An Investigation of $d$-Separable, $\bar{d}$-Separable, and $d$-Disjunct Binary Matrices

Greg Bálint$^1$, Melanie Bloch$^2$, Robert Ellis$^1$, Gregory Monson$^3$, Abraham Schulte$^4$, Allie Schultz$^5$, Micky Soule$^6$, and David Vaden$^7$

$^1$Illinois Institute of Technology  
$^2$University of Colorado at Denver  
$^3$University of California at Berkeley  
$^4$Northwestern University  
$^5$University of Notre Dame  
$^6$San Diego State University  
$^7$Southwestern University

August 9, 2013
## Contents

1 Properties of $d$-separable, $\bar{d}$-separable, and $d$-disjunct binary matrices 2  
1.1 Introduction ......................................................... 2  
1.2 Separable binary matrices ........................................... 4  
  1.2.1 Restrictions on column weights for $\bar{d}$-separable binary matrices ... 5  
  1.2.2 Restrictions on dimensions for $\bar{d}$-separable binary matrices ....... 7  
1.3 Disjunct binary matrices ............................................. 8  
  1.3.1 Restrictions on column weights for $d$-disjunct binary matrices ...... 9  
  1.3.2 Restrictions on dimensions for $d$-disjunct binary matrices ......... 11  
1.4 Counting collisions and covers .................................... 15  
  1.4.1 $s$-collision and $s$-cover matrices .................................... 15  
  1.4.2 Row weight inequalities ........................................... 17  
  1.4.3 A lower bound on the number of collisions ............................ 18  
  1.4.4 Using a generalization of Sperner’s theorem to count covers ............ 19  
  1.4.5 Applying outside problems to disjunctness .......................... 20  
1.5 Existence of a $10 \times 14$ $2$-disjunct binary matrix ................ 21  
1.6 A conjecture on $d$-disjunct matrices with optimal dimensions .......... 23  
1.7 Other items of interest .............................................. 24  

2 Constructing efficient decodable pooling matrices 28  
2.1 On $d$-separable binary matrices with time optimal analysis algorithms ... 28  
  2.1.1 The Construction ................................................... 29  
  2.1.2 Determining the number of defectives .................................. 30  
  2.1.3 Recovering the Defective Values .................................... 30  
  2.1.4 Runtime Analysis .................................................. 32  
  2.1.5 Comparison with other Matrices .................................... 33  
2.2 A new construction of $d$-disjunct matrices with $K (d + 1) q^t$ rows .... 33  
2.3 Fast Decoding Using Reed Solomon Matrices ............................. 38  
2.4 Appendix ............................................................... 41
Chapter 1

Properties of $d$-separable, $\bar{d}$-separable, and $d$-disjunct binary matrices

1.1 Introduction

**Notation.** We denote the rows of a $t \times n$ matrix $A$ as $R_1, \ldots, R_t$ or as $R(1), \ldots, R(t)$. We denote the $j^{th}$ entry of row vector $R_m$ as $R_m[j]$. Similarly, we denote the columns of $A$ as $C_1, \ldots, C_n$ or as $C(1), \ldots, C(n)$. We denote the $i^{th}$ entry of column vector $C_m$ as $C_m[i]$. Thus, we denote the entry of $A$ which lies in the $i^{th}$ row and $j^{th}$ column as $A_{i,j}$, $R_i[j]$, or $C_j[i]$.

**Definition.** We call an entry of $a$ in a matrix an $a$-entry. We call the sum of entries in row $R$ the **row weight** of $R$. We denote the weight of $R$ as $|R|$. We define and denote the **column weight** of a column similarly. For a binary vector $v$, $|v|$ is the number of 1-entries in $v$.

**Definition.** We call a binary row vector (binary column vector) $v$ a zero-row (zero-column) if $|v| = 0$. We call $v$ a full-row (full-column) if $|v|$ is equal to the length of $v$. That is, $v$ is a zero-row (zero-column) if and only if $v$ has only 0-entries and $v$ is a full-row (full-column) if and only if $v$ has only 1-entries.

**Definition.** We denote the Boolean sum of binary vectors $v_1$ and $v_2$ as $v_1 \oplus v_2$. We call the Boolean sum of $s$ vectors a $s$-**sum** or a **sum**. We denote the Boolean product of binary vectors $v_1$ and $v_2$ as $v_1 \otimes v_2$. We call the Boolean product of $s$ vectors a $s$-**product** or **product**. We say that a column vector $C_i$ is contained in a $s$-sum $C_1 \oplus \cdots \oplus C_s$ or $s$-product $C_1 \otimes \cdots \otimes C_s$ if $i$ is in the set of indices $\{1, \ldots, s\}$.
**Definition.** Let \( C \) be a \( s \)-sum of vectors. We say that \( C_k \) is a sub-sum or a \( k \)-sub-sum of \( C \) if \( C_k \) is a \( k \)-sum of vectors contained in \( C \), where \( k \leq s \). We define sub-product or \( k \)-sub-product similarly.

**Notation.** Let \( C \) be a sum (product) of rows or columns of a binary matrix \( A \). We denote the number of columns contained in \( C \) as \( \mu(C) \). Thus, \( C \) is a \( \mu(C) \)-sum (\( \mu(C) \)-product) of the rows or columns of \( A \).

**Definition.** We say that two rows (columns) of a matrix \( A \) are distinct if they are not the same row (column). Thus, two rows (columns) can be both equal and distinct. We call two sums (products) of the rows or columns of \( A \) distinct if each sum (product) contains at least one row or column, accordingly, not contained in the other.

**Notation.** We denote the maximum and minimum weights of any row of a matrix \( A \) as \( P(A) \) and \( \rho(A) \), respectively. We denote the maximum and minimum weights of any column of \( A \) as \( \Gamma(A) \) and \( \gamma(A) \), respectively. When it is clear, we may omit \( (A) \) and use \( P, \rho, \Gamma, \gamma \).

**Notation.** We denote the maximum and minimum weights of any \( s \)-sum of rows in a matrix \( A \) as \( P_s^\oplus(A) \) and \( \rho_s^\oplus(A) \), respectively. We denote the maximum and minimum weights of any \( s \)-sum of columns in \( A \) as \( \Gamma_s^\oplus(A) \) and \( \gamma_s^\oplus(A) \), respectively. Similarly, we denote the maximum and minimum weights of any \( s \)-product of rows and columns of \( A \) as \( P_s^\otimes(A), \rho_s^\otimes(A), \Gamma_s^\otimes(A), \gamma_s^\otimes(A) \), respectively. When it is clear, we may omit the \( (A) \) and use \( P_s^\oplus, \rho_s^\oplus, \Gamma_s^\oplus, \gamma_s^\oplus, P_s^\otimes, \rho_s^\otimes, \Gamma_s^\otimes, \gamma_s^\otimes \).

**Lemma 1.1.1.** For any binary matrix

\[
\begin{align*}
\Gamma_s^\otimes &= t \implies P_t^\otimes \geq s \\
P_s^\otimes &= t \implies \Gamma_t^\otimes \geq s \\
\gamma_s^\otimes &= t \implies \rho_t^\otimes \geq s \\
\rho_s^\otimes &= t \implies \gamma_t^\otimes \geq s
\end{align*}
\]

**Proof.** We prove the first implication. The proofs for the other implications are similar.

Let \( A \) be a binary matrix such that \( \Gamma_s^\otimes = t \). Then for some \( s \) columns of \( A \), there are associated \( t \) rows \( R \) such that every column of \( C \) has a 1-entry in every row of \( R \). Thus, each of the \( t \) rows of \( R \) has a 1-entry in each column of \( C \). So their \( t \)-product is at least \( s \). Thus, \( P_t^\otimes \geq s \).

**Lemma 1.1.2.** Let \( A \) be a \( t \times n \) binary matrix. Let \( \gamma \) be the minimum column weight of \( A \). Let \( m_s \) be given recursively as:

\[
m_0 = \left\lceil \frac{|A|}{t} \right\rceil, m_s = \left\lceil \frac{(\gamma - s)m_{s-1}}{t - s} \right\rceil, 1 \leq s \leq \gamma
\]

(1.1)
If \( \sigma \leq m_s \), then \( \Gamma^\otimes_\sigma > s \), where \( \Gamma^\otimes_\sigma \) is the maximum weight of the \( \sigma \)-products of the columns of \( A \).

Notice that increasing \( \gamma \) does not decrease \( m_s \) for any \( 1 \leq s \leq \gamma \).

**Proof.** We prove by induction.

Base case. By the pigeonhole principle, at least one row \( R \) has at least \( \left\lceil \frac{|A|}{t} \right\rceil = m_0 \) 1-entries. Let \( C_0 \) be the set of columns with a 1-entry in \( R \). Notice that there are at least \( m_0 \) columns in \( C_0 \), so \( | \otimes C_0 | > 0 \). Thus, \( \Gamma^\otimes_{m_0}(A) > 0 \).

Inductive case. Assume \( \Gamma^\otimes_{m_{s-1}}(A) \geq s \). Then there exists a \( m_{s-1} \)-product \( C_{s-1} \) such that \( |C_{s-1}| \geq s \). Since \( \gamma(A) < t \), \( m_s \leq m_{s-1} \). Thus, for any \( m_s \)-product \( C_s \) of the columns contained in \( C_{s-1} \), we have \( | \otimes C_s | \geq s \). Let \( A' \) be the submatrix of \( A \) formed by deleting the \( s \) rows where this \( m_s \)-product contains a 1-entry and keeping only the columns of \( C_{s-1} \). Notice that \( A' \) is a \((t-s) \times m_{s-1}\) matrix. Since each column of \( A \) has weight at least \( \gamma(A) \), \( \gamma(A') \geq \gamma(A) - s \). Thus, there are at least \( (\gamma(A) - s)m_{s-1} \) total 1-entries in \( A' \). Thus, by the pigeonhole principle, at least one row of \( A' \) has at least \( \left\lceil \frac{(\gamma(A)-s)m_{s-1}}{t-s} \right\rceil \) 1-entries. This row corresponds with a row in \( A \) which must have \( \frac{(\gamma(A)-s)m_{s-1}}{t-s} \) 1-entries in the original \( m_s \) columns of \( C_s \). Since \( A \) contains \( s \) rows not contained in \( A' \), each with 1-entries in the columns of \( C_s \), it follows that the \( m_s \)-product of the columns contained in \( C_s \) has weight greater than \( s \). Thus, \( \Gamma^\otimes_{m_s}(A) > s \). \( \square \)

**Lemma 1.1.3.** Let \( A \) be a \( t \times n \) binary matrix. Let \( \rho \) be the minimum row weight of \( A \). Let \( m_s \) be given recursively as:

\[
m_0 = \left\lceil \frac{|A|}{n} \right\rceil, \quad m_s = \left\lceil \frac{(\rho - s)m_{s-1}}{n - s} \right\rceil, \quad 1 \leq s \leq \rho
\]

If \( \sigma \leq m_s \), then \( P^\otimes_{\sigma} > s \), where \( P^\otimes_{\sigma} \) is the maximum weight of the \( \sigma \)-products of rows of \( A \).

Notice that increasing \( \rho \) does not decrease \( m_s \) for any \( 1 \leq s \leq \rho \).

**Proof.** The proof is similar to the proof of Lemma 1.1.2. \( \square \)

### 1.2 Separable binary matrices

The definitions and claims for separability and disjunctness can be found in [1]

**Definition.** Let \( C_1 \) and \( C_2 \) be distinct column sums of a binary matrix. If \( C_1 = C_2 \), we say the unordered set \( \{C_1, C_2\} \) forms a collision. If \( \mu(C_1) = \mu(C_2) = s \), we say \( \{C_1, C_2\} \) is a \( s \)-collision.

**Definition.** Let \( A \) be a binary matrix. Let \( d \) be a natural number. We say that \( A \) is \( d \)-separable if \( A \) has no \( d \)-collisions. We say that \( A \) is \( \overline{d} \)-separable if no collision \( \{C_1, C_2\} \) in \( A \) is such that \( \mu(C_1), \mu(C_2) \leq d \). Thus, a matrix is \( \overline{d} \)-separable if any distinct \( s \)-sum,
s₂-sum are not equal, where s₁, s₂ ≤ d. Notice that a binary matrix that is \( \overline{d} \)-separable is \( d \)-separable.

**Claim.** The binary matrix that results from deleting a column from a \( \overline{d} \)-separable binary matrix is \( \overline{d} \)-separable. The binary matrix that results from adding a column to a binary matrix that is not \( \overline{d} \)-separable is not \( \overline{d} \)-separable. Deleting or adding a zero-row or a full-row to a binary matrix preserves separability.

**Claim.** If a binary matrix \( A \) is \( \overline{d} \)-separable, then \( A \) is \( \overline{s} \)-separable for any natural number \( s \leq d \).

**Lemma 1.2.1.** Let \( A \) be a \( t \times n \) binary matrix. Let \( R \) be a row of \( A \). Let \( A' \) be the \((t - 1) \times (n - |R|)\) submatrix of \( A \) that results from deleting row \( R \) and the columns of \( A \) with 1-entries in \( R \). If \( A \) is \( \overline{d} \)-separable, then \( A' \) is \( \overline{d} \)-separable.

**Proof.** Let \( A \) be \( \overline{d} \)-separable. Notice that after deleting the columns of \( A \) where \( R \) has 1-entries, \( R \) will be a zero-row. Thus, after deleting that zero-row, the resulting submatrix, \( A' \), will be \( \overline{d} \)-separable.

**Lemma 1.2.2.** Let \( A \) be a \( t \times n \) binary matrix. Let \( s < d \) be a natural number. Let \( \mathcal{C} \) be a \( s \)-sum of columns of \( A \). Let \( A' \) be the \((t - |\mathcal{C}|) \times (n - s)\) submatrix of \( A \) that results from deleting the rows of \( A \) in which \( \mathcal{C} \) has a 1-entry and the columns contained in \( \mathcal{C} \). If \( A \) is \( \overline{d} \)-separable, then \( A' \) is \( \overline{d} - s \)-separable.

**Proof.** Let \( A \) be \( \overline{d} \)-separable. Let \( \mathcal{C}_0' \) denote a sum of columns in \( A' \) which corresponds with a sum \( \mathcal{C}_0 \) of columns in \( A \) and vice versa. Assume, by way of contradiction, that \( A' \) is not \( \overline{d} - s \)-separable. Then there exist distinct sums \( \mathcal{C}_1', \mathcal{C}_2' \) in \( A' \), where \( \mu(\mathcal{C}_1'), \mu(\mathcal{C}_2') \leq d - s \), such that \( \mathcal{C}_1' = \mathcal{C}_2' \). Notice \( \mathcal{C}_1' \oplus \mathcal{C} = \mathcal{C}_2' \oplus \mathcal{C} \). Since \( \mathcal{C}_1', \mathcal{C}_2' \) are distinct, \( \mathcal{C}_1, \mathcal{C}_2 \) are distinct. Thus, \( \{\mathcal{C}_1 \oplus \mathcal{C}, \mathcal{C}_2 \oplus \mathcal{C}\} \) forms a collision in \( A \), a contradiction, since \( \mu(\mathcal{C}_1 \oplus \mathcal{C}), \mu(\mathcal{C}_2 \oplus \mathcal{C}) \leq d \) and \( A \) is \( \overline{d} \)-separable. Thus \( A' \) is \( \overline{d} - s \)-separable.

**Definition.** Let \( A \) be a \( \overline{d} \)-separable binary matrix. We call \( A \) \( \overline{d} \)-reducible if there is some submatrix of \( A \) that results from deleting a row and a column of \( A \) that is \( \overline{d} \)-separable. We call \( A \) \( \overline{d} \)-irreducible if there is no such submatrix.

### 1.2.1 Restrictions on column weights for \( \overline{d} \)-separable binary matrices

**Theorem 1.2.1.** Let \( A \) be a \( t \times n \) binary matrix. Let \( s < d \) be a natural number. If \( A \) is \( \overline{d} \)-separable, then

\[
\Gamma_s^\oplus \leq t - \log_2 \left( \sum_{i=0}^{d-s} \binom{n-s}{i} \right)
\]
Proof. Let $A$ be $d$-separable. Let $C$ be a $s$-sum of the columns of $A$. Let $\mathcal{R}$ be the set of $t - |\mathcal{C}|$ rows with 0-entries in $\mathcal{C}$. Notice that there are $\sum_{i=0}^{d-s} \binom{n-s}{i}$ ways to choose columns for 0-sums, 1-sums, \ldots, $(d-s)$-sums from the $n-s$ columns not contained in $\mathcal{C}$. Notice there are $2^{t-w}$ ways to choose the entries for $\mathcal{R}$ of the Boolean sums. Since $A$ is $d$-separable, these Boolean sums are unique. Thus, $2^{t-|\mathcal{C}|} \geq \sum_{i=0}^{d-s} \binom{n-s}{i}\\nRightarrow t - |\mathcal{C}| \geq \log_2 \left( \sum_{i=0}^{d-s} \binom{n-s}{i} \right)\\nRightarrow |\mathcal{C}| \leq t - \log_2 \left( \sum_{i=0}^{d-s} \binom{n-s}{i} \right) \quad \Box$

Lemma 1.2.3. Let $A$ be a $t \times n$ $d$-separable binary matrix. Let $s < d$ be a natural number. Let $\mathcal{C}$ be a $s$-sum of the columns of $A$. Let $\mathcal{C}_{s-1}$ be a $(s-1)$-sum of some $s-1$ columns contained in $\mathcal{C}$. Let $C$ be the column contained in $\mathcal{C}$ not contained in $\mathcal{C}_{s-1}$. If $A$ is $d$-irreducible, then $|\mathcal{C}| > |\mathcal{C}_{s-1}| + d - s$. That is, if $A$ is $d$-irreducible, then $C$ has more than $d-s$ 1-entries such that $\mathcal{C}_{s-1}$ has 0-entries in the rows of those 1-entries.

Proof. Let $A$ be $d$-irreducible. We prove by contradiction.

Suppose $|\mathcal{C}| \leq |\mathcal{C}_{s-1}| + d - s$. Then $C$ has at most $d-s$ 1-entries such that $\mathcal{C}_{s-1}$ has 0-entries in the rows of those 1-entries. Let $\mathcal{R}$ be the collection of rows with those 1-entries. Assume, by way of contradiction, that each row of $\mathcal{R}$ has a 1-entry in some column not contained in $\mathcal{C}$. Then there is a collection of at most $d-s$ columns not contained in $\mathcal{C}$ whose sum $\mathcal{C}_0$ has a 1-entry in each row of $\mathcal{R}$. Notice that $\{\mathcal{C}_0 \oplus \mathcal{C}, \mathcal{C}_0 \oplus \mathcal{C}_{s-1}\}$ forms a collision, a contradiction, since $\mu(\mathcal{C}_0 \oplus \mathcal{C}), \mu(\mathcal{C}_0 \oplus \mathcal{C}_{s-1}) \leq d$ and $A$ is $d$-separable. Thus, there is some row $R$ in $\mathcal{R}$ such that the columns contained in $\mathcal{C}$ have the only 1-entry in $R$. Thus, $C$ is the only column to have a 1-entry in $R$. Notice that the $(t-1) \times (n-1)$ submatrix of $A$ that results from deleting $C$ and the resulting 0-row $R$ is $d$-separable, a contradiction, since $A$ is $d$-irreducible. Thus, $|\mathcal{C}| > |\mathcal{C}_{s-1}| + d - s$ \quad \Box

Theorem 1.2.2. Let $A$ be a $t \times n$ $d$-separable binary matrix. Let $s < d$ be a natural number. If $A$ is $d$-irreducible, then

$$\gamma_s^\oplus \geq sd - \frac{s(s-1)}{2}$$

Proof. Let $A$ be $d$-irreducible. We prove by induction.

Base case. We prove $\gamma \geq d$. Let $\mathcal{C}$ be a 1-sum of the columns of $A$. Notice that $\mathcal{C}$ is equal to a column of $A$. By Lemma 1.2.3, $\mathcal{C}$ has at least $d$ 1-entries such that any 0-sum from the columns contained in $\mathcal{C}$ has 0-entries in the rows of those 1-entries. Since any 0-sum of $A$ has only 0-entries, $\mathcal{C}$ has at least $d$ 1-entries. Thus, the columns of $A$ each have at least $d$ 1-entries.
Inductive case. Assume $\gamma^\oplus_{s-1} \geq (s-1)d - \frac{(s-1)(s-2)}{2}$, where $s < d$. We prove $\gamma^\oplus_s \geq sd - \frac{s(s-1)}{2}$.

Notice $\gamma^\oplus_{s-1} \geq (s-1)d - \sum_{i=1}^{s-2} i$. Let $\mathcal{C}$ be a $s$-sum of the columns of $A$. Let $\mathcal{C}_{s-1}$ be the $(s-1)$-sum of some $(s-1)$ columns contained in $\mathcal{C}$. Since $A$ is $d$-irreducible, by Lemma 1.2.3, $|\mathcal{C}| \geq |\mathcal{C}_{s-1}| + d - s + 1$. Thus, $|\mathcal{C}| \geq (s-1)d - \sum_{i=1}^{s-2} i + d - (s-1) = sd - \sum_{i=1}^{s-1} i = sd - \frac{s(s-1)}{2}$.

\[ |\mathcal{C}| \geq (s-1)d - \sum_{i=1}^{s-2} i + d - (s-1) = \gamma^\oplus_{s-1} + d - s + 1. \]

### 1.2.2 Restrictions on dimensions for $d$-separable binary matrices

**Theorem 1.2.3.** Let $A$ be a $t \times n$ binary matrix. If $A$ is $d$-separable, then

$$ t \geq \log_2 \left( \sum_{i=0}^{d} \binom{n}{i} \right) $$

**Proof.** Let $A$ be $d$-separable. Notice that there are $\sum_{i=0}^{d} \binom{n}{i}$ ways to choose columns contained in distinct 0-sums, 1-sums, \ldots, $d$-sums. Notice that there are at most $2^t$ ways to choose the entries of the Boolean sums. Since $A$ is $d$-separable, these sums are unique. Thus, $\sum_{i=0}^{d} \binom{n}{i} \leq 2^t \implies t \geq \log_2 \left( \sum_{i=0}^{d} \binom{n}{i} \right)$. \qed

**Theorem 1.2.4.** Let $A$ be a $t \times n$ $d$-separable binary matrix. Let $s < d$ be a natural number. If $A$ is $d$-irreducible, then

$$ t \geq \log_2 \left( \sum_{i=0}^{d-s} \binom{n-s}{i} \right) + sd - \frac{s(s-1)}{2} $$

**Proof.** Let $A$ be $d$-irreducible. By Theorems 1.2.1 and 1.2.2

$$ sd - \frac{s(s-1)}{2} \leq \gamma^\oplus_s \leq \Gamma^\oplus_s \leq t - \log_2 \left( \sum_{i=0}^{d-s} \binom{n-s}{i} \right) $$

$$ \implies sd - \frac{s(s-1)}{2} \leq t - \log_2 \left( \sum_{i=0}^{d-s} \binom{n-s}{i} \right) $$

$$ \implies t \geq \log_2 \left( \sum_{i=0}^{d-s} \binom{n-s}{i} \right) + sd - \frac{s(s-1)}{2} \quad \Box $$

**Corollary 1.2.1.** Let $A$ be a $t \times (t+1)$ $d$-separable binary matrix. Let $s < d$ be a natural number. If $A$ is $d$-irreducible, then

$$ t \geq \log_2 \left( \sum_{i=0}^{d-s} \binom{t+1-s}{i} \right) + sd - \frac{s(s-1)}{2} $$

Tables outlining the results of this corollary may be found in the appendix.
1.3 Disjunct binary matrices

Definition. Let $C_1, C_2$ be distinct sums of columns of a binary matrix $A$. We say that $C_2$ covers $C_1$ if $C_2 = C_1 \oplus C_2$. We use the notation $C_1 \subseteq C_2$. If $\mu(C_1) = \mu(C_2) = s$ and $C_1 \subseteq C_2$ then we say the ordered set $\{C_1, C_2\}$ is a $s$-cover in $A$. Notice that a collision $\{C_1, C_2\}$ corresponds with two covers: $\{C_1, C_2\}$ and $\{C_2, C_1\}$.

Definition. We say that a matrix $A$ is $d$-disjunct if $A$ has no $d$-covers.

Claim. If $A$ is a $d$-disjunct binary matrix, then $A$ is $s$-disjunct for any natural number $s \leq d$.

Definition. Let $A$ be a $d$-disjunct binary matrix. We call $A$ $d$-reducible if there is some submatrix of $A$ that results from deleting a row and a column of $A$ that is $d$-disjunct. We call $A$ $d$-irreducible if there is no such submatrix.

Claim. The binary matrix that results from deleting a column from a $d$-disjunct binary matrix is $d$-disjunct. The binary matrix that results from adding a column to a binary matrix that is not $d$-disjunct is not $d$-disjunct. Deleting or adding a zero-row or a full-row to a binary matrix preserves disjunctness. A matrix with a zero-column or a full-column is not $d$-disjunct for any natural number $d$. A binary matrix with two identical columns $d$-disjunct for any natural number $d$.

Claim. Any binary matrix that is $d$-disjunct is $\overline{d}$-separable. The binary matrix that results from deleting a row of a $d$-disjunct matrix is $\overline{d}$-separable. A binary matrix that is $\overline{d}$-separable is $(d - 1)$-disjunct.

Claim. If there is no $t \times n$ $d$-disjunct binary matrix, then there is no $(t - 1) \times (n - 1)$ $d$-disjunct binary matrix.

Lemma 1.3.1. Let $A$ be a $t \times n$ binary matrix. Let $R$ be a row of $A$. Let $A'$ be the $(t - 1) \times (n - |R|)$ submatrix of $A$ that results from deleting row $R$ and the columns of $A$ with 1-entries in $R$. If $A$ is $d$-disjunct, then $A'$ is $d$-disjunct.

Proof. The proof is analogous to the proof of Lemma 1.2.1.

The following definition, Sperner’s Theorem, and the LYM Inequality can be found in [6].

Definition. A Sperner Family is a collection of subsets of a set such that no subset is contained in any other subset.

Sperner’s Theorem. Let $\mathcal{K}$ be a set of $k$ elements. Suppose $\mathcal{S}$ is a Sperner Family of $\mathcal{K}$. Then $\mathcal{S}$ contains at most $\left(\binom{k}{\lfloor \frac{k}{2} \rfloor}\right)$ subsets of $\mathcal{K}$.
The LYM Inequality. If $\mathcal{F}$ is a Sperner family of a set of size $t$, then

$$
\sum_{X \in \mathcal{F}} \frac{1}{|X|} \leq 1
$$

Definition. Let $C$ be a binary vector. The subset version of $C$ is the set of indices where $C$ has a 1-entry.

Definition. Let $A$ be a $t \times n$ matrix. Let $A_d^\oplus$ be the $t \times \left(\binom{n}{d}\right)$ matrix whose columns are the distinct $d$-sums of columns of $A$. We call $A_d^\oplus$ the $d$-sum-matrix of $A$.

Lemma 1.3.2. A binary matrix $A$ is $d$-disjunct if and only if $A_d^\oplus$ is 1-disjunct.

Proof. Suppose $A$ is $d$-disjunct. If any $C_1$ column in $A_d^\oplus$ covers another column $C_2$ in $A_d^\oplus$, then the $d$-sum $C_1$ of columns of $A$ corresponding to $C_1$ in $A_d^\oplus$ must cover the $d$-sum $C_2$ of columns of $A$ corresponding to $C_2$ in $A_d^\oplus$, a contradiction, since $A$ is $d$-disjunct. Thus, $A_d^\oplus$ must be 1-disjunct.

Suppose $A$ is not $d$-disjunct. Then there is some $d$-sum $C_1$ of $A$ which covers another $d$-sum $C_2$ of $A$. Thus, there is a column in $A_d^\oplus$ which covers another column in $A_d^\oplus$. Thus, $A_d^\oplus$ is not 1-disjunct.

1.3.1 Restrictions on column weights for $d$-disjunct binary matrices

Notation. Let $n, d$ be natural numbers, where $n \geq d$. We define the function $f(n, d) = m$, where $m$ is the largest integer satisfying $\left(\frac{m}{2}\right) < \left(\frac{n}{d}\right)$.

Theorem 1.3.1. Let $A$ be a $t \times n$ binary matrix. If $A$ is $d$-disjunct, then for any natural number $s < d$,

$$
\Gamma_s^\oplus < t - f(n - s, d - s).
$$

where $\Gamma_s^\oplus$ is the maximum weight of the $s$-sums of columns of $A$.

Proof. Let $A$ be $d$-disjunct. We prove by contradiction

Suppose $\Gamma_s^\oplus \geq t - f(n - s, d - s)$. There is a $s$-sum $\mathcal{C}_0$ of columns such that $|\mathcal{C}_0| = \Gamma_s^\oplus$. Let $k$ be the number of 0-entries in $\mathcal{C}_0$. Notice that $k \leq f(n - s, d - s)$. Let $A'$ be the submatrix of $A$ that results from deleting all rows in which $\mathcal{C}_0$ has a 1-entry and each column contained in $\mathcal{C}_0$. Notice that $A'$ is a $k \times (n - s)$ binary matrix. We denote the column of $A'$ that results from a column $C_i$ of $A$ as $C_i'$ and the sum of columns of $A'$ that results from a sum of columns $\mathcal{C}$ in $A$ as $\mathcal{C}'$. Notice that $\left(\frac{k}{\binom{n}{s}}\right) \leq \left(\frac{f(n-s,d-s)}{\binom{n-s}{d-s}}\right) < \left(\frac{n-s}{d-s}\right)$.

So by Theorem 1.3.4, $A'$ is not $(d - s)$-disjunct. Thus, there is some $(d - s)$-sum $\mathcal{C}_1'$ in $A'$ which covers a column $C'$ in $A'$. The $d$-sum $\mathcal{C}_0 \oplus \mathcal{C}_1$ in $A$ covers $C$ in $A$, a contradiction, since $A$ is $d$-disjunct.

9
Theorem 1.3.2. Let $A$ be a $t \times n$ $d$-disjunct binary matrix. Let $\gamma$ be the minimum column weight in $A$. If $A$ is $d$-irreducible, then

$$\gamma > d$$

where $\gamma$ is the minimum weight of the columns of $A$.

Proof. Let $A$ be $d$-irreducible. We prove by contradiction.

Suppose $\gamma \leq d$. There is a column $C_j$ with weight $\gamma$. Let $\mathcal{R}$ be the collection of rows where $C_j$ has 1-entries. Assume, by way of contradiction, that each row in $\mathcal{R}$ has a 1-entry in some column other than $C_j$. Then there is a $s$-cover of $C_j$, where $s \leq \gamma \leq d$, a contradiction, since $A$ is $d$-disjunct. Thus, there is some row $R_i$ in $\mathcal{R}$ such that $R_i[j]$ is the only 1-entry of $R_i$. The submatrix of $A$ that results from deleting $C_j$ and then deleting the resulting 0-row $R_i$ is $d$-disjunct, a contradiction, since $A$ is $d$-irreducible. Thus $\gamma > d$. \qed

Lemma 1.3.3. Let $d \geq 2$ be a natural number. Let $A$ be a $t \times n$ $d$-disjunct binary matrix. Let $C_j, C_k$ be two columns of $A$. If $A$ is $d$-irreducible, then $|C_j \oplus C_k| \geq |C_j| + d$ and $|C_j \oplus C_k| \leq |C_k| - d$.

Proof. Let $A$ be $d$-irreducible. We prove by contradiction.

Suppose $|C_j \oplus C_k| < |C_j| + d$. Then by the inclusion-exclusion principle, $|C_j \oplus C_k| > |C_k| - d$. Notice that $|C_j \oplus C_k| \leq |C_k| - 2$, since $A$ is $d$-disjunct. Thus, there are at least 2 and at most $|C_j| - |C_k| + 2d - 2$ rows such that one but not both of the columns have a 1-entry in that row. Notice that there are less than $|C_j| - |C_k| + d$ rows such that the 1-entry is in $C_j$, and at most $d - 1$ rows such that the 1-entry is in $C_k$. Let $\mathcal{R}_A$ be the set of rows such that there is a 1-entry in $C_k$ but not $C_j$. Suppose for each row of $\mathcal{R}_A$, there is a column, not $C_k$, such that there is a 1-entry in that column for that row. Let $\mathcal{C}_A$ be the sum of these columns. Notice that $1 \leq \mu(\mathcal{C}_A) < d$. Notice that the $d$-sum of $\mathcal{C}_A$ with $C_j$ covers $C_k$. But this is a contradiction, since $A$ is $d$-disjunct. Thus, for some row $R_i$, $R_i[k]$ is the only 1-entry of $R_i$. Let $A'$ be the submatrix of $A$ that results from deleting $C_k, R_i$. Since after deleting $C_k, R_i$ will be a 0-row, $A'$ is $d$-disjunct, a contradiction, since $A$ is irreducible. Thus, $|C_j \oplus C_k| \geq |C_j| + d$. By applying the inclusion-exclusion principle, $|C_j \oplus C_k| \leq |C_k| - d$. \qed

Lemma 1.3.4. Let $A$ be a $t \times n$ $d$-disjunct binary matrix. Let $s < d$ be a natural number. Let $\mathcal{C}$ be a $s$-sum of the columns of $A$. Let $\mathcal{C}_{s-1}$ be a $(s-1)$-sum of some $s-1$ columns contained in $\mathcal{C}$. Let $C$ be the column contained in $\mathcal{C}$ not contained in $\mathcal{C}_{s-1}$. If $A$ is $d$-irreducible, then $|\mathcal{C}| > |\mathcal{C}_{s-1}| + d - s$. That is, if $A$ is $d$-irreducible, then $C$ has more than $d - s$ 1-entries such that $\mathcal{C}_{s-1}$ has 0-entries in the rows of those 1-entries.

Proof. Since $A$ is $d$-disjunct, $A$ is $d$-separable. Thus, the results follow from Lemma 1.2.3. \qed
**Theorem 1.3.3.** Let $A$ be a $t \times n$ $d$-disjunct binary matrix. Let $1 < s < d$ be a natural number. If $A$ is $d$-irreducible, then

$$
\gamma_s^{\oplus} \geq sd - \frac{s(s-1)}{2} + 2
$$

where $\gamma_s^{\oplus}$ is the minimum weight of the $s$-sums of columns of $A$.

**Proof.** Let $A$ be $d$-irreducible. We prove by induction.

Base case. We prove $\gamma_2^{\oplus} \geq 2d + 1$. This follows directly from Theorem 1.3.2 and Lemma 1.3.3.

Inductive case. Assume $\gamma_{s-1}^{\oplus} \geq (s - 1)d - \frac{(s-1)(s-2)}{2} + 2$, where $2 < s < d$. We prove $\gamma_s^{\oplus} \geq sd - \frac{s(s-1)}{2} + 2$. Notice that $\gamma_{s-1}^{\oplus} \geq (s - 1)d - \sum_{i=1}^{s-2} i + 2$. Let $\mathcal{C}$ be a $s$-sum of $A$. Let $\mathcal{C}_{s-1}$ be the $(s - 1)$-sum of $(s - 1)$ columns forming $\mathcal{C}$. Since $A$ is $d$-irreducible, by Lemma 1.3.4, $|\mathcal{C}| \geq |\mathcal{C}_{s-1}| + d - s + 1$. Thus, $|\mathcal{C}| \geq (s - 1)d - \sum_{i=1}^{s-2} i + d - (s - 1) + 2 = sd - \sum_{i=1}^{s-1} i + 2 = sd - \frac{s(s-1)}{2} + 2$. □

**Lemma 1.3.5.** Let $A$ be a $t \times n$ binary matrix, where $n > t$. If $A$ is $d$-disjunct, then $A$ has less than $t - 2d$ columns of weight at most $d$.

**Proof.** Let $A$ be $d$-disjunct. We prove by contradiction.

Suppose $A$ has at least $t - 2d$ columns of weight at most $d$. Call this set of columns $\mathcal{C}$. We examine $C_j$ in $\mathcal{C}$. Let $\mathcal{R}$ be the set of at most $d$ rows where $C_j$ has a 1-entry. Assume, by way of contradiction, that each row of $\mathcal{R}$ has a 1-entry in a column other than $C_j$. Then there is a $d$-cover of $C_j$, a contradiction, since $A$ is $d$-disjunct. Thus, for each column $C_k \in \mathcal{C}$, there is a corresponding $R_i$ such that $C_j$ is the only column with a 1-entry in that row. The submatrix that results from deleting each column $C_k \in \mathcal{C}$ of $A$, and the resulting 0-rows $R_i$ corresponding to each $C_k$ is a $(t - |\mathcal{C}|) \times (n - |\mathcal{C}|)$ $d$-disjunct binary matrix. Since $|\mathcal{C}| \geq t - 2d$, there exists a $m \times (m + 1)$ $d$-disjunct binary matrix, where $m \leq 2d$, a contradiction, since by Lemma 1.3.6, no such matrix exists. Thus, $A$ has less than $t - 2d$ columns of weight at most $d$. □

### 1.3.2 Restrictions on dimensions for $d$-disjunct binary matrices

**Theorem 1.3.4.** Let $A$ be a $t \times n$ binary matrix. Let $s \leq d$ be a natural number. If $A$ is $d$-disjunct, then

$$
\binom{n}{s} \leq \binom{t}{\lfloor \frac{t}{2} \rfloor}
$$
Proof. Let $A$ be $d$-disjunct. Consider the collection $\mathcal{S}$ of all $s$-sums of the columns $A$. Since $A$ is $d$-disjunct, the subset versions of the $s$-sums of $\mathcal{S}$ must be a Sperner family of $\{1, \ldots, t\}$. There are \( \binom{t}{s} \) possible $s$-sums. Notice that for each $s$-sum $X$ contained in $\mathcal{S}$, \( \binom{t}{|X|} \leq \binom{\lfloor \frac{t}{2} \rfloor}{s} \). By the LYM Inequality, $1 \geq \sum_{X \in \mathcal{S}} \frac{1}{\binom{|X|}{s}} \geq \binom{\lfloor \frac{t}{2} \rfloor}{s} \implies \binom{t}{s} \leq \binom{\lfloor \frac{t}{2} \rfloor}{s}$

Lemma 1.3.6. Let $A$ be a $t \times (t+1)$ binary matrix. Let $d \geq 2$ be a natural number. If $A$ is $d$-disjunct, then $t > 3d$.

Proof. Let $A$ be $d$-disjunct. We prove by contradiction.

Suppose $t \leq 3d$. We prove exhaustively.

Suppose $t = 3d$. Notice that \( \binom{\lfloor \frac{2d-1}{2} \rfloor}{d-1} < \binom{\lfloor \frac{3d}{2} \rfloor}{d-1} \). So by Theorem 1.3.1, $\Gamma < d + 1$. But by Lemma 1.3.5, $\Gamma \geq d + 1$, a contradiction. Thus, $t < 3d$. That is, there is no $3d \times (3d + 1)$ $d$-disjunct binary matrix. Thus, there is no $(3d - 1) \times 3d$ $d$-disjunct binary matrix. Thus, there is no $(3d - 1) \times 3d$ $d$-disjunct binary matrix, $\ldots$, there is no $(d - 1) \times d$ $d$-disjunct binary matrix. Thus, $t > 3d$.

Lemma 1.3.7. Let $a, b, c$ be natural numbers such that $a \geq 2$, $b \geq 3c$. If $a \leq f(b, c)$, then $a - 1 \leq f(b - 1, c)$.

Proof. We prove by cases.

Suppose $a$ is even. Then \( \binom{\lfloor \frac{a-1}{2} \rfloor}{\lfloor \frac{a}{2} \rfloor} = \frac{1}{2} \binom{\frac{a}{2}}{\frac{a}{2}} \). By assumption, \( \binom{\lfloor \frac{a-1}{2} \rfloor}{\lfloor \frac{a}{2} \rfloor} = \frac{1}{2} \binom{\frac{a}{2}}{\frac{a}{2}} < \frac{1}{2} \binom{\frac{b}{2}}{\binom{\frac{b}{2}}{\frac{b}{2}}} \binom{\frac{b-1}{2}}{\frac{b-1}{2}} \). Since $b \geq 3c$, \( \frac{1}{2} \binom{\frac{b}{b-c}}{\frac{b}{b-c}} \binom{\frac{b-1}{b-c}}{\frac{b-1}{b-c}} \leq 1 \). So $a - 1 \leq f(b-1, c)$.

Suppose $a$ is odd. Then \( \binom{\lfloor \frac{a-1}{2} \rfloor}{\lfloor \frac{a}{2} \rfloor} = \frac{a+1}{2a} \binom{\frac{a}{2}}{\frac{a}{2}} \). By assumption, \( \binom{\lfloor \frac{a-1}{2} \rfloor}{\lfloor \frac{a}{2} \rfloor} = \frac{a+1}{2a} \binom{\frac{a}{2}}{\frac{a}{2}} < \frac{a+1}{2a} \binom{\frac{a+1}{b-b}}{\frac{b-1}{b-b}} \binom{\frac{b-1}{b-c}}{\frac{b-1}{b-c}} \). Since $b \geq 3c$, \( \frac{a+1}{b-b} \binom{\frac{a+1}{b-c}}{\frac{b-1}{b-c}} \binom{\frac{b-1}{b-c}}{\frac{b-1}{b-c}} \leq 1 \), since $a \geq 3$. So $a - 1 \leq f(b-1, c)$.

Theorem 1.3.5. Let $d \geq 2$ be a natural number. Let $A$ be a $t \times n$ binary matrix, where $n > t > 3d$. If $A$ is $d$-disjunct, then

\[ t > f(n-1, d-1) + d + 1 \]

Proof. We first prove that if $A$ is $d$-irreducible, then the inequality holds. We then prove that if $A$ is $d$-disjunct, then the inequality holds, regardless of $d$-irreducibility.

Let $A$ be $d$-irreducible. We prove by contradiction.

Suppose $t \leq f(n-1, d-1) + d + 1$. Then $t - f(n-1, d-1) \leq d + 1$. By Theorem 1.3.1, $\Gamma < t - f(n-1, d-1)$. Thus, $\Gamma \leq d$. Suppose $\gamma < d$. Then by the contrapositive of Theorem 1.3.2, $A$ is $d$-reducible, a contradiction. Thus, $\gamma > d$, a contradiction, since $\Gamma \leq d$. Thus, $t > d + f(n-1, d-1) + 1$.

Let $A$ be $d$-disjunct. We prove by contradiction.

Let $m_1 = f(n-1, d-1)$. Suppose $t \leq m_1 + d + 1$. By the contrapositive of the first part of this proof, there is a $(t-1) \times (n-1)$ $d$-disjunct binary matrix. Since $(n-1) \geq 3(d-1)$,
by Lemma 1.3.7, $m_2 = m_1 - 1 \leq f(n-2,d-1)$. Notice that if $t \leq d + m_1 + 1$, then $t - 1 \leq d + m_2 + 1$, so by the contrapositive of the first part of this proof, there is a $(t-2) \times (n-2)$ $d$-disjunct binary matrix. We know that for any natural number $x \leq t - 3d$, $n - x \geq 3(d - 1)$, since $n - x < 3(d - 1) \implies x > n - 3d + 3 \implies x > t - 3d + 3$, a contradiction, since $x \leq t - 3d$. Thus, this continues until, since $n - (t - 3d) \geq 3(d - 1)$, by Lemma 1.3.7, $m_{t-3d} = m_1 - (t - 3d - 1) \leq f(n - (t - 3d), d - 1)$. So by the contrapositive of the first part of this proof, there is a $(t - (t - 3d)) \times (n - (t - 3d))$ $d$-disjunct binary matrix. Thus, there exists a $3d \times (3d + 1)$ $d$-disjunct binary matrix, a contradiction, since by Lemma 1.3.6, no such matrix exists. Thus, $t > f(n - 1, d - 1) + d + 1$.

**Corollary 1.3.1.** Let $d \geq 2$ be a natural number. Let $A$ be a $t \times (t + 1)$ binary matrix, where $t > 3d$. If $A$ is $d$-disjunct, then

$$t > f(t,d - 1) + d + 1$$

**Theorem 1.3.6.** Let $d \geq 3$ be a natural number. Let $A$ be a $t \times n$ $d$-disjunct binary matrix. Let $1 < s < d$ be a natural number. If $A$ is $d$-irreducible, then

$$t > sd - \frac{s(s - 1)}{2} + 2 + f(n - s, d - s)$$

**Proof.** We first prove that if $A$ is $d$-irreducible, then the inequality holds. We then prove that if $A$ is $d$-disjunct, then the inequality holds, regardless of $d$-irreducibility.

Let $A$ be $d$-irreducible. By Theorems 1.3.3 and 1.3.1, $sd - \frac{s(s - 1)}{2} + 2 \leq \gamma_s^{\oplus} \leq \Gamma_s^{\oplus} < t - f(n - s, d - s) \implies sd - \frac{s(s - 1)}{2} + 2 < t - f(n - s, d - s) \implies t > sd - \frac{s(s - 1)}{2} + 2 + f(n - s, d - s)$. 

**Corollary 1.3.2.** Let $d \geq 3$ be a natural number. Let $A$ be a $t \times (t + 1)$ $d$-disjunct binary matrix, where $t > 3d$. If $A$ is $d$-irreducible, then

$$t > sd - \frac{s(s - 1)}{2} + 2 + f(t + 1 - s, d - s)$$

**Claim.** If $A$ is a $t \times n$ binary matrix, then $t \leq \frac{\Gamma}{\gamma}$, where $\Gamma$ is the maximum column weight of $A$ and $\gamma$ is the minimum row weight of $A$.

**Claim.** If $A$ is a $t \times n$ binary matrix, then $n \leq \frac{P}{\gamma}$, where $P$ is the maximum row weight of $A$ and $\gamma$ is the minimum column weight of $A$.

**Proposition 1.3.1.** Let $A$ be a $t \times n$ $d$-disjunct binary matrix with $\gamma = \Gamma = d + 1$. If $A$ is $d$-irreducible, then $n \leq \frac{\lceil \gamma + 1 \rceil}{d + 1}$. 

13
Proof. Let $A$ be $d$-irreducible. By Lemma 1.3.3, no column can have a 2-product of weight greater than 1 with any other column. Assume, by way of contradiction, that $P > \left\lfloor \frac{t-1}{d} \right\rfloor$. Then there exists a row $R$ such that $|R| > \left\lfloor \frac{t+d-1}{d} \right\rfloor$. Thus, there is a collection of $|R|$ columns $\mathcal{C}$ such that each column has a 1-entry in $R$. Notice that each column of $\mathcal{C}$ has $d$ 1-entries not in $R$. Notice that there are $t-1$ rows not $R$ in $A$. Since $\left\lceil \frac{t+d-1}{d} \right\rceil \geq 2$, by the pigeonhole principle, there is at least one row not $R$ of $A$ such that two columns of $\mathcal{C}$ have a 1-entry in that row. These two columns have a 2-product of weight greater than 1, a contradiction. Thus, $P \leq \left\lfloor \frac{t-1}{d} \right\rfloor$. Thus, $n \leq \frac{t-1}{d+1}$.

Proposition 1.3.2. There exists no $8 \times 9$ 2-disjunct binary matrix.

Proof. Let $A$ be a $8 \times 9$ binary matrix. We prove by contradiction. Suppose $A$ is 2-disjunct. By Corollary 1.3.1, $A$ is 2-irreducible. Thus, by Theorem 1.3.2, $\gamma > 2$. By Theorem 1.3.1, $\Gamma < 4$. Thus, the weight of all columns of $A$ is 3. Thus, by the contrapositive of Proposition 1.3.1, $A$ is not 2-irreducible, a contradiction. Thus, $A$ is not 2-disjunct.

Tables outlining the results of Corollary 1.3.2 and Proposition 1.3.2, may be found in the appendix.

Proposition 1.3.3. There exists no $9 \times 13$ 2-disjunct binary matrix.

Proof. Let $A$ be a $9 \times 13$ binary matrix. We prove by contradiction. Suppose $A$ is 2-disjunct. By Proposition 1.3.2, $A$ is 2-irreducible. Thus, by Theorem 1.3.2, $\gamma > 2$. By Theorem 1.3.1, $\Gamma < 4$. Thus, the weight of all columns of $A$ is 3. Thus, by the contrapositive of Proposition 1.3.1, $A$ is not 2-irreducible, a contradiction. Thus, $A$ is not 2-disjunct.
1.4 Counting collisions and covers

**Definition.** We denote the Boolean sum of two matrices $A$, $B$ as $A \oplus B$, where

$$(A \oplus B)[i,j] = \begin{cases} 
1, & \text{if } A[i,j] = 1 \text{ or if } B[i,j] = 1 \\
0, & \text{if } A[i,j] = B[i,j] = 0.
\end{cases}$$

Similarly, we denote the Boolean product of two matrices $A$, $B$ as $A \otimes B$, where

$$(A \otimes B)[i,j] = \begin{cases} 
1, & \text{if } A[i,j] = B[i,j] = 1 \\
0, & \text{if } A[i,j] = 0 \text{ or if } B[i,j] = 0.
\end{cases}$$

**Definition.** We define the weight of a matrix $A$, denoted $|A|$, to be the sum of the entries of $A$. In a binary matrix, the weight is the number of 1-entries.

1.4.1 $s$-collision and $s$-cover matrices

Let $A$ be a matrix and $s$ a natural number. Let $A_s^\oplus$ be the $s$-sum-matrix of $A$. Let $C_1$ and $C_2$ be distinct columns of $A_s^\oplus$ (so they are distinct $s$-sums of $A$). If $C_1 = C_2$, then $\{C_1, C_2\}$ forms a $s$-collision of $A$.

**Notation.** We denote the number of $s$-collisions of $A$ as $Z_s^\oplus(A)$. The number of $s$-collisions will be denoted $Z_s^\ominus(A)$. We denote the number of $s$-covers of a matrix $A$ as $Z_s^\subseteq(A)$.

**Claim.** $A$ is $d$-separable if and only if $Z_s^\ominus(A) = 0$. $A$ is $d$-separable if and only if $Z_s^\oplus(A) = 0$ for each $s \leq d$. $A$ is $d$-disjunct if and only if $Z_s^\subseteq(A) = 0$.

**Definition.** Let $A$ be a $t \times n$ matrix and let $s$ be a natural number. Let $A_s^=\oplus$ be the $\binom{n}{s} \times \binom{n}{s}$ matrix defined by

$$A_s^=\oplus[i,j] = \begin{cases} 
1, & \text{if } C_j = C_i \text{ in } A_s^\oplus \\
0, & \text{otherwise}.
\end{cases}$$

We call $A_s^=\oplus$ the $s$-collision matrix of $A$.

**Definition.** Let $A$ be a $t \times n$ matrix. Let $A_s^\ominus$ be the $s$-sum-matrix of $A$. Define $A_s^\subseteq$ to be the $\binom{n}{s} \times \binom{n}{s}$ binary matrix whose entries are given by:

$$A_s^\subseteq[i,j] = \begin{cases} 
1, & \text{if } C_i \subseteq C_j \text{ in } A_s^\oplus \\
0, & \text{otherwise}.
\end{cases}$$

We call $A_s^\subseteq$ the $s$-cover-matrix of $A$. 

15
Proposition 1.4.1. Let $A$ be a $t \times n$ matrix. Then we have

$$3^\mathcal{S}_\mathcal{S}(A) = \frac{1}{2} \left| A^\mathcal{S}_\mathcal{S} \right| - \binom{n}{s}$$

and

$$3^C(A) = \left| A^C \right| - \binom{n}{s}.$$  

Proof. The 1-entries on the diagonal of $A^\mathcal{S}_\mathcal{S}$ represent the same sum paired with itself, which is not a $s$-collision, so we must subtract these $\binom{n}{s}$ 1-entries. Each $s$-collision $C_i$, $C_j$ is represented twice in $A^\mathcal{S}_\mathcal{S}$, in $A^\mathcal{S}_\mathcal{S}[i,j]$ and $A^\mathcal{S}_\mathcal{S}[j,i]$, so we must divide by 2. A similar proof works for the second equation, but since covers are ordered pairs, we don’t divide by 2. □

Lemma 1.4.1. Let $A$ be a $t \times n$ binary matrix. Let $d$ be a natural number. Let $A(1), \ldots, A(m)$ be submatrices of $A$ such that every row of $A$ is a row in at least one of $A(1), \ldots, A(m)$ (so each $A(i)$ has $n$ columns). If $|A(1)_{d}^\cap \cdots \cap A(m)_{d}^\cap| > \binom{n}{d}$, then $A$ is not $d$-separable. If $|A(1)_{d}^\cap \cdots \cap A(m)_{d}^\cap| > \binom{n}{d}$, then $A$ is not $d$-disjunct.

Proof. Notice $|A(1)_{d}^\cap \cdots \cap A(m)_{d}^\cap| - \binom{n}{d}$ is the number of $d$-collisions which are $d$-collisions in each of $A(1), \ldots A(m)$. If there are distinct $d$-sums $\mathcal{C}_1$ and $\mathcal{C}_2$ in $A$ such that $\mathcal{C}_1(i) = \mathcal{C}_2(i)$ in $A(i)$ for each $i$, this corresponds with a $d$-collision in $A$. If $A$ has a $d$-collision, then $A$ cannot be $d$-separable. Similarly $|A(1)_{d}^\cap \cdots \cap A(m)_{d}^\cap|$ is the number of $d$-covers which are $d$-covers in each of $A(1), \ldots A(m)$. If there is a $d$-sum $\mathcal{C}_1(i)$ which covers $d$-sum $\mathcal{C}_2(i)$ in $A(i)$ for each $i$, this corresponds with a $d$-cover in $A$. If $|A(1)_{d}^\cap \cdots \cap A(m)_{d}^\cap| > \binom{n}{d}$, then $A$ has at least one $d$-cover, and therefore $A$ is not $d$-disjunct. □

Lemma 1.4.2. Let $A$ be a $t \times n$ matrix, and let $A(1)$ and $A(2)$ be submatrices of $A$ such that every row of $A$ is a row in either $A(1)$ or in $A(2)$, possibly both. If $A$ is $d$-separable, then we have

$$|A(1)_{d}^\cap| + |A(2)_{d}^\cap| \leq \left( \frac{n}{d} \right)^2 + \left( \frac{n}{d} \right).$$

Furthermore, if $A$ is $d$-disjunct, then we have

$$|A(1)_{d}^\cap| + |A(2)_{d}^\cap| \leq \left( \frac{n}{d} \right)^2 + \left( \frac{n}{d} \right).$$

Proof. We have $|A(1)_{d}^\cap| + |A(2)_{d}^\cap| = |A(1)_{d}^\cap \circ \circ A(2)_{d}^\cap| + |A(1)_{d}^\cap \circ A(2)_{d}^\cap| = |A(1)_{d}^\cap \circ A(2)_{d}^\cap| + |A_{d}^\cap| \leq \binom{n}{d}^2 + \binom{n}{d}$, applying Lemma 1.4.1. A similar proof works for the disjunct case. □

Theorem 1.4.1. Let $A$ be a $t \times n$ $d$-separable binary matrix, and let $A(1), \ldots, A(m)$ be submatrices of $A$ such that every row of $A$ is a row in at least one of $A(1), \ldots, A(m)$. Then we have

$$\sum_{i=1}^{m} |A(i)_{d}^\cap| \leq (m - 1) \left( \frac{n}{d} \right)^2 + \left( \frac{n}{d} \right).$$
Furthermore, if \( A \) is \( d \)-disjunct, then we have
\[
\sum_{i=1}^{m} |A(i)\subseteq\bar{d}| \leq (m - 1) \left( \binom{n}{d} \right)^2 + \binom{n}{d}.
\]

**Proof.** We prove the separable case. The disjunct case is proved analogously. We prove by induction on \( m \). By Lemma 1.4.2 we know the inequality is true for \( m = 2 \). Suppose it is true for \( m \geq 2 \). Let \( A(1), \ldots, A(m+1) \) be submatrices of \( A \) such that every row of \( A \) is a row in at least one of \( A(1), \ldots, A(m+1) \). Consider the submatrix of \( A \) which results from deleting only the rows contained in \( A(m+1) \). Call this submatrix \( \bar{A} \). Then \( A(1), \ldots, A(m) \) are submatrices of \( \bar{A} \) satisfying the conditions of our inductive hypothesis, so we know
\[
\sum_{i=1}^{m} |A(i)\subseteq\bar{d}| \leq (m - 1) \left( \binom{n}{d} \right)^2 + \binom{n}{d}.
\]
Thus we have
\[
\sum_{i=1}^{m+1} |A(i)\subseteq\bar{d}| = \left[ \sum_{i=1}^{m} |A(i)\subseteq\bar{d}| \right] + |A(m+1)\subseteq\bar{d}| \leq (m - 1) \left( \binom{n}{d} \right)^2 + \binom{n}{d} + \left( \binom{n}{d} \right)^2 + \binom{n}{d}^2 = m \left( \binom{n}{d} \right)^2 + \binom{n}{d}
\]
as desired. \( \square \)

### 1.4.2 Row weight inequalities

Consider a \( t \times n \) matrix \( A \), and consider the rows \( R(1), \ldots, R(t) \) of \( A \). If we think of the rows as submatrices of \( A \), we can apply Theorem 1.4.1 to get the following result:

**Lemma 1.4.3.** If \( A \) is \( d \)-separable, we have
\[
\sum_{i=1}^{t} |R(i)\subseteq\bar{d}| \leq (t - 1) \left( \binom{n}{d} \right)^2 + \binom{n}{d}.
\]

If we also have that \( A \) is \( d \)-disjunct, then
\[
\sum_{i=1}^{t} |R(i)\subseteq\bar{d}| \leq (t - 1) \left( \binom{n}{d} \right)^2 + \binom{n}{d}.
\]

**Lemma 1.4.4.** Let \( R \) be a \( 1 \times n \) binary matrix. Then
\[
|R_s\subseteq| = \left( \binom{n}{s} \right)^2 + 2 \left( \binom{n-|R|}{s} \right)^2 - 2 \left( \binom{n}{s} \right) \left( \binom{n-|R|}{s} \right)
\]

**Proof.** If the weight of \( R \) is \( |R| \), then there are \( |R| \) 1-entries of \( R \) and \( (n - |R|) \) 0-entries. Consider the \( 1 \times \binom{n}{s} \) \( s \)-sum-matrix \( R_s\subseteq \). There are \( \binom{n-|R|}{s} \) 0-entries and \( \binom{n}{s} - \binom{n-|R|}{s} \) 1-entries in \( R_s\subseteq \). This corresponds to have exactly \( \left( \binom{n-|R|}{s} \right)^2 + \left[ \binom{n}{s} - \binom{n-|R|}{s} \right]^2 \) 1-entries in \( A_s\subseteq \), which reduces to the desired result. \( \square \)

**Theorem 1.4.2.** Let \( A \) be a \( t \times n \) \( d \)-separable matrix with rows \( R(1), \ldots, R(t) \). Then we have
\[
\left( \binom{n}{d} \right)^2 - \frac{n}{d} \leq 2 \sum_{i=1}^{t} \left[ \binom{n}{d} \left( \frac{n-|R(i)|}{d} \right) - \left( \frac{n-|R(i)|}{d} \right)^2 \right]^2.
\]
Proof. The proof follows directly from Lemma 1.4.4 and Lemma 1.4.3.

**Lemma 1.4.5.** Let $R$ be a $1 \times n$ binary matrix. Then

$$|R^<_s| = \binom{n}{s}^2 - \binom{n-|R|}{s} \binom{n}{s} + \left(\binom{n-|R|}{s}\right)^2$$

Proof. If the weight of $R$ is $|R|$, then there are $|R|$ 1-entries of $R$ and $(n - |R|)$ 0-entries. Consider the $1 \times \binom{n}{s}$ $s$-sum-matrix $R^<_s$. There are $\binom{n-|R|}{s}$ 0-entries and $\binom{n}{s} - (\binom{n-|R|}{s})$ 1-entries in $R^<_s$. This corresponds to $\binom{n-|R|}{s} \cdot \binom{n}{s} + \left[\binom{n}{s} - (\binom{n-|R|}{s})\right]^2$ 1-entries in $A^>_d$, which reduces to the desired result.

**Theorem 1.4.3.** Let $A$ be a $t \times n$ $d$-disjunct matrix with rows $R(1), \ldots R(t)$. Then we have

$$\binom{n}{d}^2 - \binom{n}{d} \leq \sum_{i=1}^{t} \left[\binom{n}{d} \left(n - |R(i)|\right) + \left(n - |R(i)|\right)^2\right]$$

(1.3)

Proof. The proof follows directly from Lemma 1.4.5 and Lemma 1.4.3.

### 1.4.3 A lower bound on the number of collisions

**Lemma 1.4.6.** Let $A$ be a $t \times n$ matrix, and let $A^>_s$ be the $s$-sum-matrix of $A$. Suppose $A^>_s$ has exactly $v$ nonidentical columns $C_1, \ldots, C_v$ such that each column in $A^>_s$ is equal to one of the columns in $C_1, \ldots, C_v$. Let $\chi_i$ denote the number of columns of $A^>_s$ which are equal to $C_i$. Let $A^>_s$ be the $s$-collision matrix of $A$. Then we have:

$$|A^>_s| = \sum_{i=1}^{v} \chi_i^2$$

Proof. Each pair of identical columns corresponds to exactly one 1-entry in $A^>_s$. This includes repeated columns and any order, so there are $\chi_i^2$ 1-entries for each set of $\chi_i$ identical columns.

**Theorem 1.4.4.** Let $A$ be a $t \times n$ matrix, and let $A^>_s$ be the $s$-sum-matrix of $A$. Suppose $A^>_s$ has exactly $v$ nonidentical columns $C_1, \ldots, C_v$ such that each column in $A^>_s$ is equal to one of the columns in $C_1, \ldots, C_v$. Let $\mathfrak{B}_s(A)$ be the number of $s$-collisions of $A$. Then we have:

$$\mathfrak{B}_s(A) = \frac{1}{2} \sum_{i=1}^{v} \chi_i^2 - \binom{n}{s}$$
**Theorem 1.4.5.** Let $A$ be a $t \times n$ matrix, and let $A^\oplus_s$ be the $s$-sum-matrix of $A$. Suppose $A^\oplus_s$ has exactly $v$ nonidentical columns $C_1, \ldots, C_v$ such that each column in $A^\oplus_s$ is equal to one of the columns in $C_1, \ldots, C_v$. Let $k$ and $r$ be nonnegative integers such that $t \choose s = kv + r$, and $r < v$. Then we have:

$$|A^\oplus_s| \geq (v - r) \left( \frac{s}{v} \right)^2 + r \left( \frac{s}{v} \right)^2$$

**Proof.** Consider the function $F(x_1, \ldots, x_v) = \sum_{i=1}^{v} x^2$ on the nonnegative integers, subject to the restriction $\sum_{i=1}^{v} x_i = \binom{n}{s}$. This function takes a minimum when each $x_i$ is as equal as possible, which occurs when $x_1 = \cdots = x_{v-r} = \left\lfloor \frac{s}{v} \right\rfloor$, and $x_{v-r+1} = \cdots = x_v = \left\lceil \frac{s}{v} \right\rceil$. This function models $|A^\oplus_s|$, and for any choice of $\chi_1, \ldots, \chi_v$, we have $F(\chi_1, \ldots, \chi_v) \geq (v - r) \left( \frac{s}{v} \right)^2 + r \left( \frac{s}{v} \right)^2$. \hfill \Box

### 1.4.4 Using a generalization of Sperner’s theorem to count covers

The following definition and theorem can be found in [8].

**Definition.** A **multifamily** of a set $\mathcal{I}$ is a collection of subsets of $\mathcal{I}$ where repetitions are allowed. We use two notations. We can list the elements of our multifamily $\mathcal{M} = \{Y_1, \ldots, Y_n\}$, where each $Y_i$ is a distinct (but possibly equal) subset. Or we can write $\mathcal{M} = \{(\chi_1, m_1), \ldots, (\chi_q, m_q)\}$, where $\chi_i \neq \chi_j$ for $i \neq j$, but each subset $Y_i$ is equal to some $\chi_j$. We call the set of representative subsets $\{\chi_1, \ldots, \chi_q\}$ the **support** of $\mathcal{M}$. We call the set $\{m_1, \ldots, m_q\}$ the set of **multiplicities**. Each $m_i$ represents the number of subsets in $\mathcal{M}$ which are equal to $\chi_i$. So we have $\sum_{i=1}^{q} m_i = n$.

**Notation.** Let $\mathcal{M} = \{Y_1, \ldots, Y_n\}$ be a multifamily of a set $\mathcal{T}$. We denote the set of ordered pairs $(i, j)$ such that $Y_i \subseteq Y_j$ as $\phi(\mathcal{M})$.

**A Generalization of Sperner’s Theorem.** Let $\mathcal{M}$ be a multifamily of a $t$-element set $\mathcal{T}$. Suppose there are $n$ subsets in $\mathcal{M}$. Let $k$, $r$ be the nonnegative integers satisfying $n = k \binom{t}{\frac{t}{2}} + r$ and $r < \binom{t}{\frac{t}{2}}$. Then we have

$$|\phi(\mathcal{M})| \geq k(k - 1) \binom{t}{\frac{t}{2}} + 2kr + n.$$  

If $k = 0$, or if $k = 1$ and $r = 0$, then equality holds if and only if $\mathcal{M}$ is a Sperner family. For $k \geq 1$, equality holds if the support of $\mathcal{M}$ consists of all subsets of $\mathcal{T}$ of size $\lfloor \frac{t}{2} \rfloor$, or all subsets of $\mathcal{T}$ of size $\lceil \frac{t}{2} \rceil$, and if all multiplicities are $k$ or $k + 1$. If $k \geq 1$ and $t \geq 4$, then no other multifamilies achieve this bound.
Claim. Let $A$ be a $t \times n$ binary matrix. Let $s$ be a natural number, let $A^\oplus_s$ be the $s$-sum-matrix of $A$, and let $A^\subseteq_s$ be the $s$-cover matrix of $A$. If you consider the columns of $A^\oplus_s$ as subsets of $[t]$, then we have $|A^\subseteq_s| = |\phi(A^\oplus_s)|$.

**Proposition 1.4.2.** Let $A$ be a $t \times n$ binary matrix. Let $s$ be a natural number. Let $k$, $r$ be the nonnegative integers satisfying $\binom{n}{s} = k\left(\binom{\frac{t}{2}}{\frac{s}{2}}\right) + r$ and $r < \binom{\frac{t}{2}}{\frac{s}{2}}$. Then we have

$$3^\subseteq_s \geq k(k-1)\left(\binom{\frac{t}{2}}{\frac{s}{2}}\right) + 2kr.$$

Equality is achieved if and only if $A^\oplus_s$ has columns of equal weight, either $|C| = \binom{\frac{t}{2}}{\frac{s}{2}}$ or $|C| = \binom{\frac{t}{2}}{\frac{s}{2}}$ for each column $C$ of $A^\oplus_s$.

### 1.4.5 Applying outside problems to disjunctness

The following definition and theorem can be found in [4].

**Definition.** Let $\mathcal{M}$ be a multifamily of subsets of $[t]$. We say the parameters of $\mathcal{M}$ is the set of numbers $\{p_0, \ldots, p_t\}$, where $p_i$ represents the number of subsets of $\mathcal{M}$ of cardinality $i$.

**Existence Theorem for Sperner Families.** Let $\mathcal{S}$ be a Sperner family of subsets of $[t]$. Let $\{p_0, \ldots, p_t\}$ be the parameters of $\mathcal{S}$. Then there is a Sperner family $\mathcal{Y}$ on $[t]$ with parameters $\{q_0, \ldots, q_t\}$, where $q_i = 0$ for $0 \leq i < \frac{t}{2}$, $q_i = p_{i-1} + p_i$ for $\frac{t}{2} < i \leq t$, and when $t$ is even, $q_{\frac{t}{2}} = p_{\frac{t}{2}}$.

**Definition.** Let $A$ be a $t \times n$ matrix. We call the parameters of $A$ the set of numbers $\{p_0, \ldots, p_t\}$, where $p_i$ represents the number of columns of $A$ with weight $i$.

**Proposition 1.4.3.** Let $A$ be a $t \times n$ $d$-disjunct matrix. Let $\{p_0, \ldots, p_t\}$ be the parameters of $A^\oplus$. Then there exists a $t \times \binom{n}{d}$ $1$-disjunct matrix with parameters $\{q_0, \ldots, q_t\}$, where $q_i = 0$ for $0 \leq i < \frac{t}{2}$, $q_i = p_{i-1} + p_i$ for $\frac{t}{2} < i \leq t$, and when $t$ is even, $q_{\frac{t}{2}} = p_{\frac{t}{2}}$.

The following definition and theorem can be found in [7].

**Definition.** A Sperner family $\mathcal{S}$ is called flat if for all $S \in \mathcal{S}$, $|S| = x$ or $|S| = x + 1$ for some nonnegative integer $x$.

**The Flat Antichain Theorem.** If $\mathcal{S}$ is a Sperner family of $[t]$, then there exists a flat Sperner family $\mathcal{Y}$ of $[t]$ with the same number of subsets as $\mathcal{S}$, and the same average set size as $\mathcal{S}$.

**Lemma 1.4.7.** Let $A$ be a $t \times n$ $d$-disjunct matrix. Suppose $|A^\oplus| = k\binom{n}{d} + r$ for nonnegative integers $k$, $r$, with $r < \binom{n}{d}$. Then there exists a $1$-disjunct $t \times \binom{n}{d}$ matrix $B$ with $\binom{n}{d} - r$ columns of weight $k$ and $r$ columns of weight $r + 1$. 
Theorem 1.4.6. Let $A$ be a $t \times n$ $d$-disjunct matrix, and let $A_{d}^\oplus$ be the $d$-sum-matrix of $A$. Let $k = \left\lceil \frac{|A_{d}^\oplus|}{\binom{n}{d}} \right\rceil$. Then we have

$$\left( \frac{n}{d} \right) \leq \max \left( \binom{t}{k}, \binom{t}{k+1} \right).$$

Proof. Let $M = \max \left( \binom{t}{k}, \binom{t}{k+1} \right)$, and suppose $\left( \frac{n}{d} \right) > M$. Let $k = \left\lfloor \frac{|A_{d}^\oplus|}{\binom{n}{d}} \right\rfloor$, and let $r$ be the nonnegative integer satisfying $|A_{d}^\oplus| = k\binom{n}{d} + r$. Then by Lemma 1.4.7, we know there exists 1-disjunct $t \times \binom{n}{d}$ matrix $B$ with $\binom{n}{d} - r$ columns of weight $k$ and $r$ columns of weight $k + 1$. Applying the LYM Inequality, we see that $\frac{\binom{n}{d}}{M} \leq \frac{(\binom{n}{d} - r)}{\binom{t}{k}} + \frac{r}{\binom{t}{k+1}} \leq 1$. Therefore we have $\left( \frac{n}{d} \right) \leq M$ and the proof is complete. \hfill \Box

Example. Suppose there exists a $10 \times 14$ 2-disjunct matrix, $A$. Let $x$ denote the average column weight of $A_{d}^\oplus$. Then $3 \leq x \leq 7$.

1.5 Existence of a $10 \times 14$ 2-disjunct binary matrix

Suppose there is a $10 \times 14$ 2-disjunct binary matrix. Call this matrix $\mathcal{A}$.

Notation. We denote the number of columns of a matrix $A$ with weight $w$ as $\aleph_w(A)$. When it is clear, we may use $\aleph_w$.

Proposition 1.5.1. $\aleph_4(\mathcal{A}) \geq 5$.

Proof. We prove by contradiction, by cases.

Suppose $\mathcal{A}$ has less than 4 columns of weight 4. By Theorem 1.3.1, each column of $\mathcal{A}$ has weight at most 4. Since $\mathcal{A}$ is $d$-irreducible, by Theorem 1.3.2, each column of $\mathcal{A}$ has weight at least 3. By Lemma 1.1.3, $\mathcal{A}$ has a row of weight at least 5. Since $\aleph_4 \leq 3$ by Lemma 1.7.1 there are greater than 10 rows in $\mathcal{A}$, a contradiction. Thus, $\mathcal{A}$ has at least 4 columns of weight 4.

Suppose $\mathcal{A}$ has 4 columns of weight 4. Since $\mathcal{A}$ is $d$-irreducible, by Theorem 1.3.2, $\gamma(A) = 3$. Let $\mathcal{A}'$ be the submatrix of $\mathcal{A}$ taken by deleting a 4-column of $\mathcal{A}$. Notice that $\mathcal{A}$ has 10 3-columns and 3 4-columns. By Lemma 1.1.3, $P(\mathcal{A}') \geq 5$. Thus, by Lemma 1.7.1, there are greater than 10 rows in $\mathcal{A}'$, a contradiction. Thus, $\mathcal{A}'$ has at least 5 columns of weight 4.

Notation. Let $C_1, C_2, \ldots, C_j$ be a collection of columns. We call the set of indices $i$ such that $(C_i \oplus C_2 \oplus \cdots \oplus C_j)[i] = 0$, the zero-set of $C_1, C_2, \ldots, C_j$, and use the notation $Z[1, 2, \ldots, j]$. That is, the zero-set of a number of columns is the set of rows where all the columns have 0-entries. We call the set of indices where at least one of the columns has a 1-entry the unit-set of $C_1, C_2, \ldots, C_j$ and use the notation $U[1, 2, \ldots, j]$. 21
Proposition 1.5.2. \( \mathcal{A} \) has at least one pair of 4-columns whose 2-product has weight 2.

Proof. We prove by contradiction.

Suppose \( \mathcal{A} \) has no 4-columns whose 2-product with another 4-column has weight 2. By Proposition 1.5.1, \( \mathcal{A} \) has at least 5 columns of weight 4. Without loss of generality, assume \( C_1, \ldots, C_5 \) are 4-columns. Suppose for some two 4-columns, the weight of their 2-product is 0. Without loss of generality, assume these are \( C_1, C_2 \). Suppose for some other 4-column, \( C_Z, C_Z \) has two 1-entries in \( Z[1,2] \). Then since \( |C_1 \oplus C_2 \oplus C_Z| = 10 \), by the pigeonhole principle any other 4-column must have a 2-product with at least one of \( C_1, C_2, C_Z \) with weight at least 2, a contradiction, since no such column exists. Thus no other 4-column has two 1-entries in \( Z[1,2] \). Suppose a 4-column has one 1-entry in \( Z[1,2] \). Such a column has three 1-entries in \( U[1,2] \), and thus has a 2-product with at least one of \( C_1, C_Z \) with weight at least 2, a contradiction, since no such column exists. Since \( \mathcal{A} \) is 2-disjunct, there cannot be any column of \( A \) with zero 1-entries in \( Z[1,2] \). Thus, since there are no columns with 0, 1, 2 1-entries in \( Z[1,2] \), and there are two rows \( Z[1,2] \), there is a contradiction. Thus, there are no two 4-columns of \( \mathcal{A} \) with a 2-product of weight 0. Thus, each 4-column of \( \mathcal{A} \) has a 2-product with any other 4-column with weight 1.

Suppose at least 2 4-columns of \( \mathcal{A} \) have a 3-product with \( C_1 \) of weight 1. Notice the 2-sum of these two columns must have a 1-entry in each of the 6 rows of \( Z[1] \). Then by the pigeonhole principle, since these three columns have at least one 1-entry in each row of \( \mathcal{A} \), any other 4-column must have a 2-product with one of these three columns of weight at least 2, a contradiction. Thus no 2 4-columns of \( \mathcal{A} \) have a 3-product with \( C_1 \) of weight 1.

Suppose there are at least six 4-columns of \( \mathcal{A} \). Then by the pigeonhole principle, at least 2 4-columns of \( \mathcal{A} \) have a 3-product with \( C_1 \) of weight 1, a contradiction, since no such columns exist. Thus, there are exactly five 4-columns of \( \mathcal{A} \).

Notice that there are nine 3-columns in \( \mathcal{A} \). Suppose at least four of these columns have zero 1-entries in \( U[1] \). Notice that the submatrix \( \mathcal{A}' \) with 6 rows and at least 4 columns formed by taking these columns and the rows of \( Z[1] \) must be 2-disjunct. Thus, by Theorem 1.3.1 \( \Gamma(\mathcal{A}') < 3 \), a contradiction, since \( \Gamma(\mathcal{A}) = 3 \). Thus, at least six 3-columns have one 1-entry in \( U[1] \).

By the pigeonhole principle, at least two of these 3-columns have a 1-entry in the same row of \( U[1] \), \( R_i \). Thus, the other two 1-entries for each of these columns must occupy a distinct four rows of \( Z[1] \).

Since each 4-column not \( C_1 \) has a 1-entry in \( U[1] \), and no two of them have that 1-entry in that same row, one of the 4-columns has a 1-entry in \( R_i \). Since there are only two rows not occupied by the two 3-columns in \( Z[1] \), the 4-column must have a 2-product with one of the 3-columns of weight at least 2, a contradiction. Thus, \( \mathcal{A} \) has at least 2 4-columns whose 2-product has weight 2. \( \square \)
1.6 A conjecture on $d$-disjunct matrices with optimal dimensions

**Definition.** For a given number of rows $t$, we call a $t \times n$ $d$-disjunct binary matrix $d$-optimal if there exists no $t \times (n + 1)$ $d$-disjunct binary matrix. We also say the matrix is of $d$-optimal dimensions.

**Conjecture.** For a given number of rows, there exists a binary matrix $A$ of $d$-optimal dimensions such that $A$ only has columns whose weight is of the form $nd + 1$, where $n$ is a natural number.

**Evidence:** Every best known matrix for $d=2$ that we’ve seen (or seen at least seen the Steiner system they’re based off of) have column weights of 1, 3, 5, or 7 ($t=9, 10, 11, 12, 13, 16, 17, 21, 22, 23, 26$). All columns of weight $d + 1$ can overlap with any other column of weight $d + 1$ in any one place independently of other columns, all columns of weight $2d + 1$ can overlap with any other column of weight $2d + 1$ in any two places independently of other columns, etc. For $d=2$ and a given 4-column $C$, there are only at most three ways other 4-columns can overlap by two with $C$, and there are at most ten ways 5-columns can overlap by two with $C$.

**Lemma 1.6.1.** Let $A$ be a $t \times n$ 2-disjunct binary matrix. If $A$ has exactly one column of weight 4 and $n - 1$ columns of weight 3, then there exists a $t \times n$ 2-disjunct binary matrix with all columns of weight 3.

**Proof.** Let $A$ be a $t \times n$ 2-disjunct binary matrix with one column of weight 4 and $n - 1$ columns of weight 3. Without loss of generality, assume $C_1$ is the 4-column. Let $\hat{A}$ be the matrix which results by changing exactly one 1-entry in $C_1$ to a 0-entry. Let $\hat{C}_j$ denote the column of $\hat{A}$ which results from column $C_j$ of $A$. Notice that $\hat{C}_j = C_j$ for all $1 < j \leq n$. Suppose $\hat{A}$ is not 2-disjunct. Clearly, this can only happen if $\hat{C}_1$ is covered by two columns of $\hat{A}$. Without loss of generality, assume these are $\hat{C}_2, \hat{C}_3$. Thus, $\hat{C}_1 \otimes \hat{C}_2 + \hat{C}_1 \otimes \hat{C}_3 \geq 3$. Then $C_1 \otimes C_2 + C_1 \otimes C_3 \geq 3$. So by the pigeonhole principle, at least one of $C_2, C_3$ has a 2-product with $C_1$ with weight at least 2, a contradiction, since by Lemma 1.3.3, the 2-product is at most 1. Thus $A$ is 2-disjunct.

**Lemma 1.6.2.** Let $A$ be a $t \times n$ 2-disjunct binary matrix. If $A$ has exactly two columns of weight 4 and $n - 2$ columns of weight 3, then there exists a $t \times n$ 2-disjunct binary matrix with all columns of weight 3.

**Proof.** Let $A$ be a $t \times n$ 2-disjunct binary matrix with two columns of weight 4 and $n - 2$ columns of weight 3. Without loss of generality, assume $C_1, C_2$ are the 4-columns. We prove by cases.

Suppose $C_1 \otimes C_2 \geq 1$. Then there is at least one row where both $C_1, C_2$ have a 1-entry. Without loss of generality, assume this is $R_1$. Let $\tilde{A}$ be the matrix which results by changing
\(C_1[1], C_2[1]\) to 0-entries. Let \(\hat{C}_j\) denote the column of \(\hat{A}\) which results from column \(C_j\) of \(A\). Notice \(\hat{C}_j = C_j\) for all \(3 \leq j \leq n\). Suppose \(\hat{A}\) is not 2-disjunct. Clearly, this can only happen if \(C_1\) or \(C_2\) is covered. Without loss of generality, assume \(\hat{C}_1\) is covered. Suppose \(\hat{C}_1\) is covered by \(\hat{C}_2\) with some other column \(\hat{C}_k\) of \(\hat{A}\). Then \(C_1\) is covered by \(C_2 \oplus C_k\), a contradiction, since \(A\) is 2-disjunct. Suppose \(\hat{C}_1\) is covered by two columns \(\hat{C}_l, \hat{C}_m\) not \(\hat{C}_2\) of \(\hat{A}\). Then \(\hat{C}_1 \oplus \hat{C}_l + \hat{C}_1 \oplus \hat{C}_m \geq 3\). So by the pigeonhole principle, at least one of \(C_l, C_m\) has a 2-product with \(C_1\) with weight at least 2, a contradiction, since by Lemma 1.3.3, the 2-product is at most 1. Thus \(\hat{A}\) is 2-disjunct.

Suppose \(C_1 \times C_2 = 0\). Let \(\hat{A}\) be the matrix which results by changing exactly one 1-entry in both \(C_1, C_2\) to a 0-entry. Let \(\hat{C}_j\) denote the column of \(\hat{A}\) which results from column \(C_j\) of \(A\). Notice that \(\hat{C}_j = C_j\) for all \(3 \leq j \leq n\). Suppose \(\hat{A}\) is not 2-disjunct. Clearly, this can only happen if \(\hat{C}_1\) or \(\hat{C}_2\) is covered. Without loss of generality, assume \(\hat{C}_1\) is covered. Notice that \(C_1\) cannot be covered by the 2-sum of \(C_2\) with any other column of \(\hat{A}\). Thus, \(\hat{C}_1\) must be covered by two columns \(\hat{C}_k, \hat{C}_l\) not \(\hat{C}_2\) of \(\hat{A}\). Then \(\hat{C}_1 \oplus \hat{C}_k + \hat{C}_1 \oplus \hat{C}_l \geq 3\). So by the pigeonhole principle, at least one of \(C_k, C_l\) has a 2-product with \(C_1\) with weight at least 2, a contradiction, since by Lemma 1.3.3, the 2-product is at most 1. Thus \(\hat{A}\) is 2-disjunct.

1.7 Other items of interest

Lemma 1.7.1. Let \(A\) be an \(t \times n\) binary matrix where \(\Gamma \leq \gamma + 1 = d + 2\). Let \(R_i\) be a row of weight \(P\). Let \(\aleph_w\) denote the number of \(w\)-columns with a 1-entry in \(R_i\). We define \(\alpha_a\) and \(\beta_b\) as follows:

\[
\alpha_a = a \cdot d, 0 \leq a \leq \aleph_\gamma
\]

\[
\psi_b = \frac{b(2\Gamma - 1) - b^2}{2}
\]

\[
\omega_0 = 0, \omega_b = \min \{ \varphi : \exists \varphi \times b \text{ d-disjunct binary matrix with weights all } \Gamma - 1\}
\]

\[
\beta_b = \max_{0 \leq b \leq \aleph_i} \{ \psi_b, \omega_b\}
\]

If \(A\) is d-disjunct, then \(t > \min_{a+b=P} \{ \alpha_a + \beta_b \}\).

Proof. Let \(A\) be an \(d\)-disjunct \(t \times n\) \(d\)-disjunct binary matrix, \(\Gamma \leq \gamma + 1 = d + 2\). Let \(R_i\) be a row of weight \(P\). Let the \(\mathcal{C}\) be the collection of columns with a 1-entry in \(R_i\). Let \(\mathcal{C}_s\) be the subset of \(s\)-columns of \(\mathcal{C}\).

Let \(\alpha_a = a \cdot d, 0 \leq a \leq \aleph_\gamma\). We show that if \(|\mathcal{C}_s| = a\), then \(\mathcal{C}_s\) has 1-entries in \(\alpha_a\) rows not \(R_i\).

Suppose \(|\mathcal{C}_s| = a\). By the inclusion-exclusion principle and Lemma 1.3.3, each column of \(\mathcal{C}_s\) has a 2-product with any other column of \(\mathcal{C}\) of weight at most 1. Since each column of \(\mathcal{C}_s\) has a 1-entry in \(R_i\), it must be that the other \(d\) entries of each column of \(\mathcal{C}_s\) are the
only 1-entries in the row among the columns of $\mathcal{C}$. Thus, there are at least $a \cdot d$ such rows not $R_i$. Call these rows $\mathcal{R}_\gamma$. 

Let $\psi_0 = 0, \psi_b = \psi_{b-1} + \Gamma - b, 1 \leq b \leq \aleph_\Gamma$. We show by induction that if $|\mathcal{C}_\Gamma| = b$, then $\mathcal{C}_\Gamma$ has 1-entries in at least $\psi_b$ rows not $R_i$.

Base case. Suppose $|\mathcal{C}_\Gamma| = 0$. Then clearly, $\mathcal{C}_\Gamma$ has 1-entries in at least 0 rows not $R_i$.

Inductive case. Assume that if $|\mathcal{C}_\Gamma| = b - 1$, $\mathcal{C}_\Gamma$ has 1-entries in at least $\psi_{b-1}$ rows not $R_i$. Notice that by the inclusion-exclusion principle and Lemma 1.3.3, the 2-product of any columns of $\mathcal{C}_\Gamma$ is at most $\Gamma - d$. Thus, for any pair of columns of $\mathcal{C}_\Gamma$, there is at most $\Gamma - d - 1$ rows not $R_i$ such that both columns have a 1-entry in that row. Thus, the $b^{th}$ column of $\mathcal{C}_\Gamma$ has at least $\Gamma - 1 - (b - 1)(\Gamma - d - 1) \geq \Gamma - b$ rows such that the column is the only column of $\mathcal{C}_\Gamma$ to have a 1-entry in that row. Thus, $\mathcal{C}_\Gamma$ has 1-entries in at least $\psi_{b-1} + \Gamma - b$ rows not $R_i$. That is, $\mathcal{C}_\Gamma$ has 1-entries in at least $\psi_b$ rows not $R_i$. Notice that this sequence can be given by $\psi_b = b(2\Gamma - 1) - b^2 / 2, 0 \leq b \leq \aleph_\Gamma$.

Let $\omega_0 = 0, \omega_b = \min \{ \varphi : \exists \varphi \times b \text{ d-disjunct binary matrix with weights all } \Gamma - 1 \}, 0 \leq b \leq \aleph_\Gamma$. We prove by contradiction that if $|\mathcal{C}_\Gamma| = b$, then $\mathcal{C}_\Gamma$ has 1-entries in at least $\omega_b$ rows not $R_i$.

Let $b = |\mathcal{C}_\Gamma|$. Suppose $\mathcal{C}_\Gamma$ has 1-entries in less than $\omega_b$ rows not $R_i$. Notice the submatrix $A'$ of $A$ from the columns of $\mathcal{C}_\Gamma$ is d-disjunct, since $A$ is d-disjunct. Notice that the rows without 1-entries for the columns of $\mathcal{C}$ correspond to 0-rows in $A'$ and, $R_i$ corresponds to a full-row in $A'$. Thus, the submatrix $A''$ with less than $\omega_b$ rows and $b$ columns formed by deleting these corresponding rows in $A'$ is d-disjunct, a contradiction, since $\omega_b$ is such that no such matrix exists. Thus, $\mathcal{C}_\Gamma$ has 1-entries in at least $\omega_b$ rows not $R_i$.

Thus $\mathcal{C}_\Gamma$ has 1-entries in at least max$\{\psi_b, \omega_b\}$ rows not $R_i$. Call these rows $\mathcal{R}_\Gamma$.

Notice that the rows of $\mathcal{R}_\gamma$ are distinct from the rows of $\mathcal{R}_\Gamma$. Since neither $\mathcal{R}_\gamma$ nor $\mathcal{R}_\Gamma$ contain $R_i$, the total number of rows of $A$ is greater than $\min_{a+b=\Gamma} \{ \alpha_a + \beta_b \}$.

Lemma 1.7.2. Let $A$ be an $d$-irreducible binary matrix with $t$ rows. Let $C_a, C_b, C_c$ be columns of $A$ such that, for each row of $A$, at least one of these three columns has a 1-entry. Let $x$ be the number of rows where $C_a$ has a 1-entry and $C_b$ has 0-entries. Let $y$ be similar for $C_b$ with $C_a$. Let $z$ be the number of rows where $C_a$ and $C_b$ both have 1-entries. Let $n = \min \{ b, x, y, z \}$.

If $A$ is 2-disjunct, then $A$ has at most $n + 3$ columns.

Proof. Let $A$ be an $d$-irreducible 2-disjunct binary matrix with $t$ columns, three columns such that their 3-sum has a 1-entry in each row of $A$. Without loss of generality, assume these are $C_1, C_2, C_3$. Notice that by Theorem 1.3.2 $\gamma \geq 3$. Let $x$ be the number of rows where $C_1$ has a 1-entry and $C_2, C_3$ have 0-entries. Let $y$ be similar for $C_2$. Let $z$ be the
number of rows where $C_1$ and $C_2$ both have 1-entries. Let $R_x, R_y, R_z$ be the rows that $x, y, z$ correspond to, respectively. Let $R_w$ be the collection of rows where $C_3$ has a 1-entry and $C_1, C_2$ have 0-entries. Notice that each other column of $A$ must have at least one 1-entry in each of $R_x, R_y, R_w$.

Fix some rows of $R_x, R_y$. Let $t$ be the total selected number of these rows. Let $C$ denote the columns not $C_1, C_2, C_3$ with 1-entries in all of these rows, 0-entries in all other rows corresponding to $x, y$. Notice that each column of $C$ has at most $\Gamma - t - 1$ 1-entries in $R_z$. Notice that if the 1-entries in $R_z$ for any column $C_a$ of $C$ contain the 1-entries in $R_z$ for another column $C_b$ of $C$, then $C_a \oplus C_3$ covers $C_b$, a contradiction, since $A$ is 2-disjunct. Thus, no column of $C$ has 1-entries in $R_z$ which contain the 1-entries in $R_z$ for another column of $C$. That is, the 1-entries in $R_z$ for the columns of $C$ form a Sperner family. So by Sperner’s Theorem, there are at most $\binom{\lceil \frac{\Gamma}{2} \rceil}{2}$ columns in $C$. Additionally, since there are at most $\Gamma - t - 1$ 1-entries in $R_z$ for any column in $C$, if $\Gamma - t - 1 < \lceil \frac{\Gamma}{2} \rceil$, that is, if $t \geq \Gamma - \lceil \frac{\Gamma}{2} \rceil$, then by the LYM Inequality there are at most $\binom{\lceil \frac{\Gamma}{2} \rceil}{2} - 1$ columns in $C$.

Thus, for the choice of any $i$ rows of $R_x$, $j$ rows of $R_y$, if $i + j \geq \Gamma - \lceil \frac{\Gamma}{2} \rceil$, then there are an associated $\binom{\lceil \frac{\Gamma}{2} \rceil}{2} - 1$ columns, and if $i + j \leq \Gamma - \lceil \frac{\Gamma}{2} \rceil - 1$, then there are an associated $\binom{\lceil \frac{\Gamma}{2} \rceil}{2}$ columns.

Notice that for any column, $C_c$ not $C_1, C_2, C_3$, since there is at least one 1-entry in $R_y, R_w$ there are at most $\Gamma - 2$ 1-entries in $R_x$. Additionally, there is at most $\Gamma - x - 1$ 1-entries in $R_x$, since if there are $x$ 1-entries, $C_c \oplus C_2$ covers $C_1$, a contradiction, since $A$ is 2-disjunct. Thus, for any column not $C_1, C_2, C_3$, there are at most $\min \{ \Gamma - 2, x - 1 \}$ 1-entries in $R_x$. By similar reasoning, there are at most $\min \{ \Gamma - 2, x - 1 \}$ 1-entries in $R_y$. This implies that there are at most

$$n = \sum_{i=1}^{\min \{ \Gamma - 2, x - 1 \}} \left( \sum_{j=1}^{\Gamma - \lceil \frac{\Gamma}{2} \rceil - i - 1} \binom{x}{i} \binom{y}{j} \binom{z}{\lceil \frac{\Gamma}{2} \rceil} + \sum_{j=\Gamma - \lceil \frac{\Gamma}{2} \rceil - i}^{\min \{ \Gamma - 2, y - 1 \}} \binom{x}{i} \binom{y}{j} \binom{z}{\lceil \frac{\Gamma - i - j - 1}{2} \rceil} \right)$$

$$= \sum_{i=1}^{\min \{ \Gamma - 2, x - 1 \}} \left( \binom{x}{i} \binom{z}{\lceil \frac{\Gamma}{2} \rceil} \sum_{j=1}^{\Gamma - \lceil \frac{\Gamma}{2} \rceil - i - 1} \binom{y}{j} + \sum_{j=\Gamma - \lceil \frac{\Gamma}{2} \rceil - i}^{\min \{ \Gamma - 2, y - 1 \}} \binom{y}{j} \binom{z}{\lceil \frac{\Gamma - i - j - 1}{2} \rceil} \right)$$

columns not $C_1, C_2, C_3$. Thus, $A$ has at most $n + 3$ columns. \hfill \Box

**Lemma 1.7.3.** Let $A$ be an $d$-irreducible binary matrix with $t$ rows. Let $C_a, C_b, C_c$ be columns of $A$ such that, for each row of $A$, at least one of these three columns has a 1-entry. Let $x$ be the number of rows where $C_a$ has a 1-entry and $C_b$ has 0-entries. Let $y$ be similar for $C_b$ with $C_a$. Let $z$ be the number of rows where $C_a$ and $C_b$ both have 1-entries.

---

26
Let 
\[ n = \min\{\Gamma - 2, x - 1\} \sum_{i=1} \left( \binom{x}{i} \left\lfloor \frac{y}{\frac{\Gamma}{2}} \right\rfloor^{\Gamma - \left\lfloor \frac{y}{\frac{\Gamma}{2}} \right\rfloor - i} \sum_{j=0}^{\min\{\Gamma - 3, z\}} \left( \binom{z}{j} \left( \frac{y}{\Gamma - j - 1} \right) \right) \right) \]

If \( A \) is 2-disjunct, then \( A \) has at most \( n + 3 \) columns.

Proof. This proof follows exactly as the proof for Lemma 1.7.2. However, we instead fix rows of \( R_x, R_z \).

For any column not \( C_1, C_2, C_3 \), there are at least zero (not one), 1-entries in \( R_z \). Since there is at least one 1-entry in \( R_x, R_y, R_w \), there is at most \( \Gamma - 3 \) 1-entries in \( R_z \). There is also at most \( z \) 1-entries in \( R_z \). Thus, for any column not \( C_1, C_2, C_3 \), there are at most \( \min\{\Gamma - 3, z\} \) 1-entries in \( R_z \). This leads to the different bounds from Lemma 1.7.2 on the inner summations in the expression for \( n \). \qed
Chapter 2

Constructing efficient decodable pooling matrices

2.1 On $d$-separable binary matrices with time optimal analysis algorithms

Combinatorial group testing is a well known problem which has seen substantial research during the last decade (see for example [insert citations here]). Suppose we have $N$ items, some of which are 'defective'. Suppose we can perform tests on subsets of these $N$ items which will tell us whether some defective exists in that subset or not. Our goal is to identify exactly which items are defective, and which are not. Naturally, we could test each item one by one, performing $N$ tests, but in practice there are much more efficient strategies for 'pooling' objects together to minimize the number of required tests.

Testing strategies can be represented as the binary incidence matrix where we place a 1 in the $i,j$ entry if item $j$ is contained in test $i$, and 0 otherwise. Being able to recover the indices of defective items has motivated the definition of $d$-separable matrices, which are precisely the matrices that can be decoded if there are at most $d$ defectives.

More precisely, a $t \times N$ binary matrix induces a function from the subsets $S$ of $[N]$ of size less than or equal to $d$ to $(0,1)^t$ by taking the boolean sum (note that some authors refer to this as union) of the columns corresponding to the elements of $S$. If $S$ is the set of at most $d$ defectives, then the image of $S$ under this map is the test result vector, i.e. the vector representing the outcomes of each test. If this function is injective, we say that the matrix is $d$-separable.

Although $d$-separable matrices always can be decoded, decoding such matrices is often intractable if the matrix is very large. The naive decoding algorithm would be to calculate every possible boolean sum of at most $d$ columns until we find the one that matches the test result vector, a process which requires time $O(N^d)$. This has led to the study of $d$-disjunct matrices, which are $d$-separable but have a decoding algorithm which takes time $O(tN)$. 

28
For very large matrices, using even the \( d \)-disjunct decoding algorithm may be intractable. The best we can possibly do is recover the defective indices by looking at the results of each test, doing a constant amount of work at each step, regardless of the number of columns. Such an algorithm runs in time \( O(t) \) and is referred to as a time optimal analysis algorithm.

Eppstein, Goodrich, and Hirschberg [5] constructed a \( 3 \)-separable matrix with time-optimal analysis algorithm with dimensions \( 2^{2\left(q,\frac{q}{2}\right)} \times 2^q \) for any positive integer \( q \). We extend their construction, techniques, and proofs to construct a \( 2^{d-1}\left(q,\frac{q}{d-1}\right) \times 2^q \) \( d \)-separable matrix for each \( d \geq 3 \) and every positive integer \( q \) together with a time optimal analysis algorithm.

### 2.1.1 The Construction

Suppose there are \( N = 2^q \) items where \( q \in \mathbb{N} \). We will express an item index \( X \in \{0, 1, \ldots, N-1\} \) in binary notation so that \( X = X_{q-1}X_{q-2} \ldots X_0 \) where each \( X_p \in \{0, 1\} \). For this paper we will let \( p_i \in \{0, 1, \ldots, q-1\} \), \( v_i \in \{0, 1\} \) be radix positions and binary values respectively, for each \( i \). For convenience we define \( I_k \) to be the indexing set \( \{1, 2, \ldots, k\} \). We write \( \oplus \) to represent the Boolean sum.

Let \( M_d \) be the \( 2^{d-1}\left(q,\frac{q}{d-1}\right) \times 2^q \) matrix formed by associating each row with an unordered collection of \( d-1 \) distinct position values \( p_1, p_2, p_{d-1} \) together with binary values for each position \( v_1, v_2, \ldots, v_{d-1} \). We denote a row index by the set \( \{(p_i, v_i)\}_{i \in I_{d-1}} \). We define \( M_d[\{(p_i, v_i)\}_{i \in I_{d-1}}, X] = 1 \) if for each \( i \), \( X_{p_i} = v_i \), and 0 otherwise.

We remark that each column of \( M_d \) will have weight \( \left(q,\frac{q}{d-1}\right) \), since, given a column index \( X \), for every set of distinct positions \( \{p_1, \ldots, p_{d-1}\} \), there is exactly one tuple \( (v_1, \ldots, v_{d-1}) \) such that \( M_d[\{(p_i, v_i)\}_{i \in I_{d-1}}, X] = 1 \).

The weight of each row \( \{(p_i, v_i)\}_{i \in I_{d-1}} \) of \( M_d \) is \( 2^{q-d+1} \), which is the number of indices \( X \in \{0, \ldots, 2^q\} \) such that \( X_{p_i} = v_i \) for each \( i \).

We define the following tests for use in our decoding algorithm:

- \( \text{test}_{M_d}[\{(p_i, v_i)\}_{i \in I_{d-1}}] \) is 1 if the test result for that row is 1, and 0 otherwise. Equivalently, this test is positive if there is some defective \( D \) such that \( D_{p_i} = v_i \) for each \( i \).

- \( \text{test}_{1,M_d}(p_1, \ldots, p_{d-1}) \) is the number of distinct ordered \( (d-1) \)-tuples \( (v_1, v_2, \ldots, v_{d-1}) \) of values present among defectives at positions \( p_1, \ldots, p_{d-1} \). We calculate it as follows:

\[
\text{test}_{1,M_d}(p_1, \ldots, p_{d-1}) = \sum_{(v_1, \ldots, v_{d-1}) \in \{0, 1\}^{d-1}} \text{test}_{M_d}[\{(p_i, v_i)\}_{i \in I_{d-1}}]
\]
2.1.2 Determining the number of defectives

For $P \subseteq \{0, 1, \ldots, q-1\}$ and item $X \in \{0, 1\}^q$, define $X[P] \subseteq \{0, 1\}^{|P|}$ by choosing exactly those bits from $X$ corresponding to the positions in $P$.

Lemma 2.1.1. Given $\{D^1, \ldots, D^k\}$, a set of $k$ distinct items, there is some $P \subseteq \{0, 1, \ldots, q-1\}$ with $|P| = k-1$ such that $\{D^1[P], D^2[P], \ldots, D^k[P]\}$ are all distinct.

Proof. We induct on $k$. Base case $k = 1$ is trivial, so assume $k > 1$. By induction there exists $P'$ such that $D^1[P'], D^2[P'], \ldots, D^{k-1}[P']$ are distinct. Let $T = \{D^1[P'], D^2[P'], \ldots, D^{k-1}[P']\}$. If $D^k[P'] \notin T$, we may arbitrarily extend $P'$ to $P$. If not, there is exactly one defective $D^i \in T$ such that $D^i[P'] = D^k[P']$. Since $D^i$ and $D^k$ are distinct, there must be some position $p \notin P$ such that $D^i([p]) \neq D^k([p])$, so we take $P = P' \cup \{p\}$. \hfill $\square$

Proposition 2.1.1. If there are at most $d$ defectives, the exact number of defectives is given by

$$\max_{p_1, \ldots, p_{d-1}} \ (\text{test}_{1,M_d}(p_1, \ldots, p_{d-1}))$$

Proof. Let $d' \leq d$ be the number of defectives present. For any choice of distinct $p_1, \ldots, p_{d-1}$, $\text{test}_{1,M_d}(p_1, \ldots, p_{d-1}) \leq d'$, since each defective can contribute at most one ordered $(d-1)$-tuple of values at the positions $p_1, \ldots, p_{d-1}$. Thus we need only show that there exist positions $p_1, \ldots, p_{d-1}$ such that $\text{test}_{1,M_d}(p_1, \ldots, p_{d-1})$ attains $d'$. Suppose the defectives are $D^1, \ldots, D^{d'}$. By Lemma 2.1.1, there exists a set of positions $P' = \{p_1, \ldots, p_{d'-1}\}$ such that $D^1[P'], \ldots, D^{d'}[P']$ are all distinct. Since $|P'| = d' - 1 \leq d - 1$, we may arbitrarily extend it to some $P = \{p_1, \ldots, p_{d-1}\}$, and then $\text{test}_{1,M_d}(p_1, \ldots, p_{d-1}) = d'$.

$\square$

2.1.3 Recovering the Defective Values

To identify the defectives, it is sufficient to determine the binary values of each defective index at each radix position. We say that a set of position-value pairs $\{(p_1, v_1), \ldots, (p_k, v_k)\}$ distinguishes a defective $D$ if it is the only defective such that for each $i$, $D_{p_i} = v_i$. We also say that a set $P = \{p_1, \ldots, p_k\}$ of positions distinguishes $D$ if there exist such values $v_1, \ldots, v_k$ that $\{(p_1, v_1), \ldots, (p_k, v_k)\}$ distinguishes $D$.

The general strategy will be to find for each defective $D$ a set of position-value pairs $S$ with $|S| = d - 2$ that distinguishes $D$. Then letting $p$ be arbitrary, we calculate $\text{test}_{M_d}(S \cup \{(p, 1)\})$. If this is 1, we may conclude that the defective has value 1 at position $p$, and otherwise it must have value 0 at $p$.

We remark that it is sufficient to find $S'$ that distinguishes $D$ with $|S'| \leq d - 2$ since we can arbitrarily extend $S'$ to some $S$ with $|S| = d - 1$ and simply cycle through the $2^{d-|S|}$ possibilities of values for the additional positions until $\text{test}_{M_d}(S) = 1$, thereby identifying the values at each position in $S$. We may then use those position-value pairs to efficiently find the others as above.
When it is not possible to find positions and values that distinguish each defective, we will make use of the following lemma:

**Lemma 2.1.2.** Suppose there are \(d\) defectives. If there is a set of positions \(\{p_1, \ldots, p_{d-2}\}\) such that for each \(i \in I_{d-2}\), position \(p_i\) distinguishes defective \(D^i\), and there is a defective \(E\) whose digits are known, then the digits of the last defective \(F\) can be computed.

**Proof.** For each \(i \in I_{d-2}\) let \(v_i = D^i_{p_i}\). Since each of these defectives is distinguished by \(p_i\), we may conclude that \(F_{p_i} = v_i\) for each \(i \in I_{d-2}\). For any other position \(p\), compute \(E_p = v\). If \(test_{M_d}(\{(p_i, v_i)\}_{i \in I_{d-2}} \cup \{p, v\}) = 1\) we may conclude that \(F_p = v\) since no other defectives are present in this test. However, if it is 0, we may conclude that \(F_p = v\) since otherwise it would have caused the test to be positive.

We now have the tools to prove the following:

**Theorem 2.1.1.** For \(d \geq 3\), a \(2^{d-1} \binom{q}{d-1} \times 2^q\) binary \(d\)-separable matrix with time-optimal analysis algorithm can be constructed for each positive integer \(q\).

**Proof.** Suppose that \(d = 3\). Then, by Proposition 2.1.1 we may find positions \(p_1, p_2\) such that \(test_{M_3}(p_1, p_2)\) is the number of defectives present.

Case 1: Suppose there is exactly one defective. If we pick any position \(p\), clearly the defective is distinguished by \(p\), as it is the only defective.

Case 2: Suppose that there are exactly two defectives. Since \(test_{M_3}(p_1, p_2) = 2\), the defectives must not agree in at least one of the positions, say \(p\). Then both defectives are distinguished by \((p, 0)\) and \((p, 1)\), respectively.

Case 3: Suppose that there are exactly three defectives, \(D^1, D^2, D^3\). All three defectives cannot have the same value at the same position, else the other position would have to distinguish all three of them. Say then that \(D^1\) is distinguished by \(p_1\). \(D^1\) cannot then be distinguished by \(p_2\), else \(D^2\) and \(D^3\) would agree at both positions, so say \(D^2\) is distinguished by \(p_2\). Then the digits of \(D^2\) can be determined. By Lemma 2.1.2, we may then determine the digits of \(D^3\).

Since only a constant amount of work was required at each step, \(M_3\) has a time-optimal analysis algorithm. We give credit to Eppstein, et al. [5] for the construction of our base case.

Now, suppose that \(M_{d-1}\) has a time-optimal analysis algorithm. We note that

\[
\text{test}_{M_{d-1}}(\{(p_i, v_i)\}_{i \in I_{d-2}}) = \bigoplus_{v_{d-1} \in \{0, 1\}} \text{test}_{M_d}(\{(p_i, v_i)\}_{i \in I_{d-1}})
\]
Hence, if there are \( d' < d \) defectives present, by our inductive hypothesis we may recover the values of each defective in time \( O(t) \) by computing the tests of \( M_{\text{max}}(3,d') \). Thus we need only consider the case when there are exactly \( d \) defectives present.

Suppose that there are exactly \( d \) defectives, \( D^1, \ldots, D^d \). By Proposition 2.1.1 we may find positions \( p_1, \ldots, p_{d-1} \) such that \( \text{test}_1 M_d(p_1, \ldots, p_{d-1}) = d \). Let \( P = \{ p_1, p_2, \ldots, p_{d-1} \} \) and for each \( i, P_i = \{p_1, p_2, \ldots, p_{d-1}\} \setminus \{ p_i \} \) one of its subsets of order \( p - 2 \). If for each defective one of the \( P_i \) distinguishes it, we are done, so assume that there is some defective, say \( D^d \) that cannot be distinguished by any of the \( P_i \). Then for each \( P_i \), \( D^d \) must agree with at least one other defective. However it must agree with exactly one since otherwise one position would have to distinguish three defectives. Furthermore, \( D^d \) cannot agree with the same defective on different \( P_i \) since then they would agree on all of \( P \), contradicting the fact that \( \text{test}_1 M_d(p_1, \ldots, p_{d-1}) = d \). Then \( D^d \) agrees with each of the other defectives on exactly one of the \( P_i \). Without loss of generality say \( D^d \) agrees with \( D^i \) on \( P_i \). Then \( D^d_1 = D^i_1 = D^d_2 = \cdots = D^d_{i-1} \). But \( D^d_1 \neq D^d_i \) since otherwise \( D^1 \) and \( D^d \) would agree on all of \( P \). Thus, \( D^1 \) is distinguished by \( p_i \). Similarly, \( D^i \) is distinguished by \( p_i \) for each \( i \in I_{d-1} \). Applying Lemma 2.1.2, we may also compute the values of \( D^d \) at each position.

Only a constant amount of work was required for each step, so \( M_d \) has a time-optimal analysis algorithm.

\[ \square \]

### 2.1.4 Runtime Analysis

We now give a more detailed analysis of the runtime of this algorithm. Pick \( q, d \in \mathbb{N} \), let \( N = 2^q \) and \( t = 2^{d-1} \binom{q}{d-1} \), and \( M_d \) the \( t \times N \) matrix as described above.

The number of operations required to determine the number of defectives is at most \( 2^{d-1} \binom{q}{d-1} + \binom{q}{d-1} \), as computing \( \text{test}_1 p_1, \ldots, p_{d-1} \) for every possible choice of \( p_1, \ldots, p_{d-1} \) requires us to compute \( \binom{q}{d-1} \) sums of \( 2^{d-1} \) elements each, and then we must compare each sum to the previously found maximum. If there are exactly \( d \) defectives, we may require even fewer operations as we may stop once one of the sums is \( d \).

Once the a maximum value is found, we also immediately obtain the witnessing set \( S = \{ p_1, \ldots, p_{d-1} \} \) such that \( \text{test}_1 p_1, \ldots, p_{d-1} \) requires us to compute \( \binom{q}{d-1} \) sums of \( 2^{d-1} \) elements each, and then we must compare each to the previously found maximum. If there are exactly \( d \) defectives, we may require even fewer operations as we may stop once one of the sums is \( d \).

To find the sets \( P_i \subset S, i \in I_d \), with \( |P_i| = d - 2 \) where \( P_i \) distinguishes defective \( D^i \), requires at most \( d(d-1)^2 \) operations. For example, if we look at each of the \( d - 1 \) subsets \( P_i \subset S \) with \( |P_i| = d - 2 \), we simply look at the values of each of the \( d \) defectives at \( P_i \), compare it to each of the other \( d - 1 \) defectives, and thus determine if it is distinguished from the others at \( P_i \). We note that this bound could be improved by a more efficient algorithm, but since \( d \) is generally small, \( d(d-1)^2 \) is a sufficient bound for our needs.

Once a set \( P \) with \( |P| = d - 2 \) is found that distinguishes a defective, \( q-d+1 \) operations are required to determine the remaining digits of that defective, since all that is required
is looking up the value of a single test for each digit.

Thus far we have described the runtime of determining up to \( d - 1 \) defectives. If all \( d \) are present, computing the last one requires \( 2(q - d + 2) \) as all we must do for each unknown digit is look up the value of one of the previously computed defectives, and then look up the value of a single test.

Letting \( \log \) represent the base-2 logarithm, analysis of the runtime analysis of each stage of our algorithm shows that the overall runtime is bounded above by

\[
2^{d-1} \left( \frac{q}{d-1} \right) + \left( \frac{q}{d-1} \right) + d + d(d-1)^2 + (d-1)(q-d+1) + 2(q-d+2)
\]

\[
\leq 2^{d-1} \left( \frac{q}{d-1} \right) + \left( \frac{q}{d-1} \right) + d^3 + (d+1)q
\]

\[
\leq \left( 2^{d-1} + 1 \right) q^{d-1} + (d+1)q + d^3
\]

\[
= (2^{d-1} + 1) (\log N)^{d-1} + (d+1) \log (N) + d^3
\]

### 2.1.5 Comparison with other Matrices

\( d \)-disjunct matrices have been studied extensively due to the fact that they can be decoded in \( \Theta(Nt) \), but exist with desirable dimensions, in that they can be found such that the number of columns is exponential in the number of rows. Table 1 lists the number of tests required by our construction for various values of \( d \) and \( N \) compared with \( t_d \), one of the best known upper bounds on the minimum number of rows of a \( d \)-disjunct matrix due to Cheng, et al [2]. We note that our construction requires many more tests than a \( d \)-disjunct matrix with the same number of items, though we also remark that ours is time-optimal.

<table>
<thead>
<tr>
<th>( N )</th>
<th>( 10^3 )</th>
<th>( 10^6 )</th>
<th>( 10^{10} )</th>
<th>( 10^{20} )</th>
<th>( 10^{30} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( M_3 )</td>
<td>180</td>
<td>760</td>
<td>2244</td>
<td>8844</td>
<td>19,800</td>
</tr>
<tr>
<td>( t_3 )</td>
<td>192</td>
<td>383</td>
<td>639</td>
<td>1277</td>
<td>1915</td>
</tr>
<tr>
<td>( M_4 )</td>
<td>960</td>
<td>9120</td>
<td>47,872</td>
<td>383,240</td>
<td>1,293,600</td>
</tr>
<tr>
<td>( t_4 )</td>
<td>312</td>
<td>624</td>
<td>1039</td>
<td>2077</td>
<td>3116</td>
</tr>
<tr>
<td>( M_5 )</td>
<td>3360</td>
<td>77,520</td>
<td>742,016</td>
<td>12,263,680</td>
<td>62,739,600</td>
</tr>
<tr>
<td>( t_5 )</td>
<td>461</td>
<td>921</td>
<td>1535</td>
<td>3069</td>
<td>4604</td>
</tr>
</tbody>
</table>

### 2.2 A new construction of \( d \)-disjunct matrices with \( K(d+1)q^t \) rows

**Definition.** Let \( C \) be a column of a binary matrix. We call the set of row indices where \( C \) has a 1-entry the **support** of \( C \).
**Definition.** Let $A$ be a binary matrix. We call a matrix whose columns are the supports of the columns of $A$ the *helper matrix* of $A$.

**Construction 2.2.1.** Let $d \geq 2$ be a natural number. Let $q \geq d + 1$ be prime. Let $t$ be a natural number. We construct a $(d+1) \times q^{2t}$ matrix whose pairs of columns share at most one entry.

Let $M$ be the 0-indexed $(d+1) \times q^{2t}$ matrix whose columns are constructed as:

$$C_i = \begin{bmatrix}
\left( 0 \left\lfloor \frac{i}{q} \right\rfloor + i \right) \mod q^t \\
\left( 1 \left\lfloor \frac{i}{q} \right\rfloor + i \right) \mod q^t + q^t \\
\vdots \\
\left( d \left\lfloor \frac{i}{q} \right\rfloor + i \right) \mod q^t + dq^t 
\end{bmatrix}$$

where $0 \leq i < q^t$ and $d = \frac{q^t - 1}{q - 1}$.

Any two columns of $M$ share at most one entry.

**Proof.** We prove by contradiction.

Suppose two columns, $C_j, C_k$ of $M$ share at least two entries. Notice that each entry in $R_0$ of $M$ is in $\{0, 1, 2, \ldots, q^t - 1\}$, each entry in $R_1$ of $M$ is in $\{q^t, q^t + 1, \ldots, 2q^t - 1\}$, \ldots, each entry in $R_d$ of $M$ is in $\{dq^t, dq^t + 1, \ldots, (d+1)q^t - 1\}$. Thus, any shared entry of $C_j, C_k$ must be in the same row of $M$. So $C_j, C_k$ have the same entry in at least two rows. So there are elements $a, b$ in both column $C_j$ and column $C_k$ such that

$$a = \left( m \left\lfloor \frac{j}{q} \right\rfloor + j \right) \mod q^t + mq^t = \left( m \left\lfloor \frac{k}{q} \right\rfloor + k \right) \mod q^t + mq^t \text{ and } b = \left( n \left\lfloor \frac{j}{q} \right\rfloor + j \right) \mod q^t + nq^t = \left( n \left\lfloor \frac{k}{q} \right\rfloor + k \right) \mod q^t + nq^t$$

for $0 \leq m, n \leq d$ with $m \neq n$, where $m, n$ indicate which rows $a, b$ are found in, and where $0 \leq j, k < q^t$ with $j \neq k$. Thus,

$$m \left\lfloor \frac{j}{q} \right\rfloor + j \equiv m \left\lfloor \frac{k}{q} \right\rfloor + k \pmod{q^t} \text{ and } n \left\lfloor \frac{j}{q} \right\rfloor + j \equiv n \left\lfloor \frac{k}{q} \right\rfloor + k \pmod{q^t}$$

$$\Rightarrow m \left( \left\lfloor \frac{j}{q} \right\rfloor - \left\lfloor \frac{k}{q} \right\rfloor \right) \equiv k - j \pmod{q^t} \text{ and } n \left( \left\lfloor \frac{j}{q} \right\rfloor - \left\lfloor \frac{k}{q} \right\rfloor \right) \equiv k - j \pmod{q^t}$$

$$\Rightarrow m \left( \left\lfloor \frac{j}{q} \right\rfloor - \left\lfloor \frac{k}{q} \right\rfloor \right) \equiv n \left( \left\lfloor \frac{j}{q} \right\rfloor - \left\lfloor \frac{k}{q} \right\rfloor \right) \pmod{q^t}$$

$$\Rightarrow (m - n) \left( \left\lfloor \frac{j}{q} \right\rfloor - \left\lfloor \frac{k}{q} \right\rfloor \right) \equiv 0 \pmod{q^t}$$

Since $m \neq n$, $0 \leq m, n < q < q^t, m - n \neq 0 \pmod{q^t}$, and $q$ prime, $\left\lfloor \frac{j}{q} \right\rfloor - \left\lfloor \frac{k}{q} \right\rfloor \equiv 0 \pmod{q^t}$. Thus,

$$\left\lfloor \frac{j}{q} \right\rfloor \equiv \left\lfloor \frac{k}{q} \right\rfloor \pmod{q^t}$$

$$\Rightarrow \left\lfloor \frac{j}{q} \right\rfloor = \left\lfloor \frac{k}{q} \right\rfloor$$

since $j, k < q^{2t}$ and $|j - k| < q^t$ (since $|j - k| \geq q^t$ would imply $\left\lfloor \frac{j}{q} \right\rfloor \neq \left\lfloor \frac{k}{q} \right\rfloor$).

$$\Rightarrow j \equiv k \pmod{q^t}$$

$$\Rightarrow \left( m \left\lfloor \frac{j}{q} \right\rfloor + j \right) = \left( m \left\lfloor \frac{k}{q} \right\rfloor + k \right) \pmod{q^t}$$

$$\Rightarrow j = k + iq^t \text{ for some } i \in \mathbb{Z}$$

34
\[ \Rightarrow j = k, \text{ since } |j - k| < q', \text{ a contradiction. Thus, any two columns of } M \text{ share at most one entry.} \]

**Claim.** Let \( C \) be a column vector of length \( \ell \). If each entry in \( C \) is distinct, then any column for which \( C \) is the support will have weight \( \ell \).

**Construction 2.2.2.** Let \( M \) be the \((d + 1) \times q^{2t}\) matrix constructed as in Construction 2.2.1. Let \( \mu = \left\lfloor \frac{q^t}{d+1} \right\rfloor \). Let \( \tilde{M} \) be the \((d + 1) \times (\mu(d + 1))\) matrix whose columns are constructed as:

\[
C_l = \begin{bmatrix}
    lq + \left( (d + 1) \left\lfloor \frac{l}{d+1} \right\rfloor \right) \\
    lq + 1 + \left( (d + 1) \left\lfloor \frac{l}{d+1} \right\rfloor \right) \\
    \vdots \\
    lq + d + \left( (d + 1) \left\lfloor \frac{l}{d+1} \right\rfloor \right)
\end{bmatrix} \text{ where } 0 \leq l < \mu (d + 1)
\]

Let \( \tilde{M} \) be the \((d + 1) \times (q^{2t} + \mu(d + 1))\) matrix formed by appending the \( q^{2t} \) columns of \( M \) to the \( \mu(d + 1) \) columns of \( \tilde{M} \). Let \( \hat{M} \) be the helper matrix of a 0-indexed \((d + 1)q^t \times (q^{2t} + \mu(d + 1))\) binary matrix \( A \).

\( A \) is \( d \)-disjunct and has columns of weight \( d + 1 \).

**Proof.** Recall that any two columns of \( M \) share at most one entry. Notice that column \( C_m \) of \( \tilde{M} \) contains \( d + 1 \) of the \( q \) elements from row \( R_m \) of \( M \), whose elements are distinct from the other rows of \( M \). Thus, none of the columns of \( \tilde{M} \) share any entries and each of the \( q^{2t} \) columns from \( M \) shares at most one entry with any of the columns of \( \tilde{M} \). Thus, each column in \( \tilde{M} \) will share at most one entry with any other column of \( \tilde{M} \).

Since no two columns in \( \tilde{M} \) have more than one entry in common, any column in \( A \) will have a product of weight at most \( d \) with any \( d \)-sum of any of the other columns in \( A \). Since every column in \( A \) has weight \( d + 1 \), no column in \( A \) can be covered by the \( d \)-sum of any of the other columns. Thus, \( A \) is \( d \)-disjunct.

Notice that for any column \( C_n \) of \( \tilde{M} \), the entries of \( C_n \) are distinct. So each column of \( A \) has weight \( d + 1 \).

**Corollary 2.2.1.** Let \( M \) be the matrix constructed as in Construction 2.2.1. If \( M \) is a helper matrix for a binary matrix \( A \), then \( A \) is \( d \)-disjunct and has columns of weight \( d + 1 \).

**Lemma 2.2.1.** Let \( t \) be a natural number. For odd \( q \), \( q^{2t} \equiv 1 \) or \( 3 \pmod{6} \).

**Proof.** We prove by induction.

Base case. Let \( q \) be odd. We prove \( q^2 \equiv 1 \) or \( 3 \pmod{6} \).

Notice \( q = 2m + 1 \) for some \( m \in \mathbb{Z} \). Thus, \( q^2 = 4m^2 + 4m + 1 \). Assume, by way of contradiction, that \( 4m^2 + 4m + 1 \equiv 5 \pmod{6} \). Thus,
$4m^2 + 4m \equiv 4 \pmod{6}$

$\implies 4(m^2 + m - 1) \equiv 0 \pmod{6}$

$\implies (m^2 + m - 1) \in 3\mathbb{Z}$

$\implies m^2 + m - 1 = 3n$ for some $n \in \mathbb{Z}$

$\implies m^2 + m - 1 \equiv 0 \pmod{3}$, a contradiction, since $0^2 + 0 - 1 \equiv 2 \pmod{3}$, $1^2 + 1 - 1 \equiv 1 \pmod{3}$, and $2^2 + 2 - 1 \equiv 2 \pmod{3}$. Thus, $q^2 \not\equiv 5 \pmod{6}$. Since $q^2$ is odd, $q^2 \equiv 1$ or $3 \pmod{6}$.

Inductive case. Assume $q^{2t} \equiv 1$ or $3 \pmod{6}$, where $t$ is a natural number. We prove

$q^{2(t+1)} \equiv 1$ or $3 \pmod{6}$ by cases.

Case 1: Suppose $q^{2t} \equiv 1 \pmod{6}$. Then $q^{2(t+1)} \equiv q^{2t}q^2 \equiv \begin{cases} 1(1) \equiv 1 \pmod{6} & \text{if } q^2 \equiv 1 \pmod{6} \\ 1(3) \equiv 3 \pmod{6} & \text{if } q^2 \equiv 3 \pmod{6} \end{cases}$

Case 2: Suppose $q^{2t} \equiv 3 \pmod{6}$. Then $q^{2(t+1)} \equiv q^{2t}q^2 \equiv \begin{cases} 3(1) \equiv 3 \pmod{6} & \text{if } q^2 \equiv 1 \pmod{6} \\ 3(3) \equiv 3 \pmod{6} & \text{if } q^2 \equiv 3 \pmod{6} \end{cases}$

The following fact is well known, and can be found in [3].

**Fact 1.** Let $v$ be a natural number. There exists a Steiner Triple System from a set of size $v$ if and only if $v \equiv 1$ or $3 \pmod{6}$.

**Lemma 2.2.2.** Let $q \geq 3$ be odd. If $t$ is a natural number, then there exists a $q^{2t} \times \frac{q^{4t} - q^{2t}}{6}$ 2-disjunct binary matrix whose columns are all of weight 3. This matrix may be formed from the incidence matrix of a Steiner Triple System from a set of size $q^{2t}$.

**Proof.** Let $q \geq 3$ be odd. Let $t$ be a natural number. By Lemma 2.2.1 $q^{2t} \equiv 1$ or $3 \pmod{6}$. By Fact 1, there exists a Steiner Triple System from a set of size $q^{2t}$. Notice that the incidence matrix of this Steiner Triple System will be a 2-disjunct binary matrix $A$ with $q^{2t}$ rows such that every pair of rows has a 2-product of weight 1 and all columns are weight 3. Thus, $A$ must have a uniform row weight of $\frac{\text{number of rows} - 1}{\text{column weight}} = \frac{q^{2t} - 1}{2}$.

Notice $\frac{(\text{number of rows})(\text{row weight})}{\text{column weight}} = \frac{(q^{2t})(\frac{q^{2t} - 1}{2})}{q^{2t} - q^{2t}} = \frac{q^{4t} - q^{2t}}{6} = (\text{number of columns})$. $lacksquare$

**Proposition 2.2.1.** Let $q \geq 3$ be prime. If $t$ is a natural number, then there exists a $3q^{2t} \times \left( q^{4t} + \frac{(q^{4t} - q^{2t})}{2} \right)$ 2-disjunct binary matrix whose columns are of weight 3.

**Proof.** Let $q \geq 3$ be prime. Let $t$ be a natural number. Let $M$ be the $3 \times q^{4t}$ constructed as in Construction 2.2.1. Let $M$ be the helper matrix for a $3q^{2t} \times q^{4t}$ binary matrix $A$. By Fact 1, since each row of $M$ contains $q^{2t}$ different elements and none of these elements are in any other row of $M$, there exists a Steiner Triple System from the elements of each row of $M$. Let $S_1, S_2, S_3$ be the three matrices representing these three Steiner Triple Systems. Notice from the proof of Lemma 2.2.2, $S_1, S_2, S_3$ are of dimensions $3 \times \frac{q^{4t} - q^{2t}}{6}$. Let $M' \subseteq \mathbb{Z}$ be the matrix formed by appending the $3 \left( \frac{q^{4t} - q^{2t}}{6} \right) = \frac{q^{4t} - q^{2t}}{2}$ columns from $S_1, S_2, S_3$ to
Notice that none of the columns of $M^\subseteq$ will share more than one entry with any other column of $M^\subseteq$.

Let $M^\subseteq$ be the helper matrix of a $3q^{2t} \times \left( q^{4t} + \frac{q^{4t} - q^{2t}}{2} \right)$ 0-indexed binary matrix $A$. Notice that each column of $A$ will be of weight 3. Since no two columns of $M^\subseteq$ share more than one entry, any column in $A$ will have a product of weight at most 2 with the 2-sum of any two of the other columns of $A$. Thus, no column in $A$ is covered by the 2-sum of any of the other columns of $A$. Thus, $A$ is 2-disjunct.

Notice that for any column $C_j$ of $M^\subseteq$, the entries of $C_j$ are distinct. So each column of $A$ has weight 3.

**Lemma 2.2.3.** If $\rho \geq 2$ for some $d$-disjunct binary matrix $A$, no column of $A$ can have all but one of its entries covered by any sum of $d - 1$ columns.

**Proof.** We prove by contradiction. Let’s say a column $C$ in $A$ had all but one of its entries covered by a sum of $d - 1$ columns in $A$. Let the row that contains this entry be row $R$. Since $\rho \geq 2$, there must be a column other than $C$ that has a 1-entry in row $R$. Therefore, a sum of $d$ columns can cover $C$, which implies a contradiction. \(\square\)

**Theorem 2.2.1.** Given a $d$-disjunct $(n,q,k)$ Reed-Solomon construction $M$ that used an identity matrix for its inner matrix, choose a $d$-disjunct $r \times c$ matrix $A$ with $r = q$. If $\gamma(A) \geq d + 1$ and $\rho(A) \geq 2$, then $n \times c$ columns can be concatenated with the original Reed-Solomon matrix to form a new matrix that will be $d$-disjunct.

**Proof.** Let $M$ be a $d$-disjunct binary matrix created by an $(n,q,k)$ Reed Solomon construction that used a $q \times q$ identity matrix for it’s inner matrix. Now separate $M$ into $n$ block matrices where $M_1$ is the first set of $q$ rows of $M$, $M_2$ is the second set of $q$ rows of $M$, and so on, so that

$$M = \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_n \end{bmatrix}$$

Each of the columns in these individual block matrices will have column weights of one, since each block matrix refers to a specific row of the Reed-Solomon codeword matrix used to construct it and an identity matrix was used for the inner matrix for the construction. Now, choose a $d$-disjunct $r \times c$ matrix $A$ with $r = q$ such that $\gamma(A) \geq d + 1$ and $\rho(A) \geq 2$. Since $A$ is $d$-disjunct, it’s clear that the matrix

$$\bar{A} = \begin{bmatrix} A & 0 & \ldots & 0 \\ 0 & A & \ldots & 0 \\ \vdots \\ 0 & 0 & \ldots & A \end{bmatrix}$$
will be \(d\)-disjunct as well where 0 represents an \(r \times c\) zero matrix. Now we will show that \(M^* = [M \ A]\) is also \(d\)-disjunct. Note that \(M\) and \(A\) have the same number of rows \((nq)\) by nature of construction.

Since \(\rho(A) = 2\), Lemma 2.2.3 implies that none of the columns of \(A\) are covered by a \(d\)-sum of any of the columns of \(M^*\), since each column of \(M\) can only cover one 1-entry of \(A\). Each column of \(A\) can only cover at most one 1-entry from any column in \(M\), and each row in every section \(M_i\) (where \(1 \leq i \leq n\)) has a row weight of at least two by the nature of the Reed-Solomon construction. Therefore, no column that wasn’t already covered in \(M\) can be covered in \(M^*\), and \(M^*\) is \(d\)-disjunct. By the nature of the construction, \(A\) has dimensions \(nr \times nc\), so \(M^*\) has dimensions \(nq \times q^k + nc\) (since \(r = q\)) and is \(d\)-disjunct. \(\square\)

Note that for constructing matrices with the Reed-Solomon process for a biological application, using an identity matrix for the inner matrix is typically necessary to maintain a low enough column weight. The above method of column concatenation will enable us to improve best known dimensions for lower column weights pretty easily.

### 2.3 Fast Decoding Using Reed Solomon Matrices

We make use of the celebrated Reed Solomon concatenation technique to describe a fast decoding algorithm for a group testing regime.

First, we define the notion of separability and disjunctness for \(q\)-ary matrices. A matrix is said to be \(q\)-ary if its entries are among some set \(S\) with \(|S| = q\). By a convenient abuse of notation we associate each entry value with the singleton set containing that value. We define the union of two columns by taking the union entry by entry, i.e.

\[
\begin{bmatrix}
a_1 \\
\vdots \\
a_k \\
b_1 \\
\vdots \\
b_k \\
\end{bmatrix} \cup 
\begin{bmatrix}
b_1 \\
\vdots \\
b_k \\
\end{bmatrix} = 
\begin{bmatrix}
\{a_1\} \\
\vdots \\
\{a_k\} \\
\{b_1\} \\
\vdots \\
\{b_k\} \\
\end{bmatrix} \cup 
\begin{bmatrix}
\{a_1\} \cup \{b_1\} \\
\vdots \\
\{a_k\} \cup \{b_k\} \\
\end{bmatrix}
\]

Similarly we define the finite union of columns. For two vectors \(v_1, v_2\), we say \(v_1 \subseteq v_2\) if each entry of \(v_1\) is a subset of the corresponding entry of \(v_2\).

Equivalently we may view \(q\)-ary matrices as the subset of \(M(\mathcal{P}(S))\) whose matrices have singletons as entries. Note that this coincides with our definition of binary matrices if for a binary matrix we associate 0 with the empty set and 1 with \(\{1\}\).

With these definitions, the notions of \(d\)-disjunctness and of \(d\)-separability are precisely the same as in the binary case, namely, a matrix is said to be \(d\)-separable if and only if the union of \(d\) columns (not necessarily distinct) is equal to no other union of \(d\) columns. A
matrix is $d$-disjunct if and only if the union of any $d$ columns does not contain any other column.

We now turn our attention to Reed Solomon codes. An $(n, q, k)$ Reed Solomon codeword matrix $M$ is an $n \times q^k$-ary matrix where $k < n \leq q$, produced via the following construction: Associate to each column a distinct polynomial $p(x) \in GF(q)[x]$ of degree less than $k$ and to each row a distinct field element $a \in GF(q)$. The entries $M_{a,p(x)}$ are given by the evaluation of $p(x)$ at $a$, i.e., $M_{a,p(x)} = p(a)$.

We also define the concatenation operation on a Reed Solomon outer matrix and a binary inner matrix. Let $M_{\text{out}}$ be an $n \times q^k$-ary Reed Solomon matrix whose entries are from some set $S$ with $|S| = q$ and let $M_{\text{in}}$ be a $t \times q$ binary matrix. Define a bijection $\phi$ from $S$ to the set of columns of $M_{\text{in}}$ and construct a $tn \times q^k$ binary matrix by replacing each entry $a$ in $M_{\text{out}}$ with the column $\phi(a)$. Denote this matrix by $M_{\text{out}} \circ M_{\text{in}}$.

**Theorem 2.3.1.** If $M_{\text{out}}$ is a $q$-ary $d$-separable matrix and $M_{\text{in}}$ is a binary $d$-separable matrix with $q$ columns, then the concatenation $M_{\text{out}} \circ M_{\text{in}}$ is $d$-separable.

**Proof.** Let $M = M_{\text{out}} \circ M_{\text{in}}$. For a column $c$ of $M$, let $c[k]$ be the restriction of $c$ to the $k^{th}$ row section of $M$ corresponding to the $k^{th}$ row of $M_{\text{out}}$. Let $c_{\text{out}}$ be the column of $M_{\text{out}}$ that corresponds to $c$ and $c_{\text{out}}[k]$ the element in row $k$ of $c_{\text{out}}$.

Suppose we have two sets of $d$ columns in $M$, $\{v^1, \ldots, v^d\}$ and $\{u^1, \ldots, u^d\}$, such that

$$\bigcup_1^d v^i = \bigcup_1^d u^i$$

This must also hold for each row section, so for each $k$ we have

$$\bigcup_1^d v^i[k] = \bigcup_1^d u^i[k]$$

Now each $v^i[k]$ and $u^i[k]$ is a column of $M_{\text{in}}$, so by definition, since $M_{\text{in}}$ is $d$-separable, there are no two distinct unions of $d$ columns that are equal so we must have

$$\{v^1[k], \ldots, v^d[k]\} = \{u^1[k], \ldots, u^d[k]\}$$

As there is a bijection between the possible entries of $M_{\text{out}}$ and the columns of $M_{\text{in}}$, we have

$$\{v^1_{\text{out}}[k], \ldots, v^d_{\text{out}}[k]\} = \{u^1_{\text{out}}[k], \ldots, u^d_{\text{out}}[k]\}$$

Thus,

$$\bigcup_1^d v^i_{\text{out}}[k] = \bigcup_1^d u^i_{\text{out}}[k]$$
Since this is true for every \( k \), it follows that
\[
\bigcup_{1}^{d} v_{\text{out}}^{i} = \bigcup_{1}^{d} u_{\text{out}}^{i}
\]
But since \( M_{\text{out}} \) is \( d \)-separable, we must have
\[
\{v_{\text{out}}^{1}, \ldots, v_{\text{out}}^{d}\} = \{u_{\text{out}}^{1}, \ldots, u_{\text{out}}^{d}\}
\]
Consequently, \( \{v^{1}, \ldots, v^{d}\} = \{u^{1}, \ldots, u^{d}\} \). Thus, \( M \) is \( d \)-separable.

In general, we will choose our outer matrix to be \( d \)-disjunct and the inner matrix to be \( d \)-separable, though there is some work that could possibly be done to describe \( d \)-separable \( q \)-ary matrices. We now present our construction of \( d \)-separable matrices using Reed Solomon matrices.

By using Reed Solomon Concatenation, we can form large \( d \)-separable matrices which decode faster than the naive disjunct decoding algorithm on a similarly sized matrix. We begin by choosing \( M_{\text{in}} \), a \( d \)-separable, \( t \times q \) inner matrix with analysis time \( O(f(t,q)) \) for some function \( f(t,q) \). We concatenate using \( M_{\text{out}} \), the \((n,q,k)\) Reed Solomon codeword matrix with parameters \( n \) and \( k \) chosen to ensure \( d \)-disjunctness; for optimal dimensions, we fix \( k \) and set \( n = d(k-1) + 1 \). Let \( M = M_{\text{out}} \circ M_{\text{in}} \).

We now describe the analysis algorithm of \( M \). Denote any vector \( v \) restricted to the \( i \)th row section by \( v[i] \). Let \( r \) be the result vector. Then \( r \) is the union of some \( d \) columns (not necessarily distinct) of \( M \); denote these vectors as \( c^{1}, c^{2}, \ldots, c^{d} \). We wish to identify the indices of these columns.

Since \( \bigcup_{j} c^{j} = r \), it is also true that for each row section \( i \), we have \( \bigcup_{j} c^{j}[i] = r[i] \). As each \( c^{j}[i] \) is a column of the inner matrix, let us, decode \( M_{\text{in}}v_{i} = r[i] \) for each row section \( i \) where \( v_{i} \) is the unknown defective vector. This takes time \( O(nf(t,q)) \), and gives sets \( S^{i}_{\text{in}} \) of size at most \( d \) of candidate column types for defectives for each row section.

We now characterize the defective items in \( M \).

**Theorem 2.3.2.** A column \( c \) of \( M \) is defective if and only if \( c[i] \in S^{i}_{\text{in}} \) for all \( i \).

**Proof.** If \( c \) is defective, then there are other columns \( c^{2}, \ldots, c^{d} \) such that \( c \cup \bigcup_{j=2}^{d} c^{j} = r \), so certainly for each row section \( i \) we have \( c[i] \cup \bigcup_{j=2}^{d} c^{j}[i] = r[i] \). By separability of the inner matrix, we must have \( c[i] \in S^{i}_{\text{in}} \) for each \( i \).

Suppose \( c \) is not defective. Denote the defectives as \( c^{1}, \ldots, c^{d} \) so that \( \bigcup_{j} c_{j} = r \). We naturally lift this equation to the outer matrix: \( \bigcup_{j} c_{\text{out}}^{j} = r_{\text{out}} \). By disjunctness of the outer matrix, we have \( c_{\text{out}} \notin \bigcup_{j} c_{\text{out}}^{j} \), or equivalently \( c_{\text{out}} \notin r_{\text{out}} \), so for some row \( i \) we must have \( c_{\text{out}}[i] \notin r_{\text{out}}[i] \). The entries of \( r_{\text{out}} \) are sets of at most \( d \) symbols that correspond to the defective column types in the row sections of \( M \), so lifting back to \( M \) we have \( c[i] \notin S^{i}_{\text{in}} \).  

40
This theorem is useful because to identify the defectives in $M$, we need only find those
columns with defective column types in every row section.

The column types in the row sections naturally correspond to symbols of the outer
matrix so we will define $S^i_{\text{out}}$ to be the set of symbols corresponding to the column types
in $S^i_{\text{in}}$.

We now recover the indices of the defectives in $M_{\text{out}}$, which are the same as the defective
indices of $M$.

First, pick row sections (possibly overlapping) in $M_{\text{out}}$ of size $k$ so that every row is
contained in some row section. Since $n = d(k - 1) + 1$, we can do so with $d$ different row
sections. For each of these row sections, choose an element of the corresponding $S^i_{\text{out}}$ for
every row. There are $d^k$ choices for each section. By Lagrange Interpolation we can find
the unique polynomial over $GF(q)$ of degree less than $k$ that attains those values at the
field elements corresponding to the chosen rows. This takes time $O(d^{k+1}k^2)$.

Each row section gives us a set of polynomials, and the intersection of these sets over
the $d$ sections gives precisely the polynomials that attain defective symbols in every row.
By the above theorem, these are precisely the defectives. Intersecting $d$ sets of size $d^k$ using
hashes takes time $O(d^{k+1})$. Since $M_{\text{out}}$ is indexed by these polynomials, we are done!

The overall time complexity is $O(n f(t, q) + d^{k+1}k^2 + d^{k+1})$, and if we fix $k$, $d$, and $n$,
it is $O(f(t, q))$, which is generally much faster than naive disjunct decoding.

### 2.4 Appendix

The following tables gives lower bounds on the maximum number of rows, $t$, for which
there is no $t \times (t + 1)$ $d$-separable binary matrix.

<table>
<thead>
<tr>
<th>$d$</th>
<th>$t$</th>
<th>$d$</th>
<th>$t$</th>
<th>$d$</th>
<th>$t$</th>
<th>$d$</th>
<th>$t$</th>
<th>$d$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>12</td>
<td>88</td>
<td>22</td>
<td>271</td>
<td>32</td>
<td>551</td>
<td>42</td>
<td>931</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>13</td>
<td>102</td>
<td>23</td>
<td>294</td>
<td>33</td>
<td>585</td>
<td>43</td>
<td>974</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>14</td>
<td>117</td>
<td>24</td>
<td>319</td>
<td>34</td>
<td>619</td>
<td>44</td>
<td>1019</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>15</td>
<td>133</td>
<td>25</td>
<td>344</td>
<td>35</td>
<td>655</td>
<td>45</td>
<td>1064</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
<td>16</td>
<td>149</td>
<td>26</td>
<td>371</td>
<td>36</td>
<td>691</td>
<td>46</td>
<td>1111</td>
</tr>
<tr>
<td>7</td>
<td>34</td>
<td>17</td>
<td>167</td>
<td>27</td>
<td>398</td>
<td>37</td>
<td>729</td>
<td>47</td>
<td>1158</td>
</tr>
<tr>
<td>8</td>
<td>43</td>
<td>18</td>
<td>186</td>
<td>28</td>
<td>427</td>
<td>38</td>
<td>767</td>
<td>48</td>
<td>1206</td>
</tr>
<tr>
<td>9</td>
<td>53</td>
<td>19</td>
<td>206</td>
<td>29</td>
<td>457</td>
<td>39</td>
<td>807</td>
<td>49</td>
<td>1256</td>
</tr>
<tr>
<td>10</td>
<td>63</td>
<td>20</td>
<td>226</td>
<td>30</td>
<td>487</td>
<td>40</td>
<td>847</td>
<td>50</td>
<td>1306</td>
</tr>
<tr>
<td>11</td>
<td>75</td>
<td>21</td>
<td>248</td>
<td>31</td>
<td>519</td>
<td>41</td>
<td>888</td>
<td>51</td>
<td>1358</td>
</tr>
<tr>
<td></td>
<td></td>
<td>52</td>
<td>1410</td>
<td>53</td>
<td>1463</td>
<td>54</td>
<td>1518</td>
<td>55</td>
<td>1573</td>
</tr>
<tr>
<td></td>
<td></td>
<td>56</td>
<td>1630</td>
<td>57</td>
<td>1687</td>
<td>58</td>
<td>1745</td>
<td>59</td>
<td>1805</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>1865</td>
<td>61</td>
<td>1926</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The following tables gives lower bounds on the maximum number of rows, $t$, for which
there is no $t \times (t + 1)$ $d$-disjunct binary matrix.

41
<table>
<thead>
<tr>
<th>d</th>
<th>t</th>
<th>d</th>
<th>t</th>
<th>d</th>
<th>t</th>
<th>d</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8</td>
<td>22</td>
<td>276</td>
<td>32</td>
<td>556</td>
<td>42</td>
<td>936</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>23</td>
<td>299</td>
<td>33</td>
<td>590</td>
<td>43</td>
<td>980</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>24</td>
<td>324</td>
<td>34</td>
<td>624</td>
<td>44</td>
<td>1024</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>25</td>
<td>349</td>
<td>35</td>
<td>660</td>
<td>45</td>
<td>1070</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>26</td>
<td>376</td>
<td>36</td>
<td>696</td>
<td>46</td>
<td>1116</td>
</tr>
<tr>
<td>7</td>
<td>39</td>
<td>27</td>
<td>404</td>
<td>37</td>
<td>734</td>
<td>47</td>
<td>1163</td>
</tr>
<tr>
<td>8</td>
<td>48</td>
<td>28</td>
<td>432</td>
<td>38</td>
<td>772</td>
<td>48</td>
<td>1212</td>
</tr>
<tr>
<td>9</td>
<td>57</td>
<td>29</td>
<td>462</td>
<td>39</td>
<td>812</td>
<td>49</td>
<td>1261</td>
</tr>
<tr>
<td>10</td>
<td>68</td>
<td>30</td>
<td>492</td>
<td>40</td>
<td>852</td>
<td>50</td>
<td>1312</td>
</tr>
<tr>
<td>11</td>
<td>80</td>
<td>31</td>
<td>524</td>
<td>41</td>
<td>894</td>
<td>51</td>
<td>1363</td>
</tr>
</tbody>
</table>

Tables and Data regarding Fast Decoding with Reed Solomon Matrices:

For $d = 2$:

Inner Matrix: $8 \times 11$ from $S(3,2,13)$, Outer Matrix: $(n, 11, k)$ R-S codeword matrix.

<table>
<thead>
<tr>
<th>k</th>
<th>n</th>
<th>Dimensions</th>
<th>D</th>
<th>Best known disjunct dimensions</th>
<th>C.W. $\leq$</th>
<th>R.W. $\leq$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>$24 \times 121$</td>
<td>0.1212</td>
<td>$24 \times 253$</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>$40 \times 1331$</td>
<td>0.0128</td>
<td>$40 \times 1170$</td>
<td>15</td>
<td>499</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>$56 \times 14641$</td>
<td>0.0016</td>
<td>$60 \times 8100$</td>
<td>21</td>
<td>5490</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>$72 \times 161051$</td>
<td>0.0002</td>
<td>$75 \times 74088$</td>
<td>27</td>
<td>60394</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>$88 \times 1771561$</td>
<td>0.0000</td>
<td>$90 \times 371293$</td>
<td>33</td>
<td>664335</td>
</tr>
</tbody>
</table>

Inner Matrix: $12 \times 25$ from $S(3,2,13)$, Outer Matrix: $(n, 25, k)$ R-S codeword matrix.

<table>
<thead>
<tr>
<th>k</th>
<th>n</th>
<th>Dimensions</th>
<th>D</th>
<th>Best known disjunct dimensions</th>
<th>C.W. $\leq$</th>
<th>R.W. $\leq$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>$36 \times 625$</td>
<td>0.0432</td>
<td>$36 \times 730$</td>
<td>9</td>
<td>156</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>$60 \times 15625$</td>
<td>0.0018</td>
<td>$60 \times 8100$</td>
<td>15</td>
<td>3906</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>$84 \times 390625$</td>
<td>0.0001</td>
<td>$85 \times 314432$</td>
<td>21</td>
<td>97656</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>$108 \times 9765625$</td>
<td>0.0000</td>
<td>$108 \times 3200000$</td>
<td>27</td>
<td>2411406</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>$132 \times 244140625$</td>
<td>0.0000</td>
<td>$132 \times 6400000$</td>
<td>33</td>
<td>61035156</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>$156 \times 6103515625$</td>
<td>0.0000</td>
<td>$153 \times 5159780352$</td>
<td>39</td>
<td>1525878906</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>$180 \times 152587890625$</td>
<td>0.0000</td>
<td>$180 \times 25600000000$</td>
<td>45</td>
<td>38146972656</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>$204 \times 3814697265625$</td>
<td>0.0000</td>
<td>$204 \times 512000000000$</td>
<td>51</td>
<td>953674316406</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>$228 \times 95367431640625$</td>
<td>0.0000</td>
<td>$228 \times 10240000000000$</td>
<td>57</td>
<td>23841857910156</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>$252 \times 2384185791015625$</td>
<td>0.0000</td>
<td>unknown</td>
<td>63</td>
<td>596046447753906</td>
</tr>
<tr>
<td>12</td>
<td>23</td>
<td>$276 \times 59604644775390625$</td>
<td>0.0000</td>
<td>unknown</td>
<td>69</td>
<td>1490161193847656</td>
</tr>
<tr>
<td>13</td>
<td>25</td>
<td>$300 \times 149016119384765625$</td>
<td>0.0000</td>
<td>unknown</td>
<td>75</td>
<td>372529029846191406</td>
</tr>
</tbody>
</table>

For $d = 3$:

Inner Matrix: $15 \times 19$ from $S(4,2,16)$, Outer Matrix: $(n, 19, k)$ R-S codeword matrix.

42
<table>
<thead>
<tr>
<th>k</th>
<th>n</th>
<th>Dimensions</th>
<th>D</th>
<th>Best known disjunct dimensions</th>
<th>C.W. ≤</th>
<th>R.W. ≤</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>60 × 361</td>
<td>0.0707</td>
<td>65 × 520</td>
<td>16</td>
<td>96</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>105 × 6859</td>
<td>0.0045</td>
<td>unknown</td>
<td>28</td>
<td>1829</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>150 × 130321</td>
<td>0.0004</td>
<td>unknown</td>
<td>40</td>
<td>34752</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>195 × 2476099</td>
<td>0.0000</td>
<td>unknown</td>
<td>52</td>
<td>660293</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>240 × 47045881</td>
<td>0.0000</td>
<td>unknown</td>
<td>64</td>
<td>12545568</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>285 × 893871739</td>
<td>0.0000</td>
<td>unknown</td>
<td>76</td>
<td>238365797</td>
</tr>
</tbody>
</table>

Inner Matrix: 24 × 49 from S(4,2,25), Outer Matrix: (n, 49, k) R-S codeword matrix.

<table>
<thead>
<tr>
<th>k</th>
<th>n</th>
<th>Dimensions</th>
<th>D</th>
<th>Best known disjunct dimensions</th>
<th>C.W. ≤</th>
<th>R.W. ≤</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>96 × 2401</td>
<td>0.0221</td>
<td>unknown</td>
<td>16</td>
<td>400</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>168 × 117649</td>
<td>0.0005</td>
<td>unknown</td>
<td>28</td>
<td>19608</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>240 × 5764801</td>
<td>0.0000</td>
<td>unknown</td>
<td>40</td>
<td>960800</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>312 × 282475249</td>
<td>0.0000</td>
<td>unknown</td>
<td>52</td>
<td>47079208</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>384 × 13841287201</td>
<td>0.0000</td>
<td>unknown</td>
<td>64</td>
<td>2306881200</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>456 × 678223072849</td>
<td>0.0000</td>
<td>unknown</td>
<td>76</td>
<td>113037178808</td>
</tr>
<tr>
<td>8</td>
<td>22</td>
<td>528 × 33232930569601</td>
<td>0.0000</td>
<td>unknown</td>
<td>88</td>
<td>5538821761600</td>
</tr>
</tbody>
</table>

Notice that Proposition 1.3.1 gives us a maximum number of columns for a given number of rows if column weight is \(d + 1\). Note that Proposition 1.3.1 implies that the matrices derived from the affine family of Steiner Systems with \(n = 2\) are the first matrices to beat the identity whenever \(d + 1\) is a prime power, if you intuitively assume that the first matrix to beat the identity will have column weights of \(d + 1\), which we suspect but have been unable to fully prove. See how it compares to the best known \(d\)-disjunct matrix dimensions for a given number of rows once we’ve beaten the identity:

<table>
<thead>
<tr>
<th>d</th>
<th>Rows</th>
<th>Maximum Columns</th>
<th>Best known</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>35</td>
<td>42</td>
</tr>
</tbody>
</table>

We believe the \(15 \times 35\) has a column weight of 5, hence it beats the bound set by Proposition 1.3.1.

Below is a similar table for \(d = 3\), but we aren’t as convinced that the matrices listed as best-known truly are the best known. If so, we can beat various best knowns easily with Reed-Solomon constructions, but these constructions will all have a higher column weight than \(d + 1\), making Proposition 1.3.1 irrelevant.
<table>
<thead>
<tr>
<th>d</th>
<th>Rows</th>
<th>Maximum Columns</th>
<th>Best known</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>17</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>38</td>
<td>37</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>26</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>54</td>
<td>54</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>77</td>
<td>76</td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>3</td>
<td>33</td>
<td>82</td>
<td>82</td>
</tr>
<tr>
<td>3</td>
<td>34</td>
<td>93</td>
<td>92</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>99</td>
<td>99</td>
</tr>
<tr>
<td>3</td>
<td>37</td>
<td>111</td>
<td>111</td>
</tr>
<tr>
<td>3</td>
<td>38</td>
<td>114</td>
<td>114</td>
</tr>
<tr>
<td>3</td>
<td>39</td>
<td>117</td>
<td>117</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td>3</td>
<td>41</td>
<td>133</td>
<td>133</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
<td>136</td>
<td>136</td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>150</td>
<td>149</td>
</tr>
<tr>
<td>3</td>
<td>44</td>
<td>154</td>
<td>154</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>157</td>
<td>157</td>
</tr>
<tr>
<td>3</td>
<td>46</td>
<td>172</td>
<td>171</td>
</tr>
<tr>
<td>3</td>
<td>47</td>
<td>176</td>
<td>176</td>
</tr>
<tr>
<td>3</td>
<td>48</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td>3</td>
<td>49</td>
<td>196</td>
<td>196</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>200</td>
<td>350</td>
</tr>
</tbody>
</table>
Bibliography


