On sets forming Boolean algebras and partite hypergraphs

D. S. Gunderson
Bielefeld, Germany

V. Rödl
Emory University, Atlanta

A. Sidorenko
Courant Institute, New York

Abstract

Three classes of finite structures are related by extremal properties: d-partite hypergraphs, d-dimensional affine cubes of integers, and families of 2^d sets forming a d-dimensional Boolean algebra. We review extremal results for each of these classes and derive new ones for Boolean algebras and hypergraphs, many obtained by employing relationships between the three classes. The similarity in bounds for extremal problems in different classes is remarkable.

1 Introduction

The original purpose of our research was to determine extremal results for Boolean algebras of sets. We employed theorems and techniques from two other areas, extremal aspects of d-partite d-uniform hypergraphs, and some old, somewhat hard questions concerning integers. We review some of the facts used about integers, and develop extremal results for Boolean algebras and hypergraphs. Each area contains many interesting open questions, solutions to some of which would immediately yield improvements in each of the other two.

In this introduction, we outline the results in this paper. In the next section on hypergraphs, we establish some statements for later employment in theorems for both affine cubes and for Boolean algebras, and briefly survey some related facts. Of independent interest, new upper bounds for extremal numbers for partite hypergraphs are proved using a technique involving affine spaces. Section 3 outlines known related results concerning integers, some obtained by hypergraph proofs, some of which are used later for extremal problems on Boolean algebras. Section 4 is on Boolean algebras of sets, containing Ramsey and density theorems.

Let X be a finite set; $\mathcal{P}(X) = \{Y : Y \subseteq X\}$ denotes the power set of X and $[X]^s = \{S \subset X : |S| = s\}$. It will often be convenient to use the notation

$[n] = [1, n] = \{1, 2, \ldots, n\}$ and use $X = [n]$. For many mathematicians, the term "Boolean algebra" has a very specific meaning. For a finite set X, the collection $\mathcal{P}(X)$ is often referred to as the *Boolean algebra on* X. We will use the term "Boolean algebra" in a slightly different manner, however, the analogy should be clear.

Definition 1.0.1 *A collection $\mathcal{B} \subseteq \mathcal{P}(X)$ forms a d-*dimensional Boolean algebra *if and only if there exist pairwise disjoint sets $X_0, X_1, \ldots, X_d \in \mathcal{P}(X)$, all non-empty with perhaps the exception of X_0, so that*

$$\mathcal{B} = \left\{ X_0 \cup \bigcup_{i \in I} X_i : I \subseteq [1, d] \right\}.$$

The Boolean algebra generated by the sets $X_0, X_1, \ldots, X_d$ may be viewed as a lattice (of inclusion) with meet X_0, join $\cup_{i=0}^{d} X_i$, and d atoms $X_0 \cup X_1, \ldots, X_0 \cup X_d$. We refer to any d-dimensional Boolean algebra of sets as simply a $\mathcal{B}(d)$. From the definition, it follows that a $\mathcal{B}(d)$ contains 2^d elements. It is not uncommon to view a power set as a lattice with inclusion as the relation; under such an interpretation, the set $\mathcal{P}(X)$ of subsets of an n-element set X is the standard Boolean algebra of sets (which has, by our definition, dimension n).

Definition 1.0.2 *Given an n-element set X and positive integer d, define $r(d, n)$ to be the largest integer so that for every partition of a set X,*

$$\mathcal{P}(X) = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \ldots \cup \mathcal{F}_{r(d,n)},$$

into $r(d, n)$ color classes, one color class contains a $\mathcal{B}(d)$.

In Theorem 4.1.1 we make the observation that $r(1, n) = n$. The bounds for $r(2, n)$ are tight up to a constant; in Theorem 4.1.2, using a well known extremal result for graphs and a statement about Singer sets, we show that for n sufficiently large,

$$\frac{1}{\sqrt{2}} n^{1/2} \leq r(2, n) \leq (1 + o(1)) n^{1/2}.$$

For general d, we have Theorem 4.1.3, stating

$$cn^{1/2^d} \leq r(d, n) \leq n^{\frac{d}{2^d - 1}(1+o(1))},$$

where the lower bound follows from Theorem 4.3.1, a density result, and the upper bound uses a partition result for certain families of integers.

Definition 1.0.3 *Given an n-element set X and a positive integer d, define $b(n, d)$ to be the maximum size of a family $\mathcal{F} \subset \mathcal{P}(X)$ which does not contain a $\mathcal{B}(d)$.*

Note that a $\mathcal{B}(1)$ is simply a pair of sets, one contained in the other. By Sperner's theorem (see [39] for but one of the many elegant proofs),

$$b(n,1) = \binom{n}{\lfloor n/2 \rfloor} \sim (\sqrt{2/\pi}) n^{-1/2} \cdot 2^n.$$

Erdős and Kleitman [14] found that there exist constants c_1 and c_2 so that for n sufficiently large,

$$c_1 n^{-1/4} \cdot 2^n \leq b(n,2) \leq c_2 n^{-1/4} \cdot 2^n. \quad (1)$$

Voigt [49] asked about a general bound for $b(n,d)$; in this paper we show that for each $d \geq 1$ there exist positive constants c_1, c_2, so that for n sufficiently large,

$$c_1 n^{-\frac{d}{2^{d+1}-2}(1-o(1))} \cdot 2^n \leq b(n,d) \leq c_2 n^{-1/2^d} \cdot 2^n. \quad (2)$$

The lower bound in (2) appears as Theorem 4.2.1. The upper bound in (2) appears as Theorem 4.3.1, and is proved probabilistically; an old result (Theorem 2.2.4, due to Erdős) concerning extremal d-partite hypergraphs is used.

The proof of the lower bound given for $b(n,d)$ involves some famous results (and refinements thereof) regarding of integers and certain sum-sets, which we present in Section 3. Essentially, we use families of subsets determined only by their sizes, and these sizes are determined, in part, by families of integers containing no arithmetic progressions, and by *affine cubes*, which we now define.

Definition 1.0.4 *For $d+1$ positive integers $x_0, x_1, \ldots, x_d$, the collection*

$$H(x_0, x_1, \ldots, x_d) = \left\{ x_0 + \sum_{i \in I} x_i : I \subseteq [d] \right\}$$

is called a d-dimensional affine cube, or simply, an affine d-cube.

Hilbert [35] showed that for every r, d, there exists a least number $h(d,r)$ so that for every coloring $\chi : [1, h(d,r)] \to [1,r]$, there exists an affine d-cube monochromatic under χ. Szemerédi [46] (Lemma $p(\delta, l)$, p. 93) gave a density version of Hilbert's theorem (see also problem 14.12 in [38] for two proofs) and implicitly showed that if $A \subset [n]$ is chosen with at least $cn^{1-2^{1-d}}$ elements, then A contains an affine d-cube. In [8] it was shown that $h(2,r) = (1+o(1))r^2$. Also in [8], it was noted that there exist constants c_1 and c_2 so that

$$r^{c_1 d} \leq h(d,r) \leq r^{c_2^d}, \quad (3)$$

where $c_2 \sim 2.6$ follows from Hilbert's proof (using Fibonacci numbers). In [32] bounds (for $d > 2$) in (3) are improved to

$$r^{(1-o(1))(2^d-1)/d} \leq h(d,r) \leq cr^{2^{d-1}}. \quad (4)$$

The upper bound in (4) appears here as Corollary 3.5.1; the lower bound in (4) is Theorem 3.3.2. In [32] and [33], hypergraphs are also used to give different proofs of Szemerédi's density thoerem.

Definition 1.0.5 *A d-dimensional* uniform *Boolean algebra is a Boolean algebra of sets $\mathcal{B} = \{B_0 \cup \bigcup_{i \in I} B_i : I \subseteq [1, d]\}$ (where the B_i's are pairwise disjoint) which satisfies $|B_1| = |B_2| = \ldots = |B_d|$. Define $b_u(n, d)$ to be the maximum size of a family in $\mathcal{P}([n])$ which contains no d-dimensional uniform Boolean algebra.*

In Theorem 4.4.1, we show that for n large enough,

$$\frac{c}{n^{o(1)}} 2^n \leq b_u(n, d) \leq \epsilon \cdot 2^n. \qquad (5)$$

Again, the lower bound of (5) is proved by choosing sets according only to their size, and these sizes are determined by a Behrend set. The main tool used in the upper bound of (5) is a density version (proved using dynamical systems by Furstenberg and Katznelson [26]) of the celebrated Hales-Jewett theorem [34].

2 Density results for d-partite hypergraphs

Density results for d-partite hypergraphs have many applications. For example, we will use such results in the proof of an upper bound for $b(n, d)$, and in [32] and [33] they are employed in proving related results for affine d-cubes. Extremal results for hypergraphs are also of great independent interest. In this section, we examine some known results and methods and, introducing a technique which slightly improves some probabilistic 'constructions', give new bounds.

2.1 Notation

A *d-uniform hypergraph* is a pair $G = (V, \mathcal{E}) = (V(G), \mathcal{E}(G))$, with vertex set V and hyperedge set $\mathcal{E} \subset [V]^d$. Note that by this definition, each d-set from V may occur only once as a hyperedge; that is, we deal only with *simple* hypergraphs. An ordinary graph is a 2-uniform hypergraph, with hyperedges called, simply, edges.

For pairwise disjoint sets $X_1, X_2, \ldots, X_d$, let

$$G = (X_1, X_2, \ldots, X_d, \mathcal{E}(G))$$

denote a d-uniform hypergraph on vertex set $V(G) = \cup_{i=1}^d X_i$ and edge set $\mathcal{E}(G) \subseteq [V(G)]^d$, where for each $E \in \mathcal{E}(G)$ and each $i = 1, \ldots, d$, $|E \cap X_i| = 1$ holds. In this case, G is called a d-partite d-uniform hypergraph, and the sets $X_1, \ldots, X_d$ are called *partite* sets.

Let $K^{(d)}(n_1, n_2, \ldots, n_d)$ denote the complete d-partite d-uniform hypergraph on $\sum_{i=1}^{d} n_i$ vertices, partitioned into sets of sizes $n_1, n_2 \ldots, n_d$, and having $\prod_{i=1}^{d} n_i$ edges, each edge containing exactly one vertex from each partite set. (The "(d)" in the notation is not redundant, it depicts the number of vertices per hyperedge.) The complete d-partite d-uniform hypergraph with two vertices in each partite set, $K^{(d)}(2, 2, \ldots, 2)$, is of particular interest; the graph C_4, also denoted $K^{(2)}(2, 2)$ or $K_{2,2}$, is such an example.

For any d-uniform hypergraph $\mathcal{H}$, we let $\mathrm{ex}(n, \mathcal{H})$ denote the maximum number of d-hyperedges in a hypergraph on n vertices which does not contain a copy of $\mathcal{H}$. This number is also called the *Turan number*, or *extremal number* for $\mathcal{H}$. Turan's theorem gives the extremal numbers for complete graphs, and the Erdős-Stone theorem gives a good approximation of extremal numbers for graphs with chromatic number greater than two.

Few non-trivial bounds are known for extremal numbers of graphs which are 2-chromatic. Erdős and Simonovits [17] showed $\mathrm{ex}(n, Q^3) < cn^{8/5}$ (where Q^3 denotes the graph on 8 vertices and 12 edges corresponding to a cube); also close to our present interests, Erdős and Spencer [18] proved $\mathrm{ex}(n, K_{t,t}) \geq (1/2)n^{2-2/(t+1)}$. For more references, see any of many good surveys in, for example, [5], [22], or [44]. In what follows, we will be particularly interested in extremal numbers $\mathrm{ex}(n, K^{(d)}(2, 2, \ldots, 2))$.

2.2 Upper bounds for $\mathrm{ex}(n, K^{(d)}(2, 2, \ldots, 2))$

In 1938 Erdős and Klein [11] (see [13]) proved that $\mathrm{ex}(n, K_{2,2}) = \Theta(n^{3/2})$, and in 1966, Erdős, Rényi, and Sós [15] and Brown [7] showed $\mathrm{ex}(n, K_{2,2}) = \frac{1}{2}n^{3/2} + O(n^{4/3})$. We might mention that in 1954, Kővári, Sós, and Turán [37] proved that $\mathrm{ex}(n, K_{2,2}) \leq 1 + n + \frac{1}{2}n^{3/2}$, so it is interesting to see the "$O(n^{4/3})$" term in a later result.

By what is now a standard argument (for example, see [38]), counting pairs of vertices in neighborhoods, the following is easily shown.

Lemma 2.2.1 *For every $n \geq 4$, $ex(n, K_{2,2}) \leq \frac{n}{4}(1 + \sqrt{4n-3})$.*

So we see that $(1 + o(1))\frac{1}{2}n^{3/2}$ edges in an n-vertex graph suffices to produce a $K_{2,2}$. If a graph is equibipartite on n vertices, then only $(1 + o(1))\frac{1}{2\sqrt{2}}n^{3/2}$ edges is needed, as we see in the next lemma (whose proof uses the same idea as used for Lemma 2.2.1).

Lemma 2.2.2 *If G is an equibipartite graph on n vertices ($n/2$ in each partite set), and $|E(G)| > \frac{n}{4}(1 + \sqrt{2n-3})$, then G contains a $K_{2,2}$.*

Proof: Write $G = (X_1, X_2, E)$, where $|X_1| = n/2$, and let D be the average degree of vertices in X_1 (which is the same as in X_2).

If $\sum_{x\in X_1}\binom{\deg(x)}{2} > \binom{n/2}{2}$, one pair of vertices in X_2 is in two different neighborhoods of vertices from X_1. By Jensen's inequality,

$$\sum_{x\in X_1}\binom{\deg(x)}{2} > \frac{n}{2}\binom{D}{2},$$

and so it suffices to have

$$\frac{n}{2}\binom{2|E|/n}{2} > \binom{n/2}{2} \tag{6}$$

to guarantee a $K_{2,2}$. Simple calculations shows that $|E| > \frac{n}{4}(1+\sqrt{2n-3})$ is sufficient to guarantee (6). □

Increasing slightly the constant $\frac{1}{2\sqrt{2}}$ from Lemma 2.2.2 guarantees many copies of $K_{2,2}$; this can be seen probabilistically, however we choose to show a double counting argument used by [13] in his proof of Theorem 2.2.4. A similar technique was used in [33] in a restricted extremal result for $K^{(d)}(2,\dots,2)$'s, which in turn was used to show extremal results for affine d-cubes of integers (and implicitly, then could be used for Boolean algebras).

Lemma 2.2.3 *For any $\delta > 0$, and $n \geq \max\{4, 1/(4\delta^2)\}$ if G is an equibipartite graph on n vertices and*

$$|\mathcal{E}(G)| \geq (1+\delta)\frac{1}{2}n^{3/2}$$

then G contains $\binom{n/2}{2}$ copies of $C_4 \cong K_{2,2}$.

Proof: Let $G = (X_1, X_2, \mathcal{E})$ be an equibipartite graph with $|X_1| = |X_2| = a$, where $a = n/2 \geq 1/(8\delta^2)$. Assume that $|\mathcal{E}(G)| \geq (1+\delta)\sqrt{2}a^{3/2}$. For any $x \in V(G) = X_1 \cup X_2$, $\deg(x)$ denotes the degree of x in G; for $i,j \in V(G)$, let $\deg(i,j)$ denote the pairwise degree of i and j, that is, the number of common neighbors to i and j. The number of copies of $K^{(2)}(2,2) \cong C_4$ in G is

$$\begin{aligned}
\sum_{\{i,j\}\in[X_1]^2}\binom{\deg(i,j)}{2} &\geq \binom{a}{2}\binom{\sum_{[X_1]^2}\deg(i,j)/\binom{|X_1|}{2}}{2} \quad \text{(by Jensen's ineq.)},\\
&= \binom{a}{2}\binom{\frac{1}{\binom{a}{2}}\sum_{x\in X_2}\binom{\deg(x)}{2}}{2} \quad \text{(counting from } X_2\text{)}\\
&\geq \binom{a}{2}\binom{\frac{a}{\binom{a}{2}}\binom{|\mathcal{E}(G)|/|X_2|}{2}}{2} \quad \text{(again by Jensen's)},\\
&= \binom{a}{2}\binom{\frac{2}{a-1}\binom{(1+\delta)\sqrt{2a}}{2}}{2}
\end{aligned}$$

$$\geq \binom{a}{2},$$

where the last line follows because $a \geq \max\{2, 1/(8\delta^2)\}$.□

The following theorem due to Erdős was proved by induction, and in a more general setting (for arbitrary l, not just $l = 2$ as we state it here, see also [22], equation (4.2).) This result is a critical tool used in later giving an upper bound for $b(n, d)$, thereby increasing our interest in studying similar extremal results.

Theorem 2.2.4 (Erdős [13]) *For each d and n sufficiently large,*

$$ex(n, K^{(d)}(2, 2, \ldots, 2)) \leq n^{d - \frac{1}{2^{d-1}}}.$$

Examining the inductive step of the proof of Theorem 2.2.4 shows that if one can somehow improve the upper bound for some particular d, we get an improvement for every $d' > d$ thereafter.

Corollary 2.2.5 ([31]) *If for some $d > 2$, f is so that for n sufficiently large,*

$$ex(n, K^{(d-1)}(2, 2, \ldots, 2)) \leq n^{d-1-f},$$

then

$$ex(n, K^{(d)}(2, 2, \ldots, 2)) \leq n^{d-f/2}$$

holds for n sufficiently large.

In particular, it is well known (see, for example, [7]) that $\mathrm{ex}(n, C_4) = \frac{(1+o(1))}{2} n^{3/2}$, so hopes of actually employing Corollary 2.2.5 are, at present at least, remote.

2.3 Restricted upper bounds

We now consider a special class of d-partite hypergraphs.

Definition 2.3.1 *For positive integers $d \geq 2$ and a, let $\mathcal{G}(d, a)$ be the class of d-partite d-uniform hypergraphs $G = (X_1, X_2, \ldots, X_d, \mathcal{E}(G))$ which satisfy $|X_1| = a$ and for each $i = 2, \ldots, d$, $|X_i| = a^{2^{i-2}}$ (hence $|X_2| = a$ as well). Define $p(1, a) = \binom{a}{2}$, and for $d \geq 2$, define*

$$p(d, a) = \binom{a}{2} \prod_{i=2}^{d} \binom{a^{2^{i-2}}}{2},$$

the number of ways to pick two vertices from each partite set in a graph from $\mathcal{G}(d, a)$.

The following theorem has a coloring analogue, appearing in [28] (Lemma 5.6), which also yields a density result.

Theorem 2.3.2 ([33]) *For each integer $d \geq 2$ and real number $\epsilon > 0$, there exists a_0 so that for $a \geq a_0$, and any $G \in \mathcal{G}(d,a)$, if*

$$|\mathcal{E}(G)| \geq (1+\epsilon)2^{d-3/2} \cdot a^{(2^d-1)/2}$$

then G contains $p(d-1,a) = \Theta\left(\frac{a^{2^{d-1}}}{2^{d-1}}\right)$ $(a \to \infty)$ copies of $K^{(d)}(2,\ldots,2)$. Up to a multiplicative constant, this result is sharp, that is, there exists a constant c so that for every $d \geq 2$, and sufficiently large a, there exists a hypergraph $G \in \mathcal{G}(d,a)$ with $|\mathcal{E}(G)| = ca^{(2^d-1)/2}$ which is $K^{(d)}(2,\ldots,2)$-free.

Theorem 2.3.2 is proved by induction with the base case being Lemma 2.2.3. The sharpness aspect is proved by extending every edge in a $K_{2,2}$-free bipartite graph to a hyperedge in every possible way.

2.4 Lower bounds for ex$(n, K^{(d)}(2,2,\ldots,2))$

In [13], it was stated that there is a universal constant C so that for any integers $l > 1$ and $r > 1$ and n sufficiently large, $\mathrm{ex}(n; K^{(r)}(l,\ldots,l)) \geq n^{r-C/l^{r-1}}$. Unfortunately, the proof for this claim has not been found (cp. [22], p. 259). In what follows, we examine arguments for lower bounds on this extremal number, some standard and well known, some new.

For certain choices of n, the (upper) bound in Lemma 2.2.1 is nearly tight, as was found by Reiman [42]. In [38], (Problem 10.36), a standard proof of the following is given using polars in projective planes.

Lemma 2.4.1 *If $n = q^2+q+1$ for some prime power q, then*

$$ex(n, K_{2,2}) > (1/2)(n^{3/2} - n^{1/2}).$$

Another proof for a lower bound for $\mathrm{ex}(n, K_{2,2})$ is standard, and a bit simpler, though the constant is different. Briefly, the construction is as follows. Fix a finite projective plane of order q with points P and lines $\mathcal{P}$. Let $n = q^2+q+1$ and form the equibipartite graph on $2n$ vertices, where the points P are in one part and the lines $\mathcal{P}$ are the vertices of the other part; a point is connected to a line if and only if the point is on the line. In this case, we get a C_4-free (regular) graph with $(q^2+q+1)(q+1) \sim n^{3/2}$ edges on $2n$ vertices. This shows that the constant in Lemma 2.2.2 is best possible.

The lower bound in Lemma 2.4.1 only applies for certain values of n (although with some interpolation, a similar statement follows for all n) and it is not quite equal to the upper bound found in Lemma 2.2.1. Other techniques were sought for the lower bound (many exact values are known—see [22] for references—Füredi completing those for $n = q^2+q+1$, q a prime power). The number of C_4-free graphs on n vertices with $cn^{3/2}$ edges is well studied (see [36], for example).

A natural choice for a method by which to prove general lower bounds (for $\mathrm{ex}(n, K_{2,2})$ and other extremal numbers) is the probabilistic method; in this case, a fairly simple argument may give a more general result however considerably weaker than some known constructions. The following is a popular probabilistic argument, one we refer to as the "deletion method" (see [12] for perhaps the first instantiation of this method; see also [18]). In later applications of the deletion method (for example, Lemma 3.2.5) we only sketch the proof; the first time we give considerably more detail.

Before we begin, we review two inequalities and some notation. For a random variable X, we use $E(X)$ to denote the expectation of X, and $Var(X) = E(X^2) - E(X)^2$ to denote the variance of X. For $t > 0$, Markov's inequality (e.g., see [40]) states that for a non-negative random variable X,

$$\mathrm{Prob}(X \geq t) < \frac{E(X)}{t}, \tag{7}$$

and for a random variable Y; Chebychev's inequality is

$$\mathrm{Prob}(|Y - E(Y)| \geq t) \leq \frac{Var(Y)}{t^2}. \tag{8}$$

Lemma 2.4.2 *There is a constant c so that for n sufficiently large,*

$$ex(n, K_{2,2}) > cn^{4/3}.$$

Proof: Let $G \in \mathcal{G}_{n,p}$ be a random graph on n vertices where edges are chosen independently with probability p. Define the random variable $X = X(G)$ to be the number of copies (not necessarily induced) of C_4 in G. The expected number of copies of C_4 in G is

$$E(X) = 3\binom{n}{4}p^4 \sim \frac{n^4p^4}{8}.$$

For $0 < \epsilon < 1$, by Markov's inequality (7), $\mathrm{Prob}(X \geq \frac{E(X)}{1-\epsilon}) < 1 - \epsilon$, and so $\mathrm{Prob}(X < \frac{1}{1-\epsilon}3\binom{n}{4}p^4) > \epsilon$.

Define the random variable $Y = Y(G) = |\mathcal{E}(G)|$, the number of edges in G. It is not too difficult to see that Y has a binomial distribution with $E(Y) = \binom{n}{2}p$ and variance $Var(Y) = \binom{n}{2}p(1-p)$. By Chebychev's inequality (8), for $t > 0$,

$$\mathrm{Prob}(|Y - \binom{n}{2}p| \geq t) \leq \frac{\binom{n}{2}p(1-p)}{t^2},$$

and so

$$\mathrm{Prob}(Y \geq \binom{n}{2}p - t) > 1 - \frac{\binom{n}{2}p(1-p)}{t^2} > 1 - \frac{\binom{n}{2}p}{t^2}.$$

Using $t = \frac{1}{2}\binom{n}{2}p$, we have $\mathrm{Prob}(Y \geq \frac{1}{2}\binom{n}{2}p) > 1 - \frac{4}{\binom{n}{2}p}$.

If

$$\epsilon + 1 - \frac{4}{\binom{n}{2}p} > 1, \tag{9}$$

then there is a graph G with both at least $\frac{1}{2}\binom{n}{2}p$ edges and at most $\frac{1}{1-\epsilon}3\binom{n}{4}p^4$ copies of $C_4 \cong K_{2,2}$. Observe that for ϵ fixed and n sufficiently large, equation (9) holds whenever $n^2p \to \infty$ (as $n \to \infty$). Suppose that ϵ and p have been chosen so that equation (9) holds and G is a witness with $|\mathcal{E}(G)| \geq \frac{1}{2}\binom{n}{2}p$ and at most $\frac{1}{1-\epsilon}3\binom{n}{4}p^4$ C_4's. We wish to construct a C_4-free G' from G by deleting an edge from each C_4 in G. If there are half as many C_4's as edges, we have

$$\frac{1}{4}\binom{n}{2}p = \frac{1}{1-\epsilon}3\binom{n}{4}p^4,$$

which is satisfied for $p = (1 + o(1))(1 - \epsilon)^{1/3}n^{-2/3}$. With this choice of p, we have on the order of $n^2p/8 = \Theta(n^{4/3})$ edges remaining after deletion, as claimed in the statement of the theorem. It remains to observe that with this choice of p, $n^2p \to \infty$. $\Box$

2.5 Lower bounds for ex$(n, K^{(3)}(2, 2, 2))$

We review some methods for proving the lower bound for ex$(K^{(3)}(2, 2, 2))$, the first following naturally from Lemma 2.4.1, the second a deletion argument similar to Lemma 2.4.2, and the third an extension based on affine spaces. Even though each of the successive results imply the previous, we include them all here for the record. We believe that each technique may be of independent value.

The first bound is based on a straightforward observation of Füredi [23] (this idea was also used in showing the sharpness of the bound in Theorem 2.3.2).

Lemma 2.5.1 *There is a constant c so that for n sufficiently large,*

$$ex(n, K^{(3)}(2, 2, 2)) > cn^{5/2}.$$

Proof: Let X_1, X_2, X_3 be disjoint vertex sets, each with n vertices. We construct the 3-regular 3-partite hypergraph $G = (X_1, X_2, X_3, \mathcal{E})$ on $3n$ vertices and edges $\mathcal{E}$ as follows. By Lemma 2.4.1 on vertices $X_1 \cup X_2$ impose a graph G_{12} (an ordinary graph) which is C_4-free and has $cn^{3/2}$ edges, (where $c > 1$ is a constant independent of n). Now define the edge set of G by

$$\mathcal{E} = \{(x_1, x_2, x_3) : (x_1, x_2) \in \mathcal{E}(G_{12}), x_3 \in X_3\}.$$

So what we have done is extend every edge of G_{12} n times to hyperedges containing elements from X_3, yielding $|\mathcal{E}(G)| = n|\mathcal{E}(G_{12})| = cn^{5/2}$ edges. $\Box$

One can improve the bound found in Lemma 2.5.1 by imitating the proof of Lemma 2.4.2, the standard deletion technique. We show this in the first proof of the next lemma. In the second proof, a slight variation of this technique is employed yielding the same bound (up to at most a constant multiple). The second proof is given because it is the foundation for yet another technique using affine subspaces.

Lemma 2.5.2 *There is a constant c so that for n sufficiently large,*

$$ex(n, K^{(3)}(2,2,2)) > cn^{18/7}.$$

First proof of Lemma 2.5.2: We only outline the calculations; the method is the deletion method as used in Lemma 2.4.2.

Let $G = G(n,p)$ be a 3-regular hypergraph whose hyperedges are chosen independently at random with probablity p. The expected number of edges in G is on the order of n^3p.

The expected number of $K^{(3)}(2,2,2)$'s in G is on the order of n^6p^8. In order to delete one hyperedge from each copy of $K^{(3)}(2,2,2)$, it suffices to have

$$n^6p^8 \sim n^3p.$$

In this case, $p \sim n^{-3/7}$. Deleting these edges gives $cn^{18/7}$ edges remaining. $\Box$

Second proof of Lemma 2.5.2: Let X, Y, Z be disjoint vertex sets, each with n vertices. We construct the 3-regular 3-partite hypergraph $G = (X, Y, Z, \mathcal{E})$ on $3n$ vertices and hyperedges $\mathcal{E}$ as follows. For each pair (x, y), $x \in X$, $y \in Y$, let S_{xy} be a random subset of Z with elements chosen independently with probability p. Define $\mathcal{E}' = \{(x,y,z) : x \in X, y \in Y, z \in S_{xy}\}$. The expected number of edges in $\mathcal{E}'$ is n^3p.

We now count the expected number of $K^{(3)}(2,2,2)$'s. For each 4-tuple of vertices

$$\{x_1, x_2, y_1, y_2 : x_1, x_2 \in X, y_1, y_2 \in Y\},$$

the expected number of pairs common to $S_{x_1y_1} \cap S_{x_1y_2} \cap S_{x_2y_1} \cap S_{x_2y_2}$ is $\binom{n}{2}p^8$. Hence the expected number of $K^{(3)}(2,2,2)$'s is $\binom{n}{2}^2\binom{n}{2}p^8 = \Theta(n^6p^8)$.

From this point on, we only outline the calculations behind the deletion method (as in Lemma 2.4.2). Fix a hypergraph with these expected values. Now from each copy of $K^{(3)}(2,2,2)$ delete one hyperedge. In order for this to be possible,

$$n^6p^8 = \Theta(n^3p)$$

suffices. In this case, $p = \Theta(n^{-3/7})$, giving $cn^{18/7}$ edges remaining. $\Box$

To demonstrate a technique using affine subspaces, we use the following well known lemma (for example, see [47], p. 292).

Lemma 2.5.3 *If V is an s-dimensional vector space on l^s points, the number of k-dimensional affine subspaces contained in V is given by the Gaussian coeffiecient*

$$\left[\begin{array}{c} s \\ k \end{array}\right]_l = \frac{l^s(l^s-1)(l^s-l)\cdots(l^s-l^{k-1})}{l^k(l^k-1)(l^k-l)\cdots(l^k-l^{k-1})} = (1+o(1))l^{(s-k)(k+1)}$$

as $l \to \infty$.

Lemma 2.5.4 *There is a constant c so that for n sufficiently large,*

$$ex(n, K^{(3)}(2,2,2)) > cn^{13/5}.$$

Proof: Let X, Y, Z be disjoint vertex sets, each with n vertices. We construct the 3-regular 3-partite hypergraph $G = (X, Y, Z, \mathcal{E})$ on $3n$ vertices and edges $\mathcal{E}$ as follows. For the moment, fix integers $2 \le r < s$ and let $l = n^{1/s}$, and so $n = l^s$. Let $Z = l^s$, viewed as the elements of the vector space $l^s = \{0, 1, \dots, l-1\}^s$.

Let $\mathcal{R}$ be the collection of r-dimensional affine subspaces of l^s. Each element $R \in \mathcal{R}$ has l^r elements. First we define the hyperedge set $\mathcal{E}'$, from which we will delete some hyperedges to yield $\mathcal{E}$. For each $x \in X$, $y \in Y$ select at random $R_{xy} \in \mathcal{R}$

$$\mathcal{E}' = \{(x, y, z) : x \in X, y \in Y, z \in R_{xy}\}.$$

For the moment, fix two vertices in X and two in Y, and let R_1, R_2, R_3, and R_4 be the four r-spaces thereby determined.

The number of lines (one dimensional affine subspaces) in l^s is $\binom{l^s}{2}/\binom{l}{2} \sim l^{2s-2}$ (as $l \to \infty$). For a fixed line $L \subset Z = l^s$, the probability that L is contained in one R is the total number of lines in R divided by the total number of lines, that is l^{2r-2}/l^{2s-2}, and so

$$\text{Prob}\,[L \subseteq R_1 \cap R_2 \cap R_3 \cap R_4] = \left(\frac{l^{2r-2}}{l^{2s-2}}\right)^4 l^{2s-2}.$$

Since any two lines intersect in at most one point, any two points from Z determining a copy of $K^{(3)}(2,2,2)$ must come from the same line, and with l points per line, the expected number of edges contributing to copies of $K^{(3)}(2,2,2)$ is

$$\binom{n}{2}^2 4 \left(\frac{l^{2r-2}}{l^{2s-2}}\right)^4 l^{2s-2} l.$$

If we delete all of these edges, it suffices that

$$l^{4s} \left(\frac{l^{2r-2}}{l^{2s-2}}\right)^4 l^{2s-2} \cdot l \le \frac{l^{2s} l^r}{2}. \tag{10}$$

This implies that $7r - 4s \le 1$. With $r = 3$ and $s = 5$, this inequality is satisfied and in this case the number of original edges is at least $n^{13/5}$. The resulting hypergraph formed by deleting edges then has $cn^{13/5}$ edges and contains no copy of $K^{(3)}(2,2,2)$. $\square$

It is interesting to note that in the proof of Lemma 2.5.4, if we instead use planes in the role of lines, naively one would arrive at

$$l^{4s} \left(\frac{l^{3r-3}}{l^{3s-3}}\right)^4 l^{3s-3} \cdot l^2 \le \frac{l^{2s} l^r}{2},$$

yielding $11r - 7s \leq 1$, and with $r = 2$, $s = 3$, one would hope to obtain $n^{8/3}$ remaining edges. However, this neglects to take into account that four r-spaces may intersect in only a line, not just a plane, and in that case, two points from such a line will determine a $K^{(3)}(2,2,2)$ unaccounted for. This idea is not to be forsaken however.

One can extend proof techniques used in Lemma 2.5.4 to forbidding copies of $K^{(3)}(2,2,t)$, where t is large. We only hint at the proof.

Lemma 2.5.5 *For every $\delta > 0$ and for t sufficiently large, there is a constant $c = c(t,\delta)$ so that for all n sufficiently large,*

$$ex(n, K^{(3)}(2,2,t)) > cn^{8/3-\delta}.$$

Proof idea: Take some sphere of points in the common intersection of planes, guaranteeing that the troublesome lines hit at most two points in the intersection.

2.6 Extensions to $\mathbf{ex}(n, K^{(d)}(2,2,\ldots,2))$

We can use the standard deletion technique as in Lemmas 2.4.2 and 2.5.2 to give a lower bound for $\mathrm{ex}(n, K^{(d)}(2,2,\ldots,2))$.

Lemma 2.6.1 *For each $d \geq 3$, there exists $c = c(d)$ and $n_0 = n_0(d)$ so that for $n \geq n_0$ sufficiently large,*

$$ex(n, K^{(d)}(2,2,\ldots,2)) > cn^{d-\frac{d}{2^d-1}}.$$

Proof: We only outline the calculations; the proof follows the deletion method given in Lemma 2.4.2. If we examine a random d-uniform hypergraph $G = G(n,p)$ on n vertices and d-hyperedges chosen with probability p, the expected number of copies of $K^{(d)}(2,2,\ldots,2)$ is of the order of $n^{2d}p^{2^d}$ and the expected number of hyperedges is of the order $n^d p$. Equating these numbers yields $p = n^{\frac{-d}{2^d-1}}$, and hence the desired number of edges. $\square$

We again apply the affine-space technique (as in Lemma 2.5.4) to d-uniform hypergraphs. The following is presently (cf. [22]) the best known lower bound for $\mathrm{ex}(n, K^{(d)}(2,2,\ldots,2))$, not much of an improvement over that given in Lemma 2.6.1.

Lemma 2.6.2 *For each $d \geq 3$, if there exists a (smallest) positive integer s so that $\frac{sd-1}{2^d-1}$ is an integer, then there exists $c = c(d)$ and $n_0 = n_0(d)$ so that for $n \geq n_0$ sufficiently large,*

$$ex(n, K^{(d)}(2,2,\ldots,2)) > cn^{d-\frac{d-1/s}{2^d-1}},$$

Proof: For $n = l^s$, and $2 \leq r < s$, duplicate the construction in the proof of Lemma 2.5.4, but for arbitrary d. Corresponding to (10), we need (counting edges used in copies of $K^{(d)}(2, 2, \ldots, 2)$),

$$l^{2(d-1)s} \left(\frac{l^{2r-2}}{l^{2s-2}} \right)^{2^{d-1}} \cdot l^{2s-2} l \ \sim \ l^{(d-1)s} l^r. \tag{11}$$

For (11) to hold, it suffices to have

$$r = \frac{s(d + 1 - 2^d s) - 1}{1 - 2^d},$$

which yields

$$l^{s(d-1+\frac{d+1-2^d}{1-2^d})-1} = n^{d - \frac{d-1/s}{2^d-1}}$$

edges. $\Box$.

Notice that if the s in the statement of Lemma 2.6.2 does not exist, then we can let s be arbitrarily large, say kl for some constant k, and then since $l^{-1/s}$ can be bounded by a constant, we get the bound stated in Lemma 2.6.1. In the case $d = 3$, $s = 5$ was found to satisfy the condition in Lemma 2.6.2. For $d = 4$, the assignments $s = 4$ and $r = 3$ work for the lower bound, and together with the upper bound obtained from Erdős' Theorem 2.2.4, we have:

Corollary 2.6.3 *There is a constant c so that for n sufficiently large,*

$$n^{15/4} \leq ex(n, K^{(4)}(2, 2, 2, 2) \leq cn^{31/8}.$$

One obvious condition is that if d is a power of 2, then the desired s in Lemma 2.6.2 exists. Perhaps the result can be restated with $s = 4$ or $s = 3$, but even so, one obtains only a minimal improvement (for large d) over Lemma 2.6.1, still far from the upper bound.

3 Integers and cubes

3.1 Sidon sets

A *Sidon set* is a collection of integers whose pairwise sums $a + b$, $(a \neq b)$ are all distinct; these are also referred to as B_2-sets. In the proof of the upper bound in Theorem 4.1.2, we use the following result due to Singer [45] to produce a partition into Sidon sets. See also [19] for a simple construction of a single Sidon set, but with different bounds.

Theorem 3.1.1 (Singer) *Let m be a prime power. There exist $m+1$ integers*

$$0 \leq x_1 < x_2 \ldots < x_{m+1} < m^2 + m + 1$$

so that the $m^2 + m$ differences $x_i - x_j, 1 \leq i \neq j \leq m + 1$, are distinct modulo $m^2 + m + 1$.

For example, with $m = 4$, the integers 0, 1, 6, 8, and 18 have distinct differences modulo 21.

	0	1	6	8	18
0		20	15	13	3
1	1		16	14	4
6	6	5		19	9
8	8	7	2		11
18	18	17	12	10	

3.2 Affine cubes

As with the Boolean algebras, we will be concerned with both coloring and density results for Hilbert sets (as defined in the introduction), also called affine d-cubes. In what follows, r, d, and $x_1, \ldots, x_d$ are positive integers and x_0 is non-negative. Many results here are contained in [32] and [33], and so appear without proof.

Theorem 3.2.1 (Hilbert [35]) *For every r, d, there exists a least number $h(d,r)$ so that for every coloring $\chi : [h(d,r)] \to [r]$, there exists an $H(d)$ monochromatic under χ.*

Recall that, by definition, for any $H(d) = H(x_0, x_1, \ldots, x_d)$, $|H(d)| \le 2^d$ trivially holds, and that the $H(d)$-set is *full* if $|H(d)| = 2^d$, that is, if all the sums defining $H(d)$ are distinct.

The following has a simple proof, but is essential in showing known bounds for partitions of $[n]$ into sets not containing affine d-cubes.

Lemma 3.2.2 ([32]) *If a finite collection X of distinct non-negative integers does not contain any full affine d-cubes and does not contain any arithmetic progressions of length three, then X does not contain any affine d-cubes.*

Apart from the original source, one can find the proof of the following in [28], (lower bound of their Theorem 6.6).

Theorem 3.2.3 (Behrend [4]) *There exists a constant c so that for m sufficiently large, there exists $B \subset [m]$ not containing any arithmetic progressions of length three, and satisfying*

$$|B| \ge me^{-c\sqrt{\ln m}} = m^{1-o(1)}.$$

A modification of the proof of Behrend's Theorem yields the following partition result.

Theorem 3.2.4 ([32]) *For sufficiently large n, there exists $q < e^{3\sqrt{\ln n}}$ and a partition $[n] = X_1 \cup \ldots \cup X_q$ so that each X_i does not contain an arithmetic progression of length 3 and $|X_i| \le n/e^{\ln 2\sqrt{\ln n}}$.*

The proof of the following result employs the standard deletion technique using the probabilistic method (as in Lemma 2.4.2).

Lemma 3.2.5 ([32]) *For each d, and every set X of positive integers, there exists $A \subset X$ not containing a full affine d-cubes and*

$$|A| \geq \frac{1}{8}|X|^{1-\frac{d}{2^d-1}}.$$

The following lemma is a combination of Theorem 3.2.3 and a special case of Lemma 3.2.5, stated separately for later use in the proof of Theorem 4.2.1.

Lemma 3.2.6 *For every d there is a constant c so that for every k and every m, there is a set $S \subset [k+1, k+m]$ containing no full affine d-cubes nor containing any arithmetic progression of length 3, yet has at least*

$$|S| \geq cm^{1-\frac{d}{2^d-1}(1-o(1))}$$

elements.

Proof: By Theorem 3.2.3, let $B \subset [1, m]$ containing no arithmetic progression, and with

$$|B| \geq me^{-\sqrt{\ln m}} = m^{1-o(1)}.$$

The translation of B, $B_k = \{b + k : b \in B\}$ also has no arithmetic progression. With B_k playing the role of X, Lemma 3.2.5 yields S as desired. □

3.3 Lower coloring bound for affine d-cubes

Lemma 3.3.1 ([32]) *For each $d \geq 2$, there exists a constant c so that for any sufficiently large set X of positive integers, there exists a partition $X = A_0 \cup A_1 \cup \ldots \cup A_r$ into $r + 1 \leq c|X|^{\frac{d}{2^d-1}}$ colors so that no color class contains a full affine d-cube.*

The two partition results, Theorem 3.2.4 and Lemma 3.3.1, together with Lemma 3.2.2 give a coloring result for affine d-cubes.

Theorem 3.3.2 ([32]) *For each $d \geq 2$, and r sufficiently large,*

$$r^{\frac{2^d-1}{d}(1-o(1))} \leq h(d, r).$$

3.4 Density for affine d-cubes—upper bounds

Szemerédi [46] (Lemma $p(\delta, l)$, p. 93) gave a density version of Hilbert's theorem.

Theorem 3.4.1 (Szemerédi [46]) *For each d there exists a constant c so that for n sufficiently large, if $A \subseteq [1,n]$ satisfies $|A| \geq n^{1-\frac{1}{2^d}}$, then A contains an affine d-cube.*

In [32] and [33], other proofs of Theorem 3.4.1 have been given using hypergraphs. In [33], each edge of a graph $G \in \mathcal{G}(d,a)$ (recall Definition 2.3.1) were related to an integer, thereby deriving a density result for full affine d-cubes, and so for affine d-cubes in general.

Theorem 3.4.2 ([33]) *For each integer $d \geq 2$, for every real number $\epsilon > 0$, there exists an n_0 so that for every $n \geq n_0$, if $A \subset [1,n]$ satisfies*

$$|A| \geq (1+\epsilon)2^{d-3/2} n^{1-1/2^d}(1+o(1)),$$

then A contains $(1-o(1))n^2/2^{d-1}$ full affine d-cubes.

3.5 Upper coloring bound for affine d-cubes

An upper bound for the number $h(d,r)$ is the trivial one obtained by the associated density result, say Theorem 3.4.1 or 3.4.2.

Corollary 3.5.1 ([32]) *For every $d \geq 2$, there exists a constant c so that*

$$h(d,r) \leq cr^{2^d}.$$

4 Boolean Algebras

4.1 Bounds on $r(d,n)$; partition results

Theorem 4.1.1 *For any positive integer n, $r(1,n) = n$.*

Proof of Theorem 4.1.1: Coloring subsets of $[n]$ according to the $n+1$ different set sizes shows that $r(1,n) < n+1$. An easy proof by induction yields $r(1,n) \geq n$. Let $\chi : \mathcal{P}([1,n+1]) \to [1,n+1]$ be given. Without loss of generality, assume that $\chi([1,n+1]) = n+1$. If any other set also received the color $n+1$, then, since that set must be contained in $[1,n+1]$, we have a monochromatic $\mathcal{B}(1)$; if not, then every set in $\mathcal{P}([1,n])$ receives one of at most n colors, in which case the inductive hypothesis yields a monochromatic $\mathcal{B}(1)$. □

Theorem 4.1.2 *For n sufficiently large,*

$$\frac{1}{\sqrt{2}} n^{1/2} \leq r(2,n) \leq (1+o(1))n^{1/2}.$$

Before giving the proof of Theorem 4.1.2, we briefly mention what we need to show. To prove the lower bound, we need to demonstrate that for every coloring with fewer than $n^{1/2}/\sqrt{2}$ colors, one color class will contain a $\mathcal{B}(2)$. This will be done with a surprisingly simple trick, appearing (in some sense) first in [14], again in [16] (as referred to in [3]).

To see the upper bound, it suffices to give a $(1+o(1))n^{1/2}$-coloring which 'kills' every $\mathcal{B}(2)$. This will be done in a manner very similar to that used in [9] (or summarized in [27]).

Proof of lower bound in Theorem 4.1.2: Let $\epsilon > 0$ and $n \geq n(\epsilon)$ and fix a coloring

$$\mathcal{P}([n]) = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \ldots \cup \mathcal{F}_r,$$

where $r \leq \frac{n^{1/2}}{\sqrt{2}(1+\epsilon)}$.

Without loss of generality, suppose that n is even, and consider *only* the $n^2/4$ sets

$$S(i,j) = \{1, 2, \ldots, i\} \cup \{n/2+1, n/2+2, \ldots, n/2+j\},$$

where $1 \leq i, j \leq n/2$. By the pigeon-hole principle, there is one family $\mathcal{F}_k$ containing at least

$$\frac{n^2/4}{\frac{n^{1/2}}{\sqrt{2}(1+\epsilon)}} = \frac{1}{2\sqrt{2}}(1+\epsilon)n^{3/2}$$

sets $S(i,j)$. Consider the bipartite graph $G = (V_1, V_2, \mathcal{E})$, where

$$V_1 = \{[1], [2], \ldots, [n/2]\},$$

$$V_2 = \{\{n/2+1\}, \ldots, [n/2+1, n]\},$$

and $([i], [n/2, n/2+j]) \in \mathcal{E}$, if and only if $S(i,j) \in \mathcal{F}_k$. By Lemma 2.2.2, since

$$|\mathcal{E}| = |\mathcal{F}_k| \geq \frac{1}{2\sqrt{2}}(1+\epsilon)n^{3/2},$$

G contains a rectangle (copy of C_4), determined by say, $S(i,j)$, $S(i,j')$, $S(i',j)$ and $S(i',j')$. These sets clearly form a $\mathcal{B}(2)$. $\square$

The proof of the upper bound in Theorem 4.1.2 uses a construction duplicating those employed in [9] and [8]. One property of a 2-dimensional Boolean algebra

$$\mathcal{B}(2) = \{X_0, X_0 + X_1, X_0 + X_2, X_0 + X_1 + X_2\} = \{A, B, C, D\}$$

we will use is that since $B \backslash A = X_1 = D \backslash C$, (or $C \backslash A = X_2 = D \backslash B$), set sizes of differences of sets (in the Boolean algebra) are repeated.

Proof of upper bound in Theorem 4.1.2: Let m be a prime power and let $0 \leq x_1 < x_2 \ldots < x_{m+1} < m^2 + m + 1$ be as in Singer's theorem. For each $j = 1, \ldots, m+1$, define

$$Y_j = \{x_i - x_j \pmod{m^2 + m + 1} : 1 \leq i \leq m+1, i \neq j\} \subset [1, m^2 + m],$$

and put $Y_0 = \{0\}$. A simple calculation shows that if $\{a, b, c, d\} \subset Y_j$ for some j, and $a + b = c + d$, then $\{a, b\} = \{c, d\}$. Furthermore, it is not too difficult to see that the Y_j's partition the set $[0, m^2 + m]$ and for each $j \neq 0$, $|Y_j| = m$.

For each $j = 0, 1, 2, \ldots, m+1$, define

$$\mathcal{S}_j = \{X \subset [1, m^2 + m] : |X| \in Y_j\}.$$

This defines a decomposition of the power set of $[1, m^2 + m]$ into $m + 2$ classes. If for some j, there were sets $A, B, C, D \in \mathcal{S}_j$ with $|A| + |B| = |C| + |D|$, then $\{|A|, |B|\} = \{|C|, |D|\}$, and so these four sets do not form a $\mathcal{B}(2)$ (cf. [14]).

Now for a given n, let $m = m(n)$ be the smallest prime power so that $n \leq m^2 + m$. Since the ratio between consecutive prime powers tends to one, (as $n \to \infty$) the minimum number of color classes required to prevent a monochromatic $\mathcal{B}(2)$ is at most

$$m + 2 = (1 + o(1))\sqrt{m^2 + m} = (1 + o(1))\sqrt{n}. \ \Box$$

Theorem 4.1.3 *For $d > 2$, and n sufficiently large,*

$$cn^{1/2^d} \leq r(d, n) \leq n^{\frac{d}{2^d - 1}(1 + o(1))}.$$

Proof: If we color $\mathcal{P}([n])$ with fewer than $cn^{1/2^d}$ colors, then one color class contains $c_1 n^{-1/2^d}$ elements, and so by the density result, Theorem 4.3.1, one class contains a $\mathcal{B}(d)$.

For a number r, to prove that $r(d, n) < r$, we need to produce a partition $\mathcal{P}([n]) = \mathcal{F}_1 \cup \ldots \cup \mathcal{F}_r$ so that each $\mathcal{F}_i$ is $\mathcal{B}(d)$-free. It follows from the proof of Theorem 3.3.2 that there exists a partition $[n] = S_1 \cup S_2 \cup \ldots \cup S_r$, with $r = n^{\frac{d}{2^d - 1}(1 + o(1)}$, each class containing no full affine d-cube, nor any arithmetic progressions. For each $i = 1, \ldots, d$, put

$$\mathcal{F}_i = \{F \subset [n] : |F| \in S_i\}.$$

Suppose, in hopes of a contradiction, that $\mathcal{B} \subset \mathcal{F}_i$ is a d-dimensional Boolean algebra, that is, there exist pairwise disjoint sets $B_0, B_1, \ldots, B_d$, so that

$$\mathcal{B} = \left\{ B_0 \cup \bigcup_{i \in I} B_i : I \subset [1, d] \right\},$$

For each $i = 0, 1, \ldots, d$, put $|B_i| = x_i$. If all of the sets in $\mathcal{B}$ are different sizes, then the set

$$\{|B| : B \in \mathcal{B}\} = \left\{ x_0 + \sum_{i \in I} x_i : I \subset [1, d] \right\} \subset S_i$$

is a full affine d-cube, a contradiction.

So there must be two elements of $\mathcal{B}$ with the same size. Suppose that $C, D \in \mathcal{B}$ satisfy $|C \cap D| = a$ and $|C| = |D| = a + b$. Since $\mathcal{B}$ is a Boolean algebra, the sets $C \cap D$, C, and $C \cup D$ are contained in $\mathcal{B}$, but in this case, the respective sizes, (which are members of S_i) a, $a + b$, $a + 2b$ form an arithmetic progression, another contradiction.

We conclude that each $\mathcal{F}_i$ can not contain any d-dimensional Boolean algebra, ending the proof. $\Box$

4.2 Lower bound for $b(n, d)$, a density result

Theorem 4.2.1 *For each $d > 2$, there exists $c_1 = c_1(d)$ so that for n sufficiently large,*

$$c_1 2^n n^{-\frac{d}{2^{d+1}-2}(1-o(1))} \leq b(n, d).$$

To streamline the proof of the theorem, we provide the following simple estimate regarding the number of subsets of a set which are close to the average size.

Lemma 4.2.2 *For n sufficiently large and each i satisfying $-\sqrt{n}/2 \leq i \leq \sqrt{n}/2$,*

$$\binom{n}{n/2 + i} \geq \frac{1}{\sqrt{e}} \binom{n}{n/2}.$$

Proof: It suffices to show the result for $i = \sqrt{n}/2$, which we shall assume is an integer (as well as $n/2$).

$$\begin{aligned}
\frac{\binom{n}{n/2 - \sqrt{n}/2}}{\binom{n}{n/2}} &= \prod_{j=0}^{\sqrt{n}/2 - 1} \frac{n/2 - j}{n/2 + \sqrt{n}/2 - j} \\
&= \prod_{j=0}^{\sqrt{n}/2 - 1} \left(1 - \frac{\sqrt{n}}{n + \sqrt{n} - 2j}\right) \\
&\geq \left(1 - \frac{1}{\sqrt{n}}\right)^{\sqrt{n}/2} \\
&\geq e^{\frac{-1/\sqrt{n}}{1 - 1/\sqrt{n}} \frac{\sqrt{n}}{2}} \\
&> e^{-1/2},
\end{aligned}$$

where the penultimate inequality follows from $1 - x \geq e^{-x/(1-x)}$. $\square$

Proof of Theorem 4.2.1: Fix n and let X be a set of n elements. We will construct a large family $\mathcal{F}$ of subsets of X which contains no d-dimensional algebra.

Applying Lemma 3.2.6 with $m = \sqrt{n}$ and $k = \frac{n}{2} - \frac{\sqrt{n}}{2} - 1$, let

$$S \subset \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]$$

be a collection of

$$|S| = n^{\frac{1}{2} - \frac{d}{2^{d+1}-2}(1-o(1))}$$

integers that contains no full $H(d)$ and no arithmetic progression of length three. Define

$$\mathcal{F} = \{Y \subset X : |Y| \in S\}.$$

Calculating the size of $\mathcal{F}$,

$$\begin{aligned} |\mathcal{F}| &= \sum_{s \in S} \binom{n}{s} \\ &> |S| e^{-1/2} \binom{n}{n/2} \qquad \text{(by Lemma 4.2.2)} \\ &\sim |S| \frac{1}{\sqrt{\pi e n}} 2^n \\ &= c 2^n n^{-\frac{d}{2^{d+1}-2}(1-o(1))}. \end{aligned}$$

So $\mathcal{F}$ contains the desired number of elements; the fact that $\mathcal{F}$ does not contain a d-dimensional Boolean algebra follows as in the proof of the upper bound in Theorem 4.1.3. $\square$

4.3 Upper bound for $b(n, d)$

In [43], Rödl proved a weak version of the following density result; this proof is based on similar ideas.

Theorem 4.3.1 *For each $d \geq 1$ there exists a constant c_2 so that*

$$b(n, d) \leq c_2 n^{-1/2^d} \cdot 2^n.$$

First we give a preparatory discussion of chains in Boolean lattices, then give the proof of Theorem 4.3.1 which relies both on these notions and a result from Section 2 on hypergraphs.

Let Y be a set of t vertices. A collection $\mathcal{C} \subseteq \mathcal{P}(Y)$ of subsets of Y is a *chain* if and only if for every $A, B \in \mathcal{C}$, either $A \subset B$ or $B \subset A$. A chain $\mathcal{C} \subseteq \mathcal{P}(Y)$ is *symmetric* if for every $C \in \mathcal{C}$ there exists $C' \in \mathcal{C}$ so that $\{|C|, |C'|\} = \{\lceil t/2 \rceil + i, \lfloor t/2 \rfloor - i\}$ for some $i \geq 0$. A chain is *convex* if whenever $A \subset B \subset C$ and both A and C are in the chain, then so is B.

There are a number of methods by which a t-dimensional Boolean lattice can be partitioned into $\binom{t}{\lfloor t/2 \rfloor})\}$ disjoint symmetric convex chains. One proof is a fairly easy by induction, likely due to de Bruin, known since the early 1950's. In [1] (or [2], p. 439), Aigner uses "lexicographic matchings", duplicating that of the so-called "upper-lower neighbor construction" (see for example, [2], p. 436). In [29], (or see [30], p. 30) a construction by Greene and Kleitman called "parenthesization" gives an explicit construction of such partition. The problem of partitioning a linear lattice (the lattice of subfields of a finite dimensional field) into symmetric chains has only recently been solved by Vogt and Voigt [48].

Let $\boldsymbol{\mathcal{C}} = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_{\binom{t}{\lfloor t/2 \rfloor}}\}$ be a decomposition of $\mathcal{P}(Y)$ into disjoint symmetric convex chains, and let $\boldsymbol{\mathcal{C}}^{>2i} \subset \boldsymbol{\mathcal{C}}$ denote the subcollection of those chains having length greater than $2i$.

Since each chain $\mathcal{C} \in \boldsymbol{\mathcal{C}}^{>2i}$ contains a different set with $\lfloor t/2 \rfloor - i$ vertices, it follows that

$$|\boldsymbol{\mathcal{C}}^{>2i}| = \binom{t}{\lfloor t/2 \rfloor - i}.$$

For any permutation $\pi : Y \to Y$ of the vertices of Y and for any chain $C \in \boldsymbol{\mathcal{C}}$, the collection

$$\pi(\mathcal{C}) = \{\pi(C) : C \in \mathcal{C}\}$$

is also a chain, so

$$\pi(\boldsymbol{\mathcal{C}}) = \{\pi(\mathcal{C}) : \mathcal{C} \in \boldsymbol{\mathcal{C}}\}$$

is also a symmetric chain decomposition of $\mathcal{P}(Y)$, with $\pi(\boldsymbol{\mathcal{C}}^{>2i}) \subset \pi(\boldsymbol{\mathcal{C}})$.

Lemma 4.3.2 *Let Y be a set of t elements. Fix $D \subset Y$ and let*

$$\boldsymbol{\mathcal{C}} = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_{\binom{t}{\lfloor t/2 \rfloor}}\}$$

be a fixed decomposition of the power set $\mathcal{P}(Y)$ into disjoint symmetric convex chains. If $\pi : Y \to Y$ is a permutation chosen randomly from the set of $t!$ permutations of Y, then

$$\textit{prob}(D \in \pi(\mathcal{C}) \textit{ for some } \mathcal{C} \in \boldsymbol{\mathcal{C}}^{>2i}) > \left(1 - \frac{2i+2}{t}\right)^i.$$

Proof: If $|D| \leq \lfloor t/2 \rfloor - i$ or $|D| \geq \lfloor t/2 \rfloor + i$, then since $\pi(\cup \boldsymbol{\mathcal{C}}^{>2i})$ contains all sets of these sizes, then $D \in \pi(\mathcal{C})$ for some $C \in \boldsymbol{\mathcal{C}}^{>2i}$, that is, the probability is 1.

Now fix $D \subset Y$ with $\lfloor t/2 \rfloor - i < |D| < \lfloor t/2 \rfloor + i$. Set $\mathcal{S} = \cup \boldsymbol{\mathcal{C}}^{>2i} \cap [Y]^{|D|}$. Let $\pi : Y \to Y$ be a random permutation. For a fixed $S \subset Y$ chosen with $|S| = |D|$,

$$\mathrm{Prob}(\pi^{-1}(D) = S) = \frac{|D|!(t - |D|!)}{t!} = \frac{1}{\binom{t}{|D|}}.$$

Hence,

$$\begin{aligned}
\mathrm{Prob}(D \in \pi(\mathcal{S})) &= \mathrm{Prob}(\pi^{-1}(D) \in \mathcal{S}) \\
&= \sum_{S \in \mathcal{S}} \mathrm{Prob}(\pi^{-1}(D) = S) \\
&= \frac{|\mathcal{S}|}{\binom{t}{|D|}} \\
&= \frac{|\boldsymbol{\mathcal{C}}^{>2i}|}{\binom{t}{|D|}} \\
&\geq \frac{\binom{t}{\lfloor t/2 \rfloor - i}}{\binom{t}{\lfloor t/2 \rfloor}} \\
&= \prod_{j=0}^{i-1} \frac{\lfloor t/2 \rfloor - j}{\lceil t/2 \rceil + i - j} \\
&\geq \prod_{j=0}^{i-1} \frac{t/2 - 1/2 - j}{t/2 + 1/2 + i - j} \\
&\geq \prod_{j=0}^{i-1} \left(1 - \frac{i+1}{t/2 + 1/2 + i - j}\right) \\
&> \left(1 - \frac{2i+2}{t}\right)^i . \ \square
\end{aligned}$$

The following fact follows from a simple averaging argument; we omit the proof.

Lemma 4.3.3 *Let $H = (V_1, \ldots, V_d, \mathcal{E}(H))$ be a given d-partite d-uniform hypergraph and let $v \leq \min_{1 \leq i \leq d}\{|V_i|\}$. For each $i = 1, \ldots, d$, there exist vertex sets $W_i \subseteq V_i$, $|W_i| = v$, so that the the subgraph H' induced by $\cup_{i=1}^d W_i$ has edge density at least that of H, that is,*

$$\frac{|\mathcal{E}(H')|}{v^d} \geq \frac{|\mathcal{E}(H)|}{|V_1| \cdot \ldots \cdot |V_d|}.$$

We are now prepared to prove an upper bound for $b(n,d)$.

Proof of Theorem 4.3.1: Let X be a set of n elements and fix a positive integer d. Let $\mathcal{F} \subset \mathcal{P}(X)$ satisfy

$$|\mathcal{F}| \geq c_1 n^{-1/2^d} 2^n, \tag{12}$$

where, with c from Theorem 2.2.4,

$$c_1 = c \cdot 10^d 2^{-1/2^{d-1}} d^{d-1/2^d}. \tag{13}$$

We will show that $\mathcal{F}$ contains a Boolean algebra of dimension d.

Partition $X = X_1 \cup X_2 \cup \ldots \cup X_d$ into d sets, each with size $\lfloor n/d \rfloor \leq |X_j| \leq \lceil n/d \rceil$. For each $j = 1, \ldots d$, fix $\boldsymbol{\mathcal{C}}_j$, a symmetric chain decomposition of $\mathcal{P}(X_j)$; for i to be determined later, let $\boldsymbol{\mathcal{C}}_j^{>2i} \subseteq \boldsymbol{\mathcal{C}}_j$ be the subcollection of those chains longer than $2i$. For each $j = 1, \ldots, d$ let $\pi_j : X_j \to X_j$ denote a permutation of X_j chosen randomly from the collection of all $|X_j|!$ permutations on X_j (the permutations $\pi_1, \pi_2, \ldots, \pi_d$ are chosen independently). Let $\mathcal{F}_{\pi_1,\ldots,\pi_d} \subset \mathcal{F}$ be a random subset of $\mathcal{F}$ defined by

$$\mathcal{F}_{\pi_1,\ldots,\pi_d} = \{F \in \mathcal{F} : \forall j = 1, \ldots, d, \exists \mathcal{D}(j) \in \pi_j(\boldsymbol{\mathcal{C}}_j^{>2i}) \text{ with } F \cap X_j \in \mathcal{D}(j)\}. \tag{14}$$

By Lemma 4.3.2, for any $F \in \mathcal{F}$,

$$\text{prob}\,(F \in \mathcal{F}_{\pi_1,\ldots,\pi_d}) > \prod_{j=1}^{d} \left(1 - \frac{2i+2}{|X_j|}\right)^i. \tag{15}$$

Fix $i = \lfloor \sqrt{n/d} \rfloor$, sufficient for our purpose in what follows. Then for sufficiently large n, as $|X_j| \geq \lfloor n/d \rfloor$, the right hand side of (15) can be further bounded from below by

$$\left(1 - \frac{2\sqrt{\lfloor n/d \rfloor} + 2}{\lfloor n/d \rfloor}\right)^{\lfloor \sqrt{n/d} \rfloor d} > \left(1 - \frac{2.1}{\sqrt{n/d}}\right)^{(\sqrt{n/d})d} > \left(\frac{1}{e^{2.1}}\right)^d > (.1)^d.$$

Hence the expected number of sets in $\mathcal{F}_{\pi_1,\ldots,\pi_d}$ is

$$\mathrm{E}(|\mathcal{F}_{\pi_1,\ldots,\pi_d}|) > (.1)^d |\mathcal{F}|. \tag{16}$$

Fix a choice of $\hat{\pi}_1, \ldots, \hat{\pi}_d$ for which (16) is realized. For each $j = 1, \ldots, d$, set $\boldsymbol{\mathcal{D}}_j = \hat{\pi}_j(\boldsymbol{\mathcal{C}}_j^{>2i})$, the family of disjoint chains in X_j longer than $2i$, and write

$$\boldsymbol{\mathcal{D}}_j = \left\{\mathcal{D}_{j,k_j} : 1 \leq k_j \leq \binom{|X_j|}{\lfloor |X_j|/2 \rfloor - i}\right\}.$$

Put $\mathcal{G} = \mathcal{F}_{\hat{\pi}_1,\ldots,\hat{\pi}_d}$. Note that by (14) and (16),

$$\mathcal{G} = \{F \in \mathcal{F} : \forall j = 1, \ldots, d, \exists \mathcal{D}(j) \in \boldsymbol{\mathcal{D}}_j \text{ with } F \cap X_j \in \mathcal{D}(j)\},$$

and

$$|\mathcal{G}| > (.1)^d |\mathcal{F}|. \tag{17}$$

For each choice of $k_1, \ldots, k_d$ (the k_i's not necessarily distinct), define the set system

$$\mathcal{D}_{1,k_1} \otimes \cdots \otimes \mathcal{D}_{d,k_d} = \{\cup_{j=1}^d D_{j,k_j} : D_{j,k_j} \in \mathcal{D}_{j,k_j}\},$$

and also define

$$\mathcal{D} = \bigcup_{k_1,,\ldots,k_d} (\mathcal{D}_{1,k_1} \otimes \cdots \otimes \mathcal{D}_{d,k_d}),$$

where now we have $\mathcal{G} = \mathcal{F} \cap \mathcal{D}$. Also for each $j = 1, \ldots, d$, set $s_j = |\cup \boldsymbol{\mathcal{D}}_j|$ the number of sets in chains in $\boldsymbol{\mathcal{D}}_j$. Furthermore, put

$$\mathcal{G}_{k_1,\ldots,k_d} = \mathcal{F} \cap (\mathcal{D}_{1,k_1} \otimes \cdots \otimes \mathcal{D}_{d,k_d}).$$

We observe that by (14) and (17),

$$|\mathcal{G}| = \left| \bigcup_{k_1,\ldots,k_d} \mathcal{G}_{k_1,\ldots,k_d} \right| = |\mathcal{F} \cap \mathcal{D}| > (.1)^d |\mathcal{F}|.$$

Since

$$\sum_{k_1,\ldots,k_d} |\mathcal{D}_{1,k_1}| \cdots |\mathcal{D}_{d,k_d}| = s_1 \cdots s_d < 2^n,$$

we infer that there is a choice of $\hat{k}_1, \ldots, \hat{k}_d$ so that

$$\frac{|\mathcal{G}_{\hat{k}_1,\ldots,\hat{k}_d}|}{|\mathcal{D}_{1,\hat{k}_1}| \cdots |\mathcal{D}_{d,\hat{k}_d}|} \geq \frac{|\mathcal{G}|}{s_1 \cdots s_d} > \frac{(.1)^d |\mathcal{F}|}{2^n}. \tag{18}$$

Using (12), we obtain from (18),

$$|\mathcal{G}_{\hat{k}_1,\ldots,\hat{k}_d}| > c_1 (.1)^d \cdot n^{-1/2^d} |\mathcal{D}_{1,\hat{k}_1}| \cdots |\mathcal{D}_{d,\hat{k}_d}|. \tag{19}$$

By Lemma 4.3.3, for each $j = 1, \ldots, d$, choose $\mathcal{D}^*_{j,\hat{k}_j} \subset \mathcal{D}_{j,\hat{k}_j}$ with $|\mathcal{D}^*_{j,\hat{k}_j}| = 2\lfloor \sqrt{n/d} \rfloor$ so that for

$$\mathcal{G}^*_{\hat{k}_1,\ldots,\hat{k}_d} = \mathcal{F} \cap (\mathcal{D}^*_{1,\hat{k}_1} \otimes \cdots \otimes \mathcal{D}^*_{d,\hat{k}_d}),$$

the corresponding inequality to (19) holds, namely,

$$\begin{aligned} |\mathcal{G}^*_{\hat{k}_1,\ldots,\hat{k}_d}| &> c_1 (.1)^d \cdot n^{-1/2^d} |\mathcal{D}^*_{1,\hat{k}_1}| \cdots |\mathcal{D}^*_{d,\hat{k}_d}| \\ &= c_1 (.1)^d \cdot n^{-1/2^d} \cdot (2\lfloor \sqrt{n/d} \rfloor)^d. \end{aligned}$$

For $m = d \cdot 2\lfloor\sqrt{n/d}\rfloor$, then $n \geq (m/2)^2 \cdot 1/d$, and hence

$$\begin{aligned} |\mathcal{G}^*_{\hat{k}_1,\ldots,\hat{k}_d}| &> c_1(.1)^d \left(\left(\frac{m}{2}\right)^2 \frac{1}{d}\right)^{-1/2^d} \left(\frac{m}{d}\right)^d \\ &= c_1(.1)^d 2^{1/2^{d-1}} d^{-d+2^{-d}} m^{d-1/2^{d-1}}. \end{aligned} \tag{20}$$

By the choice of c_1, (13) and (20) yield

$$|\mathcal{G}^*_{\hat{k}_1,\ldots,\hat{k}_d}| > c_2 \cdot m^{d-1/2^{d-1}}. \tag{21}$$

For each $j = 1, \ldots, d$ consider $Y_j = \mathcal{D}^*_{j,\hat{k}_j}$ as a vertex set, vertices being subsets of X_j in the chain $\mathcal{D}^*_{j,\hat{k}_j}$. Using the d-partite d-uniform hypergraph

$$\mathcal{H} = (Y_1, \ldots, Y_d, \mathcal{G}^*_{\hat{k}_1,\ldots,\hat{k}_d}),$$

then Theorem 2.2.4 (with $d \cdot 2\lfloor\sqrt{n/d}\rfloor$ as the number of vertices) and (21) imply that there is a copy of $K^{(d)}(2, 2, \ldots, 2)$ in $\mathcal{H}$. That is, for each $j = 1, \ldots, d$, there are $A_j^0, A_j^1 \in \mathcal{D}^*_{j,\hat{k}_j}$ with $A_j^0 \neq A_j^1$, say $A_j^0 \subset A_j^1$, so that for any choice of $(\delta_1, \ldots, \delta_d) \in \{0,1\}^d$,

$$A_1^{\delta_1} \cup \ldots \cup A_d^{\delta_d} \in \mathcal{G}^*_{\hat{k}_1,\ldots,\hat{k}_d} \subset \mathcal{F}.$$

In this case,

$$\{A_1^{\delta_1} \cup \ldots \cup A_d^{\delta_d} : (\delta_1, \ldots, \delta_d) \in \{0,1\}^d\}$$

is the desired d-dimensional Boolean algebra (cf. [28], Lemma 5.7) (with meet $(A_1^0 \cup \ldots \cup A_d^0)$ and join $(A_1^1 \cup \ldots \cup A_d^1)$) completing the proof. □

4.4 Uniform Boolean algebras

Theorem 4.4.1 *For any fixed d, $\epsilon > 0$, there exists n_0 and a constant c so that for every $n \geq n_0$,*

$$\frac{c}{n^{o(1)}} 2^n \leq b_u(n, d) \leq \epsilon \cdot 2^n.$$

By Theorem 4.3.1, if a subset of $\mathcal{P}([n])$ is chosen with $cn^{-1/2^d} 2^n$ elements, then this subset contains a d-dimensional Boolean algebra. In this section, we show that (for n sufficiently large), if a subset of $\mathcal{P}([n])$ is chosen with $\epsilon \cdot 2^n$ elements, then it contains a d-dimensional uniform Boolean algebra. The main tool used here is a density version of the Hales-Jewett theorem, which we now briefly describe.

Let $A = \{a_1, a_2, \ldots, a_t\}$ be an alphabet of t distinct letters. Let $A^m = \{f : [m] \to A\}$ denote the set of words $f = (f(1), f(2), \ldots, f(m))$ of length m formed by letters from A. A collection $\mathcal{L} = \{g_1, \ldots, g_t\} \subset A^m$ is a *combinatorial line* if there exists a partition of the coordinates $[m] = F \cup M$ [*F-ixed* and *M-oving*] so that for every $g_p, g_q \in \mathcal{L}$,

$$\begin{aligned} g_p(i) &= g_q(i) \quad \text{for each } i \in F, \text{ and} \\ g_p(j) &= a_p \quad \text{for each } j \in M. \end{aligned}$$

A density version of the Hales-Jewett theorem [34] was proved by Furstenburg and Katznelson [26] (or see [25] for survey paper):

Theorem 4.4.2 *For any $\epsilon > 0$ and any alphabet A, $|A| = t$, there exists m_0 so that for $m \geq m_0$, if $\mathcal{S} \subset A^m$ satisfies $|\mathcal{S}| \geq \epsilon t^m$, then $\mathcal{S}$ contains a combinatorial line.*

Proof of upper bound in Theorem 4.4.1: Put

$$A = \mathcal{P}([d]) = \{\emptyset, \{1\}, \ldots, \{d\}, \{1, 2\}, \ldots, \{1, 2, \ldots, d\}\},$$

$t = 2^d = |A|$, and without loss, assume that $m = n/d$ is an integer. Any word from A^m has the form $f = (S_1, \ldots, S_m)$, where for each $i = 1, \ldots, m$, $S_i \subset [d]$. We will use special notation to describe subsets of $[n] = [md]$. For each $i = 1, \ldots, m$, let $[d]_i = \{1_i, 2_i, \ldots, d_i\}$ be a copy of $[d]$; write

$$[n] = [md] = \cup_{i=1}^{m} [d]_i,$$

the union of m disjoint copies of $[d]$. Consider the bijection

$$\psi : A^m \to \mathcal{P}(md)$$

defined by

$$\psi((S_1, \ldots, S_m)) = \cup_{i=1}^{m} \{s_i : s \in S_i\}.$$

For example, with $d = 2$, $m = 6$ and $f = (\emptyset, \{2\}, \{1, 2\}, \emptyset, \{2\}, \{1\})$, we have

$$\psi(f) = \{2_2, 1_3, 2_3, 2_5, 1_6\}.$$

Now fix $\epsilon > 0$ and d and let m be so large that Theorem 4.4.2 applies, and let $\mathcal{L} = \{f_1, f_2, \ldots, f_t\}$ be a combinatorial line in A^m with fixed coordinates $F \subset [m]$ and moving coordinates $M \subset [m]$. We claim that the family

$$\psi(\mathcal{L}) = \{\psi(f_j) : j = 1, \ldots, t\}$$

is a d-dimensional uniform Boolean algebra.

Let B_0 be the union of those subsets of $[d]_i$'s determined by the fixed coordinates; to be precise, for each $i = 1, \ldots, m$, put $f_1(i) = S_i$ and

$$B_0 = \bigcup_{i \in F} \phi(f_1(i)) = \bigcup_{i \in F} \{s_i : s \in S_i\}.$$

Thus, B_0 can be interpreted as $\psi(f_1)$ provided f_1 is chosen so that $f_1(j) = \emptyset$ (or $\emptyset_j$) for each $j \in M$. For each $j = 1, 2, \ldots, d$, put

$$B_j = \{j_i : i \in M\}.$$

Clearly $|B_1| = |B_2| = \ldots = |B_d| = |M|$, and all the B_j's are disjoint. Now, since for any set $J \subset [d]$, there is a word $f \in \mathcal{L}$ so that for every $i \in M$, $f(i) = J$, we see that for each $J \subset [1, d]$,

$$B_0 \cup \bigcup_{j \in J} B_j \in \psi(\mathcal{L}). \ \Box$$

4.5 Lower bound for $b_u(n, d)$

Proof of lower bound in Theorem 4.4.1: Essentially, one duplicates the proof of Theorem 4.2.1, except without mention of the Hilbert set.

Let $S \subset [\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}]$ which contains no arithmetic progression of length 3 and is as large as possible. By Behrend's theorem (Theorem 3.2.3, using $m = \sqrt{n}$, and then translating the set B by $\frac{n}{2} - \frac{\sqrt{n}}{2} - 1$), we can have $|S| = (n^{1/2})^{1-o(1)} = n^{1/2-o(1)}$. Defining $\mathcal{S} = \{X \subset [n] : |X| \in S\}$,

$$|\mathcal{S}| \geq \binom{n}{n/2 - \sqrt{n}} |S| \sim c \frac{2^n}{\sqrt{n}} |S| = c \frac{2^n}{n^{o(1)}}.$$

It is now not difficult to see that $\mathcal{S}$ contains no d-dimensional uniform Boolean algebra. $\Box$

5 Conclusion

One more observation with respect to Lemma 2.6.1 may be in order. The bound in Lemma 2.6.1 is similar to that of one in Theorem 4.2.1, so it may be reasonable to conclude a (so far inscrutable) relationship between lower bounds for $\mathrm{ex}(n, K^{(d)}(2, 2, \ldots, 2))$ and $b(n, d)$—after all, upper bounds are analogously similar, and one is used to prove the other (see the proof of Theorem 4.3.1). Efforts to find the putative correspondence have failed as of yet.

If one can improve a known bound on $\mathrm{ex}(n, K^{(d)}(2, 2, \ldots, 2))$ for some $d > 2$, then one immediately improves other results in this paper (e.g., Theorem 4.3.1). In the proof of Theorem 3.4.2, (the variation of Szemeredi's density theorem for cubes) Theorem 2.3.2 was employed. A much simpler proof using instead Theorem 2.2.4, also works, however yielding a weaker result. It may be that Theorem 4.3.1 can be improved by applying Theorem 3.4.2 rather than Theorem 2.2.4 (preliminary calculations look hopeful, yet imposing).

There are many fields of study closely related to the questions studied in this paper. One could ask extremal questions for families of sets forbidding only certain types of substructures in a Boolean algebra, as in union-free families; instead of investigating these here, we refer the reader to [3] and [20] for an introduction and further references. We may also consider the work here as one kind of extension of Sperner's Lemma; many other interesting extensions of Sperner's Lemma have been made in similar directions, for example, [21], [24] and [10]. Another perspective on this work may be taken from the point of hypercubes and extremal questions thereof; see [6] for recent references and results.

For comparison and contrast purposes, we list some of the bounds mentioned in this paper. For $d = 2$:

$$(1 - o(1))\frac{1}{2n^{1/2}} \leq \frac{\mathrm{ex}(n, K_{2,2})}{n^2} \leq (1 + o(1))\frac{1}{2n^{1/2}};$$

$$(1 - o(1))r^2 \leq h(2, r) \leq (1 + o(1))r^2;$$

$$\frac{1}{\sqrt{2}}n^{1/2} \leq r(2, n) \leq (1 + o(1))n^{1/2};$$

$$c_1 n^{-1/4} \leq \frac{b(n, 2)}{2^n} \leq c_2 n^{-1/4}.$$

For $d \geq 3$:

$$\frac{c}{n^{\frac{d}{2^d-1}}} \leq \frac{\mathrm{ex}(n, K^{(d)}(2, 2, \ldots, 2))}{n^d} \leq \frac{1}{n^{\frac{1}{2^{d-1}}}};$$

$$r^{(1-o(1))\frac{2^d-1}{d}} \leq h(d, r) \leq cr^{2^{d-1}};$$

$$cn^{\frac{1}{2^d}} \leq r(d, n) \leq n^{\frac{d}{2^d-1}(1+o(1))};$$

$$\frac{c_1}{n^{\frac{d}{2^{d+1}-2}(1-o(1))}} \leq \frac{b(n, d)}{2^n} \leq \frac{c_2}{n^{1/2^d}}.$$

References

[1] M. Aigner, Lexicographic matching in Boolean algebras, *J. Combin. Th. Ser. B* **14** (1973), 187–194.

[2] M. Aigner, *Combinatorial Theory*, Grundlehren der mathematischen Wissenschaften **234**, Springer-Verlag, New York, 1979.

[3] M. Aigner, D. Duffus, and D. Kleitman, Partitioning the power set into union-free classes, *Discrete Math.* **88** (1991), 113–119.

[4] F. A. Behrend, On sets of integers which contain no three elements in arithmetic progression, *Proc. Nat. Acad. Sci.* **23** (1946), 331-332.

[5] B. Bollobás, *Extremal Graph Theory*, Academic Press, New York, 1979.

[6] P. Brass, H. Harborth, and H. Nienborg, On the maximum number of edges in a C_4-free subgraph of Q_n, *J. Graph Theory*, **19** (1995), 17–23.

[7] W. G. Brown, On graphs that do not contain a Thomsen graph, *Canad. Math. Bull.* **9** (1966), 281–285.

[8] T. C. Brown, F. R. K. Chung, P. Erdős, and R. L. Graham, Quantitative forms of a theorem of Hilbert, *J. Combin. Th. Ser. A* **38** (1985), 210–216.

[9] F. R. K. Chung and R. L. Graham, On multicolor Ramsey numbers for complete bipartite graphs, J. Combin. Th. Ser. A **18** (1975), 164–169.

[10] E. Damiani, O. D'Antona, and D. Loeb, Decompositions of $\mathcal{B}_n$ and Π_n using symmetric chains, *J. Combin. Th. Ser. A* **65** (1994), 151–157.

[11] P. Erdős, On sequences of integers no one of which divides the product of two others and on some related problems. *Izv. Nauk Mat. i Mech. Tomsk* **2** (1938), 74–82.

[12] P. Erdős, Graph theory and probability, *Canadian J. Math.* **11**, 34–38.

[13] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183–190. (Also reprinted in: P. Erdős, *The art of counting; selected writings*, J. Spencer ed., MIT Press, Cambridge, Massachusetts, 1973.)

[14] P. Erdős and D. Kleitman, On collections of subsets containing no 4-member Boolean algebra, *Proc. Amer. Math. Soc.* **28** (1971), 87–90.

[15] P. Erdős, A. Rényi, and V. Sós, On a problem of graph theory, *Studia Sci. Math. Hungar.* **1**, 215–235.

[16] P. Erdős and S. Shelah, On problems of Moser and Hanson, *Graph Theory and Applications*, Lecture Notes in Math. **303** (1972), 75–79.

[17] P. Erdős and M. Simonovits, Some extremal problems in graph theory, *Proc. Colloq. Math. Soc. János Bolyai* **4**, Combinatorial Theory and its Appl. I, (1970), 378–392.

[18] P. Erdős and J. Spencer, *Probabilistic methods in combinatorics*, Academic Press, London, New York, and Akadémiai Kiadó, Budapest, 1974.

[19] P. Erdős and P. Turán, On a problem of Sidon in additive number theory, and on some related problems, *Journal of the London Mathematical Society* **16** (1941), 212–215. (Also see addendum, ibid, **19** (1944), 208.)

[20] P. Frankl and Z. Füredi, Union-free families of sets and equations over fields, *J. of Number Theory*, **23** (1986), 210–218.

[21] Z. Füredi, A Ramsey-Sperner theorem, *Graphs and Combinatorics,* **1** (1985), 51–56.

[22] Z. Füredi, Turán type problems, in *Surveys in Combinatorics, 1991*, (ed. A. D. Keedwell), London Math. Soc. Lecture Notes **166**, Cambridge University Press, Cambridge (1991), 253–300.

[23] Z. Füredi, personal communication with V. Rödl, 1994.

[24] Z. Füredi, J. R. Griggs, A. M. Odlyzko, and J. B. Shearer, *Ramsey-Sperner Theory*, preprint (1986?).

[25] H. Furstenberg, Recurrent ergodic structures and Ramsey theory, in *Proceedings of the International Congress of Mathematicians*, Vol II, (Kyoto, 1990), 1057–1069, Math. Soc. Japan, Tokyo, 1991.

[26] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem, *Journal d'Analyse Mathematique* **57** (1991), 64–119.

[27] R. L. Graham, *Rudiments of Ramsey theory,* Regional conference series in mathematics, No. 45, American Mathematical Society, 1981 (reprinted with corrections, 1983).

[28] R. L. Graham and V. Rödl, Numbers in Ramsey theory, in *Surveys in Combinatorics 1987* (C. Whitehead, ed.), 111–153, London Math. Soc. Lecture Note Series **123**, Cambridge University Press, 1987.

[29] C. Greene and D. J. Kleitman, Strong versions of Sperner's theorem, *J. Combin. Th. Ser. A* **20** (1976), 80–88.

[30] C. Greene and D. J. Kleitman, Proof techniques in the theory of finite sets, in *Studies in combinatorics*, C. Rota, ed., MAA Studies in Mathematics **17**, 1978.

[31] D. S. Gunderson, *Extremal properties for Boolean algebras, sum-sets, and hypergraphs*, Dissertation, Emory University, Atlanta, 1995.

[32] D. S. Gunderson and V. Rödl, Extremal properties for affine cubes and hypergraphs, submitted.

[33] D. S. Gunderson and V. Rödl, An alternate proof of Szemerédi's cube lemma using extremal hypergraphs, submitted to SFB Preprint series, Universität Bielefeld.

[34] A. W. Hales and R. I. Jewett, Regularity and positional games, *Trans. Amer. Math. Soc.* **106** (1963), 22-229.

[35] D. Hilbert, Über die Irreduzibilatät ganzer rationaler Funktionen mit ganzzahligen Koeffizienten, (On the irreducibility of entire rational functions with integer coefficients) *J. Reine Angew. Math.* **110** (1892), 104–129.

[36] D. J. Kleitman and K. J. Winston, On the number of graphs without 4-cycles, *Discrete Math.* **41** (1982), 167–172.

[37] T. Kővári, V. T. Sós and P. Turán, On a problem of K. Zarankiewicz, *Coll. Math.* **3** (1954), 50–57.

[38] L. Lovász, *Combinatorial problems and exercises*, North-Holland Publishing Company, New York, 1979.

[39] D. Lubell, A short proof of Sperner's lemma, *J. Combin. Th.* **7** (1966), 299. (Also reprinted in: *Classic papers in combinatorics*, I. Gessel, G-C. Rota, eds., Birkhauser, Boston, 1987.)

[40] E. M. Palmer, *Graphical evolution*, John Wiley & Sons, New York, 1985.

[41] H. J. Prömel, *Ramsey theory for discrete structures*, Habilitationsschrift, Bonn (1987).

[42] I. Reiman, Uber ein Problem von K. Zarankiewicz, *Acta Math. Acad. Sci. Hung.* **9** (1959), 269–279.

[43] V. Rödl, Note on finite Boolean algebras, *Acta Polytechnica*, (1982), 47–49.

[44] M. Simonovits, Extremal graph theory, in *Selected Topics in Graph Theory 2*, (eds. L. W. Beineke, R. J. Wilson), Academic Press, New York, 1983, 161–200.

[45] J. Singer, A theorem in finite projective geometry, *Trans. Amer. Math. Soc.* **43** (1938), 377–385.

[46] E. Szemerédi, On sets of integers containing no four elements in arithmetic progression, *Acta Math. Acad. Sci. Hungar.* **20** (1969), 199-245.

[47] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, Cambridge University Press, 1992.

[48] F. Vogt and B. Voigt, Symmetric chain decompositions of linear lattices, Preprint Nr. 1633, *Fachbereich Mathematik*, Technische Hochschule Darmstadt, April 1994.

[49] B. Voigt, Personal communication with V. Rödl, Prague, 1981.