Pairwise Intersections and Forbidden Configurations

Richard P. Anstee, Peter Keevash

EuroComb 2005 DMTCS proc. **AE**, 2005, 17–20

Pairwise Intersections and Forbidden Configurations

R. P. Anstee[1 †] and Peter Keevash[2‡]

[1]*Mathematics Department, The University of British Columbia, Vancouver, B.C., Canada V6T 1Z2*
[2]*Department of Mathematics, California Institute of Technology Pasadena, California, USA*

Let $f_m(a,b,c,d)$ denote the maximum size of a family $\mathcal{F}$ of subsets of an m-element set for which there is no pair of subsets $A, B \in \mathcal{F}$ with

$$|A \cap B| \geq a, \quad |\bar{A} \cap B| \geq b, \quad |A \cap \bar{B}| \geq c, \quad \text{and} \quad |\bar{A} \cap \bar{B}| \geq d.$$

By symmetry we can assume $a \geq d$ and $b \geq c$. We show that $f_m(a,b,c,d)$ is $\Theta(m^{a+b-1})$ if either $b > c$ or $a, b \geq 1$. We also show that $f_m(0,b,b,0)$ is $\Theta(m^b)$ and $f_m(a,0,0,d)$ is $\Theta(m^a)$. This can be viewed as a result concerning forbidden configurations and is further evidence for a conjecture of Anstee and Sali. Our key tool is a strong stability version of the Complete Intersection Theorem of Ahlswede and Khachatrian, which is of independent interest.

Keywords: forbidden configurations, extremal set theory, intersecting set systems, uniform set systems, (0,1)-matrices

---

Let $f_m(a,b,c,d)$ denote the maximum size of a family $\mathcal{F}$ of subsets of an m-element set for which there is no pair of subsets $A, B \in \mathcal{F}$ with

$$|A \cap B| \geq a, \quad |\bar{A} \cap B| \geq b, \quad |A \cap \bar{B}| \geq c, \quad \text{and} \quad |\bar{A} \cap \bar{B}| \geq d.$$

By symmetry we can assume $a \geq d$ and $b \geq c$.

Theorem 1 *Suppose $a \geq d$ and $b \geq c$. Then $f_m(a,b,c,d)$ is $\Theta(m^{a+b-1})$ if either $b > c$ or $a, b \geq 1$. Also $f_m(a,0,0,d)$ is $\Theta(m^a)$ and $f_m(0,b,b,0)$ is $\Theta(m^b)$.*

Some motivation for studying this function comes from the forbidden configuration problem for matrices popularised by the first author. We can identify a family $\mathcal{A} = \{A_1, \cdots, A_n\}$ of subsets of $[m]$ with an $m \times n$ $(0,1)$-matrix A determined by incidence, i.e. A_{ij} is 1 if $i \in A_j$, otherwise 0. Such a matrix is *simple*, by which we mean it has no repeated columns. Let F be a $(0,1)$-matrix (not necessarily simple). We define forb(m, F) to be the largest n for which there is a simple $m \times n$ $(0,1)$-matrix A that does not contain an F configuration, i.e. a submatrix which is a row and column permutation of F. If we interpret

---

†Research is supported in part by NSERC
‡Research was done partly while visiting the first author with the support of the first author's NSERC grant

A, F as incidence matrices of systems $\mathcal{A}, \mathcal{F}$ (the latter possibly having sets with multiplicity) then A has an F configuration exactly when $\mathcal{A}$ has $\mathcal{F}$ as a *trace*, i.e. $\mathcal{F} \subset \{A \cap X : A \in \mathcal{A}\}$ for some $X \subset [m]$.

The first forbidden configuration result was obtained independently by Sauer [6], Perles, Shelah [7], Vapnik and Chervonenkis [8]. When F is the $k \times 2^k$ $(0,1)$-matrix with all possible distinct columns they showed that $\mathrm{forb}(m, F) = \sum_{i=0}^{k-1} \binom{m}{i}$. For a general k-row matrix F, Füredi obtained an $O(m^k)$ upper bound on $\mathrm{forb}(m, F)$, but it seems hard to determine the order of magnitude of $\mathrm{forb}(m, F)$ for each F. This was achieved when F has 2 rows by Anstee, Griggs and Sali [2] and for 3 rows by Anstee and Sali [3], but is open in general.

It is not hard to see that if F consists of a single column with s 0's and t 1's then $\mathrm{forb}(m, F)$ is $\Theta(m^{\max\{s-1,t-1\}})$. In this paper we solve the problem when F has two columns. Let F_{abcd} be the $(a+b+c+d) \times 2$ $(0,1)$-matrix which has a rows of [11], b rows of [10], c rows of [01], d rows of [00]. Then $\mathrm{forb}(m, F_{abcd}) = f_m(a, b, c, d)$ as defined above.

In [3] a conjecture was made for the asymptotic behaviour of $\mathrm{forb}(m, F)$ as a function of m and F. In particular, a restricted set of constructions of simple matrices were described in [3] that were conjectured to predict the asymptotics of $\mathrm{forb}(m, F)$. These were used in this paper to predict the asymptotics in Theorem 1 as well as to provide construction. This is further evidence for the conjecture in [3].

Our key tool is a strong stability version of the Complete Intersection Theorem of Ahlswede and Khachatrian [1], which is of independent interest. Strong stability results have been employed with success by the second author, for example in [4],[5]. First we recall some notation. Let numbers k, r_1, r_2 be given and suppose G and H are disjoint sets with $|G| = k - r_1 + r_2$. We define $\mathcal{I}^k_{r_1,r_2}$ on the pair (H, G) to be the family consisting of all sets of size k in $G \cup H$ that intersect G in at least $k - r_1 = |G| - r_2$ points. Note that any two sets in $\mathcal{I}^k_{r_1,r_2}$ have at least $|G| - 2r_2 = k - r_1 - r_2$ points in common, i.e. $\mathcal{I}^k_{r_1,r_2}$ is $(k - r)$-intersecting, where $r = r_1 + r_2$.

We also define $\mathcal{F}^k_{r_1,r_2}$ on the pair (H, G) to be the family consisting of all sets of size k in $G \cup H$ that intersect G in exactly $k - r_1 = |G| - r_2$ points. Clearly this is a subsystem of $\mathcal{I}^k_{r_1,r_2}$ and $|\mathcal{I}^k_{r_1,r_2} \backslash \mathcal{F}^k_{r_1,r_2}|$ is of a lower order of magnitude than $|\mathcal{I}^k_{r_1,r_2}|$ and $|\mathcal{F}^k_{r_1,r_2}|$. In particular, if the systems are defined on the ground set $[m]$ with $k = \Theta(m)$ then $|\mathcal{I}^k_{r_1,r_2}|$ and $|\mathcal{F}^k_{r_1,r_2}|$ are $\Theta(m^r)$, whereas $|\mathcal{I}^k_{r_1,r_2} \backslash \mathcal{F}^k_{r_1,r_2}| < m^{r-2}$. The Complete Intersection Theorem, conjectured by Frankl, and proved by Ahlswede and Khachatrian [1], is that any k-uniform, $(k-r)$-intersecting family of maximum size on a given ground set is isomorphic to $\mathcal{I}^k_{r-p,p}$, for some $0 \le p \le r$, which depends on the size of the ground set. We prove the following result.

Theorem 2 *Suppose $\mathcal{A}$ is a k-uniform $(k - r)$-intersecting set system on $[m]$ of size at least $(5r)^{5r} m^{r-1}$. Then $\mathcal{A} \subset \mathcal{I}^k_{r-p,p}$ for some $0 \le p \le r$.*

We use this theorem in our proofs of the upper bounds in Theorem 1 in cases where $\mathcal{A}$ is a k-uniform $(k - r)$-intersecting set system satisfying some additional properties. If $|\mathcal{A}|$ is small, we can ignore it for the purposes of upper bounds. If $|\mathcal{A}|$ is large enough to matter for the upper bounds, we can use the fact that $\mathcal{A} \subset \mathcal{I}^k_{r-p,p}$ to deduce structure in $\mathcal{A}$ (e.g. the partition G, H above) which we can exploit in our proofs.

References

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

[2] R.P. Anstee, J.R. Griggs and A. Sali, Small Forbidden Configurations, *Graphs and Combinatorics*, **13** (1997), 97–118.

[3] R.P. Anstee and A. Sali, Small Forbidden Configurations IV: The 3-rowed case, *Combinatorica*, to appear.

[4] P. Keevash and D. Mubayi, Stability theorems for cancellative hypergraphs, *J. Combin. Theory Ser. B* **92** (2004), 163–175.

[5] P. Keevash and B. Sudakov, The Turán number of the Fano plane, *Combinatorica*, to appear.

[6] N. Sauer, On the density of families of sets, *J. Combin. Th. Ser A* **13**(1973), 145-147.

[7] S. Shelah, A combinatorial problem: Stability and order for models and theories in infinitary languages, *Pac. J. Math.* **4**(1972), 247-261.

[8] V.N. Vapnik, A.Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Th. Prob. and Applics.* **16**(1971), 264-280.