Improved bounds and algorithms for hypergraph two-coloring*

Jaikumar Radhakrishnan† Aravind Srinivasan‡

Abstract

We show that for all large n, every n-uniform hypergraph with at most $0.7\sqrt{n/\ln n} \times 2^n$ edges can be 2-colored. This makes progress on a problem of Erdős (1963), improving the previous-best bound of $n^{1/3-o(1)} \times 2^n$ due to Beck (1978). We further generalize this to a "local" version, improving on one of the first applications of the Lovász Local Lemma. We also present fast randomized algorithms that output a proper 2-coloring with high probability for n-uniform hypergraphs with at most $0.7\sqrt{n/\ln n} \times 2^n$ edges, for all large n. In addition, we derandomize and parallelize these algorithms, to derive NC^1 versions of these results.

1 Introduction

A *hypergraph* $H = (V, E)$ consists of a set V and a collection E of subsets of V. The elements of V and E are respectively called *vertices* and *edges*, and we only consider finite hypergraphs here. Hypergraph coloring is a generalization of graph coloring: H is said to be c-colorable iff there is a function $V \to \{1, 2, \ldots, c\}$ such that no edge is monochromatic. In contrast with graphs, deciding if a given hypergraph is 2-colorable is NP-complete, even if all edges have cardinality at most 3 (Lovász [20], Garey & Johnson [16]). Hypergraph 2-colorability is a central problem in combinatorics that has been studied since the early part of this century. It has also been studied by computer scientists due to its connections to the graph coloring and satisfiability problems. In this work, we make progress on an extremal problem of Erdős on 2-colorable hypergraphs, improving on a result of Beck from 1978; we further generalize this to a "local" version, improving on one of the first applications of the Lovász Local Lemma [12]. Furthermore, our first result translates to fast sequential and parallel (deterministic) algorithms.

History. The property of hypergraph 2-colorability, also called *Property B*, has been studied for long (Bernstein [8], Miller [23]: see Jensen & Toft [17]). Much work has been done on proving hypergraph families 2-colorable and on the corresponding algorithmic questions [1, 2, 6, 7, 12, 21, 22, 24, 26, 29]. Inspired in part by recent work on approximate graph coloring via semidefinite

*Part of this work was done when the authors attended the *Workshop on Randomized Algorithms* preceding the Annual Conference on Foundations of Software Technology and Theoretical Computer Science, 1997. A preliminary version of this research appears in the *Proc. IEEE Symposium on Foundations of Computer Science*, 1998.

†School of Technology and Computer Science, Tata Institute for Fundamental Research, Mumbai 400005, India. jaikumar@tcs.tifr.res.in

‡Current address: Room 2C-302A, Bell Laboratories, Lucent Technologies, 700 Mountain Avenue, Murray Hill, NJ 07974-0636, USA. This work was done: (i) at the School of Computing, National University of Singapore, Singapore 119260; and (ii) while visiting DIMACS (Center for Discrete Mathematics and Theoretical Computer Science).
DIMACS is an NSF Science and Technology Center, funded under contract STC-91-19999; and also receives support from the New Jersey Commission on Science and Technology. srin@research.bell-labs.com.

programming (Karger, Motwani & Sudan [18]), Alon, Kelsen, Mahajan & Ramesh [3], Chen & Frieze [9], and Krivelevich & Sudakov [19] have provided approximation algorithms for coloring 2-colorable hypergraphs. They present polynomial-time algorithms to c-color a given 2-colorable hypergraph $H = (V, E)$, where c is a function of $|V|$ and $\max_{f \in E} |f|$. There is also a natural maximization version of 2-coloring: color V with two colors such that a maximum possible number of edges are non-monochromatic. Approximation algorithms with a performance ratio of $0.724\cdots$ for this problem, have been provided by Andersson & Engebretsen [5].

We now present the setting of our first main result. $H = (V, E)$ is called an *ℓ-uniform* hypergraph if each edge is of cardinality ℓ. In an approach that is now oft-used in the context of, e.g., the maximum satisfiability problem, Erdős showed in 1963 that any n-uniform hypergraph with less than 2^{n-1} edges is 2-colorable: color the vertices Red and Blue uniformly at random and independently, and observe that the expected number of monochromatic edges is smaller than 1 [10]. This prompted one of his extremal problems: what is the least $m(n)$ such that there is an n-uniform hypergraph with $m(n)$ edges that is *not* 2-colorable? (This is problem 15.1 in [17].) Clearly, his above result shows that $m(n) \geq 2^{n-1}$. In yet another approach that is now much used in computer science, Erdős then used the probabilistic method to construct a "random n-uniform hypergraph" for which no 2-coloring exists: he showed that $m(n) < n^2 2^{n+1}$ [11]. See [17] for bounds on $m(n)$ for small n, as well as recurrence relations for $m(n)$.

It was conjectured by Erdős and Lovász in their seminal paper [12] that $m(n)$ may be $\Theta(n2^n)$. An elegant result of Beck showed that $m(n) > n^{1/3 - o(1)} 2^n$; i.e., for a certain function $g(n)$ that tends to 0 as n increases, Beck showed (using a randomized algorithm) that any n-uniform hypergraph with $n^{1/3 - g(n)} 2^n$ edges is 2-colorable [6]. (Of course, in such positive results, the n-uniformity condition can be weakened to each edge having *at least* n vertices: just restrict each edge to an arbitrary n-element subset of it.) A simpler, probabilistic and algorithmic, proof of Beck's result was presented by Spencer [26]. Recall from above that $m(n) = O(n^2 2^n)$. To quote Spencer from the second edition of [27],

> It may appear then that the bounds on $m(n)$ are close together. But from a probabilistic viewpoint a factor of 2^{n-1} may be considered a unit. We could rewrite the problem as follows. Given a family F let X denote the number of monochromatic sets under a random coloring. What is the maximal $k = k(n)$ so that, if F is a family of n-sets with $E(X) \leq k$, then $\Pr[X = 0] > 0$? In this formulation $cn^{1/3} < k(n) < cn^2$ and the problem clearly deserves more attention.

1.1 Contributions of this work

(i) Improved lower bounds for $m(n)$. By building on the Beck-Spencer approach, we improve on Beck's result to prove that $m(n) = \Omega(\sqrt{n/\ln n} \times 2^n)$. We show that if $H = (V, E)$ is an n-uniform hypergraph with at most $(1/10)\sqrt{n/\ln n} \times 2^n$ edges, then H is 2-colorable; for sufficiently large n, this bound can be improved to $0.7\sqrt{n/\ln n} \times 2^n$. We, in fact, present fast randomized algorithms that output a proper 2-coloring with high probability for such hypergraphs. We also derandomize and parallelize these algorithms, to derive NC^1 versions of these results. See Theorems 2.1 and 3.1.

(ii) 2-coloring hypergraphs with small overlap. Next, we generalize result (i) to a "local" version, using the Lovász Local Lemma. A useful parameter of H is its *overlap* D, defined to be the maximum number of edges, including itself, that any edge of H intersects. One of the first and major applications of the powerful Lovász Local Lemma [12]—abbreviated LLL here—was to show that any n-uniform hypergraph with $D \leq 2^{n-1}/e$ is 2-colorable. (Here, as usual, e denotes the base of the natural logarithm.) The Local Lemma was also applied to this parameterization based on D to show, e.g., that for any $d \geq 9$, any d-uniform hypergraph in which each vertex appears in at most d edges, is 2-colorable [12]; see [2, 29, 22] for further improvements. The "$D \leq 2^{n-1}/e$" result was major progress on the extremal question: what is the least $D^* = D^*(n)$ for which there exists an n-uniform hypergraph with overlap at most D^* that is *not* 2-colorable? The result of [11] described previously, shows that $D^*(n) = \mathrm{O}(n^2 2^n)$. By applying Theorem 4.1 to our approach for proving result (i), we show in Theorem 4.2 that $D^*(n) > 0.17\sqrt{n/\ln n}2^n$ for sufficiently large n. This improves on the above-seen $\Omega(2^n)$ lower bound of [12]. In other words, we show that for sufficiently large n, any n-uniform hypergraph with overlap D at most $0.17\sqrt{n/\ln n}2^n$ is 2-colorable. Modulo constant factors, this is easily seen to be a generalization of our result (i), since $|E| \geq D$ trivially.

(iii) Hypergraphs with "small" intersections. What could be a possible avenue for improving our result (i)? One answer is to start by considering some restricted but interesting family $\mathcal{F}$ of uniform hypergraphs. In Section 5, we build on a lead suggested by the work of [12] and present a candidate family $\mathcal{F}$. We show that the upper bound $\mathrm{O}(n^2 2^n)$ of [11] holds even for $m(n)$ restricted to this family. We then show that for all $\epsilon > 0$, $m(n) \geq \Omega(n^{1-\epsilon}2^n)$ for this family, by analyzing a modification of our algorithm for result (i).

At a high level, our main contribution is the following. An important branch of the basic probabilistic method is the *method of alteration*: one starts with an appropriate random construction, which, however, may (or will) contain some "blemishes". To correct the blemishes, we first argue that we do not expect too many blemishes, and then proceed to make a (hopefully) small alteration to correct these blemishes. See [4, 27] for several concrete instances of this methodology. In "correcting the blemish", two popular approaches are to proceed deterministically, or to alter several components of the current structure independently, with low probability. The above-mentioned result of Beck is essentially an example of the latter idea [6]. More precisely, one starts with a random coloring, and then independently flips the colors of vertices lying in monochromatic edges, with a small probability. Our main idea is to *slow down* this recoloring process: in the random recoloring process, we recolor the vertices lying in monochromatic edges in *random order*, processing these vertices one-by-one. The probabilistic recoloring of a vertex will take place only if a certain condition necessitates it: the advantage is that the effect of previously-processed vertices may have hopefully falsified this condition. The reader is referred to Section 2 for a precise description and analysis. Can such "lazy alteration" be applied to other probabilistic (alteration) arguments?

The rest of this paper is organized as follows. Section 2 presents the main result, and a derandomized parallel version is shown in Section 3. Sections 4 and 5 study hypergraphs with small overlap and small intersections, respectively.

2 Slow recoloring

We now prove our first main result—result (i) of the introduction. Edges of H will be denoted by f, f', h, h' etc.

For a coloring $\chi : V \to \{\text{Red}, \text{ Blue}\}$, let

$$\mathcal{M}(\chi) = \{f \in E : f \text{ is monochromatic in } \chi\},$$

and for $v \in V$, let

$$\mathcal{M}(v, \chi) = \{f \in \mathcal{M}(\chi) : v \in f\}.$$

For $S \subseteq V$, let $R(S, \chi)$ denote the event 'S is completely red in χ', and $B(S, \chi)$ the event 'S is completely blue in χ'. Sometimes, instead of $R(S, \chi)$ and $B(S, \chi)$ we write 'S is red in χ' and 'S is blue in χ' respectively.

The Algorithm.

Phase 1. Generate a random coloring $\chi_0 : V \to \{\text{Red}, \text{ Blue}\}$ by choosing $\chi_0(v)$ to be Red or Blue with probability 1/2, independently for each vertex $v \in V$.

Phase 2. In this phase, we flip the colors of some of the vertices to make edges in $\mathcal{M}(\chi_0)$ non-monochromatic, and hope that this does not make the already non-monochromatic edges monochromatic. In the method of [26], colors of all vertices that belonged to at least one monochromatic set were flipped (independently, with a certain probability q) *simultaneously*. In our proof we will not recolor all vertices at the same time. Instead, each vertex will be assigned a *delay*, which will be real number between 0 and 1, and the vertices will be processed in the order of their delays.

Formally, we pick the delay function $\mathsf{delay} : V \to [0, 1]$ by choosing $\mathsf{delay}(v)$ uniformly at random in $[0, 1]$ and independently for each $v \in V$. (Note that with probability 1, no two vertices are assigned the same delay.) Next, pick $b : V \to \{0, 1\}$ by choosing $b(v) = 1$ with probability p, and $b(v) = 0$ with probability $1-p$, independently for each $v \in V$. (Appropriate values for p will be presented later.) Using delay and b, we will recolor in $|V|$ steps as follows.[1] Let $v_1, v_2, \ldots$ be the vertices of H written in the order of their delays (v_1 has the smallest delay).

Step 1. If $\mathcal{M}(v_1, \chi_0) \neq \emptyset$ and $b(v_1) = 1$, then flip the color of v_1. Let the resulting coloring be χ_1.

Step 2. If some edge in $\mathcal{M}(v_2, \chi_0)$ continues to be monochromatic in χ_1 and $b(v_2) = 1$, then flip the color of v_2. Let the resulting coloring be χ_2.

$$\vdots$$

Step i. If some edge in $\mathcal{M}(v_i, \chi_0)$ continues to be monochromatic in χ_{i-1} and $b(v_i) = 1$, then flip the color of v_i. Let the resulting coloring be χ_i.

Let χ^* be the coloring obtained after all vertices have been considered.

This *slow recoloring* is the key to our improvement: one expects that a vertex that gets a "large" delay will have a "low" probability of having to be recolored due to the effect of the vertices with smaller delays. Intuitively, this helps increase the probability that an edge that was non-monochromatic at some point, remains so.

[1] The previous version of this algorithm used delays in the (discrete) range $1, \ldots, r$, where r was about $\ln n$. Joel Spencer showed that choosing real-valued delays from the range $[0, 1]$ considerably simplifies the calculations.

Simplification. Ravi Boppana (personal communication) has shown that the second phase of the algorithm can be considerably simplified, as follows. We pick delays and let $v_1, v_2, \ldots$ be the vertices written in the order of their delays, as before. In step i, if some edge in $\mathcal{M}(v_i, \chi_0)$ continues to be monochromatic in χ_{i-1}, then we flip the color of v_i. Let the resulting coloring be χ_i. Boppana has shown that this algorithm also achieves the performance bounds we claim for our algorithm.

Indeed, one of the referees pointed out that this can be seen readily by comparing what happens in Phase 2 of the two algorithms. Let N be the number of vertices that belong to some monochromatic set after Phase 1, and let X be a random variable having Binomial Distribution $\mathrm{Bin}(N, p)$. The claim is that if Boppana's algorithm is stopped after the first X (of the N) vertices have been considered, then it corresponds exactly to the algorithm we have presented above. In our algorithm, the number of vertices v with $b(v) = 1$ has the same distribution as X, and these vertices are considered in a random order. On the other hand, in Boppana's algorithm (with stopping after X) we pick a random order order on all vertices and stop after X vertices have been considered. Clearly, the two are the same. Thus, Boppana's algorithm when stopped after X vertices have been considered does at least as well as our algorithm; it could do no worse (and might even do better) if it is not stopped early. Note that the bits $b(v)$ play no role in the Boppana's algorithm; they appear only in the analysis of the algorithm!

2.1 Analysis

Let $|E| = k2^n$, for some parameter $k = k(n)$. We wish to show that $\mathcal{M}(\chi^*) = \emptyset$ with non-zero probability if $k = k(n)$ is not "too large". Consider an edge f of H. We will estimate the probability that f is monochromatic in χ^*. We have two cases based on whether or not at least one vertex of f changed its color during the recoloring phase.

Case 1. If f was blue in both χ_0 and χ^*, then we say that event $\mathcal{A}_{Blue}(f)$ took place; that is

$$\mathcal{A}_{Blue}(f) \equiv B(f, \chi_0) \wedge B(f, \chi^*).$$

Similarly, we have $\mathcal{A}_{Red}(f) \equiv R(f, \chi_0) \wedge R(f, \chi^*)$, which states that the edge f was red in χ_0 and this was not rectified while recoloring.

Case 2. Suppose f was not blue in χ_0 but *became* blue during the recoloring phase. That is, during the recoloring phase every red (in χ_0) vertex of f changed its color. Let w be the last red vertex of f to change its color. Why was it necessary to recolor w? There must be an edge $f' \neq f$ such that $w \in f'$, and f' was red in χ_0 and continued to be red until w was considered. That is, f became blue in χ^* because w had to be recolored to rectify the improper coloring of f'. When this happens, we say that f *blames* f' for making it (i.e., f) blue; this event is denoted by $\mathcal{B}_{Blue}(f, f')$. Note that a fixed f can blame more than one f', that is, $\mathcal{B}_{Blue}(f, f')$ might hold for for more than one f'.

To account for the possibility that f was not red in χ_0 but became red in χ^*, we interchange the roles of red and blue in the above discussion. We then arrive at the event $\mathcal{B}_{Red}(f, f')$, which is true exactly when f blames f' for making it red. We thus have the following lemma.

Lemma 2.1 *If $f \in E$ is blue in χ^*, then at least one of $\mathcal{A}_{Blue}(f)$ or $\mathcal{B}_{Blue}(f, f')$ takes place for some $f' \in E$ ($f \neq f'$). If f is red in χ^*, then at least one of $\mathcal{A}_{Red}(f)$ or $\mathcal{B}_{Red}(f, f')$ takes place for some $f' \in E$ ($f \neq f'$).*

Thus, to bound the probability that there is some monochromatic set in χ^* it is enough to bound the probabilities of the events

$$\exists f \in E : (\mathcal{A}_{Blue}(f) \vee \mathcal{A}_{Red}(f)) \qquad \text{and} \qquad \exists f, f' \in E : (\mathcal{B}_{Blue}(f, f') \vee \mathcal{B}_{Red}(f, f')).$$

The next three claims will help us estimate the probabilities of these events.

Claim 2.1 $\Pr[\mathcal{A}_{Blue}(f)] = \Pr[\mathcal{A}_{Red}(f)] = 2^{-n}(1-p)^n$.

Proof: Now $\mathcal{A}_{Blue}(f) \equiv B(f, \chi_0) \wedge (\forall v \in f : b(v) = 0)$. Thus, $\Pr[\mathcal{A}_{Blue}(f)] = 2^{-n}(1-p)^n$. Similarly, $\Pr[\mathcal{A}_{Red}(f)] = 2^{-n}(1-p)^n$. ∎

Claim 2.2 *If* $|f \cap f'| > 1$, *then* $\Pr[\mathcal{B}_{Blue}(f, f')] = \Pr[\mathcal{B}_{Red}(f, f')] = 0$.

Proof: Suppose f blames f' for making it blue. Then, the red vertex of f that was recolored last, say w, lies in f'. That is, all other vertices of $f \cap f'$ became blue before w was considered. But then f' could not have been red just before w was considered for recoloring. Thus, $\Pr[\mathcal{B}_{Blue}(f, f')] = 0$. Similarly, $\Pr[\mathcal{B}_{Red}(f, f')] = 0$. ∎

Let $f \cap f' = \{w\}$. For $S \subseteq f - f'$, consider the following event.

$$\mathcal{E}^1_{Blue}(S, f, f') \equiv R(f', \chi_0) \wedge B(f - S - f', \chi_0) \wedge R(S, \chi_0) \wedge (\forall v \in (S \cup \{w\}) : b(v) = 1).$$

Suppose $\mathcal{B}_{Blue}(f, f')$ holds. Then the event $\mathcal{E}^1_{Blue}(S, f, f')$ must hold for some $S \subseteq f - f'$ (namely, S is the set of red vertices (in χ_0) of $f - f'$). Furthermore, since w was the last red vertex of f to be recolored and f' was red until then, we have that the following event also holds.

$$\mathcal{E}^2_{Blue}(S, f, f') \equiv (\forall u \in S : \mathsf{delay}(u) \leq \mathsf{delay}(w)) \wedge (\forall v \in (f' - \{w\}) : (\mathsf{delay}(v) \geq \mathsf{delay}(w) \vee b(v) = 0)).$$

Here S is the set of red vertices of f. Thus, $\mathcal{B}_{Blue}(f, f')$ implies

$$\hat{\mathcal{B}}_{Blue}(f, f') \stackrel{\text{def}}{=} \exists S \subseteq f - f' : \mathcal{E}_{Blue}(S, f, f'),$$

where $\mathcal{E}_{Blue}(S, f, f') \equiv \mathcal{E}^1_{Blue}(S, f, f') \wedge \mathcal{E}^2_{Blue}(S, f, f')$. Similarly, when considering the event $\mathcal{B}_{Red}(f, f')$ we obtain the corresponding event $\hat{\mathcal{B}}_{Red}(f, f')$. We summarize our observations as follows.

Claim 2.3 $\mathcal{B}_{Blue}(f, f')$ *implies* $\hat{\mathcal{B}}_{Blue}(f, f')$ *and* $\mathcal{B}_{Red}(f, f')$ *implies* $\hat{\mathcal{B}}_{Red}(f, f')$.

Claim 2.4 *If* $|f \cap f'| = 1$, *then* $\Pr[\hat{\mathcal{B}}_{Blue}(f, f')] = \Pr[\hat{\mathcal{B}}_{Red}(f, f')] \leq 2^{-2n+1}p$.

Proof: It is easy to check, using our definition of χ_0, delay and b, that

$$\Pr[\mathcal{E}_{Blue}(S, f, f') \mid \mathsf{delay}(w) = x] = 2^{-2n+1} p^{|S|+1} x^{|S|} (1 - xp)^{n-1}.$$

On integrating over x and summing over all S, we obtain

$$\begin{aligned}
\Pr[\hat{\mathcal{B}}_{Blue}(f,f')] &\le 2^{-2n+1}\sum_{\ell=0}^{n-1}\binom{n-1}{\ell}p^{\ell+1}\int_0^1 x^\ell(1-xp)^{n-1}\,dx \\
&= 2^{-2n+1}p\int_0^1(1-xp)^{n-1}\left[\sum_{\ell=0}^{n-1}\binom{n-1}{\ell}p^\ell x^\ell\right]\,dx \\
&= 2^{-2n+1}p\int_0^1(1-xp)^{n-1}(1+xp)^{n-1}\,dx \\
&= 2^{-2n+1}p\int_0^1(1-(xp)^2)^{n-1}\,dx \\
&\le 2^{-2n+1}p\int_0^1 dx \\
&= 2^{-2n+1}p.
\end{aligned}$$

Similarly, $\Pr[\hat{\mathcal{B}}_{Red}(f,f')] \le 2^{-2n+1}p$. ■

Recall that $|E| = k2^n$. We are now ready to show that if k is not "too large", then with constant probability $\mathcal{M}(\chi^*) = \emptyset$. First, from Claim 2.1, we have

$$\Pr[\exists f \in E : \mathcal{A}_{Blue}(f) \vee \mathcal{A}_{Red}(f)] \le k2^n \times 2 \times 2^{-n}(1-p)^n \le 2k(1-p)^n. \tag{1}$$

Remark. *By considering the positive correlation between the events considered in Claim 2.1, one can improve the bound in (1) to $1-(1-(1-p)^n)^{2k}$. The detailed argument is presented in the appendix.*

Next, from Claims 2.2, 2.3 and 2.4, we have

$$\Pr[\exists f, f' \in E : \mathcal{B}_{Blue}(f,f') \vee \mathcal{B}_{Red}(f,f')] \le k^2 2^{2n} \times 2 \times 2^{-2n+1}p = 4k^2p. \tag{2}$$

By Lemma 2.1,

$$\Pr[\mathcal{M}(\chi^*) \ne \emptyset] \;\le\; 2k(1-p)^n + 4k^2p. \tag{3}$$

For $0 < \epsilon \le 1$, $k = (1/\sqrt{2})(1-\epsilon)\sqrt{n/\ln n}$, $p = (1/2)\ln n/n$, and for all large n, this probability is at most $1-\epsilon$. (In particular, for any fixed $\epsilon > 0$ and $n \ge n_0(\epsilon)$, every n-uniform hypergraph with $(1/\sqrt{2})(1-\epsilon)\sqrt{n/\ln n} \times 2^n$ edges has a proper 2-coloring.) If ϵ is, e.g., a constant, then this "success probability" of ϵ can of course be boosted to any constant probability less than 1 by repeating our basic random process a sufficiently large constant number of times. If $n \ge 2$ is arbitrary, we can take, e.g., $k = (1/10)\sqrt{n/\ln n}$ and $p = (1/2)\ln n/n$; (3) will then imply that every n-uniform hypergraph with at most $(1/10)\sqrt{n/\ln n} \times 2^n$ edges is 2-colorable.

Theorem 2.1 *Let $H = (V, E)$ be an arbitrary n-uniform hypergraph with at most $(1/10)\sqrt{n/\ln n}\times 2^n$ edges; if n is sufficiently large, then H having up to $0.7\sqrt{n/\ln n} \times 2^n$ edges is also admissible. Then, H is 2-colorable; also, a proper 2-coloring for H can be found with high probability in $O(poly(|V| + |E|))$ time.*

A derandomized parallel version of Theorem 2.1 is presented in the next section (see Theorem 3.1).

2.2 When there are few singleton intersections

In our discussion above, pairs of edges that intersect in only one element play a special role. Call $f \in E$ *relevant* iff there is some $f' \neq f$ such that $|f \cap f'| = 1$; let $I(H)$ be the set of relevant edges of H. Suppose we can 2-color the sub-hypergraph of H that only contains the edges in $I(H)$. Then there is a simple way to start with this and 2-color H, as follows. First, if any vertex is currently uncolored (since it only occurred in edges in $E - I(H)$), color each such vertex arbitrarily. Next, repeat the following as long as there exists any monochromatic edge: choose an arbitrary monochromatic edge and flip the color of any one of its vertices. It is known and easy to see that no new monochromatic edge is ever created, and hence this process stops after considering each edge in $E - I(H)$ at most once. Thus, a simple consequence of Theorem 2.1 is

Theorem 2.2 *The consequences of Theorem 2.1 hold even if the upper bounds of Theorem 2.1 on $|E|$, are only upper bounds on $|I(H)|$.*

[In particular, this implies the known fact that if $I(H) = \emptyset$, then H is 2-colorable in polynomial time.]

3 Derandomized parallel version: recoloring with discrete delays

We show how to derandomize and parallelize our algorithm. To this end, we first present the original version of our algorithm, which is more amenable to derandomization. We then present the derandomized parallel version in Section 3.2.

3.1 Recoloring with discrete delays

The original version of our coloring algorithm uses delays chosen from the set $\{1, 2, \ldots, r\}$ (for a suitable $r = \Omega(\log n)$) instead of $[0, 1]$, and will be a useful version to base our derandomized parallel algorithm on. Since the randomized algorithm is very similar to that of Section 2, we only present a brief description now. Phase 1 is the same as the one of Section 2. In phase 2, the main difference is that we pick the delay function $\mathsf{delay} : V \to \{1, 2, \ldots, r\}$ by choosing $\mathsf{delay}(v) = i$ with probability $1/r$, independently for each $v \in V$. The bits $b(\cdot)$ are chosen the same way as before: $b(v) = 1$ with probability p, and $b(v) = 0$ with probability $1 - p$, independently for all $v \in V$.

In phase 2, we now recolor in r stages as follows. Let χ_j denote the coloring at the end of stage j. In stage i, all v *with* $\mathsf{delay}(v) = i$ perform the following action in parallel: if some edge in $\mathcal{M}(v, \chi_0)$ continues to be monochromatic in χ_{i-1} and $b(v) = 1$, then flip the color of v.

The analysis is almost identical to the analysis in Section 2.1. Here, we shall describe only the parts where there is a difference. Let events $\mathcal{A}_{Blue}(f)$, $\mathcal{A}_{Red}(f)$, $\mathcal{B}_{Blue}(f, f')$ and $\mathcal{B}_{Red}(f, f')$ be defined as before. Then, Lemma 2.1 and Claim 2.1 are still valid. For Claim 2.2 we have the following analog.

Claim 3.1 *If either $\mathcal{B}_{Blue}(f, f')$ or $\mathcal{B}_{Red}(f, f')$ takes place, then all vertices in $f \cap f'$ must have the same delay.*

Proof: Suppose $\mathcal{B}_{Blue}(f, f')$ holds, but $\mathsf{delay}(u) < \mathsf{delay}(v)$ for some $u, v \in f \cap f'$. Clearly, u and v were both made blue during the the recoloring phase, and u became blue before v. But then, when the last red vertex in f was considered the color of u was already blue. Hence, f' was not red at that stage; so f could not have blamed f' for making it blue. Thus, all vertices in $f \cap f'$ must have the same delay. By symmetry, the same conclusion follows when $\mathcal{B}_{Red}(f,f')$ holds. ∎

Claim 3.2 $\Pr[\mathcal{B}_{Blue}(f,f')] = \Pr[\mathcal{B}_{Red}(f,f')] \leq 2^{-2n+1}p(2p/r)^{|f\cap f'|-1}e^{(n-|f\cap f'|)p/r}$. *Furthermore, this calculation can be done by only considering the* χ_0, delay, *and* b *values of the vertices in* $f \cup f'$.

Proof: The proof has the same structure as that of Claim 2.4; the only difference is that we allow $|f \cap f'| > 1$. Let $f \cap f' = T$ and let $t = |T|$. For $S \subseteq f - f'$, consider the following event.

$$\mathcal{E}^1_{Blue}(S,f,f') \equiv R(f',\chi_0) \wedge B(f-S-f',\chi_0) \wedge R(S,\chi_0) \wedge (\forall v \in (S \cup T) : b(v) = 1).$$

Suppose $\mathcal{B}_{Blue}(f,f')$ holds. Then the event $\mathcal{E}^1_{Blue}(S,f,f')$ must hold for some $S \subseteq f - f'$ (namely, S is the set of red vertices (in χ_0) of $f - f'$). Also, by Claim 3.1 all vertices in $f \cap f'$ have the same delay, say d. Since the last red vertex of f to be recolored is in f', and f' was red until that vertex was taken up for recoloring, the following event also holds.

$$\mathcal{E}^2_{Blue}(S,d,f,f') \equiv (\forall u \in S : \mathsf{delay}(u) \leq d) \wedge (\forall v \in (f'-T) : (\mathsf{delay}(v) \geq d \vee b(v) = 0)).$$

Thus $\mathcal{B}_{Blue}(f,f')$ implies

$$\exists d \, \exists S \subseteq f - f' : \mathcal{E}_{Blue}(S,d,f,f'),$$

where $\mathcal{E}_{Blue}(S,d,f,f') \equiv \mathcal{E}^1_{Blue}(S,f,f') \wedge \mathcal{E}^2(S,d,f,f') \wedge \mathsf{delay}(T) = \{d\}$.

It is easy to check, using our definition of χ_0, delay and b, that $\Pr[\mathsf{delay}(T) = \{d\}] = 1/r^t$, for $d = 1,2,\ldots,r$, and that

$$\Pr[\mathcal{E}_{Blue}(S,d,f,f') \mid \mathsf{delay}(T) = \{d\}] = 2^{-2n+t}p^{|S|+t}(d/r)^{|S|}(1-(d-1)p/r)^{n-t}.$$

On summing over all d and S, we obtain

$$\begin{aligned}
\Pr[\mathcal{B}_{Blue}(f,f')] &\leq 2^{-2n+t}\sum_{\ell=0}^{n-t}\binom{n-t}{\ell}p^{\ell+t}(1/r)^t\sum_{d=1}^{r}(\frac{d}{r})^{\ell}(1-\frac{d-1}{r}p)^{n-t} \\
&= 2^{-2n+t}(p/r)^t\sum_{d=1}^{r}(1-\frac{d-1}{r}p)^{n-t}\sum_{\ell=0}^{n-t}\binom{n-t}{\ell}(\frac{dp}{r})^{\ell} \\
&= 2^{-2n+t}(p/r)^t\sum_{d=1}^{r}(1-\frac{d-1}{r}p)^{n-t}(1+\frac{d}{r}p)^{n-t} \\
&\leq 2^{-2n+t}(p/r)^t\sum_{d=1}^{r}(1+\frac{p}{r})^{n-t} \\
&\leq 2^{-2n+t}p^t(1/r)^{t-1}e^{(n-t)p/r}.
\end{aligned}$$

Similarly, we bound $\Pr[\mathcal{B}_{Red}(f,f')]$. ∎

Recall that $|E| = k2^n$. To bound the probability that $\mathcal{M}(\chi^*) \neq \emptyset$, we repeat the calculation presented at the end of Section 2, this time using Claims 3.1 and 3.2 instead of Claims 2.2, 2.3 and 2.4.

Inequality (1) still holds. To bound the probability of the event "$\exists f,f' \in E : (\mathcal{B}_{Blue}(f,f') \vee \mathcal{B}_{Red}(f,f'))$", we use Claim 3.2 and obtain that for edges f and f' ($f \neq f'$) with $|f \cap f'| = t \geq 1$,

$$\Pr[\mathcal{B}_{Blue}(f,f') \vee \mathcal{B}_{Red}(f,f')] \leq 2 \times 2^{-2n+1}p(2p/r)^{t-1}e^{p(n-t)/r} \leq 4 \times 2^{-2n}pe^{pn/r};$$

the last inequality holds because $r \geq 2$. Thus, we have

$$\Pr[\exists f, f' \in E : (\mathcal{B}_{Blue}(f, f') \vee \mathcal{B}_{Red}(f, f'))] \leq k^2 2^{2n} \times 4 \times 2^{-2n} p e^{pn/r} = 4k^2 p e^{pn/r}.$$

By Lemma 2.1,

$$\Pr[\mathcal{M}(\chi^*) \neq \emptyset] \ \leq \ 2k(1-p)^n + 4k^2 p e^{pn/r}. \tag{4}$$

For $0 < \epsilon \leq 1$, $k = (1/\sqrt{2})(1-\epsilon)\sqrt{n/\ln n}$, $p = (1/2)\ln n/n$, $r = \lceil \epsilon^{-1} \ln n \rceil$ and for all large n, this probability is at most $1 - \epsilon$.

3.2 Derandomization and parallelization

We now show how the above algorithm can be derandomized and also be made to run in NC^1. A fundamental idea in derandomization, due to Naor & Naor [25], is that randomized algorithms are typically robust to small changes in the underlying distribution. As explained in more detail below, this opens up the following avenue to potentially derandomize a given randomized algorithm. The key goal is to show that there is an efficiently constructible "small" sample space, such that sampling from this small space changes a (carefully crafted) analysis of the algorithm negligibly. Thus, the algorithm will work essentially as well, if its random choice comes from the small space. This in turn yields a deterministic algorithm, which runs the randomized algorithm (in parallel) on all possible seeds from the small space, and finally outputs the best solution found. The work of [25] presents explicit constructions of small sample spaces that "approximate" some properties of certain much larger sample spaces, in a precise sense. To specify the type of "approximation" we need, we start with a definition.

For any non-negative integer t, let $[t] = \{0, 1, \ldots, t-1\}$; an interval of $[t]$ is a set of the form $\{i : a \leq i \leq b\}$, for some $a, b \in [t]$ such that $a \leq b$. Let $\mathcal{J}_t$ be the set of all intervals of $[t]$.

Extending the work of [25], Even, Goldreich, Luby, Nisan & Veličković defined the following [13, 14]. Suppose $X_1, X_2, \ldots, X_N \in [t]$ are *independent* random variables with *arbitrary* individual distributions and joint distribution D; let $\vec{X}$ denote the vector $(X_1, X_2, \ldots, X_N)$. Call a set $A \subseteq [t]^N$ an (ℓ, δ)-approximation for D if, for $\vec{Y} = (Y_1, \ldots, Y_N)$ sampled uniformly at random from A,

- for all index sets $I \subseteq \{1, 2, \ldots, N\}$ with $|I| \leq \ell$, and
- for all $J : I \to \mathcal{J}_t$,

we have

$$|\Pr[\bigwedge_{i \in I}(Y_i \in J(i))] - \prod_{i \in I} \Pr_D[X_i \in J(i)]| \leq \delta.$$

(We call any event of the form "$\bigwedge_{i \in I}(X_i \in J(i))$" an *interval event w.r.t.* $\vec{X}$.) Among other results, it was shown in [13] that such a set A with cardinality $\text{poly}(2^\ell, \log N, \delta^{-1})$ can be constructed explicitly using $\text{poly}(|A| + N)$ processors in $O(\log(|A| + N))$ time on an EREW PRAM. (See [14] for the journal version of [13].)

Some remarks on this construction of [13]:

- The construction possesses some stronger properties, which, however, we do not need here.
- As described in [13], the construction handles the situation where $J(i)$ is a singleton for each i, but it is easy to see that a minor modification makes it work for all interval events.

- The description in [13] does not explicitly discuss such parallel constructions, but such a parallel version of the work of [13] is immediate from a reading of [13].

- $|A|$ above does not depend on t, since we can assume without loss of generality that $t = O(\ell/\delta)$ [13].

One basic utility of such constructions to, say, derandomized parallel algorithms, is as follows. Given a randomized algorithm that uses the independent random variables $X_1, X_2, \ldots, X_N$, one first shows that the analysis is changed little by an approximation as defined above, if ℓ is sufficiently large and δ suitably small. Then, one may just exhaustively search such a sample space A (allocating an appropriate number of processors to each element of A), and thus *deterministically* find a value for $(X_1, X_2, \ldots, X_N)$ that is "good enough" for the algorithm. In our context, we get

Theorem 3.1 *For any sufficiently large n, let $H = (V, E)$ be an arbitrary n-uniform hypergraph with at most $0.7\sqrt{n/\ln n} \times 2^n$ edges. Then, H is 2-colorable; also, a proper 2-coloring for H can be found in NC^1.*

Proof: Suppose for a sufficiently large n that $H = (V, E)$ is an n-uniform hypergraph with at most $0.7\sqrt{n/\ln n} \times 2^n$ edges. Clearly, we may assume without loss of generality that $|V| \leq n|E|$. We may also assume that $|E| \geq 2^{n/500}$: if not, one can employ the NC^1 algorithm of [1] to 2-color H. So suppose $|E| \geq 2^{n/500}$. Then, the "input size" for our problem is $2^{\Theta(n)}$, and hence an NC^1 algorithm would use $2^{\Theta(n)}$ processors and $O(n)$ time.

We now show how to derandomize the algorithm of this section within these processor and time bounds. In the "approximating distributions" notation used above, the underlying independent random variables $X_1, X_2, \ldots$ are the χ_0, b and delay values of the vertices; $t = r$. Let D be the joint distribution of these independent random variables. Our plan is to proceed as follows. We will show that our failure probability bound (4) is obtainable as a sum of at most $s = 2^{O(n)}$ terms, where each term is the probability of an interval event (w.r.t. $\vec{X}$) that depends on at most $6n$ of the variables $X_1, X_2, \ldots$. So, since (4) leads to a failure probability that is bounded away from 1, we may instead generate the random variables $X_1, X_2, \ldots$ from a $(6n, \delta)$-approximation A for D, where $\delta = o(1/s)$ (say, $1/(s \log n)$); it is then easy to check that our failure probability is still bounded away from 1. Thus, we would have shown the existence of a point in A that leads to a successful 2-coloring for H. Since $|A| = 2^{\Theta(n)}$, the exhaustive search of A can be made to run in $O(n)$ time using $2^{\Theta(n)}$ processors.

So, let us study our analysis of this section, and the proof of Claim 3.2. Recall our goal from above: to show that our failure probability bound is a sum of at most $2^{O(n)}$ terms, each term being the probability of an interval event that depends on at most $6n$ of the X_i. Our failure probability was the sum of the probabilities of the events $\mathcal{A}_{Blue}(f)$, $\mathcal{A}_{Red}(f)$, $\mathcal{E}_{Blue}(S, d, f, f')$ and $\mathcal{E}_{Red}(S, d, f, f')$, for all edges f, for all $f' \neq f$ with $f' \cap f \neq \emptyset$, for all $S \subseteq (f - f')$, and for all $d \in \{1, 2, \ldots, r\}$. It is immediate that there are $2k2^n$ events of the form $\mathcal{A}_{Blue}(f)$ and $\mathcal{A}_{Red}(f)$, and that each of these is an interval event w.r.t. $\vec{X}$ that depends on $2n$ of the X_i. Next, choose distinct edges f, f' arbitrarily such that $|f' \cap f| \geq 1$. There are at most $2^{2n}\mathrm{poly}(n)$ such choices. Having fixed such a choice of f, f', choose any $S \subseteq (f - f')$ and any $d \in \{1, 2, \ldots, r\}$: there are at most $O(2^n \log n)$ such choices. We now show that $\mathcal{E}_{Blue}(S, d, f, f')$ can be expressed as the disjunction of at most 2^n *disjoint* interval events; similarly for $\mathcal{E}_{Red}(S, d, f, f')$. This would conclude the proof.

Let $f \cap f' = T$. Recall that $\mathcal{E}_{Blue}(S, d, f, f') \equiv \mathcal{E}^1_{Blue}(S, f, f') \wedge \mathcal{E}^2(S, d, f, f') \wedge \mathsf{delay}(T) = \{d\}$. It is easily seen that $\mathcal{E}_{Blue}(S, d, f, f')$ depends on at most $6n$ of the X_i. Also, $\mathcal{E}_{Blue}(S, d, f, f')$ is "almost" an interval event, except for the part "$\forall v \in (f' - T) : (\mathsf{delay}(v) \geq d \vee b(v) = 0)$". However, for any given v, the event "$\mathsf{delay}(v) \geq d \vee b(v) = 0$" can be rewritten as the disjunction of two *disjoint* interval events:

$$(\mathsf{delay}(v) \geq d \vee b(v) = 0) \;\equiv\; (b(v) = 0) \vee (b(v) = 1 \wedge \mathsf{delay}(v) \in [d, r]).$$

Thus, $\mathcal{E}_{Blue}(S, d, f, f')$ is the disjunction of at most 2^n disjoint interval events. This finishes the proof. ■

4 Generalization: A Local Version

As mentioned in the introduction, a useful parameter of H is its *overlap* $D = \max_{f \in E} |\{f' \in E : f \cap f' \neq \emptyset\}|$. We now show that any n-uniform hypergraph with $D \leq 0.17\sqrt{n/\ln n} \times 2^n$ is 2-colorable, for sufficiently large n.

Let us recall a special case of the LLL, which shows a useful sufficient condition for simultaneously avoiding a set $A_1, A_2, \ldots, A_N$ of "bad" events:

Theorem 4.1 ([12]) *Suppose events $A_1, A_2, \ldots, A_N$ are given. Let $S_1, S_2, \ldots S_N$ be subsets of $[N]$ such that for each i, A_i is independent of (any Boolean combination of) the events $\{A_j : j \in ([N] - S_i)\}$. Suppose that $\forall i \in [N]$: (i) $\Pr[A_i] < 1/2$, and (ii) $\sum_{j \in S_i} \Pr[A_j] \leq 1/4$. Then, $\Pr[\bigwedge_{i \in [N]} \overline{A_i}] > 0$.*

Remark: Often, each $j \in N$ will be an element of at least one of the sets S_i; in such cases, it clearly suffices to only verify condition (ii) of Theorem 4.1.

Suppose H is an n-uniform hypergraph with overlap $D = \lambda 2^n$. Let χ^* be the random coloring obtained by running the slow recoloring algorithm of Section 2; the value of p will be specified shortly. By Lemma 2.1, if we can simultaneously avoid the following events, then χ^* will be a valid 2-coloring of H:

$$\{\mathcal{A}_{Blue}(f), \mathcal{A}_{Red}(f) : f \in E\} \cup \{\mathcal{B}_{Blue}(f, f'), \mathcal{B}_{Red}(f, f') : f, f' \in E\}.$$

In Claims 2.2 and 2.3 we observed that the event $\mathcal{B}_{Blue}(f, f')$ holds only if $|f \cap f'| = 1$ and the event $\hat{\mathcal{B}}_{Blue}(f, f')$ holds; similarly $\mathcal{B}_{Red}(f, f')$ holds only if $|f \cap f'| = 1$ and the event $\hat{\mathcal{B}}_{Red}(f, f')$ holds. Thus, it is enough if we can simultaneously avoid the following two types of events.

- Type 1 events: $\{\mathcal{A}_{Blue}(f), \mathcal{A}_{Red}(f) : f \in E\}$.
- Type 2 events: $\{\hat{\mathcal{B}}_{Blue}(f, f'), \hat{\mathcal{B}}_{Red}(f, f') : (f, f' \in E) \wedge (|f \cap f'| = 1)\}$.

We will now show that for $\lambda \leq 0.17\sqrt{n/\ln n}$, these events satisfy the conditions of Theorem 4.1. Thus, we can simultaneously avoid these *bad* events.

Our argument rests on the following observations.

(a) Every bad event has at most two edges as its arguments.

(b) The occurrence of a bad event is completely determined once the value of $\chi_0(v)$, $\mathsf{delay}(v)$ and $b(v)$ are fixed for all vertices belonging to its arguments.

For instance, consider the bad event $\mathcal{B} = \hat{\mathcal{B}}_{Blue}(f, f')$. Suppose $\mathcal{C}$ is any collection of bad events such that for *each* event $\mathcal{E} \in \mathcal{C}$: *no* argument of $\mathcal{E}$ intersects f, and *no* argument of $\mathcal{E}$ intersects f'. Then, the above observations, together with the fact that χ_0, delay and b are chosen independently for different vertices, imply that $\mathcal{B}$ is independent of any Boolean combination of events in $\mathcal{C}$.

For a bad event $\mathcal{B}$, let $S(\mathcal{B})$ be the set of all bad events at least one of whose arguments has a non-empty intersection with an argument of $\mathcal{B}$. Thus, as discussed above, $\mathcal{B}$ is independent of any Boolean combination of the events outside $S(\mathcal{B})$. Thus, to apply Theorem 4.1, we need to bound the sum of the probabilities of the events in $S(\mathcal{B})$. To do this, we will first bound the number of events of each type in $S(\mathcal{B})$; then we will use Claims 2.1 and 2.4 to bound their probabilities.

Claim 4.1 *For all bad events $\mathcal{B}$, $S(\mathcal{B})$ has at most $4D$ events of type 1 and at most $8D^2$ events of type 2.*

Proof: Let the arguments of $\mathcal{B}$ be f and f' (we will take $f = f'$ if $\mathcal{B}$ is a type 1 event). The only events of type 1 that are in $S(\mathcal{B})$ correspond to edges that intersect either f or f'. There are at most $2D$ such edges by the definition of D; for each such edge h, there are two type 1 events—$\mathcal{A}_{Blue}(h)$ and $\mathcal{A}_{Red}(h)$. Thus, $S(\mathcal{B})$ has at most $4D$ events of type 1.

For a type 2 event with arguments (h, h') to be in $S(\mathcal{B})$, at least one of h and h' must intersect at least one of f and f'; furthermore, h and h' must themselves intersect. It follows that there are at most $4D^2$ possibilities for (h, h'). For each such argument (h, h'), there are two bad events $\hat{\mathcal{B}}_{Blue}(h, h')$ and $\hat{\mathcal{B}}_{Red}(h, h')$ in $S(\mathcal{B})$. Thus the number of type 2 events in $S(\mathcal{B})$ is at most $8D^2$. ■

Claim 4.2 *Suppose $D = \lambda 2^n$, where $\lambda \leq 0.17\sqrt{n/\ln n}$ and n is sufficiently large. Then, for any bad event $\mathcal{B}$, $\sum_{\mathcal{E}\in S(\mathcal{B})} \Pr[\mathcal{E}] \leq 1/4$.*

Proof: We use Claim 4.1 to bound the number of events of each type in $S(\mathcal{B})$; to bound the probabilities of these events, we use Claims 2.1 and 2.4. We have

$$\sum_{\mathcal{E}\in S(\mathcal{B})} \Pr[\mathcal{E}] \leq 4D \times 2^{-n}(1-p)^n + 8D^2 \times 2^{-2n+1}p = 4\lambda(1-p)^n + 16\lambda^2 p.$$

If $p = (1/2)\ln n/n$, $\epsilon > 0$ and $\lambda = (1-\epsilon)\sqrt{1/32}\sqrt{n/\ln n}$, then for all large n, this probability is less than $1/4$. (Note that $\sqrt{1/32} > 0.176$.) ■

We have thus established that condition (ii) of Theorem 4.1 holds if D is chosen suitably. As remarked before, this implies that condition (i) holds as well.

Theorem 4.2 *Suppose, for a sufficiently large n, that H is an arbitrary n-uniform hypergraph in which each edge intersects at most $0.17\sqrt{n/\ln n} \times 2^n$ other edges. Then, H is 2-colorable.*

By following the proof of Theorem 2.2, we see that Theorem 4.2 holds even if each *relevant* edge intersects at most $0.17\sqrt{n/\ln n} \times 2^n$ other *relevant* edges.

5 Almost-disjoint hypergraphs

We now explore possibilities for improving the results of Sections 2 and 4. As seen in Section 2.2, hypergraphs where pairwise edge-intersections are typically “large”, can often be 2-colored. This seems to suggest that 2-coloring may be most difficult for hypergraphs where edge-intersections are “small”. Motivated in part by this, Erdős and Lovász considered *simple* hypergraphs (also called *nearly-disjoint* hypergraphs), wherein any two distinct edges intersect in at most one vertex [12]. Recall the function $m(n)$ defined in the introduction; let $m^*(n)$ be the analog of $m(n)$ when we restrict our attention to nearly-disjoint n-uniform hypergraphs. It is shown in [12] that

$$\Omega(4^n/n^3) \leq m^*(n) \leq \mathrm{O}(n^4 4^n). \tag{5}$$

(These results of [12] were pointed out to us by Noga Alon and Jeong-Han Kim.) This lower bound on $m^*(n)$ has been further improved to $4^n/n^{1+\epsilon}$ by Szabó [28], for any fixed $\epsilon > 0$ and all sufficiently large n. This result was pointed out to us by József Beck.

Thus, $m^*(n) \gg m(n)$. So, in order to make progress on our goal of improving the result of Section 2, we need to consider a restricted family of hypergraphs such that: (i) edge-intersections are “small”, and (ii) the family is rich enough so that the restriction of $m(n)$ to the family is not much bigger than $m(n)$. This is what is done in the rest of this section.

Remark: Simple hypergraphs sometimes connote “hypergraphs with no repeated edges”. Throughout this paper, simple hypergraphs mean nearly-disjoint hypergraphs as mentioned above.

5.1 Definition and Theorems

Let F be a collection of subsets of some universe $\mathcal{A}$. For the definitions below, we do not assume that all elements of F are distinct, i.e., we allow F to be a multiset; however, when we construct hypergraphs, we will ensure that in the end all its edges are distinct. For $a \in \mathcal{A}$, let

$$d_F(a) \stackrel{\text{def}}{=} |\{f \in F : a \in f\}|.$$

Define $\binom{F}{t}$ to be the (multi-)set of all t-element subcollections of F. Let

$$\begin{aligned} \Lambda(F) &\stackrel{\text{def}}{=} \sum_{a \in \mathcal{A}} \max\{0, d_F(a) - 1\} \\ \mathcal{I}(F) &\stackrel{\text{def}}{=} 2^{\Lambda(F)} \\ \mathcal{I}_t(F) &\stackrel{\text{def}}{=} \mathop{\mathbf{E}}_{H \in \binom{F}{t}} [\mathcal{I}(H)]. \end{aligned} \tag{6}$$

[In (6), the expectation is over a H chosen uniformly at random from $\binom{F}{t}$.]

Our interest in $\Lambda(F)$ stems from the equality

$$|\bigcup_{f \in F} f| = (\sum_{f \in F} |f|) - \Lambda(F).$$

If the edges of F are disjoint, then $\Lambda(F) = 0$, and $\mathcal{I}_t(F) = 1$ for all $t \geq 1$. Suppose $F = \{f_1, f_2, \ldots\}$ and that for some s and any $i \neq j$, $|f_i \cap f_j| \leq s$. Then, by inclusion-exclusion, $|\bigcup_{f \in F} f| \geq \sum_{f \in F} |f| - s\binom{|F|}{2}$; i.e.,

$$\Lambda(F) \leq s\binom{|F|}{2}; \ \forall t \geq 1, \ \mathcal{I}_t(F) \leq 2^{s\binom{t}{2}}. \tag{7}$$

Thus, $\mathcal{I}$ being "small" can be thought of some notion of "small pairwise intersections". We wish to consider hypergraphs whose edges do not, *on an average*, intersect much. We will use $\mathcal{I}_t(F)$ to formalize this notion.

Definition 5.1 (ϵ-almost-disjoint hypergraphs) *We say that a family of hypergraphs $\{G_n = (V_n, E_n)\}$ is a family of* uniform ϵ-almost-disjoint hypergraphs *if G_n is n-uniform, and for all large n, there exists a t ($2 \le t \le (\ln n)^{1/3}$) such that*

$$\mathcal{I}_t(E_n) \le n^{\epsilon t - 3}. \tag{8}$$

So, rather than require all pairwise edge-intersections to be "small" (say, at most 1), we just need edge-intersections to be "small on average", in the sense defined above. We will show that for almost-disjoint families of hypergraphs, there is a coloring algorithm that works even if more edges are allowed. First, we will see that Definition 5.1 captures a large enough family of hypergraphs, by showing that $m(n)$ restricted to such families, for any fixed $\epsilon > 0$, is still at most $\mathrm{O}(n^2 2^n)$.

Theorem 5.1 *For any fixed $\epsilon > 0$, there is an infinite family $\{G_n\}$ of uniform ϵ-almost-disjoint hypergraphs such that: (a) G_n has at most $n^2 2^n$ edges, and (b) for all sufficiently large n, G_n is* not *2-colorable.*

We next show the main result of this section that complements Theorem 5.1. Theorem 5.2 shows that for ϵ-almost-disjoint families of hypergraphs, there is a coloring algorithm that works even if up to $n^{1-\epsilon} 2^n$ edges are allowed. [From now on, whenever we claim that an algorithm 2-colors a hypergraph "with high probability" (whp), we just mean that the algorithm succeeds with probability lower bounded by some positive constant. Since we can efficiently check if a coloring is a 2-coloring, such an algorithm can be repeated a sufficient number of times to drive down the failure probability.]

Theorem 5.2 *Let $\{G_n\}$ be a family of uniform ϵ-almost-disjoint hypergraphs. Suppose G_n has at most $n^{1-\epsilon} 2^n$ edges ($\epsilon > 0$). Then, for all large enough n (i.e., $n \ge n(\epsilon)$), G_n is 2-colorable; there is a randomized polynomial-time algorithm that constructs such a 2-coloring whp.*

5.2 Proof of Theorem 5.1: a random family is almost-disjoint

The probabilistic construction of non-2-colorable n-uniform hypergraphs in [4, page 8] proceeds as follows: take $V(G_n) = n^2/2$, and for $E(G_n)$ pick $(e(\ln 2)/4 + o(1)) n^2 2^n$ edges at random (with replacement, say, but note that the probability of picking the same n-set up in two different trials is negligible, and the hypergraph will whp have no repeated edges). As usual, e here denotes the base of the natural logarithm. Then, with probability at least 3/4, the resulting random hypergraph G_n cannot be 2-colored [4]. We will show that with probability at least 3/4 the hypergraph will be almost-disjoint; thus, the "$n^{1-\epsilon}$" in Theorem 5.2 cannot be replaced by any term larger than n^2.

We define $\exp(x) \doteq e^x$. We start with a useful claim.

Claim 5.1 *For the random family $\{G_n\}$ considered above, we have for all $t \ge 2$ and for all $n \ge 4\ln(2t)$ that $\mathbf{E}[\mathcal{I}_t(E(G_n))] \le 2\exp(4t^2)$.*

Proof: Let $F = (f_1, f_2, \ldots, f_t)$ be a random collection of n-subsets of V_n obtained by choosing t subsets at random (with replacement). We will show that if $n \ge 4\ln(2t)$, then

$$\mathbf{E}[\mathcal{I}(F)] \le 2\exp(4t^2). \tag{9}$$

It will then follow, by linearity of expectation, that $\mathbf{E}[\mathcal{I}_t(E(G_n))] \leq 2\exp(4t^2)$.

Let $N \stackrel{\text{def}}{=} |V_n|$. Instead of computing $\mathcal{I}(F)$ directly, it will be easier to relate it to $\mathcal{I}(F')$, for a slightly different family F' (for which the calculations are simpler). The family F' has t sets $f'_1, f'_2, \ldots, f'_t$, where f'_i is generated by picking each element of the universe independently with probability $p \stackrel{\text{def}}{=} 2n/N = 4/n$. If each set in F' has size at least n, pick a random n-sized subset of each; in this case, the resulting collection F'' has the same distribution as F. Now,

$$\begin{aligned}\mathbf{E}[\mathcal{I}(F')] &\geq \Pr[\forall f' \in F' \, |f'| \geq n] \times \mathbf{E}[\mathcal{I}(F'') \mid \forall f' \in F' |f'| \geq n] \\ &= \Pr[\forall f' \in F' \, |f'| \geq n] \times \mathbf{E}[\mathcal{I}(F)].\end{aligned}$$

Hence

$$\mathbf{E}[\mathcal{I}(F)] \leq \mathbf{E}[\mathcal{I}(F')] / \Pr[\forall f' \in F' \, |f'| \geq n]. \tag{10}$$

Now, our claim will follow from (10), if we show the following.

$$\mathbf{E}[\mathcal{I}(F')] \leq \exp(4t^2); \tag{11}$$

$$\Pr[\forall f' \in F' \, |f'| \geq n] \geq \frac{1}{2}. \tag{12}$$

First, consider (11). We have

$$\begin{aligned}\mathbf{E}[\mathcal{I}(F')] = \mathbf{E}[2^{\Lambda(F')}] &= \mathbf{E}[\prod_{a \in V_n} 2^{\max\{0, d_{F'}(a)-1\}}] \\ &= \prod_{a \in V_n} \mathbf{E}[2^{\max\{0, d_{F'}(a)-1\}}]. \end{aligned} \tag{13}$$

It, thus, suffices to bound $\mathbf{E}[2^{\max\{0, d_{F'}(a)-1\}}]$ separately for each a. Fix $a \in V_n$. Then,

$$\begin{aligned}\mathbf{E}[2^{\max\{0, d_{F'}(a)-1\}}] &= (1-p)^t + \sum_{j=1}^{t} 2^{j-1} \binom{t}{j} p^j (1-p)^{t-j} \\ &= [(1-p)^t + (1+p)^t]/2 \\ &\leq [\exp(-pt) + \exp(pt)]/2 \\ &\leq \exp(p^2 t^2/2).\end{aligned}$$

[For the last inequality, compare the Taylor series of the two sides.] Then, by (13), we have

$$\mathbf{E}[\mathcal{I}(F')] \leq \exp(\frac{p^2 t^2}{2} N) \leq \exp(4t^2),$$

thus establishing (11).

It remains to show (12). For $i = 1, 2, \ldots, t$, (see [4, page 238,Theorem A.13])

$$\Pr[|f'_i| < n] \leq \exp(-\frac{n^2}{4n}) \leq \exp(-\frac{n}{4}).$$

Thus, $\Pr[\exists f' \in F' \, |f'| < n] \leq t\exp(-\frac{n}{4}) \leq 1/2$, since $n \geq 4\ln(2t)$. ■

Given any $\epsilon > 0$, let n be sufficiently large. Let $\mathcal{D}$ be the event that all the edges of G_n are distinct. It is easy to see that $\Pr[D] \geq 9/10$. It follows from Claim 5.1 that

$$\mathbf{E}[\mathcal{I}_t(E(G_n)) \mid \mathcal{D}] \leq \frac{\mathbf{E}[\mathcal{I}_t(E(G_n))]}{\Pr[\mathcal{D}]} \leq \frac{20}{9}\exp(4t^2).$$

Then, by Markov's inequality we have, for each t ($2 \leq t \leq (\ln n)^{1/3}$),

$$\Pr[\mathcal{I}_t(E(G_n)) \geq 100\exp(4t^2)(\ln n)^{1/3} \mid \mathcal{D}] \leq \frac{1}{45(\ln n)^{1/3}},$$

and

$$\Pr[\exists t,\ 2 \leq t \leq (\ln n)^{1/3} :\ \mathcal{I}_t(E(G_n)) > 100\exp(4t^2)(\ln n)^{1/3} \mid \mathcal{D}] \leq \frac{1}{45}.$$

Note that if n is sufficiently large as a function of ϵ, then $100\exp(4t^2)(\ln n)^{1/3} \ll n^{\epsilon t-3}$ for any $t \leq (\ln n)^{1/3}$. Thus, with probability at least $\frac{9}{10} \times \frac{44}{45} > \frac{3}{4}$, G_n has no repeated edges and is ϵ-almost-disjoint in a strong sense: for *all* t, $2 \leq t \leq (\ln n)^{1/3}$, it is true that $\mathcal{I}_t(E(G_n)) \leq n^{\epsilon t-3}$.

5.3 Algorithm for Theorem 5.2

As before, let χ_0 be the coloring obtained after the first random coloring phase. We say that an edge $f \in E$ is *almost-red* in χ_0 if: (1) $f \notin \mathcal{M}(\chi_0)$ and (2) f has at most τ blue vertices (τ is a parameter to be chosen later). Let $\mathcal{AR}(\chi_0)$ be the set of almost-red edges with respect to χ_0. Similarly, we define *almost-blue* and $\mathcal{AB}(\chi_0)$. In the recoloring phase we pay special attention to the edges in $\mathcal{AR}(\chi_0)$ and $\mathcal{AB}(\chi_0)$.

We will modify the algorithm used in the proof of Theorem 2.1. In the analysis we present now, it is not important that the vertices be considered in a random order. So let us fix an ordering $v_1, v_2, \ldots$ of the vertices. Step i is now implemented as follows.

Step i**.** *If in* χ_{i-1}*: (a)* v_i *is the only red vertex of an edge in* $\mathcal{AB}(\chi_0)$*, or (b)* v_i *is the only blue vertex of an edge in* $\mathcal{AR}(\chi_0)$*, then skip* v_i*. Otherwise, if some edge in* $\mathcal{M}(v_i, \chi_0)$ *continues to be monochromatic in* χ_{i-1} *and* $b(v_i) = 1$*, then flip the color of* v_i*. Let the resulting coloring be* χ_i*.*

As before, let χ^* be the coloring obtained after all vertices have been considered.

5.4 Analysis of the algorithm of Section 5.3

Fix a family $\{G_n\}$ of uniform ϵ-almost-disjoint hypergraphs. Fix n large, and let $G = (V, E)$ stand for the graph $G_n = (V_n, E_n)$ of the family. Let $|E|/2^n \leq k$, where $k = n^{1-\epsilon}$. Suppose t is the positive integer satisfying (8). We will show that the above algorithm, with $\tau = t - 2$, produces a proper coloring for G with probability lower-bounded by some positive absolute constant. Note that we can assume that $\epsilon < 2/3$, for by Theorem 2.1, n-uniform hypergraphs with at most $n^{1/3}2^n$ edges are 2-colorable. Recall that the bits $b(v)$ were set to be 1 with probability p and 0 with probability $1-p$; we will take $p = (2\ln k)/n$. Let us bound the probability of failure of this algorithm. Suppose f is blue in χ^* (the other case, when f is red, is similar). We have three possibilities:

Case 0. f was red in χ_0.
Case 1. f was blue in χ_0.
Case 2. f was not completely red or blue in χ_0, but was made blue during the recoloring process.

Note that this implies that f was *not* almost-blue in χ_0, because the new recoloring phase never makes such edges blue.

We will now bound the probabilities of the three cases separately. We will show that for all large enough n,

$$\Pr[\exists f \ \mathsf{Case0}(f)] < \frac{1}{10}; \tag{14}$$

$$\Pr[\exists f \ \mathsf{Case1}(f)] < \frac{1}{10}; \tag{15}$$

$$\Pr[\exists f \ \mathsf{Case2}(f)] < \frac{1}{10}. \tag{16}$$

It follows that $\Pr[\mathcal{M}(\chi^*) \neq \emptyset] < 2(1/10 + 1/10 + 1/10) = 3/5$; therefore, G has a proper 2-coloring that can be found whp using our algorithm.

5.4.1 Case 0

It is easy to see that the probability of this happening is at most $|E|2^{-n}p^n$; for large n, this is less than 1/10, since $p \leq (2\ln n)/n$ and $|E| \leq n2^n$. Thus, (14) holds.

5.4.2 Case 1

Suppose f was blue in χ_0 and continued to be blue in χ^*. Let us examine the events that led to this. Our algorithm attempted to change the colors of vertices $v \in f$ for which $b(v) = 1$. Thus, if f continued to be blue in χ^*, then it must be that all these attempts failed. Let

$$\mathsf{Pivots}(f) = \{v \in f : b(v) = 1\}.$$

We have two cases. First, it might be that not enough attempts were made to change the color of vertices in f; that is, $\mathsf{Pivots}(f)$ was very small—an extreme case of this occurs when $\mathsf{Pivots}(f) = \emptyset$. Second, it might be that a good number of attempts were made, but each time the color of the vertex could not be flipped because it happened to be the only blue vertex of some other edge that was almost red in χ_0 (see the revised recoloring step above). We bound the probabilities of these two cases separately. We, therefore, write

$$\Pr[\mathsf{Case1}(f)] \leq (S1) + (S2), \tag{17}$$

where

$$\begin{aligned} (S1) &= \Pr[\mathsf{Case1}(f) \wedge |\mathsf{Pivots}(f)| \leq t-1]; \\ (S2) &= \Pr[\mathsf{Case1}(f) \wedge |\mathsf{Pivots}(f)| \geq t]. \end{aligned}$$

First, consider $(S1)$.

$$\begin{aligned} (S1) &\leq \Pr[B(f,\chi_0) \wedge |\mathsf{Pivots}(f)| \leq t-1] \\ &= 2^{-n} \Pr[|\mathsf{Pivots}(f)| \leq t-1] \\ &= 2^{-n} \sum_{i=0}^{t-1} \binom{n}{i} p^i (1-p)^{n-i} \end{aligned}$$

$$\leq \quad 2^{-n}(1-p)^{n-t}\sum_{i=0}^{t-1}\binom{n}{i}p^i$$

$$\leq \quad 2^{-n}(1-p)^{n-t}\sum_{i=0}^{t-1}(np)^i$$

$$\leq \quad 2^{-n}(1-p)^{n-t}(np)^t. \tag{18}$$

Since $p = (2\ln k)/n$, $t \leq (\ln n)^{1/3}$ and $k = n^{1-\epsilon}$, we have

$$(S1) \leq \frac{2^{-n}}{k^{3/2}}. \tag{19}$$

Now, we consider (S2), that is, the probability that f remains blue even when at least t attempts were made. It follows from our definition of recoloring that if an attempt at recoloring a certain vertex $v \in \mathsf{Pivots}(f)$ was unsuccessful, then v was the last blue vertex of some edge $f' \in \mathcal{AR}(\chi_0)$. This motivates the following definition.

Definition 5.2 *For $v \in f$, we say that v is* blocked *by the edge f' if*

(a) $f \cap f' = \{v\}$*;*

(b) $b(v) = 1$*;*

(c) $f' \in \mathcal{AR}(\chi_0)$*;*

(d) for all blue (in χ_0) vertices w of $f' - \{v\}$, $b(w) = 1$.

We then have the following claim.

Claim 5.2 *If $B(f,\chi_0) \wedge B(f,\chi^*)$ and $v \in \mathsf{Pivots}(f)$, then v is blocked by some $g \in E$.*

Note, in particular, that if edge g is held responsible for preventing the recoloring of vertex $v \in \mathsf{Pivots}(f)$, then $g \cap f = \{v\}$. In our case, not just one but t vertices are prevented from recoloring. Thus, there is a set of edges $F = \{f_1, f_2, \ldots, f_t\}$ where each f_i blocks a different vertex of $\mathsf{Pivots}(f)$.

Definition 5.3 *Let F be a set of t edges $\{f_1, f_2, \ldots, f_t\}$, where each f_i intersects f on exactly one vertex v_i, and these t vertices are distinct. We say that F conspires against f if*

(i) f is blue in χ_0.

(ii) f_i blocks v_i, for $i = 1, 2, \ldots, t$.

We then have the following analog of Claim 5.2.

Claim 5.3 *If $B(f,\chi_0) \wedge B(f,\chi^*)$ and $|\mathsf{Pivots}(f)| \geq t$, then $\exists F \in \binom{E}{t}$ such that F conspires against f.*

Thus,

$$(S2) \leq \sum_{F \in \binom{E}{t}} \Pr[F \text{ conspires against } f]. \tag{20}$$

Fix $F \in \binom{E}{t}$, such that each edge in F intersects f on exactly one element and different edges in F intersect f on different elements. (If F does not have this property then $\Pr[F \text{ conspires against } f] = 0$.) Let $f_1, f_2, \ldots, f_t$ be the elements of F listed in some order. Let $f_0 \stackrel{\text{def}}{=} f$; for $i = 1, 2, \ldots, t$, let $\hat{f}_i \stackrel{\text{def}}{=} f_i - \bigcup_{j=0}^{i-1} f_j$. If F conspires against f, then we must have $b(w) = 1$ for all vertices $w \in \hat{f}_i$ that were blue in χ_0. We summarize our observations as follows: if F conspires against f, then

(C1) f is blue in χ_0; and for $i = 1, 2, \ldots, t$,

(C2) if $f_i \cap f = \{v_i\}$, then $b(v_i) = 1$;

(C3) $\forall w \in \hat{f}_i \, B(w, \chi_0) \rightarrow b(w) = 1$;

(C4) $\hat{f}_i$ has at most $\tau - 1$ blue vertices in χ_0.

It follows that

$$\Pr[F \text{ conspires against } f] \leq 2^{-n} \times p^t \times \prod_{i=1}^{t} \sum_{j=0}^{\tau-1} \binom{|\hat{f}_i|}{j} p^j 2^{-|\hat{f}_j|}. \tag{21}$$

(Here the first factor is justified by (C1), the second by (C2), and the third by (C3) and (C4).) Let $f_i' = f_i - f$, for $i \geq 1$; note that $|f_i'| = n - 1$. Also let $F' = \{f_i' : 1 \leq i \leq t\}$. It is easy to check that

$$\sum_{i=1}^{t} |\hat{f}_i| = |\bigcup_{i=1}^{t} f_i'| = t(n-1) - \Lambda(F') = t(n-1) - \Lambda(F).$$

Also, since $|\hat{f}_i| \leq n$, we have $\sum_{j=0}^{\tau-1} \binom{|\hat{f}_i|}{j} p^j \leq (np)^\tau$. Thus, (21) gives

$$\Pr[F \text{ conspires against } f] \leq 2^{-n} \times p^t \times 2^{-t(n-1)+\Lambda(F)} \times (np)^{\tau t}. \tag{22}$$

Then, by (20) we have

$$\begin{aligned}
(S2) &\leq \binom{|E|}{t} \mathop{\mathbf{E}}_{F}[2^{-n} p^t 2^{-t(n-1)+\Lambda(F)} (np)^{\tau t}] \\
&\leq k^t 2^{tn} \times 2^{-n} 2^{-t(n-1)} (np)^{\tau t} p^t \mathop{\mathbf{E}}_{F}[2^{\Lambda(F)}] \\
&= 2^{-n} k^t (np)^{\tau t} (2p)^t \mathcal{I}_t(E).
\end{aligned}$$

Since $\mathcal{I}_t(E) \leq n^{\epsilon t - 3}$, $p = (2 \ln k)/n$, $k \leq n^{(1-\epsilon)t}$ and $t, \tau \leq (\ln n)^{1/3}$, we have for all large n

$$\begin{aligned}
(S2) &\leq 2^{-n} n^{(1-\epsilon)t} (2 \ln n)^{t^2} (\frac{4 \ln n}{n})^t n^{\epsilon t - 3} \\
&\leq \frac{2^{-n}}{n^2}.
\end{aligned}$$

Thus,

$$\Pr[\mathsf{Case1}(f)] \leq (S1) + (S2) \leq 2^{-n} \times [\frac{1}{k^{3/2}} + \frac{1}{n^2}],$$

and, considering all possible f,

$$\Pr[\exists f \, \mathsf{Case1}(f)] \leq \frac{2}{\sqrt{k}}.$$

Since $k = n^{1-\epsilon}$, for large n this quantity is negligible; we have thus established (15).

5.4.3 Case 2

Notation. In the following, let $(\ell)_r = \ell(\ell-1)\cdots(\ell-r+1)$.

Suppose Case 2 holds for f. As discussed before, since f is not almost-blue in χ_0, it must have had at least $\tau+1$ red vertices in χ_0. During the recoloring phase all red (in χ_0) vertices of f changed their color; hence, for each $v \in S$, there must be an edge $f_v \in E$, such that $v \in f_v \cap f$ and f_v was red in χ_0 and continued to be so when v was considered for recoloring; we say that f_v was responsible for making v change its color. (Note that an edge can be held responsible for at most one vertex.) Let T be a random $(\tau+1)$-sized subset of the red vertices of f; then f *blames* the set of edges $F = \{f_v : v \in T\}$ for making it blue. That is, for each set f with at least $\tau+1$ red vertices in χ_0, we pick a random subset $T = T(f)$ of those vertices; the sets $T(f)$ are picked *independently* for each such edge f. Thus,

$$\Pr[\mathsf{Case2}(f)] \le \sum_{F \in \binom{E}{\tau+1}} \Pr[f \text{ blames } F].$$

We have already considered the situation when f was completely red in χ_0 (this was Case 0); so in the sum above, we have a contribution only for F's that do not contain f. We will now estimate $\Pr[f \text{ blames } F]$ for fixed f and F, where $f \notin F$. Let $U = \bigcup_{f' \in F} f'$. Consider the sequence of events that resulted in f blaming F.

1. All of U was colored red in χ_0.
2. Let R be the set of red vertices of f. Then, $\forall w \in R$, $b(w) = 1$. Also, $U \cap f \subseteq R$; let $R_1 = U \cap f$ and $R_2 = R - R_1$. Let $r_1 = |R_1|$ and $r_2 = |R_2|$.
3. Let $T' = \{v \in R : f_v \in F\}$; please see a few lines above for the definition of the f_v. Then, for f to blame F, T' must coincide with the randomly chosen $(\tau+1)$-sized subset T of R.

Using these observations, we get

$$\begin{aligned}
\Pr[f \text{ blames } F] &\le 2^{-|U|} \times \sum_{r_2=0}^{n-r_1} \binom{n-r_1}{r_2} 2^{-|f-U|} p^{r_1+r_2} \binom{r_1+r_2}{\tau+1}^{-1} \\
&= 2^{-|U \cup f|} \times \sum_{r_2=0}^{n-r_1} \binom{n-r_1}{r_2} p^{r_1+r_2} \binom{r_1+r_2}{\tau+1}^{-1}.
\end{aligned}$$

Note that $r_1 \ge \tau+1$, and that the sum above is a decreasing function of r_1. Thus,

$$\begin{aligned}
\Pr[f \text{ blames } F] &\le 2^{-|U \cup f|} \times \sum_{r_2=0}^{n-(\tau+1)} \binom{n-(\tau+1)}{r_2} p^{(\tau+1)+r_2} \binom{(\tau+1)+r_2}{\tau+1}^{-1} \\
&\le 2^{-(\tau+2)n+\Lambda(F\cup\{f\})} \times \sum_{r_2=0}^{n-(\tau+1)} \binom{n-(\tau+1)}{r_2} p^{\tau+1+r_2} \binom{(\tau+1)+r_2}{\tau+1}^{-1} \\
&\le 2^{-(\tau+2)n+\Lambda(F\cup\{f\})} (1+p)^n \binom{n}{\tau+1}^{-1}. \qquad (23)
\end{aligned}$$

For the last inequality, we used the fact that for any $x \geq 0$,

$$\sum_{i=0}^{q} \binom{q}{i} \frac{x^{i+r}}{(i+r)_r} = \frac{1}{(q+r)_r} \sum_{i=0}^{q} \binom{q+r}{i+r} x^{i+r} \leq \frac{1}{(q+r)_r} \sum_{i=0}^{q+r} \binom{q+r}{i} x^i = \frac{(1+x)^{q+r}}{(q+r)_r}.$$

(To derive (23) from this, set $x = p$, $q = n - (\tau+1)$ and $r = \tau + 1$.) Summing over all $f \in E$ and $F \in \binom{E}{\tau+1}$ $(f \notin F)$, we obtain

$$\begin{aligned}
\Pr[\exists f\, \mathsf{Case2}(f)] &\leq |E| \binom{|E|-1}{\tau+1} 2^{-(\tau+2)n} (1+p)^n \binom{n}{\tau+1}^{-1} \mathop{\mathbf{E}}_{F' \in \binom{E}{\tau+2}} [2^{\Lambda(F')}] \\
&\leq k2^n \binom{k2^n}{\tau+1} 2^{-(\tau+2)n} (1+p)^n \binom{n}{\tau+1}^{-1} \mathcal{I}_{\tau+2}(E) \\
&\leq \frac{k^{\tau+4}}{(n)_{\tau+1}} \mathcal{I}_{\tau+2}(E) \qquad (24) \\
&= \frac{k^{t+2}}{(n)_{t-1}} \mathcal{I}_t(E). \qquad (25)
\end{aligned}$$

To get (24) above, we substituted $p = (2 \ln k)/n$. Since $t \leq (\ln n)^{1/3}$, $k = n^{1-\epsilon}$, $\mathcal{I}_t(E) \leq n^{\epsilon t - 3}$ and n is large, the right hand side of (25) is less than $n^{-\epsilon}$. This establishes (16), and completes the proof of Theorem 5.2.

5.5 A local version

We now use Theorem 4.1 with a modification of the analysis of Section 5.4, in order to derive a "local version" generalization of Theorem 5.2 for hypergraphs with small edge-intersections bounded by a constant a; recall that for such hypergraphs, we have the inequality (7).

Theorem 5.3 *For any fixed (a, ϵ) with $\epsilon > 0$, let $\{G_n\}$ be any family of uniform hypergraphs such that in G_n, any two distinct edges intersect in at most a vertices. Suppose further that G_n has overlap $D \leq n^{1-\epsilon}2^n$. Then, for all large enough n (i.e., $n \geq N_0(a, \epsilon)$), G_n is 2-colorable.*

Proof: Suppose $D \leq k2^n$, where $k = n^{1-\epsilon}$. We may assume that $\epsilon \leq 2/3$, for otherwise we are covered by Theorem 4.2. We will use the algorithm of Section 5.3 with $p = (2 \ln k)/n$, $t = \lceil 3/\epsilon \rceil$ and $\tau = t - 2$. As in Section 5.3, we will work with an arbitrary but fixed permutation $v_1, v_2, \ldots$ of the vertices, and consider the vertices in this order, when attempting to recolor.

As in Section 5.4, the analysis of the event "f is blue in χ^*" splits into the same three cases: Case 0, Case 1 and Case 2.

The idea once again is to define a collection of "bad" events and show that Theorem 4.1 applies to them. The following bad events, defined for each allowed choice of F, capture the event that f is blue in χ^*:

- $Z_1(f)$: $R(f, \chi_0) \wedge \mathsf{Pivots}(f) = f$. This covers the event $\mathsf{Case0}(f)$.
- $Z_2(f)$: $B(f, \chi_0) \wedge |\mathsf{Pivots}(f)| \leq t - 1$. This corresponds to the subcase of $\mathsf{Case1}(f)$ whose probability was bounded using (S1) in the previous section.

- $Z_3(f,F)$: $B(f,\chi_0)$, and F conspires against f. (Here $F \in \binom{E-\{f\}}{t}$.) This corresponds to the subcase of $\mathsf{Case1}(f)$, whose probability was bounded using (S2) in the previous section.

- $Z_4(f,F)$: f blames F. (Here $F \in \binom{E-\{f\}}{\tau+1}$.) This covers $\mathsf{Case2}(f)$.

Similarly, we define analogous events $Z_1'(f)$, $Z_2'(f)$, $Z_3'(f,F)$ and $Z_4'(f,F)$ for the case of f becoming red in χ^*.

Clearly, for each edge f,

$$\Pr[Z_1(f)] = \Pr[Z_1'(f)] \le 2^{-n}p^n. \tag{26}$$

Now, for events Z_2 and Z_2', (18) gives

$$\Pr[Z_2(f)] = \Pr[Z_2'(f)] \le 2^{-n}(1-p)^{n-t}(np)^t. \tag{27}$$

For Z_3 and Z_3', we have from (22), that

$$\Pr[Z_3(f,F)] = \Pr[Z_3'(f,F)] \le 2^{-n}p^t 2^{-t(n-1)+\Lambda(F)}(np)^{\tau t} \le 2^{-n}p^t 2^{-t(n-1)+t(t-1)a/2}(np)^{\tau t}. \tag{28}$$

Finally, for Z_4, (23) gives

$$\Pr[Z_4(f,F)] = \Pr[Z_4'(f,F)] \le 2^{-(\tau+2)n+(\tau+2)(\tau+1)a/2}e^{np}\binom{n}{\tau+1}^{-1}. \tag{29}$$

The main observation once again is that there is not much dependence between these events. All the events above, except $Z_4(f,F)$ and $Z_4'(f,F)$, depend only on the χ_0 and b values of the vertices contained in either f or one of the elements of F; in the case of $F_4(f,F)$ and $F_4'(f,F)$, we need to know, in addition, the value of the random set $T(f)$ to determine if these events hold (see conditions 3 in Case 2 of the previous section). Thus, we can proceed as in Section 4. To illustrate this, we first note that each of the events $\mathcal{E}$ above, has at most $\ell = 1+t = \mathrm{O}(1/\epsilon)$ edges as its arguments. Clearly, $\mathcal{E}$ depends on at most $2D\ell$ events of type Z_1, Z_1', and on at most $2D\ell$ events of type Z_2, Z_2'. How many events of type $Z_3(f,F)$ does $\mathcal{E}$ depend upon? Some argument f' of $\mathcal{E}$ must intersect either f or some element of F. There are at most ℓ choices for f'; fix f'. If f' intersects f, then there are at most D choices for f and once we fix f, there are at most $\binom{D}{t}$ choices for F. Otherwise, f' intersects some element f'' of F: we can first choose f'' (at most D choices), then choose f (at most D choices), and finally choose the remaining $t-1$ elements of F (at most D^{t-1} choices). So, $\mathcal{E}$ depends on at most $\mathrm{O}(\ell D^{t+1})$ events of type Z_3; similarly, $\mathcal{E}$ depends on at most $\mathrm{O}(\ell D^{\tau+2})$ events of type Z_4.

Recall that $k = n^{1-\epsilon}$, $D \le k2^n$, $p = (2\ln k)/n$, $t = \lceil 3/\epsilon \rceil$, $\tau = t-2$, and $\ell \le \mathrm{O}(1/\epsilon)$. With these values, one sees that

$$D\ell 2^{-n}(p^n + (1-p)^{n-t}(np)^t) + \ell D^{t+1}2^{-n}p^t 2^{-t(n-1)+t(t-1)a/2}(np)^{\tau t} + \frac{\ell D^{\tau+2}2^{-(\tau+2)n+(\tau+2)(\tau+1)a/2}e^{np}}{\binom{n}{\tau+1}}$$

goes to 0 as n increases. Thus, bounds (26), (29), (27) and (28), in conjunction with Theorem 4.1, complete the proof. ∎

Recall the definition of $m^*(n)$ from the first paragraph of Section 5. As an application of Theorem 5.3 we obtain the following corollary, giving a different proof of Szabó's result [28] that $m^*(n) \ge 4^n/n^{1+\epsilon}$.

Corollary 5.1 *For any fixed $\epsilon > 0$ and all sufficiently large n, $m^*(n) \geq 4^n/n^{1+\epsilon}$.*

Proof: The following proposition was shown by Erdős and Lovász [12].

Proposition 5.1 ([12]) *Suppose every* simple *t-uniform hypergraph in which each vertex lies in at most $h(t)$ edges is 2-colorable. Then, $m^*(n) \geq (h(n-1))^2/n$.*

We reproduce the proof of [12] below, but before that let us see why this implies the corollary. By Theorem 5.3, for any constant $\epsilon > 0$ and for all sufficiently large n, $h(n) \geq 2^n/n^\epsilon$. On substituting this in the proposition above, we obtain $m^*(n) \geq \Omega(4^n/n^{1+2\epsilon})$. This implies the corollary.

Proof of Proposition 5.1. Let H be an n-uniform simple hypergraph which is not 2-colorable. We wish to show that H must have many edges. Since H is n-uniform,

$$|E(H)| \geq \frac{1}{n} \sum_{v \in V(H)} d_H(v). \tag{30}$$

Thus, it suffices to show that H has many vertices of high degree. For each $e \in E(H)$, let v_e be a vertex in e with maximum degree; that is, $d_H(v_e) \geq d_H(w)$ for all $w \in e$. If there are several such maximum-degree vertices v_e, choose one arbitrarily.

Consider the following hypergraph H', with $V(H) = V(H')$ and

$$E(H') = \{e - \{v_e\} : e \in E(H)\}.$$

Clearly, H' is $(n-1)$-uniform and simple. Since H is not 2-colorable, H' is not 2-colorable. By the hypothesis of the proposition, there is a vertex v in H' with $d_{H'}(v) \geq h(n-1)$. Let $e'_1, e'_2, \ldots, e'_h$ ($h \geq h(n-1)$) be the edges of H' incident on v, and let $e_1, e_2, \ldots, e_h$ be the corresponding edges of H. Now, by the definition of H', each e_i has a vertex v_i different from v, such that $d_H(v_i) \geq d_H(v)$. Since H is simple, $v_i \neq v_j$ for $i \neq j$. We have thus obtained h distinct vertices each of degree at least h. Since $h \geq h(n-1)$, the proposition follows from (30). ■

Acknowledgments. We thank Noga Alon, Ravi Boppana, Jeong-Han Kim and Joel Spencer for their encouragement and several valuable suggestions, which have substantially improved the quality of this work. We also thank József Beck, Devdatt Dubhashi, Alexander Russell and Benny Sudakov for their helpful comments. The exposition of the Beck-Spencer approach in [4, 27] provided much of the inspiration for this work.

We thank the referees for several suggestions; in particular, the analysis of Boppana's algorithm, which is included in Section 2 of this version, was shown to us by one of the referees.

References

[1] N. Alon. A parallel algorithmic version of the Local Lemma. *Random Structures & Algorithms*, 2:367–378, 1991.

[2] N. Alon and Z. Bregman. Every 8-uniform 8-regular hypergraph is 2-colorable. *Graphs and Combinatorics*, 4:303–305, 1988.

[3] N. Alon, P. Kelsen, S. Mahajan, and H. Ramesh. Approximate hypergraph coloring. *Nordic Journal on Computing*, 3:425–439, 1996.

[4] N. Alon and J. H. Spencer. *The Probabilistic Method.* Wiley–Interscience Series, John Wiley & Sons, Inc., New York, 1992.

[5] G. Andersson and L. Engebretsen. Better approximation algorithms for SET SPLITTING and NOT-ALL-EQUAL SAT. *Inform. Proc. Lett.*, 65:305–311, 1998.

[6] J. Beck. On 3-chromatic hypergraphs. *Discrete Mathematics*, 24:127–137, 1978.

[7] J. Beck. An algorithmic approach to the Lovász Local Lemma. *Random Structures & Algorithms*, 2:343–365, 1991.

[8] F. Bernstein. Zur Theorie der trigonometrische Reihen. *Leipz. Ber.*, 60:325–328, 1908.

[9] H. Chen and A. Frieze. Coloring bipartite hypergraphs. In *Proc. MPS Conference On Integer Programming and Combinatorial Optimization*, Lecture Notes in Computer Science **1084**, Springer-Verlag, pages 345–358, 1996.

[10] P. Erdős. On a combinatorial problem, I. *Nordisk Mat. Tidskrift*, 11:5–10, 1963.

[11] P. Erdős. On a combinatorial problem, II. *Acta Mathematica of the Academy of Sciences*, Hungary, 15:445–447, 1964.

[12] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Infinite and Finite Sets*, A. Hajnal et. al., editors, Colloq. Math. Soc. J. Bolyai 11, North Holland, Amsterdam, pages 609–627, 1975.

[13] G. Even, O. Goldreich, M. Luby, N. Nisan, and B. Veličković. Approximations of general independent distributions. In *Proc. ACM Symposium on Theory of Computing*, pages 10–16, 1992.

[14] G. Even, O. Goldreich, M. Luby, N. Nisan, and B. Veličković. Efficient approximations for product distributions. *Random Structures & Algorithms*, 13: 1–16, 1998.

[15] C. M. Fortuin, J. Ginibre, and P. N. Kasteleyn. Correlational inequalities for partially ordered sets. *Communications of Mathematical Physics*, 22:89–103, 1971.

[16] M. R. Garey and D. S. Johnson. Computers and intractability: a guide to the theory of NP-completeness. W. H. Freeman and Company, New York, 1978 (printing updated in 1991).

[17] T. R. Jensen and B. Toft. Graph coloring problems. Wiley–Interscience Series, John Wiley & Sons, Inc., 1995.

[18] D. R. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *Journal of the ACM*, 45:246–265, 1998.

[19] M. Krivelevich and B. Sudakov. Approximate coloring of uniform hypergraphs. In *Proc. European Symposium on Algorithms*, pages 477–489, 1998.

[20] L. Lovász. Covering and coloring of hypergraphs. In *Proc. 4th Southeastern Conference on Combinatorics, Graph Theory and Computing*, pages 3–12, Utilitas Mathematica Publishing, Winnipeg, 1973.

[21] C. McDiarmid. A random recolouring method for graphs and hypergraphs. *Combinatorics, Probability and Computing*, 2:363–365, 1993.

[22] C. McDiarmid. Hypergraph colouring and the Lovász Local Lemma. *Discrete Mathematics*, 167/168:481–486, 1997.

[23] E. W. Miller. On a property of families of sets. *C. R. Soc. Sci. Varsovie*, 30:31–38, 1937.

[24] M. Molloy and B. Reed. Further algorithmic aspects of the Local Lemma. In *Proc. ACM Symposium on Theory of Computing*, pages 524–529, 1998.

[25] J. Naor and M. Naor. Small–bias probability spaces: efficient constructions and applications. *SIAM J. Comput.*, 22(4):838–856, 1993.

[26] J. H. Spencer. Coloring n-sets red and blue. *J. Combinatorial Theory, Series A*, 30:112–113, 1981.

[27] J. H. Spencer. *Ten Lectures on the Probabilistic Method.* SIAM, Philadelphia, 1987. (Second edition published in 1994.)

[28] Z. Szabó. An application of Lovász' Local Lemma–a new lower bound for the van der Waerden number. *Random Structures & Algorithms*, 1:343–360, 1990.

[29] C. Thomassen. The even cycle problem for directed graphs. *Journal of the American Mathematical Society*, 5:217–229, 1992.

A Appendix

Exploiting positive correlation. Let $\mathcal{A}(f)$ be the event $\mathcal{A}_{Blue}(f) \vee \mathcal{A}_{Red}(f)$. Then,

$$\mathcal{A}(f) \equiv \mathcal{M}(f, \chi_0) \wedge (\forall v \in f : b(v) = 0).$$

Claim A.1 $\Pr[\exists f \in E : \mathcal{A}(f)] \leq 1 - (1 - (1-p)^n)^{2k}$.

Proof: We have

$$\begin{aligned} \Pr[\exists f \in E :\ \mathcal{A}(f)] &= 2^{-|V|} \sum_{\chi: V \to \{\text{Red,Blue}\}} \Pr[\exists f \in \mathcal{M}(\chi)\ \forall v \in f :\ b(v) = 0] \\ &= 2^{-|V|} \sum_{\chi: V \to \{\text{Red,Blue}\}} (1 - \Pr[\forall f \in \mathcal{M}(\chi)\ \exists v \in f :\ b(v) = 1]). \quad (31) \end{aligned}$$

Fix $\chi : V \to \{\text{Red, Blue}\}$. It is easy to check using the FKG inequality [15] that we have *positive correlation*:

$$\begin{aligned} \Pr[\forall f \in \mathcal{M}(\chi)\ \exists v \in f :\ b(v) = 1] &\geq \prod_{f \in \mathcal{M}(\chi)} \Pr[\exists v \in f :\ b(v) = 1] \\ &= (1 - (1-p)^n)^{|\mathcal{M}(\chi)|}. \end{aligned}$$

Thus, by (31),

$$\begin{aligned}
\Pr[\exists f \in E :\ \mathcal{A}(f)] &\leq 2^{-|V|} \sum_{\chi : V \to \{\text{Red,Blue}\}} (1-(1-(1-p)^n))^{|\mathcal{M}(\chi)|} \\
&= 1 - 2^{-|V|} \sum_{\chi : V \to \{\text{Red,Blue}\}} (1-(1-p)^n)^{|\mathcal{M}(\chi)|} \\
&= 1 - \mathop{\mathbf{E}}_{\chi_0}[(1-(1-p)^n)^{|\mathcal{M}(\chi_0)|}] \\
&\leq 1 - (1-(1-p)^n)^{\mathbf{E}[|\mathcal{M}(\chi_0)|]};
\end{aligned}$$

the last inequality follows from Jensen's inequality, since, for any fixed $a > 0$, the function $x \mapsto a^x$ is convex (as its second derivative $a^x(\ln a)^2$ is non-negative). Thus, since $\mathbf{E}[|\mathcal{M}(\chi_0)|] = 2^{-n+1} \cdot k2^n = 2k$, we have $\Pr[\exists f \in E :\ \mathcal{A}(f)] \leq 1 - (1-(1-p)^n)^{2k}$. ■