TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 300, Number 1, March 1987

FORBIDDEN INTERSECTIONS

PETER FRANKL AND VOJTECH RODL

ABSTRACT. About ten years ago P. Erdös conjectured that if $\mathcal{F}$ is a family of subsets of $\{1, 2, \ldots, n\}$ without $F, F' \in \mathcal{F}$, $|F \cap F'| = [n/4]$, then $|\mathcal{F}| < (2-\varepsilon)^n$ holds for some positive absolute constant ε. Here this conjecture is proved in a stronger form (Theorem 1.1), which solves a \$250 problem of Erdös. Suppose $\mathcal{C}$ is a code (i.e., a collection of sequences of length n) over an alphabet of q elements, where $\frac{1}{2} > \delta > 0$ is arbitrary. Suppose further that there are no two codewords at Hamming distance d where d is a fixed integer, $\delta n < d < (1-\delta)n$, and d is even if $q = 2$. Then $|\mathcal{C}| < (q - \varepsilon)^n$, where $\varepsilon > 0$ depends only on q and δ.

The following conjecture of Erdös and Szemerédi is also proved: If $\mathcal{F}$ is a family of subsets of $\{1, 2, \ldots, n\}$ not containing a weak Δ-system of size r (cf. Definition 1.8), then $|\mathcal{F}| < (2 - \varepsilon_r)^n$, $\varepsilon_r > 0$ holds.

An old conjecture of Larman and Rogers is established in the following stronger form: Let $\mathcal{A}$ be a collection of $4n$-dimensional (± 1)-vectors, $r \geq 2$ is a fixed integer. Suppose that A does not contain r pairwise orthogonal vectors. Then $|\mathcal{A}| < (2 - \varepsilon)^{4n}$.

All these results can be deduced from our most general result (Theorem 1.16) which concerns the intersection pattern of families of partitions. This result has further implications in Euclidean Ramsey theory as well as for isometric embeddings into the Hamming space $H(n, q)$ (cf. Theorem 9.1).

1. Introduction and statement of the results. The results of the present paper can be divided into three areas: (i) extremal set theory; (ii) coding theory; and (iii) geometry.

(i) *Extremal set theory.* Let X be an n-element set—we often suppose $X = \{1, 2, \ldots, n\}$. Define $2^X \doteq \{H, H \subseteq X\}$, and

$$\binom{X}{k} = \{H \subseteq X, |H| = k\}.$$

A subset $\mathcal{F} \subset 2^X$ is called a family. If $\mathcal{F} \subset \binom{X}{k}$, then $\mathcal{F}$ is called k-uniform. The easiest result in extremal set theory states that if $F \cap F' \neq \emptyset$ holds for all F, $F' \in \mathcal{F}$, then $|\mathcal{F}| \leq 2^{n-1}$ (proof: at most one of F, $X - F$ can belong to $\mathcal{F}$). Under the additional restriction $|F| = k$ for all $F \in \mathcal{F}$, i.e., $\mathcal{F} \subset \binom{X}{k}$, the problem becomes more difficult. The best possible bound is $|\mathcal{F} \leq \binom{n-1}{k-1}$ (when $2k \leq n$) given by the Erdös-Ko-Rado theorem (see below).

What happens if we assume $|F \cap F'| \geq t$? The answer is given by the following two theorems.

Received by the editors October 24, 1985.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 05A05; Secondary 05B99.

KATONA'S THEOREM [**K**]. *Suppose $\mathcal{F} \subset 2^X$ and $|F \cap F'| \geq t$ holds for all F, $F' \in \mathcal{F}$. Then*

$$|\mathcal{F}| \leq \begin{cases} \sum_{i \geq (n+t)/2} \binom{n}{i} & \text{if } n+t \text{ is even,} \\ 2\sum_{i \geq ((n-1)+t)/2} \binom{n-1}{i} & \text{if } n+t \text{ is odd.} \end{cases}$$

Moreover, for $t \geq 2$ equality holds if and only if

$$\mathcal{F} = \begin{cases} F \subseteq X : |F| \geq (n+t)/2 & \text{if } n+t \text{ is even,} \\ F \subseteq X : |F \cap (X - \{x\})| \geq (n-1+t)/2 & \text{for some } x \in X, \text{ if } n+t \text{ is odd.} \end{cases}$$

ERDÖS-KO-RADO THEOREM [**EKR**]. *Let $\mathcal{F} \subset \binom{X}{k}$ and suppose $|F \cap F'| \geq t$ holds for all F, $F' \in \mathcal{F}$. Then $|\mathcal{F}| \leq \binom{n-t}{k-t}$ for $n \geq n_0(k,t)$. Moreover if $n > n_0(k,t)$, then $|\mathcal{F}| = \binom{n-t}{k-t}$ is possible if and only if $\mathcal{F} = \left\{F \in \binom{X}{k} : T \subset F\right\}$ for some $T \in \binom{X}{t}$.*

REMARK. The exact value of $n_0(k,t)$ is $(k-t+1)(t+1)$. It was determined by Frankl [**F1**] for $t \geq 15$ and recently by Wilson [**W**] for the remaining values of t.

These theorems lead to the following more general problems: Let $L = \{l_1, \ldots, l_s\}$ be a set of integers satisfying $0 \leq l_1 < l_2 < \cdots < l_s < n$.

DEFINITION. Let $m(n,l)$ (resp. $m(n,k,L)$) denote the maximum of $|\mathcal{F}|$, $\mathcal{F} \subset 2^X$ (resp. $\mathcal{F} \subset \binom{X}{k}$) subject to the constraint: $|F \cap F'| \in L$ holds for all distinct F, $F' \in \mathcal{F}$.

One can reformulate the above theorems in the above terminology. For example, the Erdös-Ko-Rado theorem states

$$m(n,k,\{t,t+1,\ldots,k-1\}) = \binom{n-t}{k-t} \quad \text{for } n \geq (k-t+1)(t+1).$$

Another example of a result of this type is the following recent theorem of the second author [**R**] who proved that for $1 \leq t < k < n$,

$$m(n,k,\{0,1,\ldots,t-1\}) = \binom{n}{t} \Big/ \binom{k}{t}(1+o(1)), \quad \text{where } o(1) \to 0 \text{ as } n \to \infty.$$

For a recent review of general results concerning $m(n,k,L)$ and $m(n,L)$ see [**DF**].

The special case which is central for this paper is when $L = \{k : 0 \leq k \leq n-1,\ k \neq l\}$ for some integer l.

Let us introduce the notation:

$$m(n,\{0,1,\ldots,n-1\} - \{l\}) = m(n,\bar{l}),$$
$$m(n,k,\{0,1,\ldots,k-1\} - \{l\}) = m(n,k,\bar{l}).$$

The problem of determining or estimating $m(n,\bar{l})$, $m(n,k,\bar{l})$ goes back to Erdös [**E1**].

FRANKL-WILSON THEOREM [**FW**]. *Suppose $k-l$ is a prime power. Then*

$$(\alpha) \qquad m(n,k,\bar{l}) \leq \binom{n}{k-l-1} \qquad \text{if } k \geq 2l+1;$$

$$(\beta) \qquad m(n,k,\bar{l}) \leq \binom{n}{l}\binom{2k-1-l}{k} \Big/ \binom{2k-1-l}{l} \qquad \text{if } k \leq 2l+1.$$

FRANKL-FÜREDI THEOREM [**FF1** AND **FF2**].

(a)

$$m(n,k,\bar{l}) \leq \binom{n-l-1}{k-l-1} \quad \text{if } k \geq 2l+2 \text{ and } n > n_0(k,l).$$

Moreover, equality is attained only for $\left\{F \in \binom{X}{k}: T \subset F\right\}$ *for some* $T \in \binom{X}{l+1}$.

(b) *For* $n > n_0(l)$

$$m(n,\bar{l}) = \sum_{i=0}^{l-1}\binom{n}{i} + \begin{cases} \sum_{i \geq (n+l+1)/2}\binom{n}{i} & \text{if } n+l \text{ is odd,} \\ 2\sum_{i \geq n+l}\binom{n-1}{i} & \text{if } n+l \text{ is even.}\end{cases}$$

Moreover, equality is attained only for

$$\left\{F \subseteq X: 0 \leq |F| < l \text{ or } |F| \geq \frac{n+l+1}{2}\right\}$$

if $n = 1$ *is odd, and for*

$$\left\{F \subseteq X: 0 \leq |F| < l \text{ or } |F \cap (X - \{x\})| \geq \frac{n+l}{2}\right\}$$

($x \in X$ *is fixed*) *if* $n+l$ *is even.*

The essence of this last theorem is that weakening the assumptions of the Erdös-Ko-Rado and Katona theorems by requiring only $|F \cap F'| \neq t-1$ instead of $|F \cap F'| > t$ still leads to practically the same bounds. However, the condition $n \geq n_0(l)$ makes it impossible to apply these results if $l = \alpha n$, $\alpha > 0$.

The applications of the next-to-last theorem are restricted by the condition that $k-l$ is a prime power. In fact, if this could be removed, then it would imply some of our results.

THEOREM 1.1. *Suppose* $0 < \eta < \frac{1}{4}$ *is given. Then there exists a positive constant* $\varepsilon_0 = \varepsilon_0(\eta)$ *so that* $m(n,\bar{l}) \leq (2-\varepsilon_0)^n$ *holds for all* l, $\eta n < l < \left(\frac{1}{2} - \eta\right) n$.

Making the necessary calculations one deduces

COROLLARY 1.2.

$$m(n, \overline{[n/4]}) < 1.99^n. \tag{1}$$

COROLLARY 1.3. *For a fixed constant* ρ *with* $0 < \rho < 1$,

$$m(n, \overline{\rho n}) \leq (2 - \rho^2/2 + o(\rho^3))^n.$$

Let us mention that taking $\mathcal{F} = \{F \subseteq X: |F| > (1+\rho)n/2\}$ one obtains a rather large family satisfying $|F \cap F'| > \rho n$ and therefore $|F \cap F'| \neq \lfloor \rho n \rfloor$.

Setting $\rho = \frac{1}{4}$ for $n > n_0$ one has $|\mathcal{F}| > 1.9378^n$, showing that the upper bound in Corollary 1.2 is not far from being best possible.

Similarly, one sees that $m(n, \overline{\lfloor \rho n \rfloor}) \geq (2 - \rho^2 + o(\rho^3))^n$ holds.
Theorem 1.1 and its corollaries will be deduced from the following, more general, two-family version.

THEOREM 1.4. *Suppose* $0 < \eta < \frac{1}{4}$ *and two families* $\mathcal{F}$, $\mathcal{G} \subset 2^X$ *are given which satisfy* $|F \cap G| \neq l$ *for* $F \in \mathcal{F}$, $G \in \mathcal{G}$. *If* $\eta n \leq l \leq \left(\frac{1}{2} - \eta\right) n$, *then*

$$|\mathcal{F}|\,|\mathcal{G}| \leq (4 - \varepsilon_1(\eta))^n,$$

where $\varepsilon_1(\eta)$ *is a positive constant depending only on* η.

Let us mention that two-family versions of extremal theorems are often more useful and cannot be deduced from the corresponding one-family version. For example, if $\mathcal{F} \subset 2^X$ satisfies $|F \cap F'| = l$ for all distinct F, $F' \in \mathcal{F}$, then Ryser's theorem [**Ry**] gives $|\mathcal{F}| \leq n$ when $l > 0$ (for $l = 0$, $|\mathcal{F}| \leq n + 1$ is trivial).

On the other hand, if $X = Y \cup Z$, $|Y \cap Z| = l$, then the two families $\mathcal{A} = \{A \subseteq Y : Y \cap Z \subset A\}$, $\mathcal{B} = \{B \subseteq Z : Y \cap Z \subseteq B\}$ satisfy $|A \cap B| = l$ for all $A \in \mathcal{A}$, $B \in \mathcal{B}$, and $|\mathcal{A}|\,|\mathcal{B}| = 2^{n-l}$ which is exponential in n.

In §10 we shall prove $|\mathcal{A}|\,|\mathcal{B}| \leq 2^n$ with equality only if $l = 0$ (Proposition 10.1). There we prove some similar results for the case when $|A \cap B| \equiv l \pmod p$ is required instead of $|A \cap B| = l$ for all $A \in \mathcal{A}$, $B \in \mathcal{B}$.

A short proof of a theorem of Ahlswede et al. [**AGP**] is also given in §10.

Theorem 1.4 can be generalized in the following way.

THEOREM 1.5. *Let* $\mathcal{F}$, $\mathcal{F}' \subset 2^X$ *be two families with* $|F \cap F'| \neq \lfloor \rho n \rfloor$, $0 < \rho < 1$. *For* $0 < p < 1$ *and* $F \in \mathcal{F}$, *set* $w_p(F) = p^{|F|}(1-p)^{n-|F|}$ *and* $w_p(\mathcal{F}) = \sum w(F)$ ($w_p(\mathcal{F}')$ *is defined analogously*). *Suppose further that* $\eta > 0$ *is such that either*

(i) $0 < \eta < \rho < p - \eta < p < \frac{1}{2}$, *or*

(ii) $\frac{1}{2} \leq p < 1$ *and* $2p - 1 + \eta < \rho < p - \eta$ *holds.*

Then $w_p(\mathcal{F}) w_p(\mathcal{F}') \leq (1 - \varepsilon(\eta, p))^n$, *where* $\varepsilon(\eta, p) > 0$.

By considering the system of all m-subsets of an n-set, where $m = \lfloor \alpha n \rfloor$, and realizing that for $p = \alpha$ these sets have joint weight $> O(1/n)$ we get the following result.

COROLLARY 1.6. *Let* α, $0 < \alpha < 1$, *and* $\eta < \alpha/2$ *be given reals. Let* $\mathcal{F}$ *and* $\mathcal{F}'$ *be two families of* m*-element subsets of an* n*-set* X *such that* $|F \cap F'| \neq \lfloor \rho n \rfloor$ *for all* $F \in \mathcal{F}$, $F' \in \mathcal{F}'$, *where* $m = \lfloor \alpha n \rfloor$ *and* $\max\{0, 2\alpha - 1\} + \eta \leq \rho \leq \alpha - \eta$.

Then

$$|\mathcal{F}|\,|\mathcal{F}'| \leq \binom{n}{m}^2 (1 - \bar{\varepsilon})^n, \quad \text{where } \bar{\varepsilon} = \bar{\varepsilon}(\alpha, \eta) > 0.$$

We will further prove the following strengthening of Corollary 1.2.

THEOREM 1.7. *Given* $0 < \delta < \frac{1}{2}$, *there exist positive constants* $\sigma = \sigma(\delta)$, $\varepsilon = \varepsilon(\delta)$ *so that whenever* $\mathcal{F} \subset 2^X$ *with* $|\mathcal{F}| > (2 - \varepsilon)^n$ *and* $|n/4 - l| \leq \sigma n$, *then*

$$|\{(F, F'); F, F' \in \mathcal{F}, |F \cap F'| = l\}| > |\mathcal{F}|^2 (1 - \delta)^n. \tag{1'}$$

DEFINITION 1.8. A family $\mathcal{A} = \{A_1, A_2, \ldots, A_r\}$ is called a strong Δ-system (weak Δ-system) of size r if $A_i \cap A_j = A_1 \cap A_2$ ($|A_i \cap A_j| = |A_1 \cap A_2|$) holds for all $1 \leq i < j \leq r$, respectively.

In [**ES**], P. Erdös and E. Szemerédi proved that there exists a family $\mathcal{G}$ of subsets of a given set X so that $\mathcal{G}$ does not contain a weak Δ-system of size 3, where

$$|X| = n, \qquad |\mathcal{G}| \geq n^{\log n / 4 \log\log n}.$$

On the other hand, they conjectured that if ε is sufficiently small and $\mathcal{F} \subset 2^X$ is sufficiently large, for example, $|\mathcal{F}| > (2-\varepsilon)^n$, then $\mathcal{F}$ contains a weak Δ-system. (Note that it is also proved in [**ES**] that if $(2-\varepsilon)^n$ is replaced by $2^{(1-1/10n^{1/2})n}$, then the family contains a strong Δ-system of 3 elements.) Theorem 1.7 implies the conjecture of Erdös and Szemerédi in a stronger form:

THEOREM 1.9. *Given $r \geq 3$, there exist constants $\eta = \eta(r)$, $\varepsilon = \varepsilon(r)$ so that for every l with $|n/4 - l| \leq \eta n$ and for every $\mathcal{F} \subset 2^X$ with $|\mathcal{F}| > (2-\varepsilon)^n$, there exist $F_1, F_2, \dots, F_r \in \mathcal{F}$ with $|F_i \cap F_j| = l$, $1 \leq i < j \leq r$.*

(ii) *Coding theory.* Given a finite set Q, $|Q| = q \geq 2$, called the alphabet, by a code of length n over Q we mean a family $\mathcal{C} = \{C_1, \dots, C_m\}$ with each C_i (called a codeword) being a sequence of length n of elements from Q. The Hamming distance of two codewords $C = (a_1, a_2, \dots, a_n)$, $D = (b_1, b_2, \dots, b_n)$ is defined by $d(C,D) = |\{i,\ a_i \neq b_i\}|$. Clearly $d(C,D) = 0$ if and only if $C = D$. Define $d(\mathcal{C}) = \{d(C,D); C \neq D \in \mathcal{C}\}$. A classical problem of coding theory is to give upper and lower bounds for $\max|\mathcal{C}|$, given $d(\mathcal{C})$. Delsarte [**D**] proved an upper bound in terms of $|d(\mathcal{C})|$. This was recently strengthened by Blokhuis and the first author:

THEOREM [**Bl,F2**]. *Suppose $\mathcal{C}$ is a code of length n over Q, p is a prime, and $d(C)$ is covered by t nonzero residue classes* mod p. *Then*

$$|\mathcal{C}| \leq \sum_{i=0}^{t} (q-i)^{n-i} \binom{n}{i}.$$

This theorem often gives quite accurate bounds if $|d(\mathcal{C})|$ is small with respect to n. However, for $d(\mathcal{C}) > n/q$ it becomes practically trivial. Our result concerns the case $|d(\mathcal{C})| = n-1$.

THEOREM 1.10. *Let $\mathcal{C}$ be a code of length n over Q, and let δ satisfy $0 < \delta < \frac{1}{2}$. Suppose that $\delta n < d < (1-\delta)n$, and d is even if $q = 2$. If $d \notin d(\mathcal{C})$, then $|\mathcal{C}| \leq (q-\varepsilon)^n$ with some positive constant $\varepsilon = \varepsilon(\delta, q)$.*

REMARK. If $Q = \{a_1, a_2\}$, then $\mathcal{C}_{\text{even}} = \{C = (c_1, c_2, \dots, c_n) \colon |\{i \colon c_i = a_1\}|$ is even $\}$ shows that the condition "d is even" is necessary for $q = 2$.

(iii) *Geometry.* The Frankl-Wilson theorem has had numerous geometric applications. In particular, it implied that if n-dimensional Euclidean space R^n is partitioned into fewer than 1.2^n sets, then in at least one of the sets all distances are realized; that is, for any positive real number γ there are two points in that set at distance exactly γ. Corollary 1.2—via the methods of Larman and Rogers [**LR**]—provides an exponential lower bound as well.

Larman and Rogers [**LR**, Conjecture 2] asked whether every set of $2^d/d^2$ (± 1)-vectors in R^d contains a pair of orthogonal vectors if d is of the form $4n$. Here we prove the following stronger statement:

THEOREM 1.11. *Given r, $n \geq r \geq 2$, there exists a positive constant $\varepsilon = \varepsilon(r)$ so that in any set of more than $(2-\varepsilon(r))^{4n}$ (± 1)-vectors of length $4n$ there are r pairwise orthogonal vectors.*

Note that if d is odd, then no two ± 1-vectors in R^d are orthogonal. If $d = 4n+2$, then the collection of ± 1-vectors with an even number of -1's shows that one can have $\frac{1}{2}2^d$ vectors without an orthogonal pair. Thus the assumption $d = 4n$ is necessary. This result leads to the following problem:

Problem 1.12. Let $2 \leq r \leq n$ and let μ be the usual Lebesgue measure on the unit sphere S^{n-1} in R^n. What is $\mu(n, r) = \sup\{\mu(E)/\mu(S^{n-1})\colon E \subset S^{n-1},\ E$ is measurable, and there are no r points in E such that the vectors pointing to them from the origin are pairwise orthogonal$\}$.

This problem was stated by Witsenhausen (personal communication) in the case $r = 2$. In [**FW**], $\mu(n, 2) \leq (1.13 + o(1))^{-n}$ is proved.

Taking the open cone with half-angle 45° shows $\mu(n, 2) \geq (\sqrt{2} + o(1))^{-n}$; it is conceivable that equality holds here. However, already $\mu(3, 2)$ is unknown. An easy averaging argument along with Theorem 1.11, and the obvious inequality $\mu(n, r) < \mu(n', r)$ for $n < n'$ yield:

THEOREM 1.13. *There exists a positive constant* $\varepsilon = \varepsilon(r)$ *so that* $\mu(n, r) \leq (1 + \varepsilon)^{-n}$ *holds.*

DEFINITION. For an integer l and families $\mathcal{F}$, $\mathcal{G}$ define:

$$I_l(\mathcal{F}, \mathcal{G}) = \{(F, G)\colon F \in \mathcal{F},\ G \in \mathcal{G},\ |F \cap G| = l\},$$
$$i_l(\mathcal{F}, \mathcal{G}) = |I_l(\mathcal{F}, \mathcal{G})|.$$

For $\mathcal{F} = \{F\}$ we write $i_l(F, \mathcal{G})$ and $I_l(F, \mathcal{G})$. In §§6, 7, and 8, various stronger theorems are established.

THEOREM 1.14. *Suppose* $1 \geq p_1$, p_2, δ, *and* η, *are positive constants*, $\mathcal{F}_i \subset \binom{X}{p_i n}$, *and* l *is an integer with*

$$\max\{(p_1 + p_2 - 1 + \eta)n, \eta n\} < l < \min\{(p_1 - \eta)n, (p_2 - \eta)n\}.$$

Then there exists a positive $\varepsilon = \varepsilon(p_1, p_2, \delta, \eta)$ *so that:*

$$|\mathcal{F}_1|\,|\mathcal{F}_2| > \binom{n}{p_1 n}\binom{n}{p_2 n}(1 - \varepsilon)^n$$

implies

$$i_l(\mathcal{F}_1, \mathcal{F}_2) > i_l\left(\binom{X}{p_1 n}, \binom{X}{p_2 n}\right)(1 - \delta)^n.$$

For $l_1, l_2, \ldots, l_s$ positive integers with $l_1 + l_2 + \cdots + l_s = n$, let $\binom{X}{l_1, \ldots, l_s}$ denote the set of all ordered partitions $(A_1, \ldots, A_s)$ of X with $|A_i| = l_i$. Obviously

$$\left|\binom{X}{l_1, \ldots, l_s}\right| = \binom{n}{l_1, \ldots, l_s} = \frac{n!}{l_1! \cdots l_s!}.$$

For

$$\vec{A} = (A_1, \ldots, A_s) \in \binom{X}{l_1, \ldots, l_s} \quad \text{and} \quad \vec{B} = (B_1, \ldots, B_t) \in \binom{X}{k_1, \ldots, k_t},$$

the intersection pattern is an s by t matrix M with general entry $m_{ij} = |A_i \cap B_j|$. We denote this fact by writing $|\vec{A} \cap \vec{B}| = M$. For families

$$\mathcal{A} \subseteq \binom{X}{l_1, \ldots, l_s}, \qquad \mathcal{B} \subseteq \binom{X}{k_1, \ldots, k_t},$$

let us define $I_M(\mathcal{A},\mathcal{B}) = \{(\vec{A},\vec{B}) : \vec{A} \in \mathcal{A},\ \vec{B} \in \mathcal{B},\ |\vec{A} \cap \vec{B}| = M\}$, $i_M(\mathcal{A},\mathcal{B}) = |I_m(\mathcal{A},\mathcal{B})|$.

A necessary and sufficient condition for M to occur as intersection pattern at all is: $\sum_i m_{ij} = k_j$, $\sum_j m_{ij} = l_i, \sum_{i,j} m_{ij} = n$. Our next theorem asserts that if m_{ij} is not too small and $|\mathcal{A}|$, $|\mathcal{B}|$ are large, then this condition implies that there are many pairs with intersection pattern M.

THEOREM 1.15. *Let η, γ be arbitrary positive constants, $\mathcal{A} \subset \binom{X}{l_1,\dots,l_s}$, $\mathcal{B} \subset \binom{X}{k_1,\dots,k_t}$, and let $M = (m_{ij})$ be an s by t matrix satisfying $m_{ij} \geq \eta n$. Then there exists a positive constant $\varepsilon = \varepsilon(\eta,\gamma)$ so that*

$$|\mathcal{A}||\mathcal{B}| > (1-\varepsilon)^n \binom{n}{l_1,\dots,l_s}\binom{n}{k_1,\dots,k_t}$$

implies

$$i_M(\mathcal{A},\mathcal{B}) > (1-\gamma)^n i_M\left(\binom{X}{l_1,\dots,l_s},\binom{X}{k_1,\dots,k_t}\right).$$

Instead of the intersection pattern of two partitions one can consider several partitions $(A_1^{(1)},\dots,A_{s_1}^{(1)}),\dots,(A_1^{(r)},\dots,A_{s_r}^{(r)})$ and their intersection pattern, which will be an s_1 by s_2 by $\cdots$ by s_r array with general entry $|A_{i_1}^{(1)} \cap \cdots \cap A_{i_r}^{(r)}|$.

As we shall see Theorem 1.15 easily implies the following more general result.

THEOREM 1.16. *Let η, γ, be arbitrary positive constants, $\mathcal{A}^{(i)}$ a family of partitions of X,*

$$\mathcal{A}^{(i)} \subset \binom{X}{l_1^{(i)},\dots,l_{s_i}^{(i)}}, \qquad i = 1,\dots,r.$$

Let $M = (m_{i_1 i_2 \cdots i_r})$ be an s_1 by $\cdots$ by s_r array with all entries satisfying $m_{i_1\cdots i_r} > \eta n$. There exists a positive constant $\varepsilon = \varepsilon(\eta,\gamma)$ so that

$$i_M(\mathcal{A}^{(1)},\dots,\mathcal{A}^{(r)}) \geq (1-\gamma)^n i_M\left(\binom{X}{l_1^{(1)},\dots,l_{s_1}^{(1)}},\dots,\binom{X}{l_1^{(r)},\dots,l_1^{(r)},\dots,l_{s_r}^{(r)}}\right)$$

holds whenever

$$|\mathcal{A}^{(i)}| \geq (1-\varepsilon)^n \binom{n}{l_1^{(i)},\dots,l_s^{(i)}} \quad \textit{for all } i.$$

This last theorem might look somewhat technical and unattractive; however, it has many geometric applications.

Let us recall that a finite subset $A \subset R^d$ is called *Ramsey* if for every integer $r \geq 2$ there exists $n_o = n_o(A,r)$ with the property that for every partition of R^{n_0} into r classes, one of the classes contains a congruent copy of A.

DEFINITION 1.17. A finite subset $A \subset R^d$ is called *exponentially Ramsey* if there exists a positive real $\varepsilon = \varepsilon(A)$ so that for every partition of R^n into r classes with $n \geq d$ and $r < (1+\varepsilon(A))^n$, one of the classes contains a congruent copy of A.

It should be clear that being exponentially Ramsey is a property much stronger than just being Ramsey. The concept of a set being Ramsey was defined in 1972 by Erdös, Graham, Montgomery, Rothschild, Spencer, and Straus [**EGMRSS**], who proved that the vertices of an arbitrary-dimensional rectangular parallelepiped are Ramsey.

On the other hand they showed that if A is Ramsey, then A is spherical (it can be embedded in a sphere).

The first open question therefore was what happens with obtuse triangles.

In [**FR**] it is shown that all triangles are Ramsey. With the aid of Theorem 1.16 we can prove the following stronger statement.

THEOREM 1.18. *Let A be a simplex in R^d, i.e., A consists of $d+1$ affinely independent points. Then A is exponentially Ramsey.*

Note that the special case when A is a regular simplex also follows from Theorem 1.11.

The proof of Theorem 1.18 will appear in a separate paper—along with more results on Euclidean Ramsey theory.

2. The proof of Theorem 1.4 and its corollaries. First we give a proof of the following auxiliary result analogous to a theorem of Ahlswede and Katona [**AK**].

THEOREM 2.1. *Let $\mathcal{F}^*$ and $\mathcal{G}^*$ be two families on an n-set and let β, $0 < \beta < 1$, be such that $|F \cap G| > \beta n$ for all $F \in \mathcal{F}^*$ and $G \in \mathcal{G}^*$.*

Then

$$|\mathcal{F}^*|\,|\mathcal{G}^*| \leq 2^{2nH((1+\beta)/2)},$$

where $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the entropy function.

PROOF. Let t be the largest integer such that $|F \cap G| > t$ for all $F \in \mathcal{F}^*$ and $G \in \mathcal{G}^*$. Suppose, without loss of generality, that $|\mathcal{F}^*| \leq |\mathcal{G}^*|$. We clearly have

$$\sum_{i=0}^{a} \binom{n}{i} \geq |\mathcal{F}^*| > \sum_{i=0}^{a-1} \binom{n}{i}$$

for some $a \leq \frac{1}{2}n$.

Let $\sigma^t(\mathcal{F})$ be the set of all subsets of the underlying set having Hamming distance at most t from some member of $\mathcal{F}$. Then, by the theorem of Harper [**H**],

$$|\sigma^t(\mathcal{F}^*)| \geq \sum_{i=0}^{a-1+t} \binom{n}{i}.$$

The choice of t implies that $\sigma^t(\mathcal{F}^*) \cap \mathcal{G}^{*c} = \emptyset$, where $\mathcal{G}^{*c} = \{X - G; G \in \mathcal{G}^*\}$. Consequently,

$$|\mathcal{G}^*| \leq \sum_{i=a+t}^{n} \binom{n}{i}$$

and hence

$$|\mathcal{F}^*|\,|\mathcal{G}^*| \leq \sum_{i=0}^{a} \binom{n}{i} \sum_{i=a+t}^{n} \binom{n}{i}.$$

Moreover, by the Chernoff inequality [**C**] we have

$$\sum_{i=0}^{a} \binom{n}{i} \leq 2^{H(a/n)\cdot n} \quad \text{and} \quad \sum_{i=a+t}^{n} \binom{n}{i} \leq 2^{H((a+t)/n)n},$$

provided that $(a+t)/n \geq \frac{1}{2}$. Since $(1-x-\beta/2)-(\frac{1}{2}-\beta/2) = (\frac{1}{2}-\beta/2)-(x-\beta/2)$ and $H'(x) = \log((1-x)/x)$ is monotone decreasing, we have

$$\begin{aligned} H(x+\beta/2) - H(\tfrac{1}{2}+\beta/2) &= H(1-x-\beta/2) - H(\tfrac{1}{2}-\beta/2) \\ &\leq H(\tfrac{1}{2}-\beta/2) - H(x-\beta/2). \end{aligned}$$

Thus, we see that the function $H(x-\beta/2) + H(x+\beta/2)$ attains its maximum for $x = \frac{1}{2}$. Hence, for $a+t \geq n/2$,

$$|\mathcal{F}^*|\,|\mathcal{G}^*| \leq 2^{(H(a/n)+H((a+t)/n))n} \leq 2^{2nH(1/2+\beta/2)}.$$

If $a+t < n/2$, we have (similar to the above)

$$|\mathcal{F}^*|\,|\mathcal{G}^*| \leq 2^{nH(a/n)}2^n \leq 2^{(H(1/2-\beta)+H(1/2))n} \leq 2^{2nH(1/2+\beta/2)}. \qquad \square$$

We need one more auxiliary result.

THEOREM 2.2. *Let $\mathcal{A}$ and $\mathcal{B}$ be families on an n-set and let κ, $0 < \kappa < \frac{1}{2}$, be such that*

$$|A \cap B| < (\tfrac{1}{2} - \kappa)n \quad \text{for all } A \in \mathcal{A},\ B \in \mathcal{B}.$$

Then for an arbitrary λ satisfying $0 \leq \lambda < \kappa$ we have

$$|\mathcal{A}|\,|\mathcal{B}| \leq \max\{2^n 2^{H(1/2-\lambda)n+1}, 2^{2(H((1+\kappa-\lambda)/2)n+1)}\}. \tag{2}$$

PROOF. Set $\mathcal{A}_0 = \{A \in \mathcal{A} : |A| \leq (\frac{1}{2}-\lambda)n\}$. Then $|\mathcal{A}_0| \leq 2^{H(1/2-\lambda)n}$. Thus if $|\mathcal{A}_0| \geq \frac{1}{2}|\mathcal{A}|$, the statement follows.

Arguing similarly with $\mathcal{B}$, we may therefore suppose that at least half the members of both $\mathcal{A}$ and $\mathcal{B}$ have size greater than $(\frac{1}{2}-\lambda)n$. Let $\mathcal{A}^*$ and $\mathcal{B}^*$ be the corresponding subfamilies and note that $|\mathcal{A}|\,|\mathcal{B}| \leq 4|\mathcal{A}^*|\,|\mathcal{B}^*|$.

Define $\mathcal{C}^* = \{X - B : B \in \mathcal{B}^*\}$. Then for $C \in \mathcal{C}^*$ and $A \in \mathcal{A}^*$ we have

$$|A \cap C| = |A \cap (X-B)| = |A| - |A \cap B| > (\kappa - \lambda)n.$$

Applying Theorem 2.1 to $\mathcal{C}^*$ and $\mathcal{A}^*$ and noting $|\mathcal{C}^*| = |\mathcal{B}^*|$, the statement follows. $\square$

For the proof of Theorem 1.4 we will introduce some notation. Let $\mathcal{F}$, $\mathcal{G}$ be two set-systems on an n-set X. Let a, b, $0 \leq a \leq b \leq n$, be positive integers. We will write that $(\mathcal{F}, \mathcal{G}) \in \mathcal{P}(n, [a,b])$ iff $|F \cap G| \notin [a,b]$ for all $F \in \mathcal{F}$, $G \in \mathcal{G}$. For a fixed element $x \in X$ set

$$\begin{aligned} \mathcal{F}_0 &= \mathcal{F}_0(x) = \{F; x \notin F \in \mathcal{F}\}, \\ \mathcal{F}_1 &= \mathcal{F}_1(x) = \{F - \{x\}; x \in F \in \mathcal{F}\}, \\ \mathcal{G}_0 &= \mathcal{G}_0(x) = \{G; x \notin G \in \mathcal{G}\}, \\ \mathcal{G}_1 &= \mathcal{G}_1(x) = \{G - \{x\}; x \in G \in \mathcal{G}\}. \end{aligned}$$

Below we shall use the following easy

Observation. Let $(\mathcal{F}, \mathcal{G}) \in \mathcal{P}(n, [a,b])$. Then

$$\begin{aligned} &(\mathcal{F}_1, \mathcal{G}_1) \in \mathcal{P}(n-1, [a-1, b-1]), \\ &(\mathcal{F}_0, \mathcal{G}_0 \cup \mathcal{G}_1) \in \mathcal{P}(n-1, [a,b]), \\ &(\mathcal{F}_1, \mathcal{G}_0 \cap \mathcal{G}_1) \in \mathcal{P}(n-1, [a-1, b]). \end{aligned}$$

For set systems $\mathcal{F}$ and $\mathcal{G}$ we will write

$$\mathcal{P}(\mathcal{F}) = \frac{|\mathcal{F}|}{2^{|\cup \mathcal{F}|}}, \qquad \mathcal{P}(\mathcal{G}) = \frac{|\mathcal{G}|}{2^{|\cup \mathcal{G}|}}.$$

We will also use the following proposition.

PROPOSITION 2.3. *Suppose x, y, δ are positive reals with $\delta \le 1/10$ satisfying:*
(i) $(1+y)^2 \le 1+\delta$,
(ii) $(1+x)(1-y) \le 1+\delta$.
Then
(iii) $(1+y)(1-x) > 1-\delta-2\delta^2$ *holds.*

PROOF. From (ii), using also $y < \delta/2 < 1$ (which follows from (i)), we have $x \le (\delta+y)/(1-y)$. Since the LHS of (iii) is monotone decreasing in x we may assume $x = (\delta+y)/(1-y)$. Substituting this into (iii), multiplying by $1-y$, and rearranging we obtain

$$2\delta^2(1-y) \ge 2(\delta+y)y.$$

To verify the latter inequality note that the LHS is monotone decreasing, the RHS is increasing in y for $y < \delta/2$, in view of (i), and for $y = \delta/2$ we obtain

$$2\delta^2(1-\delta/2) > 3\delta^2/2$$

which holds because $\delta < \frac{1}{2}$.

Now we can give the proof of Theorem 1.4. The proof is based on the following

ALGORITHM. (a) Set $m = n$, $a = l$, $b = l$, and fix

$$\delta = \delta(\eta) \tag{3}$$

sufficiently small positive real.

(b) Check whether $a = 0$. If yes, terminate; if not go to (c).

(c) Check whether $b = m$. If yes, terminate; if not go to (d).

(d) Check whether $p(\mathcal{F}_1)p(\mathcal{G}_1) > (1+\delta)p(\mathcal{F})p(\mathcal{G})$. If yes, set $\mathcal{F} = \mathcal{F}_1$, $\mathcal{G} = \mathcal{G}_1$, $a = a-1$, $b = b-1$, and go to (h); if not go to (e).

(e) Choose $\mathcal{F}_1$ or $\mathcal{G}_1$ (say $\mathcal{F}_1$) with $p(\mathcal{F}_1) \le \sqrt{1+\delta}\,p(\mathcal{F})$ and go to (f).

(f) Check whether $p(\mathcal{G}_0 \cup \mathcal{G}_1)p(\mathcal{F}_0) > (1+\delta)p(\mathcal{G})p(\mathcal{F})$. If yes, set $\mathcal{F} = \mathcal{F}_2$, $\mathcal{G} = \mathcal{G}_0 \cup \mathcal{G}_1$, and go to (h); if not go to (g).

(g) Set $\mathcal{F} = \mathcal{F}_1$, $\mathcal{G} = \mathcal{G}_0 \cap \mathcal{G}_1$, $a = a-1$, and go to (h).

(h) Set $m = m-1$ and go to (b).

REMARK. Note that in (d) and (f) we obtain families with

$$p(\mathcal{F}')p(\mathcal{G}') > (1+\delta)p(\mathcal{F})p(\mathcal{G}).$$

Before applying (e) we have

$$\frac{p(\mathcal{F}_1)}{p(\mathcal{F})} \le \sqrt{1+\delta} \quad \text{and} \quad \frac{p(\mathcal{G}_0 \cup \mathcal{G}_1)}{p(\mathcal{G})} \cdot \frac{p(\mathcal{F}_0)}{p(\mathcal{F})} \le 1+\delta.$$

Set

$$\frac{p(\mathcal{F}_1)}{p(\mathcal{F})} = 1+y \quad \text{and} \quad \frac{p(\mathcal{G}_0 \cup \mathcal{G}_1)}{p(\mathcal{G})} = 1+x.$$

Suppose for now that $y < 0$, i.e., $p(\mathcal{F}_1) < p(\mathcal{F}_0)$. Then since

$$(1+x)(1-y) + (1-x)(1+y) > 2,$$

we infer that $(1-x)(1+y) > 1-\delta$, i.e.,

$$p(\mathcal{G}_0 \cap \mathcal{G}_1)p(\mathcal{F}_1) > (1-\delta)p(\mathcal{F})p(\mathcal{G})$$

holds. If $y \ge 0$, then the assumptions of the proposition are satisfied and we get in a similar way

$$p(\mathcal{G}_0 \cap \mathcal{G}_1)p(\mathcal{F}_1) \ge (1-\delta-2\delta^2)p(\mathcal{F})p(\mathcal{G}).$$

Summarizing: If $(\mathcal{F}, \mathcal{G}) \in \mathcal{P}(m, [a, b])$ then either

$$p(\mathcal{F}')p(\mathcal{G}') > (1+\delta)p(\mathcal{F})p(\mathcal{G}), \tag{4}$$

where

$$(\mathcal{F}', \mathcal{G}') \in \mathcal{P}(m-1, [a,b]) \cup \mathcal{P}(m-1, [a-1, b-1])$$

or

$$p(\mathcal{F}')p(\mathcal{G}') > (1-\delta-2\delta^2)p(\mathcal{F})p(\mathcal{G}) \tag{5}$$

with

$$(\mathcal{F}', \mathcal{G}') \in \mathcal{P}(m-1, [a-1, b]).$$

Let l, $\eta n \le l \le (\frac{1}{2} - \eta)n$, be the integer from the statement of Theorem 1.4. Set $l = \rho n$ and let γn be the number of steps before the algorithm terminates. Let αn be the number of steps for which (4) holds and βn such that (5) holds. Let $\mathcal{F}^*, \mathcal{G}^*$ be the families with which the algorithm terminates. We have

$$1 \ge p(\mathcal{F}^*)p(\mathcal{G}^*) \ge (1+\delta)^{\alpha n}(1-\delta-2\delta^2)^{\beta n} p(\mathcal{F})p(\mathcal{G}). \tag{6}$$

Suppose now that

$$p(\mathcal{F})p(\mathcal{G}) \ge (1-\delta^2)^n. \tag{7}$$

This yields

$$\alpha \ln(1+\delta) + \beta \ln(1-\delta-2\delta^2) + \ln(1-\delta^2) \le 0$$

and hence,

$$\alpha - \beta \le \frac{\beta \ln(1/(1+\delta)(1-\delta-2\delta^2)) + \ln(1/(1-\delta^2))}{\ln(1+\delta)}.$$

Using the fact that $x \ge \ln(1+x) \ge x - x^2/2$ holds for any $x \ge 0$, we conclude

$$\alpha - \beta \le \left(\beta \frac{3\delta + 2\delta^2}{(1+\delta)(1-\delta-2\delta^2)} + \frac{\delta}{1-\delta^2}\right)\frac{1}{1-\delta/2}$$

which gives

$$\alpha - \beta \le 3\delta \tag{8}$$

if $\delta \le 1/10$. Now we distinguish two cases according to where the algorithm terminates. Suppose that $a = 0$; this means that $\mathcal{F}^*$, $\mathcal{G}^*$ are such that $|\mathcal{F}^* \cap \mathcal{G}^*| \ge b$ for all $F^* \in \mathcal{F}^*$, $G^* \in \mathcal{G}^*$, and some $b \ge 0$. Since in each step for which (5) holds, the interval $[a, b]$ gets longer by one and this happens only if (5) holds, we conclude that $b = \beta/n$. Both $\mathcal{F}^*$ and $\mathcal{G}^*$ are families of subsets of a $(1-\gamma)n$-element set. Take the set y of the remaining γn elements and consider

$$\tilde{\mathcal{F}} = \{F^* \cup \tilde{Y}; F^* \in \mathcal{F}^*, \tilde{Y} \subset Y\},$$
$$\tilde{\mathcal{G}} = \{C^* \cup \tilde{Y}; G^* \in \mathcal{G}^*, \tilde{Y} \subset Y\}.$$

Clearly $|F \cap G| \ge b$ for every $F \in \tilde{\mathcal{F}}$, $G \in \tilde{\mathcal{G}}$. Thus, according to Theorem 2.1, we get

$$4^{(H((1+\beta)/2)-1)n} \ge p(\tilde{\mathcal{F}})p(\tilde{\mathcal{G}}). \tag{9}$$

Moreover, we have

$$p(\tilde{\mathcal{F}})p(\tilde{\mathcal{G}}) = p(\mathcal{F}^*)p(\mathcal{G}^*)$$

and hence by (6) and (9),

(10) $$4^{(H((1+\beta)/2)-1)n} \geq p(\mathcal{F}^*)p(\mathcal{G}^*) \geq (1+\delta)^{\alpha n}(1-\delta-2\delta^2)^{\beta n}p(\mathcal{F})p(\mathcal{G}).$$

Since $\alpha+\beta=\gamma\geq\rho$, we infer (using (8)) that

(11) $$\beta \geq \rho/2 - 3\delta/2,$$

Now (10) and (11) yields

(12) $$\begin{aligned}(p(\mathcal{F})p(\mathcal{G}))^{1/n} &\leq \frac{4^{H((1+\beta)/2)-1}}{(1+\delta)^{\alpha}(1-\delta-2\delta^2)^{\beta}}\\ &\leq \frac{4^{H(1/2+\rho/4-3\delta/4)-1}}{(1+\delta)^{\alpha}(1-\delta-2\delta^2)^{\beta}}.\end{aligned}$$

Let $\delta_1=\delta_1(\eta)$ be the supremum of all $\delta>0$ such that

$$4(1+\delta)^{\alpha}(1-\delta-2\delta^2)^{\beta}(1-\delta^2) \geq 4^{H(1/2+\rho/4-3\delta/4)}$$

hold for all ρ, $\eta\leq\rho\leq\frac{1}{2}-\eta$. We clearly have $\delta_1(\eta)>0$ by an easy continuity argument. It follows now from (7) and (12) that

$$(p(\mathcal{F})p(\mathcal{G}))^{1/n} \leq 1-\delta_1^2.$$

Suppose now that the second possibility occurs, i.e., $b=m$. Similar to the preceding case, βn is the length of the interval $[a,b]$ at which the algorithm terminates and, since $m=b$ we get

(13) $$a = m - \beta n.$$

Further, since the total number of steps of the algorithm equals (by definition) $(\alpha+\beta)n$ and also equals $n-m$ we find that

(14) $$(a+\beta)n = n-m.$$

Comparing (14) and (9) we get

$$(2\beta+3\delta)n > n-m$$

and since $(\frac{1}{2}-\eta)n \geq \rho n \geq m$, we infer that

(15) $$\beta \geq 1/4+\eta/2-3\delta/2.$$

As in the preceding case, consider the two families $\tilde{\mathcal{F}}$, $\tilde{\mathcal{G}}$ now defined in the following way: Let Y be a set of $(\frac{1}{2}-\eta)n-m$ elements and set

$$\tilde{\mathcal{F}} = \{F^*\cup\tilde{Y}; F^*\in\mathcal{F}^*, \tilde{Y}\subset Y\},$$
$$\tilde{\mathcal{G}} = \{G^*\cup\tilde{Y}; G^*\in\mathcal{G}^*, \tilde{Y}\subset Y\}.$$

We again have $p(\tilde{\mathcal{F}})p(\tilde{\mathcal{G}}) = p(\mathcal{F}^*)p(\mathcal{G}^*)$ and, because of (13), also

$$|\tilde{F}\cap\tilde{G}| \leq (\tfrac{1}{2}-\eta-\beta)n \leq \tfrac{1}{2}(\tfrac{1}{2}-\eta)n-(\eta-\tfrac{3}{2})n.$$

Assume without loss of generality assume that $\frac{3}{2}\delta<\eta$. Set $\kappa=(\eta-\frac{3}{2}\delta)/(1/2-\eta)$, $n'=(\frac{1}{2}-\eta)n$, and take λ_0 to minimize the RHS of (2). Let us define $\xi=\xi(\eta,\delta)$ by

$$2^{n'}2^{H(1/2-\lambda_0)n'+1} = 2^{2(H((1+\kappa-\lambda_0)/2)n'+1)} = \xi^n.$$

Then, by (10) and Theorem 2.2 (since $\delta < \frac{2}{3}\eta$ we have $\kappa > 0$) we infer that

$$(p(\mathcal{F})p(\mathcal{G}))^{1/n} \le \frac{(p(\tilde{\mathcal{F}})p(\tilde{\mathcal{G}}))^{1/n}}{(1+\delta)^\alpha(1-\delta-2\delta^2)^\beta} \le \frac{\xi(\eta,\delta)}{(1+\delta)^\alpha(1-\delta-2\delta^2)^\beta},$$

Let δ_2 be the supremum of all $\delta < \frac{2}{3}\eta$ such that

$$(1-\delta-2\delta^2)^\beta(1+\delta)^\alpha(1-\delta^2) \ge \xi(\eta,\delta)$$

holds. Again, by a continuity argument, we obtain $\delta_2 > 0$. Set $\delta = \min\{\delta_1, \delta_2\}$ (cf. (3)). Now $\varepsilon_1 = 2\delta^2$ yields $|\mathcal{F}|\,|\mathcal{G}| < (2-\varepsilon_1)^n$. $\square$

Theorem 1.1 follows now by setting $\mathcal{F} = \mathcal{G}$.

PROOF OF COROLLARY 1.2. Here, one goes through the proof of Theorem 1.4 again. Set $\rho = \frac{1}{4} = \eta$ and $\delta = 0.1$. One checks that in Proposition 2.3(iii) the improved lower bound $(1+y)(1-x) \ge 0.884728$ holds. Suppose first that the algorithm terminates with $a = 0$. Then we can use the next, improved version of (12).

$$(p(\mathcal{F})p(\mathcal{G}))^{1/n} \le 4^{H((1+\beta)/2)-1}1.1^{-\alpha}0.884728^{-\beta}.$$

Noting that the RHS is a decreasing function of α and since $\alpha + \beta \ge \frac{1}{4}$, we can suppose $\alpha = \frac{1}{4} - \beta$ and derive

$$(p(\mathcal{F})p(\mathcal{G}))^{1/n} \le 4^{H((1+\beta)/2)-1}1.1^{-1/4}1.24332^{\beta}.$$

Maximizing the RHS for $0 \le \beta \le \frac{1}{4}$, one finds $p(\mathcal{F})p(\mathcal{G}) \le 0.99^n$, which yields (2).

Suppose next that the algorithm terminates with $b = m$. Since $b \le n/4$, we infer $\gamma = \alpha + \beta \ge \frac{3}{4}$. Now using (6) instead of (12) with the improved lower bound from Proposition 2.3(iii), we infer

$$(p(\mathcal{F})p(\mathcal{G}))^{1/n} \le 1.1^{-3/4}1.24332^{\beta}.$$

Since $\beta \le \frac{1}{4}$, this yields $p(\mathcal{F})p(\mathcal{G}) < 0.984^n$, concluding the proof of (1). $\square$

COROLLARY 2.4. *If $\rho \ll 1$, then $(p(\mathcal{F})p(\mathcal{G}))^{1/n} \lesssim 1-\rho^2/4$ for any two families $\mathcal{F}$, $\mathcal{G}$ with the property $|\mathcal{F} \cap \mathcal{G}| \ne \lfloor \rho n \rfloor$ for all $F \in \mathcal{F}$, $G \in \mathcal{G}$.*

PROOF. Setting $\delta = \frac{3}{2}$ and recalling that α, $\beta \le \rho$, we have

$$H\left(\frac{1+\beta}{2}\right) - 1 = -\frac{\beta^2}{2}\log_2 e + O(\rho^3),$$

$$(1+\delta)^\alpha = 1 + \alpha\delta + O(\rho^3), \qquad (1-\delta-2\delta^2)^\beta = 1 - \beta\delta + O(\rho^3).$$

Thus, (12) gives

$$(p(\mathcal{F})p(\mathcal{G}))^{1/n} \le \frac{e^{-\beta^2+O(\rho^3)}}{1+(\alpha-\beta)\delta+O(\rho^3)} = \frac{1-\beta^2+O(\rho^3)}{1+(\alpha-\beta)\delta+O(p^3)}.$$

Using $\alpha + \beta \ge \rho$, we infer

$$(p(\mathcal{F})p(\mathcal{G}))^{1/n} \le \frac{1-\beta^2}{1+(\rho-2\beta)\rho/2} + O(\rho^3).$$

To conclude the proof, we show that for $\beta \le \rho$ the RHS is bounded from above by $1 - \rho^2/4 + O(\rho^3)$, or equivalently,

$$1-\beta^2 \le \left(1-\frac{\rho^2}{4}\right)\left(1+\frac{\rho^2}{2}-\beta\rho\right) + O(\rho^3) = 1 - \beta\rho + \frac{\rho^2}{4} + O(\rho^3).$$

Comparing the two extreme sides gives

$$\frac{\rho^2}{4} - \rho\beta + \beta^2 = \left(\frac{\rho}{2} - \beta\right)^2 \geq 0. \qquad \square$$

Now Corollary 1.3 follows again if we set $\mathcal{F} = \mathcal{G}$.

3. Proof of Theorem 1.5. We indicate the proof of (i). (Going over to the complements of sets in $\mathcal{F}$ and $\mathcal{F}'$, we can get (ii) from (i)). For (i) one needs the same proof as for Theorem 1.4 with $p(\mathcal{F})$ replaced by $w(\mathcal{F})$ (observe that for $p = \frac{1}{2}$, $p(\mathcal{F}) = w(\mathcal{F})$) and with Theorem 2.1 replaced by the following

THEOREM 3.1. *Let $\mathcal{F}^*$ and $\mathcal{G}^*$ be two families on an n-set and let β and p be such that:*

(1) $0 < \beta < p \leq \frac{1}{2}$, and

(2) $|F \cap G| > \beta n$ for any $F \in \mathcal{F}^*$ and $G \in \mathcal{G}^*$.

Then

$$w_p(\mathcal{F}^*)w_p(\mathcal{G}^*) \leq 2^{2[H((1+\beta)/2)-1]n}.$$

PROOF. First we realize that if $w(\mathcal{F}^*)w(\mathcal{G}^*)$ is maximal subject to $|F \cap G| > \beta n$, then $\mathcal{F}^*$, $\mathcal{G}^*$ are up-sets (an up-set is a family $\mathcal{F}$ such that $F \in \mathcal{F}$ and $F \subset F'$ implies $F' \in \mathcal{F}$). Suppose now that $\mathcal{F}$ is an up-set. We prove by induction on n that $w_{p'}(\mathcal{F}) \geq w_p(\mathcal{F})$ whenever $p' \geq p$. The statement is trivial for $n = 1$. To prove the induction step consider

$$w_p(\mathcal{F}) = pw_p(\mathcal{F}_1) + (1-p)w_p(\mathcal{F}_0).$$

Since $\mathcal{F}$ is an up-set we have $\mathcal{F}_0 \subset \mathcal{F}_1$ and thus $w(\mathcal{F}_0) \leq w(\mathcal{F}_1)$. Hence

$$\begin{aligned} pw_p(\mathcal{F}_1) + (1-p)w_p(\mathcal{F}_0) &\leq pw_{p'}(\mathcal{F}_1) + (1-p)w_{p'}(\mathcal{F}_0) \\ &\leq p'w_{p'}(\mathcal{F}_1) + (1-p')w_{p'}(\mathcal{F}_0) = w_{p'}(\mathcal{F}) \end{aligned}$$

(the first inequality follows by induction). Thus for $p < \frac{1}{2}$ we have by the above fact and Theorem 3.1,

$$w_p(\mathcal{F}^*)w_p(\mathcal{G}^*) \leq p(\mathcal{F}^*)p(\mathcal{G}^*) \leq 2^{2[H((1+\beta)/2)-1]n}. \qquad \square$$

Note that instead of Theorem 3.1 one can also use the following result of the first author [**F4**].

THEOREM 3.2. *If $\mathcal{A} \subset \binom{X}{k}$, $\mathcal{B} \subset \binom{X}{l}$, $|X| = n$, and $|A \cap B| \geq t$ for any $A \in \mathcal{A}$ and $B \in \mathcal{B}$, then either $|\mathcal{A}| \leq \binom{n}{k-t}$ or $|\mathcal{B}| \leq \binom{n}{l-t}$.*

Actually, for p small, we obtain a better bound in this way.

4. Proof of Theorem 1.7. First let us give the proof of a well-known simple fact which we will often need in what follows.

LEMMA 4.1 (THE BIPARTITE GRAPH COUNTING PRINCIPLE). *Suppose that G is a bipartite graph with vertex sets A and B, $|A| = a$, $|B| = b$. Assume further that G is regular on both sides, of respective degrees e and f. Also, let A_0 be a subset of A, $|A_0|/|A| = c$, and let G_0 be the subgraph of G spanned by A_0 and B. Then there are at least $c\,|B|/2$ vertices in B with degree at least $cf/2$. Moreover these vertices cover at least half of the edges of G_0.*

PROOF. The number of edges in G_0 is $|A_0|e = c|A|e = c|B|f$. Let B_0 be the set of vertices in B having degrees more than $cf/2$ in G_0. Then the number of edges covered by vertices in $B - B_0$ is at most $|B - B_0|cf/2 \le |B|cf/2$. That is, at least half the edges are incident to vertices in B_0, proving the second statement. Since no vertex in B_0 has degree more than f, $|B_0|f \ge |B|cf/2$ follows. Thus $|B_0| \ge c|B|/2$. □

Now we turn to the proof of Theorem 1.7.

We can clearly suppose that all members of $\mathcal{F}$ have the same cardinality k. We can further suppose that $2k = n$ (if $2k < n$ we find by an averaging argument a $2k$-subset of the underlying n-set that contains at least $\binom{2k}{k}(1-\varepsilon/2)^n$ members of $\mathcal{F}$ and consider this subsystem); if $2k \ge n$, add $2k - n$ new points—we denote the family which we get in both cases by $\mathcal{F}^*$. Clearly $|\mathcal{F}^*| \ge 2^{2k}(1-\varepsilon_0)^{2k}$, where $\varepsilon_0 \to 0$ as $\varepsilon \to 0$. We can further clearly suppose that the cardinality of the underlying set X is divisible by four. (If $|X| = 4n + 2$, then we add two new points x, y and extend all members of the family $\mathcal{F}^*$ by a point x, i.e. consider the system $\{F \cup \{x\}; F \in \mathcal{F}^*\}$.) Summarizing: it suffices to prove the statement for $|X| = 4m$ and $|F| = 2m$ for all $F \in \mathcal{F}$.

Set $|\mathcal{F}| = \binom{4m}{2m} f(m)$. We prove that there exists $\varepsilon > 0$ and $\sigma > 0$ such that $f(m) > (1-\varepsilon)^m$ and $|1 - l/m| < 4\sigma$ imply (1′).

For every $A \in [X]^{2m}$, denote by x_A the number of all sets $F \in \mathcal{F}$ with the property $|F \cap A| = m$. We clearly have

$$\sum\{x_A \colon A \in [X]^{2m}\} = |\mathcal{F}| \binom{2m}{m}^2 .$$

Since

$$\sum x_A \binom{4m}{2m}^{-1} = \binom{2m}{m}^2 f(m),$$

we have

$$\sum \left\{ x_A; x_A \ge \frac{1}{2} \binom{2m}{m}^2 f(m) \right\} \ge \frac{1}{2} |\mathcal{F}| \binom{2m}{m}^2 = \frac{1}{2} \binom{4m}{2m} \binom{2m}{m}^2 f(m).$$

Fix $\alpha > 0$ such that $(e/\alpha)^{4\alpha} \le (1-\delta/4)/(1-\delta/2)$ and denote by y_A the number of pairs F, $F' \in \mathcal{F}$ satisfying

$$(16)\qquad |F \cap A| = m, \quad |F' \cap A| = m, \quad |F \cap F' \cap A| = \alpha m, \quad |F \cap F'| = l.$$

Then

$$\sum\{y_A; A \in [X]^{2m}\} \le \binom{m}{\alpha m}^4 \binom{4m}{2m} \binom{2m}{m}^2 (1-\delta)^m (1+\xi)^m,$$

where $\xi \to 0$ as $\sigma \to 0$, and hence for $\sigma \le \sigma(\delta)$

$$(17)\qquad \sum y_A \binom{4m}{2m}^{-1} \le \binom{m}{\alpha m}^4 \binom{2m}{m}^2 \left(1 - \frac{\delta}{2}\right)^m .$$

Combining (16) and (17) we see that there exists $A_0 \in [X]^{2m}$ such that $x_{A_0} \geq \frac{1}{2}\binom{2m}{m}^2 f(m)$ and $y_{A_0} \leq \frac{1}{4}x_{A_0}$, since otherwise,

$$\begin{aligned}\sum\left\{y_A; x_A \geq \frac{1}{2}\binom{2m}{m}^2 f(m)\right\} &\geq \frac{1}{8}\binom{4m}{2m}\binom{2m}{m}^2 f(m)\\ &\geq \binom{m}{\alpha m}^4\binom{4m}{2m}\binom{2m}{m}^2\left(1-\frac{\delta}{2}\right)^m\\ &\geq \sum\{y_A; A \in [X]^{2m}\}\end{aligned}$$

which yields a contradiction. Note that the second inequality follows for $n \geq n_0$ since $\varepsilon < \delta/4$ and $(e/\alpha)^{4\alpha} \leq (1-\delta/4)/(1-\delta/2)$. Let $\mathcal{F}_{A_0}$ be the family consisting of those $F \in \mathcal{F}$ for which $|F \cap A_0| = m$ holds. Take the family $\overline{\mathcal{F}}_{A_0}$ which we get from $\mathcal{F}_{A_0}$ after deleting those $F \in \mathcal{F}_{A_0}$ for which there exists $F' \in \mathcal{F}_{A_0}$ such that (16) holds. Set $\mathcal{B} = \{B \in [A_0]^m$; there exists at least $\frac{1}{8}\binom{2m}{m}f(m)$ different $F \in \overline{\mathcal{F}}_{A_0}$ with $B \subset F\}$. Since

$$|\overline{\mathcal{F}}_{A_0}| \geq x_{A_0} - 2y_{A_0} \geq \frac{1}{2}x_{A_0} \geq \frac{1}{4}\binom{2m}{m}^2 f(m),$$

we find that there exists $B,\ B' \in \mathcal{B}$ with $|B \cap B'| = \alpha n$ (this holds for $\varepsilon < \varepsilon_0$, where $\varepsilon_0(\eta)$, $\eta = \alpha/2$ is from Theorem 1.2). Set $\mathcal{F}_B = \{F - B; B \subset F \in \overline{\mathcal{F}}_{A_0}\}$, $\mathcal{F}_{B'} = \{F - B';\ B' \subset F \in \overline{\mathcal{F}}_{A_0}\}$. Both $\mathcal{F}_B$ and $\mathcal{F}_{B'}$ are families on the $2m$-set $X - A_0$, and

$$|\mathcal{F}_B| \geq \frac{1}{4}\binom{2m}{m}f(m) \quad \text{and} \quad |\mathcal{F}_{B'}| \geq \frac{1}{4}\binom{2m}{m}f(m).$$

Setting

$$\eta = \min\left\{\frac{l}{2m} - \frac{\alpha}{2}, \frac{1}{2} + \frac{\alpha}{2} - \frac{l}{2m}\right\},$$

and taking $\varepsilon_1 = \varepsilon_1(\eta)$ sufficiently small with $0 < \varepsilon < \varepsilon_1$, we can use Theorem 1.4 to obtain $(F - B) \in \mathcal{F}_B$ and $(F' - B') \in \mathcal{F}_{B'}$ so that $|(F-B)\cap(F'-B')| = l - \alpha m$ and thus $|F \cap F'| = l$. Consequently, $F,\ F' \in \overline{\mathcal{F}}_{A_0}$ satisfy (16), a contradiction. □

Proof of Theorems 1.9 and 1.11.

PROOF OF THEOREM 1.9. We apply induction on r. The case $r = 2$ is covered by Theorem 1.7. Suppose that the statement is true for r with $\eta(r), \varepsilon(r)$.

We want to apply Theorem 1.7 with $\delta = \frac{1}{2}\varepsilon(r)$. Set

$$\eta(r+1) = \min\{\sigma(\delta), \eta(r)\}, \qquad \varepsilon(r+1) = \min\{\varepsilon(\delta), \tfrac{1}{2}\varepsilon(r)\}.$$

Then Theorem 1.7 guarantees the existence of $F_0 \in \mathcal{F}$ so that

$$|I_l(F_0, \mathcal{F})| > (1 - \tfrac{1}{2}\varepsilon(r))^n|\mathcal{F}| > (1 - \tfrac{1}{2}\varepsilon(r))^n(2 - \varepsilon(r+1))^n > (2 - \varepsilon(r))^n.$$

By the induction hypothesis one finds r sets $F_1, \dots, F_r$ so that $(F_0, F_i) \in I_l(F_0, \mathcal{F})$ and $|F_i \cap F_j| = l$ for $1 \leq i < j \leq r$. Thus $F_0,\ F_1, \dots, F_r$ are the desired sets. □

PROOF OF THEOREM 1.11. Let $\mathcal{V}$ be a family of $(2 - \varepsilon(r))^{4n}$ ± 1-vectors of length $4n$. To each $\vec{v} \in \mathcal{V}$ assign a subset $S(\vec{v})$ of $\{1, 2, \dots, 4n\}$ consisting of the positions of the entries of $\vec{v}$ which are equal to 1. It is easy to check that $\vec{v}$ and

$\vec{v}'$ are orthogonal iff $|S(\vec{v})\Delta S(\vec{v}')| = 2n$. Choose k, $1 < k < 4n$, so that $\mathcal{S}_k = \{S(\vec{v}); \vec{v} \in \mathcal{V}, |S(\vec{v})| = k\}$ has maximal size. Then $|\mathcal{S}_k| \geq |\mathcal{V}|/4n \geq (2-\varepsilon(r))^{4n}/4n$. Consequently $k = 2n(1+o(1))$. For S, $S' \in \mathcal{S}_k$, $|S\Delta S'| = 2n$ is equivalent to $|S \cap S'| = k - n$. Now Theorem 1.9 can be applied to complete the proof. □

REMARK. In [**F3**] for n an odd prime, it is shown that the maximum number of ± 1-vectors of length $4n$ without two of them being orthogonal is exactly $4\sum_{0\leq i<n}\binom{4n-1}{i}$. It is conjectured there that the same holds for all n. The methods of [**F3**] are completely powerless for $r \geq 3$.

6. The proof of Theorem 1.14: a special case. In this section we prove a special case of Theorem 1.14. The general case is deduced in the next section.

THEOREM 6.1. *Let H be a $2m$-element set, $\mathcal{G}_1, \mathcal{G}_2 \subset \binom{H}{m}$, and let δ, η be arbitrary positive constants. Suppose further that k is an integer, $\eta m < k < (1-\eta)m$. Then there exists $\varepsilon = \varepsilon(\delta, \eta) > 0$ such that*

$$|\mathcal{G}_1|\,|\mathcal{G}_2| > \binom{2m}{m}^2 (1-\varepsilon)^m$$

implies

$$i_k(\mathcal{G}_1, \mathcal{G}_2) > i_k\left(\binom{H}{m}, \binom{H}{m}\right)(1-\delta)^m. \tag{18}$$

PROOF. During the proof we will use various constants. It is supposed that $1 < \delta \gg \alpha \gg \sigma \gg \varepsilon$, where $\delta \gg \alpha$ means that δ is incomparably larger than α. Define

$$\mathcal{S}_i = \left\{ S \in \binom{H}{2k-\sigma m} : i_k(S, \mathcal{G}_i) > i_k\left(S, \binom{H}{m}\right)\frac{(1-\varepsilon)^m}{2} \right\}.$$

By the bipartite graph counting principle (Lemma 4.1) we have

$$|\mathcal{S}_i| \geq \binom{2m}{2k-\sigma m}\frac{(1-\varepsilon)^m}{2} \quad \text{for } i = 1, 2.$$

Set $K = i_k\left(S, \binom{H}{m}\right)(1-\varepsilon)^m/2$.

CLAIM. There are at least

$$|\mathcal{S}_1| \Big/ 2\binom{2k}{\sigma m} \geq \binom{2m}{2k-\sigma m}(1-\varepsilon)^m \Big/ 4\binom{2k}{\sigma m}$$

$2k$-sets $A \subset H$ with the property that $k \leq |A \cap G_i| \leq k + \sigma m$ holds for at least K members G_i of $\mathcal{G}_i$ $(i = 1, 2)$.

PROOF OF THE CLAIM. Define $\mathcal{S}_1' \subset \mathcal{S}_1$ to be the collection of those members of $\mathcal{S}_1$ which after adding σm elements in an appropriate way contain some member of $\mathcal{S}_2$. In other words, $\mathcal{S}_1' = \{S_1 \in \mathcal{S}_1 : \exists S_2 \in \mathcal{S}_2,\ |S_2 - S_1| \leq \sigma m\}$. Then for $S_1 \in (\mathcal{S}_1 - \mathcal{S}_1')$, $S_2 \in \mathcal{S}_2$ one has $|S_1 \cap S_2| < |S_1| - \sigma m$ and, in particular, $|S_1 \cap S_2| \neq 2k - 2\sigma m$. Using Corollary 1.6 and $\varepsilon \ll \sigma$, we infer that $|\mathcal{S}_1 - \mathcal{S}_1'|$ is very small with respect to $\mathcal{S}_1$. Consequently, $|\mathcal{S}_1'| > \frac{1}{2}|\mathcal{S}_1|$ holds.

Now associate with each $S \in \mathcal{S}_1'$ a $2k$-set $A = A(S) \subset H$, so that $S \subset A$ and A contains some member of $\mathcal{S}_2$ as well. Since the same A can be associated with at most $\binom{2k}{\sigma m}$ sets S, the number of distinct A's is at least $|\mathcal{S}_1|/2\binom{2k}{\sigma m}$ and by the definition of $\mathcal{S}_i$ it follows that the A's have the desired property, proving the claim.

Let $\mathcal{A}$ be the collection of $2k$-sets defined by the claim. For $A \in \mathcal{A}$ define

$$\mathcal{G}_A^{(i)} = \{G \in \mathcal{G}_i : k \le |A \cap G| \le k + \sigma m\}, \qquad |\mathcal{G}_A^{(i)}| = x_A^{(i)}$$

and

$$y_A = \{(G_1, G_2) \cdot |G_1 \cap G_2| = k,\ |G_1 \cap G_2 \cap A| = k - \alpha m,$$
$$G_i \in \mathcal{G}_i,\ k \le |G_i \cap A| \le k + \sigma m,\ i = 1, 2\},$$

where α is a constant, $\sigma \ll \alpha \ll \delta$. By the definitions we have

$$\sum_{A \in \mathcal{A}} y_A = i_k(G_1, G_2) \binom{k}{k - \alpha m}$$
$$\cdot \sum_{0 \le i, j \le \alpha m} \binom{m-k}{i + \alpha m} \binom{m-k}{j + \alpha m} \binom{k}{k - i - j - \alpha m}.$$

Using $\alpha \ll \delta$, we obtain

$$\sum_{A \in \mathcal{A}} y_A \le i_k(\mathcal{G}_1, \mathcal{G}_2) \left(1 + \frac{\delta}{2}\right)^m.$$

Assuming to the contrary that $i_k(\mathcal{G}_1, \mathcal{G}_2) \le \binom{2m}{m}\binom{m}{k}^2 (1-\delta)^m$, and using the identity $\binom{2m}{m}\binom{m}{k}^2 = \binom{2m}{2k}\binom{2k}{k}\binom{2m-2k}{m-k}$ we infer

$$\sum_{A \in \mathcal{A}} y_A < \binom{2m}{2k}\binom{2k}{k}\binom{2m-2k}{m-k}\left(1 - \frac{\delta}{2}\right)^m.$$

By the claim,

$$|\mathcal{A}| \ge \binom{2m}{2k - \sigma m}(1-\varepsilon)^m \Big/ 4\binom{2k}{\sigma m} > \binom{2m}{2k} \Big/ \left(1 + \frac{\delta}{6}\right)^m.$$

Thus there exists $A_0 \in \mathcal{A}$ satisfying

$$y_{A_0} < \binom{2k}{k}\binom{2m-2k}{m-k}\left(1 - \frac{\delta}{3}\right)^m.$$

By definition,

$$x_{A_0}^{(i)} \ge K = \binom{2k - \sigma m}{k}\binom{2m - 2k + \sigma m}{m - k}\frac{(1-\varepsilon)^m}{2}.$$

Consequently, $x_{A_0}^{(i)} > 2y_{A_0}$.

Let $\mathcal{D}^{(i)}$ consist of those members of $\mathcal{G}_{A_0}^{(i)}$ which do not contribute to y_{A_0}. Then $|\mathcal{D}^{(i)}| > \frac{1}{2}|\mathcal{G}_{A_0}^{(i)}|$ and there are no $G_i \in \mathcal{D}^{(i)}$ with

$$|G_1 \cap G_2| = k, \qquad |G_1 \cap G_2 \cap A_0| = k - \alpha m.$$

Recall that

$$|\mathcal{D}^{(i)}| > \frac{1}{2}|\mathcal{G}^{(i)}_{A_0}| \geq \frac{K}{2} = \binom{2k-\sigma m}{k}\binom{2m-2k+\sigma m}{m-k}\frac{(1-\varepsilon)^m}{4}.$$

Define $\mathcal{D}_i^* = \{B \subset A_0 : |\{G \in \mathcal{D}^{(i)}, G \cap A_0 = B\}| > K/2^{2k+2}\}$. Since for $G_i \in \mathcal{D}^{(i)}$ there are fewer than $2^{|A_0|} = 2^{2k}$ choices for $G_i \cap A_0$ (which satisfy $k \leq |G_i \cap A_0| \leq k + \sigma m$), it follows by an easy averaging argument for $i = 1, 2$, that

$$|\mathcal{D}_i^*| > \frac{K}{2^{2m+2}} = \binom{2k-\sigma m}{k}(1-\varepsilon)^m \binom{2m-2k+\sigma m}{m-k}\Big/ 2^{2m+3}.$$

Since ε and σ are incomparably smaller than α, Theorem 1.4 applies and provides us with $B_i \in \mathcal{B}_i^*$ such that $|B_1 \cap B_2| = k - \alpha m$.

Applying the same theorem to the two families $\mathcal{F}_i = \{G - A_0 : G \in \mathcal{D}^{(i)}, G \cap A_0 = B_i\}$, we obtain $F_i \in \mathcal{F}_i$ such that $|F_1 \cap F_2| = \alpha m$. That is, $(F_i \cup B_i) \in \mathcal{D}^{(i)}$, $|(F_1 \cup B_1) \cap (F_2 \cup B_2)| = k$, $|(F_1 \cup B_1) \cap (F_2 \cup B_2) \cap A_0| = k - \alpha m$, a contradiction, proving the theorem. □

7. The proof of Theorem 1.14: the general case. We successively reduce the general case to the special case proved in the last section.

PROPOSITION 7.1. *It is sufficient to consider the case $p_1 + p_2 \leq 1$.*

Suppose $p_1 \leq p_2$, $p_2 > \frac{1}{2}$. Define $\mathcal{F}_2^C = \{X - F_2 : F_2 \in \mathcal{F}_2\}$ and note that $|F_1 \cap F_2| = l \leftrightarrow |F_1 \cap (X - F_2)| = p_1 n - l$. That is, $i_l(\mathcal{F}_1, \mathcal{F}_2) = i_{p_1 n - l}(\mathcal{F}_1, \mathcal{F}_2^c)$ and also

$$i_l\left(\binom{X}{p_1 n}, \binom{X}{p_2 n}\right) = i_{p_1 n - l}\left(\binom{X}{p_1 n}, \binom{X}{n - p_2 n}\right).$$

To conclude the proof of the proposition we must show:

$$\max\{(p_1 + (1-p_2) - 1 + \eta)n, pn\} < p_1 n - l < \min\{(p_1 - \eta)n, (n - p_2 - \eta)n\}.$$

Since $p_1 \leq p_2$, $p_1 + 1 - p_2 - 1 = p_1 - p_2 \leq 0$, and the first inequality is equivalent to $l < (p_1 - \eta)n$. As to the second inequality, the first part follows from $l > \eta n$. The second is simply $l > (p_1 + p_2 - 1 + \eta)n$. □

We need an auxiliary averaging result.

PROPOSITION 7.2. *Given positive reals ρ, a, b, $a + b < 1$, and two families $\mathcal{A} \subset \binom{X}{an}$, $\mathcal{B} \subset \binom{X}{bn}$, there exists $\varepsilon = \varepsilon(\rho) > 0$ so that whenever*

$$|\mathcal{A}|\,|\mathcal{B}| > (1-\varepsilon)^n \binom{n}{an}\binom{n}{bn},$$

there exists a family

$$\mathcal{C} \subset \binom{X}{(a+b)n}, \qquad |\mathcal{C}| > \binom{n}{(a+b)n}(1-\rho)^n,$$

with the property that every $C \in \mathcal{C}$ contains at least $\frac{1}{2}\binom{(a+b)n}{bn}(1-\varepsilon)^n$ members of $\mathcal{A}$ and at least $\frac{1}{2}\binom{(a+b)n}{bn}(1-\varepsilon)^n$ members of $\mathcal{B}$.

PROOF OF PROPOSITION 7.2. Let ε, σ be positive constants with $\varepsilon \ll \sigma \ll \rho$. Define

$$\mathcal{F}(\mathcal{A}) = \left\{ F \in \binom{X}{(a+b-\sigma)n} : |\{A \in \mathcal{A}, A \subset F\}| \right.$$
$$\left. > \frac{1}{2}\binom{(a+b-\sigma)n}{n}(1-\varepsilon)^n \right\},$$

i.e., $\mathcal{F}(\mathcal{A})$ consists of those $(a+b-\sigma)n$-element subsets of X which contain many members of $\mathcal{A}$; $\mathcal{F}(\mathcal{B})$ is defined analogously. Since $|\mathcal{A}| \geq \binom{n}{an}(1-\varepsilon)^n$, $|\mathcal{B}| \geq \binom{n}{bn}(1-\varepsilon)^n$, the bipartite graph averaging principle yields

$$|\mathcal{F}(\mathcal{A})|, |\mathcal{F}(\mathcal{B})| \geq \frac{1}{2}\binom{n}{(a+b-\sigma)n}(1-\varepsilon)^n.$$

Now define

$$\mathcal{F}^* = \left\{ F \in \mathcal{F}(\mathcal{A}) \text{: there exists } C \in \binom{X}{\sigma n}, G \in \mathcal{F}(\mathcal{B}) \text{ with } G \subset F \cup C \right\}.$$

Note that for all $F \in (\mathcal{F}(\mathcal{A}) - \mathcal{F}^*)$ and $G \in \mathcal{F}(\mathcal{B})$ one must have $|F \cup G| > (a+b)n$, whence $|F \cap G| \neq (a+b-2\sigma)n$.

We can apply Corollary 1.6 to $\mathcal{F}(\mathcal{A}) - \mathcal{F}^*$ and $\mathcal{F}(\mathcal{B})$. Since $\varepsilon \ll \sigma$, we infer

$$|\mathcal{F}(\mathcal{A}) - \mathcal{F}^*| < \frac{1}{4}\binom{n}{(a+b-\sigma)n}(1-\varepsilon)^n$$

and thus,

$$|\mathcal{F}^*| > \frac{1}{4}\binom{n}{(a+b-\sigma)n}(1-\varepsilon)^n.$$

For each $F \in \mathcal{F}^*$ let us fix $C(F) \in \binom{X}{\sigma n}$, such that $C(F) \cap F = \emptyset$ and $F \cup C(F)$ contains some member of $\mathcal{F}(\mathcal{B})$. Define $\mathcal{C} = \{F \cup C(F) \colon F \in \mathcal{F}^*\}$. Then

$$|\mathcal{C}| \geq |\mathcal{F}^*| \binom{(a+b)n}{\sigma n}^{-1} > \frac{1}{4}\binom{n}{(a+b-\sigma)n}(1-\varepsilon)^n \binom{(a+b)n}{\sigma n}^{-1}$$
$$> \binom{n}{(a+b)n}(1-\rho)^n$$

because $\varepsilon \ll \sigma \ll \rho$. $\square$

PROPOSITION 7.3. *It is sufficient to prove Theorem* 1.14 *in the case* $p_1 + p_2 = 1$.

PROOF. Assuming this special case true with $\delta' = \delta/2$ and using Proposition 7.2 with $\rho \leq \delta/2$, we find that each $C \in \mathcal{C}$ contains at least

$$\binom{(p_1+p_2)n}{p_1 n}\binom{p_1 n}{l}\binom{p_2 n}{p_2 n - l}\left(1 - \frac{\delta}{2}\right)^n$$

pairs (A, B) with $A \in \mathcal{F}_1$, $B \in \mathcal{F}_2$, and $|A \cap B| = l$. Since there are $|\mathcal{C}|$ choices for $C \in \mathcal{C}$ and a pair (A, B) is counted at most

$$\binom{n - |A \cup B|}{(p_1 + p_2)n - |A \cup B|} = \binom{(1 - p_1 - p_2)n + l}{l}$$

times we infer

$$\begin{aligned} i_l(\mathcal{F}_1, \mathcal{F}_2) &\geq \binom{(p_1 + p_2)n}{p_1 n} \binom{p_1 n}{l} \binom{p_2 n}{l} \left(1 - \frac{\delta}{2}\right)^n \binom{n}{(p_1 + p_2)n} (1 - \rho)^n \\ &\quad \cdot \binom{(1 - p_1 - p_2)n + l}{l}^{-1} \\ &= i_l\left(\binom{X}{p_1 n}, \binom{X}{p_2 n}\right) \left(1 - \frac{\delta}{2}\right)^n (1 - \rho)^n \\ &> i_l\left(\binom{X}{p_1 n}, \binom{X}{p_2 n}\right) (1 - \delta)^n. \qquad \square \end{aligned}$$

Thus we have reduced Theorem 1.14 to the case $\mathcal{F}_1 \subset \binom{X}{p_1 n}$, $\mathcal{F}_2 \subset \binom{X}{p_2 n}$, $p_1 + p_2 = 1$. Suppose $p_1 < p_2$ and replace $\mathcal{F}_2$ by $\mathcal{F}_2^c = \{X - F : F \in \mathcal{F}_2\}$. As in the proof of Proposition 7.1 this leads to the case

$$\mathcal{G}_1 \subset \binom{X}{pn}, \quad \mathcal{G}_2 \subset \binom{X}{pn}, \qquad p = p_1 = 1 - p_2 \leq \frac{1}{2} \ (\mathcal{G}_1 = \mathcal{F}_1, \mathcal{G}_2 = \mathcal{F}_2^c).$$

Now applying Proposition 7.3 to the families $\mathcal{G}_1$, $\mathcal{G}_2$, we infer that it is sufficient to consider the special case $|X| = 2pn$, that is, the case which we dealt with in the preceding section. This concludes the proof of Theorem 1.14.

8. The proof of Theorems 1.15 and 1.16. We apply induction on $t + s$. Note first that the case $t = 2$, $s = 2$ is covered by Theorem 1.14. Next we deal with the case $t = 2$, $s > 2$.

Let us define $\mathcal{B}^1 = \{B_1 : (B_1, X - B_1) \in \mathcal{B}\}$ and

$$\mathcal{A}^1 = \{A_1 : (A_1, \ldots, A_s) \in \mathcal{A}\}, \qquad \mathcal{A}_{A_1} = \{(A_2, \ldots, A_s) : (A_1, \ldots, A_s) \in \mathcal{A}\}.$$

Define

$$\mathcal{G} = \left\{ A \in \binom{X}{l_1} : |\mathcal{A}_A| \geq \binom{n - l}{l_2, \ldots, l_s} \frac{(1 - \varepsilon)^n}{2} \right\}.$$

Obviously,

$$|\mathcal{A}| = \sum_{A \in \mathcal{A}_1} |\mathcal{A}_A| \geq \binom{n}{l_1, \ldots, l_s} (1 - \varepsilon)^n.$$

Using

$$|\mathcal{A}_A| \leq \binom{n - l_1}{l_2, \ldots, l_s} \quad \text{and} \quad \binom{n}{l_1, \ldots, l_s} = \binom{n}{l_1} \binom{n - l_1}{l_2, \ldots, l_s}$$

the bipartite graph counting principle again yields

$$|\mathcal{G}| \geq \binom{n}{l_1} \frac{(1-\varepsilon)^n}{2}.$$

Let $\delta \ll \gamma$ be a small constant. For $\varepsilon \leq \varepsilon(\delta, \eta)$ one can apply Theorem 1.14 and deduce

$$\begin{aligned} i_{m_{11}}(\mathcal{G}, \mathcal{B}^1) &\geq i_{m_{11}}\left(\binom{X}{l_1}, \binom{X}{k_1}\right)(1-\delta)^n \\ &= \binom{n}{l_1}\binom{l_1}{m_{11}}\binom{n-l_1}{k_1-m_{11}}(1-\delta)^n. \end{aligned}$$

Define

$$\mathcal{G}^* = \left\{G \in \mathcal{G}\colon i_{m_{11}}(G, \mathcal{B}) > i_{m_{11}}\left(G, \binom{X}{k_1}\right)\frac{(1-\delta)^n}{2}\right\}.$$

By Lemma 4.1, $|\mathcal{G}^*| \geq \binom{n}{l_1}(1-\delta)^n/2$ holds.

For $G \in \mathcal{G}^*$ and $C \in \binom{G}{m_{11}}$ define

$$\begin{aligned} &\mathcal{B}(G, C) = \{B - G\colon B \in \mathcal{B}\colon B \cap G = C\}, \\ &\mathcal{C}_{\mathcal{G}} = \left\{C\colon |\mathcal{B}(G, C)| \geq i_{m_{11}}\left(G, \binom{X}{k_1}\right)\frac{(1-\delta)^n}{4}\binom{l_1}{m_{11}}\right\}. \end{aligned}$$

By the bipartite graph averaging principle we have

$$|\mathcal{C}_{\mathcal{G}}| > \binom{l_1}{m_{11}}\frac{(1-\delta)^n}{4}.$$

Now apply the induction hypothesis with $\gamma' = \gamma/2$ to the families $\mathcal{A}_{\mathcal{G}}$ and $\mathcal{B}(G, C)$, $C \in \mathcal{C}_{\mathcal{G}}$. Let M' be the $(s-1)$ by 1 matrix with general entry m_{i1}, $i = 2, 3, \ldots, s$. This gives at least

$$i_{M'}\left(\binom{X-G}{l_2, \ldots, l_s}, \binom{X-G}{k_1 - m_{11}}\right)\left(1 - \frac{\gamma}{2}\right)^n$$

pairs $(A_2, \ldots, A_s) \in \mathcal{A}_{\mathcal{G}}$, $B' \in \mathcal{B}(G, C)$ with intersection pattern M'. Each such pair gives rise to a pair $(G, A_2, \ldots, A_s) \in \mathcal{A}$, $B = B' \cup C \in \mathcal{B}$ with intersection pattern M. This gives a total of not less than

$$\left(\binom{n}{l_1}\frac{(1-\delta)^n}{2}\right)\left(\binom{l_1}{m_{11}}\frac{(1-\delta)^n}{4}\right) \cdot i_{M'}\left(\binom{X-G}{l_2, \ldots, l_s}\binom{X-G}{k_1 - m_{11}}\right)\left(1 - \frac{\gamma}{2}\right)^n.$$

Since $\delta \ll \gamma$, this expression is greater than

$$i_M\left(\binom{X}{l_1, \ldots, l_s}, \binom{X}{k_1}\right)(1-\gamma)^n$$

which concludes the proof for the case $t = 2$.

Thus we have proved the desired result for the case $t = 2$ or equivalently $s = 2$. Now we turn to the case $t > 2$. The family $\mathcal{G}^*$ is defined as above except that $i_{m_{11}}\left(\binom{X}{l_1}, \binom{X}{k_1}\right)$ is replaced by $i_{M_1}\left(\binom{X}{l_1}, \binom{X}{k_1,\dots,k_t}\right)$, where M_1 is the row vector $(m_{11}, \dots, m_{1t})$.

For $(C_1, C_2, \dots, C_t)$, a partition of G with $|C_i| = m_{1i}$, define

$$\begin{aligned}\mathcal{B}(G, (C_1, \dots, C_t)) \\ &= \{(B_1 - G, \dots, B_t - G) \colon (B_1, \dots, B_t) \in \mathcal{B}, B_j \cap G = C_j, j = 1, \dots, t\},\end{aligned}$$

$$\mathcal{C}_{\mathcal{G}} = \left\{ (C_1, \dots, C_t) \colon |\mathcal{B}(G, (C_1, \dots, C_t))| \right.$$
$$\left. \geq i_{M_1}\left(G, \binom{X}{k_1, \dots, k_t}\right)(1-\delta)^n \Big/ 4 \binom{l_1}{m_{11}, \dots, m_{1t}} \right\}.$$

By the bipartite graph averaging principle we obtain

$$|\mathcal{C}_{\mathcal{G}}| > \binom{l_1}{m_{11}, m_{12}, \dots, m_{1t}} \frac{(1-\delta)^n}{4}. \tag{19}$$

Let M_2 be the matrix M with its first row deleted. We want to apply the induction hypothesis to $\mathcal{B}(G, (C_1, \dots, C_t))$ and $\mathcal{A}_{\mathcal{G}}$. Now the induction hypothesis applied for $\gamma' = \gamma/2$ gives

$$\begin{aligned}&i_{M_2}(\mathcal{A}_{\mathcal{G}}, \mathcal{B}(G, (C_1, \dots, C_t))) \\ &\quad \geq i_{M_2}\left(\binom{X-G}{l_2, \dots, l_s}, \binom{X-G}{k_1 - m_{11}, \dots, k_t - m_{1t}}\right)\left(1 - \frac{\gamma}{2}\right)^n.\end{aligned}$$

Every pair $(A_2, \dots, A_s) \in \mathcal{A}_{\mathcal{G}}$, $(F_1, \dots, F_t) \in \mathcal{B}(G, (C_1, \dots, C_t))$ with intersection pattern M_2 gives rise to a pair $(G, A_2, \dots, A_t) \in \mathcal{A}$, $(F_1 \cup C_1, \dots, F_t \cup C_t) \in \mathcal{B}$ with intersection pattern M. This leads via (19) to

$$\begin{aligned}i_M(\mathcal{A}, \mathcal{B}) &\geq |\mathcal{G}^*| \frac{1}{4} \binom{l_1}{m_{11}, \dots, m_{1t}} (1-\delta)^n \\ &\quad \cdot i_{M_2}\left(\binom{X-G}{l_2, \dots, l_s}, \binom{X-G}{k_1 - m_{11}, \dots, k_t - m_{1t}}\right)\left(1 - \frac{\gamma}{2}\right)^n.\end{aligned}$$

Using the lower bound for $\mathcal{G}^*$ and the fact that δ, ε are incomparably smaller than γ we obtain

$$\begin{aligned}i_M(\mathcal{A}, \mathcal{B}) &\geq \binom{n}{l_1}\binom{l_1}{m_1, \dots, m_t}\left(\prod_{i=2}^{s}\binom{n - \sum_{\nu=1}^{i-1} l_\nu}{l_i}\binom{l_i}{m_{i1}, \dots, m_{it}}\right)(1-\gamma)^n \\ &= i_M\left(\binom{X}{l_1, \dots, l_s}, \binom{X}{k_1, \dots, k_t}\right)(1-\gamma)^n. \qquad \square\end{aligned}$$

PROOF OF THEOREM 1.16. We apply induction on r. The case $r = 2$ is just Theorem 1.15. Suppose $r \geq 3$ and let γ_0 be the value of $\varepsilon(\eta, \gamma)$ for $r - 1$ and let $\varepsilon = \min\{\gamma_0, \varepsilon(s_3 \cdots s_r, \eta, \gamma_0) \text{ for } r = 2\}$. Define the s_1 by s_2 matrix $M^* = (m^*_{ij})$ by

$$m^*_{ij} = \sum_{i_3 \cdots i_r} m_{i,j,i_3,\dots,i_r}, \qquad 1 \leq i \leq s_1,\ 1 \leq j \leq s_2.$$

Note that $m^*_{ij} > s_3 \cdots s_r \eta n$. In view of Theorem 1.15,

$$i_{M^*}(\mathcal{A}, \mathcal{B}) > i_{M^*}\left(\binom{X}{l_1^{(1)}, \dots, l_{s_1}^{(1)}}, \binom{X}{l_1^{(2)}, \dots, l_{s_2}^{(2)}}\right)(1-\gamma_0)^n.$$

Note, however, that if $(A, B) \in i_{M^*}(\mathcal{A}, \mathcal{B})$, then they define uniquely a partition $A \cap B = (A_1 \cap B_1, \dots, A_1 \cap B_{s_2}, A_2 \cap B_1, \dots, A_{s_1} \cap B_{s_2})$ and

$$i_{M^*}\left(\binom{X}{l_1^{(1)}, \dots, l_{s_1}^{(1)}}, \binom{X}{l_1^{(2)}, \dots, l_{s_2}^{(2)}}\right) = \binom{X}{m^*_{11}, m^*_{12}, \dots, m^*_{s_1 s_2}}.$$

Therefore $\mathcal{A}_{12} = \{A \cap B : (A, B) \in I_{M^*}(\mathcal{A}, \mathcal{B})\}$ satisfies

$$|\mathcal{A}_{12}| > (1-\varepsilon_0)^n \binom{X}{m^*_{11}, \dots, m^*_{s_1 s_2}}.$$

Let M' be the $s_1 s_2$ by s_3 by $\cdots$ by s_r array which has general entry $m_{ti_3 \cdots i_r} = m_{i_1 i_2 \cdots i_r}$, where $t = (i_1 - 1)s_1 + i_2$. Now the desired statement follows by induction applied to the $r-1$ families $A_{12}, A_3, \dots, A_r$ with intersection pattern array M'. $\square$

9. The proof of Theorem 1.10. In fact, Theorem 1.10 follows almost trivially from Theorem 1.16. Therefore we only sketch the proof. For a codeword $C = (a_1, \dots, a_n)$ over $Q = \{1, 2, \dots, q\}$, define its *weight* $w(C) = (l_1, l_2, \dots, l_q)$, where $l_i = |\{j : a_j = i\}|$. There are fewer than $\binom{n}{q-1}$ choices for $(l_1, \dots, l_q)$; therefore there exists $(l_1, \dots, l_q)$ so that at least $|\mathcal{C}|/\binom{n}{q-1}$ codewords have weight $(l_1, \dots, l_q)$. It will be sufficient to consider these codewords. They are in one-to-one correspondence with ordered partitions $A(C) = (A_1, \dots, A_q)$, where $A_i = \{j : a_j = i\}$.

Now for C, $C' \in \mathcal{C}$ and $A(C) = (A_1, \dots, A_q)$, $A(C') = (A'_1, \dots, A'_q)$, one has $d(C, C') = n - \sum_{1 \le i \le q} |A_i \cap A'_i|$. Fix an intersection pattern matrix $M = (m_{ij})$ with $\sum_{1 \le i \le q} m_{ii} = n - d$ and (assume) $m_{ij} \ge \delta n / 2q^2$. Applying Theorem 1.15 with $\mathcal{A} = \mathcal{B} = \{A(C) : C \in \mathcal{C}, w(C) = (l_1, \dots, l_s)\}$, $\eta = \delta/2q^2$, and $\gamma < 1$, the statement of Theorem 1.10 follows. $\square$

REMARK. By applying Theorem 1.16 to codes, we can obtain much stronger statements. For example, it enables one to find r codewords having pairwise prescribed distance as long as there exists an intersection pattern array realizing these distances and without very small entries. We mention explicitly only one example. Let, as usual, $H(t, q)$ denote the metric space of all codewords of length t over Q. Let S be an arbitrary metric subspace of $H(t, q)$. For integers m, $b \ge 0$, $mS + b$ denotes the metric space over the same pointset with distance $d'(s, s') = md(s, s') + b$.

THEOREM 9.1. *Suppose η is a positive constant, $S \subset H(t, q)$, and m, $b > \eta n$ with $mt + b < (1-\eta)n$. Then there exists $\varepsilon = \varepsilon(\eta, q) > 0$ so that for all $T \subset H(n, q)$ satisfying $|T| > (q-\varepsilon)^n$, there exists $S' \subset T$ so that S' and $mS+b$ are isometric.* $\square$

10. More intersection theorems for two families. Let us start with the following simple result.

PROPOSITION 10.1. *Suppose that $\mathcal{A}$, $\mathcal{B} \subset 2^X$ satisfy $|A \cap B| = l$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$, where l is a fixed, nonnegative integer. Then $|\mathcal{A}|\,|\mathcal{B}| \le 2^n$ holds with equality if and only if for some partition $X = Y \cup Z$ one has $\mathcal{A} = 2^Y$, $\mathcal{B} = 2^Z$.*

This statement will be deduced from the following theorem.

THEOREM 10.2. *Suppose that $\mathcal{A}$, $\mathcal{B} \subset 2^X$ satisfy $|A \cap B| = i$ (mod 2) for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$, $i = 0$ or 1. Then either* (i) *or* (ii) *holds.*
(i) $i = 0$ *and* $|\mathcal{A}|\,|\mathcal{B}| \le 2^n$,
(ii) $i = 1$ *and* $|\mathcal{A}|\,|\mathcal{B}| \le 2^{n-1}$.

PROOF OF THEOREM 10.2. For $F \subset \{1, 2, \ldots, n\}$ let $\chi(F)$ be the characteristic vector of F, i.e., $\chi(F) = (\varepsilon_1, \ldots, \varepsilon_n)$, where $\varepsilon_i = 1$ if $i \in F$ and $\varepsilon_i = 0$ otherwise. Let us consider first case (i).

Let $V = \langle \chi(A) : A \in \mathcal{A} \rangle$ be the vector space over $GF(2)$ generated by the characteristic vectors of the sets in A. Set $W = \langle \chi(B) : B \in \mathcal{B} \rangle$.

Our assumption implies that $W \le V^{\perp}$, the orthogonal complement of V. Hence $\dim W + \dim V \le n$. This yields $|\mathcal{A}|\,|\mathcal{B}| \le |V|\,|W| \le 2^n$.

If one has equality, then $\mathcal{A} = V$ and $\mathcal{B} = W$. In particular, $\emptyset \in \mathcal{A} \cap \mathcal{B}$.

Consider next case (ii). For $F \subset \{1, 2, \ldots, n\}$ let $\tilde{\chi}(F)$ be the extended characteristic vector of F: $\tilde{\chi}(F)$ has length $n + 1$, it agrees with $\chi(F)$ in the first n positions, and its last entry is 1. Now define $\tilde{V} = \langle \tilde{\chi}(A) : A \in \mathcal{A} \rangle$, $\tilde{W} = \langle \tilde{\chi}(B) : B \in \mathcal{B} \rangle$. Again, $\tilde{V}$ and $\tilde{W}$ are orthogonal subspaces, leading to the inequality $\dim \tilde{V} + \dim \tilde{W} \le n + 1$.

Since for $A \in \mathcal{A}$ the vector $\tilde{\chi}(A)$ has 1 in the last position, $|\mathcal{A}| \le |\tilde{V}|/2$ holds. This leads to $|\mathcal{A}|\,|\mathcal{B}| \le |\tilde{V}|\,|\tilde{W}|/4 \le 2^{n-1}$, as desired. □

PROOF OF PROPOSITION 10.1. The upper bound is explicitly contained in Theorem 10.2. If one has equality, then l must be even, and—as pointed out in the above proof—the emtpy set must be among the members of $\mathcal{A}$. Thus $l = 0$. Consequently, $(\bigcup \mathcal{A}) \cap (\bigcup \mathcal{B}) = \emptyset$, implying the statement. □

One can extend Theorem 10.2 to odd primes, as well.

THEOREM 10.3. *Suppose $\mathcal{A}$, $\mathcal{B} \subset 2^X$, p is a prime, $0 \le i < p$, and for all $A \in \mathcal{A}$, $B \in \mathcal{B}$ one has $|A \cap B| \equiv i \pmod{p}$. Then either* (i) *or* (ii) *holds:*
(i) $i = 0$ *and* $|\mathcal{A}|\,|\mathcal{B}| \le 2^n$,
(ii) $0 < i < p$ *and* $|\mathcal{A}|\,|\mathcal{B}| \le 2^{n-1}$.

For the proof we need the following slight extension of a result of Odlyzko [**O**]. For a field Γ let Γ^n denote the standard n-dimensional vector space over Γ. For γ, $\delta \in \Gamma$, a vector $(x_1, \ldots, x_n) \in \Gamma^n$ is said to be a $(\gamma - \delta)$-vector if $x_i = \gamma$ or $x_i = \delta$ holds for all $i = 1, \ldots, n$.

PROPOSITION 10.4. *Suppose that U is a k-dimensional affine subspace of Γ^n. Then U contains at most 2^k $(\gamma - \delta)$-vectors.*

PROOF. Let $U = U_0 + \nu$, where U_0 is a (vector) subspace. Then U_0 has a basis of the form (IM) where I is the identity matrix of order k and M is some k by $n - k$ matrix. Let $u_1, \ldots, u_k$ be the vectors of this basis. Suppose that $v + \sum \alpha_i u_i$ is a $(\gamma - \delta)$-vector. Let β_i be the ith entry of v. Then $\alpha_i = \gamma - \beta_i$ or $\alpha_i = \delta - \beta_i$ holds. This leaves altogether 2^k possibilities for the choice of $\alpha_1, \ldots, \alpha_k$. □

PROOF OF THEOREM 10.3. Define V and W as in the case of Theorem 10.2, except that now these are vector spaces over $GF(p)$. Again $\dim V + \dim W \le n$. By Proposition 10.4 one has $|\mathcal{A}| \le 2^{\dim V}$, $|\mathcal{B}| \le 2^{\dim W}$, which yields the statement.

To prove (ii) extend the characteristic vectors $\chi(A)$ by 1 in the $(n+1)$th position and $\chi(B)$ by $-i$. Otherwise the argument is the same. □

Let us use now a similar approach to give a short proof for a slightly improved version of a result of Ahlswede, El Gamal, and Pang [**AGP**].

THEOREM 10.5. *Suppose that d is a fixed integer and $\mathcal{A}$, $\mathcal{B} \subset 2^X$ satisfy $|A \triangle B| = d$ for all $A \in \mathcal{A}$, $B \in \mathcal{B}$.*

Then $|\mathcal{A}|\,|\mathcal{B}| \leq 2^n$. Moreover if $n \neq 2d$ then $|\mathcal{A}|\,|\mathcal{B}| \leq 2^{n-1}$ holds.

REMARK. The condition $n \neq 2d$ cannot be removed as it is seen from the following example.

Let X be the disjoint union of d subsets $Y_1, \dots, Y_d$ of size 2 and define:

$$\mathcal{A} = \{A \subseteq X : |A \cap Y_i| = 1,\ i = 1, \dots, d\},$$
$$\mathcal{B} = \{B \subseteq X : |B \cap Y_i| = 0 \text{ or } 2,\ i = 1, \dots, d\}.$$

PROOF THEOREM 10.5. For $D \subset X = \{1, 2, \dots, n\}$, let $P(D)$ be the vector (point) in R^n which has ith coordinate $+1$ or -1 according to whether $i \in D$ or $i \notin D$ holds.

Let $V(\mathcal{A})$ $(V(\mathcal{B}))$ denote the affine span of $\mathcal{A}$ ($\mathcal{B}$, respectively).

The assumption $|A \triangle B| = d$ translates into $d(P(A), P(B)) = 2\sqrt{d}$, i.e., every point $P(A)$ is at the same distance from *all* points $P(B)$ with $B \in \mathcal{B}$.

This implies that $V(A)$ is orthogonal to $V(B)$, and in particular, $\dim V(A) + \dim V(B) \leq n$.

By Proposition 10.4, $|\mathcal{A}| \leq 2^{\dim V(A)}$, $|\mathcal{B}| \leq 2^{\dim V(B)}$, which implies

$$|\mathcal{A}|\,|\mathcal{B}| \leq 2^{\dim V(A) + \dim V(B)} \leq 2^n.$$

This inequality proves the first part of the theorem, and also the second unless $\dim V(A) + \dim V(B) = n$.

Let us note that the zero vector $\vec{0} = (0, \dots, 0)$ is at distance $2\sqrt{n}$ from all $P(D)$ with $D \subseteq X$. Thus the subspace generated by $V(A)$ and $\vec{0}$ is orthogonal to $V(B)$.

Therefore $\vec{0} \notin V(A)$ would lead to the desired inequality $(\dim V(A) + 1) + \dim V(B) \leq n$. One argues similarly in the case $\vec{0} \notin V(B)$.

The only remaining case is when $V(A)$ and $V(B)$ are orthogonal subspaces through the origin, that is, $(P(A), P(B)) = 0$ for $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Since $(P(A), P(B)) = n - 2|A \triangle B|$, $n = 2d$ follows $\square$

For different proofs and extensions of the result of [**AGP**] see [**DP** and **P**].

11. A linear lower bound to a problem of Galvin. Recently Ron Graham told us about the following problem of F. Galvin. What is the minimum number $m(k)$ of subsets in a family $\mathcal{F} \subset \binom{X}{2k}$, $|X| = 4k$, which has the following property:

For every $G \in \binom{X}{2k}$ there exists $F \in \mathcal{F}$ with $|G \cap F| = k$. Taking

$$\mathcal{F} = \{[i, i + k - 1] : i = 1, \dots, 2k\}$$

shows that $m(k) \leq 2k$. On the other hand, $m(k) \geq \binom{4k}{2k} \Big/ \binom{2k}{k}^2$ is clear by a counting argument. Therefore $m(k) > c\sqrt{k}$ for some positive c.

THEOREM 11.1. *Suppose that k is odd. Then $m(k) > ck$ for some positive absolute constant c.*

PROOF. Let $V = \langle \chi(F) : F \in \mathcal{F} \rangle$ be the vector space over $GF(2)$ generated by the characteristic vectors of members of $\mathcal{F}$. The assumption says that for every $G \in \binom{X}{2k}$ there exists $F \in \mathcal{F}$ with $(\chi(G), \chi(F)) = 1 \neq 0$. In particular, $G \notin V^{\perp}$. That is, $V^{\perp}$ contains no vector of weight $2k$.

Since $V^\perp$ is a subspace, it follows that $|A \triangle B| \neq 2k$ whenever $\chi(A), \chi(B) \in V^\perp$.

Now Theorem 1.1 (or Theorems 1.10 and 1.11) imply $|V^\perp| < (2-\varepsilon)^{4k}$ and thus $\dim V^\perp \leq 4k - ck$ with some positive constant c. This leads to

$$|\mathcal{F}| \geq \dim V \geq 4k - \dim V^\perp \geq ck. \qquad \square$$

Let us mention that for $k = p^\alpha$, p odd, one can use a theorem of [**F3**], mentioned in the introduction to get $m(k) \geq 0.8k$. In view of this, it seems likely that $m(k) = 2k$ holds. For k odd this would follow from

CONJECTURE 11.2. Suppose $U < GF(2)^{4k}$, $\dim U = 2k+1$. Then U contains a vector of weight $2k$.

Let us note that this problem is due to Ito [**I**], although it is not stated explicitly.

Very recently, Alon [**A**] proved this conjecture when k is a power of 2.

ADDED IN PROOF. Markent and West showed $m(2) = 3$ and $m(4) \leq 7$. Conjecture 11.2 is proved in [**EFIN**].

REFERENCES

[**A**] N. Alon (personal communication).

[**AGP**] R. Ahlswede, A. El Gamal and K. F. Pang, *A two-family extremal problem in Hamming space*, Discrete Math. **49** (1984), 1–5.

[**AK**] R. Ahlswede and G. O. H. Katona, *Contributions to the geometry of Hamming spaces*, Discrete Math. **17** (1977), 1–22.

[**Bl**] A. Blokhuis, *Few-distance sets*, CWI Tract, vol. 7, Math. Centrum, Amsterdam, 1984.

[**C**] H. Chernoff, *A measure of asymptotic efficiency for tests of hypothesis based on the sum of observations*, Ann. Math. Statist. **23** (1952), 493–509.

[**D**] P. Delsarte, *On the four principal parameters of a code*, J. Inform. Control **23** (1973), 407–438.

[**DP**] P. Delsarte and P. Piret, *An extension of an inequality by Ahlswede, El Gamal and Pang for pairs of binary codes*, Discrete Math. **55** (1985), 313–315.

[**DF**] M. Deza and P. Frankl, *The Erdös-Ko-Rado theorem 22 years later*, SIAM J. Algebraic Discrete Methods **4** (1983), 419–431.

[**EFIN**] H. Enomoto, P. Frankl, N. Ito and K. Nomura, *Codes with given distances*, Graphs Combin. (to appear).

[**EKR**] P. Erdös, C. Ko and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math. Oxford Ser. **12** (1961), 313–320.

[**E1**] P. Erdös, *Problems and results in graph theory and combinatorial analysis*, Proc. Fifth British Comb. Conf. 1975 Aberdeen, Congressus Numerantium, 15-Utilitas Math., Winnipeg, 1976.

[**E2**] ——, *The combinatorial problems which I would most like to see solved*, Combinatorica **1**(1981), 25–42.

[**EGMRSS**] P. Erdös, R. L. Graham, P. Montgomery, B. L. Rothschild, J. H. Spencer and E. G. Straus, *Euclidean Ramsey theorems*, J. Combin. Theory Ser. A **14** (1973), 341–363.

[**ES**] P. Erdös, and E. Szemerédi, *Combinatorial properties of systems of sets*, J. Combin. Theory Ser. A **24** (1978), 303–313.

[**F1**] P. Frankl, *The Erdös-Ko-Rado theorem is true for $n = ckt$*, Proc. Fifth Hungar. Combin. Colloq., North-Holland, Amsterdam, 1978, pp. 365–375.

[**F2**] ——, *Bounding the size of a family knowing the differences*, Acta Math. Acad. Sci. Hungar. **19** (1984).

[**F3**] ——, *Orthogonal vectors and codes with missing distances*, Combinatorica **6** (1986).

[**F4**] ——, *Generalizations of theorems of Katona and Milner*, Acta Math. Hungar. **27** (1976), 353–363.

[**FF1**] P. Frankl and Z. Füredi, *On hypergraphs without two edges intersecting in a given number of vertices*, J. Combin. Theory Ser. A **36** (1984), 230–236.

[**FF2**] ——, *Forbidding just one intersection*, J. Combin. Theory Ser. A **39** (1985), 160–176.

[**FR**] P. Frankl and V. Rödl, *All triangles are Ramsey*, Trans. Amer. Math. Soc. **292** (1986), 277–279.

[**FW**] P. Frankl and R. M. Wilson, *Intersection theorems with geometric consequences*, Combinatorica **1** (1981), 357–368.

[**H**] L. H. Harper, *Optimal numberings and isoperimetric problems on graphs*, J. Combin. Theory **1** (1966), 385–393.

[**I**] N. Ito, *Hadamard graphs*. I, Graphs Combin. **1** (1985), 57–64.

[**K**] G. O. H. Katona, *Intersection theorems for systems of finite sets*, Acta Math. Acad. Sci. Hungar. **15** (1964), 329–337.

[**LR**] D. G. Larman and C. A. Rogers, *The realization of distances within sets in Euclidean space*, Mathematika **19** (1972), 1–24.

[**O**] A. M. Odlyzko, *On the ranks of some* $(0,1)$*-matrices with constant row sums*, J. Austral. Math. Soc. Ser. A **31** (1981), 193–201.

[**P**] C. L. M. van Pul, *Constant distance code pairs*, Nederl. Akad. Wetensch. Proc. (to appear).

[**Ry**] H. J. Ryser, *An extension of a theorem of de Bruijn and Erdös on combinatorial designs*, J. Algebra **10** (1968), 246–261.

[**R**] V. Rödl, *On a packing and covering problem*, European J. Combin. **6** (1985), 69–78.

[**W**] R. M. Wilson, *The exact bound in the Erdös-Ko-Rado theorem*, Combinatorica **4** (1984), 247–260.

AT & T BELL LABORATORIES, MURRAY HILL, NEW JERSEY 07974

Permanent address (P. FRANKL): CNRS, 15 QUAI ANATOLE FRANCE, PARIS, 75007, FRANCE

Permanent address (V. RÖDL): FJFI CVUT, HUSOVA 5, PRAGUE, CZECHOSLOVAKIA