More on (2,2)-separating systems

Gérard Cohen
ENST, 46 rue Barrault, 75634 Paris Cedex 13, France, cohen@enst.fr
Sylvia Encheva
HSH, Bjørnsonsg. 45, 5528 Haugesund, Norway, sbe@hsh.no
Hans Georg Schaathun
Dept of Informatics, UiB, HIB, N-5020, Bergen, Norway, georg@ii.uib.no

Abstract - The theory of separating systems has been applied in different areas of science and technology such as automata synthesis, technical diagnosis and authenticating ownership claims. Constructions of (2,2)-separating systems derived from error-correcting codes are given, together with bounds on their parameters based on distance considerations.

Index Terms- Separating systems, error-correcting codes, copyright protection, watermarking.

1 Introduction

The case of $(2, 2)$-separation was introduced by Sagalovich in the context of automata: two such systems transiting simultaneously from state a to a' and from b to b' respectively should be forbidden to pass through a common intermediate state. He has written a long series of papers since the sixties, a fairly recent survey can be found in [10]. States are simply binary n-tuples and only shortest paths are allowed during transitions; in other words, the only 'moves' permitted while transiting from a to a' are complementing the $d(a, a')$ bits where a and a' differ (one at a time). Clearly if the separation property holds, no two such minimal-length paths between a and a', and b and b' will intersect.

The design of self-checking asynchronous network has been a challenging problem. Friedman et al. [6] have shown that some of the unicode single-transition-time asynchronous state assignments correspond to (2,2)-separating systems.

Separating partition systems have been studied by Friedman and Komlós [7]. The authors used information theory to derive nonexistence bounds for separating partition systems in two special cases - systems of perfect hash functions and (i, j)-separating systems.

The problem of perfect hash functions was generalized by Körner and Symonyi [8]. They improve on earlier results for (i, j)-separating systems of partitions and give new treatment of the problem about the minimum number of partitions in any (i, j)-separating partition systems for a set of given size.

The digitalization of our world expanded our concept of watermark to include immaterial, digital impressions for use in authenticating ownership claims and protecting proprietary interests. A digital watermark is a signal or pattern inserted into a digital "document" (e.g. text, graphics, multimedia presentations). Digital fingerprints are characteristics of an object in a digital format that tend to distinguish it from other similar objects. Codes were introduced in [1] (see also [11]) as a method of "digital fingerprinting" which prevents a coalition of a given size from forging a copy with no member of the coalition being caught.

The outline of the paper is as follows. Definitions and basic results are presented in Section 2. We then derive sufficient conditions on the code distances to insure separation . In Section 4, we use concatenation to provide good constructions of linear separating families over small alphabets. Finally, we present asymptotic results and a table of rates.

2 Definitions and basic results

We use the notation of [11] for fingerprinting issues and of [9] for codes and Hadamard matrices.

For any positive real number x we shall denote by $\lfloor x \rfloor$ its integer part and by $\lceil x \rceil$ the smallest integer at least equal to x. A set $\Gamma \subseteq GF(q)^n$ is called an (n, M, d)-*code* if $|\Gamma| = M$ and the minimum Hamming distance between two of its elements (codewords) is d. Suppose $\mathcal{C} \subseteq \Gamma$. For any position i define the *projection* $P_i(\mathcal{C}) = \{a_i | a \in \mathcal{C}\}$, and the *feasible set* of $\mathcal{C}$ by

$$F(\mathcal{C}) = \{x \in GF(q)^n : \forall i, x_i \in P_i(\mathcal{C})\}.$$

The feasible set $F(\mathcal{C})$ represents the set of all possible n-tuples (descendants) that could be produced by the coalition $\mathcal{C}$ by comparing the codewords they jointly hold. Observe that $\mathcal{C} \subseteq F(\mathcal{C})$ for all $\mathcal{C}$, and $F(\mathcal{C}) = \mathcal{C}$ iff $|\mathcal{C}| = 1$.

We denote by $C[n, k, d_1]_q$ (or simply $C[n, k]_q$ when d_1 is irrelevant) a *linear* code of length n, dimension k over $GF(q)$ and minimum distance d_1. By m_1 we denote the maximum distance of a code. In the nonlinear case, $(n, M)_q$

is a code of length n with M codewords. The subscript is omitted in the binary case. The *rate* of C is defined as $R(C) = R = n^{-1} \log_q |C|$. We refer to [9] for all undefined notions on codes.

Definition 1 *We say that a code C is (t, t')-separating if, for any pair (T, T') of disjoint subsets of C where $|T| = t$ and $|T'| = t'$, the feasible sets are disjoint, i.e. $F(T) \cap F(T') = \emptyset$.*

In earlier works on watermarking, (t, t)-separating codes have been called t-partially identifying codes [5] or t-secure frameproof [12, 11]. The current terminology is older though [10].

Definition 2 *A (t, t')-configuration is a pair (T, T') of disjoint vector sets of sizes t and t' respectively. We say that (T, T') is separated if $F(T) \cap F(T') = \emptyset$, and otherwise it is non-separated. A (t, t')-NSC is a non-separated (t, t')-configuration.*

A code is (t, t')-separating if and only if it contains no (t, t')-NSC. Obviously, if C is (t, t')-separating, then it is also (t', t)-separating, and (t''', t'')-separating for all $t''' \leq t$ and all $t'' \leq t'$.

3 Bounds based on distances

Proposition 1 *If C is a linear, binary $(2, 2)$-separating code of dimension k, then $m_1 < n - 2(k - 2)$.*

Proof: If $k \leq 1$, the result is trivial. For $k = 2$, it only says that the all-one codeword $\mathbf{1}$ cannot be in the code C, lest $(\mathbf{0}, \mathbf{1}; \mathbf{c}, \mathbf{c} + \mathbf{1})$ form a $(2, 2)$-NSC for $\mathbf{c} \in C \backslash \{\mathbf{0}, \mathbf{1}\}$.

We then turn to the case $k \geq 3$. We shall prove that if $n - m_1 \leq 2(k - 2)$, then C cannot be $(2, 2)$-separating. Consider a codeword $\mathbf{c}$ of maximum weight. Since the code is linear, for every set of $k - 2$ coordinate positions, there exist at least three non-zero codewords which are zero on these positions, and thus at least one which is not $\mathbf{c}$. In particular, there is a non-zero codeword $\mathbf{a}$ which is zero on half the positions not in the support of $\mathbf{c}$, and one $\mathbf{b}$ which is zero on the other half. Thus $(\mathbf{0}, \mathbf{c}; \mathbf{a}, \mathbf{b})$ is a $(2, 2)$-NSC.

□

Let $\mathbf{c}', \mathbf{c}, \mathbf{a}, \mathbf{b}$ be distinct vectors such that $(\mathbf{c}', \mathbf{c}; \mathbf{a}, \mathbf{b})$ is a $(2,2)$-NSC. From this assumption, we will derive some statements on the minimum and maximum weights of any code which is not $(2,2)$-separating. This will give a sufficient condition for a code to be $(2,2)$-separating in Theorem 1. The binary linear case of Theorem 1 has previously been stated by Sagalovich [10]. We now make some preliminary observations and then prove a lemma.

Remark 1 *If $\pi : GF(q)^n \to GF(q)^n$ is an automorphism, then (T, T') is a (t,t')-NSC if and only if $(\pi(T), \pi(T'))$ is a (t,t')-NSC.*

Remark 2 *If (T, T') is a (t,t')-NSC, then so is $(T + \mathbf{c}, T' + \mathbf{c})$ for any $\mathbf{c} \in GF(q)^n$. If $T + \mathbf{c}, T' + \mathbf{c} \subset C$, then $T, T' \subset C'$ for some code C' equivalent to C. If C is linear and $T, T' \subset C$, then $T + \mathbf{c}, T' + \mathbf{c} \subset C$.*

By Remark 2, we can assume that $\mathbf{c}' = \mathbf{0}$.

We write

$$\begin{aligned} \mathbf{c} &= (c_1, c_2, \ldots, c_n), \\ \mathbf{a} &= (a_1, a_2, \ldots, a_n), \\ \mathbf{b} &= (b_1, b_2, \ldots, b_n). \end{aligned}$$

Since $(\mathbf{0}, \mathbf{c}; \mathbf{a}, \mathbf{b})$ is a $(2,2)$-NSC, there is no coordinate i such that both $a_i \notin \{0, c_i\}$ and $b_i \notin \{0, c_i\}$.

We consider the sum

$$\begin{aligned} \Sigma &:= d(\mathbf{0}, \mathbf{a}) + d(\mathbf{0}, \mathbf{b}) + d(\mathbf{c}, \mathbf{a}) + d(\mathbf{c}, \mathbf{b}) \\ &= w(\mathbf{a}) + w(\mathbf{b}) + w(\mathbf{a} - \mathbf{c}) + w(\mathbf{b} - \mathbf{c}). \end{aligned}$$

We have trivially that

$$4d_1 \leq \Sigma \leq 4m_1.$$

Consider now the matrix with rows $\mathbf{0}, \mathbf{c}, \mathbf{a}, \mathbf{b}$. Let $\mathbf{x}_i$ be the i-th column in this matrix. We assume, with no loss of generality, that c=(1,..1,0,...0); this is clearly achievable by permuting columns and multiplying columnwise by the appropriate non zero element. Then, we are left with four main types of columns:

$$\begin{array}{lcll} \text{Type } 0 & : & \mathbf{x}_i = (0,0,0,0), & \\ \text{Type I} & : & \mathbf{x}_i \in \{(0,0,0,\alpha), (0,0,\alpha,0)\}, & \alpha \neq 0, \\ \text{Type IIa} & : & \mathbf{x}_i \in \{(0,1,0,0), (0,1,1,1)\}, & \\ \text{Type IIb} & : & \mathbf{x}_i \in \{(0,1,0,1), (0,1,1,0)\}, & \\ \text{Type III} & : & \mathbf{x}_i \in \{(0,1,0,\beta), (0,1,\beta,0), (0,1,1,\beta), (0,1,\beta,1)\}, & \beta \notin \{0,1\}. \end{array}$$

No other possibility exists because the rows form a $(2,2)$-NSC. We have now that

$$\Sigma = \sum_{i=1}^{n} \sigma(\mathbf{x}_i),$$

where $\sigma(\mathbf{x}_i)$ is 0 for Type 0, 2 for Types I and II, and 3 for Type III. Let v_{X} denote the number of columns of Type X. Then we get

$$n = v_0 + v_{\mathrm{I}} + v_{\mathrm{II}} + v_{\mathrm{III}}, \tag{1}$$
$$\Sigma = 2v_{\mathrm{I}} + 2v_{\mathrm{II}} + 3v_{\mathrm{III}}. \tag{2}$$

Lemma 1 *If* $(\mathbf{0}, \mathbf{c}; \mathbf{a}, \mathbf{b})$ *is a* $(2,2)$*-NSC, then*

$$\Sigma = w(\mathbf{c}) + w(\mathbf{a} - \mathbf{b}) + w(\mathbf{a} + \mathbf{b} - \mathbf{c}).$$

Proof: We have trivially that

$$n - w(\mathbf{c}) = v_0 + v_I. \tag{3}$$

Define two vectors

$$\begin{aligned} \mathbf{y} &= (y_1, y_2, \ldots, y_n) := \mathbf{a} + \mathbf{b} - \mathbf{c}, \\ \mathbf{z} &= (z_1, z_2, \ldots, z_n) := \mathbf{a} - \mathbf{b}. \end{aligned}$$

We have that

$$\begin{array}{llllllll} \mathbf{x_i} & \text{of} & \text{Type} & 0 & \Rightarrow y_i & = 0 & \wedge & z_i = 0, \\ \mathbf{x_i} & \text{of} & \text{Type} & \text{I} & \Rightarrow y_i & = \alpha & \wedge & z_i = \pm\alpha, \\ \mathbf{x_i} & \text{of} & \text{Type} & \text{IIa} & \Rightarrow y_i & = \pm\alpha & \wedge & z_i = 0, \\ \mathbf{x_i} & \text{of} & \text{Type} & \text{IIb} & \Rightarrow y_i & = 0 & \wedge & z_i = \pm\alpha, \\ \mathbf{x_i} & \text{of} & \text{Type} & \text{III} & \Rightarrow y_i & \neq 0 & \wedge & z_i \neq 0. \end{array}$$

This gives

$$\begin{aligned} n - w(\mathbf{a} + \mathbf{b} - \mathbf{c}) &= n - w(\mathbf{y}) = v_0 + v_{\mathrm{IIb}}, \\ n - w(\mathbf{a} - \mathbf{b}) &= n - w(\mathbf{z}) = v_0 + v_{\mathrm{IIa}}. \end{aligned}$$

By adding together the two equations above as well as (3), we get

$$\begin{aligned} &3n - (w(\mathbf{c}) + w(\mathbf{a} - \mathbf{b}) + w(\mathbf{a} + \mathbf{b} - \mathbf{c})) \\ &= 3v_0 + v_{\mathrm{IIa}} + v_{\mathrm{IIb}} + v_{\mathrm{I}}. \end{aligned}$$

From (2) and (1) we get that

$$\begin{aligned}\Sigma &= 3n - (3v_0 + v_{\mathrm{IIa}} + v_{\mathrm{IIb}} + v_{\mathrm{I}}) \\ &= w(\mathbf{c}) + w(\mathbf{a}-\mathbf{b}) + w(\mathbf{a}+\mathbf{b}-\mathbf{c}),\end{aligned}$$

as required. □

We observe that $d(\mathbf{a},\mathbf{b}) = w(\mathbf{a}-\mathbf{b})$ and $d(\mathbf{0},\mathbf{c}) = w(\mathbf{c})$ are distances in the code; hence they are bounded by m_1. If C is linear, then $w(\mathbf{a}+\mathbf{b}-\mathbf{c})$ is also a distance in the code and is thus bounded by m_1. If C is non-linear, we still have $w(\mathbf{a}+\mathbf{b}-\mathbf{c}) \leq n$. This gives directly the following theorem.

Theorem 1 *If a code satisfies $4d_1 > 2m_1 + n$, or if $4d_1 > 3m_1$ and it is linear, then it is $(2,2)$-separating.*

Example 1 *Take a linear projective code over $GF(p^2)$ [2] with length*

$$n = \frac{(p^{k_1} - (-1)^{k_1})(p^{k_1-1} - (-1)^{k_1-1})}{p^2-1},$$

dimension k_1 and weights $w_1 = p^{2k_1-3}, w_2 = p^{2k_1-3} + (-p)^{k_1-2}$.

In the case $p = 2$ for $k_1 = 4$ it gives a $[45,4,32]_4$ code, which is $(2,2)$-separating since $m_1 = w_2 = 39$ and it satisfies $4d_1 > 3 = m_1$.

Example 2 *A three-weight code over $GF(p)$ is given in [2] with length*

$$n = p + 1 + p^2(p^{k_1-1} - (-1)^{k_1-1})(p^{k_1-2} - (-1)^{k_1-2})/(p-1),$$

dimension $k = 2k_1$ and weights $w_1 = p^{2k_1-2} - (-p)^{k_1} - (-p)^{k_1-1}, w_2 = p^{2k_1-2}, w_3 = p^{2k_1-2} - (-p)^{k_1}$.

In the binary case for $k = 6$ it gives a $[39,6,20]$ code, which is $(2,2)$-separating since $m_1 = w_3 = 24$ and it satisfies $4d_1 > 3m_1$.

Corollary 1 *All linear, equidistant codes are $(2,2)$-separating. A non-linear, equidistant code is $(2,2)$-separating if $2d_1 > n$.*

The non-linear case of the corollary was proved in [5] by other methods. Note that it is tight:

$$C = \{(1000), (0100), (0010), (0001)\}$$

is an equidistant $(4,4)$ code with distance 2, but not separating. The linear case of the theorem is also tight, as the following example shows.

Example 3 *From the proposition we get that if* $(\mathbf{0}, \mathbf{c}; \mathbf{a}, \mathbf{b})$ *is a binary* $(2,2)$*-NSC and* $4d_1 = 3m_1$*, then*

$$\begin{aligned} w(\mathbf{c}) &= w(\mathbf{a}-\mathbf{b}) = w(\mathbf{a}+\mathbf{b}-\mathbf{c}) = m_1 = 4l, \\ w(\mathbf{a}) &= w(\mathbf{b}) = w(\mathbf{a}-\mathbf{c}) = w(\mathbf{b}-\mathbf{c}) = d_1 = 3l. \end{aligned}$$

It turns out that the only possible $(2,2)$*-NSC is the following, or replications thereof:*

$$\begin{aligned} \mathbf{0} &= 000000 \\ \mathbf{c} &= 111100 \\ \mathbf{a} &= 110010 \\ \mathbf{b} &= 101001. \end{aligned}$$

Note that the linear code $\langle \mathbf{a}, \mathbf{b}, \mathbf{c} \rangle$ *has also* $d_1 = 3$ *and* $m_1 = 4$*.*

By m_2 we denote the maximum support size of the union of two codewords.

Proposition 2 *If* C *is binary, linear and* $2d_1 > m_2$*, then it is* $(2,2)$*-separating.*

Proof: Let $(\mathbf{0}, \mathbf{c}; \mathbf{a}, \mathbf{b})$ be a $(2,2)$-NSC. We consider first the case where $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are linearly independent. Then $\mathbf{a}+\mathbf{b}$, $\mathbf{a}+\mathbf{b}+\mathbf{c}$, and $\mathbf{c}$ are the three non-zero codewords in some 2-dimensional subcode D. Thus we get that

$$w(\mathbf{c}) + w(\mathbf{a}-\mathbf{b}) + w(\mathbf{a}+\mathbf{b}-\mathbf{c}) = 2w(\mathbf{a}+\mathbf{b}, \mathbf{c}) \leq 2m_2, \tag{4}$$

and by Lemma 1 that

$$4d_1 \leq \Sigma = w(\mathbf{c}) + w(\mathbf{a}-\mathbf{b}) + w(\mathbf{a}+\mathbf{b}-\mathbf{c}) \leq 2m_2. \tag{5}$$

If $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are linearly dependent, then $\mathbf{a}+\mathbf{b}+\mathbf{c} = \mathbf{0}$, and (4) becomes

$$w(\mathbf{c}) + w(\mathbf{a}-\mathbf{b}) + w(\mathbf{a}+\mathbf{b}-\mathbf{c}) \leq 2m_1,$$

which is stronger than (5). □

It is easy to show that $m_2 \leq \lfloor 3m_1/2 \rfloor$, which is a maximum support weight analogue of the Griesmer bound. If this bound is not met with equality, then the above result is stronger than that of Theorem 1.

4 Concatenation

The ternary construction we now introduce employs three ingredient codes and applies twice the classical concatenation method, using the following easy result.

Proposition 3 *[4] If Γ_1 is a (t,t')-separating, M'-ary (n_1, M) code and Γ_2 a (t,t')-separating, q-ary, (n_2, M') code, then the concatenated code $\Gamma := \Gamma_2 \circ \Gamma_1$ is a (t,t')-separating $(n_1 n_2, M)_q$ code.*

The first seed is the remarkable $[4,2,3]_3$ tetracode $\mathcal{T}$, defined by the generator matrix

$$\begin{bmatrix} 1110 \\ 0121 \end{bmatrix}$$

Both $\mathcal{T}$ and $\mathcal{R}_1$, the $[9,3,7]_{3^2}$ Reed-Solomon code, are $(2,2)$-separating by Theorem 1.

The concatenated code $\mathcal{T} \circ \mathcal{R}_1$ has parameters $[36,6]_3$, and is $(2,2)$-separating by Proposition 3. In order to produce infinite families of separating codes, we need the following constructive result from Tsfasman [13].

Proposition 4 *For any $\alpha > 0$ there is an infinite family of codes $\mathcal{U}(N)$ with parameters $[N, NR, N\delta]_q$ for $N \geq N_0(\alpha)$ and*

$$R + \delta \geq 1 - (\sqrt{q} - 1)^{-1} - \alpha.$$

We should note that the rate of $\mathcal{U}(N)$ is interesting only for large q, but $\mathcal{T} \circ \mathcal{R}_1$ allows for concatenation with $\mathcal{U}(N)$ over $GF(3^6)$, which is acceptable. Thus, consider the family of $[N, K, D = \lceil 3N/4 \rceil + 1]_{3^6}$ codes $\mathcal{U}(N)$, which has rate $R' \approx 1/4 - (3^3 - 1)^{-1} = 11/52$. The concatenated code $\mathcal{T} \circ \mathcal{R}_1 \circ \mathcal{U}(N)$ gives an infinite family of linear, ternary $(2,2)$-separating and codes with rate $R'/6 \approx 0.0352$.

Example 4 *We sketch a construction with a few other values of q. As in the ternary case, we concatenate three codes to build the infinite family. Each one has $d/n > 3/4$ and thus is again $(2,2)$-separating by Theorem 1. The first two are Reed-Solomon codes or their extensions; the last one is a $\mathcal{U}(N)$ of length N and distance $\lceil 3N/4 \rceil + 1$.*

For $q = 4$, take successively:

1. $C_1[5,2,4]_4$*;*

2. $C_2[17,5,13]_{4^2}$*, getting* $C_1 \circ C_2[85,10]_4$*;*

3. $\mathcal{U}(N)[N,K,D] = \lceil 3N/4 \rceil + 1]_{4^{10}}$ *with rate approximately* $1/4 - (4^5 - 1)^{-1} \approx 1/4$.

The final outcome is an infinite constructive family of linear quaternary $(2,2)$*-separating codes with rate approximately* $1/34 \approx 0.029$.

For $q = 5$, $C_1[5,2,4]_5$, $C_2[25,7,19]_{5^2}$, *with overall rate* $14/500$*; for* $q = 7$, $C_1[7,2,6]_7$, $C_2[49,13,37]_{7^2}$, *with overall rate* $13/686$*; for* $q = 8$, $C_1[9,3,7]_8$, $C_2[65,17,49]_{8^2}$, *with overall rate* $17/780$.

All these results are summarized in Table 1.

5 Non-constructive bounds

Below we present existence proofs of linear separating codes over different fields. The first lemma is fairly well-known, and can be found in [9].

Lemma 2 *Asymptotically, for almost all linear codes, the number of codewords of weight* i, A_i, *satisfies*

$$A_i = \frac{\left(\begin{array}{c} n \\ i \end{array} \right) (q-1)^i}{q^{n(1-R)}} \approx \frac{e^{nH(i/n)} e^{i \ln(q-1)}}{e^{n(1-R)\ln q}},$$

where H *is the natural entropy function and* $R = k/n$ *is the rate.*

Since we are dealing with the asymptotical case, we normalize by setting $i = n\omega$, and we define the function $f(\omega, R, q)$ by

$$A_{\omega n} = e^{nf(\omega,R,q)}.$$

From Lemma 2, we get

$$f(\omega, R, q) = H(\omega) + \omega \ln(q-1) - (1-R)\ln q. \tag{6}$$

Note that, for a given A_i, there are two solutions for i. Setting $A_i \approx 1$, the two solutions will be the minimum and the maximum weights. These are of course also the zeroes of f.

Let $\delta = d_1/n$ and $\mu = m_1/n$ be respectively the minimum and maximum normalized weights. Because μ and δ are the zeroes of f, we get

$$H(\delta) + \delta \ln(q-1) = H(\mu) + \mu \ln(q-1),$$

or

$$\begin{aligned}(\delta - \mu)\ln(q-1) = \delta \ln \delta + (1-\delta)\ln(1-\delta) \\ - \mu \ln \mu - (1-\mu)\ln(1-\mu).\end{aligned} \tag{7}$$

Lemma 3 (Varshamov-Gilbert) *For almost all linear codes, the rate and the normalized minimum distance are related by the following equation*

$$H(\delta) + \delta \ln(q-1) = (1-R)\ln q.$$

Proof: This follows from equating $f(\omega, R, q) = 0$ as in (6). □

We know from Theorem 1 that if $\delta > 3/4$, then the code is $(2,2)$-separating. Hence we can, by substituting $\delta = 3/4$ in the Varshamov-Gilbert equation, get rates for which almost any code is $(2,2)$-separating asymptotically. The rates such obtained are presented under "Technique I" in Table 1. By the Plotkin bound, this gives nothing over small fields.

Technique II in the table is an improvement based on Theorem 1, which says that every code with $4\delta > 3\mu$ is $(2,2)$-separating. We insert $\delta = 4\mu/3$ in (7)

and get

$$\begin{aligned}\frac{\delta}{3}\ln(q-1) = \delta \ln \delta + (1-\delta)\ln(1-\delta) \\ - \frac{4\delta}{3}\ln\frac{4\delta}{3} - (1-\frac{4\delta}{3})\ln(1-\frac{4\delta}{3}).\end{aligned} \tag{8}$$

We have solved this equation numerically for the smallest fields, and the results are given in Table 1. Of course, we will always have

$$0 \leq \delta \leq \mu \leq 1,$$

which will bound $\delta \leq 3/4$ in (8).

This results in no real solution of (8) for $q \geq 11$.

		Technique I		*Technique II*		*Constructions*
q	$\delta_{\max}$	δ	*Rate*	δ	*Rate*	*Rate*
2	0.5000	0.75	N/A	0.4286	0.01477	0.026
3	0.6667	0.75	N/A	0.5695	0.01859	0.0352
4	0.7500	0.75	N/A	0.6385	0.02206	0.029
5	0.8000	0.75	0.00459	0.6786	0.02532	0.028
7	0.8571	0.75	0.02043	0.7218	0.03153	0.019
8	0.8750	0.75	0.02774	0.7340	0.03457	0.021
9	0.8889	0.75	0.03427	0.7426	0.03766	
11	0.9091	0.75	0.04530	N/A	N/A	
13	0.9231	0.75	0.05417	N/A	N/A	
16	0.9375	0.75	0.06464	N/A	N/A	
17	0.9412	0.75	0.06757	N/A	N/A	
19	0.9474	0.75	0.07279	N/A	N/A	

Table 1: Rates for which there exist (asymptotically) linear $(2, 2)$-separating codes. The number $\delta_{\max} = (q-1)/q$ is Plotkin upper bound.

Note that, in Table 1, the best results are obtained by the "Constructions" for $q \leq 5$, then by "Technique 2" for $7 \leq q \leq 9$, and finally by "Technique 1" for higher values of q. In the binary case, R=0.0642 can be achieved nonconstructively [10].

6 Acknowledgement

The authors wish to thank the referees for their useful remarks.

References

[1] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data", *Springer-Verlag LNCS* 963, pp. 452-465, 1995.

[2] I. M. Chakravarti, "Families of codes with few distinct weights from singular and non-singular Hermitian varieties and quadrics in projective

geometries and Hadamard difference sets and designs associated with two-weight codes", *Coding theory and design theory, Part 1*, Springer verlag, pp. 35-50, 1990.

[3] G. Cohen, S. Encheva, "Efficient constructions of frameproof codes", *Electronics Letters*, vol. 36, no. 22, pp.1840-1842, 2000.

[4] G. Cohen, S. Encheva and H.G. Schaathun, "On Separating Codes", *Technical reports in informatics, the Centre National de la Recherche Scientifique and Télécom Paris, Département INF*, 2001.

[5] S. Encheva and G. Cohen, "Identifying Codes for Copyright Protection", *Technical reports in informatics, the Centre National de la Recherche Scientifique and Télécom Paris, Département INF*, 2001.

[6] A.D. Friedman, R.L. Graham and J.D. Ullman, "Universal single transition time asynchronous state assignments", *IEEE Trans. Comput.*, vol. C-18, pp. 541-547, June 1969.

[7] M.L. Friedman and J. Komlós, "On the size of separating systems and families of perfect hash functions", *SIAM J. Alg. Disc. Meth.*, vol. 5, no. 1, pp. 61-68, March 1984.

[8] J. Körner and G. Symonyi, *SIAM J. Disc. Math.*, vol. 1, no. 3, pp. 355-359, Aug. 1988.

[9] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, 1977.

[10] Yu.L. Sagalovich, "Separating systems", *Problems of Information Transmission*, vol. 30 (2), pp. 105-123, 1994.

[11] D.R. Stinson, Tran Van Trung and R. Wei, "Secure Frameproof Codes, Key Distribution Patterns, Group Testing Algorithms and Related Structures", *J. Stat. Planning and Inference*, vol. 86 (2), pp. 595-617, 2000.

[12] D.R. Stinson and R. Wei, "Combinatorial properties and constructions of traceability schemes and frameproof codes", *SIAM J. Discrete Math*, 11, 41-53, 1998.

[13] M.A. Tsfasman, "Algebraic-geometric codes and asymptotic problems", *Discrete Appl. Math.*, vol. 33, pp. 241-256, 1991.