FINGERPRINTS FOR COPYRIGHT PROTECTION

Sylvia Encheva

Stord/Haugesund College

Børnsonsg. 45, 5528 Haugesund

Norway

sbe@hsh.no

Abstract If Γ is a q -ary code of length n and $a, b, \dots, e$ are codewords, then c is called a descendant of $a, b, \dots, e$ if $c_i \in \{a_i, b_i, \dots, e_i\}$ for $i = 1, \dots, n$. Codes Γ with the property that coalitions of a limited size have restrictions on their descendants are studied. Namely, we consider codes with one of the following properties, referred to later as 2-partially identifying (2-PIC) and 3-frameproof (3-FPC) codes, respectively:

- any two non intersecting coalitions of size two have no common descendant
- the descendant of any coalition of size at most 3 does not belong to the code.

Keywords: Frameproof codes, partially identifying codes, copyright protection.

1 Introduction

The ability to resolve ownership disputes and copyright infringement is difficult in the worldwide digital age. There is an increasing need to develop techniques that protect the owner of digital data. Digital watermarking is a technique used to embed a known piece of digital data within another one. The embedded piece acts as a fingerprint for the owner, allowing the protection of copyright, authentication of the data, and tracing of illegal copies.

A publisher embeds a unique fingerprint pattern into each distributed copy of a document, keeping a database of sold copies and their corresponding fingerprints. If, later on, an illegally distributed copy is discovered, he may trace that copy back to the offending user by comparing its fingerprint to the database. We consider the attack which results when two users collude and compare their independently marked copies. Then

---

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35413-2_36](https://doi.org/10.1007/978-0-387-35413-2_36)

they can detect and locate the differences, and combine their copies into a new one whose fingerprint differs from all the users’.

Codes were introduced in [2] (see also [13]) 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, or from framing an innocent user.

The outline of the paper is as follows. Definitions and basic results are presented in Section 2. In Section 3 we first prove that Sylvester type matrices are 2-PIC. Although, they can accommodate a limited number of users, their length 2^i make them suitable for practical applications and concatenation.

In Section 4, Hadamard matrices of order $4(2k + 1)$ are shown to be 3-FPC. Such codes are also good for practical applications since a large number of Hadamard matrices of order $4(2k + 1)$ have been constructed (there are already 486 Hadamard matrices of order 28 with Hall sets). Section 5 presents asymptotic constructions *via* concatenation, for which seeds described in Sections 3 and 4 can be used. The last section is devoted to q -ary s -PIC. In particular we show that an equidistant code with $2w > n$ is 2-PIC. Thus we improve the requirement $4d > 3n$ in [4]. We also improve the sufficient condition $d/n > 1 - 1/s^2$ for s -PIC [13].

2 Definitions and basic results

We use the notation of [13] for fingerprinting issues and of [15] for codes and Hadamard matrices.

We identify a vector with its *support*, set of its non-zero positions; in the binary case for a pair of vectors (u, v) , we denote by $u \cap v$ the vector having for support the intersection of the supports of u and v . 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*

$$P_i(\mathcal{C}) = \bigcup_{a \in \mathcal{C}} a_i.$$

Define 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 $F(\mathcal{C}) = \mathcal{C}$ if $|\mathcal{C}| = 1$.

Now, if there is a codeword $a \in F(C) \setminus C$, then the user who owns codeword a can be “framed” because the coalition C can actually construct a .

Thus the following definition from [2] to forbid this situation.

Definition 1 An (n, M) -code C is called a s -frameproof code if, for every $C \subseteq C$ such that $|C| \leq s$, we have $F(C) \cap C = C$.

If two non-intersecting coalitions can produce the same descendant, it will be impossible to trace with certainty even one guilty user. This motivates the following reworded definition from [13].

Definition 2 An (n, M) -code Γ is called a s -partially identifying code (s -PIC for short) if, for every couple of coalitions $C, C' \subseteq \Gamma$ such that $|C| \leq s, |C'| \leq s$ and $C \cap C' = \emptyset$, we have $F(C) \cap F(C') = \emptyset$.

The identification property can be rephrased as follows when $q = 2$:

For any ordered $2s$ -tuple of codewords written as columns, there is a coordinate where the $2s$ -tuple $(1..10..0)$ of weight s or its complement occurs.

We denote by $C[n, k, d]_q$ (or simply $C[n, k]_q$ when d is irrelevant) a linear code of length n , dimension k over $GF(q)$ and minimum distance d . The subscript is omitted in the binary case. The rate of C is $R(C) = R = k/n$. In the nonlinear case, the rate is defined analogously as $n^{-1} \log_q M$. We refer to [15] for all undefined notions on codes.

3 Binary 2-PIC

Theorem 1 [1]

If H_n is an n times n Hadamard matrix with $n > 1$, then n is even and for any two distinct rows of H there are precisely $\frac{n}{2}$ columns in which the entries in the two rows agree. Further, if $n > 2$ then n is divisible by 4, and for any three distinct rows of H there are precisely $\frac{n}{4}$ columns in which the entries in all three rows agree. The same statements hold for columns of H .

Let H_n be a Sylvester type matrix of order $n = 2^i$ [15]. If +1's are replaced by 0's and -1's by 1's, H_n is changed into the binary Hadamard matrix A_n [15].

Theorem 2 The matrix A_n is a $(n = 2^i, M = 2^i)$ 2-PIC.

Proof: Set $n = 4m$ with $m = 2^{i-2}$. We distinguish between two cases, according to whether or not one of the four codewords is the first row of A_{4m} (the all 1 vector).

Case 1. Let $c_1, c_2, c_3, c_4 \in A_{4m}$ be all different from the first row in A_{4m} . By [15] Ch. 2, since any two rows in A_{4m} are orthogonal, they agree in $2m$ places and differ in $2m$ places. Any such two rows contain m columns of the four possible types $(11)^T, (00)^T, (01)^T, (10)^T$. Then A_{4m} will not be 2-PIC if the following occurs:

$$\begin{array}{lcl}
 c_1 : & \overbrace{1\dots 1}^{\beta} & \overbrace{1\dots\dots\dots 1}^{\alpha} & \overbrace{1\dots 1}^{\lfloor \frac{\alpha}{2} \rfloor} \overbrace{0\dots 0}^{\lceil \frac{\alpha}{2} \rceil} & \overbrace{* \dots * \dots * \dots *}_{3m-\alpha} \\
 c_2 : & \overbrace{1\dots 1}^{\beta} & \overbrace{1\dots\dots\dots 1}^{\alpha} & \overbrace{0\dots 0}^{\lfloor \frac{\alpha}{2} \rfloor} \overbrace{1\dots 1}^{\lceil \frac{\alpha}{2} \rceil} & \overbrace{* \dots * \dots * \dots *}_{3m-\alpha} \\
 c_3 : & \overbrace{1\dots 1}^{\beta} & \overbrace{1\dots 1}^{\lfloor \frac{\alpha}{2} \rfloor} \overbrace{0\dots 0}^{\lceil \frac{\alpha}{2} \rceil} & \overbrace{1\dots\dots\dots 1}^{\alpha} & \overbrace{* \dots * \dots * \dots *}_{3m-\alpha} \\
 c_4 : & \overbrace{1\dots 1}^{\beta} & \overbrace{0\dots 0}^{\lfloor \frac{\alpha}{2} \rfloor} \overbrace{1\dots 1}^{\lceil \frac{\alpha}{2} \rceil} & \overbrace{1\dots\dots\dots 1}^{\alpha} & \overbrace{* \dots * \dots * \dots *}_{3m-\alpha}
 \end{array}$$

where the number of columns $(11)^T$ in any two rows of A_{4m} is $\alpha + \beta = m$, a '*' being indifferently a 0 or 1. If α is an odd number then c_2, c_4 have more than $1 + \alpha + \beta = 1 + m$ columns $(11)^T$, a contradiction. Thus α is even, say $\alpha = 2\alpha'$. Then A_{4m} will fail to be 2-PIC if the following occurs:

$$\begin{array}{lcl}
 c_1 : & \overbrace{1\dots 1}^{\beta} & \overbrace{1\dots\dots\dots 1}^{2\alpha'} & \overbrace{1\dots 1}^{\alpha'} \overbrace{0\dots 0}^{\alpha'} & \overbrace{1\dots\dots\dots 1}^{2m-\beta-3\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} \\
 c_2 : & \overbrace{1\dots 1}^{\beta} & \overbrace{1\dots\dots\dots 1}^{2\alpha'} & \overbrace{0\dots 0}^{\alpha'} \overbrace{1\dots 1}^{\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} & \overbrace{1\dots\dots\dots 1}^{2m-\beta-3\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} \\
 c_3 : & \overbrace{1\dots 1}^{\beta} & \overbrace{1\dots 1}^{\alpha'} \overbrace{0\dots 0}^{\alpha'} & \overbrace{1\dots\dots\dots 1}^{2\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} & \overbrace{1\dots\dots\dots 1}^{2m-\beta-3\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} \\
 c_4 : & \overbrace{1\dots 1}^{\beta} & \overbrace{0\dots 0}^{\alpha'} \overbrace{1\dots 1}^{\alpha'} & \overbrace{1\dots\dots\dots 1}^{2\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} & \overbrace{0\dots\dots\dots 0}^{2m-\beta-3\alpha'} & \overbrace{1\dots\dots\dots 1}^{2m-\beta-3\alpha'}
 \end{array}$$

Columns in $[c_1, c_2, c_3, c_4]$ are written this way up to position $\beta + 4\alpha'$ since any two rows in A_{4m} have columns $(11)^T$ in m positions. The number of $(01)^T$ and $(10)^T$ columns are equal, otherwise two of the rows c_1, c_2, c_3, c_4 would have more than m columns $(11)^T$. Therefore rows c_1, c_2, c_3, c_4 must have disjoint supports from position $\beta + 4\alpha' + 1$ on, since any two of them contain already m columns $(11)^T$ on the first $\beta + 4\alpha'$ positions. We now upperbound by $4m$ the number of columns containing at least one 1, getting $\beta + 4\alpha' + 4(2m - \beta - 3\alpha') \leq 4m$, or $3\beta + 8\alpha' \geq 4m$. On the other hand, for any two rows, the number of

$(11)^T$ columns is $m = \beta + 2\alpha'$. This leads to $2\alpha' = m, \beta = 0$, which is impossible since the first column of A_{4m} contains only 1's.

Case 2 now. Suppose that c_1 is the first row in A_{4m} , i.e. $wt(c_1) = 4m$. Then A_{4m} will not be 2-PIC if the following occurs.

$$\begin{array}{cccc}
 c_1 : & \overbrace{1\dots 1}^m & \overbrace{1\dots 1}^m & \overbrace{1\dots 1}^m & \overbrace{1\dots 1}^m \\
 & m & m & m & m \\
 c_2 : & \overbrace{1\dots 1}^m & \overbrace{1\dots 1}^m & \overbrace{0\dots 0}^m & \overbrace{0\dots 0}^m \\
 & m & m & m & m \\
 c_3 : & \overbrace{1\dots 1}^m & \overbrace{0\dots 0}^m & \overbrace{1\dots 1}^m & \overbrace{0\dots 0}^m \\
 & m & m & m & m \\
 c_4 : & \overbrace{10\dots 0}^m & \overbrace{1\dots 1}^m & \overbrace{01\dots 1}^m & \overbrace{0\dots 0}^m
 \end{array}$$

The row c_4 should have a 1 in the first position since all rows of A_{4m} do. The support of c_4 in positions $\{m+1, m+2, \dots, 2m\}$ has size m by the assumption that A_{4m} is not 2-PIC. In this case $[c_2, c_4]$ will have $m+1$ columns $(11)^T$, again reaching a contradiction. Therefore A_m is a 2-PIC. (Note that the actual support of c_4 in positions $\{2m+1, \dots, 3m\}$ is irrelevant to our proof).

□

Remark 1 A_{4m} is NOT a 3-FPC. For $m = 2$, a counterexample is given below (the fourth row, e.g., can be “framed” by the first three, since neither (1110) nor (0001) appear as a column of the first four rows). It easily extends to all m .

11111111
 10101010
 11001100
 10011001
 11110000
 10100101
 11000011
 10010110

4 The case $n = 4m$, m odd

Let $H_{4(2k+1)}$ be the matrix obtained from a Hadamard matrix [15] of order $4(2k+1)$ when +1's are replaced by 0's and -1's by 1's.

Theorem 3 $H_{4(2k+1)}$ is a binary 3-FPC($4(2k+1), 4(2k+1)$).

Proof: Suppose $H_{4(2k+1)}$ is not a binary 3-FPC($4(2k+1), 4(2k+1)$), i.e. there are four rows $c_1, c_2, c_3, c_4 \in H_{4(2k+1)}$ yielding no column of either

type $(0001)^T$ or $(1110)^T$. Again, two cases may occur. If no c_i is the all 1 vector, then, w.l.o.g. we may assume that these four rows form the following array:

$$\begin{array}{cccc}
 c_1 & \overbrace{1 \dots 1}^{2k+1} & \overbrace{1 \dots 1}^{2k+1} & \overbrace{0 \dots 0}^{2k+1} & \overbrace{0 \dots 0}^{2k+1} \\
 & \overbrace{}^{2k+1} & \overbrace{}^{2k+1} & \overbrace{}^{2k+1} & \overbrace{}^{2k+1} \\
 c_2 & \overbrace{1 \dots 1}^{2k+1} & \overbrace{0 \dots 0}^{2k+1} & \overbrace{1 \dots 1}^{2k+1} & \overbrace{0 \dots 0}^{2k+1} \\
 c_3 & \overbrace{1 \dots 1}^m \overbrace{0 \dots 0}^{2k+1-m} & \overbrace{1 \dots 1}^{2k+1-m} \overbrace{0 \dots 0}^m & \overbrace{1 \dots 1}^{2k+1-m} \overbrace{0 \dots 0}^m & \overbrace{1 \dots 1}^m \overbrace{0 \dots 0}^{2k+1-m} \\
 & \overbrace{}^m \overbrace{}^{2k+1-m-x} & \overbrace{}^y \overbrace{}^{2k+1-m-y} \overbrace{}^z \overbrace{}^{m-z} & \overbrace{}^t \overbrace{}^{2k+1-m-t} \overbrace{}^\alpha \overbrace{}^{m-\alpha} & \overbrace{}^\beta \overbrace{}^{m-\beta} \overbrace{}^{2k+1-m} \\
 c_4 & \overbrace{1 \dots 1}^m \overbrace{1 \dots 1}^x \overbrace{0 \dots 0}^{2k+1-m-x} & \overbrace{1 \dots 1}^y \overbrace{0 \dots 0}^{2k+1-m-y} \overbrace{1 \dots 1}^z \overbrace{0 \dots 0}^{m-z} & \overbrace{1 \dots 1}^t \overbrace{0 \dots 0}^{2k+1-m-t} \overbrace{1 \dots 1}^\alpha \overbrace{0 \dots 0}^{m-\alpha} & \overbrace{1 \dots 1}^\beta \overbrace{0 \dots 0}^{m-\beta} \overbrace{0 \dots 0}^{2k+1-m}
 \end{array}$$

provided any row has weight $2(2k+1)$.

By Theorem 1 any three rows in $H_{4(2k+1)}$ agree in $2k+1$ positions. Applying Theorem 1 to the rows c_2, c_3, c_4 we obtain $m+t=z$. Since $z \leq m$ we conclude that $t=0$ and $m=z$.

In a similar way, considering c_1, c_2, c_4 gives $x=0$ and $m=\beta$, and c_1, c_3, c_4 gives $m=\alpha$. Any two rows of weight $2(2k+1)$ in $H_{4(2k+1)}$ have $2k+1$ columns of type $(11)^T$. Therefore c_4 should have $2k+1$ 1's in positions $\{2(2k+1)+1, \dots, 4(2k+1)\}$, which leads to $t+\alpha+\beta=2k+1$. This contradicts our conclusion that $\alpha=\beta=m$ and $t=0$. Therefore there is a column of either type $(0001)^T$ or $(1110)^T$ in $H_{4(2k+1)}$.

Assuming now (second case) that one of c_1, c_2, c_3, c_4 is the all 1 vector, we obtain a contradiction in a similar way. This concludes the proof that $H_{4(2k+1)}$ is a 3-FPC.

□

5 Asymptotic constructions

The following classical concatenation method is quite useful for obtaining binary FP or PIC codes from p -ary ones, $p=2^k$. Let C_1 be an (N, M_p) code over $GF(p)$; let C be an (n, M) binary code with $M \geq p$. We map arbitrarily $GF(p)$ onto a subset C_2 of C of size p . Denote by $C_1 \star C_2$ the *concatenation* of C_1 and C_2 , obtained by replacing every occurrence of a p -ary symbol in codewords of C_1 by its image (a binary n -tuple of C_2).

5.1 Infinite constructions of 3-FPC

We have the following result from [5].

Proposition 1 $C_1 \star C_2$ is an (Nn, M_p) binary code. If C_1 and C_2 are both 3 - FPC (over $GF(p)$ and $GF(2)$ respectively), then $C_1 \star C_2$ is a 3 - FPC binary code.

The efficiency of concatenation is due to the fact that, for large alphabets $GF(p)$, the condition $3d > 2n$ ([14]) on distance ensures that the code is a 3-FPC. We shall state the results for *linear* codes, since they are more practical, although they hold for unrestricted codes as well.

Let $p = 2^{2^m}$. For any $\alpha > 0$, and any N_0 , there exist (see [16]) constructions of codes $C_1(N)$ with parameters $[N, NR, N\delta]$ over $GF(p)$ with $N \geq N_0$ and $R + \delta \geq 1 - (2^m - 1)^{-1} - \alpha$.

Choosing $m = 3$, $\delta = 2/3 + \epsilon$ and concatenating $C_1(N)$ and C_2 , a binary (64, 64) code obtained from Section 3, yields a constructive infinite sequence $(C_1(N) \star C_2)_N$ of binary 3-FP codes with rates arbitrarily close to $1/56$.

5.2 Infinite constructions of 2-PIC

Proposition 2 (6) If C_1 and C_2 are both 2-PIC (over $GF(p)$ and $GF(2)$ respectively), then $C_1 \star C_2$ is a 2-PIC binary code.

Let $p = 2^{2^m}$ again. In the previous construction of codes $C_1(N)$ $[N, NR, N\delta]$ over $GF(p)$ with $N \geq N_0$ and $R + \delta \geq 1 - (2^m - 1)^{-1} - \alpha$, choose now $m = 7$ and $\delta = 3/4 + \epsilon$; then concatenate $C_1(N)$ and C_2 , where C_2 is a binary [126, 14, 55] code of rate $1/9$ (see [6]). This yields a constructive infinite sequence $(C_1(N) \star C_2)_N$ of binary linear 2-PIC codes with rates arbitrarily close to 0.026.

6 On s -PIC

Positions where two codewords coincide are denoted $\overbrace{*\dots*}$, and positions where they have different coordinates are denoted $\underbrace{\dots}$. We now improve on a result from [13].

Theorem 4 Let Γ be a code with $d/n \geq 1 - 1/s^2$; then its codewords make up the elements of a s -partially identifying code.

Proof: For simplicity we shall consider only $s = 2$. The same idea carries over to any s . Suppose Γ is not 2-PIC. Then, possibly after some column

permutation, there are codewords $c_1, c_2, c_3, c_4 \in \Gamma$, such that:

c_1, c_3 coincide on positions $1, 2, \dots, x + \alpha; x + \alpha + y + 1, \dots, x + \alpha + y + x_1 + \alpha_1$

c_1, c_4 coincide on positions $x + 1, \dots, x + \alpha + y; x + \alpha + y + x_1 + 1, \dots, x + \alpha + y + x_1 + \alpha_1 + y_1$

c_2, c_3 coincide on positions $1, 2, \dots, x + \alpha; x + \alpha + y + x_1 + \alpha_1 + y_1 + 1, \dots, x + \alpha + y + x_1 + \alpha_1 + y_1 + x_2 + \alpha_2$

c_2, c_4 coincide on positions $x + 1, \dots, x + \alpha + y; x + \alpha + y + x_1 + \alpha_1 + y_1 + x_2, \dots, n$, i.e.

$$\begin{array}{lcl}
 c_1 : & \begin{array}{c} x \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} x_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} x_2 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha_2 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y_2 \\ \underbrace{*\dots*} \end{array} \\
 c_2 : & \begin{array}{c} x \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} x_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} x_2 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha_2 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y_2 \\ \underbrace{*\dots*} \end{array} \\
 c_3 : & \begin{array}{c} x \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} x_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} x_2 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha_2 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y_2 \\ \underbrace{*\dots*} \end{array} \\
 c_4 : & \begin{array}{c} x \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} x_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y_1 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} x_2 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} \alpha_2 \\ \underbrace{*\dots*} \end{array} & \begin{array}{c} y_2 \\ \underbrace{*\dots*} \end{array} .
 \end{array}$$

Combining $n = x + \alpha + y + x_1 + \alpha_1 + y_1 + x_2 + \alpha_2 + y_2$ with the number of positions where (c_1, c_3) , (c_2, c_3) , (c_1, c_4) and (c_2, c_4) coincide, i.e.

$$x + \alpha + x_1 + \alpha_1 + x + \alpha + x_2 + \alpha_2 + \alpha + y + \alpha_1 + y_1 + \alpha + y + \alpha_2 + y_2 \leq 2^2(n - d),$$

we get $4d < 3n$. Thus, Γ is 2-PIC if $2^2d \geq (2^2 - 1)n$.

□

Example 1 The ternary hexacode $[4, 2, 3]$ is 2-PIC by Theorem 4.

Proposition 3 Let C be an equidistant q -ary code with $\text{dist}(c_i, c_j) = w$. If $2w > n$, then C is 2-PIC.

Proof: Suppose C is not 2-PIC. Then w.l.o.g. we may assume that there are four codewords $c_1, c_2, c_3, c_4 \in C$ such that c_1, c_2 and c_3, c_4

can produce a common descendant. Let

c_1, c_2 coincide on positions $1, 2, \dots, x + \alpha + y = n - w$,

c_1, c_3 coincide on positions $1, 2, \dots, x + \alpha; x + \alpha + y + 1, \dots, x + \alpha + 2y$

c_2, c_3 coincide on positions $1, 2, \dots, x + \alpha; x + \alpha + 2y + 1, \dots, x + \alpha + 3y$,

c_1, c_4 coincide on positions $x + 1, \dots, x + \alpha + y + m_1; x + \alpha + 3y + 1, \dots, 2x + \alpha + 3y - m_1$.

c_2, c_4 coincide on positions $x + 1, \dots, x + \alpha + y; x + \alpha + 2y + 1, \dots, x + \alpha + y + m_2$;

$$2x + 3y + \alpha - m_1 + 1, \dots, 3x + \alpha + 3y - m_1 - m_2 = n.$$

Thus the 4 times n array, obtained from c_1, c_2, c_3, c_4 , may be described as follows

$$\begin{array}{l}
 c_1 : \begin{array}{cccccccc} \overbrace{*\dots*}^x & \overbrace{*\dots*}^\alpha & \overbrace{*\dots*}^y & \overbrace{**}^{m_1} & \overbrace{**}^{y-m_1} & \overbrace{\dots}^{m_2} & \overbrace{\dots}^{y-m_2} & \overbrace{*\dots*}^{x-m_1} & \overbrace{\dots\dots}^{x-m_2} \end{array} \\
 c_2 : \begin{array}{cccccccc} \overbrace{*\dots*}^x & \overbrace{*\dots*}^\alpha & \overbrace{*\dots*}^y & \overbrace{\dots}^{m_1} & \overbrace{\dots}^{y-m_1} & \overbrace{**}^{m_2} & \overbrace{**}^{y-m_2} & \overbrace{\dots\dots}^{x-m_1} & \overbrace{*\dots*}^{x-m_2} \end{array} \\
 c_3 : \begin{array}{cccccccc} \overbrace{*\dots*}^x & \overbrace{*\dots*}^\alpha & \overbrace{\dots\dots}^y & \overbrace{**}^{m_1} & \overbrace{**}^{y-m_1} & \overbrace{**}^{m_2} & \overbrace{**}^{y-m_2} & \overbrace{\dots\dots}^{x-m_1} & \overbrace{\dots\dots}^{x-m_2} \end{array} \\
 c_4 : \begin{array}{cccccccc} \overbrace{\dots\dots}^x & \overbrace{*\dots*}^\alpha & \overbrace{*\dots*}^y & \overbrace{**}^{m_1} & \overbrace{\dots}^{y-m_1} & \overbrace{**}^{m_2} & \overbrace{\dots}^{y-m_2} & \overbrace{*\dots*}^{x-m_1} & \overbrace{*\dots*}^{x-m_2} \end{array} .
 \end{array}$$

Combining $\alpha + m_1 + m_2 = n - w$, (c_3, c_4 coincide on $n - w$ positions) and $\alpha + x + y = n - w$, (c_1, c_2 coincide on $n - w$ positions) we obtain $m_1 + m_2 = x + y$. The last equality together with

$$m_1 \leq x, m_1 \leq y, m_2 \leq x, m_2 \leq y$$

leads to $x = y = m_1 = m_2$.

Looking at the positions where c_1, c_3 coincide we get

$$x + y + \alpha = 2y + \alpha = n - w.$$

Replacing $2y$ by w gives $\alpha + 2w = n$. This is possible only for $\alpha = 0, 2w = n$. By assumption $2w > n$, therefore C is 2-PIC.

□

Example 2 *There exists an equidistant ternary (13, 39, 8) code [3]: by Proposition 3, it is 2-PIC.*

*References

- [1] Th. Beth, D. Jungnickel and H. Lenz, *Design Theory*, Wissenschaftsverlag, Berlin, (1985).
- [2] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data", *Springer-Verlag LNCS* 963 (1995), pp. 452-465.
- [3] A.E. Brouwer, H. O. Hömölöinen, P.R.J. Östergård, "Bounds on mixed binary/ternary codes", *IEEE Transactions on Information Theory*, vol. 44 (1998) pp. 140-161.
- [4] B. Chor, A. Fiat and M. Naor, "Tracing traitors", *Springer-Verlag LNCS* 839 (1994), pp. 257-270.
- [5] G. Cohen, S. Encheva, "Efficient constructions of frameproof codes", *Electronics Letters*, 2000.
- [6] G. Cohen, S. Encheva and S. Litsyn, "Intersecting codes and partially identifying codes", to appear in *Discrete Math*.

- [7] G. Cohen, S. Encheva and G. Zemor, "Copyright Protection for Digital Data". *IEEE Communications Letters*, vol. 4, issue 15, 2000.
- [8] G. Cohen and G. Zémor, "Intersecting codes and independent families", *IEEE Transactions on Information Theory*, vol. 40 (1984) pp. 1872-1881.
- [9] S. Encheva and G. Cohen, "Constructions of Intersecting Codes", *IEEE Trans. Inform. Theory*, May 1999.
- [10] H.D.L. Hollmann, J.H. van Lint, J.-P. Linhartz and L.M.G.M. Tolhuizen, "On codes with the identifiable parent property", *J. Combin. Th.*, Series A, 82 (1998), 121-133.
- [11] H. Kimura, "Hadamard Matrices and Dihedral Groups", *Designs, Codes and Cryptography*, 9, pp. 71-77, 1996.
- [12] J. Körner and G. Simonyi, "Separating partition systems and locally different sequences", *SIAM J. Discrete Math.*, vol. 1, pp. 355-359, 1988.
- [13] 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)(2000), pp. 595-617.
- [14] D.R. Stinson and R. Wei, "Combinatorial properties and constructions of traceability schemes and frameproof codes", *SIAM J. Discrete Math.*, 11 (1998), 41-53.
- [15] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, (1977).
- [16] M.A. Tsfasmann, "Algebraic-geometric codes and asymptotic problems" *Discrete Appl. Math.*, vol. 33, 1991, pp. 241-256.