The Chromatic Numbers of Random Hypergraphs

Michael Krivelevich * Benny Sudakov †

Abstract

For a pair of integers $1 \leq \gamma < r$, the γ-chromatic number of an r-uniform hypergraph $H = (V, E)$ is the minimal k, for which there exists a partition of V into subsets $T_1, \ldots, T_k$ such that $|e \cap T_i| \leq \gamma$ for every $e \in E$. In this paper we determine the asymptotic behavior of the γ-chromatic number of the random r-uniform hypergraph $H_r(n, p)$ for all possible values of γ and for all values of p down to $p = \Theta(n^{-r+1})$.

1 Introduction

A *hypergraph* H is an ordered pair $H = (V, E)$, where V is a finite set (the *vertex set*), and E is a family of distinct subsets of V (*the edge set*). A hypergraph $H = (V, E)$ is r-uniform if all edges of H are of size r. In this paper we consider only r-uniform hypergraphs. Our terminology follows that of [3].

A random r-uniform hypergraph $H_r(n, p)$ is an r-uniform hypergraph on n labeled vertices $V = [n] = \{1, \ldots, n\}$, in which every subset $e \subset V$ of size $|e| = r$ is chosen to be an edge of H randomly and independently with probability p, where p may depend on n. Thus, for $r = 2$ this model reduces to the well known and thoroughly studied model $G(n, p)$ of random graphs. The reader is referred to the paper of Karoński and Łuczak [6] for additional information and survey of the state of the art in random hypergraphs. In this paper we study some asymptotic properties of $H_r(n, p)$, that is, we think of n as tending to infinity while r is kept fixed.

One of the most interesting parameters of a random hypergraph is its chromatic number. Actually, a family of chromatic numbers can be defined as the reader will immediately see from the definitions below. For an integer $1 \leq \gamma \leq r - 1$, a *γ-independent set* in a hypergraph $H = (V, E)$ is a subset $V_0 \subseteq V$ such that $|e \cap V_0| \leq \gamma$ for every $e \in E$. The *γ-independence number* $\alpha_\gamma(H)$ of H is

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: krivelev@math.tau.ac.il. Research supported in part by a DIMACS Graduate Fellowship for a short term visit and by a Charles Clore Fellowship.

†Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: sudakov@math.tau.ac.il. Research supported in part by a DIMACS Graduate Fellowship for a short term visit and forms part of a Ph.D thesis written by the author under the supervision of Prof. N. Alon.
Mathematics Subject Classification (1991): 05C80, 05C15, 05C65.

the maximal size of a γ-independent set. A *γ-coloring* of H is a partition of the vertex set of H into γ-independent sets (colors). The *γ-chromatic number* $\chi_\gamma(H)$ of H is the minimal number of colors in a γ-coloring of H. In particular, for $\gamma = r-1$ we require that every edge of H is not monochromatic, the corresponding chromatic number is usually called *the weak chromatic number* of H. In another extreme case $\gamma = 1$ the vertices of every edge of H should be colored by distinct colors, the corresponding chromatic number is called *the strong chromatic number* of H. The notions of weak and strong chromatic numbers have been used in particular in [11]

This paper is devoted to the investigation of the asymptotic behavior of the chromatic numbers of a random hypergraph. For the case of random graphs ($r = 2$), this problem has been studied intensively during the last twenty years and finally has been completely solved by Bollobás [4] for the case of dense graphs and by Łuczak [7] for all remaining values of probability. The key ingredient of both proofs was the use of martingales. However, for every $r \geq 3$ and for every value of γ the situation was far from being clear. Only partial results have been known so far. Schmidt, Shamir and Upfal in [11] considered the weak chromatic number and obtained lower and upper bounds which differ by a factor of two. Schmidt [10] treated the case of a general γ and got lower and upper bounds whose ratio is bounded by an absolute constant, thus establishing the asymptotic order of magnitude of the γ-chromatic number. Finally, Shamir [12], again using martingale techniques, found the asymptotic value of the γ-chromatic number for the dense case (that is, when $(n^{r-1}p)^{1/\gamma} \geq n^{1-\epsilon}$ for some fixed $\epsilon > 0$). For other values of p the problem remains unsolved. Moreover, even the asymptotic behavior of the γ-independence number was not known for these values of p. More details can be found in [6].

The main result of this paper is the following theorem.

Theorem 1 *For every $1 \leq \gamma \leq r-1$ there exists a constant d_0 such such that if $d^{(\gamma)} = d^{(\gamma)}(n,p) = \gamma\binom{r-1}{\gamma}\binom{n-1}{r-1}p \geq d_0$ and $d^{(\gamma)} = o(n^\gamma)$, then almost surely*[1]

$$\left(\frac{d^{(\gamma)}}{(\gamma+1)\log d^{(\gamma)}}\right)^{\frac{1}{\gamma}} \leq \chi_\gamma(H_r(n,p)) \leq \left(\frac{d^{(\gamma)}}{(\gamma+1)\log d^{(\gamma)}}\left(1+\frac{1}{\log^{0.1} d^{(\gamma)}}\right)\right)^{\frac{1}{\gamma}} .$$

This statement combined with the result of Shamir solves the problem completely. Note that for $r = 2$ the above theorem essentially coincides with the theorem of Łuczak [7]. It is worth mentioning here that we do not make any attempt to optimize an error term in the upper bound for $\chi_\gamma(H_r(n,p))$.

Theorem 1 implies immediately (the difficult half of) the following corollary about the asymptotic value of the γ-independence number of $H_r(n,p)$.

Corollary 1 *For every $1 \leq \gamma \leq r-1$ there exists a constant d_0 such that if $d^{(\gamma)} = d^{(\gamma)}(n,p) = \gamma\binom{r-1}{\gamma}\binom{n-1}{r-1}p \geq d_0$ and $d^{(\gamma)} = o(n^\gamma)$, then a.s.*

$$n\left(\frac{d^{(\gamma)}}{(\gamma+1)\log d^{(\gamma)}}\left(1+\frac{1}{\log^{0.1} d^{(\gamma)}}\right)\right)^{-\frac{1}{\gamma}} \leq \alpha_\gamma(H_r(n,p)) \leq n\left(\frac{d^{(\gamma)}}{(\gamma+1)\log d^{(\gamma)}}\right)^{-\frac{1}{\gamma}} .$$

[1]An event $\mathcal{E}_n$ holds *almost surely* (a.s.) in $H_r(n,p)$ if the probability of $\mathcal{E}_n$ tends to 1 as n tends to infinity.

The rest of the paper is organized as follows. The main idea of the proof is described briefly in Section 2. This section presents most of the crucial ingredients of the proof, whereas the somewhat complicated details are postponed to the subsequent sections. In Section 3 we state an upper bound for the γ-independence number of $H_r(n,p)$. In Section 4 we prove a technical lemma bounding the γ-chromatic number of the subhypergraphs of $H_r(n,p)$ spanned by relatively small subsets of V. It turns out that establishing the asymptotic value of the weak chromatic number ($\gamma = r-1$) plays a key role in dealing with the other values of γ. We treat the weak chromatic number in Section 5, following mainly the proof of Łuczak. Finally, in Section 6 we prove Theorem 1 for $\gamma < r-1$.

Based on the results of Shamir [12] and Łuczak [7], we may assume that $r \geq 3$ and that $(n^{r-1}p)^{1/\gamma} \leq n^{1-\epsilon}$ for some $\epsilon > 0$. Throughout the paper, we omit occasionally the floor and ceiling signs for the sake of convenience. All logarithms are natural.

We use the following notation.

$$\begin{aligned}
H_U^i &= \{e \cap U : e \in E(H), |e \cap U| = i\} , \\
H_U^{\geq i} &= \bigcup_{j \geq i} H_U^i , \\
d^{(\gamma)} &= \gamma \binom{r-1}{\gamma} \binom{n-1}{r-1} p , \\
d &= d(n,p,r) = (r-1)\binom{n-1}{r-1} p , \\
d^* &= n^{r-1} p .
\end{aligned}$$

2 The main idea

In this section we describe briefly the core idea of the proof. We omit all technicalities and calculations, postponing them to the next sections.

First, it is quite easy to get a lower bound for the γ-chromatic number by upper bounding the γ-independence number. This in turn can be done by a straightforward first moment argument. Therefore, most of our efforts will be devoted to proving an upper bound.

It turns out that the weak chromatic number, that is, the case $\gamma = r-1$ is much more tractable than the γ-chromatic number for other values of γ. The main reason of this phenomenon originates from the fact, that the weak chromatic number is vertex Lipschitz (see, e.g., [2], Ch. 7, for relevant definitions and discussion). This means that if two hypergraphs H, H' with the same vertex set differ only in the edges containing some particular vertex v, then their weak chromatic numbers differ by at most one. This makes the situation in this case very similar to that in the random graph $G(n,p)$, thus enabling the use of martingales and the application of the main ideas of the proof of Łuczak [7].

However, for every $\gamma < r-1$ the corresponding chromatic number ceases to be vertex Lipschitz. Therefore, we need to develop a different approach to tackle this case. Fortunately, one can use the upper bound for the weak chromatic number to cope with this task. We illustrate this by presenting an outline of the proof for the case $r = 3$, $\gamma = 1$, that is, for the strong chromatic number of a 3-uniform random hypergraph. For this case $d = d(n, p, 3) \approx n^2 p$ and we need to show that the strong chromatic number of $H_3(n, p)$ is $(1 + o(1))d/2\log d$.

Let $s = \log^4(n^2 p)$. We fix a partition of the vertex set V into s parts $V_1, \ldots, V_s$ of equal size $|V_i| = n/s = n_0$. Let $H_i = H^{\geq 2}_{V_i}$. For every $1 \leq i \leq s$ we find a strong coloring of H_i, using pairwise disjoint sets of colors for different values of i.

Consider the hypergraph H_i for some i. It is important to note that most of the edges of H_i will be of size 2. These edges determine the asymptotic behavior of $\chi_1(H_i)$.

For $j = 2, 3$ we denote $H_{i,j} = H^j_{V_i}$. A crucial observation in the whole proof is that for every subset $e \subset V_i$ of size $|e| = j$ the probability of the event "e is an edge of $H_{i,j}$" is exactly $p_j = 1 - (1-p)^{\binom{n-n_0}{3-j}} \approx p\binom{n-n_0}{3-j}$. Moreover, all these events are completely independent. Therefore, each of the subhypergraphs $H_{i,j}$ can be treated as a random hypergraph $H_j(n_0, p_j)$.

Consider first the hypergraph $H_{i,2}$. As explained above, this hypergraph is actually a random graph $G(n_0, p_2)$. Therefore it can be colored a.s. by $(1 + o(1))n_0 p_2/2\log(n_0 p_2)$ colors. (For a general γ, at this stage we find a γ-coloring of $H_{i,\gamma+1}$ which is by definition a *weak* coloring of this $(\gamma + 1)$-uniform subhypergraph.) Fix one such coloring, by Lemma 3.1 all color classes are of size at most, say, $4\log(n_0 p_2)/p_2$. Now we expose the edges of $H_{i,3}$. We call an edge $e \in H_{i,3}$ bad if it has at least two vertices of the same color. Denote by X_i the number of bad edges. A calculation of the expectation of X_i gives that $EX_i \leq cn_0/\log^3(n_0 p_2)$ for some absolute constant $c > 0$. Since X_i is distributed binomially, we get that with high probability $X_i \leq 2cn_0/\log^3(n_0 p_2)$. Now we delete from H_i the union of all bad edges. Combining the colorings of all hypergraphs H_i without the bad edges, we obtain a strong coloring of all but at most $s(6cn_0/\log^3(n_0 p_2)) = (1 + o(1))6cn/\log^3(n^2 p)$ vertices in $(1 + o(1))s(n_0 p_2/2\log(n_0 p_2)) = (1 + o(1))n^2 p/2\log(n^2 p)$ colors. The remaining $(1+o(1))6cn/\log^3(n^2 p)$ vertices can be colored by a much smaller number of additional colors using a simple greedy-type algorithm based on the degrees. Thus the total number of colors is

$$(1 + o(1))\frac{n^2 p}{2\log(n^2 p)} = (1 + o(1))\frac{d}{2\log d} .$$

This finishes our argument.

3 Bounding the γ-independence number

In this section we state an upper bound on the γ-independence number of $H_r(n, p)$. This bound is easily obtained by computing the first moment of an appropriate random variable. We cite it from Lemma 6.3 of [10].

Lemma 3.1 *With probability $1-O(1/n^4)$ the γ-independence number of $H_r(n,p)$ satisfies the following inequality:*

$$\alpha_\gamma(H_r(n,p)) \le n\left(\frac{d^{(\gamma)}}{(\gamma+1)\log d^{(\gamma)}}\right)^{-\frac{1}{\gamma}} .$$

It is obvious that $\chi_\gamma(H) \ge |V(H)|/\alpha_\gamma(H)$ for every hypergraph $H=(V,E)$. Hence we get immediately the following lower bound on $\chi_\gamma(H_r(n,p))$.

Corollary 3.2 *With probability $1-O(1/n^4)$ the γ-chromatic number of $H_r(n,p)$ is bounded from below as follows:*

$$\chi_\gamma(H_r(n,p)) \ge \left(\frac{d^{(\gamma)}}{(\gamma+1)\log d^{(\gamma)}}\right)^{\frac{1}{\gamma}} .$$

4 Coloring small subsets

In this technical section we bound from above the γ-chromatic number of all subhypergraphs of $H_r(n,p)$, spanned by subsets of V of relatively small size.

Lemma 4.1 *Let $d^*=n^{r-1}p$ and let $1\le\gamma\le r-1$. For every fixed $c>0$ with probability $1-o(1/n)$ in $H_r(n,p)$ the following holds: for every subset $U\subset V$ of size $|U|\le cn/\log^3 d^*$ the subhypergraph $H_U^{\ge\gamma+1}$ is γ-colorable by at most $(d^*/\log^{2\gamma-0.7} d^*)^{1/\gamma}$ colors.*

The lemma will follow easily from a sequence of claims.

Claim 4.2 *For every fixed $c>0$ with probability $1-o(1/n^2)$ in $H_r(n,p)$ the following holds: for every integer $s\le cn/\log^3 d^*$ and every subset $U\subset V$ of size $|U|=s$ the subhypergraph $H_U^{\ge\gamma+1}$ has less than $rsd^*/\log^{3\gamma-0.5} d^*$ edges.*

Proof. For every fixed set U of size $|U|=s$ and every subset $e\subseteq U$ of size $|e|=i$, where $\gamma+1\le i\le r$, the probability that e is an edge of $H_U^{\ge\gamma+1}$ is $1-(1-p)^{\binom{n-s}{r-i}}\le n^{r-i}p$. Moreover, all such events "e is an edge of $H_U^{\ge\gamma+1}$" are mutually independent. Also, if $H_U^{\ge\gamma+1}$ has at least $rsd^*/\log^{3\gamma-0.5} d^*$ edges, then some of the subhypergraphs H_U^i, $\gamma+1\le i\le r$, has at least $sd^*/\log^{3\gamma-0.5} d^*$ edges. Therefore, the probability of the existence of a set U violating the claim can be bounded from above by:

$$\sum_{s\le\frac{cn}{\log^3 d^*}}\binom{n}{s}\sum_{i=\gamma+1}^{r}\binom{\binom{s}{i}}{\frac{sd^*}{\log^{3\gamma-0.5} d^*}}(n^{r-i}p)^{\frac{sd^*}{\log^{3\gamma-0.5} d^*}}$$

$$\le \sum_{s\le\frac{cn}{\log^3 d^*}}\binom{n}{s}\sum_{i=\gamma+1}^{r}\left[\frac{O(1)s^{i-1}\log^{3\gamma-0.5} d^*}{d^*}n^{r-i}p\right]^{\frac{sd^*}{\log^{3\gamma-0.5} d^*}}$$

$$\leq \sum_{s\leq \frac{cn}{\log^3 d^*}} \left[O(1)\frac{n}{s}\left(\frac{O(1)s^{\gamma}n^{r-\gamma-1}p\log^{3\gamma-0.5}d^*}{d^*}\right)^{\frac{d^*}{\log^{3\gamma-0.5}d^*}}\right]^s$$

$$\leq \sum_{s\leq \frac{cn}{\log^3 d^*}} \left[O(1)\left(\frac{s}{n}\right)^{\gamma-\frac{\log^{3\gamma-0.5}d^*}{d^*}}\log^{3\gamma-0.5}d^*\right]^{\frac{sd^*}{\log^{3\gamma-0.5}d^*}} .$$

Denote the s-th summand of the above sum by a_s. Then, if $s \geq \sqrt{n}$, we have

$$a_s \leq \left[O(1)\left(\frac{c}{\log^3 d^*}\right)^{\gamma-\frac{\log^{3\gamma-0.5}d^*}{d^*}}\log^{3\gamma-0.5}d^*\right]^{\sqrt{n}} = o\left(\frac{1}{n^3}\right) ,$$

while if $s \leq \sqrt{n}$, we have

$$a_s \leq \left(\frac{\log^{3\gamma-0.5}d^*}{n^{1/3}}\right)^{\frac{d^*}{\log^{3\gamma-0.5}d^*}} = o\left(\frac{1}{n^3}\right) ,$$

thus establishing the claim. □

Claim 4.3 *With probability* $1-o(1/n)$ *in* $H_r(n,p)$ *the following holds: for every integer* $s \leq n/d^*$ *and every subset* $W \subset V$ *of size* $|W| = s$ *the subhypergraph* $H_W^{\geq\gamma+1}$ *has less than* $rs\log d^*$ *edges.*

Proof. Using arguments similar to those of the proof of Claim 4.2, we can estimate from above the probability of the existence of a set W violating the claim by:

$$\sum_{s\leq \frac{n}{d^*}} \binom{n}{s} \sum_{i=\gamma+1}^{r} \binom{\binom{s}{i}}{s\log d^*} (n^{r-i}p)^{s\log d^*}$$

$$\leq \sum_{s\leq \frac{n}{d^*}} \binom{n}{s} \sum_{i=\gamma+1}^{r} \left(\frac{O(1)s^{i-1}n^{r-i}p}{\log d^*}\right)^{s\log d^*}$$

$$\leq \sum_{s\leq \frac{n}{d^*}} \left(O(1)\left(\frac{s}{n}\right)^{\gamma-\frac{1}{\log d^*}}\frac{d^*}{\log d^*}\right)^{s\log d^*} .$$

Denote the s-th summand of the above sum by b_s. Then, if $s \geq n^{\epsilon/2}$, where $\epsilon > 0$ is a fixed constant defined in the introduction and satisfying $(d^*)^{1/\gamma} \leq n^{1-\epsilon}$, we can estimate b_s from above as follows:

$$b_s \leq \left(O(1)\frac{d^{*-\gamma+1+\frac{1}{\log d^*}}}{\log d^*}\right)^{s\log d^*} \leq \left(\frac{O(1)}{\log d^*}\right)^s = o\left(\frac{1}{n^2}\right) ,$$

and if $s \leq n^{\epsilon/2}$, relying on our assumption that $d^* \leq n^{(1-\epsilon)\gamma}$, we get:

$$b_s \leq \left(\frac{O(1) n^{\left(-1+\frac{\epsilon}{2}\right)\left(\gamma - \frac{1}{\log d^*}\right)} d^*}{\log d^*} \right)^{\log d^*} \leq \left(\frac{O(1) n^{-\gamma + \frac{2\epsilon}{3}\gamma + (1-\epsilon)\gamma}}{\log d^*} \right)^{\log d^*} = o\left(\frac{1}{n^2}\right) . \quad \square$$

Claim 4.4 *For any fixed $c > 0$ with probability $1 - o(1/n^2)$ in $H_r(n,p)$ the following holds. For every subset $U \subset V$ of size $|U| = s \leq \frac{cn}{\log^3 d^*}$ consider the subhypergraph $H_U^{\geq \gamma+1}$. Then there exists a γ-coloring of all but at most $\frac{s}{\log^{0.1} d^*}$ vertices of $H_U^{\geq \gamma+1}$ in at most $\left\lceil \left(\frac{4r^3 d^*}{\log^{3\gamma-0.6} d^*} \right)^{\frac{1}{\gamma}} \right\rceil$ colors.*

Proof. This claim follows deterministically from the assertion of Claim 4.2.

Fix a subset $U \subset V$ of size $|U| = s \geq \frac{cn}{\log^3 d^*}$ and consider the subhypergraph $H(U) = H_U^{\geq \gamma+1}$. Denote

$$M = \left\lceil \left(\frac{4r^3 d^*}{\log^{3\gamma-0.6} d^*} \right)^{\frac{1}{\gamma}} \right\rceil .$$

According to Claim 4.2, with probability $1 - o(1/n^2)$ the number of vertices of degree more than $r^2 d^* / \log^{3\gamma-0.6} d^*$ in $H(U)$ is at most $\left(\frac{rsd^*}{\log^{3\gamma-0.5} d^*} r \right) \Big/ \frac{r^2 d^*}{\log^{3\gamma-0.6} d^*} = s / \log^{0.1} d^*$. Let U_0 be the set of all vertices of degree at most $r^2 d^* / \log^{3\gamma-0.6} d^*$ in $H(U)$. Then $|U \setminus U_0| \leq s/\log^{0.1} d^*$. Let $H(U_0) = H_{U_0}^{\geq \gamma+1}$. We will prove that the $H(U_0)$ is γ-colorable by at most M colors. Every edge of $H(U_0)$ intersects at most $r \cdot r^2 d^* / \log^{3\gamma-0.6} d^* = r^3 d^* / \log^{3\gamma-0.6} d^* \leq M^\gamma/4$ other edges of $H(U_0)$. Now consider a random coloring of the vertices of U_0 in colors $1, \ldots, M$, obtained by assigning to each vertex $v \in U_0$ color i independently and with probability $1/M$, where $1 \leq i \leq M$. For $e \in E(H(U_0))$ denote by A_e the event "e is monochromatic". Then the probability of A_e is $(1/M)^{|e|-1} \leq (1/M)^\gamma$. Also, the event A_e is mutually independent of all other events $A_{e'}$, but those for which $e \cap e' \neq \emptyset$. The number of such events is at most $M^\gamma/4$. Then, applying the symmetric version of the Lovász Local Lemma (see, e.g., [2], Ch. 5), we get: $P[\bigwedge_{e \in E(H(U_0))} \overline{A_e}] > 0$, thus ensuring the existence of a desired coloring. $\square$

Claim 4.5 *With probability $1 - o(1/n)$ in $H_r(n,p)$ the following holds: for every integer $s \leq n/d^*$ and every subset $W \subset V$ of size $|W| = s$ the subhypergraph $H_W^{\geq \gamma+1}$ is γ-colorable by at most $2r \log d^*$ colors.*

Proof. This claim follows deterministically from the assertion of Claim 4.3. Recall that a graph G is called *d-degenerate* if every subgraph of it contains a vertex of degree at most d. It is very easy to see that a d-degenerate graph G is $(d+1)$-colorable.

Fix a subset $W \subset V$ of size $|W| = s \leq n/d^*$. Clearly it is enough to prove that the strong chromatic number of $H_W^{\geq \gamma+1}$ is at most $2r \log d^*$. Define an auxiliary graph G with vertex set W and two vertices $u, v \in W$ being connected by an edge if and only if there exists an edge $e \in E(H)$ such that $u, v \in e$. Then it is easy to see that the chromatic number of G is equal to the strong chromatic number of $H_W^{\geq \gamma+1}$. By Claim 4.3 with $\gamma = 1$, every subset $W_0 \subseteq W$ of size $|W_0| = s_0 \leq s$ spans less than $rs_0 \log d^*$ edges in G, and therefore the induced subgraph $G[W_0]$ has a vertex of degree less than $2r \log d^*$. This implies that G is $(2r \log d^* - 1)$-degenerate and thus can be colored by $2r \log d^*$ colors, yielding the desired result. $\Box$

Proof of Lemma 4.1. For a fixed subset U, first apply Claim 4.4 recursively, starting with U, each time using at most $\lceil (4r^3 d^* / \log^{3\gamma - 0.6} d^*)^{1/\gamma} \rceil$ fresh colors and decreasing the size of the current subset by a factor of $\log^{0.1} d^*$. Then, after at most $\log(sd^*/n)$ iterations, we get a subset W of U of size at most n/d^*, to which we apply Claim 4.5. The total number of colors used is at most

$$\left\lceil \left(\frac{4r^3 d^*}{\log^{3\gamma - 0.6} d^*} \right)^{\frac{1}{\gamma}} \right\rceil \log\left(\frac{sd^*}{n}\right) + 2r \log d^* \leq \left(\frac{d^*}{\log^{2\gamma - 0.7} d^*} \right)^{\frac{1}{\gamma}} . \qquad \Box$$

5 The weak chromatic number

In this section we establish the asymptotic behavior of the weak chromatic number of $H_r(n,p)$. Our argument is essentially an adaptation of the proof of Łuczak [7], with some changes incorporated.

Theorem 5.1 *There exists a constant d_0 such that if $d = d(n,p,r) = (r-1)\binom{n-1}{r-1}p > d_0$ and $d = o(n^{r-1})$ then, with probability $1 - o(1/n)$*

$$\left(\frac{d}{r \log d} \right)^{\frac{1}{r-1}} \leq \chi_{r-1}(H_r(n,p)) \leq \left(\frac{d}{r \log d} \left(1 + \frac{28r \log\log d}{\log d} \right) \right)^{\frac{1}{r-1}} .$$

The lower bound follows immediately from Corollary 3.2. Thus it is enough to show only the second inequality.

Lemma 5.2 *Let $k = n(\frac{r(\log d - 3\log\log d)}{d})^{\frac{1}{r-1}}$. There exists a constant d_0 such that whenever $d > d_0$ and $d = o(n^{r-1})$, then with probability at least $1 - n^{-3}$, $H_r(n,p)$ contains a subset with at least $n \log^{-5} d$ vertices which can be properly colored using at most $n \log^{-5} d/k$ colors.*

Proof. To prove the lemma we use Talagrand's inequality as suggested by the referee. First we will describe an adaptation of Talagrand's inequality, convenient for combinatorial applications, as presented by Spencer in [14]. Let $\Omega = \prod_{i=1}^n \Omega_i$ be a product probability space and let $h : \Omega \to R$ be a real-valued random variable. We call h *Lipschitz* if $|h(x) - h(y)| \leq 1$ for all $x, y \in \Omega$ which differ in only one coordinate. For a fixed function $f : N \to N$ we say that h is f-certifiable if whenever $h(x) \geq s$ for some $x \in \Omega$ there is a set of at most $f(s)$ indices $I \subseteq \{1, \ldots, n\}$ that certify $h(x) \geq s$

in the sense that $h(y) \geq s$ for all $y \in \Omega$ that agree with x on I. Let m be a median of the random variable $h(x)$. Then as shown in [14], Talagrand's inequality implies

$$Pr(|h(x) - m| \geq t) \leq 2\exp(-t^2/4f(m)). \tag{1}$$

Now consider the probability space $H_r(n,p)$ as a product space, where each Ω_i corresponds to the r-tuples of $\{1,\ldots,n\}$ containing vertex i and contained in $\{1,\ldots,i\}$. Let X be the size of the largest $n\log^{-5} d/k$-colorable subset of $H_r(n,p)$ and let X^* be a random variable defined by $X^* = \min(5n\log^{-5} d, X)$. Then by the definition X^* is Lipschitz and always bounded from above by $5n\log^{-5} d$. Therefore it is also $5n\log^{-5} d$-certifiable, since it is enough to expose the edges from at most $5n\log^{-5} d$ vertices to certify the value of X^*. Denote by $m(X^*)$ the median of X^*, obviously $m(X^*) \leq 5n\log^{-5} d$. Then inequality (1) with $t = n/2\log^{6.1} d$ will imply that

$$Pr\left(|X^* - m(X^*)| > \frac{n}{2\log^{6.1} d}\right) \leq 2\exp\left(-\frac{n}{80\log^{7.2} d}\right) \tag{2}$$

Thus to prove Lemma 5.2, it is enough to show that the probability that $H_r(n,p)$ contains a $n\log^{-5} d/k$-colorable subset with more than $n(\log^{-5} d + \log^{-6.1} d)$ elements is greater than $2\exp(-n/80\log^{7.2} d)$. Indeed, in this case, by inequality (2) we obtain that the median $m(X^*)$ should be at least $n(\log^{-5} d + \log^{-6.1} d) - n/2\log^{6.1} d = n\log^{-5} d + n/2\log^{6.1} d$. Therefore

$$Pr\left(X^* < \frac{n}{\log^5 d}\right) \leq Pr\left(X^* - m(X^*) < -\frac{n}{2\log^6 .1d}\right) \leq 2\exp\left(-\frac{n}{80\log^{7.2} d}\right) < n^{-3}.$$

Let Y be the number of subsets of mk_0 elements, where $m = n\log^{-5} d/k$ and

$$k\frac{\log^{-5} d + \log^{-6.1} d}{\log^{-5} d} \leq k_0 = n\left(\frac{r(\log d - 2\log\log d)}{d}\right)^{\frac{1}{r-1}},$$

which can be split into exactly m independent sets, each of size k_0. Then the event "$Y > 0$" implies that $X^* \geq mk_0 \geq n(\log^{-5} d + \log^{-6.1} d)$. On the other hand, to bound from below the probability that Y is positive we can use the following inequality (see, e.g., [5], p. 3)

$$Pr(Y > 0) \geq \frac{(EY)^2}{EY^2}.$$

Then

$$\frac{EY^2}{(EY)^2} \leq \prod_{i=1}^{m} \sum_{\substack{k_1,\ldots,k_{m+1} \\ \sum_{j=1}^{m+1} k_j = k_0}} \frac{\binom{k_0}{k_1}\cdots\binom{k_0}{k_m}\binom{n-(i-1)k_0}{k_{m+1}}}{\binom{n-(i-1)k_0}{k_0}(1-p)^{\sum_{j=1}^{m}\binom{k_j}{r}}}$$

$$\leq \left[\sum_{l=0}^{k_0} \frac{a_l\binom{n-mk_0}{k_0-l}}{\binom{n-mk_0}{k_0}}\right]^m$$

$$\leq \left[\sum_{l=0}^{k_0} \frac{a_l}{(n-(m+1)k_0)^l} \frac{k_0!}{(k_0-l)!} \right]^m ,$$

where

$$a_l = \sum_{\substack{k_1,\ldots,k_m \\ \sum_{j=1}^m k_j = l}} \binom{k_0}{k_1} \ldots \binom{k_0}{k_m} (1-p)^{-\sum_{j=1}^m \binom{k_j}{r}}.$$

Let $k_{i_1}, \ldots, k_{i_t}$ be those from $k_1, \ldots, k_m$ which are greater than $n(\frac{r \log\log d}{d})^{\frac{1}{r-1}}$. Since $\sum_{j=1}^m k_j = l \leq k_0 < n(\frac{r \log d}{d})^{\frac{1}{r-1}}$, so $t < \log^{\frac{1}{r-1}} d$. Thus the number of terms with different sequences $k_{i_1}, \ldots, k_{i_t}$ is less than

$$(mk_0)^{\log^{\frac{1}{r-1}} d} < n^{\log^{\frac{1}{r-1}} d}. \tag{3}$$

Moreover, for every k', k'' such that $k' \geq k'' \geq n(\frac{r \log\log d}{d})^{\frac{1}{r-1}}$ and $k' + k'' \leq l \leq k_0$ we have

$$\frac{\binom{k_0}{k'}\binom{k_0}{k''}(1-p)^{-\binom{k'}{r}-\binom{k''}{r}}}{\binom{k_0}{k'+k''}(1-p)^{-\binom{k'+k''}{r}}} < 1.$$

Indeed, when $n(\frac{r \log\log d}{d})^{\frac{1}{r-1}} \leq k' + k'' \leq 0.7k_0$ then

$$\frac{\binom{k_0}{k'}\binom{k_0}{k''}(1-p)^{-\binom{k'}{r}-\binom{k''}{r}}}{\binom{k_0}{k'+k''}(1-p)^{-\binom{k'+k''}{r}}} \leq \frac{(k_0)_{k''}}{(k_0-k')_{k''}} \binom{k'+k''}{k''} \exp\left\{ -p\left(\binom{k'+k''}{r} - \binom{k'}{r} - \binom{k''}{r} \right) \right\}$$

$$\leq \left(\frac{k_0-k''}{k_0-k'-k''} \frac{e(k'+k'')}{k''} \exp\left\{ -p\binom{k'}{r-1} \right\} \right)^{k''} \leq \left(\frac{100}{\log\log d} \right)^{k''} < 1,$$

whereas for $k' + k'' \geq 0.7k_0$ we have

$$\frac{\binom{k_0}{k'}\binom{k_0}{k''}(1-p)^{-\binom{k'}{r}-\binom{k''}{r}}}{\binom{k_0}{k'+k''}(1-p)^{-\binom{k'+k''}{r}}} \leq 2^{k_0} 2^{k_0} \exp\left\{ -pk''\binom{k'}{r-1} \right\} < 1.$$

Hence

$$\binom{k_0}{k_1} \ldots \binom{k_0}{k_m} (1-p)^{-\sum_{j=1}^m \binom{k_j}{r}} \leq \binom{k_0}{l} (1-p)^{-\binom{l}{r}} \leq \binom{k_0}{l} \exp\left((1+p) \frac{dl^r}{r(r-1)n^{r-1}} \right). \tag{4}$$

Furthermore, for every choice of $k'_1, \ldots, k'_s$, one can easily get the following inequality

$$\sum_{\substack{k'_1,\ldots,k'_s,\ \max_{1\le j\le s} k'_j=f \\ \sum_{j=1}^s k'_j=l}} \binom{k_0}{k'_1}\cdots\binom{k_0}{k'_s}(1-p)^{-\sum_{j=1}^s \binom{k'_j}{r}} \le \binom{sk_0}{l}\exp\left((1+p)\frac{ldf^{r-1}}{r(r-1)n^{r-1}}\right). \qquad (5)$$

Now we divide the sum in the definition of a_l into two parts. The first part covers the case where all k_j are at most $n(\frac{r\log\log d}{d})^{\frac{1}{r-1}}$. For this part we use estimate (5). The second part covers the case where at least one of k_j is greater than $n(\frac{r\log\log d}{d})^{\frac{1}{r-1}}$. In this case we denote by i the sum of all such k_j's and use estimate (4) and (5). This way we get the following inequality.

$$a_l \le \binom{mk_0}{l}\log^{\frac{2l}{r-1}} d + n^{\log^{\frac{1}{r-1}} d}\sum_{i=n(\frac{r\log\log d}{d})^{\frac{1}{r-1}}}^{l}\binom{k_0}{i}\exp\left((1+p)\frac{di^r}{r(r-1)n^{r-1}}\right)\binom{mk_0}{l-i}\log^{\frac{2(l-i)}{r-1}} d.$$

For $i_0 = n(\frac{r\log\log d}{d})^{\frac{1}{r-1}} \le i \le l$, set

$$b_{i,l} = \binom{k_0}{i}\exp\left((1+p)\frac{di^r}{r(r-1)n^{r-1}}\right)\binom{mk_0}{l-i}\log^{\frac{2(l-i)}{r-1}} d.$$

Then

$$\begin{aligned}
\frac{b_{i+1,l}}{b_{i,l}} &= \frac{k_0-i}{i+1}\exp\left(\frac{(1+p)d(ri^{r-1}+\binom{r}{2}i^{r-2}+\ldots+1)}{r(r-1)n^{r-1}}\right)\frac{l-i}{mk_0-l+i+1}\log^{\frac{-2}{r-1}} d \\
&= \left(\frac{l+1}{i+1}-1\right)\frac{k_0-i}{mk_0-l+i+1}\exp\left(\frac{(1+p)d(ri^{r-1}+\binom{r}{2}i^{r-2}+\ldots+1)}{r(r-1)n^{r-1}}\right)\log^{\frac{-2}{r-1}} d. \qquad (6)
\end{aligned}$$

Let $i=\alpha k_0$. Since the second factor in (6) is at most $2d^{\frac{-1}{r-1}}\log^6 d$ and the exponential factor is of order $d^{\frac{(1+o(1))r\alpha^{r-1}}{r-1}}$, we get that $b_{i+1,l}/b_{i,l} < 1$ for $0 \le \alpha \le 0.4 < \left(\frac{(1+o(1))}{r}\right)^{\frac{1}{r-1}}$. Thus $b_{i_0,l}$ is maximal for this interval of values of α, and an upper bound for $b_{i_0,l}$ is given by

$$\begin{aligned}
b_{i_0,l} &= \binom{k_0}{i_0}\exp\left((1+p)\frac{di_0^r}{r(r-1)n^{r-1}}\right)\binom{mk_0}{l-i_0}\log^{\frac{2(l-i_0)}{r-1}} d \\
&\le \binom{(m+1)k_0}{l}\log^{\frac{2l}{r-1}} d \le \left(\frac{3mk_0\log^{\frac{2}{r-1}} d}{l}\right)^l.
\end{aligned}$$

Now let $i=\alpha l, 0.4 \le \alpha \le 0.99^{\frac{1}{r-1}}$. Then

$$
\begin{aligned}
b_{i,l} &= \binom{k_0}{\alpha l} \exp\left((1+p)\frac{d(\alpha l)^r}{r(r-1)n^{r-1}}\right)\binom{mk_0}{(1-\alpha)l} \log^{\frac{2(1-\alpha)l}{r-1}} d \\
&\le \left[\left(\frac{ek_0}{\alpha l}\right)^{\alpha} \exp\left((1+p)\frac{\alpha^r l^{r-1} d}{r(r-1)n^{r-1}}\right)\left(\frac{emk_0}{(1-\alpha)l}\right)^{(1-\alpha)} \log^{\frac{2(1-\alpha)}{r-1}} d\right]^l \\
&\le \left[\frac{k_0 d^{\frac{(1-\alpha)+(1+p)\alpha^r}{r-1}} \log^{\frac{2}{r-1}} d}{l}\right]^l \le \left(\frac{3mk_0 \log^{\frac{2}{r-1}} d}{l}\right)^l .
\end{aligned}
$$

Now we are going to bound $b_{i,l}$ from above for $0.99^{\frac{1}{r-1}} l \le i \le l$. We consider two cases. First, let $l \le (0.9)^{\frac{1}{r-1}} k_0$. Then

$$
\begin{aligned}
\max\{b_{i,l} : 0.99^{\frac{1}{r-1}} l \le i \le l\} &\le \binom{2k_0}{l} \exp\left(\frac{(1+p)dl^r}{r(r-1)n^{r-1}}\right)\left(\frac{emk_0}{(1-0.99^{\frac{1}{r-1}})l}\right)^{(1-0.99^{\frac{1}{r-1}})l} \log^{\frac{0.02l}{r-1}} d \\
&\le \left(\frac{2ek_0}{l} \exp\left((1+p)\frac{dl^{r-1}}{r(r-1)n^{r-1}}\right)\left(\frac{n}{l}\right)^{0.01} \log^{\frac{0.02}{r-1}} d\right)^l \\
&\le \left(\frac{6k_0 d^{\frac{0.92}{r-1}}}{l}\right)^l \le \left(\frac{3mk_0 \log^{\frac{2}{r-1}} d}{l}\right)^l .
\end{aligned}
$$

For $l \ge (0.9)^{\frac{1}{r-1}} k_0$, in order to maximize $b_{i,l}$ it is enough to consider $i \ge l - ld^{\frac{-0.3}{r-1}}$. Indeed, for $0.99l \le i \le l - ld^{-0.3}$ the first factor in $\frac{b_{i+1,l}}{b_{i,l}}$ is at least $d^{\frac{-0.3}{r-1}}$, the second one is at least $d^{\frac{-1.3}{r-1}}$ and the exponent is greater then $d^{\frac{0.88r}{r-1}}$, thus $\frac{b_{i+1,l}}{b_{i,l}} > 1$. Furthermore the following inequality holds

$$
\begin{aligned}
\max\{b_{i,l} : l - ld^{\frac{-0.3}{r-1}} \le i \le l\} &\le \binom{k_0}{l - ld^{\frac{-0.3}{r-1}}} \exp\left(\frac{(1+p)dl^r}{r(r-1)n^{r-1}}\right)\left(\frac{emk_0}{ld^{\frac{-0.3}{r-1}}}\right)^{ld^{\frac{-0.3}{r-1}}} \log^{\frac{ld^{\frac{-0.3}{r-1}}}{r-1}} d \\
&\le \binom{k_0}{k_0 - l} \exp\left((1+p)\frac{dl^r}{r(r-1)n^{r-1}}\right) d^{\frac{3}{r-1} k_0 d^{\frac{-0.3}{r-1}}} .
\end{aligned}
$$

Hence

$$
\sum_{l=0}^{k_0} \frac{a_l}{(n-(m+1)k_0)^l} \frac{k_0!}{(k_0-l)!} \le n^{2\log^{\frac{1}{r-1}} d} \sum_{l=0}^{k_0} \left(\frac{2k_0}{n}\right)^l \left(\frac{3mk_0 \log^{\frac{2}{r-1}} d}{l}\right)^l + n^{2\log^{\frac{1}{r-1}} d} d^{\frac{3}{r-1} k_0 d^{\frac{-0.3}{r-1}}} \times
$$

$$
\sum_{j=0}^{\left(1-0.9^{\frac{1}{r-1}}\right)k_0} \frac{k_0!}{(n-(m+1)k_0)^{k_0}} \frac{n^j}{j!} \binom{k_0}{j} \exp\left((1+p)\frac{dk_0{}^r}{r(r-1)n^{r-1}}\right) \exp\left((1+p)\frac{d((k_0-j)^r - k_0{}^r)}{r(r-1)n^{r-1}}\right)
$$

$$\leq n^{2\log^{\frac{1}{r-1}} d} \sum_{l=0}^{k_0} \left(\frac{6mk_0{}^2 \log^{\frac{2}{r-1}} d}{ln} \right)^l + n^{2\log^{\frac{1}{r-1}} d} d^{\frac{3}{r-1} k_0 d^{\frac{-0.3}{r-1}}} \left(\frac{k_0}{en(1-\log^{-4.9} d)} \exp\left(\frac{(1+p)dk_0{}^{r-1}}{r(r-1)n^{r-1}} \right) \right)^{k_0}$$

$$\times \sum_{j=0}^{\left(1-0.9^{\frac{1}{r-1}}\right)k_0} \left(\frac{9k_0 n}{j^2} \exp\left(-\frac{(1+p)d(k_0{}^{r-1} + k_0{}^{r-2}(k_0-j) + \ldots + (k_0-j)^{r-1})}{r(r-1)n^{r-1}} \right) \right)^j .$$

All terms of the first sum are less than $\exp\left(6mk_0{}^2 \log^{\frac{2}{r-1}} d/n\right)$, for d large enough, whereas for the second sum we have

$$\frac{k_0}{en(1-\log^{-4.9} d)} \exp\left(\frac{(1+p)dk_0{}^{r-1}}{r(r-1)n^{r-1}} \right) < \frac{1}{e(1-\log^{-4.9} d)} \left(\frac{d}{\log^2 d} \right)^{\frac{1+p}{r-1}} \left(\frac{r\log d}{d} \right)^{\frac{1}{r-1}} < \frac{4d^{\frac{p}{r-1}}}{e\log^{\frac{1}{r-1}} d} < 1$$

and

$$\exp\left(-\frac{(1+p)d(k_0{}^{r-1} + k_0{}^{r-2}(k_0-j) + \ldots + (k_0-j)^{r-1})}{r(r-1)n^{r-1}} \right)$$

$$< \exp\left(-\frac{(1+p)dk_0{}^{r-1}(1 + 0.9^{\frac{1}{r-1}} + \ldots + 0.9^{\frac{r-1}{r-1}})}{r(r-1)n^{r-1}} \right) < \left(\frac{d}{\log^2 d} \right)^{-\frac{1+0.9(r-1)}{r-1}} .$$

Therefore

$$d^{\frac{3}{r-1} k_0 d^{\frac{-0.3}{r-1}}} \sum_{j=0}^{\left(1-0.9^{\frac{1}{r-1}}\right)k_0} \left(\frac{9k_0 n \log^{\frac{2+1.8(r-1)}{r-1}} d}{j^2 d^{\frac{1+0.9(r-1)}{r-1}}} \right)^j < \exp\left(3k_0 d^{\frac{-0.25}{r-1}} + nd^{\frac{-1.3}{r-1}}\right)$$

Thus

$$\sum_{l=0}^{k_0} \frac{a_l}{(n-(m+1)k_0)^l} \frac{k_0!}{(k_0-l)!} \leq n^{2\log^{\frac{1}{r-1}} d} \exp\left(\frac{6mk_0{}^2 \log^{\frac{2}{r-1}} d}{n} \right) + n^{2\log^{\frac{1}{r-1}} d} \exp\left(nd^{\frac{-1.2}{r-1}}\right)$$

and, finally we arrive at

$$Pr(Y>0) \geq \frac{(EY)^2}{EY^2} \geq \left[\exp\left(\frac{7mk_0{}^2 \log^{\frac{2}{r-1}} d}{n} \right) \right]^{-m}$$

$$\geq \exp\left(-7n\log^{\frac{2}{r-1}} d \left(\frac{k_0 m}{n} \right)^2 \right) \geq \exp\left(-n\log^{-7.5} d\right) .$$

This completes the proof of Lemma 5.2. $\square$

The next lemma and its proof are shaped after Lemma 2 of [7].

Lemma 5.3 *There is a constant d_0 such that if $d > d_0$ and $d = o(n^{r-1})$, then with probability $1 - o(1) - \log^{-1} d$, more than $n - 2n\log^{-3} d$ vertices of $H_r(n,p)$ can be properly colored with less than*

$$\left(\frac{d}{r\log d} \left(1 + \frac{(28r-23)\log\log d}{\log d} \right) \right)^{\frac{1}{r-1}}$$

colors.

Proof. In the proof we shall use the "expose-and-merge" technique introduced by D. Matula [9]. For $A \subset [n]$ define $[A]^r = \{|S| = r | S \subset A\}$ and let

$$\overline{k_0} = n\left(\frac{r(\log d - (28r-24)\log\log d)}{d}\right)^{\frac{1}{r-1}}, \quad l_0 = n/(\overline{k_0}\log^{33} d).$$

Consider the following algorithm:

Algorithm

$E := \emptyset$

$F_0 := \emptyset$

$W_0 := \emptyset$

for $i = 1$ **to** $\log^{33} d - \log^{30} d$ **do**

  begin

  choose randomly $A_i \subset [n] \setminus W_i$ with $|A_i| = n\log^{-28} d$;

  define $\mathcal{H}_i$ as the hypergraph with the set of vertices A_i and the set of edges E_i, where

    each $e \in [A_i]^r$ belongs to E_i randomly and independently with probability p;

  choose a family $\{\mathcal{R}_1^i, \dots, \mathcal{R}_{l_0}^i\}$ of disjoint independent sets from A_i, such that

    $\sum_{l=0}^{l_0} |\mathcal{R}_l^i| = n\log^{-33} d$ - if it is not possible FAIL;

  $E'_i := E_i \setminus (E_i \cap F_{i-1})$;

  $E := E \cup E'_i$;

  $F_i := F_{i-1} \cup [A_i]^r$;

  $W_i := W_{i-1} \cup \bigcup_{l=1}^{l_0} \mathcal{R}_l^i$;

  end

$\overline{F} := [n]^r \setminus \bigcup_{i=1}^{\log^{33} d - \log^{30} d} F_i$;

form $\overline{E} \subset \overline{F}$ by choosing each $e \in \overline{F}$ to belong to $\overline{E}$ randomly and independently with probability p;

$E := E \cup \overline{E}$;

output E; $\{\mathcal{R}_1^1, \mathcal{R}_2^1, \dots, \mathcal{R}_{l_0}^{\log^{33} d - \log^{30} d}\}$;

end

Let us first observe that the probability that $e \in E$ is equal to p for each $e \in [n]^r$, thus the hypergraph $\mathcal{H}$ with the set of vertices $[n]$ and the set of edges E may be treated as $H_r(n,p)$.

Obviously, we may consider each $\mathcal{H}_i$ as $H_r(\overline{n}, p)$, where $\overline{n} = n\log^{-28} d$. Let $\overline{d} = d(\overline{n}, p) = (r-1)\binom{\overline{n}-1}{r-1}p = (1+o(1))d\log^{-28(r-1)} d$. Then

$$\overline{k_0} = n\left(\frac{r(\log d - (28r-24)\log\log d)}{d}\right)^{\frac{1}{r-1}} < \overline{n}\left(\frac{r(\log \overline{d} - 3\log\log \overline{d})}{\overline{d}}\right)^{\frac{1}{r-1}}.$$

Thus from Lemma 5.2, the probability that $\mathcal{H}_i$ contains no subset with $n\log^{-33} d < \overline{n}\log^{-5}\overline{d}$ elements which can be properly colored using $\overline{n}\log^{-5}\overline{d}/\overline{k_0}$ colors is less than $\overline{n}^{-3}$, so the probability of FAIL in the Algorithm is less than n^{-2}.

Thus, with probability at least $1-n^{-2}$, the Algorithm finds

$$(\log^{33} d - \log^{30} d) l_0 = \frac{n - n\log^{-3} d}{\overline{k_0}} < \left(\frac{d}{r\log d}\left(1 + \frac{(28r-23)\log\log d}{\log d}\right)\right)^{\frac{1}{r-1}}.$$

disjoint sets $\mathcal{R}_1^1, \mathcal{R}_2^1, \ldots, \mathcal{R}_{l_0}^{\log^{33} d - \log^{30} d}$ such that

$$\sum_i \sum_l |\mathcal{R}_l^i| = (\log^{33} d - \log^{30} d) n \log^{-33} d = n - n\log^{-3} d.$$

Note that although $\mathcal{R}_l^i$ is an independent set in $\mathcal{H}_i$ it is not necessarily independent as a subset of $\mathcal{H}$. Let X denote the number of edges of $\mathcal{H}$ contained in $\mathcal{R}_l^i$ for some $1 \le i \le \log^{33} d - \log^{30} d, 1 \le l \le l_0$. We shall estimate X from above.

Let $e \in [n]^r$ be such that $e \in E$ and also $e \subset \mathcal{R}_l^i$ for some i, l. This can occur only if e was chosen as an edge in $\mathcal{H}_j$ for $j < i$. Thus we have that $e \subset A_i$ and also $e \subset A_j$. Since for all i we choose A_i from the set of vertices of size at least $n\log^{-3} d$, the probability that for given i and j the edge e is contained in both A_i and A_j is less than $(\log^{-25} d)^{2r}$. Now observe that for any $1 \le l \le l_0$ each subset of A_j containing a $|\mathcal{R}_l^j|$ elements is equally likely to be chosen as $\mathcal{R}_l^j$(this event depends only on the structure of $\mathcal{H}_j$ which is symmetric with respect to the labeling of vertices). Due to Lemma 3.1, we may assume that $|\mathcal{R}_l^j| < n\left(\frac{r\log d}{d}\right)^{\frac{1}{r-1}}$ for all i, l. Thus, since $|A_j| = n\log^{-28} d$, the probability that all the vertices of the edge e are in the same set $\mathcal{R}_l^j$ for some l is less than

$$\left(\frac{n\left(\frac{r\log d}{d}\right)^{\frac{1}{r-1}}}{n\log^{-28} d}\right)^{r-1} \le \frac{r\log^{(28r-27)} d}{d}.$$

The probability that the edge e appears in the hypergraph $\mathcal{H}_j$ is equal to p, so finally we have the following upper bound on the expectation of X:

$$EX \le \binom{n}{r}\binom{\log^{33} d - \log^{30} d}{2} \cdot \log^{-50r} d \cdot p \cdot \frac{r\log^{(28r-27)} d}{d} < n\log^{-5} d.$$

Therefore, from Markov's inequality,

$$Pr(X > \frac{1}{r} n\log^{-3} d) < \log^{-1} d.$$

Now for all i, l delete from $\mathcal{R}_l^i$ all the vertices of the edges of $\mathcal{H}$ that are contained in $\mathcal{R}_l^i$ and denote the obtained sets by $\overline{\mathcal{R}_l^i}$. Then

$$\sum_{i,l} |\overline{\mathcal{R}_l^i}| \ge \sum_{i,l} |\mathcal{R}_l^i| - rX = n - n\log^{-3} d - rX$$

so

$$Pr\left(\sum_{i,l} |\overline{\mathcal{R}_l^i}| \ge n - 2n\log^{-3} d\right) > 1 - n^{-2} - \log^{-1} d\ ,$$

and the assertion of the lemma follows. $\Box$

Proof of Theorem 5.1. ¿From Lemmas 5.3 and 4.1 we have the following estimate:

$$Pr\left(\chi(H_r(n,p)) < \left(\frac{d}{r\log d}\left(1+\frac{(28r-22)\log\log d}{\log d}\right)\right)^{\frac{1}{r-1}}\right) > 1-o(1)-\log^{-1}d > 1/2\ .$$

In order to show that essentially the same upper bound on $\chi_\gamma(H_r(n,p))$ holds with much higher probability we use the argument of Frieze. Define $Y(H_r(n,p))$ to be the minimal size of a set of vertices S, such that the induced subhypergraph $H[V\backslash S]$ can be colored by $w = \left(\frac{d}{r\log d}\left(1+\frac{(28r-22)\log\log d}{\log d}\right)\right)^{\frac{1}{r-1}}$ colors. Then Y is a random variable that satisfies the Lipschitz condition (see [2], Ch. 7 for more details). Denote $E(Y)=\mu$ and apply the vertex exposure martingale on $H_r(n,p)$. Then by Azuma's inequality

$$Pr\left(Y \le \mu - \lambda\sqrt{n-1}\right) < e^{-\lambda^2/2}\ ,$$
$$Pr\left(Y \ge \mu + \lambda\sqrt{n-1}\right) < e^{-\lambda^2/2}\ .$$

Let λ satisfy $e^{-\lambda^2/2}=n^{-2}$. Since $Pr(Y=0)>1/2$ we have that $\mu \le \lambda\sqrt{n-1} = c\sqrt{n\log n}$ for some constant $c>0$. Thus $Pr(Y \ge 2\lambda\sqrt{n-1}) < n^{-2}$. Therefore with probability $1-n^{-2}$ there is a w-coloring of all but at most $c\sqrt{n\log n}$ vertices. By Lemma 4.1 with probability $1-o(1/n)$ these vertices can be colored by $\left(\frac{d}{\log^{2(r-1)-0.7}d}\right)^{\frac{1}{r-1}}$ colors. Combining these two colorings we get that with probability $1-o(1/n)$ the hypergraph $H_r(n,p)$ can be colored by at most

$$\left(\frac{d}{r\log d}\left(1+\frac{28r\log\log d}{\log d}\right)\right)^{\frac{1}{r-1}}$$

colors. This finishes the proof of Theorem 5.1. $\Box$

6 The γ-chromatic numbers, $\gamma < r-1$

In this section we finish the proof of Theorem 1. Due to the results of the previous section it remains to treat the case $\gamma < r-1$. For reader's convenience, we restate the formulation of Theorem 1 for this case.

Theorem 6.1 *For every $1 \le \gamma \le r-2$ there exists a constant d_0 such that if $d^{(\gamma)} = d^{(\gamma)}(n,p) = \gamma\binom{r-1}{\gamma}\binom{n-1}{r-1}p \ge d_0$ but $d^{(\gamma)} = o(n^\gamma)$, then almost surely*

$$\left(\frac{d^{(\gamma)}}{(\gamma+1)\log d^{(\gamma)}}\right)^{\frac{1}{\gamma}} \le \chi_\gamma(H_r(n,p)) \le \left(\frac{d^{(\gamma)}}{(\gamma+1)\log d^{(\gamma)}}\left(1+\frac{1}{\log^{0.1}d^{(\gamma)}}\right)\right)^{\frac{1}{\gamma}}\ .$$

Proof of Theorem 6.1. The lower bound follows immediately from Corollary 3.2, so it is enough to prove the upper bound. The proof of the upper bound relies on the upper bound for the weak chromatic number, given by Theorem 5.1.

Throughout the proof, the symbols $C_1, C_2, \ldots$ denote positive constants dependinng only on r. Let

$$u_0 = u_0(n_0, p_{\gamma+1}, \gamma+1) = \left(\frac{d(n_0, p_{\gamma+1}, \gamma+1)}{(\gamma+1)\log d(n_0, p_{\gamma+1}, \gamma+1)} \left(1 + \frac{28(\gamma+1)\log\log d(n_0, p_{\gamma+1}, \gamma+1)}{\log d(n_0, p_{\gamma+1}, \gamma+1)} \right) \right)^{\frac{1}{\gamma}} ,$$

where $d(n_0, p_{\gamma+1}, \gamma+1) = \binom{n_0-1}{\gamma}\gamma p_{\gamma+1}$ and the exact value of the edge probability $p_{\gamma+1}$ will be defined later. Note that according to Theorem 5.1 the expression in the definition of u_0 is an upper bound on the weak chromatic number of a random $\gamma+1$-uniform hypergraph on n_0 vertices with edge probability $p_{\gamma+1}$.

Let $s = \log^4 d^*$. We fix a partition of the vertex set V into s disjoint parts $V_1, \ldots, V_s$ of sizes $\lfloor \frac{n}{s} \rfloor$ or $\lceil \frac{n}{s} \rceil$. For the sake of simplicity we assume that n/s is integer and then $|V_i| = n/s$ for $1 \leq i \leq s$. Let $n_0 = n/s$.

For $1 \leq i \leq s$, let $H_i = H_{V_i}^{\geq \gamma+1}$. A key step in the proof of Theorem 6.1 is the following lemma.

Lemma 6.2 *Let $p_{\gamma+1} = 1 - (1-p)^{\binom{n-n_0}{r-\gamma-1}}$. There exists a constant C_1 such that, for every fixed i, $1 \leq i \leq s$, with probability $1 - o(s/n)$ all but at most $C_1 n_0 / \log^3 d^*$ vertices of the hypergraph H_i can be γ-colored by at most $u_0 = u_0(n_0, p_{\gamma+1}, \gamma+1)$ colors.*

Proof. We represent

$$H_i = \bigcup_{j=\gamma+1}^{r} H_{i,j} \ ,$$

where $H_{i,j} = H_{V_i}^{j}$. It is easy to see that the edges of $H_{i,\gamma+1}$ constitute a great part of the edges of H_i, so the most important task for us will be to color properly the subhypergraph $H_{i,\gamma+1}$. For each $\gamma+1 \leq j \leq r$ denote

$$p_j = 1 - (1-p)^{\binom{n-n_0}{r-j}}$$

and note that

$$\left(\frac{n}{2(r-j)} \right)^{r-j} p \ \leq p_j \ \leq \binom{n}{r-j} p \ . \tag{7}$$

Now observe crucially that for every subset $e \subset V_i$ of size $|e| = j$, where $\gamma+1 \leq j \leq r$, the probability of the event "e is an edge of H_i" is exactly p_j. Moreover, all such events are mutually independent. This enables us to treat each of the subhypergraphs $H_{i,j}$ as a random hypergraph from the probability space $H_j(n_0, p_j)$.

Let us first expose the edges of $H_{i,\gamma+1}$. Note that a γ-coloring of this subhypergraph is a *weak* coloring, since all edges of $H_{i,\gamma+1}$ have size $\gamma+1$. Recalling that $H_{i,\gamma+1}$ is an element of the probability space $H_r(n_0, p_{\gamma+1})$, we can use the result of Theorem 5.1 in order to claim that with probability $1-o(s/n)$ the γ-chromatic number of $H_{i,\gamma+1}$ is at most u_0. Moreover, due to Lemma 3.1 the γ-independence number of $H_{i,\gamma+1}$ is at most, say, $2n_0/u_0$ with probability $1-o(s/n)$.

Fix some γ-coloring $f: V_i \to \{1,\dots,u_0\}$ of $H_{i,\gamma+1}$ by u_0 colors with color classes $T_1,\dots,T_{u_0}$, each of size at most $2n_0/u_0$. Now we expose the edges of the subhypergraphs $H_{i,j}$ for $j>\gamma+1$. We call an edge $e \in \bigcup_{j=\gamma+2}^{r} E(H_{i,j})$ *bad* if it improperly γ-colored by f, that is, there exists a color class T_l such that $|e\cap T_l| \geq \gamma+1$. Let X_i be the number of bad edges in H_i. Note that each bad edge has at least $\gamma+2$ vertices in common with V_i. Therefore the expectation of X_i can be estimated as follows.

$$\begin{aligned}
EX_i &\leq p\sum_{l=1}^{u_0}\left(\binom{|T_l|}{\gamma+1}n_0\binom{n}{r-\gamma-2}+\sum_{j=\gamma+2}^{r}\binom{|T_l|}{j}\binom{n}{r-j}\right)\\
&\leq u_0p\left(\binom{\frac{2n_0}{u_0}}{\gamma+1}n_0\binom{n}{r-\gamma-2}+(r-\gamma-1)\binom{\frac{2n_0}{u_0}}{\gamma+2}\binom{n}{r-\gamma-2}\right)\\
&\leq 2u_0p\binom{\frac{2n_0}{u_0}}{\gamma+1}n_0\binom{n}{r-\gamma-2} \leq \frac{2u_0p\left(\frac{2n_0}{u_0}\right)^{\gamma+1}n_0n^{r-\gamma-2}}{(\gamma+1)!} \leq \frac{4n^{r-\gamma-2}n_0^{\gamma+2}p}{u_0^{\gamma}}\ .
\end{aligned}$$

Substituting the expression from the definition of u_0 and using estimate (7) for $p_{\gamma+1}$, we get

$$\begin{aligned}
EX_i &\leq \frac{4n^{r-\gamma-2}n_0^{\gamma+2}p}{\frac{d(n_0,p_{\gamma+1},\gamma+1)}{(\gamma+1)\log d(n_0,p_{\gamma+1},\gamma+1)}} \leq \frac{4(\gamma+1)n^{r-\gamma-2}n_0^{\gamma+2}\log(n^{r-1}p)p}{\left(\frac{n_0-1}{\gamma}\right)^{\gamma}\gamma\left(\frac{n}{2(r-\gamma-1)}\right)^{r-\gamma-1}p}\\
&\leq C_2\frac{n_0^2\log d^*}{n} = C_2\frac{n_0}{\log^3 d^*}\ .
\end{aligned}$$

The random variable X_i is naturally represented as a sum of independent indicator random variables and thus standard exponential bounds on the tails of X_i can be applied to show that with probability $1-o(s/n)$ we have $X_i \leq 2C_2n_0/\log^3 d^*$. Take a union U_i of all bad edges, it contains at most $2rC_2n_0/\log^3 d^*$ vertices. The subhypergraph of H_1 spanned by $V_i \setminus U_i$ is clearly γ-colorable in u_0 colors, thus proving the claim of the lemma with $C_1 = 2rC_2$. □

Applying Lemma 6.2 to each of the subsets V_i for $1 \leq i \leq s$ and using distinct colors for every i, we get the following intermediate result:

Corollary 6.3 *There exists a constant C_1 such that with probability $1-o(s^2/n)$ all but at most $C_1n/\log^3 d^*$ vertices of $H_r(n,p)$ can be γ-colored in at most su_0 colors.*

Now we plug in Lemma 4.1 and get the following upper bound on the γ-chromatic number of $H_r(n,p)$.

Corollary 6.4 *With probability $1-o(s^2/n)$ the γ-chromatic number of $H_r(n,p)$ is at most $su_0+(d^*/\log^{2\gamma-0.7} d^*)^{1/\gamma}$.*

The proof is almost finished. It remains only to write the bound from Corollary 6.4 in a more "visible" form. This can be done by performing routine (but rather tedious) arithmetic manipulations. Below we present them in a somewhat scketchy way. Let A denote the bound of Corollary 6.4. We will use the following trivial inequalities: $1/(a-b) \leq (1+2b/a)/a$ for $0 < b \leq a/2$; and $(1+a)(1+b) \leq 1+3a$ for $0 < b \leq a \leq 1$. Substituting the definition of u_0 and denoting for brevity $d(n_0) = d(n_0, p_{\gamma+1}, \gamma+1)$ we get:

$$\begin{aligned} A &\leq s\left(\frac{d(n_0)}{(\gamma+1)\log d(n_0)}\left(1+\frac{28(\gamma+1)\log\log d(n_0)}{\log d(n_0)}\right)\right)^{\frac{1}{\gamma}} + \left(\frac{d^*}{\log^{2\gamma-0.7} d^*}\right)^{\frac{1}{\gamma}} \\ &\leq s\left(\frac{d(n_0)}{(\gamma+1)\log d(n_0)}\left(1+\frac{1}{\log^{0.2} d(n_0)}\right)\right)^{\frac{1}{\gamma}} . \end{aligned}$$

Now we bound $d(n_0)$ as follows:

$$d(n_0) \leq \binom{n_0-1}{\gamma}\gamma\binom{n-n_0}{r-\gamma-1}p \leq \frac{n_0^\gamma}{\gamma!}\gamma\frac{n^{r-\gamma-1}}{(r-\gamma-1)!}p = \frac{\gamma n^{r-1}p}{\gamma!(r-\gamma-1)!s^\gamma} .$$

On the other hand, $d(n_0) \geq n_0^\gamma n^{r-\gamma-1}p/C_3 = d^*/C_3 s^\gamma$. Then $\log d(n_0) \geq \log d^* - \log(C_3 s^\gamma)$ and thus $1/\log d(n_0) \leq (1 + C_4 \log\log d^*/\log d^*)/\log d^*$. We get

$$A \leq \left(\frac{\gamma}{(\gamma+1)!(r-\gamma-1)!}\frac{d^*}{\log d^*}\left(1+\frac{C_5}{\log^{0.2} d^*}\right)\right)^{\frac{1}{\gamma}} .$$

Now, we use the following inequalities: $n^{r-1} \leq \binom{n-1}{r-1}(r-1)!(1+2r^2/n)$ and also $\log d^* \geq \log d^{(\gamma)} - \log C_6$. Then

$$\begin{aligned} A &\leq \left(\frac{\gamma}{\gamma+1}\frac{(r-1)!}{\gamma!(r-\gamma-1)!}\frac{d^*}{(r-1)!\log d^*}\left(1+\frac{C_5}{\log^{0.2} d^*}\right)\right)^{\frac{1}{\gamma}} \\ &\leq \left(\frac{\gamma}{\gamma+1}\binom{r-1}{\gamma}\binom{n-1}{r-1}p\left(1+\frac{C_7}{\log^{0.2} d^*}\right)\right)^{\frac{1}{\gamma}} \\ &\leq \left(\frac{d^{(\gamma)}}{(\gamma+1)\log d^{(\gamma)}}\left(1+\frac{1}{\log^{0.1} d^{(\gamma)}}\right)\right)^{\frac{1}{\gamma}} . \qquad \square \end{aligned}$$

7 Concluding remarks

We have established an asymptotic behavior of the γ-chromatic number of a random r-uniform hypergraph $H_r(n,p)$ for all values of the parameter γ and for all values of the edge probability $p = p(n)$ down to the case $p = cn^{-r+1}$ for some constant $c > 0$. As in the graph case, it turned out that the γ-chromatic number of the random hypergraph is asymptotically equal to the ratio of the number of vertices n and its γ-independence number. Though this paradigm can be carried over from the graph case to that of hypergraphs, the proof is similar only for the case of the weak chromatic number, corresponding to $\gamma = r-1$. The reason for it is that the standard martingale based techniques, successfully used to establish the asymptotic value of the chromatic number of the random graph $G(n,p)$, do not seem to be directly applicable for $\gamma < r-1$. This is due to the fact that the γ-chromatic number of a hypergraph H is a vertex Lipschitz function only for the case of $\gamma = r-1$. We succeeded to bypass this difficulty by partitioning the vertex set $V(H)$ into $s = \log^4(n^{r-1}p$ parts $V_1, \dots, V_s$ and coloring each part separately. This partition enabled us to use the arguments from the case of the weak chromatic number to color each subhypergraph $H[V_i]$ and thus to reduce a general case to that of the weak chromatic number.

An interesting related problem, for which the above mentioned difference between the case $\gamma = r-1$ and the other cases may also play an important role, is that of determining a concentration of the γ-chromatic number of $H_r(n,p)$. For a hypergraph theoretic function $X(H)$ and probability space $H_r(n,p)$, we say that $X(H_r(n,p))$ is *concentrated in width* $s = s(n,p)$ if there exists a function $u = u(n,p)$ so that

$$\lim_{n\to\infty} Pr(u \le X(H_r(n,p)) \le u+s) = 1 \ .$$

The question of estimating the width of concentration of the chromatic number of a random graph $G(n,p)$ is studied in papers [13], [8], [1]). All these papers rely heavily on the fact that the chromatic number of a graph is vertex Lipschitz. In contrast, nobody seems to have addressed the corresponding hypergraph question. Similarly to the problem of determining the asymptotic value of the γ-chromatic number of $H_r(n,p)$, the case of the weak chromatic number should be quite similar to the graph case, and most of the results about the concentration of the chromatic number of $G(n,p)$ can be transferred to this special hypergraph case. However, for every $\gamma < r-1$ a simple adaptation of the graph arguments does not seem to be possible. New ideas are required to tackle this case.

Acknowledgment. The authors would like to thank the anonymous referees for many helpful comments and in particular for suggesting the use of Talagrand's inequality for proving Lemma 5.2.

References

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

[2] N. Alon and J.Spencer, **The probabilistic method**, Wiley, New York, 1992.

[3] C. Berge, **Hypergraphs**, North-Holland, Amsterdam, 1989.

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

[5] B. Bollobás, **Random graphs**, Academic Press, London, 1985.

[6] M. Karoński and T. Łuczak, *Random hypergraphs*, in Combinatorics, Paul Erdős is eighty, D. Miklós et al. Eds., Bolyai Soc. Math. Studies Vol. 2, 283–293.

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

[8] T. Łuczak, *A note on the sharp concentration of the chromatic number of random graphs*, Combinatorica 11 (1991), 45–54.

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

[10] J. P. Schmidt, *Probabilistic analysis of strong hypergraph coloring algorithms and the strong chromatic number*, Discrete Math. 66 (1987), 259–277.

[11] J. Schmidt-Pruzan, E. Shamir and E. Upfal, *Random hypergraph coloring algorithms and the weak chromatic number*, J. Graph Theory 8 (1985), 347–362.

[12] E. Shamir,*Chromatic number of random hypergraphs and associated graphs*, in Advances in Computing Research 5 (1989), 127–142.

[13] E. Shamir and J. Spencer, *Sharp concentration of the chromatic number of random graphs $G_{n,p}$*, Combinatorica 7 (1987), 121–130.

[14] J.Spencer, *Probabilistic methods in combinatorics*, in Proceedings Int. Congress of Math., Zürich, 1994, 1375–1383.