

Direct Constructions of d -disjunct matrices

Review

Macula's Method

$M: \binom{n}{d} \times \binom{n}{k}$ matrix, indexed by $\binom{N}{d}$ and $\binom{N}{k}$ where $|N| = n$.

e.g.

$$N = 7, d = 2, k = 3. \quad N = \mathbb{Z}_7.$$

	012	013	014	015	...
01	1	1			
02	1	0			
03	0	1			
		0			
	0				
12	1	0			
13	0	1			
		0			
56	0				

56

$$\binom{7}{2} \times \binom{7}{3}$$

$$21 \times 35$$

———— (1)

$$\begin{cases} M(i,j) = 1 \text{ if the } i^{\text{th}} \text{ } d\text{-subset is contained in } j^{\text{th}} \text{ } k\text{-subset of } N \\ M(i,j) = 0 \text{ otherwise} \end{cases}$$

$\Rightarrow M$ is d -disjunct. (So, the above pooling design is 2-disjunct)

Proof. Consider $d+1$ distinct columns, i.e., $d+1$ distinct

k -subsets of N , $C_1, C_2, \dots, C_{d+1}$. Since they are distinct

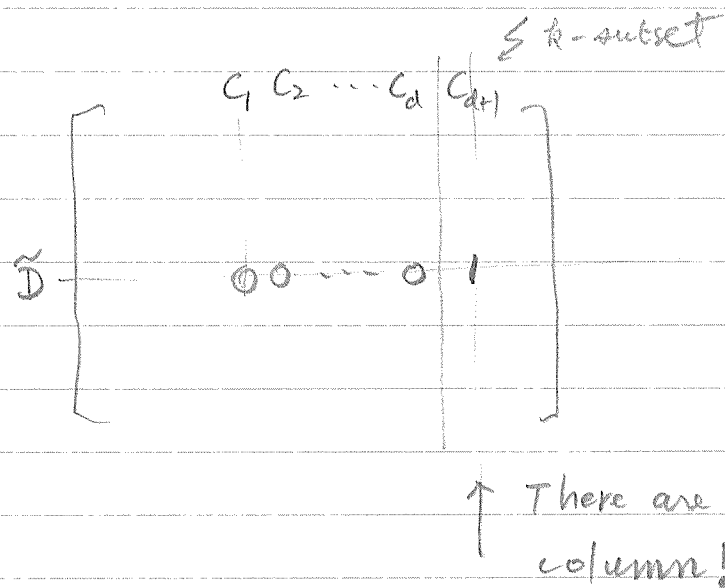
k -subsets of N , for each $j \in \{1, 2, \dots, d\}$, $C_{d+1} \setminus C_j \neq \emptyset$, let $x_j \in C_{d+1} \setminus C_j$

$D =$
Now, $\{x_1, x_2, \dots, x_d\}$ is a multi-set, i.e., some element may occur

more than once. Let D' be a subset of D . Then, $D' \subseteq C_{d+1}$.
(not a multiset)

Moreover, $D' \subseteq \tilde{D}$ where $\tilde{D}$ is a d -subset of N occurred in the i^{th} row of M . Now, $\tilde{D} \not\subseteq C_j$ for each $j \in \{1, 2, \dots, d\}$. Hence,

$C_{d+1} \not\subseteq \bigcup_{j=1}^d C_j$. This concludes the proof. $\blacksquare$



Observation 1 Any two columns can have at most $\binom{k-1}{d}$ common 1's.

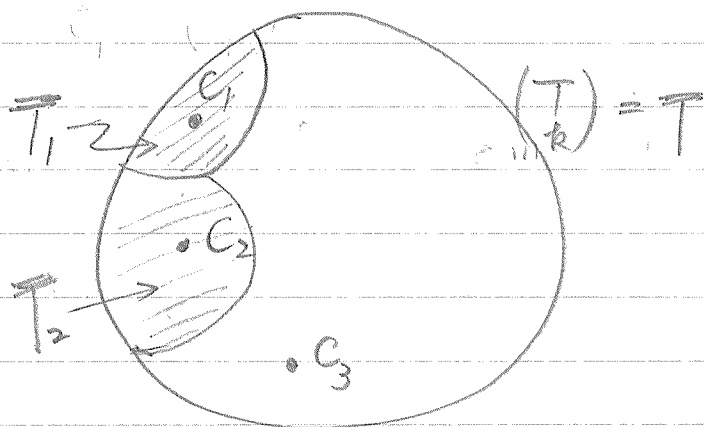
e.g. In (1), $\textcircled{015}$ and $\textcircled{014}$, do have a common one in row indexed by $\textcircled{01}$.

Observation 2 If we plan to cover a column by the other columns, then we need at least $\left\lceil \frac{\binom{k}{d}}{\binom{k-1}{d}} \right\rceil$ columns.

In general, this idea is called "Intersection Property" or column-intersection method.

Algorithm

Let $|T| = t$ and $\binom{T}{k} = \{k\text{-subsets of } T\}$. Define $r = \lceil t/(16d^2) \rceil$, $k = 4dr$ and $m = 4r$.



Choose C_1 .

Choose $C_2 \in \overline{T} \setminus T_1$

$$\overline{T}_1 = \{C \mid C \in T \text{ and } |C \cap C_1| \geq m\}.$$

$$T_2 = \{C \mid C \in T, C \notin T_1, \text{ and } |C \cap E_2| \geq m\}$$












Choose $C_n \in \mathbb{T} \setminus \bigcup_{i=1}^{n-1} (U_i, T_i)$

Theorem Let $C_1, C_2, \dots, C_n$ be the columns and $1, 2, \dots, t$ be the rows-indices. Then, we have a $t \times n$ d -disjunct matrix

with $n \geq 2.3 \left(\frac{t}{16d^2} - 1 \right)$

Proof. Since any two columns intersect in at most $m-1$ rows, (elements)

to cover a column, it takes at least $\left\lceil \frac{k}{4r-1} \right\rceil$ columns, which
($k=4dr$)

is larger than d . Hence the pooling design is d -disjunct.

Now, we estimate n .

$$\text{For each } T_j, j=1, 2, \dots, n, |T_j| \leq \sum_{i=m}^k \binom{k}{i} \cdot \binom{t-k}{k-i} \rightarrow \text{Not in common}$$

$$\text{This implies that } n \geq \frac{\binom{t}{k}}{\sum_{i=m}^k b_i}.$$

common elements

Observe that, for $3r+1 \leq i < k=4dr$,

$$\frac{b_i}{b_{i-1}} = \frac{\binom{k}{i} \binom{t-k}{k-i}}{\binom{k}{i-1} \binom{t-k}{k-i+1}}$$

$$= \frac{\binom{k}{i} \cdot \binom{t-k}{k-i}}{\binom{k}{i-1} \cdot \binom{t-k}{k-i+1}}$$

$$= \frac{\frac{k!}{i!(k-i)!}}{\frac{k!}{(i-1)!(k-i+1)!}} \cdot \frac{\frac{(t-k)!}{(k-i)!(t-k+i)!}}{\frac{(t-k)!}{(k-i+1)!(t-k+i-1)!}}$$

$$= \frac{k-i+1}{i} \cdot \frac{(k-i+1)}{(t-2k+i)}$$

$$= \frac{(k-i+1)^2}{(i)(t-2k+i)} \leq \frac{(4dr-3r)^2}{3r \cdot (16d^2-8dr+3r)} \quad (i=3r+1)$$

$$= \frac{(4d-3)^2}{3(16d^2-8d+3)} < \frac{1}{3}.$$

$$b_i < \left(\frac{1}{3}\right)^{i-3r} \cdot b_{3r}, \quad \forall 3r+1 \leq i < k = 4dr.$$

$$\begin{aligned} \sum_{i=m}^k b_i &= \left(\frac{1}{3}\right)^{-3r} \cdot b_{3r} \cdot \sum_{i=m}^k \left(\frac{1}{3}\right)^i \\ &= \left(\frac{1}{3}\right)^{-3r} \cdot b_{3r} \cdot \left(\frac{1}{3}\right)^m \cdot \left(1 + \frac{1}{3} + \dots + \left(\frac{1}{3}\right)^{k-m-1}\right) \xrightarrow{<} \frac{1}{1-\frac{1}{3}} \\ &< b_{3r} \cdot \left(\frac{1}{3}\right)^{m-3r-1} \xrightarrow{>} = 2b_{3r} \cdot 3^{1-r} \end{aligned}$$

Since $\binom{k}{r} = \sum_{i=0}^k b_i > \binom{k}{r} \cdot \left[\frac{\binom{k}{r}}{\sum_{i=m}^k b_i} \right] > 2 \cdot 3^{r-1} \geq 2 \cdot 3^{\left(\frac{k}{16d^2} - 1\right)}$

Too small!

So Solve for k .

$$n \geq 2 \cdot 3^{\frac{k}{16d^2} - 1}$$

$$\log_3 n \geq \log_3 2 + \frac{k}{16d^2} - 1$$

$$k \leq 16d^2 (\log_3 n - \log_3 2 + 1).$$

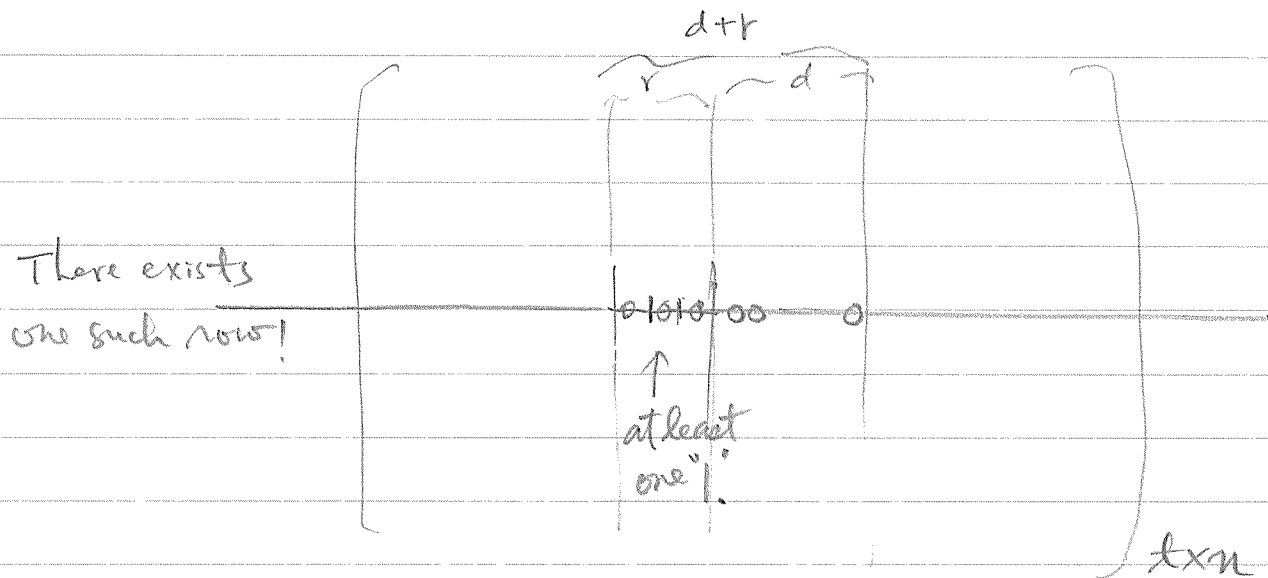
$$k(d, n) \leq 16d^2 (\log_3 n - \log_3 2 + 1).$$

$$k(d, n) \leq \underline{16d^2 (\log_3^2)} \underline{\log n (1+o(1))}.$$

(Non-constructive Method!)

With the help of designs, we can do better.

Definition A $t \times n$ binary matrix is (d, r) -disjunct if for $d+r$ arbitrary columns, the union of any r columns is not contained in the union of the other d columns.



D'yachkov and Rykov, A survey of superimposed code theory, Problems, Control Inform. Theory, 12 (1983), 1-13.

A $t \times n$ (d, r) -disjunct matrix is equivalent to the existence of a superimposed $(d, \textcircled{n}, r)$ -code of length $\textcircled{t}$.

Definition (Superimposed code) $\begin{matrix} t & n \\ \uparrow & \uparrow \\ & \nearrow \end{matrix}$ d (positives)

A superimposed code (binary) $SIC(n, N, m)$ consists of N

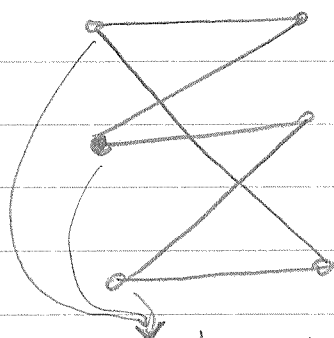
binary codewords of length n , with the property that from the Boolean sum of any m -subset we are able uniquely determine the individual codewords

Construction of (d, r) -disjunct matrices

Definition An (ℓ, d, ε) -dispenser is a bipartite graph $G = (A, B)$ satisfying the following conditions:

(1) $\forall S \subseteq A$ with $|S| \geq \ell$ is adjacent to at least $(1-\varepsilon)|B|$ vertices of B ; and
 $(|N(S)| \geq (1-\varepsilon)|B|)$

(2) $\forall v \in A, \deg_G(v) = d$.

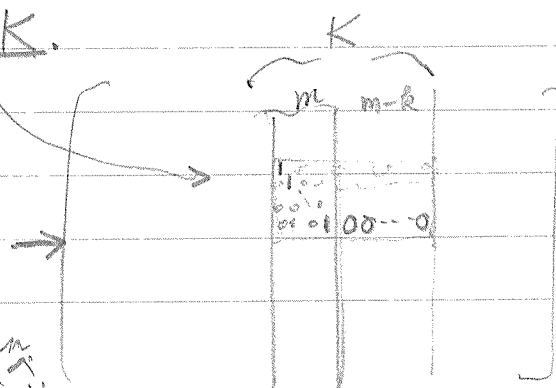


$(2, 2, \frac{1}{10})$ -dispenser

Review

Definition A (k, m, n) -selector is a binary matrix with n columns such that for any k -set K of columns, at least m of them are isolated in K .

A column is called isolated if there exists a row incident to it but not to any other column.



Expander Graph

An expander graph is a sparse graph (in edges) that has strong connectivity properties, quantified using edges, vertices or spectral expansion.

(1) Edge expansion (the most common version of expander graph)

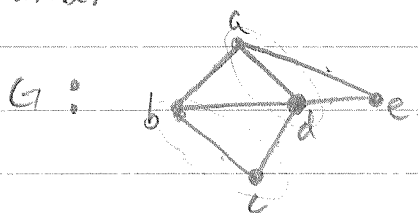
Let G be a connected graph with $|G| = n$. Let ∂S denote (edge boundary)

the set of edges $\langle S, V(G) \setminus S \rangle = \{ uv \mid u \in S \text{ and } v \in V(G) \setminus S \}, S \subseteq V(G),$

Then, the edge expansion, $h(G)$, of G is defined as

$$h(G) = \min_{0 < |S| \leq \frac{n}{2}} \frac{|\partial S|}{|S|}$$

Cheeger constant
Isoperimetric number
e.g.



$$S = \{d\}, \quad \frac{|\partial S|}{|S|} = 4$$

$$S = \{a, d\}, \quad |\partial S| = 5, \quad \frac{|\partial S|}{|S|} = 2.5$$

$$S = \{b, c\}, \quad |\partial S| = 3, \quad \frac{|\partial S|}{|S|} = 1.5$$

$$h(G) = 1.5$$

min.(?)

(2) Vertex expansion (Vertex isoperimetric number)

$$h_{\text{out}}(G) = \min_{0 < |S| \leq \frac{n}{2}} \frac{|\partial_{\text{out}}(S)|}{|S|}$$

outer boundary

$$N_G(S) = \partial_{\text{out}}(S)$$

$$S = \{b, c\}, \quad |\partial_{\text{out}}(S)| = 2, \quad h_{\text{out}}(G) = 1(?)$$

Fact 1 If G is a d -regular graph, then

$$h_{\text{out}}(G) \leq h(G) \leq d \cdot h_{\text{out}}(G).$$

$\partial_{\text{in}}(S)$: Inner boundary

↳ the set of vertices in S with at least one neighbor in $V(G) \setminus S$.

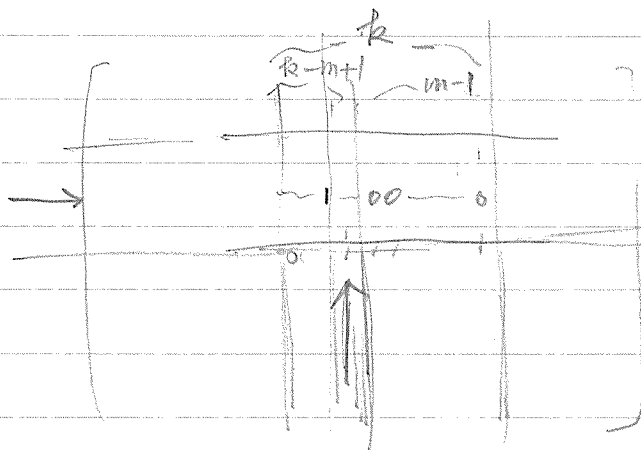
$$h_{\text{in}}(G) = \min_{0 < |S| \leq \frac{n}{2}} \frac{|\partial_{\text{in}}(S)|}{|S|}.$$

How about Petersen Graph?

Is $h(G) = 2$? $h_{\text{out}}(G) = 1$?

Fact A (k, m, n) -selector is $(m-1, k-m+1)$ -disjunct.

Proof.



The union of any $k-m+1$ columns is not contained in the union of $m-1$ other columns since the $n-k+1$ columns contain at least one column (isolated) in K .

(*) Note, if $k-m+1=1$, i.e. $k=m$, then the matrix is in fact a $(k-1)$ -disjunct matrix. Therefore, in any k columns, one of them is not contained in the union of the other $k-1$ columns.

So, consider $d=m-1$ and $r=k-m+1$. A (d, r) -disjunct matrix can be obtained by constructing a (k, m, n) -selector.

© B.S. Chlebus and D. R. Kowalski, Almost optimal explicit selector,

LNCS 3623 (Springer-Verlag, 2005) 270-280.

(*) Let M be an $s \times n$ (h, h, n) -selector $((h-1)$ -disjunct)

and $H = (A, B)$ a (k, d, ε) -dispenser.

$A = \mathbb{Z}_n$, $B = \mathbb{Z}_g' = \{0', 1', 2', \dots, g-1'\}$ and rows of M are indexed by $\mathbb{Z}_s$. Define an $sg \times n$ binary matrix M^* as follows:

$\forall v \in A$ and $i = as + b$ where $a \geq 0$ and $0 \leq b < s$, $M^*(i, v) = 1$

if and only if v is neighbor of vertex a' in B and $M(b, v) = 1$.

Then, M^* is a (k, h, n) -selector provided $(1-\varepsilon)gh > kd$.
(See figure in next page!)

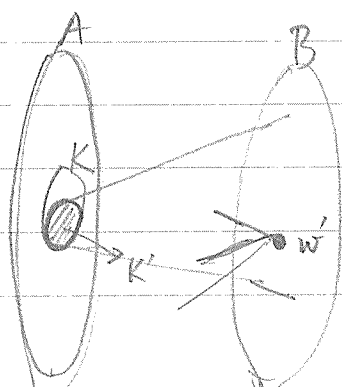
Proof. Suppose not. Then, there exists a k -set K of columns

and a $(k-h+1)$ -subset $K' \subseteq K$ such that every column in K' is

not isolated in K . (Since we need h isolated columns, we have

at most $k-h$ non-isolated columns in K .)

Now, consider the (k, d, ε) -dispenser H . K' (and K) can be "recognized" as a vertex subset of A . Let $w' \in B \cap N_H(K')$.

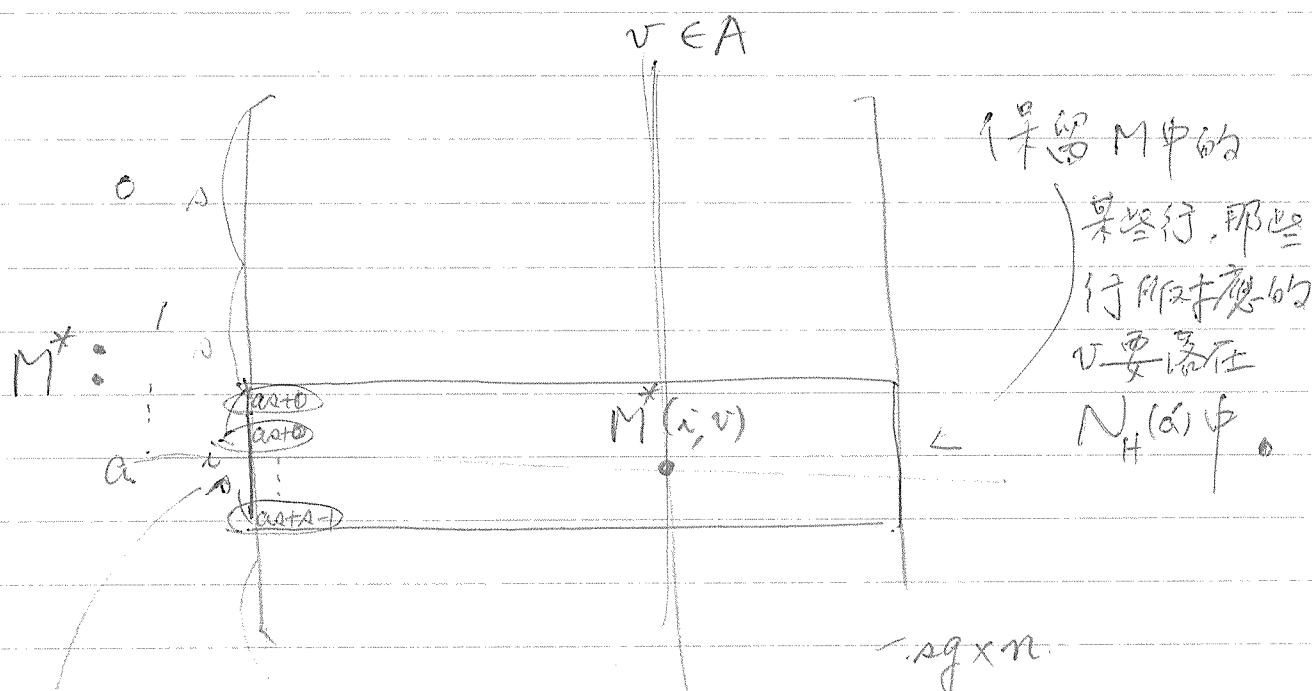
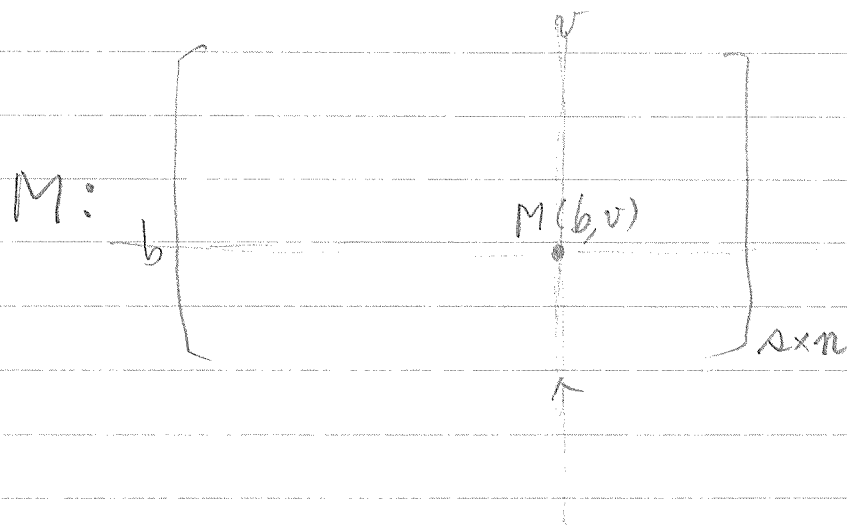


Claim:

w' has more than h neighbors in K .

Suppose at most h .

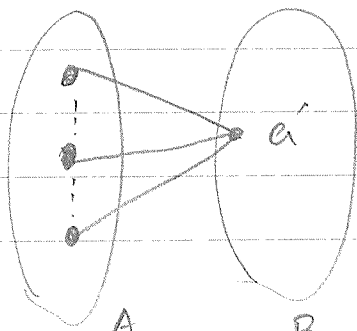
Let $N_K(w') = \{v \in K, v \sim_H w'\}$.



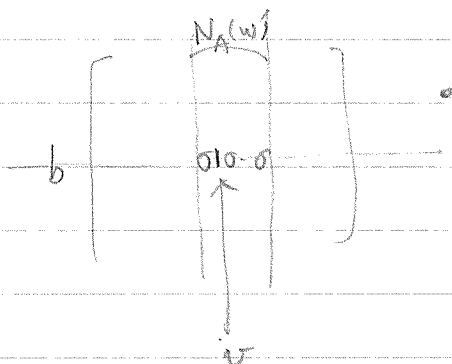
$$i = as + b, \quad a \in \mathbb{Z}_q \text{ and } 0 \leq b < s.$$

$$M^*(i, v) = 1 \text{ iff } M(b, v) = 1 \text{ and } a' \in B \text{ (corresponding to } a \text{).}$$

$v \in N_H(a')$ where



Since M is an (h, h, n) -selector, there exists a ^{row} a, b in M such that



Now, the $(ws+b)$ -th row of M^* intersects K with exactly one column (an isolated column), $\rightarrow \leftarrow K'$.

Since H is a (R, d, ε) -dispenser, $|N_H(K')| \geq (1-\varepsilon)|V_2| = (1-\varepsilon)g$.

Thus, $|KH, N_H(K')| > (1-\varepsilon)gh$. However, $\forall v \in A$, $\deg_H(v) = d$,

we have at most dR edges which is less than $(1-\varepsilon)gh$ edges,

a contradiction. ■

~

More properties of a (d, r) -disjunct matrix.

① In a d -disjunct matrix, $t(d, 1, n) \geq \log_2 \binom{n}{d} - \log_2 \binom{d+1}{d}$.

Proof.

