Design Theory and some Forbidden Configurations

R.P. Anstee[*], Farzin Barekat[†], Zachary Pellegrin[‡]
Mathematics Department
The University of British Columbia
Vancouver, B.C. Canada V6T 1Z2
anstee@math.ubc.ca

September 27, 2019

Abstract

In this paper we relate t-designs to a forbidden configuration problem in extremal set theory. Let $\mathbf{1}_t\mathbf{0}_\ell$ denote a column of t 1's on top of ℓ 0's. Let $q \cdot \mathbf{1}_t\mathbf{0}_\ell$ denote the $(t+\ell) \times q$ matrix consisting of t rows of q 1's and ℓ rows of q 0's. We consider extremal problems for matrices avoiding certain submatrices. Let A be a (0,1)-matrix forbidding any $(t+\ell) \times (\lambda+2)$ submatrix $(\lambda+2) \cdot \mathbf{1}_t\mathbf{0}_\ell$. Assume A is m-rowed and only columns of sum $t+1, t+2, \ldots, m-\ell$ are allowed to be repeated. Assume that A has the maximum number of columns subject to the given restrictions assume m is sufficiently large. Then A has each column of sum $0, 1, \ldots, t$ and $m-\ell+1, m-\ell+2, \ldots, m$ exactly once and, given the appropriate divisibility condition, the columns of sum $t+1$ correspond to a t-design with block size $t+1$ and parameter λ. The proof derives a basic upper bound on the number of columns of A by a pigeonhole argument and then a careful argument, for large m, reduces the bound by a substantial amount down to the value given by design based constructions. We extend in a few directions.

Keywords: design theory, t-designs, extremal set theory, (0,1)-matrices, forbidden configurations

1 Introduction

We explore a connection between block designs and extremal set theory. Combinatorial objects can be defined by forbidden substructures. Let $[m] = \{1, 2, \ldots, m\}$. Also for any

[*]Research supported in part by NSERC
[†]Research supported in 2007 by NSERC of first author
[‡]Research supported in 2019 by NSERC USRA and NSERC of first author

finite set S, let $\binom{S}{k}$ denote all k-subsets of S. Thus $|\binom{[m]}{k}| = \binom{m}{k}$. An $m \times n$ (0,1)-matrix A encodes a multiset $\mathcal{F}$ on $[m]$ consisting of n sets (counted by multiplicity) where each column of A is the incidence vector of a set in $\mathcal{F}$. In this paper we consider certain (0,1)-submatrices (called *configurations*) as the forbidden substructures of interest.

Consider a multiset $\mathcal{F}$ where each element of $\mathcal{F}$ is a called a *block*. We say $\mathcal{F}$ is t-(m,k,λ) *design* if each set $B \in \mathcal{F}$ has $B \in \binom{[m]}{k}$ (namely blocks of size k) and for each t-set $S \in \binom{[m]}{t}$ there are exactly λ sets (blocks) in $\mathcal{F}$ containing S, with sets counted according to multiplicity . It is usual in the study of block designs to use v instead of m and to allow repeated blocks. Recent results of Keevash [8] yield that (for fixed t, k, λ and m large) t-designs exist assuming the easy *divisibility conditions*:

$$\binom{k-i}{t-i} \text{ divides } \lambda \binom{m-i}{t-i} \text{ for } i = 1, 2, \ldots, t-1. \tag{1}$$

Moreover, Keevash shows that we can require that there are no repeated blocks i.e. $\mathcal{F}$ is a set. In that we case we call $\mathcal{F}$ a *simple* design.

Theorem 1.1 *(Keevash [8]) Let t, λ, k be given. Assume m is sufficiently large and satisfies the divisibility conditions (1). Then a simple t-(m,k,λ) design exists.*

Our initial investigations [2] only considered 2-$(m,3,\lambda)$ designs because we needed simple designs. A result of Dehon [7] establishes the existence of simple 2-$(m,3,\lambda)$ designs. The result of Keevash [8] above establishes the existence of simple t-designs, for $t \geq 3$ and it suggested seeking greater generality than in [2].

In the paper we use a (0,1)-matrix interpretation of sets. Let $\|A\|$ denote the number of columns of A (counted by multiplicity if that is relevant). For an m-rowed matrix A and a set $S \subseteq [m]$, let $A|_S$ denote the submatrix of A formed by the rows S. Let $\mathbf{1}_t\mathbf{0}_\ell$ denote the $(t+\ell) \times 1$ vector of t 1's on top of ℓ 0's. For a $s \times 1$ vector $\mathbf{v}$, Let $t \cdot \mathbf{v}$ denote the $s \times t$ matrix of t copies of $\mathbf{v}$.

Theorem 1.2 *Let t, λ, k, m be given. Let A be a (0,1)-matrix with column sums k. Assume that for each each t-set $S \in \binom{[m]}{t}$ that $A|_S$ contains $\lambda \cdot \mathbf{1}_t$ and does not contain $(\lambda+1) \cdot \mathbf{1}_t$. Then $\|A\| = \lambda\binom{m}{t}/\binom{k}{t}$ and A is the incidence matrix of a t-(m,k,λ) design.*

Theorem 1.2 corresponds to the usual definition of a design. We could state a version of this by only requiring $A|_S$ contains $\lambda \cdot \mathbf{1}_t$ but also requiring $\|A\| = \lambda\binom{m}{t}/\binom{k}{t}$.

Our motivation for studying these problems came from extremal set theory. An $m \times n$ (0,1)-matrix A can be thought of a multiset of n subsets of $[m]$. For an $m \times 1$ (0,1)-column α, we define

$$I(\alpha) = \{i \in [m] \,:\, \alpha \text{ has 1 in row } i\}. \tag{2}$$

From this we define the natural multiset system $\mathcal{A}$ associated with the matrix A:

$$\mathcal{A} = \{I(\alpha_i) \,:\, \alpha_i \text{ is column } i \text{ of } A\}. \tag{3}$$

Similarly, if we are given a multiset system $\mathcal{A}$, we can form its *incidence* matrix A, as long as we don't care about column order. We define a *simple* matrix A as a (0,1)-matrix with no repeated columns. In this case $\mathcal{A}$ yields a *set* system and it is in this setting that extremal set theory problems are typically stated.

The property of forbidding a submatrix is usually extended to forbidding any row and column permutation of the submatrix. Let A and F be (0,1)-matrices. We say that A has F as a *configuration* and write $F \prec A$ if there is a submatrix of A which is a row and column permutation of F. For the configuration $F = s \cdot \mathbf{1}_t\mathbf{0}_\ell$ only row permutations actually matter but we are motivated by the study of forbidden configurations where row and column permutations matter [1].

Theorem 1.3 *Let t, λ, k, m be given. Let A be an m-rowed (0,1)-matrix with column sums k. Assume that $(\lambda+1)\cdot \mathbf{1}_t \not\prec A$. Then*

$$\|A\| \leq \lambda \binom{m}{t} / \binom{k}{t}.$$

Moreover in the case of equality the columns of columns in A form the incidence matrix of a t-(m,k,λ) design.

Proof: Let A be a $m \times \lambda\binom{m}{t}/\binom{k}{t}$ incidence matrix of a t-(m,k,λ) design. Then A provides a construction yielding the lower bound. The upper bound follows from a straightforward pigeonhole argument. Each column of sum k has $\binom{k}{t}$ t-subsets of rows containing t 1's. We can have at most λ such columns for a given t-set in A. ■

Our extremal matrices, under a forbidden configuration restriction, yield a design. The following result is the analogue of Theorem 1.3 where the hypotheses are altered and weakened. The theorem relates interesting forbidden configurations (i.e. $(\lambda+2)\cdot \mathbf{1}_t\mathbf{0}_\ell$) to designs. Some repeated columns are allowed in the manner of Design Theory investigations. Our main result is:

Theorem 1.4 *Let t, ℓ, λ be given with $t > \ell$. Let A be an m-rowed (0,1)-matrix with no repeated columns of sum $0, 1, \ldots, t$ nor column sums $m-\ell+1, m-\ell+2, \ldots, m$. Assume that $(\lambda+2)\cdot \mathbf{1}_t\mathbf{0}_\ell \not\prec A$. Then there exist an M so that for $m \geq M$,*

$$\|A\| \leq \sum_{i=0}^{t-1}\binom{m}{i} + \left(1+\frac{\lambda}{t+1}\right)\binom{m}{t} + \sum_{i=m-\ell+1}^{m}\binom{m}{i}. \tag{4}$$

Moreover in the case of equality the columns of column sum $t+1$ in A form the incidence matrix of a t-$(m, t+1, \lambda)$ design and all columns of sum $1, 2, \ldots, t$ and $m-\ell+1, m-\ell+2, \ldots, m$ are present. ■

Of course, a similar result holds for $\ell > t$. We chose the multiplier for $\mathbf{1}_t\mathbf{0}_\ell$ to be $\lambda+2$ so that we would end up with a t-$(m, t+1, \lambda)$ design and connect with Design Theory notation. We will prove this in Section 3. General Lemmas are in Section 2.

This specializes to a forbidden configuration result. Define

$$\text{Avoid}(m, F) = \{A : A \text{ is an } m\text{-rowed matrix with } F \nprec A\}.$$

Then our extremal set theory problem is:

$$\text{forb}(m, F) = \max_A\{\|A\| : A \in \text{Avoid}(m, F), A \text{ is simple }\}.$$

These problems have been extensively investigated [1]. Exact results have been rare for non-simple configurations F. One exception is $\text{forb}(m, q \cdot K_t) = \text{forb}(m, q \cdot \mathbf{1}_t)$ where the design constructions achieve equality [3]. We would like to handle all cases $F = (\lambda + 2) \cdot (\mathbf{1}_t\mathbf{0}_\ell)$ but this paper only succeeds when $t > \ell$ (or, of course, $t < \ell$ by taking (0,1)-complements) or the special case $t = \ell = 2$ (see Theorem 4.3). Define K_k to be the $k \times 2^k$ incidence matrix of all subsets of $[k]$ and define K_k^s to be the $k \times \binom{k}{s}$ incidence matrix of $\binom{[k]}{s}$. We specialize Theorem 1.4 to simple matrices to obtain the following.

Corollary 1.5 *Let t, ℓ, λ be given with $t > \ell$. Then there exists an M so that for $m > M$:*

$$\textit{forb}(m, (\lambda + 2) \cdot \mathbf{1}_t\mathbf{0}_\ell) \leq \sum_{i=0}^{t-1} \binom{m}{i} + \left(1 + \frac{\lambda}{t+1}\right)\binom{m}{t} + \sum_{i=m-\ell+1}^{m} \binom{m}{i}. \qquad (5)$$

Equality is only achieved when (1) is satisfied (with $k = t + 1$) and for a matrix $[K_m^0 K_m^1 \cdots K_m^t\, T\, K_m^{m-\ell+1} \cdots K_m^m]$ where T is the incidence matrix of a simple t-$(m, t + 1, \lambda)$ design.

We have some alternate constructions for the case of equality for small m. For $m = \lambda + t + \ell$ we have $A = [K_m^0 K_m^1 K_m^2 \cdot K_m^t K_m^{t+1} K_m^{m-\ell+1} K_m^{m-\ell+2} \cdots K_m^m] \in \text{Avoid}(m, (\lambda + 2) \cdot \mathbf{1}_t\mathbf{0}_\ell)$. Choose some subset $S \in \binom{[m]}{t}$. We check that there are exactly $m - t$ columns in K_m^{t+1} that are all 1's on rows S and one further column of K_m^t that is all 1's on rows S. Moreover on the remaining $m - t = \lambda + \ell$ rows we have $(\lambda + 1)\mathbf{0}_\ell \nprec K_{\lambda+\ell}^1$. Thus our construction is in $\text{Avoid}(m, (\lambda + 2) \cdot \mathbf{1}_t\mathbf{0}_\ell)$. We check that the total number of columns is bigger than our design construction bound (4) by $\binom{m}{t+1} - \frac{\lambda}{t+1}\binom{m}{t} = \frac{\ell}{t+1}\binom{\lambda+t+\ell}{t}$ (given $m = \lambda + t + \ell$). Thus we need some condition on m being large in order to obtain our result. In essence, the pigeonhole argument explored in Lemma 2.2 is insufficient to prove our bounds. The following would be a version of Theorem 1.4 more in keeping with Theorem 1.2.

Theorem 1.6 *Let t, ℓ, λ be given with $t > \ell$. There exists an M so that for $m > M$, if A is an $m \times n$ (0,1)-matrix with column sums in $\{t + 1, t + 2, \ldots, m - 1\}$ and $A \in \text{Avoid}(m, (\lambda + 1) \cdot \mathbf{1}_t\mathbf{0}_\ell)$ then*

$$n \leq \frac{\lambda}{t+1}\binom{m}{t} \qquad (6)$$

and we have equality if and only if the columns of A correspond to the $(t + 1)$-sets of a t-$(m, t + 1, \lambda)$ design. ■

The cases with $t = \ell$ would be more difficult since $\mathbf{1}_t\mathbf{0}_t$ is self-complementary (under (0,1)-complements) and matrices in $\text{Avoid}(m, \mathbf{1}_t\mathbf{0}_t)$ could easily have large column sums. Theorem 4.2 in Section 4 is an example of this. For small m such as $m = \lambda + 3$, the construction $A = [K_m^0 K_m^1 K_m^2 K_m^3 K_m^{m-2} K_m^{m-1} K_m^m] \in \text{Avoid}(m, (\lambda+2)\cdot \mathbf{1}_2\mathbf{0}_2)$ and exceeds the bound (21) below much as is true above for $(\lambda + 2)\cdot \mathbf{1}_t\mathbf{0}_\ell$. Since we were unable to generalize Theorem 4.2 to $\mathbf{1}_t\mathbf{0}_t$ for $t \geq 3$, we will not prove Theorem 4.2 here but state it for completeness in Section 4 with the unrefereed proof in the arXiv [2].

The proof of Theorem 4.2 uses Turán's bound (Theorem 2.1). If there were improvements on Turán's bound [6] for $t \geq 3$, they might help handle the configurations $(\lambda + 3)\cdot \mathbf{1}_t\mathbf{0}_t$ but there has been no recent improvements in the bounds.

2 Basic Lemmas for $\text{Avoid}(m, (\lambda + 2)\cdot \mathbf{1}_t\mathbf{0}_\ell)$

First we state an important result we use. The following bounds were proven by Turán [9] for $t = 2$ and by de Caen [6] for general t. Perhaps better bounds are possible for general t.

Theorem 2.1 *Turán Bounds. Let k, t, m be given with $k \leq m$. Let G be a collection of n distinct sets in $\binom{[m]}{t}$. For n satisfying*

$$n \geq \binom{m}{t} - \frac{m-k+1}{m-t+1}\binom{m}{t}\Big/\binom{k-1}{t-1},$$

there exists a set $S \subset [m]$ with $|S| = k$ so that all $\binom{k}{t}$ t-subsets $\binom{S}{t}$ are in G. ∎

We follow some of the arguments noted in [2] as well as new arguments to obtain Theorem 1.4. The Lemmas below consider an m-rowed matrix $A \in \text{Avoid}(m, (\lambda + 2)\cdot \mathbf{1}_t\mathbf{0}_\ell)$. Assume A has the property that the column sums are restricted to $\{t, t+1, \ldots, m-\ell\}$ and that columns of sum t are not repeated. Note that columns of sum $1, 2, \ldots, t-1$ and $m-\ell+1, m-\ell+2, \ldots, m$ do not contribute to the forbidden configuration $(\lambda+2)\cdot \mathbf{1}_t\mathbf{0}_\ell$. Also note that if we allowed repeated columns of sum t, then we would get a less interesting result. For $i = t, t+1$, let a_i denote the number of columns of column sum i in A and let $a_{\geq t+2}$ denote the number of columns of column sum at least $t+2$ in A. Given the nature of the extremal matrices yielding equality in the bound (4), we expect $a_t = \binom{m}{t}$ and $a_{\geq t+2} = 0$. We will be proving Theorem 1.4 by contradiction and will assume that

$$a_t + a_{t+1} + a_{\geq t+2} > \left(1 + \frac{\lambda}{t+1}\right)\binom{m}{t}. \tag{7}$$

Our first Lemma extends a pigeonhole principle.

Lemma 2.2 *Let m, t, ℓ, λ be given with $t > \ell$. Let A be an $m \times n$ matrix with no $(\lambda+2)\cdot(\mathbf{1}_t\mathbf{0}_\ell)$, columns sums in $\{t, t+1, \ldots, m-\ell\}$ and with no repeated columns of sum t. Assume $m \geq t+\ell+\lambda+2$ and (7). Then*

$$\binom{t}{t}\binom{m-t}{\ell}a_t + \binom{t+1}{t}\binom{m-t-1}{\ell}a_{t+1} + \binom{t+2}{t}\binom{m-t-2}{\ell}a_{\geq t+2}$$

$$\leq \binom{m}{t+\ell}\binom{t+\ell}{\ell}(\lambda+1). \tag{8}$$

There exists positive constants c_1, c_2 so that

$$\binom{m}{t} - c_1 m^{t-1} \leq a_t \leq \binom{m}{t} \tag{9}$$

$$\text{and} \quad a_{\geq t+2} \leq c_2 m^{t-1}. \tag{10}$$

Proof: We note that a column of column sum k has $\binom{k}{t}\binom{m-k}{\ell}$ configurations $\mathbf{1}_t\mathbf{0}_\ell$ and note that $\binom{k}{t}\binom{m-k}{\ell} \geq \binom{t+2}{t}\binom{m-t-2}{\ell}$ for $t+2 \leq k \leq m-1$. Counting the configurations $\mathbf{1}_t\mathbf{0}_\ell$ (which can appear on $t+\ell$ rows in up to $\binom{t+\ell}{\ell}$ orderings) and using the pigeonhole argument yields (8)

For $m \geq 3\ell+t+1$ we have $\binom{t+1}{t}\binom{m-t-1}{\ell} < \binom{t+2}{t}\binom{m-t-2}{\ell}$. Hence

$$\binom{m-t}{\ell}a_t + (t+1)\binom{m-t-1}{\ell}(a_{t+1} + a_{\geq t+2}) \leq \binom{m}{t+\ell}\binom{t+\ell}{\ell}(\lambda+1).$$

From (7), we have $a_{t+1} + a_{\geq t+2} \geq \left(1 + \frac{\lambda}{t+1}\right)\binom{m}{t} - a_t$. We substitute and obtain

$$\binom{m-t-1}{\ell}\binom{m}{t}(t+1+\lambda) - \binom{m}{t+\ell}\binom{t+\ell}{\ell}(\lambda+1)$$

$$\leq \left((t+1)\binom{m-t-1}{\ell} - \binom{m-t}{\ell}\right)a_t,$$

We deduce that there is a constant c_1 (will depend on λ, t, ℓ) so that first half of (9) holds. The second half of (9) follows from the fact that no column of sum t is repeated. In a similar way we have

$$\binom{m-t}{\ell}a_t + (t+1)\binom{m-t-1}{\ell}\left(1 + \frac{\lambda}{t+1}\binom{m}{t} - a_t - a_{\geq t+2}\right)$$

$$+\binom{t+2}{t}\binom{m-t-2}{\ell}a_{\geq t+2}, \leq \binom{m}{t+\ell}\binom{t+\ell}{\ell}(\lambda+1)$$

and when we substitute the upper bound of (9), we deduce that there is a constant c_2 (will depend on λ, t, ℓ) so that (10) holds. ■

Partition A into three parts: A_t consists of the columns of column sum t, A_{t+1} is the columns of column sum $t+1$ and $A_{\geq t+2}$ is the columns of column sum greater or equal than $t+2$. We construct $\mathcal{A}_t$, $\mathcal{A}_{t+1}$ from A_t and A_{t+1} using the notations of (2) and (3). Note that $\mathcal{A}_t$ is a set given that there are no repeated columns of sum t while $\mathcal{A}_{t+1}$ is a multiset. Let $S = \{i_1, ..., i_t\} \in \binom{[m]}{t}$. Then define:

$$\mu(S) = \begin{cases} 1 & \text{if } S \in \mathcal{A}_t \\ 0 & \text{if } S \notin \mathcal{A}_t \end{cases} \quad , \qquad E = \{S \in \binom{[m]}{t} \,:\, \mu(S) = 0\} = \binom{[m]}{t} \setminus \mathcal{A}_t \quad . \tag{11}$$

Thus E denotes the t-sets missing from $\mathcal{A}_t$. We expect $E = \emptyset$. Now

$$a_t = \sum_{S \in \binom{[m]}{t}} \mu(S) = \binom{m}{t} - |E|. \tag{12}$$

We deduce from (9) that $|E| \leq c_1 m^{t-1}$. We use hypergraph degree definitions applied to the multiset $\mathcal{A}_{t+1}$. For $S \in \binom{[m]}{t}$, define

$$d(S) = |\{x \in [m] \,:\, S \cup x \in \mathcal{A}_{t+1}\}|, \tag{13}$$

where we count the sets $S \cup x$ with their multiplicity in $\mathcal{A}_{t+1}\}$. For example, if we have $t = 2$ and sets $\{1,2,3\}$, $\{1,2,4\}$, $\{1,2,4\}$ in $\mathcal{A}_3$ ($\mathcal{A}_3$ may have repeated columns) then $d(\{1,2\}) = 3$. Thus

$$(t+1) \cdot a_{t+1} = \sum_{S \in \binom{[m]}{t}} d(S). \tag{14}$$

Since $m > \lambda + \ell + t + 1$ and we are avoiding $(\lambda+2) \cdot (\mathbf{1}_t \mathbf{0}_\ell)$ in A_{t+1} then $|\{G \,:\, G \in \mathcal{A}_{t+1}$ and $S \subset G\}| < \lambda + \ell + 1$,where we count G's according to the multiplicity in $\mathcal{A}_{t+1}\}$.

Lemma 2.3 *Let A satisfy hypotheses of Lemma 2.2. Then*

$$d(S) \leq (\lambda + 1) - \mu(S). \tag{15}$$

Proof: Recall (13) for which we are counting by multiplicity the $(t+1)$-sets containing a given t-set S. We proceed to a contradiction by assuming the opposite of (15), namely we have an $S \in \binom{[m]}{t}$ with $d(S) + \mu(S) \geq \lambda + 2$. Let $B_1, B_2, \ldots, B_{\lambda+2}$ denote $\lambda + 2$ sets in $\mathcal{A}_t \cup \mathcal{A}_{t+1}$, each containing the t-set S. Thus there are at most $\lambda + 2$ elements in $\cup_i B_i$ which are not already in S. Thus for $m \geq \lambda + 2 + t + \ell$, we will have ℓ elements of $[m]$ not in any B_i yielding the configuration $(\lambda+1) \cdot \mathbf{1}_t \mathbf{0}_\ell$, a contradiction. ∎

Our hypothetical extremal construction for A avoiding $(\lambda+2) \cdot \mathbf{1}_t \mathbf{0}_\ell$ is $[K_m^t | T]$ where T is the incidence matrix of a t-$(m, t+1, \lambda)$ design. In that case all t-sets are present exactly once and for any t-set S, the number of sets, apart from the t-set $S \in \binom{[m]}{t}$, containing S is λ and they are of size $t+1$. Let Y denote these 'typical' t-sets :

$$Y = \{S \in \binom{[m]}{t} \,:\, d(S) = \lambda \text{ and } \mu(S) = 1\} \tag{16}$$

We wish to show $Y = \binom{[m]}{t}$ in our proof of Theorem 1.4. The following Lemma is a step in that direction.

Lemma 2.4 *Let A satisfy hypotheses of Lemma 2.2. There exists a constant c_3 so that*

$$|Y| \geq \binom{m}{t} - c_3 m^{t-1} \tag{17}$$

Proof: We partition the $\binom{m}{t}$ t-sets S into 3 parts: Y, E and the rest. By Lemma 2.3, for each $S \in E$ we have $d(S) \leq \lambda + 1$. Note that for $S \notin Y \cup E$, we have $\mu(S) = 1$ and so $d(S) \leq \lambda - 1$ (else $S \in Y$). Thus from (14) and Lemma 2.3,

$$(t+1)a_{t+1} = \sum_{S \in \binom{[m]}{t}} d(S) \leq \left(\lambda|Y| + (\lambda+1)|E| + (\lambda-1)\left(\binom{m}{t} - |Y| - |E|\right)\right)$$

Hence

$$a_{t+1} \leq \frac{1}{t+1}\left(\lambda\binom{m}{t} + |E| - \binom{m}{t} + |Y| + |E|\right) \tag{18}$$

Substituting estimates of a_t, a_{t+1}, $a_{\geq t+2}$ from (9), (14), (10) into (7), we have

$$\binom{m}{t} - |E| + \frac{1}{t+1}\left(\lambda\binom{m}{t} + 2|E| - \binom{m}{t} + |Y|\right) + c_2 m^{t-1} > \left(1 + \frac{\lambda}{t+1}\right)\binom{m}{t}$$

We deduce $\left(\frac{2}{t+1} - 1\right)|E| + \frac{1}{t+1}|Y| + c_2 m^{t-1} > \frac{1}{t+1}\binom{m}{t}$ and so there exists a constant $c_3 = (t+1)c_2$ so that (17) holds. ∎

Lemma 2.5 *Let k be given. Use the notations of Section 2. There exists an M so that for $m \geq M$, there exists a set of k rows B such that for any t-set $S \in \binom{B}{t}$ then $S \in Y$.*

Proof: Form a t-hypergraph G of m vertices corresponding to the rows of A and with edge S if and only if $S \in Y$. Thus by Lemma 2.4, the number of edges (t-sets) of G is at least $\binom{m}{t} - c_3 m^{t-1}$. We apply Theorem 2.1, by a result of de Caen [6]. Thus there exists an M so that for $m \geq M$, there is a $B \subset [m]$ with $|B| = k$ so for any $S \in \binom{B}{t}$ we have $S \in Y$. Hence for $S \in \binom{B}{t}$ we have $d(S) = \lambda$ and $\mu(S) = 1$. ∎

3 Exact Bound for $(\lambda+2)\cdot(\mathbf{1}_t\mathbf{0}_\ell)$

The following two lemmas provide useful counting inequalities. Our main idea is that if we have column sums t, whether repeated or not, that avoid $(\lambda+2)\cdot\mathbf{1}_{t-1}$ we may use a straightforward pigeonhole bound that the number of columns is at most a constant times m^{t-1}.

Lemma 3.1 *Let A_{t+1} be the columns of sum $t+1$ in A. Given any row $r\in[m]$, let A^r_{t+1} be the submatrix of A_{t+1} formed by the columns having a 1 in row r, then*

$$a^r_{t+1}=||A^r_{t+1}||\le\frac{\lambda+1}{t}\binom{m-1}{t-1}.$$

Proof: Since $m>t+l+\lambda+2$, any matrix with column sums $t+1$ containing $(\lambda+2)\cdot\mathbf{1}_t$ must also contain $(\lambda+2)\cdot\mathbf{1}_t\mathbf{0}_\ell$, therefore A^r_{t+1} must avoid $(\lambda+2)\cdot\mathbf{1}_t$.

Since each column in A^r_{t+1} has a 1 in row r, on the remaining $m-1$ rows the matrix must avoid $(\lambda+2)\cdot\mathbf{1}_{t-1}$ since adding in row r would create $(\lambda+2)\cdot\mathbf{1}_t$ and therefore $(\lambda+2)\cdot\mathbf{1}_t\mathbf{0}_\ell$.

The bound for a matrix of column sum t avoiding $(\lambda+2)\cdot\mathbf{1}_{t-1}$ on $m-1$ rows is $\frac{\lambda+1}{t}\binom{m-1}{t-1}$ (see Theorem 1.6) , thus

$$||A^r_{t+1}||\le\frac{\lambda+1}{t}\binom{m-1}{t-1}\qquad\blacksquare$$

Lemma 3.2 *Let A_{t+1} be the columns of sum $t+1$ in A. Given any set of rows $R\in[m]$, let A^R_{t+1} be the submatrix of A_{t+1} formed by the columns having a 1 in any row $r\in R$ then*

$$a^R_{t+1}=||A^R_{t+1}||\le|R|\cdot\frac{\lambda+1}{t}\binom{m-1}{t-1}.$$

Proof: By definition, A^R_{t+1} is all columns which are in A^r_{t+1} for some $r\in R$, therefore

$$||A^R_{t+1}||\le\sum_r||A^r_{t+1}||\le|R|\cdot\frac{\lambda+1}{t}\binom{m-1}{t-1}\qquad\blacksquare$$

Lemma 3.3 *Let $R\in[m]$ be any set of rows of constant size $|R|=\rho$. Define A^R_{t+1} according to the notation of Lemma 3.2 and construct $\mathcal{A}^R_{t+1}$ from A^R_{t+1} using the notations of (2) and (3). Let*

$$W_R=\{S\in\binom{[m]}{t}\ :\ \exists\, x\in[m]\ s.t.\ S\cup x\in\mathcal{A}^R_{t+1}\}.$$

Let

$$Z_R=Y\setminus W_R\ .$$

Then, for m large enough, $|Z_R|>0$.

Proof: Each $t+1$-set in $\mathcal{A}_{t+1}^R$ contributes $t+1$ t-sets to W_R, therefore

$$|W_R| \le (t+1)\cdot|\mathcal{A}_{t+1}^R| \le (t+1)\rho\cdot\frac{\lambda+1}{t}\binom{m-1}{t-1}.$$

Thus there will exist a constant c_4 with

$$|Z_R| = |Y| - |W_R| \ge \binom{m}{t} - c_3m^{t-1} - (t+1)\rho\cdot\frac{\lambda+1}{t}\binom{m-1}{t-1} \ge \binom{m}{t} - c_4m^{t-1}.$$

For m sufficiently large, $\binom{m}{t} > c_4m^{t-1}$, therefore $|Z_R| > 0$ $\blacksquare$

In the following lemma, we use the result $Z_R \neq \emptyset$.

Lemma 3.4 *A has no column with fewer than $\lambda+\ell$ 0's*

Proof: We assume for contradiction that there exists some column α with fewer than $\lambda+\ell$ 0's. Let $R\subset[m]$ be the set of rows on which α is 0. By assumption $|R| < (\lambda+\ell)$ but we also note $R > \ell$ by assumptions on A. Construct Z_R according to Lemma 3.3. Then since $|R|$ is bounded by a constant, by Lemma 3.3, $|Z_R| > 0$.

Thus there exists some t-set $S\in Z_R$ such that $d(S)=\lambda$, $\mu(S)=1$ since $Z_R\subseteq Y$ and using (16). Each of the $\lambda+1$ columns contributing to this count has 0's in all rows of R since this count consists of a column of sum t which is 0 outside of S and columns of sum $t+1$ which do not contribute to W_R and are therefore not in A_{t+1}^R. Thus these columns must have no 1 in R. Since the columns of sum $t+1$ have no 1 in rows R we must have that $S\cap R=\emptyset$ and therefore the column of sum t is also 0's in rows R. Also, since α is 1 outside of R, α has only 1's in rows S and is also 0's in rows R.

Therefore we have λ columns of sum $t+1$ with 1's in S, a column of sum t with 1's in S and a column α has 1's in rows S. Thus we have $(\lambda+2)\cdot\mathbf{1}_t$ in rows S in these $\lambda+2$ columns. Additionally as argued above, each of these columns has 0's in rows R. Recalling that $|R|\ge\ell$, this creates the forbidden object $(\lambda+2)\cdot\mathbf{1}_t\mathbf{0}_\ell$ on these columns, a contradiction.

Thus no such column α can exist and so A has no column with fewer than $\lambda+\ell$ 0's. $\blacksquare$

Lemma 3.5 *Let $A\in\text{Avoid}(m,(\lambda+2)\cdot\mathbf{1}_t\mathbf{0}_\ell)$ with column sums in $\{t,t+1,\ldots,m-\ell\}$. Assume no column of sum t is repeated. Then*

$$||A|| \le \left(1+\frac{\lambda}{t+1}\right)\binom{m}{t}$$

Proof: Assume for contradiction that A exceeds this bound then A must exceed the bound for avoiding $(\lambda+2)\cdot 1_t$, so must contain this object on some set of $\lambda+2$ columns. Let D be the matrix of these columns. Therefore A contains the rows of 1's of the forbidden object and it remains to show that A must also contain the rows of 0's in appropriately chosen columns. If there are at least ℓ rows of 0's in D then A contains the forbidden object. Otherwise D must have at most $\ell - 1$ rows of 0's and all other rows of D have at least one 1.

By Lemma 2.5, with $k = (\lambda+2)\cdot(t+1)+\ell$, there exists some clique of rows $B \subset [m]$ with $|B| = (\lambda+2)\cdot(t+1)+\ell$ for which any t-set is in Y. Recall that D must have at most $\ell-1$ rows of 0's and all other rows have at least one 1 and hence $D|_B$ has at least $(\lambda+2)\cdot(t+1)+1$ 1's. By pigeonhole principle, there must be some column β in D with $t+2$ 1's in the rows of B. Take any t-set from these $t+2$ rows. This set of t rows of B must be in Y. Therefore, since α has column sum at least $t+2$, there exist λ columns of sum $t+1$ with 1's in these rows and a column of sum t with 1's in these rows. These along with β create the rows of 1's in the forbidden object.

All other rows of the column of sum t are 0 and the columns of sum $t+1$ have at most λ rows in which they are not all 0. By Lemma 3.4 β has at least $\lambda+\ell$ 0's. Therefore, there are at least ℓ rows in which all of these columns are 0 creating the forbidden object, a contradiction.

Thus A must satisfy this bound. $\blacksquare$

Proof of Theorem 1.4: The upper bound follows from Lemma 3.5. We now consider the case of equality. We can repeat the previous arguments with equality and Lemma 2.5 and Lemma 3.4 with be true. Using the same arguments as Lemma 3.5 we see that if A contains $(\lambda+2)\cdot \mathbf{1}_t$ and equals the bound then A must contain our forbidden object $(\lambda+2)\cdot \mathbf{1}_t\mathbf{0}_\ell$. Therefore a matrix A achieving this bound avoids $(\lambda+2)\cdot \mathbf{1}_t\mathbf{0}_\ell$ if and only if A avoids $(\lambda+2)\cdot \mathbf{1}_t$. Thus in the case of equality we must have that $a_t = \binom{m}{t}$, $a_{\geq t+2} = 0$, $a_{t+1} = \frac{\lambda}{t+1}\binom{m}{t}$ and the columns of sum $t+1$ correspond to a $t-(m, t+1, \lambda)$ design. $\blacksquare$

4 The cases $F = q\cdot \mathbf{1}_1\mathbf{0}_1$, $F = q\cdot \mathbf{1}_2\mathbf{0}_2$ and further problems

Use the notation $[A|B]$ to denote the matrix formed by concatenating A with B. A related problem is attempting to compute $\text{forb}(m, [\mathbf{1}_{t+\ell}|\lambda\cdot \mathbf{1}_t\mathbf{0}_\ell])$. We have shown $\text{forb}(m, [\mathbf{1}_{t+1}|2\cdot \mathbf{1}_t\mathbf{0}_1]) = \text{forb}(m, 3\cdot \mathbf{1}_t)$ [5]. Note that $3\cdot \mathbf{1}_t \prec [\mathbf{1}_{t+1}|2\cdot \mathbf{1}_t\mathbf{0}_1]$. At this point we do not know but conjecture that $\text{forb}(m, [\mathbf{1}_{t+1}|\lambda\cdot \mathbf{1}_t\mathbf{0}_1]) = \text{forb}(m, (\lambda+1)\cdot \mathbf{1}_t)$ for large m. Note that $K_4 \in \text{Avoid}(m, [\mathbf{1}_3|3\cdot \mathbf{1}_2\mathbf{0}_1)$ but $\|A\| = 1 + \frac{5}{3}\binom{m}{2} + \binom{m}{1} + \binom{m}{0}$. We are using our intuition about what drives the bound forb.

The following results appear in the unrefereed manuscript [2]. In [4] we showed that

$$\left\lfloor \frac{q+1}{2}m \right\rfloor + 2 \leq \text{forb}(m, q\cdot(\mathbf{1}_1\mathbf{0}_1)) \leq \left\lfloor \frac{q+1}{2}m + \frac{(q-3)m}{2(m-2)} \right\rfloor + 2$$

where the upper bound obtained by a pigeonhole argument is achieved for $m = q-1$ by taking $A = [K_m^0 K_m^1 K_m^2 K_m^{m-1} K_m^m]$. For m with $m \geq \max\{3q+2, 8q-19\}$, we are able to show that the lower bound is correct and slice $\frac{(q-3)m}{2(m-2)} \approx \frac{q-3}{2}$ off a pigeonhole bound. It is likely that our bound is valid for smaller $m > q-1$. The case $q = 4$, is Lemma 3.1 in [5] and took a page to establish. The unrefereed manuscript [2] contains the following.

Theorem 4.1 *Let $q \geq 3$ be given. Then for $m \geq \max\{3q+2, 8q-19\}$,*

$$forb(m, q \cdot \mathbf{1}_1\mathbf{0}_1) = \lfloor \frac{q+1}{2} m \rfloor + 2. \qquad \blacksquare \tag{19}$$

The lower bound is easy. For m even or $q-3$ even, let G be a (simple) graph on m vertices for which all the degrees are $q-3$ and for $m, q-3$ odd let G be a graph for which $m-1$ vertices have degree $q-3$ and one vertex has degree $q-4$. Such graphs are easy to construct. Let H be the vertex-edge incidence matrix associated with G, namely for each edge $e = (i,j)$ of G, we add a column to H with 1's in rows i, j and 0's in other rows. Thus H is a simple m-rowed matrix with $\lfloor \frac{(q-3)m}{2} \rfloor$ columns each of column sum 2. The simple matrix $A = [K_m^0\, K_m^1\, H\, K_m^{m-1}\, K_m^m]$ has $\lfloor \frac{(q+1)m}{2} \rfloor + 2$ columns and no configuration $q \cdot (\mathbf{1}_1\mathbf{0}_1)$ which establishes $\mathrm{forb}(m, q \cdot (\mathbf{1}_1\mathbf{0}_1)) \geq \lfloor \frac{(q+1)m}{2} \rfloor + 2$.

To prove Corollary 4.3 and Theorem 4.4, we would prove the following:

Theorem 4.2 *[2] Let $\lambda > 0$, m, be given. Let A be an $m \times n$ (0,1)-matrix so that no column of sum 0,1,2, $m-2$, $m-1$ or m is repeated. Assume $A \in \mathrm{Avoid}(m, (\lambda+3)\cdot(\mathbf{1}_2\mathbf{0}_2)$. Then there exists a constant M so that for $m > M$,*

$$n \leq 2 + 2m + \left(2 + \frac{\lambda}{3}\right)\binom{m}{2} \tag{20}$$

with equality for $m \equiv 1, 3 (mod\ 6)$. If A is an $m \times forb(m, (\lambda+3) \cdot (\mathbf{1}_2\mathbf{0}_2))$ matrix with $m > M$ and $m \equiv 1, 3 (mod\ 6)$, then A consists of all possible columns of sum 0, 1, 2, $m-2$, $m-1$ and m once each and there are two positive integers a, b satisfying $a+b = \lambda$ with the columns of column sum 3 correspond to a $2-(m,3,a)$ design and the columns of sum $m-3$ correspond to the complements in $[m]$ of the blocks of a $2-(m,3,b)$ and A has no further columns. $\blacksquare$

Specializing to simple matrices we obtain the following:

Corollary 4.3 *[2] Let $\lambda > 0$ be given. There exists a constant $M = M(q)$ so that for $m > M$,*

$$forb(m, (\lambda+3) \cdot (\mathbf{1}_2\mathbf{0}_2)) \leq 2 + 2m + \left(2 + \frac{\lambda}{3}\right)\binom{m}{2}, \tag{21}$$

We have equality in (21) for $m \equiv 1, 3 (mod\ 6)$. If A is an $m \times forb(m, (\lambda+3) \cdot (\mathbf{1}_2\mathbf{0}_2))$ simple matrix with $m > M$ and $m \equiv 1, 3 (mod\ 6)$, then there exist positive integers a, b with $a + b = \lambda$ so that A consists of all possible columns of sum 0, 1, 2, $m-2$, $m-1$, m and with the columns of column sum 3 correspond to a $2-(m,3,a)$ design and the columns of sum $m-3$ correspond to the complements in $[m]$ of the blocks of a $2-(m,3,b)$ design and A has no further columns. $\blacksquare$

A design theory version of this is:

Theorem 4.4 *[2] Let λ and m be given integers. There exists an M so that for $m > M$, if A is an $m \times n$ (0,1)-matrix with column sums in $\{3, 4, \ldots, m-3\}$ and $A \in$* $\text{Avoid}(m, (\lambda+1) \cdot \mathbf{1}_2\mathbf{0}_2)$ *then*

$$n \leq \frac{\lambda}{3}\binom{m}{2}. \tag{22}$$

We have equality in (4.4) if and only if there are positive integers a, b satisfying $a+b=\lambda$ and there are $\frac{a}{3}\binom{m}{2}$ columns of A of column sum 3 corresponding to the blocks of a $2-(m,3,a)$ design and there are $\frac{b}{3}\binom{m}{2}$ columns of A of column sum $m-3$ of $m-3$-sets whose complements (in $[m]$) corresponding to the blocks of a $2-(m,3,b)$ design. ■

References

[1] R.P. Anstee, A survey of forbidden configuration results, *Electronic J. Combin.* **20**(2013), DS20

[2] R.P. Anstee, Farzin Barekat, Design Theory and Some Non-simple Forbidden Configurations, ArXiv 1909.07580

[3] R.P. Anstee, Z, Füredi, Forbidden Submatrices, *Discrete Mathematics* **62**(1986), 225-243.

[4] R.P. Anstee, R. Ferguson, A. Sali, Small Forbidden Configurations II, *Electronic J. Combin.* **8**(2001), R4 (25pp)

[5] R.P. Anstee, N. Kamoosi, Small Forbidden Configurations III, *Electronic J. Combin.* **14**(2007), R79 (34pp)

[6] D. de Caen, Extension of a theorem of Moon and Moser on complete subgraphs *Ars Combin.* **16**(1983), 5-10.

[7] Michel Dehon, On the existence of 2-designs $S_\lambda(2,3,v)$ without repeated blocks, *Discrete Math.* **43**(1983), 155-171.

[8] P. Keevash, The Existence of Designs, ArXiv: 1401.3665

[9] P. Turán, Eine Extremalaufgabe aus der Graphentheorie, *Math. Fiz Lapok.* **48**(1941), 436-452.