Electronic Colloquium on Computational Complexity, Report No. 44 (1998)

Bounds on Pairs of Families with Restricted Intersections

Jiří Sgall*

July 31, 1998

Abstract

We study pairs of families $\mathcal{A}, \mathcal{B} \subseteq 2^{\{1,\ldots,r\}}$ such that $|A \cap B| \in L$ for any $A \in \mathcal{A}$, $B \in \mathcal{B}$. We are interested in the maximal product $|\mathcal{A}| \cdot |\mathcal{B}|$, given r and L. We give asymptotically optimal bounds for L containing only elements of $s < q$ residue classes modulo q, where q is arbitrary (even non-prime) and s is a constant. As a consequence, we obtain a version of Frankl-Rödl result about forbidden intersections for the case of two forbidden intersections. We also give tight bounds for $L = \{0, \ldots, k\}$.

1 Introduction

Throughout the paper we work with a universe $R = \{1, \ldots, r\}$ of size r. The family of all the subsets of R is denoted by 2^R. We study pairs of families such that the intersections of the sets in them have only some restricted values.

Definition 1.1 *Let $L \subseteq \{0, \ldots, r\}$. Let $\mathcal{A}, \mathcal{B} \subseteq 2^R$ be two families of subsets of R. We say that $(\mathcal{A}, \mathcal{B})$ is an* L-intersecting pair of families *if $|A \cap B| \in L$ for any $A \in \mathcal{A}$ and $B \in \mathcal{B}$. The* size *of a pair $(\mathcal{A}, \mathcal{B})$ is the product $|\mathcal{A}| \cdot |\mathcal{B}|$.*

For any given r and L, we are interested to find the maximal size $|\mathcal{A}| \cdot |\mathcal{B}|$ over all L-intersecting pairs of families $(\mathcal{A}, \mathcal{B})$. Known results, both ours (in bold) and previous, are summarized by Table 1. First we give some auxiliary definitions.

*E-mail: `sgall@math.cas.cz`, `http://www.math.cas.cz/~sgall/`. Mathematical Institute, AS CR, Žitná 25, CZ-11567 Praha 1, Czech Republic and Dept. of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Prague. Partially supported by grant A1019602 of GA AV ČR, postdoctoral grant 201/97/P038 of GA ČR, and cooperative research grant INT-9600919/ME-103 from the NSF (USA) and the MŠMT (Czech republic).

ISSN 1433-8092

Definition 1.2 *Let* $\mathbf{N}$ *denote the non-negative integers. For* $L \subseteq \mathbf{N}$*, let*

$$L \bmod q = \{i \bmod q : i \in L\}$$
$$L - 1 = \{i - 1 : i \in L - \{0\}\}$$

Define

$$\binom{r}{\leq s} = \binom{r}{0} + \binom{r}{1} + \cdots + \binom{r}{s}.$$

Let $H(x)$ *be the entropy function* $H(x) = -x \log_2 x - (1-x)\log_2(1-x)$.

	L	Upper bound	Lower bound
(1)	$\lvert L \bmod q \rvert = s < q$	$2^s \cdot \binom{r}{\leq s-1} 2^r$	$\binom{r}{\leq s-1} 2^r$
(2)	$\{k, k+1\}$, independent of k	$(2r-1)2^r$	$(r+1)2^r$
(3)	$\{k, k+1\}$	$(2r-1)2^r$	$\Omega(r2^r/\sqrt{k})$
(4)	$\{0, \ldots, s\}$	$\binom{r}{\leq s-1} 2^r$	$\binom{r}{\leq s-1} 2^r$
(5)	$\{0, 1\}$	$(r+1)2^r$	$(r+1)2^r$
(6)	$\{k\}$, independent of k	2^r	2^r
(7)	$\{k, l\}, \lvert k - l \rvert > 1$, indep. of k and l	2^r	2^r
(8)	$\{k\}, k \geq 1$	2^{r-1}	$\Omega(2^r/\sqrt{k})$
(9)	$\{0, \ldots, r\} - \{\lceil r/4 \rceil\}$	$1.99^{2r} \approx 3.96^r$	$1.937^{2r} \approx 3.755^r$
(10)	$\{0, \ldots, r\} - \{\lceil r/5 \rceil, \lceil 2r/5 \rceil\}$	$4 \cdot 1.980^{2r} \approx 4 \cdot 3.921^r$	$1.842^{2r} \approx 3.393^r$

Table 1: *The bounds on the maximal size of* L*-intersecting families on* r*-element universe.*

(1): The upper bound, Corollary 3.4, is a consequence of our main result, which is slightly more general. The case of restricted number of intersections modulo p of course implies also the case when the cardinality of L is bounded. It is important to note that our bounds are valid for any q, not necessarily prime or prime power. The lower bound, Example 4.1, simply takes one family to be all sets and the other family to be all small sets.

(2), (3): The upper bound, Corollary 3.5, is motivated by Conjecture 6.2 related to communication complexity (see Section 6). This is also the reason why we are interested in an upper bound independent of k, i.e., the sizes of all the intersections are restricted to be two consecutive integers, be it $\{1, 2\}$ or $\{n/2, n/2+1\}$. A lower bound dependent on k is given in Example 4.3.

(4), (5): Using correlation inequalities, a tight upper bound can be proved for L containing all small integers, see Theorem 5.2. The lower bound is Example 4.1, mentioned above.

(6), (7), (8): The upper bounds are from [5]. In fact, they prove a stronger statement, namely that the bounds hold also for $L = \{l : l \equiv k \pmod p\}$ for some k and a prime p. This follows easily from linear algebra, since $\mathcal{A}$ and $\mathcal{B}$ are contained in orthogonal affine spaces, if taken as sets of $0, 1$ vectors over the field $GF(p)$. It follows that the case of two consecutive integers is the only case with $|L| \le 2$ where the size of the pair can be more than 2^r.

(9), (10): As their main result, Frankl and Rödl [5] prove that for $L = \{1, \ldots, k-1, k+1, \ldots, r\}$, where $\lambda r \le k \le (1/2 - \lambda)r$, any L-intersecting pair has size at most $(4 - \varepsilon)^n$, where ε depends only on λ. We are able to prove a theorem in the same spirit when we forbid special two intersections instead of one. The lower bounds are obtained by taking both families equal, containing all sets with cardinality larger than $5r/8$, or $7r/10$.

Our main results should be compared to the similar theorems for single families, namely variants of Ray-Chaudhuri–Wilson theorem, see [12, 6], or a survey by Babai and Frankl [3]. One difference is that our bounds has an extra factor of 2^r: for single families, the bounds are polynomial, $\binom{r}{s}$ or $\binom{r}{\le s}$ for uniform or non-uniform families (in the modular version of one-family theorems there is an extra requirement of the set size being different residue class than the intersections). This shows a different nature of the two problems. Perhaps a more important difference is that the modular theorems work only for prime moduli and to some extent for prime powers. In fact it is known that there is a significant difference between prime and composite moduli: some examples for special values are given in [3], and Grolmusz [7] proves that the bounds for any composite number have to be significantly super-polynomial. In contrast, our results are valid for *arbitrary* moduli.

2 Definitions and notation

Recall that our families are always families of subsets of $R = \{1, \ldots, r\}$, i.e., $\mathcal{A}, \mathcal{B} \subseteq 2^R$.

Definition 2.1 *Given a family $\mathcal{A}$, we define the* signature of a set B *to be the set $L_B^{\mathcal{A}}$ of all intersection sizes of B with elements of $\mathcal{A}$, i.e.,*

$$L_B^{\mathcal{A}} = \{|A \cap B| : A \in \mathcal{A}\}.$$

With this definition, a pair is $\{k, k+1\}$-intersecting if the signature of all elements of B is contained in $\{k, k+1\}$. For the inductive proof it is essential to define a relaxation of the intersecting condition which is no longer symmetric in $\mathcal{A}$ and $\mathcal{B}$. This approach is motivated by the work of Ahlswede et al [1], who use it for signatures of cardinality 1, with an additional restriction that the single element in all signatures must be at least δr for some fixed $\delta > 0$.

In our proof we proceed by an induction on the "height" of signatures. For the application to $\{k, k+1\}$ intersecting pairs, height is simply the cardinality of the set (or signature), moreover only height 1 and 2 is relevant. With this replacement, it is possible to skip directly to the last definition in this section. However, to formulate our results in a more general form, we work with a more general notion than the cardinality of the set. The following definition extracts the properties we need.

Definition 2.2 *We say that a function* $\|\cdot\| : 2^{\mathbf{N}} \to \mathbf{N} \cup \{\infty\}$ *is a* height-function *if the following holds:*

1. $\|L\| = 0$ *if and only if* $L = \emptyset$,
2. *if* $\|L\| = s < \infty$ *and* $L' \subseteq L$ *then* $\|L'\| \leq s$,
3. *if* $\|L\| = s < \infty$ *and* $L' \subseteq L - 1$ *then* $\|L'\| \leq s$,
4. *if* $\|L'\|, \|L\| \leq s < \infty$ *then either* $\|L' \cap L\| \leq s - 1$ *or* $\|L' \cap (L-1)\| \leq s - 1$.

It is easy to note that the height of a set has to be at least the cardinality of the largest contiguous interval contained in the set:

Lemma 2.3 *For any height function and for any* $a, b \in \mathbf{N}$, $a \leq b$, *and* $L \supseteq \{a, \ldots, b\}$ *we have* $\|L\| \geq b - a + 1$.

Proof. By induction on $b - a$. For $b = a$ the claim follows from the first property of the height-function, since L is non-empty. For $b > a$, let $L' = L$. Both $L' \cap L$ and $L' \cap (L-1)$ contain $\{a, \ldots, b-1\}$, hence their height is at least $b - a$ by the induction assumption. Using the contrapositive of the last property it follows that $\|L\| \geq b - a + 1$. □

The next lemma gives the most important examples of height-functions.

Lemma 2.4

1. *The cardinality of a set, i.e.,* $\|L\| = |L|$ *is a height-function.*
2. *For any* $q > 1$, *the function defined by*

$$\|L\| = \begin{cases} \infty & \text{if } |L \bmod q| = q, \\ |L \bmod q| & \text{otherwise,} \end{cases}$$

   is a height-function.

Note that excluding the sets that contain all numbers modulo q is necessary, to satisfy the last condition of the definition. The previous lemma does not cover all possible height-functions. It is possible to find a height function such that $\|\{0, 1, 3, 5\}\| = 2$, even though the basic height functions always give height at least 3 (the construction is left as an exercise).

Definition 2.5 *A pair of families* $(\mathcal{A}, \mathcal{B})$ has height s *if there exists a height-function* $\|\cdot\|$ *such that for all* $B \in \mathcal{B}$ *we have* $\|L_B^{\mathcal{A}}\| \leq s$.

Any $\{k, k+1\}$ intersecting pair has height 2.

3 The upper bound

Lemma 3.1 *Let* $x, x', y, y', m, M \geq 0$ *satisfy*

$$\begin{aligned} xy &\leq m, \\ x'y &\leq M, \\ xy' &\leq M, \\ x'y' &\leq M. \end{aligned}$$

Then

$$(x + x')(y + y') \leq 2(M + m)$$

Proof. If $x = 0$ or $y = 0$ then the statement is trivial. Otherwise let $X = x'/x$, $Y = y'/y$, and $Z = M/(xy)$. Now the assumptions give $0 \leq X \leq Z$, $0 \leq Y \leq Z$, and $0 \leq XY \leq Z$. It follows that $X + Y \leq 1 + Z$ (since for a fixed product XY, the sum $X + Y$ increases with the distance of X and Y). Therefore

$$\begin{aligned} (x + x')(y + y') &= (1 + X)(1 + Y)xy = (1 + (X + Y) + XY)xy \\ &\leq (1 + (1 + Z) + Z)xy \leq 2m + 2M. \end{aligned}$$

□

Definition 3.2 *Let* $f(r, s)$ *be the maximal size* $|\mathcal{A}| \cdot |\mathcal{B}|$ *of a pair of height* s *on a universe of* r *elements.*

Theorem 3.3 *If a pair* $(\mathcal{A}, \mathcal{B})$ *on a universe with* r *elements has height* s *then*

$$|\mathcal{A}| \cdot |\mathcal{B}| \leq 2^{r+s} \binom{r}{\leq s-1} \leq 2^{r+s+H(s/r)\cdot r}.$$

For $s = 2$ *and any* $r \geq 2$ *we have*

$$|\mathcal{A}| \cdot |\mathcal{B}| \leq (2r - 1)2^r.$$

Proof. We derive a recurrent bound for $f(r, s)$. Let $(\mathcal{A}, \mathcal{B})$ be a pair of families on the universe $\{1, \ldots, r+1\}$ with height s. For a set $X \subseteq \{1, \ldots, r\}$ and a family $\mathcal{X} \subseteq 2^{\{1,\ldots,r+1\}}$ we define

$$\begin{aligned} X^+ &= X \cup \{r + 1\} \\ \mathcal{X}_0 &= \{X \in 2^{\{1,\ldots,r\}} : X \in \mathcal{X}\}, \\ \mathcal{X}_1 &= \{X \in 2^{\{1,\ldots,r\}} : X^+ \in \mathcal{X}\}, \end{aligned}$$

i.e., $\mathcal{X}_1$ consists of the elements in $\mathcal{X}$ containing $r+1$, with this element removed, and $\mathcal{X}_0$ is the rest of $\mathcal{X}$. Obviously

$$|\mathcal{X}| = |\mathcal{X}_0| + |\mathcal{X}_1| = |\mathcal{X}_0 \cup \mathcal{X}_1| + |\mathcal{X}_0 \cap \mathcal{X}_1|.$$

Therefore

$$|\mathcal{A}| \cdot |\mathcal{B}| = (|\mathcal{A}_0| + |\mathcal{A}_1|)(|\mathcal{B}_0 \cup \mathcal{B}_1| + |\mathcal{B}_0 \cap \mathcal{B}_1|).$$

We will bound the right-hand side of this equality using Lemma 3.1. To do this, we need to bound the four products suitably.

For any $B \subseteq \{1, \ldots, r\}$,

$$\begin{aligned} L_B^{\mathcal{A}_0} &\subseteq L_B^{\mathcal{A}}, \text{ and} \\ L_B^{\mathcal{A}_0} &= L_{B+}^{\mathcal{A}_0} \subseteq L_{B+}^{\mathcal{A}}, \end{aligned} \tag{1}$$

since $\mathcal{A}_0 \subseteq \mathcal{A}$ and for any $A \in \mathcal{A}_0$, we have $A \cap B = A \cap B^+$. Thus the pair $(\mathcal{A}_0, \mathcal{B}_0 \cup \mathcal{B}_1)$ has height s. Similarly, for any $B \subseteq \{1, \ldots, r\}$,

$$\begin{aligned} L_B^{\mathcal{A}_1} &\subseteq L_B^{\mathcal{A}}, \text{ and} \\ L_B^{\mathcal{A}_1} &= L_{B+}^{\mathcal{A}_1} \subseteq L_{B+}^{\mathcal{A}} - 1, \end{aligned} \tag{2}$$

since for any $A \in \mathcal{A}_1$, we have $A^+ \in \mathcal{A}$ and $|A \cap B| = |A \cap B^+| = |A^+ \cap B^+| - 1$. Therefore the pair $(\mathcal{A}_1, \mathcal{B}_0 \cup \mathcal{B}_1)$ has height s as well. Thus we obtain, for $c = 0, 1$,

$$|\mathcal{A}_c| \cdot |\mathcal{B}_0 \cap \mathcal{B}_1| \leq |\mathcal{A}_c| \cdot |\mathcal{B}_0 \cup \mathcal{B}_1| \leq f(r, s) \tag{3}$$

Now we cover $\mathcal{B}_0 \cap \mathcal{B}_1$ by two families. Let

$$\begin{aligned} \mathcal{C} &= \{B \in \mathcal{B}_0 \cap \mathcal{B}_1 : \|L_B^{\mathcal{A}} \cap L_{B+}^{\mathcal{A}}\| \leq s - 1\}, \\ \mathcal{D} &= \{B \in \mathcal{B}_0 \cap \mathcal{B}_1 : \|L_B^{\mathcal{A}} \cap (L_{B+}^{\mathcal{A}} - 1)\| \leq s - 1\}. \end{aligned}$$

For any $B \in \mathcal{B}_0 \cap \mathcal{B}_1$, we have $B, B^+ \in \mathcal{B}$, and thus $\|L_B^{\mathcal{A}}\|, \|L_{B+}^{\mathcal{A}}\| \leq s$. Hence by the last condition in the definition of a height-function either $B \in \mathcal{C}$ or $B \in \mathcal{D}$, and $\mathcal{B}_0 \cap \mathcal{B}_1 = \mathcal{C} \cup \mathcal{D}$. ($\mathcal{C}$ and $\mathcal{D}$ are not necessarily disjoint.) For any $B \in \mathcal{C}$, using (1) we obtain $L_B^{\mathcal{A}_0} \subseteq L_B^{\mathcal{A}} \cap L_{B+}^{\mathcal{A}}$, therefore $\|L_B^{\mathcal{A}_0}\| \leq s - 1$ and the pair $(\mathcal{A}_0, \mathcal{C})$ has height $s - 1$. Similarly, for any $B \in \mathcal{D}$, using (2), we obtain $L_B^{\mathcal{A}_1} \subseteq L_B^{\mathcal{A}} \cap (L_{B+}^{\mathcal{A}} - 1)$, hence $\|L_B^{\mathcal{A}_1}\| \leq s - 1$ and the pair $(\mathcal{A}_1, \mathcal{D})$ has height $s - 1$.

Thus we have (see Fig. 1)

$$\begin{aligned} |\mathcal{A}_0| \cdot |\mathcal{C}| &\leq f(r, s - 1), \\ |\mathcal{A}_1| \cdot |\mathcal{D}| &\leq f(r, s - 1). \end{aligned} \tag{4}$$

	$\mathcal{B}_0 \cup \mathcal{B}_1$	$\mathcal{C}$	$\mathcal{D}$
$\mathcal{A}_0$	$f(r,s)$	$f(r,s-1)$	$f(r,s)$
$\mathcal{A}_1$	$f(r,s)$	$f(r,s)$	$f(r,s-1)$

Figure 1: Summarizing inequalities (3) and (4).

Let $\mathcal{A}_c$ be the smaller of the sets $\mathcal{A}_0$ and $\mathcal{A}_1$, i.e., choose $c = 0, 1$ such that $|\mathcal{A}_c| \leq |\mathcal{A}_0|, |\mathcal{A}_1|$. We have

$$\begin{aligned} |\mathcal{A}_c| \cdot |\mathcal{B}_0 \cap \mathcal{B}_1| &= |\mathcal{A}_c| \cdot |\mathcal{C}| + |\mathcal{A}_c| \cdot |\mathcal{D}| \\ &\leq |\mathcal{A}_0| \cdot |\mathcal{C}| + |\mathcal{A}_1| \cdot |\mathcal{D}| \\ &\leq 2f(r,s-1). \end{aligned} \tag{5}$$

	$\mathcal{B}_0 \cup \mathcal{B}_1$	$\mathcal{B}_0 \cap \mathcal{B}_1$
$\mathcal{A}_c$	$f(r,s)$	$2f(r,s-1)$
$\mathcal{A}_{1-c}$	$f(r,s)$	$f(r,s)$

Figure 2: Summarizing inequalities (3) and (5).

Using Lemma 3.1 for $x = |A_c|$, $x' = |A_{1-c}|$, $y = |\mathcal{B}_0 \cap \mathcal{B}_1|$, $y' = |\mathcal{B}_0 \cup \mathcal{B}_1|$, $m = 2f(r, s-1)$, and $M = f(r,s)$, we obtain

$$f(r+1,s) \leq 2f(r,s) + 4f(r,s-1).$$

It is easy to obtain boundary conditions for f: $f(r,0) = 0$ for any $r \geq 0$, and $f(0,s) = 1$ for any $s \geq 1$.

For $s = 1$ we obtain $f(r,1) = 2^r$. For $s = 2$, $f(2,2) \leq 3 \cdot 4$, since otherwise both families contain all sets and the signature $\{0,1,2\}$ appears; hence $f(r,2) \leq (2r-1)2^r$ for $r \geq 2$.

For general s, let $g(r,s) = f(r,s)/2^{r+s}$. Then the recurrence has form

$$g(r+1,s) \leq g(r,s) + g(r,s-1),$$

hence $g(r,s) \leq \binom{r}{\leq s-1}$ for any s. (For a constant s, the bound can be slightly improved using the fact that the boundary condition for $g(0,s)$ is not tight, and a case analysis for small r similar to $s = 2$; however, the leading term of the bound will be the same.) $\square$

Corollary 3.4 *Suppose that for some $q > s \geq 1$ the pair $(\mathcal{A}, \mathcal{B})$ has only s different intersections modulo q. Then $|\mathcal{A}| \cdot |\mathcal{B}| \leq 2^{r+s}\binom{r}{\leq s-1} \leq 2^{r+s+H(s/r)\cdot r}$.*

Corollary 3.5 *Any $\{k, k+1\}$ intersecting pair on a universe $\{1, \ldots, r\}$, $r \geq 2$ has size $|\mathcal{A}| \cdot |\mathcal{B}| \leq (2r-1)2^r$.*

Our largest example of a $\{k, k+1\}$ intersecting pair has size $(r+1)2^r$ (one family contains everything, the other all singletons and the empty set). We conjecture that this is optimal even among all pairs of height 2, but our upper bound is away by a factor of 2. This would follow from the following stronger conjecture.

Conjecture 3.6 $f(r+1, s) = 2f(r, s) + 2f(r, s-1)$.

Corollary 3.7 *Conjecture 3.6 implies that the size of any pair of height s is at most*

$$f(r, s) = \binom{r}{\leq s-1} 2^r.$$

Proof. Straightforward from the recurrence in the conjecture. □

This bound would be tight as shown by the pair with one family containing all sets and the other family containing all sets of cardinality smaller than s. Interestingly, this bound would imply the famous result of Frankl and Rödl [5] on forbidden intersections (in fact, the numerical bound would be even significantly better than 1.99^{2r} from [5]):

Corollary 3.8 *Conjecture 3.6 implies that if any pair $(\mathcal{A}, \mathcal{B})$ satisfies that for any $A \in \mathcal{A}$ and $B \in \mathcal{B}$, $|A \cap B| \neq \lceil r/4 \rceil$, then $|\mathcal{A}| \cdot |\mathcal{B}| \leq 4 \cdot 3.876^r \approx 4 \cdot 1.969^{2r}$.*

Proof. Let $\mathcal{A}' = \{A \in \mathcal{A} : |A| < 5r/8\}$, $\mathcal{B}' = \{B \in \mathcal{B} : |B| < 5r/8\}$, i.e. families of not too big sets in $\mathcal{A}$ and $\mathcal{B}$. The pair $(\mathcal{A}', \mathcal{B}')$ has height $3r/8$, since the intersection size is never equal to $\lceil r/4 \rceil$ modulo $\lceil 3r/8 \rceil$ (the intersection cannot be $\lceil r/4 \rceil + \lceil 3r/8 \rceil \geq 5r/8$, since we removed all big sets). Hence the conjecture implies that

$$|\mathcal{A}'| \cdot |\mathcal{B}'| \leq \binom{r}{\leq \frac{3}{8}r} 2^r \leq 2^{(1+H(3/8))r},$$

where H is the entropy function. Now, since we can bound the number of removed sets by the entropy,

$$\begin{aligned} |\mathcal{A}| \cdot |\mathcal{B}| &\leq (|\mathcal{A}'| + 2^{H(3/8)r})(|\mathcal{B}'| + 2^{H(3/8)r}) \\ &\leq |\mathcal{A}'| \cdot |\mathcal{B}'| + 3 \cdot 2^{(1+H(3/8))r} \leq 4 \cdot 2^{(1+H(3/8))r} \approx 4 \cdot 2^{1.9544r}, \end{aligned}$$

which gives the bounds in the theorem. □

We can prove a weaker version of the result of Frankl and Rödl [5] with two forbidden intersections already from our main result. Using sizes modulo $\lfloor r/5 \rfloor$, Corollary 3.4, and similar entropy considerations as in the previous proof we obtain the following corollary.

Corollary 3.9 *If any pair* $(\mathcal{A}, \mathcal{B})$ *satisfies that for any* $A \in \mathcal{A}$ *and* $B \in \mathcal{B}$, $|A \cap B| \notin \{\lfloor r/5 \rfloor, 2\lfloor r/5 \rfloor\}$, *then* $|\mathcal{A}| \cdot |\mathcal{B}| \le 4 \cdot 3.921^r \approx 4 \cdot 1.980^{2r}$, *for some constant* $\varepsilon > 0$.

Replacing $1/5$ by approximately 0.209 improves the bound to $3.864^r \approx 1.967^{2r}$, which is our best bound for two restricted intersections.

4 Examples for lower bounds

The following example is the only extreme example for $L = \{0, \ldots, k\}$. This example for $L = \{0, 1\}$ is also the best one we know for $L = \{k, k+1\}$, k arbitrary.

Example 4.1 $L = \{0, \ldots, k\}$. *Take* $\mathcal{A} = \{B \subseteq R : |B| \le k\}$, $\mathcal{B} = 2^R$. *I.e.,* $\mathcal{A}$ *contains all the sets of size at most* k *and* $\mathcal{B}$ *contains all* 2^r *sets. We get* $|\mathcal{A}| \cdot |\mathcal{B}| = \sum_{i=0}^{k} \binom{r}{i} 2^r$

If small intersections are not allowed, the most natural example is to include k elements into all of the sets, both in $\mathcal{A}$ and in $\mathcal{B}$, and then continue as above. This gives an example of size 2^{r-k} for $L = \{k\}$ and of size $(r-k+1)2^{r-k}$ for $L = \{k, k+1\}$. The following examples improve upon this simple bound significantly.

Example 4.2 ([1]) $L = \{k\}$. *Take* $\mathcal{A} = \{\{1, \ldots, \min\{2k, r\}\}\}$, $\mathcal{B} = \{B \subseteq R : |B \cap \{1, \ldots, 2k\}| = k\}$. *I.e.,* $\mathcal{A}$ *contains a single set of size* $2k$ *(or* r *if* $k > r/2$*), and* $\mathcal{B}$ *contains all the sets which intersect it in exactly* k *elements. There are* $\binom{2k}{k} 2^{r-2k} = \Theta(2^r/\sqrt{k})$ *such sets for* $k \le r/2$. *If* $k > r/2$ *and* $\varepsilon = r - 2k$ *then the size is* $2^{H(\varepsilon)n + o(n)}$.

Example 4.3 $L = \{k, k+1\}$, $k \le r/2$. *Take* $\mathcal{A} = \{A \subseteq R : \{1, \ldots, 2k\} \subseteq A \wedge |A| \le 2k+1\}$, $\mathcal{B} = \{B \subseteq R : |B \cap \{1, \ldots, 2k\}| = k\}$. *I.e.,* $\mathcal{A}$ *contains a single set of size* $2k$ *and all its extensions by one element, and* $\mathcal{B}$ *contains all the sets which intersect the* $2k$ *element sets in exactly* k *elements. We get* $|\mathcal{A}| \cdot |\mathcal{B}| = \Theta(r2^r/\sqrt{k})$ *such sets. (For* $k > r/2$, *the bound is again asymptotically smaller.)*

Ahlswede, Cai, and Zhang [1] show that their example for $L = \{k\}$ is asymptotically optimal and conjecture that they are optimal even absolutely. This conjecture would imply that the maximal size decreases with k.

5 The upper bound for $L = \{0, \ldots, k\}$

The upper bound for $\{0, \ldots, k\}$-intersecting pair follows easily from Kleitman's lemma proved in [8], see also [4, Ch. 19] or [2, Ch. 6]. This lemma can be proved by induction on the size of the universe.

Lemma 5.1 *Let* $\mathcal{A}, \mathcal{B}$ *be two monotone decreasing families of subsets of* R *(i.e., if* $A' \subseteq A$ *and* $A \in \mathcal{A}$ *then* $A' \in \mathcal{A}$*, and similarly for* $\mathcal{B}$*). Then*

$$|\mathcal{A}| \cdot |\mathcal{B}| \le |\mathcal{A} \cap \mathcal{B}| \cdot 2^r.$$

Theorem 5.2 *Any $\{0,\ldots,k\}$-intersecting pair $(\mathcal{A},\mathcal{B})$ has size at most $\sum_{i=0}^{k}\binom{r}{i}2^r$.*

Proof. We can assume that the families are monotone decreasing, since if we add into $\mathcal{A}$ a subset of $A \in \mathcal{A}$ then for any $B \in \mathcal{B}$, $|A' \cap B| \leq |A \cap B| \leq k$. All elements of $\mathcal{A} \cap \mathcal{B}$ must have size at most k, from the intersection property. The bound now follows directly from Lemma 5.1. $\square$

The example of $\mathcal{A} = \{A \subseteq R : |A| \leq k\}$, $\mathcal{B} = 2^R$ is the unique extremal example (up to exchanging the families). First, both families have to contain all sets of size at most k, to achieve equality in Lemma 5.1. Second, inspecting the proof of the lemma, it follows that for any $i \in R$, either $A \in \mathcal{A}$ implies $A \cup \{i\} \in \mathcal{A}$, or $B \in \mathcal{B}$ implies $B \cup \{i\} \in \mathcal{B}$; only the example above satisfies this.

6 Motivation

Our motivation comes from communication complexity, for a general reference on communication complexity see [9]. For a $0,1$ matrix $\mathbf{M}$, let $CC(\mathbf{M})$ and $\mathrm{rank}(\mathbf{M})$ be its deterministic communication complexity and its rank over the field of reals, respectively. It is known that $\log \mathrm{rank}(\mathbf{M}) \leq CC(\mathbf{M}) \leq \mathrm{rank}(\mathbf{M})$. In [10, 11], it is conjectured that

Conjecture 6.1 $CC(\mathbf{M}) = (\log \mathit{rank}(\mathbf{M}))^{O(1)}$.

Even a weaker upper bound would be interesting, since no better upper bound than $\mathrm{rank}(\mathbf{M})$ is known. Nisan and Wigderson [11] constructed an example of a matrix such that $CC(\mathbf{M}) \geq (\log \mathrm{rank}(\mathbf{M}))^{1.63\ldots}$, which is the largest known gap between $\log \mathrm{rank}(\mathbf{M})$ and $CC(\mathbf{M})$. Their method cannot give examples exhibiting larger than quadratic gap between $CC(\mathbf{M})$ and $\log \mathrm{rank}(\mathbf{M})$, and it would be particularly interesting to break this barrier, esp. because the quadratic gap is very common in the relation among $\log \mathrm{rank}(\mathbf{M})$ and variants of the deterministic and nondeterministic communication complexity, cf. [9].

Nisan and Wigderson [11] also show that to prove Conjecture 6.1, it would be sufficient to prove that every $0,1$ matrix of rank r has a large monochromatic submatrix, where large means that its area is at least $1/2^{(\log r)^{O(1)}}$ fraction of the original area. (In fact, it would be sufficient to show that there exists a large submatrix with rank at most cr for some $c < 1$.) This is related to the following conjecture.

Conjecture 6.2 *For any $\{k, k+1\}$-intersecting pair $(\mathcal{A},\mathcal{B})$ of size $S = |\mathcal{A}| \cdot |\mathcal{B}|$ there exists a pair of subfamilies $\mathcal{A}' \subseteq \mathcal{A}$ and $\mathcal{B}' \subseteq \mathcal{B}$ of size $|\mathcal{A}'| \cdot |\mathcal{B}'| \geq S/2^{(\log r)^{O(1)}}$ such that $(\mathcal{A}',\mathcal{B}')$ is either a $\{k\}$-intersecting or a $\{k+1\}$-intersecting pair.*

Any $\{k\}$-intersecting pair has size at most 2^r, for an arbitrary k. Hence Conjecture 6.2 implies that any $\{k, k+1\}$-intersecting family has size at most $2^{(\log r)^{O(1)}} \cdot 2^r$. We have proved a much stronger bound in Corollary 3.5, which gives some support for Conjecture 6.2.

On the other hand, Conjecture 6.2 implies the matrix property mentioned above for a special kind of rank r matrices. Any families $(\mathcal{A}, \mathcal{B})$ can be represented as a matrix with rows indexed by $A \in \mathcal{A}$, columns indexed by $B \in \mathcal{B}$ and entry in row A and column B defined as $|A \cap B|$. This matrix is a sum of r $0,1$ matrices of rank 1 (each corresponding to one element); conversely each sum of r $0,1$ matrices of rank 1 can be represented by some families (A, B). If the families (A, B) are $\{k, k+1\}$-intersecting then the corresponding matrix has entries k and $k+1$ only, and if we subtract k from each entry, we obtain a $0,1$ matrix of rank $r+1$.

Conclusions and open problems

We have proved number of results concerning the size of pairs of families with restricted intersection. The main open problem is to close the gap of (about) 2^s in our bounds. It is possible that pairs of height s can be larger than the corresponding intersecting pairs; in particular, does there exist a pair of height 2 and size more than $(r+1)2^r$?

For $L = \{k\}$ and $L = \{k, k+1\}$, it would be interesting to prove or disprove that the bound decreases with k; if true, this would imply a tight upper bound matching our lower bounds.

Given our motivation from communication complexity, an interesting open problem is to decide if any $\{k, k+1\}$-intersecting pair must contain a big $\{k\}$- or $\{k+1\}$-intersecting pair of subfamilies, see Conjecture 6.2.

Last, it would be nice to have a better understanding of the reasons why in our modular theorem we can allow arbitrary modulus, while in similar theorems for single families the modulus is restricted to be a prime or a prime power.

Acknowledgements

I am grateful to Avi Wigderson for bringing this problem to my attention, to Vojta Rödl for many comments, and to Pavel Pudlák for encouraging me to continue to work on this problem – also by his partial results – and for pointing me to the paper [1].

References

[1] R. Ahlswede, N. Cai, and Z. Zhang. A general 4-words inequality with consequences for 2-way communication complexity. *Advances in Applied Mathematics*, 10:75–94, 1989.

[2] N. Alon and J. H. Spencer. *The Probabilistic Method.* Wiley, 1992.

[3] L. Babai and P. Frankl. Linear algebra methods in combinatorics, preliminary version 2. Manuscript, Dept. of Computer Science, The Univesity of Chicago, 1992.

[4] B. Bollobás. *Combinatorics — Set Systems, Hypergraphs, Families of Vectors, and Combinatorial Probability.* Cambridge University Press, 1986.

[5] P. Frankl and V. Rödl. Forbidden intersections. *Transactions of the AMS*, 300(1):259–286, 1987.

[6] P. Frankl and R. M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1:313–368, 1981.

[7] V. Grolmusz. On set systems with restricted intersections modulo a composite number. In *Proceedings of the 3rd Ann. International Computing and Combinatorics Conference, Lecture Notes in Computer Science 1276*, pages 82–90. Springer-Verlag, 1997.

[8] D. J. Kleitman. Families of non-disjoint subsets. *Journal Combinatorial Theory*, 1:153–155, 1966.

[9] E. Kushilevitz and N. Nisan. *Communication Complexity.* Cambridge University Press, 1996.

[10] L. Lovász and M. Saks. Lattices, Möbius functions, and communication complexity. In *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science*, pages 81–90. IEEE, 1988.

[11] N. Nisan and A. Wigderson. On rank vs. communication complexity. In *Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science*, pages 831–836. IEEE, 1994.

[12] D. K. Ray-Chaudhuri and R. M. Wilson. On t-designs. *Osaka J. Math.*, 12:737–744, 1975.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'