Improved bounds for separating hash families

Marjan Bazrafshan
Tran van Trung
Institut für Experimentelle Mathematik
Universität Duisburg-Essen
Ellernstrasse 29
45326 Essen, Germany
{marjan,trung}@iem.uni-due.de

Abstract

An $(N; n, m, \{w_1, \ldots, w_t\})$-separating hash family is a set $\mathcal{H}$ of N functions $h: X \longrightarrow Y$ with $|X| = n$, $|Y| = m$, $t \geq 2$ having the following property. For any pairwise disjoint subsets $C_1, \ldots, C_t \subseteq X$ with $|C_i| = w_i$, $i = 1, \ldots, t$, there exists at least one function $h \in \mathcal{H}$ such that $h(C_1), h(C_2), \ldots, h(C_t)$ are pairwise disjoint. Separating hash families generalize many known combinatorial structures such as perfect hash families, frameproof codes, secure frameproof codes, identifiable parent property codes. In this paper we present new upper bounds on n which improve many previously known bounds. Further we include constructions showing that some of these bounds are optimal.

Keywords. Separating hash family, perfect hash family, frameproof code, 2-IPP code.

1 Introduction

Let h be a function from a set X to a set Y and let $C_1, C_2, \ldots, C_t \subseteq X$ be t pairwise disjoint subsets. We say that h *separates* $C_1, C_2, \ldots, C_t$ if $h(C_1), h(C_2), \ldots, h(C_t)$ are pairwise disjoint. Let $|X| = n$ and $|Y| = m$. We call a set $\mathcal{H}$ of N functions from X to Y an $(N; n, m, \{w_1, \ldots, w_t\})$-*separating hash family*, and we shall also write as an $\mathsf{SHF}(N; n, m, \{w_1, w_2, \ldots, w_t\})$, if for all pairwise disjoint subsets $C_1, \ldots, C_t \subseteq X$ with $|C_i| = w_i$, for $i = 1, \ldots, t$, there exists at least one function $h \in \mathcal{H}$ that separates $C_1, C_2, \ldots, C_t$. The multiset $\{w_1, w_2, \ldots, w_t\}$ is the *type* of the separating hash family. To exclude trivial cases we always assume that $n > m$, $t \geq 2$ and $\sum_{i=1}^{t} w_i \leq n$. Obviously, we have $t \leq m$. Separating hash family with $t = 2$ was introduced in [21] and the general case in [23]. Separating hash families include various well-studied objects. For example, if $w_1 = \cdots = w_t = 1$, an $\mathsf{SHF}(N; n, m, \{1, 1, \ldots, 1\})$ is called a *perfect hash family* which is usually denoted by $\mathsf{PHF}(N; n, m, t)$. Perfect hash families have been studied extensively, see for instance, [2], [9], [6], [18], [17], [25], [20]. *w-frameproof codes* are separating hash families of type $\{1, w\}$ [10], [19] and *w-secure frameproof codes* are separating hash families of type $\{w, w\}$ [21]. Further, *codes with identifiable parent property (2-IPP codes)* are separating hash families of type $\{1, 1, 1\}$ and $\{2, 2\}$ simultaneously [15], [19], [24].

An $\mathsf{SHF}(N; n, m, \{w_1, w_2 \ldots, w_t\})$ can be depicted as an $N \times n$ array A in which the columns are labeled by the elements of X, the rows by the functions $h_i \in \mathcal{H}$ and the $(i,j)-$ entry of the array is the value $h_i(j)$. Thus, an $\mathsf{SHF}(N; n, m, \{w_1, w_2 \ldots, w_t\})$ is equivalent to an $N \times n$ array with entries from a set of m symbols such that for all disjoint sets of columns $C_1, C_2, \ldots, C_t$ of A with $|C_i| = w_i$, for $i = 1, 2, \ldots, t$, there exists at least one row r of A such that

$$\{\mathsf{A}(r,x) : \ x \in C_i\} \cap \{\mathsf{A}(r,y) : \ y \in C_j\} = \emptyset.$$

for all $i \neq j$. We call A the *matrix representation* or *array representation* of the hash family.

One of the main problems in studying separating hash families is to maximize n, when the other parameters N, m and $\{w_1, w_2, \ldots, w_t\}$ are given. The determination of bounds for n has been subject of much research recently [5], [16], [19], [22], [23], [3].

Often orthogonal arrays are used to construct certain good classes of separating hash families. An *orthogonal array* $\mathsf{OA}(t, N, m)$ is an $N \times m^t$ array A with entries from a set of $m \geq 2$ symbols such that within any t rows of A every possible $t-$tuple of symbols occurs exactly once. This property is equivalent to the fact that every two columns of A agree in at most $t-1$ rows. A classical construction of orthogonal arrays is as follows [12]. Let q be a prime power and $t \geq 2$. Let $\mathcal{P} = \{P_1, P_2, \ldots, P_{q^t}\}$ be the set of all polynomials of degree at most $t-1$ over the finite field $\mathbb{F}_q$. Now let $\mathcal{R}$ be a subset of elements of $\mathbb{F}_q \cup \{\infty\}$. Define an $|\mathcal{R}| \times q^t$ array A in which the entry $\mathsf{A}(u,j)$ is $P_j(u)$ if $u \in \mathcal{R} \setminus \{\infty\}$ and is a_{t-1} when $P_j(x) = \sum_{i=0}^{t-1} a_i x^i$ and $u = \infty$. Then A is an $\mathsf{OA}(t, |\mathcal{R}|, q)$.

This paper contains new results on bounds for separating hash families. The bounds obtained improve the known bounds in the literature. Section 2 presents some basic results and known bounds for separating hash families. Section 3 contains a new bound for SHF of general type $\{w_1, \ldots, w_t\}$ with $t \geq 3$. Section 4 deals with bounds for SHF when $N = w_1 + \cdots + w_t$. At first, new bounds for SHF of type $\{1, w\}$ and type $\{2, 2\}$ are proved, which are then used to prove a bound for SHF of type $\{w_1, w_2\}$ by the method of induction. A bound for SHF of general type is derived from this bound. Section 5 shows an optimal bound for SHF of type $\{1, 2\}$ when N is odd. Section 6 presents constructions of optimal or asymptotically optimal SHF for several types.

2 Basic results and known bounds on separating hash families

In the following we present some basic results and several known best bounds for separating hash families.

The following results are basic and useful, see for instance [23].

Lemma 1 *If an* $\mathsf{SHF}(N; n, m, \{w_1, w_2, \ldots, w_t\})$ *exists, then an* $\mathsf{SHF}(N; n, m, \{w_1', w_3, \ldots, w_t\})$ *with* $w_1' = w_1 + w_2$ *exists.*

Lemma 2 *Let* $c \geq 2$ *be an integer. Suppose there exists an* $\mathsf{SHF}(N; n, m, \{w_1, \ldots, w_t\})$. *Then there exists an* $\mathsf{SHF}(\lceil \frac{N}{c} \rceil; n, m^c, \{w_1, \ldots, w_t\})$.

Proof. Let $\mathcal{H} = \{h_1, h_2, \ldots, h_N : X \longrightarrow Y\}$ be an $\mathsf{SHF}(N; n, m, \{w_1, \ldots, w_t\})$. Let $d := \lceil \frac{N}{c} \rceil$. Consider d subsets $A_1, \ldots, A_d$ of $\{1, 2, \ldots, N\}$ such that $|A_u| = c$ for $u = 1, \ldots, d$ and $A_1 \cup \ldots \cup A_d =$

$\{1, 2, \ldots, N\}$. Define a hash family $\mathcal{H}' = \{h'_1, h'_2, \ldots, h'_d : X \longrightarrow Y^c\}$, where $h'_u(x) = (h_i(x) : i \in A_u)$. We see that $\mathcal{H}'$ is an $\mathsf{SHF}(d; n, m^c, \{w_1, \ldots, w_t\})$. This is because if the sets $h_{i_0}(C_j)$ and $h_{i_0}(C_k)$ are disjoint, where $i_0 \in A_u$ and $u \in \{1, \ldots, d\}$, then the sets $h'_u(C_j)$ and $h'_u(C_k)$ are also disjoint. For if we have $h'_u(C_j) \cap h'_u(C_k) \neq \emptyset$, then there are $x \in C_j$ and $y \in C_k$ such that $h'_u(x) = h'_u(y)$. This implies that $h_i(x) = h_i(y)$ for all $i \in A_u$, contradicting the fact that $h_{i_0}(x) \neq h_{i_0}(y)$ as $h_{i_0}(C_j)$ and $h_{i_0}(C_k)$ are disjoint. □

In the following we only record some known best bounds for separating hash families. Further bounds for SHF are found in recent papers, see [5], [22], [23], [16], [13], [21], [15], [4].

For special type $\{1, 1, 2\}$ Stinson, Wei and Chen have proved the following strong bound.

Theorem 1 ([23]) *Suppose there is an* $\mathsf{SHF}(N; n, m, \{1, 1, 2\})$*. Then* $n \leq 3m^{\lceil \frac{N}{3} \rceil} + 2 - 2\sqrt{3m^{\lceil \frac{N}{3} \rceil} + 1}$.

A strong bound for SHF of general type is obtained by Blackburn, Etzion, Stinson and Zaverucha [5].

Theorem 2 ([5]) *If there is an* $\mathsf{SHF}(N; n, m, \{w_1, \ldots, w_t\})$ *, then*

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

where $u = \sum_{i=1}^{t} w_i$*,* $\gamma = (w_1.w_2 + u - w_1 - w_2)$ *and* $w_1, w_2 \leq w_i$ *for* $i = 3, \ldots, t$*.*

However, this bound has been improved by the following recent results in [3].

Theorem 3 ([3]) *Suppose there exists an* $\mathsf{SHF}(N; n, m, \{w_1, \ldots, w_t\})$*. Let* $u = \sum_{i=1}^{t} w_i$*. Then*

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

Theorem 4 ([3]) *Let* $t \geq 3$ *be an integer. Suppose there exists an* $\mathsf{SHF}(N; n, m, \{w_1, w_2, \ldots, w_t\})$*. Let* $u = \sum_{i=1}^{t} w_i$*. Then*

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

3 An improved bound for SHF of general type $\{w_1, \ldots, w_t\}$ with $t \geq 3$

In this section we present a new bound for SHF of general type. The proof for this bound uses the method of induction. We first prove a useful lemma about the structure of a general SHF. This crucial lemma is then used for the induction step of the proof.

Lemma 3 *Suppose there exists an* $\mathsf{SHF}(N; n, m, \{w_1, \ldots, w_t\})$ *with* $w_1 + \cdots + w_t \geq 3$*,* $w_t \geq w_i$ *for* $i = 1, \ldots, w_{t-1}$ *and* $n - m \geq w_1 + \cdots + w_t - 1$*. Then there exists an* $\mathsf{SHF}(N-1; n_1, m, \{w_1, \ldots, w_{t-1}, w_t - 1\})$ *with* $n_1 \geq n - m$*.*

Proof. Let A be the matrix representation of an $\mathsf{SHF}(N; n, m, \{w_1, \ldots, w_t\})$ with $w_1 + \cdots + w_t \geq 3$. Let m_1 denote the number of symbols that appear in the first row of A. Since permuting the columns of A does not change the separation property, we may assume that the first row of A has pairwise different symbols in the first m_1 columns. Let A_1 denote the $(N-1) \times (n - m_1)$ matrix obtained from A by ignoring the first row and the first m_1 columns of A. Set $n_1 := n - m_1$. Then $n_1 \geq n - m \geq w_1 + \cdots + w_{t-1} + w_t - 1$. We claim that A_1 is an $\mathsf{SHF}(N-1; n_1, m, \{w_1, \ldots, w_{t-1}, w_t - 1\})$. Assume that A_1 is not an $\mathsf{SHF}(N-1; n_1, m, \{w_1, \ldots, w_{t-1}, w_t - 1\})$. Then there are column sets $C_1, \ldots, C_{t-1}, C_t$ with $|C_i| = w_i$, $i = 1, \ldots, t-1$ and $|C_t| = w_t - 1$, that are not separated in any row of A_1. Note that if $w_t = 1$ (i.e. $w_1 = w_2 = \ldots = w_t = 1$), then we consider $|C_t|$ temporarily as an empty set. Let a be a symbol appearing in some column of C_1 in the first row of A. Then in the first m_1 columns of A there is a column c having symbol a in the first row. Add this column c to C_t. Now it is easily checked that $C_1, \ldots, C_{t-1}, C_t \cup \{c\}$ are not separated in A, which contradicts the separation property of A. □

We now prove a new bound for $\mathsf{SHF}(u-1; n, m, \{w_1, w_2, w_3\})$.

Theorem 5 *Suppose there exists an* $\mathsf{SHF}(u-1; n, m, \{w_1, w_2, w_3\})$*, where* $u = w_1 + w_2 + w_3$ *and* $w_3 \geq 2$*. Then*

$$n \leq (u-1)m + 2 - 2\sqrt{3m+1}.$$

Proof. We prove the theorem by induction on u. Note that $u \geq 4$. Let A be the matrix representation of an $\mathsf{SHF}(u-1; n, m, \{w_1, w_2, w_3\})$. Assume $u = 4$. Then $w_1 = w_2 = 1$ and $w_2 = 2$. By Theorem 1 of Stinson, Wei and Chen we have $n \leq 3m + 2 - 2\sqrt{3m+1}$. Hence, the statement is valid. Assume by induction that the statement $n \leq (u-1)m + 2 - 2\sqrt{3m+1}$ is valid for all $u = 4, \ldots, k-1$, with $k - 1 \geq 4$. Suppose now that there exists an $\mathsf{SHF}(k-1; n, m, \{w_1, w_2, w_3\})$ with $w_3 \geq 2$ such that $n > (k-1)m + 2 - 2\sqrt{3m+1}$, where $k = w_1 + w_2 + w_3$. Since $k \geq 5$, $m \geq 3$ and $n - m > (k-2)m + 2 - 2\sqrt{3m+1}$, we have $n - m > k - 1$ and therefore $n - m > w_1 + w_2 + w_3 - 1$. By Lemma 3 there exists an $\mathsf{SHF}(k-2; n_1, m, \{w_1, w_2, w_3 - 1\})$ with $n_1 \geq n - m > (k-2)m + 2 - 2\sqrt{3m+1}$, which contradicts the assumption of the induction. This completes the proof. □

The next corollary is an immediate consequence of Theorem 5.

Corollary 1 *Suppose there exists an* $\mathsf{SHF}(u-1; n, m, \{w_1, \ldots, w_t\})$*, where* $u = \sum_{i=1}^{t} w_i \geq 4$ *and* $t \geq 3$*. Then*

$$n \leq (u-1)m + 2 - 2\sqrt{3m+1}.$$

Proof. The corrolary follows by Lemma 1, because any $\mathsf{SHF}(u-1; n, m, \{w_1, \ldots, w_t\})$ with $t \geq 4$ is also an $\mathsf{SHF}(u-1; n, m, \{w_1, w_2, w_3'\})$ with $w_3' = w_3 + \cdots + w_t$. □

We now use Lemma 2 and Corollary 1 to derive a new bound for SHF.

Theorem 6 *Suppose there exists an* $\mathsf{SHF}(N; n, m, \{w_1, w_2, \ldots, w_t\})$ *with* $t \geq 3$ *and* $u = \sum_{i=1}^{t} w_i \geq 4$*. Then*

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

Proof. Assume, by contradiction, that there exists an $\mathsf{SHF}(N; n, m, \{w_1, \dots, w_t\})$ with $t \geq 3$ and $u = \sum_{i=1}^{t} w_i \geq 4$ such that $n = (u-1)m^{\lceil \frac{N}{(u-1)} \rceil} + 2 - 2\sqrt{3m^{\lceil \frac{N}{(u-1)} \rceil} + 1} + 1$. By Lemma 2 there exists an $\mathsf{SHF}(\lceil \frac{N}{c} \rceil; n, m^c, \{w_1, \dots, w_t\})$ with $c := \lceil \frac{N}{(u-1)} \rceil$. Observe that if there exists an $\mathsf{SHF}(N; n, m, \{w_1, w_2, \dots, w_t\})$ with matrix representation A. Then for any $N' > N$ there exists an $\mathsf{SHF}(N'; n, m, \{w_1, w_2, \dots, w_t\})$ obtained by adding $N' - N$ arbitrary new rows using the same symbol set to A. Now, as $\lceil \frac{N}{c} \rceil \leq u - 1$, the observation says that there is an $\mathsf{SHF}(u-1; n, m^c, \{w_1, \dots, w_t\})$ with $n = (u-1)m^c + 2 - 2\sqrt{3m^c + 1} + 1$, which contradicts Theorem 5. □

Note that for any given type $\{w_1, \dots, w_t\}$ with $t \geq 3$ and $u \geq 4$ and a large value m the bound in Theorem 6 improves previously known bounds.

Remark 3.1 A method used to establishing bounds for SHF as shown in the literature and also in this section is that one first finds a bound for SHF with $N = u - 1$, where $u = \sum_{i=1}^{t} w_i$ and then apply Lemma 2 to obtain a bound for any N. It seems that the general bounds obtained in Theorems 3, 4 and 6 are strong when $(u-1)|N$. There are hints showing that the constant $(u-1)$ in the bound $n \leq (u-1)m$ for an $\mathsf{SHF}(u-1; n, m, \{w_1, \dots, w_t\})$ is the best possible when m is getting to infinity, in spite of the fact that this bound can be further improved. For instance, we have proved the following results.

(i) In an $\mathsf{SHF}(3; n, m, \{2, 2\})$ with $m \geq 7$, we have $n < 3m - 6$;
for $m \in \{2, \dots, 8\}$ we have $n \leq 2m$ if m is even, and $n \leq 2m - 1$ if m is odd.

(ii) In an $\mathsf{SHF}(4; n, m, \{3, 2\})$ with $m > 3$ we have $n \leq 4m - 6$.

(iii) In an $\mathsf{SHF}(5; n, m, \{3, 3\})$ with $m > 11$ we have $n < 5m - 13$.

4 A further improved bound for $\mathsf{SHF}(u; n, m, \{w_1, \dots, w_t\})$

From the general bounds for $\mathsf{SHF}(N; n, m\{w_1, \dots, w_t\})$ in Theorems 3, and 4 and 6 we see that n is roughly equal to $(u-1)m^2$ when $N = u$, where $u := \sum_{i=1}^{t} w_i$. In this section we present further improvement of the bounds for this case. More precisely, we prove that the constant $(u-1)$ can be reduced to a value of size at most 1.

The proof of the statement is as follows. Firstly, we prove the claim for SHF of types $\{1, w\}$ and $\{2, 2\}$, which is then used for an induction proof of the statement for type $\{w_1, w_2\}$. Finally, we apply Lemma 1 to obtain the result for the general type $\{w_1, \dots, w_t\}$.

4.1 A strong bound for $\mathsf{SHF}(w+1; n, m, \{1, w\})$

In this section we prove a strong bound on n for $\mathsf{SHF}(u; n, m, \{1, w\})$ with $u = 1 + w$. It turns out that this bound is optimal when $m \geq u$ as it will be shown in a subsequent section.

To begin with, we state a structural lemma for SHF of type $\{1, w\}$.

Lemma 4 *Suppose there is an* $\mathsf{SHF}(N; n, m, \{1, w\})$ *with* $N \geq w + 1$. *Let* $\mathsf{A} = (a_{i,j})$ *be its matrix representation. If two distinct columns* j_1 *and* j_2 *of* A *agree in at least* $N - (w - 1)$ *positions, then there are rows* i_1 *and* i_2 *such that* a_{i_ℓ, j_ℓ} *appears only once in row* i_ℓ *for* $\ell = 1, 2$.

Proof. By permuting the rows of A we can assume that the two columns j_1 and j_2 agree in the first ℓ positions, $\ell \geq N - (w - 1)$. If for each row $i = \ell + 1, \ldots, N$, there is a column j_i, $j_i \neq j_1$, such that $a_{i,j_1} = a_{i,j_i}$, then the two sets $C_1 = \{j_1\}$ and $C_2 = \{j_2, j_{\ell+1}, \ldots, j_N\}$, ($|C_2| \leq w$), cannot be separated. □

Theorem 7 *Suppose there is an* $\mathsf{SHF}(u; n, m, \{1, w\})$ *with* $u = 1 + w$. *Then we have*

(i) $n \leq m^2$, *if* $u \leq m$,

(ii) $n \leq um$, *if* $u > m$.

Proof. Le A be the matrix representation of an $\mathsf{SHF}(u; n, m, \{1, w\})$. Let C denote the set of columns of A. Divide C into two different parts A_1 and A_2, where A_1 consists of the columns which have the Hamming distance at least w to all other columns (i.e. any column in A_1 agrees in at most one position to all other columns). Define $\mathsf{A}_2 = \mathsf{C} - \mathsf{A}_1$. Thus, for any column $c_1 \in \mathsf{A}_2$ there exists at least one further column $c_2 \in \mathsf{A}_2$ such that c_1 and c_2 agree in at least two positions. By Lemma 4, if $c \in \mathsf{A}_2$, then there is some symbol in some row of c which cannot appear anywhere else in that row. For each column c in part A_2 consider one of these symbols and denote it by a_c. We construct pairwise disjoint sets of columns $\mathsf{C}_1, \mathsf{C}_2, \ldots, \mathsf{C}_u$, each C_i consisting of the columns c having a_c in row i, $1 \leq i \leq u$. If $|\mathsf{C}_i| = \ell_i, 1 \leq i \leq w + 1$, then $|\mathsf{A}_2| = \ell_1 + \ell_2 + \cdots + \ell_{w+1}, 0 \leq \ell_i \leq m$. On the other hand, $|\mathsf{C}_i| = \ell_i$ means that A_1 can have at most $m - \ell_i$ symbols in row i. Let $\ell_{max} = max\{\ell_i : 1 \leq i \leq u\}$ and $\ell_{min} = min\{\ell_i : 1 \leq i \leq u\}$. So there is some $i, 1 \leq i \leq u$ with $\ell_{min} = \ell_i$. Thus in A_1, the i'th row has the maximum possible number of symbols and each symbol can appear at most $m - \ell_{max}$ times, otherwise two columns in A_1 agree in more than one position. So we have:

$$\begin{aligned} n \;=\; |\mathsf{A}_1| + |\mathsf{A}_2| \;&\leq\; (m - \ell_{min})(m - \ell_{max}) + \ell_1 + \cdots + \ell_{w+1} \\ &\leq\; m^2 - m.\ell_{min} - m.\ell_{max} + \ell_{min}.\ell_{max} + (w + 1).\ell_{max}. \end{aligned}$$

As $\ell_{max} \leq m$ we have $\ell_{max}.\ell_{min} - m.\ell_{min} \leq 0$ and hence

$$n \leq m^2 + (w + 1 - m)\ell_{max},$$

which is less than or equal to um for $u > m$ and is bounded by m^2 for $u \leq m$. □

4.2 A new bound for $\mathsf{SHF}(4; n, m, \{2, 2\})$

Although the bound $n \leq m^2$ for $\mathsf{SHF}(u; n, m, \{1, w\})$ of type $\{1, w\}$ in previous section turns out to be optimal when $m \geq u$, as shown in section 6, we show in this section, however, that n is even less than m^2 for $\mathsf{SHF}(4; n, m, \{2, 2\})$.

Theorem 8 *If there exists an* $\mathsf{SHF}(4; n, m, \{2, 2\})$ *with* $m \geq 4$, *then* $n < m^2$.

Proof. Assume by contradiction that there exists an $\mathsf{SHF}(4; m^2, m, \{2,2\})$. Let A be its matrix representation. We first prove that every two columns of A can agree in at most one position.

Assume that there are two columns of A agreeing in the first two rows. If there are also two columns agreeing in the last two rows, then we have the following forbidden configuration in which the sets of columns $\{1,3\}$ and $\{2,4\}$ are not separable:

$$\begin{array}{cccc} a & a & * & * \\ b & b & * & * \\ * & * & c & c \\ * & * & d & d \end{array}$$

Therefore in the last two rows of A any two columns have at most one agreement. As there are m^2 columns in A, any pair of symbols appears in the last two rows as column exactly once. This implies that in the last two rows any symbol appears exactly m (≥ 4) times. Assume that the first two columns of A have the form.

$$\begin{array}{cc} a & a \\ b & b \\ x & z \\ y & t \end{array}$$

As x and t appear in the corresponding rows at least four times, we get the following forbidden configuration in which the sets of columns $\{1,4\}$ and $\{2,3\}$ are not separable.

$$\begin{array}{cccc} a & a & * & * \\ b & b & * & * \\ x & z & x & * \\ y & t & * & t \end{array}$$

This shows that any two columns of A agree in at most one row and every symbol appears in each row exactly m times.

Next we show that in an $\mathsf{SHF}(4; m^2, m, \{2,2\})$ there exist two columns having no agreement in all four rows.

Assume, by contradiction, that every two columns agree in at least one position (i.e. exactly one position). Consider a submatrix of A consisting of four columns having the same symbol in the first row (as $m \geq 4$, such columns exist) together with a fifth column having a different symbol in the first row. So we have the following configuration:

$$\begin{array}{ccccc} a & a & a & a & b \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{array}$$

where $b \neq a$. Each of the first four columns should agree with the last column at least in one row from the last three rows. Thus by the pigeonhole principle, two columns from the first four columns agree with the last column in the same row. It implies that two columns from the first four columns agree in at least two positions which contradicts the above argumentation.

Now we can prove that A is not an SHF of type $\{2,2\}$. Consider two columns of A having no agreement in all four rows.

$$\begin{matrix} x & \alpha \\ y & \beta \\ z & \gamma \\ t & \delta \end{matrix}$$

As each pair appears exactly once in every two rows and two columns agree in at most one position, we have the following submatrix

$$\begin{matrix} x & \alpha & x & * \\ y & \beta & \beta & * \\ z & \gamma & * & z \\ t & \delta & * & \delta \end{matrix}$$

in which the sets of columns $\{1,2\}$ and $\{3,4\}$ are not separable. This completes the proof. □

With Theorems 7 and 8 we are now in a position to prove a bound for SHF of type $\{w_1, w_2\}$.

Theorem 9 *Suppose there exists an* $\mathsf{SHF}(u; n, m, \{w_1, w_2\})$*, where* $u = w_1 + w_2$ *and* $m \geq u$*. Then* $n \leq m^2$*.*

Proof. The proof uses induction on u. The case $w_1 = 1 \leq w_2$ has been proved by Theorem 7 and the case $w_1 = w_2 = 2$ by Theorem 8. We may assume for the induction that $w_1, w_2 \geq 2$. Now assume as an induction step that the statement of the theorem is valid for $u \geq 4$. Assume, for a contradiction, that an $\mathsf{SHF}(u; m^2+1, m, \{w_1, w_2\})$ exists with A as the matrix representation. Let A_1 be the $(u-2) \times (m^2+1)$ matrix obtained from A by ignoring the first two rows of A. By the induction assumption A_1 is not an $\mathsf{SHF}(u-2; m^2+1, m, \{w_1-1, w_2-1\})$. Therefore there are two disjoint sets of columns C_1 and C_2 with $|\mathsf{C}_1| = w_1 - 1$ and $|\mathsf{C}_2| = w_2 - 1$, which are not separated in A_1.

Denote by M_i the set of all elements in the row i appearing in C_i for $i = 1, 2$. Also let C be the set of all columns of A. Define $\mathsf{C}' = \mathsf{C} \setminus (\mathsf{C}_1 \cup \mathsf{C}_2)$. By considering the first two rows and the columns of C', we see that one of the following two cases has to occur.

(i) There are columns $c_1, c_2 \in \mathsf{C}'$ with $c_1 \neq c_2$ such that the element in the second row of c_1 is a member of M_2 and the element in the first row of c_2 is a member of M_1. Then the sets $\mathsf{C}_1 \cup \{c_1\}$ and $\mathsf{C}_2 \cup \{c_2\}$ are not separated in A, a contradiction.

(ii) Either c_1 (or c_2) as defined in (i) does not exist or there exists only one column in which the element in the first row belongs to M_1 and the element in the second row belongs to M_2. W.l.o.g. we assume that c_1 does not exist. This implies that C' has at most $(m-1)$ elements in row 1 to fill at least $m^2 + 1 - (w_1 - 1 + w_2 - 1 + 1)$ columns. As $m \geq u$ we have $m^2 + 1 - (w_1 - 1 + w_2 - 1 + 1) \geq m^2 - m + 2$. Thus there is an element appearing at least $m+1$ times in row 1. Hence there are two columns $c_1, c_2 \in \mathsf{C}'$ agreeing in the first two rows. It means that $\mathsf{C}_1 \cup \{c_1\}$ and $\mathsf{C}_2 \cup \{c_2\}$ are not separated in A, a contradiction.

This completes the proof. □

As an immediate consequence of Theorem 9 we have the following corollary.

Corollary 2 *Suppose there exists an* $\mathsf{SHF}(u; n, m, \{w_1, \ldots, w_t\})$*, where* $u = \sum_{i=1}^{t} w_i$ *and* $m \geq u$*. Then* $n \leq m^2$*.*

Proof. The corollary follows from Theorem 9 and Lemma 1. □

When $w_1 = w_2 = w$ we can prove a slightly stronger result. The proof makes use of Theorem 8 and of a similar argument as that of Theorem 9, therefore we omit it.

Theorem 10 *Suppose there exists an* $\mathsf{SHF}(u; n, m, \{w, w\})$*, where* $u = w + w = 2w \geq 4$ *and* $m \geq u$*. Then* $n < m^2$*.*

5 An optimal bound for $\mathsf{SHF}(N; n, m, \{1, 2\})$

For the small type $\{1, 2\}$ we can prove a tight bound for SHF, when N is odd. For general bounds of $\mathsf{SHF}(N; n, m, \{1, w\})$ we refer the reader to [8].

Precisely, we prove the following result.

Theorem 11 *For any* $\mathsf{SHF}(2d + 1; n, m, \{1, 2\})$ *we have* $n \leq m^{d+1}$*.*

Proof. The following simple observation (O) is relevant for our proof. Let A be any $\mathsf{SHF}(2d + 1; n, m, \{1, 2\})$. If there are two columns of A agreeing in the first $(d + 1)$ rows (resp. in the last $(d + 1)$ rows), then the corresponding two $d-$tuples in the last d rows (resp. in the first d rows) of these two columns are unique. Since, otherwise these two columns could not be separated by a column having a repeated d-tuple in the last d rows (resp. in the first d rows).

Now assume there is an $\mathsf{SHF}(2d+1; m^{d+1}+1, m, \{1, 2\})$. Let A be its matrix representation. Since A has $m^{d+1} + 1$ columns, there are two columns agreeing in $d + 1$ first rows. So, from observation (O) the d-tuples of symbols in the last d rows of these two columns are unique.

Removing these two columns from A gives rise to an array B with $m^{d+1} - 1$ columns having only $m^d - 2$ (d)-tuples of symbols distributed in the last d rows. If each d-tuple of symbols appears at most m times in the last d rows, then we can fill only $(m^d - 2)m = m^{d+1} - 2m$ columns. So, there are $m^{d+1} - 1 - (m^{d+1} - 2m) = 2m - 1$ columns in which certain d-tuples of symbols in the last d rows are repeated at least $m+1$ times. This is to say that there are at least $2m - 1 + 1 = 2m$ $(d+1)$-tuples of symbols that have to repeat in the last $d + 1$ rows, as there are m symbols altogether. These $2m$ repeated $(d + 1)$-tuples (in the last $(d + 1)$ rows), provide $2m$ unique d-tuples in the first d rows by observation (O). Removing these $2m$ columns having unique d-tuples of symbols in the first d rows from A, gives rise to an array C with $m^{d+1} + 1 - 2m$ columns having $m^d - 2m$ different d-tuples in the first d rows. If each of these $m^d - 2m$ (d)-tuples appears at most m times, then again we can fill at most $(m^d - 2m)m = m^{d+1} - 2m^2$ columns. So there are $m^{d+1} - 2m + 1 - (m^{d+1} - 2m^2) = 2m^2 - 2m + 1$ columns with d-tuples in the first d rows that have to repeat at least $m + 1$ times. This gives us

$(2m^2-2m+1)+1 = 2m^2-2m+2$ repeated $(d+1)$-tuples in the first $d+1$ rows. Therefore, observation (O) provides $2m^2-2m+2$ unique (d)-tuples in the last d rows.

Now removing these $2m^2-2m+2$ columns from B we obtain an array D with $m^{d+1}-1-(2m^2-2m+2)$ columns having $m^d-2-(2m^2-2m+2) = m^d-2m^2+2m-4$ (d)-tuples in the last d rows. Again, if each of these d-tuples appear at most m times, only at most $(m^d-2m^2+2m-4)m = m^{d+1}-2m^3+2m^2-4m$ columns of D can be filled. Thus, there are $m^{d+1}-1-(2m^2-2m+2)-(m^{d+1}-2m^3+2m^2-4m) = 2m^3-4m^2+6m-3$ (d)-tuples of symbols in the last d rows repeated at least $m+1$ times. This implies that there are at least $2m^3-4m^2+6m-3+1 = 2m^3-4m^2+6m-2$ repeated $d+1$-tuples in the last $d+1$ rows. Hence, observation (O) shows that the corresponding d-tuples in the first d rows of these $d+1$-tuples must be unique.

We see that the number of unique $d-$tuples is increasing at each step. Continuing this argument after d steps will lead to a negative number of d-tuples available for a positive number of columns, which is a contradiction. □

6 Constructions for SHF

In this section we give several constructions for "good" SHF. First we present a construction for $\mathsf{SHF}(3; n, m, \{1,1,1,1\})$'s, whose value n is close to the bound given in Theorem 5. We then give two constructions for SHF derived from orthogonal arrays showing that the bounds in Theorems 7 and 11 are tight.

6.1 A Construction of $\mathsf{SHF}(3; n, m, \{1,1,1,1\})$

In this section we present a general construction of a good class of $\mathsf{SHF}(3; n, m, \{1,1,1,1\})$. We believe that for large values of m the number of columns n obtained from this construction is close to an optimal bound. To be more precise, the construction provides separating hash families with roughly $n \cong 3(m-2\lfloor\sqrt{m}\rfloor)$ columns. Thus $\lim_{m\to\infty}\tilde{n}/m = 3$, where $\tilde{n}$ is value of n such that an $\mathsf{SHF}(3; n, m, \{1,1,1,1\})$ exists. This implies that $\gamma = 3$ is asymtotically the best possible minimum value for constant γ such that $n < \gamma(m-c)$ for any fixed number $c > 0$. The construction that will be described also includes a construction for $\mathsf{SHF}(3; n, m, \{1,1,2\})$ and $\mathsf{SHF}(3; n, m, \{2,2\})$, since the existence of an $\mathsf{SHF}(3; n, m, \{1,1,1,1\})$ is equivalent to the existence of an SHF of types $\{1,1,2\}$ by Lemma 3.17 [23] and implies the existence of an SHF of type $\{2,2\}$ as well.

Let $m \geq 4$ be an integer. We write $m = m_1 + 2m_2$, where $m_2 = \lfloor\sqrt{m}\rfloor$. Let

$$V = V_1 \cup V_2 \cup V_3$$

be a set of m symbols consisting of a union of three disjoint sets
$V_1 = \{x_1, \ldots, x_{m_1}\}, \quad V_2 = \{y_1, \ldots, y_{m_2}\}, \quad \text{and} \quad V_3 = \{z_1, \ldots, z_{m_2}\}.$

Construction

Let $r \geq 1$, $\delta, c \geq 0$ be integers such that

a) $r \leq m_2$ and
b) $0 \leq m_1 - r(m_2 - \delta) := c \leq m_2$, i.e., $r(m_2 - \delta) + c = m_1$.

Define the following $(1 \times m_1)$ arrays:

$$\mathsf{X} = [x_1 \dots x_{m_1}]$$
$$\mathsf{Y}_1 = [\underbrace{y_1 \dots y_1}_{r} \; \underbrace{y_2 \dots y_2}_{r} \; \cdots \; \underbrace{y_{m_2-\delta} \cdots y_{m_2-\delta}}_{r} \; \underbrace{y_{m_2-\delta+1} \cdots y_{m_2-\delta+1}}_{c}]$$
$$\mathsf{Y}_2 = [\underbrace{y_1 y_2 \dots y_r \quad y_1 y_2 \dots y_r \quad \dots \quad y_1 y_2 \dots y_r}_{m_2-\delta} \quad y_1 y_2 \dots y_c]$$
$$\mathsf{Z}_1 = [\underbrace{z_1 \dots z_1}_{r} \; \underbrace{z_2 \dots z_2}_{r} \; \cdots \; \underbrace{z_{m_2-\delta} \cdots z_{m_2-\delta}}_{r} \; \underbrace{z_{m_2-\delta+1} \cdots z_{m_2-\delta+1}}_{c}]$$
$$\mathsf{Z}_2 = [\underbrace{z_1 z_2 \dots z_r \quad z_1 z_2 \dots z_r \quad \dots \quad z_1 z_2 \dots z_r}_{m_2-\delta} \quad z_1 z_2 \dots z_c]$$

Now define an $3 \times 3m_1$ array A.

$$\mathsf{A} \; = \; \begin{array}{|c|c|c|} \hline \mathsf{X} & \mathsf{Y}_1 & \mathsf{Z}_1 \\ \hline \mathsf{Y}_1 & \mathsf{X} & \mathsf{Z}_2 \\ \hline \mathsf{Y}_2 & \mathsf{Z}_2 & \mathsf{X} \\ \hline \end{array}$$

We show that A is an $\mathsf{SHF}(3; 3m_1, m, \{1,1,1,1\})$.

We divide the columns of A into 3 blocks S_1, S_3 and S_3. The first m_1 columns form the block S_1, the next m_1 columns the block S_2 and the last m_1 columns the block S_3. The following observation is useful.

-Two different columns from each block S_i, $i = 1, 2, 3$, agree in at most one row.

-Two columns from two different blocks S_i and S_j do not agree in any row.

Let $\{c_1, c_2, c_3, c_4\}$ be a given set of four columns of A. We need to consider the following cases.

(i) c_1, c_2, c_3, c_4 belong to one block. Then the row having elements of X separates these columns.

(ii) c_1, c_2, c_3, c_4 are distributed in two blocks. W.l.o.g. we need to consider only the following two cases.

(a) $c_1, c_2, c_3 \in S_1$ and $c_4 \in S_2$. Then the first row separates c_1, c_2, c_3, c_4.

(b) $c_1, c_2 \in S_1$ and $c_3, c_4 \in S_2$. If c_3, c_4 are separated in the first row, then the first row separates the four columns. Assume that c_3, c_4 are not separated in the first row. Then c_3, c_4 are separated in the second and the third row. Since c_1, c_2 have to be separated either in the second row or in the third row, it follows that c_1, c_2, c_3, c_4 are separated either by the second or the third row.

(iii) c_1, c_2, c_3, c_4 are distributed in three blocks. Wlog we may assume $c_1, c_2 \in S_1$, $c_3 \in S_2$ and $c_4 \in S_3$. It is obvious that the first row separates the four columns.

Thus A is a separating hash family of type $\{1,1,1,1\}$.

We record the result of the construction in the next theorem.

Theorem 12 *There exists an* $\mathsf{SHF}(3; 3m - 6\lfloor\sqrt{m}\rfloor, m, \{1,1,1,1\})$ *for any integer* $m \geq 4$*.*

Especially, if m is of the form $m = v^2 + 2v$, we choose $r = v$, $\delta = c = 0$ and obtain the following result. Note that this result has appeared in a paper of Hollmann et al.[15], in Blackburn [7] and in [23].

Proposition 1 *There is an* $\mathsf{SHF}(3; 3v^2, v^2 + 2v, \{1,1,1,1\})$ *for any integer* $v \geq 1$*.*

Observe that the construction above still leaves room for slight improvement depending on the values of m. For instance, assume $m = v^2$. Then we have $m_2 = v$, $m_1 = v^2 - 2v$. If we choose $r = v - 1$, $\delta = 2$ and $c = v - 2$, then one symbol in V_2 and one in V_3 are not used in the construction. These two free symbols are then used to form an $\mathsf{SHF}(3; 4, 2, \{1,1,1,1\})$. In this way we can construct 4 more columns. Hence we have the following.

Proposition 2 *There is an* $\mathsf{SHF}(3; 3(v^2 - 2v) + 4, v^2, \{1,1,1,1\})$ *for any integer* $v \geq 2$*.*

6.2 Optimal $\mathsf{SHF}(1 + w; n, m, \{1, w\})$ and $\mathsf{SHF}(N; n, m, \{1, 2\})$

The following results show that the bounds in Theorems 7 and 11 are tight.

Theorem 13 *For any prime power* m *and any integer* w *with* $w + 1 \leq m$*, there is an optimal* $\mathsf{SHF}(w + 1; m^2, m, \{1, w\})$*.*

Proof. Let m be a prime power such that $w+1 \leq m$. Let $\mathcal{R} \subseteq \mathbb{F}_m$ with $|\mathcal{R}| = w+1$. Consider the classical orthogonal array $\mathsf{OA}(2, |\mathcal{R}|, m)$ which is an $(w + 1) \times m^2$ array A. Now any two different columns of A agree in at most one row. It follows that for given two disjoint subsets of columns C_1 and C_2 of A with $|\mathsf{C}_1| = 1$ and $|\mathsf{C}_2| = w$, there is at least one row that separates C_1 and C_2. Hence A is an optimal $\mathsf{SHF}(w + 1; m^2, m, \{1, w\})$ according to Theorem 7.

For an arbitrary integer m we provide a further direct construction of optimal SHF for Theorem 7 from mutually orthogonal Latin squares (MOLS). A *Latin square of order* m is an $m \times m$ array consisting of elements of an $m-$set, say S, with the property that each row and each column of the array is a permutation of S. Two $m \times m$ Latin squares are *orthogonal* if no ordered pair occurs more than once when they are superimposed. A set of $t \geq 2$ Latin squares is said to be *mutually orthogonal*, or a set of MOLS, if any two of t squares are orthogonal. Let $\{L_i : 1 \leq i \leq s\}$ be a set of s MOLS on symbols $\{0, 1, \ldots, m - 1\}$. Form an $(s + 2) \times m^2$ array $\mathsf{A} = (a_{ij})$ whose columns are $(i, j, L_1(i, j), L_2(i, j), \ldots, L_s(i, j))^T$ for $0 \leq i, j < m$. Then A is an orthogonal array, $\mathsf{OA}(2, s+2, m)$. Now any two columns of A agree in at most one row, therefore A forms an $\mathsf{SHF}(s+2; m^2, m, \{1, s+1\})$ which is optimal by Theorem 7 when $s + 2 \leq m$. We have the following.

Theorem 14 *Suppose there are* $w - 1$ MOLS *of order* m *with* $w + 1 \leq m$*. Then there exists an optimal* $\mathsf{SHF}(w + 1; m^2, m, \{1, w\})$*.*

Example 1 Let us consider several small values for m that are not prime powers. It is well-known that there are at least two MOLS of order 10, five MOLS of order 12, three MOLS of order 14 and four MOLS of order 15, see for instance, [14]. Hence Theorem 14 provides the following optimal separating hash families:
$\mathsf{SHF}(w+1; 10^2, 10, \{1, w\})$ for $w = 2, 3$,
$\mathsf{SHF}(w+1; 12^2, 12, \{1, w\})$ for $w = 2, 3, 4, 5, 6$,
$\mathsf{SHF}(w+1; 14^2, 14, \{1, w\})$ for $w = 2, 3, 4$,
$\mathsf{SHF}(w+1; 15^2, 15, \{1, w\})$ for $w = 2, 3, 4, 5$.

By [14] (3.81 Table, page 175) it is known there are at least 6 MOLS of order m for all $m \geq 75$. Thus we have the following theorem.

Theorem 15 *For any integer* $m \geq 75$ *there is an optimal* $\mathsf{SHF}(w+1; m^2, m, \{1, w\})$ *for* $w = 2, 3, 4, 5, 6, 7$.

For type $\{1, 2\}$ we have the following results.

Theorem 16 *If* m *is a prime power, then there is an optimal* $\mathsf{SHF}(2d+1; m^{d+1}, m, \{1, 2\})$ *with* $2d+1 \leq m+1$.

Proof. Let A be a classical $\mathsf{OA}(d+1, 2d+1, m)$. So, A is a $(2d+1) \times m^{d+1}$ array with entries from $\mathbb{F}_m$ and any two columns of A agree in at most d rows. Therefore A is an $\mathsf{SHF}(2d+1; m^{d+1}, m, \{1, 2\})$. This separating hash family achieves the bound of Theorem 11 and is therefore optimal. □

Theorem 16 requires that m is a prime power, however if $d = 1$, we can remove this restriction.

Theorem 17 *For any integer* $m \geq 2$*, there is an optimal* $\mathsf{SHF}(3; m^2, m, \{1, 2\})$.

Proof. It is well-known that an $\mathsf{OA}(2, 3, m)$ exists for any $m \geq 2$. An easy construction of such an OA is the zero sum construction: taking all triples $[a, b, c] \in \mathbb{Z}_m^3$ with $a + b + c = 0$ in $\mathbb{Z}_m$ as columns of the array. This orthogonal array is also an $\mathsf{SHF}(3; m^2, m, \{1, 2\})$. □

For any integer $m \geq 2$ we have the following result.

Theorem 18 *Let* $m = p_1^{e_1} p_2^{e_2} \dots p_s^{e_s}$ *be a prime power factorization of an integer* $m \geq 2$ *such that* $p_1^{e_1} < p_2^{e_2} < \dots < p_s^{e_s}$. *Then there exists an optimal* $\mathsf{SHF}(2d+1; m^{d+1}, m, \{1, 2\})$ *for any positive integer* d *with* $2d \leq p_1^{e_1}$.

Proof. It is known by a result of Bush (see [11] or [14], 7.20 Theorem, page 226) that there is an $\mathsf{OA}(d+1, k, m)$ for $d+1 < p_1^{e_1}$ and $k \leq p_1^{e_1} + 1$. If we choose $k = 2d+1$, then an $\mathsf{OA}(d+1, 2d+1, m)$ provides an optimal $\mathsf{SHF}(2d+1; m^{d+1}, m, \{1, 2\})$. □

7 Concluding remarks

We have presented new bounds for SHF of general type and specific types. As well, we showed further improved bounds for the case $N = u$. We presented constructions showing that several bounds of specific types are optimal or asymptotically optimal. There are hints indicating that if $N = u - 1$, the leading constant γ in a bound of the form $n \leq \gamma(m - c)$, where c is a constant, cannot be smaller than $(u - 1)$ for almost all values of m. Our bounds for SHF of type $\{1, w\}$ when $m \geq u$ and $N = u$, as well as for SHF of type $\{1, 2\}$ and arbitrary odd N are optimal. For $N = u$ it remains a challenging problem to find optimal bounds on n for type $\{w_1, w_2\}$ with $w_1, w_2 \geq 2$.

References

[1] N. Alon, Explicit construction of exponential sized families of k-independent sets, *Discrete Math.* **58** (1986), 191–193.

[2] M. Atici, S. S. Magliveras, D. R. Stinson, and W.-D. Wei, Some recursive constructions for perfect hash families, *J. Combin. Des.* 4 (1996), 353–363.

[3] Marjan Bazrafshan and Tran van Trung, Bounds for separating hash families, *J. Combin. Theory Ser. A* 118 (2011), 1129–1135.

[4] Marjan Bazrafshan, Separating Hash Families, PhD thesis, University of Duisburg-Essen, 2011.

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

[6] S. R. Blackburn, Perfect hash families: probabilistic methods and explicit constructions, *J. Combin. Theory Ser. A* 92 (2000), 54–60.

[7] S. R. Blackburn, Perfect hash families with few functions, Unpublished manuscript, 2000.

[8] S. R. Blackburn, Frameproof codes, *SIAM J. Discrete Math.* 16 (2003), 499–510.

[9] S. R. Blackburn and P. R. Wild, Optimal linear perfect hash families, *J. Combin. Theory Ser. A* 83 (1998), 1897-1905.

[10] D. Boneh, J. Shaw, Collusion-free fingerprinting for digital data, *IEEE Trans. Inform. Theory* 44 (1998), 1897–1905.

[11] K. A. Bush, A generalization of a theorem due to MacNeish, *Ann. Math. Stat.* 23 (1952) 293–295.

[12] K. A. Bush, Orthogonal arrays of index unity, *Ann. Math. Stat.* 23 (1952) 426–434.

[13] G. D. Cohen, S. B. Encheva and H. G. Schaathun, On separating codes, Technical report 2001D003 (2001), TELECOM ParisTech - Ecole Nationale Superieure des Telecommunications.

[14] C. J. Colbourn and J. H. Dinitz, editors. *The CRC Handbook of Combinatorial Designs* Chapman and Hall/CRC, Boca Raton, FL, 2nd edition, 2007.

[15] H. D. L. Hollmann, J. H. van Lint, J.-P. Linnartz and L. M. G. M. Tolhuizen, On codes with the identifiable parent property, *J. Combin. Theory Ser. A* 82 (1998), 121–133.

[16] P. C. Li, R. Wei and G. H. J. van Rees, Constructions of 2-cover-free families and related separating hash families, *J. Combin. Des.* 14 (2006), 423–440.

[17] S. S. Martirosyan, Tran van Trung, Explicit constructions for perfect hash families, *Des. Codes Cryptogr.* 46 (2008), 97–112.

[18] K. Mehlhorn, *Data Structures and Algorithms 1: Sorting and Searching,* Springer-Verlag, Berlin, 1984.

[19] J. N. Staddon, D. R. Stinson and R. Wei, Combinatorial properties of frameproof and traceability codes, *IEEE Transaction on Information Theory* 47 (2001), 1042-1049.

[20] D. R. Stinson, R. Wei and L. Zhu, New constructions for perfect hash families and related structures using combinatorial designs and codes, *J. Combin. Des.* 8 (2000), 189–200.

[21] D. R. Stinson, Tran van Trung and R. Wei, Secure frameproof codes, key distribution patterns, group testing algorithms and related structures, *J. Statist. Plann. Inference* 86 (2000), 595–617.

[22] D. R. Stinson, G. M. Zaverucha, Some improved bounds for secure frameproof codes and related separating hash families, *IEEE Transaction on Information Theory* 54 (2008), 2508–2514.

[23] D. R. Stinson, R. Wei and K. Chen, On Generalized Separating Hash Families, *J. Combin. Theory Ser. A* 115 (2008), 105-120.

[24] Tran van Trung, S. S. Martirosyan, New constructions for IPP codes, *Des. Codes Cryptogr.* 35 (2005), 227–239.

[25] R. A. Walker II and C. J. Colbourn, Perfect hash families: Constructions and Existence, *J. Math. Crypt.* 1 (2007), 125–150.