A bound 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

This paper aims to present new upper bounds on the size of separating hash families. These bounds improve previously known bounds for separating hash families.

Key words. Separating hash family, perfect hash family, frameproof code, w-IPP code.

1 Introduction

Let h be a function from a set A to a set B and let $C_1, C_2, \ldots, C_t \subseteq A$ 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 $|A| = n$ and $|B| = m$. We call a set $\mathcal{H}$ of N functions from A to B an $(N; n, m)$-*hash family*. We say that $\mathcal{H}$ is *an* $(N; n, m, \{w_1, w_2, \ldots, w_t\})$ *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 $C_1, C_2, \ldots, C_t \subseteq A$ with $|C_i| = w_i$, for $i = 1, 2, \ldots, t$, 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. Obviously, we have $2 \leq t \leq m$ and $\sum_{i=1}^{t} w_i \leq n$. Separating hash family with $t = 2$ was introduced in [13] and the general case in [16]. It is worth remarking that various well-known combinatorial objects may be viewed as special cases of separating hash families. For example, if $w_1 = w_2 = \ldots = w_t = 1$, 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, [1, 3, 5, 9, 10, 12, 18]. A *w-frameproof code* is a separating hash family of type $\{1, w\}$ [6, 11, 4] and a *w-secure frameproof code* is a separating hash family of type $\{w, w\}$ [13] . Further, a *w-IPP code* (code with identifiable parent property) [7, 11, 17], is necessarily a PHF with $t = w + 1$ and an SHF of type $\{w, w\}$.

An $\mathsf{SHF}(N; n, m, \{w_1, w_2 \ldots, w_t\})$ can be depicted as an $N \times n$ array $\mathcal{A}$ in which the columns are labeled by the elements of A, 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 $\mathcal{A}$

with $|C_i| = w_i$, for $i = 1, 2, \ldots, t$, there exists at least one row r of $\mathcal{A}$ such that

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

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

In general, for given N, m, $\{w_1, w_2, \ldots, w_t\}$ we want to maximize n. The determination of bounds for n has been subject of much research recently [2, 8, 11, 14, 15, 16].

The best known upper bounds on n for separating hash families of type $\{w_1, w_2\}$ are the following.

Theorem 1 ([5],[11]) *Suppose there exists an* $\mathsf{SHF}(N; n, m, \{1, w\})$ *with* $w \geq 2$. *Then* $n \leq w(m^{\lceil \frac{N}{w} \rceil} - 1)$.

Theorem 2 ([16]) *Suppose there is an* $\mathsf{SHF}(N; n, m, \{2, 2\})$. *Then* $n \leq 4m^{\lceil \frac{N}{3} \rceil} - 3$.

For the special case $\{w_1, w_2, w_3\} = \{1, 1, 2\}$ we have the following strong bound.

Theorem 3 ([16]) *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 general bound for SHF of type $\{w_1, \ldots, w_t\}$) has been obtained by Stinson and Zaverucha in [14]. In [2] Blackburn, Etzion, Stinson and Zaverucha introduce a new method to establish a significant bound for SHF of type $\{w_1, \ldots, w_t\}$, which considerably improves the bound in [14], when $w_i \geq 2$ for all $i = 1, \ldots, t$. We record this bound for SHF of type $\{w_1, \ldots, w_t\}$) in the following theorem.

Theorem 4 ([2]) *Suppose an* $\mathsf{SHF}(N; n, m, \{w_1, \ldots, w_t\})$ *exists. Let* $u = \sum_{i=1}^{t} w_i$. *Then*

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

where $\gamma = (w_1 w_2 + u - w_1 - w_2)$, *and* w_1 *and* w_2 *are the smallest two of the integers* w_i.

Note that the constant γ in Theorem 4 depends on $w_1, w_2, \ldots, w_t$. If we take $\gamma = \binom{u}{2}$ for the theorem, we obtain a bound derived from the graph theoretical method [2], and if we take $\gamma = 2(u - w_1)w_1 - w_1$, where w_1 is the smallest of the integers w_i, we have the bound in [14].

It should be noted that there exist further bounds for type $\{w_1, w_2\}$ and for general type $\{w_1, w_2, \ldots, w_t\}$ [14, 15]. However as those bounds have been improved by the bound of Theorem 4, they are not included here.

To date, Theorem 4 presents the best known bound for SHF of general type $\{w_1, \ldots, w_t\}$.

In this paper we present new strong bounds for SHF which improve the Blackburn-Etzion-Stinson-Zaverucha bound of Theorem 4.

2 A bound for SHF of type $\{w_1, \ldots, w_t\}$

We aim to prove the following results.

Theorem 5 *Suppose there exists an* $\mathsf{SHF}(N; n, m, \{w_1, w_2\})$*. Let* $u = w_1 + w_2$*. Then*

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

Theorem 6 *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*

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

Theorem 5 is an immediate consequence of the subsequent Lemma 1 and Theorem 7. And Theorem 6 is derived from Lemma 1 and Theorem 8.

We first include a basic but useful lemma that can be found, for example, in [2].

Lemma 1 *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. □

2.1 A bound for $\mathsf{SHF}(u-1; n, m, \{w_1, w_2\})$

We begin with a lemma that is necessary to the proof of Theorem 7.

Lemma 2 *Suppose there exists an* $\mathsf{SHF}(N; n, m, \{w_1, w_2\})$ *with* $n - m \geq w_1 + w_2 - 1$ *and* $w_2 \geq 2$*. Then there exists an* $\mathsf{SHF}(N-1; n_1, m, \{w_1, w_2 - 1\})$ *with* $n_1 \geq n - m$*.*

Proof. Let $\mathcal{A}$ be the matrix representation of an $\mathsf{SHF}(N; n, m, \{w_1, w_2\})$ with $w_2 \geq 2$. Let m_1 denote number of symbols that appear in the first row of $\mathcal{A}$. Since permuting the columns of $\mathcal{A}$ does not change the separation property, we may assume that the first row of $\mathcal{A}$ has pairwise different symbols in the first m_1 columns. Let $\mathcal{A}_1$ denote the $(N-1) \times (n - m_1)$ matrix obtained from $\mathcal{A}$ by ignoring the first row and the first m_1 columns of $\mathcal{A}$. Set $n_1 := n - m_1$. Then $n_1 \geq n - m \geq w_1 + w_2 - 1$. We claim that $\mathcal{A}_1$ is an $\mathsf{SHF}(N-1; n_1, m, \{w_1, w_2 - 1\})$. Assume that $\mathcal{A}_1$ is not an $\mathsf{SHF}(N-1; n_1, m, \{w_1, w_2 - 1\})$. Then there are two column sets $\mathcal{C}_1$ and $\mathcal{C}_2$ with $|\mathcal{C}_1| = w_1$ and $|\mathcal{C}_2| = w_2 - 1$, that are not separated in any row of $\mathcal{A}_1$. Let a be a symbol appearing in some column of $\mathcal{C}_1$ in the first row of $\mathcal{A}$. Then in the first m_1 columns of $\mathcal{A}$ there is a column c having symbol a in the first row. Add this column c to $\mathcal{C}_2$. Now it is easily checked that $\mathcal{C}_1$ and $\mathcal{C}_2 \cup \{c\}$ are not separated in $\mathcal{A}$, which contradicts the separation property of $\mathcal{A}$. $\square$

Theorem 7 *Suppose there exists an* $\mathsf{SHF}(u-1; n, m, \{w_1, w_2\})$*, where* $u = w_1 + w_2$*. Then* $n \leq (u-1)m$*.*

Proof. We prove the theorem by induction on u. Note that $u \geq 2$. Let $\mathcal{A}$ be the matrix representation of an $\mathsf{SHF}(u-1; n, m, \{w_1, w_2\})$. Assume $u = 2$. Then $w_1 = w_2 = 1$ and $\mathcal{A}$ is an $1 \times n$ matrix. Hence, all n symbols in the unique row of $\mathcal{A}$ must be pairwise different, i.e. $n \leq m$. Now assume, as an inductive hypothesis, that the statement $n \leq (u-1)m$ is valid for all $u = 2, \ldots, k-1$, with $k-1 \geq 2$. Suppose now that there exists an $\mathsf{SHF}(k-1; n, m, \{w_1, w_2\})$ such that $n > (k-1)m$, where $k = w_1 + w_2$. As $k \geq 3$, we may assume $w_2 \geq 2$. From $m \geq 2$ and $n - m > (k-2)m$ we have $n - m > k - 1$, therefore $n - m > w_1 + w_2 - 1$. By Lemma 2 there exists an $\mathsf{SHF}(k-2; n_1, m, \{w_1, w_2 - 1\})$ with $n_1 \geq n - m > (k-2)m$, which contradicts the assumption of the induction. This completes the proof. $\square$

Using Lemma 1 and Theorem 7 we obtain Theorem 5.

Proof. [*of Theorem 5*] Assume, by contradiction, that there exists an $\mathsf{SHF}(N; n, m, \{w_1, w_2\})$ with $n = (u-1)m^{\lceil \frac{N}{(u-1)} \rceil} + 1$. By Lemma 1 there exists an $\mathsf{SHF}(\lceil \frac{N}{c} \rceil; n, m^c, \{w_1, w_2\})$ with $c := \lceil \frac{N}{(u-1)} \rceil$. We make use of a simple observation. Suppose there exists an $\mathsf{SHF}(N; n, m, \{w_1, w_2, \ldots, w_t\})$ with matrix representation $\mathcal{A}$. Then for any $N' > N$ there exists an $\mathsf{SHF}(N'; n, m, \{w_1, w_2, \ldots, w_t\})$ obtained by adding $N' - N$ arbitrary new rows using the same symbol set to $\mathcal{A}$. Now, as $\lceil \frac{N}{c} \rceil \leq u-1$, the observation says that there is an $\mathsf{SHF}(u-1; n, m^c, \{w_1, w_2\})$ with $n = (u-1)m^{\lceil \frac{N}{(u-1)} \rceil} + 1$, which contradicts Theorem 7. $\square$

2.2 A bound for $\mathsf{SHF}(u-1; n, m, \{w_1, \ldots, w_t\})$ with $t \geq 3$

In this section we first prove a new bound for SHF with $u-1$ rows for the general type $\{w_1, w_2, \ldots, w_t\}$ with $t \geq 3$. This bound is slighly stronger than the bound of Theorem 7. Observe that any $\mathsf{SHF}(N; n, m, \{w_1, w_2, \ldots, w_t\})$ with $t \geq 3$ yields an $\mathsf{SHF}(N; n, m, \{w_1, w_2, w_3'\})$ where $w_3' = w_3 + \ldots + w_t$. So, the proof of Theorem 8 can be reduced to the case of $\mathsf{SHF}(u-1; n, m, \{w_1, w_2, w_3\})$. However, as the proof uses a new idea and is constructive, we think it would be useful to present it for the general type $\{w_1, w_2, \ldots, w_t\}$.

Theorem 8 *Let* $t \geq 3$ *be an integer. Suppose there exists an* $\mathsf{SHF}(u-1; n, m, \{w_1, w_2, \ldots, w_t\})$*, where* $u = \sum_{i=1}^{t} w_i$*. Then* $n \leq (u-1)(m-1) + 1$*.*

Proof. Assume, for a contradiction, that there exists an $\mathsf{SHF}(u-1; n, m, \{w_1, w_2, \ldots, w_t\})$ with $n = (u-1)(m-1) + 2$. Wlog we assume that w_1 and w_2 are the smallest two of the integers $w_1, w_2, \ldots, w_t$. Let $\mathcal{A} = (a_{i,j})$ be its matrix representation and let $\mathcal{C}$ denote the set of columns of $\mathcal{A}$. The proof describes a procedure how to construct disjoint subsets $C_1, C_2, \ldots, C_t \subseteq \mathcal{C}$ with $|C_i| \leq w_i$ that are not separated by any row of $\mathcal{A}$. We begin with a simple counting of the number of columns having at least one unique symbol in some row $i \in \{2, \ldots, u-1\}$. Since each row can have at most $(m-1)$ unique symbols (if there were m unique symbols, we would only have m columns), there are at most $(u-2)(m-1)$ such columns. Let $\mathcal{C}_1$ denote this set of columns. Define $\mathcal{C}_2 := \mathcal{C} \setminus \mathcal{C}_1$. Then $|\mathcal{C}_2| \geq m+1$. The set $\mathcal{C}_2$ has the following property: for each column $j \in \mathcal{C}_2$ and for each row $i \in \{2, \ldots, u-1\}$ the symbol $a_{i,j}$ appears in row i at least twice. As $|\mathcal{C}_2| \geq m+1$, it follows that there are two columns $j_1, j_2 \in \mathcal{C}_2$ having the same symbol in the first row and having non-unique symbols in all other rows.

We now describe how to construct the subsets $C_1, \ldots, C_t$ of $\mathcal{C}$ we are seeking. We start with $C_i = \emptyset$ for $i = 1, \ldots, t$ and then construct C_i's using the following four steps.

Step 1: Add j_1 to C_1 and j_2 to C_2. We will focus on the specified columns j_1 and j_2 in the following steps to construct $C_1, C_2, C_3, \ldots, C_t$.

Step 2: This step starts building sets C_i for $i = 3, \ldots, t$.

Consider all the rows $k = 2, \ldots, u - w_1 - w_2 + 1$ of $\mathcal{A}$. For each such row k, the symbol a_{k,j_2} appears in at least one more column, say j, other than j_2 (i.e. $j \neq j_2$).

(i) If $j \in \bigcup_{i=3}^{t} C_i \cup C_1$, then do nothing.

(ii) If $j \notin \bigcup_{i=3}^{t} C_i \cup C_1$ and if $|C_i| < w_i$ for some $i = 3, \ldots, t$, then add column j to set C_i.

We eventually obtain subsets $C_3, \ldots, C_t$ with $|C_i| \leq w_i$ that are not separated from column j_2 in any row $k = 2, \ldots, u - w_1 - w_2 + 1$. Note that after Step 2 all sets $C_3, \ldots, C_t$ could remain empty, this would be the case if column j is unique and $j = j_1$ for all k.

Step 3: This step continues to construct the sets $C_3, \ldots, C_t$ as long as it is still possible, otherwise it constructs the set C_2.

Consider all the rows $k = u - w_1 - w_2 + 2, \ldots, u - w_1$ (i.e. $w_2 - 1$ rows). In each row k there exists a column j with $j \neq j_1$ such that $a_{k,j} = a_{k,j_1}$ (as the symbol a_{k,j_1} is repeated).

(i) If column $j \in \bigcup_{i=3}^{t} C_i$, then do nothing.

(ii) If column $j \notin \bigcup_{i=3}^{t} C_i \cup C_2$ and if $\sum_{i=3}^{t} |C_i| < w_3 + \ldots + w_t$, then add j to one of C_i with $|C_i| < w_i$, $i \geq 3$.

(iii) If column $j \notin \bigcup_{i=3}^{t} C_i \cup C_2$ and if $\sum_{i=3}^{t} |C_i| = w_3 + \ldots + w_t$, then add j to C_2.

(iv) If column $j \in C_2$, then do nothing.

Note that before Step 3 we have $C_2 = \{j_2\}$. In Step 3 for each of $w_2 - 1$ considered rows we add at most one column to C_2. So we have $|C_2| \leq w_2$ after Step 3.

The process in Step 3 is characterized by the following property: By finishing Step 3, if $|C_2| \geq 2$, then $\sum_{i=3}^{t} |C_i| = w_3 + \ldots + w_t$ (i.e. $|C_i| = w_i$ for all $i = 3, \ldots, t$).

It is clear that $C_1, C_2, C_3, \ldots, C_t$ are not separated in any row $k = u - w_1 - w_2 + 2, \ldots, u - w_1$.

Define a set D_2 as follows: D_2 is the set of columns j obtained from (i) and (ii) of Step 3 after it is finished. Note here that $D_2 \cup C_2$ is the set of columns that are responsible for the non-separation of C_1 from $C_2, C_3, \ldots, C_t$ in the rows $k = u - w_1 - w_2 + 2, \ldots, u - w_1$. Define $D_1 := \bigcup_{i=3}^{t} C_i \setminus D_2$.

Step 4: This step essentially deals with the extension of C_1 by using rows $k = u - w_1 + 1, \ldots, u - 1$. A crucial point of this step is that we might need to modify the so far constructed sets $C_2, C_3, \ldots, C_t$. To make the description clearer we consider two cases.

Case A: $|C_2| = 1$ (i.e. $C_2 = \{j_2\}$).

For each $k = u - w_1 + 1, \ldots, u - 1$, there exists a column $j \neq j_2$ such that $a_{k,j} = a_{k,j_2}$, as the symbol a_{k,j_2} is repeated.

(a) If $j \in \bigcup_{i=3}^{t} C_i$, do nothing.
(b) If $j \notin \bigcup_{i=3}^{t} C_i$, add j to C_1

It can be checked that the constructed $C_1, C_2, C_3, \ldots, C_t$ are not separated in any row $k = u - w_1 + 1, \ldots, u - 1$.

Case B: $|C_2| \geq 2$.

Suppose $|C_2| := \alpha \geq 2$. As just described in Step 3 this case implies that $|C_i| = w_i$ for all $i = 3, \ldots, t$. Moreover, we have $\bigcup_{i=3}^{t} C_i = D_1 \cup D_2$ as defined in Step 3.

Since $\alpha - 1$ columns are added to C_2 in Step 3, we have $|D_2| = w_2 - 1 - (\alpha - 1) = w_2 - \alpha$. Further, as

$$w_2 \leq w_3 \leq |\bigcup_{i=3}^{t} C_i| = w_3 + \ldots + w_t = |D_1| + |D_2| = |D_1| + w_2 - \alpha,$$

we have

$$|D_1| \geq \alpha.$$

We now use this fact to construct C_1 or possibly to modify the so far constructed $C_2, C_3, \ldots, C_t$.

For each row $k = u - w_1 + 1, \ldots, u - 1$, there exists a column $j \neq j_2$ such that $a_{k,j} = a_{k,j_2}$, as the symbol a_{k,j_2} is repeated.

(i) If $j \in \bigcup_{i=3}^{t} C_i$, do nothing.
(ii) If $j \notin \bigcup_{i=3}^{t} C_i \cup C_2$, add j to C_1.
(iii) If $j \in C_2$ (i.e. cases (i) and (ii) do not happen), then we do the following operation: Move one column $j' \in D_1$ to C_1 and substitute j' with j. We observe that this step can always be done, as $|D_1| \geq \alpha$. Note that the size of C_2 is reduced by one each time this operation is applied.

Note also that before Step 4 we have $C_1 = \{j_1\}$. In Step 4 for each of $w_1 - 1$ considered rows we add at most one column to C_1. Hence, $|C_1| \leq w_1$ after Step 4.

Now it is not difficult to check that the constructed column subsets $C_1, C_2, C_3, \ldots, C_t$ cannot be separated by any row of $\mathcal{A}$. This can be seen as follows. After Steps 1,2,3 the so far constructed $C_1, C_2, C_3, \ldots, C_t$ are not separated by any of the first $(u - w_1)$ rows of $\mathcal{A}$, (i.e. rows $k = 1, \ldots, u - w_1$). The key observation being that any operation in Step 4, namely adding a new column to C_1 or moving one column from D_1 to C_1 and replace it by a column from C_2, does not change the non-separation property of the newly constructed sets $C_1, C_2, C_3, \ldots, C_t$ in rows $k = 1, \ldots, u - w_1$. Moreover, the construction in Step 4 makes clear that the column sets $C_1, C_2, C_3, \ldots, C_t$ are not separated by any of the last $(w_1 - 1)$ rows, i.e. rows $k = u - w_1 + 1, \ldots, u - 1$. This completes the proof. □

Now using Lemma 1 and Theorem 8 we obtain Theorem 6 by a similar argumentation as given in the proof for Theorem 5 above.

3 Discussion

The new bounds in Theorem 5 and Theorem 6 improve the Blackburn-Etzion-Stinson-Zaverucha bound for any type $\{w_1, \ldots, w_t\}$ with $w_i \geq 2$ for all i. For example, when $t = 2$ and $w_1 = w_2 = w \geq 2$, the bound in Theorem 5 provides $n \leq (2w-1)m^{\lceil \frac{N}{(u-1)} \rceil}$, whereas the bound in Theorem 4 gives $n \leq (w^2)m^{\lceil \frac{N}{(u-1)} \rceil}$. From observing the constant $(u-1)$ in Theorem 7 and Theorem 8, an interesting question arises:

Question *Is there any type $\{w_1, w_2, \ldots, w_t\}$ for which the constant $(u-1)$ in Theorem 7 or Theorem 8 can be replaced by another constant c strictly smaller than $(u-1)$?*

For certain types we know the answer to the question. For instance, there are constructions for $\mathsf{SHF}(3; n, m, \{2, 2\})$, for which $\lim_{m\to\infty} n/m = 3$, see for example [7]. This implies that $u - 1 = 3$ is the smallest value γ such that $n \leq \gamma m$ for all m. Another example is an $\mathsf{SHF}(2; n, m, \{1, 1, 1\})$. Such an SHF is, in fact, a perfect hash family $\mathsf{PHF}(2; n, m, 3)$ for which a result in [9, 18] shows that $n \leq 2m - 2$ and there exists a $\mathsf{PHF}(2; 2(m-1), m, 3)$ for very m. This again shows that $u-1 = 2$ cannot be further improved. Although it is not known whether the leading constant $u-1$ in Theorem 7 or Theorem 8 can be improved, it is expected that the bounds in these theorems may further be improved when all $w_i \geq 2$. For example we have proved that $n < 3m - 6$ for any $\mathsf{SHF}(3; n, m, \{2, 2\})$ with $m > 7$, despite the fact that the leading constant 3 cannot be improved for every m.

Acknowledgments

The authors would like to thank the anonymous referees for useful comments to improve the paper.

References

[1] 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.

[2] 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.

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

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

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

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

[7] 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.

[8] 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.

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

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

[11] 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.

[12] 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.

[13] 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.

[14] D. R. Stinson, G. M. Zaverucha, New bounds for generalized separating hash families, *CACR Technical Report 2007-21*, University of Waterloo.

[15] 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.

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

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

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