Splitting Systems and Separating Systems

Alan C. H. Ling[*]
Department of Computer Science
University of Vermont
Burlington, Vermont
U.S.A 05405
aling@emba.uvm.edu

P. C. Li[†] G. H. J. van Rees[‡]
Department of Computer Science
University of Manitoba
Winnipeg, Manitoba
Canada R3T 2N2
lipakc@cs.umanitoba.ca vanrees@cs.umanitoba.ca

March 26, 2003

Abstract

Suppose m and t are integers such that $0 < t \leq m$. An (m, t) splitting system is a pair $(X, \mathcal{B})$ where $|X| = m$, $\mathcal{B}$ is a set of $\lfloor \frac{m}{2} \rfloor$ subsets of X, called blocks such that for every $Y \subseteq X$ and $|Y| = t$, there exists a block $B \in \mathcal{B}$ such that $|B \cap Y| = \lfloor \frac{t}{2} \rfloor$ or $|(X \setminus B) \cap Y| = \lfloor \frac{t}{2} \rfloor$. We will give some results on splitting systems for $t = 2$ or 4 which often depend on results from uniform separating systems. Suppose that m is an even integer, t_1, t_2 are integers such that $t_1 + t_2 \leq m$. A

[*] research supported by ARO grant 19-01-1-0406 and a DOE grant
[†] research supported by URGP grant 431-1725-60
[‡] research supported by NSERC grant OGP# 0003558

uniform (m, t_1, t_2)*-separating system* is an ordered pair $(X, \mathcal{B})$ where $|X| = m$, $\mathcal{B}$ is a set of subsets of X of size $\frac{m}{2}$, called blocks, such that for every $P \subseteq X, Q \subseteq X$ where $|P| = t_1, |Q| = t_2$ and $P \cap Q = \emptyset$, there exists a block $B \in \mathcal{B}$ for which either $P \subseteq B$, $Q \cap B = \emptyset$ or $Q \subseteq B$, $P \cap B = \emptyset$. We also give new results for separating systems.

1 Introduction

Recently, splitting systems were used by Stinson [5] in baby-step giant-step algorithms for the low hamming weight discrete logarithm problem. The smaller the splitting system the better the algorithms are. Stinson gave only a few constructions for large systems. We find the problem of determining a splitting system with the fewest blocks interesting in its own right. So we plan to study splitting systems in this paper. First we define them.

Definition 1.1 *Suppose* m *and* t *are integers where* $0 < t \leq m$. *An* (m, t) *splitting system is a pair* $(X, \mathcal{B})$ *where* $|X| = m$, $\mathcal{B}$ *is a set of subsets of size* $\lfloor \frac{m}{2} \rfloor$ *of* X, *called blocks such that for every* $Y \subseteq X$ *with* $|Y| = t$, *there exists a block* $B \in \mathcal{B}$ *such that* $|B \cap Y| = \lfloor \frac{t}{2} \rfloor$ *or* $|(X \setminus B) \cap Y| = \lfloor \frac{t}{2} \rfloor$. *We say that* B *t-splits* Y. *We will also say that* $(X, \mathcal{B})$ *is a t-splitting system.*

We will let $\mathrm{SS}(N; m, t)$ denote an (m, t) splitting system with N blocks. We will let $S(m, t)$ denote the minimum number of blocks over all (m, t) splitting systems. We say that an $\mathrm{SS}(N; m, t)$ is *optimal* if $N = S(m, t)$. We limit our investigations to m even and $t = 2, 4$. Clearly, under these conditions, any block of the splitting set may be replaced by its complement in X.

The following Lemma is due to Coppersmith as cited in Stinson [5]. We include the proof for completeness.

Lemma 1.2 *For all even integers* m *and* t *with* $0 < t \leq m$, *there exists an* $SS(\frac{m}{2}; m, t)$.

Proof: Let $X = \mathbb{Z}_m$ and define $B_i = \{i + j \bmod m : 0 \leq j \leq \frac{m}{2} - 1\}$ for $i \in \mathbb{Z}_m$. Let $\mathcal{B} = \{B_i : 0 \leq i \leq \frac{m}{2} - 1\}$. We will show that $(X, \mathcal{B})$ is an (m, t) splitting system.

Fix any subset $Y \subseteq X$ such that $|Y| = t$. For $i \in \mathbb{Z}_m$, define $v(i) = |B_i \cap Y| - |(\mathbb{Z}_m \setminus B_i) \cap Y|$. If $v(0) = 0$, then we are done, so assume that $v(0) \neq 0$. It is easy to show that $v(i)$ is even for all i, $v(m/2) = -v(0)$, and $|v(i+1) - v(i)| \in \{-2, 0, 2\}$ for all i. Therefore there is some integer i such that $0 < i < m/2$ and $v(i) = 0$.□

We give some simple examples of splitting systems now. The last three examples are optimal.

Example 1.3 *The blocks* $\{1,2\}$*,* $\{1,3\}$ *and* $\{2,3\}$ *form a* $(4,2)$ *splitting system.*

Example 1.4 *The blocks* $\{1,2,3\}$*,* $\{1,2,6\}$*, and* $\{1,3,5\}$ *form a* $(6,4)$ *splitting system.*

Example 1.5 *The blocks* $\{1,2,3,4\}$*,* $\{1,2,5,6\}$*,* $\{1,3,5,7\}$ *form an* $(8,2)$ *splitting system but not a* $(8,4)$ *splitting system.*

Example 1.6 *The blocks* $\{1,2,3,4\}$*,* $\{1,2,5,6\}$*, and* $\{1,2,7,8\}$ *form an* $(8,4)$ *splitting system but not an* $(8,2)$ *splitting system.*

Most constructions of splitting systems, including the next one, use the incidence matrix of the splitting system so we define that now.

Definition 1.7 *The incidence matrix,* M*, of an* (m,t) *splitting system on* b *blocks is an* $m \times b$ *array where* $M(i,j) = 1$ *if element* i *occurs in block* B_j *and is 0 otherwise.*

The exact bound for $S(m,2)$, where m is even is given by the following result that was proven by Rényi [4]. We follow the proof in Freidman et al. [2].

Theorem 1.8 $S(m,2) = \lceil \log_2 m \rceil$ *for each positive even integer* m*.*

Proof Consider the incidence matrix of an $(m,2)$ splitting system. Label the rows from 0 to $m-1$. Let row i be i in binary, then clearly the rows i and j have a column where they differ and hence $\{i,j\}$ is split. If there were fewer columns, then two rows would be the same and hence $\{i,j\}$ would not be split. □

Next we prove a lower bound for $(m,4)$ splitting systems.

Theorem 1.9 $S(m,4) \geq \lceil \log_2 m \rceil$*, for even* $m > 4$*.*

Proof : Assume there exists an incidence matrix, M of an $(m, 4)$ splitting system, S, on $b < \lceil \log_2 m \rceil$ blocks. In the first $b-1$ columns of M there are only 2^{b-1} distinct binary rows so that there must be 3 identical rows, say i, j, k. To split $\{i, j, k, l\}$, column b must contain two 0's and two 1's, say $M(i, b) = M(j, b) = 1$. Since $m > 4$, there must be another 1 in column b, say in row m. Then $\{i, j, k, m\}$ is not split which is a contradiction. □.

This paper is mainly concerned with splitting systems where $t = 4$. But we found that most splitting constructions required separating systems as ingredients. So in the next section we discuss the constructions involved with separating systems. In Section 3 we construct $(m, 4)$ splitting systems using Hadamard matrices and perfect hashing functions. We conclude the paper in Section 4 with a table containing bounds on the size of optimal 4-splitting systems.

2 Constructions with separating systems

In this section, we want to use the separating systems constructions used by Freidman, et al. [2]. They did not specify the block size in their definition of separating systems but most of their recursive constructions actually produced separating systems of uniform block size given that uniform ones were put in as ingredients. We now define uniform separating systems.

Definition 2.1 *Suppose that m is an even integer, t_1, t_2 are integers such that $t_1 + t_2 \leq m$. A uniform (m, t_1, t_2)-separating system is an ordered pair $(X, \mathcal{B})$ where $|X| = m$, $\mathcal{B}$ is a set of subsets of X of size $\frac{m}{2}$, called blocks, such that for every $P \subseteq X, Q \subseteq X$ where $|P| = t_1, |Q| = t_2$ and $P \cap Q = \emptyset$, there exists a block $B \in \mathcal{B}$ for which either $P \subseteq B$, $Q \cap B = \emptyset$ or $Q \subseteq B$, $P \cap B = \emptyset$. We also say that $(X, \mathcal{B})$ is a (t_1, t_2) separating system.*

Suppose $\mathcal{B}$ is a collection of blocks $\mathcal{B}$, and that P and Q are two sets where $|P| = t_1, |Q| = t_2$ and $P \cap Q = \emptyset$. If there exists $B \in \mathcal{B}$ for which either $P \subseteq B$, $Q \cap B = \emptyset$ or $Q \subseteq B$, $P \cap B = \emptyset$, we say that P and Q are *separated* by B or $\{P; Q\}$ is separated by B. We will abuse this notation and often write P and Q as a list of elements.

Separating systems were studied extensively in the seminal paper by Freidman, Graham and Ullman [2]. Good lower bounds were established by Fredman and Komlos [1]. In more recent times, separating systems have been constructed based on techniques used to construct perfect hash families. These results can be found in Stinson, van Trung and Wei [6]. This paper is mostly interested in systems that are both (2,1) separating and 4-splitting but many of the results improve and generalize results on (2,1) separating systems.

We will denote by $\mathrm{TT}(N; m, 4)$ an $(m, 4)$ splitting system on N blocks that is also a (2,1) separating system and denote by $T(m, 4)$ the minimum N over all $\mathrm{TT}(n; m, 4)$.

The following lemmas are easy to prove.

Lemma 2.2 *If $(X, \mathcal{B})$ is a uniform (n, t_1, t_2)-separating system, then it is a uniform (n, t_2, t_1)-separating system.*

Lemma 2.3 *If $t_3 \leq t_2$ and if $(X, \mathcal{B})$ is a uniform (n, t_1, t_2) -separating system, then it is a uniform (n, t_1, t_3)-separating system.*

Lemma 2.4 *If $(X, \mathcal{B})$ is a uniform (n, t, t)-separating system, then it is an $(n, 2t)$-splitting system.*

The following two specific lemmas will prove useful for generating splitting systems.

Lemma 2.5 *If $(X, \mathcal{B})$ is a uniform $(n, 2, 2)$ separating system on b blocks, then there is an $(n, 4)$ splitting system on $b - 2$ blocks.*

Proof Consider the set $\{i, j, k, l\}$. The pairs of sets $(\{i, j\}, \{k, l\})$, $(\{i, k\}, \{j, l\}), (\{i, l\}, \{j, k\})$ must be separated in the separating system in three distinct blocks. Then if only two blocks are deleted, the system must still split $\{i, j, k, l\}$. □

Lemma 2.6 *If $(X, \mathcal{B})$ is a uniform $(n, 2, 2)$ separating system on b blocks with $n > 4$, then there is an $(n, 2, 1)$ separating system on $b - 1$ blocks.*

Proof Delete one of the blocks of the $(n, 2, 2)$ separating system. If i and j are in the deleted block and k is in the complement or if k is in the deleted block and i and j are in the complement of the deleted block, then $\{i, j; k\}$ may no longer separated by the new system. However, let l be another element in the block containing i and j, then the separating system must separate $\{i, j; k, l\}$ in some block other than the deleted block. In this block, $\{i, j; k\}$ is separated in the new system. □

Using the previous two theorems we get the following result.

Lemma 2.7 *If $(X, \mathcal{B})$ is a uniform $(n, 2, 2)$ separating system on b blocks with $n > 4$, then there is an $(n, 4)$ splitting system which is also an $(n, 2, 1)$ separating system on $b - 1$ blocks.*

We now state a simple construction of $(n, 4)$ splitting systems which are also $(n, 2, 1)$ separating systems. Basically it is Freidman's [2] Lemma 2.

Theorem 2.8 *If there exists a uniform $(n, 2, 1)$-separating system on b blocks and which is also an $(n, 4)$ splitting system and there exists an $(n, 2)$ splitting system on c blocks, then there exists a $(2n, 4)$ splitting system on b blocks and a $(2n, 4)$ splitting system that is also a $(2n, 2, 1)$ separating system on $b + c + 1$ blocks.*

Proof Let T be the incidence matrix of the $(n, 4)$ splitting system which is also an $(n, 2, 1)$ separating system on b blocks and let S be the incidence matrix of the $(n, 2)$ splitting system on c blocks. Let $\overline{S}$ be the incidence matrix S in which 0's and 1's have been interchanged. Then we claim that the following matrix R is the incidence matrix of a $(2n, 4)$ splitting set which is also a $(2n, 2, 1)$ separating set on $b+c+1$ blocks and that the leftmost b columns of R is the incidence matrix of a $(2n, 4)$ splitting system.

$$\mathrm{R} \quad = \quad \left[\begin{array}{c|c|c} T & S & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline T & \overline{S} & \begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} \end{array}\right]$$

Let the rows (i.e. elements) of R be labelled a_1 to a_n and b_1 to b_n from top to bottom. Let there be three types of columns in R. Type I columns are the first b columns from the left and type II columns are the next c columns and the type III column is the last column on the right. Let i, j, k, l be distinct integers between 1 and n, inclusive. We will now prove that every set of size 4 is split by some block/column. 4-sets of the form $\{a_i, a_j, a_k, a_l\}$ are split in the type I columns. Since a_l and b_l are identical in type I columns $\{a_i, a_j, a_k, b_l\}$ is also split in type I columns as is the set $\{a_i, a_j, b_k, b_l\}$. $\{a_i, a_j, a_k, b_i\}$ is split in type I columns as $\{a_1; a_2, a_3\}$ is (2,1) separated there and rows a_1 and b_1 are identical in those columns. Sets like $\{a_i, a_j, b_i, b_j\}$ are split in type I columns as $\{a_i, a_j\}$ is split there. Finally $\{a_i, a_j, b_i, b_k\}$ is split in type I columns as $\{a_i; a_j, a_k\}$ is separated there. So the first b columns of R form the incidence matrix of a $(2n, 4)$ splitting system.

We will now prove that R is the incidence matrix of a uniform $(2n; 2, 1)$ separating system. Sets $\{a_i, a_j, a_k\}$ and $\{a_i, a_j, b_k\}$ are (2,1) separated all 3 ways in type I columns. Set $\{a_j; a_i, b_i\}$ is separated in type I columns as $\{a_i, a_j\}$ is separated there. $\{b_i; a_i, a_j\}$ is separated in the type III columns while $\{a_i; a_j, b_i,\}$ is separated in type II columns as $\{a_i, a_j\}$ is split there and a_i and b_i have opposite values there. These are all the distinct cases. □

Corollary 2.9 *$T(2^n q, 4) \leq \frac{1}{2}n(n+1) - 1 + (n-1)\lceil \log_2 q \rceil + T(2q, 4)$ for q odd.*

Proof From Theorem 2.8, we get the following recursion: $T(2^n q, 4) \leq T(2^{n-1} q, 4) + S(2^{n-1} q, 2) + 1$ for $n \geq 2$. Since we know, by Lemma 1.8, that $S(m, 2) = \lceil \log_2 m \rceil$, we solve the recurrence to get the result. □

If $q = 1$ and we let $T(2, 4) = 1$, then we get the following result.

Corollary 2.10 *$T(2^n, 4) \leq \frac{1}{2}n(n+1)$.*

In the proof of Theorem 2.8 we needed only the type I blocks to split every 4-set in R, so we can state the following result.

Lemma 2.11 *If there exists a uniform $(n, 2, 1)$-separating system on b blocks and which is also an $(n, 4)$ splitting system, then there exists a $(2n, 4)$ splitting system on b blocks or $S(2n, 4) \leq T(n, 4)$.*

Applying Lemma 2.11 and Corollary 2.10, we get the following result.

Corollary 2.12 *$S(2^n, 4) \leq \frac{1}{2}n(n-1)$.*

One obvious question to ask is whether every $(n, 2, 1)$ separating system is also an $(n, 4)$ splitting system? The answer is no. The blocks $\{1, 2, 3, 4\}$, $\{1, 2, 3, 5\}$, $\{1, 2, 3, 6\}$, $\{1, 2, 4, 5\}$, $\{1, 2, 4, 6\}$, $\{1, 2, 4, 7\}$, $\{1, 2, 4, 8\}$, $\{1, 3, 4, 7\}$, $\{1, 3, 4, 8\}$, $\{2, 3, 4, 7\}$, $\{2, 3, 4, 8\}$ form an $(8, 2, 1)$ separating system, but is not an $(8, 4)$ splitting system, since the 4-set $\{1, 2, 3, 4\}$ is not split by any of the blocks stated.

The following theorem gives a doubling construction for splitting systems. It is the well-known doubling construction used for Hadamard designs.

Lemma 2.13 *If there exists an $(X, \mathcal{B})$ which is an $(n, 4)$ splitting system on b blocks, then there exists a $(2n, 4)$ splitting system on $2b$ blocks.*

Proof Let X be the incidence matrix of the $(n, 4)$ splitting system on b blocks. Then the following matrix R is the incidence matrix of a $(2n, 4)$ splitting system on $2b$ blocks.

$$\mathrm{R} \quad = \quad \left[\begin{array}{c|c} X & X \\ X & \overline{X} \end{array}\right]$$

It is easy to check that R is a splitting system. □

We have no good direct constructions for (2,1) separating and 4-splitting when the number of elements is 2 times a prime. So we give a recursive construction based on the set system with two fewer elements. This is a variant on the above doubling construction.

Lemma 2.14 *If, there exists an $(n, 4)$ splitting system with b blocks that is also an $(n, 2, 1)$ separating system, then there is on $2b+2$ blocks an $(n + 2, 4)$ splitting system which is also an $(n + 2, 2, 1)$ separating system.*

Proof Let S be an $(n, 4)$ splitting and $(n, 2, 1)$ separating system. Form the blocks of the new system as follows: First, to the blocks of the old system add the new element a. Second, to the complements of the old blocks add the new element a. Third, add two new blocks containing the new elements a and b and add $n/2 - 1$ old elements to the two new blocks ensuring that the old elements in these two new blocks are distinct.

We need to check that this collection of blocks form both an $(n + 2, 4)$ splitting system and an $(n + 2, 2, 1)$ separating system. Let the old elements be $1, 2, \ldots, n$. These are split and separated already. Consider the 4-set $\{i, j, k, a\}$. It is split as $\{i; j, k\}$ is separated in the old system and a occurs with every old block and every complement of an old block. $\{i; j, a\}$ is separated as $\{i; j\}$ is separated in the old design. $\{a; i, j\}$ is separated as $\{k; i, j\}$ is separated in the old system. Since a is symmetric to b we do not need to check with b alone. So consider the 4-set $\{i, j, a, b\}$. It is split as $\{i; j\}$ is separated in the old design. The set $\{a; i, b\}$ is separated as the set $\{a; b\}$ is separated and b occurs with every old element. Finally, $\{i; a, b\}$ is separated in one of the two new blocks. $\square$

Next, we generalize a theorem of Freidman et al. [2] on separating systems. Of course we want our systems to be 4-splitting as well.

Theorem 2.15 *If, on b_1 blocks, there exists an $(n_1, 4)$ splitting system that is also an $(n_1, 2, 1)$ separating system and if, on b_2 blocks, there exists an $(n_2, 4)$ splitting system that is also an $(n_2, 2, 1)$ separating system such that $pn_1 = jn_2$ for $2 < p \leq n_2$ for integers j, n_1, n_2, p , then there exists an $(pn_1, 4)$ splitting system on $b_1 + b_2$ blocks and on $b_1 + b_2(1 + \lceil \log_{n_2} j \rceil)$ blocks there is a $(pn_1, 4)$ splitting system that is also a $(pn_1, 2, 1)$ separating system.*

Proof: Let S_1 and S_2 be the incidence matrices of the systems on n_1 and n_2 elements, respectively. We will construct the incidence matrix, R of the required system. The incidence matrix will consist of 3 types of columns. The type I columns contain the matrices $S_{11}, S_{21}, \ldots, S_{n_1 1}$, where S_{i1} contains p copies of the i^{th} row of S_1. There are b_1 columns of type I. The type II columns contain j copies of the $n_2 \times b_2$ incidence matrix S_2. The other $b_2 \lceil \log_{n_2} j \rceil$ columns are formed as follows. Start with the $j \times \lceil \log_{n_2} j \rceil$ matrix M whose row m is the value m written in base n_2 with enough leading zeroes to have exactly $\lceil \log_{n_2} j \rceil$ digits. Then digit f of m is replaced by the $n_2 \times b_2$ matrix $\theta^f S_2$, where $\theta^f S_2$ is the matrix S_2 with the rows of S_2 cycled upward f times (first row cycles to the last row). The sets of n_2 rows in the type III columns will be called $N_1, N_2, \ldots, N_j$.

$$\text{Let } R \quad = \quad \begin{bmatrix} S_{11} & S_2 & N_1 \\ S_{21} & S_2 & N_2 \\ \vdots & \vdots & \vdots \\ S_{n_1 1} & S_2 & N_j \end{bmatrix}$$

If row m can be represented in base n_2 as $d_{\lceil \log_{n_2} j \rceil} \ldots d_3 d_2 d_1$ then

$$N_m = [\theta^{d_{\lceil \log_{n_2} j \rceil}} S_2] \ldots [\theta^{d_3} S_2][\theta^{d_2} S_2][\theta^{d_1} S_2]$$

Because $pn_1 = jn_2$, each block of the finished design has the same number of entries, i.e. half the entries. Note that S_{i1} has the same or fewer rows than does S_2, whereas S_2 has the same number of rows as N_i does. We now have to check that the system is 4-splitting and (2,1) separating.

We will first prove that the Type I and Type II columns of R form an incidence matrix of a 4-splitting system. Let i, j, k, l be distinct elements/rows of R. Case 1: If the 4 rows come from different S_{d1}, then they are split in Type I columns as any 4 distinct rows/elements of S_1 are split in S_1. Case 2: If i and j are the only elements from the same S_{d1}, then since $\{i; k, l\}$ is separated in S_1, the 4 elements are split in type I blocks. Case 3: Let i, j, k be in one S_{d1} and let l be in another, say S_{e1}. Since $p \leq n_2$, rows i, j, k limited to the type II columns must be distinct rows of S_2, say i', j', k'. If row l in the type II columns is identical to one of the other rows in the type II columns, say i', then since $\{i'; j', k'\}$ is separated in S_2, $\{i, j, k, l\}$ is split in type II blocks. If row l is distinct from the other rows in the type II columns, then let it be equal to row l' in S_2. Since $\{i', j', k', l'\}$ is split in S_2, then $\{i, j, k, l\}$ is split in the type II columns. Case 4: If i, j, k, l are all in the same S_{d1}, then rows i, j, k, l limited to the type II columns must again be distinct rows of of S_2, say i', j', k', l'. Since $\{i', j', k', l'\}$ are split in S_2, $\{i, j, k, l\}$ are split in type II blocks.

Secondly let us see if $\{i, j; k\}$ is separated by a column of R. There are 4 cases. Case 1: If i, j and k are in different S_{d1}, then they are separated in all 3 ways in type I columns as any 3 distinct elements of S_1 are separated in all 3 ways. Case 2: Let i and j be in one S_{d1} and k be in another, say S_{e1}. Since a (2,1) separating system is a (1,1) separating system, $\{i; k\}$ is separated in type I columns and then so must $\{i, j; k\}$ be separated in type I columns. Case 3: If i, j and k are all in the same S_{d1}, then they must be distinct rows of S_2 in the type II columns as $p \leq n_2$. Since any three distinct rows of S_2 are separated in all 3 ways then $\{i, j; k\}$ are separated in type II columns. Case 4: Let i and k be in the same S_{d1} and let j be in a different one, say S_{e1}. Let rows i, j, k in the type II columns be rows i', j', k' in S_2. If j' is not equal to i' or k' or if j' is equal to i' then the separation occurs in the type II columns. If j' is equal to k' then we must use type III columns. By the way M and then N were constructed there must be a set of b_2 columns in the type III columns where row k is in $\theta^t S_2$ and row j is something else i.e. $\theta^r S_2$ where $r \neq t$. That is, row k' is a different row from either row j' or row i' in the rows of S_2 in those b_2 columns of type III. Then as before, $\{i, j; k\}$ must be split in those b_2 type III columns. □

The previous theorem can be iterated starting with Example 1.3 which is also a (2,1) separating system to produce the following result.

Corollary 2.16 $T(2^{2^n}, 4) \leq 3^n$ *or equivalently* $T(m, 4) \leq (\log_2 m)^{1.59}$ *for* m *of the appropriate form.*

We will now show the existence of good splitting and separating systems.

Theorem 2.17 *There exists an* $(n, 4)$ *splitting system on* $b = 5.9 \log_2 n$ *blocks.*

Proof Any system of blocks can choose its blocks from a set of $\binom{n}{n/2}$ distinct blocks. So there are $\binom{n}{n/2}^b$ systems on b blocks. Since there are $\binom{4}{2}\binom{n-4}{(n-4)/2}$ blocks that split a particular 4-set, there are

$$r = \binom{n}{n/2} - \binom{4}{2}\binom{n-4}{(n-4)/2}$$

blocks that do not. So there are r^b systems that do not split a particular 4-set. Since there are $\binom{n}{4}$ such 4-sets, there are $\binom{n}{4} r^b$ systems that do not split some 4-set. Therefore the number of systems that split every 4-set is:

$$I = \binom{n}{n/2}^b - \binom{n}{4} r^b.$$

If $I > 0$, then we have a 4-splitting set on b blocks. This condition becomes:

$$1 - \binom{n}{4}\left(1 - \frac{6\binom{n-4}{(n-4)/2}}{\binom{n}{n/2}}\right)^b > 0,$$

or

$$\binom{n}{4}\left(1 - \frac{6n(n-2)}{16(n-1)(n-3)}\right)^b < 1.$$

Since

$$1 - \frac{6n(n-2)}{16(n-1)(n-3)} < 5/8 \text{ and } \binom{n}{4} < n^4,$$

there exists a splitting system on b blocks if $n^4(5/8)^b < 1$. Taking logarithms we get:

$$b > 4\frac{\log_2 n}{\log_2 (8/5)}$$

or

$$b > 5.9 \log_2 n.$$

□

So we have $S(n, 4) \leq 5.9 \log_2 n$. The next theorem tackles $T(n, 4)$ in the same way.

Theorem 2.18 $T(n, 4) \leq 9.6 \log_2 n$.

Proof We need to count the number of uniform systems that do not separate a particular 1-set from a particular 2-set. There are

$$\binom{n-3}{(n-2)/2} + \binom{n-3}{(n-4)/2} = \binom{n-2}{(n-2)/2}$$

blocks that do the separating so

$$s = \binom{n}{n/2} - \binom{n-2}{(n-2)/2}$$

blocks that do not separate them. So there are s^b systems that do not separate a particular 1-set from a particular 2-set. Since there are $3\binom{n}{3}$ such pairs of 1-sets and 2-sets, there are $3\binom{n}{3}s^b$ systems that do not separate some pair of 1-set from 2-set. So if we want a system that is 4-splitting and (2,1) separating, we need:

$$\binom{n}{n/2}^b - \binom{n}{4}r^b - 3\binom{n}{3}s^b > 0$$

which becomes

$$\binom{n}{4}\left(1 - \frac{6n(n-2)}{16(n-1)(n-3)}\right)^b + 3\binom{n}{3}\left(1 - \frac{n}{4(n-1)}\right)^b < 1.$$

So we get a 4-splitting and (2,1) separating system if:

$$n^4(5/8)^b + n^3(3/4)^b < 1$$

when $n \geq 6$, which simplifies to

$$n^4(3/4)^b < 1$$

Taking logarithms we get $b > 9.6 \log_2 n$. □

We have $\log_2 n$ as a lower bound for splitting sets. The argument used can also give us a lower bound of $\log_2 n$ for (2,1) separating systems. There is a big gap between $\log_2 n$ and either $5.9 \log_2 n$ or $9.2 \log_2 n$. Although we have some constructions, it would be useful to have concrete constructions for these systems even if they have more blocks than the existence result. So in the next section, we give different kind of results.

3 Other Constructions

In this section we describe some other constructions for splitting systems. The first one uses Hadamard Designs.

Definition 3.1 *A Hadamard matrix, H, of order n is an $n \times n$ matrix whose entries are either 1 or -1 with the property that $HH^t = nI_n$.*

Lemma 3.2 *If a Hadamard matrix of order n, $n \geq 8$, exist, then there exists an $(n, 4)$ splitting system on $n-3$ blocks, a $(2n, 4)$ splitting system on $n-3$ blocks, an $(n, 4)$ splitting system that is also an $(n, 2, 1)$ separating set on $n-2$ blocks and a uniform $(2, 2)$ separating system on $n-1$ blocks.*

Proof Standardize the Hadamard matrix, H, by multiplying each column by -1 if the entry in the first row of that column is -1. This ensures that the entries in the first row of the matrix are 1. Consider any 4 columns of H. Let x_i be the number of rows in these 4 columns that can be considered as the number i in binary where 1 is considered as 1 and -1 is considered as 0. Then, 6 equations can be written down with these variables using the fact that any two columns of the four columns must be orthogonal, i.e. the number of (1,1) and (-1,-1) pairs in a specified pair of columns must equal the number of (-1,1) and (1,-1) pairs in the specified pair of columns. This number must be $n/2$. We get the following equation.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{14} \\ x_{15} \end{bmatrix} = \begin{bmatrix} n/2 \\ n/2 \\ n/2 \\ \vdots \\ n/2 \\ n/2 \end{bmatrix}$$

Solving these equations we get the following solution:
$x_1 + x_{14} = x_2 + x_{13} = x_4 + x_{11} = x_7 + x_8$ and
$x_0 + x_{15} = x_3 + x_{12} = x_5 + x_{10} = x_6 + x_9$.

Since $x_{15} \geq 1$, we have that there is at least one occurrence where either any two columns are (1,1) and the other two columns are (-1,-1) or the other way around. If we delete the first row of the Hadamard matrix and change the -1's to 0's we get the incidence matrix of a set system. And we have just proved that it is (2,2) separating on $n-1$ blocks. Using Lemma 2.5 and Lemma 2.6, we can construct the systems on n elements and using Theorem 2.8 we can construct the system on $2n$ elements. □

This construction is very useful to give many good input set systems for the recursive constructions of Freidman et al. [2].

The next class of construction, produce systems for a large number of elements. They are based on balanced perfect hash functions(BPHF) and Linear codes. See Stinson [5] for more details and proofs of some of the theorems. We will define these objects via their incidence matrices which are oriented the opposite way of splitting and separating systems.

Definition 3.3 *A balanced perfect hash function, $(N; n, m, w)$, is an $N \times n$ matrix whose entries are from a set of m symbols such that each symbol occurs n/m times in each row and for each subset of w columns, there is at least one row such that the w columns are distinct.*

Definition 3.4 *An (N, K, D, q) linear code is a set of K vectors in $(\mathbb{F}_q^N)$ such that the sum of every two vectors is a vector and the Hamming distance between codewords (number of coordinate positions in which codewords disagree) is at least D.*

The following is well known.

Theorem 3.5 *If a linear code (N, K, D, q) code exists with $D/N > 1 - 1/\binom{w}{2}$, then there exists a BPHF$(N; K, q, w)$.*

Using Reed-Solomon Codes, we get very good large BPHF.

Theorem 3.6 *Let q be a prime power, $0 < p < q$ is an integer such that $q-1 > (p-1)\binom{w}{2}$, then there exists a BPHF$(q-1; q^p, q, w)$.*

Proof We know by MacWilliams and Sloane [3] that there is a Reed-Solomon code with the following parameters: $(q-1, q^p, q-p, q)$ where q is a prime power. Now apply Theorem 3.5. □

We can use the good BPHF to produce good splitting systems as in the next basic theorem.

Theorem 3.7 *If there exists a BPHF$(N_0; n, m, t)$ and a SS$(N_1; m, t)$, then there exists an SS$(N_0N_1; n, t)$.*

Here is the basic construction for splitting systems.

Theorem 3.8 *Let q be a prime power. If $0 < p < q$ is an integer such that $q-1 > (p-1)\binom{w}{2}$, and if there exists an SS$(N; q, w)$, then there exists a SS$(N(q-1); q^p, w)$.*

Proof Theorem 3.6 ensures that the BPHF$(q-1, q^p, q-p, q)$ exists and then Lemma 3.7 ensures the result. □

If we put $q = 8, p = 2, w = 4$ and use an SS$(3; 8, 4)$, we get the following result.

Corollary 3.9 *$S(8^2, 4) \leq 21$.*

Now suppose we have an SS$(N_0; q, 4)$ where $N_0 < N \log_2 q$ for some integer N. If q is an even prime power greater than or equal to 64, then there exists BPHF$(q-1, q^{\lceil q/6 \rceil}, q, 4)$, and hence we can construct an SS$(N_0(q-1); q^{\lceil q/6 \rceil}, 4)$. If $r = q^{\lceil q/6 \rceil}$ then the number of blocks in the system is approximately $6N \log_2 r$. If we iterate the construction, we get an SS$(N_1, n, 4)$ where N_1 is approximately $N \log_2^* n \log_2 n$ where $\log_2^* n$ is the iterated logarithm. Of course, the number of iterations is small if the number of elements in the splitting set is smaller than the number of atoms in the universe.

More concretely, let us assume that there is an SS$(q/2; q, 4)$ under the same conditions and let us do the construction just once. We get an SS$(q(q-1)/2; q^{\lceil q/6 \rceil}, 4)$ where if we let $r = q^{\lceil q/6 \rceil}$ then the number of blocks is about $(\log_q r)^2$. We can do a bit better using Corollary 2.16's result that there exists an SS$(n^{1.59}; 2^n, 4)$. With this we can 4-split $r = q^{\lceil q/6 \rceil}$ elements with $\mathrm{O}(\log_q r (\log_q \log_q r)^{1.59})$ blocks. Unfortunately this is not $\mathrm{O}(\log r)$. Nevertheless, it is the best result known for t=4.

4 Table

In this section we put our constructions together to produce bounds on the number of blocks in the minimal SS(n,4)'s and TT(n,4)'s. These small values help to give a feel for the problem. Note that $S(n,4)$ is not monotone. It would be helpful if we could get a general construction for separating systems like the Coppersmith construction for splitting systems. Finally, the tables would be improved if a better bound was found for $S(4n+2,4)$ and $T(4n+2,4)$ when $4n+2$ is twice a prime number.

The following is the meaning of the letters in the Table 1 :
b=Coppersmith Construction Lemma 1.2
c=optimal, found by computer
d=doubling construction: Theorem 2.8 or Lemma 2.11
g=generalized multiplication: Theorem 2.15
h=Hadamard matrix exists and Lemma 3.2
i=incremental construction: Theorem 2.14
We also include a column to show the parameters for the generalized multiplication construction.

References

[1] M. L. Fredman, J. Komlós. *On the size of Seperating Systems and Families of Perfect Hash Functions*, SIAM J. Alg. Disc. Meth., **5** # 1 (1984) 61-68.

[2] A. D. Friedman, R. L. Graham, J. D. Ullman. *Universal Single Transition Time Asynchronous State Assignments*, IEEE Transactions of Computers, Vol. **C-18**, 6 (1969) 541-547.

[3] F. J. MacWilliams, N. J. A. Sloane. **The Theory of Error Correcting Codes**, North Holland, New York; 1977.

[4] A. Rényi. *On random generating elements of a finite Boolean algebra*, Acta. Sci. Math. Szeged **22** (1961) 75-81.

[5] D. R. Stinson. *Some Baby-Step Giant-Step Algorithms for the Low Hamming Weight Discrete Logarithm Problem*, Mathematics of Computation, **71** (2002) 379-391.

n	$S(n,4)$	p	n_1	j	n_2	$\mathrm{T}(n,4)$
4	1^c					3^c
6	3^c					6^c
8	3^c					$6^c h$
10	5^c					8^c
12	6^c					9^g
14	7^c					20^i
16	6^c					9^g
18	9^b					18^g
20	8^d					13^d
22	11^b					28^i
24	9^d					14^d
26	13^b					30^i
28	14^b					25^d
30	12^g	5	6	5	6	18^g
32	9^d					14^d
36	12^g	3	12	9	4	18^g
40	11^g	4	10	10	4	17^g
42	21^b					36^i
44	22^b					38^d
48	12^g	3	16	12	4	18^g
50	16^g	5	10	5	10	24^g
56	12^g	7	8	7	8	18^g
60	14^g	10	6	6	10	22^g
64	12^g	8	8	8	8	18^g
70	16^g	7	10	7	10	24^g
72	15^g	9	8	6	12	24^g
80	14^g	10	8	8	10	22^g
90	16^g	9	10	9	10	24^g
96	15^g	12	8	8	12	24^g
100	16^g	10	10	10	10	24^g
112	15^g	14	8	7	16	24^g
128	15^g	16	8	8	16	24^g
256	18^g	16	16	16	16	27^g
512	23^g	16	32	32	16	
512		8	64	64	8	36^g

Table 1: Upper Bounds for $S(n,4)$ and $T(n,4)$ for some small n

[6] D. R. Stinson, T. van Trung, R. Wei. *Secure frameproof codes, key distribution patterns, group testing algorithms and related structures*, J. of Stats., Plan. and Inf. **86** (2000) 595-617.