On a Generalization of the Erdős-Szekeres Theorem

Tibor Szabó [1]

University of Illinois at Urbana-Champaign
1409 W Green St, Urbana, IL 61801
e-mail: tszabo@math.uiuc.edu

Gábor Tardos [2]

Alfréd Rényi Mathematical Institute
Reáltanoda u. 13–15, H-1053 Budapest, Hungary
e-mail: tardos@math-inst.hu

Abstract

We consider a generalization of the theorem of Erdős and Szekeres on monotone subsequences.

1 Introduction

The classic lemma of Pál Erdős and György Szekeres [2] about monotone subsequences can also be formulated as a Ramsey-type coloring statement.

Theorem 1 *[2] Let H_0 and H_1 be linear orderings of the n-element set V. Define a 2-coloring of the edges of the complete graph on vertex set V by coloring the edge uv blue if the order of u and v in H_0 agrees with that in H_1, and coloring it red otherwise. Then there exists a monochromatic clique of size $\lceil\sqrt{n}\rceil$.*

Moreover, this result is best possible. That is there exist linear orderings H_0, H_1 such that in the corresponding coloring the largest monochromatic clique is of size $\lceil\sqrt{n}\rceil$.

In the present paper we consider a generalization of this Theorem.

[1]Research supported in part by the Alfred P. Sloan Foundation Grant 96-6-2, the State of New Jersey, and NSF Grant DMS-9970270.

[2]Research supported in part by DIMACS Center and the grants OTKA T-020914, T-030059 and FKFP 0607/1999.

Let $\mathcal{H} = (H_0, \ldots, H_d)$ be a list of $d+1$ linear orderings on a finite set V. Let us 2^d-color the edges of the complete graph on the vertex set V by coloring the edge uv with the color $(c_1, \ldots, c_d) \in \{0,1\}^d$, where $c_i = 0$ if H_i agrees with H_0 on $\{u, v\}$, and $c_i = 1$ otherwise.

The first natural generalization of Theorem 1 coming into mind is about determining the size of the largest monochromatic subset one can guarantee. As it turns out this question is solved easily with repeated applications of the Erdős-Szekeres Theorem.

Instead, we concentrate on the property of a monochromatic subset in a 2-edge-coloring, that it does *not* contain *all* the colors. We will try to determine the size of the largest subset missing at least one of the 2^d colors. This generalization was raised in connection with some problems in analysis (specifically the problem whether any compact set of positive Lebesgue measure in d-space admits a contraction onto a ball [5]) and we found it interesting on its own combinatorial right as well.

Let us make things more precise by introducing a few definitions. For $\vec{c} \in \{0,1\}^d$ we call a subset $U \subseteq V$ *$\vec{c}$-free* if U spans a subgraph with no edge of color $\vec{c}$. We define $m_{\vec{c}}(V, \mathcal{H})$ to be the size of the maximal $\vec{c}$-free subset in V. Let $m(V, \mathcal{H}) = \max_{\vec{c}} m_{\vec{c}}(V, \mathcal{H})$ and $m(n,d) = \min m(V, \mathcal{H})$ where the maximum ranges over the colors $\vec{c} \in \{0,1\}^d$ and the minimum ranges over the n element sets V and the lists $\mathcal{H}$ of $(d+1)$ linear orderings of V.

PROBLEM: Determine $m(n,d)$. Find the order of magnitude for fixed $d \geq 0$.

All orders of magnitude, and all the O, o, Θ, and Ω notations in this paper are in the variable n with respect to a fixed d unless otherwise stated.

We trivially have $m(n,0) = 1$. For $d = 1$ our problem reduces to the theorem of Erdős and Szekeres and one gets $m(n,1) = \lceil \sqrt{n} \rceil$. For $d > 1$ however, the problem starts to get interesting. A trivial lower bound is $m(n,d) \geq m(n,1) = \lceil \sqrt{n} \rceil$ for $d \geq 1$.

For an easy upper bound of $m(n,d) = O(n^{\frac{d}{d+1}})$ one can generalize the construction usually associated with the second part of Theorem 1.

Construction 1

Let $n = n_0^{d+1}$, let V consist of the $(d+1)$-tuples from $\{0, \ldots, n_0 - 1\}$ let $\mathcal{H} = (H_0, \ldots, H_d)$ such that H_i extends the natural ordering according to the i^{th} coordinate (Label the coordinates from 0 to d). Let $\vec{c} \in \{0,1\}^d$ be a color. We define a partition of V into at most $(3n_0 - 2)^d$ monochromatic subsets $R_{\vec{a}}$, where $\vec{a} \in \{-(n_0 - 1), \ldots, 0, 1, \ldots 2n_0 - 2\}^d$. Let

$$R_{\vec{a}} = \{(x_0, x_1, \ldots, x_d) \in V : \vec{a} = (x_1, \ldots, x_d) + x_0(2\vec{c} - (1, \ldots, 1))\}.$$

It is clear that each $R_{\vec{a}}$ is monochromatic in the color $\vec{c}$, thus a $\vec{c}$-free subset does not contain more than one element of it. This implies $m_{\vec{c}}(V,\mathcal{H}) \leq (3n_0-2)^d$ and since $\vec{c}$ was arbitrary we have $m(n,d) \leq m(V,\mathcal{H}) \leq (3n_0-2)^d = O(n^{\frac{d}{d+1}})$. $\square$

The problem of obtaining a decent lower bound resisted our attempts so far. Currently we do not know anything better than $m(n,d) \geq \sqrt{n}$. This is immediate from the Erdős-Szekeres Theorem and proves unnecessarily too much. It provides a subset of size $\lceil\sqrt{n}\rceil$ free of not just one, but half of the colors.

With the hope that it might shed some light on the problem, L. Pósa (see in [4]) suggested a related simpler question. Instead of $d+1$ linear orderings consider a d-tuple $\mathcal{P} = (\mathcal{P}_1,\ldots,\mathcal{P}_d)$ of partitions of a base set V. With the aid of these partitions, we define a 2^d-edge-coloring of the complete graph on vertex set V by letting $\vec{c} = (c_1,c_2,\ldots,c_d) \in \{0,1\}^d$ be the color of the edge uv, where $c_i = 0$ if u and v are in the same class of $\mathcal{P}_i$ and $c_i = 1$ otherwise. We define the analogue of $m_{\vec{c}}(V,\mathcal{H}), m(V,\mathcal{H})$, and $m(n,d)$ for this coloring. For $\vec{c} \in \{0,1\}^d$ let $r_{\vec{c}}(V,\mathcal{P})$ be the size of the maximal $\vec{c}$-free subset of V. Let $r(V,\mathcal{P}) = \max_{\vec{c}} r_{\vec{c}}(V,\mathcal{P})$, and $r(n,d) = \min r(V,\mathcal{P})$, where the maximum ranges over the colors $\vec{c} \in \{0,1\}^d$ and the minimum ranges over the n element sets V and the d-tuples $\mathcal{P}$ of partitions of V. Our goal is also similar: the determination of $r(n,d)$.

In Section 2 we consider this problem on partitions. As it turns out Construction 1 can be translated into the language of partitions. We also prove a matching lower bound in Theorem 2, thus obtain a precise answer for many values of the parameter n: $r(n_0^{d+1},d) = n_0^d$. This implies the asymptotic characterization of $r(n,d)$: $r(n,d) \approx n^{\frac{d}{d+1}}$. This result was independently proved by L. Pósa (for $d=2$) and Gy. Petruska [4].

In Section 5 we consider the random version of the original problem. We show that Construction 1 is very typical in the following sense. If we choose $d+1$ linear orderings independently and uniformly from all linear orderings of the set V, then almost always $m(V,\mathcal{H}) = \Theta(n^{\frac{d}{d+1}})$. We also prove a stronger concentration result about the value of $m(V,\mathcal{H})$.

After considering a simpler variant, and the random version of our original problem it might seem plausible to conjecture that $m(n,d) = \Theta(n^{\frac{d}{d+1}})$. As we show in Section 4 this, however, is *not* the case. We improve on Construction 1 to obtain $m(n,d) = O(n^{e_d})$, with an exponent satisfying $e_d < \frac{d}{d+1}$ for $d \geq 2$. For example we have $m(n,2) = O(n^{5/8}) \ll n^{2/3}$. For large values of d we have $e_d = 1 - 2/d + o(1/d)$, where the o bound is in the variable d.

2 Partitions

Theorem 2 *Let $\mathcal{P}_1, \ldots, \mathcal{P}_d$ be d partitions of the n element set V. Let us define a 2^d-edge-coloring of K_n on vertex set V by letting $\vec{c} = (c_1, c_2, \ldots, c_d) \in \{0,1\}^d$ to be the color of the edge uv, where $c_i = 0$ if u and v are in the same class of $\mathcal{P}_i$ and $c_i = 1$ otherwise.*

There exists a color $\vec{c}$ and a $\vec{c}$-free subset $B \subseteq V$, with $|B| \geq n^{\frac{d}{d+1}}$.

Theorem 2 was also proved independently by Pósa for $d = 2$, and Petruska [4] for arbitrary d. Their proofs are different from ours.

We obtain Theorem 2 as a consequence of a stronger statement.

Theorem 3 *In the setting of Theorem 2 there exist subsets $B_0, B_1, \ldots, B_d$ of V, where B_i is $(0, \ldots, 0, \underbrace{1, 1 \ldots, 1}_{i})$-free, and*

$$\frac{1}{d+1} \sum_{i=0}^{d} |B_i| \geq n^{\frac{d}{d+1}}.$$

Proof of Theorem 3. We proceed by induction on d. The case $d = 1$ is immediate. Now let us assume that the statement is true for $d - 1$.

Consider the restrictions of $\mathcal{P}_2, \ldots, \mathcal{P}_d$ to the classes $S_1, \ldots, S_k$ of $\mathcal{P}_1$. By the induction hypothesis there exist subsets $B_0^j, \ldots, B_{d-1}^j$ of S_j for every j $(1 \leq j \leq k)$, such that $\frac{1}{d} \sum_{i=0}^{d-1} |B_i^j| \geq |S_j|^{1-1/d}$ and B_i^j does not use the color $(0, \ldots, 0, \underbrace{1, \ldots, 1}_{i}) \in \{0,1\}^{d-1}$. We define $B_i := \cup_{j=1}^k B_i^j \subseteq V$ for $0 \leq i \leq d-1$. The set B_i of size $|B_i| = \sum_{j=1}^{k} |B_i^j|$ is $(\overbrace{0, \ldots, 0, \underbrace{1 \ldots 1}_{i}}^{d})$-free.

Let B_d be one of the largest classes in $\mathcal{P}_1$ and let $t = |B_d|$. B_d is clearly $(1, \ldots, 1)$-free. We obtain

$$\sum_{i=0}^{d} |B_i| = \sum_{i=0}^{d-1} |B_i| + t = \sum_{i=0}^{d-1} \sum_{j=1}^{k} |B_i^j| + t = \sum_{j=1}^{k} \sum_{i=0}^{d-1} |B_i^j| + t \geq d \sum_{j=1}^{k} |S_j|^{\frac{d-1}{d}} + t.$$

Using $|S_j| \leq t$ we estimate $|S_j|^{\frac{d-1}{d}} \geq |S_j| t^{-1/d}$ to get

$$\sum_{i=0}^{d} |B_i| \geq d \sum_{j=1}^{k} |S_j|^{\frac{d-1}{d}} + t \geq d t^{-1/d} \sum_{j=1}^{k} |S_j| + t = d n t^{-1/d} + t \geq (d+1) n^{\frac{d}{d+1}}.$$

Here the last inequality follows because $dnt^{-1/d} + t$ attains its minimum in t at $n^{\frac{d}{d+1}}$. $\square$

Recall the definition of $r(n, d)$ from the Introduction.

Corollary 4 *For positive integers d and n_0 we have*

$$r(n_0^{d+1}, d) = n_0^d.$$

For arbitrary n and fixed d we have

$$r(n,d) = (1+o(1))n^{\frac{d}{d+1}}.$$

Proof. Theorem 2 provides the lower bound on $r(n,d)$.

Construction 1 can be transformed into a construction of partitions and provides the upper bound. Let $n = n_0^{d+1}$, and $V = \{0,1,\ldots,n_0-1\}^{d+1}$. For each $i = 1,2,\ldots,d$ we define the partition $\mathcal{P}_i$ using the i^{th} coordinates: two elements of V are in the same class of $\mathcal{P}_i$ if they have the same i^{th} coordinate. (Coordinates are labeled from 0 to d.)

Take a color $\vec{c} = (c_1,\ldots,c_d)$. One can partition V into $\vec{c}$-monochromatic subsets $R_{\vec{a}}$ of size n_0 ($\vec{a} \in \{0,1,\ldots n_0\}^d$). Indeed, let

$$R_{\vec{a}} = \{(i,\vec{x}) : i \in \{0,1\ldots,n_0-1\}, \vec{x} = \vec{a} + i\vec{c}\},$$

where the sum in the coordinates is computed modulo n_0.

Thus for any color $\vec{c}$, the size of the largest $\vec{c}$-free subset is at most n_0^d.

□

3 The Random Orders

We prove the following Theorem about the random version of our problem.

Theorem 5 *Let V be a set of n elements. Let $H_0, H_1, \ldots H_d$ be linear orderings of V chosen independently and uniformly out of all possible linear orderings of V. Let $\vec{c} = (0,0,\ldots,0)$. Then almost always*

$$m_{\vec{c}}(V,\mathcal{H}) = \Theta(n^{\frac{d}{d+1}}).$$

That is for any d there exist constants $C_1 = C_1(d)$ and $C_2 = C_2(d)$ such that

$$\lim_{n\to\infty} Pr(C_1 n^{\frac{d}{d+1}} < m_{\vec{c}}(V,\mathcal{H}) < C_2 n^{\frac{d}{d+1}}) = 1.$$

Corollary 6 *Let V be a set of n elements. Let $H_0, H_1, \ldots H_d$ be linear orderings of V chosen independently and uniformly out of all possible linear orderings of V. Then almost always*

$$m(V,\mathcal{H}) = \Theta(n^{\frac{d}{d+1}}).$$

Proof. It follows from Theorem 5 using the symmetry of colors.
□

Proof of Theorem 5. We say that $x \in V$ *dominates* $y \in V$ if x precedes y in each of the orderings $H_0, \ldots, H_d$. In this case we also say that x and y form a *dominating pair.* A subset of V is called *dominating* if every pair of elements of it is dominating. A subset is called *domination-free* if it contains no dominating pairs. With these definitions $m_{\vec{c}}(V, \mathcal{H})$ is just the maximum size of a domination-free subset.

Obtaining the lower bound is the easier part of the proof. The probability of a certain fixed r-subset $R \subseteq V$ being dominating is $(1/r!)^d$, since each of $H_1, H_2, \ldots$ and H_d has to agree with H_0 on R. Thus

$$Pr(\exists \text{ a dominating } R \subseteq V, |R| = r) \leq \binom{n}{r}\left(\frac{1}{r!}\right)^d < \left(\frac{ne}{r}\right)^r \left(\frac{e}{r}\right)^{rd} = \left(e^{d+1}\frac{n}{r^{d+1}}\right)^r.$$

Thus we have

$$Pr(\exists \text{ a dominating } R \subseteq V, |R| > 3n^{\frac{1}{d+1}}) = o(1).$$

Domination defines a partial ordering on V, where dominating sets correspond to chains, and domination-free subsets correspond to antichains. By Dilworth's Theorem the product of the maximum chain size and the maximum antichain size is at least n. Thus

$$Pr(\exists \text{ a domination-free } F \subseteq V, |F| > \frac{1}{3}n^{\frac{d}{d+1}}) \to 1.$$

In the rest of this proof we prove the upper bound. We produce the random orderings of V by choosing independently n points, one for each element of V, in the unit hypercube $K = [0,1)^{d+1}$, according to a uniform distribution. Define H_i to be the linear ordering given by the i^{th} coordinates of the points. (We label the coordinates by $0, 1 \ldots, d$.) There are no "ties" with probability 1.

Let l be the integer with $2^{l(d+1)} \leq n < 2^{(l+1)(d+1)}$ and let $k = 2^l$. We partition K into k^{d+1} hypercubes of sidelength $1/k$. We refer to these hypercubes as *smallcubes* and label each smallcube with a vector in $\{0, \ldots, k-1\}^{d+1}$: the point $(x_0, x_1, \ldots, x_d) \in K$ is in the smallcube with label $(\lfloor kx_0 \rfloor, \ldots, \lfloor kx_d \rfloor)$.

The *line of a smallcube* $(a_0, \ldots, a_d)$ is defined to be the set of all smallcubes with label $(a_0 + b, \ldots, a_d + b)$ for some integer b. It is clear that two points of V chosen from different smallcubes of the same line form a dominating pair. Thus a domination-free set can contain points from only one smallcube of each line. The number of lines is $k^{d+1} - (k-1)^{d+1} = O(n^{\frac{d}{d+1}})$. It is a routine task to prove that almost always there are $O(\log n / \log\log n)$ elements of V in each smallcube. Thus we get $m_{\vec{c}}(V,\mathcal{H}) = O(n^{\frac{d}{d+1}}\log n/\log\log n)$ almost always. In order to get rid of the logarithmic factor we need a further idea.

Let $\mathcal{K}_0 = \{K\}$. For each $i = 1,\dots l$ let $\mathcal{K}_i$ contain all subcubes of sidelength $1/2^i$, which are obtained by partitioning the elements of $\mathcal{K}_{i-1}$ into 2^{d+1} subcubes each. Thus $|\mathcal{K}_i| = 2^{d+1}|\mathcal{K}_{i-1}| = 2^{(d+1)i}$. The elements of $\mathcal{K}_l$ are the subcubes we called smallcubes. Let $\mathcal{K} = \cup_{i=0}^{l}\mathcal{K}_i$.

Let C be a large constant to be chosen later. We assign each point of V to a member of $\mathcal{K}$ it is contained in. We do this one by one (in some arbitrary fixed order). In the process we do not assign more than C points to any particular member of $\mathcal{K}$, except maybe to K. If a subcube is assigned to C points we say it is *full*. We always assign a point to the smallest possible subcube, which is not yet full.

In the simple proof of the weaker upper bound above we basically assigned the points to the elements of $\mathcal{K}_l$. This way we could not guarantee that each subcube is assigned to only a constant number of points. With the use of subcubes of larger sizes this shortcoming can be fixed.

We have seen above, that a domination-free subset S of V has got points in at most $(d+1)2^{ld}$ members of $\mathcal{K}_l$. It can be seen similarly that S has got points in at most $(d+1)2^{id}$ members of $\mathcal{K}_i$ for every $i = 1,\dots l$. Since all cubes in $\mathcal{K}\setminus\{K\}$ are assigned to at most C points, S contains at most $C(d+1)\sum_{i=1}^{l} 2^{id} = O(n^{\frac{d}{d+1}})$ points besides the ones we assigned to K.

Our goal is to show that we can choose a $C = C(d)$ such that the probability of at least C points being assigned to K is $o(1)$. This will finish the proof of Theorem 5.

Suppose that at least C points are assigned to K. The reason these points are assigned to K is that the smaller subcubes containing them were already full. Let us define $\mathcal{L}_i \subseteq \mathcal{K}_i$ by recursion for $i = 0,\dots,l$. We set $\mathcal{L}_0 = \{K\}$ and for $i > 0$ we let $\mathcal{L}_i$ consist of the members of $\mathcal{K}_i$ that are *full* subcubes of a cube in $\mathcal{L}_{i-1}$. Let $\mathcal{L} = \cup_{i=0}^{l}\mathcal{L}_i$ and set $s = |\mathcal{L}|$. All the elements of $\mathcal{L}$ are full, thus they contain at least Cs points. Notice that by the rule we used when assigning points to cubes all these points are contained in $\cup\mathcal{L}_l$. The probability of the event that some fixed set $\mathcal{L}_l$ of at most s smallcubes contains at least Cs out of the n uniformly chosen random points is at most $\binom{n}{Cs}(s/k^{d+1})^{Cs} < (2^{d+1}e/C)^{Cs}$.

$\mathcal{K}$ can be considered the full 2^{d+1}-ary tree of depth l. In this setting $\mathcal{L}$ is a subtree of size s containing the root. To bound the number of possible sets $\mathcal{L}$ we can use the formula for the number of rooted subtrees of size s in an *infinite* $D = 2^{d+1}$-ary tree. The bound 2^{Ds} is straightforward. The exact number of the latter subtrees is known to be $\binom{Ds}{s}/((D-1)s+1)$, see [3]. Either bound suffice for our proof.

Thus we have

$$Pr(K \text{ is full}) < \sum_{s=l}^{\infty} \binom{2^{d+1}s}{s} \left(\frac{2^{d+1}e}{C}\right)^{Cs} < \sum_{s=l}^{\infty} (2^{d+1}e)^s \left(\frac{2^{d+1}e}{C}\right)^{Cs} < \sum_{s=l}^{\infty} (1/2)^s = 2/2^l = o(1).$$

Above we used that $\mathcal{L}$ has $s \geq l$ elements, and that $2^{d+1}e(2^{d+1}e/C)^C < 1/2$ provided C is large enough.

□

Corollary 6 shows that the median of $m(V, \mathcal{H})$ is $\Theta(n^{\frac{d}{d+1}})$. Using standard technic involving Talagrand's Inequality one can improve on Corollary 6 and obtain that $m(V, \mathcal{H})$ is very strongly concentrated around its median. In particular the expected value of $m(V, \mathcal{H})$ is also $\Theta(n^{\frac{d}{d+1}})$.

Theorem 7 *Let m be the median of $m(V, \mathcal{H})$ and let $\omega(n) \to \infty$ arbitrarily slowly. Then*

$$Pr(|m(V, \mathcal{H}) - m| > \omega(n) n^{\frac{1}{2} - \frac{1}{2(d+1)}}) = o(1).$$

Proof. We don't include the details here. The proof is a standard application of Talagrand's Inequality. For a clear explanation of this powerful probabilistic tool see for example [1]. □

4 The construction

In this section we present a generalization of Construction 1 to improve on the exponent of the upper bound for $m(n, d)$ provided $d \geq 2$.

Theorem 8 *We have $m(n, d) = O(n^{e_d})$, with*

$$e_d = 1 - \max_{i \leq d+1} \frac{\sum_{j=0}^{i-1} \binom{d}{j}}{i 2^d}.$$

In particular $m(n, 2) = O(n^{5/8})$ and there exists an absolute constant $c > 0$, such that for every fixed d we have

$$m(n, d) < O\left(n^{1 - \frac{2}{d} + \frac{c\sqrt{\log d}}{d^{3/2}}}\right).$$

Proof. We start by defining a product operation. Let $\mathcal{H}' = (H'_0, \ldots, H'_d)$ be $d+1$ linear orderings on the finite set V' and let $\mathcal{H}'' = (H''_0, \ldots, H''_d)$ be $d+1$ linear orderings on the finite set V''. We define $V = V' \times V''$ to be the Cartesian product and $\mathcal{H}' \times \mathcal{H}'' = \mathcal{H} = (H_0, \ldots, H_d)$ to be $d+1$ orderings on V, where H_i is the lexicographic ordering of V using the orderings H'_i and H''_i on the coordinates $(i = 0, 1, \ldots, d)$.

Lemma 9 *For any color $\vec{c} \in \{0,1\}^d$ we have $m_{\vec{c}}(V,\mathcal{H}) = m_{\vec{c}}(V',\mathcal{H}') \cdot m_{\vec{c}}(V'',\mathcal{H}'')$.*

Proof. For the $\geq$ direction take subsets S' of V', and S'' of V'', which are both $\vec{c}$-free and notice that $S' \times S'' \subseteq V$ is also $\vec{c}$-free.

For the $\leq$ direction of the claim take a $\vec{c}$-free subset $S \subseteq V$. Note that the projection S' of S to V' is $\vec{c}$-free, and the slice $S_a = \{b \in V'' | (a,b) \in S\}$ is also $\vec{c}$-free for any $a \in V'$.

□

In what follows we modify Construction 1 to obtain a list $\mathcal{H} = (H_0, \ldots, H_d)$ of linear orderings such that $m_{\vec{c}}(V,\mathcal{H})$ is substantially lower for most of the colors but it is very high (in fact n) for the remaining colors. Then we use the product construction of Lemma 9 for averaging.

Let $1 \leq i \leq d+1$ and let us take an i dimensional subspace W of $I\!R^{d+1}$ in general position with respect to the coordinate axes. In the following we consider d, i and W fixed, and use the O notation with respect to n. Consider (a rotation of) the i dimensional unit square grid in W. Let V be the n points of this grid closest to the origin. Notice that the diameter of V is $O(n^{1/i})$. For $i = 0, 1 \ldots, d$ define H_i to be the linear ordering given by the ordering of the i^{th} coordinates of the points. (Coordinates are labeled from 0 to d.)

The color of the edge $\vec{u}\vec{v}$ ($\vec{u}, \vec{v} \in V$) depends on which of the 2^{d+1} space orthant contains the vector $\vec{u} - \vec{v}$. An orthant Q is associated with the same color as $-Q$, hence the $2^{d+1}/2$ colors. Let us denote by $Q_{\vec{c}}$ the union of the two orthants associated with the color $\vec{c}$. In the following claim we show that the magnitude of $m_{\vec{c}}(V,\mathcal{H})$ depends only on whether $Q_{\vec{c}} \cap W$ is trivial or not.

Claim 10 *If $Q_{\vec{c}} \cap W = \{0\}$ then $m_{\vec{c}}(V,\mathcal{H}) = n$. If $Q_{\vec{c}} \cap W \neq \{0\}$, then $m_{\vec{c}}(V,\mathcal{H}) = O(n^{\frac{i-1}{i}})$.*

Proof. The first statement simply follows from the definition of $Q_{\vec{c}}$ and from the fact that W is closed under subtraction.

To prove the second statement choose a vector $\vec{v} \in W$ that lies in the interior of $Q_{\vec{c}}$. Let $S \subseteq V$ be a subset not containing the color $\vec{c}$.

Let us project S in the direction of $\vec{v}$ onto the subspace of W orthogonal to $\vec{v}$ and call the projected point set S_0. Recall that $\vec{v}$ is in the interior of $Q_{\vec{c}}$ and thus there exists a positive angle α with the property that any vector within angle at most α from $\vec{v}$ is in $Q_{\vec{c}}$. Clearly for any two vectors $\vec{u}$ and $\vec{w} \in S$ the distance of their projections $\vec{u_0}, \vec{w_0} \in S_0$ is at least $\sin\alpha$ times the distance of $\vec{u}$ and $\vec{w}$, as otherwise the difference $\vec{u} - \vec{w}$ (or $\vec{w} - \vec{u}$) is within angle α from $\vec{v}$, making $\vec{c}$ the color of the edge $\vec{u}\vec{v}$, a contradiction. Thus the minimum distance in S_0 is constant, while the diameter is at most the diameter of S, which is $O(n^{1/i})$. A simple volume calculation shows that $|S| = |S_0| = O(n^{\frac{i-1}{i}})$. □

Suppose that in the construction above we have $l = l(i)$ colors with $m_{\vec{c}}(V, \mathcal{H}) = O(n^{1-1/i})$ and $2^d - l$ colors with $m_{\vec{c}}(V, \mathcal{H}) = n$. There are 2^d ways to reverse some of the linear orderings $H_1, \ldots, H_d$ and obtain a different construction $\mathcal{H}^j$, $j = 1, 2, \ldots, 2^d$. For each of them the set of l colors with $m_{\vec{c}}(V, \mathcal{H}^j) = O(n^{\frac{i-1}{i}})$ might be different. Because of symmetry any fixed color $\vec{c}$ occurs l times out of the 2^d with $m_{\vec{c}}(V, \mathcal{H}^j) = O(n^{1-1/i})$.

Now we use the product construction to multiply these 2^d systems. For the resulting family $(V^*, \mathcal{H}^*)$ the values $m_{\vec{c}}(V^*, \mathcal{H}^*)$ average out for every color $\vec{c}$,

$$m_{\vec{c}}(V^*, \mathcal{H}^*) = \prod_{j=1}^{2^d} m_{\vec{c}}(V, \mathcal{H}^j) = O\left(n^{l\frac{i-1}{i} + (2^d - l)}\right) = O\left(N^{1-\frac{l}{i2^d}}\right),$$

where $N = |V^*| = n^{2^d}$

Assuming the next lemma on the value of l (as a function of i) the first statement of Theorem 8 follows. For the last statement of the theorem notice that with the choice $i = \lfloor d/2 + 10\sqrt{d \log d} \rfloor$ Chernoff bound gives $\sum_{j=0}^{i-1} \binom{d}{j}/(i2^d) = 2/d - O(\sqrt{\log d / d^3})$ where the O is with respect to d. □

Lemma 11 *If $W \subseteq \mathbb{R}^{d+1}$ is an i-dimensional subspace in general position with respect to the coordinate axes, then W nontrivially intersects exactly $2\sum_{j=0}^{i-1} \binom{d}{j}$ of the 2^{d+1} orthants of $\mathbb{R}^{d+1}$.*

Proof. The intersections of the $d+1$ coordinate hyperplanes of $\mathbb{R}^{d+1}$ with W are $d+1$ $(i-1)$-dimensional subspaces of W in general position. Thus counting the $(d+1)$-dimensional orthants intersected nontrivially by W is the same as counting the connected parts $\mathbb{R}^i$ is cut by $d+1$ subspaces of dimension $i-1$ in general position.

Our formula for this number can easily be established by a recurrence relation. We tried to find the oldest reference instead. In 1852 L. Schläfli [6] proved that j affine hyperplanes in general position partition the Euclidean k-space into $a(j, k) = \sum_{t=0}^{k} \binom{j}{t}$ parts. Notice that he uses affine subspaces and we use linear subspaces. We partition the i-space by $d+1$ linear subspaces of dimension $i-1$. Fix one of the subspaces S and consider affine hyperplanes S_1 and S_2 parallel to S that lie on different sides of S. Clearly each part of the i-space intersects exactly one of S_1 or S_2, thus the number of parts in our partition is the total number of parts S_1 and S_2 is partitioned by the other d of our subspaces. As the subspaces different from S intersect S_1 and S_2 in affine subspaces in general position we have that both S_1 and S_2 is partitioned into $a(i-1, d)$ parts, proving the theorem. □

In Theorem 8 the value of e_d is defined with the help of the dimension parameter i. The construction in the proof works for each value $1 \leq i \leq d+1$ and choosing i optimally yields the exponent e_d. Observe that the choice $i = d + 1$ provides a version of Construction 1 from the Introduction. By choosing $i = d$ we obtain a construction beating the random one even for small values of d. For $d = 2$, 3, and 4 this is the optimal choice for i and we get $m(n, 2) = O(n^{5/8}), m(n, 3) = O(n^{17/24})$, and $m(n, 4) = O(n^{49/64})$. For $d = 5$, the optimal choice is $i = 4$ that yields $m(n, 5) = O(n^{51/64})$. For large values of d the optimal choice for i is $d/2+O(\sqrt{d \log d})$ yielding $e_d = 1-\frac{2}{d}+O\left(\frac{\sqrt{\log d}}{d^{3/2}}\right)$, where the O notation refers to asymptotics in d.

5 Concluding Remarks and Open Problems

Remarks.

1. It would seem natural to try to prove the result of Theorem 3 for the average of $|B_{\vec{c}}|$'s over *all* the 2^d colors $\vec{c}$, where $B_{\vec{c}} \subseteq V$ is a $\vec{c}$-free subset of maximum size. This is not possible because there are counterexamples with two partitions, where $\sum_{i=1}^{4} |B_i| < 4n^{2/3}$. An $\Omega(n^{1-1/(k+1)})$ bound trivially follows from Theorem 3 for the average of all the 2^d colors.

2. Our argument gives a similar result for a coloring induced by 2 linear orderings and $d - 1$ partitions.

3. Since the proof of the current best lower bound for $m(n, 2)$ provides a subset of size $n^{1/2}$ containing only 2 of the 4 colors, it seems reasonable to conjecture, that $\lim_{n\to\infty} m(n, 2)/n^{1/2} = \infty$. We don't know anything better for large d either. As a first step it would even be interesting to see whether there is a constant d for which $\lim_{n\to\infty} m(n, d)/n^{1/2} = \infty$.

ACKNOWLEDGMENT Enlightening conversations with Noga Alon, Michael Krivelevich, János Pach, and Ákos Seress are greatly appreciated.

References

[1] N. Alon, J. Spencer, *The Probabilistic Method*, 2nd Edition, Wiley & Sons, (2000)

[2] P. Erdős, G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* **2**, (1935), 463-470.

[3] D. Knuth, *The Art of Computer Programming*, Vol. I, Addison Wesley, London, 1969, p. 396 (Exercise 11).

[4] G. Petruska, manuscript.

[5] D. Preiss, personal communication.

[6] L. Schläfli, Theorie der vielfachen Kontinutät (Berne, 1852), Ges. Math. Abh. Vol. I, p. 209 (Basel, 1950).