Intersecting families are essentially contained in juntas

Irit Dinur * Ehud Friedgut †

January 4, 2007

Abstract

A family $\mathcal{J}$ of subsets of $\{1,\ldots,n\}$ is called a j-junta if there exists $J \subseteq \{1,\ldots,n\}$, with $|J| = j$, such that the membership of a set S in $\mathcal{J}$ depends only on $S \cap J$.

In this paper we provide a simple description of intersecting families of sets. Let n and k be positive integers with $k < n/2$, and let $\mathcal{A}$ be a family of pairwise intersecting subsets of $\{1,\ldots,n\}$, all of size k. We show that such a family is essentially contained in a j-junta $\mathcal{J}$ where j does not depend on n but only on the ratio k/n and on the interpretation of "essentially".

When $k = o(n)$ we prove that every intersecting family of k-sets is almost contained in a dictatorship, a 1-junta (which by the Erdős-Ko-Rado theorem is a maximal intersecting family): for any such intersecting family $\mathcal{A}$ there exists an element $i \in \{1,\ldots,n\}$ such that the number of sets in $\mathcal{A}$ that do not contain i is of order $\binom{n-2}{k-2}$ (which is approximately $\frac{k}{n-k}$ times the size of a maximal intersecting family).

Our methods combine traditional combinatorics with results stemming from the theory of Boolean functions and discrete Fourier analysis.

1 Introduction

1.1 Notation

Let us begin with some notation. Let $[n] = \{1, 2, \ldots n\}$ and $\binom{[n]}{k} = \{T \in [n] : |T| = k\}$. We shall refer to the elements of $\binom{[n]}{k}$ simply as k-sets. If J is a ground set we denote the family of all subsets of J by 2^J. General families of subsets of $[n]$ are denoted by script letters $\mathcal{A}, \mathcal{B}, \mathcal{J}$, etc. and subsets of $[n]$ by capital letters S, T, etc. Two sets will be called intersecting if their intersection is nonempty and a family of sets will be called intersecting if its elements are pairwise intersecting. Two families $\mathcal{A}$ and $\mathcal{B}$ will be called cross intersecting if every member of $\mathcal{A}$ intersects every member of $\mathcal{B}$. The following definition is important enough in this paper to deserve separate indentation.

Definition 1.1 *A family of sets $\mathcal{J}$ will be called a j-junta if there exists a subset $J \subseteq [n]$ with $|J| = j$ such that membership of a set in $\mathcal{A}$ is determined only by its intersection with J, in other words there exists a family $\mathcal{J}^* \subseteq 2^J$ such that $\mathcal{J} = \{T \subseteq [n] : (T \cap J) \in \mathcal{J}^*\}$. In this case we will say that $\mathcal{J}$ is the junta generated by $\mathcal{J}^*$.*

AMS 2000 subject classification: 05d05

Key words and phrases: Independent sets, intersecting families, Product graphs.

*School of Computer Science and Engineering, Hebrew University, Jerusalem, Israel. Email: dinuri at cs.huji.ac.il.
Research supported in part by the Binational Science Foundation and by the Israel Science Foundation.

†Institute of Mathematics, Hebrew University, Jerusalem, Israel. Email: ehudf at math.huji.ac.il.
Research supported in part by the Israel Science Foundation, grant no. 0329745.

1.2 Results

Questions regarding a family of sets given some simple information regarding their intersection pattern are among the most fundamental in combinatorics. Perhaps the first and most basic of these questions is answered by the Erdős-Ko-Rado theorem [4] which deals with an *intersecting family* of sets. If $\mathcal{A} \subset \binom{[n]}{k}$ is intersecting, with $k \leq n/2$, the EKR theorem states that $\mathcal{A}$ is of size no more than $\binom{n-1}{k-1}$ and that if k is strictly less than $n/2$ this bound is attained only by a family of all k-subsets containing some fixed element. Many other theorems in extremal set theory provide such elegant characterizations of the largest possible family of sets given certain information on their intersections. A notable prototype of such a theorem is the Ahlswede-Khachatrian theorem [1] which deals with a family of k-sets given that the pairwise intersections of its elements are of size at least r for some fixed integer r. There, once again, the extremal families are j-juntas where j can be bounded by some function of r and k/n.

The principal principle that we prove in this paper shows that such results should come as no surprise since (roughly speaking) *every* intersecting family of sets is contained in a family with such a simple description, namely every intersecting family is essentially contained in a junta.

To make this more concrete, we want to state a theorem that says that for every intersecting family $\mathcal{A} \subset \binom{[n]}{k}$ and for every $\varepsilon > 0$ there exists an integer j which depends only on ε and on k/n, a set $J \subset [n]$ with $|J| = j$ and a non-trivial family $\mathcal{J}^* \subset 2^J$ so that if $\mathcal{J}$ is the junta generated by $\mathcal{J}^*$ then $|\mathcal{A} \setminus \mathcal{J}| \leq \varepsilon\binom{n}{k}$. However, when $k = pn$ for constant p, this holds rather trivially by taking a very large junta: if we take $j \gg \log_{k/n}(\varepsilon)$, take an arbitrary $J \subset [n]$ of size j, and define $\mathcal{J}^* = 2^J \setminus \emptyset$ and $\mathcal{J}$ the generated junta this already implies $|\mathcal{A} \setminus \mathcal{J}| < \varepsilon\binom{n}{k}$, for simple statistical reasons that have nothing to do with the structure of $\mathcal{A}$.

Such a trivial hindrance to providing an accurate description of intersecting families does not exist for $k = o(n)$ and at this point we break our treatment of intersecting families contained in $\binom{[n]}{k}$ into two cases: $k = pn$ for some constant $0 < p < 1/2$ and $k = o(n)$.

The case $k = pn$, p constant, $0 < p < \frac{1}{2}$. As illustrated above, finding a junta that contains our intersecting family up to an ε fraction is hardly an accomplishment. It would be far more gratifying if we could find a j-junta $\mathcal{J}$ which is itself intersecting or, in other words, if we could guarantee that $\mathcal{J}^*$ is an intersecting family. This would, in a sense, provide a simple explanation to the fact that $\mathcal{A}$ is intersecting. We conjecture that this in fact is true, see Conjecture 1.3 below. But for now we make do with proving that $\mathcal{J}^*$ is almost-intersecting. The definition of being *almost-intersecting* is subtle and we defer it to Section 3.2. Our main theorem in this section is as follows,

Theorem 1.2 *There exists a function $j(p, \varepsilon)$ such that the following holds. Let $0 < p < 1/2$ and let $\varepsilon > 0$. Let $j = j(\varepsilon, p)$ and let k and n be positive integers with $j \ll k = pn$. Then for every intersecting family $\mathcal{A} \subset \binom{[n]}{k}$ there exists a j-junta $\mathcal{J}$ generated by a family $\mathcal{J}^*$ such that*

1. *$|\mathcal{A} \setminus \mathcal{J}| \leq \varepsilon\binom{n}{k}$.*
2. *$\mathcal{J}^*$ is ε-almost-intersecting in $\mu_p(\{0,1\}^j)$.*

In other words there is a simple (almost) necessary condition for a set to belong to $\mathcal{A}$, where the condition depends only on j elements. Note that finding such a simple sufficient condition is hopeless since one can take arbitrarily complicated subfamilies of $\mathcal{A}$ (e.g. a random subfamily) which, of course, will still be intersecting families. As stated above we hope that Theorem 1.2 can be improved:

Conjecture 1.3 *The family $\mathcal{J}^*$ guaranteed by Theorem 1.2 can be taken to be an intersecting family.*

Since the Erdős-Ko-Rado theorem breaks down when $k > n/2$ it comes as no surprise that the function $j(\varepsilon, p)$ guaranteed by Theorem 1.2 explodes as p approaches $1/2$. However when p approaches 0 the situation is simpler.

The case $k = o(n)$. As $\frac{k}{n}$ goes to 0, i.e. $k = o(n)$, we can state a theorem that is perhaps more satisfying and surprising. It turns out that when k is small compared to n then *every* intersecting family of k-sets is practically contained in a maximal intersecting family, a family of all sets containing one fixed element. To better appreciate the results that follow let us first calibrate the quantities that appear in the statements. If $k = o(n)$ then, denoting $p = k/n$,

$$\binom{n}{k} \gg \binom{n-1}{k-1} \gg \binom{n-2}{k-2} \gg \ldots \gg \binom{n-r}{k-r},$$

as $\frac{\binom{n-r}{k-r}}{\binom{n}{k}} \approx p^r$. The following theorems hold for any natural numbers $k < n/2$ but are only interesting when $k = o(n)$ otherwise the hidden constants can be set to make the theorems hold trivially. As usual we denote $p = k/n$.

Theorem 1.4 *There exists a constant $c(1)$ such that if $\mathcal{A} \subset \binom{[n]}{k}$ is an intersecting family then there exists an element $i \in [n]$ such that all but $c(1) \cdot \binom{n-2}{k-2}$ of the sets in $\mathcal{A}$ contain i.*

In other words every intersecting family is practically contained in a maximal intersecting family.
Remarks:

- Note that up to the value of $c(1)$ the theorem is sharp as shown by the example of all k-sets that contain at least two elements out of a given set of three elements.
- Theorem 1.4 is a special case of Theorem 1.5 below, but we state it separately because we find it to be aesthetically pleasing that the junta involved in this case is a dictatorship, a 1-junta.

The following theorem offers a higher degree of precision at the inevitable price of needing a larger junta to capture the intersecting family.

Theorem 1.5 *There exists functions $j(r), c(r)$ such that for any integers $1 < j(r) < k < n/2$, if $\mathcal{A} \subset \binom{[n]}{k}$ is an intersecting family with $|\mathcal{A}| > c(r) \cdot \binom{n-(r+1)}{k-(r+1)}$ then there exists an intersecting j-junta $\mathcal{J}$ with $j \leq j(r)$ and*

$$|\mathcal{A} \setminus \mathcal{J}| \leq c(r) \cdot \binom{n-(r+1)}{k-(r+1)} = O\left(p^{r+1} \cdot \binom{n}{k}\right).$$

Note that unlike the $k = pn$ case we are able to ensure that the junta $\mathcal{J}$ is in fact intersecting.

The above two theorems also have cross-intersecting-families versions. Although the simple versions of the theorems are slightly more concise we state the cross intersecting versions not only because they are generalizations but because we actually need them for our induction based proofs.

Theorem 1.6 *There exists a constant $c(1)$ such that if $k + \ell < n$ and $\mathcal{A} \subset \binom{[n]}{k}$ with $|\mathcal{A}| > c(1) \cdot \binom{n-2}{k-2}$ and $\mathcal{B} \subset \binom{[n]}{\ell}$ with $|\mathcal{B}| > c(1) \cdot \binom{n-2}{\ell-2}$ are cross intersecting families then there exists an element $i \in [n]$ such that all but $c(1) \cdot \binom{n-2}{k-2}$ of the sets in $\mathcal{A}$ contain i, and similarly, all but $c(1) \cdot \binom{n-2}{\ell-2}$ of the sets in $\mathcal{B}$ contain i.*

Theorem 1.7 *There exist functions $j(r), c(r) \geq 1$ such that the following holds. Let $n, k, \ell \in \mathbb{N}$ with $1 < j(r) < \ell \leq k$ and $k + \ell < n$ and let $\mathcal{A} \subset \binom{[n]}{k}$ and $\mathcal{B} \subset \binom{[n]}{\ell}$ be cross-intersecting families such that $|\mathcal{A}| > c(r) \cdot \binom{n-(r+1)}{k-(r+1)}$ and $|\mathcal{B}| > c(r) \cdot \binom{n-(r+1)}{\ell-(r+1)}$. Then there exist cross-intersecting families $\mathcal{J}$ and $\mathcal{I}$ such that*

1. *$\mathcal{J}$ and $\mathcal{I}$ are both j-juntas with $j \leq j(r)$.*

2. *$|\mathcal{A} \setminus \mathcal{J}| \leq c(r) \cdot \binom{n-(r+1)}{k-(r+1)}$ and $|\mathcal{B} \setminus \mathcal{I}| \leq c(r) \cdot \binom{n-(r+1)}{\ell-(r+1)}$.*

Remark: Clearly the juntas $\mathcal{J}$ and $\mathcal{I}$ cannot be trivial, i.e. neither of them is equal to $\emptyset$ or $2^{[n]}$. They cannot be equal to $\emptyset$ since $\mathcal{A}$ and $\mathcal{B}$ are too large to be approximated by the empty family, and neither of them can be $2^{[n]}$ since $2^{[n]}$ cross intersects only with the empty family. The other side of this coin is that without assuming the lower bound on the size of ,say, $\mathcal{A}$ we can say nothing interesting about the structure of $\mathcal{B}$ because in this case we may approximately capture $\mathcal{A}$ by the empty family and $\mathcal{B}$ by $2^{[n]}$.

1.3 Methods

Although the theorems stated in this paper deal with families of sets restricted to a fixed size the starting point of our proofs is always embedding them in the discrete cube $\{0,1\}^n$ (identified with $2^{[n]}$) endowed with a product measure, and considering the filter generated by them. The measure we choose is in some cases concentrated (by Chernoff inequalities) near the layer of the cube where the original family of sets resides.

We will rely on some results that are proven via discrete harmonic analysis on $\{0,1\}^n$. In the case $k = pn$, Theorem 1.2, our result will be a relatively simple corollary of the main result in [2]. That, in turn, relies crucially on the relatively new invariance principle of Mossel, O'Donnell and Oleszkiewicz, [11]. This principle enables one to deduce, under certain conditions, when two real functions defined on a product space depend jointly in a non-negligible manner on a fixed coordinate. This, of course, is the name of the game in the current paper: proving that a family of sets is essentially defined by some few fixed elements.

In the case $k = o(n)$ we need a simple condition that guarantees that a family $\mathcal{A} \subset 2^{[n]}$ may be approximated by a junta. This is supplied by a result from [6]. We then utilize the technique of [3] to show that, using an appropriate measure, the condition may be fulfilled, hence there exists an approximating junta $\mathcal{J}$. Returning to the original measure that we are interested in we can no longer deduce that $\mathcal{J}$ is close to $\mathcal{A}$ but rather that it approximately contains $\mathcal{A}$, as desired.

1.4 Outline of the paper

In Section 2 we gather our main combinatorial and analytical tools stating and proving our main lemmas. In Section 3 we then prove our main theorems: Theorems 1.6 and 1.7 in subsection 3.1 and Theorem 1.2 in subsection 3.2.

2 Preliminaries

We proceed with additional definitions. We will identify any $S \subseteq [n]$ with its characteristic vector $s \in \{0,1\}^n$ defined by $s_i = 1$ if and only if $i \in S$. Hence, for any $p \in [0,1]$ we will consider the product measure on the discrete cube to induce a measure on $2^{[n]}$ defined by $\mu_p(S) = p^{|S|}(1-p)^{n-|S|}$ and, naturally, $\mu_p(\mathcal{A}) = \sum_{A \in \mathcal{A}} \mu_p(A)$.

If $J \subset [n]$ and $\mathcal{A} \subseteq 2^{[n]}$ then for every $S \subseteq J$ we will be interested in the family $\mathcal{A}_S \subset 2^{[n]\setminus J}$ defined as follows.

$$\mathcal{A}_S = \mathcal{A}_S(J) = \{T \in [n] \setminus J : (T \cup S) \in \mathcal{A}\}.$$

A family $\mathcal{A} \subseteq 2^{[n]}$ is monotone if $A \in \mathcal{A}$ and $A' \supseteq A$ implies $\mathcal{A}' \in \mathcal{A}$. For any$\mathcal{A} \subseteq 2^{[n]}$ we define its monotone closure $\mathcal{A}^{\uparrow}$ to be the family

$$\mathcal{A}^{\uparrow} = \bigcup_{A \in \mathcal{A}} \left\{ A' \in 2^{[n]} \ : \ A' \supseteq A \right\}.$$

It is useful to note that if $\mathcal{A}$ is intersecting then so is $\mathcal{A}^{\uparrow}$.

The upper shadow of a set $A \in \binom{[n]}{k}$ is the family of all $k+1$-element sets A' that contain A. The upper shadow of a family of sets $\mathcal{A} \subseteq \binom{[n]}{k}$ is the union of all upper shadows of its elements:

$$\partial \mathcal{A} = \bigcup_{A \in \mathcal{A}} \left\{ A' \in \binom{[n]}{k+1} \ : \ A' \supset A \right\}.$$

The famous Kruskal-Katona theorem, [9], [8], says that among all families $\mathcal{A} \subset \binom{[n]}{k}$ with $|\mathcal{A}| = m$ the size of the upper shadow of $\mathcal{A}$ is minimized when $\mathcal{A}$ consists of the first m sets in $\binom{[n]}{k}$ according to the lexicographical order. The following immediate corollary of the Kruskal-Katona theorem will be useful to us. Define $\partial^i \mathcal{A}$ inductively by $\partial^1 \mathcal{A} = \partial \mathcal{A}$ and $\partial^i \mathcal{A} = \partial(\partial^{i-1}\mathcal{A})$ for $i > 1$.

Corollary 2.1 *Let $\mathcal{A} \subseteq \binom{[n]}{k}$, and let $0 < r < k$ and $i \leq n-k$. If $|\mathcal{A}| \geq \binom{n-r}{k-r}$ then $\left|\partial^i \mathcal{A}\right| \geq \binom{n-r}{(k+i)-r}$*

Another observation that follows from a simple counting argument will also be needed:

Proposition 2.2 *For any $\mathcal{A} \subseteq \binom{[n]}{k}$ and for all $1 \leq i \leq n-k$*

$$\frac{|\mathcal{A}|}{\binom{n}{k}} \leq \frac{|\partial^i \mathcal{A}|}{\binom{n}{k+i}}.$$

Corollary 2.1 provides an easy derivation of the following generalization of the Erdős-Ko-Rado theorem. For this and more on cross-intersecting families see e.g. [7] and [5] where this result is proven. We include the proof for sake of self containedness.

Theorem 2.3 *Let n, k, ℓ be integers such that $k + \ell < n$. Let $\mathcal{A} \subseteq \binom{[n]}{k}$ and $\mathcal{B} \subseteq \binom{[n]}{\ell}$ be cross intersecting. Then either $|\mathcal{A}| \leq \binom{n-1}{k-1}$ or $|\mathcal{B}| \leq \binom{n-1}{\ell-1}$.*

Proof: Assume that $|\mathcal{A}| > \binom{n-1}{k-1}$ and $|\mathcal{B}| > \binom{n-1}{\ell-1}$. For every $i > 0$, $\partial^i \mathcal{A}$ and $\mathcal{B}$ are cross intersecting. In particular, $\mathcal{B}$ cross-intersects with $\mathcal{A}' = \partial^{n-\ell-k}\mathcal{A}$ which consists of $n-\ell$-element subsets. By corollary 2.1, $|\mathcal{A}| > \binom{n-1}{k-1}$ implies $|\mathcal{A}'| \geq \binom{n-1}{n-\ell-1}$.

Let $\mathcal{B}^c = \{[n] \setminus S \ : \ S \in \mathcal{B}\} \subseteq \binom{[n]}{n-\ell}$, then clearly $|\mathcal{B}^c| = |\mathcal{B}| > \binom{n-1}{\ell-1}$. We conclude that $|\mathcal{B}^c| + |\mathcal{A}'| > \binom{n-1}{\ell-1} + \binom{n-1}{n-\ell-1} = \binom{n}{n-\ell}$ so there must be some $T \in \mathcal{B}^c \cap \mathcal{A}'$ but this contradicts the fact that $\mathcal{A}'$ and $\mathcal{B}$ are cross-intersecting. ■

We now study some claims regarding the measure of monotone families.

The following easy claim is slightly reminiscent of Theorem 2.3.

Claim 2.4 *Let $\mathcal{A} \subseteq 2^{[n]}$ and $\mathcal{B} \subseteq 2^{[n]}$ be monotone cross intersecting families, then for any $p \leq 1/2$, $\mu_p(\mathcal{A}) + \mu_p(\mathcal{B}) \leq 1$.*

Proof: The monotonicity of $\mathcal{A}$ and $\mathcal{B}$ imply that $\mu_p(\mathcal{A})$ and $\mu_p(\mathcal{B})$ are non-decreasing functions of p therefore $\mu_p(\mathcal{A}) + \mu_p(\mathcal{B}) \le \mu_{\frac{1}{2}}(\mathcal{A}) + \mu_{\frac{1}{2}}(\mathcal{B})$. However, $\mu_{\frac{1}{2}}(\mathcal{A}) + \mu_{\frac{1}{2}}(\mathcal{B}) \le 1$ since $\mu_{\frac{1}{2}}$ is simply the uniform measure over subsets, and if $A \in \mathcal{A}$ then $([n] \setminus A) \notin \mathcal{B}$. ■

The following claim relates the size of $\mathcal{A}$ to the measure of $\mathcal{A}^{\uparrow}$.

Claim 2.5 *Let $n > k > r$, $\mathcal{A} \subseteq \binom{[n]}{k}$, and $\mathcal{A}^{\uparrow}$ be the monotone closure of $\mathcal{A}$. Let $p \in (0,1)$ and $\varepsilon > 0$ be such that $p(1-\varepsilon) > \frac{k-r}{n-r}$. If $|\mathcal{A}| \ge \binom{n-r}{k-r}$ then $\mu_p(\mathcal{A}^{\uparrow}) > p^r[1 - \exp(-\varepsilon^2 pn/2)]$.*

Proof: This follows from corollary 2.1, and a tail bound.

$$\begin{aligned}\mu_p(\mathcal{A}^{\uparrow}) = \sum_{i=0}^{n-k} p^{k+i}(1-p)^{n-k-i}\left|\partial^i \mathcal{A}\right| \quad &\ge \quad \sum_{i=0}^{n-k} p^{k+i}(1-p)^{n-k-i}\binom{n-r}{k+i-r} \\ &= \quad p^r \cdot \mu_p\left\{S \in 2^{[n-r]} \ : \ |S| \ge k-r\right\}\end{aligned}$$

A standard Chernoff bound implies that $\Pr_{S\sim\mu_p}[|S| < (1-\varepsilon)pn] < \exp(-\varepsilon^2 pn/2)$, which gives the lemma. ■

We will also need another claim relating the size of a family of sets on the kth level of $\{0,1\}^n$ with its upper closure.

Lemma 2.6 *Let j, k and n be integers with $1 \le j \ll k \le n/2$. Let $\mathcal{A} \subset \binom{[n]}{k}$ and let $\mathcal{J}$ be a j-junta. Set $p = k/n$. Then*

$$|\mathcal{A} \setminus \mathcal{J}| \le (2 + o(1))\binom{n}{k}\mu_p(\mathcal{A}^{\uparrow} \setminus \mathcal{J}).$$

Proof: Let $\mathcal{J}^* \subseteq 2^J$ be the family that generates $\mathcal{J}$. For $S \subseteq J$ Let

$$\varepsilon_S = \frac{|\mathcal{A}_S|}{\left|\binom{[n]}{k}_S\right|} = \frac{|\mathcal{A}_S|}{\left|\binom{n-j}{k-|S|}\right|}.$$

Then

$$|\mathcal{A} \setminus J| = \sum_{S \in 2^J \setminus \mathcal{J}^*} |\mathcal{A}_S| = \sum_{S \in 2^J \setminus \mathcal{J}^*} \varepsilon_S \binom{n-j}{k-|S|} \tag{1}$$

Note that $\frac{\binom{n-j}{k-|S|}}{\binom{n}{k}}$ is the probability that a set of $k = pn$ elements, chosen randomly without replacement from $[n]$, intersects J precisely at S. By the asymptotic equality of binomial and hypergeometric random variables, this is approximately $p^{|S|}(1-p)^{j-|S|} = \mu_p(\mathcal{N}_S)$ where $\mathcal{N}_S = \{T \subset [n] \ : \ T \cap J = S\}$. Therefore, (1) equals

$$\begin{aligned} &= \quad (1 + o(1))\binom{n}{k} \sum_{S \in 2^J \setminus \mathcal{J}^*} \varepsilon_S \cdot \mu_p\left(\mathcal{N}_S\right) &(2) \\ &\le \quad (2 + o(1))\binom{n}{k} \sum_{S \in 2^J \setminus \mathcal{J}^*} \varepsilon_S \cdot \mu_p\left(\bigcup_{\ell \ge k} \mathcal{N}_S \cap \binom{[n]}{\ell}\right) &(3) \\ &\le \quad (2 + o(1))\binom{n}{k}\mu_p(\mathcal{A}^{\uparrow} \setminus \mathcal{J}) &(4)\end{aligned}$$

where in (3) we used the fact that the probability that the binomial random variable $B(n-j, k/n)$ exceeds k is asymptotically $1/2$, and (4) follows from Proposition 2.2. ■

To conclude this section we need, for the case $k = o(n)$, a lemma allowing us to relate a pair of families with an initial pair of juntas that we will use to construct the required capturing juntas.

In [6] it is shown that a Boolean function where the sum of influences is bounded is close to a junta. This translates via the well known Margulis-Russo Lemma ([12],[10]) to the following.

Lemma 2.7 *There exists a function $j(p, K, \varepsilon) = 2^{O(K/\varepsilon)}$ such that the following holds. Let $p_0 \in [0, 1]$ and let $\mathcal{A}$ be a monotone family of subsets of $[n]$. Assume that*

$$\left.\frac{d\mu_p(\mathcal{A})}{dp}\right|_{p=p_0} \leq K.$$

Then for every $\varepsilon > 0$ there exists a j-junta $\mathcal{J}$, with $j = j(p_0, K, \varepsilon)$ such that

$$\mu_{p_0}(\mathcal{A} \triangle \mathcal{J}) \leq \varepsilon.$$

What follows now is a special case of Lemma 1.2 in [3]. Let us present the essentials of the reasoning there. Consider the quantity $0 \leq \mu_p(\mathcal{A}) \leq 1$. Viewing it as a function of p its derivative cannot be too large throughout the interval $[0, 1]$, hence there must exist values of p where Lemma 2.7 can be applied. This implies the slightly surprising result that for any family $\mathcal{A}$ one can always find values of p for which $\mathcal{A}$ is approximable by a junta according to the corresponding measure. In [3] this is applied simultaneously to several families, but we need only consider the case of two families $\mathcal{A}$ and $\mathcal{B}$.

Lemma 2.8 *Let $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$ be monotone families, and let $(a, a + \delta) \subseteq [0, 1]$. Then for every $\epsilon > 0$ there exists $p \in (a, a + \delta)$ and j-juntas $\mathcal{J}$ and $\mathcal{I}$,with $j = 2^{O(1/\delta\varepsilon)}$, such that*

$$\mu_p(\mathcal{A} \triangle \mathcal{J}) \leq \varepsilon$$

and

$$\mu_p(\mathcal{B} \triangle \mathcal{I}) \leq \varepsilon.$$

Proof: The set of points for which $\frac{d\mu_p(\mathcal{A})}{dp} > 2/\delta$ has measure at most $\delta/2$, and a similar statement holds for $\frac{d\mu_p(\mathcal{B})}{dp}$. The corollary follows. ■

By the way, for readers who are worried by issues of measurability we mention that $\frac{d\mu_p(\mathcal{A})}{dp}$ is in fact a polynomial in p and hence well behaved.

3 Proofs of main theorems

3.1 The case $k = o(n)$

In this section we prove Theorems 1.6 and 1.7 concerning cross-intersecting families, and hence also Theorems 1.4 and 1.5 which are special cases. Given families $\mathcal{A} \subset \binom{[n]}{k}$ and $\mathcal{B} \subset \binom{[n]}{l}$ we will use Lemma 2.8 to find juntas that approximate their monotone closures according to some measure μ_p. The surprising thing is that p, being a constant, is much larger than $\frac{k}{n}$ and $\frac{\ell}{n}$ and hence, supposedly, this approximation is not strongly indicative of the structure of the families on the kth and ℓth levels of the discrete cube. We then resort to use the corollary of the Kruskal-Katona theorem, Corollary 2.1, to project back down to the kth and ℓth levels and prove our assertions.

Throughout what follows we think of n as tending to infinity, and $k, \ell = o(n)$. However we will want k and ℓ to be sufficiently large for our approximations (such as the application of Lemma 2.8 in the proof of Lemma 3.2 below) to hold.

We begin with the following notion.

Definition 3.1 *Let $m, k, n \in \mathbb{N}$ and let $\mathcal{A} \subseteq \binom{[n]}{k}$ be a family of subsets. We say that a subset $J \subset [n]$ m-*captures *$\mathcal{A}$ if the number of subsets in $\mathcal{A}$ that miss J is at most m,*

$$|\mathcal{A}_{\emptyset}(J)| = |\{R \in \mathcal{A} \,:\, R \cap J = \emptyset\}| \leq m.$$

We will think of $|J|$ as fixed as n tends to infinity, and $m \ll |\mathcal{A}|$ so the definition is non-trivial. Indeed a set of size k misses J with probability $\frac{\binom{n-|J|}{k}}{\binom{n}{k}}$ which is ≈ 1 when $k/n = o(1)$. It's easy to calculate that if we fix $|J|$ and chose a random $\mathcal{A} \subseteq \binom{[n]}{k}$ then even if we take $m < \frac{1}{2}|\mathcal{A}|$ with high probability (over the choice of $\mathcal{A}$) there is no subset J that m-captures $\mathcal{A}$.

The following lemma asserts, however, that if $\mathcal{A} \subseteq \binom{[n]}{k}$ and $\mathcal{B} \subseteq \binom{[n]}{\ell}$ are cross intersecting families and $k, \ell = o(n)$, then there do exist subsets J, I of bounded size that capture $\mathcal{A}, \mathcal{B}$ respectively.

Lemma 3.2 *There exists a function $j_0(r) \geq 1$ such that for every $r \geq 1$ the following holds. Let $\mathcal{A} \subseteq \binom{[n]}{k}, \mathcal{B} \subseteq \binom{[n]}{\ell}$ be cross intersecting families such that $|\mathcal{A}| > \binom{n-(r+1)}{k-(r+1)}$ and $|\mathcal{B}| > \binom{n-(r+1)}{\ell-(r+1)}$. There are sets $J, I \subset [n]$ with $|J|, |I| \leq j_0(r)$ such that J $\binom{n-(r+1)}{k-(r+1)}$-captures $\mathcal{A}$ and I $\binom{n-(r+1)}{\ell-(r+1)}$-captures $\mathcal{B}$.*

Proof: Let $\mathcal{A}^{\uparrow}$ be the monotone closure of $\mathcal{A}$, and $\mathcal{B}^{\uparrow}$ the monotone closure of $\mathcal{B}$. According to Lemma 2.8, for every $\varepsilon > 0$ there exists $0.3 < p < 0.5$ and $s(\varepsilon)$-juntas $\mathcal{J}$ and $\mathcal{I}$ with $s(\varepsilon) = 2^{O(1/\varepsilon)}$ such that

$$\mu_p(\mathcal{A}^{\uparrow} \triangle \mathcal{J}) < \epsilon \quad \text{and} \quad \mu_p(\mathcal{B}^{\uparrow} \triangle \mathcal{I}) < \epsilon. \tag{5}$$

Now set $\varepsilon = 100^{-r}$ and apply Lemma 2.8 to get $p \in (0.3, 0.5)$ and juntas $\mathcal{J}$ and $\mathcal{I}$ defined on coordinate sets J and I respectively. Set $j_0(r) = s(\varepsilon)$, the bound on $|J|$ and $|I|$. We now prove that J and I are the required capturing sets. Assume J does not $\binom{n-(r+1)}{k-(r+1)}$-capture $\mathcal{A}$, i.e.,

$$|\mathcal{A}_{\emptyset}(J)| = |\{R \in \mathcal{A} \,:\, R \cap J = \emptyset\}| > \binom{n-(r+1)}{k-(r+1)}.$$

For any $S \subseteq J$ and family of sets $\mathcal{C}$ recall the notation

$$\mathcal{C}_S = \mathcal{C}_S(J) = \{T \in [n] \setminus J : (T \cup S) \in \mathcal{C}\}.$$

Let $\mathcal{J}^*$ be the family generating $\mathcal{J}$. By our assumption $|\mathcal{A}_{\emptyset}| > \binom{n-(r+1)}{k-(r+1)}$. Applying Claim 2.5 to $\mathcal{A}_{\emptyset}^{\uparrow}$, viewed as a family of subsets of $[n] \setminus J$, we get

$$\mu_p(\mathcal{A}_{\emptyset}^{\uparrow}) \geq (1 - o(1))p^{r+1} > \frac{p^{r+1}}{2}.$$

From the monotonicity of $\mathcal{A}^{\uparrow}$ it follows that $\mathcal{A}_S^{\uparrow} \supseteq \mathcal{A}_{\emptyset}^{\uparrow}$ for every $S \subseteq J$. Hence $\mu_p(\mathcal{A}_S^{\uparrow}) \geq \mu_p(\mathcal{A}_{\emptyset}^{\uparrow}) > \frac{p^{r+1}}{2}$. Therefore,

$$\frac{\mu_p\left(\left\{S \cup T \,:\, T \in \mathcal{A}_S^{\uparrow}\right\}\right)}{p^{|S|}(1-p)^{|J \setminus S|}} = \mu_p(\mathcal{A}_S^{\uparrow}) \geq \frac{p^{r+1}}{2},$$

where the first μ_p refers to $\{0,1\}^{[n]}$, and the second μ_p refers to $\{0,1\}^{[n]\setminus J}$. Hence

$$\begin{aligned} \epsilon \;&\geq\; \mu_p(\mathcal{A}^{\uparrow} \setminus \mathcal{J}) = \sum_{S \notin \mathcal{J}^*} \mu_p\left(\left\{S \cup T \,:\, T \in \mathcal{A}_S^{\uparrow}\right\}\right) \\ &\geq\; \sum_{S \notin \mathcal{J}^*} \frac{p^{r+1}}{2} p^{|S|}(1-p)^{|J \setminus S|} = \frac{p^{r+1}}{2} \cdot (1 - \mu_p(\mathcal{J})) \end{aligned}$$

which implies

$$1 - \mu_p(\mathcal{J}) \leq \frac{2\epsilon}{p^{r+1}}. \tag{6}$$

Since $|\mathcal{B}| > \binom{n-(r+1)}{\ell-(r+1)}$ it follows from Claim 2.5 that $\mu_p(\mathcal{B}^{\uparrow}) > p^{r+1} - o(1)$. By Claim 2.4 $\mu_p(\mathcal{A}^{\uparrow}) + \mu_p(\mathcal{B}^{\uparrow}) \leq 1$ so

$$1 - \mu_p(\mathcal{A}^{\uparrow}) \geq \mu_p(\mathcal{B}^{\uparrow}) \geq p^{r+1} - o(1)\,.$$

On the other hand, since $\mu_p(\mathcal{A}^{\uparrow} \setminus \mathcal{J}) \leq \epsilon$ it follows that

$$1 - \mu_p(\mathcal{J}) \geq 1 - \mu_p(\mathcal{A}^{\uparrow}) - \varepsilon \geq p^{r+1}/2 > \frac{2\epsilon}{p^{r+1}},$$

which, by our choice of ε, contradicts (6). ■

Having proven the existence of sets capturing cross-intersecting families we now proceed to use these sets to construct the desired juntas.

Proof: The proof is by induction on r, the case $r = 1$ corresponding to Theorem 1.6 and the case of $r > 1$ corresponding to Theorem 1.7.

Base Case, $r = 1$: Set $c(1) = 2^{j_0(1)} + 1$ where j_0 is the function defined in Lemma 3.2. Let$\mathcal{A} \subset \binom{[n]}{k}$ and $\mathcal{B} \subset \binom{[n]}{\ell}$ be cross intersecting with $|\mathcal{A}| > c(1) \cdot \binom{n-2}{k-2}$ and $|\mathcal{B}| > c(1) \cdot \binom{n-2}{\ell-2}$.

We apply Lemma 3.2 with $r = 1$, and find $J, I \subset [n]$ such that J $\binom{n-2}{k-2}$-captures $\mathcal{A}$ and I $\binom{n-2}{\ell-2}$-captures $\mathcal{B}$, and such that $|J|, |I| \leq j_0(1)$. Let

$$\mathcal{A}_S = \mathcal{A}_S(J) = \{U \subseteq [n] \setminus J \,:\, U \cup S \in \mathcal{A}\} \qquad \text{and} \qquad \mathcal{B}_T = \mathcal{B}_T(I) = \{W \subseteq [n] \setminus I \,:\, W \cup T \in \mathcal{B}\}\,,$$

and

$$\mathcal{J}^* = \left\{S \subseteq J \,:\, |\mathcal{A}_S| > \binom{n-2}{k-2}\right\} \qquad \text{and} \qquad \mathcal{I}^* = \left\{T \subseteq I \,:\, |\mathcal{B}_T| > \binom{n-2}{\ell-2}\right\}.$$

The fact that J and I are capturing sets means precisely that $\emptyset \notin \mathcal{J}^*$ and $\emptyset \notin \mathcal{I}^*$. Now let $\mathcal{J}$ and $\mathcal{I}$ be the juntas defined by $\mathcal{J}^*$ and $\mathcal{I}^*$ respectively. Defining $j(1) := j_0(1)$ they are both $j(1)$-juntas. Also

$$|\mathcal{A} \setminus \mathcal{J}| = \sum_{S \notin \mathcal{J}^*} |\mathcal{A}_S| < \sum_{S \notin \mathcal{J}^*} \binom{n-2}{k-2} < c(1) \cdot \binom{n-2}{k-2},$$

as there are at most $2^{j_0(1)} < c(1)$ summands. Similarly

$$|\mathcal{B} \setminus \mathcal{I}| < c(1) \cdot \binom{n-2}{\ell-2}.$$

Now, if $|S| > 1$ then $S \notin \mathcal{J}^*$ because $|\mathcal{A}_S|$ can be no more than the number of $(k - |S|)$-subsets of $[n] \setminus J$, which is less than $\binom{n-2}{k-2}$. A similar argument applies to $\mathcal{I}^*$ hence all sets in $\mathcal{J}^*, \mathcal{I}^*$ have cardinality precisely 1. The point is that if $\{a\} \in \mathcal{J}^*$ and $\{b\} \in \mathcal{I}^*$ then $a = b$. Assume otherwise, and observe that $\mathcal{A}_{\{a\}}$ and $\mathcal{B}_{\{b\}}$ must be cross intersecting. By Theorem 2.3 either $\left|\mathcal{A}_{\{a\}}\right| \leq \binom{n-|J|-1}{k-2}$ or $\left|\mathcal{B}_{\{b\}}\right| \leq \binom{n-|I|-1}{k-2}$ implying either $\{a\} \notin \mathcal{J}^*$ or $\{b\} \notin \mathcal{I}^*$.

Inductive step, $r \geq 2$**:** Assume we have proven the theorem for values up to $r-1 \geq 1$ and proceed to prove it for r. The proof begins by applying Lemma 3.2 with precision parameter r and finding sets $J_0 \subset [n]$ that $\binom{n-(r+1)}{k-(r+1)}$-captures $\mathcal{A}$ and $I_0 \subset [n]$ that $\binom{n-(r+1)}{\ell-(r+1)}$-captures $\mathcal{B}$. As in the case above, let

$$\mathcal{A}_S = \mathcal{A}_S(J_0) = \{U \subseteq [n] \setminus J_0 \,:\, U \cup S \in \mathcal{A}\} \qquad \text{and} \qquad \mathcal{B}_T = \mathcal{B}_T(I_0) = \{W \subseteq [n] \setminus I_0 \,:\, W \cup T \in \mathcal{B}\}\,,$$

and

$$\mathcal{J}_0^* = \left\{S \subseteq J \,:\, |\mathcal{A}_S| > \binom{n-(r+1)}{k-(r+1)}\right\} \qquad \text{and} \qquad \mathcal{I}_0^* = \left\{T \subseteq I \,:\, |\mathcal{B}_T| > \binom{n-(r+1)}{\ell-(r+1)}\right\}.$$

Clearly, as before, $\emptyset \notin \mathcal{J}_0^*$ and $\emptyset \notin \mathcal{I}_0^*$. Now let $\mathcal{J}_0$ and $\mathcal{I}_0$ be the juntas generated by $\mathcal{J}_0^*$ and $\mathcal{I}_0^*$ respectively. They are our first approximation to $\mathcal{J}, \mathcal{I}$ and we would now be done except for the fact that they are not necessarily cross-intersecting. Indeed we have a bound on the sizes of J_0 and I_0, they are both no larger than $j_0(r)$. Also, using the same reasoning as in the base case $r = 1$ we see that

$$|\mathcal{A} \setminus \mathcal{J}_0| < 2^{|J_0|} \cdot \binom{n-(r+1)}{k-(r+1)} \quad \text{and} \quad |\mathcal{B} \setminus \mathcal{I}_0| < 2^{|I_0|} \cdot \binom{n-(r+1)}{\ell-(r+1)}.$$

One may ask, if this is the case, why do we not make do with a theorem that guarantees approximating juntas, and does not assert that they are cross intersecting. The (slightly self referential) answer is that the stronger statement is needed as the induction hypothesis in order to push the proof through (and that it is nicer to prove a stronger statement).

Returning to $\mathcal{J}_0$ and $\mathcal{I}_0$ we will now mend the potential flaw of them not being cross-intersecting by iteratively refining them, adding more elements to J_0 and I_0 and redefining $\mathcal{J}^*, \mathcal{I}^*$. This process will terminate after a finite number of steps, eventually making the approximating families cross-intersecting.

We rely on the inductive hypothesis to prove the following claim.

Claim 3.3 *Let $S \in \mathcal{J}_0^*$ and $T \in \mathcal{I}_0^*$ be disjoint. Then there exist two sets $J^{(S)} \subseteq [n]$ and $I^{(T)} \subseteq [n]$ and two non trivial juntas $\mathcal{J}^{(S)}$ generated by $\mathcal{J}^{(S)*} \subset 2^{J^{(S)}}$ and $\mathcal{I}^{(T)}$ generated by $I^{(T)*} \subset 2^{I^{(T)}}$ with the following attributes.*

1. $|J^{(S)}|, |I^{(T)}| \leq j_0(r-1)$, *and* $J^{(S)} \cap S = \emptyset$ *and* $I^{(T)} \cap T = \emptyset$.
2. $\mathcal{J}^{(S)}$ *and* $\mathcal{I}^{(T)}$ *are cross-intersecting.*
3. $\mathcal{J}^{(S)}$ *and* $\mathcal{I}^{(T)}$ *approximate* $\mathcal{A}_S$ *and* $\mathcal{B}_T$ *well:*

$$\left|\mathcal{A}_S \setminus \mathcal{J}^{(S)}\right| \leq \binom{n-(r+1)}{k-(r+1)}$$

and

$$\left|\mathcal{B}_T \setminus \mathcal{I}^{(T)}\right| \leq \binom{n-(r+1)}{\ell-(r+1)}.$$

Proof: Clearly $\mathcal{A}_S$ and $\mathcal{B}_T$ are cross intersecting. These are families of $(k-|S|)$-element sets and $(\ell-|T|)$-element sets respectively. We apply the inductive hypothesis to these two cross-intersecting families, with $r-1$. To do so we note that

$$|\mathcal{A}_S| > \binom{n-(r+1)}{k-(r+1)} > c(r-1)\binom{n-(r+1)}{(k-|S|)-(r+1)}$$

and

$$|\mathcal{B}_T| > \binom{n-(r+1)}{\ell-(r+1)} > c(r-1)\binom{n-(r+1)}{(\ell-|T|)-(r+1)}$$

and deduce that $\mathcal{J}^{(S)}$ and $\mathcal{I}^{(T)}$ exist as required, with

$$\left|\mathcal{A}_S \setminus \mathcal{J}^{(S)}\right| \leq c(r-1)\binom{n-(r+1)}{(k-|S|)-(r+1)} \leq \binom{n-(r+1)}{k-(r+1)} \tag{7}$$

and

$$\left|\mathcal{B}_T \setminus \mathcal{I}^{(T)}\right| \leq c(r-1)\binom{n-(r+1)}{(\ell-|T|)-(r+1)} \leq \binom{n-(r+1)}{\ell-(r+1)}. \tag{8}$$

Finally, observe that without loss of generality $J^{(S)} \cap S = \emptyset$ and $I^{(T)} \cap T = \emptyset$ since every member of $\mathcal{A}_S$ is disjoint from S, and likewise for $\mathcal{B}_T$. ∎

We now apply the following iterative algorithm to generate $\mathcal{J}^*$ and $\mathcal{I}^*$:

Step 1: Set $J \leftarrow J_0, I \leftarrow I_0$ and $\mathcal{J}^* \leftarrow \mathcal{J}_0^*, \mathcal{I}^* \leftarrow \mathcal{I}_0^*$.
Step 2: Repeat this step until there are no subsets $S \in \mathcal{J}^*$ and $T \in \mathcal{I}^*$ such that

- $S \cap T = \emptyset$, and
- $|\mathcal{A}_S| > \binom{n-(r+1)}{k-(r+1)}$ and $|\mathcal{B}_T| > \binom{n-(r+1)}{\ell-(r+1)}$.

Select a pair of such subsets S, T and let $J^{(S)}$ and $I^{(T)}$ be the two subsets guaranteed by Claim 3.3 and let $\mathcal{J}^{(S)*} \subseteq 2^{J^{(S)}}$ and $\mathcal{I}^{(T)*} \subseteq 2^{I^{(T)}}$ be the corresponding families, and $\mathcal{J}^{(S)}$ and $\mathcal{I}^{(T)}$ the juntas generated by them. Replace J by $J \cup J^{(S)}$ and update $\mathcal{J}^*$ by removing S and replacing it with $S \cup S'$ for all sets $S' \in \mathcal{J}^{(S)*}$. Similarly for I and $\mathcal{I}^*$:

$$J \leftarrow J \cup J^{(S)}, \quad I \leftarrow I \cup I^{(T)}$$

and

$$\mathcal{J}^* \leftarrow \mathcal{J}^* \cup \left\{S \cup S' : S' \in \mathcal{J}^{(S)*}\right\} \setminus \{S\}, \quad \mathcal{I}^* \leftarrow \mathcal{I}^* \cup \left\{T \cup T' : T' \in \mathcal{J}^{(T)*}\right\} \setminus \{T\}.$$

Step 3: Remove from $\mathcal{J}^*$ all sets S such that $|\mathcal{A}_S| \leq \binom{n-(r+1)}{k-(r+1)}$, and similarly for $\mathcal{I}^*$.

We first observe that the procedure ends after a finite number of steps that depends only on r. To see this note that whenever Step 2 is applied, and a set $S \in J^*$ is replaced by new sets, it is replaced by at most $2^{j(r-1)}$ sets, which are all of size strictly larger than it. Also, if $|S| \geq r+1$ then $S \notin J^*$ because $|\mathcal{A}_S| \leq \binom{n-(r+1)}{k-(r+1)}$. Hence every original set in $\mathcal{J}_0^*$ has at most $2^{r \cdot j(r-1)}$ descendants in this process, resulting altogether in a total of at most $2^{j_0(r)} \cdot 2^{r \cdot j(r-1)}$ sets in each family. Since all sets are of size at most r we can define $j(r) = r \cdot 2^{j_0(r)} \cdot 2^{r \cdot j(r-1)}$ The resulting families are clearly intersecting since each pair of subsets $S \in \mathcal{J}^*, T \in \mathcal{I}^*$ would have undergone replacement if they were disjoint. It remains to show that $\mathcal{A}$ and $\mathcal{B}$ are indeed almost contained in the resulting families. Let $S \in \mathcal{J}^*$ at some stage of the algorithm. Consider the update operation where S is removed from $\mathcal{J}^*$ and replaced by $\{S \cup S' : S' \in \mathcal{J}^{(S)*}\}$, and $\mathcal{J}$ is updated accordingly. There is a 1-1 correspondence between the sets thus added to $\mathcal{A} \setminus \mathcal{J}$ and the sets in $\mathcal{A}_S \setminus \mathcal{J}^{(S)}$. From (7) we have that the number of these is no more than $\binom{n-(r+1)}{k-(r+1)}$. Hence we may accumulate throughout the process at most $c(r)\binom{n-(r+1)}{k-(r+1)}$ such sets where $c(r) := 2^{j_0(r)} \cdot (2^{r\cdot j(r-1)} + 1)$. A similar estimate holds for the approximation of $\mathcal{B}$ by $\mathcal{I}$. This completes the proof. ■

Remark: Clearly the recursive definition given above of the function $j(r)$ gives a tower like growth rate. However, using some combinatorial reasoning, one can deduce that the true behavior is no more than exponential. The final approximating junta may be represented, after discarding any variables that do not play an actual role, as a critical intersecting hypergraph with all edge sizes at most r. Here critical means that for every vertex v there exists a pair of edges E, F such that $E \cap F = \{v\}$. It follows from results of Tuza [13] that the number of vertices in such a hypergraph is at most $O(4^r)$ and this is tight.

3.2 Independent sets in graph products and the case $k = pn$

In this subsection we prove Theorem 1.2, the proof being an almost immediate corollary of results from [2], where independent sets in graph products are studied.

An intersecting family is nothing else than an independent set in a Kneser graph: the graph whose vertices are $\binom{[n]}{k}$, and edges connect pairs of disjoint subsets. The Kneser graph is not a graph product per se, but there is a natural 'smoothed' analog of it from which Theorem 1.2 follows.

We begin with some definitions. The n-th *power* of an undirected graph $G = (V, E)$, denoted by G^n, is defined as follows: the vertex set is V^n, and two vertices $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ are adjacent in G^n if and only if u_i is adjacent to v_i in G for every i. This is, in graph-theoretic terms, the n-fold weak product of G with itself. An alternative description is that the adjacency matrix of G^n is the n-fold tensor product of the adjacency matrix of G with itself. In the spirit of our previous definition of a junta we will say that a set of vertices $\mathcal{J}$ in $V(G^n)$ is a j-junta generated by $\mathcal{J}^*$ if there exists a set of coordinates $J \subset [n]$, and a family $\mathcal{J}^* \subset V^J$ such that $\mathcal{J} = \{v : v|_J \in \mathcal{J}^*\}$, where $v|_J$ denotes the restriction of the vector v to the coordinates in J. For any graph G let $M = M(G)$ be the normalized adjacency matrix of G,

$$M_{u,v} = \begin{cases} 1/\deg(u) & u \sim v \\ \\ 0 & \text{otherwise}\,. \end{cases}$$

Let μ_G be the unique stationary measure of the simple random walk on $V(G)$ (the Markov chain described by M),

$$\mu_G(u) = \frac{\deg(u)}{2|E|}.$$

For a set $\mathcal{A} \subset V(G)$ define the sparseness of $\mathcal{A}$, denoted by $Sp(\mathcal{A})$, to be the proportion of the edges in G that are spanned by $\mathcal{A}$. This is equivalent to the following definition which is the correct one for purpose of the generalization that will follow below.

$$Sp(\mathcal{A}) = Pr\{v \in \mathcal{A}, w \in \mathcal{A}\} \tag{9}$$

where v and w are two consecutive steps in the simple random walk on $V(G)$ with the starting point chosen according to μ_G. Note that if $\mathcal{A}$ is independent then $Sp(\mathcal{A}) = 0$. Also note that in a product graph G^n if $\mathcal{J}$ is a j-junta generated by $\mathcal{J}^*$ then $Sp(\mathcal{J}) = Sp(\mathcal{J}^*)$, where the sparseness of $\mathcal{J}^*$ is measured in G^j.

If the eigenvalues of $M(G)$ are $1 = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$ then we define

$$\lambda(G) = \max\{|\lambda_i|, i > 1\}.$$

Note that if G is connected and nonbipartite then $\lambda(G) < 1$. Also, $\lambda(G^n) = \lambda(G)$.

We now can state the main theorem of [2] which asserts that if G is connected and nonbipartite then every independent set in G^n is almost contained in a sparse junta.

Theorem 3.4 ([2]) *There exists a function $j'(\varepsilon, \lambda)$ such that the following holds. Let G be a connected and non-bipartite graph with $\lambda(G) = \lambda < 1$. Let μ be the stationary measure of the simple random walk on $V(G^n)$. Let $\mathcal{A} \subset V(G^n)$ be an independent set. Then for every $\varepsilon > 0$ there exists a j'-junta $\mathcal{J}$, with $j' = j'(\varepsilon, \lambda)$ such that*

1. $\mu(\mathcal{A} \setminus \mathcal{J}) \leq \varepsilon$.
2. $Sp(\mathcal{J}) \leq \varepsilon$.

As noted in [2] the proof of the theorem applies also to more general reversible, irreducible, aperiodic Markov chains (corresponding to connected non-bipartite graphs with weighted edges.) We now wish to apply the theorem to the setting of intersecting families in $\{0,1\}^n$. To this end we fix $p \in (0, 1/2)$ and consider the Markov chain on the two point space $\{0,1\}$ defined by the matrix

$$M = \begin{pmatrix} \frac{1-2p}{1-p} & \frac{p}{1-p} \\ 1 & 0 \end{pmatrix}. \tag{10}$$

Note that this Markov chain, when in state 1, always moves to 0. When tensored the result has the following nice attributes.

1. The stationary measure of the Markov chain defined by $M^{\otimes n}$ on $\{0,1\}^n$ is precisely the product measure μ_p defined by $\mu_p(x) = p^{\Sigma x_i}(1-p)^{\Sigma(1-x_i)}$.
2. The probability to move from a state $s \in \{0,1\}^n$ to a state $t \in \{0,1\}^n$ is nonzero only if the subsets corresponding to s and t are disjoint. Hence this chain can be viewed as a smoothed analog of the Kneser graph.
3. A Chernoff type bound implies that almost all of the weight of the transitions in the chain is supported by vertices corresponding to sets of size close to $k = pn$.

To verify this it suffices to study the case $n = 1$. We leave this to the reader.

Finally, let us define the notion of ε-almost-intersecting which appears in the statement of Theorem 1.2. Recalling that we identified each $S \subseteq [n]$ with its characteristic vector $s \in \{0,1\}^n$ we henceforth think of the Markov chain defined by $M^{\otimes n}$ as a Markov chain on $2^{[n]}$.

Definition 3.5 *Let $p \in [0, 1/2]$, and let $\mathcal{A} \subseteq 2^{[n]}$. We say that $\mathcal{A}$ is ε-almost intersecting if*

$$Sp(\mathcal{A}) \leq \varepsilon$$

where sparseness is with respect to the chain $M^{\otimes n}$ (with M as defined in (10).)

Remark: An alternative way to think about the definition of the Markov chain is this: from a state corresponding to a set T the chain moves to a set S which is disjoint from T by choosing a random subset of the compliment of T, picking each element independently with probability $\frac{p}{1-p}$. Recalling that the stationary measure of this chain is μ_p the sparseness of $\mathcal{A}$ or the fact that $\mathcal{A}$ is ε-almost-intersecting may be interpreted as follows. When one selects at random two disjoint subsets of $[n]$ such that the marginal distribution of each is μ_p then the sparseness of $\mathcal{A}$ is precisely the probability that they both belong to $\mathcal{A}$. Specifically, let $p \in [0, 1/2]$, and let $\mathcal{A} \subseteq 2^{[n]}$. Partition

$[n]$ into three sets S, T, R at random by adding each element independently to S with probability p, to T with probability p and to R with probability $1-2p$. Then $\mathcal{A}$ is ε-almost-intersecting if

$$\Pr[S \in \mathcal{A} \text{ and } T \in \mathcal{A}] \leq \varepsilon.$$

Going back to the chain M, the eigenvalues of M are 1 and $-\frac{p}{1-p}$ so for $p < 1/2$ we have $\lambda(M) = \frac{p}{1-p} < 1$ and we can apply Theorem 3.4 and deduce the following.

Corollary 3.6 *There exists a function $j'(\varepsilon, p)$ such that the following holds. Let $p \in (0, 1/2)$ and let $\mathcal{A} \subset 2^{[n]}$ be an intersecting family. Then for every $\varepsilon > 0$ there exists a j'-junta $\mathcal{J}$, with $j' = j'(\varepsilon, p)$ such that*

1. *$\mu_p(\mathcal{A} \setminus \mathcal{J}) \leq \varepsilon$.*
2. *$\mathcal{J}$ is ε-almost-intersecting in $\mu_p(\{0,1\}^n)$.*

Theorem 1.2 now follows easily from this corollary.

Proof: Let $\mathcal{A} \subset \binom{[n]}{k}$ be an intersecting family, and let $\mathcal{A}^{\uparrow} \subset \{0,1\}^n$ be its monotone closure. Set $p = k/n$, $\varepsilon' = \varepsilon/3$ and define $j = j(p, \varepsilon) := j'(p, \varepsilon')$. Corollary 3.6 guarantees the existence of an ε-almost-intersecting junta $\mathcal{J}$ in $\mu_p(\{0,1\}^n)$ such that $\mu_p(\mathcal{A}^{\uparrow} \setminus \mathcal{J}) \leq \varepsilon'$. Lemma 2.6 then implies that $|\mathcal{A} \setminus \mathcal{J}| \leq \varepsilon \binom{n}{k}$ as required. ■

Remark: Note that choosing any $p' \neq p$ would not work in this proof. The application of Lemma 2.6 utilizes the fact that μ_p focuses on $\binom{[n]}{k}$.

References

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

[2] I. Dinur, E. Friedgut, and O. Regev. Independent sets in graph powers are almost contained in juntas. Preprint.

[3] I. Dinur and S. Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 162(1):439–485, 2005.

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

[5] P. Frankl and N. Tokushige. Some inequalities concerning cross-intersecting families. *Combinatorics Probability and Computing*, 7:247–260, 1998.

[6] E. Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998.

[7] Z. Füredi. Cross-intersecting families of finite sets. *J. Combin. Theory Ser. A*, 72:332–339, 1995.

[8] G. O. H. Katona. A theorem of finite sets. In P. Erdo, editor, *Theory of Graphs*. Akade'miai Kiado' and Academic Press, 1968.

[9] J. Kruskal. The number of simplices in a complex. In R. Bellman, editor, *Mathematical Optimization Techniques*, pages 251–278. University of California Press, 1963.

[10] G. Margulis. Probabilistic characteristics of graphs with large connectivity. *Prob. Peredachi Inform.*, 10:101–108, 1974.

[11] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. In *Proc. 46th IEEE Symp. on Foundations of Computer Science*, 2005.

[12] L. Russo. An approximate zero-one law. *Z. Wahrsch. werw. Gebiete*, 61:129–139, 1982.

[13] Z. Tuza. Critical hypergraphs and intersecting set-pair systems. *J. Combin. Theory Ser. B*, 39(2), 1985.