The Two Possible Values of the Chromatic Number of a Random Graph

Dimitris Achlioptas
optas@microsoft.com

Assaf Naor
anaor@microsoft.com

Abstract

Given $d \in (0, \infty)$ let k_d be the smallest integer k such that $d < 2k \log k$. We prove that the chromatic number of a random graph $G(n, d/n)$ is either k_d or $k_d + 1$ almost surely.

1 Introduction

The classical model of random graphs, in which each edge on n vertices is chosen independently with probability p, is denoted by $G(n, p)$. This model, introduced by Erdős and Rényi in 1960, has been studied intensively in the past four decades. We refer to the books [3, 5, 9] and the references therein for an account of many remarkable results on random graphs, as well as for their connections to various areas of mathematics. In the present paper we consider random graphs of bounded average degree, i.e., $p = d/n$ for some fixed $d \in (0, \infty)$.

One of the most important invariants of a graph G is its chromatic number $\chi(G)$. Since the mid-1970s, work on $\chi\left(G(n, p)\right)$ has been in the forefront of random graph theory, motivating some of the field's most significant developments. Indeed, one of the most fascinating facts known [11] about random graphs is that for every $d \in (0, \infty)$ there exists an integer k_d such that almost surely $\chi(G(n, d/n))$ is either k_d or $k_d + 1$. The value of k_d itself, nevertheless, has remained a mystery. The best known estimate [10] for $\chi(G(n, d/n))$ asserts that there exists a constant d_0 such that for all $d > d_0$, with probability that tends to 1 as $n \to \infty$,

$$\chi\left(G(n, d/n)\right) \in \left[\frac{d}{2\log d} + \frac{1}{2} \cdot \frac{d\ \log\log d}{(\log d)^2},\ \ \frac{d}{2\log d} + 15 \cdot \frac{d\ \log\log d}{(\log d)^2}\right] .$$

Here we prove

Theorem 1. *Given $d \in (0, \infty)$, let k_d be the smallest integer k such that $d < 2k \log k$. With probability that tends to* 1 *as $n \to \infty$,*

$$\chi\left(G(n, d/n)\right) \in \{k_d, k_d + 1\} .$$

Indeed, we determine $\chi\left(G(n, d/n)\right)$ *exactly* for roughly half of all $d \in (0, \infty)$.

Theorem 2. *If $d \in ((2k - 1)\log k, 2k \log k)$, then with probability that tends to* 1 *as $n \to \infty$,*

$$\chi\left(G(n, d/n)\right) = k + 1.$$

The first questions regarding the chromatic number of $G(n, d/n)$ were raised in the original Erdős-Rényi paper [7] from 1960. It was only until the 1990's, though, that any progress was made in that direction. By the mid 1970s, the expected value of $\chi(G(n,p))$ was known up to a factor of two for the case of fixed p, due to the work of Bollobás and Erdős [6] and Grimmett and McDiarmid [8]. This gap remained in place for another decade until, in a celebrated paper, Bollobás [4] proved that for every constant $p \in (0,1)$, almost surely $\chi(G(n,p)) = \frac{n}{2\log n}\log\left(\frac{1}{1-p}\right)(1+o(1))$. Łuczak [10] later extended this result to all $p > d_0/n$, where d_0 is a universal constant.

Questions regarding the concentration of the chromatic number were first examined in a seminal paper of Shamir and Spencer [12] in the mid-80s. They showed that $\chi(G(n,p))$ is concentrated in an interval of length $O(\sqrt{n})$ for all p and on an interval of length 5 for $p < n^{-1/6-\varepsilon}$. Łuczak [11] showed that, for $p < n^{-1/6-\varepsilon}$ the chromatic number is, in fact, concentrated on an interval of length 2. Finally, Alon and Krivelevich [2] extended 2-value concentration to all $p < n^{1/2-\varepsilon}$.

Our proof of Theorem 1 is largely analytic, breaking with the more traditional combinatorial arguments. Specifically, we establish that $G(n, d/n)$ is k-colorable by first passing to a slightly different model of random (multi)graphs and then estimating the second moment of the size of a certain subclass of their k-colorings. After an elementary probabilistic computation this estimate reduces to a certain entropy-energy inequality over $k \times k$ doubly stochastic matrices. We establish this inequality as a particular case of a general optimization principle that we formulate (Theorem 7). We believe that this principle will find further applications, for example in probability and statistical physics, as moment estimates are often characterized by similar tradeoffs.

2 Preliminaries

It will be convenient to first establish Theorem 1 in a slightly different model of random graphs. Specifically, let $G(n,m)$ denote a random (multi)graph on n vertices with precisely m independent edges, each edge formed by selecting two vertices uniformly, with replacement. We will say that a sequence of events $\mathcal{E}_n$ occurs with high probability (w.h.p.) if $\lim_{n\to\infty}\Pr[\mathcal{E}_n] = 1$ and with uniformly positive probability (w.u.p.p.) if $\liminf_{n\to\infty}\Pr[\mathcal{E}_n] > 0$. Throughout, we will consider k to be arbitrarily large but fixed, while n tends to infinity. In particular, all asymptotic notation is with respect to $n \to \infty$.

Let Y be the number of k-colorings of a random graph $G(n,m)$. Trivially, $\Pr[Y > 0] \le \mathbb{E}[Y]$. Also, $\mathbb{E}[Y] \le k^n\left(1-\frac{1}{k}\right)^m$ as, in any fixed k-partition, a random edge is monochromatic with probability at least $1/k$. Thus, $G(n, m = cn)$ is w.h.p. non-k-colorable for $c > u_k \equiv \frac{\log k}{\log k - \log(k-1)}$. We prove that this elementary bound, first suggested by Devroye, is nearly tight. Observe that $u_k < (k-1/2)\log k$ and let

$$c_k \equiv k \log k \ .$$

Theorem 3. *A random graph $G(n, m = cn)$ is w.h.p. k-colorable for $c < c_{k-1}$.*

Thus, w.h.p. $\chi(G(n,cn)) = k+1$ for $c \in (u_k, c_k)$ and $\chi(G(n,cn)) \in \{k, k+1\}$ for all $c \in [c_{k-1}, c_k)$.

A random graph $G(n,m)$ may contain $q \ge 0$ self-loops and multiple edges. Nevertheless, removing any such blemishes leaves a graph on n vertices whose edge set is uniformly random among all edge sets of size $m - q$. Moreover, note that if $m \le cn$ for any constant c, then w.h.p. $q = o(n)$. Finally, note that the edge-set of a random graph $G(n, p = 2c/n)$ is uniformly random conditional on its size, and that w.h.p. this size is in the range $cn \pm n^{2/3}$. Thus, if A is any monotone decreasing property that holds with probability at least $\theta > 0$ in $G(n, m = cn)$, then A must hold with probability at least $\theta - o(1)$ in $G(n, p = d/n)$ for any constant $d < 2c$. Similarly, for increasing properties and $d > 2c$. Thus, from our discussion above we readily see that $G(n, p = d/n)$ is w.h.p. non-k-colorable for $d \geq (2k-1)\log k > 2u_k$.

To prove Theorem 3 it will be convenient to work with random graphs $G(n,m)$ where n is a multiple of k. In particular, we will prove that for every fixed $k \geq 3$ and $c < c_{k-1}$, such graphs are k-colorable w.u.p.p. To get Theorems 1 and 3 we then argue as follows. Let n' be the smallest multiple of k greater than n. Observe that if k-colorability holds with probability θ in $G(n', c/n')$ then it must hold with probability at least θ in $G(t, c/n')$ for all $t \leq n'$. Moreover, for $n \leq t \leq n'$, $c/n' = (1 - o(1))c/t$. Thus, if $G(n', m = cn')$ is k-colorable w.u.p.p. for n' that are multiples of k, then $G(n, p = d/n)$ must be k-colorable with uniformly positive probability for all $d < 2c$. To conclude the proof we employ the following result of Achlioptas and Friedgut [1].

Theorem 4. *If $G(n, d^*/n)$ is k-colorable w.u.p.p. then for $d < d^*$, $G(n, d/n)$ is k-colorable w.h.p.*

Thus, $G(n, p = d/n)$ is w.h.p. k-colorable for all $d < 2c_{k-1}$ (and this implies Theorem 3). Combined with the fact that $G(n, p = d/n)$ is w.h.p. non-k-colorable for $d \geq (2k-1)\log k$ this implies Theorems 1 and 2.

3 The Second Moment Method and Stochastic Matrices

In the following we will only consider random graphs $G(n, m = cn)$ where n is a multiple of k. We will say that a partition of n vertices into k parts is balanced if each part contains precisely n/k vertices. Let Z be the number of balanced k-colorings. Observe that each balanced partition is a valid k-coloring with probability $(1 - 1/k)^m$. Thus, by Stirling's approximation,

$$\mathbb{E}Z = \frac{n!}{[(n/k)!]^k}\left(1 - \frac{1}{k}\right)^m = \Omega\left(\frac{1}{n^{(k-1)/2}}\right)\left[k\left(1 - \frac{1}{k}\right)^c\right]^n . \tag{1}$$

Observe that the probability that a k-partition is a valid k-coloring is maximized when the partition is balanced. Therefore, $\mathbb{E}Z$ differs by only a polynomial factor from the expected number of all k-colorings, while focusing on balanced partitions simplifies some of the ensuing calculations. We will show that $\mathbb{E}Z^2 < C \cdot (\mathbb{E}Z)^2$ for some $C = C(k,c) < \infty$. By (1) this reduces to proving

$$\mathbb{E}Z^2 = O\left(\frac{1}{n^{k-1}}\right)\left[k\left(1 - \frac{1}{k}\right)^c\right]^{2n} .$$

This will conclude the proof since, by the Cauchy-Schwartz inequality, $\Pr[Z > 0] \geq (\mathbb{E}Z)^2/\mathbb{E}Z^2$.

Since Z is the sum of $n!/[(n/k)!]^k$ indicator variables, one for each balanced partition, we see that to calculate $\mathbb{E}Z^2$ it suffices to consider all pairs of balanced partitions and, for each pair, bound the probability that both partitions are valid colorings. For any fixed pair of partitions σ and τ, since edges are chosen independently, this probability is the mth power of the probability that a random edge is bichromatic in both σ and τ. If ℓ_{ij} is the number of vertices with color i in σ and color j in τ, this single-edge probability is

$$1 - \frac{2}{k} + \sum_{i=1}^{k}\sum_{j=1}^{k}\left(\frac{\ell_{ij}}{n}\right)^2 .$$

Observe that the second term above is independent of the ℓ_{ij} only because σ and τ are balanced.

Denote by $\mathcal{D}$ the set of all $k \times k$ matrices $L = (\ell_{ij})$ of non-negative integers such that the sum of each row and each column is n/k. For any such matrix L observe that there are $n!/(\prod_{i,j} \ell_{ij}!)$ corresponding pairs of balanced partitions. Therefore,

$$\mathbb{E}Z^2 = \sum_{L \in \mathcal{D}} \frac{n!}{\prod_{i=1}^{k} \prod_{j=1}^{k} \ell_{ij}!} \cdot \left[1 - \frac{2}{k} + \sum_{i=1}^{k} \sum_{j=1}^{k} \left(\frac{\ell_{ij}}{n}\right)^2\right]^{cn} . \tag{2}$$

To get a feel for the sum in (2) observe that the term corresponding to $\ell_{ij} = n/k^2$ for all i, j, alone, is $\Theta(n^{-(k^2-1)/2}) \cdot [k(1-1/k)^c]^{2n}$. In fact, the terms corresponding to matrices for which $\ell_{ij} = n/k^2 \pm O(\sqrt{n})$ already sum to $\Theta((\mathbb{E}Z)^2)$. To establish $\mathbb{E}Z^2 = O((\mathbb{E}Z)^2)$ we will show that for $c \le c_{k-1}$ the terms in the sum decay exponentially in their distance from $(\ell_{ij}) = (n/k^2)$ and apply Lemma 1 below. The lemma amounts to applying the classical Laplace method of asymptotic analysis to the Birkhoff polytope $\mathcal{B}_k$, i.e., to the set of all $k \times k$ doubly stochastic matrices. Its proof is presented in Section 6. For a matrix $A \in \mathcal{B}_k$ we denote by ρ_A the square of its 2-norm, i.e. $\rho_A \equiv \sum_{i,j} a_{ij}^2 = \|A\|_2^2$. Moreover, let $\mathcal{H}(A)$ denote the entropy of A, which is defined as

$$\mathcal{H}(A) \equiv -\frac{1}{k} \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} \log a_{ij}. \tag{3}$$

Finally, let $J_k \in \mathcal{B}_k$ be the constant $\frac{1}{k}$ matrix.

Lemma 1 (Laplace method for the Birkhoff polytope). *Assume that for $\varphi : \mathcal{B}_k \to \mathbb{R}$ there exist a constant $\beta > 0$ such that for every $A \in \mathcal{B}_k$,*

$$\mathcal{H}(A) + \varphi(A) \le \mathcal{H}(J_k) + \varphi(J_k) - \beta(\rho_A - 1) \ .$$

Then there exists a constant $C = C(\beta, k) > 0$ such that

$$\sum_{L \in \mathcal{D}} \frac{n!}{\prod_{i=1}^{k} \prod_{j=1}^{k} \ell_{ij}!} \cdot \exp\left[n \cdot \varphi\left(\frac{k}{n} L\right)\right] \le \frac{C}{n^{k-1}} \cdot \left(k^2 e^{\varphi(J_k)}\right)^n . \tag{4}$$

Let $\mathcal{S}_k$ denote the set of all $k \times k$ singly (row) stochastic matrices. For $A \in \mathcal{S}_k$ define

$$g_c(A) = -\frac{1}{k} \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} \log a_{ij} + c \log\left(1 - \frac{2}{k} + \frac{1}{k^2} \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij}^2\right) \equiv \mathcal{H}(A) + c\,\mathcal{E}(A).$$

The heart of our analysis is the following inequality. Recall that $c_{k-1} = (k-1)\log(k-1)$.

Theorem 5. *For every $A \in \mathcal{S}_k$ and $c \le c_{k-1}$, $g_c(J_k) \ge g_c(A)$.*

Theorem 5 is a consequence of a general optimization principle, proved in Section 4, which is of independent interest. Before passing to the proof of Theorem 5, we show how it implies $\mathbb{E}Z^2 = O((\mathbb{E}Z)^2)$.

For any $A \in \mathcal{B}_k \subset \mathcal{S}_k$ and any $c < c_{k-1}$ we have

$$g_c(J_k) - g_c(A) = g_{c_{k-1}}(J_k) - g_{c_{k-1}}(A) + (c_{k-1} - c) \log\left(1 + \frac{\rho_A - 1}{(k-1)^2}\right) \ge (c_{k-1} - c) \frac{\rho_A - 1}{2(k-1)^2} \ ,$$

where for the inequality we applied Theorem 5 for $c = c_{k-1}$ and used that $\rho_A \le k$ so that $\frac{\rho_A - 1}{(k-1)^2} \le \frac{1}{2}$. We thus get that for every $c < c_{k-1}$ and every $A \in \mathcal{B}_k$

$$g_c(A) \le g_c(J_k) - \frac{c_{k-1} - c}{2(k-1)^2} \cdot (\rho_A - 1) \ . \tag{5}$$

Setting $\beta = (c_{k-1} - c)/(2(k-1)^2)$ and applying Lemma 1 with $\varphi(\cdot) = c\,\mathcal{E}(\cdot)$ yields $\mathbb{E}Z^2 = O((\mathbb{E}Z)^2)$.

One can interpret the maximization of g_c geometrically by recalling that the Birkhoff polytope has the $k!$ permutation matrices as its vertices (each such matrix having one non-zero element in each row and column) and J_k as its barycenter. By convexity, J_k is the *maximizer* of the entropy over $\mathcal{B}_k$ and the *minimizer* of the 2-norm. By the same token, the permutation matrices are *minimizers* of the entropy and *maximizers* of the 2-norm. The constant c is, thus, the control parameter determining the relative importance of each term. Indeed, it is not hard to see that for sufficiently small c, g_c is maximized by J_k and that for sufficiently large c it is not. The pertinent question is when does the transition occur, i.e., how big does c need to be before there are doubly stochastic matrices for which the norm gain away from J_k makes up for the entropy loss. Probabilistically, this is the point where the second moment explodes (relative to the square of the expectation), as the dominant contribution stops corresponding to uncorrelated k-colorings, i.e., to J_k.

The generalization from $\mathcal{B}_k$ to $\mathcal{S}_k$ is motivated by the desire to exploit the product structure of the polytope $\mathcal{S}_k$. Indeed, Theorem 5 is optimal with respect to c (up to an additive constant). Moreover, Theorem 5 does not even hold for $A = \frac{1}{k-1} J_k + \frac{k-2}{k-1} I \in \mathcal{B}_k$ when $c = c_k - 1$. Thus, 2-point concentration is optimal for the second moment method on balanced k-colorings.

4 Optimization on products of simplices

In this section we will prove an inequality which is the main step in the proof of Theorem 5. This will be done in a more general framework since the greater generality, beyond its intrinsic interest, actually leads to a simplification over the "brute force" argument.

In what follows we denote by Δ_k the k-dimensional simplex and by $S^{k-1} \subset \mathbb{R}^k$ the unit Euclidean sphere. Recall that $\mathcal{S}_k$ denotes the set of all $k \times k$ stochastic matrices. For $1 \le \rho \le k$ we denote by $\mathcal{S}_k(\rho)$ the set of all $k \times k$ stochastic matrices with 2-norm $\sqrt{\rho}$, i.e., $\mathcal{S}_k(\rho) = \left\{A \in \mathcal{S}_k;\ ||A||_2^2 = \rho\right\}$.

Definition 6. *For $\frac{1}{k} \le r \le 1$, let $s^*(r)$ be the unique vector in Δ_k of the form $(x, y, \ldots, y)$ having 2-norm $\sqrt{r}$. Observe that*

$$x = x_r \equiv \frac{1 + \sqrt{(k-1)(kr-1)}}{k} \qquad \text{and} \qquad y = y_r \equiv \frac{1 - x_r}{k-1} \ .$$

Given $h : [0,1] \to \mathbb{R}$ and an integer $k > 1$ we define the function $f : [1/k, 1] \to \mathbb{R}$ as

$$f(r) = h\left(x_r\right) + (k-1) \cdot h\left(y_r\right) \ . \tag{6}$$

Our main inequality provides a sharp bound for the maximum of entropy-like functions over stochastic matrices with a given 2-norm. In particular, in Section 5 we will prove Theorem 5 by applying Theorem 7 below to the function $h(x) = -x \log x$.

Theorem 7. *Fix an integer $k > 1$ and let $h : [0,1] \to \mathbb{R}$ be a continuous strictly concave function which is six times differentiable on $(0,1)$. Assume that $h'(0^+) = \infty$, $h'(1^-) > -\infty$ and $h^{(3)} > 0$, $h^{(4)} < 0$, $h^{(6)} < 0$ point-wise. Given $1 \le \rho \le k$, for $A \in \mathcal{S}_k(\rho)$ define*

$$H(A) = \sum_{i=1}^{k} \sum_{j=1}^{k} h(a_{ij}).$$

Then, for f as in (6)

$$H(A) \le \max\left\{ m \cdot k\, h\left(\frac{1}{k}\right) + (k-m)\cdot f\left(\frac{k\rho - m}{k(k-m)}\right);\ 0 \le m \le \frac{k(k-\rho)}{k-1} \right\}. \tag{7}$$

To understand the origin of the right hand side in (7) consider the following. Given $1 \le \rho \le k$ and an integer $0 \le m \le \frac{k(k-\rho)}{k-1}$ let $B_\rho(m) \in \mathcal{S}_k(\rho)$ be the matrix whose first m rows are the constant $1/k$ vector and the remaining $k-m$ rows are the vector $s^*\left(\frac{k\rho-m}{k(k-m)}\right)$. Define $Q_\rho(m) = H(B_\rho(m))$. Theorem 7 then asserts that $H(A) \le \max_m Q_\rho(m)$, where $0 \le m \le \frac{k(k-\rho)}{k-1}$ is real.

To prove Theorem 7 we observe that if ρ_i denotes the squared 2-norm of the i-th row then

$$\max_{A \in \mathcal{S}_k(\rho)} H(A) \ = \max_{(\rho_1,\ldots,\rho_k)\in\rho\Delta_k} \sum_{i=1}^{k} \max\left\{\hat{h}(s); s \in \Delta_k \cap \sqrt{\rho_i}\, S^{k-1}\right\} \ , \tag{8}$$

where $\hat{h}(s) = \sum_{j=1}^{k} h(s_j)$. The crucial point, reflecting the product structure of $\mathcal{S}_k$, is that to maximize the sum in (8) it suffices to maximize $H(\cdot)$ in each row independently. The maximizer of each row is characterized by the following proposition:

Proposition 2. *Fix an integer $k \ge 1$ and let $h : [0,1] \to \mathbb{R}$ be a continuous strictly concave function which is three times differentiable on $(0,1)$. Assume that $h'(0^+) = \infty$,and $h''' > 0$ point-wise. Fix $\frac{1}{k} \le r \le 1$ and assume that $s = (s_1,\ldots,s_k) \in \Delta_k \cap (\sqrt{r}\, S^{k-1})$ is such that*

$$\hat{h}(s) \equiv \sum_{i=1}^{k} h(s_i) = \max\left\{ \sum_{i=1}^{k} h(t_i);\ (t_1,\ldots,t_k) \in \Delta_k \cap \sqrt{r}\, S^{k-1} \right\},$$

Then, up to a permutation of the coordinates, $s = s^*(r)$ where $s^*(r)$ is as in Definition 6.

Thus, if ρ_i denotes the squared 2-norm of the i-th row of $A \in \mathcal{S}_k$, Proposition 2 implies that $H(A) \le F(\rho_1,\ldots,\rho_k) \equiv \sum_{i=1}^{k} f(\rho_i)$, where f is as in (6). Hence, to prove Theorem 7 it suffices to give an upper bound on $F(\rho_1,\ldots,\rho_k)$, where $(\rho_1,\ldots,\rho_k) \in \rho\Delta_k \cap [1/k,1]^k$. This is another optimization problem on a symmetric polytope and had f been concave it would be trivial. Unfortunately, f is not generally concave (in particular, not when $h(x) = -x\log x$). Nevertheless, the conditions of Theorem 7 on h suffice to impart some structure on f:

Lemma 3. *Let $h : [0,1] \to \mathbb{R}$ be six times differentiable on $(0,1)$ such that $h^{(3)} > 0$, $h^{(4)} < 0$ and $h^{(6)} < 0$ point-wise. Then the function f defined in* (6) *satisfies $f^{(3)} < 0$ point-wise.*

The following lemma is the last ingredient in the proof of Theorem 7 as it will allow us to exploit Lemma 3 to bound F.

Lemma 4. *Let $\psi : [0,1] \to \mathbb{R}$ be continuous on $[0,1]$ and three times differentiable on $(0,1)$. Assume that $\psi'(1^-) = -\infty$ and $\psi^{(3)} < 0$ point-wise. Fix $\gamma \in (0,k]$ and let $s = (s_1,\ldots,s_k) \in [0,1]^k \cap \gamma\Delta_k$. Then*

$$\Psi(s) \equiv \sum_{i=1}^{k} \psi(s_i) \le \max\left\{ m\psi(0) + (k-m)\psi\left(\frac{\gamma}{k-m}\right);\ m \in [0, k-\gamma] \right\}.$$

To prove Theorem 7 we define $\psi : [0,1] \to \mathbb{R}$ as $\psi(x) = f\left(\frac{1}{k} + \frac{k-1}{k}x\right)$. Lemma 3 and our assumptions on h imply that ψ satisfies the conditions of Lemma 4 (the assumption that $h'(0^+) = \infty$ implies that $\psi'(1^-) = -\infty$). Hence, applying Lemma 4 with $\gamma = \frac{k(\rho-1)}{k-1}$ yields Theorem 7, i.e.,

$$F(A) = \sum_{i=1}^{k} \psi\left(\frac{k\rho_i - 1}{k-1}\right) \le \max\left\{ m\,\psi(0) + (k-m)\psi\left(\frac{k(\rho-1)}{(k-1)(k-m)}\right);\ m \in \left[0, k - \frac{k(\rho-1)}{k-1}\right] \right\}.$$

4.1 Proof of Proposition 2

When $r = 1$ there is nothing to prove, so assume that $r < 1$. We begin by observing that $s_i > 0$ for every $i \in \{1, \dots, k\}$. Indeed, for the sake of contradiction, assume without loss of generality (since $r < 1$) that $s_1 = 0$ and $s_2 \geq s_3 > 0$. Fix $\varepsilon > 0$ and denote

$$\mu(\varepsilon) = \frac{s_2 - s_3 + \varepsilon - \sqrt{(s_2 - s_3 - \varepsilon)^2 + 4\varepsilon(s_3 - \varepsilon)}}{2} \quad \text{and} \quad \nu(\varepsilon) = -\mu(\varepsilon) - \varepsilon \ .$$

Denote $v(\varepsilon) = (\varepsilon, s_2 + \mu(\varepsilon), s_3 + \nu(\varepsilon), s_4, \dots, s_k)$. This choice of $\mu(\varepsilon)$ and $\nu(\varepsilon)$ ensures that for ε small enough $v(\varepsilon) \in \Delta_k \cap (\sqrt{r} \cdot S^{k-1})$. Recall that, by assumption, $h'(0) = \infty$ and $h'(x) < \infty$ for $x \in (0, 1)$. When $s_2 > s_3$ it is clear that $|\mu'(0)| < \infty$ and, thus, $\frac{d}{d\varepsilon}\hat{h}(v(\varepsilon))\Big|_{\varepsilon=0} = \infty$. On the other hand, when $s_2 = s_3 = s$ it is not hard to see that $\frac{d}{d\varepsilon}\hat{h}(v(\varepsilon))\Big|_{\varepsilon=0} = h'(0^+) - h'(s) + sh''(s) = \infty$. Thus, in both cases, we have $\frac{d}{d\varepsilon}\hat{h}(v(\varepsilon))\Big|_{\varepsilon=0} = \infty$ which contradicts the maximality of $\hat{h}(s)$.

Since $s_i > 0$ for every i (and, therefore, $s_i < 1$ as well), we may use Lagrange multipliers to deduce that there are $\lambda, \mu \in \mathbb{R}$ such that for every $i \in \{1, \dots, k\}$, $h'(s_i) = \lambda s_i + \mu$. Observe that if we let $\psi(u) = h'(u) - \lambda u$ then $\psi'' = h''' > 0$, i.e., ψ is strictly convex. It follows in particular that $|\psi^{-1}(\mu)| \leq 2$. Thus, up to a permutation of the coordinates, we may assume that there is an integer $1 \leq m \leq k$ and $a, b \in (0, 1)$ such that $s_i = a$ for $i \in \{1, \dots, m\}$ and $s_i = b$ for $i \in \{m+1, \dots, k\}$. Without loss of generality $a \geq b$ (so that in particular $a \geq 1/k$ and $b \leq 1/k$). Since $ma + (k-m)b = 1$, and $ma^2 + (k-m)b^2 = r$ it follows that

$$a = \frac{1}{k} + \frac{1}{k}\sqrt{\frac{k-m}{m}(kr-1)} \quad \text{and} \quad b = \frac{1}{k} - \frac{1}{k}\sqrt{\frac{m}{k-m}(kr-1)} \ .$$

(The choice of the minus sign in the solution of the quadratic equation defining b is correct since $b \leq 1/k$.) Define $\alpha, \beta : [1, r^{-1}] \to \mathbb{R}$ by:

$$\alpha(t) = \frac{1}{k} + \frac{1}{k}\sqrt{\frac{k-t}{t}(kr-1)} \quad \text{and} \quad \beta(t) = \frac{1}{k} - \frac{1}{k}\sqrt{\frac{t}{k-t}(kr-1)} \ .$$

Additionally, set $\varphi(t) = t \cdot h(\alpha(t)) + (k-t) \cdot h(\beta(t))$, so that $\hat{h}(s) = \varphi(m)$.

The proof will be complete once we verify that φ is strictly decreasing. Observe that

$$\begin{cases} t\alpha(t) + (k-t)\beta(t) = 1 \\ t\alpha(t)^2 + (k-t)\beta(t)^2 = r \ . \end{cases}$$

Differentiating these identities we find that

$$\begin{cases} \alpha(t) + t\alpha'(t) - \beta(t) + (k-t)\beta'(t) = 0 \\ \alpha(t)^2 + 2t\alpha(t)\alpha'(t) - \beta(t)^2 + 2(k-t)\beta(t)\beta'(t) = 0 \end{cases}$$

implying

$$\alpha'(t) = -\frac{\alpha(t) - \beta(t)}{2t} \quad \text{and} \quad \beta'(t) = -\frac{\alpha(t) - \beta(t)}{2(k-t)} \ .$$

Hence,

$$\begin{aligned} \varphi'(t) &= h(\alpha(t)) - h(\beta(t)) + t\alpha'(t)h'(\alpha(t)) + (k-t)\beta'(t)h'(\beta(t)) \\ &= h(\alpha(t)) - h(\beta(t)) - \frac{\alpha(t) - \beta(t)}{2}[h'(\alpha(t)) + h'(\beta(t))] \ . \end{aligned}$$

Therefore, in order to show that $\varphi'(t) < 0$ it is enough to prove that for every $\alpha, \beta \in [0, 1]$ such that $\alpha > \beta$,

$$h(\alpha) - h(\beta) - \frac{\alpha - \beta}{2}[h'(\alpha) + h'(\beta)] < 0 \ .$$

Fix β and define $\zeta : [\beta, 1] \to \mathbb{R}$ by $\zeta(\alpha) = h(\alpha) - h(\beta) - \frac{\alpha-\beta}{2}[h'(\alpha) + h'(\beta)]$. Now,

$$\zeta'(\alpha) = \frac{\alpha - \beta}{2}\left(\frac{h'(\alpha) - h'(\beta)}{\alpha - \beta} - h''(\alpha)\right).$$

By the Mean Value Theorem there is $\beta < \theta < \alpha$ such that

$$\zeta'(\alpha) = \frac{\alpha - \beta}{2}[h''(\theta) - h''(\alpha)] < 0,$$

since $h''' > 0$. This shows that ζ is strictly decreasing. Since $\zeta(\beta) = 0$ it follows that for $\alpha \in (\beta, 1]$, $\zeta(\alpha) < 0$, which concludes the proof.

4.2 Proof of Lemma 3

If we make the linear change of variable $z = (k-1)(kx-1)$ then our goal is to show that the function $g : [0, (k-1)^2] \to \mathbb{R}$ given by:

$$g(z) = h\left(\frac{1}{k} + \frac{\sqrt{z}}{k}\right) + (k-1)h\left(\frac{1}{k} - \frac{\sqrt{z}}{k(k-1)}\right),$$

satisfies $g''' < 0$ point-wise. Direct differentiation gives

$$\begin{aligned} 8kz^{5/2}g'''(z) \quad = \quad & \frac{z}{k^2}\left[h'''\left(\frac{1}{k} + \frac{\sqrt{z}}{k}\right) - \frac{1}{(k-1)^2}h'''\left(\frac{1}{k} - \frac{\sqrt{z}}{k(k-1)}\right)\right] \\ & -\frac{3\sqrt{z}}{k}\left[h''\left(\frac{1}{k} + \frac{\sqrt{z}}{k}\right) + \frac{1}{k-1}h''\left(\frac{1}{k} - \frac{\sqrt{z}}{k(k-1)}\right)\right] \\ & +3\left[h'\left(\frac{1}{k} + \frac{\sqrt{z}}{k}\right) - h'\left(\frac{1}{k} - \frac{\sqrt{z}}{k(k-1)}\right)\right]. \end{aligned}$$

Denote $a = \frac{\sqrt{z}}{k}$ and $b = \frac{\sqrt{z}}{k(k-1)}$. Then $8kz^{5/2}g'''(z) = \psi(a) - \psi(-b)$ where:

$$\psi(t) = t^2h'''\left(\frac{1}{k} + t\right) - 3th''\left(\frac{1}{k} + t\right) + 3h'\left(\frac{1}{k} + t\right).$$

Now:

$$\psi'(t) = t^2h''''\left(\frac{1}{k} + t\right) - th'''\left(\frac{1}{k} + t\right).$$

The assumptions on h''' and h'''' imply that $\psi'(t) < 0$ for $t > 0$, and since $a \geq b$, it follows that $\psi(a) \leq \psi(b)$. Since $8kz^{5/2}g'''(z) = \psi(a) - \psi(-b) = \big[\psi(a) - \psi(b)\big] + \big[\psi(b) - \psi(-b)\big]$, it suffices to show that for every $b > 0$, $\zeta(b) = \psi(b) - \psi(-b) < 0$. Since $\zeta(0) = 0$, this will follow once we verify that $\zeta'(b) < 0$ for $b > 0$. Observe now that $\zeta'(\beta) = b\chi(b)$, where:

$$\chi(b) = b\left[h''''\left(\frac{1}{k} + b\right) + h''''\left(\frac{1}{k} - b\right)\right] - \left[h'''\left(\frac{1}{k} + b\right) - h'''\left(\frac{1}{k} - b\right)\right].$$

Our goal is to show that $\chi(b) < 0$ for $b > 0$, and since $\chi(0) = 0$ it is enough to show that $\chi'(b) < 0$. But:

$$\chi'(b) = b\left[h^{(5)}\left(\frac{1}{k} + b\right) - h^{(5)}\left(\frac{1}{k} - b\right)\right],$$

so that the required result follows from the fact that $h^{(5)}$ is strictly decreasing.

4.3 Proof of Lemma 4

Before proving Lemma 4 we require one more preparatory fact.

Lemma 5. *Fix $0 < \gamma < k$. Let $\psi : [0,1] \to \mathbb{R}$ be continuous on $[0,1]$ and three times differentiable on $(0,1)$. Assume that $\psi'(1^-) = -\infty$ and $\psi''' < 0$ point-wise. Consider the set $A \subset \mathbb{R}^3$ defined by:*

$$A = \{(a,b,\ell) \in (0,1] \times [0,1] \times (0,k];\ b < a \text{ and } \ell a + (k-\ell)b = \gamma\}\,.$$

Define $g : A \to \mathbb{R}$ by $g(a,b,\ell) = \ell\psi(a) + (k-\ell)\psi(b)$. If $(a,b,\ell) \in A$ is such that $g(a,b,\ell) = \max_{(a,b,\ell)\in A} g(a,b,\ell)$ then $a = \gamma/\ell$.

Proof of Lemma 5. Observe that if $b = 0$ or $\ell = k$ we are done. Therefore, assume that $b > 0$ and $\ell < k$. We claim that $a < 1$. Indeed, if $a = 1$ then $b = \frac{\gamma-\ell}{k-\ell} < 1$, implying that for small enough $\varepsilon > 0$, $w(\varepsilon) \equiv \left(1-\varepsilon, b + \frac{\ell\varepsilon}{k-\ell}, \ell\right) \in A$. But $\frac{d}{d\varepsilon} g(w(\varepsilon))\big|_{\varepsilon=0} = -\ell\psi'(1^-) + \ell\psi'(b) = \infty$, which contradicts the maximality of $g(a,b,\ell)$.

Since $a \in (0,1)$ and $\ell \in (0,k)$ we can use Lagrange multipliers to deduce that there is $\lambda \in \mathbb{R}$ such that $\ell\psi'(a) = \lambda\ell$, $(k-\ell)\psi'(b) = \lambda(k-\ell)$ and $\psi(a) - \psi(b) = \lambda(a-b)$. Combined, these imply

$$\psi'(a) = \psi'(b) = \frac{\psi(a)-\psi(b)}{a-b}\quad .$$

By the Mean Value Theorem, there exists $\theta \in (b,a)$ such that $\psi'(\theta) = \frac{\psi(a)-\psi(b)}{a-b}$. But, since $\psi''' < 0$, ψ' cannot take any certain value three times, yielding the desired contradiction. $\square$

Let $s \in [0,1]^k \cap \gamma\Delta_k$ be such that $\Psi(s)$ is maximal. If $s_1 = \cdots = s_k = 1$ then we are done, so we assume that there exists i for which $s_i < 1$. Observe that in this case $s_i < 1$ for every $i \in \{1,\ldots,k\}$. Indeed, assuming the contrary we may also assume without loss of generality that $s_1 = 1$ and $s_2 < 1$. For every $\varepsilon > 0$ consider the vector $u(\varepsilon) = (1-\varepsilon, s_2+\varepsilon, s_3, \ldots, s_k)$. For ε small enough $u(\varepsilon) \in [0,1]^k \cap \gamma\Delta_k$. But $\frac{d}{d\varepsilon}\Psi(u(\varepsilon))\big|_{\varepsilon=0} = \infty$, which contradicts the maximality of $\Psi(s)$.

Without loss of generality we can further assume that $s_1, \ldots, s_q > 0$ for some $q \le k$ and $s_i = 0$ for all $i > q$. Consider the function $\tilde{\Psi}(t) = \sum_{i=1}^q \psi(t_i)$ defined on $[0,1]^q \cap \gamma\Delta_q$. Clearly, $\tilde{\Psi}$ is maximal at $(s_1,\ldots,s_q)$. Since $s_i \in (0,1)$ for every $i \in \{1,\ldots,q\}$, we may use Lagrange multipliers to deduce that there is $\lambda \in \mathbb{R}$ such that for every $i \in \{1,\ldots,q\}$, $\psi'(s_i) = \lambda$. Since $\psi''' < 0$, ψ' is strictly concave. It follows in particular that the equation $\psi'(y) = \lambda$ has at most two solutions, so that up to a permutation of the coordinates we may assume that there is an integer $0 \le \ell \le q$ and $0 \le b < a \le 1$ such that $s_i = a$ for $i \in \{1,\ldots,\ell\}$ and $s_i = b$ for $i \in \{\ell+1,\ldots,q\}$. Now, using the notation of Lemma 5 we have that $(a,b,\ell) \in A$ so that

$$\begin{aligned}\Psi(s) &= (k-q)\psi(0) + g(a,b,\ell)\\ &\le (k-q)\psi(0) + \max\left\{\theta\psi(0) + (q-\theta)\psi\left(\frac{\gamma}{q-\theta}\right);\ \theta \in [0, q-\gamma]\right\}\quad .\end{aligned}$$

5 Proof of Theorem 5

Let $h(x) = -x\log x$ and note that $h'(x) = -\log x - 1$, $h'''(x) = \frac{1}{x^2}$, $h^{(4)}(x) = \frac{-2}{x^3}$ and $h^{(6)}(x) = \frac{-24}{x^5}$, so that the conditions of Theorem 7 are satisfied in this particular case. Using Theorem 7 it is, thus, enough to show that for $c \le c_{k-1} = (k-1)\log(k-1)$,

$$\frac{m\log k}{k} + \frac{k-m}{k} f\left(\frac{k\rho - m}{k(k-m)}\right) + c\log\left(1 - \frac{2}{k} + \frac{\rho}{k^2}\right) \le \log k + 2c\log\left(1-\frac{1}{k}\right), \tag{9}$$

for every $1 \le \rho \le k$ and $0 \le m \le \frac{k(k-\rho)}{k-1}$. Here f is as in (6) for $h(x) = -x\log x$. Inequality (9) simplifies to

$$c\log\left(1+\frac{\rho-1}{(k-1)^2}\right) \le \left(1-\frac{m}{k}\right)\left[\log k - f\left(\frac{k\rho-m}{k(k-m)}\right)\right]. \tag{10}$$

Denoting $t = m/k$, $s = \rho - 1$ and using the inequality $\log(1+a) \le a$, it suffices to demand that for every $0 \le t \le 1 - \frac{s}{k-1}$ and $0 \le s \le k-1$,

$$\frac{cs}{(k-1)^2} \le (1-t)\left[f\left(\frac{1}{k}\right) - f\left(\frac{1}{k}+\frac{s}{k(1-t)}\right)\right]. \tag{11}$$

To prove (11) we define $\eta : (0, 1-1/k] \to \mathbb{R}$ by

$$\eta(y) = \frac{f\left(\frac{1}{k}\right) - f\left(\frac{1}{k}+y\right)}{y} \quad ,$$

and $\eta(0) = -f'\left(\frac{1}{k}\right) = \frac{k}{2}$, making η continuous on $[0, 1-1/k]$. Observe that (11) reduces to

$$c \le \frac{(k-1)^2}{k} \cdot \eta\left(\frac{s}{k(1-t)}\right).$$

Now, $\eta'(y) = \frac{\zeta(y)}{y^2}$, where $\zeta(y) = f\left(\frac{1}{k}+y\right) - f\left(\frac{1}{k}\right) - yf'\left(\frac{1}{k}+y\right)$. Observe that $\zeta'(y) = -yf''\left(\frac{1}{k}+y\right)$, so by Lemma 3, ζ can have at most one zero in $\left(0, 1-\frac{1}{k}\right)$. A straightforward computation gives that $\zeta\left(\frac{(k-2)^2}{k(k-1)}\right) = 0$, so η achieves its global minimum on $\left[0, 1-\frac{1}{k}\right]$ at $y \in \left\{0, \frac{(k-2)^2}{k(k-1)}, 1-\frac{1}{k}\right\}$. Direct computation gives $\eta\left(1-\frac{1}{k}\right) = \frac{k}{k-1}\cdot\log k$, $\eta\left(\frac{(k-2)^2}{k(k-1)}\right) = \frac{k-1}{k-2}\cdot\log(k-1)$ and, by definition, $\eta(0) = \frac{k}{2}$. Hence

$$\begin{aligned}\frac{(k-1)^2}{k}\cdot\eta\left(\frac{s}{k(1-t)}\right) \quad &\ge \quad \frac{(k-1)^2}{k}\cdot\min\left\{\frac{k}{2}, \frac{k-1}{k-2}\cdot\log(k-1), \frac{k}{k-1}\cdot\log k\right\} \\ &= \quad \frac{(k-1)^3}{k(k-2)}\cdot\log(k-1) > c_{k-1} \quad , \end{aligned} \tag{12}$$

where (12) follows from elementary calculus.

Remark: The above analysis shows that Theorem 5 is asymptotically optimal. Indeed, let A be the stochastic matrix whose first $k-1$ rows are the constant $1/k$ vector and whose last row is the vector $s^*(r)$, defined in Definition 6, for $r = \frac{1}{k} + \frac{(k-2)^2}{k(k-1)}$. This matrix corresponds to $m = k-1$ and $\rho = 1 + \frac{(k-2)^2}{k(k-1)}$ in (10), and a direct computation shows that any c for which Theorem 5 holds must satisfy $c < c_{k-1} + 1$.

6 Appendix: Proof of Lemma 1

If (ℓ_{ij}) are non-negative integers such that $\sum_{i,j}\ell_{ij} = n$, standard Stirling approximations imply

$$\frac{n!}{\prod_{i=1}^k\prod_{j=1}^k \ell_{ij}!} \le \left[\prod_{i=1}^k\prod_{j=1}^k\left(\frac{\ell_{ij}}{n}\right)^{-\ell_{ij}/n}\right]^n \cdot \min\left\{3\sqrt{n}, \left[(2\pi n)^{k^2-1}\prod_{i=1}^k\prod_{j=1}^k\frac{\ell_{ij}}{n}\right]^{-1/2}\right\} \quad . \tag{13}$$

Since $|\mathcal{D}| \le (n+1)^{(k-1)^2}$, the contribution to the sum in (4) of the terms for which $\rho_{\frac{k}{n}L} > 1 + 1/(4k^2)$ can, thus, be bounded by

$$3\sqrt{n}(n+1)^{(k-1)^2}\left(e^{\mathcal{H}\left(\frac{k}{n}L\right)+\log k+\varphi\left(\frac{k}{n}L\right)}\right)^n \le 3n^{k^2}\left(k^2e^{\varphi(J_k)}\right)^n \cdot e^{-\frac{\beta n}{4k^2}} = O\left(n^{-k^2}\right)\left(k^2e^{\varphi(J_k)}\right)^n. \quad (14)$$

On the other hand, if $L \in \mathcal{D}$ is such that $\rho_{\frac{k}{n}L} \le 1 + \frac{1}{4k^2}$, then for every $1 \le i,j \le k$ we have

$$\left(\frac{k}{n}\ell_{ij} - \frac{1}{k}\right)^2 \le \sum_{s=1}^{k}\sum_{t=1}^{k}\left(\frac{k}{n}\ell_{st} - \frac{1}{k}\right)^2 = \rho_{\frac{k}{n}L} - 1 \le \frac{1}{4k^2} \ .$$

Therefore, for such L we must have $\ell_{ij} \ge n/(2k^2)$ for every i,j. Combining with (13), (14) we get

$$\sum_{L\in\mathcal{D}} \frac{n!}{\prod_{i=1}^{k}\prod_{j=1}^{k}\ell_{ij}!} \cdot \exp\left[n\varphi\left(\frac{k}{n}L\right)\right] \le \frac{C(\beta,k)}{n^{(k^2-1)/2}} \cdot \left(k^2e^{\varphi(J_k)}\right)^n \cdot \sum_{L\in\mathcal{D}} e^{-\beta n\left(\frac{k^2}{n^2}\rho_L - 1\right)} \ . \quad (15)$$

Denote by $M_k(\mathbb{R})$ the space of all $k\times k$ matrices over $\mathbb{R}$ and let F be the subspace of $M_k(\mathbb{R})$ consisting of all matrices $X = (x_{ij})$ for which the sum of each row and each column is 0. The dimension of F is $(k-1)^2$. Denote by B_∞ the unit cube of $M_k(\mathbb{R})$, i.e. the set of all $k\times k$ matrices $A = (a_{ij})$ such that $a_{ij} \in [-1/2, 1/2]$ for all $1 \le i,j \le k$. For $L \in \mathcal{D}$ we define $T(L) = L - \frac{n}{k}J_k + (F \cap B_\infty)$, i.e., the tile $F\cap B_\infty$ shifted by $L - \frac{n}{k}J_k$.

Lemma 6. *For every $L \in \mathcal{D}$,*

$$e^{-\beta n\left(\frac{k^2}{n^2}\rho_L - 1\right)} \le e^{\frac{k^4\beta}{4n}} \cdot \int_{T(L)} e^{-\frac{k^2\beta}{2n}\|X\|_2^2} dX \ .$$

Proof. By the triangle inequality, we see that for any matrix X

$$\left\|L - \frac{n}{k}J_k\right\|_2^2 \ \ge\ \frac{1}{2}\|X\|_2^2 - \left\|L - \frac{n}{k}J_k - X\right\|_2^2 \ \ge\ \frac{1}{2}\|X\|_2^2 - k^2\left\|L - \frac{n}{k}J_k - X\right\|_\infty^2 \ . \quad (16)$$

Thus, for $X \in T(L)$ we have $\|X\|_2^2 \le 2\left(\frac{k^2}{n^2}\rho_L - 1\right)\left(\frac{n}{k}\right)^2 + \frac{k^2}{2}$, since $\left\|L - \frac{n}{k}J_k\right\|_2^2 = \left(\frac{k^2}{n^2}\rho_L - 1\right)\left(\frac{n}{k}\right)^2$ and $\left\|L - \frac{n}{k}J_k - X\right\|_\infty^2 \le \frac{1}{4}$. Therefore,

$$\int_{T(L)} e^{-\frac{k^2\beta}{2n}\|X\|_2^2}dX \ \ge\ \int_{T(L)} e^{-\frac{k^4\beta}{4n}} \cdot e^{-\beta n\left(\frac{k^2}{n^2}\rho_L - 1\right)}dX \ =\ e^{-\frac{k^4\beta}{4n}} \cdot e^{-\beta n\left(\frac{k^2}{n^2}\rho_L - 1\right)}\mathrm{vol}\left(F\cap B_\infty\right),$$

since $\mathrm{vol}\left(E\cap B_\infty\right) \ge 1$ for every subspace E, by a theorem of Vaaler [13]. □

Thus, to bound the second sum in (15) we apply Lemma 6 to get

$$\begin{aligned}
\sum_{L\in\mathcal{D}} e^{-\beta n\left(\frac{k^2}{n^2}\rho_L - 1\right)} &\le e^{\frac{k^4\beta}{4n}} \sum_{L\in\mathcal{D}} \int_{T(L)} e^{-\frac{k^2\beta}{2n}\|X\|_2^2}\, dX \\
&\le e^{\frac{k^4\beta}{4n}} \int_F e^{-\frac{k^2\beta}{2n}\|X\|_2^2}\, dX \\
&= e^{\frac{k^4\beta}{4n}} \int_{\mathbb{R}^{(k-1)^2}} e^{-\frac{k^2\beta}{2n}\|X\|_2^2}\, dX \\
&= e^{\frac{k^4\beta}{4n}} \left(\frac{2\pi n}{\beta k^2}\right)^{(k-1)^2/2}
\end{aligned}$$

where we have used the fact that the interiors of the "tiles" $\{T(L)\}_{L\in\mathcal{D}}$ are disjoint, that the Gaussian measure is rotationally invariant and that F is $(k-1)^2$ dimensional.

References

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

[2] N. Alon and M. Krivelevich. The concentration of the chromatic number of random graphs. *Combinatorica*, 17:303–313, 1997.

[3] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2000. With an appendix on the life and work of Paul Erdős.

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

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

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

[7] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[8] G. R. Grimmett and C. J. H. McDiarmid. On colouring random graphs. *Math. Proc. Cambridge Philos. Soc.* 77, 313-324 (1975).

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

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

[11] T. Łuczak. A note on the sharp concentration of the chromatic number of random graphs. *Combinatorica*, 11(3):295–297, 1991.

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

[13] J. D. Vaaler. A geometric inequality with applications to linear forms. *Pacific J. Math.*, 83(2):543–553, 1979.