Improved Bounds on Families Under k-wise Set-Intersection Constraints

Weiting Cao*, Kyung-Won Hwang†, Douglas B. West‡

August, 2005; revised November, 2006

Abstract

Let p be a prime, and let L be a set of s congruence classes modulo p. Let $\mathcal{H}$ be a family of subsets of $[n]$ such that the size modulo p of each member of $\mathcal{H}$ is not in L, but the size modulo p of every intersection of k distinct members of $\mathcal{H}$ is in L. We prove that $|\mathcal{H}| \le (k-1)\sum_{i=0}^{s}\binom{n-1}{i}$, improving the bound due to Grolmusz and generalizing results proved for $k=2$ by Snevily.

1 Introduction

Finding the largest family $\mathcal{H}$ of subsets of an n-element set subject to various conditions on intersections of its members is a classical theme in extremal set theory.

For example, Frankl and Wilson [3] proved that $|\mathcal{H}| \le \sum_{i=0}^{s}\binom{n}{i}$ when the sizes of pairwise intersections are restricted to a set L of s nonnegative integers. Deza, Frankl, and Singhi [1] obtained the same result by a different method. Snevily [9] improved the conclusion to $|\mathcal{H}| \le \sum_{i=0}^{s}\binom{n-1}{i}$ when L is a set of s positive integers (in [6] he had proved that the bound holds when n is sufficiently large, and in [8] he had proved it when the values in L are consecutive).

*Mathematics Dept., Univ. of Illinois, 1409 W. Green St., IL 61801 U.S.A, *wcao1@math.uiuc.edu*.

†Department of Mathematics, Kookmin University, 861-1 Jeongneung-dong Seongbuk-gu, Seoul 136-702, Korea, *khwang7@kookmin.ac.kr*.

‡Mathematics Dept., Univ. of Illinois, 1409 W. Green St., IL 61801 U.S.A, *west@math.uiuc.edu*. Work supported in part by the NSA under Award No. MDA904-03-1-0037.

RUNNING HEAD: Bounds Under k-wise Constraints

In [6], Snevily also proved a modular version of the result. That is, if the pairwise intersection sizes lie in s congruence classes modulo p, where p is a prime, and the sizes of the members of $\mathcal{H}$ themselves avoid these s classes, then $|\mathcal{H}| \le \sum_{i=0}^{s} \binom{n-1}{i}$. Our main result generalizes Snevily's result to the setting of restrictions on the intersections of k distinct members of the family; his result is the special case $k = 2$.

Throughout the paper, p is a fixed prime number, L is a fixed set of s congruence classes modulo p, and $c_p(S)$ is the congruence class of $|S|$ modulo p. Also let $\mathbf{2}^n$ be the family of all subsets of $[n]$. For $k \in \mathbb{N}$, a *k-wise p-modular L-intersecting family* in $\mathbf{2}^n$ is a family $\mathcal{H}$ of subsets of $[n]$ such that $c_p(H) \notin L$ for all $H \in \mathcal{H}$, but $c_p(H_1 \cap \cdots \cap H_k) \in L$ whenever $H_1, \ldots, H_k$ are distinct members of $\mathcal{H}$.

We study the maximum size of such families. When $k = 1$, the maximum size is 0, trivially. Using matrices, Grolmusz [4] proved a general upper bound.

Theorem 1 (Grolmusz [4]) *If $\mathcal{H}$ is a k-wise p-modular L-intersecting family in $\mathbf{2}^n$ and $|L| = s$, then*

$$|\mathcal{H}| \le (k-1) \sum_{i=0}^{s} \binom{n}{i}.$$

Grolmusz and Sudakov [5] gave another proof of this bound using multilinear polynomials. Furthermore, under the additional restriction that $c_p(H) \in S$ for all $H \in \mathcal{H}$, where S is a set of size t contained in $\{s-t, \ldots, p-1\} - L$, they proved the tighter bound $|\mathcal{H}| \le (k-1) \sum_{i=s-t+1}^{s} \binom{n}{i}$. Szabo and Vu [10] proved that if $\sum_{i=0}^{s} \binom{n}{i} \ge \lfloor \log_p (k-1) \rfloor$, then $|\mathcal{H}| \le (k-1) \left(\sum_{i=0}^{s} \binom{n}{i} - \lfloor \log_p (k-1) \rfloor + 1 \right)$. Also, Füredi and Sudakov [2] proved that $|\mathcal{H}| \le \frac{k+s-1}{s+1} \binom{n}{s} + \sum_{i \le s-1} \binom{n}{i}$ when n is sufficiently large, and this is asymptotically sharp.

In this paper, we improve the main bound of Theorem 1 by replacing $\binom{n}{i}$ with $\binom{n-1}{i}$ in the sum. We also further improve the upper bound to $(k-1)\binom{n}{s}$ when $L = \{0, 1, \ldots, s-1\}$; this improves the bound by keeping only the top two terms of the sum. Snevily [9] conjectured for the case $k = 2$ that the improved bound of $\binom{n}{s}$ also holds for every L. Our proofs follow the techniques of Snevily and of Grolmusz and Sudakov, again relying on multilinear polynomials and dimension of vector spaces.

A non-modular version follows by taking $p > n$. If L is a set of s nonnegative integers, and $\mathcal{H} \subseteq \mathbf{2}^n$ satisfies $|H| \notin L$ for $H \in \mathcal{H}$ and $|H_1 \cap \cdots \cap H_k| \in L$ whenever $H_1, \ldots, H_k$ are distinct sets in $\mathcal{H}$, then the same bound $|\mathcal{H}| \le (k-1) \sum_{i=0}^{s} \binom{n-1}{i}$ holds, with improvement to $|\mathcal{H}| \le (k-1)\binom{n}{s}$ when $L = \{0, \ldots, s-1\}$. Snevily [7] proved the special case of the last improvement for $k = 2$; again, our proofs extend his methods.

When n is sufficiently large, the Füredi–Sudakov bound of $\frac{k+s-1}{s+1}\binom{n}{s}+o(n^s)$ is stronger, and it does not need the condition that $|H|\notin L$ for $H\in\mathcal{H}$. The Grolmusz–Sudakov bound of $(k-1)\sum_{i=0}^{s}\binom{n}{i}$ in the non-modular case also does not need the condition that $|H|\notin L$. Note, however, that when $k=2$ and $L=\{0,\ldots,s-1\}$, the family $\mathcal{H}$ consisting of all sets in $\mathbf{2}^n$ with size at most s has intersection sizes in L and has size $\sum_{i=0}^{s}\binom{n}{i}$. Thus the improvement to $\binom{n}{s}$ is not available without some restriction, such as $|H|\notin L$.

Our bound of $(k-1)\sum_{i=0}^{s}\binom{n-1}{i}$ is achieveable when $k=2$ if the condition $c_p(H)\notin L$ is dropped, by letting $L=\{1,\ldots,s\}$ and letting $\mathcal{H}$ consist of all sets of size at most $s+1$ containing the element 1. We do not know whether our upper bound is ever sharp when the family also satisfies the condition $c_p(H)\notin L$ that is needed for our proof of the upper bound.

Finally, we mention one more conjecture by Snevily [9] that would extend some of these results (for $k=2$) in a "cross-intersection" setting. He conjectured for any set L of positive integers that if $H_1,\ldots,H_m$ and $H'_1,\ldots,H'_m$ are two families in $\mathbf{2}^n$ such that $|H_i\cap H'_j|\in L$ for $i\neq j$ but $|H_i\cap H'_i|=0$ for all i, then $m\le\binom{n}{|L|}$.

2 The Proof

Applications of the polynomial method to extremal set theory problems often use the "triangular argument". The polynomials may have many variables.

Proposition 2 *Polynomials $f_1,\ldots,f_m$ are linearly independent on a domain S if there are points $y^{(1)},\ldots,y^{(m)}$ in S such that $f_i(y^{(j)})$ is nonzero when $i=j$ and zero when $i>j$.*

Proof. Assume that some linear combination $\sum_{i=1}^{m}c_if_i$ is identically zero on S. Evaluate successively at the points $y^{(1)},\ldots,y^{(m)}$. When evaluating at $y^{(j)}$, assume that the previous evaluations have shown that $c_1=\cdots=c_{j-1}=0$. By hypothesis, $c_if_i(y^{(j)})=0$ for $i>j$, but $f_j(y^{(j)})\neq 0$. Hence $\sum_{i=1}^{m}c_if_i(y^{(j)}=0$ requires $c_j=0$. Induction yields that all coefficients are 0. Hence there is no equation of dependence. □

In the technical lemma that we use to prove the main theorem, we invoke the triangular argument three times.

Lemma 3 *Let $A_1,\ldots,A_m$ and $B_1,\ldots,B_m$ be subsets of $[n]$, and let p be a prime and s be a positive integer. For $1\le i\le m$, let $L_i=\{c_p(A_{i+1}\cap B_i),\ldots,c_p(A_m\cap B_i)\}$. If*

1) Each L_i has size at most s and does not contain $c_p(A_i\cap B_i)$, and

2) There exist $r\in[m]$ and $z\in[n]$ with $z\notin A_i\cup B_i$ for $i\le r$ and $z\in A_i\cap B_i$ for $i>r$, then $m\le\sum_{i=0}^{s}\binom{n-1}{i}$.

Proof. We perform all computations in the field $\mathbb{F}_p$ of order p, and hence we write congruence as equality. Let $u^{(j)}$ and $v^{(i)}$ denote the incidence vectors of A_j and B_i, respectively. Note that $u^{(j)} \cdot v^{(i)} = |A_j \cap B_i|$. These vectors are n-dimensional; we also use x as a vector of n variables $(x_1, \ldots, x_n)$.

For $1 \le i \le m$, define a polynomial f_i by $f_i(x) = \prod_{l \in L_i} (x \cdot v^{(i)} - l)$. By the hypotheses,

$$f_i(u^{(i)}) = \prod_{l \in L_i} (|A_i \cap B_i| - l) \neq 0 \text{ for } 1 \le i \le m,$$
$$f_i(u^{(j)}) = \prod_{l \in L_i} (|A_j \cap B_i| - l) = 0 \text{ for } 1 \le i < j \le m.$$

Let $C_1, \ldots, C_t$ be the family of all subsets of $[n]$ with size at most $s-1$ that do not contain element z, indexed in nondecreasing order of size. Let $w^{(i)}$ be the incidence vector of C_i. Note that $t = \sum_{j=0}^{s-1} \binom{n-1}{j}$. For $1 \le i \le t$, define

$$h_i(x) = \prod_{j \in C_i} x_j \qquad \text{and} \qquad g_i(x) = (x_z - 1) h_i(x).$$

Note that $g_1(x) = (x_z - 1)$, since the product over an empty set is 1.

Each f_i and g_i is a product of at most s linear factors. On the restricted domain $\{0,1\}^n$, exponentiation of variables has no effect. Hence on this domain, each term in the expansion of such a polynomial can be expressed as a constant multiple of a product of at most s distinct variables. We conclude that the restrictions of these polynomials to $\{0,1\}^n$ lie in a space of dimension $\sum_{j=0}^{s} \binom{n}{j}$. Since $t = \sum_{j=1}^{s} \binom{n-1}{j-1}$ and $\binom{n}{j} - \binom{n-1}{j-1} = \binom{n-1}{j}$, to prove the desired bound on m it suffices to show that the set $\{f_1, \ldots, f_m\} \cup \{g_1, \ldots, g_t\}$ is linearly independent on $\{0,1\}^n$.

The triangular argument (with indices in reverse order) shows that $\{f_1, \ldots, f_m\}$ is independent, because $f_i(u^{(j)}) \neq 0$ if $j = i$ and $f_i(u^{(j)}) = 0$ if $j > i$.

To show that $\{g_1, \ldots, g_t\}$ is independent, it suffices to show independence for $\{h_1, \ldots, h_t\}$. Again we use the triangular argument. Since h_i is the product of the positions in the incidence vector corresponding to C_i, we have $h_i(w^{(i)}) = 1 \neq 0$. If $i > j$, then C_i contains an element not in C_j. Hence $h_i(x)$ has as a factor a variable that is 0 in $w^{(j)}$, and $h_i(w^{(j)}) = 0$.

Finally, to show independence of the full set, let

$$\sum_{i=1}^{m} \alpha_i f_i(x) + \sum_{j=1}^{t} \beta_j g_j(x) = 0$$

be an equation of dependence valid on $\{0,1\}^n$. Recall that $z \in A_i$ if and only if $i > r$. Let $A_i' = A_i \cup \{z\}$, with incidence vector $y^{(i)}$. On each $y^{(i)}$, the contribution of the second sum is always 0. Also $f_i(y^{(j)}) = f_i(u^{(j)})$ for $i \le j$, since $A'_j \cap B_i = A_j \cap B_i$ if $i \le j$ (we have $A'_j = A_j$ if $j > r$, and $z \notin B_i$ if $j \le r$). Hence $f_i(y^{(j)}) \neq 0$ if $j = i$ and $f_i(y^{(j)}) = 0$ if $j > i$. By the triangular argument, $\alpha_1 = \cdots = \alpha_m = 0$. Since $\{g_1, \ldots, g_t\}$ is linearly independent, we now also have $\beta_1 = \cdots = \beta_t = 0$. □

We need a slight variation for the improvement when $L = \{0, ..., s-1\}$. The proof is even simpler than for Lemma 3, so we merely stress the differences and give fewer details.

Lemma 4 *Let $A_1, \ldots, A_m$ and $B_1, \ldots, B_m$ be subsets of $[n]$, and let p be a prime and s be a positive integer. Let $L = \{0, \ldots, s-1\}$. If $c_p(A_j \cap B_i) \in L$ for $j > i$ and $c_p(A_i \cap B_i) \notin L$, then $m \le \binom{n}{s}$.*

Proof. Define $u^{(i)}$, $v^{(i)}$, and f_i as in the proof of Lemma 3, with $L_i = L$ for all i. Also define $C_1, \ldots, C_t$ and $w^{(1)}, \ldots, w^{(t)}$ and $h_1, \ldots, h_t$ as before, but without excluding any element of $[n]$ from presence in $C_1, \ldots, C_t$. Thus now $t = \sum_{j=0}^{s-1} \binom{n}{j}$, and proving independence of $\{h_1, \ldots, h_t\} \cup \{f_1, \ldots, f_m\}$ leaves $m \le \binom{n}{s}$.

Independence follows from a single application of the triangular argument. Evaluate in order at $w^{(1)}, \ldots, w^{(t)}$ and then $u^{(1)}, \ldots, u^{(m)}$ (also index $h_1, \ldots, h_t$ and then $f_m, \ldots, f_1$). As before, $h_i(w^{(j)})$ is 1 if $i = j$ and 0 if $i > j$. Since $|C_j| < s$, we have $w^{(j)} \cdot y \in L$ whenever $w^{(j)}$ is the incidence vector of C_j and $y \in \{0,1\}^n$; thus $f_i(w^{(j)}) = 0$ for all i and j. Finally, the hypotheses again yield $f_i(u^{(j)}) = 0$ for $j > i$ and $f_i(u^{(j)}) \neq 0$ for $j = i$. □

In the proof of the main theorem, the factor $k-1$ arises in the bound from a procedure that extracts up to $k-1$ members of $\mathcal{H}$ at a time. This idea originates with Grolmusz and Sudakov [5].

Theorem 5 *If $\mathcal{H}$ is a k-wise p-modular L-intersecting family in $\mathbf{2}^n$, then*

$$|\mathcal{H}| \le (k-1) \sum_{i=0}^{s} \binom{n-1}{i}.$$

Furthermore, if $L = \{0, \ldots, s-1\}$, then $|\mathcal{H}| \le (k-1)\binom{n}{s}$.

Proof. We repeatedly delete subsets of $\mathcal{H}$ until $\mathcal{H}$ is empty. The ith step deletes at most $k-1$ members of $\mathcal{H}$ and produces subsets A_i and B_i of $[n]$. Hence we obtain $A_1, \ldots, A_m$ and $B_1, \ldots, B_m$ with $m \ge |\mathcal{H}|/(k-1)$. The claim then follows by showing that these families satisfy the conditions of Lemmas 3 and 4 in the general and special cases.

If $\mathcal{H}$ remains nonempty after $i-1$ steps, then move some sets from $\mathcal{H}$ to a family $\mathcal{G}$ as follows. First choose a remaining set $H_1 \in \mathcal{H}$ such that $1 \notin H_1$, if any exists; otherwise, choose any remaining member of $\mathcal{H}$ as H_1. Delete H_1 from $\mathcal{H}$ and initialize $\mathcal{G} = \{H_1\}$.

Having moved $H_1, \ldots, H_{j-1}$ from $\mathcal{H}$ to $\mathcal{G}$, move some $H_j \in \mathcal{H}$ if $c_p(H_1 \cap \cdots \cap H_j) \notin L$. Continue until no remaining member of $\mathcal{H}$ has this property. Since $\mathcal{H}$ is a k-wise p-modular L-intersecting family, fewer than k sets move. With $\mathcal{G} = \{H_1, \ldots, H_d\}$ at the end, let $A_i = H_1$ and $B_i = \bigcap_{j=1}^{d} H_j$. This completes the ith step. Note that if $1 \in A_i$, then also $1 \in B_i$, since $1 \in H_1$ implies that all remaining members of $\mathcal{H}$ contain the element 1.

Due to the termination condition for the ith step, every remaining $H \in \mathcal{H}$ after the ith step has the property that $c_p(H \cap B_i) \in L$. Thus $c_p(A_j \cap B_i) \in L$ for $j > i$, and $L_i \subseteq L$. Also, $c_p(A_i \cap B_i) = c_p(B_i) \notin L$; when $d > 1$ this holds by construction, and when $d = 1$ it holds by the hypothesis that $c_p(H) \notin L$ for all $H \in \mathcal{H}$.

We have ensured that $c_p(A_i \cap B_i) \notin L$ and that $c_p(A_j \cap B_i) \in L$ for $j > i$. Also, the preference for members of $\mathcal{H}$ avoiding 1 ensures the existence of r such that $1 \notin A_i \cup B_i$ for $i \leq r$ and $1 \in A_i \cap B_i$ for $i > r$; after the rth step, all remaining members of $\mathcal{H}$ contain element 1. Thus the conditions of Lemmas 3 and 4 hold in the general and special cases, respectively, as desired. □

Acknowledgement

The authors thank Zoltán Füredi for bringing the problem to our attention and the referees for their insightful comments.

References

[1] M. Deza, P. Frankl, and N. M. Singhi, *On functions of strength t*, Combinatorica **3** (1983), 331–339.

[2] Z. Füredi and B. Sudakov, *Extremal set-systems with restricted k-wise intersections*, J. Comb. Theory Ser. A **105** (2004), 143–159.

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

[4] V. Grolmusz, *Set systems with restricted multiple intersections*, Electron. J. Combin. **9** (2002) #R8, 10pp.

[5] V. Grolmusz and B. Sudakov, *On k-wise set-intersections and k-wise Hamming-distances*, J. Comb. Theory. Ser. A **99** (2002), 180–190.

[6] H. S. Snevily, *On generalizations of the deBruijn-Erdös Theorem*, J. Comb. Theory. Ser. A **68** (1994), 232–238.

[7] H. S. Snevily, *A generalization of the Ray-Chaudhuri–Wilson Theorem*, J. Comb. Designs **3** (1995), 349–352.

[8] H. S. Snevily, *A generalization of Fisher's Inequality*, J. Comb. Theory. Ser. A **85** (1999), 120–125.

[9] H. S. Snevily, *A sharp bound for the number of sets that pairwise intersect at k positive values*, Combinatorica **23** (2003), 527–532.

[10] T. Szabó and V. H. Vu, *Exact k-wise intersection theorems*, Graphs and Combinatorics **21** (2005), 247–261.