Large Subsets of Points with all Pairs (Triples) Having Different Distances (Areas)

William Gasarch *
Univ. of MD at College Park

Sam Zbarsky †
Montgomery Blair High School

January 22, 2013

Abstract

Let $\{p_1, \ldots, p_n\} \subseteq \mathsf{R}^d$. We think of $d \le n$. How big is the largest subset X of points such that all of the distances determined by elements of $\binom{X}{2}$ are different? We show that X is at least $\Omega((n^{1/(6d)}(\log n)^{1/3})/d^{1/3})$. Assume that no three of the original points are collinear. How big is the largest subset X of points such that all of the areas determined by elements of $\binom{X}{3}$ are different? We show that, if $d = 2$ then X is at least $\Omega((\log\log n)^{1/2901})$, and if $d = 3$ then X is at least $\Omega((\log\log n)^{1/27804})$. All of our results use variants of the canonical Ramsey theorem and some geometric lemmas.

1 Introduction

Let $\{p_1, \ldots, p_n\} \subseteq \mathsf{R}^d$. We think of $d \le n$. How big is the largest subset X of points such that all of the distances determined by elements of $\binom{X}{2}$ are different? Assume that no three of the original points are collinear. How big is the largest subset X of points such that all of the areas determined by elements of $\binom{X}{3}$ are different?

Def 1.1 Let $a \ge 1$. Let $h_{a,d}(n)$ be the largest integer so that if $p_1, \ldots, p_n$ are any set of n distinct points in R^d, no a in the same $(a-2)$-dimensional space, then there exists a subset X of $h_{a,d}(n)$ points for which all of the volumes determined by elements of $\binom{X}{a}$ are different. The $h_{a,d}(n)$ *problem* is the problem of establishing upper and lower bounds on $h_{a,d}(n)$.

There does not seem to be much known about $h_{a,d}(n)$. Below we summarize all of the references we found.

*University of Maryland at College Park, Department of Computer Science, College Park, MD 20742. gasarch@cs.umd.edu

†Montgomery Blair High School, Silver Spring, MD, 20901 sa_zbarsky@yahoo.com

1. Erdős considered the $h_{a,d}(n)$ problem 1957 [6] and 1970 [7]. The former paper is in Hungarian and not available online so we do not know what is in it. In the latter paper he notes that $h_{2,2}(7) = 3$ [9] and $h_{2,3}(9) = 3$ [3]. Erdős conjectured that $h_{2,1}(n) = (1 + o(n))n^{1/2}$ and notes that $h_{2,1}(n) \leq (1 + o(n))n^{1/2}$ [13].

2. Erdős considered the $h_{a,d}(n)$ problem in 1986 [8]. He states *It is easy to see that $h_{2,d}(n) > n^{\epsilon_d}$ but the best possible value of ϵ_d is not known. $\epsilon_1 = \frac{1}{2}$ follows from a result of Ajtai, Komlos, Sulyok and Szemeredi [19].*

3. Erdős and Purdy [10] commented on the $h_{a,d}(n)$ problem in 1995; however, they do not state anything that was not already known in 1986.

4. Erdős [5] showed that the number of distinct differences in the $\sqrt{n} \times \sqrt{n}$ grid is $\leq O(\frac{n}{\sqrt{\log n}})$. Therefore $h_{2,2}(n) \leq O\left(\sqrt{\frac{n}{\sqrt{\log n}}}\right)$. For $a \geq 3$ We do not know of any nontrivial upper bounds on $h_{a,d}$ (Note that the $n^{1/d} \times \cdots \times n^{1/d}$, which yields has three (actually many) points collinear and hence cannot be used to obtain an upper bound.)

Note 1.2 The problem of $h_{2,2}$ is similar to but distinct from the *Erdős Distance Problem*: give a set of n points in the plane how many distinct distances are guaranteed. For more on this problem see [15, 16]. The problem of $h_{3,2}$ is similar to but distinct from the problem of determining, given n points in the plane no three collinear, how many distinct triangle-areas are obtained (see [4] and references therein). We do not know of any reference to a higher dimensional analog of these problems.

In this paper we show

- $h_{2,d}(n) \geq \Omega((n^{1/(6d)}(\log n)^{1/3})/d^{1/3})$.
- $h_{3,2}(n) \geq \Omega((\log \log n)^{1/2901})$.
- $h_{3,3}(n) \geq \Omega((\log \log n)^{1/27804})$.

Our proof has two ingredients: (1) upper bounds on variants of the canonical Ramsey numbers, and (2) geometric lemmas about points in R^d.

In Section 2,3, and 4 we define terms, prove lemmas, and finally prove an upper bound on a variant of the canonical Ramsey Theorem. Our proof uses some ideas from the upper bound on the standard canonical Ramsey number, $ER(k, k)$, due to Lefmann and Rodl [21]. In Section 5 we prove a geometric lemma about points in R^d. In Section 6 we use our upper bound and our geometric lemma to prove lower bounds on $h_{2,d}(n)$. In Section 7 we show lower bounds on $h_{3,2}(n)$ and $h_{3,3}(n)$. In Section 8 we speculate about lower bounds for $h_{a,d}$ for $a \geq 4$. In Section 9 we list open problems.

2 Variants of the Canonical Ramsey Theorem

Notation 2.1 Let $n \in \mathsf{N}$.

1. $[n]$ is the set $\{1, \ldots, n\}$.
2. If A is a set and $0 \le a \le |A|$ then $\binom{A}{a}$ is the set of all a-sized subsets of A.
3. K_n is the graph (V, E) where

$$\begin{aligned} V = &\ [n] \\ E = &\ \binom{[n]}{2} \end{aligned}$$

We will identify $\binom{[n]}{2}$ with K_n. We may refer to vertices and edges of $\binom{[n]}{2}$.

We define terms and then state the canonical Ramsey theorem (for graphs). It was first proven by Erdős and Rado [11]. The best known upper bounds on the canonical Ramsey numbers are due to Lefmann and Rodl [21].

Def 2.2 Let $COL : \binom{[n]}{2} \to \omega$. Let $V \subseteq [n]$.

1. The set V is *homogenous* (henceforth *homog*) if every edge in $\binom{V}{2}$ is colored the same.
2. The set V is *min-homogenous* (henceforth *min-homog*) if for all $a < b$ and $c < d$

$$COL(a, b) = COL(c, d) \text{ iff } a = c.$$

3. The set V is *max-homogenous* (henceforth *max-homog*) if for all $a < b$ and $c < d$

$$COL(a, b) = COL(c, d) \text{ iff } b = d.$$

4. The set V is *rainbow* if every edge in $\binom{V}{2}$ is colored differently.

Notation 2.3 We will state many results in terms of functions from $\binom{[n]}{a}$ to ω. When we use these results we will actually use functions from $\binom{[n]}{a}$ to R. The change in our results to accommodate this is only a change of notation.

Theorem 2.4 *For all k there exists n such that, for all colorings of $\binom{[n]}{2}$ there is either a homog set of size k, a min-homog set of size k, a max-homog set of size k, or a rainbow set of size k. We denote the least value of n that works by $ER(k, k)$.*

We now state the asymmetric canonical Ramsey Theorem.

Theorem 2.5 *For all k_1, k_2 there exists n such that, for all colorings of $\binom{[n]}{2}$, there is either a homog set of size k_1, a min-homog set of size k_1, a max-homog set of size k_1, or a rainbow set of size k_2. We denote the least value of n that works by $ER(k_1, k_2)$.*

We will actually need a variant of the asymmetric canonical Ramsey Theorem which is weaker but gives better upper bounds.

Def 2.6 Let $COL : \binom{[n]}{2} \rightarrow \omega$. Let $V \subseteq [n]$. The set V is *weakly homogenous* (henceforth *whomog*) if there is a way to linear order V (not necessarily the numerical order),

$$V = \{x_1, x_2, \ldots, x_L\},$$

such that, for all for all $1 \le a \le L-3$, for all $a < b < c \le L$,

$$COL(x_a, x_b) = COL(x_a, x_c).$$

Informally, the color of (x_i, x_j), where $i < j$, depends only on i. (We intentionally have $1 \le a \le L-3$. We do not care if $COL(x_{L-2}, x_{L-1}) = COL((x_{L-2}, x_L).)$

Note 2.7 When presenting a whomog set we will also present the needed linear order.

The following theorem follows from 2.5.

Theorem 2.8 *For all k_1, k_2 there exists n such that, for all colorings of $\binom{[n]}{2}$ there is either a whomog set of size k_1, or a rainbow set of size k_2. We denote the least value of n that works by $WER(k_1, k_2)$.*

In Theorem 4.1 we will show

$$WER(k_1, k_2) \le \frac{(Ck_2)^{6k_1-18}}{(\log k_2)^{2k_1-6}}.$$

3 Lemma to Help Obtain Rainbow Sets

The next definition and lemmas gives a way to get a rainbow set under some conditions.

Def 3.1 Let $COL : \binom{[m]}{2} \rightarrow \omega$. If c is a color and $v \in [m]$ then $\deg_c(v)$ is the number of c-colored edges with an endpoint in v.

The following result is due to Alon, Lefmann, and Rodl [1].

Lemma 3.2 *Let $m \ge 3$.*

1. *Let $COL : \binom{[m]}{2} \rightarrow \omega$ such that, for all $v \in [m]$ and all colors c, $\deg_c(v) \le 1$. Then there exists a rainbow set of size $\ge \Omega((m \log m)^{1/3})$.*
2. *There exists a coloring of $\binom{[m]}{2}$ such that for all $v \in [m]$ and all colors c, $\deg_c(v) \le 1$ and all rainbow sets are of size $\le O((m \log m)^{1/3})$.*

The following easily follows:

Lemma 3.3 *Let $m \geq 3$. Let $COL : \binom{[m]}{2} \rightarrow \omega$ such that, for all $v \in [m]$ and all colors c, $\deg_c(v) \leq 1$. If $m = \Omega(\frac{k^3}{\log k})$ then there exists a rainbow set of size k.*

The following definitions and lemmas will be used to achieve the premise of Lemma 3.3

Def 3.4 Let $COL : \binom{[m]}{2} \rightarrow \omega$. Let c be a color and let $x \in [m]$.

1. $\deg_c(x)$ is the number of c-colored edges (x, y).
2. A *bad triple* is a triple a, b, c such that a, b, c does not form a rainbow K_3.

The next two lemmas show us how to, in some cases, reduce the number of bad triples.

Lemma 3.5 *Let $COL : \binom{[m]}{2} \rightarrow \omega$ such that, for every color c and vertex v, $\deg_c(v) \leq d$. Then the number of bad triples is less than $\frac{dm^2}{2}$.*

Proof: We assume that d divides $m-1$. We leave the minor adjustment needed in case d does not divide $m-1$ to the reader.

Let b be the number of bad triples. We upper bound b by summing over all v that are the point of the triple with two same-colored edges coming out of it.

$$\begin{aligned} b &\leq \textstyle\sum_{v\in[m]} \sum_{c\in\mathsf{N}} \text{ Num of bad triples } \{v, u_1, u_2\} \text{ with } COL(v, u_1) = COL(v, u_2) = c \\ &\leq \textstyle\sum_{v\in[m]} \sum_{c\in\mathsf{N}} \binom{\deg_c(v)}{2} \end{aligned}$$

Note that we are not assuming $v < u_1, u_2$. Also note that this upper bound may be an overcount since there could be a homog K_3.

We bound the inner summation. Since v is of degree $m-1$ we can renumber the colors as $1, 2, \ldots, m-1$. Note that $\sum_{c=1}^{m-1} \deg_c(v) = m-1$ and $(\forall c)[\deg_c(v) \leq d]$. This sum is maximized when $d = \deg_1(v) = \deg_2(v) = \cdots = \deg_{(m-1)/d}(v)$ and the rest of the $\deg_c(v)$'s are 0. Hence we have

$$b \leq \sum_{v\in[m]} \sum_{c=1}^{m-1} \binom{\deg_c(v)}{2} \leq \sum_{v\in[m]} \frac{m-1}{d}\binom{d}{2} < m\frac{m}{d}\frac{d^2}{2} = \frac{dm^2}{2}.$$

∎

Lemma 3.6 *Let $COL : \binom{[m]}{2} \rightarrow \omega$ that has b bad triples. Let $1 \leq m' \leq m$. There exists an m'-sized set of vertices with $\leq b\left(\frac{m'}{m}\right)^3$ bad triples.*

Proof: Pick a set X of size m' at random. Let E be the expected number of bad triples. Note that

$$E = \sum_{\{v_1,v_2,v_3\} \text{ bad}} \text{Prob that } \{v_1, v_2, v_3\} \subseteq X \ .$$

Let $\{v_1, v_2, v_3\}$ be a bad triple. The probability that all three nodes are in X is bounded by

$$\frac{\binom{m-3}{m'-3}}{\binom{m}{m'}} \le \frac{m'(m'-1)(m'-2)}{m(m-1)(m-2)} \le \left(\frac{m'}{m}\right)^3.$$

Hence the expected number of bad triples is $\le b(\frac{m'}{m})^3$. Therefore there must exist some X that has $\le b(\frac{m'}{m})^3$ bad triples.

∎

Note 3.7 The above theorem presents the user with an interesting tradeoff. She wants a large set with few bad triples. If m' is large then you get a large set, but it will have many bad triples. If m' is small then you won't have many bad triples, but m' is small. We will need a Goldilocks-m' that is just right.

4 The Asymmetric Weak Canonical Ramsey Theorem

Theorem 4.1 *There exists C such that, for all k_1, k_2,*

$$WER(k_1, k_2) \le \frac{(Ck_2)^{6k_1-18}}{(\log k_2)^{2k_1-6}}.$$

Proof:

Let n, m, m', m'', δ be parameters to be determined later. They will be functions of k_1, k_2. Let $COL : \binom{[n]}{2} \to \omega$.

Intuition: In the usual proofs of Ramsey's Theorem (for two colors) we take a vertex v and see which of $\deg_{RED}(v)$ or $\deg_{BLUE}(v)$ is large. One of them must be at least half of the size of the vertices still in play. Here we change this up:

- Instead of taking a particular vertex v we ask if there is *any* v and any color c such that $\deg_c(v)$ is large.
- What is large? Similar to the proof of Ramsey's theorem it will be a fraction of what is left, a fraction δ which we will pick later. Unlike the proof of Ramsey's theorem δ will depend on k_2.
- In the proof of Ramsey's theorem we were guaranteed that one of $\deg_{RED}(v)$ or $\deg_{BLUE}(v)$ is large. Here we have no such guarantee. We may fail. In that case something else happens and leads to a rainbow set!

CONSTRUCTION
Phase 1:
Stage 0:

1. $V_0 = \emptyset$. The set V_0 will be vertices such that the edges from them to all vertices remaining are the same color.
2. $N_0 = [n]$.
3. COL' is not defined on any points.

Stage i: Assume that V_{i-1}, and N_{i-1} are already defined.

If there exists $x \in N_{i-1}$ and c a color such that $\deg_c(x) \geq \delta N_{i-1}$ then do the following:

$$\begin{aligned} V_i &= V_{i-1} \cup \{x\} \\ N_i &= \{v \in N_{i-1} : COL(x,v) = c\} \\ x_i &= x \\ COL'(x_i) &= c \end{aligned}$$

Note that $|N_i| \geq \delta |N_{i-1}|$ and $|V_i| = i$. If $i = k_1 - 3$ then goto Phase 2.

If no such x, c exist then goto Phase 3.
End of Phase 1

Phase 2: Since we are in Phase 2 we have $V = V_{k_1-3}$. Let

$$V = \{x_1, x_2, \ldots, x_{k_1-3}\}.$$

(This is the order the elements came into V, not the numeric order.) By construction V is a whomog set of size $k_1 - 3$. Note that for all elements $x \in N_{k_1-3}$, $COL(x_i, x) = c_i$.

We will need $|N_{k_1-3}| \geq 3$. Since $|N_{k_1-3}| \geq \delta^{k_1-3} n$ we have the constraint

$$n \geq \frac{3}{\delta^{k_1-3}}.$$

Let x_{k_1-2}, x_{k_1-1}, and x_{k_1} be three points from N_{k_1-3}. Let H be (in this order)

$$H = \{x_1, x_2, \ldots, x_{k_1-3}, x_{k_1-2}, x_{k_1-1}, x_{k_1}\}.$$

H is clearly whomog. (Recall that in a whomog set of size k_1 we do not care if $COL(x_{k_1-2}, x_{k_1-1}) = COL(x_{k_1-2}, x_{k_1})$.)
End of Phase 2

Phase 3: Let i be the stage where you entered Phase 3. Clearly $|N_i| \geq \delta^i n$. We will need that $|N_i| \geq m$. Since $i \leq k_1 - 4$ we have the constraint $m \leq \delta^{k_1-4} n$. We satisfy it by insisting that

$$n \geq \frac{m}{\delta^{k_1-4}}.$$

Recall that we also want $n \geq \frac{3}{\delta^{k_1-3}}$. In order for the $n \geq \frac{m}{\delta^{k_1-4}}$ to imply the inequality $n \geq \frac{3}{\delta^{k_1-3}}$ we have the constraint

$$m \geq \frac{3}{\delta}.$$

Combining $n \geq \frac{m}{\delta^{k_1-4}}$ and $m \geq \frac{3}{\delta}$ we obtain

$$n \geq \frac{3}{\delta^{k_1-3}}.$$

Let $|N| = m_0$. Let COL be the coloring restricted to $\binom{N}{2}$. We can assume the colors are a subset of $\{1, \ldots, \binom{m_0}{2}\}$.

Since we are in Phase 3 we know that, for all $v \in N$, for all colors c, $\deg_c(v) \leq \delta m_0$. Note also that, for any vertex $v \in N$,

$$m_0 - 1 = \sum_{c=1}^{\binom{m_0}{2}} \deg_c(v) \leq \sum_{c=1}^{\binom{m_0}{2}} \delta m \leq \delta m_0^2 m.$$

Hence we need

$$\delta > \frac{m_0 - 1}{m_0^2 m}.$$

Since $m_0 \geq m$ the constraint $m \geq \frac{3}{\delta}$ already implies this.

Note that COL is a coloring of $\binom{N}{2}$ such that for every $v \in N$ and color c, $\deg_c(v) \leq \delta m_0$. Hence, by Lemma 3.5, there are at most

$$\frac{\delta m_0 \times m_0^2}{2} \leq \delta m_0^3$$

bad triples (we ignore the denominator of 2 since it makes later calculations easier and only affects the constant).

By Lemma 3.6 there exists $X \subseteq N$ of size m' that has

$$b < \delta m_0^3 \times \left(\frac{m'}{m_0}\right)^3 = \delta(m')^3$$

bad triples. Note also that this number is independent of m_0 which will enable us to pick our parameters without knowing m_0.

We set m' such that the number of bad triples is so small that we can just remove one point from each. We will have a set X of size m'' with *no* bad triples.

Recall that the number of bad triples is $\delta(m')^3$. Hence we want

$$m' - \delta(m')^3 \geq m''$$

so we have the constraint

$$\delta \le \frac{m' - m''}{(m')^3}.$$

We will now set the parameters. Since we will use Lemma 3.3 it would be difficult to optimize the parameters. Hence we pick parameters that are easy to work with.

We will use Lemma 3.3 on X to obtain a rainbow set of size k_2. Hence we take

$$m'' = \frac{Ak_2^3}{\log k_2}$$

where A is chosen to (1) make m'' large enough to satisfy the premise of Lemma 3.3, (2) make m'' an integer, and (3) make m', m, n, which will be functions of m'', integers.

We take

$m' = 1.5m''$. This is the value that minimize δ though this does not matter .

$$\delta = \frac{m' - m''}{(m'')^3} = \frac{1}{B(m'')^2}$$

where B is an appropriate constant.

$$m = \frac{3}{\delta}$$

$$n = \frac{3}{\delta^{k_1-3}} = 3(Bm'')^{2(k_1-3)} = \frac{(Ck_2)^{6k_1-18}}{(\log k_2)^{2k_1-6}}$$

where C is an appropriate constant. ∎

5 Lemmas from Geometry

Def 5.1 Let $d \in \mathsf{N}$.

1. If $p, q \in \mathsf{R}^d$ then let $|p - q|$ be the Euclidean distance between p and q.
2. Let $p_1, \ldots, p_n$ be points in R^d. $(p_1, \ldots, p_n)$ is a *cool sequence* if for all $1 \le i \le n-3$, for all $i < j \le n$ $|p_i - p_j|$ is determined solely by p_i. (Formally: for all $1 \le i \le n-3$ there exists L_i such that, for all $i+1 \le j \le n$, $|p_i - p_j| = L_i$.) We intentionally have $1 \le i \le L - 3$. We do not care if $|p_{n-2} - p_{n-1}| = |p_{n-2} - p_n|$.
3. The sphere with center $x \in \mathsf{R}^{d+1}$ and radius $r \in \mathsf{R}^+$ is the set

$$\{y \in \mathsf{R}^{d+1} : |x - y| = r\}.$$

If the sphere is completely contained in an $(n+1)$-dimensional plane then the sphere is called an n-sphere.

Note that if $(p_1, \ldots, p_n)$ is cool then $(p_2, \ldots, p_n)$ is cool. We use this implicitly without mention.

The following lemma is well known.

Lemma 5.2 *Let S be an n-sphere. Let $x \in S$ and $r \in \mathsf{R}^+$. The set*

$$\{y \in S : |x - y| = r\}$$

is either an $(n-1)$-sphere or is empty.

Lemma 5.3 *Let $d \in \mathsf{N}$ and let $n < d$. There does not exist a cool sequence $p_1, \ldots, p_{n+3}$ on an n-sphere.*

Proof: We prove this by induction on n.
Base Case $n = 0$: Assume, by way of contradiction, that (p_1, p_2, p_3) form a cool sequence on a 0-sphere. A 0-sphere is a set of two points, hence this is impossible. (Note that being a cool sequence did not constraint (p_1, p_2, p_3) at all.)
Induction Hypothesis: The theorem holds for $n - 1$ and $n \geq 2$.
Induction Step: We prove the theorem for n. Assume, by way of contradiction, that $(p_1, \ldots, p_{n+2})$ form a cool sequence on an n-sphere. Since $|p_1 - p_2| = |p_1 - p_3| = \cdots = |p_1 - p_{n+2}|$ we know, by Lemma 5.2, that $p_2, p_3, \ldots, p_{n+2}$ are on an $(n-1)$-sphere. Since $p_2, \ldots, p_{n+2}$ is a cool sequence this is impossible by the induction hypothesis. ∎

Note 5.4 The following related statement seems to be well known: *if there are $d + 2$ points in R^d then it is not the case that all $\binom{d+2}{2}$ distances are the same.* We have not been able to locate this result in an old fashion journal (perhaps its behind a paywall); however, there is a proof at mathoverflow.net here:
`http://mathoverflow.net/questions/30270/`
`maximum-number-of-mutually-equidistant-points-`
`in-an-n-dimensional-euclidean-space`

Notation Warning: The n we use in the next lemma is not connected to the n we used in the definitions or lemma above.

Lemma 5.5 *Let $d \in \mathsf{N}$. Let $p_1, \ldots, p_n$ be points in R^d. Color $\binom{[n]}{2}$ via $COL(i, j) = |p_i - p_j|$. This coloring has no whomog set of size $d + 3$.*

Proof: Assume, by way of contradiction, that there exists a whomog set of size $d + 3$. By renumbering we can assume the whomog set is $\{1, \ldots, d + 3\}$. Clearly $p_1, \ldots, p_{d+3}$ form a cool sequence. Note that our not-caring about $COL(d + 1, d + 2) = COL(d + 1, d + 3)$ is reflected in our not-caring about $|p_{d+1} - p_{d+2}| = |p_{d+1} - p_{d+3}|$.

Since $|p_1 - p_2| = |p_1 - p_3| = \cdots = |p_1 - p_{d+3}|$, $p_2, \ldots, p_{d+3}$ are on the $(d-1)$-sphere (centered at p_1). This contradicts Lemma 5.3. ∎

6 Lower Bound on $h_{2,d}(n)$

Theorem 6.1 *For all* $d \geq 1$, $h_{2,d}(n) = \Omega((n^{1/(6d)}(\log n)^{1/3})/d^{1/3})$.

Proof: Let $P = \{p_1, \ldots, p_n\}$ be n points in R^d. Let $COL : \binom{[n]}{2} \rightarrow \mathsf{R}$ defined by $COL(i,j) = |p_i - p_j|$.

Let k be the largest integer such that $n \geq WER(k, d+3)$. By Theorem 4.1 there is either a whomog set of size $d+3$ or a rainbow set of size k. By Lemma 5.5 there cannot be a whomog set of size $d+3$, hence must be a rainbow set of size k. This is the set we seek.

How big does k have to be? Rewrite Theorem 4.1 as

$$WER(k_1, k_2) \leq \left(\frac{Ck_2^3}{\log k_2}\right)^{2k_1-6}.$$

We need

$$n \leq \left(\frac{Ck^3}{\log k}\right)^{2d}$$

to guarantee a rainbow set of size k or a whomog set of size $d+3$ (which can't happen). Clearly $k = \Omega((n^{1/(6d)}(\log n)^{1/3})/d^{1/3})$ suffices.

▮

7 Lower Bounds on $h_{3,2}$ and $h_{3,3}$

For the problem of $h_{2,d}$ we used (1) upper bounds on the asymmetric weak canonical Ramsey Theorem and (2) a geometric lemma. Here we will use (1) upper bounds on the standard asymmetric canonical Ramsey Theorem and (2) a geometric lemma. This is because we have not been able to get other versions of the canonical Ramsey theorem to yield better bounds.

7.1 The Asymmetric 3-ary Canonical Ramsey Theorem

Def 7.1 Let $COL : \binom{[n]}{a.} \rightarrow \omega$. Let $V \subseteq [n]$.

1. Let $I \subseteq \{1, \ldots, a\}$. The set V is *I-homogenous* (henceforth *I-homog*) if for all $x_1 < \cdots < x_a \in \binom{[n]}{a}$ and $y_1 < \cdots < y_a \in \binom{[n]}{a}$,

$$(\forall i \in I)[x_i = y_i] \text{ iff } COL(x_1, \ldots, x_a) = COL(y_1, \ldots, y_a).$$

   Informally, the color of an element of $\binom{[n]}{a}$ depends exactly on the coordinates in I.

2. The set V is *rainbow* if every edge in $\binom{V}{a}$ is colored differently. Note that this is just an I-homog set where $I = \{1, \ldots, a\}$.

We will need the asymmetric hypergraph Ramsey numbers and a-ary Erdős-Rado numbers.

Def 7.2 Let $a \ge 1$. Let $k_1, k_2, \ldots, k_c \ge 1$.

1. Let $COL : \binom{[n]}{a} \rightarrow [c]$. (Note that there is a bound on the number of colors.) Let $V \subseteq [n]$. The set V is *homog with color* i if COL restricted to $\binom{V}{a}$ always returns i.
2. $R_a(k_1, k_2, \ldots, k_c)$ is the least n such that, for all $COL : \binom{[n]}{a} \rightarrow [c]$, there exists $1 \le i \le c$ and a homog set of size k_i with color i. $R_a(k_1, k_2, \ldots, k_c)$ is known to exist by the hypergraph Ramsey Theorem.
3. $ER_a(k_1, k_2)$ is the least n such that, for all $COL : \binom{[n]}{a} \rightarrow \omega$, there exists either (1) an $I \subset [a]$ and an I-homog set of size k_1, or (2) a rainbow set of size k_2. $ER_a(k_1, k_2)$ is known to exist by the a-ary canonical Ramsey theorem.

Lemma 7.3 *Let* $a \ge 3$, $c \ge 2$, *and* $k_1, \ldots, k_c \ge 1$. *Let* $P = k_1 \cdots k_{c-1}$ *and* $S = k_1 + \cdots + k_{c-1}$.

1. $R_a(k_1, k_2, \ldots, k_c) \le c^{R_{a-1}(k_1-1, k_2-1, \ldots, k_c-1)^{a-1}}$.
2. $$ER_3(k_1, k_2) \le R_4\left(6, 6, 6, 6, k_1, k_1, k_1, k_1, \left\lceil \frac{k_1^3}{4} \right\rceil, \left\lceil \frac{k_1^3}{4} \right\rceil, 2k_1^3, \left\lceil \frac{k_2^5}{36} \right\rceil\right).$$
3. *Let* $$Z = \{\sigma \in [c]^* : \textit{ for all } i \in [c],\ \sigma \textit{ contains at most } k_i - 1\ i\textit{'s} \}.$$ *Then* $$\sum_{\sigma \in Z} |\sigma| \le P(k_c + S)^{S+2}.$$
4. $R_3(k_1, \ldots, k_c) \le c^{P(k_c+S)^{S+2}}$.
5. $R_4(k_1, \ldots, k_c) \le c^{c^{3P(k_c+S-c)^{S+2-c}}}$.
6. *For almost all* k, $ER_3(e, k) \le 2^{2^{k^{12.5e^3+20e+81}}}$.

Proof:
1) Erdős-Rado [12, 17, 18] showed that $R_a(k, k) \le 2^{\binom{R_{a-1}(k-1,k-1)+1}{a-1}} + a - 2$. This can be modified to show
$$R_a(k_1, k_2, \ldots, k_c) \le c^{\binom{R_{a-1}(k_1-1,\ldots,k_c-1)}{a-1}+a-2}.$$
Our result easily follows.

2) Lefmann and Rodl [20] obtained a bound on $ER_3(k,k)$ in terms of 4-hypergraph Ramsey numbers. Our result is obtained by a straightforward analysis and modification of their proof.

3) Clearly

$$\begin{aligned}\sum_{\sigma\in Z}|\sigma| &= \sum_{j_1=0}^{k_1-1}\cdots\sum_{j_c=0}^{k_c-1}(j_1+\ldots+j_c)\tfrac{(j_1+\ldots+j_c)!}{j_1!\cdots j_c!} \le \sum_{j_1=0}^{k_1-1}\cdots\sum_{j_c=0}^{k_c-1}(k_c+S)\tfrac{(k_c+S)!}{k_c!}\\ &\le P\sum_{j_c=0}^{k_c-1}(k_c+S)^{S+1} \le Pk_c(k_c+S)^{S+1} \le P(k_c+S)^{S+2}\end{aligned}$$

4) Conlon, Fox, and Sudakov [2] have the best known upper bounds on $R_3(k,k)$. Gasarch, Parrish, Sandow [17] have done a straightforward analysis of their proof to extend it to c colors. A modification of that proof yields

$$R_3(k_1,\ldots,k_c) \le c^{\sum_{\sigma\in Z}|\sigma|}.$$

Our result follows.
5) This follows from parts 1 and 4. We could obtain a better result by replacing P by $(k_1-1)\cdots(k_c-1)$ but that would not help us later.
6) Note that, using $\lceil x\rceil \le x+1$,

$$6+6+6+6+e+e+e+e+\left\lceil\frac{e^3}{4}\right\rceil+\left\lceil\frac{e^3}{4}\right\rceil+2e^3 \le 2.5e^3+4e+26$$

Let $s(e)=2.5e^3+4e+26$. Let $p(e)$ be the product of these terms. By parts 2 and 5, for k large, we have the following.

$$ER_3(e,k) \le 12^{12^{3p(e)(k^5+s(e)-12)^{s(e)-10}}} \le 12^{12^{3p(e)(2k^5)^{s(e)-10}}} \le 12^{12^{(6p(e)k^5)^{s(e)-10}}} \le 2^{2^{(36p(e)k^5)^{s(e)-10}}}$$

Let $f(e)=(36p(e))^{s(e)-10}$. Then we have

$$ER_3(e,k) \le 2^{2^{f(e)k^{5s(e)-50}}} \le 2^{2^{k^{5s(e)-49}}} \le 2^{2^{k^{12.5e^3+20e+81}}}.$$

∎

7.2 Geometric Lemmas

Def 7.4 Let $d\in\mathsf{N}$. If $p,q,r\in\mathsf{R}^d$ then let $AREA(p,q,r)$ be the area of the triangle with vertices p,q,r.

We will need Lemma 4 of [4] whose proof is in the appendix of that paper. They credit [14], which is unavailable, with the proof.

Lemma 7.5 *Let C_1,C_2,C_3 be three cylinders with no pair of parallel axis in Then $C_1\cap C_2\cap C_3$ consists of at most 8 points.*

Lemma 7.6

1. *Let $p_1, \ldots, p_n$ be points in R^2, no three collinear. Color $\binom{[n]}{3}$ via $COL(i,j,k) = AREA(p_i, p_j, p_k)$. If $I \subset \{1,2,3\}$ then this coloring has no I-homog set of size 6.*

2. *Let $p_1, \ldots, p_n$ be points in R^3, no three collinear. Color $\binom{[n]}{3}$ via $COL(i,j,k) = AREA(p_i, p_j, p_k)$. If $I \subset \{1,2,3\}$ then this coloring has no I-homog set of size 13.*

Proof:
1) Assume, by way of contradiction, there exists an I-homog set of size 6. By renumbering we can assume the I-homog set is $\{1,2,3,4,5,6\}$.

Case 1: $I = \{1\}$, $\{1,2\}$, or $\{2\}$.

We have $AREA(p_1, p_2, p_4) = AREA(p_1, p_2, p_5)$. Thus p_4 and p_5 are either on a line parallel to p_1p_2 or are on different sides of p_1p_2. In the later case the midpoint of p_4p_5 is on p_1p_2.

We have $AREA(p_1, p_3, p_4) = AREA(p_1, p_3, p_5)$. Thus p_4 and p_5 are either on a line parallel to p_1p_3 or are on different sides of p_1p_3. In the later case the midpoint of p_4p_5 is on p_1p_3.

We have $AREA(p_2, p_3, p_4) = AREA(p_2, p_3, p_5)$. Thus p_4 and p_5 are either on a line parallel to p_2p_3 or are on different sides of p_2p_3. In the later case the midpoint of p_4p_5 is on p_2p_3.

One of the following must happen.

- Two of these cases have p_4, p_5 on the same side of the line. We can assume that p_4, p_5 are on a line parallel to both p_1p_2 and p_1p_3. Since p_1, p_2, p_3 are not collinear there is no such line.

- Two of these cases have p_4, p_5 on opposite sides of the line. We can assume that the midpoint of p_4p_5 is on both p_1p_2 and p_1p_3. Since p_1, p_2, p_3 are not collinear the only point on both p_1p_2 and p_1p_3 is p_1. So the midpoint of p_4, p_5 is p_1. Thus p_4, p_1, p_5 are collinear which is a contradiction.

Note that for $I = \{1\}$, $\{1,2\}$, or $\{2\}$ we used the line-point pairs

$$\{p_1p_2, p_1p_3, p_2p_3\} \times \{p_4, p_5\}.$$

For the rest of the cases we will just specify which line-point pairs to use.

Case 2: $I = \{3\}$ or $\{2,3\}$. Use

$$\{p_4p_5, p_3p_5, p_3p_4\} \times \{p_1, p_2\}.$$

Case 3: $I = \{1,3\}$ Use

$$\{p_1p_4, p_1p_5, p_1p_6\} \times \{p_2, p_3\}.$$

This is the only case that needs 6 points.

2) Assume, by way of contradiction, that there exists an I-homog set of size 13. By renumbering we can assume the I-homog set is $\{1, \ldots, 13\}$.

Case 1: $I = \{1\}$, $\{1, 2\}$, or $\{2\}$.

We have $AREA(p_1, p_2, p_4) = AREA(p_1, p_2, p_5) = \cdots = AREA(p_1, p_2, p_{12})$. Hence $p_4, \ldots, p_{12}$ are all on a cylinder with axis p_1p_2.

We have $AREA(p_1, p_3, p_4) = AREA(p_1, p_3, p_5) = \cdots = AREA(p_1, p_3, p_{12})$. Hence $p_4, \ldots, p_{12}$ are all on a cylinder with axis p_1p_3.

We have $AREA(p_2, p_3, p_4) = AREA(p_2, p_3, p_5) = \cdots = AREA(p_2, p_3, p_{12})$. Hence $p_4, \ldots, p_{12}$ are all on a cylinder with axis p_2p_3.

Since p_1, p_2, p_3 are not collinear the three cylinders mentioned above satisfy the premise of Lemma 7.5. By that lemma there are at most 8 points in the intersection of the three cylinders. However, we just showed there are 9 such points. Contradiction.

Note that for $I = \{1\}$, $\{1, 2\}$, or $\{2\}$ we used the line-point pairs

$$\{p_1p_2, p_1p_3, p_2p_3\} \times \{p_4, \ldots, p_{12}\}.$$

For the rest of the cases we will just specify which line-point pairs to use.

Case 2: $I = \{3\}$ or $\{2, 3\}$. Use

$$\{p_{11}p_{12}, p_{10}p_{12}, p_{10}p_{11}\} \times \{p_1, \ldots, p_9\}.$$

Case 3: $I = \{1, 3\}$ Use

$$\{p_1p_{11}, p_1p_{12}, p_1p_{13}\} \times \{p_2, \ldots, p_{10}\}.$$

This is the only case that needs 13 points.

■

7.3 Lower Bounds on $h_{3,2}(n)$ and $h_{3,3}(n)$

Theorem 7.7

1. $h_{3,2}(n) \geq \Omega((\log \log n)^{1/2901})$.
2. $h_{3,3}n) \geq \Omega((\log \log n)^{1/27804})$.

Proof:

a) Let $P = \{p_1, \ldots, p_n\}$ be n points in R^2. Let COL be the coloring of $\binom{[n]}{3}$ defined by $COL(i, j, k) = AREA(p_i, p_j, p_k)$.

Let k be the largest integer such that

$$n \geq ER_3(6, k).$$

By Lemma 7.3 it will suffice to take $k = \Omega((\log \log n)^{1/2901})$. By the definition of $ER_3(6, k)$ there is either a I-homog set, with $I \subset \{1, 2, 3\}$, of size 6 or a rainbow set of size k. By Lemma 7.6.a there cannot be such an I-homog set, hence must be a rainbow set of size k.

b) Let $P = \{p_1, \ldots, p_n\}$ be n points in R^3. Let COL be the coloring of $\binom{[n]}{3}$ defined by $COL(i, j, k) = AREA(p_i, p_j, p_k)$.

Let k be the largest integer such that

$$n \geq ER_3(13, k).$$

By Lemma 7.3 it will suffice to take $k = \Omega((\log \log n)^{1/27804})$. By the definition of $ER_3(13, k)$ there is either a I-homog set, with $I \subset \{1, 2, 3\}$, of size 13 or a rainbow set of size k. By Lemma 7.6.b there cannot be such an I-homog set, hence must be a rainbow set of size k. ▮

Note 7.8 A more careful analysis of the bound on $ER_3(k)$ from [20], perhaps using some weak version, may lead to larger exponents in the lower bounds for $h_{3,2}$ and $h_{3,3}$; however, such an analysis will not lead to an improvement from $\log \log n$ to $\log n$.

To obtain similar bounds on $h_{3,d}(n)$ we have the needed bounds on the asymmetric canonical Ramsey numbers but do not have the needed geometric lemmas. We make the following conjecture.

1. There exists a function $f(d)$ such that the following is true: Let $p_1, \ldots, p_n$ be points in R^d, no three collinear. Color $\binom{[n]}{3}$ via $COL(i, j, k) = AREA(p_i, p_j, p_k)$. If $I \subset \{1, 2, 3\}$ then this coloring has no I-homog set of size $f(d)$.

2. There exists a function $\epsilon(d)$ such that, for all d, $h_{3,d}(n) \geq \Omega((\log \log n)^{\epsilon(d)})$. (This follows from the conjecture above.)

8 Speculation about Higher Dimensions

To get lower bounds on $h_{a,d}(n)$ using our approach you need the following:

- Upper bounds on $ER_a(k_1, k_2)$. The upper bound on $ER_a(k, k)$ involves R_{2a-1} which lead to a tower of height $2a - 1$. The bound on $ER(k_1, k_2)$ (if $k_1 \ll k_2$ which is our case) leads to a tower of height $2a - 2$. So this ingredient is already known though perhaps could be improved upon.

- The following geometric lemma: There exists a function $f(a, d)$ such that the following is true: Let $p_1, \ldots, p_n$ be points in R^d, no a in the same $(a - 2)$-dimensional space. Color $\binom{[n]}{a}$ via $COL(i_1, \ldots, i_a) = VOLUME(p_{i_1}, \ldots, p_{i_a})$. If $I \subset \{1, \ldots, a\}$ then this coloring has no I-homog set of size $f(a, d)$.

We conjecture that the geometric lemma needed is true. If that is the case then the following is true: For all a, d there is a constant $\epsilon_{a,d}$ such that

$$(\forall a \geq 4)[h_{a,d}(n) = \Omega((\log^{(2a-2)} n)^{\epsilon_{a,d}})].$$

Our results on $h_{2,d}$ (called $h_{2,d}$ throughout this paper) and $h_{3,d}$ are much better than the conjecture. There are several reasons for this. We explain them and speculate on which ones can be used for $h_{a,d}$.

1. There are already very good bounds on $ER_2(k,k)$ and $ER_3(k,k)$. Hence whatever modifications we did had a good starting point. For $a \geq 4$, the best bound for $ER_a(k,k)$ involves R_{2a-1}. Finding better bounds on $ER_a(k,k)$ for $a \geq 4$ seems like a hard problem.

2. For $h_{2,d}$ we looked at $WER(k_1,k_2)$. There are two aspects to $WER(k_1,k_2)$ that made for better bounds.

   (a) The definition of $WER(k_1,k_2)$ only had the implication about the coloring going in one direction. This modification of the definition of ER_a may lead to better bounds, perhaps replacing R_{2a-1} with R_{a+c} for some constant c.

   (b) The definition of $WER(k_1,k_2)$ used a different definition of homog. We doubt this can be adapted to the a-ary case even for $a = 3$.

The upshot is that looking at a definition of $WER_a(k_1,k_2)$ that involves only one side of the coloring implication may lead to better bounds. However, the geometric lemmas are also needed.

9 Open Questions

1. The upper and lower bounds for $h_{2,d}$, $h_{3,2}$, $h_{3,3}$ are very far apart. Close the gap!

2. We obtain $h_{2,1}(n) = \Omega(n^{1/6}(\log n)^{1/3})$. The known result, $h_{2,1}(n) = \Theta(n^{1/2})$, has a rather difficult proof. It would be of interest to obtain an easier proof of either the known result or a weaker version of it that is stronger than what we have. An easy probabilistic argument yields $h_{2,1}(n) = \Omega(n^{1/4})$.

3. Obtain the geometric lemmas needed to get nontrivial lower bounds on (1) $h_{3,d}$ for $d \geq 4$, and (2) $h_{a,d}$ for $d \geq 4$, and $a \geq d$. Combining this with known results would get lower bounds of the form $\Omega((\log^{(2a-1)}(n))^{\epsilon_{a,d}})$. Obtain upper bounds on some weak version of canonical Ramsey numbers that will lead to less iterated logs in the final result. See Section 8 for more thoughts on this.

4. Look at a variants of $h_{a,d}(n)$ with different metrics on R^d or in other metric spaces entirely.

5. Look at a variant of $h_{a,d}(n)$, which we call $h'_{a,d}$, where the only condition on the points is that they are not all on the same $(a-2)$-dimensional space. Using the $n^{1/d}\times\cdots\times n^{1/d}$ grid it is easy to show that, for $a,d \ll n$, $h'_{a,d}(n) \le O(n^{(a-1)/a})$.

10 Acknowledgments

We would like to thank Tucker Bane, Andrew Lohr, Jared Marx-Kuo, and Jessica Shi for helpful discussions. We would like to thank David Conlon for discussion of the bounds on the asymmetric hypergraph Ramsey Theorem.

References

[1] N. Alon, H. Lefmann, and V. Rodl. On an anti-Ramsey type result. In *Sets, graphs, and numbers*, pages 9–22, 1991. Coll. Math. Soc. János Bolyai 60.

[2] D. Conlon, J. Fox, and B. Sudakov. Hypergraph Ramsey numbers. *Journal of the American Mathematical Society*, pages 247–266, 2010. `www.dpmms.cam.ac.uk/~dc340`.

[3] H. Croft. 9 points and 7 point configurations in 3-space. *Proceedings of the London Mathematical Society*, 12:400–424, 1962. `plms.oxfordjournals.org/content/s1-16/1/local/back-matter.pdf`.

[4] A. Dumitrescu, M. Sharir, and C. Toth. Extremal problems on triangle area in two and three dimensions, 2008. `http://arxiv.org/pdf/0710.4109v1.pdf`.

[5] P. Erdős. On sets of distances of n points. *The American Mathematical Monthly*, 53:248–250, 1946. `http://www.cs.umd.edu/~gasarch/erdos\_dist/erdos\_dist.html` or `http://www.jstor.org/`.

[6] P. Erdős. Nehany geometriai problemeraol (Hungarian). *Mat. Lapok*, 8:86–92, 1957.

[7] P. Erdős. On sets of distances of n points. *The American Mathematical Monthly*, 77:738–740, 1970.

[8] P. Erdős. On some metric and combinatorial geometric problems. *Discrete Mathematics*, 60:147–153, 1986.

[9] P. Erdős and L. Kelly. Elementary problem 735. *The American Mathematical Monthly*, 54:227–229, 1947.

[10] P. Erdős and G. Purdy. Extremal problems in combinatorial geometry. In G. and-Grotschel and Lovasz, editors, *Handbook of Combinatorics*. North Holland, 1995.

[11] P. Erdős and R. Rado. A combinatorial theorem. *Journal of the London Mathematical Society*, 25:249–255, 1950. `http://jlms.oxfordjournals.org/`.

[12] P. Erdős and R. Rado. Combinatorial theorems on classifications of subsets of a given set. *Proceedings of the London Mathematical Society*, 2:417–439, 1952.

[13] P. Erdős and P. Turan. On a problem of Sidon. *Journal of the London Mathematical Society*, 16:292–296, 1941.

[14] Ezra, Pach, and Sharir. Regular vertices on the union of cylinders, 2006.

[15] J. Gariabldi, A. Iosevich, and S. Senger. *The Erdős distance problem.* American Mathematics Society, 2010.

[16] W. Gasarch. A webpage on the Erdős distance problem. `http://www.cs.umd.edu/~gasarch/erdos_dist/erdos_dist.html`.

[17] W. Gasarch, A. Parrish, and S. Sinai. Three proof of the hypergraph Ramsey theorem (an expsosition), 2012. `http://arxiv.org/abs/1206.4257`.

[18] R. Graham, B. Rothschild, and J. Spencer. *Ramsey Theory.* Wiley, 1990.

[19] J. Komlos, Sulyok, and E. Szemeredi. Linear problems in combinatorial number theory. *Acta. Math. Acad. Soc. Hung.*, 26:113–121, 1975.

[20] H. Lefmann and V. Rodl. A canonical Ramsey numbers for coloring three-elements sets. In *Finite and infinite combinatorics in sets and logic*, volume 411, pages 237–247, 1991. Nato science series.

[21] H. Lefmann and V. Rodl. On Erdos-Rado numbers. *Combinatorica*, 15:85–104, 1995. `http://www.springer.com/new+%26+forthcoming+titles+%28default%29/journa%l/493`.