

# Ahlswede–Khachatrian Theorems: Weighted, Infinite, and Hamming

Yuval Filmus\*

April 4, 2017

## Abstract

The seminal complete intersection theorem of Ahlswede and Khachatrian gives the maximum cardinality of a  $k$ -uniform  $t$ -intersecting family on  $n$  points, and describes all optimal families. We extend this theorem to several other settings: the weighted case, the case of infinitely many points, and the Hamming scheme.

The weighted Ahlswede–Khachatrian theorem gives the maximal  $\mu_p$  measure of a  $t$ -intersecting family on  $n$  points, where  $\mu_p(A) = p^{|A|}(1-p)^{n-|A|}$ . As has been observed by Ahlswede and Khachatrian and by Dinur and Safra, this theorem can be derived from the classical one by a simple reduction. However, this reduction fails to identify the optimal families, and only works for  $p < 1/2$ . We translate the two original proofs of Ahlswede and Khachatrian to the weighted case, thus identifying the optimal families in all cases. We also extend the theorem to the case  $p > 1/2$ , using a different technique of Ahlswede and Khachatrian (the case  $p = 1/2$  is Katona’s intersection theorem). We then extend the weighted Ahlswede–Khachatrian theorem to the case of infinitely many points.

The Ahlswede–Khachatrian theorem on the Hamming scheme gives the maximum cardinality of a subset of  $\mathbb{Z}_m^n$  in which any two elements  $x, y$  have  $t$  positions  $i_1, \dots, i_t$  such that  $x_{i_j} - y_{i_j} \in \{-(s-1), \dots, s-1\}$ . We show that this case corresponds to  $\mu_p$  with  $p = s/m$ , extending work of Ahlswede and Khachatrian, who considered the case  $s = 1$ . We also determine the maximum cardinality families. We obtain similar results for subsets of  $[0, 1]^n$ , though in this case we are not able to identify all maximum cardinality families.

## 1 Introduction

### 1.1 Background

The Erdős–Ko–Rado theorem [13], a basic result in extremal combinatorics, states that when  $k \leq n/2$ , a  $k$ -uniform intersecting family on  $n$  points contains at most  $\binom{n-1}{k-1}$  sets; and furthermore, when  $k < n/2$  the only families achieving these bounds are *stars*, consisting of all sets containing some fixed point.

The original paper of Erdős, Ko and Rado gave birth to an entire area in extremal combinatorics. Erdős–Ko–Rado theorems have been proved in many other domains, such as vector spaces [16], permutations [12], weighted families [18], and many others (see the recent monograph [19]).

A different research direction is extending the original theorem to  $t$ -intersecting families, in which every two sets are guaranteed to contain  $t$  points in common rather than just one. Extending earlier results of Frankl [15], Wilson [28] showed that when  $n \geq (t+1)(k-t-1)$ , a  $t$ -intersecting family contains at most  $\binom{n-t}{k-t}$  sets, and for  $n > (t+1)(k-t-1)$  this is achieved uniquely by  $t$ -stars. When  $n < (t+1)(k-t-1)$ , there exist  $t$ -intersecting families which are larger than  $t$ -stars. Frankl [15] conjectured that for every  $n, k, t$ , the maximum families are always of the form

$$\mathcal{F}_{t,r} = \{S : |S \cap [t+2r]| \geq t+r\}.$$

Frankl’s conjecture was finally proved by Ahlswede and Khachatrian [3, 5], who gave two different proofs (see also the monograph [1]). They also determined the maximum families under the condition that the intersection of all sets in the family is empty [2], as well as the maximum  $t$ -intersecting families

---

\*Department of Computer Science, Technion — Israel Institute of Technology, Israel.

without a restriction on the size [6] (“Katona’s theorem”). They also proved an analogous theorem for the Hamming scheme [4].

A different generalization of the Ahlswede–Khachatrian theorem appears in the work of Dinur and Safra [10] on the hardness of approximating vertex cover. Dinur and Safra were interested in the maximum  $\mu_p$  measure of a  $t$ -intersecting family on  $n$  points, where the  $\mu_p$  measure is given by  $\mu_p(A) = p^{|A|}(1-p)^{n-|A|}$ . When  $p \leq 1/2$ , they related this question to the setting of the original Ahlswede–Khachatrian theorem with parameters  $K, N$  satisfying  $K/N \approx p$ . A similar argument appears in work of Ahlswede–Khachatrian [6, 4] in different guise. The  $\mu_p$  setting has since been studied in many works, and has been used by Friedgut [18] and by Keller and Lifshitz [23] to prove stability versions of the Ahlswede–Khachatrian theorem.

## 1.2 Our results

While not stated explicitly in either work, the methods of Dinur–Safra [10] and Ahlswede–Khachatrian [4] give a proof of an Ahlswede–Khachatrian theorem in the  $\mu_p$  setting for all  $p < 1/2$ , without any constraint on the number of points. More explicitly, let  $w(n, t, p)$  be the maximum  $\mu_p$ -measure of a  $t$ -intersecting family on  $n$  points, and let  $w_{\text{sup}}(t, p) = \sup_n w(n, t, p)$ . The techniques of Dinur–Safra and Ahlswede–Khachatrian show that when  $\frac{r}{t+2r-1} \leq p \leq \frac{r+1}{t+2r+1}$ ,  $w_{\text{sup}}(t, p) = \mu_p(\mathcal{F}_{t,r})$ . This theorem is incomplete, for three different reasons:

1. The theorem describes  $w_{\text{sup}}(t, p)$  rather than  $w(n, t, p)$ . If  $n < t + 2r$  then the family  $\mathcal{F}_{t,r}$  does not fit into  $n$  points, and so  $w(n, t, p) < w_{\text{sup}}(t, p)$ .
2. Whereas the classical Ahlswede–Khachatrian theorem describes all optimal families, the methods of Dinur–Safra and Ahlswede–Khachatrian cannot produce such a description in the  $\mu_p$  case.
3. The theorem only handles the case  $p < 1/2$ .

Katona [21] solved the case  $p = 1/2$ , which became known as “Katona’s theorem”. Ahlswede and Khachatrian gave a different proof [6], and their technique applies also to the case  $p > 1/2$ . We complete the picture by finding  $w(n, t, p)$  for all  $n, t, p$  and determining all families achieving this bound. We do this by rephrasing the two original proofs [3, 5] of the Ahlswede–Khachatrian theorem in the  $\mu_p$  setting. Essentially the same proofs allow us to obtain the tighter results. Repeating the proofs in the easier setting of  $\mu_p$  has the additional advantage of simplifying some of the notations and calculations in the proofs. Curiously, whereas the classical Ahlswede–Khachatrian theorem can be proven using either of the techniques described in [3, 5], our proof needs to use both. Part of these results appear in the author’s PhD thesis [14].

We also generalize Katona’s circle argument [22] to the  $\mu_p$  setting, thus giving a particularly simple proof of the Erdős–Ko–Rado theorem in the  $\mu_p$  setting (the work of Dinur and Friedgut [9] contains a similar but much longer proof, and Friedgut [17] gives a more complicated proof in the same style).

The methods of Dinur–Safra and Ahlswede–Khachatrian calculate the maximum  $\mu_p$  measure of a  $t$ -intersecting family on an unbounded number of points. It is natural to ask what happens when we allow *infinitely* many points rather than just an unbounded number. Formally, let  $w_{\infty}(p, t)$  be the supremum  $\mu_p$ -measure of a  $t$ -intersecting subset of  $\{0, 1\}^{\mathbb{N}^0}$ ; we only consider  $\mu_p$ -measurable subsets. When  $p < 1/2$ , we obtain  $w_{\infty}(p, t) = w_{\text{sup}}(p, t)$ , and moreover families achieving this depend (up to measure zero) on a finite number of points, and thus must be of the form  $\mathcal{F}_{t,r}$ . We also have  $w_{\infty}(1/2, 1) = w_{\text{sup}}(1/2, 1) = 1/2$ , and this is achieved only on families depending (up to measure zero) on a finite number of points. When  $t > 1$ , we have  $w_{\infty}(1/2, t) = w_{\text{sup}}(1/2, t) = 1/2$ , but no family achieves this. Finally, when  $p > 1/2$  we have  $w_{\infty}(p, t) = w_{\text{sup}}(p, t) = 1$ , and in fact there is a family of  $\mu_p$ -measure 1 in which any two sets have infinite intersection.

Ahlswede and Khachatrian [4] considered a natural generalization of their theorem to the Hamming scheme. In this variant, a family is a subset of  $\mathbb{Z}_m^n$ , and a family  $\mathcal{F}$  is  $t$ -intersecting if any two vectors in  $\mathcal{F}$  agree on at least  $t$  coordinates. Ahlswede and Khachatrian show, in effect, that this corresponds to  $p = 1/m$ . This setting thus constitutes a combinatorial analog of  $\mu_p$  for  $p$  of the form  $1/m$ . It is natural to ask whether we can find similar combinatorial analogs for other values of  $p$ . We show that it is indeed the case, by describing a combinatorial analog for  $p = s/m$ , where  $s \leq m/2$ . In this analog, we still consider subsets of  $\mathbb{Z}_m^n$ , but the definition of  $t$ -intersecting is different: two vectors  $x, y$  are  $t$ -intersecting

if there are  $t$  positions satisfying  $x_i - y_i \in \{-(s-1), \dots, s-1\}$ , the subtraction taking place in  $\mathbb{Z}_m$ . We show that the maximum cardinality of such a family is  $w(n, t, s/m)m^n$ . Moreover, when  $s/m < 1/2$ , all optimal families arise from families of the form  $\mathcal{F}_{t,r}$ .

A similar setting, which works for all  $p < 1/2$ , is that in which  $\mathbb{Z}_m$  is replaced by the unit circle  $[0, 1]$ . We consider (Lebesgue) measurable subsets of  $[0, 1]^n$  in which any two vectors  $x, y$  have  $t$  coordinates satisfying  $x_i - y_i \in [-p, p]$ , the subtraction taking place on the circle (that is, modulo 1). We show that the maximum  $\mu_p$  measure of such a family is  $w(n, t, p)$ . Unfortunately, our argument is not strong enough to determine all optimal families.

**Organization of the paper** The main part of the paper is Section 2, in which we prove the Ahlswede–Khachatrian theorem in the  $\mu_p$  setting. Following an exploration of Katona’s circle argument in Section 3 (which is needed for subsequent sections), we prove the infinite version of the Ahlswede–Khachatrian theorem in Section 4, and a version for the Hamming scheme in Section 5. Finally, we describe in Section 6 the arguments of Dinur–Safra and Ahlswede–Khachatrian, which lift the Ahlswede–Khachatrian theorem to the  $\mu_p$  setting.

**Notation** We will use  $[n]$  for  $\{1, \dots, n\}$ , and  $\binom{[n]}{k}$  for all subsets of  $[n]$  of size  $k$ . We also use the somewhat unorthodox notation  $\binom{[n]}{\geq k}$  for all subsets of  $[n]$  of size at least  $k$ . The set of all subsets of a set  $A$  will be denoted  $2^A$ . The notation  $\text{Bin}(n, p)$  is used for the binomial distribution with  $n$  trials and success probability  $p$ .

## 2 Weighted Ahlswede–Khachatrian theorem

In this section we prove an Ahlswede–Khachatrian theorem for families of sets, measured according to the  $\mu_p$  measures. We start with several basic definitions.

**Definition 2.1.** A *family on  $n$  points* is a collection of subsets of  $[n]$ . A family  $\mathcal{F}$  is  *$t$ -intersecting* if any  $A, B \in \mathcal{F}$  satisfy  $|A \cap B| \geq t$ . Two families  $\mathcal{F}, \mathcal{G}$  are *cross- $t$ -intersecting* if any  $A \in \mathcal{F}$  and  $B \in \mathcal{G}$  satisfy  $|A \cap B| \geq t$ .

We will refer to a 1-intersecting family or a pair of cross-1-intersecting families as an *intersecting* family or a pair of *cross-intersecting* families.

**Definition 2.2.** For any  $p \in (0, 1)$  and any  $n$ , the product measure  $\mu_p$  is a measure on the set of subsets of  $[n]$  given by

$$\mu_p(A) = p^{|A|}(1-p)^{n-|A|}.$$

**Definition 2.3.** A family  $\mathcal{F}$  on  $n$  points is *monotone* if whenever  $A \in \mathcal{F}$  and  $B \supseteq A$  then  $B \in \mathcal{F}$ . Given a family  $\mathcal{F}$ , its *up-set*  $\mathcal{G}$  is the smallest monotone family containing  $\mathcal{F}$ , namely

$$\mathcal{G} = \{B \supseteq A : A \in \mathcal{F}\}.$$

We also denote  $\mathcal{G}$  by  $\langle \mathcal{F} \rangle$ .

Note that if  $\mathcal{F}$  is  $t$ -intersecting then so is its up-set. It follows that if  $\mu_p(\mathcal{F}) = w(n, t, p)$  then  $\mathcal{F}$  is monotone.

One main goal of this section is to determine the two quantities  $w(n, t, p)$  and  $w_{\text{sup}}(t, p)$ , which are defined as follows.

**Definition 2.4.** For  $n \geq t \geq 1$  and  $p \in (0, 1)$ , the parameter  $w(n, t, p)$  is the maximum of  $\mu_p(\mathcal{F})$  over all  $t$ -intersecting families on  $n$  points.

**Definition 2.5.** For  $t \geq 1$  and  $p \in (0, 1)$ , the parameter  $w_{\text{sup}}(t, p)$  is given by

$$w_{\text{sup}}(t, p) = \sup_n w(n, t, p).$$

We can also define  $w_{\text{sup}}(t, p)$  as a limit instead of a supremum.

**Lemma 2.1.** For  $t \geq 1$  and  $p \in (0, 1)$ ,

$$w_{\text{sup}}(t, p) = \lim_{n \rightarrow \infty} w(n, t, p).$$

*Proof.* If  $\mathcal{F}$  is a  $t$ -intersecting family on  $n$  points then  $\mathcal{F}' = \{S \subseteq [n+1] : S \cap [n] \in \mathcal{F}\}$  is a  $t$ -intersecting family on  $n+1$  points having the same  $\mu_p$ -measure. Therefore the sequence  $w(n, t, p)$  is non-decreasing, and the lemma follows.  $\square$

As mentioned in the introduction,  $w_{\text{sup}}(t, p)$  for  $p < 1/2$  can be derived from the classical Ahlswede–Khachatryan theorem via the methods of Dinur–Safra and Ahlswede–Khachatryan, derivations that we describe in Section 6.

The other main goal of this section is finding all  $t$ -intersecting families on  $n$  points which achieve the bound. The families, first described by Frankl [15], are the following.

**Definition 2.6.** For  $t \geq 1$  and  $r \geq 0$ , the  $(t, r)$ -Frankl family on  $n$  points is the  $t$ -intersecting family

$$\mathcal{F}_{t,r} = \{A \subseteq [n] : |A \cap [t+2r]| \geq t+r\}.$$

A family  $\mathcal{F}$  on  $n$  points is *equivalent* to a  $(t, r)$ -Frankl family if there exists a set  $S \subseteq [n]$  of size  $t+2r$  such that

$$\mathcal{F} = \{A \subseteq [n] : |A \cap S| \geq t+r\}.$$

We will prove the following theorem.

**Theorem 2.2.** Let  $n \geq t \geq 1$  and  $p \in (0, 1)$ . If  $\mathcal{F}$  is  $t$ -intersecting then

$$\mu_p(\mathcal{F}) \leq \max_{r: t+2r \leq n} \mu_p(\mathcal{F}_{t,r}).$$

Moreover, unless  $t = 1$  and  $p \geq 1/2$ , equality holds only if  $\mathcal{F}$  is equivalent to a Frankl family  $\mathcal{F}_{t,r}$ .

When  $t = 1$  and  $p > 1/2$ , the same holds if  $n+t$  is even, and otherwise  $\mathcal{F} = \mathcal{G} \cup \left(\geq \frac{[n]}{\frac{n+t+1}{2}}\right)$  where  $\mathcal{G} \subseteq \left(\frac{[n]}{\frac{n+t-1}{2}}\right)$  contains exactly  $\binom{n-1}{\frac{n+t-1}{2}}$  sets.

When  $t = 1$  and  $p = 1/2$  there are many optimal families. For example, the families  $\mathcal{F}_{1,r}$  all have  $\mu_{1/2}$ -measure  $1/2$ , as does the family  $(\{S : 1 \in S\}) \setminus \{\{1\}\} \cup \{\{2, \dots, n\}\}$ .

Similarly, when  $t = 1$ ,  $p > 1/2$  and  $n+1$  is odd there are many optimal families. For example, both of the following families are optimal:

$$\begin{aligned} & \{S \subseteq [n] : |S| \geq \frac{n}{2} + 1\} \cup \{S \subseteq [n] : |S| = \frac{n}{2} \text{ and } 1 \notin S\}, \\ & \{S \subseteq [n] : |S| \geq \frac{n}{2} + 1\} \cup \{S \subseteq [n] : |S| = \frac{n}{2} \text{ and } 1 \in S\}. \end{aligned}$$

Our proof actually implies the following more detailed corollary, which can also be deduced from the theorem using Lemma 2.5 below.

**Corollary 2.3.** Let  $n \geq t \geq 1$ . Define  $r^*$  as the maximal integer satisfying  $t+2r^* \leq n$ .

If  $t = 1$  then

$$w(n, 1, p) = \begin{cases} p & p \leq \frac{1}{2}, \\ \mu_p(\mathcal{F}_{1,r^*}) & p \geq \frac{1}{2}. \end{cases}$$

Furthermore, if  $\mathcal{F}$  is an intersecting family of  $\mu_p$ -measure  $w(n, 1, p)$  for  $p \in (0, 1)$  then:

- If  $p < \frac{1}{2}$  then  $\mathcal{F}$  is equivalent to  $\mathcal{F}_{1,0}$ .
- If  $p > \frac{1}{2}$  and  $n$  is odd then  $\mathcal{F}$  is equivalent to  $\mathcal{F}_{1, \frac{n-1}{2}}$ .
- If  $p > \frac{1}{2}$  and  $n$  is even then  $\mathcal{F} = \mathcal{G} \cup \left(\geq \frac{[n]}{\frac{n}{2}+1}\right)$ , where  $\mathcal{G}$  contains half the sets in  $\binom{[n]}{\frac{n}{2}}$ : exactly one of each pair  $A, [n] \setminus A$ .

If  $t \geq 2$  then

$$w(n, t, p) = \begin{cases} \mu_p(\mathcal{F}_{t,r}) & \frac{r}{t+2r-1} \leq p \leq \frac{r+1}{t+2r+1} \text{ for some } r < r^*, \\ \mu_p(\mathcal{F}_{t,r^*}) & \frac{r^*}{t+2r^*-1} \leq p. \end{cases}$$

Furthermore, if  $\mathcal{F}$  is a  $t$ -intersecting family of  $\mu_p$ -measure  $w(n, t, p)$  for  $p \in (0, 1)$  then:

- If  $\frac{r}{t+2r-1} < p < \frac{r+1}{t+2r+1}$  for some  $r < r^*$  then  $\mathcal{F}$  is equivalent to  $\mathcal{F}_{t,r}$ .
- If  $\frac{r^*}{t+2r^*-1} < p$  then  $\mathcal{F}$  is equivalent to  $\mathcal{F}_{t,r^*}$ .
- If  $p = \frac{r+1}{t+2r+1}$  for some  $r < r^*$  then  $\mathcal{F}$  is equivalent to  $\mathcal{F}_{t,r}$  or to  $\mathcal{F}_{t,r-1}$ .

As a corollary, we can compute  $w(t, p)$ .

**Corollary 2.4.** *We have*

$$w_{\text{sup}}(1, p) = \begin{cases} p & p \leq \frac{1}{2}, \\ 1 & p > \frac{1}{2} \end{cases}$$

For  $t \geq 2$ , we have

$$w_{\text{sup}}(t, p) = \begin{cases} \mu_p(\mathcal{F}_{t,r}) & \frac{r}{t+2r-1} \leq p \leq \frac{r+1}{t+2r+1}, \\ \frac{1}{2} & p = \frac{1}{2}, \\ 1 & p > \frac{1}{2}. \end{cases}$$

*Proof.* We start with the case  $p < 1/2$ . If  $t = 1$  then  $w(n, 1, p) = p$  for all  $n \geq 1$ , implying the formula for  $w_{\text{sup}}(1, p)$ . If  $t \geq 2$  and  $\frac{r}{t+2r-1} \leq p \leq \frac{r+1}{t+2r+1}$  then  $w(n, t, p) = \mu_p(\mathcal{F}_{t,r})$  for all  $n \geq t + 2r$ , implying the formula for  $w(t, p)$ .

Suppose now that  $p \geq 1/2$ . In this case we have  $w(t + 2r, t, p) = w(t + 2r + 1, t, p) = \mathcal{F}_{t,r}$ , and  $\mu_p(\mathcal{F}_{t,r}) = \Pr[\text{Bin}(t + 2r, p) \geq t + r]$ . If  $p = 1/2$  then

$$\mu_{\frac{1}{2}}(\mathcal{F}_{t,r}) = \Pr[\text{Bin}(t + 2r, \frac{1}{2}) \geq t + r] = \Pr[\text{Bin}(t + 2r, \frac{1}{2}) \geq \frac{t}{2} + r] - o_r(1) = \frac{1}{2} - o_r(1),$$

using the central limit theorem, and so  $w_{\text{sup}}(t, 1/2) = 1/2$ . If  $p > 1/2$  then for any  $q \in (1/2, p)$ , for large enough  $r$  we have

$$\mu_p(\mathcal{F}_{t,r}) = \Pr[\text{Bin}(t + 2r, p) \geq t + r] \geq \Pr[\text{Bin}(t + 2r, p) \geq q(t + 2r)] \xrightarrow{r \rightarrow \infty} 1,$$

using the central limit theorem. □

Figure 1 illustrates Corollary 2.3 and Corollary 2.4.

**Organization of the section** Section 2.1 gives a brief overview of the entire proof. Section 2.2 analyzes the  $\mu_p$ -measures of the families  $\mathcal{F}_{t,r}$ . The proof of the main theorem in the cases  $p < 1/2$  and  $p = 1/2$  appears in Section 2.6 and Section 2.7 (respectively), using techniques developed in Sections 2.3–2.5. Finally, the case  $p > 1/2$  is handled in Section 2.8.

## 2.1 Proof overview

The cases  $p \leq 1/2$  and  $p > 1/2$  have completely different proofs, which we describe separately.

When  $p > 1/2$ , the proof uses a shifting technique due to Ahlswede and Khachatrian [6], which they developed with the case  $p = 1/2$  (Katona's theorem) in mind. Their shifting is parametrized by two disjoint sets  $A, B$  of sizes  $|A| = s$  and  $|B| = s + 1$ . If a set  $S$  in the family contains  $A$  and is disjoint from  $B$ , the shifting replaces  $S$  with  $(S \setminus A) \cup B$ , unless the latter is already in the family. This shifting increases the  $\mu_p$ -measure of the family when  $p > 1/2$ , and if done correctly maintains the property of being  $t$ -intersecting. Eventually we reach a family which is stable under all such shifts. A simple argument shows that in a stable  $t$ -intersecting family, all sets have cardinality at least  $\frac{n+t-1}{2}$ . When  $n + t$  is even, this directly yields the desired result. When  $n + t$  is odd, the argument is a bit more

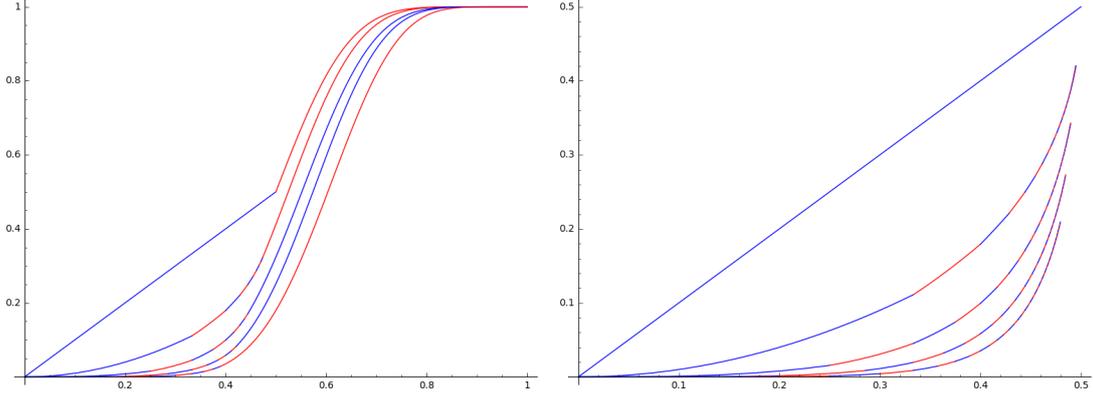


Figure 1: The function  $w(20, t, p)$  for  $1 \leq t \leq 5$  (left) and the function  $w_{\text{sup}}(t, p)$  for  $1 \leq t \leq 5$  (right). In both cases, larger functions correspond to smaller  $p$ . The colors switch at each of the breakpoints  $\frac{r}{t+2r-1}$  for  $r \leq r^*$  (left) or for each  $r$  (right).

subtle, involving a characterization of all maximum size  $t$ -intersecting subsets of  $\binom{[t+2r+1]}{t+r}$ . The entire argument is described in Section 2.8.

The case  $p \leq 1/2$  combines two different techniques of Ahlswede and Khachatrian: the technique of generating sets [3], and the pushing-pulling technique [5]. In the classical setting, each of these techniques suffices by itself to prove the Ahlswede–Khachatrian, using the fact that if  $\mathcal{F} \subseteq \binom{[n]}{k}$  is  $t$ -intersecting then  $\{[n] \setminus S : S \in \mathcal{F}\} \subseteq \binom{[n]}{n-k}$  is  $(n - 2k + t)$ -intersecting. In our setting, this argument doesn't seem to work, and so we need to use both techniques together.

**Shifting** The first step is to apply classical shifting to the family to make it left-compressed: if  $A$  is in the family and  $i < j$  are such that  $j \in A$  and  $i \notin A$ , then  $(A \setminus \{j\}) \cup \{i\}$  is also in the family. The shifting maintains the property of being  $t$ -intersecting and doesn't change the  $\mu_p$ -measure. This (standard) step is described in Section 2.3.

**Generating sets** The first technique, that of generating sets, is used to show that an optimal family depends on few coordinates. We can assume without loss of generality that the family  $\mathcal{F}$  in question is monotone. If  $\mathcal{G}$  consists of all inclusion-minimal sets of  $\mathcal{F}$  (also known as *minterns* or, following Ahlswede and Khachatrian, *generating sets*), then  $\mathcal{F}$  simply consists of all supersets of sets in  $\mathcal{G}$ . We can thus work with  $\mathcal{G}$  instead of  $\mathcal{F}$ . The idea is that since  $\mathcal{F}$  is left-compressed, the number of coordinates that  $\mathcal{F}$  depends on is exactly the maximal number appearing in a set of  $\mathcal{G}$ . Denote this number by  $m$ , and let  $\mathcal{G}^*$  consist of all sets in  $\mathcal{G}$  containing  $m$ .

Suppose that we would like to prove that if  $\mathcal{F}$  has maximum measure then it depends on at most  $t + 2r$  coordinates. If  $m > t + 2r$  then we would like to modify  $\mathcal{F}$  to a  $t$ -intersecting family of larger measure. Since the optimal family depends on fewer than  $m$  coordinates, it is natural to modify  $\mathcal{F}$  by getting rid of its dependency on  $m$ . One way of doing this is by deleting  $m$  from sets in  $\mathcal{G}^*$ . Case-by-case analysis shows that the only reason the new family will not be  $t$ -intersecting is if we removed  $m$  from two sets  $A, B \in \mathcal{G}^*$  which intersect in *exactly*  $t$  elements. Moreover, this can only happen if  $|A| + |B| = m + t$ . In view of this, we partition  $\mathcal{G}^*$  into sets  $\mathcal{G}_0^*, \dots, \mathcal{G}_m^*$  according to their size.

Going ahead with the idea of deleting  $m$  from sets in  $\mathcal{G}^*$  while keeping in mind the need for the family to remain  $t$ -intersecting, we operate on pairs  $\mathcal{G}_a^*, \mathcal{G}_b^*$  for  $a + b = m + t$ . If  $a \neq b$  then we can delete  $m$  from all sets in  $\mathcal{G}_a^*$  and, in order to guarantee that the result is  $t$ -intersecting, *remove* all sets in  $\mathcal{G}_b^*$ . Alternatively, we can delete  $m$  from all sets in  $\mathcal{G}_b^*$ , and remove all sets in  $\mathcal{G}_a^*$ . Calculation shows that one of these two families has measure larger than  $\mathcal{F}$  whenever  $p < 1/2$ .

We cannot execute the preceding operation on  $\mathcal{G}_a^*$  when  $a = \frac{m+t}{2}$ . Instead, we choose an element  $i < m$ , delete  $m$  from all sets in  $\mathcal{G}_a^*$  *not* containing  $i$ , and remove all sets in  $\mathcal{G}_a^*$  containing  $i$ . The resulting family is  $t$ -intersecting for all  $i < m$ , and when  $p < \frac{m-t}{2(m-1)}$ , one of the new families has measure larger than  $\mathcal{F}$ . This case cannot happen when  $m = t + 2r + 1$ , and when  $m = t + 2r + 2$ , the critical probability

is  $\frac{m-t}{2(m-1)} = \frac{r+1}{t+r+1}$ , what we expect.

These arguments appear in Section 2.4.

**Pushing-pulling** The second technique, that of pushing-pulling, is used to show that an optimal family is *symmetric*, that is, invariant to permutations of coordinates it depends on; or in other words, a family of the form  $\{S : |S \cap X| \geq s\}$ . If a family is not completely symmetric then perhaps it is symmetric in at least some of its coordinates. We define  $\ell$  to be the largest number such that the family is invariant under permutations of the first  $\ell$  coordinates. The parameter  $\ell$  plays a role similar to that of the parameter  $m$  in the preceding technique.

Suppose that the family  $\mathcal{F}$  depends on  $m > \ell$  coordinates. Since  $\mathcal{F}$  is left-compressed, the only thing preventing it from being more symmetric are sets  $S \in \mathcal{F}$  such that  $(S \setminus \{i\}) \cup \{\ell + 1\} \notin \mathcal{F}$ , where  $i \in S$  and  $\ell + 1 \notin S$ . We denote the collection of all these sets by  $\mathcal{X}$ . The natural way of making  $\mathcal{F}$  more symmetric is to take some  $A \in \mathcal{X}$  and add  $(A \setminus \{i\}) \cup \{\ell + 1\}$  to  $\mathcal{F}$  for all  $i \in A$ ; we call this operation *expansion*. Case-by-case analysis shows that the only reason the new family will not be  $t$ -intersecting is if we applied this operation to two sets  $A, B \in \mathcal{X}$  having intersection *exactly*  $t$ . Moreover, this can only happen if  $|A \cap [\ell]| + |B \cap [\ell]| = \ell + t$ . In view of this, we partition  $\mathcal{X}$  into sets  $\mathcal{X}_0, \dots, \mathcal{X}_\ell$  according to the size of their intersection with  $[\ell]$ .

Mirroring our efforts in the preceding technique, we will operate on pairs  $\mathcal{X}_a, \mathcal{X}_b$  for  $a + b = \ell + t$ . If  $a \neq b$  then we can expand all sets in  $\mathcal{X}_a$  and delete all sets in  $\mathcal{X}_b$  or vice versa. In both cases the result will be  $t$ -intersecting. Calculation shows that for *all*  $p$ , one of the two new families will have measure larger than  $\mathcal{F}$ .

In order to handle  $\mathcal{X}_a$  for  $a = \frac{\ell+t}{2}$  we will need a different approach: we expand  $\mathcal{X}_a$ , but for both  $\mathcal{X}_a$  and its expansions we keep only sets containing some new common element  $s \neq \ell + 1$ . Such a common element can be found if  $m < n$  but not if  $m = n$ , and this corresponds to the fact that the optimal family  $\mathcal{F}_{t,r}$  satisfies  $t + 2r \leq n$ . This operation increases the measure as long as  $p > \frac{\ell-t+2}{2(\ell+1)}$ . Assuming that  $m = t + 2r$ , if  $\ell = m - 1$  then the new operation isn't needed (since  $\ell + t$  is odd), and if  $\ell = m - 2$  then the critical probability is  $\frac{\ell-t+2}{2(\ell+1)} = \frac{t}{t+2r-1}$ , again what we expect.

These arguments appear in Section 2.5.

**Completing the proof** When  $\frac{r}{t+2r-1} < p < \frac{r+1}{t+2r+1}$ , applying both techniques immediately shows that any left-compressed  $t$ -intersecting family of maximum measure must equal  $\mathcal{F}_{t,r}$ . An arbitrary  $t$ -intersecting family can be made left-compressed by applying a sequence of shifts. Since the shifts don't affect the measure, this shows that any  $t$ -intersecting family has measure at most  $\mu_p(\mathcal{F}_{t,r})$ . A more delicate argument (taken from [3] and appearing in Section 2.3) shows that if shifting results in a family equivalent to  $\mathcal{F}_{t,r}$  then the original family is also equivalent to  $\mathcal{F}_{t,r}$ , and this shows that any  $t$ -intersecting family of maximum measure must be equivalent to  $\mathcal{F}_{t,r}$ .

When  $\frac{r}{t+2r-1} < p \leq 1/2$  and when  $p = \frac{r}{t+2r-1}$  the application of the techniques becomes more subtle, but otherwise the approach is very similar.

The argument for  $p < 1/2$  appears in Section 2.6, and the one for  $p = 1/2$  appears in Section 2.7.

## 2.2 Frankl families

Before starting the proper part of the proof, it is prudent to study the Frankl families  $\mathcal{F}_{t,r}$ . It is easy to see that they are all  $t$ -intersecting. The following lemma analyzes their measures.

**Lemma 2.5.** *Let  $t \geq 1$  and  $r \geq 0$  be parameters, and let  $p_{t,r} = \frac{r+1}{t+2r+1}$ . If  $p < p_{t,r}$  then  $\mu_p(\mathcal{F}_{t,r}) > \mu_p(\mathcal{F}_{t,r+1})$ . If  $p = p_{t,r}$  then  $\mu_p(\mathcal{F}_{t,r}) = \mu_p(\mathcal{F}_{t,r+1})$ . If  $p > p_{t,r}$  then  $\mu_p(\mathcal{F}_{t,r}) < \mu_p(\mathcal{F}_{t,r+1})$ .*

*Proof.* Suppose for simplicity that we consider the two families  $\mathcal{F}_{t,r}$  and  $\mathcal{F}_{t,r+1}$  as families on  $t + 2r + 2$  points. It is then not too hard to check that  $\mathcal{F}_{t,r} \setminus \mathcal{F}_{t,r+1} = \binom{[t+2r]}{t+r}$  and  $\mathcal{F}_{t,r+1} \setminus \mathcal{F}_{t,r} = \binom{[t+2r]}{t+r-1} \times \{t + 2r + 1, t + 2r + 2\}$ . Therefore

$$\begin{aligned} \mu_p(\mathcal{F}_{t,r}) - \mu_p(\mathcal{F}_{t,r+1}) &= \binom{t+2r}{t+r} p^{t+r} (1-p)^{r+2} - \binom{t+2r}{t+r-1} p^{t+r+1} (1-p)^{r+1} \\ &= \binom{t+2r}{t+r-1} p^{t+r} (1-p)^{r+1} \left[ \frac{r+1}{t+r} (1-p) - p \right]. \end{aligned}$$

The bracketed expression equals  $\frac{r+1}{t+r} - \frac{t+2r+1}{t+r}p$ , and so is positive for  $p < \frac{r+1}{t+2r+1}$ , negative for  $p > \frac{r+1}{t+2r+1}$ , and vanishes when  $p = \frac{r+1}{t+2r+1}$ .  $\square$

**Corollary 2.6.** *Let  $t \geq 2$ . If  $p_{t,r-1} < p < p_{t,r}$  then the maximum of  $\mu_p(\mathcal{F}_{t,s})$  is attained at  $s = r$ . If  $p = p_{t,r}$  then the maximum of  $\mu_p(\mathcal{F}_{t,s})$  is attained at  $s = r$  and  $s = r + 1$ .*

Using this lemma one can deduce Corollary 2.3 from Theorem 2.2.

## 2.3 Shifting

The proofs of Ahlswede and Khachatrian use the technique of *shifting* to obtain families which are easier to analyze.

**Definition 2.7.** Let  $\mathcal{F}$  be a family on  $n$  points and let  $i, j \in [n]$  be two different indices. The shift operator  $\mathbb{S}_{i,j}$  acts on  $\mathcal{F}$  as follows. Let  $\mathcal{F}_{i,j}$  consist of all sets in  $\mathcal{F}$  containing  $i$  but not  $j$ . Then

$$\begin{aligned} \mathbb{S}_{i,j}(\mathcal{F}) = & (\mathcal{F} \setminus \mathcal{F}_{i,j}) \cup \{A : A \in \mathcal{F}_{i,j} \text{ and } (A \setminus \{i\}) \cup \{j\} \in \mathcal{F}\} \\ & \cup \{(A \setminus \{i\}) \cup \{j\} : A \in \mathcal{F}_{i,j} \text{ and } (A \setminus \{i\}) \cup \{j\} \notin \mathcal{F}\}. \end{aligned}$$

In words, we try to “shift” each set  $A \in \mathcal{F}_{i,j}$  by replacing it with  $A' = (A \setminus \{i\}) \cup \{j\}$ . If  $A' \notin \mathcal{F}$  then we replace  $A$  with  $A'$ , and otherwise we don't change  $A$ .

Shifting clearly preserves the  $\mu_p$ -measure of the family, a property which we record.

**Lemma 2.7.** *For any family  $\mathcal{F}$  and indices  $i, j$  and for all  $p \in (0, 1)$ ,  $\mu_p(\mathcal{F}) = \mu_p(\mathbb{S}_{i,j}(\mathcal{F}))$ .*

Shifting preserves the property of being  $t$ -intersecting. This is a well-known property which we prove for completeness.

**Lemma 2.8.** *If  $\mathcal{F}$  is  $t$ -intersecting then so is  $\mathbb{S}_{i,j}(\mathcal{F})$  for any  $i, j$ .*

*Proof.* Suppose  $A, B \in \mathbb{S}_{i,j}(\mathcal{F})$ . There are three cases to consider.

**Case 1:**  $A, B \in \mathcal{F}$ . In this case it is clear that  $|A \cap B| \geq t$ .

**Case 2:**  $A, B \notin \mathcal{F}$ . In this case  $j \in A, B$ ,  $i \notin A, B$ , and the two sets  $A' = (A \setminus \{j\}) \cup \{i\}$  and  $B' = (B \setminus \{j\}) \cup \{i\}$  are in  $\mathcal{F}$ . It is not hard to check that  $|A \cap B| = |A' \cap B'| \geq t$ .

**Case 3:**  $A \in \mathcal{F}$  and  $B \notin \mathcal{F}$ . In this case  $j \in B$ ,  $i \notin B$ , and  $B' = (B \setminus \{j\}) \cup \{i\} \in \mathcal{F}$ . Thus  $|A \cap B'| \geq t$ . If  $i \notin A$  or  $j \in A$  then  $|A \cap B| \geq |A \cap B'| \geq t$ . If  $i \in A$  and  $j \notin A$  then  $A' = (A \setminus \{j\}) \cup \{i\} \in \mathcal{F}$  (since otherwise  $A$  would have been shifted), and so  $|A \cap B| \geq |A' \cap B'| \geq t$ .  $\square$

By shifting  $\mathcal{F}$  repeatedly we can obtain a *left-compressed* family. This is another well-known property which we prove for completeness.

**Definition 2.8.** A family  $\mathcal{F}$  on  $n$  points is *left-compressed* if whenever  $A \in \mathcal{F}$ ,  $i \in A$ ,  $j \notin A$ , and  $j < i$ , then  $(A \setminus \{i\}) \cup \{j\} \in \mathcal{F}$ . (Informally, we can *shift*  $i$  to  $j$ .)

**Lemma 2.9.** *Let  $\mathcal{F}$  be a  $t$ -intersecting family on  $n$  points. There is a left-compressed  $t$ -intersecting family  $\mathcal{G}$  on  $n$  points with the same  $\mu_p$ -measure for all  $p \in (0, 1)$ . Furthermore,  $\mathcal{G}$  can be obtained from  $\mathcal{F}$  by applying a sequence of shift operators.*

*Proof.* A family is left-compressed if  $\mathbb{S}_{i,j}(\mathcal{F}) = \mathcal{F}$  for all  $j < i$ , as a simple examination of the definitions reveals. Consider now the potential function  $\Phi$ , whose value on a family  $\mathcal{H}$  is the total sum of all elements appearing in all sets of  $\mathcal{H}$ . It is easy to check that if  $\mathbb{S}_{i,j}(\mathcal{F}) \neq \mathcal{F}$  for some  $j < i$  then  $\Phi(\mathbb{S}_{i,j}(\mathcal{H})) < \Phi(\mathcal{H})$ . Since  $\Phi$  is a non-negative integer, if we apply the shifts  $\mathbb{S}_{i,j}$  for all  $j < i$  repeatedly on  $\mathcal{F}$ , eventually we will reach a family whose potential is not decreased by any of the shifts, and this family is left-compressed.  $\square$

Lemma 2.9 shows that in order to determine  $w(n, t, p)$  it is enough to focus on left-compressed families. Moreover, since the up-set of a  $t$ -intersecting family is also  $t$ -intersecting, we will assume in most of what follows that  $\mathcal{F}$  is a monotone left-compressed  $t$ -intersecting family. We will show that except for the case  $p \geq 1/2$  and  $t = 1$ , such a family can only have maximum  $\mu_p$ -measure if it is a Frankl family with the correct parameters. We will deduce that general  $t$ -intersecting families of measure  $w(n, t, p)$  are equivalent to a Frankl family using the following lemma, whose proof closely follows the argument of Ahlswede and Khachatrian [3].

**Lemma 2.10.** *Let  $\mathcal{F}$  be a monotone  $t$ -intersecting family on  $n$  points, and let  $i, j \in [n]$ . If  $\mathbb{S}_{i,j}(\mathcal{F})$  is equivalent to  $\mathcal{F}_{t,r}$  then so is  $\mathcal{F}$ .*

*Proof.* Let  $S \subseteq [n]$  be the set of size  $t + 2r$  such that

$$\mathbb{S}_{i,j}(\mathcal{F}) = \{A \subseteq [n] : |A \cap S| \geq t + r\}.$$

Suppose first that  $i, j \in S$  or  $i, j \notin S$ . If  $A \in \mathbb{S}_{i,j}(\mathcal{F})$  then  $A \in \mathcal{F}$ , since otherwise  $A$  would have originated from  $A' = (A \setminus \{j\}) \cup \{i\}$ , but that is impossible since  $A' \in \mathbb{S}_{i,j}(\mathcal{F})$ . It follows that  $\mathbb{S}_{i,j}(\mathcal{F}) \subseteq \mathcal{F}$  and so  $\mathbb{S}_{i,j}(\mathcal{F}) = \mathcal{F}$ , since shifting preserves cardinality. Therefore the lemma trivially holds.

The case  $i \in S$  and  $j \notin S$  cannot happen. Indeed, consider some set  $A \subseteq S$  containing  $i$  but not  $j$  of size  $t + r$ . Then  $A \in \mathbb{S}_{i,j}(\mathcal{F})$  and so, by definition of the shift,  $A' = (A \setminus \{i\}) \cup \{j\} \in \mathbb{S}_{i,j}(\mathcal{F})$ . However,  $|A' \cap S| = t + r - 1$ , and so  $A' \notin \mathbb{S}_{i,j}(\mathcal{F})$ , and we reach a contradiction.

It remains to consider the case  $i \notin S$  and  $j \in S$ . Suppose first that  $r = 0$ . Then  $S \in \mathbb{S}_{i,j}(\mathcal{F})$ , and so either  $S \in \mathcal{F}$  or  $S' = (S \setminus \{j\}) \cup \{i\} \in \mathcal{F}$ . In both cases, since  $\mathcal{F}$  is monotone, it contains all supersets of  $S$  or of  $S'$ . Since shifting preserves cardinality,  $\mathcal{F}$  must consist exactly of all supersets of  $S$  or of  $S'$ , and thus is equivalent to a  $(t, 0)$ -Frankl family.

Suppose next that  $r > 0$ . Let  $V$  be the collection of all subsets of  $S \setminus \{j\}$  of size exactly  $t + r - 1$ . For each  $A \in V$  we have  $A \cup \{j\} \in \mathbb{S}_{i,j}(\mathcal{F})$ , and so either  $A \cup \{j\} \in \mathcal{F}$  or  $A \cup \{i\} \in \mathcal{F}$ .

If  $\mathcal{F}$  contains  $A \cup \{j\}$  for all  $A \in V$  then  $\mathcal{F}$  contains all subsets of  $S$  of size  $t + r$  (since other subsets are not affected by the shift). Monotonicity forces  $\mathcal{F}$  to contain all of  $\mathbb{S}_{i,j}(\mathcal{F})$ , and thus  $\mathcal{F} = \mathbb{S}_{i,j}(\mathcal{F})$  as before.

If  $\mathcal{F}$  contains  $A \cup \{i\}$  for all  $A \in V$ , then in a similar way we deduce that  $\mathcal{F}$  is equivalent to the  $(t, r)$ -Frankl family based on  $(S \setminus \{j\}) \cup \{i\}$ .

It remains to consider the case in which  $\mathcal{F}$  contains  $A \cup \{i\}$  for some  $A \in V$ , and  $B \cup \{j\}$  for some other  $B \in V$ . We will show that in this case,  $\mathcal{F}$  is not  $t$ -intersecting. Consider the graph on  $V$  in which two vertices are connected if their intersection has the minimal size  $t - 1$ . This graph is a generalized Johnson graph, and we show below that it is connected. This implies that there must be two sets  $A, B$  satisfying  $|A \cap B| = t - 1$  such that  $A \cup \{i\}, B \cup \{j\} \in \mathcal{F}$ . Since  $|A \cap B| = t - 1$ , we have reached a contradiction.

To complete the proof, we prove that the graph is connected. For reasons of symmetry, it is enough to give a path connecting  $x = \{1, \dots, t + r - 1\}$  and  $y = \{2, \dots, t + r\}$ . Indeed, the vertex  $\{2, \dots, t, t + r + 1, \dots, t + 2r\}$  is connected to both  $x$  and  $y$ .  $\square$

The preceding lemmas allow us to reduce the proof of Theorem 2.2 to the left-compressed case. The final lemma of this section abstracts this argument.

**Lemma 2.11.** *Fix  $n, t, p$ . If every monotone left-compressed  $t$ -intersecting family on  $n$  points has  $\mu_p$ -measure at most  $q$ , then every  $t$ -intersecting family on  $n$  points has  $\mu_p$ -measure at most  $q$ .*

*Suppose moreover that the only monotone left-compressed  $t$ -intersecting families on  $n$  points having  $\mu_p$ -measure  $q$  are  $\mathcal{F}_{t,r}$ , where  $r$  ranges over some set  $R$ . Then every  $t$ -intersecting family on  $n$  points having  $\mu_p$ -measure  $q$  is equivalent to a Frankl family  $\mathcal{F}_{t,r}$  for some  $r \in R$ .*

*Proof.* Suppose first that every monotone left-compressed  $t$ -intersecting family on  $n$  points has  $\mu_p$ -measure at most  $q$ , and let  $\mathcal{F}$  be a  $t$ -intersecting family on  $n$  points. Lemma 2.9 gives a left-compressed  $t$ -intersecting family  $\mathcal{G}$ , obtained from  $\mathcal{F}$  by applying a sequence of shifts, which has the same  $\mu_p$ -measure as  $\mathcal{F}$ . Let  $\mathcal{H}$  be the up-set of  $\mathcal{G}$ , which is a monotone left-compressed  $t$ -intersecting family. By assumption,  $\mu_p(\mathcal{H}) \leq q$ , and so  $\mu_p(\mathcal{F}) = \mu_p(\mathcal{G}) \leq \mu_p(\mathcal{H}) \leq q$ .

Suppose next that the only monotone left-compressed  $t$ -intersecting families on  $n$  points having  $\mu_p$ -measure  $q$  are  $\mathcal{F}_{t,r}$  for  $r \in R$ , and suppose that  $\mathcal{F}$  has  $\mu_p$ -measure  $q$ . In that case necessarily  $\mathcal{G} = \mathcal{H}$ , and so  $\mathcal{G} = \mathcal{F}_{t,r}$  for some  $r \in R$ , by assumption. Repeatedly applying Lemma 2.10 (starting with the last shifting applied and working backwards), we see that  $\mathcal{F}$  is equivalent to  $\mathcal{F}_{t,r}$ .  $\square$

## 2.4 Generating sets

The goal of the first part of the proof, which follows [3], is to show that any monotone left-compressed  $t$ -intersecting family of maximum  $\mu_p$ -measure has to depend on a small number of points. We will use a representation of monotone families in which this property has a simple manifestation. Our definition is simpler than the original one, due to the simpler setting.

**Definition 2.9.** A family  $\mathcal{F}$  on  $n$  points is *non-trivial* if  $\mathcal{F} \notin \{\emptyset, 2^{[n]}\}$ .

Let  $\mathcal{F}$  be a non-trivial monotone family. A *generating set* is a set  $S \in \mathcal{F}$  such that no subset of  $S$  belongs to  $\mathcal{F}$ . The *generating family* of  $\mathcal{F}$  consists of all generating sets of  $\mathcal{F}$ . The *extent* of  $\mathcal{F}$  is the maximal index appearing in a generating set of  $\mathcal{F}$ . The *boundary generating family* of  $\mathcal{F}$  consists of all generating sets of  $\mathcal{F}$  containing its extent.

If  $\mathcal{G}$  is the generating family of  $\mathcal{F}$  then we use the notation  $\mathcal{G}^*$  for the boundary generating family of  $\mathcal{F}$ . For each integer  $a$ , we use the notation  $\mathcal{G}_a^*$  for the subset of  $\mathcal{G}^*$  consisting of sets of size  $a$ .

Generating sets are also known as *minterns*. If  $\mathcal{G}$  is the generating family of  $\mathcal{F}$  then  $\mathcal{G}$  is an antichain and  $\mathcal{F}$  is the up-set of  $\mathcal{G}$  (and this gives an alternative definition of  $\mathcal{G}$ ). If  $\mathcal{F}$  has extent  $m$  then  $\mathcal{F}$  depends only on the first  $m$  coordinates:  $S \in \mathcal{F}$  iff  $S \Delta \{i\} \in \mathcal{F}$  for all  $i > m$ . For this reason, for the rest of the section we treat a family having extent  $m$  as a family on  $m$  points.

One reason to focus on the boundary generating family of  $\mathcal{F}$  is the following simple observations.

**Lemma 2.12.** *Let  $\mathcal{F}$  be a non-trivial monotone left-compressed family of extent  $m$  with generating family  $\mathcal{G}$  and boundary generating family  $\mathcal{G}^*$ . For any subset  $G \subseteq \mathcal{G}^*$ ,*

$$\langle \mathcal{G} \setminus G \rangle = \mathcal{F} \setminus G.$$

*Proof.* Since  $\mathcal{G}$  is an antichain, no  $A \in G$  is a superset of any other set in  $\mathcal{G}$ . For this reason,  $\langle \mathcal{G} \setminus G \rangle \subseteq \mathcal{F} \setminus G$ .

On the other hand, let  $S \in \mathcal{F} \setminus G$ . If  $S$  is not a superset of any  $A \in G$  then clearly  $S \in \langle \mathcal{G} \setminus G \rangle$ . If  $S \supseteq A$  for some  $A \in G$  then since  $S \neq A$ , there is an element  $i \in S \setminus A$ . The set  $S' = S \setminus \{m\}$  is a superset of  $(A \setminus \{m\}) \cup \{i\}$ , and so  $S' \in \mathcal{F}$ . Thus  $S'$  is a superset of some  $B \in \mathcal{G}$ . Since  $m \notin S'$ , necessarily  $B \notin G$ . As  $S \supseteq B$ , we conclude that  $S \in \langle \mathcal{G} \setminus G \rangle$ .  $\square$

**Lemma 2.13.** *Let  $\mathcal{F}$  be a non-trivial monotone left-compressed family of extent  $m$  with generating family  $\mathcal{G}$  and boundary generating family  $\mathcal{G}^*$ . For any subset  $G \subseteq \mathcal{G}^*$ ,*

$$\langle (\mathcal{G} \setminus G) \cup \{A \setminus \{m\} : A \in G\} \rangle = \mathcal{F} \cup \{A \setminus \{m\} : A \in G\}.$$

*Proof.* Denote by  $\mathcal{F}'$  the left-hand side. Clearly  $\mathcal{F}' \supseteq \mathcal{F} \cup \{A \setminus \{m\} : A \in G\}$ .

On the other hand, suppose that  $S \in \mathcal{F}' \setminus \mathcal{F}$ . Then for some  $A \in G$ ,  $S$  is a superset of  $A \setminus \{m\}$  but not of  $A$ . In particular,  $m \notin S$ . We claim that  $S = A \setminus \{m\}$ . Otherwise, there exists an element  $i \in S \setminus (A \setminus \{m\})$ . Since  $\mathcal{F}$  is left-compressed,  $A' = (A \setminus \{m\}) \cup \{i\} \in \mathcal{F}$ . Since  $\mathcal{F}$  is monotone and  $S \supseteq A'$ , we conclude that  $S \in \mathcal{F}$ , contradicting our assumption. Thus  $\mathcal{F}' \setminus \mathcal{F} = \{A \setminus \{m\} : A \in G\}$ .  $\square$

The following crucial observation drives the entire approach, and explains why we want to classify the sets in the boundary generating family according to their size.

**Lemma 2.14.** *Let  $\mathcal{F}$  be a non-trivial monotone left-compressed  $t$ -intersecting family with extent  $m$  and boundary generating family  $\mathcal{G}^*$ . If  $A, B \in \mathcal{G}^*$  intersect in exactly  $t$  elements then  $|A| + |B| = m + t$ .*

*Proof.* We will show that  $A \cup B = [m]$ . It follows that

$$|A| + |B| = |A \cup B| + |A \cap B| = m + t.$$

Since  $A \cup B \subseteq [m]$  and  $m \in A \cap B$  by definition, we have to show that every element  $i < m$  belongs to either  $A$  or  $B$ . Suppose that some element  $i$  belongs to neither. Since  $\mathcal{F}$  is left-compressed, the set  $B' = (B \setminus \{m\}) \cup \{i\}$  also belongs to  $\mathcal{F}$ . However,  $|A \cap B'| = |A \cap B| - 1 = t - 1$ , contradicting the assumption that  $\mathcal{F}$  is  $t$ -intersecting.  $\square$

Our goal now is to show that if  $m$  is too large then we can remove the dependency on  $m$  while keeping the family  $t$ -intersecting and increasing its  $\mu_p$ -measure, for appropriate values of  $p$ . The idea is to remove  $m$  from sets in the boundary generating family. The only obstructions for doing so are sets  $A, B$  in the boundary generating family whose intersection contains *exactly*  $t$  elements, and here we use Lemma 2.14 to guide us: this can only happen if  $|A| + |B| = m + t$ . Accordingly, our modification will involve generating sets in  $\mathcal{G}_a^*$  and  $\mathcal{G}_b^*$  for  $a + b = m + t$ . There are two cases to consider:  $a \neq b$  and  $a = b$ . The first case is simpler.

**Lemma 2.15.** *Let  $\mathcal{F}$  be a non-trivial monotone left-compressed  $t$ -intersecting family with extent  $m$ , generating family  $\mathcal{G}$ , and boundary generating family  $\mathcal{G}^*$ . Let  $a \neq b$  be parameters such that  $a + b = m + t$  and  $\mathcal{G}_a^*, \mathcal{G}_b^*$  are not both empty. Consider the families  $\mathcal{F}_1 = \langle \mathcal{G}_1 \rangle$  and  $\mathcal{F}_2 = \langle \mathcal{G}_2 \rangle$ , where*

$$\begin{aligned}\mathcal{G}_1 &= (\mathcal{G} \setminus (\mathcal{G}_a^* \cup \mathcal{G}_b^*)) \cup \{S \setminus \{m\} : S \in \mathcal{G}_b^*\}, \\ \mathcal{G}_2 &= (\mathcal{G} \setminus (\mathcal{G}_a^* \cup \mathcal{G}_b^*)) \cup \{S \setminus \{m\} : S \in \mathcal{G}_a^*\}.\end{aligned}$$

*Both families  $\mathcal{F}_1, \mathcal{F}_2$  are  $t$ -intersecting. Moreover, if  $p < 1/2$  then  $\max(\mu_p(\mathcal{F}_1), \mu_p(\mathcal{F}_2)) > \mu_p(\mathcal{F})$ ; and if  $p = 1/2$ ,  $\max(\mu_p(\mathcal{F}_1), \mu_p(\mathcal{F}_2)) \geq \mu_p(\mathcal{F})$ , with equality only if  $\mu_p(\mathcal{F}_1) = \mu_p(\mathcal{F}_2) = \mu_p(\mathcal{F})$ .*

*Proof.* We start by showing that  $\mathcal{F}_1$  (and so  $\mathcal{F}_2$ ) is  $t$ -intersecting. Clearly, it is enough to show that its generating family  $\mathcal{G}_1$  is  $t$ -intersecting. Suppose that  $S, T \in \mathcal{G}_1$ . We consider several cases.

If  $S, T \in \mathcal{G}$  then  $|S \cap T| \geq t$  since  $\mathcal{G}$  is  $t$ -intersecting.

If  $S \in \mathcal{G}$  and  $T \notin \mathcal{G}$  then  $T' = T \cup \{m\} \in \mathcal{G}$  and so  $|T'| = b$ . If  $m \notin S$  then  $|S \cap T| = |S \cap T'| \geq t$ . If  $m \in S$  then by construction  $|S| \neq a$ , and so  $|S \cap T| = |S \cap T'| - 1 \geq t$ , using Lemma 2.14.

If  $S, T \notin \mathcal{G}$  then  $S' = S \cup \{m\} \in \mathcal{G}$  and  $T' = T \cup \{m\} \in \mathcal{G}$ , and so  $|S'| = |T'| = b$ . As in the preceding case,  $|S \cap T| = |S' \cap T'| - 1 \geq t$  due to Lemma 2.14.

Lemma 2.12 and Lemma 2.13 show that

$$\mathcal{F}_1 = (\mathcal{F} \setminus \mathcal{G}_a^*) \cup \{S \setminus \{m\} : S \in \mathcal{G}_b^*\}.$$

Since  $\mu_p(S \setminus \{m\}) = \frac{1-p}{p} \mu_p(S)$  whenever  $m \in S$ ,

$$\mu_p(\mathcal{F}_1) = \mu_p(\mathcal{F}) - \mu_p(\mathcal{G}_a^*) + \frac{1-p}{p} \mu_p(\mathcal{G}_b^*).$$

Similarly,

$$\mu_p(\mathcal{F}_2) = \mu_p(\mathcal{F}) - \mu_p(\mathcal{G}_b^*) + \frac{1-p}{p} \mu_p(\mathcal{G}_a^*).$$

Taking the average of both estimates, we obtain

$$\frac{\mu_p(\mathcal{F}_1) + \mu_p(\mathcal{F}_2)}{2} = \mu_p(\mathcal{F}) + \frac{1}{2} \left( \frac{1-p}{p} - 1 \right) (\mu_p(\mathcal{G}_a^*) + \mu_p(\mathcal{G}_b^*)).$$

When  $p < 1/2$ , the second term is positive, and so  $\max(\mu_p(\mathcal{F}_1), \mu_p(\mathcal{F}_2)) > \mu_p(\mathcal{F})$ . When  $p = 1/2$  it vanishes, and so  $\max(\mu_p(\mathcal{F}_1), \mu_p(\mathcal{F}_2)) \geq \mu_p(\mathcal{F})$ .  $\square$

When  $a = b$  the construction in Lemma 2.15 cannot be executed, and we need a more complicated construction. The new construction will only work for small enough  $p$ , mirroring the fact that the optimal families for larger  $p$  depend on more points.

**Lemma 2.16.** *Let  $\mathcal{F}$  be a non-trivial monotone left-compressed  $t$ -intersecting family with extent  $m > 1$ , generating family  $\mathcal{G}$ , and boundary generating family  $\mathcal{G}^*$ . Suppose that  $a = \frac{m+t}{2}$  is an integer and that  $\mathcal{G}_a^*$  is non-empty. For each  $i \in [m-1]$ , let  $\mathcal{G}_{a,i}^* = \{S \in \mathcal{G}_a^* : i \in S\}$ , and define*

$$\mathcal{G}_i = (\mathcal{G} \setminus \mathcal{G}_a^*) \cup \{S \setminus \{m\} : S \in \mathcal{G}_a^* \setminus \mathcal{G}_{a,i}^*\}.$$

*All families  $\mathcal{F}_i = \langle \mathcal{G}_i \rangle$  are  $t$ -intersecting. Moreover, if  $p < \frac{m-t}{2(m-1)}$  then  $\mu_p(\mathcal{F}_i) > \mu_p(\mathcal{F})$  for some  $i \in [m-1]$ .*

*Proof.* We start by showing that the families  $\mathcal{F}_i$  are  $t$ -intersecting. Clearly, it is enough to show that  $\mathcal{G}_i$  is  $t$ -intersecting. Let  $S, T \in \mathcal{G}_i$ . We consider several cases.

If  $S, T \in \mathcal{G}$  then  $|S \cap T| \geq t$  since  $\mathcal{G}$  is  $t$ -intersecting.

If  $S \in \mathcal{G}$  and  $T \notin \mathcal{G}$  then  $T' = T \cup \{m\} \in \mathcal{G}_a^*$  and  $i \notin S$ . If  $m \notin S$  then  $|S \cap T| = |S \cap T'| \geq t$ . If  $m \in S$  and  $S \notin \mathcal{G}_a^*$  then  $|S \cap T| \geq |S \cap T'| - 1 \geq t$ , according to Lemma 2.14. If  $m \in S$  and  $S \in \mathcal{G}_a^*$  then by construction  $i \notin S$ . Since  $\mathcal{F}$  is left-compressed,  $S' = (S \setminus \{m\}) \cup \{i\} \in \mathcal{F}$ . Therefore  $|S \cap T| = |S' \cap T'| \geq t$ .

If  $S, T \notin \mathcal{G}$  then  $S' = S \cup \{m\}$  and  $T' = T \cup \{m\}$  both belong to  $\mathcal{G}_a^*$ , and  $i$  belongs to neither. Since  $\mathcal{F}$  is left-compressed,  $T'' = T \cup \{i\} \in \mathcal{F}$ , and so  $|S \cap T| = |S' \cap T''| \geq t$ .

Lemma 2.12 and Lemma 2.13 show that

$$\mathcal{F}_i = (\mathcal{F} \setminus \mathcal{G}_{a,i}^*) \cup \{S \setminus \{m\} : S \in \mathcal{G}_a^* \setminus \mathcal{G}_{a,i}^*\}.$$

Since  $\mu_p(S \setminus \{m\}) = \frac{1-p}{p} \mu_p(S)$  whenever  $m \in S$ ,

$$\begin{aligned} \mu_p(\mathcal{F}_i) &= \mu_p(\mathcal{F}) - \mu_p(\mathcal{G}_{a,i}^*) + \frac{1-p}{p} (\mu_p(\mathcal{G}_a^*) - \mu_p(\mathcal{G}_{a,i}^*)) \\ &= \mu_p(\mathcal{F}) + \frac{1-p}{p} \mu_p(\mathcal{G}_a^*) - \frac{1}{p} \mu_p(\mathcal{G}_{a,i}^*). \end{aligned}$$

Averaging over all  $i \in [m-1]$ , we obtain

$$\frac{1}{m-1} \sum_{i=1}^{m-1} \mu_p(\mathcal{F}_i) = \mu_p(\mathcal{F}) + \frac{1-p}{p} \mu_p(\mathcal{G}_a^*) - \frac{1}{p(m-1)} \sum_{i=1}^{m-1} \mu_p(\mathcal{G}_{a,i}^*).$$

Since the sets in  $\mathcal{G}_a^*$  contain exactly  $a$  elements, each set is counted  $a-1$  times in  $\sum_{i=1}^{m-1} \mu_p(\mathcal{G}_{a,i}^*)$ , and so

$$\frac{1}{m-1} \sum_{i=1}^{m-1} \mu_p(\mathcal{F}_i) = \mu_p(\mathcal{F}) + \left( \frac{1-p}{p} - \frac{a-1}{p(m-1)} \right) \mu_p(\mathcal{G}_a^*).$$

When  $1-p > \frac{a-1}{m-1} = \frac{m+t-2}{2(m-1)}$ , the bracketed quantity is positive, and so  $\max_i \mu_p(\mathcal{F}_i) > \mu_p(\mathcal{F})$ .  $\square$

## 2.5 Pushing-pulling

The goal in the second part of the proof, which follows [5], is to show that any monotone left-compressed  $t$ -intersecting family of maximum  $\mu_p$ -measure is symmetric within its extent, or in other words, of the form  $\mathcal{F}_{t,r}$ .

The analog of extent in this part is the *symmetric extent*.

**Definition 2.10.** Let  $\mathcal{F}$  be a left-compressed family on  $n$  points. Its *symmetric extent* is the largest integer  $\ell$  such that  $\mathcal{S}_{ij}(\mathcal{F}) = \mathcal{F}$  for  $i, j \leq \ell$ .

If  $\ell < n$  then the *boundary* of  $\mathcal{F}$  is the collection

$$\mathcal{X} = \{A \in \mathcal{F} : \ell+1 \notin A \text{ and } (A \setminus \{i\}) \cup \{\ell+1\} \text{ for some } i \in A \cap [\ell]\}.$$

In other words,  $\mathcal{X}$  consists of those sets in  $\mathcal{F}$  preventing it from having larger symmetric extent.

The definition of symmetric extension guarantees that  $\mathcal{X}$  can be decomposed as

$$\mathcal{X} = \sum_{a=0}^{\ell} \binom{[\ell]}{a} \times \mathcal{X}_a,$$

where  $\mathcal{X}_a$  is a collection of subsets of  $[n] \setminus [\ell+1]$ , a notation we use below.

The symmetric extent of a family is always bounded by its extent, apart from one trivial case.

**Lemma 2.17.** *Let  $\mathcal{F}$  be a non-trivial monotone family on  $n$  points having extent  $m$  and symmetric extent  $\ell$ . Then  $\ell \leq m$ .*

*Proof.* The family  $\mathcal{F}$  has the general form

$$\mathcal{F} = \bigcup_{i=0}^{\ell} \binom{[\ell]}{i} \times \mathcal{F}_i,$$

where  $\mathcal{F}_1, \dots, \mathcal{F}_\ell$  are collections of subsets of  $[n] \setminus [\ell]$ . We claim that if  $m < \ell$  then all  $\mathcal{F}_i$  are equal. Indeed, let  $i < \ell$ . For each  $A \in \mathcal{F}_i$ , we have  $[i] \cup A \in \mathcal{F}$ . Since the extent of  $\mathcal{F}$  is smaller than  $\ell$ ,  $[i] \cup \{\ell\} \cup A \in \mathcal{F}$ , implying  $A \in \mathcal{F}_{i+1}$ . Similarly, for each  $A \in \mathcal{F}_{i+1}$  we have  $[i] \cup \{\ell\} \cup A \in \mathcal{F}$ , and so  $[i] \cup \{\ell\} \in \mathcal{F}$ , implying  $A \in \mathcal{F}_i$ .

We have shown that  $\mathcal{F} = 2^{[\ell]} \times \mathcal{F}_0$ . Since the extent of  $m$  is at most  $\ell$ , necessarily  $\mathcal{F} = 2^{[n]}$ .  $\square$

The following crucial observation is the counterpart of Lemma 2.14.

**Lemma 2.18.** *Let  $\mathcal{F}$  be a left-compressed  $t$ -intersecting family of symmetric extent  $\ell$  and boundary  $\mathcal{X}$ . If  $|A \cap B| = t$  for some  $A, B \in \mathcal{X}$  then  $|A \cap [\ell]| + |B \cap [\ell]| = \ell + t$ .*

*Proof.* We start by showing that  $A \cap B \subseteq [\ell]$ . Indeed, suppose that  $i \in A \cap B$  for some  $i > \ell$ . Since neither of  $A, B$  contains  $\ell + 1$ , in fact  $i > \ell + 1$ . Since  $\mathcal{F}$  is left-compressed,  $A' = (A \setminus \{i\}) \cup \{\ell + 1\} \in \mathcal{F}$ . However,  $|A' \cap B| = |A \cap B| - 1 = t - 1$ , contradicting the assumption that  $\mathcal{F}$  is  $t$ -intersecting.

Next, we show that  $A \cup B \supseteq [\ell]$ . Indeed, suppose that  $i \notin A \cup B$  for some  $i \in \ell$ . By definition of  $\mathcal{X}$ , the set  $A$  must contain some element  $j \in [\ell]$ . By definition of symmetric extent (if  $j < i$ ) or by the fact that  $\mathcal{F}$  is left-compressed (if  $j > i$ ),  $A' = (A \setminus \{j\}) \cup \{i\} \in \mathcal{F}$ . However,  $|A' \cap B| = |A \cap B| - 1 = t - 1$ , contradicting the assumption that  $\mathcal{F}$  is  $t$ -intersecting.

Finally, let  $A' = A \cap [\ell]$  and  $B' = B \cap [\ell]$ . Since  $A' \cap B' = A \cap B$  and  $A' \cup B' = [\ell]$ , we deduce that

$$|A'| + |B'| = |A' \cup B'| + |A' \cap B'| = \ell + t. \quad \square$$

Our goal now is to try to eliminate  $\mathcal{X}$ , thus increasing the symmetric extent. We do this by trying to add  $\binom{[\ell]}{a-1} \times \{\ell + 1\} \times \mathcal{X}_a$  to  $\mathcal{F}$ . The obstructions are described by Lemma 2.18, which explains why we decompose  $\mathcal{X}$  according to the size of the intersection with  $[\ell]$ . Accordingly, our modification will act on the sets in  $\mathcal{X}_a, \mathcal{X}_b$  for  $a + b = \ell + t$ . As in the preceding section, we have to consider two cases,  $a \neq b$  and  $a = b$ , and the first case is simpler.

**Lemma 2.19.** *Let  $\mathcal{F}$  be a left-compressed  $t$ -intersecting family on  $n$  points of symmetric extent  $\ell < n$ . Let  $a \neq b$  be parameters such that  $a + b = \ell + t$  and  $\mathcal{X}_a, \mathcal{X}_b$  are not both empty. Consider the two families*

$$\begin{aligned} \mathcal{F}_1 &= \left( \mathcal{F} \setminus \binom{[\ell]}{b} \times \mathcal{X}_b \right) \cup \binom{[\ell]}{a-1} \times \{\ell + 1\} \times \mathcal{X}_a, \\ \mathcal{F}_2 &= \left( \mathcal{F} \setminus \binom{[\ell]}{a} \times \mathcal{X}_a \right) \cup \binom{[\ell]}{b-1} \times \{\ell + 1\} \times \mathcal{X}_b. \end{aligned}$$

*Both families  $\mathcal{F}_1, \mathcal{F}_2$  are  $t$ -intersecting. Moreover, if  $t > 1$  then for all  $p \in (0, 1)$ ,  $\max(\mu_p(\mathcal{F}_1), \mu_p(\mathcal{F}_2)) > \mu_p(\mathcal{F})$ .*

*Proof.* We start by showing that  $\mathcal{F}_1$  (and so  $\mathcal{F}_2$ ) is  $t$ -intersecting. Suppose that  $S, T \in \mathcal{F}_1$ . We consider several cases.

If  $S, T \in \mathcal{F}$  then  $|S \cap T| \geq t$  since  $\mathcal{F}$  is  $t$ -intersecting.

If  $S \in \mathcal{F}$  and  $T \notin \mathcal{F}$  then  $T \in \binom{[\ell]}{a-1} \times \{\ell + 1\} \times \mathcal{X}_a$ . Choose  $i \in [\ell] \setminus T$  arbitrarily, and notice that  $T' = (T \setminus \{\ell + 1\}) \cup \{i\} \in \binom{[\ell]}{a} \times \mathcal{X}_a$ , and so  $T' \in \mathcal{F}$ . If  $i \notin S$  or  $\ell + 1 \in S$  then  $|S \cap T| \geq |S \cap T'| \geq t$ . Suppose therefore that  $i \in S$  and  $\ell + 1 \notin S$ . If  $S' = (S \setminus \{i\}) \cup \{\ell + 1\} \in \mathcal{F}$  then  $|S \cap T| = |S' \cap T'| \geq t$ . Otherwise, by definition of  $\mathcal{X}$ ,  $S \in \mathcal{X}$ . By definition of  $\mathcal{F}_1$ ,  $|S \cap [\ell]| \neq b$ , and so Lemma 2.18 shows that  $|S \cap T| \geq |S \cap T'| - 1 \geq t$ .

If  $S, T \notin \mathcal{F}$  then  $S, T \in \binom{[\ell]}{a-1} \times \{\ell + 1\} \times \mathcal{X}_a$ . Choose  $i \in [\ell] \setminus S$  and  $j \in [\ell] \setminus T$  arbitrarily, and define  $S' = (S \setminus \{\ell + 1\}) \cup \{i\}$  and  $T' = (T \setminus \{\ell + 1\}) \cup \{j\}$ . As before,  $S', T' \in \mathcal{X}$ , and so Lemma 2.18 shows that  $|S' \cap T'| \geq t + 1$ . Since  $S \cap T \supseteq ((S' \cap T') \setminus \{i, j\}) \cup \{\ell + 1\}$ , we see that  $|S \cap T| \geq |S' \cap T'| - 1 \geq t$ .

We calculate the measures of  $\mathcal{F}_1$  and  $\mathcal{F}_2$  in terms of the quantities  $m_a = \mu_p\left(\binom{[\ell]}{a} \times \mathcal{X}_a\right)$  and  $m_b = \mu_p\left(\binom{[\ell]}{b} \times \mathcal{X}_b\right)$ :

$$\begin{aligned} \mu_p(\mathcal{F}_1) &= \mu_p(\mathcal{F}) - m_b + \frac{\binom{\ell}{a-1}}{\binom{\ell}{a}} m_a = \mu_p(\mathcal{F}) - m_b + \frac{a}{\ell - a + 1} m_a, \\ \mu_p(\mathcal{F}_2) &= \mu_p(\mathcal{F}) - m_a + \frac{\binom{\ell}{b-1}}{\binom{\ell}{b}} m_b = \mu_p(\mathcal{F}) - m_a + \frac{b}{\ell - b + 1} m_b. \end{aligned}$$

Multiply the first inequality by  $\frac{\ell - a + 1}{\ell - t + 2}$ , the second inequality by  $\frac{\ell - b + 1}{\ell - t + 2}$ , and add; note that  $\ell - t + 2 = (\ell - a + 1) + (\ell - b + 1) > 0$ . The result is

$$\begin{aligned} \frac{\ell - a + 1}{\ell - t + 2} \mu_p(\mathcal{F}_1) + \frac{\ell - b + 1}{\ell - t + 2} \mu_p(\mathcal{F}_2) &= \mu_p(\mathcal{F}) + \left[ \frac{a}{\ell - t + 2} - \frac{\ell - b + 1}{\ell - t + 2} \right] m_a + \left[ \frac{b}{\ell - t + 2} - \frac{\ell - a + 1}{\ell - t + 2} \right] m_b \\ &= \mu_p(\mathcal{F}) + \frac{t - 1}{\ell - t + 2} (m_a + m_b), \end{aligned}$$

using  $a + b = \ell + t$ . We conclude that when  $t > 1$ ,  $\max(\mu_p(\mathcal{F}_1), \mu_p(\mathcal{F}_2)) > \mu_p(\mathcal{F})$ .  $\square$

When  $a = b$ , the construction increases the extent  $m$  (defined in the preceding section), and works only for large enough  $p$ .

**Lemma 2.20.** *Let  $\mathcal{F}$  be a non-trivial monotone left-compressed  $t$ -intersecting family on  $n$  points of extent  $m < n$  and symmetric extent  $\ell$ , and let  $s \in [n]$  be an index satisfying  $s > m$  and  $s \neq \ell + 1$  (such an element exists if  $\ell < m$  or if  $m \leq n - 2$ ). Suppose that  $a = \frac{\ell+t}{2}$  is an integer and that  $\mathcal{X}_a$  is non-empty. Let  $\mathcal{X}'_a = \{S \in \mathcal{X}_a : s \in S\}$  and define*

$$\mathcal{F}' = \left( \mathcal{F} \setminus \binom{[\ell]}{a} \times \mathcal{X}_a \right) \cup \binom{[\ell+1]}{a} \times \mathcal{X}'_a.$$

The family  $\mathcal{F}'$  is  $t$ -intersecting. Moreover, if  $p > \frac{\ell-t+2}{2(\ell+1)}$  then  $\mu_p(\mathcal{F}') > \mu_p(\mathcal{F})$ .

*Proof.* We start by showing that  $\mathcal{F}'$  is  $t$ -intersecting. Suppose that  $S, T \in \mathcal{F}'$ . We consider several cases.

If  $S, T \in \mathcal{F}$  then  $|S \cap T| \geq t$  since  $\mathcal{F}$  is  $t$ -intersecting.

If  $S \in \mathcal{F}$  and  $T \notin \mathcal{F}$  then  $T \in \binom{[\ell+1]}{a} \times \mathcal{X}'_a$  and  $\ell + 1, s \in T$ . Choose  $i \in [\ell] \setminus T$  arbitrarily, and notice that  $T' = (T \setminus \{\ell + 1\}) \cup \{i\} \in \binom{[\ell]}{a} \times \mathcal{X}_a \in \mathcal{F}$ . If  $i \notin S$  or  $\ell + 1 \in S$  then  $|S \cap T| \geq |S \cap T'| \geq t$ . Suppose therefore that  $i \in S$  and  $\ell + 1 \notin S$ . If  $S' = (S \setminus \{i\}) \cup \{\ell + 1\} \in \mathcal{F}$  then  $|S \cap T| = |S' \cap T'| \geq t$ . Otherwise,  $S \in \mathcal{X}$ . If  $|S \cap [\ell]| \neq a$  then Lemma 2.18 shows that  $|S \cap T| \geq |S \cap T'| - 1 \geq t$ . If  $|S \cap [\ell]| = a$  then by construction,  $s \in S$ . Since the extent of  $\mathcal{F}$  is  $m < s$ , also  $S' = S \setminus \{s\} \in \mathcal{F}$ . Therefore  $|S \cap T| \geq |S' \cap T'| \geq t$ , since  $s \in S \cap T$  but  $s \notin S'$ .

If  $S, T \notin \mathcal{F}$  then  $S, T \in \binom{[\ell+1]}{a} \times \mathcal{X}'_a$  and  $\ell + 1, s \in S, T$ . Choose  $i \in [\ell] \setminus S$  and  $j \in [\ell] \setminus T$ , so that  $S' = (S \setminus \{\ell + 1\}) \cup \{i\}$  and  $T' = (T \setminus \{\ell + 1\}) \cup \{j\}$  are both in  $\mathcal{F}$ . By construction,  $s$  belongs to  $S$  and  $T$  and so to  $S'$  and  $T'$ . Since the extent of  $\mathcal{F}$  is  $m < s$ ,  $S'' = S \setminus \{s\}$  and  $T'' = T \setminus \{s\}$  also belong to  $\mathcal{F}$ . Observe that  $S \cap T \subseteq ((S'' \cap T'') \setminus \{i, j\}) \cup \{\ell + 1, s\}$ , and so  $|S \cap T| \geq |S'' \cap T''| \geq t$ .

We calculate the measure of  $\mathcal{F}'$  in terms of the quantity  $m_a = \mu_p\left(\binom{[\ell]}{a} \times \mathcal{X}_a\right)$ :

$$\begin{aligned} \mu_p(\mathcal{F}') &= \mu_p(\mathcal{F}) - m_a + p \frac{\binom{\ell+1}{a}}{\binom{\ell}{a}} m_a \\ &= \mu_p(\mathcal{F}) - m_a + p \frac{\ell+1}{\ell+1-a} m_a \\ &= \mu_p(\mathcal{F}) + \frac{a - (1-p)(\ell+1)}{\ell+1-a} m_a. \end{aligned}$$

Thus  $\mu_p(\mathcal{F}') > \mu_p(\mathcal{F})$  as long as  $1 - p < \frac{a}{\ell+1} = \frac{\ell+t}{2(\ell+1)}$ .  $\square$

Lemma 2.20 cannot be applied when  $m = n$ . However, if  $n$  has the correct parity, we can combine Lemma 2.20 with Lemma 2.15 to handle this issue.

**Lemma 2.21.** *Let  $\mathcal{F}$  be a non-trivial monotone left-compressed  $t$ -intersecting family on  $n$  points of extent  $m$  and symmetric extent  $\ell$ , where either  $\ell < m$  or  $m < n$ . If  $n + t$  is even and  $\frac{\ell-t+2}{2(\ell+1)} < p \leq \frac{1}{2}$  then there exists a  $t$ -intersecting family on  $n$  points with larger  $\mu_p$ -measure.*

*Proof.* Consider first the case  $\ell < m$ . If  $m < n$  then the statement follows from Lemma 2.19 and Lemma 2.20, so suppose that  $m = n$ . Let  $\mathcal{F}' = \mathcal{F} \cup \mathcal{F} \times \{n+1\}$ , and note that this is a non-trivial monotone left-compressed  $t$ -intersecting family on  $n+1$  points. We can apply Lemma 2.20 to obtain a non-trivial monotone left-compressed  $t$ -intersecting family  $\mathcal{G}$  on  $n+1$  points satisfying  $\mu_p(\mathcal{G}) > \mu_p(\mathcal{F}') = \mu_p(\mathcal{F})$ . Since  $n+1+t$  is odd, we can apply Lemma 2.15 repeatedly to obtain a non-trivial monotone  $t$ -intersecting family  $\mathcal{H}$  on  $n+1$  points and extent  $n$  which satisfies  $\mu_p(\mathcal{H}) \geq \mu_p(\mathcal{G}) > \mu_p(\mathcal{F})$ . Since  $\mathcal{H}$  has extent  $n$ , there is a  $t$ -intersecting family on  $n$  points having the same  $\mu_p$ -measure.

Consider next the case  $\ell = m < n$ . If  $m \leq n - 2$  then the statement follows from Lemma 2.19 and Lemma 2.20, so suppose that  $m = n - 1$ . In this case  $m + t$  is odd, and so the statement follows from Lemma 2.15.  $\square$

## 2.6 The case $p < 1/2$

In this section we prove Theorem 2.2 in the case  $p < 1/2$ . In view of Lemma 2.11, it suffices to consider monotone left-compressed families.

We first settle the case  $t = 1$ , which corresponds to the classical Erdős–Ko–Rado theorem (in the weighted setting).

**Lemma 2.22.** *Let  $\mathcal{F}$  be a monotone left-compressed intersecting family on  $n$  points of maximum  $\mu_p$ -measure, for some  $p \in (0, 1/2)$ . Then  $\mathcal{F} = \mathcal{F}_{1,0}$ .*

*Proof.* Let  $m$  be the extent of  $\mathcal{F}$ . Since  $\frac{m-t}{2(m-1)} = 1/2$ , Lemma 2.15 and Lemma 2.16 together show that  $m = 1$ , and so  $\mathcal{F} = \mathcal{F}_{1,0}$ .  $\square$

The case  $t \geq 2$  requires more work.

**Lemma 2.23.** *Let  $\mathcal{F}$  be a monotone left-compressed  $t$ -intersecting family on  $n$  points of maximum  $\mu_p$ -measure, for some  $p \in (0, 1/2)$  and  $t > 1$ . Let  $r$  be the maximal integer satisfying  $p \geq \frac{r}{t+2r-1}$  and  $t + 2r \leq n$ . If  $p \neq \frac{r}{t+2r-1}$  then  $\mathcal{F} = \mathcal{F}_{t,r}$ , and if  $p = \frac{r}{t+2r-1}$  then  $\mathcal{F} \in \{\mathcal{F}_{t,r}, \mathcal{F}_{t,r-1}\}$ .*

*Proof.* Our definition of  $r$  guarantees that one of the following two alternatives holds: either  $p < \frac{r+1}{t+2r-1}$ , or  $n \leq t + 2r + 1$ .

Let  $m$  be the extent of  $\mathcal{F}$ . We claim that  $m \leq t + 2r$ . If  $n \leq t + 2r + 1$  then Lemma 2.15 shows that  $m + t$  is even, and so  $m \leq t + 2r$ . Suppose therefore that  $p < \frac{r+1}{t+2r-1}$  and  $m > t + 2r$ . Lemma 2.15 shows that in fact  $m \geq t + 2r + 2$ , and so

$$\frac{m-t}{2(m-1)} = \frac{1}{2} - \frac{t-1}{2(m-1)} \geq \frac{1}{2} - \frac{t-1}{2(t+2r+1)} = \frac{r+1}{t+2r+1}.$$

Therefore Lemma 2.15 and Lemma 2.16 contradict the assumption that  $\mathcal{F}$  has maximum  $\mu_p$ -measure.

We now turn to consider the symmetric extent  $\ell$  of  $\mathcal{F}$ . We first consider the case in which  $p > \frac{r}{t+2r-1}$ . We claim that in this case  $\ell = m$ . If  $\ell < m$  then Lemma 2.15 and Lemma 2.19 show that both  $m + t$  and  $\ell + t$  are even, and so  $\ell \leq m - 2 \leq t + 2r - 2$ . This implies that

$$\frac{\ell - t + 2}{2(\ell + 1)} = \frac{1}{2} - \frac{t-1}{2(\ell+1)} \leq \frac{1}{2} - \frac{t-1}{2(t+2r-1)} = \frac{r}{t+2r-1}.$$

Therefore Lemma 2.19 and Lemma 2.21 contradict the assumption that  $\mathcal{F}$  has maximum  $\mu_p$ -measure.

We have shown that if  $p > \frac{r}{t+2r-1}$  then  $\ell = m \leq t + 2r$ , and moreover  $m + t$  is even. Thus  $\ell = m = t + 2s$  for some  $s \leq r$ . Since  $\mathcal{F}$  is  $t$ -intersecting,  $\mathcal{F} \subseteq \mathcal{F}_{t,s}$  for some  $s \leq r$ . The fact that  $\mathcal{F}$  has maximum  $\mu_p$ -measure forces  $\mathcal{F} = \mathcal{F}_{t,s}$ . In view of Lemma 2.5, necessarily  $s = r$ .

The case  $p = \frac{r}{t+2r-1}$  is slightly more complicated. Suppose first that  $\ell = m$ . In that case, as before,  $\mathcal{F} = \mathcal{F}_{t,s}$  for some  $s \leq r$ . This time Lemma 2.5 shows that  $s \in \{r, r-1\}$ .

Suppose next that  $\ell < m$ . The same argument as before shows that  $\ell \geq m - 2$ . Lemma 2.15 and Lemma 2.19 show that both  $\ell + t$  and  $m + t$  are even, and so  $\ell = m - 2$  in this case. In the remainder of the proof, we show that this leads to a contradiction. To simplify notation, we will assume that  $m = n$ . As  $m + t$  is even, we can write  $m = t + 2s$  for some  $s \leq r$ .

Since  $\mathcal{F}$  is monotone and has symmetric extent  $m - 2$ , it can be decomposed as follows:

$$\begin{aligned} \mathcal{F} = & \binom{[t+2s-2]}{\geq a} \cup \binom{[t+2s-2]}{\geq b} \times \{t+2s+1\} \cup \\ & \binom{[t+2s-2]}{\geq c} \times \{t+2s\} \cup \binom{[t+2s-2]}{\geq d} \times \{t+2s-1, t+2s\}. \end{aligned}$$

Since the family is  $t$ -intersecting, we must have  $2d - (t + 2s - 2) + 2 \geq t$ , and so  $d \geq t + s - 2$ . If  $d \geq t + s - 1$  then monotonicity implies that  $a, b, c \geq d \geq t + s - 1$ , and so  $\mathcal{F} \subseteq \mathcal{F}_{t,s-1}$ . Since  $\mathcal{F}$  has maximum  $\mu_p$ -measure, necessarily  $\mathcal{F} = \mathcal{F}_{t,s-1}$ , in which case the extent is  $t + 2s - 2$ , contrary to assumption.

We conclude that  $d = t + s - 2$ . The fact that  $\mathcal{F}$  is  $t$ -intersecting implies that  $c + d - (t + 2s - 2) + 1 \geq t$ , and so  $c \geq t + s - 1$ . Similarly  $b \geq t + s - 1$ , and moreover  $a + d - (t + 2s - 2) \geq t$ , implying  $a \geq t + s$ . Thus  $\mathcal{F} \subseteq \mathcal{F}_{t,s}$ . Since  $\mathcal{F}$  has maximum  $\mu_p$ -measure, necessarily  $\mathcal{F} = \mathcal{F}_{t,s}$ , in which case the symmetric extent is  $t + 2s$ , contrary to assumption.  $\square$

## 2.7 The case $p = 1/2$

In this section we prove Theorem 2.2 in the case  $p = 1/2$ . This case is known as Katona's theorem, after Katona's paper [21]. In view of Lemma 2.11, it suffices to consider monotone left-compressed families.

We first settle the case  $t = 1$ .

**Lemma 2.24.** *If  $\mathcal{F}$  is intersecting then  $\mu_{1/2}(\mathcal{F}) \leq 1/2$ .*

*Proof.* Since  $\mathcal{F}$  is intersecting, for any set  $A \subseteq [n-1]$  it can contain either  $A$  or  $[n] \setminus A$ , but not both. Thus  $\mathcal{F}$  contains at most  $2^{n-1}$  sets, implying  $\mu_{1/2}(\mathcal{F}) \leq 1/2$ .  $\square$

The case  $t \geq 2$  requires more work.

**Lemma 2.25.** *Let  $\mathcal{F}$  be a monotone left-compressed  $t$ -intersecting family on  $n$  points of maximum  $\mu_{1/2}$ -measure, for some  $t > 1$ . If  $n \in \{t+2r, t+2r+1\}$  then  $\mathcal{F} = \mathcal{F}_{t,r}$ .*

*Proof.* Let  $m$  be the extent of  $\mathcal{F}$ , and  $\ell$  be its symmetric extent.

Suppose first that  $n = t+2r$ . Lemma 2.21 shows that  $\ell = m = n$ , which easily implies  $\mathcal{F} = \mathcal{F}_{t,r}$ .

Suppose next that  $n = t+2r+1$ . If  $m \leq t+2r$  then the previous case  $n = t+2r$  shows that  $\mathcal{F} = \mathcal{F}_{t,r}$ , so suppose that  $m = n$ . Since  $m+t$  is odd, Lemma 2.15 shows that there is a family  $\mathcal{H}$  of extent at most  $m-1 = t+2r$  such that  $\mu_p(\mathcal{H}) \geq \mu_p(\mathcal{F})$ . In view of the preceding case, this shows that  $\mathcal{H} = \mathcal{F}_{t,r}$ , and so  $\mu_p(\mathcal{F}) \leq \mu_p(\mathcal{F}_{t,r})$ . It remains to show that  $\mathcal{F} = \mathcal{F}_{t,r}$  when  $\mu_p(\mathcal{F}) = \mu_p(\mathcal{F}_{t,r})$ .

The family  $\mathcal{H}$  is constructed by repeatedly applying the following operation, where  $\mathcal{G}^*$  is the boundary generating family of  $\mathcal{F}$ , and  $a+b = n+t$ : remove  $\mathcal{G}_a^*$  and  $\mathcal{G}_b^*$ , and add either  $\{S \setminus \{m\} : S \in \mathcal{G}_a^*\}$  or  $\{S \setminus \{m\} : S \in \mathcal{G}_b^*\}$ . All options must lead eventually to the same family  $\mathcal{F}_{t,r}$ , and this can only happen if  $\mathcal{G}_a^* = \mathcal{G}_b^* = \emptyset$  for all  $a, b$ . However, in that case the extent of  $\mathcal{F}$  is in fact  $n-1$ , contradicting our assumption.  $\square$

Ahlswede and Khachatryan [6] give a different shifting proof of this theorem (without characterizing the equality cases).

## 2.8 The case $p > 1/2$

In this section we prove Theorem 2.2 in the case  $p > 1/2$ . The proof in this case differs from that of the other cases: it uses a different shifting argument, also due to Ahlswede and Khachatryan [6], who used it for the case  $p = 1/2$ .

The idea is to use a different kind of shifting.

**Definition 2.11.** Let  $\mathcal{F}$  be a family on  $n$  points. For two disjoint sets  $A, B \subseteq [n]$ , the shift operator  $\mathbb{S}_{A,B}$  acts on  $\mathcal{F}$  as follows. Let  $\mathcal{F}_{A,B}$  consist of all sets in  $\mathcal{F}$  containing  $A$  and disjoint from  $B$ . Then

$$\begin{aligned} \mathbb{S}_{A,B}(\mathcal{F}) = & (F \setminus \mathcal{F}_{A,B}) \cup \{S : S \in \mathcal{F}_{A,B} \text{ and } (S \setminus A) \cup B \in \mathcal{F}\} \\ & \cup \{(S \setminus A) \cup B : S \in \mathcal{F}_{A,B} \text{ and } (S \setminus A) \cup B \notin \mathcal{F}\}. \end{aligned}$$

(This is a generalization of the original shifting operator:  $\mathbb{S}_{i,j}$  is the same as  $\mathbb{S}_{\{i\},\{j\}}$ .)

A family is  $(s, s+1)$ -stable if  $\mathbb{S}_{A,B}(\mathcal{F}) = \mathcal{F}$  for any disjoint sets  $A, B$  of cardinalities  $|A| = s$  and  $|B| = s+1$ .

This kind of shift is useful when  $p > 1/2$  for the following simple reason.

**Lemma 2.26.** *If  $|B| > |A|$  then  $\mu_p(\mathbb{S}_{A,B}(\mathcal{F})) \geq \mu_p(\mathcal{F})$  for any  $p \in (1/2, 1)$ , with equality if only if  $\mathbb{S}_{A,B}(\mathcal{F}) = \mathcal{F}$ .*

*Proof.* If  $\mathbb{S}_{A,B}(\mathcal{F}) = \mathcal{F}$  then clearly both sets have the same measure. Otherwise, the shifting replaces certain sets  $S$  with the corresponding  $S' = (S \setminus A) \cup B$ . Since  $|S'| > |S|$  and  $p > 1/2$ , this increases the measure of the family.  $\square$

When done correctly,  $\mathbb{S}_{A,B}$  preserves the property of being  $t$ -intersecting.

**Lemma 2.27.** *Let  $\mathcal{F}$  be a  $t$ -intersecting family on  $n$  points, and let  $A, B \subseteq [n]$  be disjoint sets of cardinalities  $|A| = s$  and  $|B| = s+1$ . If  $\mathcal{F}$  is  $(r, r+1)$ -stable for all  $r < s$  then  $\mathbb{S}_{A,B}(\mathcal{F})$  is  $t$ -intersecting as well.*

*Proof.* If  $s = 0$  then  $\mathbb{S}_{A,B}(\mathcal{F}) \subseteq \langle \mathcal{F} \rangle$ , and so  $\mathbb{S}_{A,B}(\mathcal{F})$  is  $t$ -intersecting. Suppose therefore that  $s \geq 1$ . Let  $S, T \in \mathbb{S}_{A,B}(\mathcal{F})$ . We consider three cases.

**Case 1:** If  $S, T \in \mathcal{F}$  then  $|S \cap T| \geq t$  since  $\mathcal{F}$  is  $t$ -intersecting.

**Case 2:** If  $S, T \notin \mathcal{F}$  then  $S' = (S \setminus B) \cup A$  and  $T' = (T \setminus B) \cup A$  are both in  $\mathcal{F}$ , and so  $|S \cap T| = |S' \cap T'| - |A| + |B| \geq t + 1$ .

**Case 3:** The remaining case is when  $S \in \mathcal{F}$  and  $T \notin \mathcal{F}$ . In this case  $T' = (T \setminus B) \cup A \in \mathcal{F}$ . Since  $|S \cap T| = |S \cap T'| - |S \cap A| + |S \cap B| \geq t - |S \cap A| + |S \cap B|$ , if  $|S \cap B| \geq |S \cap A|$  then  $|S \cap T| \geq t$ . Suppose therefore that  $|S \cap B| < |S \cap A|$ .

Let  $r = \min(|S \cap A|, |B \setminus S| - 1)$ . Since  $|B \setminus S| = |B| - |S \cap B| > |B| - |S \cap A| \geq 1$ , we see that  $r \geq 0$ . Also, clearly  $r \leq s$ . If  $r = s$  then  $A \subseteq S$  whereas  $S \cap B = \emptyset$ . Therefore  $S' = (S \setminus A) \cup B \in \mathcal{F}$ , since otherwise  $S$  would have been replaced by  $S'$ . Then  $|S \cap T| = |S' \cap T'| \geq t$ .

Suppose therefore that  $r < s$ . Let  $C$  be an arbitrary subset of  $S \cap A$  of size  $r$ , and let  $D$  be an arbitrary subset of  $B \setminus S$  of size  $r + 1$ . Since  $\mathcal{F}$  is  $r$ -stable,  $S' = (S \setminus C) \cup D \in \mathcal{F}$ . Write

$$\begin{aligned} S' &= ((S \cap A) \setminus C) \cup (S \setminus A) \cup D, \\ T' &= (T \setminus B) \cup (A \cap S) \cup (A \setminus S). \end{aligned}$$

The term  $(S \cap A) \setminus C$  is a subset of  $A \cap S$ ; the term  $S \setminus A$  intersects only  $T \setminus B$ ; and  $D \subseteq B$  is disjoint from  $T'$ . Therefore

$$S' \cap T' = ((S \cap A) \setminus C) \cup ((S \setminus A) \cap (T \setminus B)).$$

Since  $T$  is disjoint from  $A$ ,  $(S \setminus A) \cap (T \setminus B) = S \cap (T \setminus B) = (S \cap T) \setminus (S \cap B)$ . Thus

$$|S' \cap T'| = |S \cap T| - |S \cap B| + |S \cap A| - |C|.$$

Since  $|S' \cap T'| \geq t$ , to complete the proof it suffices to show that  $|S \cap A| \leq |S \cap B| + r$  (recalling that  $|C| = r$ ). If  $r = |S \cap A|$  then this clearly holds, and if  $r = |B \setminus S| - 1$  then

$$|S \cap B| + r = |S \cap B| + |B \setminus S| - 1 = |B| - 1 = |A| \geq |S \cap A|. \quad \square$$

As in the case of the simpler shifting operator  $\mathbb{S}_{i,j}$ , we can convert any family to a stable family while maintaining its being  $t$ -intersecting, by repeatedly applying a shifting operation on sets  $A, B$  with minimal  $|A|$ . One way of expressing this idea is the following lemma.

**Lemma 2.28.** *Let  $p \in (1/2, 1)$ . If  $\mathcal{F}$  is a  $t$ -intersecting family on  $n$  points having maximum  $\mu_p$ -measure then  $\mathcal{F}$  is  $(s, s + 1)$ -stable for all  $s$ .*

*Proof.* Suppose that  $\mathcal{F}$  is not  $(s, s + 1)$ -stable for all  $s$ . Choose the minimal  $s$  such that  $\mathcal{F}$  is not  $(s, s + 1)$ -stable, say  $\mathbb{S}_{A,B}(\mathcal{F}) \neq \mathcal{F}$ , where  $|A| = s$  and  $|B| = s + 1$ . Lemma 2.27 shows that  $\mathbb{S}_{A,B}(\mathcal{F})$  is also  $t$ -intersecting, and Lemma 2.26 shows that  $\mu_p(\mathbb{S}_{A,B}(\mathcal{F})) > \mu_p(\mathcal{F})$ . Thus  $\mathcal{F}$  is not a  $t$ -intersecting family of maximum  $\mu_p$ -measure.  $\square$

Ahlsvede and Khachatryan [6] in fact show that if  $\mathcal{F}$  is  $(r, r + 1)$ -stable for all  $r < s$  then for  $|A| = s$  and  $|B| = s + 1$ , the family  $\mathbb{S}_{A,B}(\mathcal{F})$  is also  $(r, r + 1)$ -stable for all  $r < s$ . While this simplifies the process of compression, we do not need this result here, and refer the reader to its proof in [6].

The importance of stable families is the following simple observation.

**Lemma 2.29.** *If a  $t$ -intersecting family  $\mathcal{F}$  on  $n$  points is  $(s, s + 1)$ -stable for all  $s$  then every  $A, B \in \mathcal{F}$  satisfy  $|A| + |B| \geq n + t - 1$ .*

*Proof.* Let  $A, B \in \mathcal{F}$ . If  $A \cup B = [n]$  then  $|A| + |B| = |A \cup B| + |A \cap B| \geq n + t$ , so suppose  $|A \cup B| < n$ .

Define  $s = \min(|A \cap B|, n - |A \cup B| - 1) \geq 0$ , and choose a subset  $C \subseteq A \cap B$  of size  $s$  and a subset  $D \subseteq [n] \setminus (A \cup B)$  of size  $s + 1$ . Since  $\mathcal{F}$  is  $(s, s + 1)$ -stable,  $A' = (A \setminus C) \cup D \in \mathcal{F}$ . We have  $|A' \cap B| = |A \cap B| - |C| = |A \cap B| - s$ , showing that  $|A \cap B| \geq s + t$ . In particular,  $s = n - |A \cup B| - 1$ , and so

$$|A| + |B| = |A \cup B| + |A \cap B| \geq (n - s - 1) + (s + t) = n + t - 1. \quad \square$$

We comment that the same result holds, with the same proof, for cross- $t$ -intersecting families  $\mathcal{F}, \mathcal{G}$ : any  $A \in \mathcal{F}$  and  $B \in \mathcal{G}$  satisfy  $|A| + |B| \geq n + t - 1$ .

The bound  $n + t - 1$  is tight: when  $n = t + 2r + 1$ , the family  $\mathcal{F}_{t,r}$  is  $(s, s + 1)$ -stable for all  $s$ , and two sets  $A, B$  of size  $t + r$  satisfy  $|A| + |B| = n + r - 1$ .

We need one more lemma, on uniform families.

**Lemma 2.30.** Let  $\mathcal{F} \subseteq \binom{[t+2r+1]}{t+r}$  be a  $t$ -intersecting family of maximum size, and define a family  $\mathcal{F}'_{t,r}$ , the uniform analog of  $\mathcal{F}_{t,r}$ , as follows:

$$\mathcal{F}'_{t,r} = \left\{ S \in \binom{[t+2r+1]}{t+r} : |S \cap [t+2r]| = t+r \right\}.$$

If  $t \geq 2$  then  $\mathcal{F}$  is equivalent to  $\mathcal{F}'_{t,r}$  (that is, equals a similar family with  $[t+2r]$  possibly replaced by some other subset of  $[t+2r+1]$  of size  $t+2r$ ), and if  $t = 1$  then  $|\mathcal{F}| \leq |\mathcal{F}'_{t,r}|$ .

*Proof.* Define  $\mathcal{G} = \{\bar{A} : A \in \mathcal{F}\}$  (where  $\bar{A} = [t+2r+1] \setminus A$ ), so that  $\mathcal{G} \subseteq \binom{[t+2r+1]}{r+1}$ . Since

$$|\bar{A} \cap \bar{B}| = |\overline{A \cup B}| = t+2r+1 - |A \cup B| = t+2r+1 - |A| - |B| + |A \cap B| = |A \cap B| - (t-1),$$

we see that the condition that  $\mathcal{F}$  is  $t$ -intersecting is equivalent to the condition that  $\mathcal{G}$  is intersecting.

Since  $r+1 \leq \frac{t+2r+1}{2}$  (with equality only for  $t=1$ ), the Erdős–Ko–Rado theorem shows that  $|\mathcal{G}| \leq \binom{t+2r}{r+1} = \binom{t+2r}{t+r}$ . Moreover, when  $t \geq 2$ , equality holds only when  $\mathcal{G} = \{S \in \binom{[t+2r+1]}{r+1} : i \in S\}$  for some  $i \in [t+2r+1]$ . In that case,  $\mathcal{F} = \{S \in \binom{[t+2r+1]}{t+r} : i \notin S\}$ , and so  $\mathcal{F}$  is equivalent to  $\mathcal{F}'_{t,r}$  (in the family  $\mathcal{F}_{t,r}$  itself,  $i = t+2r+1$ ).  $\square$

We can now prove Theorem 2.2 in the case  $p > 1/2$ .

**Lemma 2.31.** Let  $\mathcal{F}$  be a  $t$ -intersecting family on  $n$  points of maximum  $\mu_p$ -measure, for some  $p \in (1/2, 1)$ . Suppose that  $n \in \{t+2r, t+2r+1\}$ .

If  $t \geq 2$  or  $n = t+2r$  then  $\mathcal{F}$  is equivalent to  $\mathcal{F}_{t,r}$ .

If  $t = 1$  and  $n = t+2r+1$  then  $\mu_p(\mathcal{F}) \leq \mu_p(\mathcal{F}_{t,r})$  and  $\mathcal{F} = \mathcal{G} \cup \binom{[t+2r+1]}{\geq t+r+1}$ , where  $\mathcal{G} \subseteq \binom{[t+2r+1]}{t+r}$  contains exactly  $\binom{t+2r}{t+r}$  sets.

*Proof.* Lemma 2.28 shows that  $\mathcal{F}$  is  $(s, s+1)$ -stable for all  $s$ , and so Lemma 2.29 shows that any  $A, B \in \mathcal{F}$  satisfy  $|A| + |B| \geq n + t - 1$ . In particular, any set  $A$  has cardinality at least  $\frac{n+t-1}{2}$ . We now consider two cases, according to the parity of  $n+t$ .

Suppose first that  $n = t+2r$ . Then  $\frac{n+t-1}{2} = t+r - \frac{1}{2}$ , and so all sets in  $\mathcal{F}$  have cardinality at least  $t+r$ . In other words,  $\mathcal{F} \subseteq \mathcal{F}_{t,r}$ . Since  $\mathcal{F}$  has maximum  $\mu_p$ -measure,  $\mathcal{F} = \mathcal{F}_{t,r}$ .

Suppose next that  $n = t+2r+1$ . Then  $\frac{n+t-1}{2} = t+r$ , and so all sets in  $\mathcal{F}$  have cardinality at least  $t+r$ . If  $|A| \geq t+r$  and  $|B| \geq t+r+1$  then  $|A \cap B| \geq |A| + |B| - n = t$ , and so the fact that  $\mathcal{F}$  has maximum  $\mu_p$ -measure shows that  $\mathcal{F} = \mathcal{G} \cup \binom{[n]}{\geq t+r+1}$ , where  $\mathcal{G} \subseteq \binom{[n]}{t+r}$  is  $t$ -intersecting.

We can now complete the proof using Lemma 2.30. If  $t \geq 2$  then  $\mathcal{G}$  is equivalent to  $\mathcal{F}'_{t,r}$ , say

$$\mathcal{G} = \left\{ S \in \binom{[t+2r+1]}{t+r} : |S \cap X| = t+r \right\},$$

where  $|X| = t+2r$ . Since any set of size at least  $t+r+1$  intersects  $X$  in at least  $t+r$  points,

$$\mathcal{F} = \{S \subseteq [n] : |S \cap X| \geq t+r\},$$

and so  $\mathcal{F}$  is equivalent to  $\mathcal{F}_{t,r}$ .

When  $t = 1$ , Lemma 2.30 shows that  $|\mathcal{G}| \leq |\mathcal{F}'_{t,r}|$ , and so the same reasoning shows that  $\mu_p(\mathcal{F}) \leq \mu_p(\mathcal{F}_{t,r})$ .  $\square$

### 3 Katona's circle argument

Katona [22] gave a particularly simple proof of the Erdős–Ko–Rado theorem, using what has become known as the *circle method*. The same proof goes through in the  $\mu_p$  setting, with much the same proof, as we show in Section 3.1. We are able to use the same argument to obtain a description of all optimal families.

The heart of Katona's argument is the following seemingly trivial observation: if  $S$  is a measurable set on the unit circumference circle in which any two points are at distance at most  $p < 1/2$ , then the length of  $S$  is at most  $p$ , the optimal sets being intervals. In Section 3.2 we give two versions of this argument for the case of two sets: one in a discrete setting, and the other in a continuous setting. These results will be used later on when we prove Ahlswede–Khachatrian theorems for families on infinitely many points (Section 4) and for the Hamming scheme (Section 5).

### 3.1 Intersecting families

The Erdős–Ko–Rado theorem, in our setting, states that if  $\mathcal{F}$  is an intersecting family then  $\mu_p(\mathcal{F}) \leq p$  for all  $p \leq 1/2$ . Katona [22] gave a particularly simple proof of the theorem in its original setting, and we adapt his proof to our setting.

**Lemma 3.1.** *Let  $n \geq 1$  and  $p \in (0, 1/2)$ . Then  $\mu_p(\mathcal{F}) \leq p$  for all intersecting families  $\mathcal{F}$ .*

*Moreover, if  $\mathcal{F}$  is an intersecting family on  $n$  points such that  $\mu_p(\mathcal{F}) = p$  then  $\mathcal{F}$  is equivalent to a  $(1, 0)$ -Frankl family. In other words, if  $\mu_p(\mathcal{F}) = p$  then for some  $i \in [n]$  we have*

$$\mathcal{F} = \{A \subseteq [n] : A \ni i\}.$$

*Proof.* The idea is to come up with a probabilistic model for the distribution  $\mu_p$ , and use it to show that  $\mu_p(\mathcal{F}) \leq p$ . Since  $\mu_p(\mathcal{F}_{1,0}) = p$ , this shows that  $w(n, 1, p) = p$ . We will then use the same probabilistic model to identify all families satisfying  $\mu_p(\mathcal{F}) = p$ .

Let  $\mathbb{T}$  be the circle of unit circumference. Choose  $n$  points  $x_1, \dots, x_n$  at random on  $\mathbb{T}$ . Choose another point  $t$  on  $\mathbb{T}$  at random, and consider the arc  $(t, t+p)$  (where  $t+p$  is taken modulo 1). The set of indices of points  $S_t$  which lie inside the arc has distribution  $\mu_p$ , and so  $\mu_p(\mathcal{F}) = \Pr[S_t \in \mathcal{F}]$ .

We will prove that for *each* setting of  $x_1, \dots, x_n$  we have  $\Pr[S_t \in \mathcal{F}] \leq p$ , the probability taken over the choice of  $t$ . Let  $I = \{t : S_t \in \mathcal{F}\}$ , and note that  $I$  is a union of intervals. The crucial observation is that if  $t_1, t_2 \in I$  then the corresponding arcs  $(t_1, t_1+p), (t_2, t_2+p)$  must intersect, and so  $d(t_1, t_2) < p$ , where  $d(\cdot, \cdot)$  is shortest distance on the circumference of the circle. Consider now any  $t_1 \in I$ . All  $t_2 \in I$  must lie in the interval  $(t_1 - p, t_1 + p)$ , and moreover for each  $s \in (0, p)$ , at most one of  $t_1 + s, t_1 - p + s$  can be in  $I$  (since  $d(t_1 + s, t_1 + s - p) = p$ ). It follows that the measure of  $I$  is at most  $p$ . In other words,  $\Pr[S_t \in \mathcal{F}] \leq p$ , implying that  $\mu_p(\mathcal{F}) \leq p$ .

For future use, we also need to identify the cases in which  $I$  has measure exactly  $p$ . We will show that  $I$  must be an interval of length  $p$ . The argument above shows that  $I$  is a union of non-empty intervals of two types,  $J_1, \dots, J_a \subseteq (t_1 - p, t_1]$  and  $K_1, \dots, K_b \subseteq [t_1, t_1 + p)$ , such that the intervals  $J_i + p, K_j$  together partition  $[t_1, t_1 + p)$ . If  $a = 0$  then  $I = [t_1, t_1 + p)$ , and if  $b = 0$  then  $I = (t_1 - p, t_1]$ . If there is a unique interval  $K_1$  which is of the form  $[t_1, t_1 + s)$  (or  $[t_1, t_1 + s]$ ) then  $I = [t_1 + s - p, t_1 + s)$  (or  $(t_1 + s - p, t_1 + s]$ ). Otherwise, there must be some interval  $K_j$  whose left end-point  $y$  is larger than  $t_1$ . There is a corresponding interval  $J_i$  whose right end-point is  $y - p$ . However, since  $p < 1/2$ , a point on  $J_i$  slightly to the left of  $y - p$  has distance larger than  $p$  from a point on  $K_j$  slightly to the right of  $y$ , contradicting our assumptions. We conclude that  $I$  must be an interval of length  $p$ .

We proceed to identify the families  $\mathcal{F}$  which achieve the upper bound, that is, satisfy  $\mu_p(\mathcal{F}) = p$ . Clearly any family  $\mathcal{F}$  equivalent to  $\mathcal{F}_{1,0}$  satisfies  $\mu_p(\mathcal{F}) = p$ . We will show that these are the only such families. First, notice that if  $\mu_p(\mathcal{F}) = p$  then  $\mathcal{F}$  must be monotone (otherwise its upset is an intersecting family of measure larger than  $p$ ). Moreover, the set  $I$  defined above is an interval of length  $p$  almost surely, with respect to the choice of  $x_1, \dots, x_n$ . In particular, there is a choice of  $x_1, \dots, x_n$ , all distinct and none at distance exactly  $p$ , such that  $I$  is an interval of length  $p$ , say  $I = [y, y + p)$ . For small  $\epsilon > 0$ , the arcs  $(y + \epsilon, y + p + \epsilon), (y + p - \epsilon, y + 2p - \epsilon)$  intersect at a small neighborhood of  $y + p$ . Since the corresponding sets  $S_{y+\epsilon}, S_{y+p-\epsilon}$  intersect, there must be some point  $x_i = y + p$ . We will show that  $\{x_i\} \in \mathcal{F}$ , and so monotonicity implies that  $\mathcal{F} = \{A \subseteq [n] : A \ni i\}$ .

Let  $A$  consist of all points in  $(y, y + p)$ . Thus  $A \cup \{x_i\} \in \mathcal{F}$  (since  $y + \epsilon \in I$  for small  $\epsilon$ ) while  $A \notin \mathcal{F}$  (since  $y - \epsilon \notin I$  for small  $\epsilon$ ). Consider now the set of configurations  $x'_1, \dots, x'_n$  in which  $x'_j \in (x'_i - p, x'_i)$  for all  $j \in A$  and  $x'_j \notin (x'_i - p, x'_i + p)$  for all  $j \notin A \cup \{i\}$ . This set of configurations has positive measure, and so there must exist one whose corresponding set  $I'$  is an interval of measure  $p$ . By construction,  $I'$  contains all points  $t$  just to the right of  $x'_i - p$  (since the corresponding  $S_t$  is  $A \cup \{i\}$ ) but not  $x'_i - p$  (since the corresponding  $S_{x'_i - p}$  is  $A$ ), and so  $x'_i - p$  is the left end-point of the interval. The right end-point is thus  $x'_i$  itself. By construction,  $S_{x'_i - \epsilon} = \{i\}$  for small enough  $\epsilon > 0$ , showing that  $\{i\} \in \mathcal{F}$ . This completes the proof.  $\square$

Other proofs of the upper bound which employ similar arguments appear in Dinur and Friedgut [9] and in Friedgut [17]. Unfortunately, it seems that Katona's idea doesn't extend to  $t$ -intersecting families for  $t > 1$ . For a discussion of this, see Howard, Károlyi and Székely [20].

## 3.2 Cross-intersecting settings

As explained in the introduction to this section, the heart of Katona's proof is a result on sets in which any two points are close. The proof of Lemma 3.1 closely follows the original argument. In this section we give alternative arguments for the cross-intersecting counterparts in both discrete and continuous settings.

### 3.2.1 Discrete setting

We start with the easier, discrete setting. First, a few definitions.

**Definition 3.1.** A set  $A \subseteq \mathbb{Z}_m$  is *s-agreeing* if every  $a, b \in A$  satisfy  $a - b \in \{-(s-1), \dots, s-1\}$ . Two sets  $A, B \subseteq \mathbb{Z}_m$  are *cross-s-agreeing* if every  $a \in A$  and  $b \in B$  satisfy  $a - b \in \{-(s-1), \dots, s-1\}$ .

**Definition 3.2.** A set  $A \subseteq \mathbb{Z}_m$  is an *interval* if it is of the form  $\{x - \ell, \dots, x + \ell\}$  (for  $2\ell + 1 < m$ ) or  $\{x - \ell, \dots, x + \ell - 1\}$  (for  $2\ell < m$ ). In the first case,  $x$  is the *center* of  $A$ , and in the second,  $x - 1/2$  is the center of  $A$ .

The following simple lemma will simplify the argument below.

**Lemma 3.2.** *Let  $s \leq m/2$ . If  $A \subseteq \mathbb{Z}_m$  is cross-s-agreeing with the non-empty interval  $\{x, \dots, y\}$  of length at most  $2s - 1$  then  $A \subseteq \{y - (s - 1), \dots, x + (s - 1)\}$ .*

*Proof.* The proof is by induction on the length of the interval. If  $y = x$  then trivially  $A \subseteq \{x - (s - 1), \dots, x + (s - 1)\}$ . Suppose now that we have already shown that if  $A$  is cross-s-agreeing with  $\{x, \dots, y\}$  then  $A \subseteq \{y - (s - 1), \dots, x + (s - 1)\}$ . If  $A$  is cross-s-agreeing with  $\{x, \dots, y + 1\}$  then

$$A \subseteq \{y - (s - 1), \dots, x + (s - 1)\} \cap \{y + 1 - (s - 1), \dots, x + 1 + (s - 1)\} = \{y + 1 - (s - 1), \dots, x + (s - 1)\}. \quad \square$$

We can now state and prove the result in the discrete setting.

**Lemma 3.3.** *Let  $m, s \geq 1$  be integers satisfying  $s < m/2$ . If  $A, B \subseteq \mathbb{Z}_m$  are non-empty s-agreeing sets then  $|A| + |B| \leq 2s$ . Moreover, if  $|A| + |B| = 2s$  then  $A, B$  are intervals centered at the same point or half-point. In particular, if  $A$  is s-agreeing then  $|A| \leq s$ , with equality only when  $A$  is an interval.*

*When  $s = m/2$  it still holds that  $|A| + |B| \leq 2s$ , with equality when  $B = \{x : x + s \notin A\}$ , and this holds even without the assumption that  $A, B$  be non-empty. In particular, if  $A$  is s-agreeing then  $|A| \leq s$ , with equality only when  $A$  contains one point out of each pair  $x, x + s$ .*

*Proof.* We start by proving that  $|A| + |B| \leq 2s$  when  $s < m/2$ . The idea is to use a shifting argument to transform  $A, B$  into intervals without decreasing  $|A| + |B|$ . We construct a sequence of non-empty s-agreeing sets, starting with  $(A_0, B_0) = (A, B)$ . Given  $(A_i, B_i)$ , note first that  $A_i \neq \mathbb{Z}_m$ , since otherwise  $B_i$  would have to be empty. Therefore there exists a point  $x \in A_i$  such that  $x + 1 \notin A_i$ . We take  $A_{i+1} = A_i \cup \{x + 1\}$ . If  $A_{i+1}, B_i$  are not cross-s-agreeing then there must be a point  $y \in B_i$  such that  $x - y \in \{-(s - 1), \dots, s - 1\}$  but  $x + 1 - y \notin \{-(s - 1), \dots, s - 1\}$ . This can only happen if  $x - y = s - 1$ , and so there is at most one such point  $y = x - (s - 1)$ . We therefore take  $B_{i+1} = B_i \setminus \{x - (s - 1)\}$ . If  $B_{i+1} = \emptyset$ , then we stop the sequence at  $(A_i, B_i)$ , and otherwise we continue. Note that  $|A_{i+1}| = |A_i| + 1$  and  $|B_{i+1}| \geq |B_i| - 1$ , and so  $|A_{i+1}| + |B_{i+1}| \geq |A_i| + |B_i|$ .

Since the size of  $A_i$  keeps increasing, there must be a last pair in the sequence, say  $(A_t, B_t)$ . Our stopping condition guarantees that  $|B_t| = 1$ , say  $B_t = \{y\}$ . This forces  $A_t \subseteq \{y - (s - 1), \dots, y + (s - 1)\}$ , and so  $|A_t| + |B_t| \leq (2s - 1) + 1 = 2s$ . Since  $|A_0| + |B_0| \leq |A_t| + |B_t| \leq 2s$ , this completes the proof that  $|A| + |B| \leq 2s$ .

Next, we show that if  $|A| + |B| = 2s$  then  $A, B$  are intervals centered at the same point. We do this by reverse induction on the sequence  $(A_0, B_0), \dots, (A_t, B_t)$ . For the base case, the argument in the preceding paragraph shows that  $|A_t| + |B_t| = 2s$  if  $A_t = \{y - (s - 1), \dots, y + (s - 1)\}$  and  $B_t = \{y\}$ , and so both sets are intervals centered at the point  $y$ .

Suppose now that  $(A_{i+1}, B_{i+1})$  are intervals centered (without loss of generality) at 0 or  $-1/2$ , depending on the parity of  $|A_{i+1}|$ . Suppose first that  $|A_{i+1}|$  is odd. Then for some  $0 \leq \ell \leq s - 1$  we have  $A_{i+1} = \{-\ell, \dots, \ell\}$  and  $B_{i+1} = \{-(s - 1 - \ell), \dots, s - 1 - \ell\}$ . By construction,  $-\ell \in A_i$ . Suppose that also  $\ell \in A_i$ , so that  $A_i = \{-\ell, \dots, x\} \cup \{x + 2, \dots, \ell\}$ . Lemma 3.2 shows that

$$B_i \subseteq \{x - (s - 1), \dots, s - 1 - \ell\} \cap \{-(s - 1 - \ell), \dots, x + 2 + (s - 1)\} \subseteq \{-(s - 1 - \ell), \dots, s - 1 - \ell\},$$

and so  $B_i \subseteq B_{i+1}$ , implying that  $|A_i| + |B_i| < 2s$ . We conclude that  $\ell \notin A_i$ , and so  $A_i = \{-\ell, \dots, \ell - 1\}$ , corresponding to  $x = \ell - 1$ . By construction,  $B_i = B_{i+1} \cup \{x - (s - 1)\} = \{-(s - \ell), \dots, s - 1 - \ell\}$ . Both sets are centered at  $-1/2$ .

The second case, when  $|A_{i+1}|$  is even, is similar. In this case for some  $1 \leq \ell \leq s - 1$  we have  $A_{i+1} = \{-\ell, \dots, \ell - 1\}$  and  $B_{i+1} = \{-(s - \ell), \dots, s - \ell - 1\}$ . By construction,  $-\ell \in A_i$ . Suppose that also  $\ell - 1 \in A_i$ , so that  $A_i = \{-\ell, \dots, x\} \cup \{x + 2, \dots, \ell - 1\}$ . Lemma 3.2 shows that

$$B_i \subseteq \{x - (s - 1), \dots, s - 1 - \ell\} \cap \{-(s - \ell), \dots, x + 2 + (s - 1)\} \subseteq \{-(s - \ell), \dots, s - 1 - \ell\},$$

and so  $B_i \subseteq B_{i+1}$ . As before, this leads to a contradiction, and we conclude that  $A_i = \{-\ell, \dots, \ell - 2\}$  is centered at  $-1$ . Moreover,  $B_i = A_i \cup \{\ell - 2 - (s - 1)\} = \{-(s - \ell) - 1, \dots, s - 1 - \ell\}$  is also centered at  $-1$ . This completes the proof.

It remains to consider the case  $s = m/2$ . The  $s$ -agreeing condition states that if  $a \in A$  and  $b \in B$  then  $a - b \neq s$ . Thus  $B \subseteq \{x : x + s \notin A\}$ , which implies  $|A| + |B| \leq m = 2s$ . This bound is tight only when  $B = \{x : x + s \notin A\}$ .  $\square$

As an easy corollary, we can derive the classical Erdős–Ko–Rado theorem.

**Corollary 3.4.** *Let  $n \geq k \geq 1$  be parameters such that  $k \leq n/2$ . If  $\mathcal{F} \subseteq \binom{[n]}{k}$  is intersecting then  $|\mathcal{F}| \leq \binom{n-1}{k-1}$ . Furthermore, if  $k < n/2$  and  $|\mathcal{F}| = \binom{n-1}{k-1}$  then  $\mathcal{F}$  consists of all sets containing some  $i \in [n]$ .*

*Proof.* The proof is very similar to the proof of Lemma 3.1. Let  $\pi$  be a random permutation of  $[n]$ , and choose  $t \in [n]$  at random. Let  $S_t = \{\pi(t + 1), \dots, \pi(t + k)\}$ , where indices are taken modulo  $n$ . Since  $S_t$  is a random set from  $\binom{[n]}{k}$ , we see that  $|\mathcal{F}|/\binom{n}{k}$  is the probability that  $S_t \in \mathcal{F}$ .

For any setting of  $\pi$ , let  $I = \{t \in [n] : S_t \in \mathcal{F}\}$ . Since  $\mathcal{F}$  is intersecting, if  $a, b \in I$  then  $\{\pi(a + 1), \dots, \pi(a + k)\}, \{\pi(b + 1), \dots, \pi(b + k)\}$  must intersect, and this implies that  $I$  is  $k$ -agreeing (in the sense of Definition 3.1). Lemma 3.3 shows that  $|I| \leq k$ , and so  $|\mathcal{F}|/\binom{n}{k} \leq k/n$ , or  $|\mathcal{F}| \leq \binom{n-1}{k-1}$ .

Suppose now that  $k < n/2$  and  $|\mathcal{F}| = \binom{n-1}{k-1}$ . Lemma 3.3 shows that for every permutation  $\pi$ , the set  $I$  must be an interval of length  $k$ . In particular, this is the case for the identity permutation. Suppose without loss of generality that in this case,  $I = \{1, \dots, k\}$ . Thus  $\{2, \dots, k + 1\}, \dots, \{k + 1, \dots, 2k\} \in \mathcal{F}$  but  $\{1, \dots, k\} \notin \mathcal{F}$  (since  $0 \notin I$ ). Let  $S \in \binom{[n]}{k}$  be any set containing  $k + 1$  but not 1, and write  $S = \{k + 1\} \cup A \cup B$ , where  $A \subseteq \{2, \dots, k\}$ . Let  $\pi$  be any permutation which starts 1,  $\{2, \dots, k\} \setminus A, A, k + 1, B$ . Since  $\{\pi(1), \dots, \pi(k)\} = \{1, \dots, k\} \notin \mathcal{F}$  whereas  $\{\pi(2), \dots, \pi(k + 1)\} = \{2, \dots, k + 1\} \in \mathcal{F}$ , the set  $I$  contains 1 but not 0, and so must be  $\{1, \dots, k\}$ . This implies that  $S \in \mathcal{F}$ .

To finish the proof, let  $T \in \binom{[n]}{k}$  be any set not containing  $k + 1$ . Let  $U$  be  $k - 1$  elements disjoint from  $T$  and not containing 1 or  $k + 1$ ; such elements exist since  $n - (k + 2) \geq (2k + 1) - (k + 2) = k - 1$ . Let  $\pi$  be any permutation starting  $T, k + 1, U$ , where we put 1 first if  $1 \in T$ . The earlier paragraph shows that  $I \supseteq \{1, \dots, k\}$ , and in particular  $T \notin \mathcal{F}$ . Thus all sets in  $\mathcal{F}$  contain  $k + 1$ , and since  $|\mathcal{F}| = \binom{n-1}{k-1}$ , it must contain all such sets. This completes the proof.  $\square$

### 3.2.2 Continuous setting

We proceed with a continuous analog of Lemma 3.3. First, the pertinent definitions.

**Definition 3.3.** The *unit circle*  $\mathbb{T}$  consists of the interval  $[0, 1]$  with its two end-points pasted. The distance between two points  $x, y \in \mathbb{T}$  is their distance on the unit circle. We denote the Lebesgue measure on  $\mathbb{T}$  by  $\mu$ .

Let  $p \leq 1/2$ . Two measurable sets  $A, B \subseteq \mathbb{T}$  are  *$p$ -agreeing* if any  $a \in A$  and  $b \in B$  are at distance less than  $p$ .

We now state and prove the analog of Lemma 3.3. We only state and prove the upper bound part, leaving the identification of optimal sets to the reader.

**Lemma 3.5.** *Let  $p < 1/2$ . If two non-empty measurable sets  $A, B \subseteq \mathbb{T}$  are  $p$ -agreeing then  $\mu(A) + \mu(B) \leq 2p$ .*

*Proof.* The proof uses an approximation argument. Let  $\epsilon > 0$  be a parameter satisfying  $p + \epsilon < 1/2$ . Since  $A$  is measurable, there is a sequence  $A_i$  of intervals of total length smaller than  $\mu(A) + \epsilon$  such that  $A \subseteq \bigcup_i A_i$ . Note that every point in  $\bigcup_i A_i$  is at distance at most  $\epsilon/2$  from a point in  $A$ , since otherwise  $\mu(\bigcup_i A_i \setminus A) \geq \epsilon$ . Similarly, there is a sequence  $B_j$  of intervals of total length at most  $\mu(B) + \epsilon$  such that  $B \subseteq \bigcup_j B_j$ , and any point in  $\bigcup_j B_j$  is at distance at most  $\epsilon/2$  from a point in  $B$ . Thus  $\bigcup_i A_i$  and  $\bigcup_j B_j$  are cross- $(p + \epsilon)$ -agreeing.

Choose  $I$  so that  $\sum_{i>I} \mu(A_i) < \epsilon$ , and set  $A^* = \bigcup_{i=1}^I A_i$ . Similarly, choose  $J$  so that  $\sum_{j>J} \mu(B_j) < \epsilon$ , and set  $B^* = \bigcup_{j=1}^J B_j$ . Thus  $A^*, B^*$  are cross- $(p + \epsilon)$ -agreeing, and each is a union of finitely many intervals.

Let  $M$  be a large integer, and define  $A_M^* = \bigcup_{x=0}^{M-1} \{[x/M, (x+1)/M] : [x/M, (x+1)/M] \subseteq A^*\}$ . Note that  $\mu(A_M^*) \geq \mu(A^*) - 2I/M$ . Define  $B_M^*$  similarly. We can view  $A_M^*, B_M^*$  as subsets of  $\mathbb{Z}_M$ . These subsets are cross- $\lfloor (p + \epsilon)M \rfloor$ -agreeing: if  $[a/M, (a+1)/M] \in A_M^*$  and  $[b/M, (b+1)/M] \in B_M^*$  then  $a/M, (b+1)/M$  are at distance less than  $p + \epsilon$ . Lemma 3.3 thus shows that  $\mu(A_M^*) + \mu(B_M^*) \leq 2(p + \epsilon)$ . On the other hand,  $\mu(A_M^*) + \mu(B_M^*) \geq \mu(A^*) + \mu(B^*) - 2(I + J)/M$ . Taking the limit  $M \rightarrow \infty$ , we deduce that  $\mu(A^*) + \mu(B^*) \leq 2(p + \epsilon)$ . Since  $\mu(A^*) + \mu(B^*) \geq \mu(A) + \mu(B) - 2\epsilon$ , we conclude that  $\mu(A) + \mu(B) \leq 2p + 4\epsilon$ . Taking the limit  $\epsilon \rightarrow 0$ , we deduce the lemma.  $\square$

As a corollary, we obtain the following useful result.

**Corollary 3.6.** *Let  $\mathcal{F}, \mathcal{G}$  be cross-intersecting families on  $n$  points. For any  $p \leq 1/2$ ,  $\mu_p(\mathcal{F}) + \mu_p(\mathcal{G}) \leq 1$ .*

*Proof.* We first settle the case  $p = 1/2$ . If  $\mathcal{F}, \mathcal{G}$  are cross-intersecting then  $\mathcal{G}$  is disjoint from  $\{\bar{A} : A \in \mathcal{F}\}$ . Since the latter set has measure  $\mu_p(\mathcal{F})$ , we conclude that  $\mu_p(\mathcal{F}) + \mu_p(\mathcal{G}) \leq 1$ .

Suppose now that  $p < 1/2$ . We will follow the argument of Lemma 3.1. We choose  $n$  points  $x_1, \dots, x_n$  at random on  $\mathbb{T}$ , and two starting points  $t_{\mathcal{F}}, t_{\mathcal{G}} \in \mathbb{T}$  at random. Let  $S_t$  be the set of points that lie inside the interval  $(t, t + p)$ , let  $I_{\mathcal{F}}$  be the set of indices  $t_{\mathcal{F}}$  such that  $S_{t_{\mathcal{F}}} \in \mathcal{F}$ , and define  $I_{\mathcal{G}}$  analogously. Thus  $\mu_p(\mathcal{F}) + \mu_p(\mathcal{G}) = \mathbb{E}[\mu(I_{\mathcal{F}}) + \mu(I_{\mathcal{G}})]$ . Since  $\mathcal{F}, \mathcal{G}$  are cross-intersecting, we see that  $I_{\mathcal{F}}, I_{\mathcal{G}}$  are cross- $p$ -agreeing. The lemma shows that  $\mu(I_{\mathcal{F}}) + \mu(I_{\mathcal{G}}) \leq 2p \leq 1$  if neither  $I_{\mathcal{F}}$  nor  $I_{\mathcal{G}}$  are empty, and  $\mu(I_{\mathcal{F}}) + \mu(I_{\mathcal{G}}) \leq 1$  trivially holds if one of the sets is empty. The corollary follows.  $\square$

We comment that the corollary remains holding if the families in question are on infinitely many points, in the sense of Section 4, with the same proof; the only difference is that we choose infinitely many points rather than just  $n$ .

## 4 Infinite Ahlswede–Khachatrian theorem

The goal of Section 2 was to calculate the quantities  $w(n, t, p)$  and  $w_{\text{sup}}(t, p)$  which are, respectively, the maximum  $\mu_p$ -measure of a  $t$ -intersecting family on  $n$  points, and the supremum  $\mu_p$ -measure of a  $t$ -intersecting family on any number of points. In this section we will consider what happens when we allow families on infinitely many points.

**Definition 4.1.** The infinite product measure  $\mu_p$  on  $\mathcal{F} \subseteq 2^{\mathbb{N}}$  is the infinite product measure extending the finite  $\mu_p$  measures.

Given  $p \in (0, 1)$ , a family on infinitely many points is a subset  $\mathcal{F} \subseteq 2^{\mathbb{N}}$  which is measurable with respect to  $\mu_p$ . The family is  $t$ -intersecting if any two sets  $A, B \in \mathcal{F}$  have at least  $t$  points in common; here  $t$  can also be  $\aleph_0$ .

A family  $\mathcal{F}$  is *finitely determined* if it is the *extension* of some family  $\mathcal{G}$  on  $n$  points:  $\mathcal{F} = \{A : A \cap [n] \in \mathcal{G}\}$ . Given  $p \in (0, 1)$ , the family  $\mathcal{F}$  is *essentially finitely determined* if it differs from a finitely determined family by a  $\mu_p$ -null set.

**Definition 4.2.** For  $t \geq 1$  and  $p \in (0, 1)$ , the parameter  $w_{\infty}(t, p)$  is the supremum of  $\mu_p(\mathcal{F})$  over all  $t$ -intersecting families on infinitely many points.

The following theorem summarizes the results proved in this section.

**Theorem 4.1.** *Let  $t \geq 1$  and  $p \in (0, 1)$ .*

*If  $p < 1/2$  then  $w_{\infty}(t, p) = w_{\text{sup}}(t, p)$ . Furthermore, if  $\mathcal{F}$  is a  $t$ -intersecting family on infinitely many points satisfying  $\mu_p(\mathcal{F}) = w_{\infty}(t, p)$  then  $\mathcal{F}$  is essentially finitely determined.*

If  $p = 1/2$  then  $w_\infty(t, p) = 1/2$ . Furthermore, if  $t > 1$  then no  $t$ -intersecting family on infinitely many points has measure  $1/2$ .

If  $p > 1/2$  then  $w_\infty(t, p) = 1$ . In fact, there is an  $\aleph_0$ -intersecting family  $\mathcal{F}$  on infinitely many points which satisfying  $\mu_p(\mathcal{F}) = 1$  for all  $p > 1/2$ .

Note that there are many intersecting families with  $\mu_{1/2}$ -measure  $1/2$ , for example  $\mathcal{F}_{1,r}$  for every  $r$ . There are also families which are not essentially finitely determined. One example is the family of sets  $S$  such that for some  $n \geq 0$ ,

- (a)  $S$  contains exactly one of  $\{2m + 1, 2m + 2\}$  for all  $m < n$ .
- (b)  $S$  contains both  $\{2n + 1, 2n + 2\}$ .

We prove the different cases in the theorem one by one, starting with the case  $p < 1/2$ . We start with the following technical proposition, which follows from the definition of  $\mu_p$ .

**Proposition 4.2.** *Let  $\mathcal{F}$  be a family on infinitely many points. For every  $p \in (0, 1)$  and for every  $\epsilon > 0$  there is a finitely determined family  $\mathcal{G}$  such that  $\mu_p(\mathcal{F} \triangle \mathcal{G}) < \epsilon$ .*

Our proof will need a stability version of Theorem 2.2, due to Keller and Lifshitz [23].

**Proposition 4.3.** *Fix  $p \in (0, 1/2)$  and  $t \geq 1$ . If  $\mathcal{F}$  is a  $t$ -intersecting family on  $n$  points of measure  $(1 - \epsilon)w(n, t, p)$  then there exists a  $t$ -intersecting family  $\mathcal{G}$  on  $n$  points of measure  $w(n, t, p)$  such that  $\mu_p(\mathcal{F} \triangle \mathcal{G}) = O(\epsilon^{\log_{1-p} p})$ . Here the hidden constant depends on  $p, t$  but not on  $n$ .*

This will allow us to settle the case  $p < 1/2$ .

**Lemma 4.4.** *Let  $t \geq 1$  and  $p \in (0, 1/2)$ . Then  $w_\infty(t, p) = w_{\text{sup}}(t, p)$ . Furthermore, if  $\mathcal{F}$  is a  $t$ -intersecting family on infinitely many points satisfying  $\mu_p(\mathcal{F}) = w_\infty(t, p)$  then  $\mathcal{F}$  is essentially finitely determined.*

*Proof.* Clearly  $w_\infty(t, p) \geq w_{\text{sup}}(t, p)$ . For the other direction, let  $\mathcal{F}$  be a  $t$ -intersecting family on infinitely many points. Let  $\epsilon \in (0, 1/4)$  be a parameter. Proposition 4.2 shows that there is a family  $\mathcal{G}$  depending on  $N$  points such that  $\mu_p(\mathcal{F} \triangle \mathcal{G}) < \epsilon$ . We can assume that  $\mathcal{G}$  depends on the first  $N$  points. For  $S \subseteq [N]$ , let  $\mathcal{F}_S = \{A \in 2^{\mathbb{N} \setminus [N]} : S \cup A \in \mathcal{F}\}$ .

Let  $\text{dist}(x, \{0, 1\}) = \min(|x|, |x - 1|)$ , and notice that

$$\mathbb{E}_{S \sim \mu_p([N])} [\text{dist}(\mu_p(\mathcal{F}_S), \{0, 1\})] \leq \mathbb{E}_{S \sim \mu_p([N])} [\mu_p(\mathcal{F}_S \triangle \mathcal{G}_S)] = \mu_p(\mathcal{F} \triangle \mathcal{G}) < \epsilon.$$

Thus with probability at least  $1 - \sqrt{\epsilon}$ ,  $\text{dist}(\mu_p(\mathcal{F}_S), \{0, 1\}) < \sqrt{\epsilon}$ .

Define now  $\mathcal{F}' = \{A \in \mathcal{F} : \mu_p(\mathcal{F}_{A \cap [N]}) > 1 - \sqrt{\epsilon}\}$ . In words,  $\mathcal{F}'$  is the subset of  $\mathcal{F}$  obtained by removing all fibers  $\mathcal{F}_S$  whose  $\mu_p$ -measure is at most  $1 - \sqrt{\epsilon}$ . In particular, we remove all fibers whose  $\mu_p$ -measure is less than  $\sqrt{\epsilon}$ , and all fibers whose measure is between  $\sqrt{\epsilon}$  and  $1 - \sqrt{\epsilon}$ . The former have measure at most  $\sqrt{\epsilon}$ , and the latter also have a measure at most  $\sqrt{\epsilon}$  due to the preceding paragraph. Thus  $\mu_p(\mathcal{F} \setminus \mathcal{F}') \leq 2\sqrt{\epsilon}$ .

If  $S, T \subseteq [N]$  are such that  $\mu_p(\mathcal{F}_S), \mu_p(\mathcal{F}_T) > 1 - \sqrt{\epsilon} > 1/2$  then  $\mathcal{F}_S, \mathcal{F}_T$  cannot be cross-intersecting, due to Corollary 3.6 (while we stated the corollary for finite families, it holds for infinite families with exactly the same proof). It follows that  $|S \cap T| \geq t$ , and so the family  $\mathcal{G} = \{A : \mu_p(\mathcal{F}_{A \cap [N]}) > 1 - \sqrt{\epsilon}\}$  containing  $\mathcal{F}'$  is  $t$ -intersecting. Since  $\mathcal{F}'$  is contained in a  $t$ -intersecting family depending on  $N$  points,  $\mu_p(\mathcal{F}') \leq w_{\text{sup}}(t, p)$ , and so  $\mu_p(\mathcal{F}) \leq w_{\text{sup}}(t, p) + 2\sqrt{\epsilon}$ . Taking the limit  $\epsilon \rightarrow 0$ , we deduce that  $\mu_p(\mathcal{F}) \leq w_{\text{sup}}(t, p)$ , and so  $w_\infty(t, p) = w_{\text{sup}}(t, p)$ .

Suppose now that  $\mu_p(\mathcal{F}) = w_{\text{sup}}(t, p)$ . For every  $\epsilon > 0$ , we have constructed above a  $t$ -intersecting family  $\mathcal{F}' = \mathcal{F}'_\epsilon$  which is contained in a  $t$ -intersecting family  $\mathcal{G}_\epsilon$  depending on  $N_\epsilon$  points and satisfies  $\mu_p(\mathcal{F}'_\epsilon) \geq \mu_p(\mathcal{F}) - 2\sqrt{\epsilon}$ . In particular,  $\mu_p(\mathcal{G}_\epsilon) \geq \mu_p(\mathcal{F}) - 2\sqrt{\epsilon}$ . For small enough  $\epsilon > 0$ , this can only happen if  $w(N_\epsilon, t, p) = w_{\text{sup}}(t, p)$ . Furthermore, Proposition 4.3 shows that  $\mu_p(\mathcal{G}_\epsilon \triangle \mathcal{H}_\epsilon) = O(\epsilon^{O(1)})$  for some  $t$ -intersecting family  $\mathcal{H}_\epsilon$  on  $N$  points of  $\mu_p$ -measure  $w_{\text{sup}}(t, p)$ . Corollary 2.4 shows that such a family is a Frankl family with appropriate parameters. In particular,  $\mathcal{H}_\epsilon$  depends on a constant number of coordinates (depending on  $t, p$ ).

Notice that  $\mu_p(\mathcal{G}_\epsilon \setminus \mathcal{F}'_\epsilon) \leq \sqrt{\epsilon}$  (since each fiber we retained had  $\mu_p$ -measure at least  $1 - \sqrt{\epsilon}$ ), and so  $\mu_p(\mathcal{G}_\epsilon \triangle \mathcal{F}) = O(\sqrt{\epsilon})$ . Thus  $\mu_p(\mathcal{F} \triangle \mathcal{H}_\epsilon) = O(\epsilon^{O(1)})$ . In particular, if  $\epsilon_1, \epsilon_2 \leq \epsilon$  we get  $\mu_p(\mathcal{H}_{\epsilon_1} \triangle \mathcal{H}_{\epsilon_2}) =$

$O(\epsilon^{O(1)})$ . Since  $\mathcal{H}_{\epsilon_1} \triangle \mathcal{H}_{\epsilon_2}$  depends on a constant number of coordinates, for small enough  $\epsilon$  this forces  $\mathcal{H}_{\epsilon_1} = \mathcal{H}_{\epsilon_2}$ . There is therefore a Frankl family  $\mathcal{H}$  satisfying  $\mu_p(\mathcal{F} \triangle \mathcal{H}) = O(\epsilon^{O(1)})$  for all small enough  $\epsilon > 0$ . Taking the limit  $\epsilon \rightarrow 0$  concludes the proof.  $\square$

We now consider the case  $p = 1/2$ . We thank Shay Moran for help with the proof of the following lemma.

**Lemma 4.5.** *For all  $t \geq 1$ ,  $w_\infty(t, 1/2) = 1/2$ . Furthermore, if  $t \geq 2$  then no  $t$ -intersecting family on infinitely many points has  $\mu_p$ -measure  $1/2$ .*

*Proof.* If  $\mathcal{F}$  is an intersecting family on infinitely many points then  $\mathcal{F}$  is disjoint from  $\{\bar{A} : A \in \mathcal{F}\}$ . Since both families have the same measure, it follows that  $\mu_p(\mathcal{F}) \leq 1/2$ . Thus  $w_\infty(t, 1/2) \leq 1/2$ . On the other hand, clearly  $w_\infty(t, 1/2) \geq w_{\text{sup}}(t, 1/2)$ , and so  $w_\infty(t, 1/2) = 1/2$  due to Corollary 2.4.

Suppose now that  $t \geq 2$  and that  $\mathcal{F}$  is a  $t$ -intersecting family on infinitely many points of  $\mu_{1/2}$ -measure  $1/2$ . By possibly taking the up-set of  $\mathcal{F}$ , we can assume that  $\mathcal{F}$  is monotone (see Definition 2.3 for the relevant definitions). Let  $\mathcal{F}_- = \{A \in 2^{\mathbb{N} \setminus \{1\}} : A \in \mathcal{F}\}$  and  $\mathcal{F}_+ = \{A \in 2^{\mathbb{N} \setminus \{1\}} : A \cup \{1\} \in \mathcal{F}\}$ . Since  $\mathcal{F}$  is monotone,  $\mathcal{F}_- \subseteq \mathcal{F}_+$ . Since  $t \geq 2$ ,  $\mathcal{F}_+$  is intersecting, and so  $\mu_{1/2}(\mathcal{F}_+) \leq 1/2$ . It follows that  $\mu_{1/2}(\mathcal{F}_+) = \mu_{1/2}(\mathcal{F}_-) = 1/2$  as well, and so  $\mathcal{F}_1 = \{A \in 2^{\mathbb{N}} : A \cap (\mathbb{N} \setminus \{1\}) \in \mathcal{F}_-\}$  also has  $\mu_{1/2}$ -measure  $1/2$ . Note that  $\mathcal{F}_1$  does not depend on 1. Note also that  $\mathcal{F}_1 \subseteq \mathcal{F}$  and that  $\mathcal{F}_1$  is  $t$ -intersecting (since  $\mathcal{F}_-$  is).

Starting with  $\mathcal{F}_1$  but working with the element 2 instead of 1, construct a  $t$ -intersecting family  $\mathcal{F}_2 \subseteq \mathcal{F}_1$  of  $\mu_{1/2}$ -measure  $1/2$ , and note that  $\mathcal{F}_2$  depends on neither 1 nor 2. Continuing in this way, we construct a sequence of families  $\mathcal{F}_n$  such that  $\mathcal{F}_n \subseteq \mathcal{F}_{n-1}$  is a  $t$ -intersecting family of  $\mu_{1/2}$ -measure  $1/2$  that does not depend on  $1, \dots, n$ . Thus  $\mathcal{F}' = \bigcap_{n \in \mathbb{N}} \mathcal{F}_n$  is a  $t$ -intersecting family of  $\mu_{1/2}$ -measure  $1/2$  which represents a tail event. However, such a family must have measure 0 or 1 due to Kolmogorov's zero-one law. This contradiction shows that the original family  $\mathcal{F}$  could not have existed.  $\square$

Finally, we dispense of the case  $p > 1/2$ .

**Lemma 4.6.** *There exists a family  $\mathcal{F}$  on infinitely many points which is  $\aleph_0$ -intersecting and satisfies  $\mu_p(\mathcal{F}) = 1$  for all  $p \in (1/2, 1)$ .*

*Proof.* Let  $f(r)$  be any integer function which is  $\omega(1)$  and  $o(r)$ , for example  $\lfloor \sqrt{r} \rfloor$ . We define  $\mathcal{F}$  to consist of all sets  $A$  such that  $|A \cap [2r + f(r)]| \geq r + f(r)$  for all large enough  $r$ . For any  $A, B \in \mathcal{F}$  it holds that for large enough  $r$ ,  $|A \cap B \cap [2r + f(r)]| \geq 2(r + f(r)) - (2r + f(r)) = f(r)$ , and so the intersection  $A \cap B$  is infinite.

Hoeffding's bound states (in one version) that  $\Pr[\text{Bin}(n, p) \leq qn] \leq e^{-2n(p-q)^2}$ . This implies that for  $A \sim \mu_p$ ,

$$\Pr[|A \cap [2r + f(r)]| < r + f(r)] \leq \exp \left[ -2(2r + f(r)) \left( p - \frac{r + f(r)}{2r + f(r)} \right)^2 \right].$$

Since  $(r + f(r))/(2r + f(r)) \rightarrow 1/2$ , there exists  $\epsilon > 0$  such that for large enough  $r$  we have  $\Pr[|A \cap [2r + f(r)]| < r + f(r)] \leq e^{-4r\epsilon^2}$ . This shows that  $\sum_r \Pr[|A \cap [2r + f(r)]| < r + f(r)]$  converges. The Borell–Cantelli lemma thus shows that almost surely, only finitely many of these “bad events” happen, and so  $\mu_p(\mathcal{F}) = 1$ .  $\square$

## 5 Ahlswede–Khachatrian theorem for the Hamming scheme

Chung, Frankl, Graham, and Shearer [8] considered the difference between *intersecting* and *agreeing* families. Let us say that a family  $\mathcal{F}$  on  $n$  points is  $\mathcal{G}$ -*intersecting* if any  $A, B \in \mathcal{F}$  satisfy  $A \cap B \in \mathcal{G}$ , where  $\mathcal{G}$  is some monotone family on  $n$  points. Thus  $t$ -intersecting families are  $\mathcal{G}$ -intersecting for  $\mathcal{G} = \{S \subseteq [n] : |S| \geq t\}$ . A family  $\mathcal{F}$  is  $\mathcal{G}$ -*agreeing* if any  $A, B \in \mathcal{F}$  satisfy  $\overline{A \triangle B} \in \mathcal{G}$ . Since  $\overline{A \triangle B} = (A \cap B) \cup (\overline{A} \cap \overline{B})$ , every  $\mathcal{G}$ -intersecting family is a fortiori  $\mathcal{G}$ -agreeing, but the converse does not hold in general. Nevertheless, Chung et al. showed that the maximum  $\mu_{1/2}$ -measure of a  $\mathcal{G}$ -agreeing family is the same as the maximum  $\mu_{1/2}$ -measure of a  $\mathcal{G}$ -intersecting family, using a simple shifting argument. This explains why many arguments, such as ones using Shearer's lemma, seem to work not

only for  $\mathcal{G}$ -intersecting families but also for  $\mathcal{G}$ -agreeing families. The latter are easier to work with since the definition is more symmetric, and this is taken to full advantage in [11], for example.

Ahlsweede and Khachatrian [4], motivated by the geometry of Hamming spaces, considered a more general question: what is the largest subset of  $\mathbb{Z}_m^n$  of diameter  $n - t$  with respect to Hamming distance? Two vectors  $x, y \in \mathbb{Z}_m^n$  have Hamming distance  $n - t$  if they agree on exactly  $t$  coordinates. Thus a subset of diameter  $n - t$  is the same as a collection of vectors, every two of which agree on at least  $t$  coordinates. They showed that this corresponds (roughly) to  $\mu_p$  for  $p = 1/m$  (though they stated this in a different language).

Motivated by the recent success of Fourier-analytic techniques to analyze questions in extremal combinatorics, Alon, Dinur, Friedgut and Sudakov [7] studied independent sets in graph products, and in particular rederived the results of [4] for the case  $t = 1$ . Shinkar [26] extended this to general  $t$  and  $m \geq t + 1$ , using methods of Friedgut [18]. Unfortunately, these spectral techniques cannot at the moment yield all results of [4].

Any bound on  $t$ -agreeing families of vectors in  $\mathbb{Z}_m^n$  readily yields matching results on the  $\mu_{1/m}$ -measure of  $t$ -intersecting families on  $n$  points. Indeed, given a family  $\mathcal{F} \subseteq \{0, 1\}^n$ , let  $\mathcal{G} \subseteq \mathbb{Z}_m^n$  consist of all vectors obtained by taking each binary vector in  $\mathcal{F}$  and replacing each 0 coordinate with one of the  $m - 1$  values  $\{2, \dots, m\}$ . The resulting family is  $t$ -agreeing, and  $|\mathcal{G}| = m^n \mu_{1/m}(\mathcal{F})$ . Therefore a bound on  $|\mathcal{G}|$  yields a bound on  $\mu_{1/m}(\mathcal{F})$ . In other words,  $t$ -agreeing families give a discrete model for  $\mu_{1/m}$ .

We extend these results in Section 5.1, by giving a discrete model for  $\mu_{s/m}$  for any integers  $s, m$  satisfying  $s/m < 1/2$ . Extending the work of [4], we prove an Ahlsweede–Khachatrian theorem in this setting. Section 5.2 describes a similar continuous model for  $\mu_p$  for any  $p < 1/2$ , and proves an Ahlsweede–Khachatrian theorem in that setting.

## 5.1 Discrete setting

We start by defining the new discrete model.

**Definition 5.1.** A family on  $\mathbb{Z}_m^n$  is a subset of  $\mathbb{Z}_m^n$ . For  $s \leq m/2$ , a family  $\mathcal{F}$  on  $\mathbb{Z}_m^n$  is  $t$ -agreeing up to  $s$  if every  $x, y \in \mathcal{F}$  have  $t$  coordinates  $i_1, \dots, i_t$  such that  $x_{i_j} - y_{i_j} \in \{-(s-1), \dots, s-1\}$  for  $1 \leq j \leq t$ .

The uniform measure on  $\mathbb{Z}_m$  is denoted by  $\mu_m$ , and the uniform measure on  $\mathbb{Z}_m^n$  by  $\mu_m^n$ .

The maximum measure of a  $t$ -agreeing up to  $s$  family on  $\mathbb{Z}_m^n$  is denoted  $w(\mathbb{Z}_m^n, t, s)$ .

Our goal is to show that  $w(\mathbb{Z}_m^n, t, s) = w(n, t, s/m)$  whenever  $s/m \leq 1/2$ , and to identify the optimal families when possible. Our candidate optimal families are obtained in the following way.

**Definition 5.2.** Fix  $n, m \geq 1$  and  $s \leq m/2$ . For  $y \in \mathbb{Z}_m^n$ , the mapping  $\sigma_y: \mathbb{Z}_m^n \rightarrow 2^{[n]}$  is given by  $\sigma_y(x) = \{i \in [n] : 1 \leq x_i - y_i \leq s\}$ .

If  $\mathcal{F}$  is a family on  $\mathbb{Z}_m^n$  and  $\mathcal{G}$  is a family on  $n$  points then  $\mathcal{F} \approx \mathcal{G}$  (or,  $\mathcal{F}$  is equivalent to  $\mathcal{G}$ ) if  $\mathcal{F} = \sigma_y^{-1}(\mathcal{G})$  for some vector  $y \in \mathbb{Z}_m^n$ . In words,  $\mathcal{F}$  arises from  $\mathcal{G}$  by going through all sets  $S \in \mathcal{G}$ , replacing each  $i \in S$  by all values in  $\{y_i + 1, \dots, y_i + s\}$ , and each  $i \notin S$  by all other values.

We say that  $\mathcal{F}$  is equivalent to a  $(t, r)$ -Frankl family if  $\mathcal{F} \approx \mathcal{G}$  for some family  $\mathcal{G}$  equivalent to a  $(t, r)$ -Frankl family (see Definition 2.6).

We will prove the following theorem.

**Theorem 5.1.** Let  $n, m, t \geq 1$  and  $s \leq m/2$ . Then  $w(\mathbb{Z}_m^n, t, s) = w(n, t, s/m)$ . Furthermore, if  $s < m/2$  and  $\mathcal{F}$  is a family on  $\mathbb{Z}_m^n$  of measure  $w(n, t, s/m)$  then  $\mathcal{F}$  is equivalent to a  $t$ -intersecting family on  $n$  points of  $\mu_{s/m}$ -measure  $w(n, t, s/m)$ . The same holds if  $m = 2$ ,  $s = 1$  and  $t > 1$ .

When  $s = m/2$  and  $s > 1$  there can be exotic families of maximum measure. For example, when  $s = 2$  and  $m = 4$  the family  $\{00, 01, 12, 13, 20, 21, 32, 33\}$  is intersecting up to 2 but doesn't arise from any family on two points. Adding another coordinate ranging over  $\{0, 1\}$ , we get a 2-intersecting up to 2 family which doesn't arise from any family on three points.

When  $m = 2$ ,  $s = 1$  and  $t = 1$ , a maximum measure family is one which contains exactly one vector of each pair of complementary vectors.

On the way toward proving the theorem, we will need to consider hybrid families in which some of the coordinates come from  $\mathbb{Z}_m$ , and others from  $\{0, 1\}$ .

**Definition 5.3.** A family on  $\mathbb{Z}_m^n \times \{0, 1\}^\ell$  is a subset of  $\mathbb{Z}_m^n \times \{0, 1\}^\ell$ . Such a family  $\mathcal{F}$  is *t-agreeing up to s* if every  $x, y \in \mathcal{F}$  have  $t$  coordinates  $i_1, \dots, i_t$  such that  $x_{i_j} - y_{i_j} \in \{-(s-1), \dots, s-1\}$  (if  $i_j \leq n$ ) or  $x_{i_j} = y_{i_j} = 1$  (if  $i_j > n$ ) for  $1 \leq j \leq t$ .

More generally, two vectors  $x, y \in \mathbb{Z}_m^n \times \{0, 1\}^\ell$  *s-agree* on a coordinate  $i \leq n$  if  $x_i - y_i \in \{-(s-1), \dots, s-1\}$ , and they *s-agree* on a coordinate  $i > n$  if  $x_i = y_i = 1$ . Thus a family on  $\mathbb{Z}_m^n \times \{0, 1\}^\ell$  is *t-agreeing up to s* if every  $x, y \in \mathcal{F}$  *s-agree* on at least  $t$  coordinates.

We measure families on  $\mathbb{Z}_m^n \times \{0, 1\}^\ell$  using the product measures  $\mu_{s,m}^{n,\ell} = \mu_m^n \times \mu_{s/m}^\ell$ .

For  $y \in \mathbb{Z}_m^n$ , the mapping  $\sigma_y: \mathbb{Z}_m^n \times \{0, 1\}^\ell \rightarrow 2^{[n+\ell]}$  is the product of  $\sigma_y$  and the identity mapping.

If  $\mathcal{F}$  is a family on  $\mathbb{Z}_m^n \times \{0, 1\}^\ell$  and  $\mathcal{G}$  is a family on  $n + \ell$  points then  $\mathcal{F} \approx \mathcal{G}$  if  $\mathcal{F} = \sigma_y^{-1}(\mathcal{G})$  for some vector  $y \in \mathbb{Z}_m^n$ . In words,  $\mathcal{F}$  arises from  $\mathcal{G}$  by going through all sets  $S \in \mathcal{F}$ , and for each  $i \leq n$ , replacing each  $i \in S$  by all values in  $\{y_i + 1, \dots, y_i + s\}$ , and each  $i \notin S$  by all other values.

We start by proving  $w(\mathbb{Z}_m^n, t, s) = w(n, t, s/m)$ . The identification of optimal families will require further refining the proof, but we present the two proofs separately for clarity. We start with a technical result about the stable set polytope.

**Proposition 5.2.** Let  $G = (V, E)$  be a graph, and let  $\mu: V \rightarrow \mathbb{R}$  be a set of weights. Consider the program

$$\begin{aligned} \max \quad & \sum_{x \in V} \mu_x v_x \\ \text{s.t.} \quad & 0 \leq v_x \leq 1 \quad \text{for all } x \in V \\ & v_x + v_y \leq 1 \quad \text{for all } (x, y) \in E \end{aligned}$$

The maximum of the program is attained (not necessarily uniquely) at some half-integral point (a point in which all entries are  $\{0, 1/2, 1\}$ ).

Moreover, if there is a unique half-integral point at which the objective is maximized, then this point is the unique maximum.

*Proof.* The stable set polytope of  $G$  is defined as the set of all  $V$ -indexed vectors which satisfy the constraints stated in the program. Every linear functional over the polytope is maximized at some vertex. Moreover, if there is a unique vertex at which the maximum is attained then this vertex is the unique maximum. Therefore the proposition follows from the fact that each vertex of the polytope is half-integral. This classical fact is due to Nemhauser and Trotter [24]. We briefly sketch the proof below, following Schrijver [25, Theorem 64.7].

Let  $v$  be a vertex of the stable set polytope. Let  $I = \{x : 0 < v_x < 1/2\}$  and  $J = \{y : 1/2 < v_y < 1\}$ . For small enough  $\epsilon > 0$ , both  $v + \epsilon(1_I - 1_J)$  and  $v - \epsilon(1_I - 1_J)$  are in the stable set polytope (where  $1_I$  is the characteristic function of  $I$ ), and so  $v$  can only be a vertex if  $I = J = \emptyset$ .  $\square$

The heart of the proof of  $w(\mathbb{Z}_m^n, t, s) = w(n, t, s/m)$  is the following lemma.

**Lemma 5.3.** Let  $\mathcal{F}$  be a family on  $\mathbb{Z}_m^n \times \{0, 1\}^\ell$  which is *t-agreeing up to s*, where  $s \leq m/2$  and  $n \geq 1$  (but possibly  $\ell = 0$ ). There is a family  $\mathcal{H}$  on  $\mathbb{Z}_m^{n-1} \times \{0, 1\}^{\ell+1}$  which is *t-agreeing up to s* and satisfies  $\mu_{s,m}^{n-1,\ell+1}(\mathcal{H}) \geq \mu_{s,m}^{n,\ell}(\mathcal{F})$ .

*Proof.* The first step is to decompose  $\mathcal{F}$  according to the values of all coordinates but the  $n$ 'th one:

$$\mathcal{F} = \bigcup_{x_1 \in \mathbb{Z}_m^{n-1}} \bigcup_{x_2 \in \{0, 1\}^\ell} \{x_1\} \times \mathcal{F}_{x_1, x_2} \times \{x_2\}.$$

It will be convenient to refer to the pair  $x_1, x_2$  as a single vector  $x = (x_1, x_2) \in \mathbb{Z}_m^{n-1} \times \{0, 1\}^\ell$ .

We now construct a graph  $G = (V, E)$  as follows. The vertices are  $V = \{x \in \mathbb{Z}_m^{n-1} \times \{0, 1\}^\ell : \mathcal{F}_x \neq \emptyset\}$ . We connect two vertices  $x, y$  (possibly  $x = y$ ) if  $\{x, y\}$  is *not t-agreeing up to s*; in other words, if the vectors  $x, y$  *s-agree* on exactly  $t - 1$  coordinates. If  $(x, y) \in E$  then the sets  $\mathcal{F}_x, \mathcal{F}_y \subseteq \mathbb{Z}_m$  must be *s-agreeing*, in the terminology of Definition 3.1.

There are now two cases to consider: the degenerate case  $n + \ell = t$ , and the non-degenerate case  $n + \ell > t$ . In the degenerate case the graph  $G$  consists of isolated vertices with self-loops, since the only way that vectors  $x, y$  of length  $t - 1$  can *s-agree* on  $t - 1$  coordinates is if  $x = y$ . Thus each set  $\mathcal{F}_x$  must be *s-agreeing*, and so  $|\mathcal{F}_x| \leq s$  by Lemma 3.3. We form the new family  $\mathcal{H}$  as follows:

$$\mathcal{H} = \bigcup_{x_1 \in \mathbb{Z}_m^{n-1}} \bigcup_{x_2 \in \{0, 1\}^\ell} \{x_1\} \times \mathcal{H}_{x_1, x_2} \times \{x_2\}, \quad \mathcal{H}_x = \begin{cases} \{1\} & \text{if } \mathcal{F}_x \neq \emptyset, \\ \emptyset & \text{otherwise.} \end{cases}$$

By construction,  $\mu_{s,m}^{1,0}(\mathcal{F}_x) \leq \mu_{s,m}^{0,1}(\mathcal{H}_x)$ : either both sets are empty, or  $\mathcal{F}_x$  has measure at most  $s/m$ , while  $\mathcal{H}_x$  has measure  $s/m$ . Thus  $\mu_{s,m}^{n,\ell}(\mathcal{F}) \leq \mu_{s,m}^{n-1,\ell+1}(\mathcal{H})$ . Moreover, it is not hard to verify that  $\mathcal{H}$  is  $t$ -agreeing up to  $s$ . This completes the proof in the degenerate case.

The proof in the non-degenerate case is more complicated. First, notice that we no longer have self-loops, since every  $x \in V$  agrees with itself on  $n - 1 + \ell \geq t$  coordinates. If  $(x, y) \in E$  then the non-empty sets  $\mathcal{F}_x, \mathcal{F}_y$  are  $s$ -agreeing, and so  $|\mathcal{F}_x| + |\mathcal{F}_y| \leq 2s$  by Lemma 3.3. In particular, if  $|\mathcal{F}_x| \geq 2s$  then  $x$  must be an isolated vertex. This suggests pruning the graph by removing all isolated vertices. We denote the resulting graph by  $G' = (V', E')$ ; note that  $V'$  could be empty. All vertices  $x \in V'$  now satisfy  $0 < |\mathcal{F}_x| < 2s$ .

Let  $v_x = |\mathcal{F}_x|/(2s)$ . Then  $0 \leq v_x \leq 1$ ,  $v_x + v_y \leq 1$  for all edges, and

$$\mu_{s,m}^{n,\ell}(\mathcal{F}) = \sum_{x \in V \setminus V'} \mu_{s,m}^{n-1,\ell}(x) \mu_m(\mathcal{F}_x) + \frac{2s}{m} \sum_{x \in V'} \mu_{s,m}^{n-1,\ell}(x) v_x.$$

Proposition 5.2 states that there exists a  $\{0, 1/2, 1\}$ -valued vector  $w$  which satisfies  $0 \leq w_x \leq 1$ ,  $w_x + w_y \leq 1$  for all edges, and  $\sum_{x \in V'} \mu_{s,m}^{n-1,\ell}(x) v_x \leq \sum_{x \in V'} \mu_{s,m}^{n-1,\ell}(x) w_x$ . We will construct  $\mathcal{H}$  according to this vector:

$$\mathcal{H} = \bigcup_{x_1 \in \mathbb{Z}_m^{n-1}} \bigcup_{x_2 \in \{0,1\}^\ell} \{x_1\} \times \mathcal{H}_{x_1, x_2} \times \{x_2\}, \quad \mathcal{H}_x = \begin{cases} \{0, 1\} & \text{if } x \in V \setminus V', \\ \{0, 1\} & \text{if } x \in V' \text{ and } w_x = 1, \\ \{1\} & \text{if } x \in V' \text{ and } w_x = 1/2, \\ \emptyset & \text{if } x \in V' \text{ and } w_x = 0, \\ \emptyset & \text{if } x \notin V'. \end{cases}$$

Note that the *support* of  $\mathcal{H}$ , which is the set of  $x \in \mathbb{Z}_m^{n-1} \times \{0, 1\}^\ell$  such that  $\mathcal{H}_x \neq \emptyset$ , is a subset of the support of  $\mathcal{F}$ .

We start by showing that  $\mathcal{H}$  is  $t$ -agreeing up to  $s$ . Let  $(x_1, \alpha, x_2), (y_1, \beta, y_2) \in \mathcal{H}$ , where  $x_1, y_1 \in \mathbb{Z}_m^{n-1}$ ,  $\alpha, \beta \in \{0, 1\}$ , and  $x_2, y_2 \in \{0, 1\}^\ell$ . If there is no edge between  $(x_1, x_2)$  and  $(y_1, y_2)$  then clearly  $(x_1, \alpha, x_2), (y_1, \beta, y_2)$   $s$ -agree on at least  $t$  coordinates. If there does exist an edge then  $x = (x_1, x_2)$  and  $y = (y_1, y_2)$  agree on  $t - 1$  coordinates, and so it suffices to show that  $\alpha = \beta = 1$ . By construction,  $0 \in \mathcal{H}_x$  if either  $x \in V \setminus V'$  or  $w_x = 1$ . In the former case, since the support of  $\mathcal{H}$  is contained in the support of  $\mathcal{F}$ ,  $x$  is isolated, and so there cannot be an edge  $(x, y)$ . In the latter case,  $w_y = 0$  (since  $w_x + w_y \leq 1$  for all edges), and so  $y$  is not in the support of  $\mathcal{H}$ . We conclude that  $\mathcal{H}$  is  $t$ -agreeing up to  $s$ .

We proceed by comparing the measures of  $\mathcal{F}$  and  $\mathcal{H}$ :

$$\begin{aligned} \mu_{s,m}^{n-1,\ell+1}(\mathcal{H}) &= \sum_{x \in \mathbb{Z}_m^{n-1} \times \{0,1\}^\ell} \mu_{s,m}^{n-1,\ell}(x) \mu_{s/m}(\mathcal{H}_x) \\ &= \sum_{x \in V \setminus V'} \mu_{s,m}^{n-1,\ell}(x) + \sum_{\substack{x \in V' \\ w_x = 1}} \mu_{s,m}^{n-1,\ell}(x) + \frac{s}{m} \sum_{\substack{x \in V' \\ w_x = 1/2}} \mu_{s,m}^{n-1,\ell}(x). \end{aligned}$$

The first term is at least as large as  $\sum_{x \in V \setminus V'} \mu_{s,m}^{n-1,\ell}(x) \mu_m(\mathcal{F}_x)$ , and the other two terms are at least  $(2s/m) \sum_{x \in V'} \mu_{s,m}^{n-1,\ell}(x) w_x$ . Therefore

$$\begin{aligned} \mu_{s,m}^{n-1,\ell+1}(\mathcal{H}) &\geq \sum_{x \in V \setminus V'} \mu_{s,m}^{n-1,\ell}(x) \mu_m(\mathcal{F}_x) + \frac{2s}{m} \sum_{x \in V'} \sum_{x \in V'} \mu_{s,m}^{n-1,\ell}(x) w_x \\ &\geq \sum_{x \in V \setminus V'} \mu_{s,m}^{n-1,\ell}(x) \mu_m(\mathcal{F}_x) + \frac{2s}{m} \sum_{x \in V'} \sum_{x \in V'} \mu_{s,m}^{n-1,\ell}(x) v_x \\ &= \mu_{s,m}^{n-1,\ell+1}(\mathcal{F}). \end{aligned}$$

This completes the proof.

We note that when  $(s, m) = (1, 2)$ , necessarily  $v_x = 1/2$  for all  $x \in V'$  (since  $0 < 2v_x < 2$  is an integer), and so we can use  $w = v$ .  $\square$

As a corollary, we can deduce the part  $w(\mathbb{Z}_m^n, t, s) = w(n, t, s/m)$  of Theorem 5.1.

**Lemma 5.4.** *Let  $n \geq t \geq 1$ ,  $m \geq 2$  and  $s \leq m/2$ . Then  $w(\mathbb{Z}_m^n, t, s) = w(n, t, s/m)$ .*

*Proof.* Let  $\mathcal{F}$  be a family on  $\mathbb{Z}_m^n$  which is  $t$ -intersecting up to  $s$ . Applying Lemma 5.3  $n$  times, we obtain a  $t$ -intersecting family  $\mathcal{H}$  on  $\{0, 1\}^n$  satisfying  $\mu_m(\mathcal{F}) \leq \mu_{s/m}(\mathcal{H}) \leq w(n, t, s/m)$ . This shows that  $w(\mathbb{Z}_m^n, t, s) \leq w(n, t, s/m)$ .

For the other direction, let  $\mathcal{H}$  be a  $t$ -intersecting family on  $n$  points satisfying  $\mu_{s/m}(\mathcal{H}) = w(n, t, s/m)$ . Define  $\mathcal{F} = \sigma_0^{-1}(\mathcal{H})$ , where  $0$  is the zero vector. It is straightforward to verify that  $\mathcal{F}$  is  $t$ -agreeing up to  $s$  and that  $\mu_m(\mathcal{F}) = \mu_{s/m}(\mathcal{H}) = w(n, t, s/m)$ . It follows that  $w(\mathbb{Z}_m^n, t, s) \geq w(n, t, s/m)$ .  $\square$

We proceed with the characterization of families achieving the bound  $w(\mathbb{Z}_m^n, t, s)$ . We start with the case  $s/m < 1/2$ . The idea is to prove a stronger version of Lemma 5.3 which states that if  $\mathcal{F}$  and  $\mathcal{H}$  have the same measure and  $\mathcal{H}$  is equivalent to a Frankl family then so is  $\mathcal{F}$ . This suggests analyzing the graph  $G = (V, E)$  constructed in the proof of Lemma 5.3 in the case of a family equivalent to a Frankl family.

**Lemma 5.5.** *Let  $n, t \geq 1$  and  $\ell, s, m$  be given, and suppose that  $s < m/2$ . Let  $\mathcal{F}$  be a family on  $\mathbb{Z}_m^n \times \{0, 1\}^\ell$  which is equivalent to some  $(t, r)$ -Frankl family, for some  $r \geq 0$ . The graph  $G' = (V', E')$  constructed in the proof of Lemma 5.3 is either empty (has no vertices) or is connected and non-bipartite.*

*The same holds when  $s = 1$  and  $m = 2$ .*

*Proof.* Fix  $m, t, r$ . We first consider the case  $s < m/2$ .

We will first prove the lemma for  $n = 1$  and  $\ell = t + 2r - 1$ . Then we will prove it for  $n = 1$  and arbitrary  $\ell$ . Finally, we will tackle the general case.

Recall that the graph  $G'$  was constructed by first constructing a larger graph  $G = (V, E)$  and then removing all isolated vertices (vertices having no edges). We will assume throughout that  $\mathcal{F}$  depends on the  $n$ th coordinate (that is, the underlying Frankl family has  $n$  in its support), since otherwise  $G$  has no edges and so  $G'$  is empty.

We will prove one more property of  $G'$ , which we call *property Z*. This property states that if  $x \in V'$  and  $x_i = 0$  for  $i > n - 1$  (that is,  $i$  is a coordinate taking values in  $\{0, 1\}$ ) then  $x$  is connected to some  $y \in V'$  with  $y_i = 1$ .

We start with the case  $(n, \ell) = (1, t + 2r - 1)$ . If  $r = 0$  then  $G$  consists of a single vertex  $[t - 1]$  with a self-loop (we identify zero-one vectors with sets), and so it is connected and non-bipartite. If  $r > 0$  then  $V = \binom{[t+2r-1]}{\geq t+r-1}$ . If  $A \in \binom{[t+2r-1]}{\geq t+r-1}$  and  $B \in \binom{[t+2r-1]}{\geq t+r}$  then  $|A \cap B| \geq (t + r - 1) + (t + r) - (t + 2r - 1) = t$ , and so  $V' \subseteq \binom{[t+2r-1]}{t+r-1}$ ; symmetry dictates that  $V' = \binom{[t+2r-1]}{t+r-1}$ . Moreover, the sequence

$$\{1, \dots, t - 1, t, \dots, t + r - 1\}, \{1, \dots, t - 1, t + r, \dots, t + 2r - 1\}, \{2, \dots, t - 1, t, \dots, t + r\}$$

corresponds to a path connecting  $\{1, \dots, t + r - 1\}$  and  $\{2, \dots, t + r\}$ . This shows that  $G'$  is connected. Since the path has even length, we conclude that there is a path of even length connecting any two vertices. In particular, there is a path of even length connecting  $\{1, \dots, t - 1, t, \dots, t + r - 1\}$  and  $\{1, \dots, t - 1, t + r, \dots, t + 2r - 1\}$ , and together with the corresponding edge, we obtain an odd cycle. This shows that  $G'$  is non-bipartite.

To prove property Z, we consider two cases. When  $r = 0$ , property Z holds vacuously, since the only vertex contains no zero coordinates. When  $r > 0$ , we can assume without loss of generality that  $x = \{2, \dots, t + r\}$  and  $i = 1$ . In that case,  $x$  is connected to  $y = \{1, \dots, t - 1, t + r, \dots, t + 2r - 1\}$ .

We now prove by induction the case  $n = 1$  and  $\ell \geq t + 2r - 1$ . Let  $G_\ell, G'_\ell$  be the graphs corresponding to a particular value of  $\ell$ , where we assume without loss of generality that  $\mathcal{F}$  depends on the first  $t + 2r$  points (we can make this assumption since  $\mathcal{F}$  has to depend on the first point). Suppose that we have shown that  $G'_\ell$  is connected and non-bipartite. Notice that  $V_{\ell+1} = V \times \{0, 1\}$  and  $E_{\ell+1} = \{((x, i), (y, j)) : (x, y) \in E_\ell, (i, j) \neq (1, 1)\}$ . This shows that  $V'_{\ell+1} = V'_\ell \times \{0, 1\}$ . Since  $G'$  is connected and non-bipartite, there is an even-length path connecting any two  $x, y \in V'$ , which lifts to even-length paths connecting  $(x, 0), (y, 0)$  and  $(x, 1), (y, 1)$  (flipping the extra coordinate at each step); and an odd-length path connecting  $x, y \in V'$  which lifts to paths connecting  $(x, 0), (y, 1)$  and  $(x, 1), (y, 0)$ . This shows that  $G'_{\ell+1}$  is connected. An odd-length path from  $x$  to itself lifts to an odd-length path from  $(x, 0)$  to itself (flipping the extra coordinate at each step but the first), showing that  $G'$  is non-bipartite.

To prove property Z, let  $x \in V'_{\ell+1}$  have  $x_i = 0$ . If  $i < \ell + 1$  then property Z for  $G'_\ell$  implies the existence of a neighbor  $y$  with  $y_i = 1$ . If  $i = \ell + 1$  then  $x$  is connected to some  $y$  (since  $x$  is not isolated), and since

$V'_{\ell+1}$  is independent of the last coordinate, we can assume that  $y_i = 1$  (the value of this coordinate does not change the number of coordinates on which  $x$  and  $y$  agree).

Suppose now that the lemma holds for families on  $\mathbb{Z}_m^n \times \{0, 1\}^\ell$ . We will show that it holds for families on  $\mathbb{Z}_m^{n+1} \times \{0, 1\}^{\ell-1}$  as well. We denote the relevant graphs  $G_n, G'_n$  and  $G_{n+1}, G'_{n+1}$ , respectively. Without loss of generality we can assume that

$$V_{n+1} = \bigcup_{x_1 \in \mathbb{Z}_m^{n-1}} \bigcup_{x_2 \in \{0, 1\}^{\ell-1}} \{x_1\} \times \sigma_0^{-1}(\{x' : (x_1, x', x_2) \in V_n\}) \times \{x_2\}.$$

In words,  $V_{n+1}$  is obtained by applying  $\sigma_0^{-1}$  to the  $n$ 'th coordinate.

We claim that  $V'_{n+1}$  is obtained from  $V'_n$  in the same way, showing that property Z holds for  $G_{n+1}$ . Indeed, suppose that  $(x_1, x', x_2) \in V'_n$  is not isolated, say it is connected to  $(y_1, y', y_2) \in V'_n$ , and let  $x'' \in \sigma_0^{-1}(x')$ . If  $x' = y' = 1$  then  $(x_1, x'', x_2)$  is connected to  $(y_1, x'', y_2)$ . If  $x' = 1$  and  $y' = 0$  then  $(x_1, x'', x_2)$  is connected to  $(y_1, x'' + s, y_2)$  (note  $s + 1 \leq x'' + s \leq 2s$ ). If  $x' = 0$  then property Z allows us to assume that  $y' = 1$ , and so  $(x_1, x'', x_2)$  is connected to  $(y_1, y'', y_2)$  for some  $1 \leq y'' \leq s$ ; indeed, Lemma 3.2 shows that the only elements  $s$ -agreeing with all of  $1, \dots, s$  are  $1, \dots, s$ .

Toward showing that  $V'_{n+1}$  is connected, we prove first that  $(x_1, a, x_2) \in V'_{n+1}$  is connected to  $(x_1, b, x_2) \in V'_{n+1}$ , by an even-length path, whenever  $\sigma_0(a) = \sigma_0(b)$ . Suppose first that  $\sigma_0(a) = \sigma_0(b) = 1$ . Let  $(y_1, y', y_2) \in V'_n$  be a neighbor of  $(x_1, 1, x_2)$  in  $G'_n$ . If  $\sigma_0(y') = 1$  then  $(x_1, a, x_2), (y_1, a, y_2), (x_1, b, x_2)$  is a path in  $G'_{n+1}$ . If  $\sigma_0(y) = 0$  then consider the following path in  $G'_{n+1}$ :

$$(x_1, 1, x_2), (y_1, s + 2, y_2), (x_1, 2, x_2), (y_1, s + 3, y_2), \dots, (x_1, s - 1, x_2), (y_1, 2s, y_2), (x_1, s, x_2).$$

This contains a sub-path connecting  $(x_1, a, x_2)$  and  $(x_1, b, x_2)$ .

Suppose next that  $\sigma_0(a) = \sigma_0(b) = 0$ . Property Z shows that  $(x_1, 0, x_2) \in V'_n$  has a neighbor  $(y_1, 1, y_2) \in V'_n$  in  $G'_n$ . Consider the following path in  $G'_{n+1}$ :

$$(x_1, s + 1, x_2), (y_1, 1, y_2), (x_1, s + 2, x_2), (y_1, 2, y_2), \dots, (y_1, s - 1, y_2), (x_1, 2s, x_2), (y_1, s, y_2).$$

This path shows that  $(x_1, a, x_2)$  and  $(x_1, b, x_2)$  are connected whenever  $s + 1 \leq a, b \leq 2s$ . Since  $(y_1, s, y_2)$  neighbors  $(x_1, c, x_2)$  for all  $2s \leq c \leq m$ , we deduce that  $(x_1, a, x_2)$  and  $(x_1, b, x_2)$  are connected for all  $s + 1 \leq a, b \leq m$ .

Consider now any two vertices  $(x_1, x, x_2), (y_1, y, y_2) \in V'_{n+1}$  such that  $(x_1, \sigma_0(x), x_2) \in V'_n$  neighbors  $(y_1, \sigma_0(y), y_2) \in V'_n$  in  $G'_n$ . We will show that  $(x_1, x, x_2)$  and  $(y_1, y, y_2)$  are connected in  $V'_{n+1}$  by an odd-length path by showing that  $(x_1, X, x_2)$  and  $(y_1, Y, y_2)$  are connected for some  $X, Y \in \mathbb{Z}_m$  satisfying  $\sigma_0(x) = \sigma_0(X)$  and  $\sigma_0(y) = \sigma_0(Y)$ . Such  $X, Y$  are given by the following table:

$\sigma_0(x), \sigma_0(y)$	0, 0	0, 1	1, 0	1, 1
$X, Y$	$s + 1, 2s + 1$	$s + 1, 1$	$1, s + 1$	$1, 1$

This shows that  $V'_{n+1}$  is connected. Moreover, any odd-length cycle in  $V'_n$  lifts to an odd-length cycle in  $V'_{n+1}$ , showing that  $G'_{n+1}$  is non-bipartite.

When  $s = 1$  and  $m = 2$ , notice that the general case differs from the case  $n = 1$  only by a translation of the coordinates. Since the proof of the case  $n = 1$  did not use the bound  $s < m/2$ , we conclude that the lemma holds even when  $s = 1$  and  $m = 2$ .  $\square$

Now we can prove the strengthening of Lemma 5.3.

**Lemma 5.6.** *Let  $n, t \geq 1$  and  $\ell, s, m$  be given, and suppose that  $s < m/2$ . Let  $\mathcal{F}$  be a family on  $\mathbb{Z}_m^n \times \{0, 1\}^\ell$  which is  $t$ -agreeing up to  $s$ , and let  $\mathcal{H}$  be the family on  $\mathbb{Z}_m^{n-1} \times \{0, 1\}^{\ell+1}$  constructed in Lemma 5.3. If  $\mu_{s,m}^{n-1, \ell+1}(\mathcal{H}) = \mu_{s,m}^{n, \ell}(\mathcal{F})$  and  $\mathcal{H}$  is equivalent to a  $(t, r)$ -Frankl family then  $\mathcal{F}$  is also equivalent to a  $(t, r)$ -Frankl family.*

*The same holds for  $(s, m) = (1, 2)$ , assuming that in the proof of Lemma 5.3 we use  $w_x = 1/2$  (see comment at the end of the proof).*

*Proof.* As in the proof of Lemma 5.3, we consider separately the degenerate case  $n + \ell = t$  and the non-degenerate case  $n + \ell > t$ .

When  $n + \ell = t$ ,  $\mathcal{H}$  must be equivalent to the family  $\{[t]\}$ . Without loss of generality, we can assume that  $\mathcal{H} = [s]^{n-1} \times \{1\}^{\ell+1}$  (this corresponds to the choice  $y = 0$  in the definition of equivalence). Thus  $\mathcal{F}$  has the form

$$\mathcal{F} = \bigcup_{x_1 \in [s]^{n-1}} \bigcup_{x_2 = \{1\}^\ell} \{x_1\} \times \mathcal{F}_{x_1, x_2} \times \{x_2\},$$

where each  $\mathcal{F}_{x_1, x_2}$  is  $s$ -agreeing. We remind the reader that the support of  $\mathcal{F}$  is the set of pairs  $(x_1, x_2)$  such that  $\mathcal{F}_{x_1, x_2} \neq \emptyset$ ; the support  $\mathcal{H}$  is defined in the same way.

Since  $\mu_{s,m}^{n-1, \ell+1}(\mathcal{H}) = \mu_{s,m}^{n, \ell}(\mathcal{F})$ , moreover  $|\mathcal{F}_{x_1, x_2}| = s$ , and so Lemma 3.3 shows that  $\mathcal{F}_{x_1, x_2}$  is an interval (when  $s = 1$  this is trivial). If  $(x_1, x_2), (y_1, y_2)$  are in the support of  $\mathcal{F}$  then  $\mathcal{F}_{x_1, x_2}, \mathcal{F}_{y_1, y_2}$  are cross- $s$ -agreeing, and so Lemma 3.3 shows that they are equal (again, when  $s = 1$  this is trivial). This shows that  $\mathcal{F}$  is equivalent to  $\{[t]\}$  as well.

Consider now the non-degenerate case  $n + \ell > t$ . Recall that in the proof of Lemma 5.3 we constructed two graphs,  $G = (V, E)$  and  $G' = (V', E')$ , and a vector  $w: V' \rightarrow \mathbb{R}$ , and defined  $\mathcal{H}$  in terms of this data. When computing the measure of  $\mathcal{H}$ , we used the estimate

$$\sum_{x \in V \setminus V'} \mu_{s,m}^{m-1, \ell}(x) \geq \sum_{x \in V \setminus V'} \mu_{s,m}^{m-1, \ell}(x) \mu_m(\mathcal{F}_x).$$

This estimate can only be tight if  $\mathcal{F}_x = \mathbb{Z}_m$  for  $x \in V \setminus V'$ . We also used the estimate

$$\sum_{\substack{x \in V' \\ w_x = 1}} \mu_{s,m}^{n-1, \ell}(x) + \frac{s}{m} \sum_{\substack{x \in V' \\ w_x = 1/2}} \mu_{s,m}^{n-1, \ell}(x) \geq \frac{2s}{m} \sum_{x \in V'} \mu_{s,m}^{n-1, \ell}(x) w_x.$$

If  $2s < m$ , this can only be tight if it is never the case that  $w_x = 1$ . Recall that  $w_x$  was a  $\{0, 1/2, 1\}$ -valued vector maximizing  $\sum_{x \in V'} \mu_{s,m}^{n-1, \ell}(x) w_x$  under the constraints  $0 \leq w_x \leq 1$  and  $w_x + w_y \leq 1$  whenever  $(x, y) \in E'$ . Since  $w_x \leq 1/2$  for all  $x \in V'$ , we see that in fact  $w_x = 1/2$  for all  $V'$  (since the constant  $1/2$  vector is always feasible). This shows that  $\mathcal{H}$  and  $\mathcal{F}$  have the same support.

Moreover (still assuming  $2s < m$ ), if there were a different  $\{0, 1/2, 1\}$ -valued vector maximizing  $\sum_{x \in V'} \mu_{s,m}^{n-1, \ell}(x) w_x$  then the construction of Lemma 5.3 would have shown that  $\mathcal{F}$  does not have maximum measure. We conclude that  $w_x$  is the unique  $\{0, 1/2, 1\}$ -valued maximizer, and so the unique maximizer according to Proposition 5.2.

When  $(s, m) = (1, 2)$ ,  $w_x = 1/2$  for all  $V'$  by assumption, and so  $\mathcal{H}$  and  $\mathcal{F}$  have the same support; the property proved in the preceding paragraph won't be needed in this case.

Lemma 5.5 shows that  $G'$  is either empty or connected. Recall that  $v_x = |\mathcal{F}_x|/(2s)$ . When computing the measure of  $\mathcal{H}$ , we used the estimate  $\sum_{x \in V'} \mu_{s,m}^{n-1, \ell}(x) w_x \geq \sum_{x \in V'} \mu_{s,m}^{n-1, \ell}(x) v_x$ . When  $s < m/2$ ,  $w$  is the unique maximizer, and so  $v_x = 1/2$  for all  $x \in V$ , that is,  $|\mathcal{F}_x| = s$  for all  $x \in V'$ ; when  $(s, m) = (1, 2)$ , the same trivially holds. For any two vertices  $x, y \in V'$  connected by an edge, the sets  $\mathcal{F}_x, \mathcal{F}_y$  must be  $s$ -agreeing, and so Lemma 3.3 shows that  $\mathcal{F}_x = \mathcal{F}_y$  (when  $s = 1$  this is trivial). Since  $G$  is connected, we see that all  $\mathcal{F}_x$  are equal. It follows that for some  $a \in \mathbb{Z}_m$ ,  $\mathcal{F}_x = \sigma_a^{-1}(\mathcal{H}_x)$  for all  $x \in V$ . Therefore  $\mathcal{F}$  is equivalent to the same  $(t, r)$ -Frankl family as  $\mathcal{H}$ .  $\square$

Finally, we can determine the maximum measure families.

**Lemma 5.7.** *Fix  $n \geq t \geq 1$ ,  $m \geq 2$ , and either  $s < m/2$  or both  $(s, m) = (1, 2)$  and  $t > 1$ . If  $\mathcal{F}$  is a family on  $\mathbb{Z}_m^n$  which is  $t$ -intersecting up to  $s$  and has measure  $w(\mathbb{Z}_m^n, t, s)$  then  $\mathcal{F}$  is equivalent to a family on  $n$  points whose  $\mu_{s/m}$ -measure is  $w(n, t, s/m)$ .*

*Proof.* Let  $\mathcal{F}_0 = \mathcal{F}, \mathcal{F}_1, \dots, \mathcal{F}_n$  be the sequence of families constructed by applying Lemma 5.3, so that  $\mathcal{F}_\ell$  is a family on  $\mathbb{Z}_m^{n-\ell} \times \{0, 1\}^\ell$ . The lemma states that  $\mu_{s,m}^{n-\ell, \ell}(\mathcal{F}_\ell) \leq \mu_{s,m}^{n-\ell-1, \ell+1}(\mathcal{F}_{\ell+1})$ . Since  $\mu_{s/m}(\mathcal{F}_n) \leq w(n, t, s/m) = w(\mathbb{Z}_m^n, t, s)$ , we conclude that  $\mu_{s,m}^{n-\ell, \ell}(\mathcal{F}_\ell) = \mu_{s,m}^{n-\ell-1, \ell+1}(\mathcal{F}_{\ell+1})$  for all  $0 \leq \ell < n$ . Also, Theorem 2.2 shows that  $\mathcal{F}_n$  is equivalent to a  $(t, r)$ -Frankl family for some  $r \geq 0$ . Applying Lemma 5.6, we conclude that the same holds for  $\mathcal{F}_0$ .  $\square$

This completes the proof of Theorem 5.1.

## 5.2 Continuous setting

The continuous analog of the setting of Section 5.1 is given by the following definitions.

**Definition 5.4.** A *continuous family on  $n$  points* is a measurable subset of  $\mathbb{T}^n$  (recall that  $\mathbb{T}$  is the unit circumference circle). For  $p \leq 1/2$ , a continuous family  $\mathcal{F}$  on  $n$  points is  *$t$ -agreeing up to  $p$*  if any two vectors  $x, y \in \mathcal{F}$  have  $t$  coordinates  $i_1, \dots, i_t$  such that the distance between  $x_{i_j}$  and  $y_{i_j}$  is less than  $p$  for all  $1 \leq j \leq t$ .

We denote the measure of a continuous family  $\mathcal{F}$  by  $\mu(\mathcal{F})$ .

The maximum measure of a continuous family on  $n$  points which is  $t$ -agreeing up to  $p$  is denoted  $w(\mathbb{T}^n, t, p)$ .

We will prove the following theorem.

**Theorem 5.8.** *Let  $n, m, t \geq 1$  and  $p < 1/2$ . Then  $w(\mathbb{T}^n, t, p) = w(n, t, p)$ .*

The proof uses a reduction to Theorem 5.1 which is very similar to the one used to deduce Lemma 3.5 from Lemma 3.3.

*Proof.* We start with the easy direction:  $w(\mathbb{T}^n, t, p) \geq w(n, t, p)$ . Let  $\mathcal{F}$  be a  $t$ -intersecting family on  $n$  points with  $\mu_p(\mathcal{F}) = w(n, t, p)$ . Define a mapping  $\tau(\mathbb{T}) \rightarrow \{0, 1\}$  by  $\tau([0, p)) = 1$  and  $\tau([p, 1)) = 0$ . It is not hard to check that  $\tau^{-1}(\mathcal{F})$  is a continuous family on  $n$  points which is  $t$ -agreeing up to  $p$  and has measure  $w(n, t, p)$ . Thus  $w(\mathbb{T}^n, t, p) \geq w(n, t, p)$ .

For the other direction, let  $\mathcal{F}$  be a continuous family on  $n$  points which is  $t$ -agreeing up to  $p$ . Let  $\epsilon > 0$  be a parameter satisfying  $p + \epsilon^{1/n} < 1/2$ . Since  $\mathcal{F}$  is measurable, there is a sequence  $F_i$  of cylindrical sets of total measure at most  $\mu(\mathcal{F}) + \epsilon$  which covers  $\mathcal{F}$ . Since an  $L_\infty$  ball of radius  $\delta$  around any point has volume  $(2\delta)^n$ , it follows that any point in  $\bigcup_i F_i$  is at  $L_\infty$ -distance  $\epsilon^{1/n}/2$  from  $\mathcal{F}$ . This implies that  $\bigcup_i F_i$  is  $t$ -agreeing up to  $p + \epsilon^{1/n}$ .

Choose  $I$  so that  $\sum_{i>I} \mu(F_i) < \epsilon$ , and let  $\mathcal{F}^* = \bigcup_{i \leq I} F_i$ . Let  $M$  be a large integer, and partition  $\mathbb{T}^n$  into  $M^n$  cubes of side length  $1/M$ . Let  $\mathcal{F}_M^*$  consist of the union of all cubes contained entirely inside  $\mathcal{F}^*$ . Thus  $\mu(\mathcal{F}_M^*) \geq \mu(\mathcal{F}^*) - O(I/M) \geq \mu(\mathcal{F}) - \epsilon - O(I/M)$ . We can view  $\mathcal{F}_M^*$  as a family on  $\mathbb{Z}_M^n$ , and this family is  $t$ -agreeing up to  $\lfloor (p + \epsilon^{1/n})M \rfloor$ . Theorem 5.1 thus shows that  $\mu(\mathcal{F}_M^*) \leq w(n, t, \lfloor (p + \epsilon^{1/n})M \rfloor)/M$ . Therefore

$$\mu(\mathcal{F}) \leq w(n, t, \lfloor (p + \epsilon^{1/n})M \rfloor)/M + \epsilon + O\left(\frac{I}{M}\right).$$

Since  $w(n, t, q)$  is continuous for  $q \leq 1/2$ , taking the limit  $M \rightarrow \infty$  we deduce that  $\mu(\mathcal{F}) \leq w(n, t, p + \epsilon^{1/n}) + \epsilon$ . Taking the limit  $\epsilon \rightarrow 0$ , we conclude that  $\mu(\mathcal{F}) \leq w(n, t, p)$ , and so  $w(\mathbb{T}^n, t, p) \leq w(n, t, p)$ .  $\square$

It is tempting to try and prove Theorem 5.8 directly, along the lines of Theorem 5.1. Besides requiring a stronger version of Lemma 3.5, one would also need a version of Proposition 5.2 for infinite polytopes, which we doubt holds.

## 6 Arguments to infinity

Suppose that we know an intersection theorem such as the Ahlswede–Khachatrian theorem in the uniform setting, that is for subsets of  $\binom{[n]}{k}$ . Can we deduce a matching theorem in related settings? This question has been considered by Ahlswede and Khachatrian [4] in their work on sets in Hamming space having small diameter (see the introduction to Section 5), and by Dinur and Safra [10] in their seminal work on the inapproximability of vertex cover (see also Tokushige [27] for a similar argument). Ahlswede and Khachatrian were interested in  $t$ -agreeing families on  $\mathbb{Z}_m^n$ , while Dinur and Safra were interested in  $t$ -intersecting families on  $n$  points, measured according to the  $\mu_p$  measure. In this section, we describe their arguments in an abstract framework, and explain their power and limitations.

We describe and motivate the abstract framework in Section 6.1. The argument of Ahlswede and Khachatrian appears in Section 6.2, and the one of Dinur and Safra in Section 6.3.

**Notation** We use the falling power notation  $n^{\underline{k}} = n!/(n-k)!$ . For a family  $\mathcal{F}$ , we denote by  $\mathcal{F}^{\equiv i}$  its subfamily consisting of all subsets of size  $i$ .

## 6.1 Abstract framework

The abstract framework described below originates in the author's PhD thesis [14].

**Definition 6.1.** Let  $\mathcal{X}$  be a monotone family on infinitely many points. A family  $\mathcal{F}$  on  $n$  points is  $\mathcal{X}$ -*intersecting* if all  $A, B \in \mathcal{F}$  satisfy  $A \cap B \in \mathcal{X}$ .

For example, a family is  $t$ -intersecting iff it is  $\mathcal{X}_t$ -intersecting, where  $\mathcal{X}_t$  is the collection of all subsets of  $\mathbb{N}$  containing at least  $t$  points.

For  $0 \leq k \neq n$ , we define  $w(n, \mathcal{X}, k)$  as the maximum of  $|\mathcal{F}| / \binom{n}{k}$  over all  $\mathcal{X}$ -intersecting  $\mathcal{F} \subseteq \binom{[n]}{k}$ .

For all  $p \in (0, 1)$ , we define  $w(n, \mathcal{X}, p)$  as the maximum of  $\mu_p(\mathcal{F})$  over all  $\mathcal{X}$ -intersecting families  $\mathcal{F}$  on  $n$  points.

Our goal is to relate  $w(n, \mathcal{X}, k)$  and  $w(n, \mathcal{X}, p)$ . We can expect such a relation in the following situation.

**Definition 6.2.** Let  $\mathcal{X}$  be a monotone family on infinitely many points, and let  $p \in (0, 1)$ . We write

$$w(n, \mathcal{X}, k) \xrightarrow{k/n \rightarrow p} w$$

if whenever  $(k_i, n_i)$  is a sequence such that  $n_i \rightarrow \infty$  and  $k_i/n_i \rightarrow p$  then  $w(n_i, \mathcal{X}, k_i) \rightarrow w$ .

As an example, consider the Ahlswede–Khachatrian theorem [3, 5].

**Definition 6.3.** Let  $t \geq 1$ ,  $r \geq 0$ , and  $n \geq k \geq t + 2r$ . The  $(t, r)$ -Frankl family  $\mathcal{F}_{t,r}^{n,k}$  is

$$\mathcal{F}_{t,r}^{n,k} = \left\{ S \in \binom{[n]}{k} : |S \cap [t + 2r]| \geq t + r \right\}.$$

A family  $\mathcal{F} \subseteq \binom{[n]}{k}$  is *equivalent* to a  $(t, r)$ -Frankl family if it equals the family in the definition of  $\mathcal{F}_{t,r}^{n,k}$  with  $[t + 2r]$  replaced by any other subset of  $[n]$  of size  $t + 2r$ .

**Proposition 6.1** (Ahlswede and Khachatrian [3, 5]). *Let  $1 \leq t \leq k \leq n$ , and suppose that  $\mathcal{F} \subseteq \binom{[n]}{k}$  is  $t$ -intersecting.*

*If*

$$\frac{r}{t + 2r - 1} < \frac{k - t + 1}{n} < \frac{r + 1}{t + 2r + 1}$$

*then  $|\mathcal{F}| \leq |\mathcal{F}_{t,r}^{n,k}|$ , with equality only if  $\mathcal{F}$  is equivalent to a  $(t, r)$ -Frankl family.*

*If*

$$\frac{k - t + 1}{n} = \frac{r + 1}{t + 2r + 1}$$

*then  $|\mathcal{F}| = |\mathcal{F}_{t,r}^{n,k}| = |\mathcal{F}_{t,r+1}^{n,k}|$ , with equality only if  $\mathcal{F}$  is equivalent to a  $(t, r)$ -Frankl family or to a  $(t, r + 1)$ -Frankl family.*

The theorem shows that the maximum measure of a  $t$ -intersecting family roughly depends only on  $k/n$ .

Indeed, as we show below, it implies that for  $\frac{r}{t+2r-1} < p < \frac{r+1}{t+2r+1}$  we have  $w(n, \mathcal{X}_t, k) \xrightarrow{k/n \rightarrow p} \mu_p(\mathcal{F}_{t,r})$ .

The following definition captures the situation depicted in the theorem more abstractly.

**Definition 6.4.** Let  $\mathcal{X}$  be a monotone family on infinitely many points, and let  $\mathcal{H}$  be an  $\mathcal{X}$ -intersecting family on  $m$  points. For  $p \in (0, 1)$ , we say that  $\mathcal{H}$  is  $p$ -*optimal* (with respect to  $\mathcal{X}$ ) if whenever  $(k_i, n_i)$  is a sequence such that  $n_i \rightarrow \infty$  and  $k_i/n_i \rightarrow p$ , for large enough  $i$  it holds that

$$w(n_i, \mathcal{X}, k_i) = \frac{1}{\binom{[n_i]}{k_i}} |\{A \in \binom{[n_i]}{k_i} : A \cap [m] \in \mathcal{H}\}|.$$

We say that  $\mathcal{H}$  is *uniquely  $p$ -optimal* if in the former situation, for large enough  $i$  the bound  $w(n_i, \mathcal{X}, k_i)$  is attained only on families of the given form, and their equivalents (obtained by applying any permutation on the  $n_i$  points).

**Lemma 6.2.** *Let  $\mathcal{X}$  be a monotone family on infinitely many points, let  $\mathcal{H}$  be an  $\mathcal{X}$ -intersecting family on  $m$  points, and suppose that  $\mathcal{H}$  is  $p$ -optimal for some  $p \in (0, 1)$ . Then*

$$w(n, \mathcal{X}, k) \xrightarrow{k/n \rightarrow p} \mu_p(\mathcal{H}).$$

Moreover,  $w(n, \mathcal{X}, p) \geq \mu_p(\mathcal{H})$  for all  $n \geq m$ .

*Proof.* It is clear that  $w(n, \mathcal{X}, p) \geq \mu_p(\mathcal{H})$ , since we can extend  $\mathcal{H}$  to a family on  $n$  points having the same measure.

Let now  $(k_i, n_i)$  be a sequence satisfying  $n_i \rightarrow \infty$  and  $k_i/n_i \rightarrow p$ . Define  $\mathcal{F}_i = \{A \in \binom{[n_i]}{k_i} : A \cap [m] \in \mathcal{H}\}$ . Without loss of generality, we can assume that  $w(n_i, \mathcal{X}, k_i) = |\mathcal{F}_i|/\binom{n_i}{k_i}$  for all  $i$  (rather than just for large enough  $i$ ). Notice that

$$\mathcal{F}_i = \bigcup_{S \in \mathcal{H}} \{S\} \times \binom{[n_i - m]}{k_i - |S|},$$

and so

$$\frac{|\mathcal{F}_i|}{\binom{n_i}{k_i}} = \sum_{S \in \mathcal{H}} \frac{\binom{n_i - m}{k_i - |S|}}{\binom{n_i}{k_i}} = \sum_{S \in \mathcal{H}} \frac{k_i^{|S|} (n_i - k_i)^{m - |S|}}{n_i^m}.$$

The right-hand side tends to

$$\sum_{S \in \mathcal{H}} p^{|S|} (1 - p)^{m - |S|} = \mu_p(\mathcal{H}). \quad \square$$

When  $p \in (0, 1/2)$  is not of the form  $\frac{r}{t+2r-1}$ , Lemma 6.2 shows that  $w(n, \mathcal{X}_t, k) \xrightarrow{k/n \rightarrow p} w(n, \mathcal{X}_t, p)$ , using both Proposition 6.1 and its weighted analog Theorem 2.2.

Lemma 6.2 gives a lower bound on  $w(n, \mathcal{X}, p)$ . The following subsections describe two matching upper bounds.

## 6.2 Argument of Ahlswede and Khachatrian

The first argument we describe originates from Ahlswede and Khachatrian [4].

**Lemma 6.3.** *Let  $\mathcal{X}$  be a monotone family on infinitely many points, let  $\mathcal{H}$  be an  $\mathcal{X}$ -intersecting family on  $m$  points, and suppose that  $\mathcal{H}$  is  $p$ -optimal for some  $p \in (0, 1)$ . Then  $w(m, \mathcal{X}, p) = \mu_p(\mathcal{H})$ .*

Moreover, for each  $m$  there is a finite set  $P_m$  such that the following holds. If  $\mathcal{H}$  is uniquely  $p$ -optimal for some  $p \notin P_m$  and  $\mathcal{F}$  is an  $\mathcal{X}$ -intersecting family on  $m$  points satisfying  $\mu_p(\mathcal{F}) = \mu_p(\mathcal{H})$  then  $\mathcal{F}$  is equivalent to  $\mathcal{H}$  (obtained from  $\mathcal{H}$  by applying a permutation).

*Proof.* In view of Lemma 6.2, it suffices that show that if  $\mathcal{F}$  is an  $\mathcal{X}$ -intersecting family on  $m$  points then  $\mu_p(\mathcal{F}) \leq \mu_p(\mathcal{H})$ .

Define a sequence  $(k_i, n_i)$  by taking  $k_i = m + i$  and  $n_i = \lfloor k_i/p \rfloor$ , so that  $n_i \rightarrow \infty$  and  $k_i/n_i \rightarrow p$  (the exact choice is arbitrary). Extend  $\mathcal{F}$  to a family  $\mathcal{F}_i$  just as in the proof of Lemma 6.2:

$$\mathcal{F}_i = \bigcup_{S \in \mathcal{F}} \{S\} \times \binom{[n_i - m]}{k_i - |S|}.$$

As in the proof of Lemma 6.2,  $|\mathcal{F}_i|/\binom{n_i}{k_i} \rightarrow \mu_p(\mathcal{F})$ . Since  $w(n_i, \mathcal{X}, k_i) \rightarrow \mu_p(\mathcal{H})$  by Lemma 6.2, we deduce that  $\mu_p(\mathcal{F}) \leq \mu_p(\mathcal{H})$ .

Suppose now that  $\mathcal{H}$  is uniquely  $p$ -optimal and  $\mu_p(\mathcal{F}) = \mu_p(\mathcal{H})$ . Then

$$\sum_{i=0}^m p^i (1 - p)^{m-i} (|\mathcal{F}^i| - |\mathcal{H}^i|) = 0.$$

If it is not the case that  $|\mathcal{F}^i| = |\mathcal{H}^i|$  for all  $i$  then  $p$  is a root of one of finitely many polynomials (whose exact number depends only on  $m$ ). Assuming that  $p$  is not one of these finitely many exceptions, we deduce that  $|\mathcal{F}^i| = |\mathcal{H}^i|$  for all  $i$ . This implies that  $|\mathcal{F}_i| = |\mathcal{H}_i|$  for all  $i$ , where  $\mathcal{H}_i$  is defined

analogously to  $\mathcal{F}_i$ . Since  $\mathcal{H}$  is uniquely  $p$ -optimal, it follows that  $\mathcal{F}_i$  is equivalent to  $\mathcal{H}$  for large enough  $i$ .

Suppose that  $\mathcal{F}_i$  is equivalent to  $\mathcal{H}$ . We can assume that  $\mathcal{H}$  depends on all  $m$  points (otherwise, replace  $\mathcal{H}$  by a family on fewer points). Let  $\mathcal{F}_i$  be obtained from the family in Definition 6.4 by applying the permutation  $\pi$  on  $[n_i]$ . We want to show that  $\pi([m]) = [m]$ , and so  $\mathcal{F}$  is equivalent to  $\mathcal{H}$ .

Indeed, consider any  $j \in [m]$ . Since  $\mathcal{H}$  depends on  $j$ , there exists an inclusion-minimal set  $S \in \mathcal{H}$  containing  $j$ . Thus  $\pi(S \cup \{m+1, \dots, m+k_i - |S|\}) \in \mathcal{F}_i$ . If  $\pi(j) > m$  then, since  $\mathcal{F}_i$  is invariant under permutations of the last  $n_i - m$  coordinates, we can find (assuming  $i$  is large enough) some coordinate  $\ell > m$  such that  $\pi(S \setminus \{j\} \cup \{m+1, \dots, m+k_i - |S|, \ell\}) \in \mathcal{F}_i$ , implying that  $S \setminus \{j\} \in \mathcal{H}$ , which contradicts the fact that  $S$  is inclusion-minimal.  $\square$

Lemma 6.3 allows us to deduce the value of  $w_{\text{sup}}(t, p)$  for all  $p \in (0, 1/2)$  not of the form  $\frac{r}{t+2r-1}$ . For  $p = \frac{r}{t+2r-1}$ , we can deduce  $w_{\text{sup}}(t, p)$  by continuity. The lemma also describes all the families of maximum measure on a large enough number of points, but only for most values of  $p \in (0, 1/2)$ ; for the countably many values  $\bigcup_{m \geq 1} P_m \cup \{\frac{r}{t+2r-1} : r \geq 0\}$ , the lemma fails to describe the families having maximum measure.

Compared to Theorem 2.2, the lifted version of Proposition 6.1 using Lemma 6.3 suffers from three shortcomings. First, it describes  $w_{\text{sup}}(t, p)$  but not  $w(n, t, p)$ . Second, it works only for  $p \leq 1/2$ . Third, it does not describe the maximum measure families for all values of  $p$ . It seems that the only way to handle these shortcomings is to prove Theorem 2.2 directly in the weighted setting.

### 6.3 Argument of Dinur and Safra

The second argument we describe originates from Dinur and Safra [10], and also appears in Tokushige [27].

**Lemma 6.4.** *Let  $\mathcal{X}$  be a monotone family on infinitely many points, let  $\mathcal{H}$  be an  $\mathcal{X}$ -intersecting family on  $m$  points, and suppose that  $\mathcal{H}$  is  $p$ -optimal for some  $p \in (0, 1)$ . Then  $w(m, \mathcal{X}, p) = \mu_p(\mathcal{H})$ .*

*Proof.* In view of Lemma 6.2, it suffices that show that if  $\mathcal{F}$  is an  $\mathcal{X}$ -intersecting family on  $m$  points then  $\mu_p(\mathcal{F}) \leq \mu_p(\mathcal{H})$ .

For a large integer  $N$ , extend  $\mathcal{F}$  to an  $\mathcal{X}$ -intersecting family  $\mathcal{F}_N$  on  $N$  points having the same  $\mu_p$ -measure. Observe that

$$\mu_p(\mathcal{F}) = \mu_p(\mathcal{F}_N) = \sum_{k=0}^N p^k (1-p)^{N-k} |\mathcal{F}_N^{\neq k}|.$$

Fix  $\epsilon > 0$ . Since  $\mathcal{H}$  is  $p$ -optimal, for large enough  $N$  we can bound  $|\mathcal{F}_N^{\neq k}| / \binom{N}{k} \leq \mu_p(\mathcal{H}) + \epsilon$  for all  $k$  satisfying  $|k/N - p| \leq 1/N^{1/3}$  (the specific error term  $1/N^{1/3}$  is arbitrary; any  $o(1/N^{1/2})$  error term would work). Hoeffding's bound implies that  $\Pr[|\text{Bin}(N, p)/N - p| > 1/N^{1/3}] < \epsilon$  for large enough  $N$ , and so

$$\mu_p(\mathcal{F}) \leq \sum_{k: |k/N - p| \leq 1/N^{1/3}} p^k (1-p)^{N-k} \binom{N}{k} (\mu_p(\mathcal{H}) + \epsilon) + \epsilon \leq \mu_p(\mathcal{H}) + 2\epsilon.$$

Taking the limit  $\epsilon \rightarrow 0$ , we deduce  $\mu_p(\mathcal{F}) \leq \mu_p(\mathcal{H})$ .  $\square$

One can also derive the second part of Lemma 6.3 using this technique, with a similar but more complicated argument which we leave to the reader.

## References

- [1] Rudolf Ahlswede and Valdimir Blinovsky. *Lectures on Advances in Combinatorics*. Springer, 2008.
- [2] Rudolf Ahlswede and Levon H. Khachatryan. The complete nontrivial-intersection theorem for systems of finite sets. *J. Comb. Theory A*, 76(1):121–138, 1996.
- [3] Rudolf Ahlswede and Levon H. Khachatryan. The complete intersection theorem for systems of finite sets. *Eur. J. Comb.*, 18(2):125–136, 1997.

- [4] Rudolf Ahlswede and Levon H. Khachatrian. The diametric theorem in Hamming spaces—optimal anticodes. *Adv. Appl. Math.*, 20:429–449, 1998.
- [5] Rudolf Ahlswede and Levon H. Khachatrian. A pushing-pulling method: New proofs of intersection theorems. *Combinatorica*, 19(1):1–15, 1999.
- [6] Rudolf Ahlswede and Levon H. Khachatrian. Katona’s intersection theorem: Four proofs. *Combinatorica*, 25(1):105–110, 2005.
- [7] Noga Alon, Irit Dinur, Ehud Friedgut, and Benny Sudakov. Graph products, Fourier analysis and spectral techniques. *Geometric and functional analysis*, 14(5):913–940, 2004.
- [8] Fan R. K. Chung, Ronald L. Graham, Peter Frankl, and James B. Shearer. Some intersection theorems for ordered sets and graphs. *J. Comb. Theory, Ser. A*, 43(1):23–37, 1986.
- [9] Irit Dinur and Ehud Friedgut. Proof of an intersection theorem via graph homomorphisms. *Electronic Journal of Combinatorics*, 13:Note #N6, 2006.
- [10] Irit Dinur and Shmuel Safra. On the hardness of approximating minimum vertex cover. *Ann. Math.*, 162(1):439–485, 2005.
- [11] David C. Ellis, Yuval Filmus, and Ehud Friedgut. Triangle-intersecting families of graphs. *Journal of the European Mathematical Society*, 14(3):841–885, 2012.
- [12] David C. Ellis, Ehud Friedgut, and Haran Pilpel. Intersecting families of permutations. *Journal of the American Mathematical Society*, 24:649–682, 2011.
- [13] Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford (2)*, 12:313–320, 1961.
- [14] Yuval Filmus. *Spectral Methods in Extremal Combinatorics*. PhD thesis, University of Toronto, 2013.
- [15] Péter Frankl. The Erdős–Ko–Rado theorem is true for  $n = ckt$ . In *Combinatorics, Proceedings of the Fifth Hungarian Colloquium on Combinatorics, Keszthely*, pages 365–375, Amsterdam, 1978. North-Holland.
- [16] Péter Frankl and Richard M. Wilson. The Erdős–Ko–Rado theorem for vector spaces. *Journal of Combinatorial Theory, Series A*, 43:228–236, 1986.
- [17] Ehud Friedgut. A Katona-type proof of an Erdős–Ko–Rado-type theorem. *Journal of Combinatorial Theory, Series A*, 111(2):239–244, 2005.
- [18] Ehud Friedgut. On the measure of intersecting families, uniqueness and stability. *Combinatorica*, 28(5):503–528, 2008.
- [19] Chris Godsil and Karen Meagher. *Erdős–Ko–Rado Theorems: Algebraic Approaches*. Cambridge University Press, 2015.
- [20] Ralph Howard, Gyula Károlyi, and László A. Székely. Towards a Katona type proof for the 2-intersecting Erdős–Ko–Rado theorem. *Electronic Journal of Combinatorics*, 8(1):Paper #R31, 2001.
- [21] Gyula O. H. Katona. Intersection theorems for systems of finite sets. *Acta Math. Acad. Sci. Hungar.*, 15:329–337, 1964.
- [22] Gyula O. H. Katona. A simple proof of the Erdős–Chao Ko–Rado theorem. *J. Comb. Theory, Ser. B*, 13:183–184, 1972.
- [23] Natan Keller and Noam Lifshitz. A tight stability version of the complete intersection theorem. *CoRR*, abs/1604.06135, 2016.
- [24] George Lann Nemhauser and Leslie Earl Trotter, Jr. Properties of vertex packing and independence system polyhedra. *Mathematical Programming*, 6(1):48–61, 1974.

- [25] Alexander Schrijver. *Combinatorial Optimization: Polyhedra and Efficiency*. Springer, 2003.
- [26] Igor Shinkar. Intersecting families, independent sets and coloring of certain graph products. Master's thesis, Weizmann Institute of Science, 2009.
- [27] Norihide Tokushige. Intersecting families — uniform versus weighted. *Ryukyu Mathematical Journal*, 18:89–103, 2005.
- [28] Richard M. Wilson. The exact bound in the Erdős-Ko-Rado theorem. *Combinatorica*, 4:247–257, 1984.