How Many Random Edges Make a Dense Hypergraph Non-2-Colorable?*

Benny Sudakov, Jan Vondrák
Department of Mathematics, Princeton University, Princeton, New Jersey 08544; e-mail: bsudakov@math.princeton.edu; jvondrak@math.princeton.edu

Received 23 August 2005; accepted 20 August 2006
Published online 14 August 2007 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/rsa.25155

ABSTRACT: We study a model of random uniform hypergraphs, where a random instance is obtained by adding random edges to a large hypergraph of a given density. The research on this model for graphs has been started by Bohman et al. (Random Struct Algorithms 22 (2003) 33–42), and continued in (Bohman et al., Random Struct Algorithms 24 (2004) 105–117) and (Krivelevich et al., Random Struct Algorithms 29 (2006), 180–193). Here we obtain a tight bound on the number of random edges required to ensure non-2-colorability. We prove that for any k-uniform hypergraph with $\Omega(n^{k-\epsilon})$ edges, adding $\omega(n^{k\epsilon/2})$ random edges makes the hypergraph almost surely non-2-colorable. This is essentially tight, since there is a 2-colorable hypergraph with $\Omega(n^{k-\epsilon})$ edges which almost surely remains 2-colorable even after adding $o(n^{k\epsilon/2})$ random edges. © 2007 Wiley Periodicals, Inc. Random Struct. Alg., 32, 290–306, 2008

Keywords: random hypergraphs; 2-colorability; smoothed analysis

1. INTRODUCTION

Research on random graphs and hypergraphs has a long history with thousands of papers and two monographs by Bollobás [9] and by Janson et al. [15] devoted to the subject and its diverse applications. In the classical Erdős–Rényi model [14], a random graph is generated by starting from an empty graph and then adding a certain number of random edges. More recently, Bohman, Frieze, and Martin [7] considered a generalized model where one starts with a fixed graph $G = (V, E)$ and then inserts a collection R of additional random edges.

Correspondence to: J. Vondrák
*Supported by NSF career award DMS-0546523; NSF (DMS-0355497); USA-Israeli BSF grant; Alfred P. Sloan fellowship.

We denote the resulting random graph by $G + R$. The initial graph G can be regarded as given by an adversary, whereas the random perturbation R represents noise or uncertainty, independent of the initial choice. This scenario is analogous to the *smoothed analysis* of algorithms proposed by Spielman and Teng [19], where an algorithm is assumed to run on the worst-case input, modified by a small random perturbation.

Usually, one investigates *monotone properties* of random graphs or hypergraphs; i.e., properties which cannot be destroyed by adding more edges, like the property of containing a certain fixed subgraph. Given a monotone property $\mathcal{A}$ of graphs on n vertices, we can ask what are the parameters for which a random graph has property $\mathcal{A}$ almost surely, i.e. with probability tending to 1 as the number of vertices n tends to infinity. In our setting, we start with a fixed hypergraph H and inquire how many random edges R we have to add so that $H + R$ has property $\mathcal{A}$ almost surely. This question is too general to get concrete and meaningful results, valid for all hypergraphs H. Therefore, rather than considering a completely arbitrary H, we start with a hypergraph from a certain natural class. One such class of graphs was considered in [7], where the authors analyze the question of how many random edges need to be added to a graph G of minimal degree at least dn, $0 < d < 1$, so that the resulting graph $G + R$ is almost surely Hamiltonian. Further properties of random graphs in this model are explored in [8].

In [16], Krivelevich et al. considered a slightly more general setting, in which one performs a small random perturbation of a graph G with at least dn^2 edges. The authors obtained tight results for the appearance of a fixed subgraph and for certain Ramsey properties in this model. In the same paper, they also considered random formulas obtained by adding random k-clauses (disjunctions of k literals) to a fixed k-SAT formula. Krivelevich et al. proved that for any formula with at least $n^{k-\epsilon}$ k-clauses, adding $\omega(n^{k\epsilon})$ random clauses of size k makes the formula almost surely unsatisfiable. This is tight, since there is a k-SAT formula with $n^{k-\epsilon}$ clauses which almost surely remains satisfiable after adding $o(n^{k\epsilon})$ random clauses. A related question, which was raised in [16], is to find a threshold for non-2-colorability of a random hypergraph obtained by adding random edges to a large hypergraph of a given density.

For an integer $k \geq 2$, a *k-uniform hypergraph* is an ordered pair $H = (V, E)$, where V is a finite non-empty set, called the set of *vertices* and E is a family of distinct k-subsets of V, called the *edges* of H. A *2-coloring* of a hypergraph H is a partition of its vertex set V into two color classes so that no edge in E is monochromatic. A hypergraph which admits a 2-coloring is called *2-colorable*.

2-colorability is one of the fundamental properties of hypergraphs, which was first introduced and studied by Bernstein [6] in 1908 for infinite hypergraphs. 2-colorability in the finite setting, also known as "Property B" (a term coined by Erdős in reference to Bernstein), has been studied extensively in the last 40 years (see, e.g., [5, 10, 11, 13, 18]). While 2-colorability of graphs is well understood being equivalent to nonexistence of odd cycles, for k-uniform hypergraphs with $k \geq 3$ it is already NP-complete to decide whether a 2-coloring exists [17]. Consequently, there is no efficient characterization of 2-colorable hypergraphs. The problem of 2-colorability of random k-uniform hypergraphs for $k \geq 3$ was first studied by Alon and Spencer [4]. They proved that such hypergraphs with $m = (c2^k/k^2)n$ edges are almost surely 2-colorable. This bound was improved later by Achlioptas et al. [1]. Recently, the threshold for 2-colorability has been determined very precisely. In [2] it was proved that the number of edges for which a random k-uniform hypergraph becomes almost surely non-2-colorable is $(2^{k-1} \ln 2 - O(1))n$.

Interestingly, the threshold for non-2-colorability is roughly one half of the threshold for k-SAT. It has been shown in [3] that a formula with m random k-clauses becomes almost surely unsatisfiable for $m = (2^k \ln 2 - O(k))n$. The two problems seem to be intimately related and it is natural to ask what is their relationship in the case of a random perturbation of a fixed instance. Recall that from [16] we know that for any k-SAT formula with $n^{k-\epsilon}$ clauses, adding $\omega(n^{k\epsilon})$ random clauses makes it almost surely unsatisfiable. In fact, the same proof yields that for any k-uniform hypergraph H with $n^{k-\epsilon}$ edges, adding $\omega(n^{k\epsilon})$ random edges destroys 2-colorability almost surely. Nonetheless, it turns out that this is not the right answer. It is enough to use substantially fewer random edges to destroy 2-colorability: roughly a square root of the number of random clauses necessary to destroy satisfiability. The following is our main result.

Theorem 1.1. *Let $k, \ell \geq 2$, $\epsilon \geq 0$ be fixed and let H be a 2-colorable k-uniform hypergraph with $\Omega(n^{k-\epsilon})$ edges. Then the hypergraph H' obtained by adding to H a collection R of $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples is almost surely non-2-colorable.*

Observe that for $\epsilon \geq 2/\ell$, the result is easy. Regardless of the hypergraph H, it is well known that a collection of $\omega(n)$ random ℓ-tuples on n vertices is almost surely non-2-colorable. So we will be only interested in the case when $\epsilon < 2/\ell$. For such ϵ we obtain the following result, which shows that the assertion of Theorem 1.1 is essentially best possible.

Theorem 1.2. *For fixed $k, \ell \geq 2$ and $0 \leq \epsilon < 2/\ell$, there exists a 2-colorable k-uniform hypergraph H with $\Omega(n^{k-\epsilon})$ edges such that its union with a collection R of $o(n^{\ell\epsilon/2})$ random ℓ-tuples is almost surely 2-colorable.*

The rest of this article is organized as follows. In the next section we present an example of the hypergraph which proves Theorem 1.2. In Section 3, we discuss some natural difficulties in proving Theorem 1.1 and describe how to deal with them in the case of bipartite graphs. This result also serves as a basis for induction which we use in Section 4 to prove the general case of 2-colorable k-uniform hypergraphs.

Remark 1.3. We have two alternative ways of adding random edges. Either we can sample a random ℓ-tuple $|R|$ times, each time uniformly and independently from the set of all $\binom{n}{\ell}$ ℓ-tuples. Or we can pick each ℓ-tuple randomly and independently with probability $p = |R|/\binom{n}{\ell}$. Since 2-colorability is a monotone property, it follows, as in Bollobás [9], Theorem 2.2 and a similar remark in [16], that if the resulting hypergraph is almost surely non-2-colorable (2-colorable) in one model then this is true in the other model as well. This observation can sometimes simplify our calculations.

Notation. Let $H = (V, E)$ be a k-uniform hypergraph. In the following, we use the notions of *degree* and *neighborhood*, generalizing their usual meaning in graph theory. For a vertex $v \in V$, we define its degree $d(v)$ to be the number of edges of H that contain v. More generally, for a subset of vertices $A \subset V$, $|A| < k$, we define its degree $d(A) = |\{e \in E : A \subset e\}|$. For a $(k-1)$-tuple of vertices A, we define its *neighborhood* as $N(A) = \{w \in V \setminus A : A \cup \{w\} \in E\}$. Also, for a $(k-2)$-tuple of vertices A, we define its *link* as $\Gamma(A) = \{\{u, v\} \in V \setminus A : A \cup \{u, v\} \in E\}$.

Throughout the article we will systematically omit floor and ceiling signs for the sake of clarity of presentation. Also, we use the notations $a_n = \Theta(b_n)$, $a_n = O(b_n)$ or $a_n = \Omega(b_n)$ for $a_n, b_n > 0$ and $n \to \infty$ if there are absolute constants C_1 and C_2 such that $C_1 b_n <$

$a_n < C_2 b_n$, $a_n < C_2 b_n$ or $a_n > C_1 b_n$, respectively. The notation $a_n = o(b_n)$ means that $a_n/b_n \to 0$ as $n \to \infty$, and $a_n = \omega(b_n)$ means $a_n/b_n \to \infty$. The parameters k, ℓ, ϵ are considered constant.

2. THE LOWER BOUND

The following example proves Theorem 1.2 and shows that our main result is essentially best possible.

Construction. Partition the set of vertices $[n]$ into three disjoint subsets X, Y, Z where $|X| = |Y| = n^{1-\epsilon/2}$. Let H be a k-uniform hypergraph whose edge set consists of all k-tuples which have exactly one vertex in X, one vertex in Y and $k-2$ vertices in Z. By construction, the number of edges in H is $\Theta(n^{k-\epsilon})$ (Fig. 1).

Claim. Color all the vertices in X by color 1 and vertices in Y by color 2. Note that no matter how we assign colors to the remaining vertices, this gives a proper 2-coloring of H. Let R be a set of $o(n^{\ell\epsilon/2})$ random ℓ-tuples. Then almost surely we can 2-color Z so that none of the ℓ-tuples in R is monochromatic, i.e., there exists a proper 2-coloring of $H + R$.

To prove this claim we transform R into another random instance R' that contains only single vertices with a fixed *prescribed color* and edges of size two which must not be monochromatic. Following Remark 1.3 we can assume that R was obtained by choosing every ℓ-tuple in $[n]$ randomly and independently with probability $p = o(n^{\ell\epsilon/2-\ell})$. First note that almost surely there is no ℓ-tuple in R whose vertices are all in X or in Y. Indeed, since $|X| = |Y| = n^{1-\epsilon/2}$, the probability that there is such an ℓ-tuple is at most $2\binom{n^{1-\epsilon/2}}{\ell} p = o(1)$.

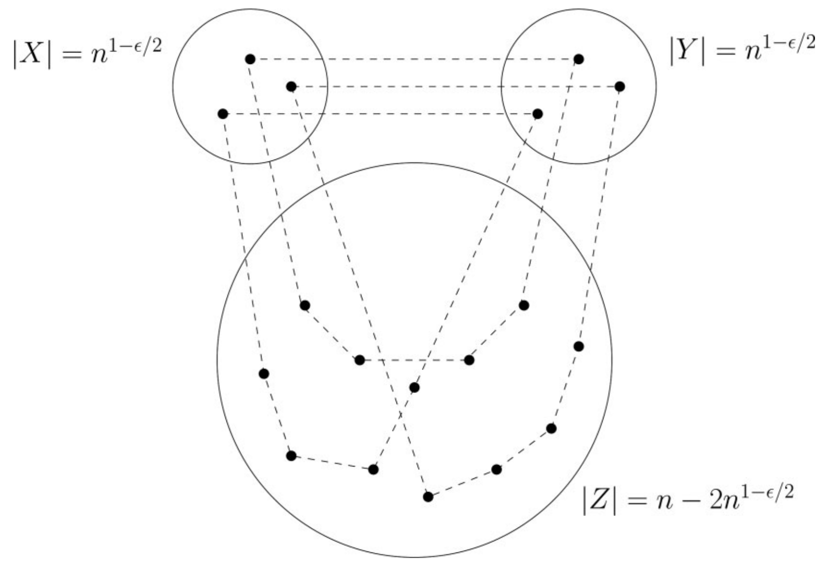

Fig. 1. Construction of the hypergraph H.

Also, every ℓ-tuple in R which has vertices in both X and Y is already 2-colored so we discard it.

For every $v \in Z$ we add it to R' with prescribed color 1 if there is a subset A of Y of size $\ell - 1$ such that $A \cup \{v\} \in R$. Since $\epsilon < 2/\ell \leq 1$, the probability of this event is

$$p_1 = \binom{|Y|}{\ell - 1} p = \binom{n^{1-\epsilon/2}}{\ell - 1} p \leq n^{(\ell-1)(1-\epsilon/2)} p = o(n^{-1+\epsilon/2}) = o(n^{-1/2}).$$

Similarly, if there is a subset B of X of size $\ell - 1$ such that $B \cup \{v\} \in R$ then we add v to R' with prescribed color 2. The probability p_2 of this event is also $o(n^{-1/2})$.

Fix an ordering $v_1 < v_2 < \dots$ of all vertices in Z. For every pair of vertices $u, w \in Z$ we add an edge $\{u, w\}$ to R' if there is an ℓ-tuple $L \in R$ such that the two smallest vertices in $L \cap Z$ are u and w. Since the number of such possible ℓ-tuples is at most $\binom{n}{\ell-2}$, and $\epsilon < 2/\ell$, the probability of this event is

$$p_3 \leq \binom{n}{\ell - 2} p = O(n^{\ell-2} p) = o(n^{\ell\epsilon/2-2}) = o(n^{-1}).$$

Also note that by definition all the above events are independent since they depend on disjoint sets of ℓ-tuples. By our construction, any 2-coloring of Z in which singletons in R' get prescribed colors and no 2-edge is monochromatic gives a proper 2-coloring of R. Therefore, to complete the proof of Theorem 1.2, it is enough to prove the following simple statement.

Lemma 2.1. *Let R' be a random instance which is obtained as follows. For $i = 1, 2$ we choose every vertex in $[n]$ with probability $p_i = o(n^{-1/2})$ (independently for $i = 1, 2$) and prescribe to it color i. In addition we choose every pair of vertices to be an edge in R' with probability $p_3 = o(n^{-1})$. Then almost surely there exists a 2-coloring of $[n]$ in which all singletons in R' get prescribed colors and no edge is monochromatic.*

Proof. Let G be the graph formed by edges from R'. The probability that there is a vertex with conflicting prescribed colors is $np_1p_2 = o(1)$. The probability that G contains a cycle is at most $\sum_{s=3}^{n} n^s p_3^s = O(n^3 p_3^3) = o(1)$. Finally the probability that there exists a path between two vertices with any prescribed color is also bounded by

$$\sum_{s=1}^{n} \binom{n}{2} (p_1 + p_2)^2 n^{s-1} p_3^s = o(n(p_1 + p_2)^2) = o(1).$$

Therefore almost surely no vertex gets prescribed conflicting colors, every connected component of G is a tree and contains at most one vertex with prescribed color. This immediately implies the assertion of the lemma, since every tree can be 2-colored, starting from the vertex with prescribed color (if any). ▪

3. BIPARTITE GRAPHS

Now let's turn to Theorem 1.1. First, consider the case of $k = \ell = 2$. Here, we claim that for any bipartite graph G with $\Omega(n^{2-\epsilon})$ edges, adding $\omega(n^\epsilon)$ random edges makes the graph almost surely non-bipartite. This will follow quite easily, since it turns out that almost surely we will insert an edge inside one part of a bipartite connected component of G, creating an odd cycle (see the proof of Proposition 3.1).

However, with the more general hypergraph case in mind, we are also interested in a scenario where random ℓ-tuples are added to a bipartite graph, and $\ell > 2$. Then we ask what is the probability that the resulting hypergraph is 2-colorable (i.e., no 2-edge and no ℓ-edge should be monochromatic). We prove the following special case of Theorem 1.1.

Proposition 3.1. *Let $\ell \geq 2$, $0 \leq \epsilon < 2/\ell$ and let G be a bipartite graph with $\Omega(n^{2-\epsilon})$ edges. Then the hypergraph H obtained by adding to G a collection R of $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples is almost surely non-2-colorable.*

Proof. Let G have $cn^{2-\epsilon}$ edges, $c > 0$ constant. Consider the connected components of G which are bipartite graphs on disjoint vertex sets (A_1, B_1), (A_2, B_2), ... (see Fig. 2). Denote $a_i = |A_i|$, $b_i = |B_i|$ and assume $a_i \geq b_i$. The number of edges in each component is at most $a_i b_i$. Since the total number of edges is $cn^{2-\epsilon}$, we have

$$\sum a_i^2 \geq \sum a_i b_i \geq cn^{2-\epsilon}.$$

Observe that for $\ell = 2$, the number of pairs of vertices inside the sets $\{A_i\}$ is $\sum \binom{a_i}{2} \geq \frac{1}{2}(cn^{2-\epsilon} - n) \geq c'n^{2-\epsilon}$, so a random edge lands inside one of these sets with probability at least $c'n^{-\epsilon}$. Consequently, the probability that none of the $\omega(n^{\epsilon})$ random edges ends up inside some A_i is at most $(1 - c'n^{-\epsilon})^{\omega(n^{\epsilon})} = o(1)$. Thus almost surely, $G + R$ contains an odd cycle.

On the other hand, in the general case we are adding $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples, which might never end up inside any vertex set A_i. The probability of hitting a specific A_i is $\binom{a_i}{\ell}/\binom{n}{\ell} = O(a_i^{\ell}/n^{\ell})$. For example, if G has n^{ϵ} components with $a_i = b_i = n^{1-\epsilon}$, then this probability is at most $O(\sum a_i^{\ell}/n^{\ell}) = O(n^{-(\ell-1)\epsilon})$. Hence we need $\omega(n^{(\ell-1)\epsilon})$ random ℓ-tuples to hit almost surely some A_i. This suggests a difficulty with the attempt to place a random ℓ-tuple in a set which is forced to be monochromatic by the original graph. We have to allow ourselves more freedom and consider sets which are monochromatic only under certain colorings.

More specifically, under any coloring, each of the sets A_i must be monochromatic and at least half of these sets must have the same color. We do not know a priori which of the sets A_i will share the same color, yet we can estimate the probability that *any* of these configurations

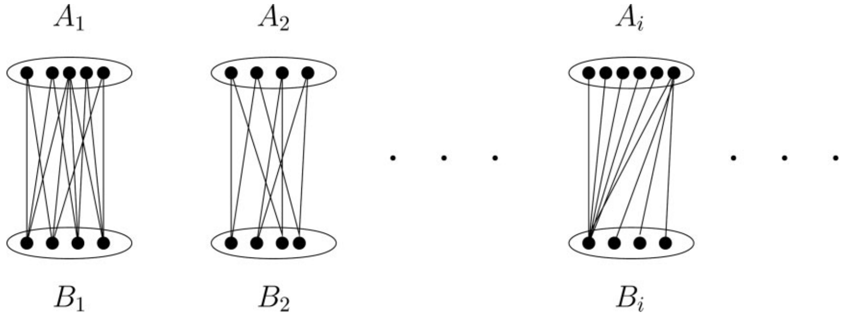

Fig. 2. Components of the bipartite graph G.

allows a feasible coloring together with the random ℓ-tuples. First, it is convenient to assume that the sets have roughly equal size, in which case we have the following claim. ▪

Lemma 3.2. *Suppose we have t disjoint subsets $A_1, \ldots, A_t$ of $[n]$ of size $\Theta(n^{1-\alpha})$. Let $\alpha \geq \epsilon/2$, $t = \Omega\big(n^{\frac{\ell}{\ell-1}(\alpha-\epsilon/2)}\big)$ and let R be a collection of $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples on $[n]$. Then the probability that R can be 2-colored in such a way that each A_i is monochromatic is at most $e^{-\omega(t)}$.*

Proof. Consider the 2^t possible colorings in which all A_i are monochromatic. For each such coloring there is a set of indices $I, |I| \geq t/2$ such that the sets $A_i, i \in I$ share the same color. Since A_i are disjoint we have $|\cup_{i\in I} A_i| \geq c_1 t n^{1-\alpha}$ for some $c_1 > 0$. The probability that one random ℓ-tuple falls inside this set is at least $\binom{c_1 tn^{1-\alpha}}{\ell}/\binom{n}{\ell} \geq c_2 (tn^{-\alpha})^\ell$ for some $c_2 > 0$. Hence

$$\Pr[\cup_{i\in I} A_i \text{ contains no } \ell\text{-tuple from } R] \leq (1-(c_2 tn^{-\alpha})^\ell)^{\omega(n^{\ell\epsilon/2})} \leq e^{-\omega(t^\ell n^{-\ell(\alpha-\epsilon/2)})} = e^{-\omega(t)}$$

where we used $t^\ell = t\cdot t^{\ell-1} = t\cdot\Omega(n^{\ell(\alpha-\epsilon/2)})$. Therefore, by the union bound over all choices of I, we get

$$\Pr[\exists I \text{ such that } \cup_{i\in I} A_i \text{ contains no } \ell\text{-tuple from } R] \leq 2^t e^{-\omega(t)} = e^{-\omega(t)}.$$

In particular, almost surely there is no 2-coloring of R in which all A_i are monochromatic. ▪

Now we can finish the proof of Proposition 3.1 for $\ell \geq 3$. Recall that G has $cn^{2-\epsilon}$ edges. Partition the components of G according to their size and let G_s contain all the components with $|A_i| \in [2^{s-1}, 2^s)$. If there is any A_i of size at least $n^{1-\epsilon/2}$, we are done immediately because one of the $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples almost surely ends up in A_i and this destroys 2-colorability. So we can assume that G_s is nonempty only for $s \leq \lfloor(1-\epsilon/2)\log_2 n\rfloor$. If we can choose a subgraph G_s with sufficiently many edges, then we can use Lemma 3.2 to finish the proof as follows.

Suppose there is $s \leq \lfloor(1-\epsilon/2)\log_2 n\rfloor$ such that G_s has $m_s \geq \frac{c}{4}2^{\frac{\ell-2}{\ell-1}s} n^{\frac{\ell}{\ell-1}(1-\epsilon/2)}$ edges. As each component of G_s has at most 2^{2s} edges, the number of components of G_s is $t_s \geq 2^{-2s} m_s = \Omega(2^{-\frac{\ell}{\ell-1}s} n^{\frac{\ell}{\ell-1}(1-\epsilon/2)})$. We set $2^s = n^{1-\alpha}$, $\alpha \geq \epsilon/2$, which means that $t_s = \Omega(n^{\frac{\ell}{\ell-1}(\alpha-\epsilon/2)})$. To summarize, we have $\Omega(n^{\frac{\ell}{\ell-1}(\alpha-\epsilon/2)})$ disjoint sets A_i of size $\Theta(n^{1-\alpha})$, each of which must be monochromatic under any feasible coloring. Thus we can apply Lemma 3.2 to conclude that for $H = G + R$, almost surely there is no feasible 2-coloring.

Finally, suppose that for any $s \leq \lfloor(1-\epsilon/2)\log_2 n\rfloor$, the number of edges in G_s is $m_s < \frac{c}{4}2^{\frac{\ell-2}{\ell-1}s} n^{\frac{\ell}{\ell-1}(1-\epsilon/2)}$ and G_s is empty for $s > \lfloor(1-\epsilon/2)\log_2 n\rfloor$. Then the total number of edges is

$$\sum_{s=1}^{\lfloor(1-\epsilon/2)\log_2 n\rfloor} m_s < \frac{c}{4}\sum_{s=1}^{\lfloor(1-\epsilon/2)\log_2 n\rfloor} 2^{\frac{\ell-2}{\ell-1}s} n^{\frac{\ell}{\ell-1}(1-\epsilon/2)}$$
$$< \frac{c}{4}\cdot\frac{n^{\frac{\ell-2}{\ell-1}(1-\epsilon/2)}}{1-2^{-\frac{\ell-2}{\ell-1}}}\cdot n^{\frac{\ell}{\ell-1}(1-\epsilon/2)} < cn^{2-\epsilon}$$

(in the last inequality, we used $\ell \geq 3$). This contradicts our assumption that G has $cn^{2-\epsilon}$ edges. ▪

4. PROOF OF THEOREM 1.1

In this section we deal with the general case of a 2-colorable k-uniform hypergraph H, to which we add a collection of random ℓ-tuples R. Our goal is to prove the main theorem which asserts that if H has $\Omega(n^{k-\epsilon})$ edges then adding to it $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples makes it almost surely non-2-colorable. The proof will proceed by induction on k. The base case when $k = 2$ follows from Proposition 3.1, so we can assume that $k \geq 3$ and that the result holds for $k - 1$. Since we have $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples available, we can divide them into a constant number of batches, where each batch still has $\omega(n^{\ell\epsilon/2})$ ℓ-tuples. We will use a separate batch for each step of the induction. We write $R = R_1 \cup R_2 \cup \ldots \cup R_k$ where $|R_j| = \omega(n^{\ell\epsilon/2})$ for each j.

We proceed in a series of lemmas which allow us to make simplifying assumptions, and eventually finish the proof of the theorem. The high-level structure of the proof is as follows.

1. If H_k contains $\Omega(n^{k-\epsilon})$ edges through $(k-1)$-tuples of degree greater than $n^{1-\epsilon/2}$, we can prove by induction that $H_k + R$ is almost surely non-2-colorable. Lemma 4.1 takes care of this case.
2. If H_k contains $\Omega(n^{k-\epsilon})$ edges through $(k-2)$-tuples of degree greater than $n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$, we can also prove by induction that $H_k + R$ is a.s. non-2-colorable. This is proved in Lemma 4.4. Also, a variant of this lemma can be used to finish the proof for $\ell = 2$ and all $k \geq 3$.
3. If neither of the first two cases apply, we can "clean up" our hypergraph (Lemma 4.5 and Lemma 4.6) to obtain a near-regular hypergraph H_α. The hypergraph provided by these lemmas satisfies the conditions of Lemma 4.9 which proves directly that $H_\alpha + R$ is almost surely non-2-colorable.

Lemma 4.1. *Let $k \geq 3$, $\ell \geq 2$ and let H_k be a k-uniform hypergraph on n vertices with $c_1 n^{k-\epsilon}$ edges. Consider all $(k-1)$-tuples $A \subset V(H_k)$ with degree greater than $n^{1-\epsilon/2}$. If there are at least $\frac{c_1}{4}n^{k-1-\epsilon}$ such $(k-1)$-tuples in H_k then H_k+R is almost surely non-2-colorable.*

Proof. For each $(k-1)$-tuple A of degree $> n^{1-\epsilon/2}$, the neighborhood $N(A)$ contains $\Omega(n^{\ell-\ell\epsilon/2})$ distinct ℓ-tuples. Therefore a random ℓ-tuple lands inside $N(A)$ with probability $\Omega(n^{-\ell\epsilon/2})$. Consequently, the probability that none of $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples from R_k ends up inside $N(A)$ is $(1-\Omega(n^{-\ell\epsilon/2}))^{\omega(n^{\ell\epsilon/2})} = o(1)$. If we have $t \geq \frac{c_1}{4}n^{k-1-\epsilon}$ such $(k-1)$-tuples, then the expected number of them, whose neighborhood does *not* contain any ℓ-tuple in R_k, is $o(t)$. Therefore, by Markov's inequality, we get almost surely at least $\frac{t}{2} \geq \frac{c_1}{8}n^{k-1-\epsilon}$ $(k-1)$-tuples with an ℓ-edge in their neighborhood. Denote by H_{k-1} the $(k-1)$-uniform hypergraph formed by these $(k-1)$-tuples.

By induction, we know that $H_{k-1} + R_1 + \ldots + R_{k-1}$ is almost surely non-2-colorable. Hence for every 2-coloring respecting $R_1 \cup \ldots \cup R_{k-1}$, there is a monochromatic $(k-1)$-tuple A in H_{k-1}. Without loss of generality assume that all vertices in A are colored by color 1. By definition, the neighborhood $N(A)$ contains an ℓ-edge $L \in R_k$. Either L is monochromatic, or one of its vertices x is colored by 1 as well. But then $A \cup \{x\}$ is a monochromatic edge of H_k. This implies that there is no feasible 2-coloring for $H_k + R_1 + \ldots + R_k$. ∎

Thus we only need to treat the case where there are at most $\frac{c_1}{4}n^{k-1-\epsilon}$ $(k-1)$-tuples with degree greater than $n^{1-\epsilon/2}$, therefore at most $\frac{c_1}{4}n^{k-\epsilon}$ edges through such $(k-1)$-tuples.

We will get rid of these high degrees by removing their edges and making all degrees of $(k-1)$-tuples at most $n^{1-\epsilon/2}$. This would also imply a bound of $n^{2-\epsilon/2}$ on the degrees of $(k-2)$-tuples, etc. However, in the following we show that for $(k-2)$-tuples we can assume an even stronger bound. More specifically, we prove that if we have many edges through $(k-2)$-tuples whose degree is at least $n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$, then we can proceed by induction. For this purpose, we first show the following.

Lemma 4.2. *Let $\ell \geq 2$ and let G be a graph on n vertices with $n^{2-\delta}$ edges. Then G contains $\frac{1}{6}n^{1-\delta}$ disjoint subsets of vertices $F_1, F_2, \ldots$ such that in every F_j, each vertex $v \in F_j$ can be assigned a set of neighbors $X(v) \subseteq N(v)$ so that $x_v = |X(v)| \geq \frac{1}{2}n^{1-\delta}$, $X(v) \cap X(w) = \emptyset$ for $v, w \in F_j$, and*

$$\sum_{v \in F_j} x_v^\ell \geq \frac{n^{\ell-(\ell-1)\delta}}{2^{\ell+2}}.$$

We construct the sets $F_1, F_2, \ldots$ by a simple algorithm. First, we show how to construct one such set for $\ell = 2$.

Claim 4.3. *Assume that G has at most n vertices and at least $\frac{1}{2}n^{2-\delta}$ edges. Then G contains a set of vertices F such that*

1. *We can assign disjoint sets of neighbors $X(v) \subseteq N(v)$ to vertices $v \in F$, so that $x_v = |X(v)| \geq \frac{1}{2}n^{1-\delta}$ and*
$$\sum_{v \in F} x_v^2 \geq \frac{1}{16}n^{2-\delta}.$$
2. *The number of edges in G incident to F is at most $3n$.*

Proof. We find F by the following procedure. We start with $F = \emptyset$ and add vertices one by one. We denote by $N(F) = \cup_{v \in F} N(v)$ the vertices connected by an edge to F and by $W = V \setminus (F \cup N(F))$ the remaining vertices. For a set S, we denote by $d_S(v)$ the number of neighbors that vertex v has in S. Intuitively, we choose a vertex from W which has a large neighborhood but not overlapping very much with $N(F)$. Note that no vertex $v \in W$ has neighbors in F, otherwise v would be in $N(F)$ itself. Hence, by our construction, F is an independent set.

We repeat the following, until $\sum_{u \in F} x_u^2 \geq \frac{1}{16}n^{2-\delta}$:

1. Find a vertex $u \in W$ maximizing
$$z_u = d_W^2(u) - \frac{1}{2}\sum_{v \in N(u) \cap W} d_W(v) - \frac{1}{4}n^{1-\delta}d_{N(F)}(u).$$
2. Set $X(u) = N(u) \cap W$, $x_u = |X(u)| = d_W(u)$ and include u in F.
3. Update $W = V \setminus (F \cup N(F))$.

We claim that as long as $\sum_{u \in F} x_u^2 < \frac{1}{16}n^{2-\delta}$, we can always find a vertex with $z_u = \Omega(n^{2-2\delta}) \geq 0$. Assuming this has been true up to a certain point, we have been choosing vertices with $z_u \geq 0$ and therefore $\sum_{v \in N(u) \cap W} d_W(v) \leq 2d_W^2(u) = 2x_u^2$ for each vertex when it was chosen. By including u in F, we increase the number of edges incident to $N(F)$ by $\sum_{v\in N(u)\cap W} d_W(v) \le 2x_u^2$. Therefore, as long as $\sum_{u\in F} x_u^2 < \frac{1}{16}n^{2-\delta}$, there are at most $\frac{1}{8}n^{2-\delta}$ edges incident to $N(F)$, and at least $\frac{3}{8}n^{2-\delta}$ edges disjoint from $N(F)$. These edges cannot touch F either (or else they would touch $N(F)$), so they are in the subgraph induced by W.

Now consider the choice of the next vertex u. Summing up z_u over all available vertices, we get

$$\begin{aligned}\sum_{u\in W} z_u &= \sum_{u\in W} d_W^2(u) - \frac{1}{2}\sum_{u\in W}\sum_{v\in N(u)\cap W} d_W(v) - \frac{1}{4}n^{1-\delta}\sum_{u\in W} d_{N(F)}(u)\\ &= \sum_{u\in W} d_W^2(u) - \frac{1}{2}\sum_{v\in W} d_W^2(v) - \frac{1}{4}n^{1-\delta}e(N(F),W)\\ &= \frac{1}{2}\sum_{u\in W} d_W^2(u) - \frac{1}{4}n^{1-\delta}e(N(F),W)\end{aligned}$$

where $e(N(F),W)$ denotes the number of edges between $N(F)$ and W. As we mentioned earlier, the number of edges induced by W is at least $\frac{3}{8}n^{2-\delta}$ which means that $\sum_{v\in W} d_W(v) \ge \frac{3}{4}n^{2-\delta}$. By Cauchy-Schwartz,

$$\sum_{u\in W} d_W^2(u) \ge \frac{1}{|W|}\left(\sum_{u\in W} d_W(u)\right)^2 \ge \frac{1}{n}\left(\frac{3}{4}n^{2-\delta}\right)^2 = \frac{9}{16}n^{3-2\delta}.$$

Also, the number of edges incident to $N(F)$ is bounded by $\frac{1}{8}n^{2-\delta}$. So we get

$$\sum_{u\in W} z_u \ge \frac{1}{2}\sum_{u\in W} d_W^2(u) - \frac{1}{4}n^{1-\delta}e(N(F),W) \ge \frac{9}{32}n^{3-2\delta} - \frac{1}{4}n^{1-\delta}\cdot\frac{1}{8}n^{2-2\delta} = \frac{1}{4}n^{3-2\delta}.$$

Thus, there must be a vertex $u \in W$ such that $z_u \ge \frac{1}{4}n^{2-2\delta}$. Consequently, we also have $x_u \ge \sqrt{z_u} \ge \frac{1}{2}n^{1-\delta}$. At the point we stop, we have $\sum_{u\in F} x_u^2 \ge \frac{1}{16}n^{2-\delta}$, as required.

Finally, we verify the number of edges incident to F. We distinguish two kinds of such edges. When we include a vertex u in F, call the edges connecting u to $N(F)$ "red", and call the other edges connecting it to W "blue". There can be at most n blue edges, because their endpoints form the disjoint sets $X(u)$. The number of red edges can be bounded in the following way. Whenever we include a vertex u in F, we have $z_u \ge d_W^2(u) - \frac{1}{4}n^{1-\delta}d_{N(F)}(u) \ge 0$. Therefore, the number of red edges contributed by vertex u at the moment when it is included in F is $d_{N(F)}(u) \le 4d_W^2(u)/n^{1-\delta} = 4x_u^2/n^{1-\delta}$. Let v^* be the last vertex which we added to F. By our construction, we have that $\sum_{u\in F\setminus\{v^*\}} x_u^2 \le \frac{1}{16}n^{2-\delta}$. Therefore the total number of red edges contributed by the vertices in $F \setminus \{v^*\}$ is

$$\sum_{u\in F\setminus\{v^*\}} d_{N(F)}(u) \le \sum_{u\in F\setminus\{v^*\}} \frac{4x_u^2}{n^{1-\delta}} < \frac{4}{n^{1-\delta}}\cdot\frac{1}{16}n^{2-\delta} = \frac{n}{4}.$$

The last vertex v^* can possibly contribute at most n red edges. So the total number of red and blue edges is at most $n + n/4 + n < 3n$. ∎

Proof of Lemma 4.2. Given the above claim, the proof of the lemma follows easily. We start with a graph $G_1 = G$. We iterate the construction of one set F, for $j = 1, 2, \dots, \frac{1}{6}n^{1-\delta}$. We apply the claim to the graph G_j and find a set F_j as required. Then we remove F_j from the graph, to obtain G_{j+1}. Since F_j is incident to at most $3n$ edges, we can iterate up to $\frac{1}{6}n^{1-\delta}$ times, and G_{j+1} still contains at least $\frac{1}{2}n^{2-\delta}$ edges. Each set F_j satisfies $\sum_{u\in F_j} x_u^2 \geq \frac{1}{16}n^{2-\delta}$. Since $x_u \geq \frac{1}{2}n^{1-\delta}$, for $\ell \geq 2$ this implies

$$\sum_{u\in F_j} x_u^\ell \geq \left(\frac{1}{2}n^{1-\delta}\right)^{\ell-2} \sum_{u\in F_j} x_u^2 \geq \frac{n^{\ell-(\ell-1)\delta}}{2^{\ell+2}}$$

■

Lemma 4.4. *Let $k \geq 3$, $\ell \geq 2$ and let H_k be a k-uniform hypergraph on n vertices with $c_1 n^{k-\epsilon}$ edges. Consider $(k-2)$-tuples of degree at least $n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$. If there are at least $\frac{c_1}{4}n^{k-\epsilon}$ edges through such $(k-2)$-tuples then $H_k + R$ is almost surely non-2-colorable.*

Proof. Consider a $(k-2)$-tuple A of degree $n^{2-\delta}$, where $\delta \leq \frac{\ell}{2(\ell-1)}\epsilon$. The link of A in H_k is a graph $\Gamma(A)$ with $n^{2-\delta}$ edges. By Lemma 4.2, we find $\frac{1}{6}n^{1-\delta}$ disjoint sets F_j such that the vertices in each F_j have disjoint sets of neighbors $X(v)$ in $\Gamma(A)$, with sizes $x_v = |X(v)|$ satisfying $\sum_{v\in F_j} x_v^\ell \geq n^{\ell-(\ell-1)\delta}/2^{\ell+2}$. We repeat this construction for each $(k-2)$-tuple of degree $D \geq n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$. Note that for each such $(k-2)$-tuple of degree D, we construct $D/(6n)$ sets as above. The sum of degrees of such $(k-2)$-tuples is at least the total number of edges through them, which is by assumption at least $\frac{c_1}{4}n^{k-\epsilon}$. Therefore, we obtain at least $\frac{c_1}{24}n^{k-1-\epsilon}$ sets F_j in total.

Now consider a set F_j chosen for a $(k-2)$-tuple A. Call it *good* if after adding the random ℓ-tuples in R_k, there is at least one vertex in F_j whose neighborhood in $\Gamma(A)$ contains a random ℓ-tuple. If this is not the case, call it *bad*. We estimate the probability that F_j is bad. By Lemma 4.2, the total number of ℓ-tuples inside the sets $X(v)$ for $v \in F_j$ is

$$\sum_{v\in F_j} \binom{x_v}{\ell} = \Omega\left(\sum_{v\in F_j} \frac{x_v^\ell}{\ell!}\right) = \Omega\left(\frac{n^{\ell-(\ell-1)\delta}}{2^{\ell+2}\ell!}\right) = \Omega(n^{\ell-\ell\epsilon/2}),$$

where we used that $\delta \leq \frac{\ell}{2(\ell-1)}\epsilon$ and ℓ is a constant. Thus the probability that a random ℓ-tuple falls inside $X(v)$ for some $v \in F_j$ is $\sum_{v\in F_j} \binom{x_v}{\ell}/\binom{n}{\ell} = \Omega(n^{-\ell\epsilon/2})$. After adding the entire batch of random ℓ-tuples R_k,

$$\Pr[F_j \text{ is bad}] = (1 - \Omega(n^{-\ell\epsilon/2}))^{\omega(n^{\ell\epsilon/2})} = o(1).$$

Consequently, the expected fraction of bad F_j's is $o(1)$. By Markov's inequality, this fraction is almost surely at most one half, which means that at least $\frac{c_1}{48}n^{k-1-\epsilon}$ sets F_j have a vertex $v \in F_j$ whose neighborhood contains some ℓ-tuple from R_k. By the construction of the sets F_j, for each one we have a set A of size $k-2$ which together with v forms a $(k-1)$-tuple $B = A \cup \{v\}$. Since the F_j's for a given $(k-2)$-tuple A are disjoint, we obtain distinct pairs (A, v) which correspond to distinct $(k-1)$-tuples with a marked vertex v. We could obtain the same $(k-1)$-tuple $B = A \cup \{v\}$ in $k-1$ different ways, but in any case we have at least $\frac{c_1}{48(k-1)}n^{k-1-\epsilon}$ $(k-1)$-tuples such that in H_k, the neighborhood of each of them contains an ℓ-tuple from R_k. Call the hypergraph of these $(k-1)$-tuples H_{k-1}.

By the induction hypothesis, $H_{k-1} + R_1 + \ldots + R_{k-1}$ is almost surely non-2-colorable. Therefore, for any 2-coloring which respects the ℓ-edges from $R_1 + \ldots + R_{k-1}$, there must be a monochromatic $(k-1)$-edge B in H_{k-1}. However, since there is an ℓ-edge from R_k in the neighborhood of B, one of its vertices should have the same color as B. This forms a monochromatic edge in H_k so there is no feasible 2-coloring for $H_k + R_1 + \ldots + R_k$. ∎

Remark. Lemma 4.4 assumes that there are $\Omega(n^{k-\epsilon})$ edges through $(k-2)$-tuples of degree at least $n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$. However, one can easily check that constant factors are not significant in the proof. In particular, if H_k is a k-uniform hypergraph on n vertices with $c_1 n^{k-\epsilon}$ edges such that there are at least $\frac{c_1}{4} n^{k-\epsilon}$ edges through $(k-2)$-tuples of degree at least $\frac{1}{4} c_1 n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$ then $H_k + R$ is almost surely non-2-colorable. Observe that in the case of $\ell = 2$, there are always at least $\frac{1}{4} c_1 n^{k-\epsilon}$ edges through $(k-2)$-tuples of degree at least $\frac{1}{4} c_1 n^{2-\frac{\ell}{2(\ell-1)}\epsilon} = \frac{1}{4} c_1 n^{2-\epsilon}$ (the remaining $(k-2)$-tuples can contribute at most $\binom{n}{k-2} \frac{1}{4} c_1 n^{2-\epsilon} \leq \frac{1}{4} c_1 n^{k-\epsilon}$ edges). Therefore in the case of $\ell = 2$ we can already conclude that $H_k + R$ is almost surely non-2-colorable.

In the following, we can assume that $\ell \geq 3$ and at most $\frac{c_1}{4} n^{k-\epsilon}$ edges go through $(k-2)$-tuples of degree greater than $n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$. Recall that from before, we can also assume that at most $\frac{c_1}{4} n^{k-\epsilon}$ edges go through $(k-1)$-tuples of degree greater than $n^{1-\epsilon/2}$. In the following step, we remove these edges so that the degrees in the hypergraph are bounded. We also make the hypergraph "k-partite" as described below.

Lemma 4.5. *Let $k, \ell \geq 3$ and let $H_k = (V, E)$ be a k-uniform hypergraph with $c_1 n^{k-\epsilon}$ edges, such that at most $\frac{c_1}{4} n^{k-\epsilon}$ edges go through $(k-1)$-tuples of degree $\geq n^{1-\epsilon/2}$ and at most $\frac{c_1}{4} n^{k-\epsilon}$ edges go through $(k-2)$-tuples of degree $\geq n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$. Then H_k contains a subhypergraph H'_k with the following properties*

1. *H'_k is k-partite, i.e. V can be partitioned into $V_1 \cup V_2 \cup \ldots \cup V_k$ so that every edge of H'_k intersects each V_i in one vertex.*
2. *Every vertex has degree at most $n^{k-1-\frac{\ell}{2(\ell-1)}\epsilon}$.*
3. *The degree of every $(k-1)$-tuple is at most $n^{1-\epsilon/2}$.*
4. *The number of edges in H'_k is at least $c_2 n^{k-\epsilon}$, $c_2 = \frac{k!}{2k^k} c_1$.*

Proof. First, remove all edges through $(k-1)$-tuples of degree greater than $n^{1-\epsilon/2}$ and through $(k-2)$-tuples of degree greater than $n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$. We get a hypergraph such that the degrees of all $(k-1)$-tuples are at most $n^{1-\epsilon/2}$ and the degrees of all $(k-2)$-tuples are at most $n^{2-\frac{\ell}{2(\ell-1)}\epsilon}$ Consequently, the degree of every vertex is at most $n^{k-3} \cdot n^{2-\frac{\ell}{2(\ell-1)}\epsilon} = n^{k-1-\frac{\ell}{2(\ell-1)}\epsilon}$. The number of remaining edges is at least $\frac{1}{2} c_1 n^{k-\epsilon}$.

Next, we use a well-known fact, proved by Erdős and Kleitman [12]. Every k-uniform hypergraph with m edges contains a k-partite subhypergraph with at least $\frac{k!}{k^k} m$ edges. (This can be achieved for example by taking a random partition $V = V_1 \cup V_2 \cup \ldots \cup V_k$ and computing the expected number of edges which intersect each V_i exactly once.) Let H'_k denote such a k-partite subhypergraph of H_k. Its number of edges is at least $c_2 n^{k-\epsilon}$ where $c_2 = \frac{k!}{2k^k} c_1$. ∎

Before the last part of the proof, we make further restrictions on the degree bounds and structure of our hypergraph, by finding a subhypergraph H_α with roughly regular $(k-1)$-degrees and sufficiently many edges. The number of edges that we can guarantee here is no longer a constant fraction of $n^{k-\epsilon}$. The statement of Lemma 4.6 may appear technical but it is exactly what we need for our final construction which finishes the proof (Lemma 4.9).

Lemma 4.6. *Let $k, \ell \geq 3$ and let H'_k be a k-uniform k-partite hypergraph on vertices $V_1 \cup V_2 \cup \ldots \cup V_k$ with $c_2 n^{k-\epsilon}$ edges, where the degrees of $(k-1)$-tuples are bounded by $n^{1-\epsilon/2}$. Then H'_k contains a subhypergraph H_α, $\alpha \geq \epsilon/2$, such that*

1. *The degree of every $(k-1)$-tuple in $V_1 \times V_2 \times \ldots \times V_{k-1}$ is either 0 or between $n^{1-\alpha}$ and $2n^{1-\alpha}$.*
2. *For some constant $c_3 = c_3(k, \ell, c_2)$, the number of edges in H_α is at least*

$$c_3\left(n^{k-\epsilon-\frac{\epsilon-\alpha}{\ell-1}} + n^{k-\epsilon-\frac{\ell-2}{\ell-1}(\alpha-\epsilon/2)}\right).$$

Proof. Consider all $(k-1)$-tuples in $V_1 \times V_2 \times \ldots \times V_{k-1}$ whose degree in H'_k is less than $\frac{1}{2}c_2 n^{1-\epsilon}$. Delete all the edges through such $(k-1)$-tuples, which is at most $\binom{n}{k-1}\frac{1}{2}c_2 n^{1-\epsilon} \leq \frac{1}{2}c_2 n^{k-\epsilon}$ edges in total. We still have at least $\frac{1}{2}c_2 n^{k-\epsilon}$ edges left. Now the degree of every $(k-1)$-tuple in $V_1 \times V_2 \times \ldots \times V_{k-1}$ is either 0 or between $\frac{1}{2}c_2 n^{1-\epsilon}$ and $n^{1-\epsilon/2}$.

We use an elementary counting argument to find the subhypergraph H_α as required. Let $n^{1-\alpha} = 2^s$ and partition $V_1 \times V_2 \times \ldots \times V_{k-1}$ into groups of $(k-1)$-tuples with degrees in intervals $[2^s, 2^{s+1})$, with s ranging between $s_1 = \log_2(\frac{1}{2}c_2 n^{1-\epsilon})$ and $s_2 = \log_2(n^{1-\epsilon/2})$. Let m_s denote the number of edges through $(k-1)$-tuples with degrees between $2^s = n^{1-\alpha}$ and $2^{s+1} = 2n^{1-\alpha}$. We prove the lemma with $c_3 = \frac{1}{16}c_2 \cdot \min\{\frac{c_2^{1/(\ell-1)}}{\ell-1}, 1\}$. Assume for the sake of contradiction that $m_s < c_3(n^{k-\epsilon-\frac{\epsilon-\alpha}{\ell-1}} + n^{k-\epsilon-\frac{\ell-2}{\ell-1}(\alpha-\epsilon/2)})$ for every s, i.e.

$$m_s < \frac{1}{16}c_2 n^{k-\epsilon}\left(\frac{c_2^{1/(\ell-1)}}{\ell-1}n^{-\frac{\epsilon-\alpha}{\ell-1}} + n^{-\frac{\ell-2}{\ell-1}(\alpha-\epsilon/2)}\right)$$

$$= \frac{1}{16}c_2 n^{k-\epsilon}\left(\frac{c_2^{1/(\ell-1)}n^{\frac{1-\epsilon}{\ell-1}}}{\ell-1}2^{-\frac{s}{\ell-1}} + \frac{2^{\frac{\ell-2}{\ell-1}s}}{n^{\frac{\ell-2}{\ell-1}(1-\epsilon/2)}}\right). \quad (1)$$

Taking a sum from $s = s_1$ to s_2, we get

$$\sum_{s=s_1}^{s_2} 2^{-\frac{s}{\ell-1}} \leq \frac{2^{-\frac{s_1}{\ell-1}}}{1-2^{-\frac{1}{\ell-1}}} = \frac{\left(\frac{1}{2}c_2 n^{1-\epsilon}\right)^{-\frac{1}{\ell-1}}}{1-2^{-\frac{1}{\ell-1}}} \leq \frac{4(\ell-1)}{c_2^{1/(\ell-1)}}n^{-\frac{1-\epsilon}{\ell-1}}$$

and

$$\sum_{i=s_1}^{s_2} 2^{\frac{\ell-2}{\ell-1}s} \leq \frac{2^{\frac{\ell-2}{\ell-1}s_2}}{1-2^{-\frac{\ell-2}{\ell-1}}} = \frac{n^{\frac{\ell-2}{\ell-1}(1-\epsilon/2)}}{1-2^{-\frac{\ell-2}{\ell-1}}} \leq 4n^{\frac{\ell-2}{\ell-1}(1-\epsilon/2)}.$$

Substituting into (1), we see that then the total number of edges would be $\sum_{s=s_1}^{s_2} m_s < \frac{1}{2}c_2 n^{k-\epsilon}$ which is a contradiction. ∎

Note that in this lemma, we lose more than a constant fraction of edges. However, from now on, we do not use induction anymore and will prove directly that $H_\alpha + R$ is almost surely non-2-colorable. We will proceed in $t = c_3 \ell^{-k} n^{\frac{\ell}{\ell-1}(\alpha-\epsilon/2)}$ stages. For each stage, we allocate a certain number of random ℓ-tuples. Namely, we set again $R = R_1 \cup R_2 \cup \ldots \cup R_k$, $|R_j| = \omega(n^{\ell\epsilon/2})$. Furthermore, we divide each R_j for $j \leq k-1$ into t parts $R_{j,1}, \ldots, R_{j,t}$ so that

$$|R_{j,i}| = \omega\left(\frac{n^{\ell\epsilon/2}}{t}\right) = \omega(n^{\ell\epsilon/2-\frac{\ell}{\ell-1}(\alpha-\epsilon/2)}).$$

The random set $R_{j,i}$ will be used for the j-th "level" of the i-th stage. The following lemma describes one stage of the construction. Finally, R_k will be used in the last step of the proof.

Lemma 4.7. *Let $k, \ell \geq 3$ and let H_α be a k-uniform k-partite hypergraph where the degree of every $(k-1)$-tuple in $V_1 \times V_2 \times \ldots \times V_{k-1}$ is either zero or is in the interval $[n^{1-\alpha}, 2n^{1-\alpha}]$, and the number of edges in H_α is at least*

$$c_3 n^{k-\epsilon-\frac{\ell-2}{\ell-1}(\alpha-\epsilon/2)}.$$

Then after adding sets of random edges $R_{1,i} + R_{2,i} + \ldots + R_{k-1,i}$ where $|R_{j,i}| = \omega(n^{\ell\epsilon/2-\frac{\ell}{\ell-1}(\alpha-\epsilon/2)})$, there exists almost surely a family of $q = \ell^{k-2}$ sets $S_1, \ldots, S_q \subseteq V_k$, $n^{1-\alpha} \leq S_i \leq 2n^{1-\alpha}$, such that for every feasible 2-coloring of $H_\alpha + R_{1,i} + \ldots + R_{k-1,i}$, at least one S_i is monochromatic.

Proof. We are going to construct an ℓ-ary tree T of depth $k-1$. We denote vertices on the j-th level by $v_{a_1 a_2 \ldots a_{j-1}}$ where $a_i \in \{1, 2, \ldots, \ell\}$. T is rooted at a vertex in V_1 and the j-th level is contained in V_j. We construct T in such a way that the vertices along every path which starts at the root and has length $k-1$ form a $(k-1)$-tuple with degree $\Theta(n^{1-\alpha})$ in H_α. The neighborhoods of all branches of length $k-1$ will be our sets S_i (not necessarily disjoint). In addition, the set of ℓ children of every node on each level $j \leq k-2$, like $\{v_{a_1 a_2 \ldots a_{j-1} 1}, v_{a_1 a_2 \ldots a_{j-1} 2}, \ldots, v_{a_1 a_2 \ldots a_{j-1} \ell}\}$, will form an edge of $R_{j,i}$ (Fig. 3).

Assuming the existence of such a tree, consider any 2-coloring of $H_\alpha + R_{1,i} + \ldots + R_{k-1,i}$. Since the children of each vertex on level $j < k-1$ form an ℓ-edge in $R_{j,i}$, every vertex has children of both colors. In particular, there is always one child with the same color as its parent. Therefore, starting from the root, we can always find a monochromatic branch A of length $k-1$. Since all the extensions of this branch to edges of H_α must be 2-colored, all the vertices in $S_i = N(A)$ must have the same color.

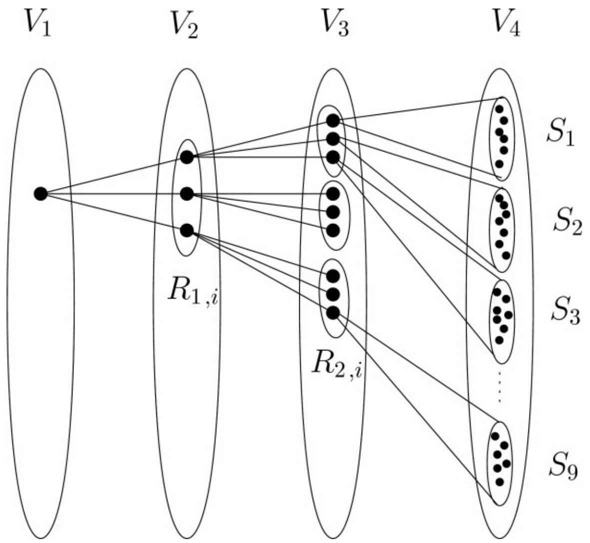

Fig. 3. Construction of the tree T, for $k = 4$ and $\ell = 3$. Branches of the tree form active $(k-1)$-tuples, with neighborhoods S_i. Each set of children on level $j+1$ forms an edge of $R_{j,i}$.

We grow the tree level by level, maintaining the property that all branches have sufficiently many extensions to edges of H_α. More precisely, we call an r-tuple in $V_1 \times \ldots \times V_r$ *active* if its degree is at least

$$\Delta_r = \frac{c_3}{2^r} n^{k-r-\epsilon-\frac{\ell-2}{\ell-1}(\alpha-\epsilon/2)}.$$

Claim 4.8. *Every active r-tuple A, $r \leq k-2$, can be extended to at least*

$$d_r = \frac{\Delta_r}{4n^{k-r-1-\alpha}} = \frac{c_3}{2^{r+2}} n^{1-\epsilon/2+\frac{1}{\ell-1}(\alpha-\epsilon/2)}$$

active $(r+1)$*-tuples* $A \cup \{x\}, x \in V_{r+1}$.

Proof. Suppose that fewer than d_r extensions of A are active. Since the degrees of $(k-1)$-tuples are at most $2n^{1-\alpha}$, we get that any $(r+1)$-tuple has degree at most $2n^{k-r-1-\alpha}$. Therefore the number of edges through all active extensions of A is smaller than $d_r \cdot 2n^{k-r-1-\alpha} = \frac{1}{2}\Delta_r$. We also have inactive extensions of A which have degrees less than Δ_{r+1}. The total number of edges through these extensions of A is smaller than $n\Delta_{r+1} = \frac{1}{2}\Delta_r$. But the total number of edges through A is at least Δ_r. This contradiction proves the claim. ∎

We start our construction from an active vertex $v \in V_1$. Since H_α has at least $n\Delta_1$ edges, such a vertex must exist. By our claim, v can be extended to at least d_1 active pairs $\{v, x\}, x \in W_2 \subset V_2$. Consider this set of d_1 vertices W_2. The probability that a random ℓ-tuple falls inside W_2 is $\binom{d_1}{\ell}/\binom{n}{\ell} = \Omega(n^{-\ell\epsilon/2+\frac{\ell}{\ell-1}(\alpha-\epsilon/2)})$. Now we use $\omega(n^{\ell\epsilon/2-\frac{\ell}{\ell-1}(\alpha-\epsilon/2)})$ random ℓ-tuples from $R_{1,i}$ that we allocated for the first level of this construction. This means that almost surely, we get an ℓ-edge $\{v_1, \ldots, v_\ell\} \in R_{1,i}$ such that $\{v, v_i\}$ is an active pair for each $i = 1, 2, \ldots, \ell$.

We continue growing the tree, using the random ℓ-tuples of $R_{j,i}$ on level j. Since we have ensured that each path from the root to the level j forms an active j-tuple, it has at least d_j extensions to an active $(j+1)$-tuple. Again, the probability that a random ℓ-tuple hits the extension vertices $W_{j+1} \subset V_{j+1}$ for a given path is $\binom{d_j}{\ell}/\binom{n}{\ell} = \Omega\big(n^{-\ell\epsilon/2+\frac{\ell}{\ell-1}(\alpha-\epsilon/2)}\big)$. Almost surely, one of the ℓ-tuples in $R_{j,i}$ will hit these extension vertices and we can extend this path to ℓ children on level $j+1$. The number of paths from the root to level j is bounded by $\ell^{j-1} \leq \ell^k$ which is a constant, so in fact we will almost surely succeed to build the entire level.

In this way, we almost surely build the tree all the way to level $k-1$. Every path from the root to one of the leaves forms an active $(k-1)$-tuple and has degree $\in [n^{1-\alpha}, 2n^{1-\alpha}]$. Define $S_1, S_2, \ldots, S_q$ to be the neighborhoods of all these $q = \ell^{k-2}$ paths. By construction, for any feasible 2-coloring of $H_\alpha + R_{1,i} + \ldots + R_{k-1,i}$, one of these paths is monochromatic which implies that the corresponding set S_i is monochromatic as well. ∎

Lemma 4.9. *Let $k, \ell \geq 3$ and let H_α be a k-uniform k-partite hypergraph where the degree of every vertex is at most $n^{k-1-\frac{\ell}{2(\ell-1)}\epsilon}$, the degree of every $(k-1)$-tuple in $V_1 \times V_2 \times \ldots \times V_{k-1}$ is either zero or is in the interval $[n^{1-\alpha}, 2n^{1-\alpha}]$, and the number of edges in H_α is at least*

$$c_3 n^{k-\epsilon-\frac{\epsilon-\alpha}{\ell-1}} + c_3 n^{k-\epsilon-\frac{\ell-2}{\ell-1}(\alpha-\epsilon/2)}.$$

Let R be a set of $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples. Then almost surely, $H_\alpha + R$ is not 2-colorable.

Proof. We apply Lemma 4.7 repeatedly in $t = c_3 \ell^{-k} n^{\frac{\ell}{\ell-1}(\alpha-\epsilon/2)}$ stages. We partition the set R as described before into $\cup_{j=1}^{k-1} \cup_{i=1}^{t} R_{j,i} \cup R_k$, where $|R_{j,i}| = \omega(n^{\ell\epsilon/2-\frac{\ell}{\ell-1}(\alpha-\epsilon/2)})$ and $|R_k| = \omega(n^{\ell\epsilon/2})$. In each stage i, we almost surely obtain $q = \ell^{k-2}$ sets $S_{i,1}, \ldots, S_{i,q}$, $n^{1-\alpha} \leq |S_{i,j}| \leq 2n^{1-\alpha}$ such that for any 2-coloring of the hypergraph $H_\alpha + R_{1,i} + \ldots + R_{k-1,i}$, one of these sets must be monochromatic. If this happens, we call such a stage "successful". After each successful stage, we remove all edges of H_α incident with any of the sets $S_{i,1}, \ldots, S_{i,q}$. Since degrees are bounded by $n^{k-1-\frac{\ell}{2(\ell-1)}\epsilon}$ and we repeat $t = c_3 \ell^{-k} n^{\frac{\ell}{\ell-1}(\alpha-\epsilon/2)}$ times, the total number of edges we remove is at most

$$\sum_{i=1}^{t} \sum_{j=1}^{q} |S_{i,j}| n^{k-1-\frac{\ell}{2(\ell-1)}\epsilon} \leq tq \cdot 2n^{1-\alpha} \cdot n^{k-1-\frac{\ell}{2(\ell-1)}\epsilon} = 2c_3 \ell^{-2} n^{k-\epsilon-\frac{\epsilon-\alpha}{\ell-1}} \leq c_3 n^{k-\epsilon-\frac{\epsilon-\alpha}{\ell-1}}.$$

In particular, before every stage we still have at least $c_3 n^{k-\epsilon-\frac{\ell-2}{\ell-1}(\alpha-\epsilon/2)}$ edges available, so we can use Lemma 4.7. Since the expected number of stages that are not successful is $o(t)$, by Markov's inequality, we almost surely get at least $t/2$ successful stages. Eventually, we obtain sets $S_{i,j}$ for $1 \leq i \leq t/2$ and $1 \leq j \leq q$ such that

- For $i_1 \neq i_2$ and any j_1, j_2, $S_{i_1,j_1} \cap S_{i_2,j_2} = \emptyset$.
- For any i and any 2-coloring of $H_\alpha + R_{1,i} + R_{2,i} + \ldots + R_{k-1,i}$, there is j_i such that S_{i,j_i} is monochromatic.

Finally, we add once again a collection R_k of $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples. We do not know a priori which selection of sets $S_{i,j}$ will be monochromatic but there is only exponential number of choices $q^{t/2} = e^{O(t)}$. For any specific choice of sets to be monochromatic, Lemma 3.2 says that the probability that after adding $\omega(n^{\ell\epsilon/2})$ random ℓ-tuples, there is a feasible 2-coloring keeping these sets monochromatic, is $e^{-\omega(t)}$. By the union bound, the probability that there exists a proper 2-coloring of $H_\alpha + \cup_{i,j} R_{j,i} + R_k$ is at most $q^{t/2} e^{-\omega(t)} = o(1)$. ▪

This completes the proof of Theorem 1.1, as outlined at the beginning of Section 4.

ACKNOWLEDGMENTS

The first author would like to thank Michael Krivelevich for helpful and stimulating discussions. We also thank both referees for helpful remarks. In particular, these remarks helped us notice a mistake in the original proof of Lemma 4.2.

REFERENCES

[1] D. Achlioptas, J. H. Kim, M. Krivelevich, and P. Tetali, Two-coloring random hypergraphs, Random Struct Algorithms 20 (2002), 249–259.

[2] D. Achlioptas and C. Moore, On the 2-colorability of random hypergraphs, In: Randomization and approximation techniques in computer science, Lecture Notes Comput Sci 2483, Springer, Berlin, 2002, pp. 78–90.

[3] D. Achlioptas and Y. Peres, The threshold for random k-SAT is $2^k \log 2 - O(k)$, J AMS 17 (2004), 947–973.

[4] N. Alon and J. Spencer, A note on coloring random k-sets, unpublished work.

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

[6] F. Bernstein, Zur Theorie der trigonometrischen Reihe, Leipz Ber 60 (1908), 325–338.

[7] T. Bohman, A. Frieze, and R. Martin, How many random edges make a dense graph Hamiltonian? Random Struct Algorithms 22 (2003), 33–42.

[8] T. Bohman, A. Frieze, M. Krivelevich, and R. Martin, Adding random edges to dense graphs, Random Struct Algorithms 24 (2004), 105–117.

[9] B. Bollobás, Random graphs, Cambridge Studies in Advanced Mathematics, 2nd ed., Vol. 73, Cambridge University Press, Cambridge, 2001.

[10] P. Erdös, On a combinatorial problem, Nordisk Mat Tidskr 11 (1963), 5–10.

[11] P. Erdös, On a combinatorial problem. II, Acta Math Acad Sci Hungar 15 (1964), 445–447.

[12] P. Erdös and D. Kleitman, On coloring graphs to maximize the proportion of multicolored k-edges, J Combinatorial Theory 5 (1968), 164–169.

[13] P. Erdös and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, In: Infinite and finite sets Vol. II, Colloq Math Soc Janos Bolyai, Vol. 10, North-Holland, Amsterdam, 1975, pp. 609–627.

[14] P. Erdős and A. Rényi, On the evolution of random graphs, Publ Math Inst Hungar Acad Sci 5 (1960), 17–61.

[15] S. Janson, T. Luczak, and A. Ruciński, Random graphs, Wiley, New York, 2000.

[16] M. Krivelevich, B. Sudakov, and P. Tetali, On smoothed analysis in dense graphs and formulas, Random Struct Algorithms 29 (2006), 180–193.

[17] L. Lovász, Coverings and colorings of hypergraphs, In: Proc. 4th Southeastern Conf. on Combinatorics, Graph Theory and Computing, Utilitas Math., Boca Raton, FL, 1973, pp. 3–12.

[18] J. Radhakrishnan and A. Srinivasan, Improved bounds and algorithms for hypergraph 2-coloring, Random Struct Algorithms 16 (2000), 4–32.

[19] D. A. Spielman and S.-H. Teng, Why the simplex algorithm usually takes polynomial time, In: Proceedings of the 33rd Annual ACM Symposium on Theory of Computing, Heraklion, Greece, 2001, pp. 296–305.