

# Group testing via residuation and partial geometries<sup>1</sup>

*Finite Geometries 2025 - Seventh Irsee Conference*

**Johan Vester Dinesen**

***Joint work with Oliver W. Gnille, Marcus Greferath and Cornelia Röbling.***

Department of Mathematics and Systems Analysis,  
Aalto University, Finland

*September 5, 2025*



**Aalto University**  
School of Science

---

<sup>1</sup>D., O.W. Gnille, M. Greferath, C. Röbling. Group testing via residuation and partial geometries. *Advances in Mathematics of Communications*, 2025, 19(2): 397-405. doi: 10.3934/amc.2024003'

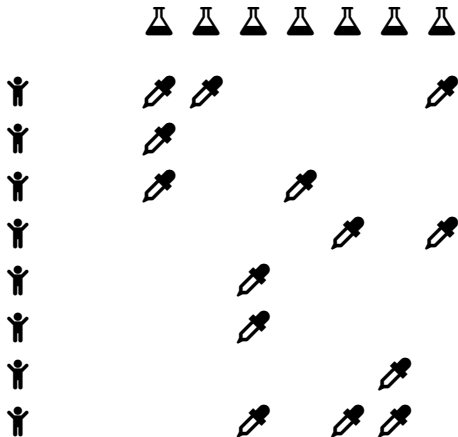
## Group testing



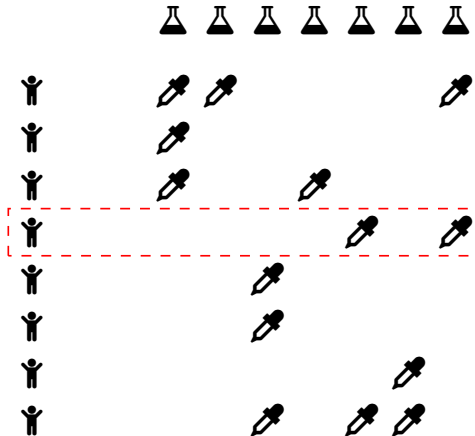
## Group testing



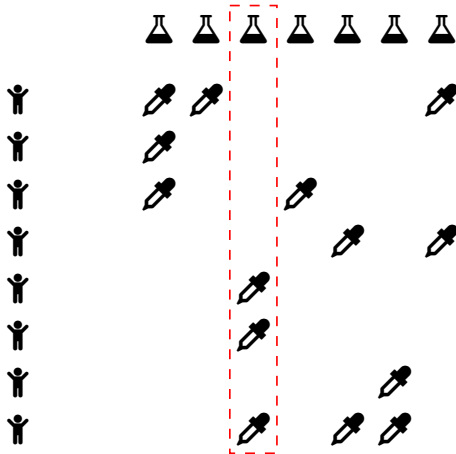
## Group testing



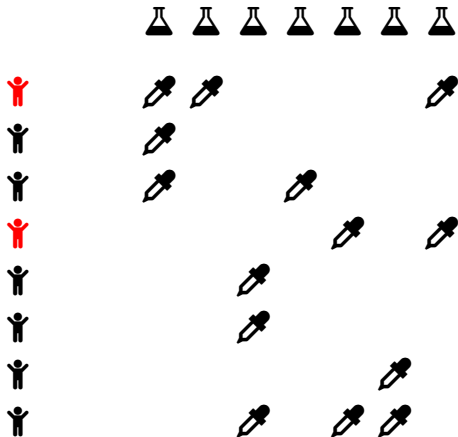
## Group testing



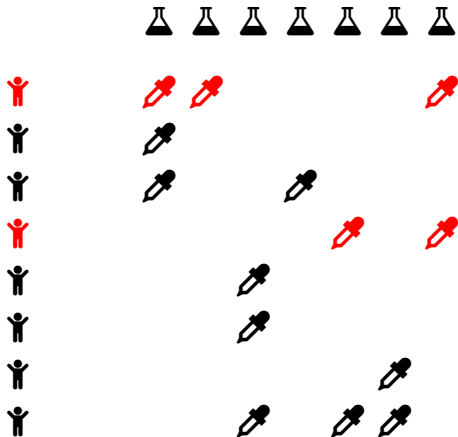
## Group testing



## Group testing

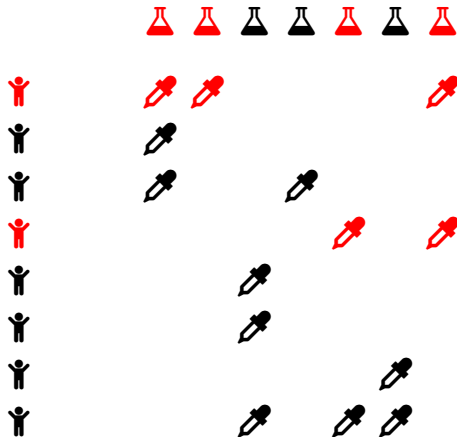


## Group testing

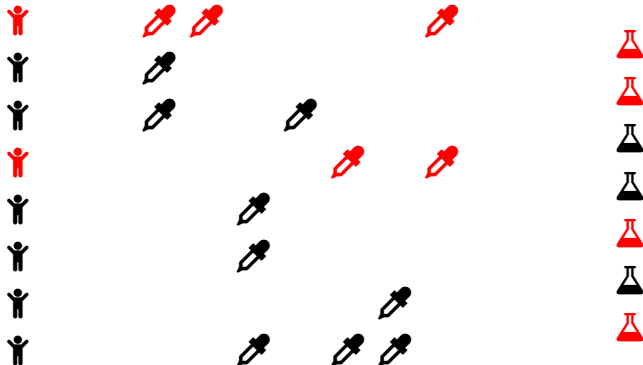




## Group testing



## Group testing



## Group testing

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}^T$$

The problem of testing  $n$  items using  $k$  tests is modelled by a Boolean  $n \times k$  matrix  $H$  by  $x^T H = y^T$ , where  $x$  denotes the items,  $H$  the test allocations,  $y$  the outcomes.

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...
- Simple problem statement yielding interesting combinatorial problems.

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...
- Simple problem statement yielding interesting combinatorial problems.
  - Well-studied through equivalent objects (superimposed codes, cover free families, disjunct systems, coverings of order-interval hypergraphs), but still has many open existence problems.

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...
- Simple problem statement yielding interesting combinatorial problems.
  - Well-studied through equivalent objects (superimposed codes, cover free families, disjunct systems, coverings of order-interval hypergraphs), but still has many open existence problems.



## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...
- Simple problem statement yielding interesting combinatorial problems.
  - Well-studied through equivalent objects (superimposed codes, cover free families, disjunct systems, coverings of order-interval hypergraphs), but still has many open existence problems.

What are some natural questions to consider motivated by compressive sensing?

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...
- Simple problem statement yielding interesting combinatorial problems.
  - Well-studied through equivalent objects (superimposed codes, cover free families, disjunct systems, coverings of order-interval hypergraphs), but still has many open existence problems.

What are some natural questions to consider motivated by compressive sensing?

We let  $H$  be an  $n \times k$  matrix over  $\mathbb{B}_2$ ,  $x \in \mathbb{B}_2^n$  and  $y \in \mathbb{B}_2^k$ .

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...
- Simple problem statement yielding interesting combinatorial problems.
  - Well-studied through equivalent objects (superimposed codes, cover free families, disjunct systems, coverings of order-interval hypergraphs), but still has many open existence problems.

What are some natural questions to consider motivated by compressive sensing?

We let  $H$  be an  $n \times k$  matrix over  $\mathbb{B}_2$ ,  $x \in \mathbb{B}_2^n$  and  $y \in \mathbb{B}_2^k$ .

- 1 For what sparseness of  $x$  can we (explicitly) construct efficient testing matrices such that we can infer  $x$ , given  $xH = y$ ?

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...
- Simple problem statement yielding interesting combinatorial problems.
  - Well-studied through equivalent objects (superimposed codes, cover free families, disjunct systems, coverings of order-interval hypergraphs), but still has many open existence problems.

What are some natural questions to consider motivated by compressive sensing?

We let  $H$  be an  $n \times k$  matrix over  $\mathbb{B}_2$ ,  $x \in \mathbb{B}_2^n$  and  $y \in \mathbb{B}_2^k$ .

- 1 For what sparseness of  $x$  can we (explicitly) construct efficient testing matrices such that we can infer  $x$ , given  $xH = y$ ?
- 2 Can we infer  $x$  efficiently?

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...
- Simple problem statement yielding interesting combinatorial problems.
  - Well-studied through equivalent objects (superimposed codes, cover free families, disjunct systems, coverings of order-interval hypergraphs), but still has many open existence problems.

What are some natural questions to consider motivated by compressive sensing?

We let  $H$  be an  $n \times k$  matrix over  $\mathbb{B}_2$ ,  $x \in \mathbb{B}_2^n$  and  $y \in \mathbb{B}_2^k$ .

- 1 For what sparseness of  $x$  can we (explicitly) construct efficient testing matrices such that we can infer  $x$ , given  $xH = y$ ?
- 2 Can we infer  $x$  efficiently?
- 3 For fixed  $d = \text{supp}(x)$  what are some lower bounds on  $n$ , given  $k$  (or the converse), such that we have guaranteed inference of  $x$ ?

## Introduction

- Group testing was first studied in 1943 by Robert Dorfman to efficiently screen draftees for syphilis during World War II.
- Lots of applications! telecommunications, information technology, data science, cyber security, engineering, genomics, forensics, logistics, software testing, material sciences...
- Simple problem statement yielding interesting combinatorial problems.
  - Well-studied through equivalent objects (superimposed codes, cover free families, disjunct systems, coverings of order-interval hypergraphs), but still has many open existence problems.

What are some natural questions to consider motivated by compressive sensing?

We let  $H$  be an  $n \times k$  matrix over  $\mathbb{B}_2$ ,  $x \in \mathbb{B}_2^n$  and  $y \in \mathbb{B}_2^k$ .

- 1 For what sparseness of  $x$  can we (explicitly) construct efficient testing matrices such that we can infer  $x$ , given  $xH = y$ ?**
- 2 Can we infer  $x$  efficiently?**
- 3 For fixed  $d = \text{supp}(x)$  what are some lower bounds on  $n$ , given  $k$  (or the converse), such that we have guaranteed inference of  $x$ ?**

## Linear Algebra on $\mathbb{B}_2$

Let  $\mathbb{B}_2$  denote the Boolean semiring ( $\mathbb{F}_2$  but  $1 + 1 = 1$ ), and  $\mathbb{B}_2^n$  the semimodule of  $n$ -tuples over  $\mathbb{B}_2$  with component-wise operations from  $\mathbb{B}_2$ . For  $x, y \in \mathbb{B}_2^n$  we impose the ordering

$$x \leq y \Leftrightarrow x + y = y \Leftrightarrow \text{supp}(x) \subseteq \text{supp}(y).$$

Lastly,  $\bar{x}$  denotes the component-wise negation on  $x$ , satisfying De Morgan's laws.

## Linear Algebra on $\mathbb{B}_2$

Let  $\mathbb{B}_2$  denote the Boolean semiring ( $\mathbb{F}_2$  but  $1 + 1 = 1$ ), and  $\mathbb{B}_2^n$  the semimodule of  $n$ -tuples over  $\mathbb{B}_2$  with component-wise operations from  $\mathbb{B}_2$ . For  $x, y \in \mathbb{B}_2^n$  we impose the ordering

$$x \leq y \Leftrightarrow x + y = y \Leftrightarrow \text{supp}(x) \subseteq \text{supp}(y).$$

Lastly,  $\bar{x}$  denotes the component-wise negation on  $x$ , satisfying De Morgan's laws.

### Definition

Let  $H \in \mathbb{B}_2^{n \times k}$  and  $d \in \mathbb{N}$ .  $H$  is

- $d$ -Rev if for any  $x \in B_{\text{Ham}}(\mathbf{0}, d) \subseteq \mathbb{B}_2^n$  and  $z \in \mathbb{B}_2^n$  where  $xH = zH$ , then  $x = z$ .
- $d$ -disjunct if the sum of any  $t \leq d$  rows of  $H$  does not contain any other row of  $H$ .



## Linear Algebra on $\mathbb{B}_2$

Let  $\mathbb{B}_2$  denote the Boolean semiring ( $\mathbb{F}_2$  but  $1 + 1 = 1$ ), and  $\mathbb{B}_2^n$  the semimodule of  $n$ -tuples over  $\mathbb{B}_2$  with component-wise operations from  $\mathbb{B}_2$ . For  $x, y \in \mathbb{B}_2^n$  we impose the ordering

$$x \leq y \Leftrightarrow x + y = y \Leftrightarrow \text{supp}(x) \subseteq \text{supp}(y).$$

Lastly,  $\bar{x}$  denotes the component-wise negation on  $x$ , satisfying De Morgan's laws.

### Definition

Let  $H \in \mathbb{B}_2^{n \times k}$  and  $d \in \mathbb{N}$ .  $H$  is

- $d$ -Rev if for any  $x \in B_{\text{Ham}}(\mathbf{0}, d) \subseteq \mathbb{B}_2^n$  and  $z \in \mathbb{B}_2^n$  where  $xH = zH$ , then  $x = z$ .
- $d$ -disjunct if the sum of any  $t \leq d$  rows of  $H$  does not contain any other row of  $H$ .

$d$ -disjunctness is equivalent to if for any  $d + 1$  rows, where you designate one, then there exists a column with a 1 in the designated row, and 0 in the other  $d$  rows.

## Linear Algebra on $\mathbb{B}_2$

Let  $\mathbb{B}_2$  denote the Boolean semiring ( $\mathbb{F}_2$  but  $1 + 1 = 1$ ), and  $\mathbb{B}_2^n$  the semimodule of  $n$ -tuples over  $\mathbb{B}_2$  with component-wise operations from  $\mathbb{B}_2$ . For  $x, y \in \mathbb{B}_2^n$  we impose the ordering

$$x \leq y \Leftrightarrow x + y = y \Leftrightarrow \text{supp}(x) \subseteq \text{supp}(y).$$

Lastly,  $\bar{x}$  denotes the component-wise negation on  $x$ , satisfying De Morgan's laws.

### Definition

Let  $H \in \mathbb{B}_2^{n \times k}$  and  $d \in \mathbb{N}$ .  $H$  is

- $d$ -Rev if for any  $x \in B_{\text{Ham}}(\mathbf{0}, d) \subseteq \mathbb{B}_2^n$  and  $z \in \mathbb{B}_2^n$  where  $xH = zH$ , then  $x = z$ .
- $d$ -disjunct if the sum of any  $t \leq d$  rows of  $H$  does not contain any other row of  $H$ .

$d$ -disjunctness is equivalent to if for any  $d + 1$  rows, where you designate one, then there exists a column with a 1 in the designated row, and 0 in the other  $d$  rows.

### Theorem

$H \in \mathbb{B}_2^{n \times k}$  is  $d$ -Rev if and only if it is  $d$ -disjunct.

## Linear Algebra on $\mathbb{B}_2$

Let  $\mathbb{B}_2$  denote the Boolean semiring ( $\mathbb{F}_2$  but  $1 + 1 = 1$ ), and  $\mathbb{B}_2^n$  the semimodule of  $n$ -tuples over  $\mathbb{B}_2$  with component-wise operations from  $\mathbb{B}_2$ . For  $x, y \in \mathbb{B}_2^n$  we impose the ordering

$$x \leq y \Leftrightarrow x + y = y \Leftrightarrow \text{supp}(x) \subseteq \text{supp}(y).$$

Lastly,  $\bar{x}$  denotes the component-wise negation on  $x$ , satisfying De Morgan's laws.

### Definition

Let  $H \in \mathbb{B}_2^{n \times k}$  and  $d \in \mathbb{N}$ .  $H$  is

- $d$ -Rev if for any  $x \in B_{\text{Ham}}(\mathbf{0}, d) \subseteq \mathbb{B}_2^n$  and  $z \in \mathbb{B}_2^n$  where  $xH = zH$ , then  $x = z$ .
- $d$ -disjunct if the sum of any  $t \leq d$  rows of  $H$  does not contain any other row of  $H$ .

$d$ -disjunctness is equivalent to if for any  $d + 1$  rows, where you designate one, then there exists a column with a 1 in the designated row, and 0 in the other  $d$  rows.

### Theorem

$H \in \mathbb{B}_2^{n \times k}$  is  $d$ -Rev if and only if it is  $d$ -disjunct.

## Residuated Mappings

### Definition

Let  $(A, \leq_A)$  and  $(B, \leq_B)$  be two partially ordered sets. For mappings  $f: A \rightarrow B$  and  $g: B \rightarrow A$  the pair  $(f, g)$  is a *residuated pair*, if there holds

$$f(x) \leq_B y \iff x \leq_A g(y), \quad \text{for all } x \in A \text{ and } y \in B.$$

We call  $f$  residuated if there exists a  $g$  such that  $(f, g)$  is residuated, and we call  $g$  its residual. A residuated pair  $(f, g)$  has many nice properties:

## Residuated Mappings

### Definition

Let  $(A, \leq_A)$  and  $(B, \leq_B)$  be two partially ordered sets. For mappings  $f: A \rightarrow B$  and  $g: B \rightarrow A$  the pair  $(f, g)$  is a *residuated pair*, if there holds

$$f(x) \leq_B y \iff x \leq_A g(y), \quad \text{for all } x \in A \text{ and } y \in B.$$

We call  $f$  residuated if there exists a  $g$  such that  $(f, g)$  is residuated, and we call  $g$  its residual. A residuated pair  $(f, g)$  has many nice properties:

- $f$  and  $g$  are monotone mappings and they uniquely determine each other.

## Residuated Mappings

### Definition

Let  $(A, \leq_A)$  and  $(B, \leq_B)$  be two partially ordered sets. For mappings  $f: A \rightarrow B$  and  $g: B \rightarrow A$  the pair  $(f, g)$  is a *residuated pair*, if there holds

$$f(x) \leq_B y \iff x \leq_A g(y), \quad \text{for all } x \in A \text{ and } y \in B.$$

We call  $f$  residuated if there exists a  $g$  such that  $(f, g)$  is residuated, and we call  $g$  its residual. A residuated pair  $(f, g)$  has many nice properties:

- $f$  and  $g$  are monotone mappings and they uniquely determine each other.
- $f \circ g \circ f = f$  and  $g \circ f \circ g = g$ , so  $g \circ f$  and  $f \circ g$  forms closure and kernel operators, respectively.

### Definition

Let  $(A, \leq_A)$  and  $(B, \leq_B)$  be two partially ordered sets. For mappings  $f: A \rightarrow B$  and  $g: B \rightarrow A$  the pair  $(f, g)$  is a *residuated pair*, if there holds

$$f(x) \leq_B y \iff x \leq_A g(y), \quad \text{for all } x \in A \text{ and } y \in B.$$

We call  $f$  residuated if there exists a  $g$  such that  $(f, g)$  is residuated, and we call  $g$  its residual. A residuated pair  $(f, g)$  has many nice properties:

- $f$  and  $g$  are monotone mappings and they uniquely determine each other.
- $f \circ g \circ f = f$  and  $g \circ f \circ g = g$ , so  $g \circ f$  and  $f \circ g$  forms closure and kernel operators, respectively.
- if  $A$  and  $B$  are complete lattices then  $f$  is residuated if and only if it is a  $\vee$ -homomorphism and  $f(0_A) = 0_B$ .

## Residuated Mappings on $\mathbb{B}_2^n$

As  $\mathbb{B}_2^n$  is a complete lattice with  $\vee = +$  and  $\wedge = \cdot$ , we easily see that

$f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  is residuated if and only if  $f(x) = xH$  for some  $H \in \mathbb{B}_2^{n \times k}$ .



## Residuated Mappings on $\mathbb{B}_2^n$

As  $\mathbb{B}_2^n$  is a complete lattice with  $\vee = +$  and  $\wedge = \cdot$ , we easily see that

$f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  is residuated if and only if  $f(x) = xH$  for some  $H \in \mathbb{B}_2^{n \times k}$ .

### Theorem (D., Gnille, Greferath, Rößing, 2025)

*Let  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  be a residuated mapping represented by the  $n \times k$  matrix  $H$ . Then the residual mapping  $g: \mathbb{B}_2^k \rightarrow \mathbb{B}_2^n$  is given by the assignment  $y \mapsto \overline{y}H^\top$ .*

## Residuated Mappings on $\mathbb{B}_2^n$

As  $\mathbb{B}_2^n$  is a complete lattice with  $\vee = +$  and  $\wedge = \cdot$ , we easily see that

$f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  is residuated if and only if  $f(x) = xH$  for some  $H \in \mathbb{B}_2^{n \times k}$ .

### Theorem (D., Gnille, Greferath, Rößing, 2025)

*Let  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  be a residuated mapping represented by the  $n \times k$  matrix  $H$ . Then the residual mapping  $g: \mathbb{B}_2^k \rightarrow \mathbb{B}_2^n$  is given by the assignment  $y \mapsto \overline{y}H^\top$ .*

### Theorem (D., Gnille, Greferath, Rößing, 2025)

*Let  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  be a residuated mapping represented by the  $n \times k$  matrix  $H$  and let  $g$  be the residual of  $f$ . Then the following are equivalent:*

- 1**  $H$  is  $d$ -disjunct.
- 2**  $(g \circ f)|_{B_{\text{Ham}}(\mathbf{0}, d)} = \text{id}.$
- 3**  $B_{\text{Ham}}(\mathbf{0}, d) \subseteq \text{im}(g).$
- 4**  $B_{\text{Ham}}(\mathbf{1}, d) \subseteq \text{colsp}(H).$

## Group testing schemes

### Definition

An  $(n, k, d)$ -group testing scheme is a residuated mapping  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  together with a decoder  $g: \mathbb{B}_2^k \rightarrow \mathbb{B}_2^n$  such that  $(g \circ f)|_{B_{\text{Ham}}(\mathbf{0}, d)} = \text{id}$ .

- Any  $d$ -disjunct  $n \times k$  matrix  $H$  will yield a  $(n, k, d)$ -group testing scheme with the residual  $g$  of  $f$  being the decoder.

## Group testing schemes

### Definition

An  $(n, k, d)$ -group testing scheme is a residuated mapping  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  together with a decoder  $g: \mathbb{B}_2^k \rightarrow \mathbb{B}_2^n$  such that  $(g \circ f)|_{B_{\text{Ham}}(\mathbf{0}, d)} = \text{id}$ .

- Any  $d$ -disjunct  $n \times k$  matrix  $H$  will yield a  $(n, k, d)$ -group testing scheme with the residual  $g$  of  $f$  being the decoder.

Furthermore, the group testing scheme also allows us to verify if a given test outcome is indeed a valid outcome given the test matrix.

### Proposition

*Let  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  be a residuated mapping. Then  $y \in \text{im}(f)$  if and only if  $(f \circ g)(y) = y$ .*

## Group testing schemes

### Definition

An  $(n, k, d)$ -group testing scheme is a residuated mapping  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  together with a decoder  $g: \mathbb{B}_2^k \rightarrow \mathbb{B}_2^n$  such that  $(g \circ f)|_{B_{\text{Ham}}(\mathbf{0}, d)} = \text{id}$ .

- Any  $d$ -disjunct  $n \times k$  matrix  $H$  will yield a  $(n, k, d)$ -group testing scheme with the residual  $g$  of  $f$  being the decoder.

Furthermore, the group testing scheme also allows us to verify if a given test outcome is indeed a valid outcome given the test matrix.

### Proposition

*Let  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  be a residuated mapping. Then  $y \in \text{im}(f)$  if and only if  $(f \circ g)(y) = y$ .*

Suggests we should consider error-correcting codes!

## Error-Correcting Codes on $\mathbb{B}_2^k$

As a residuated mapping  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  induced by a  $d$ -disjunct matrix is injective on  $B_{\text{Ham}}(\mathbf{0}, d)$ , then for any  $\delta \leq d$  we can construct codes  $\mathcal{C}_{f,\delta} \subseteq \mathbb{B}_2^k$  as  $\mathcal{C}_{f,\delta} = f(B_{\text{Ham}}(\mathbf{0}, \delta))$ , with  $|\mathcal{C}_{f,\delta}| = \sum_{i=0}^{\delta} \binom{n}{i}$ .

## Error-Correcting Codes on $\mathbb{B}_2^k$

As a residuated mapping  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  induced by a  $d$ -disjunct matrix is injective on  $B_{\text{Ham}}(\mathbf{0}, d)$ , then for any  $\delta \leq d$  we can construct codes  $\mathcal{C}_{f,\delta} \subseteq \mathbb{B}_2^k$  as  $\mathcal{C}_{f,\delta} = f(B_{\text{Ham}}(\mathbf{0}, \delta))$ , with  $|\mathcal{C}_{f,\delta}| = \sum_{i=0}^{\delta} \binom{n}{i}$ .

### Motivating Example:

Let  $H$  be the  $7 \times 7$  incidence matrix of the Fano plane, which is 2-disjunct. We can define 2 codes:

- $\mathcal{C}_{f,1}$  has minimum distance 3  $\Rightarrow$  can identify one infected sample, while also correcting a faulty test.
- $\mathcal{C}_{f,2}$  has minimum distance 2  $\Rightarrow$  can identify two infected samples, but cannot correct any faulty tests.

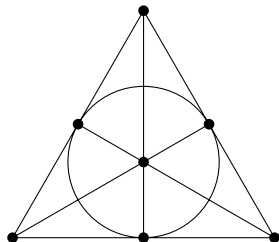


Figure: The Fano plane.

## Error-Correcting Codes on $\mathbb{B}_2^k$

As a residuated mapping  $f: \mathbb{B}_2^n \rightarrow \mathbb{B}_2^k$  induced by a  $d$ -disjunct matrix is injective on  $B_{\text{Ham}}(\mathbf{0}, d)$ , then for any  $\delta \leq d$  we can construct codes  $\mathcal{C}_{f,\delta} \subseteq \mathbb{B}_2^k$  as  $\mathcal{C}_{f,\delta} = f(B_{\text{Ham}}(\mathbf{0}, \delta))$ , with  $|\mathcal{C}_{f,\delta}| = \sum_{i=0}^{\delta} \binom{n}{i}$ .

### Motivating Example:

Let  $H$  be the  $7 \times 7$  incidence matrix of the Fano plane, which is 2-disjunct. We can define 2 codes:

- $\mathcal{C}_{f,1}$  has minimum distance 3  $\Rightarrow$  can identify one infected sample, while also correcting a faulty test.
- $\mathcal{C}_{f,2}$  has minimum distance 2  $\Rightarrow$  can identify two infected samples, but cannot correct any faulty tests.

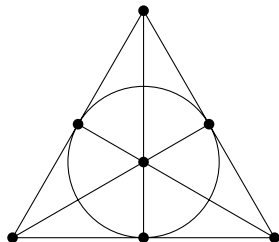


Figure: The Fano plane.

**Question:** Can we find classes of such codes with efficient decoders?



## Inducing codes from Inversive Planes

### Definition (Inversive Planes)

An inversive plane of order  $n$  is an incidence structure of points and circles, where

- 1 Any circle has  $n + 1$  points,
- 2 A unique circle is incident with any given triple of points,
- 3 If  $p_1 \in c_1$  and  $p_2 \notin c_1$ , there exists a unique circle  $c_2$  such that  $p_1, p_2 \in c_2$ ,
- 4 There exists 4 points which are not concircular.

Is also a  $3-(n^2 + 1, n + 1, 1)$ -design, or a  $S(3, n^2 + 1, n + 1)$  Steiner system.

## Inducing codes from Inversive Planes

### Definition (Inversive Planes)

An inversive plane of order  $n$  is an incidence structure of points and circles, where

- 1 Any circle has  $n + 1$  points,
- 2 A unique circle is incident with any given triple of points,
- 3 If  $p_1 \in c_1$  and  $p_2 \notin c_1$ , there exists a unique circle  $c_2$  such that  $p_1, p_2 \in c_2$ ,
- 4 There exists 4 points which are not concircular.

Is also a  $3-(n^2 + 1, n + 1, 1)$ -design, or a  $S(3, n^2 + 1, n + 1)$  Steiner system.

### Proposition

*Let  $H$  be the incidence matrix  $H$  of an inversive plane of order  $n$  and let  $f: \mathbb{B}_2^{n(n^2+1)} \rightarrow \mathbb{B}_2^{n^2+1}$  be the corresponding residuated mapping.  $H$  is then 1-disjunct, and  $\mathcal{C}_{f,1} \setminus \{0\}$  is a constant weight code of minimum distance  $n + 1$ .*

Thus, inversive planes can be used to generate testing schemes with good error-correcting capabilities, assuming the prevalence of the underlying disease is low.

## Partial Linear Spaces

### Definition

For  $s, t \in \mathbb{N}$ , a finite incidence structure  $(P, L)$  of points and lines is a *partial linear space* of order  $(s, t)$  if

- Every line is incident with  $s + 1$  points, and every point is incident with  $t + 1$  lines.
- Two different points are connected by at most one line.

## Partial Linear Spaces

### Definition

For  $s, t \in \mathbb{N}$ , a finite incidence structure  $(P, L)$  of points and lines is a *partial linear space* of order  $(s, t)$  if

- Every line is incident with  $s + 1$  points, and every point is incident with  $t + 1$  lines.
- Two different points are connected by at most one line.

### Theorem

Let  $(P, L)$  be a partial linear space of order  $(s, t)$ , and let  $\ell_1, \dots, \ell_m$  denote a collection of  $m \leq s$  distinct lines. If  $\ell \in L$  with  $\ell \subseteq \ell_1 \cup \dots \cup \ell_m$  then  $\ell = \ell_j$  for some  $j \leq m$ .

## Partial Linear Spaces

### Definition

For  $s, t \in \mathbb{N}$ , a finite incidence structure  $(P, L)$  of points and lines is a *partial linear space* of order  $(s, t)$  if

- Every line is incident with  $s + 1$  points, and every point is incident with  $t + 1$  lines.
- Two different points are connected by at most one line.

### Theorem

Let  $(P, L)$  be a partial linear space of order  $(s, t)$ , and let  $\ell_1, \dots, \ell_m$  denote a collection of  $m \leq s$  distinct lines. If  $\ell \in L$  with  $\ell \subseteq \ell_1 \cup \dots \cup \ell_m$  then  $\ell = \ell_j$  for some  $j \leq m$ .

### Proof.

Suppose for contradiction that  $\ell \neq \ell_i$  for  $i = 1, \dots, m$ . Then  $|\ell \cap \ell_i| \leq 1$ , so

$$s + 1 = |\ell| \leq \left| \ell \cap \bigcup_{i=1}^m \ell_i \right| \leq \sum_{i=1}^m |\ell \cap \ell_i| \leq m.$$



## Corollary

*An  $|L| \times |P|$  incidence matrix  $H$  of a partial linear space  $(P, L)$  of order  $(s, t)$  is  $s$ -disjunct.*

## Corollary

An  $|L| \times |P|$  incidence matrix  $H$  of a partial linear space  $(P, L)$  of order  $(s, t)$  is  $s$ -disjunct.

## Definition (Generalized quadrangle)

A partial linear space  $(P, L)$  of order  $(s, t)$  is called a generalized quadrangle,  $GQ(s, t)$ , if for any non-incident point-line pair  $(p, \ell)$ , there exists a unique point  $q$  on  $\ell$  that is connected with  $p$  by a line.

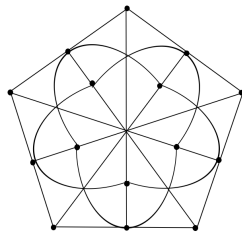


Figure:  $GQ(2,2)$ , or  $W(2)$ .

## Corollary

An  $|L| \times |P|$  incidence matrix  $H$  of a partial linear space  $(P, L)$  of order  $(s, t)$  is  $s$ -disjunct.

## Definition (Generalized quadrangle)

A partial linear space  $(P, L)$  of order  $(s, t)$  is called a generalized quadrangle,  $GQ(s, t)$ , if for any non-incident point-line pair  $(p, \ell)$ , there exists a unique point  $q$  on  $\ell$  that is connected with  $p$  by a line.

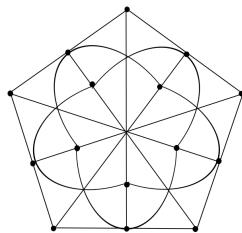


Figure:  $GQ(2,2)$ , or  $W(2)$ .

## Corollary

$GQ(s, t)$  yields a  $((s + 1)(st + 1), (t + 1)(st + 1), s)$ -group testing scheme.



## Final remarks

- Group testing is seeing actual proof of concepts, using 1-disjunct  $100 \times 20$  test matrix ([Stoltze et al., 'Combinatorial batching of DNA for ultralow-cost detection of pathogenic variants', Genome Medicine, 2023.](#)), and has received  $\geq$  €3.000.000 with the intention of clinical implementation in Denmark.

## Final remarks

- Group testing is seeing actual proof of concepts, using 1-disjunct  $100 \times 20$  test matrix (Stoltze et al., 'Combinatorial batching of DNA for ultralow-cost detection of pathogenic variants', *Genome Medicine*, 2023.), and has received  $\geq$  €3.000.000 with the intention of clinical implementation in Denmark.
- Optimal 1-disjunct matrices are solved by Sperner's Theorem, but even for  $d = 2$  and  $k \in [16]$  the maximal  $n$  is not known (A286874, OEIS).

## Final remarks

- Group testing is seeing actual proof of concepts, using 1-disjunct  $100 \times 20$  test matrix (Stoltze et al., 'Combinatorial batching of DNA for ultralow-cost detection of pathogenic variants', *Genome Medicine*, 2023.), and has received  $\geq$  €3.000.000 with the intention of clinical implementation in Denmark.
- Optimal 1-disjunct matrices are solved by Sperner's Theorem, but even for  $d = 2$  and  $k \in [16]$  the maximal  $n$  is not known (A286874, OEIS).
- Try to construct  $d$ -disjunct matrices with "nice" properties:
  - 1  $n \gg k$  and for specific  $d$ .
  - 2 Constant weight columns and rows for samples with "equal dilution".
  - 3 Sparse matrices.
  - 4 Induces codes with high minimum distance and an efficient decoder.
    - Assuming different error models; maybe false positives are more likely than false negatives, so channel should not have symmetric cross-over probability.
  - 5 ...?

Thank you for your attention!