ELSEVIER

Journal of Statistical Planning and Inference 100 (2002) 281–302

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Optimal superimposed codes and designs for Renyi's search model [☆]

Arkadii G. D'yachkov [*], Vyacheslav V. Rykov

Department of Probability Theory, Faculty of Mechanics and Mathematics, Moscow State University, Moscow 119899, Russia

Abstract

Renyi (Bull. Amer. Math. Soc. 71 (6) (1965) 809) suggested a combinatorial group testing model, in which the size of a testing group was restricted. In this model, Renyi considered the search of one defective element (significant factor) from the finite set of elements (factors). The corresponding optimal search designs were obtained by Katona (J. Combin. Theory 1 (2) (1966) 174). In the present work, we study Renyi's search model of several significant factors. This problem is closely related to the concept of binary superimposed codes, which were introduced by Kautz and Singleton (IEEE Trans. Inform Theory 10 (4) (1964) 363) and were investigated by D'yachkov and Rykov (Problems Control Inform. Theory 12 (4) (1983) 229), Erdos et al. (Israel J. Math. 51 (1–2) (1985) 75), Ruszinko (J. Combin. Theory Ser. A 66 (1994) 302) and Furedi (J. Combin. Theory Ser. A 73 (1996) 172). Our goal is to prove a lower bound on the search length and to construct the optimal superimposed codes and search designs. The preliminary results have been published by D'yachkov and Rykov (Conference on Computer Science & Engineering Technology, Yerevan, Armenia, September 1997, p. 242). © 2002 Elsevier Science B.V. All rights reserved.

MSC: 94B50

Keywords: Combinatorial group testing; Boolean sum; Binary superimposed code; Perfect hash function

1. Notations and definitions

Let $1 \leqslant s < t$, $1 \leqslant k < t$, $N > 1$ be integers and $X = ||x_i(u)||$, $i = 1, 2, \ldots, N$, $u = 1, 2, \ldots, t$, be a binary $(N \times t)$-matrix (code) with columns (codewords) $\mathbf{x}(1)$, $\mathbf{x}(2), \ldots, \mathbf{x}(t)$ and rows $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$, where $\mathbf{x}(u) = (x_1(u), x_2(u), \ldots, x_N(u))$ and $\mathbf{x}_i = (x_i(1), x_i(2), \ldots, x_i(t))$. Let

$$w = \min_u \sum_{i=1}^{N} x_i(u), \quad k = \max_i \sum_{u=1}^{t} x_i(u), \quad \lambda = \max_{u,v} \sum_{i=1}^{N} x_i(u)x_i(v)$$

[☆] The work was supported by the Russian Fundamental Research Foundation, project 98-01-00241.

[*] Corresponding author.

E-mail address: dyachkov@nw.math.msu.su (A.G. D'yachkov).

PII: S0378-3758(01)00140-9

be the *minimal weight of codewords*, the *maximal weight of rows* and the *maximal dot product of codewords*.

We say that the binary column $\mathbf{x}$ *covers* the binary column $\mathbf{y}$ if the Boolean sum $\mathbf{x} \vee \mathbf{y} = \mathbf{x}$. The code X is called (Kautz and Singleton, 1964; D'yachkov and Rykov, 1983) a *superimposed* (s,t)-*code* if the Boolean sum of any s-subset of columns of X covers those and only those columns of X which are the terms of the given Boolean sum. The code X is called (Kautz and Singleton, 1964; D'yachkov and Rykov, 1983) a *superimposed* (s,t)-*design* if all Boolean sums composed of not more than s columns of X are distinct.

Definition 1. An $(N \times t)$-matrix X is called a superimposed (s,t,k)-code (design) of *length* N, *size* t, *strength* s and *constraint* k if code X is a superimposed (s,t)-code (design) whose the maximal row weight is equal to k.

The above-mentioned constraint k was introduced by Renyi (1965) and was studied by Katona (1966) for the search designs.

2. Lower bound

Let $\lceil a \rceil$ denote the least integer $\geqslant a$.

Proposition 1. *Let $t > k \geqslant s \geqslant 2$ and $N > 1$ be integers.*

(1) *For any superimposed $(s-1,t,k)$-code ((s,t,k)-design) X of length N, the following inequality takes place*:

$$N \geqslant \left\lceil \frac{st}{k} \right\rceil. \tag{1}$$

(2) *If $k \geqslant s+1$, $st = kN$ and there exists the optimal superimposed $(s-1,t,k)$-code X of length $N = st/k$, then*
   (a) *code X is a constant weight code of weight $w = s$, for any $i = 1,2,\ldots,N$, the weight of row $\mathbf{x}_i$ is equal to k and the maximal dot product $\lambda = 1$;*
   (b) *the following inequality is true*:

$$k^2 - \frac{k(k-1)}{s} \leqslant t. \tag{2}$$

Proof. (1) It is known (D'yachkov and Rykov, 1983) that code X is a superimposed (s,t,k)-design if and only if X is superimposed $(s-1,t,k)$-code and all $\binom{t}{s}$ Boolean sums composed of s columns of X are distinct. Hence, we need to prove inequality (1) for superimposed $(s-1,t,k)$-codes only. Let $s \geqslant 2$, $1 \leqslant k < t$ be fixed integers. Consider an arbitrary superimposed $(s-1,t,k)$-code X of length N. Let n, $0 \leqslant n \leqslant t$, be the number of codewords of X having a weight $\leqslant s-1$. From definition of superimposed $(s-1,t)$-code it follows (see, Kautz and Singleton, 1964) that $n \leqslant N$ and, for each codeword of weight $\leqslant s-1$, there exists a row in which all the remaining elements, except for the element of this codeword, are 0's. We delete these n rows from X together with n codewords of weight $\leqslant s-1$. Consider the remaining $(N-n)\times(t-n)$ matrix X'. Obviously, each column of X' has a weight $\geqslant s$ and each its row contains $\leqslant k$ 1's. Since $k \geqslant s$, we have

$$s(t-n) \leqslant k(N-n), \quad ts \leqslant kN - n(k-s) \leqslant kN. \tag{3}$$

Statement (1) is proved.

(2) Let $k \geqslant s+1$, $st = kN$ and X be the optimal superimposed $(s-1,t,k)$-code of length $N = st/k$.

- Since $k \geqslant s+1$, inequality (3) has signs of equalities if and only if X is the constant weight code of weight $w = s$ and for any $i = 1,2,\ldots,N$, the weight of row $\mathbf{x}_i$ is equal to k. By contradiction, using the constant weight property $w = s$ one can easily check that the maximal dot product $\lambda = 1$.
  Statements (2)(a) is proved.
- To prove Statement (2)(b), we apply the well-known Johnson inequality

$$t\binom{w}{\lambda+1} \leqslant \binom{N}{\lambda+1},$$

which is true for any constant weight code X of length N, size t, weight w and the maximal dot product λ. In our case, $\lambda = 1$, $w = s$, $tw = kN$ and $N = st/k$. This gives

$$tw(w-1) \leqslant N(N-1), \quad k(s-1) \leqslant N - 1 = \frac{st}{k} - 1, \quad k^2(s-1) + k \leqslant st,$$

$$k^2 - \frac{k(k-1)}{s} \leqslant t.$$

Proposition 1 is proved. □

Denote by $N(s,t,k)$, $(\tilde{N}(s,t,k))$ the minimal possible length of superimposed (s,t,k)-code ((s,t,k)-design). From Proposition 1 it follows:

- if $k \geqslant s+1$, then

$$\tilde{N}(s,t,k) \geqslant N(s-1,t,k) \geqslant \left\lceil \frac{st}{k} \right\rceil.$$

- if $k \leqslant s$, then $N(s-1,t,k) = \tilde{N}(s,t,k) = t$.

3. Optimal parameters

Let $s \geqslant 2$ and $k \geqslant s+1$ be fixed integers. Denote by $q \geqslant 2$ an arbitrary integer. We shall consider the *optimal* superimposed $(s-1,kq,k)$-codes and *optimal* superimposed (s,kq,k)-designs of length $N = sq$ whose parameters satisfy (1) with the sign of equality. By virtue of (2)

- if $q \geqslant k-(k-1)/s$, then there exists a possibility to find the optimal superimposed $(s-1, kq, k)$-code of length $N = sq$;
- if $q < k-(k-1)/s$, then lower bound (1) is not achieved and the interesting *open problem* is *how to obtain a new nontrivial lower bound on* $N(s-1,t,k)$ *provided that*
$$k^2 - \frac{k(k-1)}{s} > t.$$

Some constructions of superimposed $(2, kq, k)$-designs of length $N = 2q$ and superimposed $(2, kq, k)$-codes of length $N = 3q$ were obtained in Kautz and Singleton (1964). By virtue of Proposition 1, they are optimal. We give here the parameters of these designs and codes. The following statements are true:

- if $k-1 \geqslant 2$ is a prime power and $q = k^2-k+1$, then there exists an superimposed $(2, kq, k)$-design of length $N = 2q$,
- for pair $(k = 3,\ q = 5)$ and pair $(k = 7,\ q = 25)$, there exists an superimposed $(2, kq, k)$-design of length $N = 2q$.
- if $k \geqslant 4$ and $q = k-1$, or $q = k$, then there exists an superimposed $(2, kq, k)$-code of length $N = 3q$.

The aim of this paper—to prove Theorems 1–4.

Theorem 1. *Let $s = 2$ and $k \geqslant 3$ be integers. Then*

(1) *for any integers $q \geqslant k \geqslant 3$ there exists an optimal superimposed $(1, kq, k)$-code of length $N = 2q$, i.e., $N(1, kq, k) = 2q,\ q \geqslant k$;*
(2) *for any integer $q \geqslant 2^k - 1$ there exists an optimal superimposed $(2, kq, k)$-design of length $N = 2q$, i.e., $\tilde{N}(2, kq, k) = 2q,\ q \geqslant 2^k - 1$.*

Theorem 2. *Let $s \geqslant 3$, $k \geqslant s+1$ be fixed integers and $q = k^{s-1}$. Then there exists an optimal superimposed (s, kq, k)-design X of length $N = sq$, i.e. $\tilde{N}(s, k^s, k) = sk^{s-1}$.*

Theorem 3. *Let $k = 4, 5, \ldots,$ be a fixed integer. For any integer $q \geqslant k+1$, there exists an optimal superimposed $(2, kq, k)$-code of length $N = 3q$, i.e., $N(2, kq, k) = 3q$, $q \geqslant k+1$.*

Remark. Let $s \geqslant 3$. For the case of superimposed (s, kq, k)-codes, Theorem 3 is generalized (the proof is omitted) as follows. Let p_i, $i = 1, 2, \ldots, I$, be arbitrary prime numbers and r_i, $i = 1, 2, \ldots, I$, be arbitrary integers. If
$$q = p_1^{r_1} p_2^{r_2} \cdots p_I^{r_I}, \quad 3 \leqslant s \leqslant \min_i \{p_i^{r_i}\} - 1,$$
then for any k, $s+1 \leqslant k \leqslant q+1$, the optimal length $N(s, kq, k) = (s+1)q$.

The following theorem supplements Theorem 2 if $s = 3$ and $k = 4$.

Theorem 4. *If $k=4$ and $q\geqslant 12$, then there exists an optimal superimposed $(3,kq,k)$-design of length $N=3q$, i.e.,*

$$\tilde{N}(3,4q,4)=3q, \quad q\geqslant 12.$$

To prove Theorems 1–4, we apply concatenated codes using a class of homogeneous q-nary codes of size $t=kq$. The description of cascade construction, definitions and properties of homogeneous q-nary codes will be given in Section 4. The proofs of Theorems 1–4 will be given in Sections 5–8.

The following theorem yields a different family of optimal superimposed (s,t,k)-codes. It will be proved in Section 9.

Theorem 5. *Let $s\geqslant 1$, $k\geqslant s+2$ be fixed integers. Then there exists an (s,t,k)-code of size $t=\binom{k+s}{s+1}$ and length*

$$N=\frac{(s+1)t}{k}=\frac{(s+1)\binom{k+s}{s+1}}{k}=\binom{k+s}{s},$$

i.e., the optimal length

$$N\left(s,\binom{k+s}{s+1},k\right)=\binom{k+s}{s}.$$

For Theorem 5, the optimal code constructions were invented by Macula (1996).

4. Homogeneous q-nary codes

Let $q\geqslant s\geqslant 1$, $k\geqslant 2$, $k\leqslant t\leqslant kq$, $J\geqslant 2$ be integers, $A_q=\{a_1,a_2,\ldots,a_q\}$ be an arbitrary q-nary alphabet and $B=||b_j(u)||$, $j=1,2,\ldots,J$, $u=1,2,\ldots,t$, be an q-nary ($b_j(u)\in A_q$) ($J\times t$)-matrix (code) with t columns (codewords) and J rows

$$\mathbf{b}(u)=(b_1(u),b_2(u),\ldots,b_J(u)), \quad u=1,2,\ldots,t,$$

$$\mathbf{b}_j=(b_j(1),b_j(2),\ldots,b_j(t)), \quad j=1,2,\ldots,J.$$

Denote the number of a-entries in the jth row $\mathbf{b}_j$ by $n_j(a)$, where $a\in A_q$, $j=1,2,\ldots,J$. We suppose that for any $j=1,2,\ldots,J$ and any $a\in A_q$, the value $n_j(a)\leqslant k$.

Definition 2. Let $t=kq$. Code B is called an (q,k,J)-*homogeneous* code if for any $j=1,2,\ldots,J$ and any $a\in A_q$, the number $n_j(a)=k$.

Definition 3. Code B will be called an *s-disjunct* if for any codeword $\mathbf{b}(u)$ and any s-subset of codewords $\{\mathbf{b}(u_1),\mathbf{b}(u_2),\ldots,\mathbf{b}(u_s)\}$, there exists a coordinate $j=1,2,\ldots,J$ for which $b_j(u)\neq b_j(u_i)$, $i=1,2,\ldots,s$.

For two codewords $\mathbf{b}(u)$, $\mathbf{b}(v)$, $u \neq v$, define the q-nary Hamming distance

$$D(\mathbf{b}(u);\mathbf{b}(v)) = \sum_{j=1}^{J} \chi(b_j(u);b_j(v)),$$

$$\chi(a;b) = \begin{cases} 1 & \text{if } a \neq b, \\ 0 & \text{if } a = b. \end{cases}$$

Let $D = D(B) = \min_{u \neq v} D(\mathbf{b}(u);\mathbf{b}(v)) \leqslant J$ be the Hamming distance of code B. By contradiction, one can easily prove the following statement which gives the analog of the Kautz and Singleton (1964) condition.

Proposition 2. *If* $s(J - D(B)) \leqslant J - 1$, *then code* B *is* s-*disjunct. In addition*, (q,k,s)-*homogeneous code* B *is* $(s-1)$-*disjunct code if and only if* $D(B) = s - 1$.

Let $n \leqslant t$ be a fixed integer and $\mathbf{e} = \{e_1, e_2, \ldots, e_n\}$, $1 \leqslant e_1 < e_2 < \cdots < e_n \leqslant t$ be an arbitrary n-subset of the set $[t] = \{1,2,\ldots,t\}$. For a given code B and any $j = 1,2,\ldots,J$, denote by $\mathsf{A}_j(\mathbf{e},B) \subseteq A_q$-the *set of all pairwise distinct elements of the sequence* $b_j(e_1), b_j(e_2), \ldots, b_j(e_n)$. The set $\mathsf{A}_j(\mathbf{e},B)$ is called the jth, $j = 1,2,\ldots,J$, *coordinate* set of subset $\mathbf{e} \subseteq [t]$ over code B. For its *cardinality* $|\mathsf{A}_j(\mathbf{e},B)|$, we have

$$1 \leqslant |\mathsf{A}_j(\mathbf{e},B)| \leqslant \min\{n,\ q\}.$$

Definition 4. Let $s \geqslant 1$, $n \leqslant s$, $m \leqslant s$ be arbitrary integers. Code B is called an s-*separable* code if for any two distinct subsets

$$\mathbf{e} = \{e_1, e_2, \ldots, e_n\},\ 1 \leqslant e_1 < e_2 < \cdots < e_n \leqslant t,$$

$$\mathbf{e}' = \{e'_1, e'_2, \ldots, e'_m\},\ 1 \leqslant e'_1 < e'_2 < \cdots < e'_m \leqslant t,$$

of the set $[t]$, there exists $j = 1,2,\ldots,J$, for which the corresponding coordinate sets are distinct, i.e., $\mathsf{A}_j(\mathbf{e},B) \neq \mathsf{A}_j(\mathbf{e}',B)$. In other words, for an arbitrary n-subset $\mathbf{e} = \{e_1, e_2, \ldots, e_n\}$, of the set $[t]$, there exists the *possibility to identify* this n-subset $\mathbf{e} = \{e_1, e_2, \ldots, e_n\}$ (or the corresponding n-subset of codewords $\{\mathbf{b}(e_1),\ \mathbf{b}(e_2), \ldots, \mathbf{b}(e_n)\}$ of code B) on the basis of sets:

$$\mathsf{A}_1(\mathbf{e},B), \mathsf{A}_2(\mathbf{e},B), \ldots, \mathsf{A}_J(\mathbf{e},B), \quad \mathsf{A}_j(\mathbf{e},B) \subseteq A_q.$$

Remark. In Definitions 3 and 4, we used the terminology of Du and Hwang (1993).

One can easily prove (by contradiction) the following ordering among these properties:

$$s\text{-disjunct} \Rightarrow s\text{-separable} \Rightarrow (s-1)\text{-disjunct}.$$

Definition 5. Code B is called an s-*hash* (Fridman and Komlos, 1984) if for an arbitrary s-subset

$$\mathbf{e} = \{e_1, e_2, \ldots, e_s\}, \quad 1 \leqslant e_1 < e_2 < \cdots < e_s \leqslant t,$$

of the set $[t]$, there exists a coordinate $j=1,2,\ldots,J$, where the cardinality $|\mathsf{A}_j(\mathbf{e},B)|=s$, i.e., the elements $b_j(e_1),b_j(e_2),\ldots,b_j(e_s)$ are *all different*.

Obviously, the following *ordering* takes place: s-hash ⇒ $(s-1)$-disjunct.

Definition 6. Code B is called an *s-hash&separable* if it has both of these properties.

Let q-nary alphabet $A_q=[q]=\{1,2,\ldots,q\}$. To illustrate Definitions 2–6 and the proof of Theorem 1, we give two examples of disjunct and separable codes.

Example 1. Let $k=q=2,3,\ldots$ be fixed integers. The evident $(k,k,2)$-homogeneous 1-disjunct code B of distance $D=1$ has the following $t=k^2$ columns (codewords):

$$B=\begin{pmatrix} 111\ \ldots\ 1\ 222\ \ldots\ 2\ \ldots\ kkk\ \ldots\ k \\ 123\ \ldots\ k\ 123\ \ldots\ k\ \ldots\ 123\ \ldots\ k \end{pmatrix}.$$

Example 2. For $k=3$, $q=7$, the $(7,3,2)$-homogeneous 2-hash&separable code B of distance $D=1$ has $kq=21$ codewords:

$$B=\begin{pmatrix} 111\ 222\ 333\ 444\ 555\ 666\ 777 \\ 124\ 235\ 346\ 457\ 156\ 267\ 137 \end{pmatrix}.$$

The idea of the following two examples of $(q,k,3)$-homogeneous 3-hash&separable codes will be used to prove Theorem 2.

Example 3. For $k=3$, $q=9$, the $(9,3,3)$-homogeneous 3-hash&separable code B of distance $D=2$ has $kq=27$ columns (codewords):

$$B=\left(\begin{array}{c|c|c} 111\ 222\ 333 & 444\ 555\ 666 & 777\ 888\ 999 \\ 123\ 123\ 123 & 456\ 456\ 456 & 789\ 789\ 789 \\ 123\ 456\ 789 & 123\ 456\ 789 & 123\ 456\ 789 \end{array}\right).$$

Code B contains $k=3$ groups of codewords. In the first and second rows, we use the construction idea which could be called an *alphabet separating between groups*.

Remark. Obviously, 3-separable code B from example 3 is not 3-disjunct code. Hence, in general, the ordering s-separable ⇒ s-disjunct is not true.

Example 4. For $k=4$, $q=16$, the $(16,4,3)$-homogeneous 3-hash&separable code B of distance $D=2$ has $kq=64$ columns (codewords) which are divided into $k=4$ groups:

- the first 16 codewords take the form

$$\begin{array}{llllllllll}
1111 & 2222 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 \\
1234 & 1234 & 1 & 2 & 3 & 4 & 1 & 2 & 3 & 4 \\
1234 & 5678 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16,
\end{array}$$

- the construction of the last 48 codewords of B applies the same method of alphabet separating:
  - ○ for $j=1,2$ and $u=16m+l$, $m=1,2,3$, $l=1,2,\dots,16$, the element $b_j(u)=b_j(16m+l)=b_j(l)+4m$,
  - ○ for $j=3$ and $u=16m+l$, $m=1,2,3$, $l=1,2,\dots,16$, the element $b_3(u)=b_3(16m+l)=b_3(l)=l$.

Let q-nary alphabet $A_q=[q]=\{1,2,\dots,q\}$. For code B, we denote by

$$X_B=(\mathbf{x}(1),\mathbf{x}(2),\dots,\mathbf{x}(t)), \quad k\leqslant t\leqslant kq,$$

a binary $Jq\times t$ matrix (code), whose columns (codewords) have the form

$$\mathbf{x}(u)=(\mathbf{x}^1(u),\mathbf{x}^2(u),\dots,\mathbf{x}^s(u)), \quad u=1,2,\dots,t,$$

$$\mathbf{x}^j(u)=(x_1^j(u),x_2^j(u),\dots,x_q^j(u)), \quad j=1,2,\dots,J,$$

$$x_l^j(u)=\begin{cases} 1 & \text{if } l=b_j(u), \\ 0 & \text{if } l\neq b_j(u),\ l=1,2,\dots,q. \end{cases}$$

In other words, a symbol $b\in[q]$ of q-nary matrix B is replaced by the binary q-sequence in which all elements are 0's, except for the element with number b. Obviously, each codeword $\mathbf{x}(u)$ of (code) X_B contains J 1's and $(Jq-J)$ 0's and each row $\mathbf{x}_i$ of code X_B contains $\leqslant k$ 1's. For (q,k,J)-homogeneous code B, each row $\mathbf{x}_i$ of code X_B contains k 1's and $(kq-k)$ 0'. In addition, the stated below Proposition 3 follows easily by Definitions 2–4 and Propositions 1–2.

Proposition 3. *Let $q>k\geqslant s+1$ and B be a (q,k,s)-homogeneous code. The following two statements are true*:

- *If B is a $(s-1)$-disjunct code X_B, then X_B will be the optimal superimposed $(s-1,kq,k)$-code of length $N=sq$.*
- *If B is a s-separable code, then X_B will be the optimal superimposed (s,kq,k)-design of length $N=sq$.*

Hence, to prove Theorems 1–4, it is sufficient to construct the corresponding (q,k,s)-homogeneous codes. In particularly, the constructive method of Examples 3 and 4 yields Theorem 2 for the case $s=3$, i.e., $\tilde{N}(3,k^3,k)=3k^2$, $k=4,5,\dots$.

5. Proof of Theorem 1

Let $s=2$, $q \geqslant k$, q-nary alphabet $A_q=[q]=\{1,2,\ldots,q\}$ and $B=(\mathbf{b}(1),\mathbf{b}(2),\ldots,\mathbf{b}(kq))$ be an arbitrary $(q,k,2)$-homogeneous code 1-disjunct code, i.e., B has pairwise distinct codewords $\mathbf{b}(u)=(b_1(u),b_2(u))$, $u=1,2,\ldots,kq$. Following Kautz and Singleton (1964), we introduce the *binary characteristic* $(q \times q)$*-matrix* $C=||c_i(j)||$, $i=1,2,\ldots,q$, $j=1,2,\ldots,q$, where

$$c_i(j)=\begin{cases}1 & \text{if there exists codeword } \mathbf{b}(u)=(i,j),\\ 0 & \text{otherwise.}\end{cases}$$

Example 5. For $(7,3,2)$-homogeneous code B of Example 2, the characteristic (7×7)-matrix is

$$C=\begin{pmatrix}1&1&0&1&0&0&0\\0&1&1&0&1&0&0\\0&0&1&1&0&1&0\\0&0&0&1&1&0&1\\1&0&0&0&1&1&0\\0&1&0&0&0&1&1\\1&0&1&0&0&0&1\end{pmatrix}.$$

Obviously, the 1-disjunct code B is a $(q,k,2)$-homogeneous code if and only if the *weight of each row* and the *weight of each column* of C are equal to k. It is not difficult to understand that this condition is true for any *circulant* matrix. The circulant matrix C is defined as follows:

- the first row $\mathbf{c}_1=(c_1(1),c_1(2),\ldots,c_1(q))$ of circulant matrix C is an arbitrary binary sequence of length q and weight $k \leqslant q$,
- the mth $m=2,3,\ldots,q$ row $\mathbf{c}_m=((c_m(1),c_m(2),\ldots,c_m(q))$ of C is the *cyclic shift* of the $(m-1)$th row, i.e.,

$$c_m(j)=\begin{cases}c_{m-1}(q) & \text{if } j=1,\\ c_{m-1}(j-1) & \text{if } j=2,3,\ldots,q.\end{cases}$$

The first statement of Theorem 1 is proved.

To prove the second statement of Theorem 1, we apply the evident necessary and sufficient condition of 2-separable property which is given in Kautz and Singleton (1964) *no two* 1*'s in C must occupy the same pair of rows and columns as two other* 1*'s; that is, no row of C can contain a pair of* 1*'s in the same two positions as another row.*

It is easy to check that the circulant matrix C of Example 5 satisfies this condition. Let $q \geqslant 2^k$. As the simple generalization, we consider the circulant matrix C whose first row $\mathbf{c}_1 = (c_1(1), c_1(2), \ldots, c_1(q))$ is defined as follows:

$$c_1(j) = \begin{cases} 1 & \text{if } j = 2^{n-1},\ n = 1, 2, \ldots, k, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 1 is proved.

6. Proof of Theorem 2

The following Proposition 4 gives the recurrent construction method of $(s+1)$-separable codes with the help of s-separable codes.

Proposition 4. *If there exists an (q,k,s)-homogeneous s-separable code $B(q,k,s)$ with elements from $A_q = [q]$, then there exists the $(kq,k,s+1)$-homogeneous $(s+1)$-separable code $B(kq,k,s+1)$ with elements from $A_{kq} = [kq]$.*

Proof. Let

$$B(q,k,s) = ||b_j^s(u)||, \quad b_j^s(u) \in [q],\ j = 1, 2, \ldots, s,\ u = 1, 2, \ldots, kq.$$

be an arbitrary (q,k,s)-homogeneous s-separable code. Consider the following two-step recurrent construction (cf. Examples 3 and 4) for $(kq,k,s+1)$-homogeneous code:

$$B(kq,k,s+1) = ||b_j^{s+1}(u)||, \quad b_j^{s+1}(u) \in [kq],\ j = 1, 2, \ldots, s+1,\ u = 1, 2, \ldots, k^2 q.$$

- The first kq codewords of $B(kq,k,s+1)$ have the form

$$\begin{matrix} b_1^s(1) & b_1^s(2) & \ldots & b_1^s(kq) \\ b_2^s(1) & b_2^s(2) & \ldots & b_2^s(kq) \\ \cdots & \cdots & \cdots & \cdots \\ b_s^s(1) & b_s^s(2) & \ldots & b_s^s(kq) \\ 1 & 2 & \ldots & kq, \end{matrix}$$

  i.e., for $u = 1, 2, \ldots, kq$, the element $b_j^{s+1}(u) = b_j^s(u)$, if $j = 1, 2, \ldots, s$, and $b_{s+1}^{s+1}(u) = u$.
- If the number $u = kqm + l$, $m = 1, 2, \ldots, k-1$, $l = 1, 2, \ldots, kq$, then
  - for $j = 1, 2, \ldots, s$, the element $b_j^{s+1}(u) = b_j^{s+1}(kqm + l) = b_j^{s+1}(l) + qm$,
  - for $j = s+1$, the element $b_{s+1}^{s+1}(u) = b_{s+1}^{s+1}(kqm + l) = b_{s+1}^{s+1}(l) = l$.

Note that $t = k^2 q$ codewords of $B(kq,k,s+1)$ (or the set $[k^2 q]$) could be divided into k groups of the equal cardinality kq where the mth group $G_m(q,k,s)$, $m = 1, 2, \ldots, k$, has the form

$$G_m(q,k,s)=\begin{pmatrix} b_1^s(1)+(m-1)q & b_1^s(2)+(m-1)q & \dots & b_1^s(kq)+(m-1)q \\ b_2^s(1)+(m-1)q & b_2^s(2)+(m-1)q & \dots & b_2^s(kq)+(m-1)q \\ \dots & \dots & \dots & \dots \\ b_s^s(1)+(m-1)q & b_s^s(2)+(m-1)q & \dots & b_s^s(kq)+(m-1)q \\ 1 & 2 & \dots & kq \end{pmatrix}.$$

Let

$$A_{kq}^{(m)}=\{(m-1)kq+1,(m-1)kq+2,\dots,mkq\},\quad |A_{kq}^{(m)}|=kq,\quad \bigcup_{m=1}^{k} A_{kq}^{(m)}=[k^2q]$$

be the set of numbers of codewords which belong to $G_m(q,k,s)$. Consider the $s\times kq$ matrix $B_m(q,k,s)$ composed of the first s rows of $G_m(q,k,s)$, $m=1,2,\dots,k$. Obviously, $B_m(q,k,s)$ is (q,k,s)-homogeneous code. In addition, all elements of $B_m(q,k,s)$ belong to the alphabet

$$A_q^{(m)}=\{(m-1)q+1,(m-1)q+2,\dots,mq\},\quad |A_q^{(m)}|=q,\quad \bigcup_{m=1}^{k} A_q^{(m)}=[kq],$$

and, hence, they do not may occur in $B_n(q,k,s)$, if $n\neq m$, $n=1,2,\dots,k$. On account of the s-separable property of $B(q,k,s)$, it follows the s-separable property of $B_m(q,k,s)$, $m=1,2,\dots,k$.

To prove the $(s+1)$-separable property of $B(kq,k,s+1)$, we consider an arbitrary $(s+1)$-subset of the set $[k^2q]$: $\mathbf{e}=\{e_1,e_2,\dots,\mathbf{e}_{s+1}\}$, $1\leqslant e_1<e_2<\cdots<e_{s+1}\leqslant k^2q$. Let

$$\mathsf{A}_1(\mathbf{e},B),\mathsf{A}_2(\mathbf{e},B),\dots,\mathsf{A}_s(\mathbf{e},B),\mathsf{A}_{s+1}(\mathbf{e},B)$$

be the corresponding subsets of the set $[kq]$ and

$$\mathbf{e}=\sum_{m=1}^{k}\mathbf{e}_m,\quad \mathbf{e}_m=\mathbf{e}\bigcap A_{kq}^{(m)}.$$

The above-mentioned property of groups $G_m(q,k,s)$, $m=1,2,\dots,k$ implies that for any $j=1,2,\dots,s$, the set $\mathsf{A}_j(\mathbf{e},B)$ could be written in the form

$$\mathsf{A}_j(\mathbf{e},B)=\sum_{m=1}^{k}\mathsf{A}_j(\mathbf{e}_m,B),$$

where $\mathsf{A}_j(\mathbf{e}_m,B)\subseteq A_q^{(m)}$. Hence, for any fixed $j=1,2,\dots,s$, all nonempty sets $\mathsf{A}_j(\mathbf{e}_m,B)$, $m=1,2,\dots,k$, could be identified on the basis of the set $\mathsf{A}_j(\mathbf{e},B)$.

We have two possibilities.

- There exists the unique value $m=1,2,\dots,k$ such that $\mathbf{e}_m=\mathbf{e}$, $|\mathbf{e}_m|=s+1$. It follows that for any $j=1,2,\dots,s$, the set $\mathsf{A}_j(\mathbf{e}_m,B)\neq\emptyset$ and, for any $n\neq m$, the set $\mathsf{A}_j(\mathbf{e}_n,B)=\emptyset$. Hence, one can identify the set $\mathbf{e}$ on the basis of the set $\mathsf{A}_{s+1}(\mathbf{e},B)$.

- For any $m=1,2,\ldots,k$, the cardinality $|\mathbf{e}_m|\leqslant s$. Accounting the s-separating property of $B_m(q,k,s)$, the set $\mathbf{e}_m$ could be identified on the basis of s subsets $\mathsf{A}_j(\mathbf{e}_m,B)$, $j=1,2,\ldots,s$. It follows the possibility to identify $\mathbf{e}=\sum_{m=1}^{k}\mathbf{e}_m$.

Proposition 4 is proved. □

Let an $(k,k,2)$-homogeneous 1-separable code $B(k,k,2)$ be the code from Example 1. Consider the corresponding $(k^2,k,3)$-homogeneous code $B(k^2,k,3)$ obtained from $B(k,k,2)$ on the basis of Proposition 4. For $k=3,4$, constructions of $B(k^2,k,3)$ are given in Examples 3 and 4. To prove Theorem 2, it is sufficient to establish the 3-separable property of code $B=B(k^2,k,3)$ for $k=4,5,\ldots$.

We shall use symbols which were introduced to prove Proposition 4. Note that $t=k^3$ codewords of B could be divided (in increasing order) into k groups $G_m(k,k,2)$, $m=1,2,\ldots,k$ of the equal cardinality k^2. Consider the $2\times k^2$ matrix $B_m(k,k,2)$ composed of the first 2 rows of $G_m(k,k,2)$, $m=1,2,\ldots,k$. Obviously, $B_m(k,k,2)$ is the $(k,k,2)$-homogeneous 1-separable code. In addition, all elements of $B_m(k,k,s)$ belong to the alphabet

$$A_k^{(m)}=\{(m-1)k+1,(m-1)k+2,\ldots,mk\},\quad |A_k^{(m)}|=k,\quad \bigcup_{m=1}^{k}A_k^{(m)}=[k^2],$$

and, hence, they do not may occur in $B_n(k,k,2)$, if $n\neq m$, $n=1,2,\ldots,k$.

Let $\mathbf{e}=\{e_1,e_2,e_3\}$, $1\leqslant e_1<e_2<e_3\leqslant k^3$ be an arbitrary fixed 3-subset of the set $[k^3]$ and $\{\mathbf{b}(e_1),\mathbf{b}(e_2),\mathbf{b}(e_3)\}$ be the corresponding triple of codewords of code B. To identify the codewords $\mathbf{b}(e_i)$, $i=1,2,3$, using the properties of $B_m(k,k,2)$, $m=1,2,\ldots,k$, mentioned above, it suffices to analyze the following three cases.

- There are known three numbers $1\leqslant m_1<m_2<m_3\leqslant k$ such that the codeword $\mathbf{b}(e_i)$, $i=1,2,3$ belongs to the group $G_{m_i}(k,k,2)$. In this case, $\mathbf{b}(e_i)$ could be identified on the basis of 1-separable property of $B_{m_i}(k,k,2)$.
- There is known the number $m=1,2,\ldots,k$ such that all three codewords $\mathbf{b}(e_1),\mathbf{b}(e_2)$, $\mathbf{b}(e_3)$ belong to the group $G_m(k,k,2)$. In this case, the triple $\{\mathbf{b}(e_1),\mathbf{b}(e_2),\mathbf{b}(e_3)\}$, can be identified on the basis of the set $\mathsf{A}_3(\mathbf{e},B)$ whose cardinality $|\mathsf{A}_3(\mathbf{e},B)|=3$.
- There are known two numbers $1\leqslant m<n\leqslant k$ such that (without loss of generality) codeword $\mathbf{b}(e_1)$ belongs to the group $G_m(k,k,2)$ and two other codewords $\mathbf{b}(e_2)$ and $\mathbf{b}(e_3)$ belong to the group $G_n(k,k,2)$. In this case, we have the following three-step identification:

  ○ the codeword $\mathbf{b}(e_1)=(b_1(e_1),b_2(e_1),b_3(e_1))$ is identified on the basis of 1-separable property of $B_m(k,k,2)$,

  ○ the set $\{b_3(e_2),b_3(e_3)\}$ evidently identified on the basis of symbol $b_3(e_1)$ and the set $\mathsf{A}_3(\mathbf{e},B)$,

  ○ codewords $\mathbf{b}(e_2)$ and $\mathbf{b}(e_3)$ are identified on the basis of the set $\{b_3(e_2),b_3(e_3)\}$.

Theorem 2 is proved.

7. Proof of Theorem 3

Let $q \geqslant k+1$, $k \geqslant 4$, and q-nary alphabet $A_q = [q] = \{1, 2, \ldots, q\}$. We need to construct $(q, k, 3)$-homogeneous code B of distance $D(B) = 2$. Consider the construction of $(q, k, 3)$-homogeneous code $B = ||b_j(u)||$ whose rows $\mathbf{b}_j = (b_j(1), b_j(2), \ldots, b_j(kq))$, $j = 1, 2, 3$, are defined as follows:

1. for $j = 1$, the first row $\mathbf{b}_1 = (\mathbf{b}_1^{(1)}, \mathbf{b}_1^{(2)}, \ldots, \mathbf{b}_1^{(q)})$, $\mathbf{b}_1^{(m)} = \underbrace{(m, m, \ldots, m)}_{k}$, $m = 1, 2, \ldots, q$;
2. for $j = 2$, the second row $\mathbf{b}_2 = (\mathbf{b}_2^{(1)}, \mathbf{b}_2^{(2)}, \ldots, \mathbf{b}_2^{(k)})$, $\mathbf{b}_2^{(m)} = (1, 2, \ldots, q)$, $m = 1, 2, \ldots, k$;
3. for $j = 3$, the third row $\mathbf{b}_3 = (\mathbf{b}_3^{(1)}, \mathbf{b}_3^{(2)}, \ldots, \mathbf{b}_3^{(k)})$, where the subsequence $\mathbf{b}_3^{(m)}$ of length q is the $(m-1)$-*step cyclic shift* of the sequence $(1, 2, \ldots, q)$:

$$\mathbf{b}_3^{(m)} = \begin{cases} (1, 2, \ldots, q) & \text{if } m = 1, \\ m, m+1, \ldots, q-1, q, 1, 2, \ldots, m-1 & \text{if } m = 2, 3, \ldots, k. \end{cases}$$

Obviously, this construction guarantees the distance $D(B) = 2$. From Proposition 2 it follows 2-disjunct property of B.

Theorem 3 is proved.

Example 6. As an illustration, we yield the $(6, 4, 3)$-homogeneous 2-disjunct code B with $kq = 24$ codewords

$$B = \begin{pmatrix} 111122 & 223333 & 444455 & 556666 \\ 123456 & 123456 & 123456 & 123456 \\ 123456 & 234561 & 345612 & 456123 \end{pmatrix}.$$

Remark. For 2-disjunct code B of Example 6, it is easy to check the following properties:

- The 3-subsets $\mathbf{e} = \{2, 8, 13\}$ and $\mathbf{e}' = \{2, 7, 13\}$ of the set $[24]$ have the same coordinate sets, namely: $\mathsf{A}_1 = \{1, 2, 4\}$, $\mathsf{A}_2 = \{1, 2\}$ and $\mathsf{A}_3 = \{2, 3\}$. From this it follows that the code B is not 3-separable code, i.e., in general, the ordering $(s-1)$-disjunct $\Rightarrow$ s-separable is not true.
- The 3-subset $\mathbf{e} = \{1, 2, 7\}$ of the set $[24]$ has the equal coordinate sets $\mathsf{A}_1 = \mathsf{A}_2 = \mathsf{A}_3 = \{1, 2\}$ of cardinality 2. Hence, the code B is not 3-hash code, i.e., in general, the ordering $(s-1)$-disjunct $\Rightarrow$ s-hash is not true.

8. On (*q*, *k*, 3)-homogeneous 3-separable and 3-hash codes: Proof of Theorem 4

8.1. Characteristic matrices

Consider an arbitrary $(q,k,3)$-homogeneous 2-disjunct code B. From Proposition 2 it follows that we can introduce *characteristic* $(q\times q)$-matrix $C=||c_i(j)||$, $i=1,2,\ldots,q$, $j=1,2,\ldots,q$, with elements from alphabet $A_{q+1}=\{*,[q]\}=\{*,1,2,\ldots,q\}$, where

$$c_i(j)=\begin{cases} a & \text{if there exists codeword } \mathbf{b}(u)=(a,i,j),\\ * & \text{otherwise.}\end{cases}$$

We shall say that code B *is identified* by the (characteristic) matrix C which will be called $C(q,k)$*-matrix*.

Example 7. For $k=4$, $q=6$, the $(6,4,3)$-homogeneous 2-disjunct code B of Example 6 is identified by $C(q,k)$-matrix

$$C=\begin{pmatrix} 1 & 2 & 4 & 5 & * & * \\ * & 1 & 2 & 4 & 5 & * \\ * & * & 1 & 3 & 4 & 6 \\ 6 & * & * & 1 & 3 & 4 \\ 5 & 6 & * & * & 2 & 3 \\ 3 & 5 & 6 & * & * & 2 \end{pmatrix}.$$

The evident characterization of $C(q,k)$-matrix is given by Proposition 5.

Proposition 5. *The matrix C is $C(q,k)$-matrix if and only if C has the following properties*:

- *for any $a\in[q]$, the number of a-entries in C is equal to k,*
- *for any row (column) of C, the number of $*$-entries in the row (column) is equal to $q-k$,*
- *for any $a\in[q]$ and any row (column) of C, the number of a-entries in the row (column) does not exceed* 1.

Remark. If $q=k$, then $C(q,q)$-matrix is called the *Latin square*.

Characteristic matrix C of hash, separable and hash&separable code will be called $C_{\mathrm{H}}(q,k)$-matrix, $C_{\mathrm{S}}(q,k)$-matrix and $C_{\mathrm{HS}}(q,k)$-matrix.

One can easily check the following characterization of $C_{\mathrm{H}}(q,k)$-matrix.

Proposition 6. *Matrix C is $C_{\mathrm{H}}(q,k)$-matrix if and only if C has the properties of Proposition* 5 *and the following two equivalent conditions take place*:

- *If for $i \neq m$ and $j \neq n$, the element $c_i(j) = c_m(n) = a \neq *$, then $c_i(n) = c_m(j) = *$.*
- *If for $i \neq m$, $j \neq n$ and $a \neq b$, code B contains codewords (a, i, j) and (a, m, n), then B does not contain the word (b, m, j).*

The evident characterization of $C_{\mathrm{HS}}(q,k)$-matrix is given by Proposition 7.

Proposition 7. *Let a, b and c be arbitrary pairwise distinct elements of $[q]$. Matrix C is $C_{\mathrm{HS}}(q,k)$-matrix if and only if C has properties of Propositions 5 and 6 and the following property is true. Matrix C does not contain any (3×3)-submatrix of the form*

$$\begin{pmatrix} * & a & c \\ a & * & b \\ c & b & * \end{pmatrix}, \quad \begin{pmatrix} c & a & * \\ b & * & a \\ * & b & c \end{pmatrix}, \quad \begin{pmatrix} * & c & a \\ a & b & * \\ c & * & b \end{pmatrix},$$

$$\begin{pmatrix} a & * & c \\ * & a & b \\ b & c & * \end{pmatrix}, \quad \begin{pmatrix} a & c & * \\ * & b & a \\ b & * & c \end{pmatrix}, \quad \begin{pmatrix} c & * & a \\ b & a & * \\ * & c & b \end{pmatrix}.$$

These prohibited matrices are the permutations of the same three columns.

Remark. The characterization of $C_{\mathrm{S}}(q,k)$-matrix has a tedious form and it is omitted here. Below, we give the examples of $C_{\mathrm{S}}(q,k)$-matrices which are not $C_{\mathrm{HS}}(q,k)$-matrices.

8.2. Examples of hash, separable and hash&separable codes

Let an integer $k \geqslant 2$ be fixed. How to find the minimal possible integer $q_k \geqslant k$ such that there exists $C_{\mathrm{S}}(q_k,k)$-matrix, $C_{\mathrm{H}}(q_k,k)$-matrix or $C_{\mathrm{HS}}(q_k,k)$-matrix? From Examples 3 and 4 it follows that one can put $q_k = k^2$. For $k = 2,3,4$, the following Examples 8–10 improve this result and yield $C_{\mathrm{S}}(q_k,k)$, $C_{\mathrm{H}}(q_k,k)$ and $C_{\mathrm{HS}}(q_k,k)$-matrices for which $q_k < k^2$. If $k = 2,3,4$ and $q_k < q < k^2$, then the corresponding characteristic matrices could be given also.

Example 8. For $k = 2$, $C_{\mathrm{S}}(q_2,2)$-matrix, $C_{\mathrm{H}}(q_2,2)$-matrix and $C_{\mathrm{HS}}(q_2,2)$-matrix are[1]

$$\begin{pmatrix} 1 & 2 & * \\ * & 1 & 3 \\ 3 & * & 2, \end{pmatrix}, \quad \begin{pmatrix} * & \dot{3} & \ddot{1} \\ \dot{1} & \ddot{2} & * \\ \ddot{3} & * & \dot{2} \end{pmatrix}, \quad \begin{pmatrix} 1 & * & 3 & * \\ * & 1 & * & 3 \\ 4 & * & 2 & * \\ * & 4 & * & 2 \end{pmatrix}.$$

[1] Here and below, for 3-hash codes, we mark the pairs of "bad" triples which break the 3-separable property.

The first matrix ($q_2 = 3$) identifies the separable (not hash) code. The second matrix ($q_2 = 3$) identifies the hash (not separable) code. The third matrix ($q_2 = 4$) is the particular case of Proposition 4.

Example 9. Let $k = 3$. For hash code $q_3 = 6$ and for hash&separable code $q_3 = 7$. The corresponding characteristic matrices are

$$\begin{pmatrix} * & * & \dot{1} & \ddot{2} & 3 & * \\ * & \ddot{1} & * & \dot{5} & * & 3 \\ 1 & * & * & * & 5 & 4 \\ * & \dot{2} & \ddot{5} & * & * & 6 \\ 2 & * & 4 & * & 6 & * \\ 3 & 4 & * & 6 & * & * \end{pmatrix}, \quad \begin{pmatrix} * & * & 1 & 2 & 3 & * & * \\ * & 1 & * & 5 & 7 & * & * \\ 1 & * & * & * & * & 7 & 3 \\ * & 2 & * & * & * & 5 & 4 \\ 2 & * & 7 & * & * & * & 6 \\ * & 3 & * & 4 & * & 6 & * \\ 5 & * & 4 & * & 6 & * & * \end{pmatrix}.$$

Example 10. Let $k = 4$. For hash codes, $q_4 = 8$ and for hash&separable codes, $q_4 = 13$. The corresponding characteristic $C_{\mathrm{H}}(8,4)$ and $C_{\mathrm{HS}}(13,4)$-matrices are

$$\begin{pmatrix} * & * & * & 1 & * & \dot{2} & 3 & \ddot{5} \\ * & * & 1 & * & \ddot{2} & * & 4 & \dot{6} \\ * & 1 & * & * & 3 & 4 & * & 7 \\ 1 & * & * & * & \dot{5} & \ddot{6} & 7 & * \\ * & 2 & 3 & 4 & * & * & * & 8 \\ 2 & * & 5 & 6 & * & * & 8 & * \\ 3 & 5 & * & 7 & * & 8 & * & * \\ 4 & 6 & 7 & * & 8 & * & * & * \end{pmatrix},$$

$$\begin{pmatrix} * & * & * & * & * & * & 4 & 2 & 3 & * & 8 & * & * \\ * & * & * & * & * & * & * & 7 & 13 & 5 & 1 & * & * \\ * & * & * & * & * & * & * & 9 & 10 & * & * & 1 & 8 \\ * & * & * & * & * & * & 6 & 12 & 11 & * & * & 5 & * \\ * & * & * & * & * & * & * & * & * & 10 & 11 & 13 & 3 \\ * & * & * & * & * & * & * & * & * & 9 & 12 & 7 & 2 \\ 1 & * & * & 13 & * & * & * & * & * & 6 & * & * & 4 \\ 2 & 7 & 9 & 12 & * & * & * & * & * & * & * & * & * \\ 3 & 6 & 11 & 10 & * & * & * & * & * & * & * & * & * \\ * & 5 & * & * & 11 & 9 & 13 & * & * & * & * & * & * \\ 8 & 4 & * & * & 10 & 12 & * & * & * & * & * & * & * \\ * & * & 4 & 5 & 6 & 7 & * & * & * & * & * & * & * \\ * & * & 8 & * & 3 & 2 & 1 & * & * & * & * & * & * \end{pmatrix}.$$

Open problems. (1) Is it possible to construct a $C_{\mathrm{HS}}(q,4)$-matrix if $q<13$? (2) Is it possible to construct $C_{\mathrm{HS}}(q,k)$-matrices, if $k\geqslant 5$ and $q<k^2$?

8.3. Existence of hash and hash&separable codes

The following obvious Proposition 8 can be used to construct the new characteristic matrices using the known ones.

Proposition 8 (S.M. Ekhanin, 1998). *Let $v=1,2$ and there exist $C_{\mathrm{H}}(q_v,k)$-matrix*

$$C^v=||c_i^v(j)||,\ i,j\in[q_v],\quad c_i^v(j)\in\{*,[q_v]\}.$$

Let $\tilde{C}^2=||\tilde{c}_i^2(j)||$ be the matrix whose element

$$\tilde{c}_i^2(j)=\begin{cases} q_1+c_i^2(j) & \text{if } c_i^2(j)\neq *,\\ * & \text{otherwise.}\end{cases}$$

Then matrix

$$C=\begin{pmatrix} C^1 & * \\ * & \tilde{C}^2\end{pmatrix}$$

is a $C_{\mathrm{H}}(q_1+q_2,k)$-matrix. The similar statement is also true for characteristic matrices of hash&separable codes.

With the help of the computer checking, we constructed the finite collection of "non-regular" $C_{\mathrm{H}}(q,4)$-matrices, $q\geqslant 8$, and $C_{\mathrm{HS}}(q,4)$-matrices, $q\geqslant 13$. Taking into account Proposition 8, we obtain

Proposition 9. 1. *If $q\geqslant 8$, then there exists $C_{\mathrm{H}}(q,4)$-matrix*. 2. *If $q\geqslant 13$, then there exists $C_{\mathrm{HS}}(q,4)$-matrix.*

The following statement is a generalization of the hash&separable construction of Examples 3 and 4.

Proposition 10. *If $q\geqslant k^2$, then there exists $(q,k,3)$-homogeneous 3-hash code.*

Proof. Let $k=2,3,\ldots$ and $q\geqslant k^2$. Consider the following construction of $(q,k,3)$-homogeneous code $B=||b_j(u)||$ whose rows

$$\mathbf{b}_j=(b_j(1),b_j(2),\ldots,b_j(kq)),\quad j=1,2,3$$

are defined as follows:

1. for $j=1$, the first row

$$\mathbf{b}_1=(\mathbf{b}_1^{(1)},\mathbf{b}_1^{(2)},\ldots,\mathbf{b}_1^{(q)}),\quad \mathbf{b}_1^{(m)}=\underbrace{(m,m,\ldots,m)}_{k},\quad m=1,2,\ldots,q;$$

2. for $j=2$, the second row
$$\mathbf{b}_2=(\mathbf{b}_2^{(1)},\mathbf{b}_2^{(2)},\dots,\mathbf{b}_2^{(k)}),\quad \mathbf{b}_2^{(m)}=(1,2,\dots,q),\quad m=1,2,\dots,k;$$
3. for $j=3$, the third row $\mathbf{b}_3=(\mathbf{b}_3^{(1)},\mathbf{b}_3^{(2)},\dots,\mathbf{b}_3^{(k)})$, where the subsequence $\mathbf{b}_3^{(m)}$, $m=1,2,\dots,k$ of length q is the $k(m-1)$-*step cyclic shift* of the sequence $(1,2,\dots,q)$:
$$\mathbf{b}_3^{(m)}=\begin{cases}(1,2,\dots,q) & \text{if } m=1,\\ (k(m-1)+1,k(m-1)+2,\dots,q-1,q,1,2,\dots,k(m-1)) & \\ & \text{if } m=2,3,\dots,k.\end{cases}$$

As an illustration, we yield the $(11,3,3)$-homogeneous code
$$\begin{pmatrix} 111\ 222\ 333\ 444\ 555\ 666\ 777\ 888\ 999\ aaa\ bbb\\ 123\ 456\ 789\ ab1\ 234\ 567\ 89a\ b12\ 345\ 678\ 9ab\\ 123\ 456\ 789\ ab4\ 567\ 89a\ b12\ 378\ 9ab\ 123\ 456\end{pmatrix},$$
where, for convenience of notations, we put $a=10,\ b=11$.

If $q\geqslant k^2$, then this construction of $(q,k,3)$-homogeneous code B has an evident property of *alphabet separation*, which could be formulated as follows. *Let the symbol* $\oplus$ *denote modulo kq addition and* $u=1,2,\dots,kq$ *be an arbitrary fixed integer. Then q-nary elements of the k-subsequence*
$$b_3(u),b_3(u\oplus 1),b_3(u\oplus 2),\dots,b_3(u\oplus(k-1))$$
do not may occur in the k-subsequence
$$b_3(u\oplus q),b_3(u\oplus(q+1)),b_3(u\oplus(q+2)),\dots,b_3(u\oplus(q+k-1)).$$
By virtue of the second condition of Proposition 6, it implies 3-hash property of code B. Proposition 10 is proved. □

Conjecture. *The construction of Proposition* 10 *yields hash&separable codes.*

8.4. Product of characteristic matrices

In this section, we consider a construction of homogeneous codes, which makes possible to obtain the new (more complicated) codes using the known ones.

Let $v=1,2$ and $C^v=||c_i^v(j)||$, $i,j\in[q_v]$, $c_i^v(j)\in\{*,[q_v]\}$, be $C(q_v,k_v)$-matrix of code B_v. Denote by
$$C=C^1\Diamond C^2=||c_r(u)||,\quad r,u\in[q_1q_2],\ \ c_r(u)\in\{*,[q_1q_2]\},$$
the *product of characteristic matrices of code* B_1 *and code* B_2. Matrix C is defined as follows: for arbitrary $i,j\in[q_1]$ and $l,m\in[q_2]$, put
$$r=q_2(i-1)+l,\quad u=q_2(j-1)+m,$$
$$c_r(u)=\begin{cases} q_2(c_i^1(j)-1)+c_l^2(m) & \text{if } c_i^1(j)\neq * \text{ and } c_l^2(m)\neq *,\\ * & \text{otherwise.}\end{cases}$$

Example 11. Let $k_1 = k_2 = 2$, $q_1 = q_2 = 3$, and

$$C_{\mathrm{H}}(q_1, k_1) = C_{\mathrm{H}}(q_2, k_2) = \begin{pmatrix} * & 1 & 2 \\ 1 & * & 3 \\ 2 & 3 & * \end{pmatrix}.$$

$$C_{\mathrm{H}}(q_1 q_2, k_1 k_2) = C_{\mathrm{H}}(q_1, k_1) \diamondsuit C_{\mathrm{H}}(q_2, k_2)$$

$$= \begin{pmatrix} * & * & * & * & 1 & 2 & * & 4 & 5 \\ * & * & * & 1 & * & 3 & 4 & * & 6 \\ * & * & * & 2 & 3 & * & 5 & 6 & * \\ * & 1 & 2 & * & * & * & * & 7 & 8 \\ 1 & * & 3 & * & * & * & 7 & * & 9 \\ 2 & 3 & * & * & * & * & 8 & 9 & * \\ * & 4 & 5 & * & 7 & 8 & * & * & * \\ 4 & * & 6 & 7 & * & 9 & * & * & * \\ 5 & 6 & * & 8 & 9 & * & * & * & * \end{pmatrix}.$$

Such product of matrices remains the hash property.

Example 12. Let $k_1 = k_2 = 2$, $q_1 = q_2 = 3$, and

$$C_{\mathrm{S}}(q_1, k_1) = C_{\mathrm{S}}(q_2, k_2) = \begin{pmatrix} 1 & 2 & * \\ * & 1 & 3 \\ 3 & * & 2 \end{pmatrix}.$$

The product of matrices

$$C(q_1 q_2, k_1 k_2) = C_{\mathrm{S}}(q_1, k_1) \diamondsuit C_{\mathrm{S}}(q_2, k_2)$$

$$= \begin{pmatrix} 1 & \grave{2} & * & \acute{4} & 5 & * & * & * & * \\ * & \acute{1} & 3 & * & \grave{4} & 6 & * & * & * \\ 3 & * & 2 & 6 & * & 5 & * & * & * \\ * & * & * & \grave{1} & \acute{2} & * & \bar{7} & \hat{8} & * \\ * & * & * & * & 1 & 3 & * & 7 & 9 \\ * & * & * & 3 & * & 2 & \dot{9} & * & \ddot{8} \\ 7 & \bar{8} & * & * & * & * & \hat{4} & 5 & * \\ * & \hat{7} & \ddot{9} & * & * & * & * & \bar{4} & \dot{6} \\ 9 & * & \dot{8} & * & * & * & \ddot{6} & * & 5 \end{pmatrix}$$

does not remain the separable properties of factors. In the figure, we have *marked* three pairs of "bad" triples, namely:

$$\{(\grave{1},\grave{2},\grave{4})\ (\acute{1},\acute{2},\acute{4})\},\quad \{(\bar{4},\bar{7},\bar{8})\ (\hat{4},\hat{7},\hat{8})\},\quad \{(\dot{6},\dot{8},\dot{9})\ (\ddot{6},\ddot{8},\ddot{9})\}.$$

This example shows the reason why the separable property of the product of two separable matrices is not true. To guarantee the separable property of the product of two separable matrices, *at least one of two factors should have hash&separable property*. The following proposition takes place.

Proposition 11. (1) *The product of* $C_{\rm H}(q_1,k_1)$*- and* $C_{\rm H}(q_2,k_2)$*-matrices is* $C_{\rm H}(q_1q_2,k_1k_2)$*-matrix.* (2) *The product of* $C_{\rm S}(q_1,k_1)$*- and* $C_{\rm HS}(q_2,k_2)$*-matrices is* $C_{\rm S}(q_1q_2,k_1k_2)$*-matrix. In addition, if the product of two separable matrices has the separable property, then at least one of these factors should have the hash&separable property* (S.M. Ekhanin, 1998).

To explain the second statement of Proposition 11, we give the following example.

Example 13. Let $k_1=k_2=2$, $q_1=3$, $q_2=4$, and

$$C_{\rm S}(q_1,k_1)=\begin{pmatrix} 1 & 2 & * \\ * & 1 & 3 \\ 3 & * & 2 \end{pmatrix},\quad C_{\rm HS}(q_2,k_2)=\begin{pmatrix} 1 & * & 3 & * \\ * & 1 & * & 3 \\ 4 & * & 2 & * \\ * & 4 & * & * \end{pmatrix}.$$

The product $C_{\rm S}(q_1,k_1)\Diamond C_{\rm HS}(q_2,k_2)$ has the form

$$\begin{pmatrix}
1 & * & 3 & * & 5 & * & 7 & * & * & * & * & * \\
* & 1 & * & 3 & * & 5 & * & 7 & * & * & * & * \\
4 & * & 2 & * & 8 & * & 6 & * & * & * & * & * \\
* & 4 & * & 2 & * & 8 & * & 6 & * & * & * & * \\
* & * & * & * & 1 & * & 3 & * & 9 & * & 11 & * \\
* & * & * & * & * & 1 & * & 3 & * & 9 & * & 11 \\
* & * & * & * & 4 & * & 2 & * & 12 & * & 10 & * \\
* & * & * & * & * & 4 & * & 2 & * & 12 & * & 10 \\
9 & * & 11 & * & * & * & * & * & 5 & * & 7 & * \\
* & 9 & * & 11 & * & * & * & * & * & 5 & * & 7 \\
12 & * & 10 & * & * & * & * & * & 8 & * & 6 & * \\
* & 12 & * & 10 & * & * & * & * & * & 8 & * & 6
\end{pmatrix}$$

which illustrates its separable property. The changed order product $C_{\mathrm{HS}}(q_2,k_2)\Diamond C_{\mathrm{S}}(q_1,k_1)$ also remains the separable property and has the form

$$\begin{pmatrix}
1 & 2 & \ast & \ast & \ast & \ast & 7 & 8 & \ast & \ast & \ast & \ast \\
\ast & 1 & 3 & \ast & \ast & \ast & \ast & 7 & 9 & \ast & \ast & \ast \\
3 & \ast & 2 & \ast & \ast & \ast & 9 & \ast & 8 & \ast & \ast & \ast \\
\ast & \ast & \ast & 1 & 2 & \ast & \ast & \ast & \ast & 7 & 8 & \ast \\
\ast & \ast & \ast & \ast & 1 & 3 & \ast & \ast & \ast & \ast & 7 & 9 \\
\ast & \ast & \ast & 3 & \ast & 2 & \ast & \ast & \ast & 9 & \ast & 8 \\
10 & 11 & \ast & \ast & \ast & \ast & 4 & 5 & \ast & \ast & \ast & \ast \\
\ast & 10 & 12 & \ast & \ast & \ast & \ast & 4 & 6 & \ast & \ast & \ast \\
12 & \ast & 11 & \ast & \ast & \ast & 6 & \ast & 5 & \ast & \ast & \ast \\
\ast & \ast & \ast & 10 & 11 & \ast & \ast & \ast & \ast & 4 & 5 & \ast \\
\ast & \ast & \ast & \ast & 10 & 12 & \ast & \ast & \ast & \ast & 4 & 6 \\
\ast & \ast & \ast & 12 & \ast & 11 & \ast & \ast & \ast & 6 & \ast & 5
\end{pmatrix}$$

From Propositions 9, 11 and Example 13 it follows the statement of Theorem 4.

9. Proof of Theorem 5

Let $s \geqslant 2$, $l \geqslant 1$ be fixed integers and $n > 2s + l$ be an arbitrary integer. Let $[n] = \{1,2,\ldots,n\}$ be the set of integers from 1 to n and $\mathscr{E}(s,n)$ be the collection of all $\binom{n}{s}$ s-subsets of $[n]$. Following Macula (1996), we define the binary code $X = ||x_B(A)||$, $B \in \mathscr{E}(s,n)$, $A \in \mathscr{E}(s+l,n)$, of size $t = \binom{n}{s+l}$ and length $N = \binom{n}{s}$, whose element $x_B(A) = 1$ if and only if $B \subset A$. One can easily understand that X is the constant weight code with parameters:

$$t = \begin{pmatrix} n \\ s+l \end{pmatrix}, \quad N = \begin{pmatrix} n \\ s \end{pmatrix}, \quad k = \begin{pmatrix} n-s \\ l \end{pmatrix}, \quad w = \begin{pmatrix} s+l \\ s \end{pmatrix}, \quad \lambda = \begin{pmatrix} s+l-1 \\ s \end{pmatrix},$$

where t-code size, N-code length, w-weight of columns (codewords), k-weight of rows and λ-the maximal dot product of codewords. In addition, let $A_0, A_1, \ldots, A_s$, $A_i \in \mathscr{E}(s+l,n)$ be an arbitrary $(s+1)$-collection of pairwise different $(s+l)$-subsets of $[n]$. Since $A_0 \neq A_i$, for any $i = 1,2\ldots,s$, there exists an element $a_i \in A_0$ and $a_i \notin A_i$. Hence, there exists a s-subset $B \subset A_0$ and for any $i = 1,2,\ldots,s$, $B \not\subset A_i$. It follows that X is a superimposed (s,t,k)-code. For the particular case $l = 1$, these properties yield Theorem 5.

References

Du, D.-Z., Hwang, F.K., 1993. Combinatorial Group Testing and its Applications. World Scientific, Singapore, New Jersey, London, Hong Kong.

D'yachkov, A.G., Rykov, V.V., 1983. A survey of superimposed code theory. Problems Control Inform. Theory 12 (4), 229–242.

Ekhanin, S.M., 1998. Some new constructions of optimal superimposed designs. Proc. of the 6-th International Workshop "Algebraic and Combinatorial Coding Theory", ACCT-6, Pskov, Russia, 232–235.

Fridman, M.L., Komlos, J., 1984. On the size of separating systems and families of perfect hash functions. SIAM J. Algebraic Discrete Methods 5, 538–544.

Katona, G., 1966. On separating system of a finite set. J. Combin. Theory 1 (2), 174–194.

Kautz, W.H., Singleton, R.C., 1964. Nonrandom binary superimposed codes. IEEE Trans. Inform. Theory 10 (4), 363–377.

Macula, A.J., 1996. A simple construction of d-disjunct matrices with certain constant weights. Discrete Math. 162, 311–312.

Renyi, A., 1965. On the theory of random search. Bull. Amer. Math. Soc. 71 (6), 809–828.