Non-adaptive group testing: residuation theory and disjunct matrices

Master's Thesis

Advanced Mathematical Topics with Applications

Johan Vester Dinesen

Department of Mathematical Sciences

Aalborg University

AALBORG UNIVERSITY
STUDENT REPORT

Department of Mathematical Sciences
Aalborg University
Skjernvej 4A
9220 Aalborg Øst
http://math.aau.dk

Title:

Non-adaptive group testing: residuation theory and disjunct matrices

Project:

Master's Thesis

Project Period:

February 2023 - June 2023

Project Group:

4.111 A1

Author:

Johan Vester Dinesen

Supervisors:

Matteo Bonini
Oliver Wilhelm Gnilke

Pages: 44 (45 with appendices)

Finished: 2nd of June 2023

Abstract:

This thesis presents non-adaptive group testing, which is the problem of pooling samples in tests, such that given d positive items out of n total items one minimises the amount of tests used while determining the d items of interest.

The thesis introduces the necessary background knowledge in residuation theory in order to determine properties of matrices over the Boolean semiring, specifically d-disjunct matrices, where the union of supports of any d rows does not contain the support of any other row.

Disjunct matrices provides schemes with efficient decoding for the implementation of non-adaptive group testing, and as such, we present constructions and bounds on such matrices.

Preface

The following thesis was written during the spring semester of 2023 at the Department of Mathematical Sciences at Aalborg University.

I would like to direct a thanks to my supervisors Matteo Bonini and Oliver Wilhelm Gnilke for supervision during the course of the project period, and in general for their time at Aalborg University.

Johan Vester Dinesen

Danish Abstract

Dette speciale præsenterer ikke-adaptiv gruppetestning, som er problematikken hvori man sammenlægger prøver og tester disse samtidigt. Dette gøres med henblik på at minimere antallet af tests, givet n genstande, hvoraf d er positive, hvor man desledes bestemmer de d genstande af interesse.

Specialet introducerer den nødvendige baggrundsviden indenfor residueringsteori for at kunne bestemme egenskaber ved matricer over den boolske semiring. Specifikt betragtes d-disjunkte matricer, hvor fællesmængden af støtten af d vilkårlige rækker ikke indeholder støtten af nogen anden række.

Disjunkte matricer giver anledning til effektiv dekodning ved implementeringen af ikke-adaptiv gruppetestning, og derfor præsenterer vi også diverse konstruktioner af og grænser for disse matricer.

Contents

1 Introduction

Group testing was first introduced during the Second World War by Robert Dorfman as a means to lessen the amount of tests required to detect syphilis in draftees, where instead of testing the blood of each draftee, the samples were pooled into single tests and tested simultaneously [1]. Group testing sees a wide array of use within screening for diseases, quality control in production lines, data forensics, one-way hash functions, computer diagnosis, machine learning, and many more [2].

In the following thesis we consider the case of noiseless non-adaptive group testing where all tests of the testing scheme must be conducted independently of one another, and we consider d-disjunct matrices, which are Boolean $n \times k$ matrices which, when given n items, where at most d are infected, will be guaranteed at determining these infected using k tests. Non-adaptive group testing is very similar to compressive sensing, but rather than classical vector space multiplication and addition we replace it with Boolean $\vee$ and $\wedge$ operators [3].

The thesis begins by considering residuation theory, where we consider partially ordered sets and residuated mappings between these. The main purpose of the chapter is to show that a mapping between complete lattices is a residuated mapping if and only if it has certain homomorphic properties. Lastly in this chapter, we consider Boolean semimodules and show that these are Boolean algebras, and consider the residuated mappings between such semimodules.

Lastly, we formalise the notion of non-adaptive group testing and see how this problem can be solved using the previously developed theory of d-disjunct matrices and discuss some common problems within the field. We then consider bounds on the size of $n \times k$ d-disjunct matrices by showcasing equivalency between a variety of objects, namely cover-free families, disjunct systems, superimposed codes, and coverings of order-interval hypergraphs. We also showcase a variety of constructions and combinatorial objects which lead to incidence matrices which has the d-disjunctness property, and consider some of the largest $n \times k$ 2-disjunct matrices for low values of k. The thesis ends by considering a recent project in which they developed and implemented a batch sequencing technique for DNA samples in an explorative article.

2 | Residuation Theory

In the following chapter we introduce residuation theory. The main intent of the chapter is to develop the theory of residuated mappings in relation to its later shown applicability within group testing, and as such we do not provide many general examples of this otherwise rich and interesting theory, and instead refer the reader to [4, 5], on which the chapter is also based. The theory is instead exemplified in later chapters in the context of group testing.

2.1 Ordered Sets and Order-Preserving Maps

If a binary relation R on a set E is reflexive, transitive, and anti-symmetric we say that the relation R is an ordering on E, and (E, R) is an ordered set. If no confusion can arise, E can be referred to as an ordered set, and we will often choose to denote the ordering by $\leqslant$, rather than R. Furthermore, we say that E is totally ordered, or forms a chain, if for any $x, y \in E$, then $x \leqslant y$ or $y \leqslant x$.

Additionally, any two elements x, y of the ordered set E are comparable if $x \leqslant y$ or $y \leqslant x$, denoted $x \nparallel y$. Similarly, if neither $x \leqslant y$ or $y \leqslant x$, then x and y are incomparable, denoted $x \parallel y$. For a non-empty subset F of an ordered set E, we say that F is totally unordered if all elements of F are pairwise incomparable, which is equivalent to the restriction to F of $\leqslant$ is equality. Lastly, we say that x is covered by y, denoted $x \prec y$ if $x < y$, and there exists no $z \in E$ such that $x < z < y$.

For a given relation R of E we denote its dual as $R^\top$, where $xR^\top y$ if and only if yRx, where it is clear, that if R is an ordering, then so is $R^\top$. We will denote the dual of E by $E^* = (E, R^\top)$. We denote the dual of $\leqslant$ as $\geqslant$, and similarly the dual of $<$ is $>$, where $<$ implies $x \leqslant y$ and $x \neq y$.

Based on the previous comment, that for any order R on a set E, we have that $R^\top$ is also an order on E, which provides a principle of duality. This principle states, that to any theorem concerning an ordered set E, we have a corresponding theorem on E^*, which is obtained by replacing any statement with $\leqslant$, explicitly or implicitly, by its dual.

As an example of the application of the principle of duality consider an ordered set $(E, \leqslant)$. A top element of E is then some $x \in E$ such that $y \leqslant x$ for all $y \in E$, which is unique by the anti-symmetry. Similarly, a bottom element is an element $z \in E$ such that $z \leqslant y$ for all $y \in E$, which is then unique, when it exists, by the principle of duality. If an ordered set contains both a top and a bottom element it is said to be bounded. We denote the top element of a set E as 1_E, and a bottom element as 0_E, and omit the subscript when no confusion is bound to arise.

Given an ordering we can visually represent this as a Hasse diagram, or represent the converse ordering with a dual Hasse diagram. We represent the relation $x \prec y$ by an increasing line segment from x to y, as seen in figure 2.1, or the dual Hasse diagram, which is obtained by flipping the Hasse diagram upside down.

(a) Hasse diagram. **(b)** Dual Hasse diagram.

Figure 2.1: Hasse diagram and dual Hasse diagram of the relation $x \prec y$.

We now define up-sets and down-sets from which much of the structure of a partially ordered set will be considered.

Definition 2.1. Let $(E, \leqslant)$ be an ordered set. A down-set of E is a non-empty subset $D \subseteq E$ such that if $x \in D$, and $y \leqslant x$, then $y \in D$. Furthermore, any down-set of the form

$$x^{\downarrow} = \{y \in E \mid y \leqslant x\}$$

is a principal down-set. Dually, an up-set of E is a non-empty subset $U \subseteq E$ such that if $x \in U$, and $y \geqslant x$, then $y \in U$. Furthermore, any up-set of the form

$$x^{\uparrow} = \{y \in E \mid y \geqslant x\}$$

is a principal up-set.

Definition 2.2. Let $(A, \leqslant_1)$ and $(B, \leqslant_2)$ be ordered sets. Then a mapping $f \colon A \to B$ such that

$$x \leqslant_1 y \Rightarrow f(x) \leqslant_2 f(y)$$

for any $x, y \in A$ is called isotone. Similarly, if

$$x \leqslant_1 y \Rightarrow f(x) \geqslant_2 f(y)$$

for any $x, y \in A$, then f is called antitone.

In the following result we characterise isotone mappings, but first fix a notation for the pre-image of a mapping. For two ordered sets A, B and a mapping $f \colon A \to B$, then for any non-empty subset R of B, we denote the pre-image

of R under f as the subset of A, given by

$$f^{\leftarrow}(R) = \{x \in A \mid f(x) \in R\}$$

Theorem 2.3. *Let $(A, \leqslant_1), (B, \leqslant_2)$ be ordered sets, and $f\colon A \to B$ be any mapping. Then the statements*

i. f is isotone;

ii. The pre-image of every principal down-set of B under f is either empty, or is a down-set of A;

iii. The pre-image of every principal up-set of B under f is either empty, or is an up-set of A;

are equivalent.

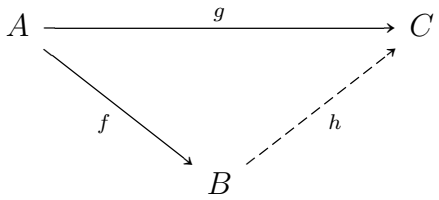
Proof. *i.* $\Rightarrow$ *ii.* Consider $x \in B$ such that $f^{\leftarrow}(x^{\downarrow})$ is non-empty, and if $y \in f^{\leftarrow}(x^{\downarrow})$, and $z \leqslant y$, we have that $f(z) \leqslant f(y) \leqslant x$, so $z \in f^{\leftarrow}(x^{\downarrow})$.

ii. $\Rightarrow$ *i.* We have that $y \in f^{\leftarrow}(f(y)^{\downarrow})$ as $f(y) \leqslant f(y)$. Thus, if $x \leqslant y$ then $x \in f^{\leftarrow}(f(y)^{\downarrow})$, so $f(x) \leqslant f(y)$.

i. $\Rightarrow$ *iii.* Let $x \in B$ such that $f^{\leftarrow}(x^{\uparrow})$ is non-empty. Now, if $y \in f^{\leftarrow}(x^{\uparrow})$, and $z \geqslant y$, then $f(z) \geqslant f(y) \geqslant x$, so $z \in f^{\leftarrow}(x^{\uparrow})$.

iii. $\Rightarrow$ *i.* We have that $y \in f^{\leftarrow}(f(y)^{\uparrow})$ as $f(y) \geqslant f(y)$. Thus, if $x \geqslant y$ then $x \in f^{\leftarrow}(f(y)^{\uparrow})$, so $f(x) \geqslant f(y)$. ■

Now, consider the following diagram of ordered sets and mappings



We are then interested in knowing when there exists an isotone mapping $h\colon B \to C$ such that $h \circ f \geqslant g$, or $h \circ f \leqslant g$. In the following results we omit subscripts for the different orderings, as it will be clear from the context.

Theorem 2.4. *Let A, B, C be ordered sets with mappings $f\colon A \to B$ and $g\colon A \to C$. Then if there exists isotone $h\colon B \to C$, then*

i. $h \circ f \geqslant g$;

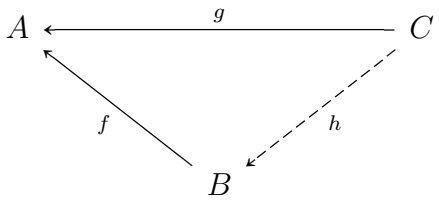
ii. $f^{\leftarrow}(x^{\downarrow}) \subseteq g^{\leftarrow}(h(x)^{\downarrow})$ for all $x \in B$;

are equivalent statements.

Proof. *i.* $\Rightarrow$ *ii.* Assume that $f^{\leftarrow}(x^{\downarrow})$ is non-empty, so for any $y \in f^{\leftarrow}(x^{\downarrow})$ then $f(y) \leqslant x$, and by the isotonicity of h we have $g(y) \leqslant (h \circ f)(y) \leqslant h(x)$, so $y \in g^{\leftarrow}(h(x)^{\downarrow})$.

ii. $\Rightarrow$ *i.* For any $y \in A$ we note $y \in f^{\leftarrow}(f(y)^{\downarrow}) \subseteq g^{\leftarrow}(h(f(y))^{\downarrow})$, so $g(y) \leqslant h(f(y))$. ■

We now consider the reversed problem of theorem 2.4, which is considering when there exists $h\colon C \to B$ such that $f \circ h \leqslant g$, or $f \circ h \geqslant g$, as in the following diagram.



Theorem 2.5. *Let A, B, C be ordered sets with mappings $f\colon B \to A$ and $g\colon A \to C$. Then*

i. There exists $h\colon C \to B$ such that $f \circ h \leqslant g$;

ii. $f^{\leftarrow}(g(x)^{\downarrow})$ is non-empty for all $x \in B$;

are equivalent statements.

Proof. *i.* $\Rightarrow$ *ii.* We have $h(x) \in f^{\leftarrow}(g(x)^{\downarrow})$ for all $x \in C$ as $f(h(x)) \leqslant g(x)$.

ii. $\Rightarrow$ *i.* Define $h\colon C \to B$ by relating each $x \in C$ with a $h(x) \in f^{\leftarrow}(g(x)^{\downarrow})$. ■

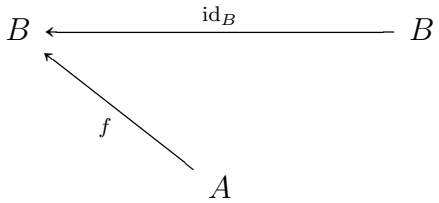
Corollary 2.6. *Let A, B be ordered sets with isotone mapping $f\colon A \to B$. Then*

i. There exists $h\colon B \to A$ such that $h \circ f \leqslant \mathrm{id}_B$;

ii. $f^{\leftarrow}(x^{\downarrow})$ is a down-set of A for all $x \in B$;

are equivalent statements.

Proof. Follows by applying theorem 2.5 on the diagram



as theorem 2.3 ensures that the down-sets are non-empty. ■

2.2 Residuated Mappings

We now turn our eyes to residuated mappings, which are of our main concern in this chapter. We first consider quasi-residuated mappings.

Definition 2.7. An isotone mapping $f\colon A \to B$ which satisfies *i.* or *ii.* of corollary 2.6 is said to be quasi-residuated.

Proposition 2.8. *A mapping $f\colon A \to B$ is quasi-residuated if and only if it is isotone and $f^{\leftarrow}(y^{\downarrow})$ is non-empty for all $y \in B$.*

Proof. Follows immediately by theorem 2.5. ■

We now characterise the mappings, from which the results of theorem 2.3 are more strict, in the sense that the pre-image of any principal down-set also yields a principal down-set.

Theorem 2.9. *Let A, B be ordered sets and $f\colon A \to B$. Then*

i. f is isotone and there exists an isotone mapping $h\colon B \to A$ such that $h \circ f \geqslant \mathrm{id}_A$ and $f \circ h \leqslant \mathrm{id}_B$;

ii. $f^{\leftarrow}(x^{\downarrow})$ is a principal down-set of A for all $x \in B$;

are equivalent statements.

Proof. *i.* $\Rightarrow$ *ii.* By theorem 2.4 we have $f^{\leftarrow}(x^{\downarrow}) \subseteq h(x)^{\downarrow}$ for all $x \in B$. Furthermore, we have for $y \in h(x)^{\downarrow}$ that $y \leqslant h(x)$, so $f(y) \leqslant f(h(x)) \leqslant \mathrm{id}_B(x) = x$, so $y \in f^{\leftarrow}(x^{\downarrow})$ by corollary 2.6, so $f^{\leftarrow}(x^{\downarrow}) = h(x)^{\downarrow}$.

ii. $\Rightarrow$ *i.* For any $x \in B$ there exists a unique $y \in A$ such that $f^{\leftarrow}(x^{\downarrow}) = y^{\downarrow}$. Thus, we construct $h\colon B \to A$ by defining $h(x) = y$, so $f^{\leftarrow}(x^{\downarrow}) = h(x)^{\downarrow}$, so $f(h(x)) \leqslant x$ for any $x \in B$. Lastly, we have $h(f(z)) \geqslant z$ for any $z \in A$ as $z \in f^{\leftarrow}(f(z)^{\downarrow}) = h(f(z))^{\downarrow}$. ■

Definition 2.10. A mapping $f\colon A \to B$ which satisfies *i.* or *ii.* of theorem 2.9 is a residuated mapping.

The following result shows uniqueness of the h given in theorem 2.9.

Proposition 2.11. *Let $f\colon A \to B$ be residuated. Then the isotone mapping $h\colon B \to A$ such that $h \circ f \geqslant \mathrm{id}_A$ and $f \circ h \leqslant \mathrm{id}_B$ is unique.*

Proof. Assume that $h, h^* : B \to A$ both satisfies the given properties. Then

$$h = \mathrm{id}_A \circ h \leqslant (h^* \circ f) \circ h = h^* \circ (f \circ h) \leqslant h^* \circ \mathrm{id}_B = h^*$$

and similarly

$$h = h \circ \mathrm{id}_B \geqslant h \circ (f \circ h^*) = (h \circ f) \circ h^* \geqslant \mathrm{id}_A \circ h^* = h^*$$

proving that $h = h^*$. ■

Definition 2.12. Let $f\colon A \to B$ be a residuated mapping. Then the unique h given by theorem 2.9 is called the residual of f, and denoted f^+.

The following proposition provides a necessary and sufficient condition for a mapping to be residuated.

Proposition 2.13. *A mapping $f\colon A \to B$ is residuated if and only if $\{x \in A \mid f(x) \leqslant y\}$ is non-empty and admits a greatest element for all $y \in B$. Furthermore, when the residual of f exists, then*

$$f^+(y) = \max\{x \in A \mid f(x) \leqslant y\}$$

for all $y \in B$

Proof. Follows by the proof of theorem 2.9. ■

We have the following important property of residuated mappings, which will see much use.

Proposition 2.14. *Let $f\colon A \to B$ be a residuated mapping. Then*

i. $f \circ f^+ \circ f = f$;

ii. $f^+ \circ f \circ f^+ = f^+$.

Proof. *i.* Since $f^+ \circ f \geqslant \mathrm{id}_A$ and $f \circ f^+ \leqslant \mathrm{id}_B$, and by the isotonocity of f, we have $f \circ f^+ \circ f \geqslant f \circ \mathrm{id}_A = f$ and $f \circ f^+ \circ f \leqslant \mathrm{id}_B \circ f = f$.

ii. Follows analogously to the proof of *i*. ■

Theorem 2.15. *Let $f\colon A \to B$ be a residuated mapping. Then*

i. $f^+ \circ f = \mathrm{id}_A$;

ii. f *is injective;*

iii. f^+ *is surjective;*

iv. For any set C, and mappings $g, h\colon C \to A$, then $f \circ g = f \circ h \Rightarrow g = h$;

are equivalent statements. Similarly,

v. $f \circ f^+ = \mathrm{id}_B$;

vi. f *is surjective;*

vii. f^+ *is injective;*

viii. For any set C*, and mappings* $g, h\colon C \to A$*, then* $g \circ f = h \circ f \Rightarrow g = h$*;*

are equivalent statements.

Proof. *i.* $\Rightarrow$ *iii.* Follows immediately.

iii. $\Rightarrow$ *i.* For any $x \in A$ there exists $y \in B$ such that $f^+(y) = x$, so

$$x = f^+(y) = (f^+ \circ f \circ f^+)(y) = (f^+ \circ f)(f^+(y)) = (f^+ \circ f)(x)$$

by proposition 2.14.

i. $\Rightarrow$ *ii.* We have that $f(x) = f(y)$ implies $x = f^+(f(x)) = f^+(f(y)) = y$.

ii. $\Rightarrow$ *iv.* If $f \circ g = f \circ h$ then $f(g(x)) = f(h(x))$ for any $x \in C$, and the result then follows by the injectivity of f.

iv. $\Rightarrow$ *i.* As $f \circ f^+ \circ f = f = f \circ \mathrm{id}_A$ then $f^+ \circ f = \mathrm{id}_A$.

The equivalencies between *v.* to *viii.* are proved analogously. ■

Theorem 2.16. *Let* A, B, C *be ordered sets, and let* $f\colon A \to B, g\colon B \to C$ *be residuated mappings. Then* $g \circ f$ *is residuated with residual* $(g \circ f)^+ = f^+ \circ g^+$.

Proof. As f and g are residuated, we have

$$(f^+ \circ g^+) \circ (g \circ f) \geqslant f^+ \circ \mathrm{id}_B \circ f = f^+ \circ f \geqslant \mathrm{id}_A$$

and

$$(g \circ f) \circ (f^+ \circ g^+) \leqslant g \circ \mathrm{id}_B \circ g^+ = g \circ g^+ \leqslant \mathrm{id}_C$$

Lastly, as composing isotone functions yields isotone functions, we have that $g \circ f$ and $f^+ \circ g^+$ are isotone, and by the uniqueness of residuals we must have that $(g \circ f)^+ = f^+ \circ g^+$. ■

Proposition 2.17. *Let* A *be an ordered set and* $f\colon A \to A$ *be a residuated mapping. Then*

i. For any $p, n \in \mathbb{N}$ *then* $f^p = f^{p+n}$ *if and only if* $(f^+)^p = (f^+)^{p+n}$*;*

ii. $f \leqslant \mathrm{id}_A$ *if and only if* $f^+ \geqslant \mathrm{id}_A$*;*

iii. $f \geqslant \mathrm{id}_A$ *if and only if* $f^+ \leqslant \mathrm{id}_A$.

Proof. *i*. By theorem 2.16 we have by induction that the residual of f^p is $(f^+)^p$, and the result then follows by uniqueness of residuals.

ii. If $f \leqslant \mathrm{id}_A$ then applying f^+ and by its isotonicity, and by it being the residual of f, we have $\mathrm{id}_A \leqslant f \circ f^+ \leqslant f^+$, with the converse argument following analogously.

iii. Follows analogously to *ii*. ■

Definition 2.18. Let A, B be ordered sets. If there exists an isotone bijection $f\colon A \to B$ such that f^{-1} is isotone, then A and B are order-isomorphic.

We note that two ordered sets $(A, \leqslant_1), (B, \leqslant_2)$ are order-isomorphic if and only if there exists a surjective isotone mapping $f\colon A \to B$.

Definition 2.19. An ordered semigroup is a semigroup S on which there exists an ordering $\leqslant$ such that if $y \leqslant z$ for $y, z \in S$, then $xy \leqslant xz$ and $yx \leqslant zx$ for any $x \in S$.

Definition 2.20. Two ordered semigroups A, B are isomorphic if and only if there exists a semigroup homomorphism $f\colon A \to B$ which is also an order isomorphism. Similarly, A, B are anti-isomorphic if and only if there exists a semigroup homomorphism which is a dual order isomorphism.

Theorem 2.21. *Let E be an ordered set, and let $\mathrm{Res}(E)$ denote the set of residuated mappings $f\colon E \to E$, and let $\mathrm{Res}^+(E)$ denote the set of residuals of $\mathrm{Res}(E)$. Then $\mathrm{Res}(E)$ and $\mathrm{Res}(E^+)$ are anti-isomorphic ordered semigroups.*

Proof. By theorem 2.16 we have that composing residuated mappings yields new residuated mappings, and thus both $\mathrm{Res}(E)$ and $\mathrm{Res}^+(E)$ are semigroups. Furthermore, these are ordered under the ordering $f \leqslant g$ if and only if $f(x) \leqslant g(x)$ for all $x \in E$ since any residuated map is also isotone.

Now consider the mapping $(+)\colon \mathrm{Res}(E) \to \mathrm{Res}^+(E)$ such that $f \mapsto f^+$. This is well-defined as residuals are unique, and since $f \leqslant g$ if and only if $g^+ \leqslant f^+$ we have that $(+)$ is an anti-isomorphism. ■

2.3 Closure Mappings

We now consider closure mappings, and their relation to residuated mappings.

Definition 2.22. An isotone mapping $f\colon A \to A$ such that $f = f \circ f \geqslant \mathrm{id}_A$ is a closure mapping. Similarly, if f satisfies $f = f \circ f \leqslant \mathrm{id}_A$, then f is a dual closure mapping.

Theorem 2.23. *Let A be an ordered set, and $f\colon A \to A$. Then*

i. f is a dual closure mapping;

ii. $f^{\leftarrow}(x^{\downarrow}) = f^{\leftarrow}(f(x)^{\downarrow})$ for all $x \in A$;

are equivalent statements.

Proof. *i.* $\Rightarrow$ *ii.* As $f(x) \leqslant f(x)$ and $f(x) \leqslant x$ for all $x \in A$ we have $f^{\leftarrow}(f(x)^{\downarrow}) \subseteq f^{\leftarrow}(x^{\downarrow})$. Conversely, if $y \in f^{\leftarrow}(x^{\downarrow})$ then $f(y) \leqslant x$, but as $f = f \circ f$, and as f is isotone we have $f(y) \leqslant f(x)$ so $y \in f^{\leftarrow}(f(x)^{\downarrow})$.

ii. $\Rightarrow$ *i.* As $x \in f^{\leftarrow}(x^{\downarrow}) = f^{\leftarrow}(f(x)^{\downarrow}) = f^{\leftarrow}(f(f(x))^{\downarrow})$ for all $x \in A$, so both $f(x) \leqslant x$ and $f(x) \leqslant (f \circ f)(x)$, so $f \leqslant \mathrm{id}_A$ and $f \leqslant f \circ f$. Furthermore, f is also isotone as $y \leqslant x$ implies $f(y) \leqslant y \leqslant x$, which yields $f(y) \leqslant f(x)$. Lastly, as $f \leqslant \mathrm{id}_A$ we also have $f \circ f \leqslant f$, so $f \circ f = f$. ■

Theorem 2.24. *Let A be an ordered set. Then $f\colon A \to A$ is a closure mapping if and only if there exists an ordered set B and a residuated mapping $g\colon A \to B$ satisfying $f = g^{+} \circ g$.*

Proof. Suppose that $f\colon A \to A$ is a closure mapping, and let F be the equivalence relation xFy if and only if $f(x) = f(y)$. Now define the relation $\preccurlyeq$ on A/F such that $[x] \preccurlyeq [y]$ if and only if $f(x) \leqslant f(y)$, which is an ordering since $\leqslant$ is an ordering. Furthermore, let $\natural\colon A \to A/F$ denote the quotient mapping, which is isotone, as

$$x \leqslant y \Rightarrow f(x) \leqslant f(y) \Leftrightarrow [x] \preccurlyeq [y]$$

We then have for any $x \in A$ that $f(x) \in [x]$ as $f(x) = f(f(x))$. Furthermore, $f(x)$ is the greatest element of $[x]$, since if not, there would exist $y \in [x]$ such that $y > f(x)$, but by the isotonicity of f we have $f(y) > f(f(x)) = f(x)$, but $f(y) = f(x)$ as $y, x \in [x]$.

We then define the mapping $h(x)\colon A/F \to A$ such that $h([x]) = f(x)$, which is well-defined by the previous arguments. It then follows, that both

$$(h \circ \natural)(x) = h(\natural(x)) = h([x]) = f(x) \geqslant x$$

$$(\natural \circ h)([x]) = \natural(f(x)) = [f(x)] = [x]$$

and as such, $\natural$ is a residuated mapping, with residual $h = \natural^{+}$, and $f = \natural^{+} \circ \natural$.

For the converse, suppose $g\colon A \to B$ is residuated, so $g^+ \circ g \geqslant \mathrm{id}_A$. Lastly, as $g = g \circ g^+ \circ g$, we have $g^+ \circ g = (g^+ \circ g) \circ (g^+ \circ g)$, so $g^+ \circ g\colon A \to A$ is a closure mapping. ■

Theorem 2.25. *Let A be an ordered set and $f\colon A \to A$ a residuated mapping. Then*

i. f is a closure mapping;

ii. f^+ is a dual closure mapping;

iii. $f = f^+ \circ f$;

iv. $f^+ = f \circ f^+$;

are equivalent statements. Similarly,

v. f is a dual closure mapping;

vi. f^+ is a closure mapping;

vii. $f = f \circ f^+$;

viii. $f^+ = f^+ \circ f$;

are equivalent statements.

Proof. *i.* $\Leftrightarrow$ *ii.* Follows by proposition 2.17.

i. $\Rightarrow$ *iii.* Holds, as $f^+ \circ f = f^+ \circ f \circ f \geqslant \mathrm{id}_A \circ f = f$ and $f = f \circ f^+ \circ f \geqslant \mathrm{id}_A \circ f^+ \circ f = f^+ \circ f$ yields $f = f \circ f^+$.

iii. $\Rightarrow$ *iv.* By proposition 2.14 we have $f \circ f^+ = f^+ \circ f \circ f^+ = f^+$.

iv. $\Rightarrow$ *ii.* Proposition 2.14 yields $f^+ \circ f^+ = f \circ f^+ \circ f \circ f^+ = f \circ f^+ = f^+$, and $f^+ = f \circ f^+ \leqslant \mathrm{id}_A$. ■

2.4 Semilattices and Lattices

In the following we consider an ordered set E and for $x \in E$ we let $\iota_x\colon x^\downarrow \hookrightarrow E$ denote the canonical injection of the given principal down-set, which is then isotone by definition.

Theorem 2.26. *Let E be an ordered set. Then*

i. for any $x \in E$ then $\iota_x\colon x^\downarrow \hookrightarrow E$ is residuated;

ii. the intersection of any two principal down-sets of E is a principal down-set of E;

are equivalent statements.

Proof. By proposition 2.13 we have that ι_x is residuated if and only if there exists some $\alpha = \max\{z \in x^{\downarrow} \mid z = \iota_x(z) \leqslant y\}$, which is equivalent to the existence of some $\alpha^{\downarrow} = x^{\downarrow} \cap y^{\downarrow}$. ■

Definition 2.27. An ordered set E which satisfies *i.* or *ii.* of theorem 2.26 is said to be an $\wedge$-semilattice, and we let $x \wedge y$ denote the element α such that $x^{\downarrow} \cap y^{\downarrow} = \alpha^{\downarrow}$.

Note that $x \wedge y$ is to be read as the meet of x and y, and E is then read as a meet semilattice. An example of a $\wedge$-semilattice is the natural numbers ordered under divisibility, where $m \wedge n = \text{hcf}\{m, n\}$.

The following result shows how any $\wedge$-semilattice can be characterised in a purely algebraic way, or how any commutative idempotent semigroup can be seen as a $\wedge$-semilattice.

Theorem 2.28. *A set E can be given the structure of a $\wedge$-semilattice if and only if it can be equipped with a binary operation $(x, y) \to x \triangle y$ which is associative, commutative and idempotent.*

Proof. The right implication follows trivially, and we now consider the converse. Assume that E is an abelian idempotent semigroup under the binary operation $(x, y) \to x \triangle y$. Now define the relation R on E such that

$$xRy \Leftrightarrow x \triangle y = x$$

which is an ordering due to the following arguments. Firstly, $x \triangle x = x$ so xRx, and similarly if xRy and yRx then

$$x = x \triangle y = y \triangle x = y$$

and lastly as xRy and yRz then

$$x = x \triangle y = x \triangle (y \triangle z) = (x \triangle y) \triangle z = x \triangle z$$

so xRz, and thus we write $\leqslant$ if instead of R. Furthermore, consider $x, y \in E$ so

$$x \triangle y = x \triangle x \triangle y = x \triangle y \triangle x$$

so $x \triangle y \leqslant x$, and switching the roles of x and y we obtain $x \triangle y \leqslant y$, so $x \triangle y \in x^{\downarrow} \cap y^{\downarrow}$, implying that the intersection of any principal down-sets is non-empty. Lastly, we have

$$\begin{aligned} z \in x^{\downarrow} \cap y^{\downarrow} &\Rightarrow z \leqslant x,\ z \leqslant y \\ &\Rightarrow z = z \triangle x,\ z = z \triangle z \triangle y \\ &\Rightarrow z = z \triangle y = z \triangle x \triangle y \\ &\Rightarrow z \leqslant x \triangle y \end{aligned}$$

proving that $x\triangle$ is the maximal element of $x^{\downarrow} \cap y^{\downarrow}$, so E is a $\wedge$-semilattice where $x \wedge y = x \triangle y$. ■

Definition 2.29. Let E be an ordered set and $F \subseteq E$ be non-empty. Then $x \in E$ is a lower of bound of F if and only if $x \leqslant y$ for any $y \in F$, and x is the greatest lower bound of F if $z \leqslant x$ for all lower bounds z of F, denoted $x = \inf F$. Similarly, if $x \geqslant y$ for any $y \in F$ it is said to be an upper bound of F, and the least upper bound of F if $z \geqslant x$ for all upper bounds z of F, denoted $y = \sup F$.

We now provide a short characterisation of the dual notions of a $\wedge$-semilattice, which is a $\vee$-semilattice, read as a join semilattice. By the dual of theorem 2.26 then the intersection of any principal up-sets results in a principal up-set, so we denote $x \vee y$ as the element β such that $x^{\uparrow} \cap y^{\uparrow} = \beta^{\uparrow}$. There is of course also a dual result to theorem 2.28 with respect to $\vee$-semilattices, where the proof uses the relation xSy if and only if $x \triangle y = y$.

Definition 2.30. An ordered set that is both a $\wedge$-semilattice and a $\vee$-semilattice is called a lattice.

Proposition 2.31. *Let E be a semilattice, and let $F = \{x_1, \ldots, x_n\}$ be a finite non-empty subset of E. Then*

i. if E is a $\wedge$-semilattice, then $\inf F = x_1 \wedge \ldots \wedge x_n$;

ii. if E is a $\vee$-semilattice, then $\sup F = x_1 \vee \ldots \vee x_n$.

Proof. Follows as $\inf\{x_1, x_2\} = x_1 \wedge x_2$ and $\sup\{x_1, x_2\} = x_1 \vee x_2$. ■

Definition 2.32. Let E, F be $\vee$-semilattices. Then $f\colon E \to F$ is a $\vee$-morphism if $f(x \vee y) = f(x) \vee f(y)$ for any $x, y \in E$, with a $\wedge$-morphism defined dually. If f is both a $\vee$-morphism and a $\wedge$-morphism, it is a lattice morphism. Furthermore, $f : E \to F$ is said to be a complete $\vee$-morphism if for every family $(x_\alpha)_{\alpha \in I}$ of elements in E such that $\bigvee_I x_\alpha$ exists in E, then $\bigvee_I f(x_\alpha)$ exists in F and

$$f\left(\bigvee_I x_\alpha\right) = \bigvee_I f(x_\alpha)$$

with a complete $\wedge$-morphism being defined dually.

Theorem 2.33. *Let E, F be $\vee$-semilattices, and let $f\colon E \to F$ be residuated. Then f is a complete $\vee$-morphism.*

Proof. Consider a family of elements $(x_\alpha)_{\alpha\in I}$ in E such that $x = \bigvee_I x_\alpha$ is an element of E. We then have $f(x) \geqslant f(x_\alpha)$ for any $\alpha \in I$. Now suppose there exists $y \in F$ such that $y \geqslant f(x_\alpha)$ for any $\alpha \in I$. Then

$$\begin{aligned} f^+(y) \geqslant f^+(f(x_\alpha)) \geqslant x_\alpha \quad &\Rightarrow \quad f^+(y) \geqslant \bigvee_I x_\alpha = x \\ &\Rightarrow \quad y \geqslant f(f^+(y)) \geqslant f(x) \end{aligned}$$

and so $\bigvee_I f(x_\alpha)$ exists and equals $f(x)$. ■

Definition 2.34. Lattices E, F are isomorphic if they are isomorphic as ordered sets.

Theorem 2.35. *Lattices E, F are isomorphic if and only if there exists a bijection $f\colon E \to F$ which is a $\vee$-morphism.*

Proof. Assume first that $E \simeq L$, so there exists a residuated bijection, which by theorem 2.33 is a $\vee$-morphism. For the converse relation assume there exists a bijection $f\colon E \to F$ which is a $\vee$-morphism. We then have for $x, y \in E$, that

$$x \leqslant y \Leftrightarrow y = x \vee y \Leftrightarrow f(y) = f(x \vee y) = f(x) \vee f(y) \Leftrightarrow f(x) \leqslant f(y)$$

proving the claim. ■

From proposition 2.31 we saw that for any $\wedge$-semilattice that the infimum exists for any finite subset, and we now extend this notion to infinite subsets.

Definition 2.36. A $\wedge$-semilattice E is $\wedge$-complete if for any subset $A = \{x_\alpha \mid \alpha \in I\}$ then A has an infimum in E, denoted $\inf_E A$ or $\bigwedge_I x_\alpha$. We define a $\vee$-complete lattice dually. A lattice that is both $\wedge$-complete and $\vee$-complete is a complete lattice.

We immediately derive the following result.

Proposition 2.37. *Every complete lattice has a top and a bottom element.*

Proof. The top element is $\sup_L L$ and the bottom element is $\inf_L L$, which exists by definition. ■

Similarly to theorem 2.28 we can also characterise lattices by the following result.

Theorem 2.38. *A set E can be given the structure of a lattice if and only if it can be equipped with two binary operations $(x, y) \to x \triangle y$ and $(x, y) \to x \triangledown y$ such that $(E, \triangle)$ and $(E, \triangledown)$ are abelian semigroups, and for any $x, y \in E$ then $x \triangle (x \triangledown y) = x = x \triangledown (x \triangle y)$.*

Proof. Suppose that E is a lattice, so there exists two binary operations satisfying the conditions, specifically $(x, y) \to x \wedge y$ and $(x, y) \to x \vee y$, and for any $x, y \in E$ we have $x \leqslant x \vee y$, so $x \wedge (x \vee y) = x$, and similarly $x \wedge y \leqslant x$, so $x \vee (x \wedge y) = x$.

Now, consider the converse and suppose E has two binary operations $\triangle, \triangledown$ satisfying the conditions. We then obtain both

$$x \triangledown x = x \triangledown (x \triangle (x \triangledown x)) = x$$
$$x \triangle x = x \triangle (x \triangledown (x \triangle x)) = x$$

so by theorem 2.28 we have that can be seen as a semilattice under both $\triangle$ and $\triangledown$. Lastly, we show that $x \triangle y = x$ if and only if $x \triangledown y = y$. Now,

$$x \triangle y = x \Rightarrow y = (x \triangle y) \triangledown y = x \triangledown y$$
$$x \triangledown y = y \Rightarrow x = x \triangle (x \triangledown y) = x \triangle y$$

proving that E is a lattice under the ordering

$$x \leqslant y \Leftrightarrow x \triangle y = x \Leftrightarrow x \triangledown y = y$$

■

Theorem 2.39. *Let E be a complete lattice, and $f \colon E \to E$ be isotone. Then f has a fixed point.*

Proof. As L is complete we have that $0 \in f(x)^{\downarrow}$, and there exists $\alpha = \sup_L f(x)^{\downarrow}$, and for any $x \in f(x)^{\downarrow}$ then $x \leqslant \alpha$, so $x \leqslant f(x) \leqslant f(\alpha)$, so $\alpha \leqslant f(\alpha)$, and $f(\alpha) \leqslant f(f(\alpha))$, so $f(\alpha) \in f(x)^{\downarrow}$, so $f(\alpha) \leqslant \alpha$ which finally implies $f(\alpha) = \alpha$. ■

We know give the central theorem of this chapter which entirely characterises residuated mappings between complete lattices.

Theorem 2.40. *Let E, F be complete lattices, and let $f \colon E \to F$. Then f is residuated if and only if it is a complete $\vee$-morphism and $f(0_E) = 0_F$.*

Proof. Assume that f is residuated, so by theorem 2.33 we have that it is a complete $\vee$-morphism. Furthermore, if $f(x) = 0_F$, then $x \geqslant 0_E$, so $f(x) = 0_F \geqslant f(0_E)$, so $f(0_E) = 0_F$ since we would otherwise obtain a contradiction.

Now let f be a complete $\vee$-morphism such that $f(0_E) = 0_F$, and fix $x \in F$ and consider $f^{\leftarrow}(x^{\downarrow})$. Then $\alpha = \sup f^{\leftarrow}(x^{\downarrow})$ exists since E is complete. Now, consider any $y \in E$ such that $y \leqslant \alpha \Leftrightarrow y \vee \alpha = \alpha$, so by $\vee$-completeness of f we obtain

$$f(\alpha) = f(y \vee \alpha) = f(y) \vee f(\alpha) \Leftrightarrow f(y) \leqslant f(\alpha)$$

so $f(y) \leqslant f(\alpha) \leqslant x$, so $y \in f^{\leftarrow}(x^{\downarrow})$ implying that $\alpha^{\downarrow} = f^{\leftarrow}(x^{\downarrow})$, so f is residuated by theorem 2.9. ■

2.5 Boolean Semimodules

In the following section we seek to introduce a specific Boolean algebra, and their residuated mappings, which will be of great use when considering group testing. We could have continued this section in the same vein as the previous, introducing distributivity and complementation in the abstract case, but refrain from doing so, and only introduce the definitions necessary for our scope. For a general reference on Boolean algebras see [4, 5].

The section is based on [6, 7], and we begin by considering the binary Boolean semiring.

Definition 2.41. $\mathbb{B}_2 = (\{0,1\}, +, \cdot)$ denotes the Boolean semiring, with addition and multiplication defined by $x + y = \max\{x, y\}$ and $xy = \min\{x, y\}$ under the natural ordering $0 \leqslant 1$.

An equivalent way to define the ordering of $\mathbb{B}_2$ is

$$x \leqslant y \Leftrightarrow x + y = y$$

for any $x, y \in \mathbb{B}_2$. We also define a negation mapping by

$$^- : \begin{cases} \mathbb{B}_2 \longrightarrow \mathbb{B}_2 \\ 0 \longmapsto 1 \\ 1 \longmapsto 0 \end{cases}$$

Note that $^-$ is an involution, and also satisfies De Morgan's laws, summarised by the following result.

Proposition 2.42.

$$\overline{x + y} = \overline{x}\,\overline{y}, \quad \overline{xy} = \overline{x} + \overline{y} \tag{2.1}$$

for any $x, y \in \mathbb{B}_2$.

Proof. Follows by direct calculations. ■

Note that any totally ordered semiring S with universal lower and upper bounds, with addition and multiplication defined by $x + y = \max\{x, y\}$ and $x \cdot y = \min\{x, y\}$ can be represented as a Boolean algebra of subsets of S, by mapping any element $x \in S$ to $x^{\downarrow}$ [8].

In the following we consider semimodules over $\mathbb{B}_2$ and let $\mathbb{B}_2^n$ denote the set of n-tuples over $\mathbb{B}_2$, which will shown is our Boolean algebra of interest.

Proposition 2.43. *$\mathbb{B}_2^n$ is a semimodule over $\mathbb{B}_2$ under component-wise addition and component-wise scalar multiplication from $\mathbb{B}_2$.*

Proof. All axioms of $\mathbb{B}_2^n$ being a semimodule are inherited from the properties of $\mathbb{B}_2$ being a commutative semiring [8]. ■

We then define multiplication over $\mathbb{B}_2^n$ as component-wise multiplication over $\mathbb{B}_2$, and also naturally extend the ordering $\leqslant$ on $\mathbb{B}_2$ to $\mathbb{B}_2^n$ as in (2.1). Lastly, we also extend our negation mapping onto $\mathbb{B}_2^n$ by component-wise negation. This then lets us view $\mathbb{B}_2^n$ as a complete lattice.

Theorem 2.44. *$(\mathbb{B}_2^n, \vee, \wedge)$ is a complete lattice for*

$$x \vee y = x + y$$
$$x \wedge y = x \cdot y$$

where $+$ is component-wise addition and $\cdot$ is component-wise multiplication, both over $\mathbb{B}_2$.

Proof. The results follows by theorem 2.38 where $\triangle = \wedge$ and $\bigtriangledown = \vee$. Furthermore, $\mathbb{B}_2^n$ is complete since it is finite, with the supremum of a subset being the join over all elements of the set, and the infimum being the meet over all elements of the set. ■

Specifically, $\mathbb{B}_2^n$ is known as a Boolean algebra, that is, a complemented distributive lattice in the sense that $\vee$ and $\wedge$ distributes over each other, and for every element $x \in \mathbb{B}_2^n$ there exists $y \in \mathbb{B}_2^n$ such that $x \vee y = 1$ and $x \wedge y = 0$, namely $y = \overline{x}$ [4, 5].

We now take well-known conventions of coding theory, and naturally impose these over $\mathbb{B}_2^n$.

Definition 2.45. The map

$$\delta\colon \begin{cases} \mathbb{B}_2^n \times \mathbb{B}_2^n & \longrightarrow & \mathbb{N} \\ (x, y) & \longmapsto & |\{i \in \{1, \dots, n\} \mid x_i \neq y_i\}| \end{cases}$$

is the Hamming distance, and

$$w\colon \begin{cases} \mathbb{B}_2^n & \longrightarrow & \mathbb{N} \\ x & \longmapsto & \delta(x, 0) \end{cases}$$

is the Hamming weight.

Note however that δ is not translation invariant as in the case of $\mathbb{F}_q$-linear codes, but it does satisfy the following property.

Proposition 2.46. *$\delta(x, y) = w(x + y) - w(x \cdot y)$ for any $x, y \in \mathbb{B}_2^n$.*

Proof. We have that

$$\begin{aligned} x_i \neq y_i \quad &\Rightarrow \quad x_i + y_i = 1 \quad \text{and} \quad x_i \cdot y_i = 0 \\ 0 = x_i = y_i \quad &\Rightarrow \quad x_i + y_i = 0 \quad \text{and} \quad x_i \cdot y_i = 0 \\ 1 = x_i = y_i \quad &\Rightarrow \quad x_i + y_i = 1 \quad \text{and} \quad x_i \cdot y_i = 1 \end{aligned}$$

which proves the claim. ■

In the following we introduce the notion of d-disjunctness, and show how this is equivalent to the matrix being injective when restricted to the Hamming ball of radius d centered on the zero element.

Definition 2.47. Let $d \in \mathbb{N}$ and $x \in \mathbb{B}_2^n$. Then

$$B_\delta(x, d) = \{z \in \mathbb{B}_2^n \mid \delta(x, z) \leqslant d\}$$

is the Hamming ball of radius d centered in x.

Definition 2.48. Let H be an $n \times k$ matrix over $\mathbb{B}_2$, and $d \in \mathbb{N}$.

i. H is d-Rev if for any $x \in \mathbb{B}_\delta(0, d)$ and $y \in \mathbb{B}_2^n$ such that $xH = yH$, then $x = y$.

ii. H is d-disjunct if for any $t \leqslant d$ and any set $T = \{x_1, \ldots, x_t\}$ of rows of H, then for any row $x \notin T$ we have $x \not\leqslant y_1 + \ldots + y_t$.

Note that we have defined d-disjunct in the transposed case, compared to most other literature which considers supersets of columns rather than rows. An equivalent way of stating d-disjunctness of a matrix is that for any $d + 1$ rows, with one of these rows being designated, is that there always exists one column with a 1 in the designated row, and 0s in the remaining d rows [9].

Theorem 2.49. *H is d-Rev if and only if it is d-disjunct.*

Proof. Assume, that H is d-Rev, and consider a sequence $h_{i_1}, \ldots, h_{i_t}$ of $t \leqslant d$ rows of H. There then exists $x \in B_\delta(0, t) \subseteq B_\delta(0, d)$ such that $xH = \sum_{j=1}^t h_{i_j}$. Consider $l \in [n] \backslash \{i_1, \ldots, i_t\}$, and define $y = e_l + x$, where e_l is the vector with 1 in the l'th position, and 0 elsewhere. Now, assume for contradiction, that $h_l \leqslant \sum_j^t h_{i_j}$, or equivalently, $\sum_j^t h_{i_j} = h_l + \sum_j^t h_{i_j}$. We then obtain

$$yH = (e_l + x)H = e_l H + xH = h_l + \sum_{j=1}^t h_{i_j} = \sum_{j=1}^t h_{i_j} = xH$$

which implies that $x = y$ as H is d-Rev, which is a contradiction as $x \neq y$.

Now, assume that H is d-disjunct, and consider $x \in B_\delta(0, d)$ and $y \in \mathbb{B}_2^n$ such that $xH = yH$, and finally assume for contradiction that $x \neq y$. We

may without loss of generality assume that there exists $i \in \operatorname{supp}(y)$ such that $i \notin \operatorname{supp}(x)$, since if not, then $y \in B_\delta(0, d)$, and the roles of x and y could then be interchanged. By the d-disjunct property of H we then have both $h_i \not\leqslant xH$ and $h_i \leqslant yH = xH$, which contradicts our assumption of $x \neq y$. ■

In the following we characterise residuated mappings from $\mathbb{B}_2^n$ to $\mathbb{B}_2^k$.

Theorem 2.50. *A mapping $f\colon \mathbb{B}_2^n \to \mathbb{B}_2^k$ is residuated if and only if $f(x) = xH$ for all $x \in \mathbb{B}_2^n$, where $H \in \mathbb{B}_2^{n\times k}$.*

Proof. Follows directly by theorem 2.40. ■

Note that as there only exists one basis for $\mathbb{B}_2^n$ we view the canonical representation of H with respect to the natural ordering of the basis vectors.

Theorem 2.51. *Let $f\colon \mathbb{B}_2^n \to \mathbb{B}_2^k$ be a residuated mapping represented by the matrix H. Then the residual $f^+\colon \mathbb{B}_2^k \to \mathbb{B}_2^n$ is given by $f^+(y) = \overline{\overline{y}H^\top}$.*

Proof. By theorem 2.9 we prove that f^+ is isotone, and satisfies $f^+ \circ f \geqslant \mathrm{id}_A$ and $f \circ f^+ \leqslant \mathrm{id}_B$. Firstly,

$$f^+(y)_i = \overline{\sum_{j\leqslant k} \overline{y_j} H_{ij}} = \prod_{j\leqslant k} \overline{\overline{y_j} H}_{ij} = \prod_{j\leqslant k} \left(\overline{H}_{ij} + y_j\right)$$

which then implies isotonicity of f^+. Now, if $y = f(x)$ then $y_j = \sum_{l\leqslant n} x_l H_{lj}$, so

$$\begin{aligned} f^+(f(x))_i = \prod_{j\leqslant k} \left(\overline{H}_{ij} + \sum_{l\leqslant n} x_l H_{ij}\right) &\geqslant \prod_{j\leqslant k} \left(\overline{H}_{ij} + x_i H_{ij}\right) \\ &\geqslant \prod_{j\leqslant k} \left(x_i \overline{H}_{ij} + x_i H_{ij}\right) \\ &= \prod_{j\leqslant k} x_i \left(\overline{H}_{ij} + H_{ij}\right) \\ &= x_i \end{aligned}$$

and similarly,

$$\begin{aligned} f(f^+(y))_j = \sum_{i\leqslant n} f^+(y)_i H_{ij} = \sum_{i\leqslant n} \left(\prod_{l\leqslant k} \left(\overline{H}_{ij} + y_l\right)\right) H_{ij} \\ \leqslant \sum_{i\leqslant n} \left(\overline{H}_{ij} + y_j\right) H_{ij} \\ = \sum_{i\leqslant n} y_j H_{ij} \\ \leqslant y_j \end{aligned}$$

finalising the proof. ■

Finally we give the following result regarding equivalent properties of these residuated mappings, which will be pivotal when discussing group testing.

Theorem 2.52. *[6] Let $f\colon \mathbb{B}_2^n \to \mathbb{B}_2^k$ be a residuated mapping represented by the $n \times k$-matrix H. Then*

i. H is d-rev;

ii. $f^+(f(x)) = x$ for all $x \in B_\delta(0,d)$;

iii. $B_\delta(0,d) \subseteq \mathrm{im}(f^+)$;

iv. $B_\delta(1,d) \subseteq \mathrm{col}(H)$;

are equivalent statements.

Proof. *i.* $\Rightarrow$ *ii.* Consider $x \in B_\delta(0,d)$ and suppose for contradiction that $f^+(f(x)) \neq x$. Then $f(f^+(f(x))) = f(x)$, but $f^+(f(x)) \neq x$ contradicting H being d-rev.

ii. $\Rightarrow$ *iii.* Follows trivially.

iii. $\Rightarrow$ *ii.* Consider $x \in B_\delta(0,d)$, so there exists $y \in \mathbb{B}_2^k$ such that $f^+(y) = x$, so $f(x) = f(f^+(y))$, and finally $f^+(f(x)) = f^+(y) = x$.

ii. $\Rightarrow$ *i.* Consider $x \in B_\delta(0,d)$ and $y \in \mathbb{B}_2^n$ such that $f(x) = f(y)$, but assume for contradiction that $x \neq y$. Without loss of generality there then exists $i \in \mathrm{supp}(y)$ such that $i \notin \mathrm{supp}(x)$. Letting e_i denote the i'th standard basis vector we see $e_i \in B_\delta(0,d)$ and $e_i \leqslant y$, while $e_i \not\leqslant x$. Then

$$f(e_i) = h_i \leqslant f(y) = f(x)$$

where h_i denotes the i'th row of H. Finally

$$f^+(f(e_i)) = e_i \leqslant f^+(f(y)) = f^+(f(x)) = x$$

which contradicts $e_i \not\leqslant x$.

iii. $\Rightarrow$ *iv.* We have $\overline{B_\delta(1,d)} = B_\delta(0,d)$, so

$$B_\delta(1,d) \subseteq \overline{\mathrm{im}(f^+)} = \left\{ yH^\top \;\middle|\; y \in \mathbb{B}_2^k \right\} = \left\{ Hy^\top \;\middle|\; y \in \mathbb{B}_2^k \right\} = \mathrm{col}(H)$$

iv. $\Rightarrow$ *iii.* Follows as

$$B_\delta(0,d) = \overline{B_\delta(1,d)} \subseteq \overline{\mathrm{col}(H)} = \left\{ \overline{yH^\top} \;\middle|\; y \in \mathbb{B}_2^k \right\} = \mathrm{im}(f^+)$$

finalising the proof. ∎

3 | Group Testing

Group testing is a method in which one aims to determine a subset of *infected* or *positive* items out of a larger set of items, by testing multiple items at once. This can be done sequentially, or adaptively, where one test is completed at a time, and based on this result one can then infer what the optimal next test is [9]. Secondly, one can consider non-adaptive group testing, in which all tests are conducted simultaneously. In the following we consider only non-adaptive group testing, and as such, when referring to group testing we imply the latter of the methods.

Specifically, a d-disjunct $n \times k$ matrix H, will represent a group testing scheme, which allows the testing of n items using k tests, and if there is at most d infected items the scheme will then be able to determine the infected items correctly. The interpretation of H is that each column represents a test, specifically, column j describes the j'th test, and if the i'th entry of the j'th column is a 1, this implies that item i is included in the j'th test. The tests are then carried out by computing $y = xH$, where $x \in \mathbb{B}_2^n$ is unknown, and $x_i = 1$ if item i is infected, and $x_i = 0$ if it is not. Lastly, one attempts to infer x given y.

We now give the formal definition of a group testing scheme based on [4, 5] in the context of residuated mappings, which we developed previously.

Definition 3.1. Let $n, k, d \in \mathbb{N}$. An (n, k, d)-group testing scheme is a residuated mapping $f\colon \mathbb{B}_2^n \to \mathbb{B}_2^k$, and a decoder $g\colon \mathbb{B}_2^k \to \mathbb{B}_2^n$ such that $g(f(x)) = x$ for all $x \in B_\delta(0, d)$.

By the previously given residuation theory we can immediately state an entire class of group testing schemes based on d-disjunct matrices.

Theorem 3.2. *Let H be an $n \times k$ d-disjunct matrix. Then the residuated mapping f represented by H, and its residual f^+, is a (n, k, d)-group testing scheme.*

Proof. The proof follows by theorem 2.52. ■

Obviously one is interested in maximising d and n, while minimising k, and while also showcasing explicit constructions. These matters are considered in the proceeding chapter, where we instead turn our attention briefly to noisy group testing.

The following is based on [7]. Given a (n, k, d)-group testing scheme with mappings f, f^+ and a resulting test $y \in \mathbb{B}_2^k$ the problem of error detection is determining if there exists an $x \in \mathbb{B}_2^n$ such that $f(x) = y$, that is, is there

a distribution of the n items, that when tested under f actually culminates in the resultant test. The following result provides a polynomial method to determine this.

Proposition 3.3. *Let $f\colon \mathbb{B}_2^n \to \mathbb{B}_2^k$ be a residuated mapping. Then*

$$y \in \operatorname{im}(f) \Leftrightarrow f(f^+(y)) = y$$

Proof. Follows by proposition 2.14. ■

With respect to the error correction capabilities of a scheme we have seen that if we consider residual mappings represented by d-disjunct matrices then f is injective when restricted to $B_\delta(0,d)$. As such, we can consider the non-linear code $f(B_\delta(0,d))$, and obtain error corrective capabilities by considering the parameters of said code. The model will then be some distorted syndrome given by $y = xH \oplus e$, where $e \in \mathbb{B}_2^k$ represents some error with entries distributed by Bernoulli(p), and $\oplus$ represents addition over $\mathbb{F}_2$.

Definition 3.4. Let $f\colon \mathbb{B}_2^n \to \mathbb{B}_2^k$ be a residuated mapping represented by a d-disjunct matrix. Then define the (n, M, d') code $C_{f,d}$ as

$$C_{f,d} := f(B_\delta(0,d)) \subseteq \mathbb{B}_2^k$$

where $M = |B_\delta(0,d)| = \sum_{i=0}^{d} \binom{n}{i}$. Furthermore, denote the distance enumerator of $C_{f,d}$

$$A_\delta(C_{f,d}) := \sum \left\{ z^{\delta(x,y)} \;\middle|\; x, y \in C_{f,d} \right\} = \sum_{i=0}^{k} A_i z^i$$

where $A_i = |\{(x,y) \in C_{f,d} \times C_{f,d} \mid \delta(x,y) = i\}|$.

The scripts used to verify d-disjunctness of any examples, or used to generate $C_{f,d}$ and distance enumerators can be found in Appendix A.

Example 3.5. Consider $f\colon \mathbb{B}_2^7 \to \mathbb{B}_2^7$ represented by H, where H is the incidence matrix of the binary Fano plane, i.e.

$$H = \begin{bmatrix} 1&0&0&0&1&1&0 \\ 0&1&0&0&0&1&1 \\ 1&0&1&0&0&0&1 \\ 1&1&0&1&0&0&0 \\ 0&1&1&0&1&0&0 \\ 0&0&1&1&0&1&0 \\ 0&0&0&1&1&0&1 \end{bmatrix}$$

which is then 2-disjunct. Thus, we can consider two different codes, namely

$C_1 := C_{f,1} = f(B_\delta(0,1))$ which is an $(7,8,3)$-code, where

$$A_\delta(C_{f,1}) = 8 + 14z^3 + 42z^4$$

and $C_2 := C_{f,2} = f(B_\delta(0,2))$ which is an $(7,29,2)$-code, where

$$A_\delta(C_{f,1}) = 29 + 294z^2 + 14z^3 + 420z^4 + 42z^5 + 42z^6$$

Thus, using C_1 one can identify one infected sample out of 7 samples through 7 tests, while also correcting an erroneous test reliably, as it has minimum distance 3. Similarly, C_2 can identify two infected samples out of 7 samples through 7 tests, but correcting flawed tests cannot be done reliably, but can be optimised under a probabilistic approach.

There are also several other models to consider in the case of non-adaptive noisy group testing, which we shall briefly mention here, with the first being an additive noise model, which is where given an infection pattern x one receives the syndrome $y = xH \vee z$ where each entry of z is drawn independently from Bernoulli(p). Thus, a test can still be positive even if all the pooled in items in the test are negative, however false-negatives are still not possible. In [2] they showed given d infected and n times there exists $k = \mathcal{O}\left(\frac{k \log n}{1-p}\right)$ such that the average error probability tends to zero as $n \to \infty$ for any fixed k.

An additional model to consider would be the case where the set of infected are distributed according to some Bernoulli distribution where the probability of an item being infected is $\frac{t}{n}$ for some $t \in \mathbb{N}$. This is a natural model to consider as often in real life one will not know an upper bound on the number of infected in the participating items, but be aware of some prevalence of the population of which the items are sampled from. For any $\varepsilon > 0$ they provide in [10] an explicit construction of $n \times k$ matrices, where $k = \mathcal{O}\left(t\frac{\log^2 n}{\log t}\right)$, that determines the set of infected with probability $1 - \varepsilon$.

Lastly, we mention the dilution model as given in [10]. Our testing result is then modelled under the binary asymmetric channel, or Z-channel, where there is a probability that a positive item in a given test can become diluted in the test. Thus, the received syndrome will be $y = Z(x)H$, where the Z-channel has probability p of mapping 1 to 0, and probability 1 of mapping 0 to 0. Given n items and d defectives [10] showed the existence of a testing matrix using $k = \mathcal{O}\left(\frac{d \log n}{(1-p)^2}\right)$ tests, where the average error probability asymptotically approaches zero.

Note that the above-mentioned models are considered independently, that is, we assume that only one model is the reason of misidentification.

4 | Bounds and Constructions

In the following chapter we consider some bounds on the parameters of n and k for a d-disjunct $n \times k$ matrix and look at equivalent objects. Furthermore, we look at combinatorial objects which gives rise to d-disjunct matrices. We also shortly consider a specific project in which they implemented combinatorial batch testing of DNA for detection of pathogenic variants [11], and we describe their construction in the setting of group testing given in this thesis. Lastly, we consider some examples of the best known 2-disjunct $n \times k$ matrices for some small values of k, according to the OEIS [12, A286874].

4.1 Basic Bounds

We now consider some simple bounds for d-disjunct matrices. Let $r(w)$ denote the number of rows of weight w, and we call a subset of $\{1, 2, \ldots, k\}$ private if it is contained in a unique row. If a row contains a private singleton subset then the row is called isolated, which implies that there exists a column intersecting only that row. We then have the following two lemmata.

Lemma 4.1. [9] *Let H be a $n \times k$ d-disjunct matrix. Then any row of weight $\leqslant d$ is isolated, and*

$$\sum_{w=1}^{d} r(w) \leqslant k$$

Proof. Suppose for contradiction, that there exists a row H_j which is not isolated, but with weight $w_j \leqslant d$. However, H_j is then contained in the union of at most d row, contradicting the assumption of H being d-disjunct. ■

Let $k(d, n)$ denote the minimum number of columns in a matrix with n row which is d-disjunct, and similarly let $n(d, k)$ denote the maximum number of rows in a matrix with k columns which is d-disjunct.

Lemma 4.2. [9] *$k(d, n) \geqslant w + k(d-1, n-1)$, where w is the weight of any row of the matrix.*

Proof. Suppose H is a d-disjunct matrix, and h is a row with weight w, and let H' be the matrix resulting from removing h, and the w columns which intersect the support of h. Now, assume for contradiction that H' is not $(d-1)$-disjunct. There then exists a row h'_i contained in the union of $d-1$ rows in H', but then h'_i must also be in the union of the same $d-1$ rows of H and the row h, which would imply that H is not d-disjunct. ■

We then obtain the following bound.

Theorem 4.3. [9]

$$k(d,n) \geqslant \min\left\{\binom{d+2}{2}, n\right\}$$

Proof. We prove the result by induction on n, where the base case of $n = 1$ clearly holds. Let H be a $n \times k$ d-disjunct matrix, and consider first the case where H has a row of weight $w \geqslant d+1$, so by lemma 4.2 and our induction hypothesis we have

$$k(d,n) \geqslant d+1+\min\left\{\binom{d+1}{2}, n-1\right\} \geqslant \min\left\{\binom{d+2}{2}, n\right\}$$

and if H does not contain such a row, then by lemma 4.1 we have

$$k(d,n) \geqslant \sum_{w=1}^{d} r(w) = n \geqslant \min\left\{\binom{d+2}{2}, n\right\}$$

■

From the previous results, we gather that if there are d infected out of our n items, such that $\binom{d+2}{2} \geqslant n$ then the minimum number of tests required is $k = n$.

Corollary 4.4. [9] *$k(d,n) = n$ for $\binom{d+2}{2} \geqslant n$.*

Proof. Clearly the $n \times n$ identity matrix is d-disjunct, so $k(d,n) \leqslant n$, and by theorem 4.3 we have $k(d,n) \geqslant n$. ■

In the following we provide a non-constructive probabilistic argument for a bound on the number of tests required.

Theorem 4.5. [9] $k(d,n) \leqslant 3(d+1)\ln\left((d+1)\binom{n}{d+1}\right)$

Proof. Let (H_{ij}) be a $n \times k$ matrix with entries generated by a Bernoulli distribution where $P(x=1) = p$. Now, for column i and rows $j_1, \ldots, j_{d+1}$ the probability that $H_{ij_1} = 1$ and $H_{ij_2} = \ldots = H_{ij_{d+1}} = 0$ will then be $p(1-p)^d$. Furthermore, denoting E as the event of no such column i existing we have

$$P(E) = \left(1 - p(1-p)^d\right)^t \tag{4.1}$$

Note that (4.1) is minimised for $p = \frac{1}{d+1}$, so under this choice of p the probability that $P(E)$ will happen for any choice of $j_1, \ldots, j_{d+1}$ is less than

$$(d+1)\binom{n}{d+1}\left(1 - \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^d\right)^k$$

as there are $(d+1)\binom{n}{d+1}$ ways of choosing $d+1$ rows, where one is designated as j_1, and the probability that E occurs to at least one column is less than the sum of the probability of E occurring to all of them. Furthermore, we have the following inequalities

$$\frac{1}{2} \geqslant \left(1 - \frac{1}{d+1}\right)^d > \frac{1}{3}, \quad d \geqslant 1$$

and

$$-\ln(1-x) \geqslant x, \quad 1 > x \geqslant 0$$

which when combined yields

$$\ln\left(1 - \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^d\right) < \frac{-1}{3(d+1)} \tag{4.2}$$

which gives that the probability of E happening to every choice of $j_1, \ldots, j_{d+1}$ is less than 1 if $k \geqslant 3(d+1)\ln\left((d+1)\binom{n}{d+1}\right)$. ■

Lastly, in [13] they state $k(d,n) \geqslant \Omega\left(\min\left\{d^2 \log_d n, n\right\}\right)$ by claiming that

$$n - \frac{d}{2} \leqslant \binom{k}{\lceil 4k/d^2 \rceil} \tag{4.3}$$

for any $n \times k$ d-disjunct matrix H, but (4.3) is trivially false in the case of H being any 2-disjunct matrix, where $n > 2$, as (4.3) would then yield $n - 1 \leqslant 1$.

4.2 Equivalent Objects

In the following section we consider objects equivalent to d-disjunct matrices, and from this derive some bounds on the parameters n and k. The following on cover-free families, disjunct set systems, and coverings of order-interval hypergraphs is based on [14].

Definition 4.6. A d-cover-free family $(X, \mathcal{F})$ is a set system such that for any d blocks $A_1, \ldots, A_d \in \mathcal{F}$, and any other block $B_0 \in \mathcal{F}$, then

$$B_0 \subsetneq \bigcup_{j=1}^{d} A_j$$

Clearly a d-cover-free family is equivalent to a d-disjunct matrix of size $|\mathcal{F}| \times |X|$. The incidence matrix of a d-cover-free family is also called a superimposed d-code in some literature [15]. Furthermore, if the d-cover-free family is w-uniform, that is, $|F| = w$ for any $F \in \mathcal{F}$ we obtain a d-disjunct matrix where each row has the same weight, or in the language of group testing, each test is equally diluted.

Definition 4.7. An (i, j)-disjunct system $(X, \mathcal{B})$ is a set system such that, for any $P, Q \subseteq X$ with $|P| \leqslant i$, $|Q| \leqslant j$ and $P \cap Q = \varnothing$, there exists $B \in \mathcal{B}$ such that $P \subseteq B$ and $Q \cap B = \varnothing$.

For an (i, j)-disjunct system $(X, \mathcal{B})$ we shall denote it as (i, j)-DS(v, b) if $|X| = v$ and $|\mathcal{B}| = b$. Similarly, for a d-cover-free family $(X, \mathcal{F})$ we shall denote it as d-CFF(k, n), where $|X| = k$ and $|\mathcal{F}| = n$.

We remind that d-disjunctness is equivalent to the statement, that for any $d + 1$ rows, where one of them is designated, there exists a column with 1 in the designated row, and 0 in the other d rows.

Theorem 4.8. *There exists a d-CFF(k, n) if and only if there exists an $(1, d)$-DS(n, k).*

Proof. Transposing the incidence matrix of the d-CFF(k, n) we obtain a matrix, where for any $d + 1$ columns, with a column designated, that is, $d + 1$ distinct points of an incidence structure, with a point designated as P, there exists a row with a 1 in that column, and 0 elsewhere. This is the same as stating there exists a block B such that $P \in B$, while the other d points are not contained in B. These arguements also holds for the converse relation. ■

As such, disjunct systems and cover-free families are dual incidence structures, and given the matrix of a cover-free family we can obtain a disjunct system by transposing said matrix, or vice versa.

We now consider order-interval hypergraphs and their coverings.

Definition 4.9. Let $P_{n;l,u} = \{Y \subseteq [n] \mid l \leqslant |Y| \leqslant u\}$, where $0 < l < u < n$, be ordered by inclusion. Let $G_{n;l,u} = (P, E)$ be the class of order-interval hypergraphs, where $P = P_{n;l,u}$, and the edges E are the maximal intervals, that is,

$$E = \left\{ I = \{C \subseteq [n] \mid Y_1 \subseteq C \subseteq Y_2\} \;\middle|\; |Y_1| = l, |Y_2| = u, Y_1, Y_2, \subseteq [n] \right\}$$

A covering of a hypergraph is a subset of points S, such that each edge of the hypergraph contains at least one point of S. This then leads to the following equivalency.

Theorem 4.10. *There exists a covering of $G_{n;l,u}$ of size b if and only if there exists a $(l, n-u)$-DS(n, b).*

Proof. A set S is a covering of $G_{n;l,u}$ if and only if for any $Y_1, Y_2 \subseteq [n]$, where $Y_1 \subset Y_2$, $|Y_1| = l, |Y_2| = u$ there exists $C \in S$ satisfying $Y_1 \subseteq C \subseteq Y_2$. This is equivalent to the statement that for any $Y_1, Y_3 \subseteq [n]$, where $|Y_1| = \ell$, $|Y_3| = n - u$, and $Y_1 \cap Y_3 = \varnothing$, there exists $C \in S$ such that $Y_1 \subseteq C$ and $Y_3 \cap C = \varnothing$. ■

In [14] they showed that the problem of determining a upper bound on k for cover-free families using an efficient algorithm is difficult. They begin by defining

$$\tau(G_{n;l,u}) = \min\{|S| \mid S \text{ is a covering of } G_{n;l,u}\}$$

and then showing that the problem of deciding $\tau(G_{n;l,u}) \leqslant m$, with input $G_{n;l,u}$ and m, is NP-complete by reducing the problem to an edge covering problem.

We now provide some bounds for cover-free families, where the following lemma will be necessary for the first bound [14].

Lemma 4.11. *Let $(X, \mathcal{B})$ be a t-uniform set system, where $|X| = w$. If for any d blocks $B_1, \ldots, B_d \in \mathcal{B}$ where $\left|\bigcup_{i=1}^{d} B_i\right| < w$ and $t \cdot d \geqslant k$, then $|\mathcal{B}| \leqslant \binom{w-1}{t}$*

Theorem 4.12. *For a w-uniform d-CFF(k, n), we have*

$$n \leqslant \binom{k}{\lceil \frac{w}{d} \rceil} \Big/ \binom{w-1}{\lceil \frac{w}{d} \rceil - 1}$$

Proof. Suppose $(X, \mathcal{F})$ is w-uniform d-CFF(k, n), and for any block $F \in \mathcal{F}$ define

$$\mathcal{N}_t(F) = \left\{F \subseteq \mathcal{F} \;\middle|\; |T| = t, \exists F' \neq F, F' \in \mathcal{F}, T \subseteq F'\right\}$$

where $t = \lceil \frac{w}{d} \rceil$. Now consider d blocks $T_1, \ldots, T_d \in \mathcal{N}_t(F)$, so $|\bigcup_{i=1}^{d} T_i| \leqslant k-1$, and as $rt \geqslant k$ we have by lemma 4.11, that

$$|\mathcal{N}_t(F)| \leqslant \binom{w-1}{t}$$

Thus, for each $F \in \mathcal{F}$ there are at least

$$\binom{w}{t} - \binom{w-1}{t} = \binom{w-1}{t-1}$$

subsets $T \subset F$, which are not contained in any $F' \in \mathcal{F}$, where $F' \neq F$. From this, we then have

$$|\mathcal{F}|\binom{w-1}{t-1} \leqslant \binom{k}{t}$$

finalising the proof as $t = \lceil \frac{w}{d} \rceil$ and $|\mathcal{F}| = n$. ■

In the language of d-disjunct matrices theorem 4.12 implies that if H is a $n \times k$ d-disjunct matrix, where $r(w) = n$, then n is bounded as in the theorem. In [16] they proved the following bound on cover-free families.

Theorem 4.13. [16] *For a d-CFF(k, n) we have*

$$n \leqslant d + \binom{k}{\left\lceil (k-d)/\binom{d+1}{2} \right\rceil}$$

Based on Sperner's theorem we can entirely determine the case for 1-cover-free families, which is the case of 1-disjunct matrices. A Sperner family, or an antichain of sets, is a family of sets such that none of the sets is a strict subset of another.

Theorem 4.14. (Sperner's theorem) [17] *Over a k-element set the largest Sperner family S is the set consisting of all subsets of size $\binom{k}{\lfloor k/2 \rfloor}$.*

Thus, the largest possible 1-disjunct matrix for fixed a value of k is the $\binom{k}{\lfloor k/2 \rfloor} \times k$ matrix consisting of all distinct rows of weight $\lfloor n/2 \rfloor$.

4.3 Kautz-Singleton Construction

In the following section we showcase the Kautz-Singleton construction based on Reed-Solomon codes, where we for any d and large enough $n \geqslant d$ can construct $n \times k$ d-disjunct matrices, where $k = \mathcal{O}\left(d^2(\log_d n)^2\right)$ [13].

For distinct $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ we shall denote the $[n, k]_q$ Reed-Solomon code over $\mathbb{F}_q$ as

$$\mathrm{RS}(n, k) = \left\{ \big(f(\alpha_1), \ldots, f(\alpha_n)\big) \;\middle|\; f \in \mathbb{F}_q[x]_{<k} \right\}$$

The Kautz-Singleton construction then takes the generator matrix of an $[n, k]_q$ Reed-Solomon code, and then replaces any entry with the corresponding basis vector indexed by said element, and it finally transposes the resultant matrix. Thus, if an entry is $i \in \mathbb{F}_q$ then it is replaced with the vector with a 1 in the

$(n-i+1)$'th entry, and 0 elsewhere. We now provide a simple example of this construction before analysing it.

Example 4.15 (Kautz-Singleton Construction)**.**

Consider the $[3,1]_3$ Reed-Solomon code with generator matrix

$$G = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

The corresponding mappings are then

$$0 \mapsto \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad 1 \mapsto \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad 2 \mapsto \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

resulting in the Kautz-Singleton matrix

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

We first provide the following sufficient condition for a matrix to be d-disjunct.

Lemma 4.16. [13] *Let H be a $n \times k$ matrix over $\mathbb{B}_2$ such that for some integers $a_{\max} \leqslant w_{\min} \leqslant t$, then every row has at least weight $w_{\min}$, and the product of any two distinct rows has at most weight $a_{\max}$. Then H is $\left\lfloor \frac{w_{\min}-1}{a_{\max}} \right\rfloor$-disjunct.*

Proof. Denote $d = \left\lfloor \frac{(w_{\min})^{-1}}{a_{\max}} \right\rfloor$, and let H_j denote the j'th row of H, and fix some $S \subseteq [n]$ such that $|S| \leqslant d$, and $j \notin S$. Now,

$$\begin{aligned} \left| \operatorname{supp}\left(H_j\right) \setminus \operatorname{supp}\left(\sum_{i \in S} H_j \cdot H_i\right) \right| &= \left| \operatorname{supp}\left(H_j\right) \right| - \left| \operatorname{supp}\left(\sum_{i \in S} H_j \cdot H_i\right) \right| \\ &\geqslant \left| \operatorname{supp}\left(H_j\right) \right| - \sum_{i \in S} \left| \operatorname{supp}\left(H_j \cdot H_i\right) \right| \\ &\geqslant w_{\min} - |S| \cdot a_{\max} \\ &\geqslant w_{\min} - d \cdot a_{\max} \\ &= w_{\min} - \frac{w_{\min} - 1}{a_{\max}} \cdot a_{\max} \\ &= 1 \end{aligned}$$

finalising the claim of H being d-disjunct. ∎

Theorem 4.17. [13] *For any integer $d \geqslant 1$ and large enough $n \geqslant d$ there exists a $n \times k$ d-disjunct matrix, where $k = \mathcal{O}\left(d^2(\log_d n)^2\right)$.*

Proof. We shall use the Kautz-Singleton construction to show the existence of such a matrix by choosing parameters that yields the sufficiency given by lemma 4.16. Let H denote the resultant matrix from the Kautz-Singleton construction from $\mathrm{RS}(q,t)$ code, which will then be a $q^t \times q^2$ matrix, where each row has exactly weight q, so $w_{\min} = q$. We now determine an expression for $a_{\max}$.

Partition the columns of H into blocks of size q and index the q^2 columns by pairs in $[q] \times [q]$, so $H_{\ell,(i,j)} = 1$ if and only if the ℓ'th codeword of the Reed-Solomon code satisfies the j'th entry being equal to j. Thus, the number of columns where the ℓ_1'th and ℓ_2'th rows both has a 1 will be the number of positions the corresponding codewords agree, but as the code is a $\mathrm{RS}(q,t)$ code we have that no two rows agree in more than $t-1$ entries, so $a_{\max} = t-1$. Now, by lemma 4.16 we have that H is d-disjunct for

$$d = \left\lfloor \frac{q-1}{t-1} \right\rfloor$$

so we need to choose a $\mathrm{RS}(q,t)$ code which satisfies the above. As $q = \mathcal{O}(td)$ and $n = q^t$, then $t = \log_q n$, which implies $q = \mathcal{O}(d \log_d n)$, and as $k = q^2$ we finally obtain $k = \mathcal{O}\left(d^2 \left(\log_d n\right)^2\right)$. ∎

The benefit of the Kautz-Singleton construction is the greatly explicit nature of it, but in the case of $d = \mathcal{O}(\mathrm{poly}(\log n))$ it is sub-optimal as shown in [18]. In [19] they showcased an explicit construction for $k = \mathcal{O}\left(d \left(\log_d n\right)^2\right)$, which is more efficient in the case of $d = \mathcal{O}(\mathrm{poly}(\log n))$.

In [18] they also showed in the regime of $d = \Omega(\log^2 n)$ that the Kautz-Singleton construction is optimal in both the noisy and noiseless case given appropriately chosen parameters q and n.

4.4 Configurations

In the following we provide the definitions of configurations, or equivalently, partial linear spaces, and prove that these incidence structures provides disjunct matrices. We also summarise some existence results for configurations. The results are based on [20, 21] unless otherwise mentioned, and any proofs omitted can be found in these.

Definition 4.18. A configuration (v_r, b_k) is a finite incidence structure such that

i. There are v points and b lines;

ii. Each line is incident with k points, and each point is incident with r lines;

iii. Two distinct lines intersect each other at most once.

If $v = b$, and thusly, $r = k$, the configuration is symmetric and denoted just by v_k. In the context of group testing we are however not interested in the symmetric case, as it will correspond to disjunct matrices where $n = k$.

Theorem 4.19. [5] *Let $L = (v_r, b_k)$ be a configuration, and consider m distinct lines $\ell_1, \ldots, \ell_m$ of L. Then, if $m \leqslant k-1$ and $\ell \in L$ such that*

$$\ell \subseteq \bigcup_{i=1}^{k-1} \ell_i$$

then $\ell = \ell_j$ for some j satisfying $1 \leqslant j \leqslant m$.

Proof. Suppose to the contrary that $\ell \neq \ell_i$ for $i = 1, \ldots, m$, while $m \leqslant k-1$, and $\ell \subseteq \bigcup_{i=1}^{k-1} \ell_i$. However, as $|\ell \cap \ell_i| \leqslant 1$ we have

$$k = |\ell| = \left|\ell \cap \bigcup_{i=1}^{m} \ell_i\right| \leqslant \sum_{i=1}^{m} |\ell \cap \ell_i| \leqslant m$$

this however implies $k < m + 1$, which is a contradiction. ■

We remind that the incidence matrix of incidence structures we consider the rows to represent the lines, and the columns to represent the points. Theorem 4.19 then yields the following result.

Corollary 4.20. *The incidence matrix of a (v_r, b_k) configuration is of size $b \times v$ and is $(k-1)$-disjunct.*

Thus, to construct good disjunct matrices we can consider the problem of constructing configurations with large k, while also aiming to maximise b and minimising v. We have the following necessary conditions for the existence of a configuration, which showcases that the aforementioned goals are conflicting.

Lemma 4.21. *Let (v_r, b_k) be a configuration. Then*

i. $v \leqslant b$ *and* $k \leqslant r$;

ii. $vr = bk$;

iii. $v \geqslant r(k-1) + 1$.

The problem of determining the existence of configurations can be considered by the existence problem of resolvable Steiner systems $S(2, k, v)$, which are $2-(n, k, 1)$ designs, where the blocks can be partitioned into set, which also forms a partition of the original point set of the design.

Theorem 4.22. *Let v be a multiple of k. If there exists a resolvable Steiner system $S(2, k, v)$ then there exists a (v_r, b_k) configuration.*

The case of $k = 3$ is entirely determined by the sufficient conditions given in 4.21.

Theorem 4.23. *There exists a configuration (v_r, b_3) if and only if $v \geqslant 2r + 1$ and $vr = 3b$.*

The case of $k = 4$ is not yet solved, as the question of whether the necessary conditions of lemma 4.21 are sufficient is still open, and no non-existence results are known [22]. There is however an analogous result to theorem 4.23.

Theorem 4.24. *If $v \equiv 4 \mod 12, v \geqslant 3r + 1$ and $vr = 4b$ then there exists a (v_r, b_4) configuration.*

There are also further results for the existence of configurations for $k = 4$ with specific restrictions on the parameters, and similarly for $k = 5$ there are even fewer existence results. We quickly summarise some of these here.

i. If $v \equiv 0 \mod 12$, $v = 3r + 3$ and $vr = 4b$, then there exists (v_r, b_4).

ii. For all $1 \leqslant \frac{b}{v} \leqslant 15$, except possibly $\frac{b}{v} = 3$ and $v = 38$, then there exists (v_r, b_4), where $r = 4\frac{b}{v}$, $v \geqslant 3r + 1$, $vr = 4b$.

iii. For all $v \geqslant 20$, v even and $b = \frac{3v}{2}$, then there exists (v_6, b_4).

iv. If $v = 4r + 4$, then there exists (v_r, b_5), where $vr = 5b$ for all $v \equiv 0 \mod 20$.

v. If $v \equiv 5 \mod 20$, $v \geqslant 4r + 1$, $vr = 5b$, and $v \geqslant 7865$, then there exists (v_r, b_5).

There are also a variety of additional configurations with additional properties one could consider, for example generalised polygons. These objects has well-known restrictions on the parameters r and k [23].

As an example of this, if one considers a configuration (v_r, b_k) which is also a generalised quadrangles then one must have $(r-1)^{\frac{1}{2}} \leqslant k - 1 \leqslant (r-1)^{\frac{1}{2}}$, and $v = k((k-1)(r-1)+1)$ and $b = r((k-1)(r-1)+1)$.

Lastly, we consider an example of a configuration which yields a 2-disjunct matrix.

Example 4.25. Consider the $(12_4, 16_3)$ configuration with incidence matrix

$$H = \begin{bmatrix} 1&0&0&0&1&0&0&0&0&0&0&1 \\ 1&0&0&0&0&1&0&1&0&0&0&0 \\ 1&0&0&0&0&0&1&0&0&1&0&0 \\ 1&0&0&0&0&0&0&0&1&0&1&0 \\ 0&1&0&1&0&0&0&0&1&0&0&0 \\ 0&1&0&0&0&1&1&0&0&0&0&0 \\ 0&1&0&0&0&0&0&1&0&0&1&0 \\ 0&1&0&0&0&0&0&0&0&1&0&1 \\ 0&0&1&1&0&0&0&0&0&1&0&0 \\ 0&0&1&0&1&0&0&0&0&0&1&0 \\ 0&0&1&0&0&0&1&0&0&0&0&1 \\ 0&0&1&0&0&0&0&1&1&0&0&0 \\ 0&0&0&1&1&1&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0&1&1 \\ 0&0&0&0&1&0&1&1&0&0&0&0 \\ 0&0&0&0&0&1&0&0&1&1&0&0 \end{bmatrix}$$

which is then 2-disjunct. Letting $f(x) = xH$ denote the corresponding residuated mapping we then have the codes

$$C_1 = f(B_\delta(0,1)), \quad C_2 = f(B_\delta(0,2))$$

with distance enumerators

$$A_\delta(C_1) = 17 + 32z^3 + 144z^4 + 96z^6$$

and

$$\begin{aligned} A_\delta(C_2) =& 121 + 354z^2 + 1116z^3 + 1800z^4 + 2322z^5 + 3084z^6 \\ &+ 2706z^7 + 2016z^8 + 870z^9 + 210z^{10} + 42z^{12} \end{aligned}$$

Thus, using C_1 we can identify one infected sample using 12 tests, while correcting an erroneous test, while C_2 can identify two infected samples, but correcting flawed tests cannot be done reliably, but can be optimised under a probalistic approach.

4.5 Inversive Planes

In the following we showcase how inversive planes yields 1-disjunct matrices, while also considering the minimum distance of the codes generated under their residuated mappings. As such, these incidence structures may be beneficial to use in the case of the set of infected items having low prevalence, while giving a relatively high error-correction.

Definition 4.26. [24] An inversive plane is a finite incidence structure consisting of points and circles, such that

i. Any circle has points,

ii. A unique circle is incident with any given triple of points,

iii. If $p_1 \in c_1$ and $p_2 \notin c_1$, then there exists a unique circle $c_2 \neq c_1$ such that $p_1 \in c_2$ and $p_2 \in c_2$,

iv. There exists 4 points which are not concircular.

Furthermore, the order of an inversive plane is an $n \in \mathbb{N}$ such that

i. The number of points is $n^2 + 1$,

ii. The number of circles is $n(n^2 + 1)$,

iii. Each circle is incident with $n + 1$ points,

iv. Each point is incident with $n(n + 1)$ circles,

v. Each circle is tangent to $n^2 - 1$ other circles,

vi. Each circle is disjoint from $n(n - 1)(n - 2)/2$ other circles.

and the inversive plane is then also a 3-$(n^2 + 1, n + 1, 1)$-design [25], or a $S(3, n^2 + 1, n + 1)$ Steiner system.

As the incidence matrix of an inversive plane is 1-disjunct we can consider the code consisting of the rows of said matrix, and the zero codeword.

Theorem 4.27. *Let H be the incidence matrix of an inversive plane of order n, and let $f\colon \mathbb{B}_2^{n(n^2+1)} \to \mathbb{B}_2^{n^2+1}$ be the corresponding residuated mapping. Then the code $f(B_\delta(0, 1))$ has minimum distance $n + 1$.*

Proof. Assume for contradiction that the minimum distance d of $f(B_\delta(0, 1))$ satisfies $d < n + 1$, and let $c_1, c_2 \in f(B_\delta(0, 1))$ such that $d = \delta(c_1, c_2)$. Clearly these are both non-zero and distinct, so $|\operatorname{supp}(c_1) \cap \operatorname{supp}(c_2)| \leqslant 2$, as any three points uniquely determines a circle in the inversive plane. Furthermore, as $|\overline{\operatorname{supp}(c_1)} \cap \overline{\operatorname{supp}(c_2)}|$ denotes the number of 0 entries shared by the two codewords we have

$$\begin{aligned} n + 1 > d &= n^2 + 1 - |\operatorname{supp}(c_1) \cap \operatorname{supp}(c_2)| - |\overline{\operatorname{supp}(c_1)} \cap \overline{\operatorname{supp}(c_2)}| \\ &\geqslant n^2 - 1 - |\overline{\operatorname{supp}(c_1)} \cap \overline{\operatorname{supp}(c_2)}| \end{aligned}$$

which implies $|\overline{\operatorname{supp}(c_1)} \cap \overline{\operatorname{supp}(c_2)}| \geqslant n^2 - n + 1$, but each codeword has a total of $n^2 - n$ zeroes so $|\overline{\operatorname{supp}(c_1)} \cap \overline{\operatorname{supp}(c_2)}| \leqslant n^2 - n$, which is a contradiction, so $d \geqslant n + 1$. Lastly, as any circle is incident with $n + 1$ points we have $\delta(0, c) = n + 1$ for any non-zero codeword $c \in f(B_\delta(0, 1))$, so $d = n + 1$. ■

Thus, it can be beneficial to use the codes generated by incidence matrices of inversive planes, as the high minimum distance implies that we will be able to recover more test errors. The following example is based on [7].

Example 4.28. Consider the inversive plane of order 3 with incidence matrix given by

$$H = \begin{bmatrix}
1&1&1&1&0&0&0&0&0&0\\
1&0&0&0&1&1&1&0&0&0\\
1&0&0&0&0&0&0&1&1&1\\
1&1&0&0&1&0&0&1&0&0\\
1&0&0&1&0&0&1&0&0&1\\
1&0&1&0&0&1&0&0&1&0\\
1&1&0&0&0&1&0&0&0&1\\
1&0&0&1&1&0&0&0&1&0\\
1&0&1&0&0&0&1&1&0&0\\
1&1&0&0&0&0&1&0&1&0\\
1&0&0&1&0&1&0&1&0&0\\
1&0&1&0&1&0&0&0&0&1\\
0&1&1&0&1&1&0&0&0&0\\
0&1&1&0&0&0&1&0&0&1\\
0&1&1&0&0&0&0&1&1&0\\
0&1&0&1&1&0&1&0&0&0\\
0&1&0&1&0&1&0&0&1&0\\
0&1&0&1&0&0&0&1&0&1\\
0&1&0&0&1&0&0&0&1&1\\
0&1&0&0&0&1&1&1&0&0\\
0&0&1&1&1&0&0&1&0&0\\
0&0&1&1&0&1&1&0&0&0\\
0&0&1&1&0&0&0&0&1&1\\
0&0&1&0&1&0&1&0&1&0\\
0&0&1&0&0&1&0&1&0&1\\
0&0&0&1&1&1&0&0&0&1\\
0&0&0&1&0&0&1&1&1&0\\
0&0&0&0&1&1&0&1&1&0\\
0&0&0&0&1&0&1&1&0&1\\
0&0&0&0&0&1&1&0&1&1
\end{bmatrix}$$

which is then 1-disjunct. Then H can identify 1 infected sample out of 30 samples through 10 tests, but as the code $C = f(B_\delta(0,1))$, where $f(x) = xH$, has minimum distance 4 we have that C, in addition to identifying one infected sample, can recover one erroneous test.

4.6 2-disjunct matrices

Consider the following table for the maximal number n of binary vectors of length k, such that the union of any 2 distinct vectors does not contain any other vector.

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$n(k,2)$	2	2	3	4	5	6	7	8	12	13	17	20	26	28	40	45

Table 4.1: Lower bounds of $n(k,2)$ according to sequence A286874 on the OEIS.

For the cases of $k = 3, 6, 7, 8, 9, 12, 13, 14$ there exists configurations that gives exactly the corresponding values in table 4.1 according to theorem 4.23, where for any other values of k the configurations will result in 2-disjunct matrices with smaller values of n. However, for $k = 15$ consider the following matrix.

$$H = \begin{bmatrix}
0&1&1&0&0&0&1&0&0&0&0&1&1&0&0\\
0&0&0&0&0&1&0&0&0&1&0&1&1&0&1\\
0&0&1&0&1&0&0&1&0&1&0&1&0&0&0\\
1&0&0&1&0&1&0&0&1&0&0&0&0&1&0\\
1&0&1&1&1&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&1&0&1&1&0&1&0&0\\
1&1&0&0&0&0&0&0&0&1&0&1&0&1&0\\
1&1&0&0&1&0&0&1&0&0&0&0&1&0&0\\
0&0&0&0&1&1&0&1&0&0&0&0&0&1&1\\
1&0&1&0&0&0&0&0&1&1&0&0&1&0&0\\
0&0&0&0&0&0&0&0&1&1&1&0&0&1&1\\
0&1&1&1&0&0&0&0&0&1&1&0&0&0&0\\
0&0&0&0&0&0&0&1&1&0&0&1&1&1&0\\
0&0&1&1&0&1&0&1&0&0&0&0&1&0&0\\
1&1&0&0&0&0&1&0&1&0&1&0&0&0&0\\
1&0&1&0&0&0&0&1&0&0&1&0&0&1&0\\
1&0&0&0&0&1&1&1&0&0&0&1&0&0&0\\
0&0&1&1&0&0&0&0&0&0&1&0&1&1&1\\
0&0&0&1&0&0&1&0&1&1&0&1&0&0&0\\
0&0&0&1&1&0&0&1&1&0&1&0&0&0&0\\
0&1&0&0&1&0&1&0&0&1&0&0&0&0&1\\
0&1&0&0&0&1&0&1&1&1&0&0&0&0&0\\
0&1&0&1&0&0&1&0&0&1&1&0&0&0&1\\
0&1&0&0&0&0&0&1&0&0&1&1&0&0&1\\
0&1&0&1&0&0&0&0&1&0&0&0&1&0&1\\
0&1&0&1&1&1&0&0&0&0&0&1&0&0&0\\
0&0&0&0&1&1&1&0&1&0&0&0&1&0&0\\
1&0&0&1&0&0&0&1&0&1&0&0&0&0&1\\
0&0&0&1&0&1&1&0&0&0&1&0&0&0&1\\
0&0&1&0&0&0&1&1&1&0&0&0&0&0&1\\
1&0&0&0&1&0&0&0&1&0&0&1&0&0&1\\
1&0&0&0&0&0&1&0&0&0&0&0&1&1&1\\
0&0&1&0&0&1&1&0&0&1&0&0&0&1&0\\
1&0&0&0&1&1&0&0&0&1&1&0&0&0&0\\
0&1&1&0&1&0&0&0&1&0&0&0&0&1&0
\end{bmatrix}$$

H is 2-disjunct and of size 42×15, implying that $n(15, 2) \geqslant 42$ improving the lower bound of $n(15, 2) \geqslant 40$ as given in table 4.1. However, H was obtained

through trial and error by appending random rows to the 40×15 2-disjunct matrix attaining the bound in table 4.1, and as such, the matrix H in itself garners no insight.

4.7 PREDiSPOSED Project

In the following section we consider the methods for combinatorial batching of DNA detection of pathogenic variants given in [11] and summarise their methods in the language of group testing as discussed in this thesis.

In [26] they exemplify their concept through the following example. They arrange 2304 individuals in a 48×48 grid, where each test consists of the samples of the people on the same row, or on the same column, resulting in 96 tests, each consisting of 48 samples.

We now generalise their construction. Let n^2 be the number of items, then the testing matrix H_n corresponding to the aforementioned scheme is the $n^2 \times 2n$ block matrix given by

$$H_n = \begin{bmatrix} A_1 & I_{n\times n} \\ \vdots & \vdots \\ A_n & I_{n\times n} \end{bmatrix}$$

where A_k is the $n \times n$ matrix such that $(A_k)_{i,j} = 1$ if $j = k$, and 0 otherwise.

Proposition 4.29. *H_n is 1-disjunct, but not 2-disjunct.*

Proof. As all rows of H_n are distinct and has weight n, so H_n is 1-disjunct. However, the sum of the k'th row and $(k+n+1)$'th row contains the $(k+1)$'th row, so H_n is not 2-disjunct. ■

Compared to a 1-disjunct matrix constructed using Sperner's Theorem this is obviously far from effective, as given $2n$ tests we can test up to $\binom{2n}{n}$ items, or equivalently we can test up to $\frac{1}{2}\prod_{j=n+1}^{2n-1} j$ times as many items given the same number of tests as their scheme. However, a testing matrix from a Sperner family obviously dilutes the samples heavily, which can have implications in implementations.

Furthermore, the rows of the testing matrix having a large weight can imply logistical issues in the form of generating enough samples of each item and distributing these, compared to a testing matrix in which the rows have relatively low weight. In their example of 2304 individuals, and to be able to test at least as many using a Sperner family we would require 14 tests, which would then let us test 3432 individuals. However, each test will then contain 1716 samples, for a total of 24024 samples, where they only require 4608 samples.

5 | References

[1] R. Dorfman, "*The Detection of Defective Members of Large Populations*," *The Annals of Mathematical Statistics*, 1943.

[2] G. K. Atia and V. Saligrama, "*Boolean Compressed Sensing and Noisy Group Testing*," *Computing Research Repository*, 2009.

[3] D. Malioutov and M. Malyutov, "*Boolean compressed sensing: LP relaxation for group testing*," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012.

[4] T. Blyth, *Lattices and Ordered Algebraic Structures*. Springer London, 2005.

[5] T. Blyth and M. F. Janowitz, *Residuation Theory*. Pergamon Press, 1972.

[6] M. Greferath and C. Rößing, "*Group testing via residuation and partial geometries*," 2022.

[7] M. Greferath and C. Rößing, "*Group Testing with Error-Correction Capability for General Pandemics*," 2022.

[8] A. E. Guterman, *Rank and determinant functions for matrices over semirings*, p. 1–33. London Mathematical Society Lecture Note Series, Cambridge University Press, 2007.

[9] D.-Z. Du and F. K. Hwang, *Combinatorial Group Testing and Its Applications*. World Scientific Publishing Company, 2nd ed., 2000.

[10] A. Mazumdar, "*Nonadaptive Group Testing With Random Set of Defectives*," *IEEE Transactions on Information Theory*, 2016.

[11] U. K. Stoltze, C. M. Hagen, and T. van Overeem Hansen *et al.*, "*Combinatorial batching of DNA for ultralow-cost detection of pathogenic variants*," *Genome Med*, 2023.

[12] OEIS Foundation Inc., "*The On-Line Encyclopedia of Integer Sequences*," 2023. Published electronically at `http://oeis.org`.

[13] V. Guruswami, A. Rudra, and M. Sudan, "*Essential Coding Theory*," 2022.

[14] R. Wei, "*On Cover-Free Families*," 2023.

[15] A. G. D'yachkov, V. V. Rykov, C. Deppe, and V. S. Lebedev, "*Superimposed Codes and Threshold Group Testing*," 2014.

[16] Z. Füredi, "*On r-Cover-free Families*," *Journal of Combinatorial Theory, Series A*, 1996.

[17] S. S. Sane, *Combinatorial Techniques*. Hindustan Book Agency Gurgaon, 2013.

[18] H. A. Inan, P. Kairouz, M. Wootters, and A. Ozgur, "*On the Optimality of the Kautz-Singleton Construction in Probabilistic Group Testing*," 2019.

[19] E. Porat and A. Rothschild, "*Explicit Non-Adaptive Combinatorial Group Testing Schemes*," 2008.

[20] H. Gropp, "*Configurations and their realization*," *Discrete Mathematics*, 1997.

[21] H. Gropp, "*Nonsymmetric configurations with natural index*," *Discrete Mathematics*, 1994.

[22] C. J. Colbourn and J. H. Dinitz, *Handbook of Combinatorial Designs*. Discrete mathematics and its applications, Chapman & Hall, 2nd ed., 2007.

[23] S. Payne and J. Thas, "*Finite Generalized Quadrangles*," 1984.

[24] F. Kárteszi, *Introduction to Finite Geometries*. North Holland, 1976.

[25] P. Dembowski, *Finite Geometries: Reprint of the 1968 Edition*. Classics in Mathematics, Springer, 1st ed., 1997.

[26] "*The PREDiSPOSED Project*." `https://sites.google.com/view/predisposed`. Accessed: 22-05-2023.

A | Python Scripts

A.1 *d*-disjunctness

```
import itertools
import numpy as np

def dDis(H,d):
    for S in itertools.combinations(range(len(H)), d):
        y=np.zeros(len(H),dtype=bool);
        for i in range(d):
            y=y|H[S[i]];
        for j in (set(range(len(H)))-set(S)):
            if (np.count_nonzero((H[j]<y)|(H[j]==y))) ==
                len(H[j]):
                print("Row number", j, " is contained in the
                       union of rows",S)
                return False
    return True
```

A.2 Code Generator

```
import itertools
import numpy as np

def GenerateCode(H,d):
    C = {}
    C[0] = np.zeros(H.shape[1],dtype=bool)
    j = 1
    for S in itertools.combinations(range(len(H)), d):
        y = np.zeros(len(H),dtype=bool)
        for i in S:
            y[i] = 1
        C[j] = np.dot(y,H)
        j = j + 1
    return C
```

A.3 Distance Enumerator

```
import numpy as np

def DistanceEnumerator(C):
    y = np.zeros(len(C[1])+1)
    for c1 in C.values():
        for c2 in C.values():
            i = sum((1*c1|1*c2) - (1*c1&1*c2))
            y[i] = y[i] + 1
    return y
```