COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA 28 (5) (2008) 503–528
DOI: 10.1007/s00493-008-2318-9

ON THE MEASURE OF INTERSECTING FAMILIES, UNIQUENESS AND STABILITY

EHUD FRIEDGUT*

Received June 15, 2006

Let $t \geq 1$ be an integer and let $\mathcal{A}$ be a family of subsets of $\{1,2,\ldots,n\}$ every two of which intersect in at least t elements. Identifying the sets with their characteristic vectors in $\{0,1\}^n$ we study the maximal measure of such a family under a non uniform product measure. We prove, for a certain range of parameters, that the t-intersecting families of maximal measure are the families of all sets containing t fixed elements, and that the extremal examples are not only unique, but also stable: any t-intersecting family that is close to attaining the maximal measure must in fact be close in structure to a genuine maximum family. This is stated precisely in Theorem 1.6.

We deduce some similar results for the more classical case of Erdős–Ko–Rado type theorems where all the sets in the family are restricted to be of a fixed size. See Corollary 1.7.

The main technique that we apply is spectral analysis of intersection matrices that encode the relevant combinatorial information concerning intersecting families. An interesting twist is that part of the linear algebra involved is done over certain polynomial rings and not in the traditional setting over the reals.

A crucial tool that we use is a recent result of Kindler and Safra [22] concerning Boolean functions whose Fourier transforms are concentrated on small sets.

1. Introduction

1.1. History and philosophy

In this paper we touch upon some fundamental questions in extremal set theory. Our focus is on families of subsets of $\{1,\ldots,n\}$ where we identify

Mathematics Subject Classification (2000): 05D05, 05C69

* Research supported in part by the Israel Science Foundation, grant no. 0329745.

them with points in the space $\{0,1\}^n$ endowed with a product measure. Analytical methods used in this setting allow us to study the measure of such families and the distance between their characteristic functions. This in turn allows us to deduce that the optimal families under certain combinatorial conditions are *stable*: a family whose size is close to that of the optimal one must also be similar to it in structure. The spectral methods we use are by no means new in this setting, and similar approaches appear e.g. in [24], [5], [25] (and many others). Indeed, in [26] Wilson proves a result that is intimately related to ours, and a close comparison of our proof with his reveals many similarities. Another closely related work, whose method we generalize, is [21] which we also use as a didactic building block.

One of the differences between the current work and those cited above is that we have at our disposal some relatively new results from the field of analysis of Boolean functions. Stability results such as [18] and its generalization [22] that we use show that Boolean functions with Fourier transform concentrated on low frequencies essentially depend on few variables, which translates back precisely to the type of combinatorial conclusions that we are interested in. To this end we focus on the product measure setting rather than on the more traditional problems concerning families of sets with a fixed size. This theme appears in several recent papers and manuscripts [2], [6], [8]. It seems that these analytical questions which appear in the study of Boolean functions are quite natural. Besides being rich in mathematical structure they turn out to be extremely useful in theoretical computer science, see, e.g., [9], [23], to mention just two examples from a multitude.

Another novel element in our proof is doing the linear algebra not only over the reals but also over certain polynomial rings. This approach arose from the need to find a *t-nilpotent element*, an element X such that $X \neq 0$ but $X^t = 0$, which naturally leads one away from fields and into the realm of rings. This method seems like a useful bookkeeping device that hopefully can be applied to other related problems. Nonetheless it is in place to remark that the specific ring related calculations can be circumvented, as pointed out to me by Ryan O'Donnell [private communication].

1.2. Notation

We begin with some notation: Let $[n] = \{1, 2, \ldots, n\}$. We will call a set $A \subseteq [n]$ with $|A| = k$ a k-set. The family of all k-subsets of $[n]$ will be denoted by $\binom{[n]}{k}$. A family of sets $\mathcal{A} \subseteq [n]$ will be called *intersecting* if the intersection of every two members of $\mathcal{A}$ is non-empty, and *t-intersecting* if every such intersection is of size at least t. We will often identify subsets of $[n]$ with elements of

$\{0,1\}^n$ by associating a set with its characteristic vector. A t-intersecting family of vectors in $\{0,1\}^n$ is therefore a set of vectors every two of which share at least t coordinates on which they equal 1. If $A \subset [n]$ then the family of all k-subsets of $[n]$ containing A is called a *principal family* defined by A. A function $f:\{0,1\}^n \to \{0,1\}$ is called *a dictatorship* if there exists $i \in [n]$ such that $f(x) = x_i$. A function $f:\{0,1\}^n \to \{0,1\}$ is called a *t-umvirate* if there exists a set $A \subset [n]$ with $|A| = t$ and $f(x) = \prod_{i \in A} x_i$. We will also call a family $\mathcal{A} \subset \{0,1\}^n$ a dictatorship or t-umvirate if its characteristic function is such. For every $p \in [0,1]$ let $q = 1-p$, and denote by μ_p the product measure on $\{0,1\}^n$ defined by $\mu_p(x) = p^{\Sigma x_i} q^{\Sigma(1-x_i)}$.

1.3. Intersecting families

The typical setting of a fundamental question in extremal set theory deals with a family of k-subsets of $[n]$ given some combinatorial information on their intersections. The most classical example is the Erdős–Ko–Rado theorem [10]:

Theorem 1.1 (EKR). *Let $k \le n/2$ and let $\mathcal{A} \subset \binom{[n]}{k}$ be an intersecting family. Then $|\mathcal{A}| \le \binom{n-1}{k-1}$. Furthermore, if $k < n/2$ then equality is attained if and only if $\mathcal{A}$ is a principal family defined by some $\{i\}$.*

It has been recognized (see, e.g., [16]) that it is quite natural to consider weighted versions of such theorems, where rather than concentrating on subsets of $\{0,1,\dots,n\}$ of a given size one considers general families of subsets of $[n]$ with a given intersection pattern, and studies their measure under the product measure μ_p on $\{0,1\}^n$. For example the measure-theoretic analogue of the EKR theorem is the following theorem (first proven perhaps in [11]).

Theorem 1.2. *Let $0 \le p \le 1/2$ and let $\mathcal{A} \subset \{0,1\}^n$ be an intersecting family. Then $\mu_p(\mathcal{A}) \le p$. Furthermore, if $0 < p < 1/2$ then equality is attained if and only if $\mathcal{A}$ is a dictatorship.*

Theorem 1.2 is quite easy to prove, and there exist quite a few (non-isomorphic) proofs of it, e.g. [16], [17], [7]. However, in this paper we approach such product-measure theorems via spectral analysis. One of the nice features of the analytical approach is that it provides adequate tools to study L^2-approximations of functions, yielding uniqueness and stability results concerning extremal families. The following is, then, a strengthening of Theorem 1.2.

Theorem 1.3. *Let $0<p<1/2$ and let $\mathcal{A}\subset\{0,1\}^n$ be an intersecting family. Then*

1. $\mu_p(\mathcal{A})\le p$.
2. **Uniqueness:** *If $\mu_p(\mathcal{A})=p$ then $\mathcal{A}$ is a dictatorship.*
3. **Stability:** *If $\mu_p(\mathcal{A})=p-\varepsilon$ then there exists a dictatorship $\mathcal{B}$ such that $\mu_p(\mathcal{A}\triangle\mathcal{B})=c\varepsilon$ where $c=c(p)$.*

Remark. For any small constant ε the function $c(p)$ given by our proof is bounded for $p\in[\varepsilon,1/2-\varepsilon]$, but it explodes as p approaches 0 or 1/2. This behavior in the vicinity of $p=1/2$ is inevitable since the theorem is false for $p=1/2$, as shown, e.g., by $\mathcal{A}=\{A:|A|>n/2\}$. However the theorem can be extended, by different techniques, to a meaningful statement for small values of p. This is done in a paper in preparation [6].

One of the proofs of Theorem 1.2 (or rather, of its generalization to 2-intersecting families) appears in a paper of Dinur and Safra, [9], where they observe that it follows from the classical EKR theorem using an asymptotic approach and some Chernoff concentration results. In this paper we wish to make deductions in the opposite direction. Although it is not clear whether Theorem 1.2 directly implies EKR we are able to use Theorem 1.3 and deduce the following stability result in the $\binom{[n]}{k}$ setting.

Corollary 1.4. *Let $0<\zeta$, let $\zeta n<k<(\frac{1}{2}-\zeta)n$ and let $\mathcal{A}\subset\binom{[n]}{k}$ be an intersecting family. If $|\mathcal{A}|\ge(1-\varepsilon)\binom{n-1}{k-1}$ then there exists a principal family $\mathcal{B}\subset\binom{[n]}{k}$ defined by some $\{i\}$ such that $|\mathcal{A}\setminus\mathcal{B}|<c\varepsilon\binom{n}{k}$ where $c=c(\zeta)$.*

As one may guess $c(\zeta)$ behaves nicely as long as ζ is bounded away from 0. For $k=o(n)$ [6] offers an even stronger statement – it turns out that for $\frac{k}{n}=o(1)$ *every* intersecting family is close to being contained in a dictatorship.

The stability corollaries that we prove in this paper, set in the usual combinatorial setting of $\binom{[n]}{k}$, seem to be quite hard to prove using elementary combinatorial reasoning (meaning that the author is not aware of any such proofs). It is interesting to compare Corollary 1.4 to the Hilton–Milner Theorem, [20], that bounds the size of an intersecting family that is not a principal family. Neither of the results implies the other despite the fact that they basically deal with similar situations. Likewise there are several results of Frankl and Füredi (also related to Corollary 1.7) concerning maximal intersecting families when one restricts the number of sets any given element can belong to, see e.g. [13], [19], [15].

1.4. t-intersecting families for $t>1$

The first natural question generalizing the EKR theorem appears already in the original EKR paper, and it is that of bounding the maximal size of a t-intersecting family of k-sets for $t>1$. Erdős, Ko and Rado proved that for every pair of positive integers $1 \le t \le k$ there exists n_0 such that for $n>n_0$ if $\mathcal{A} \subset \binom{[n]}{k}$ is a t-intersecting family then $|\mathcal{A}| \le \binom{n-t}{k-t}$, with equality if and only if $\mathcal{A}$ is a principal family defined by some t-set. In this paper we prefer changing the order of the parameters and asking, for a fixed n and t, how the size and structure of the extremal t-intersecting families of k-sets depend on k. This extensively studied question has been satisfactorily answered for families of k-sets as we shall see shortly. Our main goal in the current paper is to prove the corresponding product measure analogues including uniqueness and stability, and deduce stability corollaries for the k-set setting. Our success will be, however, limited to a certain initial segment of values of k.

Let us now add some new notation: for fixed integers n,k,t,r let

$$M(n,k,t) = \max\left\{|\mathcal{A}| : \mathcal{A} \subset \binom{[n]}{k},\ \mathcal{A} \text{ is } t\text{-intersecting}\right\}$$

and let

$$I(n,k,t,r) = \left\{A \in \binom{[n]}{k},\ |A \cap [t+2r]| \ge t+r\right\}.$$

Obviously $I(n,k,t,r)$ is a t-intersecting family, and a natural candidate to be the maximal one. Indeed, Frankl ([12]) conjectured in 1978 that for every n and k

$$M(n,k,t) = \max_r\{|I(n,k,t,r)|\}. \tag{1}$$

It turned out, historically, that the easiest cases to settle were those when the extremum is attained for $r=0$, i.e. by a principal family defined by some t-set. A relatively simple calculation shows that $|I(n,k,t,r)|$ is maximal for $r=0$ when $k-t+1<\frac{n}{t+1}$. In [12] Frankl proved that for $t\ge 15$ and for this range of k the maximum is indeed attained by principal families, and this was later extended to all values of t by Wilson in [26]. Frankl and Füredi made further progress towards settling the general question in [14]. Finally, in 1997 Ahlswede and Khachatrian presented their "complete intersection theorem", proving Frankl's conjecture (1). Dinur and Safra deduced in [9] the product measure analogue of the Ahlswede–Khachatrian theorem.

Theorem 1.5 **([9]).** *Let $0<p<1/2$, let $1\le t$ and let $\mathcal{A} \subseteq \{0,1\}^n$ be a t-intersecting family. Then $\mu_p(\mathcal{A}) \le \max_r \mu_p\left(\left\{x : \sum_{i=1}^{t+2r} x_i \ge t+r\right\}\right)$.*

We wish to extend this theorem in the same manner that Theorem 1.3 extended Theorem 1.2. For any fixed value of t we succeed in our goal for an initial segment of values of $p \in [0,1]$ which corresponds to the range of values of k for which Frankl and Wilson initially succeeded in proving (1), precisely the values of p for which a t-umvirate is the extremal example. In Section 4 we speculate as to generalizations to all $p < 1/2$.

Theorem 1.6. *Let $t \geq 1$ be an integer, let $0 < p < \frac{1}{t+1}$ and let $\mathcal{A} \subset \{0,1\}^n$ be a t-intersecting family. Then*

1. $\mu_p(\mathcal{A}) \leq p^t$.
2. **Uniqueness:** *If $\mu_p(\mathcal{A}) = p^t$ then $\mathcal{A}$ is a t-umvirate.*
3. **Stability:** *If $\mu_p(\mathcal{A}) = p^t - \varepsilon$ then there exists a t-umvirate $\mathcal{B}$ such that $\mu_p(\mathcal{A} \triangle \mathcal{B}) = c\varepsilon$ where $c = c(p)$.*

From this we deduce the following.

Corollary 1.7. *Let $t \geq 1$ be an integer, let $0 < \zeta$, $\zeta n < k < (\frac{1}{t+1} - \zeta)n$ and let $\mathcal{A} \subset \binom{[n]}{k}$ be a t-intersecting family. If $|\mathcal{A}| \geq (1-\varepsilon)\binom{n-t}{k-t}$ then there exists a principal family $\mathcal{B} \subset \binom{[n]}{k}$ defined by some B with $|B| = t$ such that $|\mathcal{A} \setminus \mathcal{B}| < c\varepsilon\binom{n}{k}$ where $c = c(\zeta)$.*

Note that both Theorem 1.6 and Corollary 1.7 deal with $t \geq 1$ so they cover the case $t = 1$, i.e. Theorem 1.3 and Corollary 1.4. We make these separate statements because, for didactic reasons, we will prove them separately.

1.4.1. Structure of the paper. The rest of the paper is organized as follows. Section 2 is the heart of the paper, and is broken into the following subsections: in 2.1 we present a proof of an old result of Hoffman. The reason we begin with this result is that it contains some of the key ingredients of our proof and, we believe, will help to provide a soft introduction to the spectral approach. In 2.2 we introduce the necessary changes in Hoffman's proof in order to deduce item 1 of Theorem 1.3. In 2.3 we upgrade our technique to deal with the case $t > 1$ and prove item 1 of Theorem 1.6. In 2.4 we prove the uniqueness and stability of the t-umvirates as stated in items 2 and 3 of Theorem 1.6. In Section 3 we deduce the stability statement for k-uniform t-intersecting families, Corollary 1.7. Finally in Section 4 we state some conjectures and suggestions as to extending the results of this paper.

2. Optimality, uniqueness and stability of t-umvirates

In this section we prove Theorems 1.3 and 1.6. Our idea is to generalize Hoffman's proof [21] from 1970 to the weighted case. We begin by recalling Hoffman's approach, as we believe that it sheds some light on our proof.

2.1. Hoffman's bound on independent sets

Let $G=(V,E)$ be an r-regular graph, and let μ denote the uniform measure on $V(G)$. This induces the usual inner product between functions on $V(G)$, and allows us to define $|f|_2=\sqrt{\langle f,f\rangle}$ for any $f:V(G)\to\mathbb{R}$. Denote by $\bar{\alpha}(G)$ the measure of the largest independent set in G. Let A be the adjacency matrix of G i.e.

$$A_{i,j} = \begin{cases} 1 & \{i,j\} \in E(G) \\ 0 & \{i,j\} \notin E(G) \end{cases}$$

Let $r=\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_{min}$ be the eigenvalues of A. Then:

Theorem 2.1 (Hoffman, [21]).

$$\bar{\alpha}(G) \leq \frac{-\lambda_{\min}}{\lambda_1 - \lambda_{min}}.$$

Remark. We refrain from substituting $\lambda_1 = r$ into the statement of the theorem in order to stress its similarity with later results in this paper.

Proof. Let $\alpha=\bar{\alpha}(G)$, let $I\subset V(G)$ be a maximal independent set and let $f:V(G)\to\{0,1\}$ be the characteristic function of I. Then $|f|_2^2=\mathrm{E}[f]=\alpha$. Since A is a symmetric real matrix it is diagonalizable by an orthonormal basis. Let $U_1,U_2,\ldots,U_m$ be a complete orthonormal set of eigenvectors of A with eigenvalues $\lambda_1,\ldots,\lambda_m$. We normalize the vectors U_i to be orthonormal as functions on $V(G)$ with the inner product defined by μ, (and not as vectors in $\mathbb{R}^m$). Furthermore, since G is r-regular, we can take U_1 to be the vector $(1,1,\ldots,1)$ and $\lambda_1=r$. Let $f=\sum a_iU_i$ be the expansion of f according to this basis. Note that

$$a_1 = |f|_1 = \alpha \tag{2}$$

and from Parseval's formula we have

$$|f|_2^2 = \sum a_i^2 = \alpha. \tag{3}$$

Next note that the fact that I is independent translates to the fact that

$$fAf^{tr} = 0. \tag{4}$$

This is because whenever $f_i = f_j = 1$ then $A_{i,j} = 0$. Expanding (4) using the expansion of f by eigenvectors we have

$$\sum \lambda_i a_i^2 = 0. \tag{5}$$

Putting (2), (3) and (5) together we get

$$0 = \sum \lambda_i a_i^2 = \lambda_1 \alpha^2 + \sum_{i \neq 1} \lambda_i a_i^2 \geq \lambda_1 \alpha^2 + (\alpha - \alpha^2)\lambda_{\min}. \tag{6}$$

Noticing that $\lambda_{\min} < 0$ (because $\sum \lambda_i = Tr(A) = 0$) we rearrange and get the desired bound on α. ∎

2.2. Adapting Hoffman's proof to intersecting families

We now proceed to manipulate Hoffman's proof to yield item 1 of Theorem 1.3. We defer the proof of items 2 and 3 to Section 2.4.

We remind the reader that we identify the subsets of $[n]$ with their characteristic vectors. In what follows we consider the elements of $\{0,1\}^n$ by reverse lexicographical order which enables us to represent functions $f : \{0,1\}^n \to \mathbb{R}$ by vectors of length 2^n indexed by the subsets of $[n]$.

To begin with we note that an intersecting family in $\{0,1\}^n$ is an independent set in the appropriate graph: define a graph whose vertices are the elements of $\{0,1\}^n$ and where two vertices are connected by an edge if and only if the corresponding sets are disjoint. Let $D = D^{(n)}$ be the adjacency matrix of this graph, the disjointness matrix of $\{0,1\}^n$. It will be essential for us to order the rows and columns of D (and the vertices of the disjointness graph) so as to reveal its product structure. We choose the reverse lexicological order, from $(0,0,\dots,0)$ to $(1,1,\dots,1)$. Trying to adapt Hoffman's proof to this graph presents us with several issues.

- To begin with, one would like to know precisely what the eigenvectors and eigenvalues of D are.
- The graph in question is not regular, and therefore the vector $(1,1,\dots,1)$ is not an eigenvector of D. This distinguished eigenvector played a central role in Hoffman's proof.
- The eigenvectors of the disjointness matrix are orthogonal with respect to the uniform measure on $\{0,1\}^n$, but we are interested in a different product measure.

We will see that the first issue is not hard to resolve due to the fact that $D^{(n)}$ is a tensor power of a 2×2 matrix. Furthermore, as we will discover shortly, there is a joint answer that addresses the other two issues simultaneously.

The key observation is that in Hoffman's proof, when using the relation $fAf^{tr} = 0$, we did not really need the fact that $A_{i,j} = 1$ when $\{i, j\}$ is an edge of G, but only that $A_{i,j} = 0$ when $\{i, j\}$ is a non-edge. The proof would work just as well if we replaced A by a *pseudo-adjacency* matrix of G, one where non-edges correspond to 0 entries with no restriction on the entries corresponding to edges. We therefore can safely replace $D^{(n)}$ by a matrix $A = A^{(n)}$, that we will define shortly, which is a pseudo-disjointness matrix for $\{0,1\}^n$, i.e. $A_{S,T} = 0$ whenever S and T have non-empty intersection. If we choose A wisely we can then take care of the other two concerns.

Let us begin with the case $n = 1$. We are looking for a matrix of the type

$$A^{(1)} = \begin{pmatrix} * & * \\ * & 0 \end{pmatrix}.$$

Solving a set of linear equations produces the matrix

$$A^{(1)} = \begin{pmatrix} \frac{q-p}{q} & \frac{p}{q} \\ 1 & 0 \end{pmatrix}$$

which has the "correct" eigenvectors: the eigenvectors are $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} \sqrt{\frac{p}{q}} \\ -\sqrt{\frac{q}{p}} \end{pmatrix}$ which form an orthonormal basis for the space $\{0,1\}$ with the measure μ_p. The corresponding eigenvalues are 1 and $\frac{-p}{q}$ respectively. Next consider the tensor products of the eigenvectors of $A^{(1)}$. For $1 \le i \le n$ define the function $\chi_{\{i\}} : \{0,1\}^n \to \mathbb{R}$ by

$$\chi_{\{i\}}(x) = \begin{cases} \sqrt{\frac{p}{q}} & \text{if } x_i = 0, \\ -\sqrt{\frac{q}{p}} & \text{if } x_i = 1 \end{cases}$$

and for any $S \subseteq [n]$ let $\chi_S = \prod_{i \in S} \chi_{\{i\}}$. The tensor structure of $\{0,1\}^n$ implies that the functions $\{\chi_S\}_{S \subseteq [n]}$ form an orthonormal basis for the space of functions on $\{0,1\}^n$ with the inner product induced by μ_p, a fact that we will use shortly. This is, in fact, the skew version of the Walsh–Fourier basis which consists of the characters of $\mathbb{Z}_2^n$ that form an orthonormal basis in the case $p = 1/2$. In spirit of the discrete Fourier analysis of the $p = 1/2$ case we will write the expansion of functions according to this basis as $f = \sum \hat{f}(S)\chi_S$.

Now let $A = A^{(n)}$ be the n-fold tensor of $A^{(1)}$ with itself. The following lemma summarizes the relevant information we need about A.

Lemma 2.2. *1. The eigenvectors of A are precisely $\{\chi_S\}_{S \subseteq [n]}$ with corresponding eigenvalues $\lambda_S = \left(\frac{-p}{q}\right)^{|S|}$.*

2. *A is a pseudo-disjointness matrix for $\{0,1\}^n$.*

Proof.

1. This is a direct consequence of the fact that A is the n-fold tensor product of $A^{(1)}$.
2. Proof by induction on n: For $n=1$ the claim holds, the only non empty intersection is of $\{1\}$ with itself, and the corresponding entry is indeed 0. Assume next that the claim holds for $n-1$. Note that $A=A^{(n)}=A^{(n-1)}\otimes A^{(1)}$ (where the factor $A^{(1)}$ actually corresponds to the element n, so we index its rows and columns by $\emptyset$ and $\{n\}$). Now let $S,T\subseteq[n]$. Note that

$$A_{S,T} = A^{(n-1)}_{S\setminus\{n\},T\setminus\{n\}} \cdot A^{(1)}_{S\cap\{n\},T\cap\{n\}}. \tag{7}$$

Assume that $S\cap T \neq \emptyset$. Then either $n\in(S\cap T)$, in which case $A^{(1)}_{S\cap\{n\},T\cap\{n\}}=0$, or S and T intersect already in $[n-1]$ and by induction $A^{(n-1)}_{S\setminus\{n\},T\setminus\{n\}}=0$. ∎

From here it is a simple matter to mimic the proof of the previous section.

Proof of Theorem 1.3 item 1. Let $\mathcal{A}\subset\{0,1\}^n$ be an intersecting family of maximal measure, and let $\alpha=\mu_p(\mathcal{A})$. Let $f:\{0,1\}^n\to\{0,1\}$ be $\mathcal{A}$'s characteristic function. Then $|f|_2^2=\mathrm{E}[f]=\alpha$. Writing $f=\sum\hat{f}(S)\chi_S$ we have as before

$$\hat{f}(\emptyset)=|f|_1=\alpha$$

and

$$|f|_2^2=\sum_{S\subset[n]}\hat{f}(S)^2=\alpha.$$

Since A is a pseudo disjointness matrix we have

$$fAf^{tr}=0$$

which yields

$$\sum\left(\frac{-p}{q}\right)^{|S|}\hat{f}(S)^2=0.$$

Note now that if $0\le p\le 1/2$ then $-1\le\frac{-p}{q}\le 0$ and the minimum of $\left(\frac{-p}{q}\right)^{|S|}$ is attained for $|S|=1$, so we finally have

$$0=\sum_S\left(\frac{-p}{q}\right)^{|S|}\hat{f}(S)^2=\alpha^2+\sum_{S\neq\emptyset}\left(\frac{-p}{q}\right)^{|S|}\hat{f}(S)^2\ge\alpha^2+(\alpha-\alpha^2)\left(\frac{-p}{q}\right).$$

Rearranging we get $\alpha\le p$ as desired. ∎

2.3. t-intersecting families for $t \geq 1$; Using polynomial rings

We now turn to prove item 1 of Theorem 1.6. Wishing to continue the momentum we have gathered in the previous section let us first have a look at $B = B^{(n)}$, the 2-disjointness matrix of $\{0,1\}^n$, defined by

$$B_{S,T} = \begin{cases} 1 & |S \cap T| < 2, \\ 0 & |S \cap T| \geq 2. \end{cases}$$

Consider, e.g. the case $n = 3$ where the columns and rows are indexed by $\emptyset, \{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}$:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Studying this matrix raises the worrying observation that it does not seem to be the tensor of smaller matrices as in the 1-intersecting case. To understand the reason for the difference let us return to Formula 7:

$$A_{S,T} = A^{(n-1)}_{S\setminus\{n\},T\setminus\{n\}} \cdot A^{(1)}_{S\cap\{n\},T\cap\{n\}}.$$

When we are only interested in the question whether T and S have non empty intersection then a simple yes/no answer of the corresponding question for the pairs $\{S\setminus\{n\}, T\setminus\{n\}\}$ and $\{S\cap\{n\}, T\cap\{n\}\}$ suffices to determine the answer for the pair $\{S,T\}$. However in the case where we are interested in the more delicate information, the size of $S \cap T$, we need the matrix $B^{(n)}$ to store this information for future generations (i.e. $B^{(n+1)}$), so it seems that a 0-1 matrix is not subtle enough for this job.

The way to remedy this is to substitute a formal variable X for 0 in the basic intersection matrix $A^{(1)}$. Define

$$M^{(1)} = \begin{pmatrix} 1 & 1 \\ 1 & X \end{pmatrix}$$

and let $M = M^{(n)}$ be the n-fold tensor of $M^{(1)}$ with itself. Now the matrix for $\{0,1\}^3$ looks like this:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & X & 1 & X & 1 & X & 1 & X \\ 1 & 1 & X & X & 1 & 1 & X & X \\ 1 & X & X & X^2 & 1 & X & X & X^2 \\ 1 & 1 & 1 & 1 & X & X & X & X \\ 1 & X & 1 & X & X & X^2 & X & X^2 \\ 1 & 1 & X & X & X & X & X^2 & X^2 \\ 1 & X & X & X^2 & X & X^2 & X^2 & X^3 \end{pmatrix}.$$

Studying the 2×2 blocks in $M^{(3)}$ the tensor structure and the combinatorial meaning emerge, and for general n it is not hard to prove the following.

Claim 2.3. *If $S,T \subseteq [n]$ then*

$$M^{(n)}_{S,T} = X^{|S\cap T|}.$$

The proof of the claim is immediate, using induction and formula (7). Furthermore, the same proof easily generalizes to pseudo-intersection matrices:

Claim 2.4. *Let $D^{(1)}$ be a matrix of the form*

$$D^{(1)} = \begin{pmatrix} * & * \\ * & X \end{pmatrix}$$

with entries from $\mathbb{R}[X]$, and let $D = D^{(n)}$ be the n-fold tensor product of $D^{(1)}$ with itself. Then $D_{T,S}$ is divisible by $X^{|T\cap S|}$.

Still trying to adhere to the proof of the previous section we define the pseudo adjacency matrix $D^{(1)}$ over $\mathbb{R}[X]$ as follows.

$$D^{(1)} = \begin{pmatrix} \frac{q-p}{q} + \frac{p}{q}X & \frac{p}{q} - \frac{p}{q}X \\ 1 - X & X \end{pmatrix}.$$

$D^{(1)}$ was designed to once again have the eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} \sqrt{\frac{p}{q}} \\ -\sqrt{\frac{q}{p}} \end{pmatrix}$ which we denoted as $\chi_\emptyset$ and $\chi_{\{1\}}$. The corresponding eigenvalues are now the ring elements 1 and $\frac{-p}{q}(1 - \frac{X}{p})$. Letting $D = D^{(n)}$ be the n-fold tensor product of $D^{(1)}$ with itself we now have a matrix that encodes the intersection information of $\{0,1\}^n$, with the eigenvectors $\{\chi_S\}_{S\subseteq[n]}$. However, we now face two new issues that we must address before proceeding with the proof. The first is that what we really need is a pseudo t-disjointness matrix for $\{0,1\}^n$, a matrix D where $D_{S,T}=0$ whenever $|S\cap T|\geq t$. The second is the more complicated nature of the new eigenvalues – the eigenvalue corresponding to χ_S is now a ring element, the polynomial $\lambda_S=\left(\frac{-p}{q}\right)^{|S|}\left(1-\frac{X}{p}\right)^{|S|}$. Recalling that the previous bounds we had involved the minimal eigenvalue we see that the current proof will be slightly more delicate.

Here is how we address these issues: first, we will work over the ring $\mathbb{R}[X]/(X^t=0)$. In this ring, because of Claim 2.3, $D=D^{(n)}$ immediately becomes a pseudo t-disjointness matrix. Next, we write the eigenvalue λ_S in the following way:

$$\lambda_S=\left(\frac{-p}{q}\right)^{|S|}\left(1-\frac{X}{p}\right)^{|S|}=\left(\frac{-p}{q}\right)^{|S|}\sum_{m=0}^{t-1}\lambda_S^{(m)}X^m$$

where

$$\lambda_S^{(m)}=\binom{|S|}{m}\left(\frac{-1}{p}\right)^m.$$

Note that the relation $X^t=0$ is the reason that m ranges from 0 only up to $t-1$. We now are in shape to reconstruct the proof from the previous section in our setting.

Proof of Theorem 1.6 item 1. Let $\mathcal{A}\subset\{0,1\}^n$ be a t-intersecting family of maximal measure, let $\alpha=\mu_p(\mathcal{A})$ and let $f=\sum\hat{f}(S)\chi_S$ be $\mathcal{A}$'s characteristic function. As in the previous cases

$$\hat{f}(\emptyset)=|f|_1=\alpha$$

and

$$|f|_2^2=\sum_{S\subset[n]}\hat{f}(S)^2=\alpha.$$

Using

$$fDf^{tr}=0$$

we have

$$\sum_S\lambda_S\hat{f}(S)^2=0$$

which we interpret as an equality between two polynomials. Comparing coefficients on both sides we get t equalities. For every $0\leq m\leq t-1$ we have

$$0=\sum_S\lambda_S^{(m)}\hat{f}(S)^2=\sum_S\left(\frac{-p}{q}\right)^{|S|}\binom{|S|}{m}\left(\frac{-1}{p}\right)^m\hat{f}(S)^2. \tag{8}$$

Now, let $Q(s)$ be any polynomial of degree $t-1$. Due to the fact that the linear space of polynomials in the variable s of degree at most $t-1$ is spanned by $\{\binom{s}{m}\}_{m=0}^{t-1}$ we can take a linear combination of the t equalities that are embodied in (8) and get

$$0 = \sum_S \left(\frac{-p}{q}\right)^{|S|} Q(|S|)\hat{f}(S)^2.$$

Plugging in the value of $\hat{f}(\emptyset)$ we get

$$\begin{aligned} 0 &= \alpha^2 Q(0) + \sum_{S\neq\emptyset} \left(\frac{-p}{q}\right)^{|S|} Q(|S|)\hat{f}(S)^2 \\ &\geq \alpha^2 Q(0) + (\alpha-\alpha^2)\min\left\{\left(\frac{-p}{q}\right)^s Q(s)\right\} \end{aligned} \tag{9}$$

where the minimum is taken over all positive integers s.

Now it is time to use some reverse engineering. We are trying to prove that the extremal function f is a t-umvirate, and accordingly that $\alpha \leq p^t$. Equation (9) will imply that bound if we can find $Q(s)$ such that $Q(0)=1$ and such that the minimal value that $\left(\frac{-p}{q}\right)^s Q(s)$ attains over positive integers s is $\frac{-p^t}{1-p^t}$. Proving the existence of such a polynomial $Q(s)$, even for a given value of t, might be a rather hard exercise in linear programming if we did not have in advance some additional information. Hoping that (9) is tight when f is a t-umvirate and observing that for that specific f the Fourier coefficients $\hat{f}(S)$ are non-zero on sets S with $|S|=s=0,1,\ldots,t$ we would like the minimum of $\left(\frac{-p}{q}\right)^s Q(s)$ to be attained on each of those values of s. Our proof will therefore be complete after proving the following lemma.

Lemma 2.5. *Let $t\geq 1$ be an integer, and let $0<p<\frac{1}{t+1}$. Then there exists a polynomial $Q(s)$ of degree $t-1$ such that if we denote $Q(s)\left(\frac{-p}{q}\right)^s$ by $F(s)$ then*

$$F(0) = 1, \quad F(1) = F(2) = \cdots = F(t) = \frac{-p^t}{1-p^t}$$

and for any integer $s>t$ it holds that $F(s)>\frac{-p^t}{1-p^t}$.

Proof. Let us assume that t is an odd integer. The proof for even t is quite similar. Let $d := \frac{-q}{p}$, $M = \frac{-p^t}{1-p^t}$ and let $Q(s)$ be the unique polynomial of degree $t-1$ such that $Q(i)=Md^i$ for $i=1,\ldots,t$. Lemma 2.5 will follow once we verify the following claim.

Claim 2.6. *1.* $Q(0)=1$.
2. $Q(x)>0$ *for all* $x\geq t$.
3. $Q(t+2)<-M(\frac{q}{p})^{t+2}$.
4. *The equation* $Q(x)=-M(\frac{q}{p})^x$ *has precisely* $t+1$ *roots (counting multiplicity), and all of them are in the interval* $[1,t+2)$.
5. $Q(x)<-M\left(\frac{q}{p}\right)^x$ *for all* $x\geq t+2$.

Item 1 in the claim implies that using Q in (9) is indeed valid. Item 2 will imply that we need only to check the value of $Q(s)$ for odd integers $s>t$, and item 5 implies that for all such values $F(s)>M$. Items 3 and 4 will help to imply item 5.

Proof of Claim 2.6.

1. Using the extrapolation form of $Q(x)$ we get

$$\begin{aligned}\frac{Q(0)}{M} &= \sum_{k=1}^{t}\prod_{j\neq k}\frac{0-j}{k-j}d^k\\ &= \sum_{k=1}^{t}\frac{t!(-1)^t(\frac{-1}{k})}{(k-1)!(t-k)!(-1)^{t-k}}d^k\\ &= \sum_{k=1}^{t}\binom{t}{k}(-d)^k = -\left[(1-d)^t-1\right] = 1-\left(\frac{1}{p}\right)^t = M^{-1}.\end{aligned}$$

2. Following the sign pattern of $Q(s)$ for $s=1,2,\ldots,t$ we see that Q changes signs $t-1$ times, hence all its roots are in the interval $[1,t]$ so $Q(x)$ must be positive for all $x>t$.
3. We want to show that $Q(t+2)/M<(d)^{t+2}$. As in item 1

$$\begin{aligned}\frac{Q(t+2)}{M} &= \sum_{k=1}^{t}\prod_{j\neq k}\frac{t+2-j}{k-j}d^k\\ &= \sum_{k=1}^{t}\frac{(t+1)!\frac{1}{t+2-k}}{(k-1)!(t-k)!(-1)^{t-k}}d^k\\ &= \sum_{k=1}^{t}(-1)^{t-k}\binom{t+1}{k-1}d^k(t+1-k)\\ &= d^{t+2}\sum_{m=2}^{t+1}\binom{t+1}{m}\left(\frac{-1}{d}\right)^m(m-1)\end{aligned}$$

$$= d^{t+2}\left[\frac{t+1}{-d}\left(\left(1-\frac{1}{d}\right)^{t}-1\right)-\left(\left(1-\frac{1}{d}\right)^{t+1}+\frac{t+1}{d}-1\right)\right]$$
$$= d^{t+2}\left[\left(1-\frac{1}{d}\right)^{t}\frac{d+t}{-d}+1\right].$$

So, the inequality holds whenever $\frac{d+t}{-d}<0$, which means that $d<-t$, or equivalently $p<\frac{1}{t+1}$.

4. It is easy to prove by induction on t that if $Q(X)$ is a polynomial of degree t then the equation $Q(x)=c^x$ has at most $t+1$ roots for any positive c. We already know that $x=1,3,5,\dots,t$ are roots. Noting that $Q(x)-\left(-M\left(\frac{q}{p}\right)^x\right)$ is negative at $x=2,4,\dots,t-1$ (because $Q(x)$ is) and also at $x=t+2$ (because of item 3) we conclude that in addition to the root at $x=1$ the equation has two roots (or a double root) in each of the intervals $(2,4),(4,6),\dots,(t-3,t-1)$ and two more roots in $(t-1,t+2)$.
5. By item 4 we know that $Q(x)-\left(-M\left(\frac{q}{p}\right)^x\right)$ has a constant sign on $[t+2,\infty)$, and by item 3 this sign is negative.

This completes the proof of Claim 2.6, of Lemma 2.5 and hence of Theorem 1.6. ∎

2.4. Uniqueness and stability

In this subsection we prove items 2 and 3 of Theorem 1.6, the uniqueness and stability of the extremal intersecting families. Both the proof of the uniqueness and of the stability will follow the same two stages. Let f be the characteristic function of a t-intersecting family $\mathcal{A}\subseteq\{0,1\}^n$. For uniqueness we will show that if $\mathrm{E}[f]=p^t$ then the support of $\hat{f}(S)$ is concentrated on sets S with $|S|\le t$, and then deduce that f is a t-umvirate. For stability we will show that if $\mathrm{E}[f]=p^t-o(1)$ then almost all of the weight of $\sum\hat{f}^2(S)$ is concentrated on sets S with $|S|\le t$, and then deduce that f is close to a t-umvirate.

First, if necessary we can, without loss of generality, enlarge $\mathcal{A}$ to be a monotone increasing family, a fact that we will use later. Recall that we proved in the previous section, in Lemma 2.5 that for any $p<\frac{1}{t+1}$ we can choose a $t-1$ degree polynomial $Q(x)$ and define $F(s)=\left(\frac{-p}{q}\right)^sQ(s)$ so that

$$\min\{F(s)\}=\frac{-p^t}{1-p^t}:=m(p)=m,$$

where the minimum is over the positive integers and is attained for $s = 1,2,\dots,t$ and for no other values of s. Furthermore, since this minimum is negative and $F(s)\to 0$ as s grows we have a "spectral gap", and we can write

$$\min\{F(s) : s > t, s \text{ is an integer}\} = m + \delta \tag{10}$$

for some $\delta > 0$.

Now, let $f = \sum \hat{f}(S)\chi_S$ be the characteristic function of a t-intersecting family $\mathcal{A} \subseteq \{0,1\}^n$ with $\mu_p(\mathcal{A}) = \alpha$. We know that $\hat{f}^2(\emptyset) = \alpha^2$, and that $\sum_{S\neq\emptyset} \hat{f}^2(S) = \alpha - \alpha^2$. Assume that this last sum breaks up as follows for some $\tau \geq 0$:

$$\sum_{0<|S|\leq t} \hat{f}^2(S) = (\alpha - \alpha^2)(1-\tau), \qquad \sum_{t<|S|} \hat{f}^2(S) = (\alpha - \alpha^2)\tau. \tag{11}$$

We now return to Equation (9), substituting $Q(0) = 1$ and utilizing (10) and (11).

$$0 = \alpha^2 + \sum_{S\neq\emptyset} F(|S|)\hat{f}(S)^2 \geq \alpha^2 + (\alpha-\alpha^2)(1-\tau)m + (\alpha-\alpha^2)\tau(m+\delta). \tag{12}$$

As long as τ is small enough that $m + \tau\delta < 0$ this simplifies to

$$\alpha \leq \frac{-[m+\tau\delta]}{1-[m+\tau\delta]}.$$

For $\tau = 0$ this gives $\alpha \leq p^t$ as before. Fixing m and δ our bound $\frac{-[m+\tau\delta]}{1-[m+\tau\delta]}$ is a continuous decreasing function of τ near $\tau = 0$. This implies that if α is close to p^t then τ is close to 0. In other words the closer $\mu_p(\mathcal{A})$ is to p^t, the less of the weight of $\sum \hat{f}^2(S)$ comes from sets S with $|S| > t$. More formally:

Lemma 2.7. *Let $0 < p < \frac{1}{t+1}$. Let $f = \sum \hat{f}(S)\chi_S$ be the characteristic function of a t-intersecting family $\mathcal{A} \subseteq \{0,1\}^n$ with $\mu_p(\mathcal{A}) = \alpha$. Let $\tau \geq 0$, let $\sum_{t<|S|} \hat{f}^2(S) = (\alpha-\alpha^2)\tau$ and let δ be defined as in equation (10). Then*

- *If $\alpha = p^t$ then $\tau = 0$.*
- *Let $\varepsilon \geq 0$. If $\alpha > p^t - \varepsilon$ then $\tau < O(\frac{\varepsilon}{\delta})$.*

Proof. Both claims follow easily from the fact that

$$\alpha \leq \frac{-[m+\tau\delta]}{1-[m+\tau\delta]}$$

and

$$\frac{d\left(\frac{-[m+\tau\delta]}{1-[m+\tau\delta]}\right)}{d\tau} = \frac{-\delta}{(1-[m+\tau\delta])^2} = \theta(-\delta). \qquad \blacksquare$$

Theorem 1.6 is easily implied by Lemma 2.7 in conjunction with Lemma 2.8 below. Recall that we may assume without loss of generality that $\mathcal{A}$ is monotone increasing.

Lemma 2.8. *Let $0<p<1/2$ and let $t\geq 1$ be an integer. Let $f:\{0,1\}^n \to \{0,1\}$ be a monotone increasing function.*

1. *If $\hat{f}(S)=0$ for all $|S|>t$ then either $f\equiv 0$ or $\mathrm{E}[f]\geq p^t$ and if $\mathrm{E}[f]=p^t$ then f is a t-umvirate.*
2. *If $\sum_{|S|>t}\hat{f}^2(S)\leq\varepsilon$ and $\mathrm{E}[f]\leq p^t$ then there exists a t-umvirate g such that $|f-g|_2^2=O(\varepsilon)$.*

Remark. For t=1, our base case, this lemma is precisely the main result of [18]: a Boolean function with almost all its Fourier weight on its linear part must be close to constant or close to a dictatorship. However for $t>1$ the assumption that f is monotone and of expectation p^t is essential, else the lemma is easily seen to be false.

Proof. The Proof of item 1 involves double induction which explains why the statement seems slightly stronger than what we really need, making the induction hypothesis stronger. The proof of item 2 relies on a deep theorem of Kindler and Safra [22].

- **Proof of 1.** We proceed by double induction on t and n. We will begin by proving the claim for $t=1$, and then assuming the claim for $n-1$ and all t we will prove it by induction on t for n. Note that the claim holds trivially for $t>n$ since $\mathrm{E}[f]\geq p^n$ for all non-zero Boolean functions f (when $p\leq 1/2$).
If $t=1$ then f is of the form $a_0+\sum a_i\chi_{\{i\}}$. Since $f=f^2$, and this expansion is unique it follows that there is at most 1 index i for which $a_i\neq 0$, else a non-zero term of the form $a_ia_j\chi_{\{i,j\}}$ would appear in $(a_0+\sum a_i\chi_{\{i\}})^2$. Hence $f(x)$ depends on at most one coordinate, so either $f\equiv 0$ or $f\equiv 1$ or $f(x)=x_i$, i.e., f is a dictatorship.
Next, assume the claim for $n-1$ and all values of t and proceed by induction on t for n. We have already proved the case $t=1$, so assume $t>1$ and that we know the claim for n and $t-1$. For $y\in\{0,1\}^{n-1}$ let $f_0(y)=f(y,0)$ and $f_1(y)=f(y,1)$. The following facts are easy to verify.

1. f_0 and f_1 are monotone.
2. For $S \subseteq [n-1]$

$$\hat{f}_1(S) = \hat{f}(S) - \sqrt{\frac{q}{p}}\hat{f}(S \cup \{n\})$$

and

$$\hat{f}_0(S) = \hat{f}(S) + \sqrt{\frac{p}{q}}\hat{f}(S \cup \{n\}).$$

3. $\hat{f}_0(S) = \hat{f}_1(S) = 0$ for all $|S| > t$.
4. $\mathrm{E}[f_0] \leq \mathrm{E}[f_1]$ with equality if and only if $f(x)$ does not depend on x_n.
5. $\mathrm{E}[f] = q\mathrm{E}[f_0] + p\mathrm{E}[f_1]$.

Now assume that $\mathrm{E}[f] \leq p^t$. Then by item 4 we have $\mathrm{E}[f_0] \leq p^t$. If $\mathrm{E}[f_0] = p^t$ then $f(x)$ does not depend on x_n so it is a function of $n-1$ variables, and by the induction hypothesis on n we are done. If $\mathrm{E}[f_0] < p^t$ then by induction on n we have $f_0 \equiv 0$. In this case, by item 5, we have $\mathrm{E}[f_1] \leq p^{t-1}$. Using the assumption that $f_0 \equiv 0$ we deduce that for $S \subseteq [n-1]$

$$\hat{f}(S) = \frac{-p}{q}\hat{f}(S \cup \{n\})$$

hence the only possible sets $T \subseteq [n]$ of size t for which $\hat{f}(T) \neq 0$ are those for which $n \in T$ and furthermore, using item 2, $\hat{f}_1(S) = 0$ not only for $|S| > t$ but also for $|S| = t$. So by induction on t we deduce that either $f_1 \equiv 0$ which implies $f \equiv 0$ or f_1 is a $(t-1)$-umvirate and f is a t-umvirate.

- **Proof of 2.** Here we rely on Theorem 3 in the yet unpublished paper [22]. We present a slightly simplified version of it.

Theorem 2.9 (Kindler, Safra). *Let $0 < p < 1$ and $t \geq 1$ an integer. Then there exist ε_0, c and K such that the following holds. Let $\varepsilon < \varepsilon_0$ and let $f : \{0,1\}^n \to \{0,1\}$ be such that $\sum_{|S|>t} \hat{f}(S)^2 = \varepsilon$. Then there exists a function $g : \{0,1\}^n \to \{0,1\}$ which depends on at most K coordinates and so that $|f - g|_2^2 < c\varepsilon$.*

This innocuous theorem is the engine that drives our whole stability approach. It belongs to the same genre as the main results in [18] (which deals with the case $t=1$) and [4] which finds the strongest possible trade-off between the rate of decay of $\hat{f}$ and f's dependence on few coordinates. The common theme to these theorems is that if the Fourier transform of a Boolean function has very little weight on the high coefficients there is a simple explanation for that, and it is that the function essentially depends on few coordinates. The extra strength of this particular theorem is that once ε, the weight of $\hat{f}^2(S)$ on sets S with $|S|>t$, is sufficiently small then the number of coordinates defining the approximating function g does not have to grow in order to get an approximation proportional to ε. It will be clear in what follows that this is essential for us in deducing that f not only depends on few coordinates, but is actually close to a t-umvirate.

Clearly we will be done if we show that for the function f in the statement of Lemma 2.8 applying the Kindler–Safra theorem yields a function g which is a t-umvirate.

Let $g^{>t}=\sum_{|S|>t}\hat{g}(S)\chi_S$, and define $f^{>t}$ similarly. Let $\mathcal{J}_K$ be the set of all functions $g:\{0,1\}^n\to\{0,1\}$ depending on at most K coordinates, and write $\mathcal{J}_K = \mathcal{J}_K^+ \cup \mathcal{J}_K^0$ where $\mathcal{J}_K^+$ is the set of all functions g for which $|g^{>t}|_2^2>0$, and $\mathcal{J}_K^0$ the set of g such that $g^{>t}\equiv 0$. Let $\mathcal{J}_K^*\subset\mathcal{J}_K^0$ be the functions which are neither a t-umvirate nor the constant 0 function. Define the following two constants.

$$\gamma = \min_{g\in\mathcal{J}_K^+}\{|g^{>t}|_2^2\}$$

and

$$\delta = \min_{g\in\mathcal{J}_K^*}\{\mathrm{E}\,[g]-p^t\}.$$

Note that, by the definition of $\mathcal{J}_K^+$, $\gamma>0$ and that by item 1 in Lemma 2.8 $\delta>0$. Next assume that $\varepsilon<\varepsilon_0$ is small enough so that $(\sqrt{c}+1)^2\varepsilon<\gamma$ and $c\varepsilon<\delta$. Applying the theorem to our f yields a function g. We claim that $g\notin\mathcal{J}_K^+$. Assume otherwise. Then since $|f^{>t}|_2^2\le\varepsilon$ we have

$$c\varepsilon > |g-f|_2^2 > |g^{>t}-f^{>t}|_2^2 \ge (|g^{>t}|_2 - |f^{>t}|_2)^2 \ge (\sqrt{\gamma}-\sqrt{\varepsilon})^2,$$

in contradiction to our choice of ε. Similarly $g\notin\mathcal{J}_K^*$ because $|f-g|_2^2\ge\delta>c\varepsilon$ for all $g\in\mathcal{J}_K^*$. Therefore g must be a t-umvirate as desired.

This completes the proof of Lemma 2.8 which together with Lemma 2.7 immediately imply Theorem 1.6. ∎∎

3. Stability in the case of *k*-uniform *t*-intersecting families when $\frac{k}{n}<\frac{1}{t+1}$

In this section we prove Corollary 1.7, which includes Corollary 1.4 as a special case. Throughout the proof we will assume that k is sufficiently large, e.g. $\sqrt{k}\gg 1/\varepsilon$, and the $o(1)$ notation will be with respect to k tending to infinity. We are applying here the converse of the nice approach of “going to infinity and back" which appears in [9] and [16]. There one wishes to prove a statement on $\mathcal{A}\subseteq\{0,1\}^n$, embeds it in a higher dimensional cube $\{0,1\}^N$ and uses asymptotic results concerning k-subsets of $[N]$. Here we deduce a result for the k-subsets from asymptotic results concerning $\mathcal{A}\subseteq\{0,1\}^N$.

Proof. Let $\mathcal{A}\subset\binom{[n]}{k}$ be a t-intersecting family with $|\mathcal{A}|>(1-\varepsilon)\binom{n-t}{k-t}$. Our strategy will be to consider the upset generated by $\mathcal{A}$, and to study it under the measure μ_p for some p slightly larger than $\frac{k}{n}$. If we choose p appropriately then the measure of the upset will be close to p^t and it will therefore be close to a t-umvirate $\mathcal{B}$. Then we will project $\mathcal{B}$ back down to the kth level and show that $\mathcal{A}\setminus\mathcal{B}$ is small.

Let $\mathcal{A}^{\uparrow}$ be the upset generated by $\mathcal{A}$,

$$\mathcal{A}^{\uparrow}=\{B:\exists A\in\mathcal{A},A\subseteq B\}$$

and let

$$\mathcal{A}^{\uparrow}_m=\mathcal{A}^{\uparrow}\cap\binom{[n]}{m}.$$

We have

$$|\mathcal{A}|=|\mathcal{A}^{\uparrow}_k|>(1-\varepsilon)\binom{n-t}{k-t}>(1-\varepsilon)\left(\frac{k}{n}\right)^t\binom{n}{k}$$

which implies by a simple counting argument that for any $m>k$

$$|\mathcal{A}^{\uparrow}_m|>(1-\varepsilon)\left(\frac{k}{n}\right)^t\binom{n}{m}. \tag{13}$$

Let $\ell=k+\sqrt{4\log\frac{\varepsilon}{k}}$ and let $p=\frac{\ell}{n}$. Then

$$p=(1+o(1))\frac{k}{n}<\frac{1}{t+1}. \tag{14}$$

We claim that $\mathcal{A}^{\uparrow}$ is close to a t-umvirate under μ_p. Note that $\mathcal{A}^{\uparrow}$ is t-intersecting because $\mathcal{A}$ is. Theorem 1.6 will therefore imply that there exists a t-umvirate $\mathcal{B}$ such that

$$\mu_p(\mathcal{A}^{\uparrow}\setminus\mathcal{B})=O(\varepsilon)$$

provided we prove the following.

Claim 3.1. $\mu_p(\mathcal{A}^{\uparrow})=(1-O(\varepsilon))p^t$.

(Note that from item 1 in Theorem 1.6 we know that $\mu_p(\mathcal{A}^{\uparrow})\le p^t$.)

Proof.

$$\mu_p(\mathcal{A}^{\uparrow}) = \sum_{m\geq k} \mu_p(\mathcal{A}^{\uparrow}_m)$$

which by (13) is

$$\geq (1-\varepsilon)\left(\frac{k}{n}\right)^t \sum_{m\geq k} \mu_p\left(\binom{[n]}{m}\right).$$

By (14) this is

$$\geq \left[(1-\varepsilon)(1-o(1))p^t\right] \sum_{m\geq k} \mu_p\left(\binom{[n]}{m}\right)$$
$$= \left[(1-\varepsilon)(1-o(1))p^t\right] \left(1-\mu_p\left(\bigcup_{m<k}\binom{[n]}{m}\right)\right).$$

The last quantity, $\mu_p\big(\bigcup_{m<k}\binom{[n]}{m}\big)$ is the probability that a binomial random variable $B(n,\frac{\ell}{n})$ will be smaller than k. Our choice of ℓ together with a simple Chernoff bound (see, e.g. [3] page 265) imply that this is bounded by ε. Since p^t is a constant in our setting this proves that $\mu_p(\mathcal{A}^{\uparrow})=(1–O(\varepsilon))p^t$ as desired. ∎

Item 1 in Theorem 1.6 now guarantees the existence of a t-umvirate $\mathcal{B}$ that approximates $\mathcal{A}^{\uparrow}$, i.e.

(15) $$\mu_p\big(\mathcal{A}^{\uparrow}\setminus\mathcal{B}\big) = O(\varepsilon).$$

We claim that $\mathcal{B}\cap\binom{[n]}{k}$ is the principal family we are seeking i.e. that $|\mathcal{A}\setminus\mathcal{B}|=O(\varepsilon)\binom{n}{k}$.

Proof. The t-umvirate $\mathcal{B}$ is defined by t coordinates, say, $1,\dots,t$. The cube $\{0,1\}^n$ is divided into 2^t subcubes of co-dimension t by these coordinates. For any $y\in\{0,1\}^t$ let C_y be the corresponding subcube,

$$C_y = \{x\in\{0,1\}^n,\ (x_1,\dots,x_t) = (y_1,\dots,y_t)\}.$$

We want to show that for $y\neq(1,1,\dots,1)$ the intersection of $\mathcal{A}$ with C_y is not too large, i.e.

$$|\mathcal{A}\cap C_y| = O(\varepsilon)\binom{n}{k}.$$

But this follows immediately from (15): the precise same calculation that proved Claim 3.1 gives that if

$$|\mathcal{A}\cap C_y| > a\binom{n}{k}$$

then

$$\mu_p\big(\mathcal{A}^{\uparrow} \cap C_y\big) > (1 - o(1))a\mu_p(C_y).$$

Since for $y \neq (1,1,\dots,1)$ we have $\mathcal{A}^{\uparrow} \cap C_y \subseteq \mathcal{A}^{\uparrow} \setminus \mathcal{B}$ this would contradict (15) if $a\mu_p(C_y) \gg \varepsilon$. Since we treat p and t as constants $\mu_p(C_y)$ is constant too, and we are done. ∎

4. Open Problems

There are many different intersection theorems and conjectures that seem to be natural candidates for attack via the methods of this paper. However, rather than list them all we will restrict ourselves to the most immediate generalizations of Theorem 1.6 and Corollary 1.7. We would like to study uniqueness and stability statements for t-intersecting families of maximal size when the answer is not a t-umvirate. Recall that the Ahlswede–Khachatrian theorem states that for $k < n/2$ the unique maximum t-intersecting families of k-sets are the families of type

$$I(n,k,t,r) = \left\{ A \in \binom{[n]}{k},\ |A \cap [t+2r]| \geq t + r \right\}.$$

The natural conjecture generalizing Corollary 1.7 is that these examples are also stable. Let $t \geq 1$ be an integer and let $0 < \zeta < 1/2$. Then there is either a unique value of r or two consecutive values that asymptotically maximize the value of $|I(n, \lfloor \zeta n \rfloor, t, r)|$. We will say that ζ is non singular for t if there is only one such value of r and in that case we denote it by $r^*(\zeta,t)$. For singular ζ let r^* and r^*+1 be the two extremal values of r.

Conjecture 4.1. Let $t \geq 1$ be an integer and let $0 < \zeta < 1/2$ be non-singular for t. Let $r^*(\zeta,t)$ be as defined above. Let $k = \lfloor \zeta n \rfloor$ and let $\mathcal{A} \subset \binom{[n]}{k}$ be a t-intersecting family. If

$$|A| \geq (1-\varepsilon)|I(n,k,t,r^*)|$$

then there exists a set $B \subset [n]$ of size $t + 2r^*$ such that

$$|\{A \in \mathcal{A} : |A \cap B| \geq t + r^*\}| \geq (1 - O(\varepsilon))|\mathcal{A}|.$$

If ζ is singular for t then either the above holds or the corresponding statement for $r^* + 1$ holds.

(Here $O(\varepsilon)$ hides a constant that depends on t and ζ but not on n.) In the spirit of Section 3 it is clear that one way to prove this conjecture is by proving the corresponding statement in the product measure setting.

Conjecture 4.2. Let $t \geq 1$ be an integer and let $0 < \zeta < 1/2$ be non-singular for t. Let $r^*(\zeta, t)$ be as defined above and let

$$\mathcal{A}^* = \left\{ x \in \{0,1\}^n : \sum_{i=1}^{r^*+2t} x_i \geq r^* + t \right\}.$$

Let $\mathcal{A} \subset \{0,1\}^n$ be a t-intersecting family. If

$$\mu_\zeta(\mathcal{A}) \geq (1-\varepsilon)\mu_\zeta(\mathcal{A}^*)$$

then there exists a set $B \subset [n]$, $|B| = 2r^* + t$ such that

$$\mu_\zeta(\mathcal{A} \triangle \mathcal{B}) = O(\varepsilon)$$

where

$$\mathcal{B} = \left\{ x \in \{0,1\}^n : \sum_{i \in B} x_i \geq r^* + t \right\}.$$

If ζ is singular for t then either the above holds or the corresponding statement for $r^* + 1$ holds.

The reader may ask herself at this point why the proof of Theorem 1.6 does not easily generalize to prove this conjecture. The reason is that, in a sense, we enjoyed a stroke of luck in the proof of the case where the extremal family is a t-umvirate. Going back to Equation (9) the proof boils down to solving a certain linear programming problem. The solution tells us something about the Fourier coefficients of the optimal function. As long as $p \leq \frac{1}{t+1}$ it turns out that the optimal solution can be realized by a *Boolean* function. However for other values of p the optimal Boolean function is not a global optimum under our constraints. As we mentioned in the introduction this was also the pre-Ahlswede–Khachatrian state of affairs, where the problem was settled first for principal families. We hope to see similar progress on this front too.

Acknowledgements. I would like to thank Guy Kindler, Ryan O'Donnell and Assaf Naor for useful conversations.

References

[1] R. Ahlswede and L. Khachatrian: The complete intersection theorem for systems of finite sets, *European J. Combin.* **18(2)** (1997), 125–136.

[2] N. Alon, I. Dinur, E. Friedgut and B. Sudakov: Graph Products, Fourier Analysis and Spectral Techniques; *G.A.F.A.* **14(5)** (2004), 913–940.

[3] N. Alon and J. Spencer: *The probabilistic method*, Second edition, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience [John Wiley and Sons], New York, 2000.

[4] J. Bourgain: On the distribution of the Fourier spectrum of boolean functions, *Israel J. Math.* **131** (2002), 269–276.

[5] P. Delsarte: The association schemes of coding theory, in: *Combinatorics (Proc. NATO Advanced Study Inst., Breukelen, 1974), Part 1: Theory of designs, finite geometry and coding theory*, pp. 139–157. Math. Centre Tracts, No. **55**, Math. Centrum, Amsterdam, 1974.

[6] I. Dinur and E. Friedgut: Intersecting families are approximately contained in juntas, in preparation.

[7] I. Dinur and E. Friedgut: A proof of an intersection theorem via graph homomorphisms, *Electron. J. Comb.* **13(1)** (2006), #N6.

[8] I. Dinur, E. Friegut and O. Regev: Independent sets in graph powers are almost contained in juntas, to appear in *G.A.F.A.* http://dx.doi.org/10.1007/s00039-008-0651-1

[9] I. Dinur and S. Safra: On the importance of being biased (1.36 hardness of approximating Vertex-Cover), *Annals of Mathematics*, to appear. *Proc. of 34th STOC*, 2002.

[10] P. Erdős, C. Ko and R. Rado: Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford, Ser. 2* **12** (1961), 313–318.

[11] P. C. Fishburn, P. Frankl, D. Freed, J. C. Lagarias and A. M. Odlyzko: Probabilities for intersecting systems and random subsets of finite sets, *SIAM J. Algebraic Discrete Methods* **7(1)** (1986), 73–79.

[12] P. Frankl: The Erdős–Ko–Rado theorem is true for $n = ckt$, in: *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976)*, Vol. **I**, pp. 365–375, Colloq. Math. Soc. János Bolyai, **18**, North-Holland, Amsterdam-New York, 1978.

[13] P. Frankl: On intersecting families of finite sets, *J. Combin. Theory Ser. A* **24(2)** (1978), 146–161.

[14] P. Frankl and Z. Füredi: Beyond the Erdős–Ko–Rado theorem, *J. Combin. Theory Ser. A* **56(2)** (1991), 182–194.

[15] P. Frankl and Z. Füredi: Finite projective spaces and intersecting hypergraphs, *Combinatorica* **6(4)** (1986), 335–354.

[16] P. Frankl and N. Tokushige: Weighted multiply intersecting families, *Studia Sci. Math. Hungar.* **40(3)** (2003), 287–291.

[17] E. Friedgut: A Katona-type proof of an Erdős–Ko–Rado-type theorem, *J. Combin. Theory Ser. A* **111(2)** (2005), 239–244.

[18] E. Friedgut, A. Naor and G. Kalai: Boolean Functions whose Fourier Transform is Concentrated on the First Two Levels, *Adv. in Appl. Math.* **29** (2002), 427–437.

[19] Z. Füredi: Erdős–Ko–Rado type theorems with upper bounds on the maximum degree, in: *Algebraic methods in graph theory*, Vol. **I**, **II** (Szeged, 1978), pp. 177–207, Colloq. Math. Soc. János Bolyai, **25**, North-Holland, Amsterdam-New York, 1981.

[20] A. J. W. Hilton and E. C. Milner: Some intersection theorems for systems of finite sets, *Quart. J. Math. Oxford, Ser. 2* **18** (1967), 369–384.

[21] A. J. Hoffman: On eigenvalues and colorings of graphs, in: *1970 Graph Theory and its Applications (Proc. Advanced Sem., Math. Research Center, Univ. of Wisconsin, Madison, Wis., 1969)*, pp. 79–91, Academic Press, New York (Reviewer: R. C. Read).

[22] G. Kindler and S. Safra: Noise-resistant boolean functions are juntas, submitted.

[23] N. Linial, Y. Mansour and N. Nisan: Constant depth circuits, Fourier transform, and learnability; *J. Assoc. Comput. Mach.* **40(3)** (1993), 607–620.
[24] L. Lovász: On the Shannon capacity of a graph, *IEEE Transactions on Information Theory* **25** (1979), 1–7.
[25] A. Schrijver: Association schemes and the Shannon capacity: Eberlein polynomials and the Erdős–Ko–Rado theorem; in: *Algebraic methods in graph theory*, Vol. **I**, **II** (Szeged, 1978), pp. 671–688, Colloq. Math. Soc. János Bolyai, **25**, North-Holland, Amsterdam-New York, 1981.
[26] R. M. Wilson: The exact bound in the Erdős–Ko–Rado theorem, *Combinatorica* **4(2–3)** (1984), 247–257.

Ehud Friedgut

Institute of Mathematics
Hebrew University
Jerusalem
Israel
ehudf@math.huji.ac.il