On the connectivity threshold for colorings of random graphs and hypergraphs

Michael Anastos and Alan Frieze*
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA 15213

August 27, 2018

Abstract

Let $\Omega_q = \Omega_q(H)$ denote the set of proper $[q]$-colorings of the hypergraph H. Let Γ_q be the graph with vertex set Ω_q and an edge $\{\sigma, \tau\}$ where σ, τ are colorings iff $h(\sigma, \tau) = 1$. Here $h(\sigma, \tau)$ is the Hamming distance $|\{v \in V(H) : \sigma(v) \neq \tau(v)\}|$. We show that if $H = H_{n,m;k}$, $k \geq 2$, the random k-uniform hypergraph with $V = [n]$ and $m = dn/k$ hyperedges then w.h.p. Γ_q is connected if d is sufficiently large and $q \gtrsim (d/\log d)^{1/(k-1)}$. Furthermore, with a few more colors, we find that the diameter of Γ_q is $O(n)$ w.h.p, where the hidden constant depends on d.

1 Introduction

In this paper, we will discuss a structural property of the set Ω_q of proper $[q]$-colorings of the random hypergraph $H = H_{n,m;k}$, where $m = dn/k$ for some large constant d. Here H has vertex set $V = V(H) = [n]$ and an edge set $E = E(H)$ consisting of m randomly chosen k-sets from $\binom{[n]}{k}$. Note that in the graph case where $k = 2$ we have $H_{n,m;2} = G_{n,m}$. A proper $[q]$-coloring is a map $\sigma : [n] \to [q]$ such that $|\sigma(e)| \geq 2$ for all $e \in E$ i.e. no edge is monochromatic. Then let us define $\Gamma_q = \Gamma_q(H)$ to be the graph with vertex set Ω_q and an edge $\{\sigma, \tau\}$ iff $h(\sigma, \tau) = 1$ where $h(\sigma, \tau)$ is the Hamming distance $|\{v \in [n] : \sigma(v) \neq \tau(v)\}|$. In the Statistical Physics literature the definition of Γ_q may be that colorings σ, τ are connected by an edge in Γ_q whenever $h(\sigma, \tau) = o(n)$. Our theorem holds a fortiori if this is the case.

*Research supported in part by NSF grant DMS1661063

Notation: $f(d) \gtrsim g(d)$ if there exists a function $\varepsilon(d) > 0$ such that $\lim_{d\to\infty} \varepsilon(d) = 0$ and $f(d) \geq (1 + \varepsilon(d))g(d)$ for d large.

Then let

$$\alpha = \left(\frac{(k-1)d}{\log d - 5(k-1)\log\log d} \right)^{\frac{1}{k-1}}, \quad \beta = 3\log^{3k} d. \tag{1}$$

We prove the following.

Theorem 1.1. *Suppose that $k \geq 2$ and $p = \frac{d}{\binom{n-1}{k-1}}$ and $m = \binom{n}{k}p$ and that $d = O(1)$ is sufficiently large. Then*

(i) If $q \geq \alpha + \beta + 1$ then w.h.p. Γ_q is connected.

(ii) If $q \geq \alpha + 2\beta + 1$ then the diameter of Γ_q is $O(\alpha\beta n)$ w.h.p.

Note that Γ_q connected implies that "The Glauber Dynamics on Ω_q is ergodic". At the moment we only know that Glauber Dynamics is rapidly mixing w.h.p. when $q \geq (1.76\ldots)d$, see Efthymiou, Hayes, Štefankovič and Vigoda [12]. So, it would seem that the connectivity of Γ_q is not likely to be a barrier to randomly sampling colorings of sparse random graphs.

We note that the lower bound for q is close to where the greedy coloring algorithm succeeds w.h.p. For the case $k = 2$ this follows from Shamir and Upfal [20]. For $k \geq 3$, the authors could not find relevant literature. Nevertheless the claim follows (partially) from the current paper. In particular, Lemmas 4.4 and 4.5 show that the greedy coloring algorithm uses at most $\alpha + \beta$ colors. Furthermore, a simple argument based on the size of an independent set selected by the greedy algorithm shows that the number of colors required is close to α.

We should note that in the case $k = 2$, that Molloy [19] has shown that w.h.p. there is no giant component in Γ_q if $q \lesssim \frac{d}{\log d}$. No corresponding result is known to the authors for $k \geq 3$. It is somewhat surprising therefore that w.h.p. Γ_q jumps very quickly from having no giant to being connected. One might have expected that $q \gtrsim \frac{d}{\log d}$ would simply imply the existence of a giant component. In Physics terminology, this implies a short *non-reconstruction* phase between *uniqueness* and *reconstruction*.

Prior to this paper, it was shown in [11] that w.h.p. $\Gamma_q, q \geq d+2$ is connected. The diameter of the *reconfiguration graph* $\Gamma_q(G)$ for graphs G has been studied in the graph theory litrature, see Bousquet and Perarnau [8] and Feghali [13]. They show that if the maximum sub-graph density of a graph is at most $d-\varepsilon$ and $q \geq d+1$ then $\Gamma_q(G)$ has polynomial diameter. Using Theorem 1 of [8] we can show a linear bound on the diameter with a small increase in the number of colors, See (ii) of Theorem 1.1.

Theorem 1.1 falls into the area of "Structural Properties of Solutions to Random Constraint Satisfaction Problems". This is a growing area with connections to Computer Science and Theoretical Physics. In particular, much of the research on the graph Γ_q has been focussed on the structure near the *colorability threshold*, e.g. Bapst, Coja-Oghlan, Hetterich, Rassman and Vilenchik [5], or the *clustering threshold*, e.g. Achlioptas, Coja-Oghlan and Ricci-Tersenghi [2], Molloy [19]. Other papers heuristically identify a sequence of phase transitions in the structure of Γ_q, e.g., Krząkala, Montanari, Ricci-Tersenghi, Semerijan and Zdeborová [18], Zdeborová and Krząkala [21] or Gabrié, Dani, Semerjian and Zdeborová [15]. The existence of these transitions has been shown rigorously for some other CSPs. One of the most spectacular examples is due to Ding, Sly and Sun [10] who rigorously showed the existence of a sharp satisfiability threshold for random k-SAT.

Section 3 describes a property (α, β)-colorability such that if H has this property then $q \geq \alpha + \beta + 1$ implies that Γ_q is connected. Section 4 proves that $H_{n,m;k}, k \geq 2$, is (α, β)-colorable for α, β defined in (1).

The paper uses some of the ideas from [4] which showed there is a giant component in $\Gamma_q(G_{n,m})$, $m = dn/2$ w.h.p. when $q \geq cd/\log d$ for $c > 3/2$.

2 Outline argument

We show that with the values $\alpha \approx ((k-1)d/\log d)^{1/(k-1)} \gg \beta$ given in (1) then w.h.p. $H = H_{n,m;k}$ has the property that **any** greedy coloring of H will need at most α maximal independent sets before being left with a graph without a β-core. (See Lemma 4.4.) We call the colorings found in this way, *good greedy colorings* and we refer to this property as (α, β)-colorability. Any good (α, β)-coloring uses at most $\alpha + \beta$ colors. It follows from this, basically using the argument from [4], that if $\sigma \in \Omega_q$ and $q \geq \alpha + \beta + 1$ then there is a *good* path in Γ_q to some good greedy coloring σ_1.

Suppose now that σ_1, τ_1 are good greedy colorings. If $q \geq \alpha + \beta + 1$ then there is a color c that is not used by σ_1. From σ_1 we move to σ_2 by re-coloring vertices colored 1 in σ_1 by c. Then we move from σ_2 to σ_3 by coloring with color 1, all vertices that have color 1 in τ_1. At this point, σ_3 and τ_1 agree on color 1. σ_3 may use more than $\alpha + \beta$ colors and so we move by a good path from σ_3 to a coloring σ_4 that uses at most $\alpha + \beta$ colors and does not change the color of any vertex currently with color 1. Here we use the fact that $H_{n,m;k}$ is (α, β)-colorable. After this, it is induction that completes the proof.

3 (α, β)-colorability

The degree of a vertex $v \in V$ in a hypergraph $H = (V, E)$ is the number of edges $e \in E$ such that $v \in e$. (For completeness, we will state several things in this short paper that one might think can be taken for granted.)

Let $H = (V, E)$. A β-core of H is a maximal subgraph of H in which every vertex has degree at least β. For every $U \subset V$, if the subgraph of H induced by U does not have a β-core then there is an ordering $\{u_1, u_2, ..., u_{|U|}\}$ of the vertices in U such that every vertex in U has at most $\beta - 1$ neighbors that precede it in that ordering.

If a hypergraph H that does not have a β-core then we can color it with at most β colors. Let $v_1, v_2, ..., v_n$ be an ordering on V where

$$\text{for every } i, \text{ there are at most } \beta - 1 \text{ edges that contain } v_i \text{ and are contained in } \{v_1, v_2, \ldots, v_i\}. \quad (2)$$

Such an ordering must exist when there is no β-core. We color the vertices in the order $v_1, v_2, \ldots, v_n$ and assign to v_i a color that is not *blocked* by the $\beta - 1$ neighbors that precede it. A color c is blocked for vertex v by vertices $w_1, w_2, \ldots, w_{k-1}$ if $e = \{v, w_1, \ldots, w_{k-1}\} \in E(H)$ and $w_1, w_2, \ldots, w_{k-1}$ have already been given color c.

Next let $V_1, V_2, \ldots, V_\alpha$ be a sequence of disjoint independent sets of H such that for each $j \geq 1$, V_j is maximal in the sub-hypergraph H_j induced by $V \setminus \bigcup_{1 \leq i < j} V_i$. We say that such a sequence is a *maximally independent sequence of length* α. Note that we allow $V_j = \emptyset$ here, in order to make our sequences of length exactly α.

Definition 3.1. *We say that a hypergraph H is (α, β)-colorable if there* **does not exist** *a maximally independent sequence of length α such that $V \setminus \bigcup_{i \leq \alpha} V_i$ has a β-core.*

The main result of this section is the following.

Theorem 3.2. *Let H be (α, β)-colorable and let $q \geq \alpha + \beta + 1$. Then $\Gamma_q(H)$ is connected.*

Later, in Section 4 we will show that $H_{n,m;k}, k \geq 2$ is (α, β)-colorable, for suitable values of m, α, β, viz. the values given in (1).

Lemma 3.3. *Let $H = (V, E)$ be an (α, β)-colorable hypergraph and $V_1 \subseteq V$ be a maximal independent set of V. Set $V' = V \setminus V_1$ and let H' be the subgraph of H induced by V'. Then H' is $(\alpha - 1, \beta)$-colorable.*

Proof. Assume that H' in not $(\alpha - 1, \beta)$-colorable. Then there exists a partition of V' into $V'_1, ..., V'_{\alpha-1}$ such that for $j \in [\alpha - 1]$, V'_j is a maximal independent set of $V' \setminus \bigcup_{\ell < j} V'_\ell$ and $W' = V' \setminus \bigcup_{\ell \leq \alpha - 1} V'_\ell$ has a β-core. For $j \in [\alpha - 1]$ set $V_{j+1} = V'_j$. Furthermore set $W = V \setminus (\bigcup_{1 \leq \ell \leq \alpha} V_\ell) = V' \setminus (\bigcup_{\ell \leq \alpha - 1} V'_\ell) = W'$. Then $V_1, ..., V_\alpha$ is a maximal independent sequence of length α and W has a β-core which contradicts the fact that H is (α, β)-colorable. $\square$

Lemma 3.4. *Let H be a hypergraph, $\alpha, \beta \geq 0$ and $q \geq \alpha + \beta + 1$. Let $W \subseteq V$ be such that the subgraph of H induced by W has no β-core. Furthermore let χ and τ be two colorings of H such that* *(i) They agree on $V \setminus W$.*

(ii) They use only α colors on the vertices in $V \setminus W$.

(iii) τ uses at most β colors on W that are distinct from the ones it uses on $V \setminus W$.

Then there exists a path from χ to τ in $\Gamma_q(H)$.

Proof. Without loss of generality we may assume that χ and τ use $[\alpha]$ to color $V \setminus W$. The proof that follows is an adaptation to hypergraphs of the proof in [4] that $\Gamma_q(G)$ is connected when a graph G has no q-core. Because W has no β-core there exists an ordering of its vertices, $v_1, v_2, ..., v_r$, such that for $i \in [r]$, v_i has at most $\beta - 1$ neighbors in $v_1, v_2, \ldots, v_{i-1}$. For $0 \leq i \leq r$ let τ_i be the coloring that agrees with τ on $\{v_1, ..., v_i\}$ and with χ on $W \setminus \{v_1, ..., v_i\}$. On $V \setminus W$ it agrees with both. Thus $\tau_0 = \chi$ and $\tau_r = \tau$. We note that $\tau_1, \tau_2, \ldots, \tau_{r-1}$ may not be proper colorings.

We proceed by induction on i to show that there is a sequence of colorings Σ_i from χ to τ_i such that (i) going from one coloring to the next in Σ_i only re-colors one vertex and (ii) all colorings in the sequence Σ_i are proper for the hypergraph induced by $V \setminus \{v_{i+1}, ..., v_r\}$. We **do not** claim that the colorings in $\Sigma_i, i < r$ are proper for H. On the other hand, taking $i = r$ we get a sequence of H-proper colorings that starts with χ, ends with τ, such that the consecutive pairs of proper colorings differ on a single vertex. Clearly, such a sequence corresponds to a path from χ to τ in $\Gamma_q(H)$.

The case $i = 1$ is trivial as we have assumed that σ, τ agree on $V \setminus W$ and so we can give v_1 the color $\tau(v_1)$. Assume that the assertion is true for $i = k \geq 1$ and let $\chi = \psi_0, \psi_1, \ldots, \psi_s = \tau_k$ be a sequence of colorings promised by the inductive ssertion. Let (w_j, c_j) denote the $(vertex, color)$ change defining the move from ψ_{j-1} to ψ_j. We construct a sequence of colorings of length at most $2s + 1$ that yields the assertion for $i = k + 1$. For $j = 1, 2, \ldots, s$, we will re-color w_j to color c_j, unless there exists a set X such that $X \cup \{w_j\} \in E$ and $\psi_{j-1}(x) = c_j, x \in X \subseteq \{v_1, v_2, \ldots, v_{k+1}\}$. The fact that ψ_j is a proper coloring of $V \setminus \{v_{k+1}, ..., v_r\}$ implies that $v_{k+1} \in X$. Because v_{k+1} has at most $\beta - 1$ neighbors in $\{v_1, ..., v_k\}$ and τ only uses colors in $[\alpha]$ to color $V \setminus W$, there exists a color $c' \neq c_j$ for v_{k+1} in $[\alpha + \beta + 1] \setminus [\alpha]$ which is not blocked by a subset of $\{v_1, v_2, \ldots, v_k\}$ and is diferent from its current color. We first re-color v_{k+1} to c' and then we re-color w_j to c_j, completing the inductive step. At the very end, i.e. at step $2s + 1$ we give v_{k+1} its color in τ. □

Definition 3.5. *A coloring with color sets $V_1, V_2, \ldots, V_{\alpha+\beta}$ is said to be a* good greedy coloring *if (i) $V_1, V_2, \ldots, V_\alpha$ is a maximally independent sequence of length α and (ii) $V \setminus \bigcup_{\ell \leq \alpha} V_\ell$ has no β-core.*

We prove Theorem 3.2 in two steps. In Lemma 3.6, we show that if $q \geq \alpha + \beta + 1$ and H is $(\alpha, \beta) - colorable$ then we can reach a good greedy coloring in $\Gamma_q(H)$ starting from any coloring. Then in Lemma 3.8, we show that if $q \geq \alpha + \beta + 1$ then any good greedy coloring τ can be reached in $\Gamma_q(H)$ from any other good greedy coloring σ.

Lemma 3.6. *Let H be an (α, β)-colorable hypergraph, $q \geq \alpha+\beta+1$ and χ be a $[q]$-coloring of H. Then there exists a good greedy coloring τ of H such that there exists a path in $\Gamma_q(H)$ from χ to τ.*

Proof. We generate the coloring τ as follows. Let $C_1, C_2, \ldots, C_q$ be the color classes of χ. Then let $V_1 \supseteq C_1$ be a maximal independent set containing C_1. In general, having defined $V_1, V_2, \ldots, V_{\ell-1}$ we let $V_{<\ell} = \bigcup_{1 \leq i < \ell} V_i$ and then we let V_ℓ be a maximal independent set in $V \setminus V_{<\ell}$ that contains $C_\ell \setminus V_{<\ell}$. Thus $V_1, V_2, \ldots, V_\alpha$ is a maximal independent sequence of length α. We now describe how we transform the coloring χ vertex by vertex into a coloring χ' in which vertices in V_i get color i for $1 \leq i \leq \alpha$. We first re-color the vertices in $V_1 \setminus C_1$ by giving them color 1, one vertex at a time. The coloring stays proper, as V_1 is an independent set. In general, having re-colored $V_1, V_2, \ldots, V_{\ell-1}$ we re-color the vertices in $V_\ell \setminus C_\ell$ with color ℓ. Again, the coloring stays proper, as V_ℓ is an independent set, containing all vertices in C_ℓ that have not been re-colored. We observe that each re-coloring of a vertex v done while turning χ into χ' can be interpreted as moving from a coloring in $\Gamma_q(H)$ to a neighboring coloring.

Let $W = V \setminus \bigcup_{1 \leq i \leq \alpha} V_i$. Because H is (α, β)-colorable, we find that W has no β-core. Because W has no β-core there exists a proper coloring τ' of the subgraph of H induced by W that uses only colors in $[\alpha+\beta+1] \setminus [\alpha]$. Set τ to be the coloring that agrees with χ' on $V \setminus W$ and with τ' on W.

Lemma 3.4 implies that there is a path from χ' to τ. Hence there is a path from χ to τ. $\square$

Remark 3.7. *In the proof of Lemma 3.6 we see that each vertex is re-colored at most twice before we apply Lemma 3.4. Thus this part of the proof yields at most α distinct sub-paths of length $O(n)$.*

Lemma 3.8. *Let H be an (α, β)-colorable hypergraph, $q \geq \alpha+\beta+1$ and let χ, τ be two good greedy colorings. Then there exists a path from χ to τ in $\Gamma_q(H)$.*

Proof. We proceed by induction on α. For $\alpha = 0$, H is $(0, \beta)$ colorable and so it does not have a β-core. Thus the base case follows directly from Lemma 3.4 by taking $W = V$.

Assume that the statement of the Lemma is true for $\alpha = k-1$ and let $\alpha = k$. There exists a maximal independent sequence $V_1, V_2, \ldots, V_k$ of length k such that if $V' = V \setminus \bigcup_{1 \leq i \leq k} V_i$ then (i) for $i \in [k]$, τ assigns the color i to $v \in V_i$ and (ii) τ assigns only colors in $[k+\beta] \setminus [k]$ to vertices in V'.

Let c be a color not assigned by χ. There is one as $q \geq k+\beta+1$. Starting from χ we recolor all vertices that are colored 1 by color c to create a coloring $\bar{\chi}$. Then we continue from $\bar{\chi}$ by recoloring all the vertices in V_1 by color 1 and we let χ' be the resulting coloring. Clearly there is a path P_1 from χ to χ' in $\Gamma_q(H)$.

We now set $H_1 = H \setminus V_1$, and set χ_1', τ_1 to be the restrictions of χ', τ on H_1. Observe that since V_1 is a maximal independent set, Lemma 3.3 implies that H_1 is $(k-1, \beta)$ colorable and in addition that τ_1 is a good greedy coloring of H_1. Lemma 3.6 implies that in $\Gamma_{q-1}(H_1)$ there is a path P_2 from χ_1' to some good greedy coloring χ_1 that uses only $k-1+\beta$ colors from $[q] \setminus \{1\}$. The induction hypothesis implies that in $\Gamma_{q-1}(H_1)$ that there is a path P_3 from χ_1 to τ_1.

Color 1 is not used in χ_1', τ_1 or in any of colorings found in the path P_2, P_3. Thus the path P_2, P_3 corresponds to a path P_4 in $\Gamma_q(H)$ from χ' to τ. Consequently the colorings χ, τ are connected in $\Gamma_q(H)$ by the path $P_1 + P_4$. □

Proof of Theorem 3.2: Let H be (α, β) colorable, $q \geq \alpha + \beta + 1$, and let χ_1, χ_2 be two colorings of H. Lemma 3.6 implies that in $\Gamma_q(H)$, there is path P_i from χ_i to a good greedy coloring τ_i for $i = 1, 2$. Lemma 3.8 implies that there is a path in $\Gamma_q(H)$ from τ_1 to τ_2. □

4 Random Hypergraphs

Theorem 1.1 follows from

Lemma 4.1. *Let $k \geq 2$ and suppose that $q \geq \alpha + \beta + 1$ and that d is sufficiently large. If $p = \frac{d}{\binom{n-1}{k-1}}$ and $m = \binom{n}{k} p$ then w.h.p. $\Gamma_q(H_{n,m;k})$ is connected.*

In the following we will assume for simplicity of notation that $d = O(1)$, so that $O(f(d)/n) = O(1/n)$. We do not know if there is an upper bound needed for the growth rate of d, but we doubt it.

To prove Lemma 4.1 we use Lemmas 4.2, 4.4, 4.5 (below) in order to deduce that w.h.p. $G_{n,dn/2}$ is (α, β) colorable. Then we apply Theorem 3.2. (Lemmas 4.2 and 4.5 are hardly new or best possible, but we prove them here for completeness.)

We will do our calculations on the random graph $H_{n,p;k}, p = d/\binom{n-1}{k-1}$ and use the fact for any hypergraph property $\mathcal{P}$, we have

$$\mathbf{Pr}(H_{n,m;k} \in \mathcal{P}) \leq O(m^{1/2})\, \mathbf{Pr}(H_{n,p;k} \in \mathcal{P}). \tag{3}$$

Lemma 4.2. *Let $p = \frac{d}{\binom{n-1}{k-1}}$ and $k \geq 2$ and d sufficiently large. Then, w.h.p. $H = H_{n,p;k}$ does not contain an independent set of size $\left(\frac{2k \log d}{(k-1)d}\right)^{\frac{1}{k-1}} n$.*

Proof. Let $u = \left(\frac{2k \log d}{(k-1)d}\right)^{\frac{1}{k-1}} n$. The probability that there exists an independent set of size u in H is bounded by

$$\binom{n}{u}(1-p)^{\binom{u}{k}} \leq \left(\frac{en}{u}\right)^u \exp\left\{-\frac{d}{\binom{n-1}{k-1}} \cdot \binom{u}{k}\right\}$$

$$
\begin{aligned}
&\leq \left(\frac{en}{u}\right)^u \exp\left\{-\frac{du}{k}\left(\frac{u}{n}\right)^{k-1}\left(1+O\left(\frac{1}{n}\right)\right)\right\} \\
&= \left(e^{k-1}\frac{(k-1)d}{2k\log d}\cdot\exp\left\{-2\log d\left(1+O\left(\frac{1}{n}\right)\right)\right\}\right)^{u/(k-1)} \\
&= \left(\frac{e^{k-1}(k-1)}{2kd\log d}\left(1+O\left(\frac{1}{n}\right)\right)\right)^{u/(k-1)} \qquad (4)\\
&= o(1).
\end{aligned}
$$

□

Notation 4.3. *We let*

$$m_0 = \frac{n}{\alpha} \text{ and } n_0 = 16 m_0 \log^2 d.$$

Furthermore, for $t \leq d$ we let

$$S_t = \left\{(s_1, s_2, ..., s_t) \in \left[\left(\frac{2k\log d}{(k-1)d}\right)^{\frac{1}{k-1}} n\right]^t : \sum_{j=1}^{t} s_i \leq \min\{tm_0, n-n_0\}\right\}.$$

Lemma 4.4. *If $k \geq 2$ and d is sufficiently large then, w.h.p.* **there does not exist** *$1 \leq t \leq d$ and disjoint sets $V_1, ..., V_t \subset V$ such that:*

(i) $V_1, V_2, \ldots, V_t$ is a maximal independent sequence of length t in $H = H_{n,p;k}$.

(ii) $(|V_1|, |V_2|, ..., |V_t|) \in S_t$.

Proof. Fix $t \in [d]$, $(s_1, ..., s_t) \in S_t$ and let $\bar{s} = \frac{1}{t}\sum_{i\in[t]} s_i$. Since $(s_1, ..., s_t) \in S_t$ we have that $\bar{s} \leq \frac{1}{t}\cdot tm_0 = m_0$. There are $\binom{n}{s_1, s_2, ..., s_t, n-t\bar{s}}$ ways to pick disjoint sets $V_1, V_2, ..., V_t \subseteq V$ of sizes $s_1, ..., s_t$ respectively. So $V_1, ..., V_t$ satisfy condition (i) of Lemma 4.4 only if for every $i \in [t]$ and every $v \in V \setminus \bigcup_{j\in[i]} V_j$, there exist $u_1, ..., u_{k-1} \in V_i$ such that $\{u_1, ..., u_{k-1}, v\} \in E(H)$. So, given $V_1, ..., V_t$ the probability that we have (i) is at most

$$p_1 = \prod_{i=1}^{t}(1-(1-p)^{\binom{s_i}{k-1}})^{n-\sum_{j=1}^{i} s_j} \leq \exp\left\{-\sum_{i=1}^{t}\left((1-p)^{\binom{s_i}{k-1}}\left(n-\sum_{j=1}^{i} s_j\right)\right)\right\}. \qquad (5)$$

Now let $t' = \max\left\{i : \sum_{j\leq i} s_j \leq n - \frac{n}{\log^2 d}\right\}$ and $s' = \sum_{i=1}^{t'} s_i$ and $\bar{s}' = \frac{s'}{t'}$. We consider 2 cases.

Case 1: $t' \geq \left(1-\frac{1}{\log d}\right)t$.

Now $t\bar{s} \geq t'\bar{s}'$ and so $\bar{s}' - \bar{s} \leq \frac{t-t'}{t}\bar{s}' \leq \frac{\bar{s}'}{\log d}$, which implies that $\bar{s}' \leq \bar{s}\left(1-\frac{1}{\log d}\right)^{-1} \leq m_0\left(1+\frac{2}{\log d}\right)$. Then,

$$\sum_{i=1}^{t}\left((1-p)^{\binom{s_i}{k-1}}\left(n-\sum_{j=1}^{i} s_j\right)\right)$$

$$\geq \sum_{i=1}^{t'} \left((1-p)^{\binom{s_i}{k-1}} \left(n - \sum_{j=1}^{i} s_j \right) \right)$$

$$\geq \frac{n}{\log^2 d} \sum_{i=1}^{t'} (1-p)^{\binom{s_i}{k-1}} \geq \frac{nt'}{\log^2 d}(1-p)^{\binom{\bar{s}'}{k-1}} \geq \frac{nt}{2\log^2 d}(1-p)^{\binom{m_0\left(1+\frac{2}{\log d}\right)}{k-1}}$$

$$\geq \frac{nt}{2\log^2 d} \exp\left\{ -(p+p^2) \binom{\left(\frac{(\log d - 5(k-1)\log\log d)}{(k-1)d}\right)^{1/(k-1)} \left(1+\frac{2}{\log d}\right) n}{k-1} \right\}$$

$$\geq \frac{nt}{2\log^2 d} \exp\left\{ -\frac{\log d - 5(k-1)\log\log d}{k-1} \cdot \left(1 + \frac{3(k-1)}{\log d}\right) \right\}$$

$$\geq \frac{nt\log^2 d}{d^{1/(k-1)}}.$$

Now

$$\binom{n}{s_1, \ldots, s_t, n - t\bar{s}} \leq \binom{n}{\bar{s}, \ldots, \bar{s}, n - t\bar{s}} \leq \prod_{i=1}^{t} \binom{n}{\bar{s}} \leq \left(\frac{en}{\bar{s}}\right)^{t\bar{s}} \leq \left(\frac{en}{m_0}\right)^{tm_0}.$$

Thus the probability that for some $t \leq d$ there exist $V_1, \ldots, V_t$ satisfying conditions (i), (ii) of Lemma 4.4 is bounded by

$$\sum_{t=1}^{d} \sum_{(s_1,\ldots,s_t)\in S_t} \binom{n}{s_1, s_2, \ldots, s_t, n - \sum_{i\in[t]} s_i} p_1$$

$$\leq \sum_{t=1}^{d} \sum_{(s_1,\ldots,s_t)\in S_t} \left(\frac{en}{m_0}\right)^{tm_0} \exp\left\{ -\frac{nt\log^2 d}{d^{1/(k-1)}} \right\} \leq \sum_{t=1}^{d} n^t \left(\frac{(e\alpha)^{(\log d)^{1/(k-1)}}}{d^{\log d}} \right)^{nt/d^{1/(k-1)}} = o(1).$$

Case 2: $t' < \left(1 - \frac{1}{\log d}\right)t$.
Thus $t - t' \geq \frac{t}{\log d}$. Observe that from Lemma 4.2 we can assume that

$$t \geq t' \geq \left(\left(1 - \frac{1}{\log^2 d}\right) \Big/ \left(\frac{2k\log d}{(k-1)d}\right)^{\frac{1}{k-1}} \right) - 1 \geq \frac{1}{4}\left(1 - \frac{1}{\log^2 d}\right)\left(\frac{d}{\log d}\right)^{\frac{1}{k-1}}. \tag{6}$$

For (6) we are using Lemma 4.2 to argue that we need at least this many independent sets to partition a set of size $n\left(1 - \frac{1}{\log^2 d}\right)$. The -1 comes from the fact that the upper bound in the definition of t' may not be tight.

Thus,

$$u = \frac{1}{t-t'} \sum_{i=t'+1}^{t} s_i \leq \frac{\log d}{t} \cdot n \left(\frac{1}{\log^2 d} + \left(\frac{2k\log d}{(k-1)d}\right)^{\frac{1}{k-1}} \right)$$

$$\leq 4\left(1 + \frac{2}{\log^2 d}\right)\left(\frac{\log d}{d}\right)^{\frac{1}{k-1}} \cdot \frac{n}{\log d} \tag{7}$$

and now with p_1 as defined in (5) we have

$$
\begin{aligned}
p_1 &\le \prod_{i=t'+1}^{t} (1-(1-p)^{\binom{s_i}{k-1}})^{n-\sum_{j=1}^{i} s_j} \le \prod_{i=t'+1}^{t} (1-(1-p)^{\binom{s_i}{k-1}})^{n_0} \\
&\le \exp\left\{-n_0 \sum_{i=t'+1}^{t} (1-p)^{\binom{s_i}{k-1}}\right\} \\
&\le \exp\left\{-n_0(t-t') \exp\left\{-(p+p^2)\binom{u}{k-1}\right\}\right\} \\
&\le \exp\left\{-n_0(t-t') \exp\left\{-d\left(\frac{u}{n}\right)^{k-1}\right\}\left(1+O\left(\frac{1}{n}\right)\right)\right\} \\
&\le \exp\left\{-n_0(t-t') \exp\left\{-\frac{4^k\left(1+\frac{2}{\log^2 d}\right)^k}{\log^{k-2} d}\right\}\right\} \\
&\le e^{-(t-t')n_0/2}.
\end{aligned}
$$

Thus the probability that for some $t \le d$ there exist $V_1, ..., V_t$ satisfying conditions (i), (ii) of Lemma 4.4 is bounded by

$$
\begin{aligned}
P = &\sum_{t=1}^{d} \sum_{(s_1,\dots,s_t)\in S_t} \prod_{i=1}^{t'} \binom{n-\sum_{j=1}^{i-1} s_j}{s_i} \prod_{i=t'+1}^{t} \binom{n-\sum_{j=1}^{i-1} s_j}{s_i} p_1 \\
&\qquad \le \sum_{t=1}^{d} \sum_{(s_1,\dots,s_t)\in S_t} \left(\frac{en}{\bar{s}'}\right)^{t'\bar{s}'} \left(\frac{en}{u}\right)^{(t-t')u} e^{-(t-t')n_0/2}.
\end{aligned}
$$

For sufficiently large d, (7) implies $u \le m_0$ and we also have that $n_0 = 16m_0 \log^2 d$. Therefore

$$
\left(\frac{en}{u}\right)^{(t-t')u} e^{-(t-t')n_0/4} \le \left(\frac{en}{m_0}\right)^{(t-t')m_0} e^{-4(t-t')m_0\log^2 d} \le e^{-3(t-t')m_0\log^2 d} \le e^{-3tm_0\log d}.
$$

Furthermore, Lemma 4.2 implies that $\bar{s}' \le \left(\frac{2k\log d}{(k-1)d}\right)^{\frac{1}{k-1}} n \le 3m_0$. Thus

$$
\left(\frac{en}{\bar{s}'}\right)^{t'\bar{s}'} e^{-(t-t')n_0/4} \le \left(\frac{en}{3m_0}\right)^{3tm_0} e^{-4(t-t')m_0\log^2 d} \le \left(\frac{en}{3m_0}\right)^{3tm_0} e^{-4tm_0\log d} \le e^{-tm_0\log d}.
$$

So,

$$
P \le dn^d e^{-4tm_0\log d} = o(1).
$$

□

Lemma 4.5. *If $k \ge 2$ and d is sufficiently large then w.h.p. every set $S \subset V$ of size at most n_0 spans fewer than $3|S|\log^{3k} d$ edges in H. Hence no subset of size at most n_0 contains a $3\log^{3k} d$ core.*

Proof. Let $L = 3\log^{3k} d$. The probability that there exists $S \subset V$ of size at most n_0 that spans at least $t = L|S|$ edges is bounded by

$$\sum_{s=1}^{n_0} \binom{n}{s} \binom{\binom{s}{k}}{t} p^t \leq \sum_{s=1}^{n_0} \left(\left(\frac{en}{s} \right)^{s/t} \cdot \frac{e\binom{s}{k}}{t} \cdot \frac{d}{\binom{n-1}{k-1}} \right)^t \leq \sum_{s=1}^{n_0} \left(\left(\frac{en}{s} \right)^{1/L} \frac{eds}{t} \left(\frac{s}{n} \right)^{k-1} \right)^t$$
$$= \sum_{s=1}^{n_0} \left(\left(\frac{s}{n} \right)^{k-1-1/L} \frac{e^{1+1/L} d}{L} \right)^t = o(1).$$

□

Proof of Theorem 1.1: Let α, β be as in (1). We argue next that the properties given by Lemmas 4.2, 4.4 and 4.5 imply that $H_{n,p;k}$ is (α, β)-colorable for d sufficiently large. Lemma 4.1 then follows directly from (3) and Theorem 3.2.

Consider a sequence of sets $V_1, V_2, \ldots, V_\alpha$ such that V_i is maximally independent in $[n] \setminus \bigcup_{j<i} V_j$ for $j \leq \alpha$ (some of these sets can be empty). It follows from Lemma 4.4 that because $\alpha m_0 = n$, we must have $\sum_{i=1}^{\alpha} |V_i| \geq n - n_0$ and then Lemma 4.5 implies that $[n] \setminus \bigcup_{i \leq \alpha} V_i$ does not have a β-core. This completes the proof of the first part of the theorem.

When $q \geq \alpha + 2\beta + 2$ we see that we have $2\beta + 2$ colors with which to color a hypergraph with no β-core and Theorem 1 of [8] implies that we need $O(\beta n)$ vertex re-colorings to do this. We will encounter at most α such hypergraphs in our re-coloring. This together with Remark 3.7 shows that there will be $O(\alpha\beta n)$ re-colorings overall and this proves the second part of the theorem. □

References

[1] D. Achlioptas and E. Friedgut, A Sharp Threshold for k-Colorability,*Random Structures and Algorithms*, 14 (1999) 63-70.

[2] D. Achlioptas, A. Coja-Oghlan and F. Ricci-Tersenghi, On the solution-space geometry of random constraint satisfaction problems, *Random Structures and Algorithms* 38 (2010) 251-268.

[3] D. Achlioptas and A. Naor, The Two Possible Values of the Chromatic Number of a Random Graph, *Annals of Mathematics* 162 (2005) 1333-1349.

[4] M. Anastos, A.M. Frieze and W. Pegden, Constraining the clustering transition for colorings of sparse random graphs, *Electronic Journal of Combinatorics* 2018.

[5] V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Rassmann and D. Vilenchik, The condensation phase transition in random graph coloring, *Communications in Mathematical Physics* 341 (2016) 543-606.

[6] V. Bapst, A. Coja-Oghlan and C. Efthymiou, Planting colourings silently, *Combinatorics, Probability and Computing* 26 (2017) 338-366.

[7] B. Bollobás and P. Erdős, Cliques in random graphs, *Mathematical Proceedings of the Cambridge Philosophical Society* 80 (1976) 419-427.

[8] N. Bousquet and G. Perarnau, Fast recoloring of sparse graphs, *European Journal of Combinatorics* 52 (2016) 1-11.

[9] A. Coja-Oghlan and D. Vilenchik, Chasing the k-colorability threshold, *Proceedings of FOCS 2013*, 380-389.

[10] J. Ding, A. Sly and N. Sun, Proof of the satisfiability conjecture for large k, arxiv.org/pdf/1411.0650.pdf.

[11] M. Dyer, A. Flaxman, A.M. Frieze and E. Vigoda, Randomly coloring sparse random graphs with fewer colors than the maximum degree, *Random Structures and Algorithms* 29 (2006) 450-465.

[12] C. Efthymiou, T. Hayes, D. Štefankovič and E. Vigoda, Sampling Colorings of Sparse random Graphs. arXiv:1707.03796.

[13] C. Feghali, Paths between colourings of sparse graphs, arxiv:1803.03950.

[14] A.M. Frieze and M. Karoński, Introduction to Random Graphs, Cambridge University Press, 2015.

[15] M. Gabrié, V. Dani, G. Semerjian and L. Zdeborová, Phase transitions in the q-coloring of random hypergraphs, *Journal of Physics A: Mathematical Theory* 50 (2017).

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

[17] S. Janson, T. Łuczak and A. Ruciński, Random Graphs, Wiley 2000.

[18] F. Krzakąla, A. Montanari, F. Ricci-Tersenghi, G. Semerijian and L. Zdeborová, Gibbs states and the set of solutins of random constraint satisfaction problems, *Proceedings of the National Academy of Sciences* 104 (2007) 10318-10323.

[19] M. Molloy, The freezing threshold for k-colourings of a random graph, *Proceedings of STOC 2012*.

[20] E. Shamir and E. Upfal, Sequential and Distributed Graph Coloring Algorithms with Performance Analysis in Random Graph Spaces, *Journal of Algorithms* 5 (1984) 488-501.

[21] L. Zdeborová and F. Krzakąla, Phase Transitions in the Coloring of Random Graphs, *Physics Review E* 76 (2007).