Fingerprinting Codes and Related Combinatorial Structures

by

Chuan Guo

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2015

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The content in Chapters 2 and 3 of this thesis was co-authored with Douglas Stinson and Tran van Trung. All other chapters in this thesis contain original work authored under the supervision of Douglas Stinson.

Abstract

Fingerprinting codes were introduced by Boneh and Shaw in 1998 as a method of copyright control. The desired properties of a good fingerprinting code has been found to have deep connections to combinatorial structures such as error-correcting codes and cover-free families. The particular property that motivated our research is called "frameproof". This has been studied extensively when the alphabet size q is at least as large as the colluder size w. Much less is known about the case $q < w$, and we prove several interesting properties about the binary case $q = 2$ in this thesis.

When the length of the code N is relatively small, we have shown that the number of codewords n cannot exceed N, which is a tight bound since the $n = N$ case can be satisfied a trivial construction using permutation matrices. Furthermore, the only possible candidates are equivalent to this trivial construction. Generalization to a restricted parameter set of separating hash families is also given.

As a consequence, the above result motivates the question of when a non-trivial construction can be found, and we give some definitive answers by considering combinatorial designs. In particular, we give a necessary and sufficient condition for a symmetric design to be a binary 3-frameproof code, and provide example classes of symmetric designs that satisfy or fail this condition. Finally, we apply our results to a problem of constructing short binary frameproof codes.

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Douglas Stinson, for graciously devoting his time and energy to introducing me to the subjects of fingerprinting codes and combinatorial design, and for his invaluable advice that are reflected in all aspects of my achievements to date. I would also like to thank my collaborators, Professor Tran van Trung, Professor Jeffrey Shallit, Professor Arseny Shur, and Professor Michael Newman.

Without scholarships from both internal and external sources, I would not have survived through my Master's career. I sincerely thank NSERC, the OGS program, and the Cheriton Scholarship Committee for dedicating resources to fund my research.

My time at the University of Waterloo has been made pleasant despite of the distressing weather by all members of the CrySP group and the David R. Cheriton School of Computer Science. Various faculties have helped me enlarge my view of the world by sharing their knowledge through teaching, particularly Professor Jeffrey Shallit and Professor Shai Ben-David, to whom I give my heartfelt gratitude.

Finally, I would like to thank my thesis committee, consisting of Professor Alfred Menezes and Professor David Jao. Their comments have tremendously aided my successful completion of this thesis.

Dedication

This is dedicated to my family for all their emotional and financial support during my continuing journey for the pursuit of knowledge. A very special thanks goes to my partner Zhuoran for her patience and devotion throughout the years.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The idea of using a fingerprinting code is not new. Back when logarithm tables were widely used, publishers would protect the copyright integrity of their tables by modifying the least significant digits of several random entries. This process generates numerous unique copies of the original table, one per customer. If an illegal copy is obtained, the publisher could match the table with their record and trace back to the customer that made the copy.

A modern formulation of this technique was given by Boneh and Shaw in 1998 [6]. Their formulation applies to digital content rather than log tables, but the fundamental idea remains the same – a secret code is embedded in the digital document, giving information to the content distributor regarding the illegal copy's original owner. More precisely, consider a digital document D, viewed as a vector of finite length from a finite alphabet. A secret codeword x is interlaced with D at fixed positions, forming a copy of D that is unique per customer.

If an illegal copy is confiscated, the content distributor could example the fixed positions that contain the secret codeword and use it to determine the pirate entity. However, if two copies were obtained, their owners could collude and determine parts of the codeword and remove or modify these parts, nullifying the code's detection property. Under this attack model, the goal of content distributors is to construct useful codes that perform well even under this collusion attack.

One weakened notion of resilience to the colluding attack is called frameproof. This property specifies that instead of attempting to trace to a pirate entity when given an illegal copy with modified codeword, we are content with tracing only unmodified codes under the guarantee that no innocent user is harmed by the scheme. In other words, we

require that no distributed codeword can be constructed from the colluding party's copies unless they possess the codeword itself. This notion will be the focus of our thesis.

The problem of constructing optimal frameproof codes has received much attention from research. It is apparent that we want the length of such codes to be small and the number of codes to be large. There have been several methods of deriving frameproof codes from existing combinatorial structures [3] [27], as well as upper bounds on the number of codes to close the gap [2] [24] [28] [32], but few tight bounds have been found so far . Furthermore, many good upper bounds require the alphabet size to be large, which motivates the question of whether a large alphabet set is required. Our thesis complements known results by proving some interesting properties about frameproof codes over the binary alphabet.

1.1 Basic Definitions

In this section, we give a mathematical formulation of the frameproof property and the closely related separating hash family.

Definition 1.1.1. Let Q be an alphabet of size q. We often take the set $Q = \{0, 1, \ldots, q-1\}$ as the canonical alphabet of size q. A set $\mathsf{C} \subseteq Q^N$ of size n is called an (N, n, q)-*code*, or simply a *code* for short. The elements of C are called codewords.

Definition 1.1.2. For an element $x \in Q^N$, we refer to the i-th entry of x by $x(i)$ whenever the notation is unambiguous. Given a code C and $P \subseteq \mathsf{C}$, we define the *descendent code* of P to be

$$\mathsf{desc}(P) = \{y \in Q^N : \text{ for each } i,\ y(i) = x(i) \text{ for some } x \in P\}.$$

For positive integer w, we say that a code C is a w-*frameproof code* if for every $P \subseteq \mathsf{C}$ with $|P| = w$, we have that $\mathsf{desc}(P) \cap \mathsf{C} = P$.

Definition 1.1.2 relates to the intuitive notion of frameproof in the following sense. We make the assumption that for a colluding party P, only parts of the codeword that differ for at least two members of the party can be detected, and the colluders may change that part of the codeword to only entries in their possession. This is called the *marking assumption* in literature. Hence the set of possible modified codewords that may result from P is $\mathsf{desc}(P)$, and we require that this set does not include any legitimate codeword from any user not in P.

Example 1.1.1. Let $Q = \{0, 1, 2\}$, $\mathsf{C} = \{011, 020, 100, 121, 200\}$, where we denote the 3-tuple (a, b, c) by abc. Then $N = 3$ and $n = 5$. For $P = \{011, 020\}$, the descendent code is

$$\mathsf{desc}(P) = \{011, 020, 010, 021\}.$$

There are variations to the marking assumption. For example, one model [21] allows all alphabet elements for detected positions in the codewords, instead of restricting selection to only alphabet elements in the colluding party's possession. Formally, the descendent set is

$$\mathsf{desc}(P) = \{y \in Q^N : \text{ for each } i,\ y(i) = x(i) \text{ if } x(i) = x'(i) \text{ for all } x, x' \in P\}$$

where $Q_P(i) = \{x(i) : x \in P\}$. For binary frameproof codes the two models coincide. Another model [6] for general fingerprinting codes allows erasing of detected symbols in the codeword, i.e.

$$\mathsf{desc}(P) = \{y \in Q^N : y(i) \in Q_P(i) \cup \{?\}\}$$

where the symbol ? represents an erased position. The definition of frameproof does not apply here since codewords modified erasures cannot be used for framing.

Note that if C is a w-frameproof code, then increasing the code length by appending arbitrary entries or removing elements from C do not affect the w-frameproof property. Hence for fixed q and w, the problem of constructing optimal codes can be formulated as either maximizing n given N or minimizing N given n.

Here is a basic construction of a w-frameproof code for every w, which gives a code of size $n = (q - 1)N$.

Construction 1 ([4])**.** $\mathsf{C} = \{x_{j,k}\}$ for $j = 1, \ldots, N$ and $k = 1, \ldots, q - 1$ where

$$x_{j,k}(i) = \begin{cases} k & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

One useful representation for codes is the *matrix representation*.

Definition 1.1.3. Given an (N, n, q)-code C, construct an $N \times n$ matrix A with entries from $Q = \{0, 1, \ldots, q - 1\}$ by designating each column of A to be a different codeword from C. A is said to be the *representation matrix* of C.

In the binary case, Construction 1 in matrix form is a permutation matrix, which we will see in Chapter 2 to be the only possible candidates when N is small.

We now introduce a related combinatorial structure known as separating hash families. These structures can be shown to be related to many other structures such as perfect hash families [20] and secure frameproof codes [26], and thus it is meaningful for us to discuss frameproof codes in this framework.

Definition 1.1.4. Let X, Y be finite sets with $|X| = n$ and $|Y| = q$. Let $\mathcal{F}$ be a family of functions from X to Y with $|\mathcal{F}| = N$. For positive integers $w_1, w_2, \ldots, w_t$, we say that $\mathcal{F}$ is a $\{w_1, w_2, \ldots, w_t\}$-*separating hash family*, denoted $\mathsf{SHF}(N; n, q, \{w_1, w_2, \ldots, w_t\})$, if for pairwise disjoint subsets $X_1, \ldots, X_t \subseteq X$ with $|X_i| = w_i$ for $i = 1, \ldots, t$, there exists at least one $f \in \mathcal{F}$ such that the sets $f(X_1), \ldots, f(X_t) \subseteq Y$ are also pairwise disjoint. Such f is said to *separate* the sets $X_1, \ldots, X_t$.

Note the parameter set $\{w_1, w_2, \ldots, w_t\}$ is a multiset. The order of the w_i's do not matter, but the number of repetitions in the multiset is important.

We may also define the *matrix representation* for families of functions.

Definition 1.1.5. Let X, Y be finite sets with $|X| = n$ and $|Y| = q$. Given a family of functions $\mathcal{F}$ from X to Y with $|\mathcal{F}| = N$, construct a matrix A by listing the functions $f \in \mathcal{F}$ on different rows and elements $x \in X$ on different columns, with $f(x)$ being the matrix entry at row f and column x. A is said to be the *representation matrix* of $\mathcal{F}$.

Separating hash families are related to frameproof codes in the following manner.

Theorem 1.1.1 ([26])**.** *Let* A *be an* $N \times n$ *matrix with entries from* $\{0, 1, \ldots, q-1\}$*. The following are equivalent:*

(i) A *is the representation matrix of an* (N, n, q) w*-frameproof code.*

(ii) A *is the representation matrix of an* $\mathsf{SHF}(N; n, q, \{1, w\})$*.*

Proof. The frameproof property requires that for any choice of column set pairs $(\{c\}, P)$ with $|P| = w$ and $c \notin P$, we have that $c \notin \mathsf{desc}(P)$. This is equivalent to the existence of a row r such that $A(r, c) \neq A(r, c')$ for all $c' \in P$. The latter is the requirement for an $\mathsf{SHF}(N; n, q, \{1, w\})$. □

Due to this equivalence, we will often discuss frameproof codes in terms of separating hash families.

Example 1.1.2. Let $Q = \{0, 1, 2\}$. Construction 1 gives the code

$$\mathsf{C} = \{100, 200, 010, 020, 001, 002\}$$

for $N = 3$ and $n = 6$. The equivalent separating hash family is $\mathcal{F} = \{f_1, f_2, f_3\}$ where the functions f_i written in vector form (i.e. ith entry is the value at element i of Y) are

$$\begin{aligned} f_1 &= (1, 2, 0, 0, 0, 0) \\ f_2 &= (0, 0, 1, 2, 0, 0) \\ f_3 &= (0, 0, 0, 0, 1, 2). \end{aligned}$$

The representation matrix for the two equivalent structures is

$$\mathsf{A} = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

Here are some basic properties of separating hash families. First note that some parameter choices are not interesting to consider. For positive integers $w_1, w_2, \ldots, w_t$, if $n < \sum_i w_i$ then any (N, n, q)-code is an $\mathsf{SHF}(N; n, q, \{w_1, w_2, \ldots, w_t\})$. Also, if $q < t$ then no (N, n, q)-code is an $\mathsf{SHF}(N; n, q, \{w_1, w_2, \ldots, w_t\})$.

Theorem 1.1.2 ([28], [9])**.** *Let $\mathcal{F}$ be an $\mathsf{SHF}(N; n, q, \{w_1, w_2, \ldots, w_t\})$ with $\sum_i w_i \leq n$ and $q \geq t$.*

(i) If $w_1' \leq w_1$ then $\mathcal{F}$ is also an $\mathsf{SHF}(N; n, q, \{w_1', w_2, \ldots, w_t\})$.

(ii) Let $w_1' = w_1 + w_2$. Then $\mathcal{F}$ is also an $\mathsf{SHF}(N; n, q, \{w_1', w_3, \ldots, w_t\})$.

(iii) For any integer $c \geq 2$, there exists an $\mathsf{SHF}(\lceil \frac{N}{c} \rceil; n, q^c, \{w_1, w_2, \ldots, w_t\})$.

(iv) If there exists an $\mathsf{SHF}(M; q, q', \{w_1, w_2, \ldots, w_t\})$ then there exists an $\mathsf{SHF}(MN; n, q', \{w_1, w_2, \ldots, w_t\})$.

Proof. (i) Let $X_1', X_2, \ldots, X_t \subset X$ be such that $|X_1'| = w_1'$ and $|X_i| = w_i$ for $i = 2, \ldots, t$. Let $Z \subseteq X \setminus (X_1' \cup X_2 \cup \ldots \cup X_t)$ be such that $|Z| = w_1 - w_1'$. Taking $X_1 = X_1' \cup Z$, we have that some $f \in \mathcal{F}$ separates $X_1, \ldots, X_t$, so it must also separate $X_1', X_2, \ldots, X_t$.

(ii) Let $X_1', X_3, \ldots, X_t \subset X$ be such that $|X_1'| = w_1'$ and $|X_i| = w_i$ for $i = 3, \ldots, t$. Let X_1, X_2 be such that $X_1' = X_1 \cup X_2$ and $|X_1| = w_1$, $|X_2| = w_2$. We have that some $f \in \mathcal{F}$ separates $X_1, \ldots, X_t$, so it must also separate $X_1', X_3, \ldots, X_t$.

(iii) For simplicity, we assume that N is a multiple of c, but the proof extends to the general case. Suppose $N = cM$, and let $\mathcal{F} = \{f_{i,j}\}$ for $i = 1, \ldots, M$ and $j = 1, \ldots, c$. For each i, consider a new function $f_i : X \to Y^c$ defined as

$$f_i(x) = (f_{i,1}(x), f_{i,2}(x), \ldots, f_{i,c}(x)).$$

The new family $\mathcal{G} = \{f_i\}$ is an $\mathsf{SHF}(\lceil \frac{N}{c} \rceil; n, q^c, \{w_1, w_2, \ldots, w_t\})$.

(iv) Let $\mathcal{G}$ be an $\mathsf{SHF}(M; q, q', \{w_1, w_2, \ldots, w_t\})$ and define $\mathcal{H} = \{g \circ f : f \in \mathcal{F}, g \in \mathcal{G}\}$. $\mathcal{H}$ is an $\mathsf{SHF}(MN; n, q', \{w_1, w_2, \ldots, w_t\})$.

□

Parts (iii) and (iv) of Theorem 1.1.2 are often used for recursive constructions and/or bounds of separating hash families [28] [29].

It is worth noting that under a different attack model, frameproof codes can be shown to have anti-collusion traitor tracing properties. Trappe et al. [30] proposed a different embedding technique for fingerprinting codes that works well with multimedia content, e.g. sound, picture, video. The property they require for traceability under a collusion attack is weaker than the frameproof property [8], hence frameproof codes can be used in this manner for anti-collusion traitor tracing.

1.2 Organization of Thesis

Our thesis is organized as follows. Chapter 2 presents some old and new results regarding binary frameproof codes. Chapter 3 will focus on the relationship between combinatorial designs and frameproof codes, featuring some known connections and new results. Chapter 4 will describe some future directions suggested by this thesis.

Chapter 2

Frameproof Codes and Separating Hash Families

A large portion of this chapter will appear in [15], which is accepted for publication.

In this chapter, we prove a tight upper bound on the number of codewords in a frameproof code. More precisely, for an $(N, n, 2)$ w-frameproof code with $N \leq 3w$, we have that $n \leq N$, and the representation matrix of an $(N, N, 2)$ w-frameproof code with $N \leq 3w$ is equivalent to a permutation matrix. Finally, we prove a tight upper bound of a similar nature for separating hash families.

2.1 Known Results

We first present some known results on frameproof codes. The following is a general upper bound for the number of codewords n in terms of N, q and w.

Theorem 2.1.1 ([24]). *In an (N, n, q) w-frameproof code, the following bound holds:*

$$n \leq w(q^{\lceil \frac{N}{w} \rceil} - 1).$$

If we let $N = w$, the above bound is tight due to Construction 1. The following stronger bound applies for restricted alphabet size.

Theorem 2.1.2 ([4]). *Let N, n, q, w, d be positive integers such that $N = wd + 1$ and $q \geq w \geq 2$. Then for any (N, n, q) w-frameproof code, we have $n \leq q^{d+1} + O(q^d)$.*

A notable improvement over Theorem 2.1.2 has been proven by van Trung [32].

Theorem 2.1.3 ([32]). *Let N, n, q, w, d be positive integers such that $N = wd + 1$ and $q \geq w \geq 2$. Then for any (N, n, q) w-frameproof code, we have $n \leq q^{d+1}$.*

It is worth mentioning that the bound is tight when q is a prime power with $q \geq wd$ via a construction using orthogonal arrays [32]. We will give a definition for orthogonal arrays in Section 3.3.

We also present some known results about general separating hash families.

Theorem 2.1.4 ([5]). *If there exists an $\mathsf{SHF}(N; n, q, \{w_1, \ldots, w_t\})$ with $w_1, w_2 \leq w_i$ for $i = 3, \ldots, t$ then*

$$n \leq \gamma q^{\lceil \frac{N}{u-1} \rceil}$$

where $u = \sum_i w_i$ and $\gamma = (w_1 w_2 + u - w_1 - w_2)$.

Theorem 2.1.5 ([2]). *If there exists an $\mathsf{SHF}(N; n, q, \{w_1, \ldots, w_t\})$ then*

$$n \leq (u-1) q^{\lceil \frac{N}{u-1} \rceil}$$

where $u = \sum_i w_i$.

Theorem 2.1.6 ([3]). *If there exists an $\mathsf{SHF}(N; n, q, \{w_1, \ldots, w_t\})$ with $t \geq 3$ and $u = \sum_i w_i \geq 4$ then*

$$n \leq (u-1) q^{\lceil \frac{N}{u-1} \rceil} + 2 - 2\sqrt{3q^{\lceil \frac{N}{u-1} \rceil} + 1}.$$

2.2 New Result

We will be discussing a new result regarding binary frameproof codes in this section. The matrix representation for separating hash families will be used extensively as it makes the analysis more clear. Since the alphabet is binary, the representation matrix will contain only the entries 0 and 1.

Here is a useful observation about binomial coefficients that we will often employ.

Lemma 2.2.1. *Let w, n be positive integers such that $w+1 \leq n$. Then for $i = 1, 2, \ldots, n-w-1$, we have $i\binom{n-i}{w} > (i+1)\binom{n-i-1}{w}$ if and only if $(i+1)(w+1) > n+1$. In particular, we have*

$$\binom{n-1}{w} > 2\binom{n-2}{w} > 3\binom{n-3}{w} > \cdots > j\binom{n-j}{w}. \tag{2.1}$$

for $j \leq n - w$ whenever $n \leq 2w$.

Proof.

$$
\begin{aligned}
i\binom{n-i}{w} > (i+1)\binom{n-i-1}{w} &\Leftrightarrow \frac{i(n-i)!}{(n-i-w)!\cdot w!} > \frac{(i+1)(n-i-1)!}{(n-i-w-1)!\cdot w!} \\
&\Leftrightarrow \frac{i(n-i)}{(n-i-w)} > i+1 \\
&\Leftrightarrow ni - i^2 > ni + n - i^2 - i - iw - w \\
&\Leftrightarrow i + iw + w > n \\
&\Leftrightarrow (i+1)(w+1) > n+1.
\end{aligned}
$$

Note that Equation 2.1 holds if and only if $i\binom{n-i}{w} > (i+1)\binom{n-i-1}{w}$ holds for $i = 1$, which corresponds to $2(w+1) > n+1$ or equivalently $n \leq 2w$. □

Definition 2.2.1. Let A be the representation matrix of an $\mathsf{SHF}(N; n, 2, \{1, w\})$. A row r of A is said to be of *weight* i if r contains exactly i entries equal to 1. Two rows r_1 and r_2 of A are said to be *overlapped* if they share a column in which both rows have an entry equal to 1. If r_1 and r_2 are not overlapped, we say that they are *disjoint*.

Definition 2.2.2. The representation matrix A of an $\mathsf{SHF}(N; n, 2, \{1, w\})$ is said to be in *standard form* if every row contains at most $\frac{n}{2}$ entries equal to 1.

For an arbitrary $\mathsf{SHF}(N; n, 2, \{1, w\})$ A, it is clear that both 0 and 1 have to occur in each row of A, otherwise that row would not contribute to the separation of any column set pair (C_1, C_2). Hence we may assume that A contains no row of weight 0 in standard form, by simply removing any such row and replacing it with an arbitrary row of weight 1.

The following observation will be used throughout this paper.

Lemma 2.2.2. *Let A be an $\mathsf{SHF}(N; n, 2, \{1, w\})$. Suppose row r of A is of weight $i \leq n/2$. If $i < w$, then row r separates exactly $i\binom{n-i}{w}$ column set pairs (C_1, C_2) where $|C_1| = 1$ and $|C_2| = w$. If $i \geq w$, then row r separates exactly $i\binom{n-i}{w} + \binom{i}{w}(n-i)$ column set pairs (C_1, C_2) where $|C_1| = 1$ and $|C_2| = w$.*

We will now prove a bound for binary frameproof codes.

Theorem 2.2.3. *Let w, N be positive integers such that $w \geq 3$ and $w+1 \leq N \leq 2w+1$. Suppose there exists an $\mathsf{SHF}(N; n, 2, \{1, w\})$. Then $n \leq N$.*

Proof. Suppose, by contradiction, that there exists an $\mathsf{SHF}(N; n, 2, \{1, w\})$ with $n = N+1$. Let A be its $N \times (N+1)$ matrix representation over the alphabet $\{0, 1\}$. Let T be the total number of pairs of disjoint column sets (C_1, C_2) of A with $|C_1| = 1$ and $|C_2| = w$ that need to be separated. Then we have $T = \binom{n}{w}(n-w) = n\binom{n-1}{w}$.

Consider the following three cases regarding the number of columns of A.

(i) $n = N + 1 \leq 2w$ (i.e. $N \leq 2w - 1$).

Using Lemma 2.2.1 we see that

$$\binom{n-1}{w} > 2\binom{n-2}{w} > 3\binom{n-3}{w} > \cdots > (w-1)\binom{n-(w-1)}{w}.$$

The term $j\binom{n-j}{w}$ in these inequalities corresponds to the number of column set pairs (C_1, C_2) separated by a row of weight j. Hence a row of weight 1 separates the largest number of column set pairs (C_1, C_2), namely $\binom{n-1}{w} = \binom{N}{w}$. Moreover, since A has N rows, the maximal number of column set pairs (C_1, C_2) that can be separated by all the rows of A is therefore $N\binom{N}{w} = (n-1)\binom{n-1}{w}$. This is a contradiction, since $(n-1)\binom{n-1}{w} < T$.

(ii) $n = N + 1 = 2w + 1$ (i.e. $N = 2w$).

Observe that we have

$$\binom{n-1}{w} = \binom{N}{w} = \binom{2w}{w} = 2\binom{2w-1}{w} = 2\binom{n-2}{w}$$

in this case. This observation together with Lemma 2.2.1 give rise to the following inequalities about the number of column set pairs (C_1, C_2) separated by a row of weight j, where $j = 1, \ldots, w$:

$$\binom{n-1}{w} = 2\binom{n-2}{w} > 3\binom{n-3}{w} > \cdots > (w-1)\binom{n-(w-1)}{w}$$

and

$$(w-1)\binom{n-(w-1)}{w} > w\binom{n-w}{w} + n - w.$$

The last inequality can be easily checked, while all other inequalities follow from Lemma 2.2.1 Note that the last term of the inequalities corresponds to the case of a row of weight w. Again, this implies that a row of A can separate at most $\binom{n-1}{w} = \binom{N}{w}$ column set pairs (C_1, C_2). Thus all N rows of A can separate at most $N\binom{N}{w} = 2w\binom{2w}{w}$ column set pairs (C_1, C_2), whereas the total number of column set pairs (C_1, C_2) that need to be separated is $T = \binom{N+1}{w}(N+1-w) = (2w+1)\binom{2w}{w}$, a contradiction.

(iii) $n = N + 1 = 2w + 2$ (i.e. $N = 2w + 1$).

In this case we have the following inequalities

$$2\binom{2w}{w} > \binom{2w+1}{w} > 3\binom{2w-1}{w} > \cdots > (w-1)\binom{w+3}{w}$$

and

$$(w-1)\binom{w+3}{w} > w\binom{w+2}{w} + (w+2) > 2(w+1)^2.$$

The last two inequalities can be easily checked, while the other inequalities follow from Lemma 2.2.1. Here the first term of the inequalities corresponds to a row of weight 2; the second term to a row of weight 1; the third term to a row of weight 3, etc., the last term corresponds to a row of weight $\lfloor n/2 \rfloor = (w+1)$.

Recall that the total number of column set pairs (C_1, C_2) is $T = \binom{n}{w}(n-w) = \binom{2w+2}{w}(w+2)$. We show that if each row of A separates a maximal number of column set pairs (C_1, C_2), then all the $N = 2w+1$ rows of A fail to separate all T column set pairs (C_1, C_2). In fact, this corresponds to the first term of the above inequalities. This is the case for which each row of A is of weight 2. So each row will separate $2\binom{2w}{w}$ column set pairs (C_1, C_2). Hence all $N = 2w + 1$ rows of A will separate at most

$$Z = 2(2w+1)\binom{2w}{w}$$

column set pairs (C_1, C_2) of A. Now using the equality $\binom{n}{m} = \frac{n}{n-m}\binom{n-1}{m}$, we see that

$$T = \binom{2w+2}{w}(w+2) = \frac{(2w+2)}{(w+2)}\frac{(2w+1)}{(w+1)}(w+2)\binom{2w}{w} = 2(2w+1)\binom{2w}{w} = Z.$$

However, if each row of A is of weight 2, then there must exist two overlapped rows, say r_1 and r_2. These rows r_1 and r_2 will then separate $\binom{2w-1}{w}$ common column set pairs (C_1, C_2). This leads to a contradiction, since all the rows of A will separate less than T column set pairs (C_1, C_2). This completes the proof.

□

Recall that a binary $N \times N$ matrix A is called a *permutation matrix* of degree N if A has precisely one entry equal to 1 in each row and each column, and 0s elsewhere. Construction 1 in matrix representation gives every permutation matrix of degree N depending on the

ordering of codewords, hence any permutation matrix of degree N is the representation matrix of an $\mathsf{SHF}(N; N, 2, \{1, w\})$ for any $N \geq w+1$. Thus the bound of Theorem 2.2.3 is tight. In the following, we prove a stronger result which states that permutation matrices are the only solutions for an $\mathsf{SHF}(N; N, 2, \{1, w\})$ with $w+1 \leq N \leq 2w+1$ and $w \geq 3$.

Theorem 2.2.4. *Let w, N be positive integers such that $w \geq 3$ and $w+1 \leq N \leq 2w+1$. Suppose there exists an $\mathsf{SHF}(N; n, 2, \{1, w\})$ with $n = N$. Then its representation matrix in standard form is a permutation matrix of degree N.*

Proof. Let A be the representation matrix of an $\mathsf{SHF}(N; N, 2, \{1, w\})$ in standard form with $w+1 \leq N \leq 2w+1$ and $w \geq 3$. Consider two cases.

(i) $n = N \leq 2w$.

Recall that the total number of column set pairs (C_1, C_2) of A that need to be separated is $T = \binom{N}{w}(N-w)$. By Lemma 2.2.1 each row of A can separate at most $\binom{N-1}{w}$ column set pairs (C_1, C_2), and this case occurs when each row is of weight 1. Thus the largest number of separated column set pairs (C_1, C_2) obtained by N rows of A is $N\binom{N-1}{w} = \binom{N}{w}(N-w)$. This number is achieved if and only if the unique entries 1 of the rows belong to the different columns, i.e., A is a permutation matrix of degree N.

(ii) $n = N = 2w+1$.

In this case we have $T = \binom{2w+1}{w}(w+1)$. A row r of A can separate at most $\binom{2w}{w}$ column set pairs (C_1, C_2). This number corresponds to r being of either weight 1 or weight 2. Further, the maximum number of possible separated column set pairs (C_1, C_2) which may be achieved by all the rows of A is $(2w+1)\binom{2w}{w}$. To achieve the maximum number $(2w+1)\binom{2w}{w}$ of separated column set pairs, any two rows of A have to separate disjoint sets of column set pairs (C_1, C_2). This implies that any two rows of A are disjoint. This is equivalent to saying that each column of A contains exactly one entry 1, otherwise if two rows r_1 and r_2 are overlapped, then these two rows separate a common non-empty subset of column set pairs (C_1, C_2), which is a contradiction. Therefore, A is a permutation matrix of degree $2w+1$.

□

The following theorem provides a useful tool for relating Theorem 2.2.3 and Theorem 2.2.4.

Theorem 2.2.5. *Let $w \geq 3$, $N \geq w+1$ and suppose that all $\mathsf{SHF}(N; N, 2, \{1, w\})$ in standard form are permutation matrices. If $\mathsf{SHF}(N; n, 2, \{1, w\})$ exists, then $n \leq N$.*

Proof. Suppose not, then there exists some $\mathsf{SHF}(N; N+1, 2, \{1, w\})$, say A. Let B be the submatrix formed by the first N columns of A. We may assume without loss of generality that B is in standard form (we may need to permute 0s and 1s in each row of A to achieve this). Then B is a permutation matrix. Thus each row of A has at most two entries of 1.

Since $N \geq w+1 \geq 4$, we have that $N/2 \geq 2$. Let C be the submatrix formed by the last N columns of A. Each row of C has at most two entries of 1 as well, so C is in standard form, and hence it is a permutation matrix. This implies the first and last columns of A are identical, which is a contradiction since $(\{1\}, \{N+1\})$ cannot be separated. □

We may use Theorem 2.2.5 to give a second proof of Theorem 2.2.3 using Theorem 2.2.4. In light of this result, it is also important to consider the question "when are permutation matrices the only representatives of $\mathsf{SHF}(N; N, 2, \{1, w\})$ in standard form?" We give an affirmative answer for $w \geq 4$ and $N \leq 3w$ through a series of lemmas below.

Lemma 2.2.6. *Let $w \geq 3$ and let A be the representation matrix of an $\mathsf{SHF}(N; N, 2, \{1, w\})$. Suppose that all $\mathsf{SHF}(N-1; N-1, 2, \{1, w\})$ in standard form are permutation matrices. If A contains a row of weight 1, then A is a permutation matrix.*

Proof. We can write A in the form

$$\mathsf{A} = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline & & & \\ & & \mathsf{B} & \\ & & & \end{array}\right)$$

Let B be the $(N-1)\times(N-1)$ matrix obtained from A by removing the first row and the first column of A. Then B is the representation matrix of an $\mathsf{SHF}(N-1; N-1, 2, \{1, w\})$. We may assume without loss of generality that B is in standard form, and hence it is a permutation matrix.

By permuting the columns of A, if necessary, we may assume that B is the identity matrix. Consider column set pairs $(C_x = \{x\}, C_{1,y,z} = \{1, y, z\})$ with $x, y, z = \{2, \dots, N\}$ and $x \neq y \neq z \neq x$. Since B is the identity matrix, a row that separates $(C_x, C_{1,y,z})$ must have entry 0 in columns $1, y, z$ and entry 1 in column x. Thus row x is the unique row separating $(C_x, C_{1,y,z})$. It follows that A is a permutation matrix. □

Lemma 2.2.7. *Let $w \geq 4$, $N \leq 3w$, and let A be the representation matrix of an $\mathsf{SHF}(N; N, 2, \{1, w\})$. Suppose the first row of A is of weight $i_0 \leq w$ with $\mathsf{A}(1,1) = 1$. Let B be the submatrix obtained by deleting the first row and first column of A. Then B is an $\mathsf{SHF}(N-1; N-1, 2, \{1, w\})$.*

Proof. If A contains a row of weight 1 then Lemma 2.2.6 applies. For the remainder of this proof, we assume that A contains no row of weight 1.

Suppose B is not an $\mathsf{SHF}(N-1; N-1, 2, \{1, w\})$. Then there exists some column set pair $(C_1 = \{x\}, C_2)$ with $|C_2| = w$ that cannot be separated by B. If x corresponds to a column of A that has an entry of 0 in the first row then C_2 contains a column of A that has an entry of 0 in the first row since $i_0 - 1 < w$. But then A also cannot separate (C_1, C_2); a contradiction. Thus x contains a 1 in the first row, and all columns of C_2 correspond to columns of A with 0's in the first row (otherwise A still cannot separate (C_1, C_2)).

Permute the columns of A so that x corresponds to column 2 and columns in C_2 correspond to columns $3, \ldots, w+2$. The matrix A is now

$$\mathsf{A} = \left(\begin{array}{ccccc|c} 1 & 1 & 0 & \cdots & 0 & \\ \hline & & & & & \\ & & & & & \end{array}\right)$$

For $1 \leq i \leq w$, let $C_i = \{3, \ldots, w+2\} \setminus \{i+2\}$. The column set pair $(\{2\}, C_i \cup \{1\})$ must be separated by A. By permuting 0's and 1's if necessary, there is some row $r_i \neq 1$ with entry 1 in column 2 and entry 0 in columns of C_i. Since $C_i \cup C_j = \{3, \ldots, w+2\}$ for $i \neq j$ and B does not separate $(\{2\}, \{3, \ldots, w+2\})$, we have that $r_i \neq r_j$ for $i \neq j$. Moreover, entry i of r_i must also be a 1. Let $R_1 = \{r_1, \ldots, r_w\}$, and by permuting the rows of A we have

$$\mathsf{A} = \left(\begin{array}{ccccccc|c} 1 & 1 & 0 & 0 & 0 & \cdots & 0 & \\ \hline 0 & 1 & 1 & 0 & 0 & \cdots & 0 & \\ 0 & 1 & 0 & 1 & 0 & \cdots & 0 & \\ \vdots & \vdots & & & & \ddots & \vdots & \\ 0 & 1 & 0 & 0 & 0 & \cdots & 1 & \\ * & * & * & * & * & \cdots & * & \\ \vdots & \vdots & & & & \ddots & \vdots & \\ * & * & * & * & * & \cdots & * & \end{array}\right)$$

Next, consider $C_i' = \{2, \ldots, w+2\} \setminus \{i+2\}$ for $i = 1, \ldots, w$. The column set pair $(\{i+2\}, C_i')$ must be separated by A with some row $r_i' \neq 1$ and $r_i' \notin R_1$. By permuting the 0's and 1's if necessary, r_i' has entry 1 in column $(i+2)$ and entry 0 in columns in C_i'. Moreover, $r_i' \neq r_j'$ for $i \neq j$. Now let $R_2 = \{r_1', \ldots, r_w'\}$, and by permuting the rows of A we have

$$\mathsf{A} = \left(\begin{array}{ccccccc|c} 1 & 1 & 0 & 0 & 0 & \cdots & 0 & \\ \hline 0 & 1 & 1 & 0 & 0 & \cdots & 0 & \\ 0 & 1 & 0 & 1 & 0 & \cdots & 0 & \\ \vdots & \vdots & & & & \ddots & \vdots & \\ 0 & 1 & 0 & 0 & 0 & \cdots & 1 & \\ \hline * & 0 & 1 & 0 & 0 & \cdots & 0 & \\ * & 0 & 0 & 1 & 0 & \cdots & 0 & \\ \vdots & \vdots & & & & \ddots & \vdots & \\ * & 0 & 0 & 0 & 0 & \cdots & 1 & \\ \hline * & * & * & * & * & \cdots & * & \\ \vdots & \vdots & & & & \ddots & \vdots & \\ * & * & * & * & * & \cdots & * & \end{array}\right)$$

We now do the following addition of rows in steps, starting with $R_3 = \emptyset$:

Step 1

Let a be the column 1 entry of r_1'. If $a = 1$, consider the column pair $(\{3\}, \{1, \ldots, w+1\} \setminus \{3\})$, which must be separated by some row $r_1'' \neq 1$ of A. Note that $r_1'' \notin R_1$ and $r_1'' \notin R_2$. Add r_1'' to R_3.

If $a = 0$, consider the column pairs $(\{3\}, C_{1,j}'' = \{2, 4, 5, w+j+2\})$ for $j = 1, \ldots, N-w-2$. Since $w \geq 4$, we have that A separates $(\{3\}, C_{1,j}'')$. If r_1' separates every such pair then r_1' is a weight 1 row; a contradiction to A having no weight 1 rows. Thus there is some j such that another row of A, call it again r_1'', that separates $(\{3\}, C_{1,j}'')$. Note that $r_1'' \neq 1$, $r_1'' \notin R_1$ and $r_2'' \notin R_2$. Add r_1'' to R_3.

Step 2

Let a be the column 1 entry of r_2'. If $a = 1$, consider the column set pair $(\{4\}, \{1, \ldots, w+1\} \setminus \{4\})$, which must be separated by some row $r_2'' \neq 1$ of A. Note that $r_2'' \notin R_1 \cup R_2$ and $r_2'' \neq r_1''$. Add r_2'' to R_3.

If $a = 0$, consider the column pairs $(\{4\}, C''_{2,j} = \{2, 3, 5, w+j+2\})$ for $j = 1, \ldots, N - w - 2$. Similar to Step 1, there exists some j for which another row of A, call it again r''_2, that separates $(\{4\}, C''_{2,j})$. Again $r''_2 \notin R_1 \cup R_2$ and $r''_2 \neq r''_1$. Add r''_2 to R_3.

Steps $i = 3, \ldots, w - 1$

Let a be the column 1 entry of r'_i. If $a = 1$, consider the column set pair $(\{i + 2\}, \{1, \ldots, w+1\} \setminus \{i+2\})$, which must be separated by some row $r''_i \neq 1$ of A. Note that $r''_i \notin R_1 \cup R_2 \cup R_3$. Add r''_i to R_3.

If $a = 0$, consider the column pairs $(\{i + 2\}, C''_{i,j} = \{2, 3, \ldots, i + 1, w + j + 2\})$ for $j = 1, \ldots, N - w - 2$. Since $|C''_{i,j}| = i + 1 \leq w$, some row of A separates $(\{i+2\}, C''_{i,j})$. Similar to Step 1, there exists some j for which another row of A, call it again r''_i, that separates $(\{i + 2\}, C''_{i,j})$. Again $r''_i \notin R_1 \cup R_2 \cup R_3$. Add r''_i to R_3.

Step w

Consider the column set pair $(\{1\}, \{2, \ldots, w+1\})$, which must be separated by some row r of A. Clearly $r \notin R_1 \cup R_2 \cup R_3$. Add r to R_3.

At the end of Step w, we have added w distinct rows to R_3, so A has at least $|R_1 \cup R_2 \cup R_3| + 1 = w + w + w + 1 = 3w + 1$ rows. This contradicts $N \leq 3w$, so Lemma 2.2.7 holds. □

Lemma 2.2.8. *Let $w \geq 4$, $w + 1 \leq N \leq 3w$, and let A be the representation matrix of an $\mathsf{SHF}(N; N, 2, \{1, w\})$. Suppose that some row of A is of weight at most w and all $\mathsf{SHF}(N - 1; N - 1, 2, \{1, w\})$ in standard form are permutation matrices. Then A is a permutation matrix.*

Proof. If A contains a row of weight 1, we can use Lemma 2.2.6 to show that A is a permutation matrix. For the remainder of this proof, we may assume that A contains no row of weight 1. Assume without loss of generality that the first row of A is of weight i_0 where $2 \leq i_0 \leq w$.

Suppose to the contrary that A is not a permutation matrix. By permuting the columns of A if necessary, we may assume that row 1 is $1^{i_0}0^{N-i_0}$. Let B be the $(N - 1) \times (N - 1)$ submatrix of A obtained by deleting the first row and first column of A.

By Lemma 2.2.7, we have that B is an $\mathsf{SHF}(N - 1; N - 1, 2, \{1, w\})$, and hence it is a permutation matrix. For row x of A, $x = 2, \ldots, N$, let c_x be the unique column of A that contains a 1 in row x. Consider the column set pair $(C_x = \{c_x\}, C'_x = C''_x \cup \{1\})$ where C''_x is some set of $w - 1$ columns not containing c_x whose entries on row 1 contains at least one

0. This is possible since $N \geq w+2 \geq i_0+2$. The only row that can separate this column set pair is row x, which forces its first entry to be a 0. Thus we have shown that

$$\mathsf{A} = \left(\begin{array}{c|cc} 1 & 1 & \\ \hline 0 & & \\ \vdots & & \mathsf{B} \\ 0 & & \end{array}\right).$$

The column set pair $(C_1 = \{1\}, C_2 = \{2,3\})$ cannot be separated by A; this is a contradiction. □

Theorem 2.2.9. *Let w, N be positive integers such that $w \geq 4$ and $2w+2 \leq N \leq 3w$. Suppose there exists an $\mathsf{SHF}(N; N, 2, \{1, w\})$. Then its representation matrix in standard form is a permutation matrix of degree N.*

Proof. The proof is by induction on $N = 2w+1, \ldots, 3w$. The base case $N = 2w+1$ is given by Theorem 2.2.4. Suppose that $N > 2w+1$ and all $\mathsf{SHF}(N-1; N-1, 2, \{1, w\})$ in standard form are permutation matrices. By Lemma 2.2.8, we only need to show that some row of weight at most w exists.

Let A be an $\mathsf{SHF}(N; N, 2, \{1, w\})$ in standard form. Fix some i where $w+1 \leq i \leq N/2$. The average number of column set pairs separated by a row is

$$\alpha = \frac{(N-w)\binom{N}{w}}{N} = \binom{N-1}{w}.$$

Let β_i be the number of column set pairs separated by a row of weight i. Then

$$\beta_i = i\binom{N-i}{w} + (N-i)\binom{i}{w} \leq N\binom{N-i}{w}.$$

Since $i \geq w+1$, we have

$$\begin{aligned}
\alpha &= \binom{N-1}{w} \\
&= \frac{(N-1)(N-2)\cdots(N-w)}{(N-w-1)(N-w-2)\cdots(N-2w)}\binom{N-w-1}{w} \\
&\geq \frac{(N-1)(N-2)\cdots(N-w)}{(N-w-1)(N-w-2)\cdots(N-2w)}\binom{N-i}{w} \\
&\geq \left(\frac{N-1}{N-w-1}\right)^w \binom{N-i}{w} \\
&\geq \left(\frac{3w+1-1}{3w+1-w-1}\right)^w \binom{N-i}{w} \\
&= \left(\frac{3}{2}\right)^w \binom{N-i}{w}.
\end{aligned}$$

For $w \geq 8$, one can check that $\left(\frac{3}{2}\right)^w > 3w \geq N$, so $\alpha > \beta_i$. It is straightforward to compute α and β_i for $4 \leq w \leq 7$ and confirm that $\alpha > \beta_i$ for all relevant values of i. Since $\alpha > \beta_i$ for every $i \geq w+1$ and A contains no row of weight $N/2+1$ or higher, there must exist some row of weight at most w. □

Finally, we give a bound similar to Theorem 2.2.3.

Theorem 2.2.10. *Let w, N be positive integers such that $w \geq 4$ and $2w+2 \leq N \leq 3w$. Suppose there exists an $\mathsf{SHF}(N; n, 2, \{1, w\})$. Then $n \leq N$.*

Proof. By Theorem 2.2.9, all $\mathsf{SHF}(N; N, 2, \{1, w\})$ in standard form are permutation matrices, hence the proof follows from Theorem 2.2.5. □

Theorems 2.2.9 and 2.2.10 can be extended to the case $w \geq 3$ via a tedious case analysis, which we exclude here. The analysis can be found in [15]. Gathering the results proven so far, we have the following main theorem.

Theorem 2.2.11. *Let w, N be positive integers such that $w \geq 3$ and $w+1 \leq N \leq 3w$. Suppose there exists an $\mathsf{SHF}(N; n, 2, \{1, w\})$. Then $n \leq N$. Furthermore, its representation matrix in standard form is a permutation matrix of degree N.*

For completeness, we include a discussion regarding the $w = 2$ case.

Theorem 2.2.12. *For every $N \geq 3$, there exists an* $\mathsf{SHF}(N; N+1, 2, \{1,2\})$.

Proof. Take the $N \times N$ identity matrix and append to it a column of 1s; call this matrix A. We will show that A is an $\mathsf{SHF}(N; N+1, 2, \{1,2\})$.

Let $(C_1 = \{x\}, C_2 = \{y, z\})$ be a column set pair. First consider $1 \leq x \leq N$. If $1 \leq y, z \leq N$ then (C_1, C_2) is clearly separated by A. Suppose w.l.o.g. that $z = N+1$, then row y has entry 1 in columns y, z and entry 0 in column x, so (C_1, C_2) is again separated.

Finally, consider $x = N+1$, so $1 \leq y, z \leq N$. Since $N \geq 3$, there is some row $w \notin \{y, z\}$, so row w has entry 0 in columns y, z and entry 1 in column x, so (C_1, C_2) is separated. □

Theorem 2.2.12 above shows that Theorem 2.2.3 does not hold when $w = 2$. We will also demonstrate that Theorem 2.2.4 and Theorem 2.2.5 do not hold when $w = 2$.

Theorem 2.2.13. *The matrix*

$$\mathsf{A} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is an $\mathsf{SHF}(4; 4, 2, \{1,2\})$.

The result in Theorem 2.2.13 can be extended to $N > 4$ by constructing the matrix

$$\mathsf{B} = \begin{pmatrix} \mathsf{A} & 0 \\ 0 & I_k \end{pmatrix}$$

where A is from Theorem 2.2.13 and I_k is the $k \times k$ identity matrix for $k = N - 4$. Observe that for every column $x, y \in \{1, 2, 3\}$, there exist rows r_x, r_y such that $r_x(x) = 1$, $r_x(y) = 0$ and $r_y(x) = 0$, $r_y(y) = 1$. It is straightforward to verify that B is indeed an $\mathsf{SHF}(N; N, 2, \{1,2\})$. Theorem 2.2.14 below covers the last case $N = 3$, and shows that Theorem 2.2.5 does hold when $w = 2$.

Theorem 2.2.14. *The representation matrix of an* $\mathsf{SHF}(3; 3, 2, \{1,2\})$ *in standard form is a permutation matrix.*

Proof. In standard form, every row is of weight 1. Two distinct rows must not overlap, so each column also has one 1. □

A recent arXiv preprint [23] has improved our results. They proved that if $w \geq 3$ and $w+1 \leq N \leq \binom{w+1}{2} - 1$ and if there exists an $\mathsf{SHF}(N; n, 2, \{1, w\})$ then $n = N$; furthermore, its representation matrix in standard form is a permutation matrix of degree N. Their bound is strictly better than that of Theorem 2.2.11 if $w \geq 6$.

Here is an interesting problem that is suggested by our work: For a given w, find the smallest N such that there exists an $\mathsf{SHF}(N; n, 2, \{1, w\})$ with $n > N$. Shangguan et al. have shown that this quantity is at most $(1 + o(1))w^2$ [23]. A closely related problem is to find the smallest N such that there exists an $\mathsf{SHF}(N; N, 2, \{1, w\})$ whose representation matrix is not a permutation matrix. We will discuss these two problems in Section 3.4.4.

2.3 Generalization

In this section, we give an extension of Theorem 2.2.3 to separating hash families of a larger type. More precisely, our result proves a tight lower bound on N for an $\mathsf{SHF}(N; n, q, \{w_1^{q-1}, w_2\})$ where w_1^{q-1} denotes the multiset consisting of $q-1$ copies of w_1. The idea is to generalize the proof of Theorem 2.2.3 by counting the total number of column q-tuples separated versus the number of column q-tuples separated by a single row. Assuming the best case of non-overlap, we can then give a bound on the number of rows needed to achieve the total. We first define the generalized notion of weight.

Definition 2.3.1. Let $x \in Q^n$ with $Q = \{0, 1, \ldots, q-1\}$. We say that x is of *weight* $(i_1, i_2, \ldots, i_{q-1})$ if the number of entries of k in x is exactly i_k for every $k = 1, \ldots, q-1$. The number of entries equal to 0 is thus $i_0 = n - \sum_{k=1}^{q-1} i_k$.

The next definition gives a simplified notation for counting the number of column set q-tuples separated by a row of weight $(i_1, i_2, \ldots, i_{q-1})$.

Definition 2.3.2. Let w_1, w_2 be positive integers with $w_1 < w_2$. For integers $i_0, i_1, \ldots, i_{q-1}$ with $i_0 \geq w_2$, $i_k \geq w_1$ for $k = 1, \ldots, q-1$ and $n \geq \sum_{k=0}^{q-1} i_k$, define

$$T^{(q-1)}_{w_1,w_2,n}(i_1, \ldots, i_{q-1}) = \binom{i_1}{w_1}\binom{i_2}{w_1}\cdots\binom{i_{q-1}}{w_1}\binom{n - \sum_{k=1}^{q-1} i_k}{w_2}.$$

Lemma 2.3.1. *Let w_1, w_2 be positive integers with $w_1 < w_2$. For integers $i_0, i_1, \ldots, i_{q-1}$ with $i_0 \geq w_2$, $w_1 \leq i_k < w_2$ for $k = 1, \ldots, q-1$ and $n \geq \sum_{i=0}^{q-1} i_k$, the number of column set q-tuples separated by a row of weight $(i_1, \ldots, i_{q-1})$ is*

$$Z = (q-1)! \cdot T^{(q-1)}_{w_1,w_2,n}(i_1, \ldots, i_{q-1}).$$

Proof. Since $w_1 \leq i_k < w_2$ for $k = 1, \ldots, q-1$, it is clear that a row of weight $(i_1, \ldots, i_{q-1})$ only separates column set q-tuples of the form $(C_1, \ldots, C_q)$ with $|C_k| = w_1$ for $k = 1, \ldots, q-1$ and $|C_q| = w_2$. The $(q-1)!$ term comes from the number of permutations of the sets $C_1, \ldots, C_{q-1}$. □

The following lemma is a generalization of Lemma 2.2.1.

Lemma 2.3.2. *Let w_1, w_2 be positive integers such that $w_1 < w_2$, and let q, n be positive integers with $q \geq 2$ and*

$$w_2 + (q-1)w_1 \leq n \leq w_2 + (q-1)w_1 + \frac{w_2}{w_1} - 1.$$

Then for every $k = 1, \ldots, q-1$, we have

$$T^{(q-1)}_{w_1,w_2,n}(i_1, \ldots, i_{q-1}) > T^{(q-1)}_{w_1,w_2,n}(i_1, \ldots, i_{k-1}, i_k + 1, i_{k+1}, \ldots, i_{q-1}).$$

In particular, $T^{(q-1)}_{w_1,w_2,n}$ obtains its global maximum at $(w_1, \ldots, w_1)$.

Proof.

$$\begin{aligned}
&T^{(q-1)}_{w_1,w_2,n}(i_1, \ldots, i_{q-1}) > T^{(q-1)}_{w_1,w_2,n}(i_1, \ldots, i_{k-1}, i_k + 1, i_{k+1}, \ldots, i_{q-1}) \\
\Leftrightarrow &\binom{i_k}{w_1}\binom{n - \sum_{k=1}^{q-1} i_k}{w_2} > \binom{i_k + 1}{w_1}\binom{n - \sum_{k=1}^{q-1} i_k - 1}{w_2} \\
\Leftrightarrow &\frac{i_k - w_1 + 1}{i_k + 1} > \frac{n - \sum_{k=1}^{q-1} i_k - w_2}{n - \sum_{k=1}^{q-1} i_k}.
\end{aligned}$$

Letting $I = \sum_{k=1}^{q-1} i_k$ and rearranging the inequality gives

$$\begin{aligned}
&(i_k + 1 - w_1)(n - I) > (n - I - w_2)(i_k + 1) \\
\Leftrightarrow &-w_1(n - I) > -w_2(i_k + 1) \\
\Leftrightarrow &n\frac{w_1}{w_2} < i_k + 1 + \frac{w_1}{w_2} I \\
\Leftrightarrow &n < i_k \frac{w_2}{w_1} + I + \frac{w_2}{w_1}
\end{aligned}$$

where the last inequality holds by the assumption $n < w_2 + (q-1)w_1 + \frac{w_2}{w_1}$ since $w_1 \leq i_k$ and $(q-1)w_1 \leq I$. □

Lemma 2.3.3. *Let q, w be positive integers with $q \geq 3$, $w \geq 2$, and let $n = 2w + q - 2$. Then*

$$(q-1)! \cdot T^{(q-1)}_{1,w,n}(1, \ldots, 1) > 2(q-2)! \cdot T^{(q-1)}_{1,w,n}(1, \ldots, 1, w).$$

Proof. Expanding the desired inequality gives

$$\begin{aligned}
&(q-1)!\binom{1}{1}^{q-1}\binom{n-q+1}{w} > 2(q-2)!\binom{1}{1}^{q-2}\binom{w}{1}\binom{w}{w} \\
\Leftrightarrow\ &(q-1)\binom{2w-1}{w} > 2w.
\end{aligned}$$

One can check that $\binom{2w-1}{w} > w$ for $w \geq 2$, and the proof follows since $q - 1 \geq 2$. □

We are now ready to prove our generalization of Theorem 2.2.3.

Theorem 2.3.4. *Let w_1, w_2 be positive integers with $w_1 < w_2$, and let q, n be positive integers with $q \geq 2$ and*

$$w_2 + (q-1)w_1 \leq n \leq w_2 + (q-1)w_1 + \frac{w_2}{w_1} - 1.$$

If there exists an $\mathsf{SHF}(N; n, q, \{w_1^{q-1}, w_2\})$ then

$$N \geq \frac{1}{(q-1)!}\binom{n}{w_1}\binom{n-w_1}{w_1}\cdots\binom{n-(q-2)w_1}{w_1}.$$

Proof. Let A be the representation matrix of an $\mathsf{SHF}(N; n, q, \{w_1, \ldots, w_1, w_2\})$. For any row r of A and $k \in \{0, 1, \ldots, q-1\}$, let i_k be the number of occurrences of symbol k on row r. By permuting the alphabet on row r if necessary, we may assume without loss of generality that $i_1 \leq i_2 \leq \ldots \leq i_{q-1} \leq i_0$. Furthermore, we may assume that $i_1 \geq w_1$ and $i_0 \geq w_2$, since otherwise row r cannot separate any column set q-tuple $(C_0, C_1, \ldots, C_{q-1})$ with $|C_k| = w_1$ for $1 \leq k \leq q-1$ and $|C_0| = w_2$.

Observe that

$$\begin{aligned}
i_{q-1} &= n - i_0 - \sum_{k=1}^{q-2} i_k \\
&\leq n - w_2 - (q-2)w_1 \\
&\leq w_1 + \frac{w_2}{w_1} - 1 \\
&\leq w_1 + (w_2 - w_1) \\
&= w_2.
\end{aligned}$$

Equality holds if and only if $w_1 = 1$, $i_k = 1$ for $k = 1, \ldots, q-2$, $i_0 = w_2$ and $n = w_2 + (q-1)w_1 + \frac{w_2}{w_1} - 1 = 2w_2 + q - 2$. We consider the following two cases.

(i) $i_{q-1} = w_2$: Let $w = w_2$. We only need to consider the case $q \geq 3$ since $q = 2$ is covered by Theorem 2.2.3. The number of column set q-tuples separated by r is exactly $2(q-2)! \cdot T^{(q-1)}_{1,w,n}(1, \ldots, 1, w)$, which is less than the number of column set q-tuples separated by a row of weight $(w_1, \ldots, w_1) = (1, \ldots, 1)$ by Lemma 2.3.1 and Lemma 2.3.3.

(ii) $i_{q-1} < w_2$: By Lemma 2.3.1, the number of column set q-tuples separated by r is

$$Z = (q-1)! \cdot T^{(q-1)}_{w_1,w_2,n}(i_1, \ldots, i_{q-1}).$$

The number of column set q-tuples separated by a row of weight $(w_1, \ldots, w_1)$ is $(q-1)! \cdot T^{(q-1)}_{w_1,w_2,n}(w_1, \ldots, w_1)$, which by Lemma 2.3.2 is greater than Z unless $i_k = w_1$ for $k = 1, \ldots, q-1$.

In either case, the number of column set q-tuples separated by r is maximal only when the row is of weight $(w_1, \ldots, w_1)$. The total number of column set q-tuples that need to be separated is

$$T = \binom{n}{w_1}\binom{n-w_1}{w_1}\cdots\binom{n-(q-1)w_1}{w_2}.$$

Thus

$$\begin{aligned} N &\geq \frac{T}{(q-1)! \cdot T^{(q-1)}_{w_1,w_2,n}(w_1, \ldots, w_1)} \\ &= \frac{1}{(q-1)!}\binom{n}{w_1}\binom{n-w_1}{w_1}\cdots\binom{n-(q-2)w_1}{w_1}. \end{aligned}$$

□

Similar to Theorem 2.2.3, the bound in Theorem 2.3.4 is also tight due to the following construction.

Construction 2. Fix positive integers n, q, w_1, w_2 with $w_1 < w_2$ and $w_2 + (q-1)w_1 \leq n$. Let

$$\mathcal{S} = \{(S_1, \ldots, S_{q-1}) : S_i \subseteq \{1, \ldots, n\} \text{ with } |S_i| = w_1 \text{ for all } i \text{ and } S_i \cap S_j = \emptyset \text{ if } i \neq j\},$$

and let

$$\mathcal{T} = \{(S_1, \ldots, S_{q-1}) \in \mathcal{S} : s_1 < s_2 < \ldots < s_{q-1} \text{ where } s_i \text{ is the smallest element of } S_i\}.$$

Now for $(S_1, \ldots, S_{q-1}) \in \mathcal{T}$, let $r_{(S_1,\ldots,S_{q-1})}$ be the vector

$$r_{(S_1,\ldots,S_{q-1})}(i) = \begin{cases} j & \text{if } i \in S_j \\ 0 & \text{otherwise} \end{cases}.$$

The representation matrix for the $\mathsf{SHF}(N; n, q, \{w_1^{q-1}, w_2\})$ contains all rows $r_{(S_1,\ldots,S_{q-1})}$ for every $(S_1, \ldots, S_{q-1}) \in \mathcal{T}$.

Example 2.3.1. Let $q = 4$, $w_1 = 2$ and $w_2 = 4$. Suppose there exists an $\mathsf{SHF}(N; 11, 4, \{2, 2, 2, 4\})$. Then

$$N \geq \frac{1}{6}\binom{11}{2}\binom{9}{2}\binom{7}{2} = \frac{1}{12}(55)(36)(21) = 6929$$

by Theorem 2.3.4. In order words, for $N \leq 6928$, we have that $n \leq 10$.
The bound for n in Theorems 2.1.4, 2.1.5 and 2.1.6 all contain a term of the form $q^{\lceil \frac{N}{u-1} \rceil}$ where $u = (q-1)w_1 + w_2 = 10$. For the value of $N = 6928$, this term is 4^{770}, which is vastly larger than the bound from Theorem 2.3.4.

Table 2.1 lists various parameter choices for q, w_1, w_2 and compares the bound in Theorem 2.3.4 to some known general bounds. The symbol Ω means the computed bound is above the Java `double` maximum value $(2 - 2^{-52}) \cdot 2^{1023}$.

Theorem 2.3.4 is particularly useful for studying the combinatorial objects known as *strong separating hash families* (denoted *SSHF*), introduced by Sarkar et al. [22], which are equivalent to $\mathsf{SHF}(N; n, q, \{1^{t_1}, t_2\})$. We can give a strong bound for the code length of SSHFs as a corollary of Theorem 2.3.4.

Corollary 2.3.5. *Let n, t_1, t_2 be positive integers with $t_1 \geq q - 1$ and $t_1 + t_2 \leq n \leq 2(t_1 + t_2) - q$. Suppose there exists an $\mathsf{SHF}(N; n, q, \{1^{t_1}, t_2\})$. Then*

$$N \geq \binom{n}{q-1}.$$

Proof. By Theorem 1.1.2, an $\mathsf{SHF}(N; n, q, \{1^{t_1}, t_2\})$ is also an $\mathsf{SHF}(N; n, q, \{1^{q-1}, t_2 + t_1 - q + 1\})$. Applying Theorem 2.3.4 gives that for $t_1 + t_2 \leq n \leq 2(t_1 + t_2) - q$, we have

$$N \geq \frac{1}{(q-1)!} n(n-1)\cdots(n-q+2),$$

as desired. □

q	w_1	w_2	$N \leq$	implies $n \leq$			
				Theorem 2.3.4	Theorem 2.1.4	Theorem 2.1.5	Theorem 2.1.6
3	1	2	9	4	243	243	213
3	1	3	20	6	2916	2916	2824
3	1	4	35	8	32805	32805	32526
3	1	5	54	10	354294	354294	353454
3	1	6	77	12	3720087	3720087	3717563
3	2	3	104	6	3.09×10^{9}	2.32×10^{9}	2.32×10^{9}
3	2	4	377	8	5.81×10^{26}	4.07×10^{26}	4.07×10^{26}
3	2	5	629	9	5.91×10^{38}	3.94×10^{38}	3.94×10^{38}
3	2	6	1484	11	7.43×10^{79}	4.77×10^{79}	4.77×10^{79}
3	3	4	2099	9	6.64×10^{112}	3.98×10^{112}	3.98×10^{112}
3	3	5	4619	10	4.84×10^{221}	2.69×10^{221}	2.69×10^{221}
3	3	6	17159	12	Ω	Ω	Ω
4	1	2	19	5	4096	4096	3987
4	1	3	54	7	2.09×10^{7}	2.09×10^{7}	2.09×10^{7}
4	1	4	118	9	6.59×10^{12}	6.59×10^{12}	6.59×10^{12}
4	1	5	219	11	1.29×10^{20}	1.29×10^{20}	1.29×10^{20}
4	1	6	362	13	3.96×10^{28}	3.96×10^{28}	3.96×10^{28}
4	2	3	1259	8	1.33×10^{96}	1.06×10^{96}	1.06×10^{96}
4	2	4	6929	10	Ω	Ω	Ω
4	2	5	13859	11	Ω	Ω	Ω
4	2	6	45044	13	Ω	Ω	Ω
4	3	4	200199	12	Ω	Ω	Ω
4	3	5	560559	13	Ω	Ω	Ω
4	3	6	3203199	15	Ω	Ω	Ω
5	1	2	33	6	390625	390625	389658
5	1	3	125	8	2.86×10^{15}	2.86×10^{15}	2.86×10^{15}
5	1	4	329	10	2.48×10^{34}	2.48×10^{34}	2.48×10^{34}
5	1	5	714	12	6.46×10^{63}	6.46×10^{63}	6.46×10^{63}
5	1	6	1364	14	1.57×10^{107}	1.57×10^{107}	1.57×10^{107}
5	2	3	17324	10	Ω	Ω	Ω
5	2	4	135134	12	Ω	Ω	Ω
5	2	5	315314	13	Ω	Ω	Ω
5	2	6	1351349	15	Ω	Ω	Ω
5	3	4	28027999	15	Ω	Ω	Ω
5	3	5	95295198	16	Ω	Ω	Ω
5	3	6	775975199	18	Ω	Ω	Ω

Table 2.1: Comparison of Bounds for $\mathsf{SHF}(N; n, q, \{w_1^{q-1}, w_2\})$

Note that if we set $q = 2$, $t_1 = 1$ and $t_2 = w$ for some positive integer w, we obtain part of the result of Theorem 2.2.3.

Example 2.3.2. Let $q = 3, t_1 = 4$ and $t_2 = 3$. Suppose there exists an $\mathsf{SHF}(N; 11, 3, \{1, 1, 1, 1, 3\})$ (Corollary 2.3.5 applies to $n = 7, 8, 9, 10$ as well). Then $N \geq \binom{11}{2} = 55$ from Corollary 2.3.5. In order words, for $N \leq 54$, we have that $n \leq 10$.
Compare this with the known results: Theorem 2.1.4 and Theorem 2.1.5 both give the bound $n \leq 6(3^9) = 118098$ for $N = 54$; Theorem 2.1.6 gives the bound

$$n \leq 6(3^9) + 2 - 2\sqrt{3 \cdot (3^9) + 1} < 118023$$

for $N = 54$.

Chapter 3

Combinatorial Designs

A large portion of this chapter will appear in [16], which is accepted for publication.

The subject of combinatorial designs has a rich history. The basic problem is to construct a family of subsets of a finite set such that it achieves certain balance properties. One of the most classic examples of a combinatorial design is the Fano plane (see Figure 3.1). We will restrict our attention to a class of designs known as *balanced incomplete block designs* (BIBDs). Many definitions and basic results from this chapter are from [25].

The main result of this chapter gives a necessary and sufficient condition for when the incidence matrix of an SBIBD is also the representation matrix of a frameproof code. We show that certain families of SBIBDs satisfy or fail this condition.

3.1 Preliminaries

In this section we recall several known results and definitions from design theory. The common theme of study in combinatorial designs is the object called *design*, defined as follows.

Definition 3.1.1. Let X be a finite set of elements (called *points*), and let $\mathcal{A}$ be a collection of non-empty subsets of X (called *blocks*). The pair $(X, \mathcal{A})$ is called a design.

Definition 3.1.2. Let v, k, λ be positive integers with $v > k \geq 2$. A (v, k, λ)-*balanced incomplete block design* (denoted (v, k, λ)-BIBD) is a design $(X, \mathcal{A})$ satisfying

(i) $|X| = v$,

(ii) $|A| = k$ for every $A \in \mathcal{A}$, and

(iii) for every $x, y \in X$, there are exactly λ blocks $A \in \mathcal{A}$ such that $x, y \in A$.

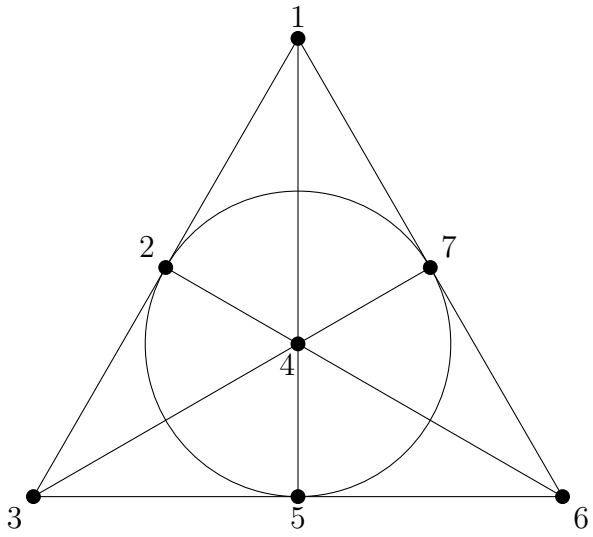

Figure 3.1: Fano plane

Example 3.1.1. A $(7, 3, 1)$-BIBD $(X, \mathcal{A})$ where

$$X = \{1, 2, 3, 4, 5, 6, 7\} \text{ and}$$

$$\mathcal{A} = \{123, 145, 167, 246, 257, 347, 356\}.$$

The blocks in $\mathcal{A}$ are abbreviated; i.e., 123 denotes the subset $\{1, 2, 3\}$. The Fano plane represents this BIBD in the following sense: the lines (and the circle) are the blocks; the points of intersection are points of X; every two points is connected by exactly one line.

There are two missing parameters from the notation (v, k, λ): the number of blocks b and the number of blocks r each point appears in. The notation (v, b, r, k, λ)-BIBD is sometimes used to record all these values. However, it turns out that b and r can be determined completely as follows.

Theorem 3.1.1. *Let* $(X, \mathcal{A})$ *be a* (v, b, r, k, λ)-BIBD. *Then*

(i) $r = \frac{\lambda(v-1)}{k-1}$

(ii) $b = \frac{vr}{k} = \frac{\lambda(v^2-v)}{k^2-k}$

When studying BIBDs, the usual convention is to consider only ones with $k \leq \frac{v}{2}$, due to the following theorem.

Theorem 3.1.2. ***(Block Complementation)*** *Let* $(X, \mathcal{A})$ *be a* (v, b, r, k, λ)-BIBD *with* $k \leq v - 2$. *Let* $\mathcal{B} = \{X \setminus A : A \in \mathcal{A}\}$. *Then* $(X, \mathcal{B})$ *is a* $v, b, b - r, v - k, b - 2r + \lambda)$-BIBD.

One particular special case of BIBDs is the *symmetric BIBD*.

Definition 3.1.3. A (v, b, r, k, λ)-BIBD is called a *symmetric BIBD* (or *SBIBD*) if $v = b$.

An equivalent definition is to require that $r = k$, and the two definitions are equivalent via a direct application of Theorem 3.1.1. The $(7, 3, 1)$-BIBD is an example of a symmetric BIBD.

Definition 3.1.4. Let $(X, \mathcal{A})$ be a (v, b, r, k, λ)-BIBD with $X = \{x_1, \ldots, x_v\}$ and $\mathcal{A} = \{A_1, \ldots, A_b\}$. Define the matrix M by

$$\mathsf{M}(i, j) = \begin{cases} 1 & \text{if } x_i \in A_j \\ 0 & \text{otherwise} \end{cases}.$$

M is called the *point-block incidence matrix* of $(X, \mathcal{A})$, and M^T is called the *block-point incidence matrix* of $(X, \mathcal{A})$.

There are several interesting implications of the symmetric requirement. It is clear that the point-block incidence matrix of an SBIBD is a square matrix, but is not necessarily symmetric as the name suggests. However, there is a notion of symmetry in the term.

Theorem 3.1.3. *Let* M *be the point-block incidence matrix of a* (v, k, λ)-SBIBD. *Then* M *is the block-point incidence matrix of a* (v, k, λ)-SBIBD.

This crucial property of SBIBDs will be exploited in Section 3.4 to prove new results. More specifically, we will see that the point-block incidence matrix of certain SBIBDs is the representation matrix of frameproof codes.

3.2 Constructions

There are many known constructions for BIBDs. Most constructions utilize results from finite field theory and are algebraic in nature. We will give several well-known constructions below.

Construction 3. Let q be a prime power, and let $\mathbb{F}_q$ be the finite field of order q. For a fixed integer $d \geq 2$, let $V = \mathbb{F}_q^{d+1}$ be the $(d+1)$-dimensional vector space over $\mathbb{F}_q$. Let $\vec{0}$ denote the zero element in V.

Let $\mathcal{V}_1$ be the set of all one-dimensional subspaces of V, and let $\mathcal{V}_d$ be the set of all d-dimensional subspaces of V. For any $W \in \mathcal{V}_d$, let

$$A_W = \{U \in \mathcal{V}_1 : U \text{ is a subspace of } W\}$$

and let $\mathcal{A} = \{A_W : W \in \mathcal{V}_d\}$.

The *d-dimensional projective geometry of order q*, denoted $\mathsf{PG}_d(q)$, is the $(\frac{q^{d+1}-1}{q-1}, \frac{q^d-1}{q-1}, \frac{q^{d-1}-1}{q-1})$-$\mathsf{SBIBD}$ $(\mathcal{V}_1, \mathcal{A})$, which can be shown as follows:

(i) For every $U \in \mathcal{V}_1$ we have $|U| = q$ and $\vec{0} \in U$. Also, if $U_1, U_2 \in \mathcal{V}_1$ with $U_1 \neq U_2$ then $U_1 \cap U_2 = \{\vec{0}\}$. Thus the set $V \setminus \{\vec{0}\}$ is partitioned by the collection

$$\{U \setminus \{\vec{0}\} : U \in \mathcal{V}_1\},$$

which gives

$$|\mathcal{V}_1| = \frac{|V|-1}{|U|-1} = \frac{q^{d+1}-1}{q-1}.$$

(ii) For every $W \in \mathcal{V}_d$ we have $|W| = q^d$ and $\vec{0} \in W$. If U_1, U_2 are subspaces of W with $U_1 \neq U_2$ then $U_1 \cap U_2 = \{\vec{0}\}$. Thus the set $W \setminus \{\vec{0}\}$ is partitioned by the collection

$$\{U \setminus \{\vec{0}\} : U \text{ is a subspace of } W\} = \{U \setminus \{\vec{0}\} : U \in A_W\},$$

which gives

$$|A_W| = \frac{|W|-1}{|U|-1} = \frac{q^d-1}{q-1}.$$

(iii) If $U_1, U_2 \in \mathcal{V}_1$ with $U_1 \neq U_2$ then there are $\frac{q^{d-1}-1}{q-1}$ d-dimensional subspace W of V that contains both U_1 and U_2 as subspaces, i.e. there exist $\frac{q^{d-1}-1}{q-1}$ block in $\mathcal{A}$ containing both U_1 and U_2.

The special case of $d = 2$ in Construction 3 has a special name, called the *projective plane of order* q, which is a $(q^2 + q + 1, q + 1, 1)$-BIBD. The $(7, 3, 1)$-SBIBD in Example 3.1.1 can be viewed as a projective plane of order 2. For any positive integer n, an $(n^2 + n + 1, n + 1, 1)$-BIBD is called a projective plane of order n. One of the most important open problems in design theory is the question of whether there exists a projective plane of non-prime power order. No such design has been found to date, and there are infinitely many n for which an impossibility result has not been given.

One important method for constructing SBIBDs is the *difference set method*.

Definition 3.2.1. Let $(G, +)$ be an additive finite group of order v with the additive identity element 0. Let k, λ be positive integers such that $v > k \geq 2$. A (v, k, λ)-*difference set* in $(G, +)$ is a subset $D \subseteq G$ with $|D| = k$ such that every element in $G \setminus \{0\}$ can be written as the difference of two distinct elements in D in λ ways.

Example 3.2.1. $G = \mathbb{Z}_{11}$ and $D = \{1, 3, 4, 5, 9\}$. It can be checked that each non-zero element in $\mathbb{Z}_{11}$ can be written as $x - y$ for $x, y \in D$ in two ways, e.g. $3 = 4 - 1 = 1 - 9$, $5 = 9 - 4 = 3 - 9$.

We will present a class of difference sets known as *Paley difference sets*.

Construction 4. Let q a prime power and let $\mathbb{F}_q$ be the finite field of order q. Let

$$\mathsf{QR}(q) = \{x^2 : x \in \mathbb{F}_q \setminus \{0\}\},$$

called the *quadratic residue* of $\mathbb{F}_q$. Suppose $q \equiv 3 \mod 4$. It can be shown that $\mathsf{QR}(q)$ is a $(q, \frac{q-1}{2}, \frac{q-3}{4})$-difference set, called the *Paley difference set* (or *quadratic residue difference set*).

Example 3.2.1 is a Paley difference set where $D = \mathsf{QR}(11) = \{1, 3, 4, 5, 9\}$.

Definition 3.2.2. Let D be a (v, k, λ)-difference set in $(G, +)$. The *development of* D, denoted $\mathsf{Dev}(D)$, is the collection

$$\mathsf{Dev}(D) = \{D + g : g \in G\}$$

where $D + g = \{x + g : x \in D\}$ is the usual right coset of D.

The notation (v, k, λ) for difference sets is not a coincidence; it corresponds to a (v, k, λ)-SBIBD in the following sense.

Theorem 3.2.1. *Let D be a (v,k,λ)-difference set of an abelian group $(G,+)$. Then $(G, \mathsf{Dev}(D))$ is a (v,k,λ)-*SBIBD.

Example 3.2.2. Let D be the $(11,5,2)$-difference set from Example 3.2.1. The point-block incidence matrix of the associated design is

$$\mathsf{M} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \end{pmatrix}$$

Both the projective plane construction and the difference set construction yield an SBIBD. We now present a method for obtaining a non-symmetric BIBD from a given symmetric BIBD.

Definition 3.2.3. Let $(X,\mathcal{A})$ be a (v,k,λ)-SBIBD. Fix $A_0 \in \mathcal{A}$. Define

$$\mathsf{Der}(X,\mathcal{A},A_0) = (A_0, \{A \cap A_0 : A \in \mathcal{A}, A \neq A_0\}),$$

called a *derived BIBD*, and

$$\mathsf{Res}(X,\mathcal{A},A_0) = (X \setminus A_0, \{A \setminus A_0 : A \in \mathcal{A}, A \neq A_0\}),$$

called a *residual BIBD*.

Theorem 3.2.2. *Let $(X,\mathcal{A})$ be a (v,k,λ)-*SBIBD*. Fix $A_0 \in \mathcal{A}$. Then $\mathsf{Der}(X,\mathcal{A},A_0)$ is a $(k, v-1, k-1, \lambda, \lambda-1)$-*BIBD *and $\mathsf{Res}(X,\mathcal{A},A_0)$ is a $(v-k, v-1, k, k-\lambda, \lambda)$-*BIBD.

Example 3.2.3. Let $(X,\mathcal{A})$ be the $(11,5,2)$-SBIBD from Example 3.2.2. Let A_0 be the first block of $\mathcal{A}$, represented by the first column of the point-block incidence matrix. The incidence matrix of the $(5,10,4,2,1)$-derived BIBD is

$$\mathsf{M}_{\mathrm{der}} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix},$$

formed by removing the first column and all rows containing a 1 in the first column. The incidence matrix of the $(6, 10, 5, 3, 2)$-residual BIBD is

$$\mathsf{M}_{\text{res}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \end{pmatrix},$$

formed by removing the first column and all rows containing a 0 in the first column.

The residual designs of a projective plane of order q has a special name, called an *affine plane*, which is a $(q^2, q^2 + q, q + 1, q, 1)$-BIBD.

3.3 Known Connections to Binary Frameproof Codes

The paper [27] gave several constructions of binary frameproof codes from t-designs and t-packing designs. We first define the two combinatorial structures.

Definition 3.3.1. Let t, v, k, λ be positive integers with $v > k \geq t$. A t-(v, k, λ) *design* is a design $(X, \mathcal{A})$ satisfying

(i) $|X| = v$,

(ii) $|A| = k$ for every $A \in \mathcal{A}$, and

(iii) for every subset $T \subseteq X$ with $|T| = t$, there are exactly λ blocks $A \in \mathcal{A}$ such that $T \subseteq A$.

Note that this definition generalizes Definition 3.1.2 in the sense that a (v, k, λ)-BIBD is a 2-(v, k, λ) design, and the latter notation is sometimes used instead.

Definition 3.3.2. Let t, v, k, λ be positive integers with $v \geq k \geq t$. A t-(v, k, λ) *packing design* is a design $(X, \mathcal{A})$ satisfying

(i) $|X| = v$,

(ii) $|A| = k$ for every $A \in \mathcal{A}$, and

(iii) for every subset $T \subseteq X$ with $|T| = t$, there are at most λ blocks $A \in \mathcal{A}$ such that $T \subseteq A$.

An obvious implication of Definition 3.3.2 is that every t-(v, k, λ) design is a t-(v, k, λ) packing design.

The incidence matrices for t-designs and t-packing designs can be defined similar to that of BIBDs. We have the following theorem stating that these two objects give rise to frameproof codes.

Theorem 3.3.1 ([27]). *Let $(X, \mathcal{A})$ be a t-(v, k, λ) packing design. Then the point-block incidence matrix of $(X, \mathcal{A})$ is an $\mathsf{SHF}(v; b, 2, \{1, w\})$ where $b = \binom{v}{t}/\binom{k}{t}$ is the number of blocks and $w = \lfloor \frac{k-1}{t-1} \rfloor$.*

The projective plane and affine planes of prime power order q together with Theorem 3.3.1 gives the following corollary.

Corollary 3.3.2 ([27]). *Let q be a prime power.*

(i) The point-block incidence matrix of the projective plane of order q is the representation matrix of an $\mathsf{SHF}(q^2 + q + 1; q^2 + q + 1, 2, \{1, q\})$.

(ii) The point-block incidence matrix of the affine plane of order q is the representation matrix of an $\mathsf{SHF}(q^2; q^2 + q, 2, \{1, q - 1\})$.

Packing designs can be obtained from orthogonal arrays, defined as follows.

Definition 3.3.3. Let t, v, k, λ be positive integers with $k \geq t \geq 2$. A t-(v, k, λ) *orthogonal array*, denoted t-(v, k, λ)-OA, is a $k \times \lambda q^t$ matrix A with entries from $Q = \{0, \ldots, q - 1\}$ such that within any t rows of A, the set Q^t is repeated exactly λ times among the columns of A.

Construction 5. Let q be a prime power and let t be a positive integer with $q > t \geq 2$. Let $P_1, \ldots, P_{q^t}$ be pairwise distinct polynomials of degree at most $t - 1$ over the finite field $\mathbb{F}_q$. Let $\mathbb{F} = \mathbb{F}_q \cup \{\infty\}$, where ∞ is an extra point. Define a $(q + 1) \times q^t$ matrix A by

$$\mathsf{A}(x, j) = \begin{cases} P_j(x) & \text{if } x \in \mathbb{F}_q \\ a_{t-1} & \text{if } x = \infty \text{ and } P_j(x) = \sum_{i=0}^{t-1} a_i x^i \end{cases}.$$

Then A is a t-$(q, q + 1, 1)$-OA.

Construction 5 is equivalent to a class of error-correcting codes called Reed-Solomon codes.

Example 3.3.1. Below is a 2-$(3, 4, 1)$-OA constructed via Construction 5.

$$\mathsf{A} = \begin{pmatrix} 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 \\ 0 & 1 & 2 & 2 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \end{pmatrix}.$$

Theorem 3.3.3 ([27]). *Let* A *be a* t-$(v, k, 1)$-OA. *Define*

$$X = \{(x, y) : 0 \leq x \leq k-1, 0 \leq y \leq v-1\},$$

and for every column $c = (y_0, y_1, \ldots, y_{k-1})^T$ *in* A, *define*

$$B_c = \{(0, y_0), (1, y_1), \ldots, (k-1, y_{k-1})\}.$$

Let $\mathcal{B} = \{B_c : c \text{ is a column of } \mathsf{A}\}$. *Then* $(X, \mathcal{B})$ *is a* t-$(vk, k, 1)$ *packing design containing* v^t *blocks.*

Combining Construction 5 and Theorem 3.3.3 we get the following corollary.

Corollary 3.3.4 ([27]). *Let* q *be a prime power. There exists a* t-$(q^2+q+1, q+1, 1)$ *packing design with* q^t *blocks.*

We get another construction for frameproof codes by Theorem 3.3.1 and Corollary 3.3.4.

Corollary 3.3.5 ([27]). *Let* q *be a prime power. There exists an* SHF$(q^2+q; q^t, 2, \{1, w\})$ *where* $w = \lfloor \frac{q}{t-1} \rfloor$ *for any positive integer* t *such that* $q > t \geq 2$.

3.4 New Result

In this section, we give a characterization of SHF$(v; v, 2, \{1, 3\})$ for all symmetric (v, k, λ)-BIBDs. For simplicity, we sometimes use the notation $\{1, w\}$-SHF when the size parameters are implied.

Theorem 3.4.1. *Let* $(X, \mathcal{B})$ *be a symmetric* (v, k, λ)-BIBD *and let* A *be its block-point incidence matrix. If* $k \geq 3\lambda + 1$ *or if* $k - \lambda$ *is odd then* A *is an* SHF$(v; v, 2, \{1, 3\})$.

Proof. Suppose that A is not an $\mathsf{SHF}(v; v, 2, \{1, 3\})$, then there exists some column set pair $(\{\mathsf{x}\}, \{\mathsf{u}, \mathsf{v}, \mathsf{w}\})$ that cannot be separated. For each $Z \subseteq \{\mathsf{u}, \mathsf{v}, \mathsf{w}\}$, partition $\mathcal{B}$ into subsets $\mathcal{A}_Z$ where $\mathcal{A}_Z = \{B \in \mathcal{B} : B \cap \{\mathsf{u}, \mathsf{v}, \mathsf{w}\} = Z\}$, and let $a_Z = |\mathcal{A}_Z|$. We obtain the following set of equations from $(X, \mathcal{B})$ being a symmetric (v, k, λ)-BIBD:

$$a_\emptyset + a_\mathsf{u} + a_\mathsf{v} + a_\mathsf{w} + a_\mathsf{uv} + a_\mathsf{vw} + a_\mathsf{uw} + a_\mathsf{uvw} = v \tag{3.1}$$
$$a_\mathsf{u} + a_\mathsf{uv} + a_\mathsf{uw} + a_\mathsf{uvw} = k \tag{3.2}$$
$$a_\mathsf{v} + a_\mathsf{uv} + a_\mathsf{vw} + a_\mathsf{uvw} = k \tag{3.3}$$
$$a_\mathsf{w} + a_\mathsf{uw} + a_\mathsf{vw} + a_\mathsf{uvw} = k \tag{3.4}$$
$$a_\mathsf{uv} + a_\mathsf{uvw} = \lambda \tag{3.5}$$
$$a_\mathsf{vw} + a_\mathsf{uvw} = \lambda \tag{3.6}$$
$$a_\mathsf{uw} + a_\mathsf{uvw} = \lambda \tag{3.7}$$

Letting $\alpha = a_\mathsf{uvw}$, we get that

$$\begin{aligned} a_\mathsf{uv} = a_\mathsf{vw} = a_\mathsf{uw} &= \lambda - \alpha \\ a_\mathsf{u} = a_\mathsf{v} = a_\mathsf{w} &= k - 2(\lambda - \alpha) - \alpha \\ &= k + \alpha - 2\lambda. \end{aligned}$$

Next, define $\mathcal{B}_Z = \{B \in \mathcal{A}_Z : \mathsf{x} \in B\}$ and let $b_Z = |\mathcal{B}_Z|$. We obtain another set of equations:

$$b_\emptyset + b_\mathsf{u} + b_\mathsf{v} + b_\mathsf{w} + b_\mathsf{uv} + b_\mathsf{vw} + b_\mathsf{uw} + b_\mathsf{uvw} = k \tag{3.8}$$
$$b_\mathsf{u} + b_\mathsf{uv} + b_\mathsf{uw} + b_\mathsf{uvw} = \lambda \tag{3.9}$$
$$b_\mathsf{v} + b_\mathsf{uv} + b_\mathsf{vw} + b_\mathsf{uvw} = \lambda \tag{3.10}$$
$$b_\mathsf{w} + b_\mathsf{uw} + b_\mathsf{vw} + b_\mathsf{uvw} = \lambda \tag{3.11}$$

Note that for every Z, we get $0 \leq b_Z \leq a_Z$. It is clear that the column set pair $(\{\mathsf{x}\}, \{\mathsf{u}, \mathsf{v}, \mathsf{w}\})$ cannot be separated if and only if $b_\emptyset = 0$ and $b_\mathsf{uvw} = \alpha$. Thus equations (3.8) – (3.11) simplify to

$$b_\mathsf{u} + b_\mathsf{v} + b_\mathsf{w} + b_\mathsf{uv} + b_\mathsf{vw} + b_\mathsf{uw} = k - \alpha \tag{3.12}$$
$$b_\mathsf{u} + b_\mathsf{v} + b_\mathsf{w} + 2(b_\mathsf{uv} + b_\mathsf{vw} + b_\mathsf{uw}) = 3(\lambda - \alpha) \tag{3.13}$$

Subtracting (3.12) from (3.13) gives

$$b_\mathsf{uv} + b_\mathsf{vw} + b_\mathsf{uw} = 3\lambda - k - 2\alpha. \tag{3.14}$$

Since $b_{\mathsf{uv}} + b_{\mathsf{vw}} + b_{\mathsf{uw}} \geq 0$, (3.14) implies that

$$0 \leq 3\lambda - k - 2\alpha. \tag{3.15}$$

Now, since $\alpha \geq 0$, we see from (3.15) that $k \leq 3\lambda$. Therefore A is an $\mathsf{SHF}(v; v, 2, \{1, 3\})$ if $k \geq 3\lambda + 1$.

Next, we multiply (3.12) by 2 and subtract (3.13), giving

$$b_{\mathsf{u}} + b_{\mathsf{v}} + b_{\mathsf{w}} = \alpha + 2k - 3\lambda. \tag{3.16}$$

Then we have

$$3(k + \alpha - 2\lambda) = a_{\mathsf{u}} + a_{\mathsf{v}} + a_{\mathsf{w}} \geq b_{\mathsf{u}} + b_{\mathsf{v}} + b_{\mathsf{w}} = \alpha + 2k - 3\lambda. \tag{3.17}$$

Therefore, from (3.17), we have

$$3\lambda - k - 2\alpha \leq 0. \tag{3.18}$$

Now, (3.15) and (3.18) together show that $3\lambda - k = 2\alpha$. This implies that $3\lambda - k$ is even, and therefore $k - \lambda$ is also even. Therefore A is an $\mathsf{SHF}(v; v, 2, \{1, 3\})$ if $k - \lambda$ is odd. □

Corollary 3.4.2. *Let $(X, \mathcal{B})$ be a symmetric (v, k, λ)-*BIBD *and let A be its block-point incidence matrix. If $k \leq 3\lambda$ and $k - \lambda$ is even then A is an $\mathsf{SHF}(v; v, 2, \{1, 3\})$ if and only if there does not exist four points $\mathsf{u}, \mathsf{v}, \mathsf{w}, \mathsf{x}$ such that*

1. *$\alpha = \frac{3\lambda - k}{2}$ blocks contain all four points $\mathsf{u}, \mathsf{v}, \mathsf{w}, \mathsf{x}$,*
2. *no block contains exactly one or three points from $\{\mathsf{u}, \mathsf{v}, \mathsf{w}, \mathsf{x}\}$, and*
3. *for any subset of two points from $\{\mathsf{u}, \mathsf{v}, \mathsf{w}, \mathsf{x}\}$, there are exactly $\lambda - \alpha$ blocks that contain these two points.*

Proof. It is clear that A is not a $\mathsf{SHF}(v; v, 2, \{1, 3\})$ if the specified four-point substructure exists. So we just need to prove the converse, namely, that the four-point substructure exists if A is not a $\mathsf{SHF}(v; v, 2, \{1, 3\})$. We use the same notation as in the proof of Theorem 3.4.1. The proof of that theorem established that $\alpha = (3\lambda - k)/2$.

For each $T \subseteq \{\mathsf{u}, \mathsf{v}, \mathsf{w}, \mathsf{x}\}$, we will compute c_T, which denotes the number of blocks B such than $B \cap \{\mathsf{u}, \mathsf{v}, \mathsf{w}, \mathsf{x}\} = T$. First, we note two relevant facts:

- The inequality in (3.17) must be an inequality, so $b_{\mathsf{u}} = a_{\mathsf{u}}$, $b_{\mathsf{v}} = a_{\mathsf{v}}$ and $b_{\mathsf{w}} = a_{\mathsf{w}}$. Now $a_{\mathsf{u}} = a_{\mathsf{v}} = a_{\mathsf{w}} = k + \alpha - 2\lambda = \lambda - \alpha$, so we obtain $b_{\mathsf{u}} = b_{\mathsf{v}} = b_{\mathsf{w}} = \lambda - \alpha$.

T	c_T	T	c_T
$\{\mathsf{x}\}$	$b_\emptyset = 0$	$\{\mathsf{u}\}$	$a_\mathsf{u} - b_\mathsf{u} = 0$
$\{\mathsf{v}\}$	$a_\mathsf{v} - b_\mathsf{v} = 0$	$\{\mathsf{w}\}$	$a_\mathsf{v} - b_\mathsf{v} = 0$
$\{\mathsf{u},\mathsf{x}\}$	$b_\mathsf{u} = \lambda - \alpha$	$\{\mathsf{u},\mathsf{v}\}$	$a_\mathsf{uv} - b_\mathsf{uv} = a_\mathsf{uv} = \lambda - \alpha$
$\{\mathsf{v},\mathsf{x}\}$	$b_\mathsf{v} = \lambda - \alpha$	$\{\mathsf{u},\mathsf{w}\}$	$a_\mathsf{uw} - b_\mathsf{uw} = a_\mathsf{uw} = \lambda - \alpha$
$\{\mathsf{w},\mathsf{x}\}$	$b_\mathsf{w} = \lambda - \alpha$	$\{\mathsf{v},\mathsf{w}\}$	$a_\mathsf{vw} - b_\mathsf{vw} = a_\mathsf{vw} = \lambda - \alpha$
$\{\mathsf{u},\mathsf{v},\mathsf{x}\}$	$b_\mathsf{uv} = 0$	$\{\mathsf{u},\mathsf{w},\mathsf{x}\}$	$b_\mathsf{uw} = 0$
$\{\mathsf{v},\mathsf{w},\mathsf{x}\}$	$b_\mathsf{vw} = 0$	$\{\mathsf{u},\mathsf{v},\mathsf{w}\}$	$a_\mathsf{uvw} - b_\mathsf{uvw} = \alpha - \alpha = 0$
$\{\mathsf{u},\mathsf{v},\mathsf{w},\mathsf{x}\}$	$b_\mathsf{uvw} = \alpha$		

Table 3.1: Block intersections with $\{\mathsf{u},\mathsf{v},\mathsf{w},\mathsf{x}\}$

- From (3.14), we see that $b_\mathsf{uv} + b_\mathsf{vw} + b_\mathsf{uw} = 0$, so $b_\mathsf{uv} = b_\mathsf{vw} = b_\mathsf{uw} = 0$.

It is now straightforward to compute the values c_T using these facts. This is done in Table 3.1. □

The following is an immediate corollary of Theorem 3.4.1. This result is in fact equivalent to a result of Kimura [18, Proposition 2.1].

Corollary 3.4.3. *The incidence matrix of a* $(4n-1, 2n-1, n-1)$-SBIBD *is an* SHF$(4n-1; 4n-1, 2, \{1, 3\})$ *if* $n > 1$ *is odd.*

3.4.1 Hadamard Designs

Theorem 3.4.1 has some interesting connections to Hadamard matrices. We first give some background information.

Definition 3.4.1. Let H be an $n \times n$ matrix with entries from $\{\pm 1\}$. We say that H is a *Hadamard matrix of order* n if $\mathsf{H}\mathsf{H}^T = nI_n$ where I_n is the $n \times n$ identity matrix.

If a Hadamard matrix has a first row and first column consisting entirely of entries equal to 1, then we say that the matrix is *standardized*. Any Hadamard matrix can be transformed into a standardized Hadamard matrix by multiplying certain rows and columns by -1.

Example 3.4.1. The following is a standardized Hadamard matrix of order 4:

$$\mathsf{H} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Clearly both (1) and (-1) are Hadamard matrices of order 1. It is not difficult to find a Hadamard matrix of order 2. The following well-known theorem says the order n of a Hadamard matrix must be divisible by 4 if $n > 2$.

Theorem 3.4.4. *Let* H *be a Hadamard matrix of order* n*. If* $n > 2$ *then* $n \equiv 0 \mod 4$.

An important open problem regarding Hadamard matrices is to determine whether or not a Hadamard matrix of order $4m$ exists for every positive integer m. It is known that if n is a power of 2 then a Hadamard matrix of order n exists. The smallest n such that a Hadamard matrix of order $n = 4m$ is not known to exist is 668.

One of the reasons why Hadamard matrices are of great importance to design theory is the following equivalence theorem.

Theorem 3.4.5. *Let* $m > 1$ *be an integer. The following are true:*

(i) Let H *be a standardized Hadamard matrix of order* $4m$*. Let* M *be the* $(4m-1)\times(4m-1)$ *binary matrix formed by removing the first row and column of* H *and replacing all entries of* -1 *with 0. Then* M *is the point-block incidence matrix of a* $(4m-1, 2m-1, m-1)$-SBIBD*.*

(ii) Let M *be the point-block incidence matrix of a* $(4m-1, 2m-1, m-1)$-SBIBD*. Let* H *be the* $4m \times 4m$ *matrix with entries from* $\{\pm 1\}$ *formed by replacing all entries of* 0 *in* M *with* -1 *and the appending a row and column of 1s as the first row and column. Then* H *is a standardized Hadamard matrix of order* $4m$*.*

As a result, we have the following definition for SBIBDs satisfying part (ii) of Theorem 3.4.5.

Definition 3.4.2. A $(4m-1, 2m-1, m-1)$-SBIBD is also called a *Hadamard design.*

There is a useful classification of Hadamard matrices in terms of substructures involving four columns; see, for example, [17]. The notion of a *type* of a Hadamard matrix is defined in [17] as follows.

Definition 3.4.3. Let H be a Hadamard matrix of order $4n$. For any non-negative integer m, let j_m denote the all 1's column vector of length m. By permuting and/or and negating rows and columns, any four columns of H may be transformed uniquely to the following form:

$$\begin{pmatrix} j_a & j_a & j_a & j_a \\ j_b & j_b & j_b & -j_b \\ j_b & j_b & -j_b & j_b \\ j_a & j_a & -j_a & -j_a \\ j_b & -j_b & j_b & j_b \\ j_a & -j_a & j_a & -j_a \\ j_a & -j_a & -j_a & j_a \\ j_b & -j_b & -j_b & -j_b \end{pmatrix}$$

where $a + b = n$ and $0 \leq b \leq \lfloor n/2 \rfloor$. A set of four columns which is transformed to the above form is said to be of *type* b. Any permutation and negation of rows and/or columns leaves the type unchanged. A Hadamard matrix is of *type* b ($0 \leq b \leq \lfloor n/2 \rfloor$) if it has a set of four columns of type b and no set of four columns of type less than b.

Lemma 3.4.6. *Suppose we construct an incidence matrix of a* $(4n-1, 2n-1, n-1)$-SBIBD *from a standardized Hadamard matrix of order* $4n > 4$ *by deleting the first row and column and replacing all occurrences of* -1*'s by* 0*'s. Then this incidence matrix is a* 3*-frameproof code if and only if the Hadamard matrix is not of type 0.*

Proof. First, suppose that the Hadamard is of type 0. Then it is obvious in the incidence matrix of the associated design that the first of the four given columns cannot be separated from the other three given columns.

Conversely, suppose that we have an incidence matrix A (of a $(4n-1, 2n-1, n-1)$-SBIBD) that is not a 3-frameproof code. From Corollary 3.4.3, n must be even for this to occur. By permuting columns of A, we can assume that column 1 cannot be separated from columns 2, 3, and 4. Now we apply Corollary 3.4.2. Looking at the first four columns of A, there must be $n/2 - 1$ occurrences of 1111 and $n/2$ occurrences of each of the other seven patterns containing an even number of 1's. When we convert A to a Hadamard matrix H of order $4n$, we change all 0's to -1's and we add an additional row of 1's. Now we multiply all rows of H that corresponded to patterns 0000, 0011, 0101 and 0110 in A by -1. We then see that these four columns in H are of type 0. □

Kimura's result (Corollary 3.4.3) is in fact a proof that a Hadamard matrix of order congruent to 4 modulo 8 is not of type 0.

Order	4	8	12	16	20	24	28
0	1	1	0	5	0	58	0
Type 1	0	0	1	0	3	1	486
2	0	0	0	0	0	1	1

Table 3.2: Number of inequivalent Hadamard matrices of different types

A classification, according to type, of (inequivalent) Hadamard matrices of small orders is given in [17]. Table 3.2 is from [17]:

We now give a family of Hadamard BIBDs that contain the forbidden substructure from Corollary 3.4.2. Hence, these designs are not $\{1,3\}$-SHFs.

Theorem 3.4.7. *For $n \geq 4$, let H_n be a standardized Hadamard matrix of order n. Let*

$$\mathsf{H} = \begin{pmatrix} \mathsf{H}_n & \mathsf{H}_n \\ \mathsf{H}_n & -\mathsf{H}_n \end{pmatrix}$$

and let A be the $(2n-1)\times(2n-1)$ submatrix of H by removing the first column and first row and replacing all -1's by 0's. Then A is the incidence matrix of a $(2n-1, n-1, \frac{n-2}{2})$-SBIBD which is not an $\mathsf{SHF}(2n-1; 2n-1, 2, \{1,3\})$.

Proof. A is a Hadamard design by construction. Let $n = 4m$, $m \geq 1$. Since H_n is a standard Hadamard matrix of order $4m$, deleting the first column gives a 2-$(2, 4m-1, m)$ orthogonal array. Hence columns 2 and 3 of H_n contain each of the pairs 00, 01, 10, 11 m times. Thus columns $2, 3, 4m+2, 4m+3$ of H contain each of the quadruples 0000, 0101, 1010, 1111 m times in rows $1, \ldots, 4m$ of H. Similarly, columns $2, 3, 4m+2, 4m+3$ of H contain each of the quadruples 0011, 0110, 1001, 1100 m times in rows $4m+1, \ldots, 8m$ of H.

Recall that the first column of H is deleted to form A. Since the first row of H consists of only 1's, we have that columns $1, 2, 4m+1, 4m+2$ of A contain each of the quadruples 0000, 0101, 1010, 0011, 0110, 1001, 1100 m times and contains 1111 $m-1$ times. Together, the eight quadruples occupy all $8m-1$ rows of A. In particular, columns $1, 2, 4m+1, 4m+2$ of A do not contain the quadruple 1000 and 0111, so $(\{1\}, \{2, 4m+1, 4m+2\})$ cannot be separated by A. □

Recall from Construction 4 that for a prime power $q \equiv 3 \bmod 4$, the Paley difference set construction yields a $(q, \frac{q-1}{2}, \frac{q-3}{4})$-SBIBD, and hence is a Hadamard design by Theorem

3.4.5. When $q > 11$ is prime, we will show that the incidence matrices of these designs are $\{1,3\}$-SHFs. The proof is similar to the main theorem in [14]; it is based on a character-theoretic bound proven by Burgess [7].

Theorem 3.4.8. *For all primes $q \equiv 3 \bmod 4$, $q > 11$, there is a $(q, \frac{q-1}{2}, \frac{q-3}{4})$-SBIBD whose block-point incidence matrix is a $\{1,3\}$-SHF.*

Proof. Let $\chi : \mathbb{Z}_q{}^* \to \{1,-1\}$ be the quadratic character, defined as follows:

$$\chi(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x \in \mathsf{QR}(q) \\ -1 & \text{otherwise} \end{cases}.$$

Let $a_1, a_2, a_3, a_4 \in \mathbb{Z}_q$ be distinct. Define

$$S = \sum_{x \in \mathbb{Z}_q} \chi(x-a_1)\chi(x-a_2)\chi(x-a_3)\chi(x-a_4). \tag{3.19}$$

By [7, Lemma 1], it immediately follows that $S \le 2\sqrt{q}+1$. For any integer $q > 11$, it is easy to see that $2\sqrt{q}+1 < q-4$. Therefore, $S < q-4$. Clearly the sum in (3.19) contains exactly four terms equal to 0. The remaining $q-4$ terms in this sum are all equal to ± 1. Since $S < q-4$, there must be a term in the sum equal to -1. That is, there exists $x \in \mathbb{Z}_q$ such that exactly one or three of the four (non-zero) values $\chi(x-a_1), \chi(x-a_2), \chi(x-a_3), \chi(x-a_4)$ are equal to 1. In the associated design, we have a block that contains an odd number of points from $\{a_1, a_2, a_3, a_4\}$. Applying Corollary 3.4.2, it follows that the incidence matrix of the design is a $\{1,3\}$-SHF. □

For all primes $q \equiv 3 \bmod 4$, $q > 1024$, it is noted in Colbourn and Kéri [11] that Paley difference sets yield covering arrays of strength four, which immediately implies that they are $\{1,3\}$-SHFs. This follows from a similar character-theoretic argument.

3.4.2 The Case $k = 3\lambda$

The case $k = 3\lambda$ is especially interesting because this corresponds to $\alpha = 0$ in Theorem 3.4.2. In this situation, the four-point substructure is an *oval*, using the terminology of Assmus and van Lint [1] (the paper [1] is a general study of ovals in symmetric BIBDs). Specializing Corollary 3.4.2 to this case, we obtain the following.

Corollary 3.4.9. *Let $(X, \mathcal{A})$ be a symmetric (v, k, λ)-*SBIBD *with $k = 3\lambda$. Then $(X, \mathcal{A})$ is not an* SHF$(v; v, 2, \{1, 3\})$ *if and only if $(X, \mathcal{A})$ contains an oval (of cardinalty 4).*

We next present some examples to show how Corollary 3.4.9 can be used to determine if a specific parameter set gives rise to $\{1, 3\}$-SHFs.

Example 3.4.2. There is a unique $(7, 3, 1)$-SBIBD up to isomorphism. As is observed in [1], the complement of any block is an oval. Therefore the $(7, 3, 1)$-SBIBD is not a $\{1, 3\}$-SHF.

Example 3.4.3. There are precisely three non isomorphic $(16, 6, 2)$-SBIBDs. It is observed in [1] that all three of these designs contain ovals. Therefore, no $(16, 6, 2)$-SBIBD is a $\{1, 3\}$-SHF.

Example 3.4.4. It is observed in [1] that there is a $(25, 9, 3)$-SBIBD that contains an oval. Therefore this SBIBD is not a $\{1, 3\}$-SHF. In fact, Denniston later showed in [13] that all 78 non isomorphic $(25, 9, 3)$-SBIBDs contain an oval, so there are no $(25, 9, 3)$-SBIBDs whose incidence matrices are $\{1, 3\}$-SHFs.

Finally, we present an infinite family of symmetric BIBDs with $k = 3\lambda$ whose incidence matrices are not $\{1, 3\}$-SHFs.

Theorem 3.4.10. *For all integers $h \geq 2$, there is a $(3^{h+1} - 2, 3^h, 3^{h-1})$-*SBIBD *whose incidence matrix is not a $\{1, 3\}$-*SHF*.*

Proof. It is shown by Tran in [31] that the Mitchell-Rajkundlia designs with the above parameters all contain ovals. (Actually, Tran showed that the Mitchell-Rajkundlia designs constructed from the Desarguesian affine planes of order 2^n all contain *maximal 2^m-arcs* for $1 \leq m \leq n$. For the specific Mitchell-Rajkundlia designs with the indicated parameters, we have $m = 1$, and the maximal 2-arcs are in fact ovals.) □

3.4.3 Further Discussion

We have a simple result which shows that certain symmetric BIBDs are $\{1, w\}$-SHF.

Theorem 3.4.11. *Suppose there exists a symmetric (v, k, λ)-*SBIBD *where $k > w\lambda$. Then the block-point incidence matrix of this SBIBD is a $\{1, w\}$-*SHF*.*

Proof. Let A be the block-point incidence matrix of the hypothesized design. Let i be one column of A and let $j_1, \ldots, j_w$ be w additional columns of A . For $1 \leq \ell \leq w$, define

$$R_\ell = \{r : \mathsf{A}(r, i) = \mathsf{A}(r, j_\ell) = 1\}.$$

Clearly $|R_\ell| = \lambda$ for all ℓ, so

$$\left| \bigcup_{\ell=1}^{w} R_\ell \right| \leq w\lambda.$$

There are k rows of A having a 1 in column i. Since $k > w\lambda$, there exists at least one row of A having a 1 in column i and 0's in columns $j_1, \ldots, j_w$. □

In the case $w = 3$, Theorem 3.4.11 provides a simple proof of the first part of Theorem 3.4.1.

Table 3.3 lists parameters for 'small' symmetric BIBDs and constructions that give rise to $\{1, 3\}$-SHFs (or not). The case of $\lambda = 1$ for $k \geq 4$ is characterized by Theorem 3.3.1 and so these parameters are omitted from the table. All other valid triples (v, k, λ) with $k \leq \frac{v}{2}$ for which the existence of a (v, k, λ)-SBIBD is known are presented in order of increasing k. The following theorem fills in an additional entry in Table 3.3.

Theorem 3.4.12. *There is a* $(39, 19, 9)$-SBIBD *whose incidence matrix is a* $\{1, 3\}$-SHF.

Proof. The website [19] includes 22 (known to date) skew Hadamard matrices of order 40. We derived Hadamard designs (i.e., $(39, 19, 9)$-SBIBDs) from all of them by standardizing with respect to a given row and column and then deleting the given row and column. Then we checked the resulting $(39, 19, 9)$-SBIBDs by computer to see if they are $\{1, 3\}$-SHF. It turned out that eight of these matrices, namely numbers 1, 5, 7, 10, 11, 13, 17 and 20, give rise to $(39, 19, 9)$-SBIBDs which are $\{1, 3\}$-SHF. Moreover, the transposes of the incidence matrices of these 22 $(39, 19, 9)$-SBIBDs give rise to eight additional $(39, 19, 9)$-SBIBDs which are $\{1, 3\}$-SHF, namely numbers 2, 3, 8, 12, 14, 16, 18 and 21. It did not matter which row/column we chose for the standardization process. □

Define $\mathcal{F}_w$ to be the set of all parameter triples (v, k, λ) such that there exists a symmetric (v, k, λ)-BIBD whose incidence matrix is a $\{1, w\}$-SHF, and define $\overline{\mathcal{F}}_w$ to be the set of all parameter triples (v, k, λ) such that there exists a symmetric (v, k, λ)-BIBD whose incidence matrix is *not* a $\{1, w\}$-SHF.

Definition 3.4.4. A parameter triple (v, k, λ) will be called a *Hadamard triple* if it has the form $(4t + 3, 2t + 1, t)$ for a positive integer t, and a *non-Hadamard triple* otherwise.

Legend	Description
T_1	Guaranteed to be {1,3}-SHFs by Theorem 3.4.1 from $k \geq 3\lambda + 1$
T_2	Guaranteed to be {1,3}-SHFs by Theorem 3.4.1 from $k - \lambda$ odd
H	Construction from Theorem 3.4.7
QR(q)	Quadratic residue difference set (Theorem 3.4.8)

v	k	λ	$\{1,3\}$-SHF	not $\{1,3\}$-SHF	Comment
7	3	1	None	All	Example 3.4.2
11	5	2	All	None	T_2
16	6	2	None	All	Example 3.4.3
15	7	3	None	All	Table 3.2
37	9	2	All	None	T_1
25	9	3	None	All	Example 3.4.4
19	9	4	All	None	T_2
31	10	3	All	None	T_1
56	11	2	All	None	T_1
23	11	5	QR(23)	H	
45	12	3	All	None	T_1
79	13	2	All	None	T_1
40	13	4	All	None	T_1
27	13	6	All	None	T_2
71	15	3	All	None	T_1
36	15	6	All	None	T_2
31	15	7	QR(31)	H	
61	16	4	All	None	T_1
49	16	5	All	None	T_1
41	16	6	[10, §II.6.9]	?	computer verified
69	17	4	All	None	T_1
35	17	8	All	None	T_2
39	19	9	Theorem 3.4.12	H	
96	20	4	All	None	T_1
85	21	5	All	None	T_1
71	21	6	All	None	T_1
43	21	10	All	None	T_2
78	22	6	All	None	T_1
47	23	11	QR(47)	H	
70	24	8	[10, §II.6.9]	?	computer verified
121	25	5	All	None	T_1
101	25	6	All	None	T_1
61	25	10	All	None	T_2
51	25	12	All	None	T_2

Table 3.3: Small Symmetric BIBDs and $\{1,3\}$-SHF

There are several parameter triples in Table 3.3 that are in $\mathcal{F}_3 \cap \overline{\mathcal{F}}_3$. However, all of these examples are Hadamard triples. We now provide an example of a non-Hadamard triple in $\mathcal{F}_3 \cap \overline{\mathcal{F}}_3$, namely $(64, 28, 12)$.

Theorem 3.4.13. *There exists a* $(64, 28, 12)$-SBIBD *whose incidence matrix is a* $\{1, 3\}$-SHF*, as well as a* $(64, 28, 12)$-SBIBD *whose incidence matrix is not a* $\{1, 3\}$-SHF*.*

Proof. We have verified by computer that the incidence matrix of the design D_1 in [12, p. 113] is a $\{1, 3\}$-SHF. Furthermore, the incidence matrix of the design constructed from the difference set in $\mathbb{Z}_4 \times \mathbb{Z}_{16}$ (see [10, p. 428]) is not a $\{1, 3\}$-SHF. □

3.4.4 Two Problems on Binary Frameproof Codes

In Section 2.2 we mentioned two open problems suggested by our work. Namely, given a positive integer w,

(i) find the smallest N such that there exists an $\mathsf{SHF}(N; N, 2, \{1, w\})$ whose representation matrix in standard form is not a permutation matrix, and

(ii) find the smallest N such that there exists an $\mathsf{SHF}(N; n, 2, \{1, w\})$ with $n > N$.

Using Theorem 2.2.11, Theorem 3.4.1 and Example 3.2.2, we have limited the range of possibilities for problem (ii) for $w = 3$ to $N = 10$ or 11. For the general case, Corollary 3.3.2 provides an upper bound to both problems.

Part (i) of Corollary 3.3.2 gives an upper bound for problem (i). More specifically, for any integer $w \geq 3$, let q be the smallest prime power such that $q \geq w$. An upper bound for the smallest N such that there exists an $\mathsf{SHF}(N; N, 2, \{1, w\})$ whose representation matrix in standard form is not a permutation is $q^2 + q + 1$, since the point-block incidence matrix of the projective plane of order q is not a permutation matrix.

Part (ii) of Corollary 3.3.2 gives an upper bound for problem (ii). That is, for any integer $w \geq 3$, let q be the smallest prime power such that $q > w$. An upper bound for the smallest N such that there exists an $\mathsf{SHF}(N; n, 2, \{1, w\})$ with $n > N$ is q^2.

Chapter 4

Future Work

We propose several interesting research problems along the line of work presented in this thesis.

1. Improve the range of applicable N for the lower bound of Theorem 2.2.11. The current best result by Shangguan et al. [23] applies to $w+1 \leq N \leq \binom{w+1}{2} - 1$. Any success would establish a tighter lower bound for problems (i) and (ii) in Section 3.4.4.

2. Can we give a result similar to Theorem 2.2.11 for non-binary SHFs? Construction 1 serves as the analogous canonical construction, giving a lower bound of $n \geq (q-1)N$. Can we prove a matching upper bound and show that any $\mathsf{SHF}(N; n, q, \{1, w\})$ with $q < w$ and $n = (q-1)N$ is equivalent to Construction 1?

3. Theorem 2.3.4 applies only to SHFs of type $\{w_1^{q-1}, w_2\}$, which seems more restrictive than what the result applies to. Could a similar result be shown for general $\{w_1, w_2, \ldots, w_q\}$-SHFs?

4. It would also be interesting to see a characterization of non-symmetric BIBDs similar to Theorem 3.4.1. For the problem of constructing large binary frameproof codes, SBIBDs do not offer much more than permutation matrices.

5. In Section 3.4.3 we defined $\mathcal{F}_w$ to be the set of all parameter triples (v, k, λ) for which there exists a symmetric (v, k, λ)-BIBD whose incidence matrix is a $\{1, w\}$-SHF, and $\overline{\mathcal{F}}_w$ to be the set of all parameter triples (v, k, λ) for which there exists a symmetric (v, k, λ)-BIBD whose incidence matrix is *not* a $\{1, w\}$-SHF. Results in Chapter 3 show that both $\mathcal{F}_3$ and $\overline{\mathcal{F}}_3$ are infinite. Is $\mathcal{F}_3 \cap \overline{\mathcal{F}}_3$ also infinite?

6. Theorem 3.4.11 can be viewed as a partial generalization of Theorem 3.4.1 in the sense that it gives a sufficient condition for the block-point incidence matrix of a (v, k, λ)-SBIBD to also be the representation matrix of a $\{1, w\}$-SHF. However, we did not find a counterpart to Corollary 3.4.2. Proving such a counterpart could allow us to establish a tighter upper bound for problems (i) and (ii) in Section 3.4.4.

7. In Table 3.3, we listed several triples (v, k, λ) for which a (v, k, λ)-SBIBD is also a $\{1, 3\}$-SHF without satisfying either the $k \geq 3\lambda + 1$ or $k - \lambda$ odd requirement of Theorem 3.4.1. We know that there are infinitely many such Hadamard triples by Theorem 3.4.8, while the two triples from [10, §II.6.9] are non-Hadamard. Is there an infinite number of non-Hadamard triples in $\mathcal{F}_3$ that do not satisfy either requirement of Theorem 3.4.1?

References

[1] E. F. Assmus Jr. and J. H. van Lint. Ovals in projective designs. *Journal of Combinatorial Theory, Series A* **27** (1979), pp. 307–324.

[2] M. Bazrafshan and Tran van Trung. Bounds for separating hash families. *Journal of Combinatorial Theory, Series A* **118** (2011), pp. 1129–1135.

[3] M. Bazrafshan and Tran van Trung. Improved bounds for separating hash families. *Designs, Codes and Cryptography* (2013), pp. 369–382.

[4] S. R. Blackburn. Frameproof codes. *SIAM Journal of Discrete Mathematics* **16** (2003), pp. 499–510.

[5] S. R. Blackburn, T. Etzion, D. R. Stinson, G. M. Zaverucha. A bound on the size of separating hash families. *Journal of Combinatorial Theory, Series A* **115** (2008), pp. 1246–1256.

[6] D. Boneh and J. Shaw. Collusion-secure fingerprinting for digital data. *IEEE Transactions on Information Theory* **44** (1998), pp. 1897–1905.

[7] D. A. Burgess. On character sums and primitive roots. *Proceedings of the London Mathematical Society* **12** (1962), pp. 179–192.

[8] M. Cheng and Y. Miao. On anti-collusion codes and detection algorithms for multimedia fingerprinting. *IEEE Transactions on Information Theory* **57** (2011), pp. 4843–4851.

[9] G. D. Cohen, S. B. Encheva and H. G. Schaathun. On separating codes. Technical Report 2001D003, TELECOM ParisTech, Ecole Nationale Superieure des Telecommunications, 2001.

[10] C. J. Colbourn and J. H. Dinitz. *Handbook of Combinatorial Designs, Second Edition*, Chapman & Hall / CRC, 2006.

[11] C. J. Colbourn and G. Kéri. Binary covering arrays and existentially closed graphs. *Lecture Notes in Computer Science* **5557** (2009), pp. 22–33 (IWCC 2009 Proceedings).

[12] D. Crnković and M. O. Pavčević. Some new symmetric designs with parameters $(64, 28, 12)$. *Discrete Mathematics* **237** (2001), pp. 109–118,

[13] R. H. F. Denniston. Enumeration of symmetric designs $(25, 9, 3)$. *Annals of Discrete Mathematics* **15** (1982), pp. 111–127.

[14] R. L. Graham and J. H. Spencer. A constructive solution to a tournament problem. *Canadian Mathematical Bulletin* **14** (1971), pp. 45–47.

[15] C. Guo, D. R. Stinson and Tran van Trung. On tight bounds for binary frameproof codes. To appear in *Designs, Codes, and Cryptography*.

[16] C. Guo, D. R. Sintson and Tran van Trung. On symmetric designs and binary 3-frameproof codes. To appear in *Springer Proceedings in Mathematics and Statistics: Algebraic Design Theory and Hadamard Matrices*.

[17] H. Kharaghani and B. Tayfeh-Rezaie. On the classification of Hadamard matrices of order 32. *Journal of Combinatorial Designs* **18** (2010), pp. 328–336.

[18] H. Kimura. Classification of Hadamard matrices of order 28. *Discrete Mathematics* **133** (1994), pp. 171–180.

[19] C. Koukouvinos. www.math.ntua.gr/people/(ckoukouv2)/hadamard.htm.

[20] K. Mehlhorn. On the program size of perfect and universal hash functions. *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science* (1982), pp. 170–175.

[21] C. Peikert, A. Shelat and A. Smith. Lower bounds for collusion-secure fingerprinting. *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (2003).

[22] P. Sarkar and D. R. Stinson. Frameproof and IPP codes. Progress in Cryptology – Indocrypt 2001, *Lecture Notes in Computer Science*, Springer, **2247** (2001), pp. 117–126.

[23] C. Shangguan, X. Wang, G. Ge and Y. Miao. New bounds for frameproof codes. Preprint, 2014. http://arxiv.org/pdf/1411.5782v1.pdf

[24] J. N. Staddon, D. R. Stinson and R. Wei. Combinatorial properties of frameproof and traceability codes, *IEEE Transactions on Informaction Theory* **47** (2001), pp. 1042–1049.

[25] D. R. Stinson. *Combinatorial Designs: Constructions and Analysis*, Springer Verlag, 2003.

[26] D. R. Stinson, Tran van Trung, and R. Wei. Secure frameproof codes, key distribution patterns, group testing algorithms and related structures. *Journal of Statistical Planning and Inference* **86** (2000), pp. 595–617.

[27] D. R. Stinson and R. Wei. Combinatorial properties and constructions of traceability schemes and frameproof codes. *SIAM Journal on Discrete Mathematics* **11** (1998), pp. 41–53.

[28] D. R. Stinson, R. Wei and K. Chen. On generalized separating hash families. *Journal of Combinatorial Theory, Series A* **115** (2008), pp. 105–120.

[29] D. R. Stinson, R. Wei and L. Zhu. New constructions for perfect hash families and related structures using combinatorial designs and codes. *Journal of Combinatorial Designs* **8** (2000), pp. 189–200.

[30] W. Trappe, M. Wu, J. Wang, and K. J. R. Liu. Anti-collusion fingerprinting for multimedia. *IEEE Transactions on Signal Processing* **51** (2003), pp. 1069–1087

[31] Tran van Trung. Maximal arcs and related designs. *Journal of Combinatorial Theory, Series A* **57** (1991), pp. 294–301.

[32] Tran van Trung. A tight bound for frameproof codes viewed in terms of separating hash families. *Designs, Codes and Cryptography* **72** (2014), pp. 713–718.