The Condensation Phase Transition in Random Graph Coloring

Victor Bapst\(^1\), Amin Coja-Oghlan\(^1\), Samuel Hetterich\(^1\), Felicia Raßmann\(^1\), Dan Vilenchik\(^2\)

\(^1\) Mathematics Institute, Goethe University, 10 Robert Mayer St, 60325 Frankfurt, Germany. 
E-mail: bapst@math.uni-frankfurt.de; acoghlan@math.uni-frankfurt.de; hetterich@math.uni-frankfurt.de; rassmann@math.uni-frankfurt.de
\(^2\) Department of Communication Systems Engineering, Ben-Gurion University of the Negev, Beer-sheba, Israel. E-mail: vilenchi@post.bgu.ac.il

Received: 31 August 2014 / Accepted: 26 June 2015 
Published online: 5 October 2015 – © Springer-Verlag Berlin Heidelberg 2015

Abstract: Based on a non-rigorous formalism called the “cavity method”, physicists have put forward intriguing predictions on phase transitions in diluted mean-field models, in which the geometry of interactions is induced by a sparse random graph or hypergraph. One example of such a model is the graph coloring problem on the Erdős–Rényi random graph \(G(n, d/n)\), which can be viewed as the zero temperature case of the Potts antiferromagnet. The cavity method predicts that in addition to the \(k\)-colorability phase transition studied intensively in combinatorics, there exists a second phase transition called the condensation phase transition (Krzakala et al. in Proc Natl Acad Sci 104:10318–10323, 2007). In fact, there is a conjecture as to the precise location of this phase transition in terms of a certain distributional fixed point problem. In this paper we prove this conjecture for \(k\) exceeding a certain constant \(k_0\).

1. Introduction and Results

1.1. Background and motivation. Since the early 2000s physicists have developed a systematic but non-rigorous formalism called the cavity method for the study of diluted mean-field models [23]. These are models of disordered systems, such as glasses or spin glasses, where the geometry of interactions is given by a sparse random graph or hypergraph. Apart from cases of immediate physical interest, such as the diluted Potts or Ising model, the cavity method has been applied to long-standing problems in combinatorics and information theory, as well as, more recently, to problems in computer science and compressive sensing [19]. The predictions obtained in this way have a very significant potential impact on all of these areas. Hence the importance of providing a

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 No. 278857-PTCC.

An extended abstract version of this work appeared in the Proceedings of the 18th International Workshop on Randomization and Computation (‘RANDOM’) 2014.
rigorous mathematical foundation for the cavity method, an effort that the present work contributes to.

The specific model that we deal with is the Potts antiferromagnet on the Erdős–Rényi random graph at zero temperature, also known as the random graph coloring problem. This problem has played a central role in combinatorics since the seminal 1960 paper of Erdős and Rényi that started the theory of random graphs [14]. In this model, the geometry is defined by the random graph $G(n, d/n)$ on $n$ vertices $V = \{1, \ldots, n\}$, any two of which are connected by an edge with probability $d/n$ independently. For an integer $k \geq 3$ we call a map $\sigma : V \to \{1, \ldots, k\}$ a $k$-coloring of $G(n, d/n)$ if $\sigma(v) \neq \sigma(w)$ for any two vertices $v, w$ that are connected by an edge. Let $Z_k$ be the number of $k$-colorings of $G(n, d/n)$. How does the (appropriately scaled) partition function $Z_k$ vary as a function of the average degree $d$ of the random graph in the thermodynamic limit $n \to \infty$?

The cavity method predicts that for any fixed $k \geq 3$ there occur two phase transitions as $d$ increases [20,21,33]. The first of these is the condensation phase transition. This phase transition is ubiquitous in physics, and is believed to hold the key to a variety of problems. For instance, the role of condensation in the context of structural glasses is a major open problem, going back to the work of Kauzmann in the 1940s [17]. In the context of diluted mean-field models, the existence of the condensation phase transition has been proved in the hypergraph 2-coloring problem [9] and the Potts antiferromagnet (at positive temperature) [10]. But thus far its location has not been determined exactly in a rigorous way. The contribution of the present paper is to establish that in the random graph coloring problem, condensation occurs at the exact point predicted by the cavity method. This is the first rigorous result of this kind in a diluted mean-field model.

As most predictions based on the cavity method, the one on condensation comes in the form of a distributional fixed point problem. Apart from studying this fixed point analytically, the key contribution of the present work is to establish an explicit link between the fixed point problem and the combinatorics of the random graph coloring problem. We expect that the technique that we develop for this purpose generalises to a variety of other models. Immediate examples that spring to mind include the random hypergraph 2-coloring problem and the $k$-NAESAT problem, which is of interest in computer science.

The second conjectured phase transition is the $k$-colorability threshold. This is the point where the random graph $G(n, d/n)$ ceases to possess a $k$-coloring. Establishing the existence and location of the $k$-colorability threshold is a major open problem in combinatorics [1], and our main result implies a slightly improved lower bound on this conjectured threshold. However, the $k$-colorability phase transition is an artefact of the zero-temperature case: it is not expected to persist in the Potts antiferromagnet at positive temperature, in contrast to the condensation phase transition.

1.2. Pinning down the condensation phase transition. Letting $Z_k(G)$ be the number of $k$-colorings of a graph $G$, we consider

$$\Phi_k(d) \equiv \lim_{n \to \infty} \mathbb{E}[Z_k(G(n, d/n))^{1/n}].$$

In the case of the graph coloring problem, $\Phi_k(d)$ is the natural scaling of the partition function $Z_k(G(n, d/n)).$\footnote{We work with the $n$th root $Z_k(G(n, d/n))^{1/n}$ instead of $\frac{1}{n} \ln Z_k(G(n, d/n))$ because $Z_k(G(n, d/n))$ may be take the value 0.}
According to physics conventions, a “phase transition” would be a point \( d_0 \) where the function \( d \mapsto \Phi_k(d) \) is non-analytic. However, the limit \( \Phi_k(d) \) is not currently known to exist for all \( d, k \).\(^2\) Hence, we need to tread carefully: for a fixed \( k \geq 3 \) we call \( d_0 \in (0, \infty) \) smooth if there exists \( \varepsilon > 0 \) such that

- for any \( d \in (d_0 - \varepsilon, d_0 + \varepsilon) \) the limit \( \Phi_k(d) \) exists, and
- the map \( d \in (d_0 - \varepsilon, d_0 + \varepsilon) \mapsto \Phi_k(d) \) has an expansion as an absolutely convergent power series around \( d_0 \).

If \( d_0 \) fails to be smooth, we say that a phase transition occurs at \( d_0 \).

For a smooth \( d_0 \) the sequence of random variables \( (Z_k(G(n, d_0/n))^{1/n})_n \) converges to \( \Phi_k(d_0) \) in probability. This follows from a concentration result for the number of \( k \)-colorings from \([2]\).

As a next step, we state (an equivalent but slightly streamlined version of) the physics prediction from \([33]\) as to the location of the condensation phase transition. As most predictions based on the “cavity method”, this one comes in terms of a distributional fixed point problem. To be specific, let \( \Omega \) be the set of probability measures on the set \([k] = \{1, \ldots, k\} \). We identify \( \Omega \) with the \( k \)-simplex, i.e., the set of maps \( \mu : [k] \to [0, 1] \) such that \( \sum_{h=1}^k \mu(h) = 1 \), equipped with the topology and Borel algebra induced by \( \mathbb{R}^k \). Moreover, we define a map \( B : \bigcup_{\gamma=1}^{\infty} \Omega^\gamma \to \Omega, (\mu_1, \ldots, \mu_\gamma) \mapsto B[\mu_1, \ldots, \mu_\gamma] \) by letting

\[
B[\mu_1, \ldots, \mu_\gamma](i) = \begin{cases} 1/k & \text{if } \sum_{h \in [k]} \prod_{j=1}^{\gamma} 1 - \mu_j(h) = 0, \\ \prod_{j=1}^{\gamma} 1 - \mu_j(h) & \text{for any } i \in [k]. \end{cases}
\]

Further, let \( \mathcal{P} \) be the set of all probability measures on \( \Omega \). For each \( \mu \in \Omega \) let \( \delta_\mu \in \mathcal{P} \) denote the Dirac measure that puts mass one on the single point \( \mu \). In particular, \( \delta_{k^{-1}1} \in \mathcal{P} \) signifies the measure that puts mass one on the uniform distribution \( k^{-1}1 = (1/k, \ldots, 1/k) \). For \( \pi \in \mathcal{P} \) and \( \gamma \geq 0 \) let

\[
Z_\gamma(\pi) = \sum_{h=1}^k \left( 1 - \int_\Omega \mu(h) d\pi(\mu) \right)^\gamma.
\]

Further, define a map \( F_{d,k} : \mathcal{P} \to \mathcal{P}, \pi \mapsto F_{d,k}[\pi] \) by letting

\[
F_{d,k}[\pi] = \exp(-d)\delta_{k^{-1}1} + \sum_{\gamma=1}^{\infty} \frac{\gamma! Z_\gamma(\pi)}{d^\gamma!} \int_{\Omega^\gamma} \left[ \sum_{h=1}^k \prod_{j=1}^\gamma 1 - \mu_j(h) \right] \delta_{B[\mu_1, \ldots, \mu_\gamma]} \prod_{j=1}^\gamma d\pi(\mu_j).
\]

\(^2\) It seems natural to conjecture that the limit \( \Phi_k(d) \) exists for all \( d, k \), but proving this might be difficult. In fact, the existence of the limit for all \( d, k \) would imply that \( d_{k,\text{col}}(n) \) converges, which is a major open problem in the theory of random graphs \([1]\).
The function $\phi_{d,k}(\pi) = \phi_{d,k}^\kappa(\pi) + \frac{1}{k} \sum_{i \in [k]} \sum_{\gamma_1, \ldots, \gamma_k = 0}^\infty \phi_{d,k}^\kappa(\pi; i; \gamma_1, \ldots, \gamma_k) \prod_{h \in [k]} \left( \frac{d}{k-1} \right)^{\gamma_h} \frac{\exp(-d/(k-1))}{\gamma_h!}$, where

$$\phi_{d,k}^\kappa(\pi; i; \gamma_1, \ldots, \gamma_k) = -\frac{d}{2k(k-1)} \sum_{h_1=1}^k \sum_{h_2 \in [k] \setminus \{h_1\}} \int_{\Omega^2} \ln \left[ 1 - \sum_{h \in [k]} \mu_1(h)\mu_2(h) \right] \frac{2}{\gamma_{h_1}} d\pi_{h_1}(\mu_1).$$

Fig. 1. The function $\phi_{d,k}$

Thus, in (1.3) we integrate a function with values in $\mathcal{P}$, viewed as a subset of the Banach space $^3$ of signed measures on $\Omega$. The normalising term $Z_{\gamma}(\pi)$ ensures that $F_{d,k}[\pi]$ really is a probability measure on $\Omega$.

The main theorem is in terms of a fixed point of the map $F_{d,k}$, i.e., a point $\pi^* \in \mathcal{P}$ such that $F_{d,k}[\pi^*] = \pi^*$. In general, the map $F_{d,k}$ has several fixed points. Hence, we need to single out the correct one. For $h \in [k]$ let $\delta_h \in \Omega$ denote the vector whose $h$th coordinate is one and whose other coordinates are 0 (i.e., the Dirac measure on $h$). We call a measure $\pi \in \mathcal{P}$ frozen if $\pi(\{\delta_1, \ldots, \delta_k\}) \geq \frac{2}{3}$; in words, the total probability mass concentrated on the $k$ vertices of the simplex $\Omega$ is at least $2/3$.

As a final ingredient, we need a function $\phi_{d,k} : \mathcal{P} \rightarrow \mathbb{R}$. To streamline the notation, for $\pi \in \mathcal{P}$ and $h \in [k]$ we write $\pi_h$ for the measure $d\pi_h(\mu) = k\mu(h)d\pi(\mu)$. With this notation, $\phi_{d,k}$ is defined in Fig. 1. The integrals in (1.4) and (1.5) are well-defined because the set where the argument of the logarithm vanishes has measure zero.

**Theorem 1.1.** There exists a constant $k_0 \geq 3$ such that for any $k \geq k_0$ the following holds. If $d \geq (2k - 1) \ln k - 2$, then $F_{d,k}$ has precisely one frozen fixed point $\pi^*_{d,k}$. Further, the function

$$\Sigma_k : d \mapsto \ln k + \frac{d}{2} \ln(1 - 1/k) - \phi_{d,k}(\pi^*_{d,k})$$

has a unique zero $d_{k,\text{cond}}$ in the interval $[(2k - 1) \ln k - 2, (2k - 1) \ln k - 1]$. For this number $d_{k,\text{cond}}$ the following three statements are true.

(i) Any $0 < d < d_{k,\text{cond}}$ is smooth and $\Phi_k(d) = k(1 - 1/k)^{d/2}$.

(ii) There occurs a phase transition at $d_{k,\text{cond}}$.

---

$^3$ To be completely explicit, the probability mass that a measurable set $A \subset \Omega$ carries under $F_{d,k}[\pi]$ is

$$F_{d,k}[\pi](A) = \exp(-d) \cdot 1_{\pi \in A}$$

$$+ \sum_{\gamma \geq 1} \frac{d^\gamma}{\gamma!} \cdot Z_\gamma(\pi) \left( \int \sum_{h=1}^k \prod_{j=1}^\gamma 1 - \mu_j(h) \right) \cdot 1_{B[\mu_1, \ldots, \mu_\gamma] \in A} \prod_{j=1}^\gamma d\tau(\mu_j),$$

where $1_{v \in A} = 1$ if $v \in A$ and $1_{v \in A} = 0$ otherwise.
(iii) If \( d > d_{k, \text{cond}} \), then
\[
\limsup_{n \to \infty} \mathbb{E}[Z_k(G(n, d/n))^{1/n}] < k(1 - 1/k)^{d/2}.
\]
Thus, if \( d \) is smooth, then \( \Phi_k(d) < k(1 - 1/k)^{d/2} \).

The key strength of Theorem 1.1 and the main achievement of this work is that we identify the precise location of the phase transition. Admittedly, this precise answer is not exactly a simple one. But that seems unsurprising, given the intricate combinatorics of the random graph coloring problem. That said, the proof of Theorem 1.1 will illuminate matters. For instance, the fixed point \( \pi^*_{d,k} \) turns out to have a nice combinatorial interpretation and, perhaps surprisingly, \( \pi^*_{d,k} \) emerges to be a discrete probability distribution.

The above formulas are derived systematically via the cavity method [23]. For example, the functional \( \phi_{d,k} \) is an installment of a generic formula, the so-called “Bethe free entropy”. Moreover, the map \( B \) is the distributional version of the Belief Propagation operator. In effect, the predictions as to the condensation phase transitions in other problems look very similar to the above. Consequently, it seems reasonable to expect that the proof technique developed in the present work carries over to many other problems.

While the main point of Theorem 1.1 is that it gives an exact answer, it is not difficult to obtain a simple asymptotic expansion of \( d_{k, \text{cond}} \) in the limit of large \( k \). Namely, \( d_{k, \text{cond}} = (2k - 1) \ln k - 2 \ln 2 + \varepsilon_k \), where \( \varepsilon_k \to 0 \) as \( k \to \infty \). This asymptotic formula was obtained in [8] by means of a much simpler argument than the one developed in the present paper. However, this simpler argument does not quite get to the bottom of the condensation phenomenon.

As per common practice, we say than an event occurs asymptotically almost surely (\( \text{a.a.s.} \)) if the probability of the event converges to one as \( n \to \infty \). We observe that the first part of Theorem 1.1 implies that \( G(n, d/n) \) has a \( k \)-coloring a.a.s. for any \( 0 < d < d_{k, \text{cond}} \). Indeed, if \( d < d_{k, \text{cond}} \), then \( \Phi_k(d) = k(1 - 1/k)^{d/2} > 0 \) and thus \( Z_k(G(n, d/n)) > 0 \) a.a.s. because \( (Z_k(G(n, d/n))^{1/n}) \) converges to \( \Phi_k(d) \) in probability.

1.3. The “entropy crisis”. Apart from predicting the location of the condensation phase transition, the cavity method also offers up a combinatorial interpretation of how this phase transition comes about, sometimes referred to as the “entropy crisis” in physics [20]. More specifically, the physics arguments predict that there is a sequence \( (\varepsilon_k)_{k>0} \) such that for \( d > (1 + \varepsilon_k) \ln k \), about a factor of two below \( d_{k, \text{cond}} \), the set of \( k \)-colorings decomposes into a large number of well-separated “clusters” a.a.s. The emergence of this decomposition goes by the name of “dynamic replica symmetry breaking”, and has been established rigorously [2,25]. In contrast to condensation, dynamic replica symmetry breaking does not mark an actual phase transition.

Now, as the average degree \( d \) increases to \( d_{k, \text{cond}} \), both the sizes of the individual clusters and the overall number \( Z_k(G(n, d/n)) \) decrease. But \( Z_k(G(n, d/n)) \) drops at a faster rate, and \( d_{k, \text{cond}} \) is expected to mark the point where the size of the largest cluster coincides with \( Z_k(G(n, d/n)) \) (up to a sub-exponential factor).

The proof of Theorem 1.1 allows us to verify this prediction. Of course, we need to formalise what we mean by “clusters” first. Thus, let \( G \) be a graph on \( n \) vertices. If \( \sigma, \tau \) are \( k \)-colorings of \( G \), we define their overlap as the \( k \times k \)-matrix \( \rho(\sigma, \tau) = (\rho_{i,j}(\sigma, \tau))_{i,j \in [k]} \) with entries

\[
\rho_{i,j}(\sigma, \tau) = \frac{1}{n} \sum_{v \in V(G)} \mathbb{1}(\sigma(v) = i, \tau(v) = j),
\]

where \( \mathbb{1}(P) \) is 1 if \( P \) is true and 0 otherwise.
Corollary 1.1. With the notation and assumptions of Theorem 1.1, the function $\Sigma_k$ is continuous, strictly positive and monotonically decreasing on $((2k-1)\ln k - 2, d_{k,\text{cond}})$, and $\lim_{d \to d_{k,\text{cond}}} \Sigma_k(d) = 0$. Further, given that $Z_k(G(n, d/n)) > 0$, let $\tau$ be a uniformly random $k$-coloring of this random graph. Then for any $d \in ((2k-1)\ln k - 2, d_{k,\text{cond}})$,

$$
\lim_{\varepsilon \to 0, n \to \infty} \mathbb{P} \left[ \frac{1}{n} \ln \frac{|C(G(n, d/n), \tau)|}{Z_k(G(n, d/n))} \leq \Sigma_k(d) + \varepsilon \mid Z_k(G(n, d/n)) > 0 \right] = 1,
$$

and

$$
\lim_{\varepsilon \to 0, n \to \infty} \mathbb{P} \left[ \frac{1}{n} \ln \frac{|C(G(n, d/n), \tau)|}{Z_k(G(n, d/n))} \geq \Sigma_k(d) - \varepsilon \mid Z_k(G(n, d/n)) > 0 \right] > 0.
$$

We emphasise that our conditioning on $Z_k(G(n, d/n)) > 0$ is necessary to speak of a random $k$-coloring $\tau$ but otherwise harmless as Theorem 1.1 implies that $G(n, d/n)$ is $k$-colorable with probability tending to one as $n \to \infty$ for any $d < d_{k,\text{cond}}$.

In other words, Corollary 1.1 shows that there is a certain function $\Sigma_k > 0$ such that the total number of $k$-colorings exceeds the number of $k$-colorings in the cluster of a randomly chosen $k$-coloring by at least a factor of $\exp[n(\Sigma_k(d) + o(1))]$ with probability tending to one. However, as $d$ approaches $d_{k,\text{cond}}$, $\Sigma_k(d)$ tends to 0, and with a non-vanishing probability the gap between the total number of $k$-colorings and the size of a single cluster is upper-bounded by $\exp[n(\Sigma_k(d) + o(1))]$.

1.4. Discussion and related work. In this section we discuss some relevant related work and also explain the impact of Theorem 1.1 on various questions that have come up in the literature.

1.4.1. The physics perspective. The original physics motivation for the study of systems in which the interactions are induced by a probability distribution was the study of materials with peculiar magnetic properties, the so-called spin glasses. A wide variety of models have been put forward, ranging from models on lattices to the well-known Sherrington–Kirkpatrick model [28], whose free energy is captured by the “Parisi formula” [26, 29]. The Sherrington–Kirkpatrick model is a fully-connected mean-field model, i.e., each variable interacts with any other (via randomly chosen couplings). By comparison, in diluted mean-field models the interactions are determined by the edges of a sparse random graph or hypergraph, rather than by a complete graph. These models thus possess a
non-trivial geometry (as opposed to fully-connected models where every pair of vertices interacts in the same way), while having only a bounded number of short cycles a.a.s. (as opposed to a finite-dimensional lattice). This makes diluted mean-field models amenable to analytic albeit non-rigorous study via the “cavity method” [23].

The condensation phase transition (which is sometimes also referred to as “static one-step replica symmetry breaking transition” or “Kauzmann transition”) has been established in a variety of models, ranging from the random energy model [12], the fully-connected $p$ spin-glass [18, 30] as well as disordered polymers on trees [13].

With respect to diluted mean-field models, Coja-Oghlan and Zdeborová [9] showed that a condensation phase transition occurs in random $r$-uniform hypergraph 2-coloring. Furthermore, [9] determines the location of the condensation phase transition up to an error $\epsilon_r$ that tends to zero as the uniformity $r$ of the hypergraph becomes large. Moreover, Contucci, Dommers, Giardina, Starr proved that a condensation phase transition occurs in the diluted mean-field $k$-spin Potts antiferromagnet at positive temperature [10], and determined the value of $d$ where it starts to occur up to an additive error of about $\ln k$.

Yet the present work is the first to fully verify the prediction of the cavity method on condensation in a diluted mean-field model; the physics prediction was derived in [20, 33]. The core of the proof of Theorem 1.1 is to establish an explicit link between the combinatorics of the graph coloring problem and the cavity formalism. In effect, in our analysis of the distributional fixed point problem we can directly incorporate some of the physics calculations from [33, Appendix C].

Finally, the problem of coloring random graphs algorithmically has received quite a bit of interest in computer science (e.g., [15]). The cavity method has inspired new “message passing” algorithms for this problem by the name of Belief/Survey Propagation Guided Decimation [6, 24]. Experiments on random graph $k$-coloring instances for small values of $k$ indicate an excellent performance of these algorithms [6, 32, 33]. While a rigorous analysis remains elusive, the physics prediction is that the performance of Belief Propagation guided decimation hinges on the location of the “condensation line” in a two-dimensional phase diagram parametrised by $d$ and a value $t$ that measures the progress of the algorithm [27]. In this notation, Theorem 1.1 identifies the location of the condensation point in the case $t = 0$. Thus, it would be interesting to extend the present techniques to $t \in (0, 1)$, and to turn this into a rigorous analysis of the algorithm.

1.4.2. The combinatorics perspective. Graph coloring is one of the most fundamental problems in combinatorics, as witnessed by the famous “four color problem”. Thus, it is unsurprising that the problem of coloring random graphs has attracted a great deal of attention since it was first posed by Erdős and Rényi [14]; see [16] for a comprehensive overview. In the case that $p = d/n$ for a fixed real $d > 0$, it is known that there exists a sharp threshold sequence $d_{k-col}(n)$ such that for any fixed $\varepsilon > 0$, the random graph $G(n, d(n)/n)$ has a $k$-coloring a.a.s. if $d(n) < (1 - \varepsilon)d_{k-col}(n)$, and fails to have a $k$-coloring a.a.s. if $d(n) > (1 + \varepsilon)d_{k-col}(n)$ [1]. It is widely conjectured but as of yet unproven that the sequence $d_{k-col}(n)$ converges to a limit $d_{k-col}$ as $n \to \infty$. If so, then $d_{k-col}$ would mark a second phase transition in the random graph coloring problem (as $\Phi_k(d) = 0$ for all $d > d_{k-col}$, while $\Phi_k(d) > 0$ for all $d < d_{k-col}$).

The best current bounds on the threshold sequence $d_{k-col}(n)$ are

$$
(2k - 1) \ln k - 2 \ln 2 + \varepsilon_k \leq \lim \inf_{n \to \infty} d_{k-col}(n) \\
\leq \lim \sup_{n \to \infty} d_{k-col}(n) \leq (2k - 1) \ln k - 1 + \delta_k,
$$

(1.7)
where $\varepsilon_k, \delta_k \to 0$ as $k \to \infty$. The upper bound is by the first moment method [7]. The lower bound rests on a second moment argument [8], which improves a landmark result of Achlioptas and Naor [4].

While Theorem 1.1 allows for the possibility that $d_{k,\text{cond}}$ is equal to the $k$-colorability threshold $d_{k,\text{-col}}$ (if it exists), the physics prediction is that these two are different. More specifically, the cavity method yields a prediction as to the precise value of $d_{k,\text{-col}}$ in terms of another distributional fixed point problem. An asymptotic expansion in terms of $k$ leads to the conjecture $d_{k,\text{-col}} = (2k - 1) \ln k - 1 + \eta_k$ with $\eta_k \to 0$ as $k \to \infty$. Thus, the upper bound in (1.7) is conjectured to be asymptotically tight in the limit $k \to \infty$.

The present work builds upon the second moment argument from [8]. Conversely, Theorem 1.1 yields a small improvement over the lower bound in (1.7). Indeed, as we saw above Theorem 1.1 implies that $\liminf_{n \to \infty} d_{k,\text{-col}}(n) \geq d_{k,\text{cond}}$, thereby determining the precise “error term” $\varepsilon_k$ in the lower bound in (1.7). In fact, $d_{k,\text{cond}}$ is the best-possible lower bound that can be obtained via the kind of second moment argument developed in [4,8]. This is because a necessary condition for the success of the second moment argument is that $\Phi_k(d) = k(1 - 1/k)^{d/2}$.

The proofs in this paper build upon some of the techniques that have been developed to study the “geometry” of the set of $k$-colorings of the random graph, and add to this machinery. Among the techniques that we harness is the “planting trick” from [2] (which, in a sense, we are going to “put into reverse”), the notion of a core [2,8,25], techniques for proving the existence of “frozen variables” (or “hard fields” in physics jargon) [25], and a concentration argument from [9]. That said, the cornerstone of the present work is a novel argument that allows us to connect the distributional fixed point problem from [33] rigorously with the geometry of the set of $k$-colorings.

### 1.5. Preliminaries and notation.
Throughout the paper we tacitly assume that $k \geq k_0$ for some large enough constant $k_0$ that is large enough for the various estimates to hold. We also implicitly assume that $n$ is sufficiently large. We use the standard $O$-notation when referring to the limit $n \to \infty$. Thus, $f(n) = O(g(n))$ means that there exist $C > 0, n_0 > 0$ such that for all $n > n_0$ we have $|f(n)| \leq C \cdot |g(n)|$. In addition, we use the standard symbols $o(\cdot), \Omega(\cdot), \Theta(\cdot)$. In particular, $o(1)$ stands for a term that tends to 0 as $n \to \infty$.

Additionally, we use asymptotic notation with respect to the limit of large $k$. To make this explicit, we insert $k$ as an index. Thus, $f(k) = O_k(g(k))$ means that there exist $C > 0, k_0 > 0$ such that for all $k > k_0$ we have $|f(k)| \leq C \cdot |g(k)|$. Further, we write $f(k) = \tilde{O_k}(g(k))$ to indicate that there exist $C > 0, k_0 > 0$ such that for all $k > k_0$ we have $|f(k)| \leq (\ln k)^C \cdot |g(k)|$.

If $L$ is an integer, then we let $[L] = \{1, \ldots, L\}$. Finally, we always set $m = \lceil dn/2 \rceil$ and we let $G(n, m)$ denote a random graph with vertex set $V = [n] = \{1, \ldots, n\}$ and with precisely $m$ edges chosen uniformly at random.

### 2. Outline
The proof of Theorem 1.1 is composed of two parallel threads. The first thread is to identify an “obvious” point where a phase transition occurs or, more specifically, a critical degree $d_{k,\text{crit}}$, where statements (i)–(iii) of the theorem are met. The second thread is to identify the frozen fixed point $\pi^*_{d,k}$ of $\mathcal{F}_{d,k}$ and to interpret it combinatorially. Finally, the two threads intertwine to show that $d_{k,\text{crit}} = d_{k,\text{cond}}$, i.e. that the “obvious” phase transition $d_{k,\text{crit}}$ is indeed the unique zero of Eq. (1.6). The first thread is an extension of ideas developed in [9] for random hypergraph 2-coloring to the (technically more
involved) random graph coloring problem. The second thread and the intertwining of the two require novel arguments.

2.1. The first thread. With the two require novel arguments.

\[ \Phi_k(d) = \lim_{n \to \infty} \mathbb{E}[Z_k(G(n, d/n))]^{1/n} \]

is difficult to calculate for general values of \( d \). However for \( d \in [0, 1) \), \( \Phi_k(d) \) is easily understood. In fact, for \( d \in [0, 1) \) the random graph \( G(n, d/n) \) decomposes into tree components and a bounded number of connected components with precisely one cycle a.a.s. \([14]\). Moreover, the number of \( k \)-colorings of a tree with \( v \) vertices and \( \mu \) edges is well-known to be \( k^v (1 - 1/k)^\mu \). Since \( G(n, d/n) \) has \( m \sim dn/2 \) edges a.a.s., we obtain

\[ Z_k(G(n, d/n))^{1/n} \sim k(1 - 1/k)^{d/2} \quad \text{for } d < 1. \tag{2.1} \]

As \( Z_k(G)^{1/n} \leq k \) for any graph on \( n \) vertices, (2.1) implies that

\[ \Phi_k(d) = \lim_{n \to \infty} \mathbb{E}[Z_k(G(n, d/n))]^{1/n} = k(1 - 1/k)^{d/2} \quad \text{for } d < 1. \tag{2.2} \]

Since \( d \mapsto k(1 - 1/k)^{d/2} \) is analytic, the least \( d > 0 \) where the limit \( \Phi_k(d) \) either fails to exist or strays away from \( k(1 - 1/k)^{d/2} \) is going to be a phase transition. Hence, we let

\[ d_{k, \text{crit}} = \sup \left\{ d \geq 0 : \text{the limit } \Phi_k(d) \text{ exists and } \Phi_k(d) = k(1 - 1/k)^{d/2} \right\}. \]

**Fact 2.1.** We have \( d_{k, \text{crit}} \leq (2k - 1) \ln k \).

**Proof.** The upper bound (1.7) on the \( k \)-colorability threshold implies that \( Z_k(G(n, d/n)) = 0 \) a.a.s. for \( d > (2k - 1) \ln k \). By contrast, \( k(1 - 1/k)^{d/2} > 0 \) for any \( d > 0 \).

Thus, \( d_{k, \text{crit}} \) is a well-defined finite number, and there occurs a phase transition at \( d_{k, \text{crit}} \). Moreover, the following proposition, which we prove in Sect. 3, yields a lower bound on \( d_{k, \text{crit}} \) and implies that \( d_{k, \text{crit}} \) satisfies the first condition in Theorem 1.1.

**Proposition 2.1.** For any \( d > 0 \) we have

\[ \limsup_{n \to \infty} \mathbb{E}[Z_k(G(n, d/n))]^{1/n} \leq k(1 - 1/k)^{d/2}. \]

Moreover,

\[ d_{k, \text{crit}} = \sup \left\{ d \geq 0 : \liminf_{n \to \infty} \mathbb{E}[Z_k(G(n, d/n))]^{1/n} \geq k(1 - 1/k)^{d/2} \right\} \]

\[ \geq (2k - 1) \ln k - 2. \]

Thus, we know that there exists a number \( d_{k, \text{crit}} \) that satisfies conditions (i)–(ii) in Theorem 1.1. Of course, to actually calculate this number we need to unearth its combinatorial “meaning”. As we saw in Sect. 1.3, if \( d_{k, \text{crit}} \) really is the condensation phase transition, then the combinatorial interpretation should be as follows. For \( d < d_{k, \text{crit}} \), the size of the cluster that a randomly chosen \( k \)-coloring \( \tau \) belongs to is smaller than \( Z_k(G(n, d/n)) \) by an exponential factor \( \exp(\Omega(n)) \) a.a.s. But as \( d \) approaches \( d_{k, \text{crit}} \), the gap between the cluster size and \( Z_k(G(n, d/n)) \) diminishes. Hence, \( d_{k, \text{crit}} \) should mark the point where the cluster size has the same order of magnitude as \( Z_k(G(n, d/n)) \).
But how can we possibly get a handle on the size of the cluster that a randomly chosen \( k \)-coloring \( \tau \) of \( G(n, d/n) \) belongs to? No “constructive” method is known for obtaining a single \( k \)-coloring of \( G(n, d/n) \) for \( d \) anywhere close to \( d_{k,\text{col}} \), let alone for sampling one uniformly at random. Nevertheless, as observed in [2], in the case that \( \Phi_k(d) = k(1 - 1/k)^{d/2} \), i.e., for \( d < d_{k,\text{crit}} \), it is possible to capture the experiment of first choosing the random graph \( G(n, d/n) \) and then sampling a \( k \)-coloring \( \tau \) uniformly at random by means of a different, much more innocent experiment.

In this latter experiment, we first choose a map \( \sigma : [n] \to [k] \) uniformly at random. Then, we generate a graph \( G(n, p', \sigma) \) on \([n]\) by connecting any two vertices \( v, w \in [n] \) such that \( \sigma(v) \neq \sigma(w) \) with probability \( p' \) independently. If \( p' = dk/(n(k - 1)) \) is chosen so that the expected number of edges is the same as in \( G(n, d/n) \) and if \( \Phi_k(d) = k(1 - 1/k)^{d/2} \), then this so-called planted model should be a good approximation to the “difficult” experiment of first choosing \( G(n, d/n) \) and then picking a random \( k \)-coloring \( \tau \). In particular, with respect to the cluster size we expect that

\[
\mathbb{E}[|C(G(n, p', \sigma), \sigma)|^{1/n}] \sim \mathbb{E}[|C(G(n, d/n), \tau)|^{1/n}],
\]

i.e., that the suitably scaled cluster size in the planted model is about the same as the cluster size in \( G(n, d/n) \). Hence, \( d_{k,\text{crit}} \) should mark the point where \( \mathbb{E}[|C(G(n, p', \sigma), \sigma)|^{1/n}] \) equals \( k(1 - 1/k)^{d/2} \). The following proposition verifies that this is indeed so.

Let us write \( G = G(n, p', \sigma) \) for the sake of brevity.

**Proposition 2.2.** Assume that \( (2k - 1) \ln k - 2 \leq d \leq (2k - 1) \ln k \) and set

\[
p' = d'/n \quad \text{with} \quad d' = \frac{dk}{k - 1}. \tag{2.3}
\]

1. If

\[
\lim_{\varepsilon \searrow 0} \liminf_{n \to \infty} \mathbb{P}
\left[
|C(G, \sigma)|^{1/n} \leq k(1 - 1/k)^{d/2} - \varepsilon
\right] = 1, \tag{2.4}
\]

then \( d \leq d_{k,\text{crit}} \).

2. Conversely, if

\[
\lim_{\varepsilon \searrow 0} \liminf_{n \to \infty} \mathbb{P}
\left[
|C(G, \sigma)|^{1/n} \geq k(1 - 1/k)^{d/2} + \varepsilon
\right] = 1, \tag{2.5}
\]

then \( \limsup_{n \to \infty} \mathbb{E}[Z_k(G(n, d/n))^{1/n}] < k(1 - 1/k)^{d/2} \). In particular, \( d \geq d_{k,\text{crit}} \).

We prove Proposition 2.2 in Sect. 4.

2.2. The second thread. Our next aim is to “solve” the fixed point problem for \( F_{d,k} \) to an extent that gives the fixed point an explicit combinatorial interpretation. This combinatorial interpretation is in terms of a certain random tree process, associated with a concept of “legal colorings”. Specifically, we consider a multi-type Galton–Watson branching process. Its set of types is

\[
\mathcal{T} = \{(i, \ell) : i \in [k], \ell \subset [k], i \in \ell\}.
\]
The intuition is that \( i \) is a “distinguished color” and that \( \ell \) is a set of “available colors”. The branching process is further parameterized by a vector \( q = (q_1, \ldots, q_k) \in [0, 1]^k \) such that \( q_1 + \cdots + q_k \leq 1 \). Let \( d' = dk/(k - 1) \) and

\[
q_{i, \ell} = \frac{1}{k} \prod_{j \in \ell \setminus \{i\}} \exp(-q_j d') \prod_{j \in [k] \setminus \ell} 1 - \exp(-q_j d') \quad \text{for } (i, \ell) \in \mathcal{T}.
\]

Then

\[
\sum_{(i, \ell) \in \mathcal{T}} q_{i, \ell} = 1.
\]

Further, for each \((i, \ell) \in \mathcal{T}\) such that \(|\ell| > 1\) we define \( \mathcal{T}_{i, \ell} \) as the set of all \((i', \ell') \in \mathcal{T}\) such that \( \ell \cap \ell' \neq \emptyset \) and \(|\ell'| > 1\). In addition, for \((i, \ell) \in \mathcal{T}\) such that \(|\ell| = 1\) we set \( \mathcal{T}_{i, \ell} = \emptyset \).

The branching process \( GW(d, k, q) \) starts with a single individual, whose type \((i, \ell) \in \mathcal{T}\) is chosen from the probability distribution \((q_{i, \ell})_{(i, \ell) \in \mathcal{T}}\). In the course of the process, each individual of type \((i, \ell) \in \mathcal{T}\) spawns a Poisson number \( Po(d' q_{i, \ell}) \) of offspring of type \((i', \ell')\) for each \((i', \ell') \in \mathcal{T}_{i, \ell}\). In particular, only the initial individual may have a type \((i, \ell)\) with \(|\ell| = 1\), in which case it does not have any offspring. Let \( 1 \leq N \leq \infty \) be the progeny of the process (i.e., the total number of individuals created).

We are going to view \( GW(d, k, q) \) as a distribution over trees endowed with some extra information. Let us define a decorated graph as a graph \( T = (V, E) \) together with a map \( \vartheta : V \to \mathcal{T} \) such that for each edge \( e = \{v, w\} \in E \) we have \( \vartheta(w) \in \mathcal{T}_{\vartheta(v)} \). Moreover, a rooted decorated graph is a decorated graph \((T, \vartheta)\) together with a distinguished vertex \( v_0 \), the root. Further, an isomorphism between two rooted decorated graphs \( T \) and \( T' \) is an isomorphism of the underlying graphs that preserves the root and the types of the vertices.

Given that \( N < \infty \), the branching process \( GW(d, k, q) \) canonically induces a probability distribution over isomorphism classes of rooted decorated trees. Indeed, we obtain a tree whose vertices are all the individuals created in the course of the branching process and where there is an edge between each individual and its offspring. The individual from which the process starts is the root. Moreover, by construction each individual \( v \) comes with a type \( \vartheta(v) \). We denote the (random) isomorphism class of this tree by \( T_{d,k,q} \). (It is most natural to view the branching process as a probability distribution over isomorphism classes as the process does not specify the order in which offspring is created.)

To proceed, we define a legal coloring of a decorated graph \((G, \vartheta)\) as a map \( \tau : V(G) \to [k] \) such that \( \tau \) is a \( k \)-coloring of \( G \) and such that for any type \((i, \ell) \in \mathcal{T}\) and for any vertex \( v \) with \( \vartheta(v) = (i, \ell) \) we have \( \tau(v) \in \ell \). Let \( Z(G, \vartheta) \) denote the number of legal colorings.

Since \( Z(G, \vartheta) \) is isomorphism-invariant, we obtain the integer-valued random variable \( Z(T_{d,k,q}) \). We have \( Z(T_{d,k,q}) \geq 1 \) with certainty because a legal coloring \( \tau \) can be constructed by coloring each vertex with its distinguished color (i.e., setting \( \tau(v) = i \) if \( v \) has type \((i, \ell)\)). Hence, \( \ln Z(T_{d,k,q}) \) is a well-defined non-negative random variable. Additionally, we write \( |T_{d,k,q}| \) for the number of vertices in \( T_{d,k,q} \).

Finally, consider a rooted, decorated tree \((T, \vartheta, v_0)\) and let \( \tau \) be a legal coloring of \((T, \vartheta, v_0)\) chosen uniformly at random. Then the color \( \tau(v_0) \) of the root is a random variable with values in \([k]\). Let \( \mu_{T, \vartheta, v_0} \in \Omega \) denote its distribution. Clearly, \( \mu_{T, \vartheta, v_0} \) is invariant under isomorphisms. Consequently, the distribution \( \mu_{T_{d,k,q}} \) of the color of
the root of a tree in the random isomorphism class $T_{d,k,q}$ is a well-defined $\Omega$-valued random variable. Let $\pi_{d,k,q} \in \mathcal{P}$ denote its distribution. Then we can characterise the frozen fixed point of $F_{d,k}$ as follows.

**Proposition 2.3.** Suppose that $d \geq (2k - 1) \ln k - 2$.

1. The function

$$q \in [0, 1] \mapsto (1 - \exp(-dq/(k - 1)))^{k-1} \quad (2.6)$$

has a unique fixed point $q^*$ in the interval $[2/3, 1]$. Moreover, with

$$q^* = k^{-1}(q^*, \ldots, q^*) \in [0, 1]^k \quad (2.7)$$

the branching process $GW(d, k, q^*)$ is sub-critical. Thus, $\mathbb{P}[\mathcal{N} < \infty] = 1$.

2. The map $F_{d,k}$ has precisely one frozen fixed point, namely $\pi_{d,k,q^*}$.

3. We have $\phi_{d,k}(\pi_{d,k,q^*}) = \mathbb{E} \left[ \ln Z(T_{d,k,q^*}) \right]$.

4. The function $\Sigma_k$ from (1.6) is strictly decreasing and continuous on $[(2k - 1) \ln k - 2, (2k - 1) \ln k - 1]$ and has a unique zero $d_{k,\text{cond}}$ in this interval.

The function (2.6) and its fixed point explicitly occur in the physics work [33]. The proof of Proposition 2.3 can be found in Sect. 5.

2.3. Tying up the threads. To prove that $d_{k,\text{cond}} = d_{k,\text{crit}}$, we establish a connection between the random tree $T_{d,k,q^*}$ and the random graph $G$ with planted coloring $\sigma$. We start by giving a recipe for computing the cluster size $|C(G, \sigma)|$, and then show that the random tree process “cooks” it.

Computing the cluster size hinges on a close understanding of its combinatorial structure. As hypothesised in physics work [23] and established rigorously in [2,7,25], typically many vertices $v$ are “frozen” in $C(G, \sigma)$, i.e., $\tau(v) = \tau'(v)$ for any two colorings $\tau, \tau' \in C(G, \sigma)$. More generally, we consider for each vertex $v$ the set

$$\ell(v) = \{\tau(v) : \tau \in C(G, \sigma)\}$$

of colors that $v$ may take in colorings $\tau$ that belong to the cluster. Together with the “planted” color $\sigma(v)$, we can thus assign each vertex $v$ a type $\vartheta(v) = (\sigma(v), \ell(v))$. This turns $G$ into a decorated graph $(G, \vartheta)$.

By construction, each coloring $\tau \in C(G, \sigma)$ is a legal coloring of the decorated graph $G$. Conversely, we will see that a.a.s. any legal coloring of $(G, \vartheta)$ belongs to the cluster $C(G, \sigma)$. Hence, computing the cluster size $|C(G, \sigma)|$ amounts to calculating the number $Z(G, \vartheta)$ of legal colorings of $(G, \vartheta)$.

This calculation is facilitated by the following observation. Let $\tilde{G}$ be the graph obtained from $G$ by deleting all edges $e = \{v, w\}$ that join two vertices such that $\ell(v) \cap \ell(w) = \emptyset$. Then any legal coloring $\tau$ of $\tilde{G}$ is a legal coloring of $G$, because $\tau(v) \in \ell(v)$ for any vertex $v$. Hence, $Z(G, \vartheta) = Z(\tilde{G}, \vartheta)$.

Thus, we just need to compute $Z(\tilde{G}, \vartheta)$. This task is much easier than computing $Z(G, \vartheta)$ directly because $\tilde{G}$ turns out to have significantly fewer edges than $G$ a.a.s. More precisely, a.a.s. $\tilde{G}$ (mostly) consists of connected components that are trees of bounded size. In fact, we shall see that in an appropriate sense the distribution of the tree components converges to that of the decorated random tree $T_{d,k,q^*}$. In effect, we obtain
Proposition 2.4. Suppose that \( d \geq (2k - 1)\ln k - 2 \) and let \( p' \) be as in (2.3). Let \( q^* \) be as in (2.7). Then the sequence \( \{1/n \ln |\mathcal{C}(G, \sigma)|\}_n \) converges to \( \mathbb{E}\left[\frac{\ln Z(\mathcal{T}_{d,k,q^*})}{|\mathcal{T}_{d,k,q^*}|}\right] \) in probability.

The proof of Proposition 2.4, which can be found in Sect. 6, is based on the precise analysis of a further, combinatorial fixed point problem called Warning Propagation.

Proof of Theorem 1.1. Combining Propositions 2.2 and 2.4, we see that \( d_{k,\text{cond}} \) is equal to \( d_{k,\text{crit}} \), which is well-defined by Proposition 2.3. Further, (2.2) implies that \( d_{k,\text{crit}} > 0 \). Assume for contradiction that \( d_{k,\text{crit}} \) is smooth. Then there is \( \varepsilon > 0 \) such that the limit \( \Phi_k(d) \) exists for all \( d \in (d_{k,\text{crit}} - \varepsilon, d_{k,\text{crit}} + \varepsilon) \) and such that the function \( d \mapsto \Phi_k(d) \) is given by an absolutely convergent power series on this interval. Moreover, Proposition 2.1 implies that \( \Phi_k(d) = k(1 - 1/k)^{d/2} \) for all \( d \in (d_{k,\text{crit}} - \varepsilon, d_{k,\text{crit}}) \). Consequently, the uniqueness of analytic continuations implies that \( \Phi_k(d) = k(1 - 1/k)^{d/2} \) for all \( d \in (d_{k,\text{crit}} - \varepsilon, d_{k,\text{crit}} + \varepsilon) \), in contradiction to the definition of \( d_{k,\text{crit}} \). Thus, \( d_{k,\text{crit}} \) is a phase transition.

2.4. Proof of Corollary 1.1. Corollary 1.1 follows rather easily from the above and the following Lemma that establishes a connection between the planted model and the Boltzmann distribution on \( G(n, d/n) \). As in Corollary 1.1, we let \( \tau \) denote a random \( k \)-coloring of \( G(n, d/n) \).

Lemma 2.1 [3]. Assume that \( d < d_{k,\text{cond}} \). Let \( \mathcal{E} \) be a set of pairs \((G, \sigma)\), where \( G \) is a graph and \( \sigma \) is a \( k \)-coloring of \( G \). Further, given that \( Z_k(G(n, d/n)) > 0 \), let \( \tau \) be a uniformly random \( k \)-coloring of \( G(n, d/n) \). Then \( \mathbb{P}[((G, \sigma) \in \mathcal{E})] = o(1) \) implies that \( \mathbb{P}[(G(n, d/n), \tau) \in \mathcal{E}|Z_k(G(n, d/n)) > 0] = o(1) \).

Proof of Corollary 1.1. The statements about the properties of the function \( \Sigma_k \) follow readily from Proposition 2.3. Now, assume that \( d \in ((2k - 1)\ln k - 2, d_{k,\text{cond}}) \). Propositions 2.3 and 2.4 show that \( 1/n \ln |\mathcal{C}(G, \sigma)| \) converges to \( \phi_{d,k}(\pi_{d,k,q^*}) \) in probability. Hence, Markov’s inequality shows that for any \( \varepsilon > 0 \),

\[
\mathbb{P}\left[\frac{1}{n} \ln |\mathcal{C}(G, \sigma)| > \phi_{d,k}(\pi_{d,k,q^*}) + \varepsilon\right] = o(1). \tag{2.8}
\]

In combination with Lemma 2.1, (2.8) entails that

\[
\mathbb{P}\left[\frac{1}{n} \ln |\mathcal{C}(G(n, d/n), \tau)| > \phi_{d,k}(\pi_{d,k,q^*}) + \varepsilon|Z_k(G(n, d/n)) > 0\right] = o(1). \tag{2.9}
\]

Further, Propositions 2.3 and 2.4 imply that for any fixed \( \varepsilon > 0 \),

\[
\mathbb{P}\left[\frac{1}{n} \ln |\mathcal{C}(G, \sigma)| < \phi_{d,k}(\pi_{d,k,q^*}) - \varepsilon\right] = o(1).
\]

Hence, Lemma 2.1 yields

\[
\mathbb{P}\left[\frac{1}{n} \ln |\mathcal{C}(G(n, d/n), \tau)| \geq \phi_{d,k}(\pi_{d,k,q^*}) - \varepsilon\right] = \Omega(1). \tag{2.10}
\]

Thus, Corollary 1.1 follows from (2.9), (2.10) and the fact that \( Z_k(G(n, d/n))^{1/n} \) converges to \( \Phi_k(d) = k(1 - 1/k)^{d/2} \) in probability.
3. Groundwork: The First and the Second Moment Method

In this section we prove Proposition 2.1 and also lay the foundations for the proof of Proposition 2.2.

3.1. The first moment upper bound. We start by deriving an upper bound on $\Phi_k(d)$ by computing the expected number of $k$-colorings. To avoid fluctuations of the total number of edges, we work with the $G(n, m)$ model.

**Lemma 3.1.** We have $E[Z_k(G(n, m))] = \Theta(k^n (1 - 1/k)^m)$.

Lemma 3.1 is folklore. We carry the proof out regardless to make a few observations that will be important later. For a map $\sigma: [n] \to [k]$ let

$$\text{Forb}(\sigma) = k \sum_{i=1}^{k} \left( |\sigma^{-1}(i)| \right)$$

be the number of “forbidden pairs” of vertices that are colored the same under $\sigma$. By convexity,

$$N - \text{Forb}(\sigma) \geq \left( 1 - \frac{1}{k} \right) N,$$

with $N = \binom{n}{2}$. (3.1)

Hence, using Stirling’s formula, we find

$$P[\sigma \text{ is a } k\text{-coloring of } G(n, m)] = \frac{\binom{N - \text{Forb}(\sigma)}{m}}{\binom{N}{m}} \leq O((1 - 1/k)^{m})$$. (3.2)

As there are $k^n$ possible maps $\sigma$ in total, the linearity of expectation and (3.2) imply

$$E[Z_k(G(n, m))] = O(k^n (1 - 1/k)^m).$$

To bound $E[Z_k(G(n, m))]$ from below, call $\sigma: [n] \to [k]$ balanced if $|\sigma^{-1}(i) - \frac{n}{k}| \leq \sqrt{n}$ for all $i \in [k]$. Let $\text{Bal} = \text{Bal}_{n,k}$ be the set of all balanced $\sigma: [n] \to [k]$. For $\sigma \in \text{Bal}$ we verify easily that $N - \text{Forb}(\sigma) = (1 - 1/k)N + O(n)$. Thus, (3.2) and Stirling’s formula yield

$$P[\sigma \text{ is a } k\text{-coloring of } G(n, m)] = \Omega((1 - 1/k)^m) \quad \text{for any } \sigma \in \text{Bal}. \quad (3.3)$$

As $|\text{Bal}| = \Omega(k^n)$ by Stirling, the linearity of expectation and (3.3) imply $E[Z_k(G(n, m))] = \Omega(k^n (1 - 1/k)^m)$, whence Lemma 3.1 follows.

Letting $Z_{k,\text{bal}}$ denote the number of balanced $k$-colorings, we obtain from the above argument

**Corollary 3.1.** For any $d \geq 0$ we have $E[Z_{k,\text{bal}}(G(n, m))] = \Theta(k^n (1 - 1/k)^m)$.

As a further consequence of Lemma 3.1, we obtain

**Corollary 3.2.** For any $c > 0$ we have

$$\limsup_{n \to \infty} E[Z_k(G(n, c/n))^{1/n}] \leq k(1 - 1/k)^{c/2}.$$
Proof. Lemma 3.1 and Jensen’s inequality yield
\[ E[Z_k(G(n, m))^{1/n}] \leq E[Z_k(G(n, m))]^{1/n} \leq k(1 - 1/k)^{d/2} + o(1) \] (3.4)
Now, let \( c > 0 \) and set \( d = c - \varepsilon \) for some \( \varepsilon > 0 \). The number of edges in \( G(n, c/n) \) is binomially distributed with mean \((1 + o(1))cn/2 = m + \Omega(n)\). Hence, by the Chernoff bound the probability of the event \( A \) that \( G(n, c/n) \) has at least \( m \) edges tends to 1 as \( n \to \infty \). Because adding further edges can only decrease the number of \( k \)-colorings and since the number of \( k \)-colorings is trivially bounded by \( k^n \), we obtain from (3.4) that
\[ E[Z_k(G(n, c/n))^{1/n}] \leq E[Z_k(G(n, c/n))]^{1/n} \cdot 1_A + P[A \text{ does not occur}] \cdot k \leq E[Z_k(G(n, m))]^{1/n} + o(1) \leq k(1 - 1/k)^{d/2} + o(1) \]
Consequently, \( \limsup E[Z_k(G(n, c/n))^{1/n}] \leq k(1 - 1/k)^{d/2} \). This holds for any \( d > c \). Hence, letting \( \varepsilon = d - c \to 0 \), we see that
\[ \limsup E[Z_k(G(n, c/n))^{1/n}] \leq k(1 - 1/k)^{c/2} \]
as desired.

3.2. The second moment lower bound. The main technical step in the article [8] that yields the lower bound (1.7) on \( d_{k\text{-col}} \) is a second moment argument for a random variable \( Z_{k,\text{tame}} \) related to the number of \( k \)-colorings. We are going to employ this second moment estimate to bound \( Z_k(G(n, d/n)) \) from below.

The random variable \( Z_{k,\text{tame}} \) counts \( k \)-colorings with some additional properties. Suppose that \( \sigma \) is a balanced \( k \)-coloring of a graph \( G \) on \( V = [n] \). We call \( \sigma \) separable if for any balanced \( \tau \in \mathcal{C}(G, \sigma) \) and any \( i \in [k] \) we have
\[ \rho_{ii}(\sigma, \tau) \geq (1 - \kappa)/k, \quad \text{where } \kappa = \ln^{20} k/k. \]
Thus, if \( \sigma \) is a balanced, separable \( k \)-coloring, then for any color \( i \) and for any other balanced \( k \)-coloring \( \tau \) in the cluster of \( \sigma \), a \( 1 - \kappa + o(1) \)-fraction of the vertices colored \( i \) under \( \sigma \) are colored \( i \) under \( \tau \) as well. In particular, the clusters of any two such colorings are either disjoint or identical.

Definition 3.1. Let \( G \) be a graph with \( n \) vertices and \( m \) edges. A \( k \)-coloring \( \sigma \) of \( G \) is tame if
T1 \( \sigma \) is balanced,
T2 \( \sigma \) is separable, and
T3 \( |\mathcal{C}(G, \sigma) \cap \text{Bal}| \leq k^n(1 - 1/k)^m \).

Let \( Z_{k,\text{tame}}(G) \) denote the number of tame \( k \)-colorings of \( G \).

Lemma 3.2 [8]. Assume that \( d > 0 \) is such that
\[ \liminf_{n \to \infty} \frac{E[Z_{k,\text{tame}}(G(n, m))]}{k^n(1 - 1/k)^m} > 0. \] (3.5)
Then
\[ \liminf_{n \to \infty} \frac{E[Z_{k,\text{tame}}(G(n, m))^2]}{E[Z_{k,\text{tame}}(G(n, m))]^2} > 0. \]
Furthermore, there exists \( \varepsilon_k = o_k(1) \) such that (3.5) is satisfied if \( d \leq (2k - 1) \ln k - 2 \ln 2 - \varepsilon_k \).
As fleshed out in [8], together with the sharp threshold result from [1], Lemma 3.2 implies that \( G(n, d/n) \) is \( k \)-colorable a.a.s. if \( d \leq (2k - 1) \ln k - 2 \ln 2 - \varepsilon_k \). Here we are going to combine Lemma 3.2 with the following variant of that sharp threshold result to obtain a lower bound on the number of \( k \)-colorings.

**Lemma 3.3** [2]. For any \( k \geq 3 \) and for any real \( \varepsilon > 0 \) there is a sequence \( d_{k, \varepsilon}(n) \) such that for any \( \varepsilon > 0 \) the following holds.

1. If \( p(n) < (1 - \varepsilon)d_{k, \varepsilon}(n)/n \), then \( Z_k(G(n, p(n))) \geq \varepsilon^n \) a.a.s.
2. If \( p(n) > (1 + \varepsilon)d_{k, \varepsilon}(n)/n \), then \( Z_k(G(n, p(n))) < \varepsilon^n \) a.a.s.

Lemmas 3.2 and 3.3 entail the following lower bound on \( d_{k, \text{crit}} \).

**Lemma 3.4.** Assume that \( d^* > 0 \) and \( \varepsilon > 0 \) are such that (3.5) holds for any \( d \in (d^* - \varepsilon, d^*) \). Then \( d_{k, \text{crit}} \geq d^* \).

**Proof.** Assume for contradiction that \( d^* \) is such that (3.5) holds for all \( d \in (d^* - \varepsilon, d^*) \) but \( d_{k, \text{crit}} < d^* \). Pick and fix a number \( \max\{d^* - \varepsilon, d_{k, \text{crit}}\} < d_* < d^* \).

Corollary 3.2 implies that \( \limsup \mathbb{E}[Z_k(G(n, d_*/n))^{1/n}] \leq k(1 - 1/k)d_*/2 \). Therefore, since \( d_* > d_{k, \text{crit}} \), there exists \( \varepsilon_* > 0 \) such that

\[
\liminf_{n \to \infty} \mathbb{E}[Z_k(G(n, d_*/n))^{1/n}] < k(1 - 1/k)d_*/2 - \varepsilon_*.
\]  

(3.6)

Further, pick and fix \( d_* < \hat{d} < d^* \) such that \( k(1 - 1/k)\hat{d}/2 > k(1 - 1/k)d_*/2 - \varepsilon_* \) and \( \xi \) such that

\[
k(1 - 1/k)d_*/2 - \varepsilon_* < \xi < k(1 - 1/k)\hat{d}/2.
\]  

(3.7)

We are going to use Lemmas 3.2 and 3.3 to establish a lower bound on \( Z_k(G(n, d_*/n)) \) that contradicts (3.6). By the Paley–Zygmund inequality and because (3.5) holds for any \( d^* - \varepsilon < d < d^* \),

\[
\mathbb{P}\left[Z_{k, \text{tame}}(G(n, m)) \geq \frac{1}{2} \mathbb{E}[Z_{k, \text{tame}}(G(n, m))]\right] \geq \frac{\mathbb{E}[Z_{k, \text{tame}}(G(n, m))]^2}{4 \cdot \mathbb{E}[Z_{k, \text{tame}}(G(n, m))]^2}
\]  

for any \( d^* - \varepsilon < d < d^* \).

(3.8)

Moreover, Lemma 3.2 and (3.8) imply

\[
\liminf_{n \to \infty} \mathbb{P}\left[Z_{k, \text{tame}}(G(n, m)) \geq \frac{1}{2} \mathbb{E}[Z_{k, \text{tame}}(G(n, m))]\right] > 0
\]  

for any \( d^* - \varepsilon < d < d^* \).

(3.9)

Further, because (3.5) is true for any \( d^* - \varepsilon < d < d^* \) and \( \xi < k(1 - 1/k)\hat{d}/2 \) for any \( d < \hat{d} < d^* \), we see that

\[
\frac{1}{2} \mathbb{E}[Z_{k, \text{tame}}(G(n, m))] = \Omega(k^n(1 - 1/k)^m) > \xi^n \text{ for any } d < \hat{d}.
\]
Hence, (3.9) implies
\[
\liminf_{n \to \infty} P\left[ Z_{k,\text{tame}}(G(n, m)) \geq \xi^n \right] > 0 \quad \text{for any } d < \hat{d}.
\] (3.10)

Since the number of edges in \( G(n, d/n) \) has a binomial distribution with mean \( m \), with probability at least 1/3 the number of edges in \( G(n, d/n) \) does not exceed \( m \). Therefore, (3.10) implies that
\[
\liminf_{n \to \infty} P\left[ Z_k(G(n, d/n)) \geq \xi^n \right] \geq \frac{1}{3} \liminf_{n \to \infty} P\left[ Z_{k,\text{tame}}(G(n, m)) \geq \xi^n \right] > 0
\]
for any \( d < \hat{d} \).
\] (3.11)

Moreover, (3.11) entails that the sequence \( d_{k,\xi}(n) \) from Lemma 3.3 satisfies \( \liminf d_{k,\xi}(n) \geq \hat{d} \). Therefore,
\[
\liminf_{n \to \infty} P\left[ Z_k(G(n, d/n)) \geq \xi^n \right] = 1 \quad \text{for any } d < \hat{d}.
\] (3.12)

Since \( d_* < \hat{d} \), (3.12) entails that
\[
\liminf_{n \to \infty} \mathbb{E}\left[ Z_{k,\text{tame}}(G(n, d_*/n))^{1/n} \right] \geq \xi.
\] (3.13)

Combining (3.6), (3.7) and (3.13) yields a contradiction, which refutes our assumption that \( d_{k,\text{crit}} < d^* \).

3.3. Proof of Proposition 2.1. We start with the following observation.

**Lemma 3.5.** Let
\[
D_* = \left\{ d > 0 : \liminf \mathbb{E}[Z_k(G(n, d/n))^{1/n}] < k(1 - 1/k)d^{d/2} \right\},
\]
\[
D^* = \left\{ d > 0 : \limsup \mathbb{E}[Z_k(G(n, d/n))^{1/n}] < k(1 - 1/k)d^{d/2} \right\}.
\]

If \( d_1 \in D_* \) and \( d_2 > d_1 \), then \( d_2 \in D_* \). Similarly, if \( d_1 \in D^* \) and \( d_2 > d_1 \), then \( d_2 \in D^* \).

**Proof.** Let \( 0 < d_1 < d_2 \) and let \( q \sim (d_2 - d_1)/n \) be such that \( d_1/n + (1 - d_1/n)q = d_2/n \). Let us denote the random graph \( G(n, d_1/n) \) by \( G_1 \). Furthermore, let \( G_2 \) be a random graph obtained from \( G_1 \) by joining any two vertices that are not already adjacent in \( G_1 \) with probability \( q \) independently. Then \( G_2 \) is identical to \( G(n, d_2/n) \), because in \( G_2 \) any two vertices are adjacent with probability \( d_1/n + (1 - d_1/n)q = d_2/n \) independently. Set \( N = \binom{n}{2} \).

Let \( e(G_i) \) signify the number of edges in \( G_i \) for \( i = 1, 2 \). Because \( e(G_i) \) is a binomial random variable with mean \( \mu_i = d_i/n \cdot N = nd_i/2 + O(1) \), the Chernoff bound implies that
\[
P_{\mu_1,\mu_2}(G_1, G_2) = P\left[ |e(G_2) - e(G_1) - (\mu_2 - \mu_1)| > n^{2/3} \right] = o(1).
\] (3.14)
Further, since $Z_k^{1/n} \leq k$ with certainty, (3.14) implies that

$$
\mathbb{E}[Z_k(G_2)^{1/n} \mid Z_k(G_1)] = \mathbb{E}[Z_k(G_2)^{1/n} \mid Z_k(G_1), \mid e(G_2) - e(G_1) - (\mu_2 - \mu_1)\mid \\
\leq n^{2/3}[(1 - p_{\mu_1, \mu_2}(G_1, G_2)) + k \cdot p_{\mu_1, \mu_2}(G_1, G_2) \\
\leq \mathbb{E}[Z_k(G_2)^{1/n} \\
\cdot 1_{|e(G_2) - e(G_1) - (\mu_2 - \mu_1)| \leq n^{2/3} \mid Z_k(G_1)] + o(1). \quad (3.15)
$$

Suppose that we condition on $e(G_1), e(G_2)$ and $|e(G_1) - \mu_1| \leq n^{2/3}, |e(G_2) - e(G_1) - (\mu_2 - \mu_1)| \leq n^{2/3}$. Assume that $\sigma$ is a $k$-coloring of $G_1$. What is the probability that $\sigma$ remains a $k$-coloring of $G_2$? For this to happen, none of the $e(G_2) - e(G_1)$ additional edges must be among the Forb$(\sigma)$ pairs of vertices with the same color under $\sigma$. Using Stirling’s formula, we see that the probability of $\sigma$ remaining a $k$-coloring in $G_2$ is bounded by

$$
\gamma = \left(\frac{N - \text{Forb}(\sigma) - e(G_1)}{e(G_2) - e(G_1)}\right) / \left(\frac{N - e(G_1)}{e(G_2) - e(G_1)}\right) \leq (1 - 1/k)^{(d_2-d_1)+o(1)n^2/2}.
$$

Hence, by (3.15), Jensen’s inequality and (3.16)

$$
\mathbb{E}[Z_k(G_2)^{1/n} \mid Z_k(G_1)] \leq \mathbb{E}\left[Z_k(G_2) \cdot 1_{|e(G_2) - e(G_1) - (\mu_2 - \mu_1)| \leq n^{2/3} \mid Z_k(G_1)\right]^{1/n} + o(1) \\
\leq \gamma^{1/n}Z_k(G_1)^{1/n} + o(1) \\
\leq (1 - 1/k)^{(d_2-d_1)/2}Z_k(G_1)^{1/n} + o(1). \quad (3.17)
$$

Averaging (3.17) over $G_1$, we obtain

$$
\mathbb{E}[Z_k(G(n, d_2/n))^{1/n}] = \mathbb{E}[Z_k(G_2)^{1/n}] \\
\leq (1 - 1/k)^{(d_2-d_1)/2}\mathbb{E}[Z_k(G_1)^{1/n} \cdot 1_{|e(G_1) - \mu_1| \leq n^{2/3}}] + k \cdot \mathbb{P}[1_{|e(G_1) - \mu_1| \leq n^{2/3}}] + o(1) \\
\leq (1 - 1/k)^{(d_2-d_1)/2}\mathbb{E}[Z_k(G(n, d_1/n))^{1/n}] + o(1) \quad \text{[due to (3.14)].}
$$

Thus, if $\mathbb{E}[Z_k(G(n, d_1/n))^{1/n}] < k(1 - 1/k)^{d_1/2} - \delta + o(1)$, then

$$
\mathbb{E}[Z_k(G(n, d_2/n))^{1/n}] \leq k(1 - 1/k)^{d_2/2} - \epsilon + o(1)
$$

for some $\epsilon = \epsilon(\delta, k, d_1, d_2) > 0$. Taking $n \to \infty$ yields the assertion.

**Proof of Proposition 2.1.** The first assertion follows from Corollary 3.2 which additionally implies that

$$
d_{k, \text{crit}} = \sup \left\{ d \geq 0 : \liminf_{n \to \infty} \mathbb{E}[Z_k(G(n, d/n))^{1/n}] \geq k(1 - 1/k)^{d/2}\right\}.
$$

Hence, the second assertion is immediate from Lemma 3.5. The third assertion follows from Lemmas 3.2 and 3.4.
4. The Planted Model

4.1. Overview. The aim in this section is to prove Proposition 2.2. The proof of the first part is fairly straightforward. More precisely, in Sect. 4.2 we are going to establish

Lemma 4.1. Assume that \((2k - 1) \ln k - 2 \leq d \leq (2k - 1) \ln k\) is such that (2.4) holds. Then \(d_{k, \text{crit}} \geq d\).

The more challenging claim is that \(d \geq d_{k, \text{crit}}\) if typically the cluster in the planted model is “too big”. To prove this, we consider a variant of the planted model in which the number of edges is fixed. More precisely, for a map \(\sigma : [n] \to [k]\) we let \(G(n, m, \sigma)\) denote a graph on the vertex set \(V = [n]\) with precisely \(m\) edges that do not join vertices \(v, w\) with \(\sigma(v) = \sigma(w)\) chosen uniformly at random. In other words, \(G(n, m, \sigma)\) is just the random graph \(G(n, m)\) conditioned on the event that \(\sigma\) is a \(k\)-coloring. The following lemma, which is a variant of the “planting trick” from [2], establishes a general relationship between \(G(n, m)\) and \(G(n, m, \sigma)\).

Lemma 4.2. Let \(d > 0\). Assume that there exists a sequence \((\mathcal{E}_n)_{n \geq 1}\) of events such that

\[
\lim_{n \to \infty} \mathbb{P}[G(n, m) \in \mathcal{E}_n] = 1 \quad \text{while} \quad \limsup_{n \to \infty} \mathbb{P}[G(n, m, \sigma) \in \mathcal{E}_n]^{1/n} < 1. \quad (4.1)
\]

Then for any \(c > d\) we have \(\limsup \mathbb{E}[\mathcal{Z}_k(G(n, c/n))^{1/n}] < k(1 - 1/k)^{c/2}\). In particular, \(d_{k, \text{crit}} \leq d\).

We prove Lemma 4.2 in Sect. 4.6. Hence, assuming that the typical cluster size in the planted model is “too big” a.a.s., we need to exhibit events \(\mathcal{E}_n\) such that (4.1) holds. An obvious choice seems to be

\[
\mathcal{E}_n(\varepsilon) = \left\{ \mathcal{Z}_k^{1/n} \leq k(1 - 1/k)^{d/2} + \varepsilon \right\}
\]

But (4.1) requires that the probability that \(\mathcal{E}_n\) occurs in \(G(n, m, \sigma)\) is exponentially small, and neither the cluster size nor \(\mathcal{Z}_k\) are known to be sufficiently concentrated to obtain such an exponentially small probability.

Therefore, we define the events \(\mathcal{E}_n\) by means of another random variable. For a graph \(G = (V, E)\) and a map \(\sigma : V \to [k]\) let \(\mathcal{H}_G(\sigma)\) be the number of edges \([v, w]\) of \(G\) such that \(\sigma(v) = \sigma(w)\). In words, \(\mathcal{H}_G(\sigma)\) is the number of edges of \(G\) that are monochromatic under \(\sigma\). Furthermore, given \(\beta > 0\) let

\[
\mathcal{Z}_{\beta, k}(G) = \sum_{\sigma : V \to [k]} \exp(-\beta \cdot \mathcal{H}_G(\sigma)),
\]

the partition function of the \(k\)-spin Potts antiferromagnet on \(G\) at inverse temperature \(\beta\).

For large \(\beta\) there is a stiff “penalty factor” of \(\exp(-\beta)\) for any monochromatic edge. Thus, we expect that \(\mathcal{Z}_{\beta, k}\) becomes a good proxy for \(\mathcal{Z}_k\) as \(\beta \to \infty\). At the same time, \(\ln \mathcal{Z}_{\beta, k}\) enjoys a Lipschitz property. Namely, suppose that we obtain a graph \(G'\) from \(G\) by either adding or removing a single edge. Then

\[
|\ln(\mathcal{Z}_{\beta, k}(G)) - \ln(\mathcal{Z}_{\beta, k}(G'))| \leq \beta. \quad (4.2)
\]

Due to this Lipschitz property, one can easily show that \(\ln \mathcal{Z}_{\beta, k}\) is tightly concentrated. More precisely, we have
Lemma 4.3. For any fixed $d > 0$, $\varepsilon > 0$ there is $\alpha > 0$ such that the following is true. Suppose that $(\sigma_n)_{n \geq 1}$ is a sequence of maps $[n] \to [k]$. Then for all large enough $n$,

$$
P\left[ |\ln(Z_{\beta,k}(G(n, p', \sigma_n))) - \mathbb{E}[\ln Z_{\beta,k}(G(n, p', \sigma_n))]| > \varepsilon n \right] \leq \exp(-\alpha n).$$

Proof. This is immediate from the Lipschitz property (4.2) and McDiarmid’s inequality [22, Theorem 3.8].

Furthermore, in Sect. 4.4 we show that Lemma 4.3 implies

Lemma 4.4. Assume that $d$ is such that (2.5) holds. Then there exist $z, \beta > 0$ such that

$$
\lim_{n \to \infty} P\left[ \frac{1}{n} \ln Z_{\beta,k}(G(n, m)) \leq z \right] = 1
$$

while

$$
\limsup_{n \to \infty} P\left[ \frac{1}{n} \ln Z_{\beta,k}(G(n, m, \sigma)) \leq z \right]^{1/n} < 1.
$$

Finally, Proposition 2.2 is immediate from Lemmas 4.1, 4.2 and 4.4.

4.2. Proof of Lemma 4.1. We use the following observation from [8].

Lemma 4.5 [8]. Suppose that $(2k-1) \ln k - 2 \leq d \leq (2k-1) \ln k$. Let $p'$ be as in (2.3). Then the planted coloring $\sigma$ is separable in $G(n, p', \sigma)$ a.a.s.

Proof of Lemma 4.1. If (2.4) holds, then there exists $\varepsilon > 0$ such that with $p'$ from (2.3) we have

$$
\lim_{n \to \infty} P\left[ |C(G(n, p', \sigma), \sigma)| \leq k^n (1 - 1/k)^m \exp(-\varepsilon n) \right] = 1. \tag{4.3}
$$

Pick a number $d^* > d$ such that with $m^* = [d^*n/2]$ we have

$$
k^n (1 - 1/k)^{m^*} \geq k^n (1 - 1/k)^m \exp(-\varepsilon n/2).
$$

We claim that if we choose $\sigma : [n] \to [k]$ uniformly at random and independently a random graph $G(n, m^*)$, then

$$
\liminf_{n \to \infty} P[\sigma \text{ is tame} | \sigma \text{ is a } k\text{-coloring of } G(n, m^*)] > 0. \tag{4.4}
$$

To see this, let $\mathcal{E}$ be the event that the random graph $G(n, p', \sigma)$ has no more than $m^*$ edges. Because the number of edges in $G(n, p', \sigma)$ is binomially distributed with mean $m - m^* = \Omega(n)$, the Chernoff bound implies that $P[\mathcal{E}] = 1 - o(1)$. Therefore, (4.3) implies

$$
\lim_{n \to \infty} P\left[ |C(G(n, p', \sigma), \sigma)| \leq k^n (1 - 1/k)^m \exp(-\varepsilon n) | \mathcal{E} \right] = 1. \tag{4.5}
$$

Further, set $d'' = kd^*/(k-1)$ and let $p'' = d''/n > p'$. Then we can think of $G(n, p'', \sigma)$ as being obtained from $G(n, p', \sigma)$ by adding further random edges. More precisely, let $\mathcal{A}$ be the event that $G(n, p'', \sigma)$ contains precisely $m^*$ edges and set

$$
p'_n = P\left[ |C(G(n, p'', \sigma), \sigma)| \leq k^n (1 - 1/k)^{m^*} | \mathcal{A} \right].$$
Since adding edges can only decrease the cluster size, (4.5) entails
\[
\lim_{n \to \infty} p_n' \geq \lim_{n \to \infty} \mathbb{P}\left[|\mathcal{C}(G(n, p', \sigma), \sigma)| \leq k^n(1 - 1/k)^m \exp(-\varepsilon n) | \mathcal{E}\right] = 1. \tag{4.6}
\]
Similarly, let \( p_n'' = \mathbb{P}\left[\sigma \text{ is separable in } G(n, p'', \sigma) | \mathcal{A}\right]. \) Then Lemma 4.5 implies
\[
\lim_{n \to \infty} p_n'' \geq \lim_{n \to \infty} \mathbb{P}\left[\sigma \text{ is separable in } G(n, p', \sigma) | \mathcal{E}\right] = 1.
\]
Further, consider \( p_n''' = \mathbb{P}[\sigma \text{ is balanced}]. \) Then by Stirling’s formula,
\[
\limsup_{n \to \infty} p_n''' > 0. \tag{4.7}
\]
Finally, let \( p_n = \mathbb{P}[\sigma \text{ is tame| } \sigma \text{ is a } k\text{-coloring of } G(n, m*)]. \) Given the event \( \mathcal{A}, \)
\( G(n, p'', \sigma) \) is just a uniformly random graph with \( m* \) edges in which \( \sigma \) is a \( k\)-coloring. Hence,
\[
p_n = \mathbb{P}[\sigma \text{ is tame| } \sigma \text{ is a } k\text{-coloring of } G(n, m*)] = \mathbb{P}[\sigma \text{ is a } k\text{-coloring of } G(n, m*)] \cdot p_n.
\]
As (4.6)–(4.7) yield \( \liminf_{n \to \infty} p_n > 0, \) we obtain (4.4).

The estimate (4.4) enables us to bound \( \mathbb{E}[Z_{k,\text{tame}}(G(n, m*))] \) from below. Indeed, by the linearity of expectation
\[
\mathbb{E}[Z_{k,\text{tame}}(G(n, m*))] = \sum_{\sigma:[n] \to [k]} \mathbb{P}[\sigma \text{ is a tame } k\text{-coloring of } G(n, m*)]
\]
\[
= k^n \cdot \mathbb{P}[\sigma \text{ is a tame } k\text{-coloring of } G(n, m*)]
\]
\[
= k^n \mathbb{P}[\sigma \text{ is a } k\text{-coloring of } G(n, m*)] \cdot p_n
\]
\[
= \mathbb{E}[Z_k(G(n, m*))] \cdot p_n.
\]
Thus, Lemma 3.1 and (4.4) yield
\[
\liminf_{n \to \infty} \frac{\mathbb{E}[Z_{k,\text{tame}}(G(n, m*))]}{k^n(1 - 1/k)^{m*}} > 0.
\]
As this holds for all \( d* \) in an interval \( (d + \eta, d + 2\eta) \) with \( \eta > 0, \) Lemma 3.4 implies that \( d_{k, \text{crit}} \geq d. \)

4.3. Proof of Lemma 4.2.

**Lemma 4.6.** Assume that \( d > 0 \) is such that \( \limsup \mathbb{E}[Z_k(G(n, m))^{1/n}] < k(1 - 1/k)^{d/2}. \) Then for any \( c > d \) we have \( \limsup \mathbb{E}[Z_k(G(n, c/n))^{1/n}] < k(1 - 1/k)^{c/2}. \)

**Proof.** Assume that \( d, \delta > 0 \) are such that \( \limsup \mathbb{E}[Z_k(G(n, m))^{1/n}] < k(1 - 1/k)^{d/2 - \delta}. \) We claim that
\[
d* \in D^* = \left\{ c > 0 : \limsup \mathbb{E}[Z_k(G(n, c/n))^{1/n}] < k(1 - 1/k)^{c/2} \right\}
\]
for any \( d^* > d. \tag{4.8} \)
Indeed, the number $e(G(n, d^*/n))$ of edges of $G(n, d^*/n)$ is binomially distributed with mean $(1 + o(1))d^*n/2$. Since $d, d^*$ are independent of $n$ and $d^* > d$, the Chernoff bound implies that

$$P\left[e(G(n, d^*/n)) \leq m\right] \leq \exp(-\Omega(n)). \quad (4.9)$$

Further, if we condition on the event that $m^* = e(G(n, d^*/n)) > m$, then we can think of $G(n, d^*/n)$ as follows: first, create a random graph $G(n, m)$; then, add another $m^* - m$ random edges. Since the addition of further random edges cannot increase the number of $k$-colorings, by $(4.9)$ we find that

$$E\left[Z_k(G(n, d^*/n))^{1/n}\right] \leq E\left[Z_k(G(n, d^*/n))^{1/n}\left|m^* > m\right\}\right] + k \cdot P\left[e(G(n, d^*/n)) \leq m\right]$$

$$\leq E\left[Z_k(G(n, m))^{1/n}\right] + o(1).$$

Taking $n \to \infty$, and assuming that $d^* > d$ is sufficiently close to $d$, we conclude that

$$\lim sup_{n \to \infty} E\left[Z_k(G(n, d^*/n))^{1/n}\right] \leq k(1 - 1/k)^{d^*/2} - \delta < k(1 - 1/k)^{d^*/2}.$$

Hence, for any $\varepsilon > 0$ there is $d^* \in (d, d + \varepsilon)$ such that $d^* \in D^*$. Thus, $(4.8)$ follows from Lemma 3.5.

**Proof of Lemma 4.2.** Assuming the existence of $d$ and $(E_n)_{n \geq 1}$ as in Lemma 4.2, we are going to argue that

$$\lim sup_{n \to \infty} E\left[Z_k(G(n, m))^{1/n}\right] < k(1 - 1/k)^{d^*/2}. \quad (4.10)$$

Then the assertion follows from Lemma 4.6.

Since $Z_k^{1/n} \leq k$ with certainty and $P\left[G(n, m) \in E_n\right] = 1 - o(1)$, Jensen’s inequality yields

$$E\left[Z_k(G(n, m))^{1/n}\right] = E\left[Z_k(G(n, m))^{1/n} \cdot 1_{E_n}\right] + o(1)$$

$$\leq E\left[Z_k(G(n, m)) \cdot 1_{E_n}\right]^{1/n} + o(1).$$

Furthermore, by the linearity of expectation,

$$E\left[Z_k(G(n, m)) \cdot 1_{E_n}\right] = \sum_{\sigma : \lceil n \rceil \rightarrow \lceil k \rceil} P\left[E_n \text{ occurs and } \sigma \text{ is a } k\text{-coloring of } G(n, m)\right]$$

$$= \sum_{\sigma : \lceil n \rceil \rightarrow \lceil k \rceil} P\left[E_n | \sigma \text{ is a } k\text{-coloring of } G(n, m)\right]$$

$$\cdot P\left[\sigma \text{ is a } k\text{-coloring of } G(n, m)\right]$$

$$= \sum_{\sigma : \lceil n \rceil \rightarrow \lceil k \rceil} P\left[G(n, m, \sigma) \in E_n\right]$$

$$\cdot P\left[\sigma \text{ is a } k\text{-coloring of } G(n, m)\right]. \quad (4.11)$$

To estimate the last factor, we use (3.1) and Stirling’s formula, which yield

$$P\left[\sigma \text{ is a } k\text{-coloring of } G(n, m)\right] \leq \binom{n}{\lceil k \rceil} - \text{Forb}(\sigma) \leq \binom{n}{\lceil k \rceil} / \binom{m}{\lceil k \rceil} \leq O((1 - 1/k)^m).$$
Plugging this estimate into (4.11) and recalling that $\sigma$ is a random map $[n] \to [k]$, we obtain
\[
\mathbb{E}[Z_k(G(n, m)) \cdot 1_{\mathcal{E}_n}] \leq O((1 - 1/k)^m) \sum_{\sigma : [n] \to [k]} \mathbb{P}[G(n, m, \sigma) \in \mathcal{E}_n] \\
= O((1 - 1/k)^m) \cdot k^n \mathbb{P}[G(n, m, \sigma) \in \mathcal{E}_n] \\
= O(\mathbb{E}[Z_k(G(n, m))]) \cdot \mathbb{P}[G(n, m, \sigma) \in \mathcal{E}_n].
\]

Finally, using our assumption that $\limsup \mathbb{P}[G(n, m, \sigma) \in \mathcal{E}_n]^{1/n} < 1$ and combining (4.11) and (4.12), we see that
\[
\limsup \mathbb{E}[Z_k(G(n, m))]^{1/n} \leq k(1 - 1/k)^{d/2} \cdot \limsup \mathbb{P}[G(n, m, \sigma) \in \mathcal{E}_n]^{1/n} < k(1 - 1/k)^{d/2},
\]
thereby completing the proof of (4.10).

### 4.4. Proof of Lemma 4.4.

**Lemma 4.7.** Let $d > 0$. For any $\epsilon > 0$ there exists $\beta > 0$ such that
\[
\frac{1}{n} \ln \mathbb{E}[Z_{\beta,k}(G(n, m))] \leq \ln k + \frac{d}{2} \ln(1 - 1/k) + \epsilon.
\]

**Proof.** For any fixed number $\gamma > 0$ we can choose $\beta(\gamma) > 0$ so large that $\ln k - \beta \gamma < 0$. Now, let $\mathcal{M}(G(n, m))$ be the set of all $\sigma : [n] \to [k]$ such that at least $\gamma n$ edges are monochromatic under $\sigma$, and let $\overline{\mathcal{M}}(G(n, m))$ contain all $\sigma \notin \mathcal{M}(G(n, m))$. Then
\[
Z_{\beta,k}(G(n, m)) \leq |\mathcal{M}(G(n, m))| \cdot \exp(-\beta \gamma n) + |\overline{\mathcal{M}}(G(n, m))| \\
\leq k^n \cdot \exp(-\beta \gamma n) + |\overline{\mathcal{M}}(G(n, m))| \leq 1 + |\overline{\mathcal{M}}(G(n, m))|.
\]
Further, if $\sigma \in \overline{\mathcal{M}}(G(n, m))$, then $\sigma$ is a $k$-coloring of a subgraph of $G(n, m)$ containing $m - \gamma n$ edges. Hence, we obtain from Stirling’s formula that for $\gamma = \gamma(\epsilon) > 0$ small enough,
\[
\mathbb{P}[\sigma \in \overline{\mathcal{M}}(G(n, m))] \leq \left(\frac{n}{\gamma n}\right) \cdot \frac{\binom{n}{\gamma n} - \text{Forb}(\sigma)}{m - \gamma n} / \binom{n}{m} \\
\leq (1 - 1/k)^m \cdot \exp(\epsilon n/2).
\]
Hence,
\[
\mathbb{E}[\overline{\mathcal{M}}(G(n, m))] \leq k^n (1 - 1/k)^m \cdot \exp(\epsilon n/2).
\]
Combining (4.13) and (4.14), we obtain
\[
\mathbb{E}[Z_{\beta,k}(G(n, m))] \leq 1 + k^n (1 - 1/k)^m \cdot \exp(\epsilon n/2) < k^n (1 - 1/k)^m \cdot \exp(\epsilon n).
\]
Taking logarithms completes the proof.
Lemma 4.8. Assume that (2.5) is true. Then there exist a fixed number \( \varepsilon > 0 \), a sequence \( \sigma_n \) of balanced maps \([n] \rightarrow [k]\) and a sequence \( \mu_n \) of numbers satisfying \( |\mu_n - n/2| \leq \sqrt{n} \) such that
\[
\lim_{n \to \infty} \mathbb{P} \left[ |C(G(n, \mu_n, \sigma_n), \sigma_n)|^{1/n} > k(1 - 1/k)^{d/2} + \varepsilon \right] = 1.
\]

Proof. Let \( A \) be the event that the number of edges in the random graph \( G(n, p', \sigma) \) differs from \( dn/2 \) by at most \( \sqrt{n} \). Let \( N = \binom{n}{2} \). For any balanced \( \sigma : [n] \rightarrow [k] \) the expected number of edges in \( G(n, p', \sigma) \) is
\[
(N - \text{Forb}(\sigma))p' = (1 - 1/k)Np' + O(1) = dn/2 + O(1).
\]

Since the number of edges in \( G(n, p', \sigma) \) is a binomial random variable, (4.15) shows together with the central limit theorem that there exists a fixed \( \gamma > 0 \) such that for sufficiently large \( n \)
\[
\mathbb{P} \left[ G(n, p', \sigma) \in A \right] \geq \gamma \quad \text{for all balanced } \sigma.
\]

Furthermore, by Stirling’s formula there is an \( n \)-independent number \( \delta > 0 \) such that for sufficiently large \( n \) we have
\[
\mathbb{P} [\sigma \in \text{Bal}] \geq \delta.
\]

Combining (4.16) and (4.17), we see that
\[
\mathbb{P} [\sigma \in \text{Bal}, G(n, p', \sigma) \in A] = \mathbb{P} [\sigma \in \text{Bal}] \cdot \mathbb{P} [G(n, p', \sigma) \in A|\sigma \in \text{Bal}] \geq \gamma \delta > 0.
\]

Thus, pick \( \sigma_n \in \text{Bal} \) and \( \mu_n \in [dn/2 - \sqrt{n}, dn/2 + \sqrt{n}] \) that maximize
\[
p(\sigma_n, \mu_n) = \mathbb{P} \left[ |C(G(n, \mu_n, \sigma_n), \sigma_n)|^{1/n} > k(1 - 1/k)^{d/2} + \varepsilon \right].
\]

Then (2.5) and (4.18) imply that \( \lim_{n \to \infty} p(\sigma_n, \mu_n) = 1 \).

Lemma 4.9. For any \( \eta > 0 \) there is \( \delta > 0 \) such that
\[
\lim_{n \to \infty} \frac{1}{n} \ln \mathbb{P} \left[ \sum_{i=1}^{k} ||\sigma^{-1}(i)| - n/k| > \eta n \right] \leq -\delta.
\]

Proof. For each \( i \in [k] \) the number \( |\sigma^{-1}(i)| \) is a binomially distributed random variable with mean \( n/k \). Moreover, if \( \sum_{i=1}^{k} ||\sigma^{-1}(i)| - n/k| > \eta n \), then there is some \( i \in [k] \) such that \( ||\sigma^{-1}(i)| - n/k| > \eta n/k \). Thus, the assertion is immediate from the Chernoff bound.

Let \( \text{Vol}_G(S) \) be the sum of the degrees of the vertices in \( S \) in the graph \( G \).

Lemma 4.10. For any \( \gamma > 0 \) there is \( \alpha > 0 \) such that for any set \( S \subset [n] \) of size \( |S| \leq \alpha n \) and any \( \sigma : [n] \rightarrow [k] \) we have
\[
\limsup \frac{1}{n} \ln \mathbb{P} [\text{Vol}_{G(n, p', \sigma)}(S) > \gamma n] \leq -\alpha.
\]
Proof. Let \((X_v)_{v \in [n]}\) be a family of independent random variables with distribution \(\text{Bin}(n, p')\). Then for any set \(S\) the volume \(\text{Vol}(S)\) in \(G(n, p', \sigma)\) is stochastically dominated by \(X_S = 2 \sum_{v \in S} X_v\). Indeed, for each vertex \(v \in S\) the degree is a binomial random variable with mean at most \(np'\), and the only correlation amongst the degrees of the vertices in \(S\) is that each edge joining two vertices in \(S\) contributes two to \(\text{Vol}(S)\). Furthermore, \(\mathbb{E}[X_S] = 2d' |S|\). Thus, for any \(\gamma > 0\) we can choose an \(n\)-independent \(\alpha > 0\) such that for any \(S \subset [n]\) of size \(|S| \leq \alpha n\) we have \(\mathbb{E}[X_S] \leq \gamma n / 2\). In fact, the Chernoff bound shows that by picking \(\alpha > 0\) sufficiently small, we can ensure that

\[
\mathbb{P} \left[ \text{Vol}(S) \geq \gamma n \right] \leq \mathbb{P} \left[ X_S \geq \gamma n \right] \leq \exp(-\alpha n),
\]
as desired.

Lemma 4.11. Assume that there exist numbers \(z > 0, \varepsilon > 0\) and a sequence \((\sigma_n)_{n \geq 1}\) of balanced maps \([n] \to [k]\) such that

\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ \ln Z_{\beta, k}(G(n, p', \sigma_n)) \right] > z + \varepsilon.
\]

Then

\[
\limsup_{n \to \infty} \mathbb{P} \left[ \ln Z_{\beta, k}(G(n, p', \sigma)) \leq nz \right]^{1/n} < 1.
\]

Proof. Let \(Y = \frac{1}{n} \ln Z_{\beta, k}\) for the sake of brevity. Suppose that \(n\) is large enough so that \(\mathbb{E}[Y(G(n, p', \sigma_n))] > z + \varepsilon / 2\). Set \(n_i = |\sigma_i^{-1}(i)|\) and let \(T\) be the set of all \(\tau : [n] \to [k]\) such that \(|\tau^{-1}(i)| = n_i\) for \(i = 1, \ldots, k\). As \(Z_{\beta, k}\) is invariant under permutations of the vertices, we have

\[
\mathbb{E} \left[ Y(G(n, p', \tau)) \right] = \mathbb{E} \left[ Y(G(n, p', \sigma_n)) \right] > z + \varepsilon / 2 \quad \text{for any } \tau \in T. \quad (4.19)
\]

Let \(\gamma = \varepsilon / (4\beta) > 0\). By Lemma 4.10 there exists \(\alpha > 0\) such that for large enough \(n\) for any set \(S \subset V\) of size \(|S| \leq \alpha n\) and any \(\sigma : [n] \to [k]\) we have

\[
\mathbb{P} \left[ \text{Vol}_{G(n, p', \sigma)}(S) > \gamma n \right] \geq 1 - \exp(-\alpha n). \quad (4.20)
\]

Pick and fix a small \(0 < \eta < \alpha / 3\) and let \(A\) be the event that \(\sum_{i=1}^k |\sigma^{-1}(i)| - n / k | \leq \eta n\). Then by Lemma 4.9 there exists an \((n\)-independent\) number \(\delta = \delta(\beta, \varepsilon, \eta) > 0\) such that for \(n\) large enough

\[
\mathbb{P} \left[ A \right] \geq 1 - \exp(-\delta n). \quad (4.21)
\]

Because \(\sigma_n\) is balanced, we have \(|n_i - n / k| \leq \sqrt{n}\) for all \(i \in [k]\). Therefore, if \(A\) occurs, then it is possible to obtain from \(\sigma\) a map \(\tau_\sigma \in T\) by changing the colors of at most \(2\eta n\) vertices. If \(A\) occurs, we let \(G_1 = G(n, p', \tau_\sigma)\). Further, let \(G_2\) be the random graph obtained by removing from \(G_1\) all edges that are monochromatic under \(\sigma\). Finally, let \(G_3\) be the random graph obtained from \(G_2\) by inserting an edge between any two vertices \(v, w\) with \(\sigma(v) \neq \sigma(w)\) but \(\tau_\sigma(v) = \tau_\sigma(w)\) with probability \(p'\) independently. Thus, the bottom line is that in \(G_3\), we connect any two vertices that are colored differently under \(\sigma\) with probability \(p'\) independently. That is, \(G_3 = G(n, p', \sigma)\).

Let \(S_\sigma\) be the set of vertices \(v\) with \(\sigma(v) \neq \tau_\sigma(v)\) and let \(\Delta\) be the number of edges we removed to obtain \(G_2\) from \(G_1\). Then \(\Delta\) is bounded by the volume of \(S_\sigma\) in \(G_1 = G(n, p', \tau_\sigma)\). Hence, (4.20) implies that

\[
\mathbb{P} \left[ \Delta \leq \gamma n |A\right] \geq 1 - \exp(-\alpha n). \quad (4.22)
\]
Since removing a single edge can reduce \( Y \) by at most \( \beta/n \), we obtain
\[
\mathbb{P}[Y(G(n, p', \sigma)) \leq z] = \mathbb{P}[Y(G_3) \leq z] \leq \exp(-\alpha n) + \mathbb{P}[Y(G_3) \leq z | A] \quad \text{[by (4.21)]}
\]
\[
\leq \exp(-\delta n) + \exp(-\alpha n) + \mathbb{P}[Y(G_3) \leq z | A, \Delta \leq \gamma n] \quad \text{[by (4.22)]}
\]
\[
\leq \exp(-\delta n) + \exp(-\alpha n) + \exp(-\beta n) \mathbb{P}[Y(G_1) \leq z + \epsilon/4 | A] \quad \text{[by the choice of } \gamma \text{ and (4.22)]}
\]
\[
\leq \exp(-\delta n) + \exp(-\alpha n) + 3 \mathbb{P}[Y(G(n, p', \sigma_n)) \leq z + \epsilon/4] \quad \text{[by (4.21)]}
\]
\[
\leq \exp(-\delta n) + \exp(-\alpha n) + 3 \mathbb{P}[Y(G(n, p', \sigma_n)) \leq \mu(n, \sigma_n)] - \epsilon/4 \quad \text{[by (4.19)]}
\]

Finally, the assertion follows from Lemma 4.3.

**Proof of Lemma 4.4.** Lemma 4.8 shows that there exist \( \varepsilon > 0 \), balanced maps \( \sigma_n : [n] \to [k] \) and a sequence \( \mu_n \) satisfying \( |\mu_n - dn/2| \leq \sqrt{n} \) such that
\[
\lim_{n \to \infty} \mathbb{P} \left[ \frac{1}{n} \ln |C(G(n, \mu_n, \sigma_n), \sigma_n)| \geq \ln k + \frac{d}{2} \ln(1 - 1/k) + \varepsilon \right] = 1. \quad (4.23)
\]

By the definition of \( Z_{\beta, k} \), (4.23) implies that
\[
\lim_{n \to \infty} \mathbb{P} \left[ \frac{1}{n} \ln Z_{\beta, k}(G(n, \mu_n, \sigma_n)) \geq \ln k + \frac{d}{2} \ln(1 - 1/k) + \varepsilon \right] = 1 \quad \text{for all } \beta > 0. \quad (4.24)
\]

By comparison, Lemma 4.7 yields \( \beta > 0 \) such that with \( z = \ln k + \frac{d}{2} \ln(1 - 1/k) + \varepsilon/8 \) we have
\[
\lim_{n \to \infty} \mathbb{P} \left[ \frac{1}{n} \ln Z_{\beta, k}(G(n, m)) \leq z \right] = 1.
\]

Thus, we aim to prove that there is \( \alpha > 0 \) such that for sufficiently large \( n \)
\[
\mathbb{P} \left[ \frac{1}{n} \ln Z_{\beta, k}(G(n, m, \sigma)) \leq z + \varepsilon/8 \right] \leq \exp(-\alpha n). \quad (4.25)
\]

Indeed, since \( |\ln Z_{\beta, k}(G(n, \mu_n, \sigma_n))| \leq \beta \mu_n = O(n) \), (4.24) implies that for large enough \( n \)
\[
\frac{1}{n} \mathbb{E}[\ln Z_{\beta, k}(G(n, \mu_n, \sigma_n))] \geq \ln k + \frac{d}{2} \ln(1 - 1/k) + \varepsilon - o(1)
\]
\[
\geq \ln k + \frac{d}{2} \ln(1 - 1/k) + \varepsilon/2.
\]

Thus, since the number of edges in \( G(n, p', \sigma_n) \) is binomially distributed with expectation \( dn/2 + O(1) \), Eq. (4.25) follows from Lemma 4.11.
5. The Fixed Point Problem

5.1. The branching process. Throughout this section we assume that \((2k-1) \ln k - 2 \leq d \leq (2k-1) \ln k\). Moreover, we recall that \(d' = kd/(k-1)\).

Lemma 5.1. Suppose that \(d \geq (2k-1) \ln k - 2\).

1. The function

\[
F_{d,k} : [0, 1]^k \rightarrow [0, 1]^k, \quad (q_1, \ldots, q_k) \mapsto \left(\frac{1}{k} \prod_{j \in [k] \setminus \{i\}} 1 - \exp(-d'q_j)\right)_{i \in [k]}
\]

has a unique fixed point \(q^* = (q_1^*, \ldots, q_k^*)\) such that \(\sum_{j \in [k]} q_j^* \geq 2/3\). This fixed point has the property that \(q_1^* = \cdots = q_k^*\). Moreover, \(q^* = kq_1^*\) is the unique fixed point of the function \((2.6)\) in the interval \([2/3, 1]\), and \(q^* = 1 - O_k(1/k)\).

2. The branching process \(GW(d, k, q^*)\) is sub-critical.

3. Furthermore, \(\frac{\partial}{\partial d} E \left[ \frac{\ln Z(T_{d,k}(q^*)^*)}{|T_{d,k}(q^*)^*|} \right] = O_k(k^{-2})\).

The proof of Lemma 5.1 requires several steps. We begin by studying the fixed points of \(F_{d,k}\).

Lemma 5.2. The function \(F_{d,k}\) maps the compact set \([\frac{2}{3k}, \frac{1}{k}]^k\) into itself and has a unique fixed point \(q^*\) in this set. Moreover, the function from \((2.6)\) has a unique fixed point \(q^*\) in the set \([2/3, 1]\) and \(q^*(1) = (q^*/k, \ldots, q^*/k)\). Furthermore,

\[
q^* = 1 - 1/k + o_k(1/k).
\]

In addition, if \(q \in [0, 1]^k\) is a fixed point of \(F_{d,k}\), then

\[
q_1 = \cdots = q_k.
\]

Proof. Let \(I = [\frac{2}{3k}, \frac{1}{k}]^k\). As a first step, we show that \(F_{d,k}(I) \subset I\). Indeed, let \(q \in I\). Then for any \(i \in [k]\)

\[
(F_{d,k}(q))_i = \frac{1}{k} \prod_{j \neq i} 1 - \exp(-d'q_j) \leq \frac{1}{k}.
\]

On the other hand, as \(d \geq (2k-1) \ln k - 2\) we see that \(d' \geq 1.99k \ln k\). Hence,

\[
(F_{d,k}(q))_i = \frac{1}{k} \prod_{j \neq i} 1 - \exp(-d'q_j) \geq \frac{1}{k} \left(1 - \exp\left(-\frac{2d'}{3k}\right)\right)^{k-1} \geq \frac{1}{k} (1 - k^{-1.1})^k = \frac{1 - o_k(1)}{k}.
\]

Thus, \(F_{d,k}(I) \subset I\).

In addition, we claim that \(F_{d,k}\) is contracting on \(I\). In fact, for any \(i, j \in [k]\)

\[
\frac{\partial}{\partial q_j} (F_{d,k}(q))_i = \frac{1_{i \neq j}}{k} \prod_{l \neq i} 1 - \exp(-d'q_l) = \frac{1_{i \neq j} d'}{k \exp(d'q_j)} \prod_{l \neq i, j} 1 - \exp(-d'q_l) = (1 + o_k(1)) \frac{1_{i \neq j} d'}{k \exp(d'q_j)}.
\]
[as $d' \geq 1.99k \ln k$ and $q_l \geq 2/3$ for all $l$]
\[ \leq k^{-1.3} \quad \text{[for the same reason].} \]

Therefore, for $q \in I$ the Jacobi matrix $DF_{d,k}(q)$ satisfies
\[ \|DF_{d,k}(q)\|_2^2 \leq \sum_{i,j \in [k]} \left( \frac{\partial}{\partial q_j} (F_{d,k}(q))_i \right)^2 \leq k^2 \cdot k^{-2.6} < 1. \]

Thus, $F_{d,k}$ is a contraction on the compact set $I$. Consequently, Banach’s fixed point theorem implies that there is a unique fixed point $q_\ast \in I$.

To establish (5.3), assume without loss that $q = (q_1, \ldots, q_k) \in [0, 1]^k$ is a fixed point such that $q_1 \leq \cdots \leq q_k$. For the trivial fixed point $q_1 = \cdots = q_k = 0$, the Eq. (5.3) obviously holds. So we assume $q_1 > 0$. Because $q$ is a fixed point, we find
\[ \frac{q_k}{q_1} = \frac{(F_{d,k}(q))_k}{(F_{d,k}(q))_1} = \frac{1 - \exp(-d'q_1)}{1 - \exp(-d'q_k)} \leq 1 \quad \text{[as $q_1 \leq q_k$]}, \]

whence (5.3) follows.

Further, we claim that the function $f_{d,k} : [0, 1] \to [0, 1], q \mapsto (1 - \exp(-dq/(k - 1)))^{k-1}$ maps the interval $[2/3, 1]$ into itself. This is because for $q \in [2/3, 1]$ we have $0 \leq \exp(-dq/(k - 1)) \leq k^{-1.3}$ due to our assumption on $d$. Moreover, the derivative of $f$ works out to be $f_{d,k}'(q) = d \exp(-dq/(k - 1))(1 - \exp(-dq/(k - 1)))^{k-2}$. Thus, for $q \in [2/3, 1]$ we find $0 \leq f_{d,k}'(q) < 1/2$. Hence, $f_{d,k}$ has a unique fixed point $q_\ast \in [2/3, 1]$. Comparing the expressions $f_{d,k}(q)$ and $F_{d,k}(q)$, we see that $(q_\ast/k, \ldots, q_\ast/k)$ is a fixed point of $F_{d,k}$. Consequently, $q_\ast = (q_\ast/k, \ldots, q_\ast/k)$.

Finally, since $f_{d,k}'(q) > 0$ for all $q$, the function $f_{d,k}$ is strictly increasing. Therefore, as $d = (2 - o_k(1))k \ln k$,
\[
q_\ast = f_{d,k}(q_\ast) \leq f_{d,k}(1) = (1 - \exp(-d/(k - 1)))^{k-1} = 1 - 1/k + o_k(1/k). \tag{5.4}
\]

Similarly, $q_\ast \geq f_{d,k}(2/3) \geq 1 - k^{-0.3}$. Hence, because $d \geq (2k - 1) \ln k - 3$, we obtain
\[
q_\ast = f_{d,k}(q_\ast) \geq f_{d,k}(1 - k^{-0.3}) = \left(1 - \exp \left(-\frac{d(1 - k^{-0.3})}{k - 1}\right)\right)^{k-1} = \left(1 - k^{-2} + o_k(k^{-2.1})\right)^{k-1} = 1 - 1/k + o_k(1/k). \tag{5.5}
\]

Combining (5.4) and (5.5), we conclude that $q_\ast = 1 - 1/k + o_k(1/k)$, as claimed.

**Remark 1.** The proofs of several statements in this section (Lemmas 5.2, 5.5, 5.6, 3.1 and Corollary 5.3) directly incorporate parts of the calculations outlined in the physics work [33] that predicted the existence and location of $d_{k, \text{cond}}$. We redo these calculations here in detail to be self-contained and because not all steps are carried out in full detail in [33].

From here on out, we let $q^\ast$ denote the fixed point of $F_{d,k}$ in $[2/(3k), 1/k]^k$ and we denote the fixed point of the function (2.6) in the interval $[2/3, 1]$ by $q^\ast$. Hence, $q^\ast = (q^\ast/k, \ldots, q^\ast/k)$. If we keep $k$ fixed, how does $q^\ast$ vary with $d'$?
Corollary 5.1. We have $\frac{dq^*}{dd} = \Theta_k (k^{-2})$.

Proof. The map $d \mapsto q^*$ is differentiable by the implicit function theorem. Moreover, differentiating (2.6) while keeping in mind that $q^* = q^*(d)$ is a fixed point, we find

$$\frac{dq^*}{dd} = \frac{d}{dd} \left( 1 - \exp(-dq^*/(k-1)) \right)^{k-1}$$

$$= (k-1) \left( 1 - \exp(-dq^*/(k-1)) \right)^{k-2} \frac{\exp(dq^*/(k-1))}{k-1 + \frac{d}{dd} dq^*}$$

Rearranging the above using $d = 2k \ln k + O_k(\ln k)$ and (5.2) yields the assertion.

Corollary 5.2. We have $q^*_{i,\ell} = \tilde{\Theta}_k (k^{-2} |\ell|-1)$ for all $(i, \ell) \in T$. Moreover, $\frac{dq^*_{i,\ell}}{dd} = \tilde{\Theta}_k (|\ell| k^{-2}|\ell|)$.

Proof. Lemma 5.2 shows that $q^*_j = q_*/k$ for all $j \in [k]$. Hence, due to (5.2) and because $d' = 2k \ln k + O_k(\ln k)$ we obtain

$$q^*_{i,\ell} = \frac{1}{k} \prod_{j \in [k] \setminus \ell} \left( 1 - \exp(-d'q^*_j) \right) \prod_{j \in \ell \setminus \{i\}} \exp(-d'q^*_j) = \tilde{\Theta}_k (k^{-2} |\ell|-1).$$

Furthermore, applying Corollary 5.1, we get

$$\frac{dq^*_{i,\ell}}{dd} = \frac{1}{k} \frac{d}{dd} \left[ \prod_{j \in [k] \setminus \ell} \left( 1 - \exp(-d'q^*_j) \right) \prod_{j \in \ell \setminus \{i\}} \exp(-d'q^*_j) \right]$$

$$= \frac{1}{k} \frac{d}{dd} \left[ (1 - \exp(-d'q^*/k))^{k-1} \exp(-d'q^*/k)^{|\ell|-1} \right]$$

$$= \frac{1}{k} \left( \frac{q_*}{k-1} + \frac{d' dq_*}{k} \right) \left[ \frac{k-|\ell|}{\exp(d'q^*/k)} (1 - \exp(-d'q^*/k))^{k-|\ell|-1} \right.$$

$$- |\ell| (1 - \exp(-d'q^*/k))^{k-|\ell|} \exp(-d'(|\ell| - 1)q^*/k)$$

$$= |\ell| O_k(k^{-2}) \exp(-d'(|\ell| - 1)q^*/k) = \tilde{\Theta}_k (|\ell| k^{-2}|\ell|).$$

Lemma 5.3. The branching process $GW(d, k, q^*)$ is sub-critical.

Proof. We introduce another branching process $GW'(d, k, q^*)$ with only three types $1, 2, 3$. The idea is that type 1 of the new process represents all types $(h, \{h\}) \in T$ with $h \in [k]$, that 2 represents all types $(h, \{j, h\}) \in T$ with $h, j \in [k], j \neq h$, and that 3 lumps together all of the remaining types. More specifically, in $GW'(d, k, q^*)$ an individual of type $i$ spawns a Poisson number $Po(M_{ij})$ of offspring of type $j$ $(i, j \in \{1, 2, 3\})$, where $M = (M_{ij})$ is the following matrix. If either $i = 1$ or $j = 1$, then $M_{ij} = 0$. Moreover,

$$M_{22} = \sum_{(i,\ell) \in T, |\ell|=2} q^*_{i,\ell} d' \quad M_{23} = \sum_{(i,\ell) \in T, |\ell|>2} q^*_{i,\ell} d'.$$
Due to the symmetry of the fixed point \( q^* \) (i.e., \( q^* = (q^*/k, \ldots, q^*/k) \)), \( M_{22} \) is precisely the expected number of offspring of type \((i, \ell)\) with \(|\ell| = 2\) that an individual of type \((i_0, \ell_0) \in T\) with \(|\ell_0| = 2\) spawns in the branching process \( GW(d, k, q^*) \). Similarly, \( M_{23} \) is just the expected offspring of type \((i, \ell)\) with \(|\ell| > 2\) of an individual with \(|\ell_0| = 2\). Furthermore, \( M_{32} \) is an upper bound on the expected offspring of type \((i', \ell')\) with \(|\ell'| = 2\) of an individual of type \((i_0, \ell_0)\) with \(|\ell_0| > 2\). Indeed, \( M_{32} \) is the expected offspring in the case that \( \ell_0 = [k] \), which is the case that yields the largest possible expectation. Similarly, \( M_{33} \) is an upper bound on the expected offspring of type \((i', \ell')\) with \(|\ell'| > 2\) in the case \(|\ell_0| > 2\). Therefore, if \( GW'(d, k, q^*) \) is sub-critical, then so is \( GW(d, k, q^*) \).

To show that this is the case, we need to estimate the entries \( M_{ij} \). Estimating the \( q^*_{i,\ell} \) via Corollary 5.2, we obtain

\[
M_{22} \leq 2kq^*_{1,\{1,2\}}d' \leq \tilde{O}_k(k^{-1}), \quad M_{23} \leq 2 \sum_{l \geq 3} l \binom{k}{l-1} q^*_l d' \leq \tilde{O}_k(k^{-2}), \\
M_{32} \leq k(k-1)q^*_{1,\{1,2\}}d' \leq \tilde{O}_k(1), \quad M_{33} \leq k \sum_{l \geq 3} l \binom{k}{l-1} q^*_l d' \leq \tilde{O}_k(k^{-1}).
\]

The branching process \( GW'(d, k, q^*) \) is sub-critical iff all eigenvalues of \( M \) are less than 1 in absolute value. Because the first row and column of \( M \) are 0, this is the case iff the eigenvalues of the \( 2 \times 2 \) matrix \( M_* = (M_{ij})_{2 \leq i, j \leq 3} \) are less than 1 in absolute value. Indeed, since the above estimates show that \( M_* \) has trace \( \tilde{O}_k(k^{-1}) \) and determinant \( \tilde{O}_k(k^{-2}) \), both eigenvalues of \( M_* \) are \( \tilde{O}_k(k^{-1}) \).

**Lemma 5.4.** We have \( \frac{d}{dT} \mathbb{E}[|T_{d,k,q^*}|^{-1} \ln \mathbb{E}(T_{d,k,q^*})] = \tilde{O}_k(k^{-2}) \).

**Proof.** Fix a number \( \delta \in [(2k - 1) \ln k - 2, (2k - 1) \ln k] \) and a small number \( \varepsilon > 0 \) and let \( \hat{d} = d + \varepsilon \). Let \( \hat{q}^* \) be the unique fixed point of \( F_{d,k} \) in \( [2/(3k), 1/k] \) and let \( \hat{q}^* \) be the unique fixed point of \( F_{\hat{d},k} \) in \( [2/(3k), 1/k] \). Set \( d' = dk/(k - 1) \) and \( \hat{d}' = \hat{d}k/(k - 1) \).

Moreover, let us introduce the shorthands \( T = T_{d,k,q^*} \) and \( \hat{T} = T_{d,k,\hat{q}^*} \). We aim to bound

\[
\Delta = \mathbb{E} \left[ \frac{\ln \mathbb{E}(T)}{|T|} \right] - \mathbb{E} \left[ \frac{\ln \mathbb{E}(\hat{T})}{|\hat{T}|} \right].
\]

To this end, we couple \( T \) and \( \hat{T} \) as follows.

- In \( T, \hat{T} \) the type \((i_0, \ell_0)\) resp. \((\hat{i}_0, \hat{\ell}_0)\) of the root \( v_0 \) is chosen from the distribution
  \( Q = (q_{i,\ell})_{(i,\ell) \in T} \) resp. \( \hat{Q} = (\hat{q}_{i,\ell})_{(i,\ell) \in T} \).

  We couple \((i_0, \ell_0), (\hat{i}_0, \hat{\ell}_0)\) optimally.

- If \((i_0, \ell_0) \neq (\hat{i}_0, \hat{\ell}_0)\), then we generate \( T, \hat{T} \) independently from the corresponding conditional distributions given the type of the root.

- If \((i_0, \ell_0) = (\hat{i}_0, \hat{\ell}_0)\), we generate a random tree \( \hat{T} \) by means of the following branching process.
the condensation phase transition in random graph coloring 573

now, let $E$ be the event that $\ell_0 = (i_0, \ell_0)$. if $A \cap \hat{A}$ occurs, then both $T, \hat{T}$ consist of a single vertex and have precisely one legal coloring. thus, $|T|^{-1} \ln Z(T) = |\hat{T}|^{-1} \ln Z(\hat{T}) = 0$. consequently,

$$
\Delta \leq \mathbb{E}\left[\ln \frac{Z(T)}{|T|} - \ln \frac{Z(\hat{T})}{|\hat{T}|} \right] \cdot \mathbb{P}\left[\neg A \lor \neg \hat{A}\right] + \mathbb{P}\left[\neg A \lor \neg \hat{A}\right] \cdot k.
$$

further, since $|T|^{-1} \ln Z(T), |\hat{T}|^{-1} \ln Z(\hat{T}) \leq \ln k$ with certainty, we obtain

$$
\Delta \leq \left(\mathbb{P}\left[\neg A \land \hat{A}\right] + \mathbb{P}\left[A \land \neg \hat{A}\right]\right) \ln k
+ \mathbb{E}\left[\ln \frac{Z(T)}{|T|} - \ln \frac{Z(\hat{T})}{|\hat{T}|} \right] \cdot \mathbb{P}\left[\neg A \land \neg \hat{A}\right].
$$

because $(i_0, \ell_0)$ and $(\hat{i}_0, \hat{\ell}_0)$ are coupled optimally and $\mathbb{P}[A] = kq^*_1, \mathbb{P}[\hat{A}] = k\hat{q}_1^*$, corollary 5.1 implies that $\mathbb{P}[\neg A \land \hat{A}], \mathbb{P}[A \land \neg \hat{A}] \leq \varepsilon \hat{O}_k(k^{-2})$. hence,

$$
\Delta \leq \varepsilon \hat{O}_k(k^{-2}) + \mathbb{E}\left[\ln \frac{Z(T)}{|T|} - \ln \frac{Z(\hat{T})}{|\hat{T}|} \right] \cdot \mathbb{P}\left[\neg A \land \neg \hat{A}\right]. \quad (5.6)
$$

now, let $E$ be the event that $\ell_0 \neq (i_0, \ell_0), \hat{\ell}_0 \neq (\hat{i}_0, \hat{\ell}_0)$ and $(i_0, \ell_0) = (\hat{i}_0, \hat{\ell}_0)$. due to corollary 5.2 and because $(i_0, \ell_0), (\hat{i}_0, \hat{\ell}_0)$ are coupled optimally, we see that

$$
\mathbb{P}\left[\neg A \land \neg \hat{A} \land \neg E\right] \leq \varepsilon \hat{O}_k(k^{-2}). \quad (5.7)
$$

combining (5.6) and (5.7), we conclude that

$$
\Delta \leq \varepsilon \hat{O}_k(k^{-2}) + \mathbb{E}\left[\ln \frac{Z(T)}{|T|} - \ln \frac{Z(\hat{T})}{|\hat{T}|} \right] \cdot \mathbb{P}[\neg A \land \neg \hat{A}] \quad (5.8)
$$
Further, since $\mathbb{P}\left[ -\mathcal{A} \land \neg \tilde{\mathcal{A}} \right] \leq \mathbb{P}\left[ -\mathcal{A} \right] \leq 1 - kq_1^* \leq O_k(1/k)$ by Lemma 5.1, (5.8) yields

$$\Delta \leq \epsilon \tilde{O}_k(k^{-2}) + O_k(1/k) \cdot \mathbb{E}\left[ \left| \ln Z(T) \right| \right] \leq \epsilon \tilde{O}_k(k^{-2}) + O_k(1/k) \cdot \mathbb{P}\left[ T \neq \hat{T} \mid \mathcal{E} \right]. \tag{5.9}$$

Thus, we are left to estimate the probability that $T \neq \hat{T}$, given that both trees have a root of the same type $(i_0, \ell_0)$ with $|\ell_0| > 1$. Our coupling ensures that this event occurs iff $s_v = 1$ for some vertex $v$ of $\tilde{T}$. To estimate the probability of this event, we observe that by Corollary 5.2

$$\lambda_{i,\ell} \leq \begin{cases} \epsilon \tilde{O}_k(1/k) & \text{if } |\ell| = 2, \\ \epsilon \tilde{O}_k(1) & \text{if } |\ell| > 2. \end{cases} \tag{5.10}$$

Now, let $\mathcal{N}_1$ be the number of vertices $v \neq v_0$ of $\tilde{T}$ such that $|\ell_v| = 2$, and let $\mathcal{N}_2$ be the number of $v \neq v_0$ such that $|\ell_v| > 2$. Then (5.9), (5.10) and the construction of the coupling yield

$$\Delta / \epsilon \leq \tilde{O}_k(k^{-2}) + \tilde{O}_k(k^{-1}) \left( k^{-1} \mathbb{E}[\mathcal{N}_1|\mathcal{E}] + \mathbb{E}[\mathcal{N}_2|\mathcal{E}] \right). \tag{5.11}$$

To complete the proof, we claim that

$$\mathbb{E}[\mathcal{N}_1|\mathcal{E}] \leq \tilde{O}_k(k^{-1}), \quad \mathbb{E}[\mathcal{N}_2|\mathcal{E}] \leq \tilde{O}_k(k^{-2}). \tag{5.12}$$

Indeed, consider the matrix $\tilde{M} = (\tilde{M}_{ij})_{i,j=1,2}$ with entries

$$\tilde{M}_{11} = \sum_{(i,\ell) \in T_{1,\{1,2\}}: |\ell| = 2} \Lambda_{i,\ell}, \quad \tilde{M}_{12} = \sum_{(i,\ell) \in T_{1,\{1,2\}}: |\ell| > 2} \Lambda_{i,\ell},$$

$$\tilde{M}_{21} = \sum_{(i,\ell) \in T_{1,|\ell|}: |\ell| = 2} \Lambda_{i,\ell}, \quad \tilde{M}_{22} = \sum_{(i,\ell) \in T_{1,|\ell|}: |\ell| > 2} \Lambda_{i,\ell}.$$

Then Corollary 5.2 entails that

$$\tilde{M}_{11} = \tilde{O}_k(k^{-1}), \quad \tilde{M}_{12} = \tilde{O}_k(k^{-2}), \quad \tilde{M}_{21} = \tilde{O}_k(1), \quad \tilde{M}_{22} = \tilde{O}_k(k^{-1}).$$

In addition, let $\xi = \left( \xi_1 \xi_2 \right)$, where $\xi_1 = 1 - \xi_2 = \mathbb{P}[ |\ell_0| = 2 | \mathcal{E} ]$. Then Corollary 5.2 shows that $\xi_2 = \tilde{O}_k(k^{-2})$. Furthermore, by the construction of the branching process and (5.2)

$$\left( \frac{\mathbb{E}[\mathcal{N}_1|\mathcal{E}]}{\mathbb{E}[\mathcal{N}_2|\mathcal{E}]} \right) \leq \sum_{t=1}^{\infty} \tilde{M}^t \xi \leq \left( \frac{\tilde{O}_k(k^{-1})}{\tilde{O}_k(k^{-2})} \right),$$

which implies (5.12).

Finally, (5.11) and (5.12) imply that $\Delta \leq \epsilon \tilde{O}_k(k^{-2})$. Taking $\epsilon \to 0$ completes the proof.

**Proof of Lemma 5.1.** The first assertion is immediate from Lemma 5.2. The second claim follows from Lemma 5.3, and the third one from Lemma 5.4.
5.2. The “hard fields”. In this section we make the first step towards proving that \( \pi_{d,k,q^*} \) is the unique frozen fixed point of \( \mathcal{F}_{d,k} \). More specifically, identifying the set \( \Omega \) with the \( k \)-simplex, we show that every face of \( \Omega \) carries the same probability mass under any frozen fixed point of \( \mathcal{F}_{d,k} \) as under the measure \( \pi_{d,k,q^*} \). Formally, let us denote the extremal points of \( \Omega \) by \( \delta_h = (1_{j=i})_{j \in [k]} \), i.e., \( \delta_h \) is the probability measure on \([k]\) that puts mass 1 on the single point \( h \in [k] \). In addition, let \( \Omega_\ell \) be the set of all \( \mu \in \Omega \) with support \( \ell \) (i.e., \( \mu(i) > 0 \) for all \( i \in \ell \) and \( \mu(i) = 0 \) for all \( i \in [k] \setminus \ell \)). Further, for a probability measure \( \pi \in \mathcal{P} \) we let \( \rho_h(\pi) = \pi(\{\delta_h\}) \) denote the probability mass of \( \delta_h \) under \( \pi \). In physics jargon, the numbers \( \rho_h(\pi) \) are called the “hard fields” of \( \pi \). In addition, recalling that \( \text{d}\pi_i(\mu) = k\mu(i)\text{d}\pi(\mu) \), we set \( \rho_{i,\ell}(\pi) = \pi_i(\Omega_\ell) \) for any \((i, \ell) \in T \).

The main result of this section is

**Lemma 5.5.** Suppose that \( d \geq (2k - 1) \ln k - 2 \). Let \( q^* \in [2/3, 1] \) be the fixed point of (2.6). If \( \pi \in \mathcal{P} \) is a frozen fixed point of \( \mathcal{F}_{d,k} \), then \( \rho_i(\pi) = q^*/k \) and \( \rho_{i,\ell}(\pi) = kq^*_\ell \) for all \((i, \ell) \in T \).

To avoid many case distinctions, we introduce the following convention when working with product measures. Let us agree that \( \Omega^0 = \{\emptyset\} \). Hence, if \( B : \Omega^0 \rightarrow \Omega \) is a map, then \( B(\emptyset) \in \Omega \). Furthermore, there is precisely one probability measure \( \pi_0 \) on \( \Omega^0 \), namely the measure that puts mass one on the point \( \emptyset \in \Omega^0 \). Thus, the integral \( \int_{\Omega^0} B(\mu) \text{d}\pi_0(\mu) \) is simply equal to \( B(\emptyset) \). If \( \pi_1, \pi_2, \ldots \) are probability measures on \( \Omega \), what we mean by the empty product measure \( \bigotimes_{y=1}^0 \pi_y \) is just the measure \( \pi_0 \) on \( \Omega^0 \).

Further, for a real \( \lambda \geq 0 \) and an integer \( y \geq 1 \) we let

\[
p_{\lambda}(y) = \lambda^y \exp(-\lambda)/y!.
\]

Moreover, for \( i \in [k] \) we let \( I_i \) be the set of all non-negative integer vectors \( y = (y_j)_{j \in [k] \setminus \{i\}} \) and for \( y \in I_i \) we set

\[
p_i(y) = \prod_{h \in [k] \setminus \{i\}} p_{\lambda,d}(y_h).
\]

We also let \( \Omega^y = \prod_{h \in [k] \setminus \{i\}} \prod_{j \in [y_h]} \Omega \) for \( y \in I_i \). The elements of \( \Omega^y \) are denoted by \( \mu_y = (\mu_{h,j})_{h \in [k] \setminus \{i\}, j \in [y_h]} \). Moreover, let

\[
\pi_{i,y} = \bigotimes_{h \in [k] \setminus \{i\}} \bigotimes_{j \in [y_h]} \pi_h.
\]

Thus, with the convention from the previous paragraph, in the case \( y = 0 \) the set \( \Omega^0 = \{\emptyset\} \) contains only one element, namely \( \mu_0 = \emptyset \). Moreover, \( \pi_{i,y} \) is the probability measure on \( \Omega^0 \) that gives mass one to the point \( \emptyset \). We recall the map \( B : \bigcup_{y \geq 1} \Omega^y \rightarrow \Omega \) from (1.1) and extend this map to \( \Omega^0 \) by letting \( B(\emptyset) = \frac{1}{k} \mathbf{1} \) be the uniform distribution on \( \Omega \). We start the proof of Lemma 5.5 by establishing the following identity.

**Lemma 5.6.** If \( \pi \) is fixed point of \( \mathcal{F}_{d,k} \), then for any \( i \in [k] \) we have

\[
\pi_i = \sum_{y \in I_i} \int_{\Omega^y} \delta_{B[\mu_y]} p_i(y) \text{d}\pi_{i,y}(\mu_y).
\]

To establish Lemma 5.6 we need to calculate the normalising quantities \( Z_y(\pi) \).
Lemma 5.7. If $\pi$ is fixed point of $\mathcal{F}_{d,k}$, then $Z_\gamma(\pi) = (k-1)^{\gamma}/k^{\gamma-1}$.

Proof. Assume that $\pi$ is fixed point of $\mathcal{F}_{d,k}$. We claim that

$$\int_\Omega \mu(h) d\pi(\mu) = 1/k \quad \text{for all } h \in [k]. \tag{5.13}$$

Indeed, set $v(h) = \int_\Omega \mu(h) d\pi(\mu)$. Then $v$ is a probability distribution on $[k]$. Since $\pi$ is a fixed point of $\mathcal{F}_{d,k}$, we find

$$v(h) = \int_\Omega \mu(h) d\mathcal{F}_{d,k}[\pi](\mu) = \sum_{\gamma=0}^\infty \frac{p_d(\gamma)}{Z_\gamma(\pi)} \int_\Omega \left[ \prod_{h=1}^k \prod_{j=1}^\gamma (1 - \mu_j(h)) \right] \cdot \mathcal{B}[\mu_1, \ldots, \mu_\gamma](h) \otimes d\pi(\mu_j)$$

$$= \sum_{\gamma=0}^\infty \frac{p_d(\gamma)}{Z_\gamma(\pi)} \int_\Omega \prod_{j=1}^\gamma (1 - \mu_j(h)) \otimes d\pi(\mu_j) \quad \text{[plugging in (1.1)]}$$

$$= \sum_{\gamma=0}^\infty \frac{p_d(\gamma)}{Z_\gamma(\pi)} \left[ \int_\Omega 1 - \mu(h) d\pi(\mu) \right]^\gamma = \sum_{\gamma=0}^\infty \frac{p_d(\gamma)}{\sum_{h' \in [k]} (1 - v(h'))^\gamma} \sum_{\gamma=0}^\infty (1 - v(h))^{\gamma} \frac{p_d(\gamma)}{\sum_{h' \in [k]} (1 - v(h'))^\gamma} = v(h). \tag{5.14}$$

Now, assume that $h_1, h_2 \in [k]$ are such that $v(h_1) \leq v(h_2)$. Then (5.14) yields

$$v(h_2) = \sum_{\gamma=0}^\infty (1 - v(h_1))^{\gamma} \frac{p_d(\gamma)}{\sum_{h' \in [k]} (1 - v(h'))^{\gamma}} \leq \sum_{\gamma=0}^\infty (1 - v(h_2))^{\gamma} \frac{p_d(\gamma)}{\sum_{h' \in [k]} (1 - v(h'))^{\gamma}} = v(h_1).$$

Hence, $v(h_1) = v(h_2)$ for all $h_1, h_2 \in [k]$, which implies (5.13). Finally, the assertion follows from (5.13) and the definition (1.2) of $Z_\gamma(\pi)$.

Proof of Lemma 5.6. If $\pi$ is a fixed point of $\mathcal{F}_{d,k}$, then by Lemma 5.7 and the definition (1.1) of the map $\mathcal{B}$ we have

$$\pi_i = \int_\Omega k \mu(i) \delta_\mu d\pi(\mu) = \int_\Omega k \mu(i) \delta_\mu d\mathcal{F}_{d,k}[\pi](\mu)$$

$$= \sum_{\gamma=0}^\infty \frac{p_d(\gamma)}{Z_\gamma(\pi)} \int_\Omega \left[ \prod_{h=1}^k \prod_{j=1}^\gamma (1 - \mu_j(h)) \right] \cdot \mathcal{B}[\mu_1, \ldots, \mu_\gamma](i) \otimes d\pi(\mu_j)$$

$$= \sum_{\gamma=0}^\infty \frac{k^\gamma p_d(\gamma)}{(k-1)^\gamma} \int_\Omega \left[ \prod_{j=1}^{\gamma} (1 - \mu_j(i)) \right] \cdot \delta_{\mathcal{B}[\mu_1, \ldots, \mu_\gamma]} \otimes d\pi(\mu_j).$$

Further, for any $\mu \in \Omega$ we have $1 - \mu(i) = \sum_{i' \neq i} \mu(i')$. Hence,

$$\pi_i = \sum_{\gamma=0}^\infty \frac{k^\gamma p_d(\gamma)}{(k-1)^\gamma} \sum_{i_1, \ldots, i_\gamma \in [k]\setminus\{i\}} \int_\Omega \left[ \prod_{j=1}^{\gamma} \mu_j(i_j) \right] \cdot \delta_{\mathcal{B}[\mu_1, \ldots, \mu_\gamma]} \otimes d\pi(\mu_j).$$
The Condensation Phase Transition in Random Graph Coloring

\[
= \sum_{\gamma=0}^{\infty} \frac{p_d(\gamma)}{(k-1)^\gamma} \sum_{i_1, \ldots, i_{\gamma} \in [k] \setminus \{i\}} \int_{\Omega^\gamma} \delta_{B[\mu_1, \ldots, \mu_\gamma]} \prod_{j=1}^{\gamma} d\pi_{i_j}(\mu_{i_j}). \quad (5.15)
\]

In the last expression, we can think of generating the sequence \(i_1, \ldots, i_\gamma\) as follows: first, choose \(\gamma\) from the Poisson distribution \(\text{Po}(d)\). Then, choose the sequence \(i_1, \ldots, i_\gamma\) by independently choosing \(i_j\) from the set \([k] \setminus \{i\}\) uniformly at random. Thus, in the overall experiment the number of times that each color \(h\) occurs has distribution \(\text{Po}(d/(k-1))\), independently for all \(h \in [k] \setminus \{i\}\), whence (5.15) implies the assertion.

**Corollary 5.3.** If \(\pi\) is fixed point of \(F_{d,k}\), then \((\rho_i(\pi))_{i \in [k]}\) is a fixed point of the function \(F_{d,k}\) from Lemma 5.1.

**Proof.** Invoking Lemma 5.6, we obtain for any \(i \in [k]\)

\[
\rho_i(\pi) = \pi([\delta_i]) = \frac{\pi_i([\delta_i])}{k} = \frac{1}{k} \sum_{y \in \Omega} \int_{\Omega^\gamma} \delta_{\pi_i} \sim B[\mu_\pi, \gamma] \prod_{j=1}^{\gamma} d\pi_{i_j}(\mu_{i_j}). \quad (5.16)
\]

A glimpse at the definition (1.1) of \(B\) reveals that \(\delta_i = B[\mu_\pi]\) iff for each \(h \in [k] \setminus \{i\}\) there is \(j \in [\gamma_h]\) such that \(\mu_{h,j} = \delta_h\). Further, in (5.16) the \(\mu_{h,j}\) are chosen independently from the distribution \(\pi_h\), and \(\pi_h(\delta_h) = k\rho_h(\pi)\). In effect, the r.h.s. of (5.16) is simply the probability that if we choose numbers \(\gamma_h\) independently from the Poisson distribution with mean \(d/(k-1)\) for \(h \neq i\) and then perform \(\gamma_h\) independent Bernoulli experiments with success probability \(k\rho_h(\pi)\), then there occurs at least one success for each \(h \neq i\). Of course, this is nothing but the probability that \(k-1\) independent Poisson variables \((\text{Po}(\rho_h(\pi)dk/(k-1)))_{h \neq i}\) are all strictly positive. Hence,

\[
\rho_i(\pi) = \frac{1}{k} \prod_{h \in [k] \setminus \{i\}} \mathbb{P}[\text{Po}(\rho_h(\pi)dk/(k-1)) > 0] = \frac{1}{k} \prod_{h \in [k] \setminus \{i\}} 1 - \exp(-\rho_h(\pi)d') \quad \text{for any } i \in [k].
\]

Consequently, \((\rho_i(\pi))_{i \in [k]} = F_{d,k}((\rho_i(\pi))_{i \in [k]}).

**Proof of Lemma 5.5.** Assume that \(\pi \in \mathcal{P}\) is a frozen fixed point of \(F_{d,k}\). Then \(\rho_1(\pi) \geq \frac{2}{3k}\) for all \(i \in [k]\). Hence, Corollary 5.3 yields \((\rho_1(\pi), \ldots, \rho_k(\pi)) \in [\frac{2}{3k}, \frac{1}{k}]^k\) is a fixed point of \(F_{d,k}\). Therefore, Lemma 5.1 implies that \(\rho_i(\pi) = q^n/k\) for all \(i \in [k]\).

To prove the second assertion, let \((i, \ell) \in T\). Then Lemma 5.6 yields

\[
\rho_{i,\ell}(\pi) = \sum_{y \in \Omega} \int_{\Omega^\gamma} 1_{B[\mu_\pi]} \in \Omega_{\ell} \prod_{j=1}^{\gamma} d\pi_{i_j}(\mu_{i_j}). \quad (5.17)
\]

Now, the definition (1.1) is such that \(B[\mu_\pi] \in \Omega_{\ell}\) iff

1. for each \(h \in [k] \setminus \ell\) there is \(j \in [\gamma_h]\) such that \(\mu_{h,j} = \delta_h\), and
2. for each \(h \in \ell \setminus \{i\}\) and any \(j \in [\gamma_h]\) we have \(\mu_{h,j} \neq \delta_h\).
Given $\gamma$, the distributions $\mu_h, j$ are chosen independently from $\pi_h$ for all $h \neq i, j \in [\gamma_h]$. Hence, for a given $\gamma$ the probability that (1) and (2) occur is precisely

$$
\eta(\gamma) = \prod_{h \in \ell \setminus \{i\}} (1 - \pi_h([\delta_h]))^{\gamma_h} \cdot \prod_{h \in [k] \setminus \ell} (1 - (1 - \pi_h([\delta_h]))^{\gamma_h} = \prod_{h \in \ell \setminus \{i\}} (1 - k\rho_h(\pi))^{\gamma_h} \cdot \prod_{h \in [k] \setminus \ell} (1 - (1 - k\rho_h(\pi))^{\gamma_h}. \tag{5.18}
$$

Thus, combining (5.17) and (5.18), we see that

$$
\rho_{i,\ell}(\pi) = \sum_{\gamma \in \Gamma_i} \eta(\gamma) p_i(\gamma) = \prod_{h \in \ell \setminus \{i\}} \left[ \sum_{\gamma_h \geq 0} (1 - k\rho_h(\pi))^{\gamma_h} p_{d_{i,\ell}}(\gamma_h) \right] \cdot \prod_{h \in [k] \setminus \ell} \left[ \sum_{\gamma_h \geq 0} (1 - (1 - k\rho_h(\pi))^{\gamma_h} p_{d_{i,\ell}}(\gamma_h) \right] = \prod_{h \in \ell \setminus \{i\}} \mathbb{P}[\text{Po}(dk\rho_h(\pi)/(k - 1) = 0)] \prod_{h \in [k] \setminus \ell} \mathbb{P}[\text{Po}(dk\rho_h(\pi)/(k - 1) > 0)] = \prod_{h \in \ell \setminus \{i\}} \exp(-d'\rho_h(\pi)) \prod_{h \in [k] \setminus \ell} 1 - \exp(-d'\rho_h(\pi)). \tag{5.19}
$$

Finally, as we already know from the first paragraph that $\rho_h(\pi) = q^*/k$, (5.19) implies that $\rho_{i,\ell}(\pi) = kq^*_{i,\ell}$.

5.3. The fixed point. The objective in this section is to establish

**Lemma 5.8.** Suppose that $d \geq (2k - 1) \ln k - 2$. Then $\pi_{d,k, q^*}$ is the unique frozen fixed point of $F_{d,k}$.

To prove Lemma 5.8, let $\mathcal{P}_\ell$ be the set of all probability measures $\pi \in \mathcal{P}$ whose support is contained in $\Omega_\ell$ (i.e., $\pi(\Omega_\ell) = 1$). For each $\pi \in \mathcal{P}$ and any $(i, \ell) \in \mathcal{T}$ we define a measure $\pi_{i,\ell}$ by letting

$$
d\pi_{i,\ell}(\mu) = \frac{1_{\mu \in \Omega_\ell}}{kq^*_i} d\pi_i(\mu) = \frac{\mu(i)}{q^*_i} 1_{\mu \in \Omega_\ell} d\pi(\mu).
$$

In addition, let $\overline{\mathcal{P}} = \prod_{(i, \ell) \in \mathcal{T}} \mathcal{P}_\ell$ be the set of all families $(\pi_{i,\ell})_{i,\ell \in \mathcal{T}}$ such that $\pi_{i,\ell} \in \mathcal{P}_\ell$ for all $(i, \ell)$.

**Lemma 5.9.** If $\pi$ if a frozen fixed point of $F_{d,k}$, then $\overline{\pi} = (\pi_{i,\ell})_{(i, \ell) \in \mathcal{T}} \in \overline{\mathcal{P}}$.

**Proof.** Let $(i, \ell) \in \mathcal{T}$. By construction, the support of $\pi_{i,\ell}$ is contained in $\Omega_\ell$. Furthermore, Lemma 5.5 implies that

$$
\pi_{i,\ell}(\Omega_\ell) = \frac{1}{kq^*_i} \int \Omega 1_{\mu \in \Omega_\ell} d\pi_i(\mu) = \frac{\pi_i(\Omega_\ell)}{kq^*_i} = \frac{\rho_{i,\ell}(\pi)}{kq^*_{i,\ell}} = 1.
$$

Thus, $\pi_{i,\ell}$ is a probability measure.
Let $\Gamma_{i,\ell}$ be the set of all non-negative integer vectors $\gamma = (\gamma_{i',\ell'})_{(i',\ell') \in T_{i,\ell}}$. For $\gamma \in \Gamma_{i,\ell}$, we let

$$p_{i,\ell}(\gamma) = \prod_{(i',\ell') \in T_{i,\ell}} p_{d'q_{i',\ell'}}(\gamma_{i',\ell'})$$

Moreover, we let $\Omega = \prod_{(i',\ell') \in T_{i,\ell}} \prod_{j \in [\gamma_{i',\ell'}]} \Omega$ and denote its points by $\mu_{\gamma} = (\mu_{i',\ell',j})_{(i',\ell',j) \in T_{i,\ell}, j \in [\gamma_{i',\ell'}]}$. In addition, if $\pi$ is a probability measure on $\Omega$ and $\gamma \in \Gamma_{i,\ell}$, we set

$$\pi_{i,\ell}(\gamma) = \prod_{(i',\ell') \in T_{i,\ell}} \pi_{i',\ell'}.$$

Further, we define for any non-empty set $\ell \subset [k]$ a map

$$B_{\ell} : \bigcup_{\gamma=1}^{\infty} \Omega \rightarrow \Omega, \quad (\mu_1, \ldots, \mu_{\gamma}) \mapsto B_{\ell}[\mu_1, \ldots, \mu_{\gamma}],$$

where

$$B_{\ell}[\mu_1, \ldots, \mu_{\gamma}](h) = \begin{cases} \frac{1_{\text{het}}}{|\ell|} & \text{if } \sum_{h' \in \ell} \prod_{j=1}^{\gamma_{i',\ell'}} 1 - \mu_{j}(h') = 0, \\
\frac{\prod_{j=1}^{\gamma_{i',\ell'}} 1 - \mu_{j}(h)}{\sum_{h' \in \ell} \prod_{j=1}^{\gamma_{i',\ell'}} 1 - \mu_{j}(h')} & \text{if } \sum_{h' \in \ell} \prod_{j=1}^{\gamma_{i',\ell'}} 1 - \mu_{j}(h') > 0. \end{cases}$$

Additionally, to cover the case $\gamma = 0$ we define $B_{\ell}[\emptyset](h) = \frac{1_{\text{het}}}{|\ell|}$. Thus, $B_{\ell}[\emptyset]$ is the uniform distribution on $\ell$.

**Lemma 5.10.** Let $\mathcal{X}$ be the set of all frozen fixed points of $F_{d,k}$. Moreover, let $\tilde{\mathcal{X}}$ be the set of all fixed points of

$$\tilde{F}_{d,k} : \tilde{\mathcal{P}} \rightarrow \tilde{\mathcal{P}}, \quad (\pi_{i,\ell})(i,\ell) \in T \mapsto \left( \sum_{\gamma \in T_{i,\ell}} \int_{\Omega} \delta_{B_{\ell}[\mu_{\gamma}]} \pi_{i,\ell}(\gamma) d\pi_{i,\ell}(\mu_{\gamma}) \right)_{(i,\ell) \in T}.$$

Then the map $\pi \in \mathcal{X} \mapsto \tilde{\pi} = (\pi_{i,\ell})(i,\ell) \in T$ induces a bijection between $\mathcal{X}$ and $\tilde{\mathcal{X}}$.

**Proof.** Suppose that $\pi \in \mathcal{X}$. Let $(i, \ell) \in T$. Then Lemma 5.6 yields

$$\pi_{i,\ell} = \int_{\Omega_{\ell}} \frac{\delta_{\mu}}{kq_{i,\ell}^n} d\pi_{i}(\mu) = \sum_{\gamma \in T_{i,\ell}} \int_{\Omega_{\ell}} \frac{1_{B[\mu_{\gamma}] \in \Omega_{\ell}}} {kq_{i,\ell}^n} d\pi_{i,\ell}(\mu_{\gamma}) = p_{i}(\gamma) d\pi_{i,\gamma}(\mu_{\gamma}).$$

Now let us fix a pair $(i, \ell) \in T$ and $(\gamma, \mu_{\gamma})$. We denote, for $h \neq i$, by $\gamma_h = \gamma_h(\mu_{\gamma})$ the number of occurrences of $\delta_h$ in the tuple $\mu_{\gamma}$. The event $B[\mu_{\gamma}] \in \Omega_{\ell}$ occurs iff

1. for each $h \in [k] \setminus \ell$ there is $j \in [\gamma_{h}]$ such that $\mu_{h,j} = \delta_h$, i.e. $\gamma_h > 0$,
2. for each $h \in \ell \setminus \{i\}$ and all $j \in [\gamma_h]$ we have $\mu_{h,j} = \delta_h$, i.e. $\gamma_h = 0$. 

Thus, Lemma 5.5 implies that
\[
\sum_{\gamma \in \Gamma_i} \int_{\Omega} \frac{1}{kq^*_i,\ell} B_{\gamma}^{[\mu_\gamma]} \frac{1}{kq^*_i,\ell} p_i(\gamma) d\pi_{i,\gamma}(\mu_\gamma) = \frac{1}{k^*_i,\ell} \prod_{h \in [k] \setminus \ell} \mathbb{P} \left[ \mathbb{P}(q^*_h d') > 0 \right] \cdot \prod_{h \in \ell \setminus [i]} \mathbb{P} \left[ \mathbb{P}(q^*_h d') = 0 \right] = 1. \tag{5.22}
\]

Furthermore, given that the event $B_{\gamma}^{[\mu_\gamma]} \in \Omega_\ell$ occurs, the measure $B_{\gamma}^{[\mu_\gamma]}$ is determined by those components $\mu_{i',\ell',j}$ with $(i', \ell') \in T_i,\ell$ only. Thus, with $\hat{\gamma} = (\hat{\gamma}_{i',\ell'})_{(i', \ell') \in T_i,\ell}$ and $\mu_{\hat{\gamma}} = (\mu_{i',\ell',j})_{(i', \ell') \in T_i,\ell, j \in [\ell_i,\ell']}$ we obtain from (5.21) and (5.22)
\[
\pi_{i,\ell} = \sum_{\gamma \in \Gamma_i,\ell} \int_{\Omega_{\hat{\gamma}}} \delta_{B_i^{[\mu_{i,\ell,j}]}(\gamma)} d\pi_{i,\ell,\gamma}(\mu_{\hat{\gamma}}).
\]

Thus, if $\pi$ is a frozen fixed point of $F_{d,k}$, then $\pi$ is a fixed point of $\hat{F}_{d,k}$.

Conversely, if $\pi = (\pi_{i,\ell})$ is a fixed point of $\hat{F}_{d,k}$, then the measure $\pi$ defined by
\[
\pi_{i,\ell} = \sum_{\gamma \in \Gamma_i,\ell} \int_{\Omega_{\hat{\gamma}}} \delta_{B_i^{[\mu_{i,\ell,j}]}(\gamma)} d\pi_{i,\ell,\gamma}(\mu_{\hat{\gamma}})
\]
is easily verified to be a fixed point of $F_{d,k}$. Moreover, for $i \in [k]$, $\rho_i(\pi) = q^*_i,\ell = q^*/k \geq 2/(3k)$ and $\pi$ is thus a frozen fixed point of $F_{d,k}$.

**Corollary 5.4.** The distribution $\pi_{d,k,q^*}$ is a fixed point of $F_{d,k}$.

**Proof.** To unclutter the notation we write $\pi = \pi_{d,k,q^*}$. Moreover, we let $T = T_{d,k,q^*}$; by Lemma 5.1 we may always assume that $T$ is a finite tree. Recall that $\pi$ is the distribution of $\mu_T$, which is the distribution of the color of the root under a random legal coloring of $T$. In light of Lemma 5.10 it suffices to show that $\pi = (\pi_{i,\ell})$ is a fixed point of $\hat{F}_{d,k}$. Thus, we need to show that for all $(i, \ell) \in T$,
\[
\pi_{i,\ell} = \sum_{\gamma \in \Gamma_i,\ell} \int_{\Omega_{\hat{\gamma}}} \delta_{B_i^{[\mu_{i,\ell,j}]}(\gamma)} \prod_{(i', \ell') \in T_i,\ell} p_{d} d_{i',\ell'}^{q^*_{i',\ell'}}(\hat{\gamma}_{i',\ell'}) \prod_{j=1}^{\ell} \pi_{i',\ell'}(\mu_{i',\ell',j}). \tag{5.23}
\]

Let us denote by $T_{i,\ell}$ the random tree $T$ given that the root has type $(i, \ell)$. We claim that $\pi_{i,\ell}$ is the distribution of $\mu_{T_{i,\ell}}$. Indeed, let $\ell' \subset [k]$. If the root $v_0$ of $T$ has type $(i, \ell)$ for some $i \in \ell$, then the support of the measure $\mu_T$ is contained in $\ell$ (because under any legal coloring, $v_0$ receives a color from $\ell$). Moreover, all children of $v_0$ have types in $T_{i,\ell}$, and if $(i', \ell') \in T_{i,\ell}$, then $|\ell'| \geq 2$. Hence, inductively we see that if $v_0$ has type $(i, \ell)$, then for any color $h \in \ell$ there is a legal coloring under which $v_0$ receives color $h$. Consequently, the support of $\mu_T$ is precisely $\ell$. Furthermore, the distribution $\mu_T$ is invariant under the following operation: obtain a random tree $T'$ by choosing a legal color $\tau$ of $T$ randomly and then changing the types $\vartheta'(v) = (i_v, \ell_v)$ of the vertices to $\vartheta'(v) = (\tau(i_v), \ell_v)$; this is because the trees $T$ and $T'$ have the same set of legal colorings. These observations imply that for any measurable set $A$ we have
\[
\mathbb{P} \left[ \mu_T \in A \mid \vartheta(v_0) = (i, \ell) \right] = \frac{\mathbb{P} \left[ \mu_T \in A, \vartheta(v_0) = (i, \ell) \right]}{\mathbb{P} \left[ \vartheta(v_0) = (i, \ell) \right]}. \]


The Condensation Phase Transition in Random Graph Coloring

TR1 Let $\pi \in \mathcal{P}$.

Proof. As before, we let $\pi \in \mathcal{P}$. We define a distribution $\tilde{\pi} = (\pi_{i,\ell}) \in \tilde{\mathcal{P}}$. We define a distribution $\tilde{\pi} = (\pi_{i,\ell}) \in \tilde{\mathcal{P}}$ by means of the following experiment. Let $(i, \ell) \in T$. Let $v_0$ denote the root of $T_{i,\ell}$ and let $\theta(v)$ signify the type of each vertex $v$.

TR1 Let $T_{i,\ell}$ be the tree obtained from $T_{i,\ell}$ by deleting all vertices at distance greater than $t$ from $v_0$.

TR2 Let $V_t$ be the set of all vertices at distance exactly $t$ from $v_0$. For each $v \in V_t$ independently, choose $\mu_v \in \Omega$ from the distribution $\pi_{i,\ell}(v)$.

TR3 Let $\mu_{i,\ell}$ be the distribution of the color of $v_0$ under a random coloring $\tau$ chosen as follows.

- Independently for each vertex $v \in V_t$ choose a color $\tau(t)(v)$ from the distribution $\mu_v$.
- Let $\tau$ be a uniformly random legal coloring of $T_{i,\ell}$ such that $\tau(v) = \tau(t)(v)$ for all $v \in V_t$; if there is no such coloring, discard the experiment.

Step TR3 of the above experiment yields a distribution $\mu_{i,\ell} \in \Omega$. Clearly $\mu_{i,\ell}$ is determined by the random choices in steps TR1–TR2. Thus, let we let $\pi_{i,\ell}$ be the distribution of $\mu_{i,\ell}$ with respect to TR1–TR2.

We now claim that for any integer $t \geq 0$ the following is true.

If $\tilde{\pi}$ is a fixed point of $\tilde{\mathcal{F}}_{d,k}^*$, then $\tilde{\pi} = \pi_t$.

(5.25)
The proof of (5.25) is by induction on $t$. It is immediate from the construction that $\pi_{i,\ell,0} = \pi_{i,\ell}$ for all $(i, \ell) \in T$. Thus, assume that $t \geq 1$. By induction, it suffices to show that $\bar{\pi}_t = \bar{\pi}_{t-1}$. To this end, let us condition on the random tree $T_{i,\ell,t-1}$. Consider a vertex $v \in V_{t-1}$ of type $\bar{\vartheta}(v) = (\bar{i}, \bar{\ell}, v)$. We obtain the random tree $T_{i,\ell,t}$ from $T_{i,\ell,t-1}$ by attaching to each such $v \in V_{t-1}$ a random number $\gamma_{i',\ell',v} = \text{Po}(d'q^*_i,\ell')$ of children of each type $(i', \ell') \in \bar{T}_{i',\ell'}$ where, of course, the random variables $\gamma_{i',\ell',v}$ are mutually independent. Further, in step TR2 of the above experiment we choose $\mu_{i',\ell',v,j} \in \Omega_{i',\ell'}$ independently from $\pi_{i',\ell'}$ for each $v \in V_{t-1}, (i', \ell') \in \bar{T}_{i',\ell}$ and $j = 1, \ldots, \gamma_{i',\ell',v}$.

Given the distributions $\mu_{i',\ell',v,j}$, suppose that we choose a legal coloring $\tau_v$ of the sub-tree consisting of $v \in V_{t-1}$ and its children only from the following distribution.

- Independently choose the colors $\tau_v(u_{i',\ell',j})$ of the children $u_{i',\ell',j}$ of $v$ of type $(i', \ell')$ from $\mu_{i',\ell',v,j}$.
- Choose a color $\tau_v(v)$ for $v$ uniformly from the set of all colors $h \in \ell$ that are not already assigned to a child of $v$ if possible.

Let $\mu_v$ denote the distribution of the color $\tau_v(v)$. Then by construction,

$$
\mu_v = \mathcal{B}_\ell([\mu_{i',\ell',v,j}(i', \ell') \in \bar{T}_{i',\ell}, j \in [\gamma_{i',\ell',v}]]).
$$

Hence, the distribution of $\mu_v$ with respect to the choice of the numbers $\gamma_{i',\ell',v}$ and the distributions $\mu_{i',\ell',v,j}$ is given by

$$
\sum_{\mathcal{Y} \in \mathcal{T}_{i,\ell}} \int_{\Omega_{i',\ell',v}} \delta_{\mathcal{B}_\ell([\mu_{i',\ell',v,j}])} \prod_{(i', \ell') \in \bar{T}_{i',\ell}} pd' q^*_i,\ell' \gamma_{i',\ell',v} \bigotimes_{j=1}^{\gamma_{i',\ell',v}} d\mu_{i',\ell',v,j}(\mu_{i',\ell',v,j}) = \pi_{i,\ell},
$$

because $\bar{\pi}$ is a fixed point of $\bar{\mathcal{F}}_{d,k}$. Therefore, the experiment of first choosing $T_{i,\ell,t}$, then choosing distributions $\mu_u$ independently from $\pi_{\bar{\theta}(u)}$ for the vertices at distance $t$, and then choosing a random legal coloring $\tau$ as in TR3 is equivalent to performing the same experiment with $t-1$ instead. Hence, $\bar{\pi}_t = \bar{\pi}_{t-1}$.

To complete the proof, assume that $\bar{\pi}, \bar{\pi}'$ are fixed points of $\bar{\mathcal{F}}_{d,k}$. Then for any integer $t \geq 1$ we have $\bar{\pi} = \bar{\pi}_t, \bar{\pi}' = \bar{\pi}'_t$. Furthermore, as $\bar{\pi}_t, \bar{\pi}'_t$ result from the experiment TR1–TR3, whose first step TR1 can be coupled, we see that for any $(i, \ell) \in T$,

$$
\|\pi_{i,\ell} - \pi'_{i,\ell}\|_{\text{TV}} = \|\pi_{i,\ell,t} - \pi'_{i,\ell,t}\|_{\text{TV}} \leq 2 \mathbb{P}[|T_{i,\ell}| \geq t].
$$

(5.26)

Because Lemma 5.1 shows that $T$ results from a sub-critical branching process, we have

$$
\lim_{t \to \infty} \mathbb{P}[|T_{i,\ell}| \geq t] = 0
$$

for any $(i, \ell) \in T$. Consequently, (5.26) shows that $\bar{\pi} = \bar{\pi}'$.

Finally, Lemma 5.8 follows directly from Lemma 5.10, Corollary 5.4 and Lemma 5.11.

5.4. The number of legal colorings. The final step of the proof of Proposition 2.3 is to relate $\phi_{d,k}(\pi_{d,k,q^*})$ to the number of legal colorings of $T_{d,k,q^*}$. The starting point for this is a formula for the (logarithm of the) number of legal colorings of a decorated tree $T, \bar{\vartheta}$. To write this formula down, we recall the map $\mathcal{B}_\ell$ from (5.20). Moreover, suppose that $\ell \subset [k]$ and $\mu_1, \ldots, \mu_\gamma \in \Omega$ are such that

$$
\exists h \in \ell \forall j \in [\gamma] : \mu_j(h) < 1.
$$

(5.27)
Then we let
\[
\phi_\ell(\mu_1, \ldots, \mu_\gamma) = \phi_\ell^v(\mu_1, \ldots, \mu_\gamma) - \frac{1}{2} \phi_\ell^e(\mu_1, \ldots, \mu_\gamma),
\]
where
\[
\phi_\ell^v(\mu_1, \ldots, \mu_\gamma) = \ln \sum_{h \in \ell} \prod_{j=1}^\gamma \left( 1 - \mu_j(h) \right),
\]
\[
\phi_\ell^e(\mu_1, \ldots, \mu_\gamma) = \sum_{j=1}^\gamma \ln \left[ 1 - \sum_{h \in \ell} \mu_j(h) B_\ell[\mu_1, \ldots, \mu_{j-1}, \mu_{j+1}, \ldots, \mu_\gamma](h) \right];
\]
the condition (5.27) ensures that these quantities are well-defined (i.e., the argument of the logarithm is positive in both instances). Additionally, to cover the case \( \gamma = 0 \) we set \( \phi_\ell(\emptyset) = \ln |\ell| \).

Further, suppose that \( T, \vartheta, v \) is a rooted decorated tree that has at least one legal coloring \( \sigma \). Let \( v_1, \ldots, v_\gamma \) be the neighbors of the root vertex \( v \) and suppose that \( \vartheta(v_j) = (i, \ell) \) and \( \vartheta(v_j) = (i_j, \ell_j) \) for \( j = 1, \ldots, \gamma \). If we remove the root \( v \) from \( T \), then each of the vertices \( v_1, \ldots, v_\gamma \) lies in a connected component \( T_i \) of the resulting forest. By considering the restrictions \( \vartheta_i \) of \( \vartheta \) to the vertex set of \( T_i \), we obtain decorated trees \( T_i, \vartheta_i \). Recall that \( \mu_{T_j, \vartheta_j, v_j} \) denotes the distribution of the color of the root in a random legal coloring of \( T_j, \vartheta_j, v_j \). Since \( \sigma \) is a legal coloring, for \( h = \sigma(v) \) for all \( j \in [\gamma] \) we have \( \mu_{T_j, \vartheta_j, v_j}(h) \leq 1 \). Thus, we can define
\[
\phi(T, \vartheta, v) = \phi_\ell(\mu_{T_1, \vartheta_1, v_1}, \ldots, \mu_{T_\gamma, \vartheta_\gamma, v_\gamma}).
\]

**Fact 5.1.** Let \( T, \vartheta \) be a decorated tree such that \( \mathcal{Z}(T, \vartheta) \geq 1 \). Then \( \ln \mathcal{Z}(T, \vartheta) = \sum_{v \in V(T)} \phi(T, \vartheta, v) \).

**Proof.** This follows from [11, Proposition 3.7]. More specifically, let \( (i_v, \ell_v) = \vartheta(v) \) be the type of vertex \( v \). In the terminology of [11] (and of the physicists “cavity method”), \( \phi(T, \vartheta, v) \) is the Bethe free entropy of the Boltzmann distribution
\[
v : [k]^{V(T)} \to [0, 1], \quad v(\tau) = \frac{1}{\mathcal{Z}(T, \vartheta)} \prod_{v \in V(T)} 1_{\tau(v) \in \ell_v} \cdot \prod_{e = [u, w] \in E(T)} 1_{\tau(u) \neq \tau(w)}.\]
Thus, \( v \) is simply the uniform distribution over legal \( k \)-colorings of \( T, \vartheta \), and \( \mathcal{Z}(T, \vartheta) \) is its partition function.

Let \( T \) denote the random rooted decorated tree \( T_{d,k,q^\ast} \). Moreover, for \( (i, \ell) \in \mathcal{T} \) we let \( T_{i,\ell} \) denote the random tree \( T \) given that the root has type \( (i, \ell) \). The starting point of the proof is the following key observation. Furthermore, if \( (T, \vartheta, v) \) is a rooted decorated tree, then we let \( (T, \vartheta, v)^* \) signify the isomorphism class of the random rooted decorated tree \( (T, \vartheta, u) \) obtained from \( (T, \vartheta, v) \) by choosing a vertex \( u \) of \( T \) uniformly at random and rooting the tree at \( u \). In other words, \( (T, \vartheta, v)^* \) is obtained by re-rooting \( (T, \vartheta, v) \) at a random vertex.

**Lemma 5.12.** Let \( T^* \) be the random rooted decorated tree obtained by re-rooting \( T \) at a random vertex. Then the distribution of \( T^* \) coincides with the distribution of \( T \).

**Proof.** This follows from the general fact that Galton–Watson trees are unimodular in the sense of [5].
Corollary 5.5. We have \( E \left[ \ln \frac{Z(T)}{|T|} \right] = E[\phi(T)]. \)

Proof. Letting \((T, \vartheta, v)\) range over rooted decorated trees, we find

\[
E \left[ \ln \frac{Z(T)}{|T|} \right] = \sum_{(T, \vartheta, v)} P[T \cong (T, \vartheta, v)] \frac{\ln Z(T, \vartheta, v)}{|V(T)|}
\]

\[
\begin{align*}
&= \sum_{(T, \vartheta, v)} \sum_{u \in V(T)} P[T \cong (T, \vartheta, u)] \frac{\ln Z(T, \vartheta, u)}{|V(T)|} \quad \text{[by Fact 5.1]} \\
&= \sum_{(T, \vartheta, v)} \sum_{u \in V(T)} P[T \cong (T, \vartheta, u)] \phi(T, \vartheta, u) \quad \text{[by Lemma 5.12]} \\
&= \sum_{(T, \vartheta, v)} P[T \cong (T, \vartheta, v)] \phi(T, \vartheta, v) = E[\phi(T)],
\end{align*}
\]

as claimed.

Lemma 5.13. We have

\[
E[\phi(T_{i,\ell})] = \sum_{\gamma \in \Gamma_{i,\ell}} p_{i,\ell}(\gamma) \int_{\Omega^r} \phi_v^\ell(\mu_\gamma) d\pi_\gamma(\mu_\gamma)
\]

\[
- \sum_{(i', \ell') \in T_{i,\ell}} \frac{q_{i', \ell'}d^d}{2} \int_{\Omega^2} \ln \left[ 1 - \sum_{h=1}^k \mu(h) \mu(h) \right] d\pi_{i',\ell'}(\mu) \otimes \pi_{i,\ell}(\mu).
\]

Proof. Writing \( \pi = \pi_{d,k,q^*} \) for the distribution of \( \mu_T \), we know from Corollary 5.4 that \( \pi_{i,\ell} \) is the distribution of \( \mu_{T_{i,\ell}} \) for any type \((i, \ell)\). Furthermore, the distribution of \( T_{i,\ell} \) can be described by the following recurrence: there is a root \( v_0 \) of type \((i, \ell)\), to which we attach for each \((i', \ell') \in T_{i,\ell}\) independently a number \( \gamma_{i', \ell'} = \text{Po}(d'q_{i', \ell'}) \) of trees \((T_{i', \ell', j})_{j=1,...,\gamma_{i', \ell'}}\) that are chosen independently from the distribution \( T_{i', \ell'} \). By independence, the distribution of the color of the root of each \( T_{i', \ell', j} \) is just an independent sample from the distribution \( \pi_{i', \ell'} \). Therefore, we obtain the expansion

\[
E[\phi(T_{i,\ell})] = \sum_{\gamma \in \Gamma_{i,\ell}} \int_{\Omega^r} \phi_v^\ell(\mu_\gamma) p_{i,\ell}(\gamma) d\pi_\gamma(\mu_\gamma).
\]

Substituting in the definition of \( \phi_v^\ell \), we obtain

\[
E[\phi(T_{i,\ell})] = I_{i,\ell} - \frac{1}{2} J_{i,\ell}, \quad \text{where}
\]

\[
I_{i,\ell} = \sum_{\gamma \in \Gamma_{i,\ell}} p_{i,\ell}(\gamma) \int_{\Omega^r} \phi_v^\ell(\mu_\gamma) d\pi_\gamma(\mu_\gamma),
\]

\[
J_{i,\ell} = \sum_{\gamma \in \Gamma_{i,\ell}} p_{i,\ell}(\gamma) \int_{\Omega^r} \phi_v^\ell(\mu_\gamma) d\pi_\gamma(\mu_\gamma).
\]

Further, by the definition of \( \phi_v^\ell \) we have

\[
J_{i,\ell} = \sum_{\gamma \in \Gamma_{i,\ell}} p_{i,\ell}(\gamma)
\]
\[
\sum_{(i, \ell) \in T_{i, \ell}} \sum_{j=1}^{\gamma_{i, \ell}} \int_{\Omega_Y} \ln \left[ 1 - \sum_{h \in \ell} \mu_{i, \ell, j}^h(h) \mathcal{B}[(\mu_{i', \ell', j})_{(i', \ell', j) \neq (i, \ell, j)}](h) \right] d\pi_Y(\mu_Y)
\]

\[
= \sum_{(i, \ell) \in T_{i, \ell}} \sum_{g \geq 1} p_{q_{i, \ell}}^g(g) \sum_{j=1}^{\gamma_{i, \ell}} p_{i, \ell}(\gamma) \mathbf{1}_{\gamma_{i, \ell} = g}
\]

\[
\cdot \int_{\Omega \times \Omega_Y} \ln \left[ 1 - \sum_{h \in \ell} \mu(h) \mathcal{B}[(\mu_{i', \ell', j})_{(i', \ell', j) \neq (i, \ell, j)}](h) \right] d\pi_{i, \ell}(\mu) \otimes d\pi_Y(\mu_Y)
\]

\[
= \sum_{(i, \ell) \in T_{i, \ell}} \sum_{g \geq 1} p_{q_{i, \ell}}^g(g) \sum_{j=1}^{\gamma_{i, \ell}} p_{i, \ell}(\gamma) \mathbf{1}_{\gamma_{i, \ell} = g-1}
\]

\[
\cdot \int_{\Omega \times \Omega_Y} \ln \left[ 1 - \sum_{h \in \ell} \mu(h) \mathcal{B}[\mu_Y](h) \right] d\pi_{i, \ell}(\mu) \otimes d\pi_Y(\mu_Y).
\]

To simplify this, we use the following elementary relation: if \( X : \mathbb{Z} \to \mathbb{R}_{\geq 0} \) is a function and \( g \) is a Poisson random variable, then \( \mathbb{E}[\mathbf{1}_{g \geq 1} X(g - 1)] = \mathbb{E}[g] \mathbb{E}[X(g)] \). Applying this observation to

\[
X(g) = \sum_{\gamma \in \Gamma_{i, \ell}} p_{i, \ell}(\gamma) \mathbf{1}_{\gamma_{i, \ell} = g-1} \int_{\Omega \times \Omega_Y} \ln \left[ 1 - \sum_{h \in \ell} \mu(h) \mathcal{B}[\mu_Y](h) \right] d\pi_{i, \ell}(\mu) \otimes d\pi_Y(\mu_Y),
\]

we obtain

\[
J_{i, \ell} = \sum_{(i, \ell) \in T_{i, \ell}} q_{i, \ell}^d \sum_{\gamma \in \Gamma_{i, \ell}} p_{i, \ell}(\gamma)
\]

\[
\cdot \int_{\Omega \times \Omega_Y} \ln \left[ 1 - \sum_{h \in \ell} \mu_{i, \ell}(h) \mathcal{B}[\mu_Y](h) \right] d\pi_{i, \ell}(\mu) \otimes d\pi_Y(\mu_Y).
\]

Now, since \( \pi \) is a fixed point of \( F_{d,k} \), the distribution of the measure \( \mathcal{B}[\mu_Y] \) is just \( \pi_{i, \ell} \). Hence,

\[
J_{i, \ell} = \sum_{(i, \ell) \in T_{i, \ell}} q_{i, \ell}^d \int_{\Omega^2} \ln \left[ 1 - \sum_{h \in \ell} \hat{\mu}(h) \mu(h) \right] d\pi_{i, \ell}(\mu) \otimes d\hat{\pi}_{i, \ell}(\hat{\mu}).
\]

Thus, we obtain the assertion.

**Lemma 5.14.** We have \( \mathbb{E}[\phi(T_{d,k,q^*})] = \phi_{d,k}(\pi_{d,k,q^*}) \).

**Proof.** Summing over all \((i, \ell) \in T\), we obtain from Lemma 5.13 that

\[
\mathbb{E}[\phi(T)] = I - \frac{1}{2} J, \quad \text{where}
\]

\[
I = \sum_{(i, \ell) \in T} q_{i, \ell}^* \sum_{\gamma \in \Gamma_{i, \ell}} p_{i, \ell}(\gamma) \int_{\Omega_Y} \phi_{i, \ell}(\mu_Y) d\pi_Y(\mu_Y),
\]

\[
J = \sum_{(i, \ell) \in T} q_{i, \ell}^d \sum_{\gamma \in \Gamma_{i, \ell}} p_{i, \ell}(\gamma) \int_{\Omega^2} \ln \left[ 1 - \sum_{h \in \ell} \hat{\mu}(h) \mu(h) \right] d\pi_{i, \ell}(\mu) \otimes d\hat{\pi}_{i, \ell}(\hat{\mu}).
\]
\[ J = d' \sum_{(i, \ell) \in T} \sum_{(j, \ell') \in T_{i, \ell}} q^{*}_{i, \ell} q^{*}_{i, \ell'} \int_{\Omega^2} \ln \left[ 1 - \sum_{h=1}^{k} \hat{\mu}(h) \mu(h) \right] d\pi_{i, \ell}(\mu) \otimes \pi_{i, \ell}'(\hat{\mu}). \]

Recalling that \( d\pi_{i, \ell}(\mu) = \frac{1_{\mu \in \Omega}}{k q_{i, \ell}} d\pi_{i}(\mu) \) and \( d\pi_{i, \ell}'(\hat{\mu}) = \frac{1_{\hat{\mu} \in \Omega}}{k q_{i, \ell}} d\pi_{i}'(\hat{\mu}) \), we get

\[ J = \frac{d'}{k^2} \sum_{(i, \ell) \in T} \sum_{(j, \ell') \in T_{i, \ell}} \int_{\Omega^2} \ln \left[ 1 - \sum_{h=1}^{k} \hat{\mu}(h) \mu(h) \right] 1_{\mu \in \Omega_{i, \ell}} 1_{\hat{\mu} \in \Omega_{i, \ell}'} d\pi_{i}(\mu) \otimes \pi_{i, \ell}'(\hat{\mu}) \]

\[ = \frac{d}{k(k-1)} \sum_{i, j \in [k] ; i \neq j} \int_{\Omega^2} \sum_{\ell, \ell'} \ln \left[ 1 - \sum_{h=1}^{k} \hat{\mu}(h) \mu(h) \right] \]

\[ \cdot 1_{\mu \in \Omega_{i, \ell}} 1_{\hat{\mu} \in \Omega_{i, \ell}'} d\pi_{i}(\mu) \otimes \pi_{i, \ell}'(\hat{\mu}) \]

\[ = \frac{d}{k(k-1)} \sum_{i, j \in [k] ; i \neq j} \int_{\Omega^2} \ln \left[ 1 - \sum_{h=1}^{k} \hat{\mu}(h) \mu(h) \right] d\pi_{i}(\mu) \otimes \pi_{i, \ell}'(\hat{\mu}) = \phi^v_{d, k}(\pi). \]

It finally remains to simplify the expression for \( I \). To do it, we introduce \( T_i = \{(i', \ell') \in T, i' \neq i\} \). We let \( \overline{T}_i \) be the set of non-negative vectors \( \overline{\gamma} = (\overline{\gamma}_{i', \ell'}) (i', \ell') \in T_i \).

Moreover, we let \( \overline{\Omega} = \prod_{(i', \ell') \in \overline{T}_i} \prod_{j \in \overline{\gamma}_{i', \ell'}} \Omega \) and denote its points by \( \mu_{\overline{\gamma}} = (\mu_{i', \ell', j}) (i', \ell', j) \in \overline{T}_i \). We note that if \( \gamma \in T_{i, \ell} \) and \( \overline{\gamma} \in \overline{T}_i \) are such that:

(a) \( \forall i' \in \ell \setminus \{i\}, \overline{\gamma}_{i', \{i]\}} = 0, \)
(b) \( \forall i' \in [k] \setminus \ell, \overline{\gamma}_{i', \{i\}} > 0 \),
(c) \( \forall (i', \ell') \in \overline{T}_{i, \ell}, \overline{\gamma}_{i', \ell'} = \overline{\gamma}_{i', \ell}, \)

and that \( \mu_{\gamma}, \overline{\mu}_{\overline{\gamma}} \) satisfy

(d) \( \forall (i', \ell') \in \overline{T}_{i, \ell}, \forall j \in [\gamma_{i', \ell'}], \mu_{i', \ell', j} = \overline{\mu}_{i', \ell', j}, \)
(e) \( \forall (i', \ell') \in \overline{T}_{i, \ell}, \forall j \in [\overline{\gamma}_{i', \ell'}], \overline{\mu}_{i', \ell', j} \in \Omega_{\ell'}. \)

then

\[ \prod_{(i', \ell') \in \overline{T}_{i, \ell}} \prod_{j \in [\overline{\gamma}_{i', \ell'}]} 1 - \overline{\mu}_{i', \ell', j}(h) = \begin{cases} 0 & \text{if } h \notin \ell, \\ \prod_{(i', \ell') \in \overline{T}_{i, \ell}} \prod_{j \in [\overline{\gamma}_{i', \ell'}]} 1 - \mu_{i', \ell', j}(h) & \text{if } h \in \ell. \end{cases} \]

Consequently

\[ \phi^v_{\ell}(\mu_{\overline{\gamma}}) = \ln \left[ \sum_{h \in [k]} \prod_{(i', \ell') \in \overline{T}_{i, \ell}} \prod_{j \in [\overline{\gamma}_{i', \ell'}]} 1 - \overline{\mu}_{i', \ell', j}(h) \right]. \] (5.28)

Moreover, choosing the \( \overline{\gamma}_{i', \ell'} \) from a Poisson distribution of parameter \( q^{*}_{i, \ell} d' \), the event “(a) and (b)” happens with probability exactly \( k q^{*}_{i, \ell} \). This allows us to write:

\[ I = \sum_{(i, \ell) \in T} q^{*}_{i, \ell} \sum_{\gamma \in T_{i, \ell}} \prod_{(i', \ell') \in \overline{T}_{i, \ell}} p_{q^{*}_{i, \ell}} d'(\gamma_{i', \ell'}) \]
\[ \int_{\Omega^t} \phi^v_\ell (\mu_{\mathcal{T}}) \bigotimes_i \bigotimes_j \pi_i^\ell \cdot \pi^\ell_j (\mu_i^\ell, \ell, j) \]

\[ = \frac{1}{k} \sum_{(i, \ell) \in \mathcal{T}_i} \sum_{\gamma_i^\ell \in \mathcal{T}_i (i^\ell, \ell^\gamma_i) \in \mathcal{T}_i} \prod_{i^\ell \in \ell^\gamma_i} p_{\gamma_i^\ell}^* \mu_{\mathcal{T}}^\ell (\gamma_i^\ell) \bigotimes_i \bigotimes_j \prod_{i^\ell \in \ell^\gamma_i} 1_{\gamma_i^\ell (\gamma_i^\ell) = 0} \prod_{i^\ell \in \ell^\gamma_i} 1_{\gamma_i^\ell (\gamma_i^\ell) > 0} \]

\[ \cdot \int_{\Omega^t} \ln \left[ \sum_{h \in [k]} \prod_{(i^\ell, \ell^\gamma_i) \in \mathcal{T}_i} 1 - \mu_i^{\ell, j} (\gamma_i^\ell) \right] \bigotimes_i \bigotimes_j \prod_{(i^\ell, \ell^\gamma_i) \in \mathcal{T}_i} \frac{1}{k} \pi_i^\ell (\mu_i^\ell, \ell, j) \]

\[ = \frac{1}{k} \sum_{i \in [k]} \sum_{\gamma_i \in \mathcal{T}_i} \prod_{i^\ell \in \ell^\gamma_i} p_{\gamma_i^\ell} (\gamma_i^\ell) \]

\[ \cdot \int_{\Omega^t} \ln \left[ \sum_{h \in [k]} \prod_{i^\ell \in \ell^\gamma_i} 1 - \mu_i^{\ell, j} (\gamma_i^\ell) \right] \bigotimes_i \bigotimes_j \prod_{(i^\ell, \ell^\gamma_i) \in \mathcal{T}_i} \frac{1}{k} \pi_i^\ell (\mu_i^\ell, \ell, j) \]

\[ = \frac{1}{k} \sum_{i \in [k]} \sum_{\gamma_i \in \mathcal{T}_i} \prod_{i^\ell \in \ell^\gamma_i} p_{\gamma_i^\ell} (\gamma_i^\ell) \]

\[ \cdot \int_{\Omega^t} \ln \left[ \sum_{h \in [k]} \prod_{i^\ell \in \ell^\gamma_i} 1 - \mu_i^{\ell, j} (\gamma_i^\ell) \right] \bigotimes_i \bigotimes_j \prod_{(i^\ell, \ell^\gamma_i) \in \mathcal{T}_i} \frac{1}{k} \pi_i^\ell (\mu_i^\ell, \ell, j). \]

We used (5.28) to go from the first to the second line, and summed over \( \ell \geq i \) to go from the second to the third. Re-indexing the vector \( \bar{\mu}_{\mathcal{T}} \) in a vector \( \mu_{\mathcal{T}}, \mathcal{T} \in \Gamma_i \) (with \( \gamma_i^\ell = \sum_{(i^\ell, \ell^\gamma_i) \in \mathcal{T}} \gamma_i^\ell, \ell^\gamma_i \)), we obtain with Lemma 5.5:

\[ I = \frac{1}{k} \sum_{i \in [k]} \sum_{\gamma_i \in \mathcal{T}_i} \prod_{i^\ell \in \ell^\gamma_i} p_{\gamma_i^\ell} (\gamma_i^\ell) \bigotimes_i \bigotimes_j \prod_{i^\ell \in \ell^\gamma_i} 1 - \mu_i^{\ell, j} (\gamma_i^\ell) \bigotimes_i \bigotimes_j \prod_{(i^\ell, \ell^\gamma_i) \in \mathcal{T}_i} \frac{1}{k} \pi_i^\ell (\mu_i^\ell, \ell, j) \]

\[ = \frac{1}{k} \sum_{i \in [k]} \sum_{\gamma_i, \ldots, \gamma_h = 0} \prod_{i^\ell \in \ell^\gamma_i} p_{\gamma_i^\ell} (\gamma_i^\ell) \bigotimes_i \bigotimes_j \prod_{(i^\ell, \ell^\gamma_i) \in \mathcal{T}_i} \frac{1}{k} \pi_i^\ell (\mu_i^\ell, \ell, j). \]

\[ \text{Proof of Proposition 2.3.} \] The first assertion is immediate from Lemma 5.1, while the second assertion follows from Lemma 5.8. The third claim follows by combining Corollary 5.5 with Lemma 5.14. With respect to the last assertion, we observe that for \( d = (2k - 1) \ln k - 2 \ln 2 + o_k(1) \) we have

\[ \ln k + \frac{d}{2} \ln (1 - 1/k) = \frac{\ln 2 + o_k(1)}{k}. \]

Moreover, as \( q^* = 1 - 1/k + o_k(1/k) \) by Lemma 5.1, one checks easily that

\[ \mathbb{E} \left[ \frac{\ln Z(T_{d,k,q^*})}{|T_{d,k,q^*}|} \right] = \frac{\ln 2 + o_k(1)}{k}. \] (5.29)
Further, by Lemma 5.1

$$\frac{\partial}{\partial d} \ln \left( \frac{\mathcal{Z}(T_{d,k}(q^*))}{|T_{d,k}(q^*)|} \right) = \tilde{O}_k(k^{-2})$$

while

$$\frac{\partial}{\partial d} \ln k + \frac{d}{2} \ln(1 - 1/k) = \Omega_k(1/k).$$

(5.30)

Combining (5.29) and (5.30) and using the third part of Proposition 2.3, we conclude that $\Sigma_k$ has a unique zero $d_k, \text{cond}$, as claimed.

6. The Cluster Size

The objective in this section is to prove Proposition 2.4. For technical reasons, we consider a variant of the “planted model” $G'(n', p', \sigma)$ in which the number of vertices is not exactly $n$ but $n - o(n)$. This is necessary because we are going to perform inductive arguments in which small parts of the random graph get removed. Thus, let $\eta = \eta(n) = o(n)$ be a non-negative integer sequence. Throughout the section, we write $n' = n - \eta(n)$. Moreover, we let $G = G(n', p', \sigma)$, where $p' = d'/n'$ with $d' = kd/(k - 1)$ as in (2.3).

By a slight abuse of notation we do not distinguish between $\sigma$ and its restriction to the vertices in $[n']$. Unless specified otherwise, all statements in this section are understood to hold for any sequence $\eta = o(n)$.

6.1. Preliminaries. Assume that $G = (V, E)$, let $\sigma$ be a $k$-coloring of $G$, let $v \in V$ and let $\omega \geq 1$ be an integer. We write $\partial^\omega_G(v)$ for the subgraph of $G$ consisting of all vertices at distance at most $\omega$ from $v$. Moreover, $|\partial^\omega_G(v)|$ signifies the number of vertices of $\partial^\omega_G(v)$. Where the reference to $G$ is clear from the context, we omit it. We begin with the following standard fact about the random graph $G$.

Lemma 6.1. Let $\omega = 10[\ln \ln \ln n]$.

1. With probability $1 - \exp(-\Omega(\ln^2 n))$ the random graph $G$ is such that $|\partial^\omega_G(v)| \leq n^{0.01}$ for all vertices $v$.
2. A.a.s. all but $o(n)$ vertices $v$ of $G$ are such that $\partial^\omega_G(v)$ is acyclic.

In addition, we need to know that the “local structure” of the random graph $G$ endowed with the coloring $\sigma$ enjoys the following concentration property.

Lemma 6.2. Let $S$ be a set of triples $(G_0, \sigma_0, v_0)$ such that $G_0$ is a graph, $\sigma_0$ is a $k$-coloring of $G_0$, and $v_0$ is a vertex of $G_0$. Let $\omega = 10[\ln \ln \ln n]$ and define a random variable $S_v = S_v(G, \sigma)$ by letting

$$S_v = 1_{(\partial^\omega_G(v), \sigma|_{\partial^\omega_G(v)}) \in S}.$$ 

Further, let $S = \sum_v S_v$. Then $S = \mathbb{E}[S] + o(n)$ a.a.s.

The proof of Lemma 6.2 is based on standard arguments. The full details can be found in Sect. 6.5.
6.2. Warning propagation. The goal in this section is to prove Proposition 2.4, i.e., to determine the cluster size $|\mathcal{C}(G, \sigma)|$. A key step in this endeavor will be to determine the sets

$$\ell(v) = \{\tau(v) : \tau \in \mathcal{C}(G, \sigma)\}$$

of colors that vertex $v$ may take under a $k$-coloring in $\mathcal{C}(G, \sigma)$. In particular, we called a vertex frozen in $\mathcal{C}(G, \sigma)$ if $\ell(v) = \{\sigma(v)\}$. To establish Proposition 2.4, we will first show that the sets $\ell(v)$ can be determined by means of a process called Warning Propagation, which hails from the physics literature (see [23] and the references therein). More precisely, we will see that Warning Propagation yields color sets $L(v)$ such that $L(v) = \ell(v)$ for all but $o(n)$ vertices a.a.s. Crucially, by tracing Warning Propagation we will be able to determine for any given type $(i, \ell)$ how many vertices of that type there are. Moreover, we will show that the cluster $\mathcal{C}(\sigma)$ essentially consists of all $k$-colorings $\tau$ of $G$ such that $\tau(v) \in L(v)$ for all $v$. In addition, the number of such colorings $\tau$ can be calculated by considering a certain reduced graph $G_{WP}(\sigma)$. This graphs turns out to be a forest (possibly after the removal of $o(n)$ vertices), and the final step of the proof consists in arguing that, informally speaking, a.a.s. the statistics of the trees in this forest are given by the distribution of the multi-type branching process from Sect. 2.

Let us begin by describing Warning Propagation on a general graph $G$ endowed with a $k$-coloring $\sigma$. For each edge $e = \{v, w\}$ of $G$ and any color $i$ we define a sequence $(\mu_{v \to w}(i, t|G, \sigma))_{t \geq 1}$ such that $\mu_{v \to w}(i, t|G, \sigma) \in \{0, 1\}$ for all $i, v, w$. The idea is that $\mu_{u \to v}(i, t|G, \sigma) = 1$ indicates that in the $t$th step of the process vertex $v$ “warns” vertex $w$ that the other neighbors $u \neq w$ of $v$ force $v$ to take color $i$. We initialize this process by having each vertex $v$ emit a warning about its original $\sigma(v)$ at $t = 0$, i.e.,

$$\mu_{v \to w}(i, 0|G, \sigma) = 1_{i = \sigma(v)}$$

for all edges $\{v, w\}$ and all $i \in [k]$. Letting $\partial v = \partial_G(v)$ denote the neighborhood of $v$ in $G$, for $t \geq 0$ we let

$$\mu_{v \to w}(i, t+1|G, \sigma) = \prod_{j \in [k] \setminus \{i\}} \max \{\mu_{u \to v}(j, t|G, \sigma) : u \in \partial v \setminus \{w\}\}. \quad (6.2)$$

That is, $v$ warns $w$ about color $i$ in step $t + 1$ iff at step $t$ it received warnings from its other neighbors $u$ (not including $w$) about all colors $j \neq i$. Further, for a vertex $v$ and $t \geq 0$ we let

$$L(v, t|G, \sigma) = \left\{ j \in [k] : \max_{u \in \partial v} \mu_{u \to v}(j, t|G, \sigma) = 0 \right\} \quad \text{and}$$

$$L(v|G, \sigma) = \bigcup_{t=0}^{\infty} L(v, t|G, \sigma).$$

Thus, $L(v, t|G, \sigma)$ is the set of colors that vertex $v$ receives no warnings about at step $t$. To unclutter the notation, we omit the reference to $G, \sigma$ where it is apparent from the context.

To understand the semantics of this process, observe that by construction the list $L(v, t|G, \sigma)$ only depends on the vertices at distance at most $t+1$ from $v$. Further, if we assume that the $t$th neighborhood $\partial^t v$ in $G$ is a tree, then $L(v, t|G, \sigma)$ is precisely the set of colors that $v$ may take in $k$-colorings $\tau$ of $G$ such that $\tau(w) = \sigma(w)$ for all vertices
$w$ at distance greater than $t$ from $v$, as can be verified by a straightforward induction on $t$. As we will see, this observation together with the fact that the random graph $G$ contains only few short cycles (cf. Lemma 6.1) allows us to show that for most vertices $v$ we have $\ell(v) = L(v|G, \sigma)$ a.a.s. In effect, the number of $k$-colorings $\tau$ of $G$ with $\tau(v) \in L(v|G, \sigma)$ for all $v$ will emerge to be a very good approximation to the cluster size $\mathcal{C}(G, \sigma)$.

Counting these $k$-colorings $\tau$ is greatly facilitated by the following observation. For a graph $G$ together with a $k$-coloring $\sigma$, let us denote by $G_{WP}(t|\sigma)$ the graph obtained from $G$ by removing all edges $\{v, w\}$ such that either $|L(v, t)| < 2$, $|L(w, t)| < 2$ or $L(v, t) \cap L(w, t) = \emptyset$. Furthermore, obtain $G_{WP}(\sigma)$ from $G$ by removing all edges $\{v, w\}$ such that $L(v) \cap L(w) = \emptyset$. We view $G_{WP}(t|\sigma)$ and $G_{WP}(\sigma)$ as decorated graphs in which each vertex $v$ is endowed with the color list $L(v, t)$ and $L(v)$ respectively. As before, we let $Z$ denote the number of legal colorings of a decorated graph. Thus, $Z(G_{WP}(\sigma))$ is the number of colorings $\tau$ of $G_{WP}(\sigma)$ such that $\tau(v) \in L(v|G, \sigma)$ for all $v$. The key statement in this section is

**Proposition 6.1.** A.a.s. we have $\ln Z(G_{WP}(\sigma)) = \ln |\mathcal{C}(G, \sigma)| + o(n)$.

We begin by proving that $Z(G_{WP}(\sigma))$ is a lower bound on the cluster size a.a.s. To this end, let us highlight a few elementary facts.

**Fact 6.1.** The following statements hold for any $G, \sigma$.

1. For all $v, w, i$ and all $t \geq 0$ we have $\mu_{v \rightarrow w}(i, t + 1) \leq \mu_{v \rightarrow w}(i, t)$.
2. We have $\sigma(v) \in L(v, t)$ for all $v, t$. Moreover, if $\mu_{v \rightarrow w}(i, t) = 1$ for some $w \in \partial v$, then $i = \sigma(v)$.
3. There is a number $t^*$ such that for any $t > t^*$ we have $\mu_{v \rightarrow w}(i, t) = \mu_{v \rightarrow w}(i, t^*)$ for all $v, w, i$.

**Proof.** We prove (1) and (2) by induction on $t$. In the case $t = 0$ both statements are immediate from (6.1). Now, assume that $t \geq 1$ and $\mu_{v \rightarrow w}(i, t) = 0$. Then there is a color $j \neq i$ and a neighbor $u \neq w$ of $v$ such that $\mu_{u \rightarrow v}(j, t - 1) = 0$. By induction, we have $\mu_{u \rightarrow v}(j, t) = 0$. Hence, (6.2) implies that $\mu_{v \rightarrow w}(i, t + 1) = 0$. Furthermore, if $\mu_{v \rightarrow w}(i, t + 1) = 1$ for some $i \neq \sigma(v)$, then $v$ has a neighbor $u \neq w$ such that $\mu_{u \rightarrow v}(\sigma(v), t) = 1$. But since $\sigma(u) \neq \sigma(v)$ because $\sigma$ is a $k$-coloring, this contradicts the induction hypothesis. Thus, we have established (1) and (2). Finally, (3) is immediate from (1).

**Fact 6.2.** If for some $t \geq 0$, $\tau$ is a coloring of $G_{WP}(t|\sigma)$ such that $\tau(v) \in L(v, t)$ for all $v$, then $\tau$ is a $k$-coloring of $G$. Moreover, if $\tau$ is a $k$-coloring of $G_{WP}(\sigma)$ such that $\tau(v) \in L(v, t)$ for all $v$, then $\tau$ is a $k$-coloring of $G$.

**Proof.** Let $\{v, w\}$ be an edge of $G$. Clearly, if $L(v, t) \cap L(w, t) = \emptyset$, then $\tau(v) \neq \tau(w)$. Thus, assume that $L(v, t) \cap L(w, t) \neq \emptyset$. Then $|L(v, t)| > 1$. Indeed, if $|L(v, t)| = 1$, then by Fact 6.1 we have $L(v, t) = \{\sigma(v)\}$ and thus $\sigma(v) \notin L(w, t)$ by (6.2). Similarly, $|L(w, t)| > 1$. Hence, the edge $\{v, w\}$ is present in $G_{WP}(t|\sigma)$, and thus $\tau(v) \neq \tau(w)$. This implies the first assertion. The second assertion follows from the first assertion and Fact 6.1, which shows that there is a finite $t$ such that $L(v, t) = L(v)$ for all $v$.

To turn Fact 6.2 into a lower bound on the cluster size, we are going to argue that a.a.s. in $G$ there are a lot of frozen vertices a.a.s. In fact, a.a.s. the number of such frozen vertices will turn out to be so large that all colorings $\tau$ as in Fact 6.2 belong to the cluster $\mathcal{C}(G, \sigma)$ a.a.s.
To exhibit frozen vertices, we consider an appropriate notion of a “core”. More precisely, assume that $\sigma$ is a $k$-coloring of a graph $G$. We denote by $\text{core}(G, \sigma)$ the largest set $V'$ of vertices with the following property.

If $v \in V'$ and $j \neq \sigma(v)$, then $|V' \cap \sigma^{-1}(j) \cap \partial v| \geq 100$.

In words, any vertex in the core has at least 100 neighbors of any color $j \neq \sigma(v)$ that also belong to the core. The core is well-defined; for if $V'$, $V''$ are two sets with this property, then so is $V' \cup V''$. The following is immediate from the definition of the core.

**Fact 6.3.** Assume that $v \in \text{core}(G, \sigma)$. Then $L(v, t) = \{\sigma(v)\}$ for all $t$.

The core has become a standard tool in the theory of random structures in general and in random graph coloring in particular. Indeed, standard arguments show that $G$ has a very large core a.a.s. More precisely, we have

**Proposition 6.2** [8] A.a.s. $G, \sigma$ are such that the following two properties hold for all sets $S \subset [n']$ of size $|S| \leq \sqrt{n}$.

1. Let $G'$ be the subgraph obtained from $G$ by removing the vertices in $S$. Then

$$|\text{core}(G', \sigma) \cap \sigma^{-1}(i)| \geq \frac{n}{k} (1 - k^{-2/3}) \quad \text{for all } i \in [k]. \quad (6.3)$$

2. If $v \in \text{core}(G', \sigma)$, then $\sigma(v) = \tau(v)$ for all $\tau \in \mathcal{C}(G, \sigma)$.

**Corollary 6.1.** A.a.s. we have $|\mathcal{C}(G, \sigma)| \geq Z(G_{\text{WP}}(\sigma))$.

**Proof.** By Proposition 6.2 we may assume that (6.3) is true for $S = \emptyset$. Let $\tau$ be a $k$-coloring of $G_{\text{WP}}(\sigma)$ such that $\tau(v) \in L(v)$ for all $v$. Then Fact 6.2 implies that $\tau$ is a $k$-coloring of $G$. Furthermore, Fact 6.3 implies that $\tau(v) = \sigma(v)$ for all $v \in \text{core}(G, \sigma)$. Hence, (6.3) entails that $\rho_{ij}(\sigma, \tau) \geq 1 - k^{-2/3} > 0.51$ for all $i \in [k]$. Thus, $\tau \in \mathcal{C}(G, \sigma)$.

While $Z(G_{\text{WP}}(\sigma))$ provides a lower bound on the cluster size, the two numbers do not generally coincide. This is because for a few vertices $v$, the list $L(v)$ produced by Warning Propagation may be a proper subset of $\ell(v)$. For instance, assume that the vertices $v_1, v_2, v_3, v_4$ induce a cycle of length four such that $\sigma(v_1) = \sigma(v_3) = 1$ and $\sigma(v_2) = \sigma(v_4) = 2$, while $v_1, v_2, v_3, v_4$ are not adjacent to any further vertices of color 1 or 2. Moreover, suppose that for each color $j \in \{3, 4, \ldots, k\}$, each of $v_1, \ldots, v_4$ has at least one neighbor of color $j$ that belongs to the core. Then Warning Propagation yields $L(v_1) = L(v_3) = \{1\}$ and $L(v_2) = L(v_4) = \{2\}$. However, $v_1, v_2, v_3, v_4$ are actually unfrozen as we might as well give color 2 to $v_1, v_3$ and color 1 to $v_2, v_4$. (A bipartite sub-structure of this kind is known as a “Kempe chain”, cf. [25].)

The reason for this problem is, roughly speaking, that we launched Warning Propagation from the initialization (6.1), which is the obvious choice but may be too restrictive. Thus, to obtain an upper bound on the cluster size we will start Warning Propagation from a different initialization. Ideally, this starting point should be such that only vertices that are frozen emit warnings. By Proposition 6.2, the vertices in the core meet this condition a.a.s. Thus, we are going to compare the above installment of Warning Propagation with the result of starting Warning Propagation from an initialization where only the vertices in the core send out warnings.

Thus, given a graph $G$ together with a $k$-coloring $\sigma$ we let

$$\mu'_{\nu \mapsto \mu}(i, 0|G, \sigma) = 1_{i = \sigma(v)} \cdot 1_{v \in \text{core}(G, \sigma)}. \quad (6.1)$$
\[ \mu_{v \rightarrow w}'(i, t + 1|G, \sigma) = \prod_{j \in [k] \setminus [i]} \max \left\{ \mu_{u \rightarrow v}'(j, t|G, \sigma) : u \in \partial v \setminus \{w\} \right\} \]

for all edges \{v, w\} of \( G \), all \( i \in [k] \) and all \( t \geq 0 \). Furthermore, let

\[
L'(v, t|G, \sigma) = \left\{ j \in [k] : \max_{u \in \partial v} \mu_{u \rightarrow v}'(j, t|G, \sigma) = 0 \right\} \quad \text{and} \quad L'(v|G, \sigma) = \bigcap_{t=0}^{\infty} L'(v, t|G, \sigma).
\]

As before, we drop \( G, \sigma \) from the notation where possible.

Similarly as before, we can use the lists \( L'(v, t) \) to construct a decorated reduced graph. Indeed, let \( G'_{WP}(t|\sigma) \) be the graph obtained from \( G \) by removing all edges \{v, w\} such that \( |L'(v, t)| < 2 \) or \( |L'(w, t)| < 2 \) or \( L'(v, t) \cap L'(w, t) = \emptyset \). We decorate each vertex in this graph with the list \( L'(v, t) \). In addition, let \( G'_{WP}(\sigma) \) be the graph obtained from \( G \) by removing all edges \{v, w\} such that \( L'(v) \cap L'(w) = \emptyset \) endowed with the lists \( L'(v) \).

**Fact 6.4.** The following statements hold for all \( G, \sigma \).

1. For all \( v \) we have \( \sigma(v) \in L'(v) \). Moreover, if there are \( j, t, w \) such that \( \mu_{v \rightarrow w}'(j, t) = 1 \), then \( j = \sigma(v) \).
2. If \( v \in \text{core}(G, \sigma) \), then \( L'(v, t) = \{\sigma(v)\} \) for all \( t \).
3. We have \( \mu_{v \rightarrow w}'(i, t + 1) \geq \mu_{v \rightarrow w}'(i, t) \).
4. There is a number \( t^* \) such that for any \( t > t^* \) we have \( \mu_{v \rightarrow w}'(i, t) = \mu_{v \rightarrow w}'(i, t^*) \) for all \( v, w, i \).

**Proof.** This follows by induction on \( t \) (cf. the proof of Fact 6.1).

**Lemma 6.3.** A.a.s. for all vertices \( v \) we have \( \ell(v) = \{\tau(v) : \tau \in \mathcal{C}(G, \sigma)\} \subset L'(v|G, \sigma) \).

**Proof.** Proposition 6.2 shows that a.a.s.

\[ \tau(v) = \sigma(v) \quad \text{for all } v \in \text{core}(G, \sigma). \tag{6.4} \]

Assuming (6.4), we are going to prove by induction on \( t \) that

\[ \ell(v) \subset L'(v, t) \quad \text{for all } v \in [n], t \geq 0. \tag{6.5} \]

By construction, for any vertex \( v \) and any color \( j \) we have \( j \in L'(v, 0) \), unless \( v \) has a neighbor \( w \in \text{core}(G, \sigma) \) such that \( \sigma(w) = j \). Moreover, if such a neighbor \( w \) exists, (6.4) implies that a.a.s. \( \tau(w) = j \) and thus \( \tau(v) \neq j \) for all \( \tau \in \mathcal{C}(\sigma) \). Hence, (6.5) is true for \( t = 0 \).

Now, assume that (6.5) holds for \( t \). Suppose that \( j \notin L'(v, t + 1) \). Then \( v \) has a neighbor \( u \) such that \( \mu_{u \rightarrow v}'(j, t + 1) = 1 \). Therefore, for each \( l \neq j \) there is \( w_l \neq v \) such that \( \mu_{w_l \rightarrow u}'(l, t) = 1 \). Consequently, \( L'(u, t) = \{j\} \). Hence, by induction we have \( \tau(u) = j \) and thus \( \tau(v) \neq j \) for all \( \tau \in \mathcal{C}(G, \sigma) \).

As an immediate consequence of Lemma 6.3 we obtain
Corollary 6.2. A.a.s. we have $|\mathcal{C}(G, \sigma)| \leq Z(G'_{\text{WP}}(\sigma))$.

Combining Corollary 6.1 and Corollary 6.2, we see that $Z(G_{\text{WP}}(\sigma)) \leq |\mathcal{C}(G, \sigma)| \leq Z(G'_{\text{WP}}(\sigma))$ a.a.s. To complete the proof of Proposition 6.1, we are going to argue that $\ln Z(G_{\text{WP}}(\sigma)) = \ln Z(G'_{\text{WP}}(\sigma)) + o(n)$ a.a.s.

To this end, we need one more general construction. Let $G$ be a graph and let $\sigma$ be a $k$-coloring of $G$. Let $t \geq 0$ be an integer. For each vertex $v$ of $G$ we define a rooted, decorated graph $T(v, t|G, \sigma)$ as follows.

- The graph underlying $T(v, t|G, \sigma)$ is the connected component of $v$ in $G_{\text{WP}}(v, t|G, \sigma)$.
- The root of $T(v, t|G, \sigma)$ is $v$.
- The type of each vertex $w$ of $T(v, t|G, \sigma)$ is $(\sigma(w), L(w, t|G, \sigma))$.

Analogously we obtain a rooted, decorated graph $T(v|G, \sigma)$ from $G_{\text{WP}}(\sigma)$, $T'(v, t|G, \sigma)$ from $G'_{\text{WP}}(t|\sigma)$ and $T'(v|G, \sigma)$ from $G'_{\text{WP}}(\sigma)$.

Of course, the total number $Z(G_{\text{WP}}(\sigma))$ of legal colorings of $G_{\text{WP}}(\sigma)$ is just the product of the number of legal colorings of all the connected components of $G_{\text{WP}}(\sigma)$. The following lemma shows that a.a.s. for all but $o(n)$ vertices the components in $G_{\text{WP}}(\sigma)$ and $G'_{\text{WP}}(\sigma)$ coincide.

Lemma 6.4. A.a.s. $G, \sigma$ is such that $T(v|G, \sigma) = T'(v|G, \sigma)$ for all but $o(n)$ vertices $v$.

The main technical step towards the proof of Lemma 6.4 is to show that a.a.s. most of the components $T'(v|G, \sigma)$ are “small” by comparison to $n$. Technically, it is easier to establish this statement for $T'(v, 0|G, \sigma)$, which contains $T'(v|G, \sigma)$ as a subgraph due to the monotonicity property Fact 6.4(3).

Lemma 6.5. For any $\varepsilon > 0$ there is a number $\omega = \omega(\varepsilon) > 0$ such that a.a.s. for at least $(1 - \varepsilon)n$ vertices $v$ the component $T'(v, 0|G, \sigma)$ contains no more than $\omega$ vertices.

The proof of Lemma 6.5, which we defer to Sect. 6.4, is a bit technical but based on known arguments. Lemma 6.1 shows that a.a.s. for most vertices $v$ such that $T'(v, 0|G, \sigma)$ contains at most, say, $\omega = \lfloor \ln \ln \ln n \rfloor$ vertices, $T'(v, 0|G, \sigma)$ is a tree. In this case, the following observation applies.

Lemma 6.6. Let $G$ be a graph and let $\sigma$ be a $k$-coloring of $G$. Assume that $T'(v, 0|G, \sigma)$ is a tree on $\omega$ vertices for some integer $\omega \geq 1$. Then for any vertex $y$ in $T'(v, 0|G, \sigma)$ we have $L(y|G, \sigma) = L'(y|G, \sigma)$. Moreover, if $T'(v, 0|G, \sigma)$ has $\omega$ vertices, then $L(y|G, \sigma) = L(y, \omega + 1|G, \sigma)$ and $L'(y|G, \sigma) = L'(y, \omega + 1|G, \sigma)$.

Proof. To get started let us recall some basic properties of the warnings

P1 If for an edge $\{x, y\}$ in $G$ we have $\mu_{x \rightarrow y}(i, 0) = 1$ or $\mu'_{x \rightarrow y}(i, 0) = 1$ then $i = \sigma(x)$.

P2 For each vertex $v \in G$ we have $\sigma(v) \in L(v, t)$ and $\sigma(v) \in L'(v, t)$ for all $t \geq 0$.

P3 For all edges $\{x, y\}$ in $G$ we have $\mu_{x \rightarrow y}(i, t) \geq \mu'_{x \rightarrow y}(i, t)$ for all $i \in [k]$.

As a first step we are going to show that for each edge $\{x, y\}$ in $T'(v, 0|G, \sigma)$ we have

$$\mu_{x \rightarrow y}(i, t) = \mu'_{x \rightarrow y}(i, t) = 0 \quad \text{for all } t > \omega \text{ and all } i \in [k]. \quad (6.6)$$
To do so, pick and fix an arbitrary vertex $y$ in $T'(v, 0)$. We define the $y$-height $h_y(x)$ of a vertex $x \neq y$ in $T'(v, 0)$ as follows. Since $T'(v, 0)$ is a tree, there is a unique path from $x$ to $y$ in $T'(v, 0)$. Let $P_y(x)$ be the neighbor of $x$ on this path. Then $h_y(x)$ is the maximum distance from $x$ to a leaf of $T'(v, 0)$ that belongs to the component of $x$ in the subgraph of $T'(v, 0)$ obtained by removing the edge $\{x, P_y(x)\}$.

Let $U$ be the set of all neighbors $u$ of $x$ that do not belong to $T'(v, 0)$, and let $U'$ be the set of all neighbors $u' \neq P_y(x)$ of $x$ in $T'(v, 0)$. We compute

$$\mu'_{x \rightarrow P_y(x)}(i, 1) = \prod_{j \in [k] \setminus \{i\}} \max \{\mu'_{u \rightarrow x}(j, 0) : u \in U\} = 0 \quad \text{for all } i \in [k]$$

where we omitted the vertices in $U'$ since by construction of core($G, \sigma$) we conclude that for all $u' \in U'$ we get $\mu'_{u' \rightarrow x}(i, 0) = 0$ for all $i \in [k]$. For each $j \in [k] \setminus L'(x, 0)$ there exists a neighbor $u \in U$ such that $\sigma(u) = j$ and $\mu'_{u \rightarrow x}(i, 0) = 1_{i = j}$ and let $U_C$ be the set of all such neighbors. By Fact 6.4 and P3 for all $u \in U_C$ we find

$$\mu_{u \rightarrow x}(i, t) = \mu'_{u \rightarrow x}(i, t) = 1_{i = \sigma(u)} \quad \text{for all } i \in [k] \text{ for all } t \geq 0. \quad (6.7)$$

By construction of $T'(v, 0)$ for all $u \in U$ the lists $L'(x, 0)$ and $L'(u, 0)$ are disjoint and by P1, P2 and (6.7) we obtain

for any $u \in U \setminus U_C$ we find $\sigma(u) \in L'(u, 0) \subset [k] \setminus L'(x, 0)$ and thus there exists a $u' \in U_C$ such that $\mu_{u \rightarrow x}(i, 0) = \mu_{u' \rightarrow x}(i, 0) = 1_{i = \sigma(u)}$ in (6.8) particular $\mu_{u \rightarrow x}(i, t) \leq \mu_{u' \rightarrow x}(i, 0) = 1_{i = \sigma(u)}$ for all $t \geq 0$.

We conclude by (6.8) that

$$\mu'_{x \rightarrow P_y(x)}(i, 1) = \prod_{j \in [k] \setminus \{i\}} \max \{\mu'_{u \rightarrow x}(j, 0) : u \in U_C\} = 0 \quad \text{for all } i \in [k]. \quad (6.9)$$

To prove (6.6) we show by induction on $h_y(x)$ that for all $i \in [k]$

$$\mu_{x \rightarrow P_y(x)}(i, t) = \mu'_{x \rightarrow P_y(x)}(i, t) = 0 \quad \text{for all } t \geq h_y(x) + 1. \quad (6.10)$$

To get started, suppose that $h_y(x) = 0$. Then $x$ is a leaf of $T'(v, 0)$. We compute

$$\mu_{x \rightarrow P_y(x)}(i, 1) = \prod_{j \in [k] \setminus \{i\}} \max \{\mu_{u \rightarrow x}(j, 0) : u \in U\}$$

$$= \prod_{j \in [k] \setminus \{i\}} \max \{\mu_{u \rightarrow x}(j, 0) : u \in U_C\} \quad \text{[by (6.8)]}$$

$$= \prod_{j \in [k] \setminus \{i\}} \max \{\mu'_{u \rightarrow x}(j, 0) : u \in U_C\} \quad \text{[by (6.7)]}$$

$$= \mu'_{x \rightarrow P_y(x)}(i, 1) = 0 \quad \text{[by (6.9)]}$$

for all $i \in [k]$. By Fact 6.1 and P3 we conclude that $\mu_{x \rightarrow P_y(x)}(i, t) = \mu'_{x \rightarrow P_y(x)}(i, t) = 0$ for all $t \geq 1$.

Now, assume that $h_y(x) > 0$. Then all $u' \in U'$ satisfy $h_y(u') < h_y(x)$. Moreover, $P_y(u') = x$. Therefore, by induction
\[ \mu_{u' \rightarrow x}(i, t) = \mu_{u' \rightarrow x}(i, h_y(x)) = 0 = \mu'_{u' \rightarrow x}(i, h_y(x)) = \mu'_{u' \rightarrow x}(i, t) \]
\[ \text{for all } u' \in U', \ i \in [k], \ t > h_y(x). \]  
\hfill (6.11)

We compute
\[ \mu_{x \rightarrow p_y(x)}(i, t) = \prod_{j \in [k]\{i\}} \max \{ \mu_{u \rightarrow x}(j, t - 1) : u \in U \cup U' \} \]
\[ = \prod_{j \in [k]\{i\}} \max \{ \mu_{u \rightarrow x}(j, 0) : u \in U_C \} \text{ by (6.8) and (6.11)} \]
\[ = \prod_{j \in [k]\{i\}} \max \{ \mu'_{u' \rightarrow x}(j, 0) : u \in U_C \} \text{ by (6.7)} \]
\[ = \mu'_{x \rightarrow p_y(x)}(i, 1) = 0 \text{ for all } i \in [k], \ t \geq h_y(x) + 1. \]

Again by Fact 6.1 and P3 we conclude that \( \mu_{x \rightarrow p_y(x)}(i, t) = \mu'_{x \rightarrow p_y(x)}(i, t) = 0 \) for all \( i \in [k] \) and \( t \geq h_y(x) + 1. \)

Finally, we observe that \( h_y(x) \leq \omega = |T'(v, 0)| \) for all \( x \). Hence, applying (6.10) to the neighbors \( x \) of \( y \) in \( T'(v, 0) \), we obtain \( \mu_{x \rightarrow y}(j, t) = \mu_{x \rightarrow y}(i, \omega + 1) = \mu'_{x \rightarrow y}(i, \omega + 1) = 0 = \mu'_{x \rightarrow y}(i, t) \) for all \( i \in [k] \) and all \( t \geq \omega \). Together with (6.7) which states that for any \( x \in T'(v, 0) \) and for any \( j \in [k]\{x\} \\backslash L'(x, 0) \) there exists a vertex \( u \notin T'(v, 0) \) that is adjacent to \( x \) in \( G \) such that \( \mu_{u \rightarrow x}(j, t) = \mu_{u' \rightarrow x}(j, t) = 1 \) for all \( t \geq 0 \) and with (6.8) which states that for any \( j \in L'(x, 0) \) there exists no vertex \( u \notin T'(v, 0) \) that is adjacent to \( x \) in \( G \) such that \( \mu_{u \rightarrow x}(j, t) = \mu_{u' \rightarrow x}(j, t) = 1 \) for any \( t \geq 0 \) we conclude that \( L(x) = L(x, \omega + 1) = L'(x, \omega + 1) = L'(x) \) as desired.

**Proof of Lemma 6.4.** Lemma 6.5 implies that for all but \( o(n) \) vertices \( v \) we have \( |T'(v, 0)| \leq \ln \ln n \) a.a.s. Together with Lemma 6.1, this implies that a.a.s. \( T'(v, 0) \) is a tree for all but \( o(n) \) vertices \( v \). Thus, assume in the following that \( v \) is such that \( T'(v, 0) \) is a tree.

It is immediate from Facts 6.1, 6.3 and 6.4 that \( L(w) \subset L'(w) \subset L'(w, 0) \) for all vertices \( w \). Therefore, \( G'_{WP}(\sigma) \subset G'_{WP}(\sigma) \subset G'_{WP}(0|\sigma) \) and thus
\[ T(v) \subset T'(v) \subset T'(v, 0). \]  
\hfill (6.12)

Conversely, Lemma 6.6 shows that \( L(x) = L'(x) \) for all vertices \( x \) in \( T'(v, 0) \). Together with (6.12), this implies that \( T(v) = T'(v) \).

**Proof of Proposition 6.1.** By Corollaries 6.1 and 6.2 we have \( Z(G_{WP}(\sigma)) \leq |C(G, \sigma)| \leq Z(G'_{WP}(\sigma)) \) a.a.s. Thus, it suffices to show that \( \ln Z(G_{WP}(\sigma)) = \ln Z(G'_{WP}(\sigma)) + o(n) \) a.a.s. Indeed, because the various connected components of \( G_{WP}(\sigma) \) can be colored independently, we find that
\[ \ln Z(G_{WP}(\sigma)) = \sum_{v \in [n']} \frac{\ln Z(T(v|G, \sigma))}{|T(v|G, \sigma)|}, \]
\[ \ln Z(G'_{WP}(\sigma)) = \sum_{v \in [n']} \frac{\ln Z(T'(v|G, \sigma))}{|T'(v|G, \sigma)|}. \]  
\hfill (6.13)
Clearly, for any vertex $v$ we have $\frac{\ln Z(T(v|G, \sigma))}{|T(v|G, \sigma)|}, \frac{\ln Z(T'(v|G, \sigma))}{|T'(v|G, \sigma)|} \leq \ln k$. Hence, Lemma 6.4 shows that a.a.s.

$$\sum_{v \in [n']} \frac{\ln Z(T(v|G, \sigma))}{|T(v|G, \sigma)|} \sim \sum_{v \in [n']} \frac{\ln Z(T'(v|G, \sigma))}{|T'(v|G, \sigma)|}. \quad (6.14)$$

Finally, the assertion follows from (6.13) and (6.14).

6.3. Counting legal colorings. Proposition 6.1 reduces the proof of Proposition 2.4 to the problem of counting the legal colorings of the reduced graph $G_{WP}(\sigma)$. Lemma 6.5 implies that a.a.s. $G_{WP}(\sigma)$ is a forest consisting mostly of trees of size, say at most $\ln \ln n$. In this section we are going to show that a.a.s. the “statistics” of these trees follow the distribution of the random tree generated by the branching process from Sect. 2. To formalise this, let $T = T_{d,k,q^*}$ with $q^*$ from (2.7) denote the random isomorphism class of rooted, decorated trees produced by the process GW($d, k, q^*$). Moreover, for a rooted, decorated tree $T$ let $H_T$ be the number of vertices $v$ in $G_{WP}(\sigma)$ such that $T(v|G, \sigma) \cong T$. In this section we prove

**Proposition 6.3.** If $T$ is such that $\mathbb{P}[T \in T] > 0$, then $(\frac{1}{n} H_T)_{n \geq 1}$ converges to $\mathbb{P}[T \in T]$ in probability.

We begin by showing that the fixed point problem $q^* = F(q^*)$ with $F$ from (5.1) provides a good approximation to the number of vertices $v$ such that $L(v|G, \sigma) = \{i\}$ for any $i$. To this end, we let

$$q^0 = (1/k, \ldots, 1/k) \quad \text{and} \quad q^t = F(q^{t-1}) \quad \text{for} \quad t \geq 1.$$

In addition, let $Q_i(t|G, \sigma)$ be the set of vertices $v$ of $G$ such that $L(v, t|G, \sigma) = \{i\}$.

**Lemma 6.7.** For any $i \in [k]$ and any fixed $t > 0$ we have $\frac{1}{n} |Q_i(t|G, \sigma)| = q^t_i + o(1)$ a.a.s.

**Proof.** We proceed by induction on $t$. To get started, we set $Q_i(-1|G, \sigma) = \sigma^{-1}(i)$ and $q^t_i = 1/k$. Then a.a.s. $\frac{1}{n} |Q_i(-1|G, \sigma)| = q^{-1}_i + o(1)$.

Now, assuming that $t \geq 0$ and that the assertion holds for $t − 1$, we are going to argue that

$$\mathbb{E}[|Q_i(t|G, \sigma)|/n] = q^t_i + o(1). \quad (6.15)$$

Indeed, let $v = n'$ be the last vertex of the random graph, and let us condition on the event that $\sigma(v) = i$. By symmetry and the linearity of expectation, it suffices to show that

$$\mathbb{P}[L(v, t|G, \sigma) = \{i\} | \sigma(v) = i] = kq^t_i + o(1). \quad (6.16)$$

To show (6.16), let $\tilde{G}$ signify the subgraph obtained from $G$ by removing $v$. Moreover, let $Q_j^{t-1}(\varepsilon)$ be the event that

$$|n^{-1} |Q_j(t-1|\tilde{G}, \sigma)| - q^t_{j-1}| < \varepsilon \quad \text{for all} \quad j \in [k]. \quad (6.17)$$
Since $\tilde{G}$ is nothing but a random graph $G(n'-1, p', \sigma)$ with one less vertex and as $n' - 1 = n - o(n)$, by induction we have

$$\Pr[Q'^{-1}(\varepsilon)] = 1 - o(1) \quad \text{for any } \varepsilon > 0. \quad (6.18)$$

Let $\mathcal{A}(i)$ be the event that for each $j \in [k] \setminus \{i\}$ there is $w \in \partial v$ such that $L(w, t - 1|G, \sigma) = \{j\}$. Given $\sigma(v) = i$, we can obtain $G$ from $\tilde{G}$ by connecting $v$ with each vertex $w \in [n'-1]$ such that $\sigma(w) \neq i$ with probability $p'$ independently. Therefore,

$$\Pr[\mathcal{A}(i)|\tilde{G}, \sigma(v) = i] = \prod_{j \neq i} 1 - (1 - p')^{\lvert Q_j(t-1|\tilde{G}, \sigma) \rvert}$$

$$\sim \prod_{j \neq i} 1 - \exp(-p'|Q_j(t-1|\tilde{G}, \sigma)|)$$

$$= \prod_{j \neq i} 1 - \exp\left[-\frac{kd}{k-1} \cdot n^{-1}|Q_j(t-1|\tilde{G}, \sigma)|\right].$$

Furthermore, for any fixed $\delta > 0$ there is an $(n$-independent) $\varepsilon > 0$ such that given that $Q'^{-1}(\varepsilon)$ occurs, we have

$$\left|q_i' - \prod_{j \neq i} 1 - \exp\left(-\frac{kd}{k-1} \cdot n^{-1}|Q_j(t-1|\tilde{G}, \sigma)|\right)\right| < \delta. \quad (6.19)$$

Combining (6.18) and (6.19), we see that for any fixed $\delta > 0$ we have

$$\left|\Pr[\mathcal{A}(i)|\sigma(v) = i] - kq_i'\right| < \delta + o(1). \quad (6.20)$$

If $v$ is acyclic, $\sigma(v) = i$ and $\mathcal{A}(i)$ occurs, then $L(v, t|G, \sigma) = \{i\}$. Therefore, (6.16) follows from (6.20) and Lemma 6.1.

Finally, the random variable $\lvert Q_i'(G, \sigma) \rvert$ satisfies the assumptions of Lemma 6.2. Indeed, the event $v \in Q_i(t|G, \sigma)$ is determined solely by the sub-graph of $G$ encompassing those vertices at distance at most $t$ from $v$. Thus, (6.15) and Lemma 6.2 imply that $\frac{1}{n}|Q_i(t|G, \sigma)| = q_i' + o(1) \text{ a.a.s., as desired.}$

As a next step, we consider the statistics of the trees $T(v, \omega|G, \sigma)$ with $\omega \geq 0$ large but fixed as $n \to \infty$. Thus, for an isomorphism class $T$ of rooted, decorated graphs we let $H_T, \omega$ be the number of vertices $v$ in $G_{WP}(\omega|\sigma)$ such that $T(v, \omega|G, \sigma) \in T$.

**Lemma 6.8.** Assume that $T$ is an isomorphism class of rooted decorated trees such that $\Pr[T = T] > 0$. Then for any $\varepsilon > 0$ there is $\omega > 0$ such that

$$\lim_{n \to \infty} \Pr\left[\left|\Pr[T = T] - \frac{1}{n} H_{T, \omega}\right| > \varepsilon\right] = 0.$$  

**Proof.** We observe that $\Pr[T = T]$ is a number that depends on $T$ but not on $n$. Furthermore, if $T_*$ is the isomorphism class of a rooted sub-tree of $T$, then $\Pr[T = T_*] \geq \Pr[T = T]$. The proof is by induction on the height of the trees in $T$. In the case that $T$ consists of a single vertex $v$ of type $(i, \{i\})$ for some $i \in [k]$, the assertion readily follows from Lemma 6.7.
Let \((i_0, \ell_0)\) be the type of the root and \(v = n'\). To this end, consider the graph \(\tilde{G}\) obtained by removing \(v\). By Lemma 6.7 the number of vertices \(w\) of \(\tilde{G}\) with \(L(w, \omega|\tilde{G}, \sigma) = \{j\}\) is \(n(q_j + o_\omega(1))\) a.a.s. for all \(j\), where \(o_\omega(1)\) signifies a term that tends to 0 in the limit of large \(\omega\). Let \(A\) be the event that this is indeed the case. Moreover, let \(B\) be the following event.

- \(\sigma(v) = i_0\).
- for each color \(j \notin \ell_0\), vertex \(v\) has a neighbor \(w\) in \(\tilde{G}\) such that \(L(w, \omega|\tilde{G}, \sigma) = \{j\}\).
- \(v\) does not have a neighbor \(w\) with \(L(w, \omega|\tilde{G}, \sigma) = \{h\}\) for any \(h \in \ell_0\).

Then

\[
\mathbb{P}[B|A] = \frac{1}{k} \prod_{j \notin \ell_0} \mathbb{P}\left[\text{Bin}(n(q_j^* + o_\omega(1)), p') > 0\right] \\
\cdot \prod_{j \in \ell_0 \setminus \{i_0\}} \mathbb{P}\left[\text{Bin}(n(q_j^* + o_\omega(1)), p') = 0\right] \\
\sim \frac{1}{k} \prod_{j \notin \ell_0} \mathbb{P}\left[\text{Po}(np'(q_j^* + o_\omega(1))) > 0\right] \\
\cdot \prod_{j \in \ell_0 \setminus \{i_0\}} \mathbb{P}\left[\text{Po}(np'(q_j^* + o_\omega(1))) = 0\right] \\
= q_{i_0, \ell_0}^* + o_\omega(1).
\]

Since \(\mathbb{P}[A] \sim 1\), we find

\[
\mathbb{P}[B] = q_{i_0, \ell_0}^* + o_\omega(1).
\] (6.21)

Let \(T_{v_0}\) be the unique tree of the isomorphism class of rooted decorated trees consisting only of the root \(v_0\). Let \(\mathcal{V}_{v}\) be the event that \(v\) has no neighbor of any type \((i', \ell')\) in \(T_{i_0, \ell_0}\). Therefore let \(q_{\emptyset}^0 = \sum_{(i', \ell') \in T_{i, \ell}} q_{i, \ell}\). We find

\[
\mathbb{P}[\mathcal{V}_{v}|B] = (1 - p')^{n(q_i + o_\omega(1))} = o_\omega(1) + \exp(-np'q_{\emptyset}) \\
= o_\omega(1) + \exp(-d'q_{\emptyset}) = o_\omega(1) + \mathbb{P}\left[T_{v_0} \in T_{i_0, \ell_0}\right].
\] (6.22)

Combining (6.21) and (6.22), we find that

\[
\mathbb{P}[B \cap \mathcal{V}_{v}] = \mathbb{P}\left[T_{v_0} \in T\right] + o_\omega(1).
\]

As for the inductive step, pick and fix one representative \(T_0 \in T\). If we remove the root \(v_0\) from \(T_0\), then we obtain a decorated forest \(T_0 - v_0\). Each tree \(T'\) in this forest contains precisely one neighbor of the root of \(T_0\), which we designate as the root of \(T'\). Let \(\mathcal{V}(T)\) be the set of all isomorphism classes of rooted decorated trees \(T'\) obtained in this way. Furthermore, for each \(\hat{T} \in \mathcal{V}(T)\) let \(y(\hat{T})\) be the number of components of the forest \(T_0 - v_0\) that belong to the isomorphism class \(\hat{T}\).

We are going to show that for \(v = n'\) and for \(\omega = \omega(T, \varepsilon)\) sufficiently large we have

\[
|\mathbb{P}[T(v, \omega|G, \sigma) \cong T_0] - \mathbb{P}[T = T]| < \varepsilon.
\]

Furthermore, for each tree \(T' \in \mathcal{V}(T)\) let \(\hat{Q}(T')\) be the set of all vertices \(w\) of \(\tilde{G}\) such that \(T(w, \omega|\tilde{G}, \sigma) \cong T'\). In addition, let \(\tilde{Q}_{\emptyset}\) be the set of all vertices \(w\) of \(\tilde{G}\) that satisfy none of the following conditions:
Further, let $q(T') = \mathbb{P}[T = T']$ and let

$$q_\emptyset(T) = q_\emptyset^0 - \sum_{T' \in \mathcal{V}(T)} q(T').$$

Let $\mathcal{Q}$ be the event that $|\tilde{Q}(T')|/n = q(T') + o_\omega(1)$ for all $T' \in \mathcal{V}(T)$ and that $|\tilde{Q}_\emptyset|/n = q_\emptyset(T) + o_\omega(1)$. Then

$$\mathbb{P}[\mathcal{Q}] \sim 1$$

by induction. Further, let $\mathcal{Y}$ be the event that for each $T' \in \mathcal{V}(T)$ we have $y(T') = |\partial v \cap \tilde{Q}(T')|$ and $\partial v \cap \tilde{Q}_\emptyset = \emptyset$. Then

$$\mathbb{P}[\mathcal{Y}|\mathcal{B}] \sim \mathbb{P}[\mathcal{Y}|\mathcal{B}, \mathcal{Q}]$$

$$= (1 - p')^{n(q_\emptyset + o_\omega(1))} \prod_{T' \in \mathcal{V}(T)} \mathbb{P}[\text{Bin}(n(q(T') + o_\omega(1)), p') = y(T')]$$

$$= o_\omega(1) + \exp(-np'q_\emptyset) \prod_{T' \in \mathcal{V}(T)} \mathbb{P}[\text{Poisson}(np'q(T')) = y(T')]$$

$$= o_\omega(1) + \exp(-d'q_\emptyset) \prod_{T' \in \mathcal{V}(T)} \mathbb{P}[\text{Poisson}(d'q(T')) = y(T')]$$

$$= o_\omega(1) + \mathbb{P}[T_0 \in T_{i_0, \ell_0}].$$

(6.23)

The last equality sign follows from the fact that in tree $T_{i_0, \ell_0}$, the root has a Poisson number of children of possible “shape” $T'$. Combining (6.21) and (6.23), we find that

$$\mathbb{P}[\mathcal{B} \cap \mathcal{Y}] = \mathbb{P}[T_0 \in T] + o_\omega(1).$$

(6.24)

Let $\mathcal{R}$ be the event that $\partial_\omega^G(v)$ is acyclic. By Lemma 6.1 we have $\mathbb{P}[\mathcal{R}] \sim 1$. Furthermore, given $\mathcal{R}$, we have $\tilde{T}(v, \omega|G, \sigma) \in T$ iff the event $\mathcal{B} \cap \mathcal{Y}$ occurs. Thus, (6.24) implies that

$$\mathbb{P}[T(v, \omega|G, \sigma) \in T] = \mathbb{P}[\mathcal{B} \cap \mathcal{Y}] + o(1) = \mathbb{P}[T = T] + o_\omega(1).$$

(6.25)

Moreover, (6.25) shows that

$$\frac{1}{n} \mathbb{E}[H_{T, \omega}] = \mathbb{P}[T = T] + o_\omega(1).$$

(6.26)

Finally, because the event $T(v, \omega|G, \sigma) \in T$ is governed by the vertices at distance at most $|T| + o$ from $v$, Lemma 6.2 implies together with (6.26) that for any $\varepsilon > 0$ there is $\omega$ such that

$$\mathbb{P}[|H_{T, \omega} - \mathbb{P}[T = T]| < \varepsilon n] = 1 - o(1).$$

This completes the induction.

**Lemma 6.9.** For any $\varepsilon > 0$ there is $\omega > 0$ such that a.a.s. all but $\varepsilon n$ vertices $v$ satisfy $T(v|G, \sigma) = T(v, \omega|G, \sigma)$. 
Lemma 6.6 implies that \( T(v|G, \sigma) = T(v, \omega + 2|G, \sigma) \), unless \( T'(v, 0|G, \sigma) \) contains at least \( \omega \) vertices. Furthermore, Lemma 6.5 implies that for any fixed \( \epsilon > 0 \) there is \( \omega = \omega(\epsilon) \) such that this holds for no more than \( \epsilon n \) vertices a.a.s.

Finally, Proposition 6.3 is immediate from Lemmas 6.8 and 6.9 and Proposition 2.4 follows from Propositions 6.1 and 6.3.

6.4. Proof of Lemma 6.5. Let \( \theta = \lceil \ln \ln n \rceil \). Moreover, for a set \( S \subseteq V \) let \( C_S \) denote the \( \sigma \)-core of the subgraph of \( G \) obtained by removing the vertices in \( S \). Further, for any vertex \( w \in S \) let \( \Lambda(w, S) \) be the set of colors \( j \in [k] \) such that in \( G \) vertex \( w \) does not have a neighbor in \( \sigma^{-1}(j) \cap C_S \). In addition, let us call \( S \) wobbly in \( G \) if the following conditions are satisfied.

W1 \( |S| = \theta \).

W2 We have \( |\Lambda(w, S)| \geq 2 \) for all \( w \in S \).

W3 The subgraph of \( G \) induced on \( S \) has a spanning tree \( T \) such that\[ \Lambda(u, S) \cap \Lambda(w, S) \neq \emptyset \quad \text{for each edge } \{u, w\} \text{ of } T. \]

Assume that \( T'(v, 0|G, \sigma) \) contains at least \( \theta \) vertices. If \( T = (S, E_T) \) is a sub-tree on \( \theta \) vertices contained in \( T'(v, 0|G, \sigma) \), then \( S \) is wobbly. Therefore, it suffices to prove that the total number \( W \) of vertices that are contained in a wobbly set \( S \) satisfies

\[
\mathbb{E}[W] \leq \sum_{S \subseteq V : |S| = \theta} \theta \cdot \mathbb{P}[S \text{ is wobbly}] = o(n). \tag{6.27}
\]

To prove (6.27), we need a bit of notation. For a set \( S \) let \( E_S \) be the event that

\[
|C_S \cap \sigma^{-1}(i)| \geq \frac{n}{k} (1 - k^{-2/3}) \quad \text{for all } i \in [k].
\]

Then Proposition 6.2 implies that for any set \( S \) of size \( \theta \) we have

\[
\mathbb{P}[E_S] \geq 1 - \exp(-\Omega(n)). \tag{6.28}
\]

Further, for a vertex \( w \in S \) and a set \( J_w \subseteq [k] \setminus \{\sigma(w)\} \) let \( L(w, J_w) \) be the event that \( \Lambda(w, S) \supseteq J_w \). Crucially, the core \( C_S \) of the subgraph of \( G \) obtained by removing \( S \) is independent of the edges between \( S \) and \( C_S \). Therefore, \( w \) is adjacent to a vertex \( x \in C_S \) with \( \sigma(x) \neq \sigma(w) \) with probability \( p' \), independently for all such vertices \( x \). Consequently,

\[
\mathbb{P}[L(w, J_w)|E_S] \leq \prod_{j \in J_w} (1 - p')^{\frac{2}{1 - k^{-2/3}}} \leq k^{-1.99|J_w|}. \tag{6.29}
\]

Moreover, due to the independence of the edges in \( G \), the events \( L(w, J_w) \) are independent for all \( w \in S \).

Let \( S \subseteq V \) be a set of size \( \theta \). Let us call a vertex \( w \in S \) rich if \( |\Lambda(w, S)| \geq \sqrt{k} \). Further, let \( R_S \) be the set of rich vertices in \( S \). To estimate the probability that \( S \) is wobbly, we consider the following events.

- Let \( A_S \) be the event that \( |R_S| \geq k^{-1/3}\theta \) and that \( G \) contains a tree \( T \) with vertex set \( S \).
– Let $\mathcal{A}_S'$ be the event that and that $G$ contains a tree $T$ with vertex set $S$ such that
\[
\sum_{w \in R_S} |\partial_T^1(w)| \geq \theta/2.
\]
(In words, the sum of the degrees of the rich vertices in $T$ is at least $\theta/2$.)

– Let $\mathcal{A}_S''$ be the event that $G$ contains a tree $T$ with vertex set $S$ such that
\[
\sum_{w \in R_S} |\partial_T^1(w)| < \theta/2.
\]

– Let $\mathcal{W}_S$ be the event that condition $\textbf{W2}$ is satisfied.

– For a given tree $T$ with vertex set $S$ let $\mathcal{W}_{S,T}'$ be the event that condition $\textbf{W3}$ is satisfied.

If $S$ is wobbly, then the event $\mathcal{A}_S \cup (\mathcal{W}_S \cap \mathcal{A}_S') \cup (\mathcal{W}_S \cap \mathcal{W}_{S,T}' \cap \mathcal{A}_S'')$ for a tree $T$ occurs. Therefore,
\[
\mathbb{P}\left[ S \text{ is wobbly} \right] \leq \mathbb{P}[\mathcal{A}_S] + \mathbb{P}[\mathcal{W}_S \cap \mathcal{A}_S' \setminus \mathcal{A}_S] + \mathbb{P}[\mathcal{W}_S \cap \mathcal{W}_{S,T}' \cap \mathcal{A}_S'' \setminus (\mathcal{A}_S \cup \mathcal{A}_S')]. \tag{6.30}
\]

In the following, we are going to estimate the three probabilities on the r.h.s. separately. With respect to the probability of $\mathcal{A}_S$, (6.28) and (6.29) yield
\[
\mathbb{P}[|R_S| \geq k^{-1/3}\theta] \leq \mathbb{P}[\neg\mathcal{E}_S] + \mathbb{P}\left[ \exists R \subset S, |R| = \lfloor k^{-1/3}\theta \rfloor : \forall w \in R : |\Lambda(w, S)| \geq \sqrt{k}|\mathcal{E}_S| \right]
\leq \exp(-\Omega(n)) + \left( \frac{\theta}{k^{-1/3}\theta} \right) \left( \frac{k}{\sqrt{k}} \right)^{-1.9\sqrt{k}}
\leq \exp(-\sqrt{k}\theta).
\]

Furthermore, by Cayley’s formula there are $\theta^{\theta - 2}$ possible trees with vertex set $S$. Since any two vertices in $S$ are connected in $G$ with probability at most $p'$, and because edges occur independently, we obtain
\[
\mathbb{P}[\mathcal{A}_S] \leq \theta^{\theta - 2} p'^{\theta - 1} \cdot \mathbb{P}[|R_S| \geq k^{-1/3}\theta] \leq \theta^{\theta - 2} p'^{\theta - 1} \exp(-\sqrt{k}\theta). \tag{6.31}
\]

To bound the probability of $\mathcal{W}_S \cap \mathcal{A}_S' \setminus \mathcal{A}_S$, let $R \subset S$. Moreover, let $e(S)$ denote the total number of edges spanned by $S$ in $G$, and let $e(R, S)$ denote the number of edges that join a vertex in $R$ with another vertex in $S$. Let $\mathcal{A}_S'(R, t)$ be the event $e(S) \geq \theta - 1$ and $e(R, S) = t$. If $\mathcal{A}_S'(R, t)$ occurs, then there exist $R \subset S$, $|R| \leq r = \lfloor k^{-1/3}\theta \rfloor$, and $t \geq \theta/4$ such that $\mathcal{A}_S'(R, t)$ occurs. Therefore, by the union bound,
\[
\mathbb{P}[\mathcal{W}_S \cap \mathcal{A}_S' \setminus \mathcal{A}_S] \leq \sum_{R \subset S : |R| \leq r} \sum_{t \geq \theta/4} \mathbb{P}[\mathcal{W}_S \cap \mathcal{A}_S'(R, t)]. \tag{6.32}
\]

Further, because the event $\mathcal{W}_S$ is independent of the subgraph of $G$ induced on $S$, (6.32) yields
\[
\mathbb{P}[\mathcal{W}_S \cap \mathcal{A}_S' \setminus \mathcal{A}_S] \leq \mathbb{P}[\mathcal{W}_S] \cdot \sum_{R \subset S : |R| \leq r} \sum_{t \geq \theta/4} \mathbb{P}[\mathcal{A}_S'(R, t)]. \tag{6.33}
\]
Because any two vertices in \( S \) are connected with probability at most \( p' \) independently, the random variable \( e(R, S) \) is stochastically dominated by a binomial distribution \( \text{Bin}(r\theta, p') \). Therefore,

\[
\mathbb{P}[e(R, S) = t] \leq \mathbb{P} \left[ \text{Bin}(r\theta, p') = t \right] \leq \left( \frac{r\theta}{t} \right) p'^t. \tag{6.34}
\]

Similarly, we find

\[
\mathbb{P}[e(S) \geq \theta - 1|e(R, S) = t] \leq \mathbb{P} \left[ \text{Bin} \left( \frac{\theta^2/2}{t}, p' \right) \geq \theta - t - 1 \right] \\
\leq \left( \frac{\theta^2/2}{\theta - t - 1} \right) p'^{\theta - t - 1}. \tag{6.35}
\]

Combining (6.34) and (6.35), we get

\[
\mathbb{P}[A'_S(R, t)] \leq \left( \frac{r\theta}{t} \right) \left( \frac{\theta^2/2}{\theta - t - 1} \right) p'^{\theta - 1} \tag{6.36}
\]

Further, plugging (6.36) into (6.33), we obtain

\[
\mathbb{P} \left[ W_S \cap A'_S \setminus A_S \right] \leq \mathbb{P}[W_S] \cdot 2^\theta p^{\theta - 1} \sum_{t \geq \theta/4} \left( \frac{r\theta}{t} \right) \left( \frac{\theta^2/2}{\theta - t - 1} \right)
\]

\[
\leq 2^{1+\theta} p^{\theta - 1} \mathbb{P}[W_S] \left( \frac{r\theta}{\theta/4} \right) \left( \frac{\theta^2/2}{3\theta/4 - 1} \right)
\]

\[
\leq 2^{1+\theta} p^{\theta - 1} \mathbb{P}[W_S] \left( \frac{er\theta}{\theta/4} \right)^{\theta/4} \left( \frac{e\theta^2/2}{3\theta/4} \right)^{3\theta/4}
\]

\[
\leq \theta^\theta p^{\theta - 1} k^{-\theta/13} \mathbb{P}[W_S]. \tag{6.37}
\]

Finally, if the event \( W_S \) occurs, then for each \( w \in S \) there is \( j \in [k] \setminus \{\sigma(w)\} \) such that \( j \in \Lambda(w, S) \). Thus, (6.28) and (6.29) yield

\[
\mathbb{P}[W_S] \leq \mathbb{P}[-E_S] + \prod_{w \in S} \sum_{j \neq \sigma(w)} \mathbb{P}[L(w, \{j\})|E_S] \\
\leq \exp(-\Omega(n)) + k^{-0.99\theta} \leq k^{-0.98\theta}. \tag{6.38}
\]

Combining (6.37) and (6.38), we arrive at

\[
\mathbb{P}[W_S \cap A'_S \setminus A_S] \leq \theta^\theta p^{\theta - 1} k^{-1.02\theta}. \tag{6.39}
\]

To bound the probability of \( A''_S \), suppose that \( T \) is a tree with vertex set \( S \), let \( U \subset S \) and denote by \( A''_S(T, U) \) the event that the following statements are true.

(i) \( T \) is contained as a subgraph in \( G \).

(ii) Let \( s_0 = \min S \) and consider \( s_0 \) the root of \( T \). Then for each \( u \in U \) the parent \( P(u) \) satisfies \( P(u) \notin R_S \).
If the event $A_S^\prime \setminus (A_S \cup A_S^\prime)$ occurs, then there exist a tree $T$ and a set $U$ of size $|U| \geq \theta/3$ such that $A_S^\prime(T, U)$ occurs. Therefore,

$$\mathbb{P} \left[ W_S \cap W'_S \cap A_S^\prime \setminus (A_S \cup A_S^\prime) \right] \leq \sum_T \sum_{U:|U| \geq \theta/3} \mathbb{P} \left[ W_S \cap W'_S \cap A_S^\prime(T, U) \right].$$

(6.40)

Fix a tree $T$ on $S$ and a set $U \subset S$, $|U| \geq \theta/3$. Since any two vertices are connected in $G$ with probability at most $p^\theta$ independently, the probability that (i) occurs is bounded by $p^{\theta^2}$. Furthermore, if (ii) occurs and $u \in U$, then $|\Lambda(P(u), S)| \leq \sqrt{k}$ because $P(u)$ is not rich. In addition, $W_3$ requires that $\Lambda(P(u), S) \cap \Lambda(u, S) \neq \emptyset$. There are two ways how this can come about: first, it could be that $\Lambda(P(u), S) \cap \Lambda(u, S) \neq \emptyset$. Then the event $\mathcal{L}(u, \{j\})$ occurs for some $j \in \Lambda(P(u), S) \setminus \{\sigma(u)\}$. Hence, due to (6.29)

$$\mathbb{P} \left[ \Lambda(P(u), S) \cap \Lambda(u, S) \setminus \{\sigma(u)\} \neq \emptyset \mid \mathcal{E}_S, |\Lambda(P(u), S)| \leq \sqrt{k} \right] \leq k^{-1.49}$$

for any $u \in U$.

Alternatively, it could be that $\sigma(u) \in \Lambda(P(u), S)$. Given that $\Lambda(P(u), S)$ has size at most $\sqrt{k}$, the probability of this event is bounded by $k^{-1/2}$ because $\sigma(u)$ is random. Additionally, by $W_2$ there is another color $j \in \Lambda(u), j \neq \sigma(u)$. Hence, the event $\mathcal{L}(u, \{j\})$ occurs and (6.29) yields

$$\mathbb{P} \left[ \sigma(u) \in \Lambda(P(u), S), \Lambda(u, S) \setminus \{\sigma(u)\} \neq \emptyset \mid \mathcal{E}_S, |\Lambda(P(u), S)| \leq \sqrt{k} \right] \leq k^{-1.49}$$

for any $u \in U$.

Combining (6.28), (6.41) and (6.42), we find

$$\mathbb{P} \left[ \forall u \in U : \Lambda(P(u), S) \cap \Lambda(u, S) \neq \emptyset \wedge |\Lambda(P(u), S)| \leq \sqrt{k} \right] \leq \exp(-\Omega(n)) + k^{-1.48|U|}.$$

(6.43)

In addition, if $w \in S \setminus U$, then $W_2$ requires that the event $\mathcal{L}(w, \{j\})$ occurs for some $j \neq \sigma(w)$ and (6.29) yields

$$\mathbb{P} \left[ \forall w \in S \setminus U : \exists j \in [k] \setminus \{\sigma(w)\} : \mathcal{L}(w, j) \mid \mathcal{E}_S \right] \leq k^{-0.99|S \setminus U|}.$$

(6.44)

Combining (6.43) and (6.44), we obtain

$$\mathbb{P} \left[ W_S \cap W'_S \cap A_S''(T, U) \mid T \subset G \right] \leq \exp(-\Omega(n)) + k^{-0.99(\theta - |U|)} \cdot k^{-1.48|U|} \leq k^{-1.1\theta}.$$

(6.45)

Further, the probability that $T$ is contained in $G$ is bounded by $p^{\theta^2}$. Thus, (6.45) implies

$$\mathbb{P} \left[ W_S \cap W'_S \cap A_S''(T, U) \right] \leq k^{-1.1\theta} p^{\theta - 1}.$$

(6.46)

Finally, combining (6.40) and (6.46) and using Cayley’s formula, we obtain

$$\mathbb{P} \left[ W_S \cap W'_S \cap A_S'' \setminus (A_S \cup A_S') \right] \leq 2^\theta \theta^2 - 2k^{-1.1\theta} p^{\theta - 1}.$$
\[ \leq \theta^{\theta - 2} p^{\theta - 1} k^{-1.09 \theta}. \quad (6.47) \]

Plugging (6.31), (6.39) and (6.47) into (6.30), we see that
\[ \theta \mathbb{P}[S \text{ is wobbly}] \leq 2 \theta^{\theta + 1} p^{\theta - 1} k^{-1.02 \theta}. \]

Hence, (6.27) yields
\[ \mathbb{E}[W] \leq 2 \theta^{\theta + 1} p^{\theta - 1} k^{-1.02 \theta} \cdot \binom{n}{2} \theta \mathbb{P}[\text{wobbly}] \leq 2 \left( \frac{en}{\theta} \right)^\theta \theta^{\theta + 1} p^{\theta - 1} k^{-1.02 \theta} \]
\[ \leq n(3np')^{\theta} k^{-1.02 \theta} \leq n(7k \ln k)^{\theta} k^{-1.02 \theta} = o(n), \]
as desired.

6.5. Proof of Lemma 6.2. The following large deviations inequality known as Warnke’s inequality facilitates the proof of Lemma 6.2.

Lemma 6.10 [31]. Let \( X_1, \ldots, X_N \) be independent random variables with values in a finite set \( \Lambda \). Assume that \( f : \Lambda^N \to \mathbb{R} \) is a function, that \( \Gamma \subset \Lambda^N \) is an event and that \( c, c' > 0 \) are numbers such that the following is true.
If \( x, x' \in \Lambda^N \) are such that there is \( k \in [N] \) such that \( x_i = x'_i \) for all \( i \neq k \), then
\[ |f(x) - f(x')| \leq \begin{cases} c, & \text{if } x \in \Gamma, \\ c', & \text{if } x \notin \Gamma. \end{cases} \quad (6.48) \]

Then for any \( \gamma \in (0, 1) \) and any \( t > 0 \) we have
\[ \mathbb{P}[|f(X_1, \ldots, X_N) - \mathbb{E}[f(X_1, \ldots, X_N)]| > t] \leq 2 \exp \left( -\frac{t^2}{2N(c + \gamma(c' - c))^2} \right) + \frac{2N}{\gamma} \mathbb{P}[(X_1, \ldots, X_N) \notin \Gamma]. \]

Proof of Lemma 6.2. The proof is based on Lemma 6.10. Of course, we can view \((G, \sigma)\) as chosen from a product space \(X_2, \ldots, X_N\) with \(N = 2n'\) where \(X_i\) is a 0/1 vector of length \(i - 1\) whose components are independent \( \text{Be}(p') \) variables for \(2 \leq i \leq n'\) and where \(X_i \in [k]\) is uniformly distributed for \(i > \binom{n'}{2}\) (“vertex exposure”). Let \( \Gamma \) be the event that \( |N^\omega(v)| \leq \lambda = n^{0.01} \) for all vertices \(v\). Then by Lemma 6.1 we have
\[ \mathbb{P}[\Gamma] \geq 1 - \exp(-\Omega(\ln^2 n)). \quad (6.49) \]
Furthermore, let \( G' \) be the graph obtained from \( G \) by removing all edges \( e \) that are incident with a vertex \( v \) such that \( |\partial^\omega_G(v)| \geq \lambda \) and let
\[ S' = \sum_v S_v(G', \sigma) = \left| \left\{ v \in [n'] : |\partial^\omega_{G'}(v), \sigma|_{\partial^\omega_{G'}(v), v} \in S \right\} \right|. \]
If \( \Gamma \) occurs, then \( S = S' \). Hence, (6.49) implies that
\[ \mathbb{E}[S'] = \mathbb{E}[S] + o(1). \quad (6.50) \]
Moreover, the random variable \( S' = f(X_2, \ldots, X_N) \) satisfies (6.48) with \( c = \lambda \) and \( c' = n' \). Indeed, altering either the color of one vertex \( u \) or its set of neighbors can only affect those vertices \( v \) that are at distance at most \( \omega \) from \( u \), and in \( G' \) there are no more than \( \lambda \) such vertices. Thus, Lemma 6.10 applied with, say, \( t = n^{2/3} \) and \( \gamma = 1/n \) and (6.49) yields

\[
P \left[ |S' - \mathbb{E}[S']| > t \right] \leq \exp(-\Omega(\ln^2 n)) = o(1).
\] (6.51)

Finally, the assertion follows from (6.50) and (6.51).

Acknowledgements. We thank Guilhem Semerjian for helpful discussions and explanations regarding the articles [20,33] and Nick Wormald for pointing us to [22, Theorem 3.8].

References

1. Achlioptas, D., Friedgut, E.: A sharp threshold for \( k \)-colorability. Random Struct. Algorithms 14, 63–70 (1999)
2. Achlioptas, D., Coja-Oghlan, A.: Algorithmic barriers from phase transitions. In: Proceedings of the 49th FOCS, pp. 793–802 (2008)
3. Babst, V., Coja-Oghlan, A., Efthymiou, C.: Planting colourings silently (2014). arXiv:1411.0610
4. Achlioptas, D., Naor, A.: The two possible values of the chromatic number of a random graph. Ann. Math. 162, 1333–1349 (2005)
5. Bordenave, C., Caputo, P.: Large deviations of empirical neighborhood distribution in sparse random graphs. Probab. Theory Relat. Fields (2013, to appear). arXiv:1308.5725
6. Braunstein, A., Mulet, R.A., Pagnani, M., Weigt, Zecchina, R.: Polynomial iterative algorithms for coloring and analyzing random graphs. Phys. Rev. E 68, 036702 (2003)
7. Coja-Oghlan, A.: Upper-bounding the \( k \)-colorability threshold by counting covers. Electron. J. Combin. 20, P32 (2013)
8. Coja-Oghlan, A., Vilenchik, D.: Chasing the \( k \)-colorability threshold. In: Proceedings of the 54th FOCS, pp. 380–389 (2013). A full version is available as arXiv:1304.1063
9. Coja-Oghlan, A., Zdeborová, L.: The condensation transition in random hypergraph 2-coloring. In: Proceedings of the 23rd SODA, pp. 241–250 (2012)
10. Contucci, P., Dommer, S., Giardina, C., Staar, S.: Antiferromagnetic Potts model on the Erdős–Rényi random graph. Commun. Math. Phys. 323, 517–554 (2013)
11. Dembo, A., Montanari, A.: Gibbs measures and phase transitions on sparse random graphs. Braz. J. Probab. Stat. 24(2), 137–211 (2010)
12. Derrida, B.: Random-energy model: an exactly solvable model of disordered systems. Phys. Rev. B 24(5), 2613–2626 (1981)
13. Derrida, B., Spohn, H.: Polymers on disordered trees, spin glasses, and traveling waves. J. Stat. Phys. 51(5–6), 817–840 (1988)
14. Erdős, P., Rényi, A.: On the evolution of random graphs. Magyar Tud. Akad. Mat. Kutato Int. Kozl. 5, 17–61 (1960)
15. Frieze, A., McDiarmid, C.: Algorithmic theory of random graphs. Random Struct. Algorithms 10, 5–42 (1997)
16. Janson, S., Łuczak, T., Ruciński, A.: Random Graphs. Wiley, New York (2000)
17. Kauzmann, W.: The nature of the glassy state and the behavior of liquids at low temperatures. Chem. Rev. 43, 219–256 (1948)
18. Kirkpatrick, T.R., Thirumalai, D.: \( p \)-spin-interaction spin-glass models: connections with the structural glass problem. Phys. Rev. B 36, 5388 (1987)
19. Krzakala, F., Mézard, M., Saussere, F., Sun, Y., Zdeborová, L.: Probabilistic reconstruction in compressed sensing: algorithms, phase diagrams, and threshold achieving matrices. J. Stat. Mech. P08009 (2012)
20. Krzakala, F., Montanari, A., Ricci-Tersenghi, F., Semerjian, G., Zdeborová, L.: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. Natl. Acad. Sci. 104, 10318–10323 (2007)
21. Krzakala, F., Pagnani, A., Weigt, M.: Threshold values, stability analysis and high-\( q \) asymptotics for the coloring problem on random graphs. Phys. Rev. E 70, 046705 (2004)
22. McDiarmid, C.: Concentration. In: Habib et al. (eds.) Probabilistic Methods for Algorithmic Discrete Mathematics, pp. 195–248. Springer, Berlin (1998)
23. Mézard, M., Montanari, A.: Information, Physics and Computation. Oxford University Press, Oxford (2009)
24. Mézard, M., Parisi, G., Zecchina, R.: Analytic and algorithmic solution of random satisfiability problems. Science 297, 812–815 (2002)
25. Molloy, M.: The freezing threshold for $k$-colourings of a random graph. In: Proceedings of the 43rd STOC, pp. 921–930 (2012)
26. Parisi, G.: A sequence of approximated solutions to the S-K model for spin glasses. J. Phys. A 13(4), L115 (1980)
27. Ricci-Tersenghi, F., Semerjian, G.: On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms. J. Stat. Mech. P09001 (2009)
28. Sherrington, D., Kirkpatrick, S.: Solvable model of a spin-glass. Phys. Rev. Lett. 35(26), 1792–1796 (1975)
29. Talagrand, M.: The Parisi formula. Ann. Math. 163, 221 (2006)
30. Talagrand, M.: Spin Glasses, a Challenge for Mathematicians. Springer, Berlin (2003)
31. Warnke, L.: On the method of typical bounded differences Comb. Prob. Comput. (2012, to appear). arXiv:1212.5796
32. Zdeborová, L.: Statistical physics of hard optimization problems. Acta Physica Slovaca 59, 169–303 (2009)
33. Zdeborová, L., Krzakala, F.: Phase transition in the coloring of random graphs. Phys. Rev. E 76, 031131 (2007)

Communicated by H.-T. Yau