

Some Improved Bounds for Secure Frameproof Codes and Related Separating Hash Families

Douglas R. Stinson* and Gregory M. Zaverucha†
David R. Cheriton School of Computer Science
University of Waterloo
Waterloo ON, N2L 3G1, Canada
{dstinson,gzaveruc}@uwaterloo.ca

February 15, 2007

Abstract

We present some improved bounds on necessary conditions for separating hash families of type $\{w, w\}$ and type $\{w, w - 1\}$. In particular, these bounds apply to secure frameproof codes, which are equivalent to separating hash families of type $\{w, w\}$. We also consider existence results for separating hash families of type $\{w_1, w_2\}$ that can be obtained from the probabilistic method. The asymptotic behaviour of these bounds is analyzed.

1 Introduction

1.1 Separating Hash Families

Informally, a *hash family* is a set of functions from a set Y to a set X . If, for any two inputs, at least one of the functions in the family gives different outputs, then we say the family is a *separating hash family* of type $\{1, 1\}$. Instead of two inputs, we can also consider two sets of inputs and ask whether the corresponding output sets are disjoint. We can also generalize to multiple input sets. The following definitions make this idea precise.

Definition 1.1. *Let X and Y be sets with $|X| = m$ and $|Y| = n$. An $(N; n, m)$ -hash family is a set \mathcal{F} of N functions from Y to X .*

Definition 1.2. *A hash family \mathcal{F} is called a separating hash family of type $\{w_1, w_2, \dots, w_t\}$ if it satisfies the following property: For any disjoint subsets*

*Research supported by NSERC grant 203114-06

†Research supported by an NSERC PGS-D postgraduate scholarship

C_1, \dots, C_t of Y with $|C_i| = w_i$ for $1 \leq i \leq t$, there exists at least one function $f \in \mathcal{F}$ such that

$$\{f(a) : a \in C_i\} \cap \{f(b) : b \in C_j\} = \emptyset$$

for every $i \neq j$. We will use the notation $SHF(N; n, m, \{w_1, \dots, w_t\})$ to denote such a hash family.

While Y can be any n -set, for simplicity we often choose $Y = \{1 \dots n\}$. The *matrix representation* of an $(N; n, m)$ -hash family is the $N \times n$ matrix A where $A_{i,j} = f_i(j)$ for all i, j . A represents an $SHF(N; n, m, \{w_1, \dots, w_t\})$ provided that the following condition is satisfied: for disjoint sets of column indices C_1, \dots, C_t where $|C_i| = w_i$ for $1 \leq i \leq t$, there exists a row r such that

$$\{A(r, a) : a \in C_i\} \cap \{A(r, b) : b \in C_j\} = \emptyset$$

for all $i \neq j$.

A third representation is also useful at times. Let A be the matrix representation of an $(N; n, m)$ -hash family. A can also be viewed as a code of size n and length N over an alphabet Y of cardinality m , where each column of A is a codeword.

The SHF studied in this paper are of types $\{w, w\}$ and $\{w, w-1\}$. SHF of type $\{w, w\}$ are of cryptographic importance for the purposes of fingerprinting digital data; in this context, they are equivalent to the codes called *secure frame-proof codes*. These connections, which are discussed in [7, 6], are briefly reviewed now. We consider the matrix representation of an $SHF(N; n, m, \{w, w\})$ and interpret the columns as codewords, as described above.

In general, suppose we have a code \mathbb{C} of length N on an alphabet Q with $|Q| = q$. Then $\mathbb{C} \subseteq Q^N$ and we will call it an (N, n, q) -code if $|\mathbb{C}| = n$. The N -tuples in \mathbb{C} are called *codewords*; each codeword has the form $x = (x_1, \dots, x_N)$, where $x_i \in Q$, $1 \leq i \leq N$.

For any subset of codewords $\mathbb{C}_0 \subseteq \mathbb{C}$, we define the set of *descendants* of \mathbb{C}_0 , denoted $\text{desc}(\mathbb{C}_0)$ by

$$\text{desc}(\mathbb{C}_0) = \{x \in Q^N : x_i \in \{a_i : a \in \mathbb{C}_0\}, 1 \leq i \leq N\}.$$

When the codewords represent fingerprints embedded in digital data, the set $\text{desc}(\mathbb{C}_0)$ consists of the N -tuples that could be produced by a coalition holding the codewords in the set \mathbb{C}_0 .

Now, let $w \geq 2$ be a positive integer. For a code \mathbb{C} , define the w -*descendant code* of \mathbb{C} , denoted $\text{desc}_w(\mathbb{C})$, as follows:

$$\text{desc}_w(\mathbb{C}) = \bigcup_{\mathbb{C}_0 \subseteq \mathbb{C}, |\mathbb{C}_0| \leq w} \text{desc}(\mathbb{C}_0).$$

Let $\mathbb{C}_i \subseteq \mathbb{C}$, $i = 1, 2, \dots, t$, be all the subsets of \mathbb{C} such that $|\mathbb{C}_i| \leq w$. (Hence $t = \sum_{j=1}^w \binom{n}{j}$.) Then we say that \mathbb{C} is a w -SFP (or w -*secure-frameproof code*) provided that for all $x \in \text{desc}_w(\mathbb{C})$, $x \in \text{desc}(\mathbb{C}_i) \cap \text{desc}(\mathbb{C}_j)$ implies that $\mathbb{C}_i \cap \mathbb{C}_j \neq \emptyset$, where $i \neq j$.

Basically, a code is w -secure frameproof if no coalition of size at most w can frame a disjoint coalition of size at most w by producing an N -tuple that could have been produced by the second coalition.

The connection between secure frameproof codes and separating hash families was shown in [7, 6].

Theorem 1.3. *Suppose that \mathcal{F} is an $(N; n, q)$ -hash family and suppose $n \geq 2w$. Then the associated (N, n, q) -code, \mathbb{C} , is a w -secure frameproof code if and only if \mathcal{F} is an SHF($N; n, q, \{w, w\}$).*

There have been several recent papers giving constructions for 2-secure frameproof code; see [3, 10, 11]. Explicit constructions for w -secure frameproof codes for arbitrary $w \geq 2$ are found in [7, 9, 5]. Existence results using the probabilistic method are provided in [7]; we will revisit these results in Section 3. Finally, a necessary condition for the existence of w -secure frameproof codes was proven in [6] (we will give a significant improvement of this result in Section 2).

1.2 Previous Results

In this section, we summarize some previous results about SHF in general and about the classes of SHF we study in this paper.

The following two well-known lemmas are elementary.

Lemma 1.4. *Suppose there exists an SHF($N; n, m, \{w_1, w_2, \dots, w_i\}$), and let $c \geq 2$ be an integer. Then there exists an SHF($\lceil \frac{N}{c} \rceil; n, m^c, \{w_1, w_2, \dots, w_i\}$).*

Lemma 1.5. *Suppose A is an SHF($N; n, m, \{w_1, w_2, \dots, w_t\}$), and let $w'_1 \leq w_1$. Then A is also an SHF($N; n, m, \{w'_1, w_2, \dots, w_t\}$).*

The next lemma was first noticed by Hollman et al. [4] and it provides a key idea for many of the proofs in this paper. When discussing matrix representations of hash families, the term “isomorphic” should be interpreted as “isomorphic up to permutation of rows and columns”.

Lemma 1.6. [4] *Let A be the matrix representation of an SHF($3; n, m, \{2, 2\}$). Then there is no submatrix of A isomorphic to the matrix*

$$\begin{array}{|c|c|c|c|} \hline a & a & * & * \\ \hline * & b & b & * \\ \hline * & * & c & c \\ \hline \end{array} .$$

Separating hash families of small types were studied by Stinson, Wei and Chen [8]. In this paper, we will generalize the strategy they used to derive necessary conditions for SHF of type $\{2, 2\}$. They first proved an upper bound on n when $N = 3$ (see Theorem 1.7). Then, using Lemma 1.4, the bound is extended to yield a bound for all N (Corollary 1.8).

Theorem 1.7. [8] *If an SHF($3; n, m, \{2, 2\}$) exists, then $n \leq 4m - 3$.*

Theorem 1.8. [8] In an SHF($N; n, m, \{2, 2\}$), $n \leq m^{\lceil \frac{N}{3} \rceil} - 3$.

The only known general bound for SHF of type $\{w, w\}$ is stated in Theorem 1.9.

Theorem 1.9. [6] Suppose there exists an SHF($N; n, m, \{w, w\}$). Then $n \leq m^{\lceil \frac{N}{w} \rceil} + 2w - 2$.

Observe that Theorem 1.7 provides a considerable improvement to Theorem 1.9 in the case $w = 2$. In Section 2, we improve the bound of Theorem 1.9 first for SHF of type $\{3, 2\}$, and then for types $\{w, w\}$ and $\{w, w - 1\}$.

2 SHF of Type $\{w, w\}$ and $\{w, w - 1\}$

2.1 Bounds For Type $\{3, 2\}$

To illustrate the technique we use, we first prove a result for SHF of type $\{3, 2\}$. We begin by looking at the special case $N = 4$.

Theorem 2.1. Suppose $m \geq 2$. If an SHF($4; n, m, \{3, 2\}$) exists, then $n \leq 7m - 6$.

Proof. Suppose A is an SHF($4; n, m, \{3, 2\}$) with $n = 7m - 5$ (then we will obtain a contradiction).

We construct a submatrix A_1 of A where all elements appearing in the first row appear at least 4 times. To do this, we delete $t_1 + t_2 + t_3$ columns from A ; namely, those in which elements appear exactly once, twice and 3 times (resp.). If $t_1 + t_2 + t_3 = m$, then A would have $3m$ columns, which is fewer than $7m - 5$ because $m \geq 2$. Thus we delete no more than $3(m - 1)$ columns, leaving A_1 with at least $7m - 5 - 3(m - 1) = 4m - 2$ columns.

Now we create A_2 from A_1 in such a way that elements in the second row of A_2 repeat three or more times. To do this, we delete $u_1 + u_2$ columns of A_1 ; namely, those with an element appearing exactly once or twice (resp.). If $u_1 + u_2 = m$, A_1 would have $2m$ columns, which is fewer than $4m - 2$, so we have that $u_1 + u_2 \leq 2(m - 1)$. We can then say that A_2 has at least $4m - 2 - 2(m - 1) = 2m$ columns.

Our final submatrix will be A_3 , where, in the third row, all elements will appear at least twice. Since A_2 has at least $2m$ columns, the number of elements appearing exactly once must be less than m . Then A_3 will have at least $2m - (m - 1) = m + 1$ columns.

Since A_3 has at least $m + 1$ columns, the fourth row of A_3 contains an element which occurs at least twice. Also, in A_3 every element in the third row occurs at least twice, so A_3 has a submatrix isomorphic to either

$$\text{case 1: } \begin{array}{|c|c|} \hline * & * \\ \hline * & * \\ \hline c & c \\ \hline d & d \\ \hline \end{array} \quad \text{or case 2: } \begin{array}{|c|c|c|} \hline * & * & * \\ \hline * & * & * \\ \hline c & c & * \\ \hline * & d & d \\ \hline \end{array} .$$

In case 1, all elements in the second row of A_2 appear at least three times. Furthermore, all elements in A_1 appear at least four times in the first row of A_1 . Therefore, A has a submatrix isomorphic to

$$\begin{array}{|c|c|c|c|} \hline a & a & * & * \\ \hline * & b & b & * \\ \hline * & * & c & c \\ \hline * & * & d & d \\ \hline \end{array}, \quad (1)$$

In case 2, when we consider the second row (which has elements repeating 3 times) there are two possibilities. The first is

$$\begin{array}{|c|c|c|c|} \hline * & * & * & * \\ \hline * & b & b & b \\ \hline * & c & c & * \\ \hline * & * & d & d \\ \hline \end{array}$$

which leads to a submatrix of A having the form

$$\begin{array}{|c|c|c|c|} \hline a & a & * & * \\ \hline * & b & b & b \\ \hline * & c & c & * \\ \hline * & * & d & d \\ \hline \end{array}, \quad (2)$$

This is easily seen because the elements in the first row of A_1 occur at least four times.

The second possibility for case 2 is

$$\begin{array}{|c|c|c|c|} \hline * & * & * & * \\ \hline b & b & * & * \\ \hline * & c & c & * \\ \hline * & * & d & d \\ \hline \end{array}.$$

After considering the first row (as a submatrix of A), there are again two possibilities that arise:

$$\begin{array}{|c|c|c|c|} \hline a & a & a & a \\ \hline b & b & * & * \\ \hline * & c & c & * \\ \hline * & * & d & d \\ \hline \end{array} \quad (3)$$

and

$$\begin{array}{|c|c|c|c|c|} \hline a & a & * & * & * \\ \hline * & b & b & * & * \\ \hline * & * & c & c & * \\ \hline * & * & * & d & d \\ \hline \end{array}, \quad (4)$$

depending on whether an a appears in columns three and four of the previous submatrix.

It is easy to see that all the possible submatrices (1), (2), (3) and (4), lead to a contradiction, because the odd columns cannot be separated from the even

columns. If (4) is a submatrix, then A is not an SHF of type $\{3, 2\}$. If (1), (2) or (3) are submatrices, then A is not an SHF of type $\{2, 2\}$. Applying Lemma 1.5, we see that A is not an SHF of type $\{3, 2\}$ in these cases, as well. In every case, we have shown that A is not an SHF of type $\{3, 2\}$, which completes the proof. \square

Corollary 2.2. *In an SHF($N; n, m, \{3, 2\}$), $n \leq 7m^{\lceil \frac{N}{4} \rceil} - 6$.*

Proof. We proceed by contradiction. Suppose there exists an SHF($N; n, m, \{3, 2\}$) where $n = 7m^{\lceil \frac{N}{4} \rceil} - 5$. Then, by Lemma 1.4 with $c = \lceil \frac{N}{4} \rceil$, there exists an SHF($4; 7m' - 5, m', \{3, 2\}$) with $m' = m^{\lceil \frac{N}{4} \rceil}$, which contradicts the previous theorem. \square

2.2 Staircases

The following definition and lemmas will be useful tools for generalizing the proof technique used in Theorem 2.1.

Definition 2.3. *An (N, t) -staircase is a matrix S with N rows of the form:*

x_1	x_1							
	x_2	x_2						
		\ddots	\ddots					
			x_{t-1}	x_{t-1}				
			x_t	x_t				
				x_{t+1}	x_{t+1}			
					\ddots	\ddots		
						x_{N-1}	x_{N-1}	
							x_N	x_N

Further,

- (i) when $t = 1$, S is a regular staircase, having $N + 1$ columns,
- (ii) when $1 < t \leq N$, S is a compressed staircase having N columns.

Example 2.4. *The first of the three matrices below is a $(4, 1)$ -staircase, the second is a $(4, 4)$ -staircase and the third is a $(4, 3)$ -staircase.*

a	a	*	*	*
*	b	b	*	*
*	*	c	c	*
*	*	*	d	d

a	a	*	*
*	b	b	*
*	*	c	c
*	*	d	d

a	a	*	*
*	b	b	*
*	c	c	*
*	*	d	d

Lemma 2.5. *Let A be a matrix with N rows. If A has submatrices $A_{N-1} \subset \dots \subset A_1 \subset A$ such that*

- *For $1 \leq i < N - 1$, all elements appearing in row i of A_i appear at least $N - i + 1$ times in row i of A_i , and*
- *The last row of A_{N-1} contains at least one element that occurs at least twice.*

Then A has a submatrix isomorphic to an (N, t) -staircase, for some $t, 1 \leq t \leq n$.

Proof. In this proof, we number the columns in increasing order from right to left. First, permute the columns of A_{N-1} so that there is a repeated element of row N appearing in columns 1 and 2 (call this element x). These two cells form the base of the staircase. We will create the staircase step by step, extending it upwards when moving from a submatrix of A_i to A_{i-1} , for $i = N-1, N-2, \dots, 1$. In row $N - 1$ of A_{N-1} , every element occurs at least twice. Letting y be the element above x in column 2, we have two possibilities (possibly after permuting columns):

$$\text{case 1: } \begin{array}{|c|c|} \hline y & y \\ \hline x & x \\ \hline \end{array} \quad \text{or case 2: } \begin{array}{|c|c|c|} \hline y & y & * \\ \hline * & x & x \\ \hline \end{array}.$$

In the first case, all elements in row $N - 2$ of the submatrix A_{N-2} occur at least three times. Our staircase is only two columns wide, so we extend it up and to the left by one step. We can continue extending leftward by one column as we consider each successive submatrix. The result is a (compressed) (N, N) -staircase.

In the second case, we consider the element in row $N - 2$ and column 3, say z . We know that z occurs at least three times in row $N - 2$ of the submatrix A_{N-2} . There are two subcases: either z occurs in columns 1, 2 and 3, or z occurs in column 3 and some new column which we can name column 4. These two subcases are analogous to subcases considered in the proof of Theorem 2.1.

In the second subcase, we have extended the staircase up and to the left by one step. In the first subcase, the staircase does not extend leftward at this stage.

This process can be continued until we reach the top of the staircase. If at any stage, we do not extend the staircase leftward, then it must be the case that all further submatrices cause the staircase to extend up and to the left. So we end up with a (N, t) -staircase, for some $t, 1 \leq t \leq n$. □

The usefulness of staircases is established in the following lemma.

Lemma 2.6. *Suppose A is an $(N; n, m)$ matrix having a submatrix isomorphic to an (N, t) -staircase S . Then A is not an SHF of type $\{h, h\}$ if $N = 2h - 1$, and A is not an SHF of type $\{h, h - 1\}$ if $N = 2h - 2$.*

Proof. When A has a submatrix isomorphic to an (N, t) -staircase, the odd and even indexed columns of the staircase are inseparable. We consider two cases, depending on the parity of N .

If $N = 2h - 1$, then S has $2h$ columns if S is regular and $2h - 1$ columns if S is compressed. If S is regular, then A is not an SHF of type $\{h, h\}$. If S is compressed, then A is not an SHF of type $\{h, h - 1\}$. By Lemma 1.5, A is not an SHF of type $\{h, h\}$.

The case when $N = 2h - 2$ is similar. S has $2h - 1$ columns if S is regular and $2h - 2$ columns if S is compressed. If S is regular, then A is not an SHF of type $\{h, h - 1\}$. If S is compressed, then A is not an SHF of type $\{h - 1, h - 1\}$. By Lemma 1.5, A is not an SHF of type $\{h, h - 1\}$. \square

2.3 The General Case

In this section, we extend the results of previous section.

Theorem 2.7. *If A is an SHF($2w - 1, n, m, \{w, w\}$), then*

$$n \leq m + (w - 1)(2w - 1)(m - 1).$$

Proof. Let $N = 2w - 1$, and suppose A had one extra column, i.e., suppose A is an SHF($N, n, m, \{w, w\}$) with $n = m + 1 + \frac{N(N-1)}{2}(m - 1)$ columns. We will derive a contradiction by showing A has a submatrix isomorphic to an (N, t) -staircase.

We will create a series of submatrices $A_{N-1} \subset \dots \subset A_1 \subset A$. Submatrix A_i has the property that elements in the i -th row repeat $N - i + 1$ or more times (in the i th row). Denote $A_0 = A$. The construction of A_{i+1} from A_i deletes all columns of A_i where elements in the $(i + 1)$ -st row appear fewer than $N - i - 1$ times. Let $|A_i|$ denote the number of columns in A_i . By repeatedly applying this construction, we claim that

$$|A_i| \geq m + 1 + \frac{1}{2}(N - i - 1)(N - i)(m - 1) \quad (5)$$

for $0 \leq i \leq N - 1$. We will prove that (5) holds by induction on i .

First note that (5) holds for A_0 . To construct A_1 from A_0 , we must delete the columns containing elements which appear fewer than N times in the first row of A_0 . Let t_q be the number of columns with elements appearing exactly q times in the first row of A_0 . We claim that $t_1 + \dots + t_{N-1} \leq m - 1$. If $t_1 + \dots + t_{N-1} = m$, then A_0 has $(N - 1)m$ columns, fewer than the number given by (5). Then we delete at most $(N - 1)(m - 1)$ columns, so

$$\begin{aligned} |A_1| &\geq |A_0| - (N - 1)(m - 1) \\ &\geq m + 1 + \frac{N(N - 1)}{2}(m - 1) - (N - 1)(m - 1) \\ &= m + 1 + \left(\frac{N(N - 1) - 2(N - 1)}{2} \right) (m - 1) \\ &= m + 1 + \frac{(N - 2)(N - 1)}{2}(m - 1), \end{aligned}$$

which shows that (5) holds for A_1 as well.

Suppose (5) holds up to A_i , suppose that $t_1 + \dots + t_{N-i-1} \leq (m-1)$, and suppose that elements in the i -th row of A_i appear at least $N-i$ times. To create A_{i+1} from A_i , we delete columns of A_i where elements in the $(i+1)$ -th row appear less than $N-(i+1) = N-i-1$ times. If $t_1 + \dots + t_{N-i-2} = m$, then A_i would have only $m(N-i-1)$ columns, fewer than the number assumed in the inductive hypothesis. Then

$$\begin{aligned}
|A_{i+1}| &\geq |A_i| - (N-(i+1))(m-1) \\
&= m+1 + \frac{(N-i-1)(N-i)}{2}(m-1) - (N-i-1)(m-1) \\
&= m+1 + \left(\frac{(N-i-1)(N-i) - 2(N-i-1)}{2} \right) (m-1) \\
&= m+1 + \left(\frac{(N-i-1)(N-i-2)}{2} \right) (m-1) \\
&= m+1 + \frac{(N-(i+1)-1)(N-(i+1))}{2}(m-1)
\end{aligned}$$

which proves (5) by induction.

The last submatrix is A_{N-1} , with

$$|A_{N-1}| \geq m+1 + \frac{(N-(N-1)-1)(N-(N-1))}{2}(m-1) = m+1.$$

Since A_{N-1} has at least $m+1$ columns, one of the elements in row N is must repeat. By applying Lemmas 2.5 and 2.6, we see that A is not an SHF, contradicting the assumption that A was an $SHF(2w-1, 4wm-4w, m, \{w, w\})$. This proves the theorem. \square

Corollary 2.8. *If an $SHF(N; n, m, \{w, w\})$ exists, it holds that*

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

Proof. We proceed by contradiction (analogous to Corollary 2.2). Suppose there exists an $SHF(N; n, m, \{w, w\})$ where $n = m+1 + (w-1)(2w-1)(m-1)$. Then by Lemma 1.4 with $c = \lceil \frac{N}{2w-1} \rceil$, there exists an $SHF(2w-1; m' + (w-1)(2w-1)(m'-1), m', \{w, w\})$ with $m' = m^{\lceil \frac{N}{2w-1} \rceil}$, which contradicts the previous theorem. \square

Theorem 2.9. *If an $SHF(2w-2; n, m, \{w, w-1\})$ exists, then*

$$n \leq m + (w-1)\left(w - \frac{3}{2}\right)(m-1).$$

Proof. Omitted. Follows the proof of Theorem 2.7 closely, with $N = 2w-2$ instead of $N = 2w-1$. \square

The proof of the next corollary is basically the same as the proof of Corollary 2.8.

Corollary 2.10. *If an SHF($N; n, m, \{w, w-1\}$) exists, it holds that*

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

3 Existence Results

In this section, we consider existence results obtained using the probabilistic method, and we analyze the asymptotic behaviour of the bounds obtained. First, we begin by stating a special case of a general bound proven in [8].

Theorem 3.1. *[8, Theorem 4.1] Let w_1 and w_2 be positive integers. Define*

$$c_T = \begin{cases} \frac{1}{w_1!w_2!} & \text{if } w_1 \neq w_2 \\ \frac{1}{2(w_1!)^2} & \text{if } w_1 = w_2. \end{cases} \quad (6)$$

Also, define

$$p_T = 1 - \frac{\chi_{w_1, w_2}(m)}{m^{w_1+w_2}}, \quad (7)$$

where $\chi_{w_1, w_2}(m)$ is the chromatic polynomial of the complete bipartite graph K_{w_1, w_2} . Then there exists an SHF($N; n, m, \{w_1, w_2\}$) provided that $n \leq B(w_1, w_2)$, where

$$B(w_1, w_2) = (1 - c_T) \left(\frac{1}{p_T} \right)^{\frac{N}{w_1+w_2-1}}. \quad (8)$$

We are interested in determining the asymptotic behaviour of the bound $B(w_1, w_2)$. In order to do this, we need to analyze the quantity $1/p_T$. This can be done by using properties of the chromatic polynomials $\chi_{w_1, w_2}(m)$. The following formula will be useful.

Theorem 3.2. *[2, Proposition 4.4] Let w_1 and w_2 be positive integers. Then*

$$\chi_{w_1, w_2}(m) = \sum_{i=1}^{w_1} S(w_1, i) m(m-1) \cdots (m-i+1)(m-i)^{w_2}, \quad (9)$$

where $S(h, i)$ denotes a Stirling number of the second kind.

The Stirling numbers of the second kind can be computed using the following well-known formula:

$$S(h, i) = \frac{1}{i!} \sum_{j=1}^i (-1)^{i-j} \binom{i}{j} j^h. \quad (10)$$

Now, the following facts about Stirling numbers can easily be verified from equation (10).

Lemma 3.3. *For all positive integers i , it holds that $S(i, i) = 1$ and $S(i+1, i) = (i+1)i/2$.*

We can now prove a useful result about the chromatic polynomials $\chi_{w_1, w_2}(m)$.

Theorem 3.4. *The coefficient of $m^{w_1+w_2-1}$ in the chromatic polynomial $\chi_{w_1, w_2}(m)$ is equal to $-w_1 w_2$.*

Proof. There are only two terms in the expansion of the formula (9) that contain monomials of degree $w_1 + w_2 - 1$, namely, the terms corresponding to $i = w_1$ and $i = w_1 - 1$. These terms are (respectively) as follows:

$$S(w_1, w_1)m(m-1)\cdots(m-w_1+1)(m-w_1)^{w_2} \quad (11)$$

and

$$S(w_1, w_1 - 1)m(m-1)\cdots(m-w_1+2)(m-w_1+1)^{w_2}. \quad (12)$$

Using (11) and (12), it is easy to compute the coefficient of $m^{w_1+w_2-1}$. From (11), we compute the coefficient of the term m^{w_1-1} in the expansion of $m(m-1)\cdots(m-w_1+1)$ and multiply it by the coefficient of the term m^{w_2} in the expansion of $(m-w_1)^{w_2}$. Then we add the coefficient of the term m^{w_1} in the expansion of $m(m-1)\cdots(m-w_1+1)$ multiplied by the coefficient of the term m^{w_2-1} in the expansion of $(m-w_1)^{w_2}$. Finally, we multiply this sum by $S(w_1, w_1) = 1$. This provides a contribution of

$$-\frac{w_1(w_1-1)}{2} - w_1 w_2. \quad (13)$$

Next, using (12), we compute the product of $S(w_1, w_1 - 1) = w_1(w_1 - 1)/2$, the coefficient of the term m^{w_1-1} in the expansion of $m(m-1)\cdots(m-w_1+2)$, and the coefficient of the term m^{w_2} in the expansion of $(m-w_1+1)^{w_2}$. This provides a contribution of

$$\frac{w_1(w_1-1)}{2}. \quad (14)$$

Adding (13) and (14), we see that the coefficient of the term $m^{w_1+w_2-1}$ is equal to $-w_1 w_2$. \square

The proof of the following lemma is an immediate consequence of Theorem 3.4.

Lemma 3.5. *For positive integers w_1 and w_2 , define p_T as in (7). Then*

$$\lim_{m \rightarrow \infty} p_T = \frac{m}{w_1 w_2}.$$

Now, the asymptotic behaviour of $B(w_1, w_2)$ can be stated.

Corollary 3.6. *For positive integers w_1 and w_2 , it holds that*

$$B(w_1, w_2) = (1 - c_T) \left(\frac{m}{w_1 w_2} + o(1) \right)^{\frac{N}{w_1 + w_2 - 1}},$$

where c_T is the constant defined in (6).

When $(w_1, w_2) = (w, w)$ or $(w, w - 1)$, the bound $B(w_1, w_2)$, which guarantees existence of the relevant SHF, has the same exponent as the necessary conditions proven in Section 2.3. In this sense our bounds can be considered to be relatively tight.

4 Summary

We have described a general method of obtaining necessary conditions for separating hash families of types $\{w, w\}$ and $\{w, w - 1\}$. These bounds are relatively close to the existence results that can be obtained using the probabilistic method. An interesting open problem would be to find good necessary conditions for separating hash families of types $\{w_1, w_2\}$ when $|w_1 - w_2| > 1$.

References

- [1] D. Boneh, and J. Shaw. Collusion-secure fingerprinting for digital data. *IEEE Transactions on Information Theory*, **44** (1998), 1897–1905.
- [2] R. Ehrenborg and S. van Willigenburg. Enumerative properties of Ferrers graphs. *Discrete and Computational Geometry*, **32** (2004), 481–492.
- [3] S. Encheva and G. Cohen. Partially identifying codes for copyright protection. *Lecture Notes in Computer Science*, **2227** (2001), 260–267 (AAECC-14).
- [4] H. Hollmann, J. van Lint, J. Linnartz, and L. Tolhuizen. On codes with the identifiable parent property. *Journal of Combinatorial Theory A*, **82** (1998), 121–133.
- [5] L. Liu and H. Shen. Explicit constructions of separating hash families from algebraic curves over finite fields. *Designs, Codes and Cryptography*, **41** (2006), 221–233.
- [6] J.N. Staddon, D.R. Stinson, and R. Wei. Combinatorial properties of frameproof and traceability codes. *IEEE Transactions on Information Theory*, **47** (2001), 1042–1049.
- [7] 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), 595–617.
- [8] D.R. Stinson, R. Wei, and K. Chen. On generalized separating hash families. Preprint. Available online: <http://www.cacr.math.uwaterloo.ca/~dstinson/pubs.html>.
- [9] 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), 189–200.
- [10] V.D. Tô, R. Safavi-Naini and Y. Wang. A 2-secure code with efficient tracing algorithm. *Lecture Notes in Computer Science*, **2551** (2002), 149–162 (INDOCRYPT 2002).

- [11] D. Tonien and R. Safavi-Naini. Explicit construction of secure frameproof codes. *International Journal of Pure and Applied Mathematics*, **6** (2003), 343–360.