Two-coloring Random Hypergraphs*

Dimitris Achlioptas[†], Jeong Han Kim[‡], Michael Krivelevich[§], and Prasad Tetali[¶]

May 31, 2001

Abstract

A 2-coloring of a hypergraph is a mapping from its vertex set to a set of two colors such that no edge is monochromatic. Let $H = H(k, n, p)$ be a random k-uniform hypergraph on a vertex set V of cardinality n, where each k-subset of V is an edge of H with probability p, independently of all other k-subsets. Let $m = p\binom{n}{k}$ denote the expected number of edges in H. Let us say that a sequence of events $\mathcal{E}_n$ holds *with high probability* (w.h.p.) if $\lim_{n\to\infty} \Pr[\mathcal{E}_n] = 1$. It is easy to show that if $m = c2^k n$ then w.h.p. H is not 2-colorable for $c > \frac{\ln 2}{2}$. We prove that there exists a constant $c > 0$ such that if $m = (c2^k/k)n$, then w.h.p. H is 2-colorable.

1 Introduction

For an integer $k \geq 2$, a *k-uniform hypergraph* H is an ordered pair $H = (V, E)$, where V is a finite non-empty set, called the set of *vertices* of H, and E is a family of distinct k-subsets of V, called the *edges* of H. For

*Research performed, in part, while Krivelevich and Tetali visited Microsoft Research.

[†]Microsoft Research, One Microsoft Way, Redmond, WA 98052, U.S.A.; research supported, in part, by an NSERC PDF.

[‡]Microsoft Research, One Microsoft Way, Redmond, WA 98052, U.S.A.

[§]Department of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel.

[¶]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, U.S.A.; research supported, in part, by NSF Grant no. 9800351.

general hypergraph terminology and background we refer the reader to [3]. A 2-*coloring* of a hypergraph $H = (V, E)$ is a partition of its vertex set V into two (color) classes, R and B (for Red and Blue, say), so that no edge in E is monochromatic. A hypergraph is *2-colorable* if it admits a 2-coloring.

The property of 2-colorability was introduced and studied by Bernstein [4] in the early 1900s for infinite hypergraphs. The 2-colorability of finite hypergraphs, also known as "Property B" (a term coined by Erdős in reference to Bernstein), has been studied for about eighty years (see, e.g. [8, 9, 2, 17]). For $k = 2$, i.e. for graphs, the problem is well understood, since graph 2-colorability is equivalent to having no odd cycle. For $k \geq 3$, though, much less is known and deciding the 2-colorability of k-uniform hypergraphs is NP-complete for every fixed $k \geq 3$ [15].

In this paper we discuss 2-colorability of random k-uniform hypergraphs for $k \geq 3$. (For the evolution of odd cycles in random graphs see [10]). Let $H(k, n, p)$ be a random k-uniform hypergraph on n labeled vertices $V = \{1, \ldots, n\}$, where each k-subset of V is chosen to be an edge of H independently and with probability $p = p(n)$. We will study asymptotic properties of $H(k, n, p)$, i.e. we will consider $k \geq 3$ to be arbitrary but fixed, while n, the number of vertices, tends to infinity. We will say that a hypergraph property A holds *with high probability* (w.h.p.) in $H(k, n, p)$ if $\lim_{n\to\infty} \Pr[H(k, n, p) \text{ has } A] = 1$. The main question in this setting is:

As p is increased, when does $H(k, n, p)$ stop being w.h.p. 2-colorable?

As it will be convenient to discuss the answer to this question in terms of the expected number of edges in $H(k, n, p)$, we denote $m = p\binom{n}{k}$.

Alon and Spencer considered the above question in [1]. They noted that by considering the expected number of 2-colorings of $H(k, n, p)$ it is easy to show that if $m = c2^k n$, where $c > \frac{\ln 2}{2}$, then w.h.p. $H(k, n, p)$ is not 2-colorable. Their main contribution was providing a lower bound on the expected number of edges necessary for $H(k, n, p)$ not to be 2-colorable w.h.p. In particular, by applying the Lovász Local Lemma, they were able to show that if $m = (c2^k/k^2)n$ then w.h.p. $H(k, n, p)$ is 2-colorable, for some small constant $c > 0$. Thus, the gap between the upper and the lower bounds of [1] is of order k^2.

It is interesting to compare the 2-colorability of random k-uniform hypergraphs with the satisfiability problem for random k-SAT formulas. For a set of n Boolean variables, let C_k denote the set of all $2^k\binom{n}{k}$ possible disjunctions of k distinct, non-complementary literals (k-clauses) on those variables. A random k-SAT formula, $F(k, n, m)$, is formed by selecting uniformly at random m clauses from C_k and taking their conjunction. The question now is "as m is increased, when does $F(k, n, m)$ stop being w.h.p. satisfiable?" Again, by considering the expected number of solutions (here, satisfying assignments), it is easy to show if $m = c2^k n$, where $c > \ln 2$, then w.h.p. $F(k, n, m)$ is unsatisfiable. In the opposite direction, Chao and Franco [6] proved that, for $k \geq 4$, a random k-SAT formula with $m = c(2^k/k)n$ clauses is w.h.p. satisfiable, if $c < 1/4$. Chvátal and Reed [7] extended the result of [6] to all $k \geq 2$ (and simplified it), while Frieze and Suen [12], inter alia, improved the constant to $c < 1$.

The similarity between the two problems is quite apparent, though probably cannot be translated into a formal statement. This similarity stimulated Alon and Spencer [1] to try and derive a result for random hypergraph 2-colorability analogous to the random k-SAT result of [7]. While their result [1], as mentioned above, falls short of that goal, the authors proposed a randomized 2-coloring algorithm similar to the one used by Chvátal and Reed [7], and conjectured that w.h.p. it 2-colors $H(k,n,p)$, as long as $m = c(2^k/k)n$, for some absolute constant $c > 0$. If true, that would reduce the gap between the upper and lower bounds for random hypergraph 2-colorability to a factor of k (from k^2).

In this paper we introduce a deterministic algorithm which is similar to the one proposed by Alon and Spencer, except for one crucial difference that simplifies the analysis greatly. (We present and compare the two algorithms in Section 2.1.) We prove that our algorithm w.h.p. finds a proper 2-coloring of $H(k,n,p)$ if $m = c(2^k/k)n$, for an absolute constant $c > 0$.

Theorem 1 *There exists a deterministic, linear-time algorithm which 2-colors $H(k,n,p)$ w.h.p. if the edge probability $p = p(n)$ satisfies*

$$p\binom{n}{k} = c\frac{2^k}{k}n \ ,$$

where $c \leq 1/50$. For $k \geq 40$, $c \leq 1/50$ can be replaced with $c \leq 1/10$.

Let us note that a recent result of Friedgut [11] can be used to show that for each $k \geq 3$, there exists a function $r_k(n)$ such that if $m = (r_k(n) - \varepsilon)n$ then w.h.p. $H(k,n,p)$ is 2-colorable, but if $m = (r_k(n) + \varepsilon)n$ then w.h.p. $H(k,n,p)$ is not 2-colorable. Naturally, $c\frac{2^k}{k} < r_k(n) < c'2^k$, for some absolute constants $c, c' > 0$. It is widely believed that one can replace $r_k(n)$ by a constant r_k. Closing the asymptotic gap in the order of $r_k(n)$ is a challenging open problem in that direction.

The rest of the paper is organized as follows. In Section 2 we present our algorithm, analyze its performance on $H(k,n,p)$ and prove Theorem 1. Section 3 is devoted to a concluding discussion. As noted before, throughout the paper we assume n to be large enough whenever needed, while keeping in mind that k is fixed. Also, for the sake of clarity of presentation we routinely omit floor and ceiling signs.

2 Proof of the main result

We will present a deterministic algorithm A for 2-coloring k-uniform hypergraphs and analyze it for $k \geq 6$. We will show that for such k, A indeed 2-colors $H(k,n,p)$ w.h.p. for $p(n)$ as in Theorem 1. We will treat the (easy) case $3 \leq k \leq 5$ separately at the end of this section.

2.1 Algorithm description

We first present a deterministic version, called dCR, of the Alon and Spencer [1] algorithm, followed by our algorithm A.

The input is always assumed to be a k-uniform hypergraph $H = (V, E)$ with $V = \{1, \ldots, n\}$. To describe the two algorithms it will be convenient to fix in advance an ordering on the vertices of V, say, the natural ordering $1, \ldots, n$. This ordering induces the corresponding lexicographic order on the subsets of V. Thus, for example, $\{1,2\} < \{1,2,3\} < \{3\}$. Also, for the sake of presentation, we will temporarily assume that the number of vertices n is even.

Both algorithms proceed in rounds, $t = 0, 1, \ldots$ Given a partial coloring of the vertices of V, we say that a k-subset of V is *i-monochromatic* if precisely i of its vertices have been colored and all i of them have received the same color.

ALGORITHM dCR

If there are $(k-2)$- or $(k-1)$-monochromatic edges
  then
    let x be the smallest uncolored vertex in the lexicographically
      smallest such edge e;
    color x so as to make e bichromatic
  else
    let x be the smallest uncolored vertex;
    color x Red or Blue, each color with probability $1/2$.

Remark: The algorithm in [1] is called CR, in reference to Chvátal and Reed, and it considers $(k-2)$-monochromatic edges only if there are no $(k-1)$-monochromatic edges. Also, e and x are chosen randomly rather than lexicographically. It is not hard to see that when applied to $H(k, n, p)$ these differences are inconsequential.

We see that dCR permanently colors one vertex in each round, giving priority to edges that are near-monochromatic, i.e. having one or two uncolored vertices. A crucial point about dCR is that the color used in each step may depend on colors used arbitrarily far into the past. More generally, to implement the first case of the If statement in the algorithm, we have to keep track of the "needed-color" by each monochromatic edge. This complicates the analysis greatly.

Our algorithm is enabled by the following simple observation: by coloring *two* vertices in each round, one can dispense with all the "needed-color" information.

ALGORITHM A

If there are $(k-3)$- or $(k-2)$-monochromatic edges
  then
    let $x < y$ be the smallest uncolored vertices in the lexicographically
      smallest such edge e;

color x Red; color y Blue

else

let $x < y$ be the smallest uncolored vertices;

color x Red; color y Blue

In particular, by "taking action" just slightly earlier than algorithm CR, algorithm A dispenses with the need to know the color used so far in e: coloring two of the remaining vertices of e in opposite colors guarantees that it is bichromatic.

Before proceeding to the analysis of the performance of A on $H(k,n,p)$, we wish to briefly compare CR and A with algorithms for random k-SAT suggested in [6, 7, 12]. All these algorithms for k-SAT set the value of one variable at a time, giving priority to variables that appear in clauses that are yet unsatisfied and have few remaining unset variables. While the algorithms differ in the exact rule for choosing which variable to set among those appearing in "short" clauses, their asymptotic performance is within a constant factor.

Similarly, both CR and A also give priority to vertices participating in short edges, where short now refers to an edge with many of its vertices already colored, all in the same color. Those edges are clearly the most dangerous, and it is thus quite natural to try to take care of them first. However, in contrast with a short clause, a short edge does not contain all the information necessary to take care of it: we need to know the color used in the remaining vertices of the original edge. Coloring two vertices in each round allows us to circumvent this problem and only affects the asymptotic performance up to a constant factor. Moreover, it's worth noting that algorithm A always produces an *equitable* 2-coloring, i.e. one where the two colors are used an equal number of times (up to parity).

2.2 Proof of Theorem 1

We will prove the statement in Theorem 1 for even values of n, with a slightly larger c' (namely, $c' = 1.01/50$). For odd n we introduce an extra (fake) vertex and consider $H(k,n+1,p)$, where $p\binom{n}{k} = c(2^k/k)n < c'(2^k/k)(n+1)$. The statement of Theorem 1 for odd n now follows by invoking our stronger statement for even n and ignoring all edges containing the extra vertex.

We need to show that w.h.p. in $H(k,n,p)$, the 2-coloring resulting from applying algorithm A contains no monochromatic edge. We will prove a slightly stronger claim.

Claim 1 *With probability $1-O(n^{-1/2})$ no edge becomes $(k-1)$-monochromatic during the algorithm's execution.*

To see why Claim 1 implies Theorem 1, note that if an edge has two of its vertices colored in the same round, then it never becomes monochromatic. Thus, to become monochromatic an edge must first become $(k-1)$-

monochromatic (and, in general, to become i-monochromatic an edge must first become $(i-1)$-monochromatic).

Proof of Claim 1. As often happens in the analysis of deterministic algorithms on random structures, it will be convenient to assume that the choice of the random hypergraph is made *in parallel* with its coloring, rather than assuming that a member of $H(k,n,p)$ is chosen before the execution of A begins.

For a family E_0 of distinct k-subsets of V, *exposing the edges from* E_0 amounts to deciding for each k-tuple $e \in E_0$, whether $e \in E(H(k,n,p))$ independently and with probability $p(n)$. Thus, if the family of all k-subsets of V is represented as a union of pairwise disjoint families $\{E_i\}$, exposing the families E_i in some order generates a random hypergraph $H(k,n,p)$.

Now let us describe how the edges of $H(k,n,p)$ are exposed as Algorithm A proceeds.

EXPOSURE PROCEDURE

1. For $0 \leq t \leq n/2 - 1$ repeat:

   Suppose we are in round t of Algorithm A, and are about to color vertex x in Red and vertex y in Blue. Let R_{t-1} be the set of vertices colored Red and B_{t-1} be the set of vertices colored Blue, in rounds $0, \ldots, t-1$.

   Expose all edges, having $k-4$ vertices in R_{t-1}, containing x and having three vertices outside $R_{t-1} \cup B_{t-1} \cup \{x, y\}$.

   Expose all edges, having $k-4$ vertices in B_{t-1}, containing y and having three vertices outside $R_{t-1} \cup B_{t-1} \cup \{x, y\}$.

2. After all vertices from V have been colored, expose any yet unexposed edges.

It is easy to see that each k-subset of V is exposed exactly once during the above exposure procedure. Hence its output is distributed according to $H(k,n,p)$. It is important to observe that in part 2 of the above exposure procedure all exposed edges have at most $k-4$ vertices of one color. In order for an edge of H to become $(k-1)$-monochromatic, it should first become $(k-3)$-monochromatic, and therefore it could only have been exposed during part 1 of the above exposure procedure. Therefore, to prove Claim 1, it is enough to restrict our attention to this first part, performed along with the execution of the algorithm.

For each edge $e \in E(H)$, exposed in round t of the algorithm, let e' be the triple of its uncolored vertices at the end of round t. We denote

$$F^{(t)} = \{e' \subset e : e \in E(H) \text{ and } e \text{ is exposed in round } t\} \ .$$

($F^{(t)}$ is a set, not a multiset, i.e. we treat multiple copies of a triple as one.) We will refer to triples from $F^{(t)}$ as t-triples and it will be notationally convenient to define $F^{(t)} = \emptyset$ for $t \geq n/2$.

For the purpose of the analysis, we group rounds into *phases*. We use t_i and $\hat{t_i}$ to denote the first and the last round of the ith phase, respectively, and the phase itself is defined as follows. The ith phase consists of the sequence of rounds $t_i, t_i+1, \ldots, \hat{t_i}$, if the number of $(k-3)$- and $(k-2)$-monochromatic edges is zero at the beginning of round t_i and round $\hat{t_i}+1$, but remains positive during the rounds t_i+1 through $\hat{t_i}$. In particular, at the beginning of a new phase there are no $(k-3)$- or $(k-2)$-monochromatic edges, and during a phase there is at least one such monochromatic edge. It will be notationally convenient to consider round $n/2-1$ as the beginning of a last, trivial phase. Notice that precisely $2t_i$ vertices are colored just before phase i starts and $2(\hat{t_i}+1)$ vertices are colored after phase i ends. For phase i, we denote

$$F_i = \bigcup_{r=t_i}^{\hat{t_i}} F^{(r)} \quad .$$

Since there are no $(k-3)$- or $(k-2)$-monochromatic edges right before phase i starts, we observe that if an edge e becomes $(k-1)$-monochromatic during phase i, then e must (a) have been exposed during phase i, and (b) if w, z are two of its vertices colored during phase i, F_i must contain two distinct triples corresponding to edges e_w, e_z (different from e) containing w, z, respectively. As there at most $n/2$ phases, to prove Claim 1 it suffices to prove the following lemma.

Lemma 1

$$\Pr[\ F_i \text{ contains } e_1, e_2, e_3 \text{ so that } e_1 \cap e_2 \neq \emptyset,\ e_1 \cap e_3 \neq \emptyset] = O(n^{-3/2}) \quad . \tag{1}$$

Proof. We first note that conditional on R_t, B_t, each triple from $V \setminus (R_t \cup B_t)$ appears in $F^{(t)}$ independently of all other such triples and with probability

$$q(t) := 1 - (1-p)^{2\binom{t}{k-4}} \quad .$$

In order to eliminate dependencies between the appearance of distinct triples in F_i, we will condition on R_{t_i} and B_{t_i}. As our argument will work for any given R_{t_i} and B_{t_i}, Lemma 1 will follow.

For $t_i \leq t \leq \hat{t_i}$, it will be convenient to define a superset $F_i^{(t)}$ of $F^{(t)}$ in which every triple in $V \setminus (R_{t_i} \cup B_{t_i})$ appears independently and with probability $q(t)$. To form $F_i^{(t)}$ we add to $F^{(t)}$ each triple in $V \setminus (R_{t_i} \cup B_{t_i})$ but not in $V \setminus (R_t \cup B_t)$ with probability $q(t)$, independently of all other such triples. Also, we will introduce an auxiliary set

$$F_i' = \bigcup_{r=t_i}^{t_i+\log^2 n-1} F_i^{(r)} \quad .$$

Note now that each triple from $V \setminus (R_{t_i} \cup B_{t_i})$ appears in F_i' independently and with probability

$$1 - \prod_{t=t_i}^{t_i + \log^2 n - 1} (1 - q(t)) \leq 1 - (1 - q(t_i + \log^2 n - 1))^{\log^2 n} \ ,$$

as $q(t)$ is increasing in t. Moreover, F_i is a subset of F_i' unless $\hat{t_i} > t_i + \log^2 n - 1$.

Letting

$$q_i := 1 - (1 - q(t_i + \log^2 n - 1))^{\log^2 n} \ ,$$

F_i' is a random set of triples from $V \setminus (R_{t_i} \cup B_{t_i})$, with each triple appearing independently and with probability *at most* q_i. Thus, recalling that F_i is a subset of F_i' unless $\hat{t_i} > t_i + \log^2 n - 1$, we see that the probability in (1) is bounded by the sum of

$$\Pr[\ F_i' \text{ contains } e_1, e_2, e_3 \text{ so that } e_1 \cap e_2 \neq \emptyset,\ e_1 \cap e_3 \neq \emptyset\ \ ] \tag{2}$$

and

$$\Pr[\hat{t_i} > t_i + \log^2 n - 1] \ . \tag{3}$$

Since $(1-a)^b \geq 1 - ab$ for all $a > 0$ and any nonnegative integer b, we have

$$\begin{aligned} q_i &\leq 2p \binom{t_i + \log^2 n - 1}{k-4} \log^2 n \\ &\leq 2 \frac{c2^k n}{k} \frac{(n-k)!k!}{n!} \frac{(t_i + \log^2 n)^{k-4}}{(k-4)!} \log^2 n \\ &\leq \frac{33ck^3 \log^2 n}{n^3} \left(\frac{2(t_i + \log^2 n)}{n} \right)^{k-4} \ . \end{aligned} \tag{4}$$

Now, as $k \geq 4$ is fixed, the probability in (2) is readily bounded from above by

$$\begin{aligned} \binom{n-2t_i}{3} \left(3 \binom{n - 2t_i - 1}{2} \right)^2 q_i^3 &= O((n-2t_i)^7 q_i^3) && (5) \\ &= O((1-\alpha)^7 \alpha^{3k-12} k^9 \log^6 n / n^2) && (6) \\ &= O(k^2 \log^6 n / n^2)) && (7) \\ &= n^{-2+o(1)} \ . \end{aligned}$$

In passing from (5) to (6) we took $2t_i = \alpha n$, while in passing from (6) to (7) we used that $(1-\alpha)^7 \alpha^{3k-12} = O(k^{-7})$.

For the probability in (3) we observe that in any round t with $t_i + 1 \leq t \leq \hat{t_i}$, at least one $(k-3)$- or $(k-2)$-monochromatic edge is 2-colored. Thus the number of such edges after round t is at most

$$\left| \bigcup_{r=t_i}^{t} F^{(r)} \right| - (t - t_i) \ ,$$

which must be positive for $t \le \hat{t_i}$. In particular, if $\hat{t_i} > t_i + \log^2 n$, then $|F_i'| \ge \log^2 n$. Note now that $|F_i'|$ is stochastically dominated by a binomial random variable Bin $\left(\binom{n-2t_i}{3}, q_i\right)$. Letting $2t_i/n = \alpha$, we get

$$\begin{aligned}
\binom{n-2t_i}{3} q_i &\le \frac{(n-2t_i)^3}{6} \frac{33ck^3 \log^2 n}{n^3} \left(\frac{2(t_i + \log^2 n)}{n}\right)^{k-4} \\
&\le \frac{17ck^3}{3}(1-\alpha)^3 \alpha^{k-4} \log^2 n \\
&\le \frac{17ck^3}{3} \left(1 - \frac{k-4}{k-1}\right)^3 \left(\frac{k-4}{k-1}\right)^{k-4} \log^2 n \\
&\le 0.9 \log^2 n \ ,
\end{aligned}$$

for $c \le 1.01/50$ and $k \ge 6$, and for $c \ge 1.01/10$ and $k \ge 40$; in the above derivation, we have also used the fact that $(1-\alpha)^3 \alpha^{k-4}$ is maximized when $\alpha = (k-4)/(k-1)$. Thus, by considering the Chernoff bound for the tail of the Binomial random variable, we see that the probability in (3) is at most $1/n^2$, concluding the proof of the lemma.

To conclude the proof of Theorem 1 we will present a deterministic algorithm which w.h.p. 2-colors $H(k, n, p)$ if $m = cn$, for $c < 1/6$. Postponing that proof for a moment, we observe that for $3 \le k \le 5$, $\frac{101}{5000} \frac{2^k}{k} < \frac{1}{6}$, which along with the bound for $k \ge 6$ above yields Theorem 1.

Let a component of a hypergraph be "bad" if it contains more than one cycle, or more than two edges sharing more than one vertex. Note that for all $k \ge 2$, if $c < \frac{1}{k(k-1)}$ then w.h.p. there are no bad components in $H(k, n, p)$ – this is the content of Theorem 3 in [13], and can also be verified by an easy first moment calculation. Our deterministic algorithm for $3 \le k \le 5$ is as follows: if H contains a bad component then exit, reporting failure. Otherwise, to color a component, repeatedly remove edges containing vertices of degree 1. Let $e_1, \ldots, e_q$ be the removed edges, in order of removal. Since $k \ge 3$ and the component is not bad, it is not hard to see that this procedure removes all the edges. Now, we add the edges back to H in reverse order, coloring vertices as follows: while adding back an edge $e_i \in E(H)$, if e_i contains two uncolored vertices, color them using distinct colors. Otherwise, take a vertex of current degree 1 in e_i and use it to make the edge bichromatic. By the ordering of the edges, one of the above two cases always happens. Finally, all uncolored vertices, which are exactly the isolated vertices of H, are colored arbitrarily.

3 Concluding remarks

- It is natural to wonder if Algorithm A in fact performs significantly better than what we have demonstrated. However, one can show that for larger values of c (e.g., $c = 1$), there exists a round t^* such that w.h.p. the number of $(k-2)$-monochromatic edges at the beginning of round t^* is greater than

$(n-2t^*)$. As a result, w.h.p. the graph induced by the uncolored vertices of these edges contains an odd cycle and hence the algorithm fails. Thus, our analysis of Algorithm A is tight up to the value of the constant c.

- Our algorithm A suggests the following algorithm for r-coloring $H(k,n,p)$ for any fixed $r \geq 2$ and $k \geq 3$.

ALGORITHM A_r

If there are $(k-r-1)$- or $(k-r)$-monochromatic edges
then
 let $x_1 < x_2 < \cdots < x_r$ be the smallest uncolored vertices in the lexicographically smallest such edge;
 color x_i with color i, for $i = 1, \ldots, r$
else
 let $x_1 < x_2 < \cdots < x_r$ be the smallest uncolored vertices;
 color x_i with color i, for $i = 1, \ldots, r$

Thus, Algorithm A is A_2. Again, the key property is that at the end of every round an equal number of vertices have been assigned each color. An analysis similar to that of A, shows that the above algorithm w.h.p. r-colors a random k-uniform hypergraph $H(k,n,p)$ if the edge probability $p = p(n)$ satisfies

$$p\binom{n}{k} = c\frac{r^k}{k}n \ ,$$

where $c < c^* = c^*(r)$. (For example, taking $c^*(r) = (r+1)!/(r+1)^{2(r+1)}$ suffices).

We note that the chromatic number of random sparse k-uniform hypergraphs $H(k,n,p)$ has been studied by Schmidt-Pruzan, Shamir, and Upfal in [18], and by Krivelevich and Sudakov in [14]. For example, it was shown in [18] that if

$$p\binom{n}{k} = \frac{(k-1)n}{Ck} \ ,$$

then $H(k,n,p)$ is 3-colorable w.h.p. for $C > e$. It is easy to see (by considering $r = 3$) that our bound improves upon the result in [18] by an exponential (in k) factor. This is actually the case for all $r \geq 3$.

- Finally, we note that using a non-rigorous technique of statistical physics, namely the replica method, in [16] it is suggested that the threshold for the satisfiability of random k-SAT formulas is at the number of clauses $m = c2^kn$ (in fact with $c = \ln 2$). We feel that improving asymptotically the easy upper bound or the existing lower bound for either the satisfiability problem or the 2-colorability problem would represent significant progress on this topic.

Acknowledgment. We would like to thank Van H. Vu for useful discussions and the anonymous referees for helpful suggestions. The last two authors

would like to thank the theory group at Microsoft Research for their visit to Microsoft during which much of this work was done.

References

[1] Alon, N., and Spencer, J. A note on coloring random k-sets. Unpublished manuscript.

[2] Beck, J. On 3-chromatic hypergraphs. *Discrete Math. 24*, 2 (1978), 127–137.

[3] Berge, C. *Hypergraphs.* North-Holland Publishing Co., Amsterdam, 1989. Combinatorics of finite sets, Translated from the French.

[4] Bernstein, F. Zur theorie der trigonometrische reihen. *Leipz. Ber. 60*, (1908), 325–328.

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

[6] Chao, M.-T., and Franco, J. Probabilistic analysis of a generalization of the unit-clause literal selection heuristics for the k-satisfiability problem. *Inform. Sci. 51*, 3 (1990), 289–314.

[7] Chvátal, V., and Reed, B. Mick gets some (the odds are on his side). In *33th Annual Symposium on Foundations of Computer Science (Pittsburgh, PA, 1992).* IEEE Comput. Soc. Press, Los Alamitos, CA, 1992, pp. 620–627.

[8] Erdős, P. On a combinatorial problem. *Nordisk Mat. Tidskr. 11* (1963), 5–10, 40.

[9] Erdős, P., and Lovász, L. Problems and results on 3-chromatic hypergraphs and some related questions. 609–627. Colloq. Math. Soc. János Bolyai, Vol. 10.

[10] Flajolet, P., Knuth, D. E., and Pittel, B. The first cycles in an evolving graph. *Discrete Math. 75*, 1-3 (1989), 167–215. Graph theory and combinatorics (Cambridge, 1988).

[11] Friedgut, E. Necessary and sufficient conditions for sharp thresholds of graph properties, and the k-SAT problem. *J. Amer. Math. Soc. 12* (1999), 1017–1054.

[12] Frieze, A. M. and Suen, S. Analysis of two simple heuristics on a random instance of k-SAT. *J. Algorithms 20*, 2 (1996), 312–355.

[13] Karoński, M. and Łuczak, T. Random hypergraphs. Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993), 283–293, Bolyai Soc. Math. Stud., 2, Janos Bolyai Math. Soc., Budapest, 1996.

[14] KRIVELEVICH, M. AND SUDAKOV, B. The chromatic numbers of random hypergraphs. *Random Struct. Alg. 12*, (1998), 381–403.

[15] LOVÁSZ, L. Coverings and coloring of hypergraphs. *Proceedings of the Fourth Southeastern Conference on Combinatorics, Graph Theory, and Computing* (Boca Raton, Florida, 1973), 3–12.

[16] MONASSON, R., AND ZECCHINA, R. Entropy of the K-satisfiability problem. *Phys. Rev. Lett. 76*, 21 (1996), 3881–3885.

[17] RADHAKRISHNAN, J., AND SRINIVASAN, A. Improved bounds and algorithms for hypergraph 2-coloring. *Random Structures Algorithms 16*, 1 (2000), 4–32.

[18] SCHMIDT-PRUZAN, J., AND SHAMIR, E., AND UPFAL, E. Random hypergraph coloring algorithms and the weak chromatic number. *Journal of Graph Theory 8*, (1985), 347–362.