}{dy'}\Bigg|_{y'=y}\right)^{-1} \stackrel{(4)}{=} \frac{t}{r_0(y)}\left(1 + O\left(\frac{1}{r_0(y)}\right)\right)$$

as required.

3 The expected number of t-stable sets of order k - proof of Theorem 2

In this section, we give an asymptotic expression for the expected number of t-stable subsets of V_n of order k in $G_{n,p}$, proving Theorem 2. In light of (2), we will consider k such that $k = k(n) = O(\ln n)$ and $k \to \infty$ as $n \to \infty$. Towards the end of the section, we will specify k and derive the upper bound of Theorem 1 by a first moment argument.

Let A be a subset of V_n that has order k. If $\alpha_t^{(k)}(G_{n,p})$ denotes the number of subsets of V_n of order k that are t-stable, then

$$\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) = \binom{n}{k}\mathbb{P}(A \in S_t).$$

Partitioning according to the number of edges that A induces, we have

$$\mathbb{P}(A \in S_t) = \sum_{m=0}^{\lfloor tk/2 \rfloor} \mathbb{P}(A \in S_t,\ e(A) = m). \tag{16}$$

By the definition of $C_{2m}(t, k)$ (given at the beginning of Section 2), it follows that

$$\mathbb{P}(A \in S_t,\ e(A) = m) \leqslant p^m(1-p)^{\binom{k}{2}-m}C_{2m}(t, k)\frac{(2m)!}{m!2^m} =: f(m). \tag{17}$$

First, we find the value of m for which the expression $f(m)$ on the right-hand side of (17) is maximised. If m^* is such that $f(m^*) = \max\{f(m) : 0 \leqslant 2m \leqslant tk\}$, it turns out that the following holds.

Lemma 7 $2m^* = tk - \sqrt{tk/bp} + o(\sqrt{k})$.

Proof Let $\lambda_m = \lambda_m(t,k) = f(m+1)/f(m)$. Thus,

$$\lambda_m = \frac{p}{1-p} \frac{C_{2m+2}(t,k)}{C_{2m}(t,k)} \frac{1}{2} \frac{(2m+2)(2m+1)}{m+1} = \frac{p}{1-p} \frac{C_{2m+2}(t,k)}{C_{2m}(t,k)} (2m+1).$$

We will estimate λ_m for all m with $0 \leqslant 2m \leqslant tk$ and treat three separate cases:

(A) $2m < tk - \sqrt{k} \ln k$;

(B) $2m > tk - \sqrt{k}/\ln k$; and

(C) $tk - \sqrt{k} \ln k \leqslant 2m \leqslant tk - \sqrt{k}/\ln k$.

We will use Theorem 5 in Case (C), as we will determine those values m for which $\lambda_m \approx 1$ within that range. In the other cases we will use a cruder argument, which is nonetheless sufficient for our purposes.

Case (A)

We will show that $\lambda_m > 1$ for any such m. We set $S_{2m}(t,k) = (2m)! C_{2m}(t,k)$. Note that this is equal to the number of ways of allocating $2m$ labelled balls into k bins so that each bin does not receive more than t balls — we also denote the set of such allocations by $\mathcal{S}_{2m}(t,k)$. We have

$$\frac{C_{2(m+1)}(t,k)}{C_{2m}(t,k)} = \frac{S_{2(m+1)}(t,k)}{S_{2m}(t,k)} \frac{1}{(2m+2)(2m+1)}. \tag{18}$$

We will obtain a lower bound on the left-hand side, by first obtaining a lower bound on the ratio $S_{2(m+1)}(t,k)/S_{2m}(t,k)$. Let us consider $2m+2$ distinct balls which we label $1, \ldots, 2m+1, 2m+2$. We construct an auxiliary bipartite graph whose parts are $\mathcal{S}_{2m}(t,k)$ and $\mathcal{S}_{2m+2}(t,k)$. If $c \in \mathcal{S}_{2m}(t,k)$ and $c' \in \mathcal{S}_{2m+2}(t,k)$, then (c,c') forms an edge in the auxiliary graph if c' restricted to balls $1, \ldots, 2m$ is c. So any $c' \in \mathcal{S}_{2m+2}(t,k)$ is adjacent to exactly one configuration $c \in \mathcal{S}_{2m}(t,k)$, that is, its degree in the auxiliary graph is equal to 1. Also, if $e(c)$ is the number of non-full bins in a configuration $c \in \mathcal{S}_{2m}(t,k)$, then c has at least $e(c)(e(c)-1)$ neighbours in $\mathcal{S}_{2m+2}(t,k)$. This is the case since there are at least $e(c)(e(c)-1)$ ways of allocating balls $2m+1$ and $2m+2$ into the non-full bins of c, therefore giving a lower bound on the number of configurations in $\mathcal{S}_{2m+2}(t,k)$ whose restriction on the first $2m$ balls is c. But $2m < tk - \sqrt{k} \ln k$ and therefore $e(c) \geqslant \sqrt{k}(\ln k)/t$. These observations imply that for k large enough

$$S_{2m+2}(t,k) \geqslant \frac{k \ln^2 k}{2t^2} S_{2m}(t,k),$$

and therefore

$$\frac{C_{2(m+1)}(t,k)}{C_{2m}(t,k)} = \frac{S_{2(m+1)}(t,k)}{S_{2m}(t,k)} \frac{1}{(2m+2)(2m+1)} \geqslant \frac{k \ln^2 k}{2(2m+2)(2m+1)} = \Omega\left(\frac{\ln^2 k}{m}\right).$$

So $\lambda_m = \Omega(\ln^2 k) > 1$ in Case (A).

Case (B)

We treat this case similarly. We consider an auxiliary bipartite graph as above. Let $c \in \mathcal{S}_{2m}(t,k)$ be a configuration of balls $1, \ldots, 2m$. Since there are at most $\sqrt{k}/\ln k$ places available in the non-full bins, there are at most $k/\ln^2 k$ ways of allocating balls $2m+1$ and $2m+2$ into the non-full bins of c. In other words, the degree of any vertex in $\mathcal{S}_{2m}(t,k)$ is at most $k/\ln^2 k$. Also, as above, the degree of any vertex/configuration $c' \in \mathcal{S}_{2m+2}(t,k)$ is equal to one. Therefore,

$$\frac{S_{2m+2}(t,k)}{S_{2m}(t,k)} \leqslant \frac{k}{\ln^2 k}.$$

Substituting this into (18), we obtain

$$\frac{C_{2(m+1)}(t,k)}{C_{2m}(t,k)} \leqslant \frac{k}{\ln^2 k} \frac{1}{(2m+2)(2m+1)}.$$

Therefore, in Case (B) we have

$$\lambda_m = O\left(\frac{k}{m \ln^2 k}\right) = O\left(\frac{1}{\ln^2 k}\right) = o(1).$$

Case (C)

In this range, we need more accurate estimates, as we will identify those m for which λ_m is approximately equal to 1. We appeal to Theorem 5 for asymptotic estimates of $C_{2m}(t,k)$ and $C_{2m+2}(t,k)$ and write $\lambda_m = (1 + o(1))\tilde{\lambda}_m$ where

$$\tilde{\lambda}_m = \frac{p}{1-p}\left(\frac{s(2m/k)}{s(2(m+1)/k)}\right)^{1/2}\left(\frac{R(r_0(2(m+1)/k))}{R(r_0(2m/k))}\right)^k \frac{r_0(2m/k)^{2m}}{r_0(2(m+1)/k)^{2m+2}}(2m+1). \tag{19}$$

Writing $2m = tk - xk$, we have $x = o(1)$. So, by (3) and (4), uniformly for every $z \in [t-x, t-x+2/k]$, we have

$$\left.\frac{dr_0}{dy}\right|_{y=z} = \frac{t}{x^2}(1+o(1));$$

thus, the Mean Value Theorem yields

$$r_0(2(m+1)/k) = r_0(2m/k) + \frac{2t}{x^2 k}(1+o(1)) \overset{(3)}{=} r_0(2m/k)\left(1 + \frac{2}{xk}(1+o(1))\right). \tag{20}$$

So, since $xk \to \infty$ as $k \to \infty$, Equation (20) and (5) yield

$$\left(\frac{s(2m/k)}{s(2(m+1)/k)}\right)^{1/2} = 1 + o(1). \tag{21}$$

To estimate the third ratio of (19), we write $r_0(2(m+1)/k) = r_0(2m/k)(1+\eta)$ where $\eta = (2/xk)(1+o(1))$ by (20). We also write

$$R(r_0(2(m+1)/k) = \frac{r_0{}^t(2(m+1)/k)}{t!} \sum_{t=0}^{t} \frac{t!}{(t-\ell)!} \frac{1}{r_0{}^\ell(2(m+1)/k)}.$$

Note that

$$\sum_{t=0}^{t} \frac{t!}{(t-\ell)!} \frac{1}{r_0{}^\ell(2(m+1)/k)} = \sum_{t=0}^{t} \frac{t!}{(t-\ell)!} \frac{(1+\eta)^{-\ell}}{r_0{}^\ell(2m/k)} = \sum_{t=0}^{t} \frac{t!}{(t-\ell)!} \frac{1 - \ell\eta(1+O(\eta))}{r_0{}^\ell(2m/k)}$$

$$= 1 + \frac{t}{r_0(2m/k)}(1-\eta) + \frac{t(t-1)}{r_0{}^2(2m/k)} + O\left(\frac{\eta^2}{r_0(2m/k)} + \frac{\eta}{r_0{}^2(2m/k)} + \frac{1}{r_0{}^3(2m/k)}\right).$$

Since this last big-O term is $o(1/k)$, it follows that

$$\frac{R(r_0(2(m+1)/k)}{r_0(2(m+1)/k)^t} = \frac{1}{t!}\left(1 + \frac{t}{r_0(2m/k)}(1-\eta) + \frac{t(t-1)}{r_0{}^2(2m/k)} + o(1/k)\right)$$

and similar calculations show that

$$\frac{R(r_0(2m/k)}{r_0(2m/k)^t} = \frac{1}{t!}\left(1 + \frac{t}{r_0(2m/k)} + \frac{t(t-1)}{r_0{}^2(2m/k)} + o(1/k)\right).$$

So the third ratio in (19) becomes

$$\left(\frac{R(r_0(2(m+1)/k))}{R(r_0(2m/k))}\right)^k = \left(\frac{r_0(2(m+1)/k)}{r_0(2m/k)}\right)^{tk} \left(1 - \frac{t\eta}{r_0(2m/k)} + o(1/k)\right)^k$$

$$= \left(\frac{r_0(2(m+1)/k)}{r_0(2m/k)}\right)^{tk} e^{-2}(1+o(1)) \qquad (22)$$

where the last equality holds by the fact that

$$\frac{t\eta k}{r_0(2m/k)} = \frac{t(2/xk)k}{t/x}(1+o(1)) = 2(1+o(1)).$$

Since $xk \to \infty$, we have by (20) and (3) that $r_0(2(m+1)/k) = r_0(2m/k)(1+o(1)) = (1+o(1))t/x$. So using (20) and (22) we can write the product of the third and the fourth terms in (19) as follows:

$$\left(\frac{R(r_0(2(m+1)/k))}{R(r_0(2m/k))}\right)^k \frac{r_0{}^{2m}(2m/k)}{r_0{}^{2m+2}(2(m+1)/k)}$$

$$= e^{-2}\left(\frac{r_0(2(m+1)/k)}{r_0(2m/k)}\right)^{tk-2m} \frac{1+o(1)}{r_0{}^2(2(m+1)/k)}$$

$$= e^{-2}\left(1 + \frac{2}{xk}(1+o(1))\right)^{xk} \frac{x^2}{t^2}(1+o(1)) \stackrel{xk\to\infty}{=} \frac{x^2}{t^2}(1+o(1)).$$

If $x \geqslant \omega(k)/\sqrt{k}$, where $\omega(k) \to \infty$, then substituting this last equation and (21) into (19) and recalling that $\lambda_m = (1 + o(1))\tilde{\lambda}_m$, we obtain

$$\lambda_m = \Omega(1)\frac{x^2}{t^2}(2m+1) = \Omega\left(\frac{\omega(k)^2 m}{k}\right) = \Omega(\omega(k)^2) \to \infty.$$

If $x \leqslant 1/(\omega(k)\sqrt{k})$, then these substitutions yield

$$\lambda_m = O(1)\frac{x^2}{t^2}(2m+1) = O\left(\frac{m}{\omega(k)^2 k}\right) = O\left(\frac{1}{\omega(k)^2}\right) = o(1).$$

Assume now that $x = \alpha/\sqrt{k}$, for some $\alpha = \Theta(1)$. In this case,

$$\lambda_m = \frac{p}{1-p}\frac{\alpha^2}{t^2 k}(tk - xk + 1)(1 + o(1)) = \frac{p}{1-p}\frac{\alpha^2}{t}(1 + o(1)) \overset{b=1/(1-p)}{=} \frac{bp\alpha^2}{t}(1 + o(1)).$$

Thus for any fixed $\alpha' < \sqrt{t/bp} < \alpha''$ and for k large enough we have $tk - \alpha''\sqrt{k} \leqslant 2m^* \leqslant tk - \alpha'\sqrt{k}$. Putting all these different cases together, we deduce that, if m^* is such that $f(m^*)$ is maximised over the set $0 \leqslant 2m \leqslant tk$, then $2m^* = tk - \sqrt{tk/bp} + o(\sqrt{k})$. This concludes the proof of Lemma 7.

Before we proceed to the proof of Theorem 2, let us use Lemma 7 to compute a precise asymptotic expression for $f(m^*)$. Recall that $b = 1/(1-p)$ and observe that

$$p^{m^*}(1-p)^{\binom{k}{2}-m^*} = b^{-\binom{k}{2}}(bp)^{tk/2 - \sqrt{tk/bp} + o(\sqrt{k})} = b^{-\binom{k}{2}}(bp)^{tk/2}\left(1 + O\left(\frac{1}{\sqrt{k}}\right)\right)^k. \quad (23)$$

For the second part of the expression for $f(m^*)$, note that, by Theorem 5,

$$C_{2m^*}(t,k) = \frac{1}{\sqrt{2\pi s(2m^*/k)}}\frac{R(r_0(2m^*/k))^k}{r_0(2m^*/k)^{2m^*}}(1 + o(1)). \quad (24)$$

By (3), we have

$$r_0(2m^*/k) = \sqrt{tbpk} + o(\sqrt{k}).$$

Thus, by (5), $s(2m^*/k) = \Theta(1/\sqrt{k})$. Now, it follows that

$$\begin{aligned}
R(r_0(2m^*/k)) &= \frac{r_0^t(2m^*/k)}{t!}\sum_{\ell=0}^{t}\frac{t!}{(t-\ell)!}\frac{1}{r_0^\ell(2m^*/k)} \\
&= \frac{r_0^t(2m^*/k)}{t!}\left(1 + \frac{t}{r_0(2m^*/k)} + O\left(\frac{1}{r_0^2(2m^*/k)}\right)\right) \\
&= \frac{r_0^t(2m^*/k)}{t!}\left(1 + \sqrt{\frac{t}{bpk}} + o\left(\frac{1}{\sqrt{k}}\right)\right);
\end{aligned}$$

therefore,

$$R(r_0(2m^*/k))^k = \frac{(r_0(2m/k))^{tk}}{t!^k} e^{\sqrt{tk/bp}+o(\sqrt{k})}.$$

Substituting this into (24), we obtain

$$C_{2m^*}(t,k) = \Theta(k^{1/4}) \frac{(r_0(2m^*/k))^{tk-2m^*}}{t!^k} e^{\sqrt{tk/bp}+o(\sqrt{k})}$$

$$= \Theta(k^{1/4}) \left(\sqrt{tbpk}+o(\sqrt{k})\right)^{\sqrt{tk/bp}+o(\sqrt{k})} e^{\sqrt{tk/bp}+o(\sqrt{k})} \frac{1}{t!^k}$$

$$= \frac{1}{t!^k} \left(1+O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k. \tag{25}$$

For the last part of the expression for $f(m^*)$, we apply Stirling's formula to obtain

$$\frac{(2m^*)!}{m^*!2^{m^*}} = \frac{(2m^*/e)^{2m^*}\sqrt{2\pi(2m^*)}e^{o(1)}}{(m^*/e)^{m^*}\sqrt{2\pi m^*}e^{o(1)}} \frac{1}{2^{m^*}} = \Theta(1)\left(\frac{2m^*}{e}\right)^{m^*}$$

$$= \Theta(1)\left(\frac{tk-\sqrt{tk/bp}+o(\sqrt{k})}{e}\right)^{tk/2-\sqrt{tk/bp}/2+o(\sqrt{k})}$$

$$= \left(\frac{tk}{e}\right)^{tk/2}\left(1+O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k. \tag{26}$$

Now, substituting (23), (25) and (26) into the expression for f (given in (17)), we obtain the following:

$$f(m^*) = b^{-\binom{k}{2}}(bp)^{tk/2}\frac{1}{t!^k}\left(\frac{tk}{e}\right)^{tk/2}\left(1+O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k$$

$$= \left(b^{-k+1}\left(\frac{tbpk}{e}\right)^t\frac{1}{t!^2}\right)^{k/2}\left(1+O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k. \tag{27}$$

Upper bound on $\mathbb{E}\left(\alpha_t^{(k)}(G_{n,p})\right)$

By (16) and (27), we deduce that

$$\mathbb{P}(A \in S_t) \leqslant \left(\frac{tk}{2}+1\right)f(m^*) = \left(b^{-k+1}\left(\frac{tbpk}{e}\right)^t\frac{1}{t!^2}\right)^{k/2}\left(1+O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k. \tag{28}$$

Thus, we obtain,

$$\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) \leqslant \binom{n}{k}\left(b^{-k+1}\left(\frac{tbpk}{e}\right)^t\frac{1}{t!^2}\right)^{k/2}\left(1+O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k$$

$$= \left(e^2n^2b^{-k+1}k^{t-2}\left(\frac{tbp}{e}\right)^t\frac{1}{t!^2}\right)^{k/2}\left(1+O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k. \tag{29}$$

Now, if we set $k = \lceil \alpha_{t,p}(n) + \varepsilon(n) \rceil$ for some function $\varepsilon(n) \gg \ln\ln n / \sqrt{\ln n}$, then, substituting this into (29), we obtain

$$\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) \leqslant \left(\left(1 + O\left(\frac{\ln\ln n}{\ln n}\right)\right) b^{-\varepsilon} \right)^{k/2} \left(1 + O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k = o(1),$$

thus proving the right-hand side inequality in Theorem 1.

Lower bound on $\mathbb{E}(\alpha_t^{(k)}(G_{n,p}))$

To derive the lower bound on $\mathbb{E}(\alpha_t^{(k)}(G_{n,p}))$, we observe

$$\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) \geqslant \binom{n}{k} \mathbb{P}(A \in S_t, \ e(A) = m^*).$$

Let $(d_1, \ldots, d_k)$ be a degree sequence such that, for every $1 \leqslant i \leqslant k$, $d_i \leqslant t$ and $\sum_i d_i = 2m^*$. By Theorem 2.16 in [3], with $\lambda := \frac{1}{m^*} \sum_i \binom{d_i}{2}$, the number of graphs with this degree sequence is

$$(1 + o(1))e^{-\lambda/2 - \lambda^2/4} \frac{(2m^*)!}{m^*!2^{m^*}}.$$

But, since $d_i \leqslant t$ for every i, then using the estimate from Lemma 7 we obtain $\lambda \leqslant t^2 k / 2m^* \leqslant 2t$ for k large enough. So the total number of graphs on k vertices, m^* edges and with maximum degree at most t is at least

$$\frac{e^{-t-t^2}}{2} C_{2m^*}(t, k) \frac{(2m^*)!}{m^*!2^{m^*}}.$$

Since $k = O(\ln n)$, we have $\binom{n}{k} = \Omega(\sqrt{1/k})(ne/k)^k$. Hence, using (27), we obtain

$$\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) \geqslant \binom{n}{k} \mathbb{P}(A \in S_t, \ e(A) = m^*) \geqslant \binom{n}{k} \frac{e^{-t-t^2}}{2} f(m^*)$$

$$= \left(e^2 n^2 b^{-k+1} k^{t-2} \left(\frac{tbp}{e}\right)^t \frac{1}{t!^2} \right)^{k/2} \left(1 + O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k. \qquad (30)$$

If $k = \lfloor \alpha_{t,p}(n) - \varepsilon(n) \rfloor$ for some function $\varepsilon(n)$ satisfying $\ln\ln n / \sqrt{\ln n} \ll \varepsilon(n) \ll \ln n$, then by (30) we obtain

$$\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) \geqslant \left(\left(1 + O\left(\frac{\ln\ln n}{\ln n}\right)\right) b^{\varepsilon(n)} \right)^{k/2} \left(1 + O\left(\frac{\ln k}{\sqrt{k}}\right)\right)^k = n^{\varepsilon(n)(1+o(1))} \to \infty.$$
$$(31)$$

In the next section, we use a sharp concentration inequality to show moreover that the following holds.

Lemma 8 *If $\varepsilon(n) \gg \ln\ln n / \sqrt{\ln n}$ is a function that satisfies $\limsup_{n\to\infty} \varepsilon(n) < 2$, then*

$$\mathbb{P}\left(\alpha_t(G_{n,p}) < \lfloor \alpha_{t,p}(n) - \varepsilon(n) \rfloor\right) = \exp\left(-n^{\varepsilon(n)(1+o(1))}\right).$$

Since the right-hand side is $o(1)$, we obtain the lower bound of Theorem 1. This lemma will also be a key step in the proof of the upper bound of Theorem 3, when we need the fact that the right-hand side tends to 0 quickly.

4 A second moment calculation - Proof of Lemma 8

Let (x_n) be a bounded sequence of real numbers such that for

$$k = 2\log_b n + (t-2)\log_b \log_b n + x_n \in \mathbb{N}$$

we have $\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) \to \infty$ as $n \to \infty$. In this section, we prove that a.a.s. there is a k-subset of V_n which is t-stable, using a second moment argument. For this, we use Janson's Inequality ([13], [14] or Theorems 2.14, 2.18 in [15]):

$$\mathbb{P}(\alpha_t^{(k)}(G_{n,p}) = 0) \leqslant \exp\left(-\frac{\mathbb{E}^2(\alpha_t^{(k)}(G_{n,p}))}{\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) + \Delta}\right), \tag{32}$$

where

$$\Delta = \sum_{A,B \subseteq V_n, k-1 \geqslant |A \cap B| \geqslant 2} \mathbb{P}(A, B \in S_t).$$

Let $p(k, \ell)$ be the probability that two k-subsets of V_n that overlap on exactly ℓ vertices are both in S_t. We write

$$\Delta = \sum_{\ell=2}^{k - \lfloor (t+3)\log_b \log_b n \rfloor} \binom{n}{k}\binom{k}{\ell}\binom{n-k}{k-\ell} p(k, \ell)$$

$$+ \sum_{\ell = k - \lfloor (t+3)\log_b \log_b n \rfloor + 1}^{k-1} \binom{n}{k}\binom{k}{\ell}\binom{n-k}{k-\ell} p(k, \ell) =: \Delta_1 + \Delta_2.$$

We conclude the proof by showing that

$$\Delta_1 = O\left(\frac{\ln^5 n}{n^2}\right) \mathbb{E}^2(\alpha_t^{(k)}(G_{n,p})) \text{ and } \Delta_2 = o(\mathbb{E}(\alpha_t^{(k)}(G_{n,p}))).$$

Thus, if $\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) = n^{\varepsilon(n)(1+o(1))}$, where $\varepsilon(n)$ satisfies $\limsup_{n\to\infty} \varepsilon(n) < 2$, then it follows that $\Delta_1 = o(\mathbb{E}(\alpha_t^{(k)}(G_{n,p})))$ and therefore

$$\frac{\mathbb{E}(\alpha_t^{(k)}(G_{n,p}))}{\mathbb{E}(\alpha_t^{(k)}(G_{n,p})) + \Delta} = 1 + o(1).$$

So Lemma 8 follows from (32) by substituting the expression for $\mathbb{E}(\alpha_t^{(k)}(G_{n,p}))$ from (31).

Bounding Δ_1

Let us begin by bounding Δ_1, first estimating $p(k,\ell)$. Let A and B be two k-subsets of V_n that overlap on exactly ℓ vertices, i.e. $|A \cap B| = \ell$. Then $p(k,\ell) = \mathbb{P}(A, B \in S_t) = \mathbb{P}(A \in S_t \mid B \in S_t)\,\mathbb{P}(B \in S_t)$.

The property of having maximum degree at most t is monotone decreasing; so if we condition on the set E of edges induced by $A \cap B$, then the conditional probability that $A \in \mathcal{S}_t$ is maximized when $E = \emptyset$. Thus,

$$\mathbb{P}\left(A \in S_t \mid B \in S_t\right) \leqslant \mathbb{P}\left(A \in S_t \mid E = \emptyset\right) \leqslant \frac{\mathbb{P}(A \in S_t)}{\mathbb{P}(E = \emptyset)} = b^{\binom{\ell}{2}}\,\mathbb{P}(A \in S_t).$$

Therefore,

$$p(k,\ell) = \mathbb{P}(A \in S_t \mid B \in S_t)\,\mathbb{P}(B \in S_t) \leqslant b^{\binom{\ell}{2}}\left(\mathbb{P}(A \in S_t)\right)^2. \tag{33}$$

On the other hand, for every $\ell \leqslant k$,

$$\binom{k}{\ell}\binom{n-k}{k-\ell} \leqslant k^\ell \frac{k^\ell}{(n-k)^\ell}\binom{n}{k}.$$

Using the estimate of (33) along with the above inequality, we have

$$\Delta_1 \leqslant \left(\binom{n}{k}\mathbb{P}(A \in S_t)\right)^2 \sum_{\ell=2}^{k-\lfloor(t+3)\log_b \log_b n\rfloor}\left(\frac{k^2}{n-k}\right)^\ell b^{\binom{\ell}{2}}$$

$$\leqslant \mathbb{E}^2(\alpha_t^{(k)}(G_{n,p})) \sum_{\ell=2}^{k-\lfloor(t+3)\log_b \log_b n\rfloor}\left(\frac{k^2}{n-k}\right)^\ell b^{\binom{\ell}{2}}. \tag{34}$$

If we set $s_\ell = (k^2/(n-k))^\ell b^{\binom{\ell}{2}}$, then $s_{\ell+1}/s_\ell = b^\ell k^2/(n-k)$. So the sequence $\{s_\ell\}$ is strictly decreasing for $\ell < \log_b(n-k) - 2\log_b k$ and is strictly increasing for $\ell > \log_b(n-k) - 2\log_b k$. So

$$\max\{s_\ell : 2 \leqslant \ell \leqslant k - \lfloor(t+3)\log_b \log_b n\rfloor\} \leqslant \max\left\{s_2, s_{\lceil 2\log_b n - 4.5\log_b \log_b n\rceil}\right\}.$$

We have that $s_2 = bk^4/(n-k)^2$, but

$$s_{\lceil 2\log_b n - 4.5\log_b \log_b n\rceil} \leqslant \left(\frac{k^2}{n-k}b^{\log_b n - 2.25\log_b \log_b n}\right)^{2\log_b n - 4.5\log_b \log_b n}$$

$$\leqslant \left(\frac{4\log_b^2 n}{\log_b^{2.25} n}\right)^{2\log_b n - 4.5\log_b \log_b n} \leqslant \left(\frac{4}{\log_b^{0.25} n}\right)^{\log_b n} = o(s_2).$$

Thus, Inequality (34) now becomes for n large enough

$$\Delta_1 \leqslant \frac{bk^5}{(n-k)^2}\mathbb{E}^2(\alpha_t^{(k)}(G_{n,p})) = O\left(\frac{\ln^5 n}{n^2}\right)\mathbb{E}^2(\alpha_t^{(k)}(G_{n,p})).$$

Bounding Δ_2

Now, we will show that $\Delta_2 = o(\mathbb{E}(\alpha_t^{(k)}(G_{n,p})))$. First, we have

$$\binom{k}{\ell}\binom{n-k}{k-\ell} \leqslant (kn)^{k-\ell}.$$

We now give a rough estimate on $p(k, \ell)$. If A, B are two k-sets of vertices that overlap on ℓ vertices (and if $\deg_S(v)$ denotes the number of neighbours of v in S), then

$$\mathbb{P}(B \in S_t \mid A \in S_t) \leqslant \mathbb{P}(\forall v \in B \setminus A, \ \deg_{A \cap B}(v) \leqslant t) \leqslant \left(\binom{\ell}{\ell-t}(1-p)^{\ell-t} \right)^{k-\ell}$$

$$\leqslant \left(k^t b^{t-\ell} \right)^{k-\ell} \leqslant b^{(t \log_b k + t - k + \lfloor (t+3) \log_b \log_b n \rfloor)(k-\ell)}$$

$$= b^{(-2 \log_b n + (t+5) \log_b \log_b n + \Theta(1))(k-\ell)} \leqslant \left(\frac{\log_b^{t+6} n}{n^2} \right)^{k-\ell}.$$

Substituting these estimates into the expression for Δ_2, we obtain

$$\Delta_2 \leqslant \binom{n}{k} \mathbb{P}(A \in S_t) \sum_{\ell=k-\lfloor (t+3) \log_b \log_b n \rfloor + 1}^{k-1} \left(kn \frac{\log_b^{t+6} n}{n^2} \right)^{k-\ell}$$

$$\leqslant \mathbb{E}(\alpha_t^{(k)}(G_{n,p})) k \left(\frac{k \log_b^{t+6} n}{n} \right) = o(\mathbb{E}(\alpha_t^{(k)}(G_{n,p}))).$$

5 The t-improper chromatic number

5.1 The upper bound

Our general approach follows Bollobás [2] — see also [23]. We revisit the analysis in order to obtain an improved upper bound to match the lower bound of Panagiotou and Steger [24]. For a fixed $0 < \varepsilon < 1$, we set $\hat{\alpha}_{t,p}(n) = \lfloor \alpha_{t,p}(n) - 1 - \varepsilon \rfloor$. First, we will show the following.

Lemma 9 *A.a.s. for all $V' \subseteq V_n$ with $|V'| \geqslant n/\ln^3 n$, we have $\alpha_t(G_{n,p}[V']) \geqslant \hat{\alpha}_{t,p}(|V'|)$.*

Proof Note that (31) implies that for any $V' \subseteq V_n$ with $|V'| \geqslant n/\ln^3 n$, we have

$$\mathbb{E}\left(\alpha_{t,p}^{(\hat{\alpha}_{t,p}(|V'|))}(G_{n,p}[V']) \right) \geqslant |V'|^{1+\varepsilon+o(1)}.$$

So, applying Lemma 8, we deduce that

$$\mathbb{P}\left(\alpha_t(G_{n,p}[V']) < \hat{\alpha}_{t,p}(|V'|) \right) = \exp\left(-|V'|^{1+\varepsilon+o(1)} \right) \leqslant \exp\left(-\left(\frac{n}{\ln^3 n} \right)^{1+\varepsilon+o(1)} \right).$$

Since there are at most 2^n choices for V', the probability that there exists a set $V' \subseteq V_n$ with $|V'| \geqslant n/\ln^3 n$ and $\alpha_t(G_{n,p}[V']) < \hat{\alpha}_{t,p}(|V'|)$ is at most $2^n \exp\left(-(n/\ln^3 n)^{1+\varepsilon+o(1)} \right) = o(1)$.

We consider the following algorithm for t-improperly colouring $G_{n,p}$. Let $V' = V_n$. While $|V'| \geqslant n/\ln^3 n$, we choose and remove a t-stable set from $G_{n,p}[V']$ of size $\hat{\alpha}_t(|V'|)$. At the end, we obtain a collection of t-stable sets and each of them will form a colour class. The above lemma implies that a.a.s. we will be able to perform this algorithm, and end up with a set of at most $n/\ln^3 n$ vertices. We give a different a colour to each of these vertices. Thus, if the above algorithm "runs" for $f(n)$ steps, then $\chi_t(G_{n,p}) \leqslant f(n) + n/\ln^3 n$.

Since $\alpha_{t,p}(s) - 1 - \varepsilon$ is strictly increasing for all s that are sufficiently large, for these s the function $\hat{\alpha}_{t,p}(s)$ is non-decreasing. It is easy to see that

$$\hat{\alpha}_{t,p}\left(\left\lceil \frac{n}{\ln^3 n} \right\rceil\right) = 2\log_b n \left(1 + O\left(\frac{\ln\ln n}{\ln n}\right)\right) = \hat{\alpha}_{t,p}(n)\left(1 + O\left(\frac{\ln\ln n}{\ln n}\right)\right).$$

Since $\hat{\alpha}_{t,p}(\lceil n/\ln^3 n\rceil) \leqslant \hat{\alpha}_{t,p}(s) \leqslant \hat{\alpha}_{t,p}(n)$ for all integers $n/\ln^3 n \leqslant s \leqslant n$,

$$\hat{\alpha}_{t,p}(s) = \hat{\alpha}_{t,p}(n)\left(1 + O\left(\frac{\ln\ln n}{\ln n}\right)\right), \tag{35}$$

and therefore

$$f(n) = \frac{n}{\hat{\alpha}_{t,p}(n)}\left(1 + O\left(\frac{\ln\ln n}{\ln n}\right)\right). \tag{36}$$

Assume that there are n_i vertices available when we have removed i t-stable sets from V_n. Thus, the t-stable set that will be picked during the $(i+1)$th iteration will have size $\hat{\alpha}_{t,p}(n_i)$. Since the colouring algorithm stops as soon as there are less than $n/\ln^3 n$ vertices available, the following inequality holds:

$$\sum_{i=0}^{f(n)-2} \hat{\alpha}_{t,p}(n_i) \leqslant n\left(1 - \frac{1}{\ln^3 n}\right) \leqslant n. \tag{37}$$

Note that for all $i \geqslant 0$, $n_i = n - \sum_{j=0}^{i-1}\hat{\alpha}_{t,p}(n_j)$. Therefore,

$$\log_b n_i = \log_b\left(n - \sum_{j=0}^{i-1}\hat{\alpha}_{t,p}(n_j)\right) = \log_b n + \log_b\left(1 - \frac{\sum_{j=0}^{i-1}\hat{\alpha}_{t,p}(n_j)}{n}\right).$$

We have[1]

$$\sum_{i=0}^{f(n)-2}\log_b\left(1 - \frac{\sum_{j=0}^{i-1}\hat{\alpha}_{t,p}(n_j)}{n}\right) = \frac{1}{\ln b}\sum_{i=0}^{f(n)-2}\frac{n}{\hat{\alpha}_{t,p}(n_i)}\ln\left(1 - \frac{\sum_{j=0}^{i-1}\hat{\alpha}_{t,p}(n_j)}{n}\right)\frac{\hat{\alpha}_{t,p}(n_i)}{n}$$

$$\overset{(35)}{=} \frac{n(1+o(1))}{\hat{\alpha}_{t,p}(n)\ln b}\sum_{i=0}^{f(n)-2}\ln\left(1 - \frac{\sum_{j=0}^{i-1}\hat{\alpha}_{t,p}(n_j)}{n}\right)\frac{\hat{\alpha}_{t,p}(n_i)}{n}$$

$$= \frac{n(1+o(1))}{\hat{\alpha}_{t,p}(n)\ln b}\int_0^1 \ln(1-x)dx = -\frac{n(1+o(1))}{\hat{\alpha}_{t,p}(n)\ln b}.$$

---

[1]Note that $\int \ln(1-x)dx = -(1-x)\ln(1-x) + 1 - x$.

So

$$\sum_{i=0}^{f(n)-2} 2\log_b n_i = (f(n)-1)2\log_b n - \frac{2n(1+o(1))}{\hat{\alpha}_{t,p}(n)\ln b}. \tag{38}$$

Also,

$$\log_b \log_b n_i \geqslant \log_b \log_b \left(\frac{n}{\ln^3 n}\right) = \log_b \log_b n + \log_b \left(1 - \frac{3\log_b \ln n}{\log_b n}\right)$$

$$= \log_b \log_b n - O\left(\frac{\ln \ln n}{\ln n}\right).$$

Moreover, $\log_b \log_b n_i \leqslant \log_b \log_b n$ so, for every $t \geqslant 0$,

$$(t-2)\sum_{i=0}^{f(n)-2} \log_b \log_b n_i \geqslant (f(n)-1)(t-2)\log_b \log_b n - O\left(\frac{f(n)\ln \ln n}{\ln n}\right). \tag{39}$$

Now, Equality (38) and Inequality (39) imply that for every $t \geqslant 0$ we have

$$\sum_{i=0}^{f(n)-2} \hat{\alpha}_{t,p}(n_i) \geqslant (f(n)-1)\left(\alpha_{t,p}(n)-\varepsilon-2\right) - \frac{2n(1+o(1))}{\hat{\alpha}_{t,p}(n)\ln b} - O\left(\frac{f(n)\ln \ln n}{\ln n}\right)$$

$$\geqslant (f(n)-1)\left(\alpha_{t,p}(n)-\varepsilon-2 - \frac{2n(1+o(1))}{f(n)\hat{\alpha}_{t,p}(n)\ln b} - O\left(\frac{\ln \ln n}{\ln n}\right)\right)$$

$$\overset{(36)}{=} (f(n)-1)\left(\alpha_{t,p}(n)-\varepsilon-2 - \frac{2}{\ln b} - o(1)\right).$$

So by (37) we obtain

$$f(n)-1 \leqslant \frac{n}{\alpha_{t,p}(n)-2/\ln b - 2 - \varepsilon - o(1)}.$$

5.2 The lower bound

This proof is the generalisation of a proof of the lower bound on the chromatic number of a dense random graph given recently by Panagiotou and Steger [24]. We let $\alpha_C(n) = 2\log_b n + (t-2)\log_b \log_b n - C$, where $C = C_n > 2\log_b n + (t-2)\log_b \log_b n - \alpha_{t,p}(n)$ is some function which is $\Theta(1)$, such that $\alpha_C(n)$ is integral. We specify C at a later stage. Let $r = r_C := \lfloor n/\alpha_C(n) \rfloor$. By Theorem 1, a.a.s. there are no t-stable sets in $G_{n,p}$ of size more than $\alpha_{t,p}(n)+1$. (In fact, according to Theorem 1, we could have used the bound $\alpha_{t,p}(n)+\varepsilon$, but this would not give any improvement.) We will estimate the expected number of t-improper colourings of $G_{n,p}$ with r colours such that each colour set has size at most $\alpha_{t,p}(n)+1$. In particular, we show that, if $C < 2\log_b n + (t-2)\log_b \log_b n - \alpha_{t,p}(n)+1+2/\ln b - \varepsilon$, then this expectation converges to zero, proving that $\chi_t(G_{n,p}) > r_C$ a.a.s.

Let $\mathcal{D}$ denote the set of r-tuples of positive integers $(k_1, \ldots, k_r)$ such that $\sum_{i=1}^r k_i = n$ and $k_i \leqslant \alpha_{t,p}(n) + 1$ for all i. For some $(k_1, \ldots, k_r) \in \mathcal{D}$, let $\mathcal{P} = (P_1, \ldots, P_r)$ denote a partition of V_n into r non-empty parts $P_1, \ldots, P_r$ such that $|P_i| = k_i$. From (28), we obtain

$$\mathbb{P}(P_i \in S_t) \leqslant \left(b^{-k_i+1} \left(\frac{tbpk_i}{e} \right)^t \frac{1}{t!^2} \right)^{k_i/2} \left(1 + O\left(\frac{\ln k_i}{\sqrt{k_i}} \right) \right)^{k_i}.$$

$$\mathbb{P}(P_i \in S_t, \ \forall i) = \prod_{i=1}^r \mathbb{P}(P_i \in S_t) \leqslant \prod_{i=1}^r \left(b^{-k_i+1} \left(\frac{tbpk_i}{e} \right)^t \frac{1}{t!^2} \right)^{k_i/2} \left(1 + O\left(\frac{\ln k_i}{\sqrt{k_i}} \right) \right)^{k_i}$$

$$= b^{-\left(\sum_{i=1}^r k_i^2/2 \right)+n/2} \left(\frac{tbp}{e} \right)^{tn/2} \left(\prod_{i=1}^r k_i^{tk_i/2} \right) \frac{1}{t!^n} (1 + o(1))^n$$

$$= \left(\frac{tb^{1+1/t}p}{et!^{2/t}} \right)^{tn/2} b^{-\sum_{i=1}^r k_i^2/2} \left(\prod_{i=1}^r k_i^{tk_i/2} \right) (1 + o(1))^n,$$

uniformly over all $(k_1, \ldots, k_r) \in \mathcal{D}$. So, if $X_{t,r} = X_{t,r}(G_{n,p})$ denotes the number of t-improper colourings with r colours and with each colour class of size at most $\alpha_{t,p}(n) + 1$, then

$$\mathbb{E}(X_{t,r}) = \frac{1}{r!} \left(\frac{tb^{1+1/t}p}{et!^{2/t}} \right)^{tn/2} \sum_{(k_1, \ldots, k_r) \in \mathcal{D}} \binom{n}{k_1 \cdots k_r} b^{-\sum_{i=1}^r \left(\frac{k_i^2}{2} - \frac{t}{2} k_i \log_b k_i \right)} (1 + o(1))^n. \quad (40)$$

We call a partition where all parts differ by at most one pairwise *balanced*. In the next subsection, we give a routine proof of the following property of balanced partitions.

Lemma 10 *For large enough n, the function*

$$h(P) := -\sum_{i=1}^r \left(\frac{k_i^2}{2} - \frac{t}{2} k_i \log_b k_i \right),$$

where $P = \{P_1, \ldots, P_r\}$ is a partition of V_n with $|P_i| = k_i$, is maximised over $\mathcal{D}$ when P is a balanced partition.

Let B be a balanced partition. Then all parts have sizes either equal to $\alpha_C(n)$ or to $\alpha_C(n) + 1$ and there are less than $\alpha_C(n)$ parts that take the latter quantity. Then

$$h(B) = -\frac{n}{\alpha_C(n)} \left(\frac{\alpha_C^2(n)}{2} - \frac{t}{2} \alpha_C(n) \log_b \alpha_C(n) \right) + o(n)$$

$$= -\frac{1}{2} n\alpha_C(n) + \frac{t}{2} n \log_b \alpha_C(n) + o(n)$$

$$= -n \log_b n - \frac{t-2}{2} n \log_b \log_b n + \frac{Cn}{2} + \frac{t}{2} n \log_b 2 + \frac{t}{2} n \log_b \log_b n + o(n)$$

$$= -n \log_b n + n \log_b \log_b n + \frac{Cn}{2} + \frac{t}{2} n \log_b 2 + o(n). \quad (41)$$

Also, for any $(k_1, \ldots, k_r) \in \mathcal{D}$, we have (for n large enough)

$$\binom{n}{k_1 \cdots k_r} \leqslant \frac{n!}{(\alpha_C(n)!)^r} = O(n^{1/2}) \frac{n^n}{(\alpha_C(n))^n \left(\sqrt{2\pi\alpha_C(n)}\right)^r} \leqslant \frac{n^n}{(\alpha_C(n))^{n+r/2}}$$

$$= b^{n\log_b n - n\log_b \alpha_C(n) - \frac{r}{2}\log_b \alpha_C(n)} = b^{n\log_b n - n\log_b 2 - n\log_b \log_b n + o(n)} \quad (42)$$

since $r\log_b \alpha_C(n) \leqslant (n/\alpha_C(n))\log_b \alpha_C(n) = o(n)$. Finally, $r! \geqslant r^r e^{-r}$ and therefore

$$\frac{1}{r!} \leqslant b^{-r\log_b r + r\log_b e} = b^{-\frac{n}{\alpha_C(n)}\log_b\left(\frac{n}{\alpha_C(n)}\right) + o(n)} = b^{-n\frac{\log_b n}{\alpha_C(n)} + o(n)} = b^{-\frac{n}{2} + o(n)}. \quad (43)$$

As there are at most $\binom{n}{r} \leqslant (en/r)^r \leqslant (2e\alpha_C(n))^r \leqslant b^{r\log_b \alpha_C(n) + O(r)} = b^{o(n)}$ summands in (40), we obtain from (41), (42) and (43) that

$$\mathbb{E}(X_{t,r}) \leqslant \left(\frac{tb^{1+1/t}p}{et!^{2/t}}\right)^{tn/2} b^{\frac{Cn}{2} + \frac{t}{2}n\log_b 2 - n\log_b 2 - \frac{n}{2}}(1 + o(1))^n$$

$$= b^{\frac{n}{2}(t\log_b(2tp/e) + t - 2\log_b t! + C - 2\log_b 2)}(1 + o(1))^n.$$

Therefore, if $C = C_n < -\log_b(t^t/t!^2) - t\log_b(2bp/e) - \log_b(1/4) - \varepsilon$, i.e. if $\alpha_C(n) > \alpha_{t,p}(n) - 2/\ln b - 1 + \varepsilon$ for an arbitrary $\varepsilon > 0$, then $\mathbb{E}(X_{t,r}) = o(1)$. Thus, a.a.s.

$$\chi_t(G_{n,p}) \geqslant \frac{n}{\alpha_{t,p}(n) - \frac{2}{\ln b} - 1 + \varepsilon}.$$

5.3   Proof of Lemma 10

Suppose $h(P)$, P, k_i are defined as in Lemma 10 and furthermore assume that the parts of P are ordered by increasing size, i.e., $k_1 \leqslant \cdots \leqslant k_r$. Let $\tilde{P} = \{\tilde{P}_1, \ldots, \tilde{P}_r\}$ be a partition of V_n where for some $v \in P_r$ we have $\tilde{P}_1 = P_1 \cup \{v\}$ and $\tilde{P}_r = P_r \setminus \{v\}$, whereas $\tilde{P}_i = P_i$ for all $1 < i < r$. In other words, we obtain $\tilde{P}$ by moving a vertex from P_r to P_1. Lemma 10 easily follows from the repeated application of the following.

Lemma 11 *For large enough n, it holds that, if $k_1 < k_r - 1$, then $h(\tilde{P}) > h(P)$.*

Proof First, $k_1 \leqslant \alpha_C(n)$ and $k_r \geqslant \alpha_C(n) + 1$, since the number of parts is $r = \lfloor n/\alpha_C(n) \rfloor$.

$$2(h(\tilde{P}) - h(P)) = -(k_1 + 1)^2 + t(k_1 + 1)\log_b(k_1 + 1) - (k_r - 1)^2 + t(k_r - 1)\log_b(k_r - 1)$$

$$+ k_1{}^2 - tk_1\log_b k_1 + k_r{}^2 - tk_r\log_b k_r$$

$$= 2(k_r - k_1 - 1) + t((k_1 + 1)\log_b(k_1 + 1) - k_1\log_b k_1)$$

$$+ t((k_r - 1)\log_b(k_r - 1) - k_r\log_b k_r). \quad (44)$$

Note that

$$(k_1 + 1)\log_b(k_1 + 1) = (k_1 + 1)\log_b k_1 + (k_1 + 1)\log_b(1 + 1/k_1)$$

$$\geqslant (k_1 + 1)\log_b k_1 + (k_1 + 1)\left(1/k_1 - 1/(2k_1{}^2)\right)$$

$$= k_1\log_b k_1 + \log_b k_1 + 1 + O(1/k_1),$$

and similarly, since $k_r \geqslant \alpha_C(n) + 1 \to \infty$ as $n \to \infty$,

$$(k_r - 1)\log_b(k_r - 1) = k_r \log_b k_r - \log_b k_r + 1 - o(1).$$

Substituting these estimates into (44), we obtain

$$2(h(\tilde{P}) - h(P)) \geqslant 2(k_r - k_1 - 1) - t(\log_b k_r - \log_b k_1) + O\left(1/k_1\right). \tag{45}$$

Assume first that $k_r - k_1 \leqslant \ln\ln n$. Then $\log_b(k_r/k_1) \leqslant \log_b(k_r/(k_r - \ln\ln n)) = \log_b(1 + \ln\ln n/(k_r - \ln\ln n)) = o(1)$. But $k_r - k_1 - 1 \geqslant 1$ and $k_1 \geqslant \alpha_C(n) + 1 - \ln\ln n$ and therefore the right-hand side of (45) is positive for n large enough. If, on the other hand $k_r - k_1 > \ln\ln n$, we write $\log_b k_r = \log_b(k_r - k_1 + k_1) = \log_b(k_r - k_1) + \log_b(1 + k_1/(k_r - k_1))$. So

$$\begin{aligned}
\log_b k_r - \log_b k_1 &= \log_b(k_r - k_1) + \log_b(1 + k_1/(k_r - k_1)) - \log_b k_1 \\
&= \log_b(k_r - k_1) + \log_b(1/k_1 + 1/(k_r - k_1)) \\
&= \log_b(k_r - k_1) + \log_b(1/k_1 + o(1)) \leqslant \log_b(k_r - k_1) + 1.
\end{aligned}$$

So

$$2(k_r - k_1 - 1) - t(\log_b k_r - \log_b k_1) + O\left(1/k_1\right) \geqslant 2(k_r - k_1 - 1) - t(\log_b(k_r - k_1) - 1) \to \infty$$

as $n \to \infty$ and, by (45), $h(\tilde{P}) - h(P) > 0$ for n large enough.

Acknowledgement

We would like to thank an anonymous referee for a number of helpful comments.

References

[1] J. A. Andrews and M. S. Jacobson. On a generalization of chromatic number. In *Proceedings of the sixteenth Southeastern international conference on combinatorics, graph theory and computing (Boca Raton, Fla., 1985)*, volume 47, pages 33–48, 1985.

[2] B. Bollobás. The chromatic number of random graphs. *Combinatorica*, 8(1):49–55, 1988.

[3] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2nd edition, 2001.

[4] B. Bollobás and P. Erdős. Cliques in random graphs. *Math. Proc. Cambridge Philos. Soc.*, 80(3):419–427, 1976.

[5] B. Bollobás and A. Thomason. The structure of hereditary properties and colourings of random graphs. *Combinatorica*, 20(2):173–202, 2000.

[6] V. Chvátal. Almost all graphs with $1.44n$ edges are 3-colorable. *Random Structures Algorithms*, 2(1):11–28, 1991.

[7] L. J. Cowen, R. H. Cowen, and D. R. Woodall. Defective colorings of graphs in surfaces: partitions into subgraphs of bounded valency. *J. Graph Theory*, 10(2):187–195, 1986.

[8] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, Cambridge, 1st edition, 2009.

[9] A. M. Frieze. On the independence number of random graphs. *Discrete Math.*, 81(2):171–175, 1990.

[10] G. R. Grimmett and C. J. H. McDiarmid. On colouring random graphs. *Math. Proc. Cambridge Philos. Soc.*, 77:313–324, 1975.

[11] F. Harary. Conditional colorability in graphs. In *Graphs and applications (Boulder, Colo., 1982)*, Wiley-Intersci. Publ., pages 127–136. Wiley, New York, 1985.

[12] F. Harary and K. F. Jones. Conditional colorability. II. Bipartite variations. In *Proceedings of the Sundance conference on combinatorics and related topics (Sundance, Utah, 1985)*, volume 50, pages 205–218, 1985.

[13] S. Janson. Poisson approximation for large deviations. *Random Structures Algorithms*, 1(2):221–229, 1990.

[14] S. Janson, T. Łuczak, and A. Ruciński. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In *Random Graphs '87 (Poznań, 1987)*, pages 73–87, Chichester, 1990. Wiley.

[15] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[16] R. J. Kang and C. J. H. McDiarmid. The t-improper chromatic number of random graphs. In *Proceedings of the 4th European Conference on Combinatorics, Graph Theory and Applications (Seville, 2007)*, volume 29, pages 419–425, 2007.

[17] R. J. Kang and C. J. H. McDiarmid. The t-improper chromatic number of random graphs. *Combinatorics, Probability and Computing*, 19:87–98, 2010.

[18] T. Łuczak. The chromatic number of random graphs. *Combinatorica*, 11(1):45–54, 1991.

[19] D. Matula and L. Kučera. An expose-and-merge algorithm and the chromatic number of a random graph. In *Random Graphs '87 (Poznań, 1987)*, pages 175–187. Wiley, Chichester, 1990.

[20] D. W. Matula. On the complete subgraphs of a random graph. In *Proceedings of the 2nd Chapel Hill Conference on Combinatorial Mathematics and its Applications (Chapel Hill, N. C., 1970)*, pages 356–369, 1970.

[21] D. W. Matula. The employee party problem. *Notices AMS*, 19(2):A–382, 1972.

[22] D. W. Matula. The largest clique size in a random graph. Tech. Rep. 1987, Department of Computer Science, Southern Methodist University, Dallas, Texas., 1976.

[23] C. J. H. McDiarmid. On the chromatic number of random graphs. *Random Structures and Algorithms*, 1(4):435–442, 1990.

[24] K. Panagiotou and A. Steger. A note on the chromatic number of a dense random graph. *Discrete Mathematics*, 309(10):3420–3423, 2009.

[25] E. R. Scheinerman. Generalized chromatic numbers of random graphs. *SIAM J. Discrete Math.*, 5(1):74–80, 1992.

[26] A. Scott. On the concentration of the chromatic number of random graphs. arXiv:0806.0178v1 [math.CO], 2008.

[27] E. Shamir and J. Spencer. Sharp concentration of the chromatic number on random graphs $G_{n,p}$. *Combinatorica*, 7(1):121–129, 1987.