EuroComb 2005 DMTCS proc. **AE**, 2005, 117–122

Color critical hypergraphs and forbidden configurations

Richard Anstee[1†], Balin Fleming[1‡], Zoltán Füredi[23§] and Attila Sali[3¶]

[1] *Mathematics Department The University of British Columbia Vancouver, B.C., Canada V6T 1Z2*
[2] *Department of Mathematics University of Illinois at Urbana-Champaign 1409 W. Green Street Urbana, Illinois 61801-2975, USA*
[3] *Alfréd Rényi Instiute of Mathematics Hungarian Academy of Sciences Budapest, P.O.Box 127 H-1364 Hungary*

The present paper connects sharpenings of Sauer's bound on forbidden configurations with color critical hypergraphs. We define a matrix to be *simple* if it is a (0,1)-matrix with no repeated columns. Let F be a $k \times l$ (0,1)-matrix (the forbidden configuration). Assume A is an $m \times n$ simple matrix which has no submatrix which is a row and column permutation of F. We define $\text{forb}(m, F)$ as the best possible upper bound on n, for such a matrix A, which depends on m and F. It is known that $\text{forb}(m, F) = O(m^k)$ for any F, and Sauer's bond states that $\text{forb}(m, F) = O(m^{k-1})$ fore *simple* F. We give sufficient condition for non-simple F to have the same bound using linear algebra methods to prove a generalization of a result of Lovász on color critical hypergraphs.

Keywords: forbidden configuration, color critical hypergraph, linear algebra method

1 Introduction

A k-uniform hypergraph $(V, \mathcal{E})$ is 3-color critical if it is not 2-colorable, but for all $E \in \mathcal{E}$ the hypergraph $(V, \mathcal{E} \setminus \{E\})$ is 2-colorable. Lovász [12] proved in 1976, that

$$|\mathcal{E}| \le \binom{n}{k-1}$$

for a 3-color critical k-uniform hypergraph. Here we prove the following that can be considered as generalization of Lovász' result.

Theorem 1 *Let $\mathcal{E} \subseteq \binom{[m]}{k}$ be a k-uniform set system on an underlying set X of m elements. Let us fix an ordering $E_1, E_2, \ldots E_t$ of $\mathcal{E}$ and a prescribed partition $A_i \cup B_i = E_i$ ($A_i \cap B_i = \emptyset$) for each member of $\mathcal{E}$. Assume that for all $i = 1, 2, \ldots, t$ there exists a partition $C_i \cup D_i = X$ ($C_i \cap D_i = \emptyset$), such that*

† Research is supported in part by NSERC
‡ Research is supported in part by NSERC
§ Research is supported in part by Hungarian National Research Fund (OTKA) numbers T034702 and T037846
¶ Research is supported in part by Hungarian National Research Fund (OTKA) numbers T034702 and T037846 and by NSERC of the first author

$E_i \cap C_i = A_i$ and $E_i \cap D_i = B_i$, but $E_j \cap C_i \neq A_j$ and $E_j \cap C_i \neq B_j$ for all $j < i$. (That is, the ith partition cuts the ith set as it is prescribed, but does not cut any earlier set properly.) Then

$$t \leq \binom{m}{k-1} + \binom{m}{k-2} + \ldots + \binom{m}{0}. \tag{1}$$

Theorem 1 was motivated by the following sharpening of Sauer's bound for forbidden configurations. Let F be a $k \times l$ 0-1 matrix, then $\mathrm{forb}(m, F)$ denotes maximum n such that there exists an $m \times n$ *simple* matrix A such that no column and/or row permutation of F is a submatrix of A. Furthermore, let K_k denote the $k \times 2^k$ simple 0-1 matrix consisting of all possible columns.

Theorem 2 *Let F be contained in $F_B = [K_k \,|\, t \cdot (K_k - B)]$ for an $k \times (k+1)$ matrix B consisting of one column of each possible column sum. Then* $\mathrm{forb}(m, F) = O(m^{k-1})$.

We explain the the connection between Theorem 1 and Theorem 2.

The study of forbidden configurations is a problem in extremal set theory. The language we use here is matrix theory which conveniently encodes the problems. We define a *simple* matrix as a (0,1)-matrix with no repeated columns. Such a matrix can be thought of a set of subsets of $\{1, 2, \ldots, m\}$ with the columns encoding the subsets and the rows indexing the elements. Assume we are give a $k \times l$ (0,1)-matrix F. We say that a matrix A has no *configuration* F if no submatrix of A is a row and column permutation of F and so F is referred to as a *forbidden configuration* (sometimes called *trace*). A variety of combinatorial objects can be defined by forbidden configurations. For a simple $m \times n$ matrix A which is assumed to have no configuration F, we seek an upper bound on n which will depend on m, F. We denote the best possible upper bound as forb(m, F). Many results have been obtained about forb(m, F) including [2],[3],[5].

At this point all values known for forb(m, F) are of the form $\Theta(m^e)$ for some integer e. We completed the classification for $2 \times l$ matrices F in [2] and for $3 \times l$ matrices F in [6]. We also put forward a conjecture on what properties of F drive the exponent e. Roughly speaking, we proposed a set of constructions and guessed that these constructions are sufficient to deduce the exponent e in the expression $\Theta(m^e)$.

We use the notation K_k to denote the $k \times 2^k$ simple matrix of all possible columns on k rows. The basic result for forb(m, F) is as follows.

Theorem 3 *[Sauer [13], Perles and Shelah [14], Vapnik and Chervonenkis [15]] We have that forb(m, K_k) is $\Theta(m^{k-1})$.*

In fact Theorem 3 is usually stated with forb$(m, K_k) = \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0}$ but the asymptotic growth of $\Theta(m^{k-1})$ was what interested Vapnik and Chervonenkis.

One easy observation is that if we let A^c denote the 0-1-complement of A then forb$(m, F^c) =$ forb(m, F). Another observation is that if F' is a submatrix of F, then forb$(m, F) \geq$ forb(m, F'). We let K_k^s denote the $k \times \binom{k}{s}$ simple matrix of all possible columns of column sum s.

We use the notation $[A|B]$ to denote the matrix obtained from concatenating the two matrices A and B. We use the notation $k \cdot A$ to denote the matrix $[A|A|\cdots|A]$ consisting of k copies of A concatenated together. We give precedence to the operation $\cdot$ (multiplication) over concatenation so that for example $[2 \cdot A|B]$ is the matrix consisting of the concatenation of B with the concatenation of two copies of A.

According to an earlier unpublished result of Füredi [10] forb$(m, F) = O(m^k)$ for *arbitrary* $k \times l$ configuration F. The goal of this paper is to give sufficient conditions that ensure forb$(m, F) = O(m^{k-1})$.

2 The boundary between m^{k-1} and m^k

Theorem 3 implies that simple configurations all have $\text{forb}(m,F) = O(m^{k-1})$, thus we investigate f's with multiple columns. First, we show that which configurations F have $\text{forb}(m,F) = \Omega(m^k)$ using the direct product construction. Let $A(k,2)$ be defined as a minimal matrix with the property that any pair of rows has $\begin{bmatrix}1\\1\end{bmatrix}$ has both with 1's in some column and such that deleting a column of $A(k,2)$ would violate this property.

Lemma 4 *Let F be a $k \times l$ configuration. forb$(m,F) = \Omega(m^k)$ if F contains $2 \cdot K_k^l$ for $2 \le l \le k-2$ and $l = 0, k$ or if F contains $[2 \cdot K_k^1 | A(k,2)]$.*

Proof: We find that $\text{forb}(m,F)$ is $\Omega(m^k)$ if F contains $2 \cdot K_k^l$ for $0 \le l \le k$ and $l \ne 1, k-1$. This follows since $2 \cdot K_k^l$ is not contained in the k-fold product of l $K_{m/k}^1$'s and $k-l$ $K_{m/k}^{(m/k)-1}$'s and so may deduce $\text{forb}(m, 2 \cdot K_k^l)$ is $\Omega(m^k)$. To verify this for $2 \le l \le k-2$, we note that any pair of rows of K_k^l has $\begin{bmatrix}1 & 0\\1 & 0\end{bmatrix}$ and so if we have a submatrix that is a row and column permutation of K_k^l, we can only choose one row from either $K_{m/k}^1$ or from $K_{m/k}^{(m/k)-1}$. The verification for K_k^0 or K_k^k is easier.

For $l = 1$ (the case $l = k-1$ is the (0,1)-complement) we can no longer assert that any pair of rows of K_k^l has $\begin{bmatrix}1 & 0\\1 & 0\end{bmatrix}$ merely $\begin{bmatrix}0\\0\end{bmatrix}$ and so can choose two rows from the copy of $K_{m/k}^1$, one row from each of $k-2$ of the $K_{m/k}^{(m/k)-1}$ terms and generate a copy of $2 \cdot K_k^1$. (Theorem 5.1 of [6] shows that $\text{forb}(m, K_k^1)$ is $\Theta(m_{k-1})$). This is fixed by considering a minimal matrix $A(k,2)$ with the property that any pair of rows has $\begin{bmatrix}1\\1\end{bmatrix}$ has both with 1's in some column and such that deleting a column of $A(k,2)$ would violate this. As above, we have that if F contains $[2 \cdot K_k^1 | A(k,2)]$, then $\text{forb}(m,F)$ is $\Omega(m^k)$. □

Lemma 4 leaves two possibilities if we want $\text{forb}(m,f)$ be bounded away from m^k. Either F is contained in a matrix $F_B = [K_k \,|\, t \cdot (K_k - B)]$ for an $k \times (k+1)$ matrix B consisting of one column of each possible column sum or F is contained in a matrix $[K_k^0 |\ t \cdot C]$ where C is a k-rowed simple matrix consisting of all columns which do not have 1's in both rows 1 and 2 and also with at least one 1. Note, that these are not mutually exclusive cases. Our main result Theorem 2 is that in the first case $\text{forb}(m,F) = O(m^{k-1})$.

Proof of Theorem 2: Let A be an $m \times n$ simple 0-1 matrix, and B be a $k \times (k+1)$ matrix consisting of one column of each possible column sum. Suppose that A does not have $F_B = [K_k \,|\, t \cdot (K_k - B)]$ as configuration. This implies that on a given k-tuple L of rows either K_k is missing, or if all possible columns of size k occur on L, then $t \cdot (K_k - B)$ must be missing. This latter means, that for some $0 \le s \le k$, two columns of column sum s occur at most $t-1$ times on L, respectively. Let $\mathcal{K}$ be the set of k-tuples of rows where the latter happens. Using Lemma 5 a set of columns of size $O(m^{k-1})$ can be removed from A to obtain A', so that for all $L \in \mathcal{K}$ a column (in fact two) is missing on L in A'. However, this implies that K_k is not a configuration in A', thus by Theorem 3 A' has at most $O(m^{k-1})$ columns. □

Let $\mathcal{K}$ be a system of k-tuples of rows such that $\forall K \in \mathcal{K}$ there are two $(k \times 1)$ columns, $\alpha_K \ne \beta_K$ specified. We say that a column x of A *violates* (K, α_K), if $x|_K = \alpha_K$, similarly, x *violates* (K, β_K), if $x|_K = \beta_K$.

Lemma 5 *Assume, that for every $K \in \mathcal{K}$ there are at most $t-1$ columns of A that violate (K, α_K), and at most $t-1$ columns of A violate (K, β_K). Then there exists a subset X of columns of A, such that $|X| = O(m^{k-1})$ and no column of $A - X$ violates any of (K, α_K) or (K, β_K).*

Proof: It can be assumed without loss of generality that for all $K \in \mathcal{K}$ $\alpha_K = \alpha$ and $\beta_K = \beta$ independent of K. Indeed, there are $2^k \times 2^k$ possible α_K, β_K pairs, that is a constant number of them .Thus, $\mathcal{K}$ can be partitioned into a constant number of parts, so that in each part $\alpha_K = \alpha$ and $\beta_K = \beta$ holds. We apply induction on k using the simplification given above. $k = 1$ is obvious.

Consider now $k \times 1$ columns $\alpha \neq \beta$. Assume first, that $\alpha \neq \overline{\beta}$. That is, there is a coordinate where α and β agree, say both have 1 as their ℓth coordinate. The case of a common 0 coordinate is similar. For the ith row of A we count how many columns have violation so that for some $K \in \mathcal{K}$ the ℓth coordinate in K is exactly row i. Let $\mathcal{K}_{i,\ell}$ be the set of these k-tuples from $\mathcal{K}$. Columns that have violation on k-tuples from $\mathcal{K}_{i,\ell}$ have 1 in the ith row, let $A_{i,1}$ denote matrix formed by the set of columns that have 1 in row i. If row i is removed from $A_{i,1}$, the remaining matrix $A'_{i,1}$ is still simple. Let $\mathcal{K}'_{i,\ell}$ denote the set of $(k-1)$-tuples obtained from k-tuples of $\mathcal{K}_{i,\ell}$ by removing their ℓth coordinate, i, furthermore let α' (β', respectively) denote the $(k-1) \times 1$ column obtained from α (β) by removing the ℓth coordinate, 1. Note, that $\alpha' \neq \beta'$. A column of A has a violation on $K \in \mathcal{K}_{i,\ell}$ iff its counterpart in $A'_{i,1}$ has a violation on the corresponding $K' \in \mathcal{K}'_{i,\ell}$. The number of those columns is at most $c\,m^{k-2}$ by the inductive hypothesis. Since $\mathcal{K} = \cup_{i=1}^{m} \mathcal{K}_{i,\ell}$, we obtain that the number of columns of A having violation on some $K \in \mathcal{K}$ is at most $m \cdot c\,m^{k-2}$.

Let us assume now, that $\alpha = \overline{\beta}$. A subset $\mathcal{J} \subseteq \mathcal{K}$ is called *independent* if there exists an ordering $J_1, J_2, \ldots J_g$ of the elements of $\mathcal{J}$ such that for every $J_i \in \mathcal{J}$ there exists an $m \times 1$ 0-1 column that violates J_i and does not violate any $J_j \in \mathcal{J}$ for $j < i$. Let us call a *maximal* independent subset $\mathcal{B}$ of $\mathcal{K}$ a *basis* of $\mathcal{K}$. If a column of A has a violation on $K \in \mathcal{K}$, then it has a violation on some $B \in \mathcal{B}$, as well. Indeed, either $K \in \mathcal{B}$ holds, or if $K \notin \mathcal{B}$, then by the maximality of $\mathcal{B}$, K cannot be added to it as a $|\mathcal{B}| + 1$st element in the order, so the column having violation on K must have a violation on $B \in \mathcal{B}$, for some B. By Theorem 1 for a basis $\mathcal{B}$ we have

$$|\mathcal{B}| \leq \binom{m}{k-1} + \binom{m}{k-2} + \ldots + \binom{m}{0},$$

since a column violating a k-tuple B_i from $\mathcal{B}$, but none of B_j for $j < i$, gives an appropriate partition of the set of rows. Thus, there could be at most $(2t-2)\left[\binom{m}{k-1} + \binom{m}{k-2} + \ldots + \binom{m}{0}\right]$ columns violating some $K \in \mathcal{K}$. □

Proof of Theorem 1: We define a polynomial $p_i(\underline{x}) \in \mathbb{R}[x_1, x_2, \ldots, x_m]$ for each E_i as follows.

$$p_i(x_1, x_2, \ldots, x_m) = \prod_{a \in A_i} (1 - x_a) \prod_{b \in B_i} x_b + (-1)^{k+1} \prod_{a \in A_i} x_a \prod_{b \in B_i} (1 - x_b) \tag{2}$$

Polynomials defined by (2) are multilinear of degree at most $k-1$, since the product $\prod_{e \in E_i} x_e$ cancels by the coefficient $(-1)^{k+1}$. Thus, they are from the space generated by monomials of type $\prod_{j=1}^{r} x_{i_j}$, for $r = 0, 1, \ldots k-1$. The dimension of this space over $\mathbb{R}$ is $\binom{m}{k-1} + \binom{m}{k-2} + \ldots + \binom{m}{0}$.

We shall prove that polynomials $p_1(\underline{x}), p_2(\underline{x}), \ldots p_t(\underline{x})$ are linearly independent over $\mathbb{R}$, which implies (1). Assume that

$$\sum_{i=1}^{t} \lambda_i p_i(\underline{x}) = 0 \tag{3}$$

is a linear combination of the $p_i(\underline{x})$'s that is the zero polynomial. Consider the partition $C_t \cup D_t = X$, and substitute $x_c = 0$ if $c \in C_t$ and $x_d = 1$ if $d \in D_t$ into (3). Then $p_t(\underline{x}) = 1$, but it is easy to see that $p_k(\underline{x}) = 0$ for $k < t$. This implies that $\lambda_t = 0$. Now assume by induction on j, that $\lambda_t = \lambda_{t-1} = \ldots = \lambda_{t-j+1} = 0$. Take the partition $C_{t-j} \cup D_{t-j} = X$ and substitute into (3) $x_c = 0$ if $c \in C_{t-j}$ and $x_d = 1$ if $d \in D_{t-j}$. Then, as before, $p_{t-j}(\underline{x}) = 1$, but $p_k(\underline{x}) = 0$ for $k < t - j$. This implies $\lambda_{t-j} = 0$, as well. Thus, all coefficients in (3) must be 0, hence the polynomials are linearly independent. $\square$

References

[1] R.P. ANSTEE, Some Problems Concerning Forbidden Configurations, preprint.

[2] R.P. ANSTEE, J.R. GRIGGS, A. SALI, Small Forbidden Configurations, *Graphs and Combinatorics* **13**(1997),97-118.

[3] R.P. ANSTEE, R. FERGUSON, A. SALI, Small Forbidden Configurations II, *Electronic J. Combin.* **8**(2001), R4 (25pp)

[4] R.P. ANSTEE, Z. FÜREDI, Forbidden Submatrices, *Discrete Math.* **62**(1986),225-243.

[5] R.P. ANSTEE, N. KAMOOSI, Small Forbidden Configurations III, preprint.

[6] R.P. ANSTEE, A. SALI Small forbidden configurations IV: The 3-rowed case, *Combinatorica*, to appear.

[7] J. BALOGH, B. BOLLOBÁS, Unavoidable Traces of Set Systems, to appear, Combinatorica.

[8] B. BOLLOBÁS Extremal graph theory. London Mathematical Society Monographs, **11** Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1978.

[9] P. ERDŐS, A.H. STONE, On the Structure of Linear Graphs, *Bull. Amer. Math. Soc.* **52**(1946), 1089-1091.

[10] Z. FÜREDI, personal communication.

[11] T. KŐVÁRI, V.T. SÓS, P. TURÁN On a problem of K. Zarankiewicz, *Colloquium Math.* **3** (1954) 50-57.

[12] L. LOVÁSZ Cromatic number of hypergraphs and linear algebra, *Studia Sci. Math. Hung.* **11** (1976) 113-114.

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

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

[15] 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.