Proper Bounded Edge-Colorings (Extended Abstract)

Claudia Bertram-Kretzberg*, Hanno Lefmann**, Vojtěch Rödl***†
and Beata Wysocka‡

Abstract. For an n-element set X and a proper coloring $\Delta\colon [X]^k \longrightarrow \{0, 1, \ldots\}$ where each color class is a matching with cardinality bounded by u, we show that there exists a totally multicolored subset $Y \subseteq X$ with

$$|Y| \geq \max \left\{ c_1 \cdot \left(n^k/u\right)^{\frac{1}{2k-1}},\ c_2 \cdot \left(n^k/u\right)^{\frac{1}{2k-1}} \cdot \left(\ln \left(u/\sqrt{n}\right)\right)^{\frac{1}{2k-1}} \right\}$$

This bound is tight up to constant factors for $u = \omega(n^{1/2+\epsilon})$ for any $\epsilon > 0$. Moreover, for fixed k, we give a polynomial time algorithm for finding such a set Y of guaranteed size.

1 Introduction

On each of $\binom{3n}{3}/n$ school days, in a school attended by $3n$ students, the students are asked to line up in n rows, each containing three students. In 1851, Kirkman asked for the existence of such a schedule that would allow each triplet of students to occupy a row on exactly one of the school days, cf. [6]. This classical problem was answered completely by Ray-Chaudhuri and Wilson [10] who proved that such a schedule exists for each $n \equiv 1, 3 \bmod 6$. Here, we investigate a somewhat related combinatorial problem. Suppose that after such a schedule was prepared, the principle of the school wants (for unrevealed purposes) to select the largest group of, say, m students with the property that no two triplets of students

* Universität Dortmund, Fachbereich Informatik, LS II, D-44221 Dortmund, Germany. bertram@ls2.informatik.uni-dortmund.de
** Universität Dortmund, Fachbereich Informatik, LS II, D-44221 Dortmund, Germany. lefmann@ls2.informatik.uni-dortmund.de
*** Emory University, Department of Mathematics and Computer Science, Atlanta, Georgia 30322, USA. rodl@mathcs.emory.edu
† Research supported by NSF grant DMS 9401559. Part of this work was done during the author's visit of Humboldt-Universität, Berlin, with a Humboldt senior-fellowship.
‡ University of North Carolina at Greensboro, Department of Mathematical Sciences, Greensboro, NC 27412, USA. beata@maths.mu.oz.au

occupy a row at the same day. Such an m must satisfy $(*)$ $c_1 \cdot n^{1/3} \cdot (\log n)^{1/3} \leq m \leq c_2 \cdot n^{2/3}$ for any schedule. While the upper bound is straightforward, the lower bound follows from [2]. Here, we give a polynomial time algorithm which finds a group of m students satisfying the lower bound in $(*)$. Moreover, there are schedules which, up to a constant factor, are the best possible. We consider the general case in which one has n students which are asked to line up in at most u rows, each containing k people. We extend earlier results from [2] and [11] where the case $u = n/k$ respectively $k = 2$ was considered.
We formulate our problem in terms of edge-colored hypergraphs: vertices correspond to students, edges to rows, and the edges are colored by the day.

Definition 1. *Let $\Delta\colon [X]^k \longrightarrow \omega$ where $\omega = \{0, 1, \ldots\}$ be a coloring of the k-element subsets of X. The coloring $\Delta\colon [X]^k \longrightarrow \omega$ with color classes $C_0, C_1, \ldots$, i.e., $\Delta^{-1}(i) = C_i$ for $i \in \omega$, is called* u-bounded *if $|C_i| \leq u$ for $i = 0, 1, \ldots$. The coloring $\Delta\colon [X]^k \longrightarrow \omega$ is called* proper *if each color class C_i, $i = 0, 1, \ldots$, is a matching, i.e., sets of the same color are pairwise disjoint, thus, $\Delta(U) = \Delta(V)$ implies $U \cap V = \emptyset$ for all distinct sets $U, V \in [X]^k$. A subset $Y \subseteq X$ is called* totally multicolored *if the restriction of the coloring Δ to the set $[Y]^k$ is a one-to-one coloring. For an n-element set X, define, minimizing over all proper u-bounded colorings $\Delta\colon [X]^k \longrightarrow \omega$, the following function*

$$f_u(n,k) = min_\Delta max\{|Y| \; ; \; Y \subseteq X \text{ is totally multicolored}\} \; .$$

The first estimates on $f_u(n,k)$ were given by Babai [4], in connection with some Sidon-type problem. He showed for the case $u = n/2$ and $k = 2$ that $c_1 \cdot n^{1/3} \leq f_{n/2}(n,2) \leq c_2 \cdot (n \cdot \ln n)^{1/3}$. In [2], the lower bound was improved by the factor $O((\ln n)^{1/3})$. Here, we will show the following:

Theorem 1. *Let $k, u \geq 2$ be fixed integers. There exist positive constants c_1, c_2, c_3 such that for n large enough,*

$$\max\left\{c_1 \cdot \left(n^k/u\right)^{1/(2k-1)}, \; c_2 \cdot \left(n^k/u\right)^{1/(2k-1)} \cdot \left(\ln\left(u/\sqrt{n}\right)\right)^{1/(2k-1)}\right\}$$
$$\leq f_u(n,k) \leq c_3 \cdot \left(n^k/u\right)^{1/(2k-1)} \cdot (\ln n)^{1/(2k-1)} \; . \qquad (1)$$

Moreover, for every u-bounded proper coloring $\Delta\colon [X]^k \longrightarrow \omega$ *with* $|X| = n$*, one can find in time* $O(u \cdot n^{2k-1})$ *a totally multicolored subset* $Y \subseteq X$ *with* $|Y| \geq \max\left\{c_1 \cdot \left(n^k/u\right)^{1/(2k-1)},\ c_2 \cdot \left(n^k/u\right)^{1/(2k-1)} \cdot (\ln(u/\sqrt{n}))^{1/(2k-1)}\right\}$.

2 The Existence

Let $\mathcal{G} = (V, \mathcal{E})$ be a hypergraph with vertex set V and edge set $\mathcal{E}$. For a vertex $v \in V$, let $d(v)$ denote the *degree* of v in $\mathcal{G}$, i.e., the number of edges $E \in \mathcal{E}$ containing v. Let $d = \sum_{v \in V} d(v)/|\mathcal{V}|$ denote the average degree of $\mathcal{G}$. The hypergraph $\mathcal{G}$ is called k*-uniform* if $|E| = k$ for each edge $E \in \mathcal{E}$. A 2*-cycle* in $\mathcal{G}$ is a pair $E, E' \in \mathcal{E}$ of distinct edges which intersect in at least two vertices. The *independence number* $\alpha(\mathcal{G})$ is the largest size of a subset $I \subseteq V$ such that the induced hypergraph contains no edges.
Here, we will prove inequality (1) of Theorem 1. Some of our arguments are based on a result of Ajtai, Komlós, Pintz, Spencer and Szemerédi, [1]. We use a modified version proved in [7], cf. [2] and [12].

Theorem 2. *Let* $k \geq 2$*. Let* $\mathcal{G}$ *be a* $(k+1)$*-uniform hypergraph on* n *vertices. If* (*i*) $\mathcal{G}$ *contains no* 2*-cycles, and* (*ii*) *the average degree satisfies* $d \leq t^k$ *where* $t \geq t_0(k)$*, then for some positive constant* $c = c(k)$*,*

$$\alpha(\mathcal{G}) \geq c \cdot n/t \cdot (\ln\ t)^{1/k}\ . \tag{2}$$

Proof. We start by showing the lower bounds in (1). Let $\Delta\colon [X]^k \to \omega$ be a u-bounded proper coloring where $|X| = n$. We construct a $2k$-uniform hypergraph $\mathcal{H} = (X, \mathcal{E})$ on X where $U \in \mathcal{E} \subseteq [X]^{2k}$ if there exist two distinct sets $S, T \in [X]^k$, $S, T \subseteq U$ so that $\Delta(S) = \Delta(T)$. As Δ is u-bounded, we infer

$$|\mathcal{E}| = \sum_{i \in \omega} \binom{|\Delta^{-1}(i)|}{2} \leq \frac{\binom{n}{k}}{u} \cdot \binom{u}{2}\ . \tag{3}$$

If $I \subseteq X$ is an independent set of $\mathcal{H}$, then I is totally multicolored w.r.t. the coloring Δ. Hence, it is enough to show that $\mathcal{H}$ contains an independent set of size $c_1 \cdot (n^k/u)^{\frac{1}{2k-1}}$. This follows by an easy probabilistic argument, i.e., choose every vertex in X independently of the other vertices with probability

$$p = (n^{k-1} \cdot u)^{-1/(2k-1)} . \tag{4}$$

By Chernoff's and Markov's inequality, we know that there exists a subset $Y \subseteq X$ with $|Y| \sim \left(n^k/u\right)^{1/(2k-1)}$, and the subhypergraph induced on Y contains at most

$$2 \cdot p^{2k} \cdot |\mathcal{E}| \le 2 \cdot p^{2k} \cdot \frac{\binom{n}{k}}{u} \cdot \binom{u}{2} \le \frac{1}{2} \cdot \left(\frac{n^k}{u}\right)^{1/(2k-1)}$$

edges. We delete one vertex from each edge in $[Y]^{2k} \cap \mathcal{E}$, and we obtain a subset $Y' \subseteq Y$ with $|Y'| \ge |Y|/2 \ge pn/2$. Then, Y' is an independent set in $\mathcal{H}$, hence Y' is totally multicolored w.r.t. $\varDelta$.

If $u = \sqrt{n} \cdot \omega(n)$, where $\omega(n) \longrightarrow \infty$ with $n \longrightarrow \infty$, we can improve this lower bound by a logarithmic factor. Let $\varDelta \colon [X]^k \to \omega$ be a u-bounded proper coloring. Consider the $2k$-uniform hypergraph $\mathcal{H} = (V, \mathcal{E})$ defined in the same way as above. Again, we want to find a large independent set in $\mathcal{H}$. The strategy is to find a random subset $Y \subseteq X$ such that the induced hypergraph has only a few 2-cycles. By deleting these 2-cycles, the desired result will follow from (2).

The number of edges of $\mathcal{H}$ satisfies inequality (3). With forsight we use a slightly larger value than in (4) for the probability p of picking vertices, namely,

$$p = \left(1/(n^{k-1} \cdot u)\right)^{1/(2k-1)} \cdot \left(u/\sqrt{n}\right)^{1/((k+1)(2k-1))} .$$

Let Y be a random subset of X obtained by choosing vertices $v \in X$ with probability p independently of the others. The expected size of Y is $E(|Y|) = p{\cdot}n$. Let $\nu_j(Y)$, for $j = 2, \ldots, 2k-1$, be random variables counting the number of $(2,j)$-cycles, i.e. the number of pairs of edges in the subhypergraph of $\mathcal{H}$ induced on Y which intersect in exactly j vertices. The random variable $\mu_2(Y) = \sum_{j=2}^{2k-1} \nu_j(Y)$ counts the total number of 2-cycles of the subhypergraph induced on Y. We will give upper bounds on the expected values $E(\nu_j(Y))$. To do so, we estimate the total number ν_j of $(2,j)$-cycles in $\mathcal{H}$. Fix an edge $E \in \mathcal{E}$. The number of pairs of distinct sets $U, V \in [X]^k$ with $\varDelta(U) = \varDelta(V)$ and $|(U \cup V) \cap E| = j$

and $1 \leq |U \cap E|, |V \cap E| \leq j-1$ is at most

$$\sum_{i=\lceil j/2 \rceil}^{j-1} \binom{2k}{i} \cdot \binom{n-2k}{k-i} \cdot \binom{2k-i}{j-i} \leq c_1 \cdot n^{k-\lceil j/2 \rceil} , \tag{5}$$

as either $|U \cap E| \geq \lceil j/2 \rceil$ or $|V \cap E| \geq \lceil j/2 \rceil$, and every color class is a matching. If $U \cap E = \emptyset$ or $V \cap E = \emptyset$, but $|(U \cup V) \cap E| = j$, then the number of such pairs U, V is bounded from above by

$$\binom{2k}{j} \cdot \binom{n-2k}{k-j} \cdot (u-1) \leq c_2 \cdot n^{k-j} \cdot u . \tag{6}$$

Now, (3), (5) and (6) imply that

$$\nu_j \leq |\mathcal{E}| \cdot \left(c_1 \cdot n^{k-\lceil j/2 \rceil} + c_2 \cdot n^{k-j} \cdot u \right) \leq c_3 \cdot u \cdot \left(n^{2k-\lceil j/2 \rceil} + n^{2k-j} \cdot u \right) . \tag{7}$$

As every color class is a matching, we have $u \leq n/k$, thus, $n^{2k-\lceil j/2 \rceil} \geq n^{2k-j} \cdot u$ for $j \geq 2$, and (7) becomes

$$\nu_j \leq c_4 \cdot u \cdot n^{2k-\lceil j/2 \rceil} . \tag{8}$$

We infer for $j = 2, \ldots, 2k-1$ that

$$\begin{aligned} E(\nu_j(Y)) &\leq p^{4k-j} \cdot c_4 \cdot u \cdot n^{2k-\lceil j/2 \rceil} = \\ &= pn \cdot c_4 \cdot u^{\frac{j-2k+\frac{1}{k+1}(4k-j-1)}{2k-1}} \cdot n^{\frac{k(j+1-2\lceil j/2 \rceil)-\lfloor j/2 \rfloor - \frac{1}{2(k+1)}(4k-j-1)}{2k-1}} . \end{aligned}$$

As $u = \sqrt{n} \cdot \omega(n) \leq n/k$, we have $\omega(n) = O(\sqrt{n})$, and hence, for j odd,

$$\begin{aligned} E(\nu_j(Y)) &\leq pn \cdot c_4 \cdot u^{\frac{j-2k+\frac{1}{k+1}(4k-j-1)}{2k-1}} \cdot n^{\frac{-(j-1)/2-\frac{1}{2(k+1)}(4k-j-1)}{2k-1}} \\ &= pn \cdot c_4 \cdot \omega(n)^{\frac{j-2k+\frac{1}{k+1}(4k-j-1)}{2k-1}} \cdot n^{\frac{-k+1/2}{2k-1}} \\ &\leq pn \cdot c_4 \cdot \omega(n)^{\frac{k-1}{(k+1)(2k-1)}} \cdot n^{\frac{-k+1/2}{2k-1}} \qquad \text{as } j \leq 2k-1 \\ &= o(pn) . \end{aligned} \tag{9}$$

Similarly, for j even, we obtain

$$E(\nu_j(Y)) = o(pn) . \tag{10}$$

By (9) and (10), we infer $E(\mu_2(Y)) = \sum_{j=2}^{2k-1} E(\nu_j(Y)) = o(pn)$. Thus, there exists a subset $Y \subseteq X$ with $|Y| = c_5 pn$ such that the induced hypergraph contains at most $c_6 p^{2k}|\mathcal{E}|$ edges and has only $o(pn)$ 2-cycles. We omit one vertex from each 2-cycle in $\mathcal{H}_0$. The remaining subhypergraph $\mathcal{H}_1$ has $(c_5 - o(1)) \cdot pn$ vertices and by (3), the average degree d^{2k-1} satisfies $d \leq c_8 \cdot (u/\sqrt{n})^{\frac{1}{(k+1)(2k-1)}}$. As $u/\sqrt{n} \longrightarrow \infty$ with $n \longrightarrow \infty$, we obtain from (2) that

$$\begin{aligned}
\alpha(\mathcal{H}) \geq \alpha(\mathcal{H}_1) &\geq c \cdot \frac{(c_1 - o(1)) \cdot p \cdot n}{c_8 \cdot (u/\sqrt{n})^{\frac{1}{(k+1)(2k-1)}}} \cdot \left[\ln \left(c_8 \cdot \left(\frac{u}{\sqrt{n}} \right)^{\frac{1}{(k+1)(2k-1)}} \right) \right]^{\frac{1}{2k-1}} \\
&\geq c' \cdot \left(n^k/u\right)^{1/(2k-1)} \cdot \left(\ln\left(u/\sqrt{n}\right)\right)^{1/(2k-1)} . \qquad \square
\end{aligned}$$

Next, we will show the upper bound in (1), extending some ideas from [4]. Let X be an n-element set where w.l.o.g. n is divisible by k. Set $m = \lceil c \cdot n^k/u \rceil$, where $c > 0$ is a constant. Let $M_1, \ldots, M_m$ be random matchings, chosen uniformly and independently from the set of all matchings of size u from $[X]^k$, and set $H_j = \bigcup_{i<j} M_i$. We define a coloring $\Delta\colon [X]^k \to \omega$ as follows: for $j = 1, \ldots, m$, color all sets in $M_j \backslash H_j$ by color j, and color all remaining elements in $[X]^k \backslash H_{m+1}$ in an arbitrary way, such that each color class is a matching. Let $Y \subseteq X$ be a fixed subset with $|Y| = x$ where $x = o(n/u^{1/k})$. We will prove that for $x \geq c_3 \cdot \left(n^k/u \cdot \ln n\right)^{1/(2k-1)}$ with probability approaching to 1 any such set Y is not totally multicolored, where $c_3 > 0$ is an appropriate constant. This will give the desired result. We split the proof into several claims.

Claim. For $j = 1, \ldots, m$ and $t = 1, 2, \ldots$,

$$\text{Prob } \left[|M_j \cap [Y]^k| \geq t \right] \leq \left(u \cdot x^k/n^k\right)^t . \tag{11}$$

Proof. The left hand side of (11) does not depend on the particular choice of Y. Thus, assume that the matching M_j is fixed. The set Y can be chosen in $\binom{n}{x}$ ways. From M_j we can choose t edges in $\binom{u}{t}$ ways, and the remaining elements of Y can be chosen in at most $\binom{n-kt}{x-kt}$ ways, hence

$$\text{Prob } \left[|M_j \cap [Y]^k| \geq t \right] \leq \binom{u}{t} \cdot \binom{n-kt}{x-kt} \Big/ \binom{n}{x} \leq \left(u \cdot x^k/n^k\right)^t . \qquad \square$$

Claim. For $t = 1, 2, \ldots$ and for large enough integers n,

$$\text{Prob } \left[|H_{m+1} \cap [Y]^k| \geq t \right] \leq \left(\frac{e \cdot (t+m) \cdot u \cdot x^k}{t \cdot n^k} \right)^t . \tag{12}$$

Proof. For $j = 1, \ldots, m$, consider the events $|M_j \cap [Y]^k| \geq t_j$. These events are independent. By Claim 2, we have Prob $\left[| M_j \cap [Y]^k | \geq t_j\right] \leq \left(u \cdot x^k/n^k\right)^{t_j}$. Since $|H_{m+1} \cap [Y]^k| \leq \sum_{j=1}^m |M_j \cap [Y]^k|$ we infer, using $\binom{n}{k} \leq (e \cdot n/k)^k$, that

$$\text{Prob } \left[|H_{m+1} \cap [Y]^k| \geq t\right] \leq \text{Prob } \left[\sum_{j=1}^m |M_j \cap [Y]^k| \geq t\right] \leq$$

$$\leq \sum_{(t_j)_{j=1}^m, t_j \geq 0, \sum_{j=1}^m t_j = t} \prod_{j=1}^m \text{Prob } \left[|M_j \cap [Y]^k| \geq t_j\right] \leq$$

$$\leq \sum_{(t_j)_{j=1}^m, t_j \geq 0, \sum_{j=1}^m t_j = t} \prod_{j=1}^m \left(u \cdot x^k/n^k\right)^{t_j} = \binom{t+m-1}{t} \cdot \left(u \cdot x^k/n^k\right)^t \leq$$

$$\leq \left(\frac{e \cdot (t+m) \cdot u \cdot x^k}{t \cdot n^k}\right)^t . \qquad \square$$

Let E_i denote the event $|H_i \cap [Y]^k| \leq c_1 \cdot x^k$, where $c_1 > 0$ is a small constant.

Claim. For large enough positive integers n,

$$\text{Prob } [E_{m+1}] \geq 1 - 2^{-c_1 \cdot x^k} .$$

Proof. For $t = c_1 \cdot x^k$ with $x = o\left(n/u^{1/k}\right)$, we have $t = o(n^k/u)$. If n is large, $m = \lceil c \cdot n^k/u \rceil$ and $ec/c_1 \leq 1/3$, then (12) is less than $(1/2)^t$, hence,

$$\text{Prob } [E_{m+1}] \geq 1 - \text{ Prob } \left[|H_{m+1} \cap [Y]^k| \geq c_1 \cdot x^k\right] \geq 1 - 2^{c_1 \cdot x^k} . \qquad \square$$

We define another random variable $Y_j = |[M_j]^2 \cap [[Y]^k \setminus H_j]^2|$ for $j = 1, \ldots, m$.

Claim. If n is a sufficiently large positive integer, then for $j = 1, \ldots, m$,

$$E(Y_j | E_j) > c_5 \cdot u^2 \cdot x^{2k}/n^{2k} .$$

Proof. Clearly, we have Prob $[E_1] = 1$. As E_j holds, it is $|[Y]^k \setminus H_j| \geq \binom{x}{k} - c_1 \cdot x^k \geq c_2 \cdot x^k$. For each set $S \in [Y]^k$, there are less than $k \cdot \binom{x-1}{k-1}$ k-element subsets of Y which are not disjoint from S. Hence, for n large, the number of sets $\{S, T\} \in [[Y]^k \setminus H_j]^2$ with $S \cap T = \emptyset$ is at least

$$1/2 \cdot c_2 \cdot x^k \cdot \left(c_2 \cdot x^k - k \cdot \binom{x-1}{k-1}\right) > c_3 \cdot x^{2k} . \qquad (13)$$

Two disjoint k-element sets S, T, are both in M_j with probability

$$\text{Prob }[S, T \in M_j] = \frac{u \cdot (u-1)}{\binom{n}{k} \cdot \binom{n-k}{k}} \geq c_4 \cdot \frac{u^2}{n^{2k}} \ . \tag{14}$$

By (13) and (14) for the conditional expected value $E(Y_j|E_j)$, we have $E(Y_j|E_j) \geq c_5 \cdot u^2 \cdot x^{2k}/n^{2k}$. □

Claim. For $j = 1, \ldots, m$, and large positive integers n, and $0 < \epsilon \ll c_5$,

$$\text{Prob }[Y_j = 1 \mid E_j] \geq (c_5 - \epsilon) \cdot u^2 \cdot x^{2k}/n^{2k} \ .$$

Proof. For $t = 1, 2, \ldots$, we claim that

$$\text{Prob }[Y_j \geq t \mid E_j] \leq \left(u \cdot x^k/n^k\right)^{\lceil \sqrt{2t+1} \rceil} \ . \tag{15}$$

Namely, the statement $Y_j \geq t$ implies $|M_j \cap [Y]^k| \geq \lceil \sqrt{2t+1} \rceil$, hence,

$$\text{Prob }[Y_j \geq t \mid E_j] \leq \text{Prob }\left[|M_j \cap [Y]^k| \geq \lceil \sqrt{2t+1} \rceil\right] \leq \left(u \cdot x^k/n^k\right)^{\lceil \sqrt{2t+1} \rceil} \ .$$

For $i = 0, 1, \ldots$, set $p_i = \text{Prob }[Y_j = i \mid E_j]$. We infer from (15), that

$$E(Y_j \mid E_j) = \sum_{i \geq 0} i \cdot p_i \leq p_1 + \sum_{i \geq 2} i \cdot \left(u \cdot x^k/n^k\right)^{\lceil \sqrt{2i+1} \rceil} =$$
$$= p_1 + O\left(\left(u \cdot x^k/n^k\right)^3\right) = p_1 + o\left(u^2 \cdot x^{2k}/n^{2k}\right) \ ,$$

as $x = o\left(n/u^{1/k}\right)$. By Claim 2, we obtain that $p_1 \geq (c_5 - \epsilon) \cdot u^2 \cdot x^{2k}/n^{2k}$ for some positive constant $\varepsilon < c_5$ and n large enough. □

Finally, let A_j denote the event ($Y_j = 0$ and E_{j+1}).

Claim.

$$\text{Prob }[A_1 \wedge \ldots \wedge A_m] \leq \exp\left(-c' \cdot u \cdot x^{2k}/n^k\right).$$

Proof. By Claim 2, we have

$$\text{Prob }(A_1) \leq \text{Prob }(Y_1 = 0 \,|E_1) \leq \text{Prob }(Y_1 \neq 1 \,|E_1) \leq$$
$$\leq 1 - (c_5 - \epsilon) \cdot u^2 \cdot x^{2k}/n^{2k} \ , \tag{16}$$

while

$$\text{Prob}\,[A_i\,|A_1 \wedge \ldots \wedge A_{i-1}] \leq \text{Prob}\,[Y_i = 0 \mid A_1 \wedge \ldots \wedge A_{i-1}]$$
$$\leq \text{Prob}\,[Y_i \neq 1\,|A_1 \wedge \ldots \wedge A_{i-1}] \leq 1 - (c_5 - \epsilon) \cdot u^2 \cdot x^{2k}/n^{2k}\,. \tag{17}$$

With $(1-x)^m \leq exp(-m \cdot x)$ and $m = \lceil c \cdot n^k/u \rceil$, inequalities (16), (17) imply

$$\text{Prob}\,[A_1 \wedge A_2 \wedge \ldots \wedge A_m] = \text{Prob}\,[A_1] \cdot \prod_{i=2}^{m} \text{Prob}\,[A_i \mid A_1 \wedge \ldots \wedge A_{i-1}] \leq$$
$$\leq \left(1 - (c_5 - \epsilon) \cdot u^2 \cdot x^{2k}/n^{2k}\right)^m \leq \exp\left(-(c_5 - \epsilon) \cdot m \cdot u^2 \cdot x^{2k}/n^{2k}\right) \leq$$
$$\leq \exp\left(-c \cdot (c_5 - \epsilon) \cdot u \cdot x^{2k}/n^k\right)\ .$$
□

Claim. The probability that there exists a totally multicolored x-element subset is at most

$$\binom{n}{x} \cdot \left(\exp\left(-c' \cdot u \cdot x^{2k}/n^k\right) + 2^{-c_1 \cdot x^k}\right)\ . \tag{18}$$

Proof. If Y is not totally multicolored, then either $A_1 \wedge \ldots \wedge A_m$ holds or E_{m+1} fails. As there are exactly $\binom{n}{x}$ x-element sets Y, by combining the estimates from Claim 3 and Claim 6, we obtain (18). □

For $x \geq c_6 \cdot \left(n^k/u\right)^{1/(2k-1)} \cdot (\ln n)^{1/(2k-1)}$, where $c_6 > 0$ is a big enough constant, expression (18) tends to 0 with $n \longrightarrow \infty$. For $n \leq n_0$, one can obtain asymptotically the same upper bound by taking an appropriately large constant.

3 The Algorithm

In this section, we sketch the idea how to find in time $O(u \cdot n^{2k-1})$ a totally multicolored subset as guaranteed by (1). The algorithm follows the probabilistic arguments given before. It is based on recent results from [5] and [8].

Let $k \geq 2$ be fixed and let $\Delta\colon [X]^k \to \omega$ with $|X| = n$ be a proper u-bounded coloring. First, we collect sets $S \in [X]^k$ of the same color. This can be done in time $O(n^k)$. We form a $2k$-uniform hypergraph $\mathcal{H} = (X, \mathcal{E})$, $\mathcal{E} \subseteq [X]^{2k}$, where $E \in \mathcal{E}$ if there exist distinct sets $S, T \in [X]^k$ with $S \cup T = E$ and $\Delta(S) = \Delta(T)$.

Then, $|\mathcal{E}| = O(u \cdot n^k)$ by (3), hence, constructing $\mathcal{H}$ can be done in time $O(u \cdot n^k)$. First, assume that $u \leq C \cdot \sqrt{n}$ for some constant $C > 0$. We use the following algorithmic version of Turán's theorem for hypergraphs, cf. [5], [13].

Lemma 1. *Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph on n vertices with average degree d^k. Then, one can find in time $O(|V| + |\mathcal{E}|)$ an independent set $I \subseteq V$ with $|I| \geq c \cdot n/d$.*

Proof. We use the method of conditional probabilities, cf. [9] and [3]. Let $V = \{v_1, \ldots, v_n\}$. Every vertex v_i will be assigned a probability p_i, $i = 1, \ldots, n$. Define a potential by

$$V(p_1, \ldots, p_n) = \sum_{i=1}^{n} p_i - \sum_{\{v_{i_1}, \ldots, v_{i_{k+1}}\} \in \mathcal{E}} \prod_{j=1}^{k+1} p_{i_j} \ .$$

In each step i, $i = 1, \ldots, n$, one after the other, we choose either $p_i = 0$ or $p_i = 1$ to maximize the value of $V(p_1, \ldots, p_n)$. As $V(p_1, \ldots, p_n)$ is linear in each p_i, for $i = 1$, for example, either $V(p_1, \ldots, p_n) < V(1, p_2, \ldots, p_n)$ or $V(p_1, \ldots, p_n) \leq V(0, p_2, \ldots, p_n)$. In the first case, we take vertex v_1, else we discard it. Choosing in the beginning, $p_1 = \ldots = p_n = p = 1/d$, the value of the potential is initially $V(p, \ldots, p) = p \cdot n - p^{k+1} \cdot n \cdot d^k/(k+1) = k/(k+1) \cdot n/d$. Finally, having chosen $p_1, \ldots, p_n \in \{0, 1\}$, the set $V' = \{v_i \in V \mid p_i = 1\}$ is an independent set in $\mathcal{G}$ of size as desired as can easily be seen. Each vertex and each edge in $\mathcal{G}$ is considered only a constant number of times, thus, the running time is $O(|V| + |\mathcal{E}|)$. □

By (3), the average degree d of $\mathcal{H}$ satisfies $d^{2k-1} \leq c' \cdot n^{k-1} \cdot u$. Thus, in $\mathcal{H} = (X, \mathcal{E})$, we can find in time $O(u \cdot n^k)$ an independent set of size at least $c \cdot \left(n^k/u\right)^{1/(2k-1)}$. This part of the algorithm can be done in time $O(u \cdot n^k)$.
Now, assume that $u = \sqrt{n} \cdot \omega(n)$ where $\omega(n) \longrightarrow \infty$ with $n \longrightarrow \infty$. First, we construct the sets $C_{2,j}$ of $(2,j)$-cycles in $\mathcal{H}$, $j = 2, \ldots, 2k-1$. Using (8), this can be done in time $O(u \cdot n^{2k-1})$. We use the following lemma, cf. [5].

Lemma 2. *Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph. Let $\mathcal{G}$ contain $\nu_j(\mathcal{G})$ many $(2,j)$-cycles for $j = 2, \ldots, k$. Then, for any real p, $0 \leq p \leq 1$, one can find in time $O(|V| + |\mathcal{E}| + \sum_{j=2}^{2k-1} \nu_j(\mathcal{G}))$ an induced subhypergraph $\mathcal{G}' = (V', \mathcal{E}')$ such* *that* $|V'| \geq p/3 \cdot |V|$, $|\mathcal{E}'| \leq 3p^{k+1} \cdot |\mathcal{E}|$ *and* $\sum_{j=2}^{k} \nu_j(\mathcal{G}') \leq 3 \cdot \sum_{j=2}^{k} p^{2k+2-j} \cdot \nu_j(\mathcal{G})$ *for* $j = 2, \ldots, k$.

Proof. Let $V = \{v_1, \ldots, v_n\}$. The proof is similar to that of Lemma 1. Step by step, one minimizes the potential $V(p_1, \ldots, p_n)$ given by

$$3^{p \cdot n/3} \cdot \prod_{i=1}^{n} (1 - 2p_i/3) + \frac{\sum_{E \in \mathcal{E}} \prod_{v_i \in E} p_i}{3 \cdot p^{k+1} \cdot |\mathcal{E}|} + \frac{\sum_{j=2}^{k} \sum_{c_j \in C_{2,j}} \prod_{v_i \in c_j} p_i}{3 \cdot \sum_{j=2}^{k} p^{2k+2-j} \cdot |C_{2,j}|},$$

which satisfies in the beginning $V(p, \ldots, p) < 1$. □

We apply Lemma 2 to the hypergraph $\mathcal{H}$ with $p = \left(1/(n^{k-1} \cdot u)\right)^{1/(2k-1)} \cdot (u/\sqrt{n})^{1/((k+1)(2k-1))}$, and we obtain in time $O(u \cdot n^{2k-1})$ an induced subhypergraph $\mathcal{H}' = (X', \mathcal{E}')$ with $|X'| \geq p/3 \cdot |X|$ and $|\mathcal{E}'| \leq 3p^{2k} \cdot |\mathcal{E}|$ and, by (9) and (10), we have $\sum_{j=2}^{2k-1} \nu_j(\mathcal{H}') \leq p/6 \cdot |X|$. In time $o\left(u \cdot n^{2k-1}\right)$, we can determine easily all 2-cycles in $\mathcal{H}'$ and delete from $\mathcal{H}'$ one vertex from each 2-cycle. The remaining hypergraph $\mathcal{H}''$ on at least $pn/6$ vertices contains at most $c \cdot p^{2k} \cdot n^k \cdot u$ edges, thus, has average degree $d^{2k-1} \leq c' \cdot p^{2k-1} \cdot n^{k-1} \cdot u$. Then, we apply the following result from [5] which gives an algorithmic version of the existence result from [7] and extends an algorithm of Fundia [8].

Theorem 3. *Let* $k \geq 2$ *be a fixed integer. Let* $\mathcal{G} = (V, \mathcal{E})$ *be a* $(k+1)$*-uniform hypergraph on* n *vertices with average degree* d^k*. If* $\mathcal{G}$ *does not contain any* 2*-cycles, then one can find for every fixed* $\delta > 0$ *in time* $O(n \cdot d^k + n^3/d^{3-\delta})$ *an independent set of size at least* $c(k, \delta) \cdot n/d \cdot (\ln d)^{1/k}$.

Using Theorem 3, we can find an independent set in $\mathcal{H}''$, hence in $\mathcal{H}$, in time $O\left(p^{2k} \cdot n^k \cdot u + n^3/\left(p \cdot n^{(k-1)/(2k-1)} \cdot u^{1/(2k-1)}\right)^{3-\delta}\right) = o\left(u \cdot n^{2k-1}\right)$.

4 Final Remarks

For $u = \omega(\sqrt{n})$, the running time of the algorithm can be reduced to $O(u^{1-2/(k+1)} \cdot n^{2k-1-1/(k+1)})$ as follows. Similarly as in Lemma 2, for $p = (\sqrt{n}/u)^{2/((k+1)(2k-1))}$ one chooses first a subhypergraph $\mathcal{H}' = (X', \mathcal{E}')$ of $\mathcal{H} = (X, \mathcal{E})$, where we do not control the 2-cycles, but where $|X'| = p/3 \cdot |X|$ and $|\mathcal{E}'| \leq 3p^{2k} \cdot |\mathcal{E}|$. Hence, $\mathcal{H}'$

has $O(u \cdot (p \cdot n)^{2k-\lceil j/2 \rceil})$ many $(2, j)$-cycles. Then, we apply Lemma 2 with a different value p' and proceed as before. Thus, we save some time as we build the 2-cycles later. We omit the details here.

It might also be interesting to find the real growth rate of $f_u(n, k)$ and a corresponding fast algorithm or to give an explicit coloring which yields our upper bound for $f_u(n, k)$ or even a better upper bound for $u = O(\sqrt{n})$.

References

1. M. Ajtai, J. Komlós, J. Pintz, J. Spencer and E. Szemerédi, Extremal Uncrowded Hypergraphs, Journal of Combinatorial Theory Ser. A 32, 1982, 321-335.
2. N. Alon, H. Lefmann and V. Rödl, On an Anti-Ramsey Type Result, Coll. Math. Soc. János Bolyai, 60. Sets, Graphs and Numbers, 1991, 9-22.
3. N. Alon and J. Spencer, The Probabilistic Method, Wiley & Sons, New York, 1992.
4. L. Babai, An Anti-Ramsey Theorem, Graphs and Combinatorics 1, 1985, 23-28.
5. C. Bertram-Kretzberg and H. Lefmann, The Algorithmic Aspects of Uncrowded Hypergraphs, 8th ACM-SIAM Symp. on Discrete Algorithms, to appear, 1997.
6. N. L. Biggs, T. P. Kirkman, Bull. London Math. Soc. 13, 1981, 97-120.
7. R. A. Duke, H. Lefmann and V. Rödl, On Uncrowded Hypergraphs, Random Structures & Algorithms 6, 1995, 209-212.
8. A. Fundia, Derandomizing Chebychev's Inequality to find Independent Sets in Uncrowded Hypergraphs, Random Structures & Algorithms 8, 1996, 131-147.
9. P. Raghavan, Probabilistic Construction of Deterministic Algorithms: Approximating Packing Integer Programs, Journal on Computer and Systems Sciences. 37, 1988, 130-143.
10. D. Ray-Chaudhuri and R. M. Wilson, Solution of Kirkman's Schoolgirl Problem, Proc. Symp. Pure Math. Vol. XIX, UCLA, AMS, Providence, 1968, 187-203.
11. H. Lefmann, V. Rödl, and B. Wysocka, Multicolored Subsets in Colored Hypergraphs, Journal of Combinatorial Theory Ser. A 74, 1996, 209-248.
12. V. Rödl and E. Šiňajová, Note on Independent Sets in Steiner Systems, Random Structures & Algorithms 5, 1994, 183-190.
13. J. Spencer, Turán's Theorem for k-Graphs, Discrete Mathematics 2, 1972, 183-186.