Upper-bounding the k-colorability threshold by counting covers

Amin Coja-Oghlan*

Goethe University
Mathematics Institute
Frankfurt am Main, Germany

acoghlan@math.uni-frankfurt.de

Submitted: May 2, 2013; Accepted: Aug 16, 2013; Published: Aug 30, 2013
Mathematics Subject Classifications: 05C80, 68Q87

Abstract

Let $G(n,m)$ be the random graph on n vertices with m edges. Let $d = 2m/n$ be its average degree. We prove that $G(n,m)$ fails to be k-colorable w.h.p. if $d > 2k \ln k - \ln k - 1 + o_k(1)$. This matches a conjecture put forward on the basis of sophisticated but non-rigorous statistical physics ideas (Krzakala, Pagnani, Weigt: Phys. Rev. E **70** (2004)). The proof is based on applying the first moment method to the number of "covers", a physics-inspired concept. By comparison, a standard first moment over the number of k-colorings shows that $G(n,m)$ is not k-colorable w.h.p. if $d > 2k \ln k - \ln k$.

1 Introduction

Let $G(n,m)$ be the random graph on $V = \{1,\ldots,n\}$ with m edges. Unless specified otherwise, we let $m = \lceil dn/2 \rceil$ for a number $d > 0$ that remains fixed as $n \to \infty$. Let $k \geqslant 3$ be an n-independent integer. We say that $G(n,m)$ has a property $\mathcal{E}$ ***with high probability*** *('w.h.p.') if $\lim_{n\to\infty} \mathrm{P}[G(n,m) \in \mathcal{E}] = 1$.*

One of the longest-standing open problems in the theory of random graphs is whether there is a phase transition for k-colorability in $G(n,m)$ and, if so, at what average degree d it occurs [1, 10, 17]. Regarding existence, Achlioptas and Friedgut [1] proved that for

*The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 278857–PTCC.

any $k \geqslant 3$ there is a *sharp threshold sequence* $d_{k-\mathrm{col}}(n)$ such that for any fixed $\varepsilon > 0$ the random graph $G(n, m)$ is k-colorable w.h.p. if $2m/n < (1-\varepsilon)d_{k-\mathrm{col}}(n)$ and non-k-colorable w.h.p. if $2m/n > (1+\varepsilon)d_{k-\mathrm{col}}(n)$. To establish the existence of an actual sharp threshold, one would have to show that the sequence $d_{k-\mathrm{col}}(n)$ converges. This is widely conjectured to be the case (explicitly so in [1]) but as yet unproven.

In any case, the techniques used to prove the existence of $d_{k-\mathrm{col}}(n)$ shed no light on its location. An upper bound is easily obtained via the *first moment method*. Indeed, a simple calculation shows that for $k \geqslant 3$ and

$$d > d_{k,\mathrm{first}} = 2k \ln k - \ln k, \tag{1}$$

the expected number number of k-colorings tends to 0 as $n \to \infty$ (e.g., [4]). Hence, Markov's inequality implies that $G(n, m)$ fails to be k-colorable for $d > d_{k,\mathrm{first}}$ w.h.p. Furthermore, Achlioptas and Naor [6] used the second moment method to prove that for any $k \geqslant 3$, $G(n, m)$ is k-colorable w.h.p. if

$$d_{k-\mathrm{col}} \geqslant d_{k,\mathrm{AN}} = 2(k-1)\ln(k-1) = 2k \ln k - 2\ln k - 2 + o_k(1). \tag{2}$$

Here and throughout the paper, we use the symbol $o_k(1)$ to hide terms that tend to zero for large k. The bound (2) was recently improved [12], also via a second moment argument, for sufficiently large k to

$$d_{k-\mathrm{col}} \geqslant d_{k,\mathrm{second}} = 2k \ln k - \ln k - 2\ln 2 + o_k(1). \tag{3}$$

This leaves an additive gap of $2\ln 2 + o_k(1)$ between the upper bound (1) and the lower bound (3).

The problem of k-coloring $G(n, m)$ is closely related to the "diluted mean-field k-spin Potts antiferromagnet" model of statistical physics. Indeed, over the past decade physicists have developed sophisticated, albeit mathematically non-rigorous formalisms for identifying phase transitions in random discrete structures, the "replica method" and the "cavity method" (see [31] for details and references). Applied to the problem of k-coloring $G(n, m)$ [27, 33, 34, 38], these techniques lead to the conjecture that

$$d_{k-\mathrm{col}} = 2k \ln k - \ln k - 1 + o_k(1). \tag{4}$$

The main result of the present paper is an improved *upper* on $d_{k-\mathrm{col}}$ that matches the physics prediction (4) (at least up to the term hidden in the $o_k(1)$).

Theorem 1. *We have* $d_{k-\mathrm{col}} \leqslant 2k \ln k - \ln k - 1 + o_k(1)$.

Theorem 1 improves the naive first moment bound (1) by about an additive 1. This proves, perhaps surprisingly, that the k-colorability threshold (if it exists) does *not* coincide with the first moment bound. Furthermore, Theorem 1 narrows the gap to the lower bound (3) to $2\ln 2 - 1 + o_k(1) \approx 0.39$.

The proof of Theorem 1 is based on a concept borrowed from the "cavity method", namely the notion of *covers*. This concept is closely related to hypotheses on the "geometry" of the set of k-colorings of the random graph, which are at the core of the cavity method [26, 27, 31, 33, 37, 38]. More precisely, let $\mathcal{S}_k(G(n,m)) \subset \{1,\ldots,k\}^n$ be the set of all k-colorings of $G(n,m)$. According to the cavity method, for average degrees $(1+o_k(1))k\ln k < d < d_{k-\mathrm{col}}$ w.h.p. the set $\mathcal{S}_k(G(n,m))$ has a decomposition

$$\mathcal{S}_k(G(n,m)) = \bigcup_{i=1}^{N} \mathcal{C}_i$$

into $N = \exp(\Omega(n))$ non-empty "clusters" $\mathcal{C}_i$ such that for any two colorings σ, τ that belong to distinct clusters we have

$$\mathrm{dist}(\sigma,\tau) = |\{v \in V : \sigma(v) \neq \tau(v)\}| \geqslant \delta n \qquad \text{for some } \delta = \delta(k,d) > 0.$$

In other words, the clusters are well-separated. Furthermore, a "typical" cluster $\mathcal{C}_i$ is characterized by a set of $\Omega(n)$ "frozen" vertices, which have the same color in all colorings $\sigma \in \mathcal{C}_i$. Roughly speaking, a cover is a representation of a cluster $\mathcal{C}_i$: the cover details the colors of (most of) the frozen vertices, while the non-frozen ones are represented by the "joker color" 0. We will define covers precisely in Section 3.

The key idea behind the proof of Theorem 1 is to apply the first moment method to the number of covers. Since, according to the cavity method, covers are in one-to-one correspondence with clusters, we basically carry out a first moment argument over the number of *clusters*. The improvement over the "classical" first moment bound for the number of k-colorings results because this approach allows us to completely ignore the cluster sizes $|\mathcal{C}_i|$. Indeed, close to the k-colorability threshold the cluster sizes are conjectured to vary wildly, as has in part been established rigorously [12]. By contrast, the "classical" first moment argument amounts to putting a rather generous uniform bound on all the cluster sizes.

The clustering and "freezing" of k-colorings of $G(n,m)$ has been studied previously [2]. Formally, let us call a set F of vertices ***δ-frozen*** in a k-coloring σ of $G(n,m)$ if any other k-coloring τ such that $\tau(v) \neq \sigma(v)$ for some vertex $v \in F$ indeed satisfies

$$|\{v \in F : \sigma(v) \neq \tau(v)\}| \geqslant \delta n.$$

There is an explicitly known sharp threshold $d_{k,\mathrm{freeze}} = (1+o_k(1))k\ln k$, about half of $d_{k-\mathrm{col}}$, such that for $d > d_{k,\mathrm{freeze}}$ w.h.p. a random k-coloring of $G(n,m)$ has $\Omega(n)$ frozen vertices [32]. The threshold $d_{k,\mathrm{freeze}}$ coincides asymptotically with the largest average degree for which efficient algorithms are known to find a k-coloring of $G(n,m)$ w.h.p. [3, 21]. In fact, it has been hypothesized that the emergence of frozen vertices causes the failure of a wide class of "local search" algorithms [2, 32].

Yet the known results [2, 32] on the freezing phenomenon only show that a *random* k-coloring of $G(n,m)$ "freezes". It is not apparent that this poses an obstacle if we merely aim to find *some* k-coloring. As an important part of the proof of Theorem 1, we show that for d close to $d_{k-\mathrm{col}}$ (but strictly below the lower bound (3)), in fact *all* k-colorings of $G(n,m)$ belong to a cluster with many frozen vertices w.h.p.

Corollary 2. *Assume that $d \geqslant 2k \ln k - \ln k - 4 + o_k(1)$. There is a number $\delta_k > 0$ such that w.h.p. every k-coloring σ of the random graph $G(n,m)$ has a set $F(\sigma)$ of δ_k-frozen vertices of size $|F(\sigma)| \geqslant (1 - o_k(1))n$.*

Due to the conjectured relationship between freezing and the demise of local-search algorithms, it would be interesting to identify the precise threshold where *all* the k-colorings of $G(n,m)$ are frozen.

Further related work. The problem of coloring $G(n,m)$ has been studied intensively over the past few decades. Improving a prior result by Matula [30], Bollobás [9] determined the asymptotic value of the chromatic number of dense random graphs. Łuczak extended this result to sparse random graphs [28]. In the case that d remains fixed as $n \to \infty$, his result yields $d_{k-\mathrm{col}} = (2 + o_k(1))k \ln k$. As mentioned above, Achlioptas and Naor [6] improved this result by obtaining the lower bound (2). In addition, Łuczak's result was sharpened in [11] for $m \ll n^{5/4}$.

The problem of locating the threshold for 3-colorability has received considerable attention as well. The best current lower bound is 4.03 [5]. Moreover, Dubois and Mandler [14] proved that $d_{3-\mathrm{col}} \leqslant 4.9364$. This improved over a stream of prior results [4, 15, 18, 20, 24].

The key idea in this line of work is to estimate the first moment of the number of "rigid" colorings: for any two colors $1 \leqslant i < j \leqslant k$, every vertex of color i must have neighbors of color j [4]. Clearly, any k-colorable graph must have a rigid k-coloring. At the same time, the number of rigid k-colorings can be expected to be significantly smaller than the total number of k-colorings, and thus one might expect an improved first-moment upper bound. However, in terms of the clustering scenario put forward by physicists, it is conceivable that many clusters contain a large (in fact, exponentially large) number of rigid k-colorings. Therefore, the idea of counting rigid k-colorings seems conceptually weaker than the approach of counting clusters pursued in the present work. In fact, the improvement obtained by counting rigid colorings appears to diminish for larger k [4].

A fairly new approach to obtaining upper bounds on thresholds in random constraint satisfaction problems is the use of the *interpolation method* [8, 19, 22, 35]. This technique gives an upper bound on, e.g., the k-colorability threshold in terms of a variational problem that is related to the statistical mechanics techniques. However, this variational problem appears to be difficult to solve. Thus, it is not clear (to me) how an *explicit* upper bound as stated in Theorem 1 can be obtained from the interpolation method.

Dani, Moore and Olson [13] studied a variant of the graph coloring problem in which each pair of (u,v) of vertices comes with a random permutation $\pi_{u,v}$ of the k possible colors. This gives rise to a concept of "permuted" k-colorings. They obtained an upper bound of $2k \ln k - \ln k - 1 + o_k(1)$ on the threshold for the existence of permuted k-colorings. The proof is based on counting the total weight of k-colorings and using an isoperimetric inequality. Moreover, as pointed out in [13], physics intuition suggests that the threshold in the permuted k-coloring problem matches the "unpermuted" k-colorability threshold.

In the context of satisfiability, Maneva and Sinclair [29] used the concept of covers to obtain a conditional upper bound on the random 3-SAT threshold. Roughly speaking, the condition that they need is that w.h.p. all satisfying assignments have frozen variables. However, verifying this condition in random 3-SAT is an open problem. (That said, it is conceivable that the approach used in [29] might yield a better upper bound on the k-SAT threshold for large k.)

2 Preliminaries

Let $[k] = \{1, 2, \ldots, k\}$. Because Theorem 1 and Corollary 2 are asymptotic statements in both k and n, we may generally assume that $k \geqslant k_0$ and $n \geqslant n_0$, where k_0, n_0 are constants that are chosen sufficiently large for the various estimates to hold.

We perform asymptotic considerations with respect to both k and n. When referring to asymptotics in k, we use the notation $O_k(\cdot)$, $o_k(\cdot)$, etc. Asymptotics with respect to n are just denoted by $O(\cdot)$, $o(\cdot)$, etc.

If G is a (multi-)graph and A, B are sets of vertices, then we let $e_G(A, B)$ denote the number of A-B-edges in G. Moreover, $e_G(A)$ denotes the number of edges inside of A. If $A = \{v\}$ is a singleton, we just write $e_G(v, B)$. The reference to G is omitted where it is clear from the context.

Working with independent edges. The random graph $G(n, m)$ consists of m edges that are chosen *almost* independently. To simplify some of the arguments below, we are going to work with a random multi-graph model $G'(n, m)$ in which edges are perfectly independent. More precisely, $G'(n, m)$ is obtained as follows: let $\mathbf{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_m) \in (V \times V)^m$ be a uniformly random m-tuple of ordered pairs of vertices. In other words, each $\mathbf{e}_i$ is chosen uniformly out of all n^2 possible vertex pairs, independently of all the others. Now, let $G'(n, m)$ be the random multi-graph comprising of $\mathbf{e}_1, \ldots, \mathbf{e}_m$ viewed as undirected edges. Thus, $G'(n, m)$ may have self-loops (if $\mathbf{e}_i = (v, v)$ for some index i) as well as multiple edges (if, for example, $\mathbf{e}_i = (u, v)$ and $\mathbf{e}_j = (v, u)$ with $1 \leqslant i < j \leqslant m$ and $u \neq v$). The two random graph models are related as follows.

Lemma 3. *For any event $\mathcal{A}$ we have* $\mathrm{P}\left[G(n, m) \in \mathcal{A}\right] \leqslant O(1) \cdot \mathrm{P}\left[G'(n, m) \in \mathcal{A}\right]$.

Proof. The random graph $G'(n, m)$ has *at most* m distinct edges. Let $\mathcal{E}$ be the event that it has *exactly* m edges. This is the case iff $\mathbf{e}_1, \ldots, \mathbf{e}_m$ induce pairwise distinct undirected edges, and no self-loops. Given the event $\mathcal{E}$, $G'(n, m)$ is identical to $G(n, m)$. Hence,

$$\mathrm{P}\left[G(n, m) \in \mathcal{A}\right] = \mathrm{P}\left[G'(n, m) \in \mathcal{A} | \mathcal{E}\right] \leqslant \mathrm{P}\left[G'(n, m) \in \mathcal{A}\right] / \mathrm{P}\left[\mathcal{E}\right] \tag{5}$$

Now,

$$\mathrm{P}\left[\mathcal{E}\right] \geqslant \prod_{i=0}^{m} \left(1 - \frac{2i + n}{n^2}\right) = \exp\left[\sum_{i=0}^{m-1} \ln(1 - (2i + n)/n^2)\right] \geqslant \exp(-d - 2d^2) = \Omega(1).$$

Thus, the assertion follows from (5). □

The Chernoff bound. We need the following Chernoff bound on the tails of a binomially distributed random variable (e.g., [23, p. 21]).

Lemma 4. *Let* $\varphi(x) = (1+x)\ln(1+x) - x$*. Let* X *be a binomial random variable with mean* $\mu > 0$*. Then for any* $t > 0$ *we have*

$$\begin{aligned} \mathrm{P}\left[X > \mathrm{E}\left[X\right] + t\right] &\leqslant \exp(-\mu \cdot \varphi(t/\mu)), \\ \mathrm{P}\left[X < \mathrm{E}\left[X\right] - t\right] &\leqslant \exp(-\mu \cdot \varphi(-t/\mu)). \end{aligned}$$

In particular, for any $t > 1$ *we have* $\mathrm{P}\left[X > t\mu\right] \leqslant \exp\left[-t\mu \ln(t/\mathrm{e})\right]$.

Balls and bins. Consider a balls and bins experiment where μ balls are thrown independently and uniformly at random into ν bins. Thus, the probability of each distribution of balls into bins equals $\nu^{-\mu}$. We will need the following well-known "Poissonization lemma" (e.g., [16, Section 2.6]).

Lemma 5. *In the above experiment let* e_i *be the number of balls in bin* $i \in [\nu]$*. Moreover, let* $\lambda > 0$ *and let* $(b_i)_{i\in[\nu]}$ *be a family of independent Poisson variables, each with mean* λ*. Then for any sequence* $(t_i)_{i\in[\nu]}$ *of non-negative integers such that* $\sum_{i=1}^{\nu} t_i = \mu$ *we have*

$$\mathrm{P}\left[\forall i \in [\nu] : e_i = t_i\right] = \mathrm{P}\left[\forall i \in [\nu] : b_i = t_i \,\middle|\, \sum_{i=1}^{\nu} b_i = \mu\right].$$

Hence, the joint distribution of $(e_i)_{i\in[\nu]}$ *coincides with the joint distribution of* $(b_i)_{i\in[\nu]}$ *given* $\sum_{i=1}^{\nu} b_i = \mu$*.*

We are typically going to use Lemma 5 to obtain an *upper* bound on the probability on the left hand side. Therefore, the following simple corollary will come in handy.

Corollary 6. *With the notation of Lemma 5, assume that* $\lambda = \mu/\nu > 0$*. Then for any sequence* $(t_i)_{i\in[\nu]}$ *of non-negative integers such that* $\sum_{i=1}^{\nu} t_i = \mu$ *we have*

$$\mathrm{P}\left[\forall i \in [\nu] : e_i = t_i\right] \leqslant O(\sqrt{\mu}) \cdot \mathrm{P}\left[\forall i \in [\nu] : b_i = t_i\right].$$

Proof. Let $b = \sum_{i=1}^{\nu} b_i = \mu$. Since the b_i are independent Poisson variables with means $\lambda = \mu/\nu$, b is Poisson with mean μ. By Stirling's formula, $\mathrm{P}\left[b = \mu\right] = \mu^{\mu} \exp(-\mu)/\mu! = \Omega(\mu^{-1/2})$. Hence, Lemma 5 yields

$$\mathrm{P}\left[\forall i \in [\nu] : e_i = t_i\right] \;=\; \frac{\mathrm{P}\left[\forall i \in [\nu] : b_i = t_i\right]}{\mathrm{P}\left[b = \mu\right]} = O(\sqrt{\mu}) \cdot \mathrm{P}\left[\forall i \in [\nu] : b_i = t_i\right],$$

as claimed. □

3 Covers

Let $G = (V, E)$ be a graph, let $k \geqslant k_0$ be an integer, and let $\sigma : V \to [k]$ be a k-coloring of G. We would like to identify a set $F \subset V$ of vertices whose colors cannot be changed easily by a "local" recoloring of a few vertices. For instance, if v is a vertex that does not have a neighbor of color j for some $j \in [k] \setminus \{\sigma(v)\}$, then v can be recolored easily. More generally, we would like to say that, recursively, a vertex can be recolored easily if there is a color j such that all its neighbors of color j can be easily recolored. To formalize this, we need the following concept.

Definition 7. Let $\zeta : V \to \{0, 1, \ldots, k\}$. We call $v \in V$ ***stable*** under ζ if $\zeta(v) \neq 0$ and if for any color $j \in [k] \setminus \{\zeta(v)\}$ there are at least two neighbors u_1, u_2 of v such that $\zeta(u_1) = \zeta(u_2) = j$.

Now, consider the following ***whitening process*** that, given a k-coloring σ of G, returns a map $\hat{\sigma} : V \to \{0, 1, \ldots, k\}$; the idea is that $\hat{\sigma}(v) = 0$ for all v that are easy to recolor.

WH1. Initially, let $\hat{\sigma}(v) = \sigma(v)$ for all $v \in V$.

WH2. While there exist a vertex $v \in V$ with $\hat{\sigma}(v) \neq 0$ that is not stable under $\hat{\sigma}$, set $\hat{\sigma}(v) = 0$.

The process **WH1–WH2** is similar to processes studied in [36, 37] in the context of random graph coloring, and in [7] in the context of random k-SAT. (The term "whitening process" stems from [36].) Clearly, the final outcome $\hat{\sigma}$ of the whitening process is independent of the order in which **WH2** proceeds.

The intuition behind the whitening process is that if we attempt to recolor some stable vertex v with another color $j \in [k] \setminus \{\hat{\sigma}(v)\}$, then we will have to recolor *two* additional stable vertices u_1, u_2 as well. Hence, any attempt to recolor a stable vertex is liable to trigger an avalanche of further recolorings (unless the graph G has an abundance of short cycles, which is well-known not to be the case in the random graph $G(n, m)$ w.h.p.).

The following definition is going to lead to a neat description of the outcome of the whitening process.

Definition 8. A k-***cover*** in G is a map $\zeta : V \to \{0, 1, \ldots, k\}$ with the following properties.

CV1. There is no edge $e = \{u, v\}$ such that $\zeta(u) = \zeta(v) \neq 0$.

CV2. If $\zeta(v) \neq 0$, then v is stable under ζ.

CV3. If $\zeta(v) = 0$, then there are $i, j \in [k]$, $i \neq j$, such that v does not have a neighbor u with $\zeta(u) = i$ and v has at most one neighbor w with $\zeta(w) = j$.

The concept of covers is very closely related and, in fact, inspired by the properties of certain fixed points of the Survey Propagation message passing procedure [31]. (To my knowledge, the term "cover" has not been used previously in the context of k-colorability, although it appears to be in common use in the context of satisfiability.)

We observe that the outcome $\hat{\sigma}$ of the whitening process is indeed a k-cover. Indeed, condition **CV1** is satisfied in $\hat{\sigma}$ because the whitening process starts from a valid k-coloring. Furthermore, **CV2** holds by construction. Finally, **CV3** is satisfied for all v because otherwise there would not have been a reason for **WH2** to set $\hat{\sigma}(v) = 0$.

Now, the outcome of $\hat{\sigma}$ is the cover characterized by the following two properties.

i. For all vertices v such that $\hat{\sigma}(v) \neq 0$ we have $\hat{\sigma}(v) = \sigma(v)$.

ii. Subject to i., $|\hat{\sigma}^{-1}(0)|$ is minimum.

Let us briefly comment on the relationship between covers and clusters of k-colorings. The intention behind the whitening process is to ensure that all vertices v with $\hat{\sigma}(v) \neq 0$ are frozen under σ. But the converse is not generally true, i.e., there are going to be some vertices v that are frozen while $\hat{\sigma}(v) = 0$. This is because step **WH2** tends to err on the side of caution: it requires that v has *two* neighbors in every color class except its own. The motivation for this is that just requiring one neighbor is not generally sufficient due to the possibility of "Kempe chains" (see [32]). In the simplest case, think of a vertex v of color $i = \sigma(v)$ that has precisely one neighbor w of color $j \neq i$, whose only neighbor of color i is v itself. Then v, w are *not* frozen because they can just swap colors. By contrast, the current construction ensures that attempting to recolor one vertex v with $\hat{\sigma}(v) \neq 0$ sets off a "chain reaction". (Nonetheless, Theorem 1 *can* be proved with the weaker notion of covers obtained by relaxing the definition of "stable" to require just that each v with $\zeta(v) \neq 0$ has at least one neighbor w with $\zeta(w) = i$ for each $i \in [k] \setminus \{\zeta(v)\}$.)

Thus, the construction is intended to ensure that each clusters really only gives rise to one k-cover. The downside is that it might lead to (a potentially exponential number of) covers that do not correspond to k-colorings. Indeed, given a k-cover ζ it is not clear that we can indeed assign a color to all the vertices v with $\zeta(v) = 0$ without creating a monochromatic edge. This motivates

Definition 9. A k-cover ζ of G is ***valid*** if G has a k-coloring σ such that $\zeta = \hat{\sigma}$.

Hence, we expect that valid k-covers correspond one-to-one to clusters of k-colorings. To prove Theorem 1, we basically perform a first moment argument over the number of valid k-covers. The main task is to show that the all-0 cover (i.e., $\zeta(v) = 0$ for all vertices v) is not a valid k-cover in $G(n, m)$ w.h.p. To this end, we need to establish a few basic properties that all k-colorings of $G(n, m)$ have w.h.p. More precisely, in Section 4 we are going to prove the following via a "standard" first moment argument over k-colorings.

Proposition 10. *Assume that $k \geqslant k_0$ for a sufficiently large constant k_0. Moreover, assume that $d = 2k \ln k - \ln k - c$, with $0 \leqslant c \leqslant 4$.*

1. Let Z be the number of k-colorings of $G(n, m)$. Then $\frac{1}{n} \ln \mathrm{E}[Z] = \frac{c + o_k(1)}{2k}$.

2. *W.h.p. all k-colorings of $G(n,m)$ satisfy $|\sigma^{-1}(i)| = (1+o_k(1))\frac{n}{k}$ for all $i \in [k]$.*

3. *In fact, w.h.p. $G(n,m)$ does not have a k-coloring σ such that $|\sigma^{-1}(i) - n/k| > n/(k\ln^4 k)$ for more than $\ln^8 k$ colors $i \in [k]$.*

Building upon Proposition 10, we will establish the following properties of valid k-covers in Section 5.

Proposition 11. *There is a number k_0 such that for $k \geqslant k_0$ and $2k\ln k - \ln k - 4 \leqslant d \leqslant 2k\ln k$ any valid k-cover ζ of $G(n,m)$ has the following properties w.h.p.*

1. *We have $|\zeta^{-1}(0)| \leqslant nk^{-2/3}$.*

2. *For all $i \in [k]$ we have $|\zeta^{-1}(i)| = (1+o_k(1))n/k$.*

3. *In fact, there are no more than $\ln^9 k$ indices $i \in [k]$ such that*

$$|\zeta^{-1}(i) - n/k| > n/(k\ln^3 k).$$

Finally, in Section 6 we perform the first moment argument over k-covers.

Proposition 12. *There is $\varepsilon_k = o_k(1)$ such that for $d \geqslant 2k\ln k - \ln k - 1 + \varepsilon_k$ w.h.p. the random graph $G(n,m)$ does not have a k-cover with properties 1.–3. from Proposition 11.*

Theorem 1 is immediate from Propositions 11 and 12. Furthermore, we will prove Corollary 2 in Section 5.

4 Proof of Proposition 10

The proof of Proposition 10 is very much based on standard arguments, reminiscent but unfortunately not (quite) identical to estimates from, e.g., [4]. Suppose $d = 2k\ln k - \ln k - c$ with $0 \leqslant c \leqslant 4$. Throughout this section we work with the random graph $G'(n,m)$ with m independent edges.

Lemma 13. *Let $\nu = (\nu_1, \ldots, \nu_k)$ be a k-tuple of non-negative integers such that $\sum_{i=1}^k \nu_i = n$. Let Z_ν be the number of k-colorings σ of $G'(n,m)$ such that $|\sigma^{-1}(i)| = \nu_i$ for all $i \in [k]$. Then*

$$\ln \mathrm{E}\left[Z_\nu\right] = o(n) + \sum_{i=1}^k \nu_i \ln(n/\nu_i) + \frac{d}{2}\ln\left[1 - \sum_{i=1}^k \left(\frac{\nu_i}{n}\right)^2\right]. \tag{6}$$

Proof. Let Σ_ν be the set of all $\sigma : V \to [k]$ such that $|\sigma^{-1}(i)| = \nu_i$ for all $i \in [k]$. By Stirling's formula,

$$\ln|\Sigma_\nu| \;\; = \;\; o(n) + \sum_{i=1}^k \nu_i \ln(n/\nu_i). \tag{7}$$

Furthermore, the probability of being a k-coloring in $G'(n,m)$ is the same for all $\sigma \in \Sigma_\nu$. In fact, due to the independence of the edges in $G'(n,m)$, this probability is $q = (1-\sum_{i=1}^k (\nu_i/n)^2)^m$, because σ is a k-coloring iff each of the color classes $\sigma^{-1}(i)$ is an independent set. As $\mathrm{E}[Z_\nu] = |\Sigma_\nu| \cdot q$, the assertion follows from (7). □

Corollary 14. *Let Z be the total number of k-colorings of $G'(n,m)$. We have*

$$\frac{1}{n}\ln \mathrm{E}[Z] = \ln k + \frac{d}{2}\ln(1-1/k) + o(1) = \frac{c}{2k} + O_k(\ln k/k^2).$$

Proof. Let $\mathcal{N}$ be the set of all k-tuples $\nu = (\nu_1,\ldots,\nu_k)$ of non-negative integers such that $\sum_{i=1}^k \nu_i = n$. Then $\mathrm{E}[Z] = \sum_{\nu\in\mathcal{N}} \mathrm{E}[Z_\nu] \leqslant n^k \max_{\nu\in\mathcal{N}} \mathrm{E}[Z_\nu]$. Hence, Lemma 13 yields

$$\frac{1}{n}\ln \mathrm{E}[Z] \;=\; o(1) + \max\left\{\sum_{i=1}^k \nu_i \ln(n/\nu_i) + \frac{d}{2}\ln\left[1-\sum_{i=1}^k\left(\frac{\nu_i}{n}\right)^2\right] : \nu\in\mathcal{N}\right\}. \qquad (8)$$

Letting $\mathcal{A}$ be the set of all k-tuples $\alpha = (\alpha_1,\ldots,\alpha_k)\in[0,1]^k$ such that $\sum_{i=1}^k \alpha_i = 1$, we obtain from (8)

$$\frac{1}{n}\ln \mathrm{E}[Z] = o(1) + \max\left\{-\sum_{i=1}^k \alpha_i\ln(\alpha_i) + \frac{d}{2}\ln\left[1-\sum_{i=1}^k \alpha_i^2\right] : \alpha\in\mathcal{A}\right\}. \qquad (9)$$

The entropy function $-\sum_{i=1}^k \alpha_i \ln(\alpha_i)$ is well-known to attain its maximum at the point $\alpha = \frac{1}{k}\mathbf{1}$ with all k entries equal to $1/k$. Furthermore, the sum of squares $\sum_{i=1}^k \alpha_i^2$ attains its minimum at $\alpha = \frac{1}{k}\mathbf{1}$ as well. Hence, the term $\frac{d}{2}\ln[1-\sum_{i=1}^k\alpha_i^2]$, and thus (9), is maximized at $\frac{1}{k}\mathbf{1}$. Consequently,

$$\begin{aligned}\frac{1}{n}\ln \mathrm{E}[Z] &= \ln k + \frac{d}{2}\ln(1-1/k) + o(1) = \ln k - \frac{d}{2}\left[\frac{1}{k}+\frac{1}{2k^2}+O_k(k^{-3})\right]\\ &= \ln k - \left[k\ln k - \frac{\ln k}{2} - \frac{c}{2}\right]\cdot\left[\frac{1}{k}+\frac{1}{2k^2}+O_k(k^{-3})\right] = \frac{c}{2k} + O(\ln k/k^2),\end{aligned}$$

as claimed. □

Corollary 15. *W.h.p. all k-colorings σ of $G'(n,m)$ satisfy $|\sigma^{-1}(i)| = (1+o_k(1))\frac{n}{k}$ for all $i\in[k]$, and there is no k-coloring σ such that $|\sigma^{-1}(i) - 1/k| > 1/(k\ln^4 k)$ for more than $\ln^8 k$ colors $i\in[k]$.*

Proof. Let $\nu = (\nu_1,\ldots,\nu_k)$ be a k-tuple of non-negative integers such that $\sum_{i=1}^k \nu_i = n$. We are going to estimate $\mathrm{E}[Z_\nu]$ in terms of how much ν deviates from the "flat" vector $(n/k,\ldots,n/k)$. To this end, we compute the first two differentials of (6). Set $\alpha = (\alpha_1,\ldots,\alpha_k) = \nu/n$ and let $f(\alpha) = -\sum_{i=1}^k \alpha_i\ln\alpha_i + \frac{d}{2}\ln(1-\sum_{i=1}^k\alpha_i^2)$. Since $\sum_{i=1}^k \alpha_i = 1$, we can eliminate the variable $\alpha_k = 1 - \sum_{i=1}^{k-1} \alpha_i$. Hence, we obtain for $i, j \in [k-1]$, $i \neq j$

$$\begin{aligned}
\frac{\partial f}{\partial \alpha_i} &= \ln(\alpha_k/\alpha_i) + \frac{d(\alpha_k - \alpha_i)}{1 - \|\alpha\|_2^2}, \\
\frac{\partial^2 f}{\partial \alpha_i^2} &= -\frac{1}{\alpha_k} - \frac{1}{\alpha_i} - \frac{2d}{1 - \|\alpha\|_2^2} - \frac{2d(\alpha_k - \alpha_i)^2}{(1 - \|\alpha\|_2^2)^2}, \qquad (10)\\
\frac{\partial^2 f}{\partial \alpha_i \partial \alpha_j} &= -\frac{1}{\alpha_k} - \frac{d}{1 - \|\alpha\|_2^2} - \frac{2d(\alpha_k - \alpha_i)(\alpha_k - \alpha_j)}{(1 - \|\alpha\|_2^2)^2}. \qquad (11)
\end{aligned}$$

In particular, the first differential Df vanishes at $\alpha = \frac{1}{k}\mathbf{1}$. At this point, the Hessian $D^2 f = (\frac{\partial^2 f}{\partial \alpha_i \partial \alpha_j})_{i,j \in [k-1]}$ is negative-definite, whence $\alpha = \frac{1}{k}\mathbf{1}$ is a local maximum. Because the rank-one matrix $((\alpha_k - \alpha_i) \cdot (\alpha_k - \alpha_j))_{i,j \in [k-1]}$ is positive semidefinite for all α, (10) and (11) show that $D^2 f$ is negative-definite for all α. In fact, due to the $-\frac{2d}{1-\|\alpha\|_2^2}$ term in (10), all its eigenvalues are smaller than $-\frac{d}{1-\|\alpha\|_2^2} \leqslant -d$. Therefore, Taylor's theorem yields that

$$f(\alpha) \leqslant f(k^{-1}\mathbf{1}) - \frac{d}{2} \left\| \alpha - k^{-1}\mathbf{1} \right\|_2^2$$

for all α. Hence, Corollary 14 implies that

$$\frac{1}{n} \ln \mathrm{E}[Z_\nu] \leqslant \frac{c}{2k} + O_k(\ln k/k^2) - \frac{d}{2} \left\| \alpha - k^{-1}\mathbf{1} \right\|_2^2. \qquad (12)$$

Since $d = (2 - o_k(1))k \ln k$, the right hand side of (12) is negative if either

- $\max_{i \in [k]} |\alpha_i - k^{-1}| > (k \ln^{1/3} k)^{-1}$, or
- there are more than $\ln^8 k$ indices $i \in [k]$ such that $|\alpha_i - 1/k| > (k \ln^4 k)^{-1}$.

Thus, Markov's inequality shows that w.h.p. there is no k-coloring with either of these properties. □

Finally, Proposition 10 is immediate from Lemma 3 and Corollaries 14 and 15.

5 Proof of Proposition 11

Suppose $d = 2k \ln k - \ln k - c$ with $0 \leqslant c \leqslant 4$. Throughout this section we work with the random graph $G'(n, m)$ with m independent edges.

5.1 The core

Let $\sigma : V \to [k]$ be a map such that $|\sigma^{-1}(i)| = (1 + o_k(1))\frac{n}{k}$ for all $i \in [k]$. Moreover, let $G'(\sigma)$ be the random multi-graph $G'(n, m)$ conditional on σ being a valid k-coloring. Thus, $G'(\sigma)$ consists of m independent random edges $\mathbf{e}_1 = (\mathbf{u}_1, \mathbf{v}_1), \ldots, \mathbf{e}_m = (\mathbf{u}_m, \mathbf{v}_m)$

such that $\sigma(\mathbf{u}_i) \neq \sigma(\mathbf{v}_i)$ for all $i \in [m]$. To prove Proposition 11 we need to show that with a very high probability, a large number of vertices of $G'(\sigma)$ will remain "unscathed" by the whitening process **WH1–WH2**. To exhibit such vertices, we consider the following construction. Let $\ell = \exp(-7)\ln k$ and assume that $k \geqslant k_0$ is large enough so that $\ell > 3$. Let $V_i = \sigma^{-1}(i)$ for $i \in [k]$.

CR1 For $i \in [k]$ let $W_i = \{v \in V_i : \exists j \neq i : e(v, V_j) < 3\ell\}$ and $W = \bigcup_{i=1}^k W_i$.

CR2 Let $U = \{v \in V : \exists j : e(v, W_j) > \ell\}$.

CR3 Set $Y = U$. While there is a vertex $v \in V \setminus Y$ that has ℓ or more neighbors in Y, add v to Y.

We call the graph $G'(\sigma) - W - Y$ obtained by removing the vertices in $W \cup Y$ the ***core*** of $G'(\sigma)$.

By construction, every vertex v in the core has at least ℓ neighbors of each color $j \neq \sigma(v)$ that also belong to the core. Indeed, by **CR1** every vertex v in the core has at least 3ℓ neighbors in every color class other than their own. Furthermore, **CR2** ensures that for every color i at most ℓ of these neighbors lie in the set W. In addition, **CR2** and **CR3** ensures that no vertex $v \in V \setminus Y$ has ℓ or more neighbors in Y. Thus, for every color $i \neq \sigma(v)$ our vertex v has at least $3\ell - 2\ell = \ell$ neighbors in $V \setminus (W \cup Y)$. In effect, if $\hat{\sigma}$ is the outcome of the whitening process applied to $G'(\sigma)$, then $\hat{\sigma}(v) = \sigma(v)$ for all vertices v in the core.

The construction **CR1–CR3** has been considered previously to show that a *random* k-coloring of the random graph $G(n, m)$ has many frozen vertices w.h.p. [2, 12]. In the present context we need to perform a rather more thorough analysis of the process **CR1–CR3** to show that w.h.p. *all* k-colorings σ of $G(n, m)$ induce a non-zero cover $\hat{\sigma}$. To obtain such a strong result, we need to control the large deviations of various quantities, particularly the sizes of the sets W, W_i and U. More precisely, in Section 5.2 we prove

Lemma 16. *With probability at least* $1 - \exp(-16n/k)$ *the random graph* $G'(\sigma)$ *has the following properties.*

1. *We have* $|W| \leqslant nk^{-0.7}$.
2. *For all* $i \in [k]$ *we have* $|W_i| \leqslant \frac{n \ln\ln k}{k \ln k}$.
3. *There are no more than* $\ln^4 k$ *indices* $i \in [k]$ *such that* $|W_i| \geqslant \frac{n}{k \ln^4 k}$.

Moreover, in Section 5.3 we are going to establish

Lemma 17. *In* $G'(\sigma)$ *we have* $\mathrm{P}\left[|U| > \frac{n \ln\ln k}{k \ln k}\right] \leqslant \exp(-10n/k)$.

To estimate the size of Y we use the following observation.

Lemma 18. *W.h.p. the random graph* $G'(n, m)$ *has the following property.*

$$\text{For any set } \mathcal{Y} \subset V \text{ of size } |\mathcal{Y}| \leqslant \lceil \tfrac{2n \ln\ln k}{k \ln k} \rceil \text{ we have } e(\mathcal{Y}) < \tfrac{\ell}{2}|\mathcal{Y}| \qquad (13)$$

Proof. For any fixed set $\mathcal{Y}$ of size $0 < yn \leqslant \lceil \frac{2n \ln\ln k}{k \ln k} \rceil$ the number $e(\mathcal{Y})$ of edges spanned by $\mathcal{Y}$ in $G'(n,m)$ is binomially distributed with mean

$$\mathrm{E}\left[e(\mathcal{Y})\right] = \frac{m(yn)^2}{n^2} = (1+o_k(1))y^2dn/2 \leqslant 2y^2nk\ln k$$

Hence, by the Chernoff bound

$$\mathrm{P}\left[e(\mathcal{Y}) \geqslant y\ell n/2\right] \leqslant \exp\left[\frac{y\ell n}{2}\ln\left(\frac{y\ell n/2}{\mathrm{e}\cdot \mathrm{E}\left[e(\mathcal{Y})\right]}\right)\right] \leqslant \exp\left[\frac{y\ell n}{3}\ln(ky)\right]. \tag{14}$$

Since $ky \leqslant 3\ln\ln k/\ln k$ and $\ell = \Omega_k(\ln k)$, (14) yields

$$\mathrm{P}\left[e(\mathcal{Y}) \geqslant y\ell n/2\right] \leqslant \exp\left[3yn\ln(y)\right]. \tag{15}$$

Further, by Stirling's formula the total number of sets $\mathcal{Y} \subset V$ of size yn is

$$\binom{n}{yn} \leqslant \exp\left[yn(1-\ln y)\right] \leqslant \exp(-2yn\ln y). \tag{16}$$

Combining (15) and (16) with the union bound, we obtain

$$\mathrm{P}\left[\exists \mathcal{Y} \subset V : |\mathcal{Y}| = yn,\ e(\mathcal{Y}) \geqslant \ell\left|\mathcal{Y}\right|/2\right] \leqslant \exp\left[yn\ln(y)\right].$$

Taking the union bound over all possible sizes yn completes the proof. □

Proof of Proportion 11. By Proposition 10 w.h.p. all k-colorings σ of the random graph $G'(n,m)$ satisfy $|\sigma^{-1}(i)| = (1+o_k(1))n/k$ for all $i \in [k]$. Let us call such a k-coloring σ of $G = G'(n,m)$ *good* if it has the following two properties (and *bad* otherwise):

G1. Step **CR1** applied to G, σ yields sets $W_1, \ldots, W_k, W$ that satisfy the three properties in Lemma 16.

G2. The set U created in step **CR2** has size $|U| \leqslant \frac{n\ln\ln k}{k\ln k}$.

Let Z_{bad} be the number of bad k-colorings of $G'(n,m)$. Since $G'(\sigma)$ is just the random graph $G'(n,m)$ conditional on σ being a k-coloring, we have

$$\begin{aligned}
\mathrm{E}\left[Z_{\mathrm{bad}}\right] &= \sum_{\sigma} \mathrm{P}\left[\sigma \text{ is a } k\text{-coloring of } G(n,m)\right] \\
&\qquad \cdot \mathrm{P}\left[\sigma \text{ is bad in } G'(n,m) | \sigma \text{ is a } k\text{-coloring}\right] \\
&= \sum_{\sigma} \mathrm{P}\left[\sigma \text{ is a } k\text{-coloring of } G(n,m)\right] \cdot \mathrm{P}\left[\sigma \text{ is bad in } G'(\sigma)\right] \\
&\leqslant 2\exp(-10n/k) \cdot \sum_{\sigma} \mathrm{P}\left[\sigma \text{ is a } k\text{-coloring of } G(n,m)\right] \quad \text{[by Lemmas 16, 17]} \\
&\leqslant 2\exp(-10n/k)\cdot \exp(cn/k + o(n)) \quad \text{[by Proposition 10]} \\
&\leqslant \exp(-6n/k + o(n)) = o(1) \quad [\text{as } c \leqslant 4].
\end{aligned}$$

Hence, w.h.p. the random graph $G'(n,m)$ does not have a bad k-coloring.

Now, consider a good k-coloring σ. By Lemma 18, we may assume that (13) holds. To bound the size of the set Y created by step **CR3**, observe that each vertex that is added to Y contributes ℓ extra edges to the subgraph spanned by Y. Thus, assume that $|Y| > \frac{2n\ln\ln k}{k\ln k}$ and consider the first time step **CR3** has got a set Y' of size $\lceil\frac{2n\ln\ln k}{k\ln k}\rceil$. Then Y' spans at least $y\ell n/2$ edges, in contradiction to (13). Hence, w.h.p. $G'(n,m)$ is such that

$$\text{for } \textit{any} \text{ good } k\text{-coloring the set } Y \text{ constructed by } \textbf{CR3} \text{ has size at most } \frac{2n\ln\ln k}{k\ln k}. \quad (17)$$

If (17) is true and $G'(n,m)$ does not have a bad k-coloring, then for any k-coloring σ the set $W\cup Y$ constructed by **CR1–CR3** has size at most $|W\cup Y| \leqslant k^{-0.7}n + \frac{2n\ln\ln k}{k\ln k} \leqslant nk^{-2/3}$ (the bound on $|W|$ follows from **G1**). This shows the first property asserted in Proposition 11, because the construction **CR1–CR3** ensures that the cover $\hat\sigma$ obtained from σ via the whitening process **WH1–WH2** satisfies $\hat\sigma(v)=\sigma(v)$ for all $v\in V\setminus(W\cup Y)$. By the same token, the second assertion follows because by **G1** and (17) for every color $i\in[k]$ we have

$$|\sigma^{-1}(i)\cap(W\cup Y)| \leqslant |W_i| + |Y| = n\cdot o_k(1/k).$$

Finally, the **G1** and (17) also imply that there cannot be more than $\ln^9 k$ indices $i\in[k]$ such that $|\sigma^{-1}(i)-n/k| > n/(k\ln^3 k)$, which is the third assertion. □

Proof of Corollary 2. We claim that the vertices in $F = V\setminus(W\cup Y)$ are δ-frozen w.h.p. for $\delta = 1/(k\ln k)$. Indeed, assume that τ is another k-coloring such that the set $\Delta = \{v\in F : \tau(v)\neq\sigma(v)\}$ has size $0<|\Delta|<\delta n$. Every vertex $v\in\Delta$ has at least ℓ neighbors in Δ. Indeed, the construction **CR1–CR3** ensures that every vertex $v\in\Delta$ has at least ℓ neighbors colored $\tau(v)\neq\sigma(v)$ in Δ. Hence, Δ violates (13). Thus, the assertion follows from Lemma 18. □

5.2 Proof of Lemma 16

We begin by estimating the number of edges between different color classes. Recall that $V_i = \sigma^{-1}(i)$ for $i\in[k]$, and that we are assuming that $|V_i| = (1+o_k(1))n/k$. Let $\nu_i = |V_i|$ for $i = 1,\dots,k$.

Lemma 19. *In $G'(\sigma)$ we have*

$$\begin{aligned} \mathrm{P}\left[\min_{1\leqslant i<j\leqslant k} e(V_i,V_j) \leqslant \frac{0.99dn}{k^2}\right] &\leqslant \exp(-11n/k) \quad \textit{and} \\ \mathrm{P}\left[\max_{1\leqslant i<j\leqslant k} e(V_i,V_j) \geqslant \frac{1.01dn}{k^2}\right] &\leqslant \exp(-11n/k). \end{aligned}$$

Proof. Because the edges $\mathbf{e}_1,\dots,\mathbf{e}_m$ are chosen independently, for any pair $1\leqslant i<j\leqslant k$ the random variable $e(V_i,V_j)$ has a binomial distribution $\mathrm{Bin}(m,q_{ij})$, where

$$q_{ij} = \frac{2\nu_i\nu_j}{n^2-\sum_{l=1}^k \nu_l^2} \geqslant \frac{2\nu_i\nu_j}{n^2}.$$

Since we are assuming that $\nu_i, \nu_j = (1 + o_k(1))\frac{n}{k}$, we have $q_{ij} \geqslant (2 + o_k(1))/k^2$. Thus, $\mathrm{E}\left[e(V_i, V_j)\right] = mq_{ij} \geqslant (1 + o_k(1))dn/k^2$. Hence, the Chernoff bound yields

$$\mathrm{P}\left[e(V_i, V_j) \leqslant \frac{0.99dn}{2k^2}\right] \leqslant \exp\left[-\frac{dn}{8 \cdot 10^4 k^2}\right] \leqslant \exp(-12n/k).$$

Finally, the first assertion follows by taking a union bound over i, j. The second assertion follows analogously. □

Proof of Lemma 16. By Lemma 19 we may disregard the case that

$$\min_{1 \leqslant i < j \leqslant k} e(V_i, V_j) \leqslant \frac{0.99dn}{k^2}.$$

Thus, fix integers $(m_{ij})_{1 \leqslant i < j \leqslant k}$ such that

$$m_{ij} \geqslant \frac{0.99dn}{k^2} \quad \text{and} \quad \sum_{1 \leqslant i < j \leqslant k} m_{ij} = m. \tag{18}$$

Let $\mathcal{M}$ be the event that $e(V_i, V_j) = m_{ij}$ for all $1 \leqslant i < j \leqslant k$.

We need to get a handle on the random variables $(e(v, V_j))_{v \in V_i}$ (i.e., the number of neighbors of v in V_j) in the random graph $G'(\sigma)$. Given that $\mathcal{M}$ occurs we know that $\sum_{v \in V_i} e(v, V_j) = e(V_i, V_j) = m_{ij}$. Furthermore, because $G'(\sigma)$ consists of m *independent* random edges $\mathbf{e}_1, \ldots, \mathbf{e}_m$, given the event $\mathcal{M}$ the m_{ij} edges between V_i and V_j are chosen uniformly and independently. Therefore, we can think of the vertices in V_i as "bins" and of the m_{ij} edges as randomly tossed "balls". In particular, the average number of balls that each bin $v \in V_i$ receives is m_{ij}/ν_i. Crucially, these balls-and-bins experiments are independent for all i, j.

To analyze them, we apply Corollary 6. Thus, consider a family $(b_{vj})_{v \in V, j \in [k] \setminus \{\sigma(v)\}}$ of mutually independent Poisson variables with means $\mathrm{E}\left[b_{vj}\right] = m_{\sigma(v)j}/\nu_{\sigma(v)}$. Then for any family $(t_{vj})_{v \in V, j \in [k] \setminus \{\sigma(v)\}}$ of integers we have

$$\mathrm{P}\left[\forall v, j : e(v, V_j) = t_{vj} | \mathcal{M}\right] \leqslant \exp(o(n)) \cdot \mathrm{P}\left[\forall v, j : b_{vj} = t_{vj}\right]. \tag{19}$$

In words, the joint probability that the random variables $(e(v, V_j))_{v \in V, j \in [k] \setminus \{\sigma(v)\}}$ take certain values given that $\mathcal{M}$ occurs is dominated by the corresponding event for the random variables (b_{vj}).

If $|W_i| > \frac{n \ln \ln k}{k \ln k}$, then there are at least $N = \frac{n \ln \ln k}{k \ln k}$ vertices $v \in V_i$ such that $\min_{j \in [k] \setminus \{i\}} e(v, V_j) < 3\ell$. Thus, let $\mathcal{W}_i$ be the number of vertices $v \in V_i$ such that $\min_{j \in [k] \setminus \{i\}} b_{vj} < 3\ell$. Then (19) yields

$$\mathrm{P}\left[|W_i| \geqslant N | \mathcal{M}\right] \leqslant \exp(o(n)) \cdot \mathrm{P}\left[\mathcal{W}_i \geqslant N\right]. \tag{20}$$

Furthermore, because the random variables $(b_{vj})_{v \in V_i, j \in [k] \setminus \{i\}}$ are mutually independent, $\mathcal{W}_i$ is a binomial random variable with mean $\mathrm{E}\left[\mathcal{W}_i\right] \leqslant \nu_i q_i$, where

$$q_i = \sum_{j \in [k] \setminus \{i\}} \mathrm{P}\left[\mathrm{Po}(m_{ij}/\nu_i) \leqslant 3\ell\right].$$

Since $\nu_i = (1 + o_k(1))n/k$ and $m_{ij} \geqslant 0.99dn/k^2$, we have $\mu_{ij}/\nu_i \geqslant 0.98d/k \geqslant 1.95 \ln k$. Recalling that $\ell = \exp(-7)\ln k$, we find $\mathrm{P}[\mathrm{Po}(m_{ij}/\nu_i) \leqslant 3\ell] \leqslant k^{-1.9}$ and thus $q_i \leqslant (k-1)k^{-1.9}$. Hence,

$$\mathrm{E}[\mathcal{W}_i] \leqslant (1 + o_k(1))k^{-1.9}n \leqslant k^{-1.8}n. \tag{21}$$

Therefore, the Chernoff bound gives

$$\mathrm{P}[\mathcal{W}_i \geqslant N] \leqslant \exp\left[-N \ln\left(\frac{k^{1.8}N}{en}\right)\right] \leqslant \exp(-20n/k). \tag{22}$$

Combining (20) and (22), we obtain

$$\mathrm{P}[|W_i| \geqslant N|\mathcal{M}] \leqslant \exp(o(n) - 20n/k) \leqslant \exp(-19n/k). \tag{23}$$

Now, consider the event that there are at least $\kappa = \lceil \ln^4 k \rceil$ classes $i_1, \ldots, i_\kappa$ such that $|W_i| \geqslant N' = \frac{n}{k \ln^4 k}$. We have

$$\mathrm{P}\left[\mathcal{W}_{i_j} \geqslant N'\right] \leqslant \exp\left[-N' \ln\left(\frac{k^{1.8}N'}{en}\right)\right] \leqslant \exp\left[-\frac{1}{2}N' \ln k\right], \tag{24}$$

Furthermore, because the random variables $\mathcal{W}_{i_1}, \ldots, \mathcal{W}_{i_\kappa}$ are independent, we obtain from (19) and (24)

$$\mathrm{P}[|\{i \in [k] : |W_i| \geqslant N'\}| \geqslant \kappa|\mathcal{M}] \leqslant \binom{k}{\kappa} \exp\left[-\frac{\kappa}{2}N' \ln k\right] \leqslant \exp(-20n/k). \tag{25}$$

With respect to the event $|W| \geqslant nk^{-0.7}$, observe that by (21) the sum $\mathcal{W} = \sum_{i=1}^k \mathcal{W}_i$ is stochastically dominated by a binomial random variable with mean $nk^{-0.8}$. Therefore, by (19) and the Chernoff bound

$$\mathrm{P}\left[|W| \geqslant nk^{-0.7}|\mathcal{M}\right] \leqslant \mathrm{P}\left[\mathcal{W} \geqslant nk^{-0.7}\right] \leqslant \exp(-nk^{-0.7}) \leqslant \exp(-20n/k). \tag{26}$$

Finally, since the estimates (23), (25), (26) hold for all $\mathcal{M}$, the assertion follows from Bayes' formula. □

5.3 Proof of Lemma 17

We begin by estimating the number of edges between the sets W_i and the color class V_j. As before, we let $V_i = \sigma^{-1}(i)$ for $i \in [k]$ and $\nu_i = |V_i| = (1 + o_k(1))n/k$ for $i = 1, \ldots, k$.

Lemma 20. *In $G'(\sigma)$ we have*

$$\mathrm{P}\left[\max_{1 \leqslant i < j \leqslant k} e(W_i, V_j) \geqslant \frac{2n \ln \ln k}{k}\right] \leqslant \exp(-11n/k).$$

Proof. Fix $1 \leqslant i < j \leqslant k$. We begin by proving the following statement.

$$\text{For any set } S \subset V_i \text{ of size } |S| \leqslant \tfrac{n \ln\ln k}{k \ln k} \text{ we have } \mathrm{P}\left[e(S, V_j) > 2n \ln\ln k / k\right] \leqslant \exp(-13n/k). \tag{27}$$

Indeed, for any set S as above the number $e(S, V_j)$ of edges $\mathbf{e}_i$ that join S to V_j has a binomial distribution $\mathrm{Bin}(m, q_{j,S})$, where

$$q_{j,S} = \frac{2\nu_j |S|}{n^2 - \sum_{l=1}^{k} \nu_l^2} \leqslant \frac{3 \ln\ln k}{k^2 \ln k};$$

the last inequality follows from our assumption that $\nu_l = (1 + o_k(1))n/k$ for all $l \in [k]$. Hence,

$$\mathrm{E}\left[e(S, V_j)\right] = m q_{j,S} \leqslant \frac{3d \ln\ln k}{2k^2 \ln k} \cdot n \leqslant \frac{3 \ln\ln k}{2k} \cdot n$$

Thus, (27) follows from the Chernoff bound. Taking the union bound over all possible sets S of size $|S| \leqslant \frac{n \ln\ln k}{k \ln k}$, we obtain from (27)

$$\begin{aligned} \mathrm{P}\left[\exists S \subset V_i,\, |S| \leqslant \frac{n \ln\ln k}{k \ln k} : e(S, V_j) > \frac{2n \ln\ln k}{k}\right] &\leqslant 2^{\nu_i} \exp(-13n/k) \\ &\leqslant \exp(-12n/k). \end{aligned} \tag{28}$$

As $\mathrm{P}\left[|W_i| > \frac{n \ln\ln k}{k \ln k}\right] \leqslant \exp(-16n/k)$ by Lemma 16, the assertion follows from (28). □

Lemma 21. *Let T_i be the number of vertices $v \in V_i$ such that $\max_{j \neq i} e(v, V_j) > 100 \ln k$ and let $T = \sum_{i \in [k]} T_i$. Then in $G'(\sigma)$ we have*

$$\mathrm{P}\left[T > \frac{n}{4k \ln k}\right] \leqslant \exp(-10n/k).$$

Proof. For an integer vector $\mathbf{m} = (m_{ij})_{1 \leqslant i < j \leqslant k}$ let $\mathcal{E}_{\mathbf{m}}$ be the event that $e(V_i, V_j) = m_{ij}$ for all $1 \leqslant i < j \leqslant k$. Set $m_{ji} = m_{ij}$ for $1 \leqslant i < j \leqslant k$. By Lemma 19 we may confine ourselves to the case that $e(V_i, V_j) \leqslant \frac{2dn}{k^2}$ for all $i \neq j$. Thus, fix any $\mathbf{m}$ such that $m_{ij} \leqslant \frac{2dn}{k^2}$ for all $i < j$. Given $\mathcal{E}_{\mathbf{m}}$, for each of the m_{ij} edges between color classes V_i, V_j the actual vertex in V_i that the edge is incident with is uniformly distributed. Thus, we can think of the vertices $v \in V_i$ as bins and of edge m_{ij} edges as balls of color j, and our goal is to figure out the probability that bin v contains more than $100 \ln k$ balls colored j for some $j \neq i$. Because we are conditioning on $\mathcal{E}_{\mathbf{m}}$, these balls-and-bins experiments are independent for all color pairs $i \neq j$.

To study these balls-and-bins experiments we use Corollary 6. Let $(b_{vj})_{v \in V, j \in [k] \setminus \{\sigma(v)\}}$ be a family of mutually independent Poisson variables such that $\mathrm{E}[b_{vj}] = m_{ij}/\nu_i$ for all $v \in V_i$, $j \in [k] \setminus \{i\}$. In addition, let $\mathcal{T}_i$ be the number of vertices $v \in V_i$ such that $\max_{j \neq i} b_{vj} > 100 \ln k$ and let $\mathcal{T} = \sum_{i=1}^{k} \mathcal{T}_i$. Then Corollary 6 gives

$$\mathrm{P}\left[T > \frac{n}{4k \ln k} \,\Big|\, \mathcal{E}_{\mathbf{m}}\right] \leqslant \exp(o(n)) \cdot \mathrm{P}\left[\mathcal{T} > \frac{n}{4k \ln k}\right] \tag{29}$$

To complete the proof we need to bound $\mathrm{P}\left[\mathcal{T} > \frac{n}{4k\ln k}\right]$. For each vertex $v \in V_i$ and each $j \neq i$ we have $\mathrm{E}[b_{vj}] = m_{ij}/\nu_i \leqslant \frac{2dn}{k^2\nu_i} \leqslant 3\ln k$. Hence, by Stirling's formula

$$\mathrm{P}[b_{vj} > 100\ln k] \leqslant \sum_{s>100\ln k} \mathrm{E}[b_{vj}]^s/s! \leqslant k^{-90}.$$

Because the random variables b_{vj} are mutually independent, $\mathcal{T}$ is a sum of independent Bernoulli random variables. Applying the union bound, we thus have

$$\mathrm{P}\left[\max_{j\neq\sigma(v)} b_{vj} > 100\ln k\right] \leqslant k^{-89} \qquad \text{for any } v \in V. \tag{30}$$

Therefore, (30) shows that $\mathcal{T}$ is stochastically dominated by a binomial random variable $\mathrm{Bin}(n, k^{-89})$. Consequently, the Chernoff bound yields

$$\mathrm{P}\left[\mathcal{T} > \frac{n}{4k\ln k}\right] \leqslant \mathrm{P}\left[\mathrm{Bin}(n, k^{-89}) > \frac{n}{4k\ln k}\right] \leqslant \exp(-20n/k). \tag{31}$$

Finally, combining (29) and (31) yields the assertion. □

Proof of Lemma 17. Let $\mathbf{d} = (d_{vj})_{v\in V, j\in[k]\setminus\{\sigma(v)\}}$ be an integer vector. Moreover, let $\mathcal{E}_{\mathbf{d}}$ be the event that $e(v, V_j) = d_{vj}$ for all $v \in V$, $j \neq \sigma(v)$. We are going to estimate the size of U given that $\mathcal{E}_{\mathbf{d}}$ occurs for a vector $\mathbf{d}$ that is "compatible" with the properties established in Lemmas 19–21. More precisely, we call $\mathbf{d}$ *feasible* if the following conditions are satisfied.

i. For all $i \neq j$ we have $m_{ij} = \sum_{v\in V_i} d_{vj} \geqslant \frac{dn}{2k^2}$. Moreover, $m_{ij} = m_{ji}$.

ii. For all $i \neq j$ we have $w_{ij} = \sum_{v\in V_i : d_{vj}\leqslant 3\ell} \leqslant \frac{2n\ln\ln k}{k}$.

iii. Let $\mathcal{T}$ be the set of all v such that $\max_{j\neq\sigma(v)} d_{vj} > 100\ln k$. Then $|\mathcal{T}| \leqslant \frac{n}{4k\ln k}$.

By Lemmas 19–21, we just need to show that for any feasible $\mathbf{d}$ we have

$$\mathrm{P}\left[|U| > \frac{n\ln\ln k}{k\ln k} \,\middle|\, \mathcal{E}_{\mathbf{d}}\right] \leqslant \exp(-10n/k). \tag{32}$$

Given the event $\mathcal{E}_{\mathbf{d}}$, the total number m_{ij} of V_i-V_j-edges is fixed. So is the number w_{ji} of W_j-V_i edges. What remains random is how these edges are matched to the vertices in V_i. More specifically, think of the W_j-V_i-edges as black balls, of the $V_j \setminus W_j$-V_i-edges as white balls, and of the vertices $v \in V_i$ as bins. Each bin v has a capacity d_{vj}. Now, the balls are tossed randomly into the bins, and our objective is to figure out the number of bins that receive more than ℓ black balls. Observe that these numbers are independent for all pairs $i \neq j$ of colors.

To estimate this probability, consider a family $(b_{vj})_{v\in V, j\in[k]\setminus\{\sigma(v)\}}$ of independent binomial random variables such that b_{vj} has distribution $\mathrm{Bin}(d_{vj}, w_{ji}/m_{ji})$. Let $\mathcal{B}$ be the event that $\sum_{v\in V_i} b_{vj} = w_{ji}$ for all $i \neq j$. Furthermore, let $\mathcal{U}$ be the number of vertices v such that $\max_{j\neq\sigma(v)} b_{vj} > \ell$. Then

$$\mathrm{P}\left[|U| > \frac{n\ln\ln k}{k\ln k} \mid \mathcal{E}_{\mathbf{d}}\right] = \mathrm{P}\left[\mathcal{U} > \frac{n\ln\ln k}{k\ln k} \mid \mathcal{B}\right] \leqslant \frac{\mathrm{P}\left[\mathcal{U} > \frac{n\ln\ln k}{k\ln k}\right]}{\mathrm{P}\left[\mathcal{B}\right]}. \tag{33}$$

The sums $\sum_{v\in V_i} b_{vj}$ are binomial random variables $\mathrm{Bin}(m_{ij}, w_{ji}/m_{ji})$. Moreover, they are independent for all $i \neq j$. Therefore, Stirling's formula yields

$$\mathrm{P}\left[\mathcal{B}\right] = \prod_{i\neq j} \mathrm{P}\left[\mathrm{Bin}(m_{ij}, w_{ij}/m_{ij}) = m_{ij}\right] = \Theta(n^{-k(k-1)/2}) = \exp(o(n)). \tag{34}$$

Let $v \in V_i$ be a vertex such that for color $j \neq i$ we have $d_{vj} \leqslant 100\ln k$. Then our assumptions i. and ii. on $\mathbf{d}$ ensure that $\mathrm{E}\left[b_{vj}\right] = \frac{w_{ji}d_{vj}}{m_{ji}} \leqslant 300\ln\ln k$. Therefore, by the Chernoff bound

$$\mathrm{P}\left[b_{vj} \geqslant \ell\right] \leqslant \exp\left[-\ell\cdot\ln\left(\frac{\ell}{\mathrm{e}\cdot\mathrm{E}\left[b_{vj}\right]}\right)\right] \leqslant k^{-100}.$$

Hence, taking the union bound, we find

$$\mathrm{P}\left[\max_{j\neq\sigma(v)} b_{vj} \geqslant \ell\right] \leqslant k^{-99} \qquad \text{if } \max_{j\neq\sigma(v)} d_{vj} \leqslant 100\ln k. \tag{35}$$

Let $\mathcal{U}'$ be the number of vertices v such that $\max_{j\neq\sigma(v)} b_{vj} \geqslant \ell$ while $\max_{j\neq\sigma(v)} d_{vj} \leqslant 100\ln k$. Because the random variables b_{vj} are independent, (35) implies that $\mathcal{U}'$ is stochastically dominated by a binomial random variable $\mathrm{Bin}(n, k^{-99})$. Therefore, the Chernoff bound gives

$$\mathrm{P}\left[\mathcal{U}' \geqslant \frac{n\ln\ln k}{2k\ln k}\right] \leqslant \mathrm{P}\left[\mathrm{Bin}(n, k^{-99}) \geqslant \frac{n\ln\ln k}{2k\ln k}\right] \leqslant \exp(-11n/k). \tag{36}$$

As $\mathcal{U} \leqslant \mathcal{T} + \mathcal{U}' \leqslant \mathcal{U}' + \frac{n}{4k\ln k}$ by our assumption iii. on $\mathbf{d}$, (36) implies that $\mathrm{P}\left[\mathcal{U} \geqslant \frac{n\ln\ln k}{k\ln k}\right] \leqslant \exp(-11n/k)$. Thus, the assertion follows from (33) and (34). □

6 Proof of Proposition 12

Throughout this section, we let $\zeta : V \to \{0, 1, \ldots, k\}$, $V_i = \zeta^{-1}(i)$ and $\nu_i = |V_i|$ for $i = 0, 1, \ldots, k$. In addition, we let $\alpha_i = \nu_i/n$. We always assume that the conditions of Proposition 12 hold, namely

Z1. $|\zeta^{-1}(0)| \leqslant nk^{-2/3}$.

Z2. $|\zeta^{-1}(i)| = (1 + o_k(1))n/k$ for all $i \in [k]$.

Z3. There are no more than $\ln^9 k$ indices $i \in [k]$ such that $|\zeta^{-1}(i) - n/k| > n/(k\ln^3 k)$.

In addition, we assume that $d = 2k\ln k - \ln k - c$ for some $0 \leqslant c \leqslant 4$.

6.1 Counting covers

To prove Proposition 12 we perform a first moment argument over the number of covers ζ. Let $\mathcal{I}_\zeta$ be the event that $V_1, \ldots, V_k$ are independent sets in $G'(n,m)$. Moreover, let $\mathcal{C}_\zeta$ be the event that ζ is a k-cover in $G'(n,m)$. Clearly, $\mathcal{C}_\zeta \subset \mathcal{I}_\zeta$, and we begin by estimating the probability of the latter event. Let $F_\zeta = \sum_{j=1}^k \alpha_j^2$.

Lemma 22. *We have* $\frac{1}{n}\ln \mathrm{P}[\mathcal{I}_\zeta] = \frac{d}{2}\ln(1-F_\zeta)$.

Proof. For each of the edges $\mathbf{e}_i$ the probability of joining two vertices in V_j is $(\nu_j/n)^2 = \alpha_j^2$. Hence, the probability that $\mathbf{e}_i$ does not fall inside any of the classes $V_1, \ldots, V_k$ is equal to $1 - F_\zeta$. Thus, the assertion follows from the independence of $\mathbf{e}_1, \ldots, \mathbf{e}_m$. □

In Section 6.2 we are going to establish the following estimate of the probability of $\mathcal{C}_\sigma$.

Lemma 23. *We have* $\frac{1}{n}\ln \mathrm{P}[\mathcal{C}_\zeta|\mathcal{I}_\zeta] \leqslant \sum_{i=0}^k \alpha_i \ln p_i + o(1)$*, where*

$$p_0 = \sum_{i,j\in[k]:i\neq j} \left(\frac{1}{2} + \frac{\alpha_j d}{1-F_\zeta}\right) \exp\left[-\frac{(\alpha_i+\alpha_j)d}{1-F_\zeta}\right],$$

$$p_i = \prod_{j\in[k]\setminus\{i\}} 1 - \left(1 + \frac{\alpha_j d}{1-F_\zeta}\right)\exp\left[-\frac{\alpha_j d}{1-F_\zeta}\right] \qquad \textit{for } i = 1, \ldots, k.$$

Proof of Proposition 12. Let A be the set of all vectors $\alpha = (\alpha_0, \ldots, \alpha_k) \in [0,1]^{k+1}$ that satisfy the following three conditions (cf. **Z1–Z3**):

i. $\sum_{i=0}^k \alpha_i = 1$,

ii. We have $\alpha_0 \leqslant k^{-2/3}$ and $\alpha_i = (1+o_k(1))/k$ for $i = 1, \ldots, k$. Indeed, there are no more than $K = \ln^9 k$ indices $i \in [k]$ such that $|\alpha_i - 1/k| > k^{-1}\ln^{-3} k$.

iii. $\alpha_i n$ is an integer for $i = 0, 1, \ldots, k$.

For $\alpha \in A$ let $\mathcal{S}_\alpha$ be the set of all maps $\zeta : V \to \{0, 1, \ldots, k\}$ such that $|\zeta^{-1}(i)| = \alpha_i n$ for all i. Then

$$\mathcal{S}_\alpha = \binom{n}{\alpha_0 n, \ldots, \alpha_k n} = \binom{n}{\alpha_0 n} \cdot \binom{(1-\alpha_0)n}{\alpha_1 n, \ldots, \alpha_k n} \leqslant \binom{n}{\alpha_0 n} \cdot k^{(1-\alpha_0)n}.$$

Hence, Stirling's formula yields

$$\frac{1}{n}\ln \mathcal{S}_\alpha \leqslant -\alpha_0 \ln \alpha_0 - (1-\alpha_0)\ln((1-\alpha_0)/k). \tag{37}$$

Lemmas 22 and 23 show that for any $\zeta \in \mathcal{S}_\alpha$,

$$\frac{1}{n}\ln \mathrm{P}[\mathcal{C}_\zeta] \leqslant o(1) + \frac{d}{2}\ln(1-F_\zeta) + \sum_{i=0}^k \alpha_i \ln p_i.$$

Given the value of α_0, the sum $F_\zeta = \sum_{i=1}^k \alpha_i^2$ is minimized if $\alpha_i = (1-\alpha_0)/k$ for all $i \in [k]$. Thus,

$$\frac{1}{n}\ln \mathrm{P}\left[\mathcal{C}_\zeta\right] \leqslant o(1) + \frac{d}{2}\ln(1-(1-\alpha_0)^2/k) + \sum_{i=0}^k \alpha_i \ln p_i. \tag{38}$$

Using the approximation $\ln(1-z) = -z - z^2/2 + O(z^3)$ and recalling that $d = 2k\ln k - \ln k - c$, we see that

$$\begin{aligned}\frac{d}{2}\ln(1-(1-\alpha_0)^2/k) &= -(1-\alpha_0)^2\ln k \\ &\quad + \frac{(1-\alpha_0)^2\ln k}{2k} + \frac{c(1-\alpha_0)^2}{2k} - \frac{(1-\alpha_0)^4\ln k}{2k} + O_k(k^{-1.9}) \\ &= -(1-2\alpha_0)\ln k + \frac{c}{2k} + o_k(1/k) \quad [\text{as } \alpha_0 \leqslant k^{-2/3} \text{ by ii.}]. \end{aligned} \tag{39}$$

Furthermore, because $F_\zeta \in (0,1)$ and as $\ln(1-z) \leqslant -z$ for all $z \in (0,1)$, we get

$$\begin{aligned}\sum_{i=1}^k \alpha_i \ln p_i &\leqslant -\sum_{i,j\in[k]:i\neq j} \alpha_i(1+\alpha_j d)\exp(-\alpha_j d) \\ &= -\sum_{j\in[k]}(1-\alpha_0-\alpha_j)(1+\alpha_j d)\exp(-\alpha_j d). \end{aligned} \tag{40}$$

Since $\alpha_j = (1+o_k(1))/k$ for all $j \in [k]$ by ii. and as $d = 2k\ln k - O_k(\ln k)$, (40) yields

$$\sum_{i=1}^k \alpha_i \ln p_i \leqslant O_k(k^{-1.9}) - (1-\alpha_0)\sum_{j\in[k]}(1+\alpha_j d)\exp(-2\alpha_j k\ln k). \tag{41}$$

Moreover, applying condition ii., we obtain from (41)

$$\begin{aligned}\sum_{i=1}^k \alpha_i \ln p_i &\leqslant O_k(k^{-1.9}) + O(K\ln k)\cdot\exp(-2(1+o_k(1))\ln k) \\ &\qquad -(1-\alpha_0)(k-K)(1+2\ln k + O_k(1/\ln^2 k))\cdot \\ &\qquad\qquad \exp(-2(1+O_k(\ln^{-3}k))\ln k) \\ &\leqslant o_k(1/k) - (1-\alpha_0)\cdot\frac{1+2\ln k}{k} \qquad [\text{as } K \leqslant k^{0.01}] \\ &\leqslant o_k(1/k) - \frac{1+2\ln k}{k} \qquad [\text{as } \alpha_0 \leqslant k^{-2/3} \text{ by ii.}] \end{aligned} \tag{42}$$

Further, again because $F_\zeta \in (0,1)$ we have

$$\begin{aligned} p_0 &\leqslant \frac{1}{2}\sum_{i,j\in[k]:i\neq j}(1+2\alpha_j d)\exp\left[-(\alpha_i+\alpha_j)d\right], \\ &\leqslant O_k(k^{-3+o_k(1)}K) + \\ &\qquad \frac{k(k-1)}{2}\left[1+4\ln k + O_k(\ln^{-2}k)\right]\exp\left[-4\ln k + O(\ln^{-2}k)\right] \ [\text{by condition ii.}] \\ &\leqslant \frac{1+4\ln k + O_k(\ln^{-1}k)}{2k^2}. \end{aligned}$$

Hence,

$$\alpha_0 \ln p_0 \;\leqslant\; \alpha_0 \ln\left(\frac{1+4\ln k}{2k^2}\right) + \alpha_0 \cdot o_k(1). \tag{43}$$

Plugging (39), (42) and (43) into (38), we obtain

$$\begin{aligned}\frac{1}{n}\ln \mathrm{P}\left[\mathcal{C}_\zeta\right] \;&\leqslant\; \frac{c}{2k} - (1-2\alpha_0)\ln k - \frac{1+2\ln k}{k} \\ &\qquad +\alpha_0 \ln\left(\frac{1+4\ln k}{2k^2}\right) + \alpha_0\cdot o_k(1) + o_k(k^{-1}) \\ &= \frac{c-2-4\ln k}{2k} - \ln k + \alpha_0 \ln\left(\frac{1+4\ln k}{2}\right) + \alpha_0 \cdot o_k(1) + o_k(k^{-1}).\end{aligned} \tag{44}$$

Finally, combining (37) and (44), we get

$$\begin{aligned}\frac{1}{n}\ln(|\mathcal{S}_\alpha|\cdot \mathrm{P}\left[\mathcal{C}_\zeta\right]) \;&\leqslant\; \frac{c-2-4\ln k}{2k} - \alpha_0 \ln\left(\frac{2k\alpha_0}{1+4\ln k}\right) \\ &\qquad -(1-\alpha_0)\ln(1-\alpha_0) + \alpha_0\cdot o_k(1) + o_k(1/k) \\ &\leqslant\; \frac{c-2-4\ln k}{2k} \\ &\qquad +\alpha_0\left[1-\ln\left(\frac{2k\alpha_0}{1+4\ln k}\right) + o_k(1)\right] + o_k(1/k).\end{aligned} \tag{45}$$

Elementary calculus shows that the function

$$\alpha_0 \in (0,1) \mapsto -\alpha_0(1 - \ln\frac{2k\alpha_0}{1+4\ln k} + o_k(1))$$

attains its maximum at $\alpha_0 = (1+o_k(1))\frac{1+4\ln k}{2k}$. Hence, (45) yields

$$\frac{1}{n}\ln(|\mathcal{S}_\alpha|\cdot \mathrm{P}\left[\mathcal{C}_\zeta\right]) \;\leqslant\; \frac{c-1+o_k(1)}{2k}. \tag{46}$$

To complete the proof, consider for any $\alpha \in A$ the number Σ_α of k-covers ζ of $G'(n,m)$ such that $|\zeta^{-1}(i)| = \alpha_i$ for all i. Then (46) implies that $\frac{1}{n}\ln \mathrm{E}[\Sigma_\alpha] \leqslant \frac{c-1}{2k} - o_k(1)$ for all $\alpha \in A$. Hence, there is $0 < \varepsilon_k = o_k(1)$ such that for $c < 1 - \varepsilon_k$ we have

$$\mathrm{E}[\Sigma_\alpha] \leqslant \exp\left[\frac{c-1}{2k} - o_k(1)\right] \leqslant \exp(-\varepsilon_k/2) = \exp(-\Omega(n)). \tag{47}$$

Since condition iii. ensures that $|A| \leqslant n^k = \exp(o(n))$, the assertion follows from (47) by taking the union bound over all $\alpha \in A$ and applying Lemma 3. □

6.2 Proof of Lemma 23

Given $\mathcal{I}_\zeta$, the pairs $\mathbf{e}_1, \ldots, \mathbf{e}_m$ that constitute the random graph $G'(n,m)$ are simply distributed uniformly and independently over the set of all $n^2(1-F_\zeta)$ possible pairs that do not join two vertices in the same class V_i for $i = 1, \ldots, k$. For each vertex v and each $j \in \{0, 1, \ldots, k\}$ let $d_{v,j}$ be the number of pairs $\mathbf{e}_i$ such that $\mathbf{e}_i$ contains v together with a vertex from V_j. Clearly, given $\mathcal{I}_\zeta$ we have $d_{v,j} = 0$ for all $v \in V_j$, $j \in [k]$.

Of course, the random variables $d_{v,j}$ are *not* independent because we know that the total number of edges is equal to m. But their dependence turns out to be "relatively mild". In fact, we are going to show that the $d_{v,j}$ can be approximated by a family of independent Poisson random variables up to an error term of $\exp(o(n))$, which is negligible for our purposes. To state this precisely, consider a family $(b_{vj})_{v \in V, j \in \{0,1,\ldots,k\}}$ of independent Poisson random variables with means

$$\mathrm{E}\left[b_{vj}\right] = \frac{\alpha_j d}{1-F_\zeta}.$$

Let $\mathcal{B}_\zeta$ be the event that

i. for any $v \in V_0$ there exist $i, j \in [k]$, $i \neq j$ such that $b_{vi} = 0$ and $b_{vj} \leqslant 1$ and

ii. for any $1 \leqslant i < j \leqslant k$ and any $v \in V_i$ we have $b_{vj} > 1$.

These two conditions mirror the conditions **CV3** and **CV2** from the definition of "cover". The key step in the proof (somewhat reminiscent of the Poisson cloning model [25]) is to establish the following.

Lemma 24. *We have* $\mathrm{P}\left[\mathcal{C}_\zeta | \mathcal{I}_\zeta\right] \leqslant \exp(o(n)) \cdot \mathrm{P}\left[\mathcal{B}_\zeta\right]$.

To prove Lemma 24 we consider a further event. Set $B_{ij} = \sum_{v \in V_i} b_{vj}$ for $i, j \in \{0, 1, \ldots, k\}$, $(i,j) \neq (0,0)$. Being sums of independent Poisson variables, the random variables B_{ij} are Poisson as well, with means

$$\mathrm{E}[B_{ij}] = \mathrm{E}\left[B_{ji}\right] = \alpha_i \alpha_j dn/(1-F_\zeta) \qquad (0 \leqslant i < j \leqslant k).$$

In addition, let B_{00} be a random variable that is independent of all of the above such that $\frac{1}{2}B_{00}$ has distribution $\mathrm{Po}(\alpha_0^2 m/(1-F_\zeta))$. (In particular, B_{00} takes even values only.) Now, let $\mathcal{V}$ be the event that

i. $B_{ij} = B_{ji}$ for all $i \neq j$ and

ii. $\frac{1}{2}B_{00} + \sum_{0 \leqslant i < j \leqslant k} B_{ij} = m$.

Lemma 25. *We have* $\mathrm{P}\left[\mathcal{V}\right] = \exp(o(n))$.

Proof. Since

$$\mathrm{E}\left[\frac{B_{00}}{2} + \sum_{1 \leqslant i < j \leqslant k} B_{ij}\right] = \frac{dn}{2(1-F(\zeta))}\left[\alpha_0^2 + \sum_{i \neq j} \alpha_i \alpha_j\right] = m,$$

there exist integers $\beta_{ij} = \mathrm{E}[B_{ij}] + O(1)$ such that $\beta_{ij} = \beta_{ji}$ and $\frac{1}{2}\beta_{00} + \sum_{0 \leqslant i < j \leqslant k} \beta_{ij} = m$. Clearly,

$$\mathrm{P}[\mathcal{V}] \geqslant \mathrm{P}[B_{ij} = \beta_{ij} \text{ for all } i, j] = \prod_{i,j} \mathrm{P}[B_{ij} = \beta_{ij}]. \tag{48}$$

Since $\beta_{ij} = \mathrm{E}[B_{ij}] + O(1)$ and B_{ij} is a Poisson variable, Stirling's formula yields

$$\mathrm{P}[B_{ij} = \beta_{ij}] = \Omega(n^{-1/2}).$$

Therefore, (48) implies $\mathrm{P}[\mathcal{V}] \geqslant \Omega(n^{-(k+1)^2/2}) = \exp(o(n))$, as claimed. □

Proof of Lemma 24. Let $\mathbf{m} = (m_{ij})_{i,j \in \{0,1,\ldots,k\}}$ be a family of non-negative integers such that

a. $m_{ij} = m_{ji}$ for all i, j,

b. $m_{ii} = 0$ for $i \in [k]$ and

c. $m_{00} + \sum_{0 \leqslant i < j \leqslant k} m_{ij} = m$.

Let $\mathcal{M}_{\mathbf{m}}$ be the event that

$$\sum_{v \in V_0} d_{v0} = 2m_{00} \quad \text{and} \quad \sum_{v \in V_i} d_{vj} = m_{ij} \text{ for any } 0 \leqslant i < j \leqslant k.$$

Analogously, let $\mathcal{M}'_{\mathbf{m}}$ be the event that

$$B_{00} = 2m_{00} \quad \text{and} \quad B_{ij} = m_{ij} \text{ for any } 0 \leqslant i < j \leqslant k.$$

We claim that for any $\mathbf{m}$ that satisfies a.–c. above we have

$$\mathrm{P}[\mathcal{C}_\zeta | \mathcal{M}_{\mathbf{m}}] = \mathrm{P}[\mathcal{B}_\zeta | \mathcal{M}'_{\mathbf{m}}]. \tag{49}$$

Indeed, let either $i = j = 0$ or $0 \leqslant i < j \leqslant k$. Given that $\mathcal{M}_{\mathbf{m}}$ occurs, we can think of the m_{ij} edges that join V_i and V_j as balls and of the vertices $v \in V_i$ as bins. Each ball is tossed into one of the bins randomly and independently, and these experiments are independent for all i, j. Thus, (49) simply follows from the Poissonization of the balls and bins experiment (Lemma 5).

To complete the proof, we need to compare $\mathrm{P}[\mathcal{M}_{\mathbf{m}} | \mathcal{I}_\zeta]$ and $\mathrm{P}[\mathcal{M}'_{\mathbf{m}} | \mathcal{V}]$. Because under the distribution $\mathrm{P}[\,\cdot\,| \mathcal{I}_\zeta]$ the pairs $\mathbf{e}_1, \ldots, \mathbf{e}_m$ are simply chosen randomly subject to the constraint that none of them joins two vertices in the same class V_i, $i \in [k]$, we see that

$$\mathrm{P}[\mathcal{M}_{\mathbf{m}} | \mathcal{I}_\zeta] = \frac{m!}{m_{00}! \prod_{0 \leqslant i < j \leqslant k} m_{ij}!} \cdot \left(\frac{\alpha_0^2}{1 - F_\zeta} \right)^{m_{00}} \prod_{0 \leqslant i < j \leqslant k} \left(\frac{2\alpha_i \alpha_j}{1 - F_\zeta} \right)^{m_{ij}}. \tag{50}$$

(The factor of 2 arises because $\mathbf{e}_1, \ldots, \mathbf{e}_m$ are *ordered* pairs.) Furthermore, because $\mathcal{V}$ provides that $B_{ij} = B_{ji}$ for all i, j, we have

$$\mathrm{P}[\mathcal{M}'_{\mathbf{m}} | \mathcal{V}] \;=\; \frac{\mathrm{P}[B_{00} = 2m_{00}] \cdot \prod_{0 \leqslant i < j \leqslant k} \mathrm{P}[B_{ij} = m_{ij}]}{\mathrm{P}[\mathcal{V}]}.$$

Thus, by Lemma 25

$$\mathrm{P}\left[\mathcal{M}'_{\mathbf{m}}|\mathcal{V}\right] = \exp(o(n)) \cdot \mathrm{P}\left[B_{00} = 2m_{00}\right] \cdot \prod_{0\leqslant i<j\leqslant k} \mathrm{P}\left[B_{ij} = m_{ij}\right]. \tag{51}$$

Since for $0 \leqslant i < j \leqslant k$ the random variables B_{ij} are Poisson with mean $\alpha_i\alpha_j dn/(1-F_\zeta)$, we have

$$\begin{aligned}\mathrm{P}\left[B_{ij} = m_{ij}\right] &= \frac{(\alpha_i\alpha_j dn/(1-F_\zeta))^{m_{ij}}}{m_{ij}!\exp(\alpha_i\alpha_j dn/(1-F_\zeta))} \\ &= \left(\frac{2\alpha_i\alpha_j}{1-F_\zeta}\right)^{m_{ij}} \frac{m^{m_{ij}}}{m_{ij}!\exp(2\alpha_i\alpha_j m/(1-F_\zeta))}.\end{aligned} \tag{52}$$

Similarly,

$$\begin{aligned}\mathrm{P}\left[B_{00} = 2m_{00}\right] &= \frac{(\alpha_0^2 m/(1-F_\zeta))^{m_{00}}}{m_{00}!\exp(\alpha_0^2 m/(1-F_\zeta))} \\ &= \left(\frac{\alpha_0^2}{1-F_\zeta}\right)^{m_{00}} \frac{m^{m_{00}}}{m_{00}!\exp(\alpha_0^2 m/(1-F_\zeta))}.\end{aligned} \tag{53}$$

Combining (50)–(53), we obtain from Stirling's formula

$$\begin{aligned}\frac{\mathrm{P}\left[\mathcal{M}_{\mathbf{m}}|\mathcal{I}_\zeta\right]}{\mathrm{P}\left[\mathcal{M}'_{\mathbf{m}}|\mathcal{V}\right]} &= \frac{m!\exp(m(1-F_\zeta)^{-1}(\alpha_0^2 + 2\sum_{0\leqslant i<j\leqslant k}\alpha_i\alpha_j))}{\exp(o(n))m^m} \\ &= \frac{m!\exp(m+o(n))}{m^m} = \exp(o(n)).\end{aligned} \tag{54}$$

Finally, combining (49) and (54) we conclude that for any $\mathbf{m}$ that satisfies a.–c. we have

$$\begin{aligned}\mathrm{P}\left[\mathcal{C}_\zeta \cap \mathcal{M}_{\mathbf{m}}|\mathcal{I}_\zeta\right] &= \mathrm{P}\left[\mathcal{C}_\zeta|\mathcal{M}_{\mathbf{m}}\right] \cdot \mathrm{P}\left[\mathcal{M}_{\mathbf{m}}|\mathcal{I}_\zeta\right] && [\text{as } \mathcal{M}_{\mathbf{m}} \subset \mathcal{I}_\zeta] \\ &= \exp(o(n))\,\mathrm{P}\left[\mathcal{B}_\zeta|\mathcal{M}'_{\mathbf{m}}\right] \cdot \mathrm{P}\left[\mathcal{M}'_{\mathbf{m}}|\mathcal{V}\right] \\ &= \exp(o(n))\,\mathrm{P}\left[\mathcal{B}_\zeta \cap \mathcal{M}'_{\mathbf{m}}\right] / \mathrm{P}\left[\mathcal{V}\right] \\ &= \exp(o(n))\,\mathrm{P}\left[\mathcal{B}_\zeta \cap \mathcal{M}'_{\mathbf{m}}\right] && [\text{due to Lemma 25}].\end{aligned}$$

Summing over all possible $\mathbf{m}$ completes the proof. □

Proof of Lemma 23. We are going to bound the probability of the event $\mathcal{B}_\zeta$. For $v \in V_0$ we have

$$\begin{aligned}\mathrm{P}\left[\exists 1 \leqslant i < j \leqslant k : b_{vi} = b_{vj} = 0\right] \leqslant &\sum_{1\leqslant i<j\leqslant k} \mathrm{P}\left[b_{vi} = b_{vj} = 0\right] \\ &+ \sum_{i,j\in[k]:i\neq j} \mathrm{P}\left[b_{vi} = 0,\, b_{vj} = 1\right] = p_0,\end{aligned}$$

because the b_{vi}, b_{vj} are independent Poisson variables. Similarly, if $v \in V_i$ for some $i \in [k]$, then

$$\mathrm{P}\left[\forall j \in [k] \setminus \{i\} : b_{vj} > 1\right] \;=\; \prod_{j \in [k] \setminus \{i\}} 1 - \mathrm{P}\left[b_{vj} \leqslant 1\right] = p_i.$$

Due to the mutual independence of the b_{vj}, we thus obtain $\mathrm{P}\left[\mathcal{B}_\zeta\right] = p_0^{\alpha_0 n} \prod_{i=1}^{k} p_i^{\alpha_i n}$. Finally, the assertion follows from Lemma 24. □

References

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

[2] D. Achlioptas, A. Coja-Oghlan: Algorithmic barriers from phase transitions. Proc. 49th FOCS (2008) 793–802.

[3] D. Achlioptas, M. Molloy: The analysis of a list-coloring algorithm on a random graph. Proc. 38th FOCS (1997) 204–212.

[4] D. Achlioptas, M. Molloy: Almost all graphs with 2.522n edges are not 3-colorable. Electronic Journal of Combinatorics **6** (1999).

[5] D. Achlioptas, C. Moore: Almost all graphs with average degree 4 are 3-colorable. Journal of Computer and System Sciences **67** (2003) 441–471.

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

[7] D. Achlioptas, F. Ricci-Tersenghi: On the solution-space geometry of random constraint satisfaction problems. Proc. 38th STOC (2006) 130–139.

[8] M. Bayati, D. Gamarnik, P. Tetali: Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. Proc. 42nd STOC (2010) 105–114.

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

[10] P. Cheeseman, B. Kanefsky, W. Taylor: Where the *really* hard problems are. Proc. IJCAI (1991) 331–337.

[11] A. Coja-Oghlan, K. Panagiotou, A. Steger: On the chromatic number of random graphs. Journal of Combinatorial Theory, Series B, **98** (2008) 980–993.

[12] A. Coja-Oghlan, D. Vilenchik: Chasing the k-colorability threshold. An extended abstract version is due to appear in the proceedings of FOCS 2013. arXiv:1304.1063 (2013).

[13] V. Dani, C. Moore, A. Olson: Tight bounds on the threshold for permuted k-colorability. Proc. 16th RANDOM (2012) 505–516.

[14] O. Dubois, J. Mandler: On the non-3-colourability of random graphs. arXiv:math/0209087 (2002).

[15] P. Dunne, M. Zito: An improved upper bound on the non-3-colourability threshold. Information Processing Letters **65** (1998) 17–23.

[16] R. Durrett: Probability and examples, 3rd edition Brooks/Cole, Belmont, CA, 2005.

[17] P. Erdős, A. Rényi: On the evolution of random graphs. Magayar Tud. Akad. Mat. Kutato Int. Kozl. **5** (1960) 17–61.

[18] N. Fountoulakis, C. McDiarmid: Upper bounds on the non-3-colourability threshold of random graphs. Discrete Mathematics and Theoretical Computer Science **5** (2002) 205–226.

[19] S. Franz, M. Leone: Replica bounds for optimization problems and diluted spin systems. J. Statist. Phys. **111** (2003) 535–564.

[20] I. Giotis, A. Kaporis, L. Kirousis: Corrigendum to "A note on the non-colorability threshold of a random graph. 2002.

[21] G. Grimmett, C. McDiarmid: On colouring random graphs. Mathematical Proceedings of the Cambridge Philosophical Society **77** (1975) 313–324

[22] F. Guerra: Broken replica symmetry bounds in the mean field spin glass model. Comm. Math. Phys. **233** (2003) 1–12.

[23] S. Janson, T. Łuczak, A. Ruciński: Random Graphs, Wiley 2000.

[24] A. Kaporis, L. Kirousis, Y. Stamatiou: A note on the non-colourability threshold of a random graph. Electronic Journal of Combinatorics **7** (2000) R29.

[25] J. Kim: Poisson cloning model for random graphs. Proc. ICM (2006) 873–897.

[26] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborova: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. National Academy of Sciences **104** (2007) 10318–10323.

[27] F. Krzakala, A. Pagnani, M. Weigt: Threshold values, stability analysis and high-q asymptotics for the coloring problem on random graphs. Phys. Rev. E **70** (2004) 046705.

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

[29] E. Maneva, A. Sinclair: On the satisfiability threshold and clustering of solutions of random 3-SAT formulas. Theoretical Computer Science **407** (2008) 359–369.

[30] D. Matula: Expose-and-merge exploration and the chromatic number of a random graph. Combinatorica **7** (1987) 275–284.

[31] M. Mézard, A. Montanari: Information, physics and computation. Oxford University Press 2009.

[32] M. Molloy: The freezing threshold for k-colourings of a random graph. Proc. 43rd STOC (2012) 921–930.

[33] R. Mulet, A. Pagnani, M. Weigt, R. Zecchina: Coloring random graphs. Phys. Rev. Lett. **89** (2002) 268701

[34] J. van Mourik, D. Saad: Random Graph Coloring - a Statistical Physics Approach. Phys. Rev. E **66** (2002) 056120.

[35] D. Panchenko, M. Talagrand: Bounds for diluted mean-fields spin glass models. Probab. Theory Relat. Fields **130** (2004) 319–336.

[36] G. Parisi: On local equilibrium equations for clustering states. arXiv:cs/0212047 (2002).

[37] L. Zdeborová: Statistical physics of hard optimization problems. PhD thesis, Charles University, 2008.

[38] L. Zdeborová, F. Krzakala: Phase transition in the coloring of random graphs. Phys. Rev. E **76** (2007) 031131.