IMA workshop "Group Testing Designs, Algorithms, and Applications to Biology",
University of Minnesota, February 13-17, 2012. Minneapolis, Minnesota 55455-0436, USA

Superimposed Codes and Designs for Group Testing Models

A. G. D'yachkov, V. V. Rykov

Moscow State University, Faculty of Mechanics and Mathematics,
Department of Probability Theory, Moscow, 119992, Russia,
agd-msu@yandex.ru, vrykov@unomaha.edu

Abstract- We will discuss superimposed codes and non-adaptive group testing designs arising from the potentialities of compressed genotyping models in molecular biology. The given lecture was motivated by the 30th anniversary of our recurrent upper bound on the rate of superimposed codes published in 1982. We were also inspired by recent results obtained for non-adaptive threshold group testing which essentially develop the theory of superimposed codes.

Index terms. Group testing, compressed genotyping, screening experiments, search designs, superimposed codes, rate of codes, rate of designs, bounds on the rate, shortened RC-code, threshold search designs.

1 Notations, Definitions and Relevant Issues

Let $[n]$ be the set of integers from 1 to n and the symbol $\triangleq$ denote definitional equalities. For integers $N \geq 2$ and $t \geq 2$, symbols $\Omega_j \subset [N]$, $j = 1, 2, \ldots, t$, denote subsets of $[N]$. Subsets Ω_j, $j \in [t]$, are identified with binary columns $\mathbf{x}(j) \triangleq (x_1(j), x_2(j), \ldots, x_N(j))$ in which

$$x_i(j) \triangleq \begin{cases} 1 & \text{if } i \in \Omega_j, \\ 0 & \text{if } i \notin \Omega_j, \quad i \in [N]. \end{cases}$$

An incidence matrix $X \triangleq \|x_i(j)\|$, $i \in [N]$, $j \in [t]$, is called a *code* with t codewords (columns) $\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(t)$ of length N corresponding to a *family of subsets* $\Omega_1, \Omega_2, \ldots, \Omega_t$.

Let $P \subset [t]$ be an arbitrary fixed subset of $[t]$ and $|P|$ is its size, i.e.,

$$P \triangleq \{p_1, p_2, \ldots, p_{|P|}\} \subset [t], \quad 1 \leq p_1 < p_2 < \cdots < p_{|P|} \leq t.$$

Denote by $\mathcal{P}(t, \leq s)$ $(\mathcal{P}(t, = s))$ the collection of all $\sum_{i=0}^{s} \binom{t}{i}$ $\left(\binom{t}{s}\right)$ subsets P of size $|P| \leq s$ $(|P| = s)$. Let $N \geq 2$ be an integer and $\mathsf{A} = \{A_1, A_1, \ldots, A_N\}$, $A_i \subset [t]$, $i \in [N]$, is a fixed family of subsets of $[t]$. Subsets A_i are identified with binary rows $\mathbf{x}_i \triangleq (x_i(1), x_i(2), \ldots, x_i(t))$ in which

$$x_i(j) \triangleq \begin{cases} 1 & \text{if } j \in A_i, \\ 0 & \text{if } j \notin A_i, \quad i \in [N],\ j \in [t]. \end{cases}$$

We will identify the family A with its incidence matrix (code) $X = \|x_i(j)\|$, $i \in [N]$, $j \in [t]$.

In the theory of *group testing* [23] (*designing screening experiments* [20]) the given, in advance, family $\mathsf{A} = \{A_1, A_1, \ldots, A_N\}$ is interpreted as a *non-adaptive search design* consisting of N group tests (experiments) A_i, $i \in [N]$. An experimenter wants to construct group tests A_i, $i \in [N]$, to carry out the corresponding experiments and then to identify an *unknown subset* $P \subset [t]$ with the help of test outcomes provided that $P \subset \mathcal{P}(t, \le s)$ or $P \subset \mathcal{P}(t, = s)$, where $s \ll t$. If for each test A_i, $i \in [N]$, its outcome *depends only on the size of intersection*

$$|P \cap A_i| = \sum_{m=1}^{|P|} x_i(p_m), \quad i \in [N],$$

then we will say that a *symmetric model* [20] of non-adaptive search design is considered.

The aim of our lecture is to present the principal combinatorial results for the symmetric search model[1]. In Sect. 2, we give a brief survey of necessary definitions and bounds on the rate of superimposed codes which are the *base for studying* of non-adaptive group testing models.

In Sect. 3, we introduce the concept of non-adaptive group testing designs arising from the potentialities of compressed genotyping models in molecular biology and establish a universal upper bound on their rate. The universal bound is prescribed by our recurrent upper bound on the rate of classical superimposed codes.

In Sect. 4, we remind our constructions of superimposed codes based on shortened RS-codes. These constructions are presented in papers [13]-[17], where we significantly extended optimal and suboptimal construction of classical superimposed codes suggested in [1]. Note that we included in [13]-[14] the detailed tables with parameters of the best known superimposed codes. We don't mention other authors because, unfortunately, we don't know any papers containing significant results, i.e., the similar or improved tables of parameters. The extension of our tables is the important open problem.

In Sect. 5, the threshold group testing model is discussed. We apply the conventional terminology of superimposed code theory to refine the description of a new upper bound on the rate of threshold designs recently obtained in [30].

2 Superimposed (z, u)-Codes

Let z and u be positive integers such that $z + u \le t$.

Definition 1. [17]. A family of subsets $\Omega_1, \Omega_2, \ldots, \Omega_t$, where $\Omega_j \subseteq [N]$, $j \in [t]$, is called an (z, u)*–cover-free family* if for any two non-intersecting subsets Z, $U \subset [t]$, $Z \cap U = \emptyset$, such that $|Z| = z$, $|U| = u$, the following condition holds:

$$\bigcap_{j \in U} \Omega_j \not\subseteq \bigcup_{j \in Z} \Omega_j.$$

An incidence matrix $X = \|x_i(j)\|$, $i \in [N]$, $j \in [t]$, corresponding to (z, u)–cover-free family is called a *superimposed* (z, u)*-code*.

The following evident necessary and sufficient condition for Definition 1 takes place.

Proposition 1. [17]. *Any binary* $(N \times t)$*-matrix* X *is a superimposed* (z, u)*-code if and only if for any two subsets* Z, $U \subset [t]$, *such that* $|Z| = z$, $|U| = u$ *and* $Z \cap U = \emptyset$ *the matrix* X *contains a row* $\mathbf{x}_i = (x_i(1), x_i(2) \ldots, x_i(t)$, *for which*

$$x_i(j) = 1 \quad \text{for all} \quad j \in U, \qquad x_i(j) = 0 \quad \text{for all} \quad j \in Z.$$

[1] We don't discuss here the general noisy symmetric model of non-adaptive search designs which can be described using the terminology of multiple access channel (MAC) [3]. An interested reader is referred to [20]. The information theory problems for non-symmetric search model are considered in [31]-[36].

Let $t(N,z,u)$ be the maximal possible size of superimposed (z,u)-codes. For fixed $1 \le u < z$, define a *rate* of (z,u)-codes:

$$R(z,u) \triangleq \overline{\lim_{N\to\infty}} \frac{\log_2 t(N,z,u)}{N}.$$

For the classical case $u=1$, superimposed $(z,1)$–codes and their applications were introduced by W.H Kautz, R.C. Singleton in [1]. Further, these codes along with new applications were investigated in [4]-[20]. In papers [4],[9],[17], we established the best known upper and lower bounds on the rate $R(z,1)$.

2.1 Recurrent Upper Bounds on $R(z,1)$ and $R(z,u)$

Let $h(\alpha) \triangleq -\alpha\log_2\alpha - (1-\alpha)\log_2(1-\alpha)$, $0<\alpha<1$, be the binary entropy. To formulate an *upper bound* on the rate $R(z,1)$, $z \ge 1$, we introduce the function [4]

$$\mathsf{f}_z(\alpha) \triangleq h(\alpha/z) - \alpha\, h(1/z), \quad z = 1,2,\ldots,$$

of argument α, $0<\alpha<1$.

Theorem 1. [4]-[5]. (Recurrent upper bound on $R(z,1)$). *If $z=1,2,\ldots$, then the rate $R(z,1) \le \overline{R}(z,1)$, where*

$$\overline{R}(1,1) = R(1,1) = 1, \quad \overline{R}(2,1) \triangleq \max_{0<\alpha<1} \mathsf{f}_2(\alpha) \;=\; 0.321928 \tag{1}$$

and sequence $\overline{R}(z,1)$, $z=3,4,\ldots$, is defined as the unique solution of recurrent equation

$$\overline{R}(z,1) = \mathsf{f}_z\left(1 - \frac{\overline{R}(z,1)}{\overline{R}(z-1,1)}\right). \tag{2}$$

Up to now, i.e., during more than 30 years, the recurrent sequence $\overline{R}(z,1)$, $z=1,2,\ldots$, defined by (1)-(2) and called a *recurrent upper bound* is the best known upper bound on the rate $R(z,1)$. The reciprocal values of $\overline{R}(z,1)$, i.e., numbers $1/\overline{R}(z,1) > z$, $z=2,3,\ldots,17$, taken from [5], are given in **Table 1.**

z	$1/\overline{R}(z,1)$	z	$1/\overline{R}(z,1)$	z	$1/\overline{R}(z,1)$	z	$1/\overline{R}(z,1)$
2	3.1063	6	12.0482	10	24.5837	14	40.3950
3	5.0180	7	14.8578	11	28.2402	15	44.8306
4	7.1196	8	17.8876	12	32.0966	16	49.4536
5	9.4660	9	21.1313	13	36.1493	17	54.2612

Table 1.

Applying Theorem 1 and the corresponding calculus arguments, we proved

Theorem 2. [4]-[5]. (Non-recurrent upper bound on $R(z,1)$). *For any $z \ge 2$, the rate $R(z,1)$ satisfies inequality*

$$R(z,1) \;\le\; \frac{2\log_2[e(z+1)/2]}{z^2}, \qquad z=2,3,\ldots,$$

which leads to the asymptotic inequality

$$R(z,1) \;\le\; \frac{2\log_2 z}{z^2}\,(1+o(1)), \quad z\to\infty.$$

Theorem 3. [21] (Recurrent inequality for $R(z,u)$). *If $z \ge u \ge 2$, then for any $i \in [z-1]$ and $j \in [u-1]$, the rate*

$$R(z,u) \le \frac{R(z-i,u-j)}{R(z-i,u-j) + \frac{(i+j)^{i+j}}{i^i \cdot j^j}}. \tag{3}$$

Recurrent inequality (3) and the known numerical values of recurrent upper bound $\overline{R}(z,1)$, $z = 1,2,\ldots$, defined by (1)-(2), give numerical values of the best known upper bound $\overline{R}(z,u)$ on the rate $R(z,u)$, $z \ge u \ge 2$. An asymptotic consequence from the given upper bound is presented by

Theorem 4. [18] *If $z \to \infty$ and $u \ge 2$ is fixed, then*

$$R(z,u) \le \overline{R}(z,u) \le \frac{(u+1)^{u+1}}{2\,e^{u-1}} \cdot \frac{\log z}{z^{u+1}} \cdot (1+o(1)).$$

2.2 Random Coding Lower Bounds on $R(z,u)$ and $R(z,1)$

Theorem 5. [17] *A random coding lower bound on the rate $R(z,u)$ has the form:*

$$R(z,u) \ge -(z+u-1)^{-1} \log_2 \left(1 - \frac{z^z\, u^u}{(z+u)^{z+u}}\right), \quad 1 \le u < z.$$

If $u \ge 1$ is fixed and $z \to \infty$, then the asymptotic form of the given lower bound is

$$R(z,u) \ge \frac{e^{-u} \cdot u^u \cdot \log_2 e}{z^{u+1}} \cdot (1+o(1)).$$

If $u = 1$, then the best known random coding lower bound on the rate $R(z,1)$ is given by

Theorem 6. [10] *For any $z = 1,2,\ldots$, the rate $R(z,1) \ge \underline{R}(z,1) \triangleq \frac{A(z)}{z}$, where*

$$A(z) \triangleq \max_{0<\alpha<1,\, 0<Q<1} \left\{ -(1-Q)\log(1-\alpha^z) + z\left(Q \log\frac{\alpha}{Q} + (1-Q)\log\frac{1-\alpha}{1-Q}\right)\right\}.$$

If $z \to \infty$, then the rate

$$R(z,1) \ge \underline{R}(z,1) = \frac{1}{z^2 \log e}(1+o(1)) = \frac{0.693}{z^2}(1+o(1)).$$

In the first and second rows of **Table 2**, we give values of $\underline{R}(s,1) < 1/s$, $s = 2,3\ldots,8$, along with the corresponding values of $\overline{R}(s,1) < 1/s$, $s = 2,3\ldots,8$, taken from **Table 1**.

s	2	3	4	5	6	7	8
$\underline{R}(s,1)$	.182	.079	.044	.028	.019	.014	.011
$\overline{R}(s,1)$	.322	.199	.140	.106	.083	.067	.056
$\overline{R}_1(\le s)$	1/2	.199	.199	.106	.106	.067	.067
$\overline{R}_2(\le s)$	-	1/3	1/4	.140	.140	.083	.083
$\overline{R}_3(\le s)$	-	-	1/4	1/5	1/6	.106	.106
$\underline{R}_1(=s)$	.302	.142	.082	.053	.037	.027	.021

Table 2

3 $\left(F^\ell, \le s\right)$–Designs, $\left(F^\ell, = s\right)$–Designs and $\mathcal{D}_s^\ell$–Codes

Using notations of Sect. 1, we give

Definition 2. Let ℓ, $1 \le \ell < s < t$ be integers and $F^\ell = F^\ell(n)$, $n = 0, 1, \ldots, \ell$, is an arbitrary fixed function of integer argument $n = 0, 1, \ldots, \ell$ such that for any $n = 0, 1, \ldots, \ell - 1$, its value $F^\ell(n) \ne F^\ell(\ell)$. Define the vector

$$\mathbf{y}^\ell(P, \mathsf{A}) \triangleq \left(y_1^\ell, y_2^\ell, \ldots, y_N^\ell\right), \quad y_i^\ell \triangleq \begin{cases} F^\ell(n) & \text{if } |P \cap A_i| = n, \quad n = 0, 1, \ldots, \ell - 1, \\ F^\ell(\ell) & \text{if } |P \cap A_i| \ge \ell, \quad i \in [N]. \end{cases}$$

or

$$\mathbf{y}^\ell(P, X) \triangleq \left(y_1^\ell, y_2^\ell, \ldots, y_N^\ell\right), \quad y_i^\ell \triangleq \begin{cases} F^\ell(n) & \text{if } \sum\limits_{m=1}^{|P|} x_i(p_m) = n, \quad n = 0, 1, \ldots, \ell - 1, \\ F^\ell(\ell) & \text{if } \sum\limits_{m=1}^{|P|} x_i(p_m) \ge \ell, \quad i \in [N]. \end{cases}$$

A code X of length N and size t is called an $\left(F^\ell, \le s\right)$*–design*, ($\left(F^\ell, = s\right)$*–design*), $1 \le \ell < s < t$, *for group testing model* if $\mathbf{y}^\ell(P, X) \ne \mathbf{y}^\ell(P', X)$ for any

$$P \ne P', \quad P \in \mathcal{P}(t, \le s),\ P' \in \mathcal{P}(t, \le s) \quad \left(P \in \mathcal{P}(t, = s),\ P' \in \mathcal{P}(t, = s)\right).$$

Remark 1. $\left(F^\ell, \le s\right)$–design and $\left(F^\ell, = s\right)$–design are examples of the important search model called *designing screening experiments* [10],[20]. For these examples, which can be interpreted as compressed genotyping [29] models in molecular biology, O. Milenkovic (2011) suggested the name "Semi-Quantitative Group Testing".

Let $1 \le \ell < s < t$ be integers. For any set $\mathcal{S} \subset [t]$ of size $|\mathcal{S}| = s$, we denote by $\binom{\mathcal{S}}{\ell}$ the collection of all $\binom{s}{\ell}$ ℓ–subsets of the set $\mathcal{S}$.

Definition 3. [6]. A family of subsets $\Omega_1, \Omega_2, \ldots, \Omega_t$ is called an $\mathcal{D}_s^\ell$*–family* if for any $\mathcal{S} \subset [t]$, $|\mathcal{S}| = s$, and any $j \notin \mathcal{S}$, the subset

$$\Omega_j \not\subseteq \bigcup_{\binom{\mathcal{S}}{\ell}} \left\{\bigcap_{k=1}^{\ell} \Omega_{j_k}\right\}, \quad \binom{\mathcal{S}}{\ell} \triangleq \{(j_1, j_2, \ldots, j_\ell) \;:\; j_i \in \mathcal{S}, \quad j_1 < j_2 < \ldots < j_\ell\},\ 1 \le \ell < s < t.$$

An incidence matrix $X = \|x_i(j)\|$, $i \in [N]$, $j \in [t]$, corresponding to $\mathcal{D}_s^\ell$– family is called a *superimposed* $\mathcal{D}_s^\ell$*-code* (briefly, $\mathcal{D}_s^\ell$*-code*).

One can easily check the following

Proposition 2. *Any binary $(N \times t)$-matrix X is a $\mathcal{D}_s^\ell$–code, $1 \le \ell < s < t$, if and only if for any collection of $s+1$ integers $j_1, j_2, \ldots, j_s, j_{s+1}$, $j_k \ne j_m$, $j_k \in [t]$, there exists $i \in [N]$ such that*

$$x_i(j_{s+1}) = 1, \qquad \sum_{k=1}^{s} x_i(j_k) \le \ell - 1.$$

For $\ell = 1$ and $s = 2, 3\ldots$, the definition of $\mathcal{D}_s^1$-code coincides with the definition of superimposed $(s, 1)$-code. In addition, if $1 \le \ell < s - 1$, then any $\mathcal{D}_s^\ell$-code is a $\mathcal{D}_s^{\ell+1}$-code.

Remark 2. For $s > \ell \ge 2$, $\mathcal{D}_s^\ell$-codes were suggested in [6] for the study of some communication systems with random multiple access.

3.1 Universal Upper Bound for $(F^\ell, \le s)$–Designs

Let $t\left(N, \mathcal{D}_s^\ell\right)$, $t\left(N, F^\ell, \le s\right)$ and $t\left(N, F^\ell, = s\right)$ be the maximal size of superimposed $\mathcal{D}_s^\ell$–codes, $(F^\ell, \le s)$–designs and $(F^\ell, = s)$–designs. For fixed $1 \le \ell < s$, define the corresponding *rates*:

$$R\left(\mathcal{D}_s^\ell\right) \triangleq \overline{\lim_{N\to\infty}} \frac{\log_2 t\left(N, \mathcal{D}_s^\ell\right)}{N}, \quad 1 \le \ell < s,$$

$$R\left(F^\ell, \le s\right) \triangleq \overline{\lim_{N\to\infty}} \frac{\log_2 t\left(N, F^\ell, \le s\right)}{N}, \qquad R\left(F^\ell, = s\right) \triangleq \overline{\lim_{N\to\infty}} \frac{\log_2 t\left(N, F^\ell, = s\right)}{N}$$

Obviously, for any $1 \le \ell < s$, the following inequalities hold:

$$t\left(N, F^\ell, \le s\right) \le t\left(N, F^\ell, = s\right), \quad R\left(F^\ell, \le s\right) \le R\left(F^\ell, = s\right) \le \frac{\log_2(\ell+1)}{s}. \tag{4}$$

Proposition 3. [6]. *If $1 \le \ell < s-1$, then any $(F^\ell, \le s)$–design is a superimposed $\mathcal{D}_{s-1}^\ell$–code, i.e.,*

$$t\left(N, F^\ell, \le s\right) \le t\left(N, \mathcal{D}_{s-1}^\ell\right), \qquad R\left(F^\ell, \le s\right) \le R\left(\mathcal{D}_{s-1}^\ell\right), \quad 1 \le \ell < s-1.$$

Proof. By contradiction. If a code $X = \|x_i(j)\|$, $i \in [N]$, $j \in [t]$ doesn't satisfy the definition of $\mathcal{D}_{s-1}^\ell$–code, then in virtue of Proposition 1, there exists a collection of s integers $j_1, j_2, \dots, j_{s-1}, j_s$, $j_k \ne j_m$, $j_k \in [t]$, such that for any $i \in [N]$,

$$x_i(j_s) = 1 \quad \Longrightarrow \quad \sum_{k=1}^{s-1} x_i(j_k) \ge \ell.$$

Hence, for $(s-1)$-subset $P \triangleq \{j_1, j_2, \dots, j_{s-1}\} \subset [t]$ and s-subset $P' \triangleq \{j_1, j_2, \dots, j_{s-1}, j_s\} \subset [t]$, the vector $\mathbf{y}^\ell(P, X) = \mathbf{y}^\ell(P', X)$. This contradicts to the definition of $(F^\ell, \le s)$–design.

Theorem 7. (De Bonis, Vaccaro [25]). *For any $1 \le \ell < s$, the rate $R\left(\mathcal{D}_s^\ell\right)$ of superimposed $\mathcal{D}_s^\ell$–codes satisfies inequality*

$$R\left(\mathcal{D}_s^\ell\right) \le R\left(\left\lfloor \frac{s}{\ell} \right\rfloor, 1\right).$$

where $R(z,1)$, $z \ge 1$, is the rate of classical superimposed $(z,1)$-codes.

Proposition 3 and Theorem 7 lead to inequalities:

$$R\left(F^\ell, \le s\right) \le R\left(\mathcal{D}_{s-1}^\ell\right) \le R\left(\left\lfloor \frac{s-1}{\ell} \right\rfloor, 1\right) \le \overline{R}\left(\left\lfloor \frac{s-1}{\ell} \right\rfloor, 1\right), \quad 1 \le \ell \le s, \tag{5}$$

where $\overline{R}(z,1)$ is the recurrent upper bound on the rate $R(z,1)$ presented by Theorem 1. For instance, if $(\ell = 3, s = 10)$ or $(\ell = 3, s = 13)$, then Table 2 shows that

$$\overline{R}(3,1) = .199 < .200 = 2/10 \quad \text{or} \quad \overline{R}(4,1) = .140 < .154 = 2/13,$$

i.e., for $\ell = 3$ and $s = 3k+1$, $k = 3, 4, \dots$, bound (5) improves the trivial bound (4).

From inequalities (4)-(5), it follows

Proposition 4. (Universal upper bound). *For any $(F^\ell, \le s)$–design, the rate*

$$R\left(F^\ell, \le s\right) \le \min\left\{\frac{\log_2 \ell}{s}; \overline{R}\left(\left\lfloor \frac{s-1}{\ell} \right\rfloor, 1\right)\right\}, \quad 1 \le \ell < s,$$

and the asymptotic inequality

$$R\left(F^\ell, \le s\right) \le \frac{2\ell^2 \log_2 s}{s^2}(1 + o(1)), \quad \ell = 1, 2, \dots, \qquad s \to \infty,$$

holds.

4 Constructions of Superimposed $(s,1)$-Codes and $\mathcal{D}_s^\ell$–Codes

4.1 Superimposed $(s,1)$-Codes and $\mathcal{D}_s^\ell$–Codes Based on Shortened Reed-Solomon Codes

Let $\mathcal{Q}$ be the set of all primes or prime powers ≥ 2, i.e.,

$$\mathcal{Q} \triangleq \{2,3,4,5,7,8,9,11,13,16,17,19,23,25,27,29,31,32,37,\ldots\}.$$

Let $q \in \mathcal{Q}$ and $2 \le k \le q+1$ be fixed integers for which there exists the q-ary Reed-Solomon code (RS-code) B of size q^k, length $(q+1)$ and the Hamming distance $d = q-k+2 = (q+1)-(k-1)$ [2]. We will identify the code B with an $\big((q+1)\times q^k\big)$–matrix whose columns, (i.e., $(q+1)$-sequences from the alphabet $\{0,1,2,\ldots,q-1\}$) are the codewords of B. Therefore, the maximal possible number of positions (rows) where its two codewords (columns) can coincide, called a *coincidence* of code B, is equal to $k-1$.

Fix an arbitrary integer $r = 0,1,2,\ldots,k-1$ and introduce the *shortened* RS-code $\tilde{B}$ of size $t = q^{k-r}$, length $n = q+1-r$ that has the same Hamming distance $d = q-k+2$. Code $\tilde{B}$ is obtained by the *shortening* of the *subcode* of B which contains $0's$ in the first r positions (rows) of B. Obviously, the coincidence of $\tilde{B}$ is equal to

$$\lambda \triangleq n-d = (q+1-r)-d = q+1-r-(q-k+2) = k-r-1. \tag{6}$$

Consider the following standard transformation of the q-ary code $\tilde{B}$, when each symbol of the q-ary alphabet $\{0,1,2,\ldots,q-1\}$ is substituted for the corresponding binary column of the length q and the weight 1, namely:

$$0 \Leftrightarrow \underbrace{(1,0,0,\ldots,0)}_{q}, \quad 1 \Leftrightarrow \underbrace{(0,1,0,\ldots,0)}_{q}, \quad \ldots \quad q-1 \Leftrightarrow \underbrace{(0,0,0,\ldots,1)}_{q}.$$

As a result we have a binary constant-weight code X of size t, length N and weight w, where

$$t = q^{k-r} = q^{\lambda+1}, \quad N = n\cdot q = (q+1-r)q, \quad w = n = q+1-r. \tag{7}$$

From Propositions 1-2 and (6), it follows

Proposition 5. *Let integers $1 \le \ell < s$ satisfy inequalities*

$$s[(k-1)-r] \;\le\; \ell\,(q+1-r)\;-\;1, \quad 2 \le k \le q+1, \quad 0 \le r \le k-1. \tag{8}$$

Then the binary constant-weight code X with parameters (7) is a $\mathcal{D}_s^\ell$–code if $2 \le \ell < s$, or X is a superimposed $(s,1)$-code if $\ell = 1$.

For $\ell = 1$, the detailed tables with parameters of the best known superimposed $(s,1)$-codes (or $\mathcal{D}_s^1$–codes) based on Proposition 5 are presented in our papers [13]-[14]. In Sect. 4.2, Table 3 gives an example of such table. In Table 3, we marked by the **boldface type** two *triples* of superimposed code parameters which were known from [1]. The rest triples of superimposed code parameters from Table 3 were obtained in [13]-[14].

Remark 3. For the general case of superimposed (z,u)-codes, $2 \le u < z$, the construction similar to Proposition 5 was developed in [17]. Another significant constructions of superimposed (z,u)-codes, $2 \le u < z$, were developed in [22].

4.2 Parameters of constant-weight superimposed $(s,1)$-codes $2 \le s \le 8$, of weight w, length N, size $t = q^{\lambda+1}$, $2^m \le t < 2^{m+1}$, $5 \le m \le 30$, based on the q-ary shortened Reed-Solomon codes.

s	2	3	4	5	6	7	8
$\underline{R}(s,1)$	.182	.079	.044	.028	.019	.014	.011
$\overline{R}(s,1)$	.322	.199	.140	.106	.083	.067	.056
m	q, λ, N	q, λ, N	q, λ, N	q, λ, N	q, λ, N	q, λ, N	q, λ, N
5	–	7, 1, 28	7, 1, 35	7, 1, 42	7, 1, 49	–	–
6	4, 2, 20	8, 1, 32	8, 1, 40	8, 1, 48	8, 1, 56	9, 1, 72	11, 1, 99
7	–	–	13, 1, 65	13, 1, 78	13, 1, 91	13, 1, 104	13, 1, 117
8	7, 2, 35	7, 2, 49	–	16, 1, 96	16, 1, 112	16, 1, 128	16, 1, 144
9	8, 2, 40	8, 2, 56	8, 2, 72	–	23, 1, 161	23, 1, 184	23, 1, 207
10	–	11, 2, 77	11, 2, 99	11, 2, 121	–	–	–
11	7, 3, 49	–	13, 2, 117	13, 2, 143	13, 2, 169	–	–
12	8, 3, 56	9, 3, 90	16, 2, 144	16, 2, 176	16, 2, 208	16, 2, 240	**16, 2, 272**
$\frac{12}{N}$	.214	.133	.083	.068	.058	.050	.044
13	–	11, 3, 110	–	23, 2, 253	23, 2, 299	23, 2, 345	23, 2, 391
14	–	13, 3, 130	13, 3, 169	–	27, 2, 351	27, 2, 405	27, 2, 459
15	8, 4, 72	–	–	–	–	32, 2, 480	32, 2, 544
16	–	16, 3, 160	16, 3, 208	16, 3, 256	19, 3, 361	–	–
17	11, 4, 99	–	–	–	–	–	–
18	13, 4, 117	13, 4, 169	–	23, 3, 368	23, 3, 437	23, 3, 506	25, 3, 625
19	–	–	–	27, 3, 432	27, 3, 513	27, 3, 594	27, 3, 675
20	11, 5, 121	16, 4, 208	**16, 4, 272**	–	32, 3, 608	32, 3, 704	32, 3, 800
$\frac{20}{N}$	.165	.096	.074	-	.034	.028	.025
21	–	–	19, 4, 323	–	–	–	41, 3, 1025
22	13, 5, 143	–	23, 4, 391	23, 4, 483	–	–	–
23	–	–	25, 4, 425	25, 4, 525	25, 4, 625	–	–
24	–	16, 5, 256	–	27, 4, 609	29, 4, 725	29, 4, 841	–
25	13, 6, 169	19, 5, 304	–	–	32, 4, 800	32, 4, 928	32, 4, 1056
$\frac{25}{N}$	.148	.082	-	-	.031	.027	.024
26	–	–	–	–	37, 4, 925	37, 4, 1073	37, 4, 1221
27	–	–	23, 5, 483	–	–	43, 4, 1247	43, 4, 1419
28	16, 6, 208	–	27, 5, 702	25, 5, 650	–	–	49, 4, 1617
29	–	19, 6, 361	29, 5, 609	29, 5, 754	31, 5, 961	–	–
$\frac{29}{N}$	–	.080	.048	.038	.030	–	–
30	–	–	–	32, 5, 832	32, 5, 992	–	–

Table 3

Table 3 also contains numerical values of the rate for several obtained codes, namely: the values of fraction $\frac{m}{N}$, $m = 12, 20, 25, 29$. The comparison with lower $\underline{R}(s,1)$ and upper $\overline{R}(s,1)$ bounds from Table 2 (their values are included in Table 3 as well) yields the following conclusions:

- if $s = 2$ and $m \le 15$, then the values $\frac{m}{N}$ exceed the random coding rate $\underline{R}(2,1) = .182$;
- if $s \ge 3$ and $m \le 30$, then the values $\frac{m}{N}$ exceed the random coding rate $\underline{R}(s,1)$.

4.3 Examples of $\mathcal{D}_s^\ell$–Codes

Example 1. If $q = 5$, then for the pair ($\ell = 2$, $s = 3$), inequalities (8) are fulfilled at $k = 5$ and $r = 2$. Therefore, the construction of Proposition 4 yields a binary constant-weight $\mathcal{D}_3^2$–code X with parameters

$$t = q^{k-r} = 5^3 = 125, \quad N = n \cdot q = (q+1-r)q = 4 \cdot 5 = 20, \quad w = n = q+1-r = 4. \tag{9}$$

Parameters (9) give the following lower bound on the maximal size: $t\left(20, \mathcal{D}_3^2\right) \geq 125$.

Example 2. If $q = 7$, then for the pair ($\ell = 2$, $s = 4$), inequalities (8) are fulfilled at $k = 6$ and $r = 3$. Therefore, the construction of Proposition 4 yields a binary constant-weight $\mathcal{D}_4^2$–code X with parameters

$$t = q^{k-r} = 7^3 = 343, \quad N = n \cdot q = (q+1-r)q = 5 \cdot 7 = 35, \quad w = n = q+1-r = 5. \tag{10}$$

Parameters (10) give the following lower bound on the maximal size: $t\left(35, \mathcal{D}_4^2\right) \geq 343$.

Example 3. If $q = 8$, then for two pairs of integers ($\ell = 2$, $s = 6$) and ($\ell = 3$, $s = 10$), inequalities (8) are fulfilled at $k = 5$ and $r = 2$. Therefore, the construction of Proposition 4 yields a binary constant-weight $\mathcal{D}_6^2$–code X and a binary constant-weight $\mathcal{D}_{10}^3$–code X with parameters

$$t = q^{k-r} = 8^3 = 512, \quad N = n \cdot q = (q+1-r)q = 7 \cdot 8 = 56, \quad w = n = q+1-r = 7. \tag{11}$$

Parameters (11) give the following lower bounds on the maximal size $t\left(N, \mathcal{D}_s^\ell\right)$ of $\mathcal{D}_s^\ell$–codes:

$$t\left(56, \mathcal{D}_6^2\right) \geq 512, \qquad t\left(56, \mathcal{D}_{10}^3\right) \geq 512.$$

For comparison, if ($u = 1$, $z = 6$) and $N = 56$, then the best known lower bound on the size of optimal superimposed $(6,1)$-codes, calculated in [13], is $t(56, 6, 1) \geq 64$. In addition, this example shows that for $\ell = 3$, the parameter $s = 10$ of $\mathcal{D}_{10}^3$–code X can exceed the corresponding code weight $w = 7$.

5 Non-Adaptive Threshold Group Testing Model

5.1 Threshold Designs and Superimposed (z, u)-Codes

Let the function $F^\ell = F^\ell(n)$ takes binary values, namely:

$$F^\ell(n) \triangleq \begin{cases} 0 & \text{if} \quad n = 0, 1, \ldots, \ell - 1, \\ 1 & \text{if} \quad n = \ell. \end{cases}$$

If $\ell \geq 2$, then the given particular case is called a *threshold group testing model* [26]. For the non-adaptive threshold group testing model which is the principal for applications [29], a *refined form* of Definition 2 can be written as follows.

Definition 4. Let ℓ, $1 \leq \ell < s < t$ be integers. For code $X = \|x_i(k)\|$, $k \in [t]$, $i \in [N]$, and a subset $P \in \mathcal{P}(t, \leq s)$, define the *$i$-th outcome of non-adaptive threshold group testing*

$$y_i^\ell\,(P, X) \;\triangleq\; \begin{cases} 0 & \text{if} \sum\limits_{k \in P} x_i(k) \leq \ell - 1, \\ 1 & \text{if} \sum\limits_{k \in P} x_i(k) \geq \ell, \quad i \in [N]. \end{cases}$$

A code X is called a *threshold* $(\ell,\le s)$*–design*, (*threshold* $(\ell,=s)$*–design*) of length N and size t if for any $P,P'\in\mathcal{P}(t,\le s)$, $P\ne P'$, and such that

$$P\setminus P'\ne\emptyset,\quad P,P'\in\mathcal{P}(t,\le s)\setminus\mathcal{P}(t,\le \ell-1)\qquad \big(P\in\mathcal{P}(t,=s),\ P'\in\mathcal{P}(t,=s)\big),$$

there exists an index $i\in[N]$, where the i-th outcome of non-adaptive threshold group testing is

$$y_i^{\ell}(P,X)=1\quad\text{and}\quad y_i^{\ell}(P',X)=0.$$

Let $t_\ell(N,\le s)$, and $t_\ell(N,=s)$ denote the maximal possible sizes of threshold $(\ell,\le s)$–designs, and threshold $(\ell,=s)$–designs. For fixed $1\le\ell<s$, define the corresponding *rates*:

$$R_\ell(\le s)\triangleq\varlimsup_{N\to\infty}\frac{\log_2 t_\ell(N,\le s)}{N},\qquad R_\ell(=s)\triangleq\varlimsup_{N\to\infty}\frac{\log_2 t_\ell(N,=s)}{N}$$

Proposition 6. ([27], [28]). *If* $1\le\ell<s$*, then any superimposed* $(s-\ell+1,\ell)$*–code is a threshold* $(\ell,\le s)$*–design, i.e.*

$$t\,(N,s-\ell+1,\ell)\ \le\ t_\ell\,(N,\le s),\qquad R(s-\ell+1,\ell)\ \le\ R_\ell\,(\le s).$$

The lower bound of Theorem 5 and Propositions 6 lead to the following lower bound on the rate of threshold $(\ell,\le s)$–designs.

Proposition 7. [27], [28]. (Random coding bound). *For any* $1\le\ell<s$*, the rate*

$$R_\ell\,(\le s)\ \ge\ R(s-\ell+1,\ell)\ge-\frac{1}{s}\,\log_2\left[1-\frac{(s-\ell+1)^{s-\ell+1}\cdot\ell^{\ell}}{(s+1)^{s+1}}\right],\quad 1\le\ell<s.$$

If $\ell\ge1$ *is fixed and* $s\to\infty$*, then the asymptotic form of the given lower bound is*

$$R_\ell\,(\le s)\ \ge\frac{e^{-\ell}\cdot\ell^{\ell}\cdot\log_2 e}{s^{\ell+1}}\cdot(1+o(1)).$$

5.2 Bounds on the Rate of Threshold $(1,\le s)$–Designs

Obviously, for any $1\le\ell<s$, the following inequalities hold:

$$t_\ell(N,\le s)\ \le\ t_\ell(N,=s),\quad R_\ell\,(\le s)\ \le\ R_\ell\,(=s)\ \le\ \frac{1}{s}.\tag{12}$$

If $\ell=1$ and $s\ge2$, then the trivial bound (12) and the universal upper bound of Proposition 4 lead to inequalities :

$$R_1\,(\le s)\ \le\ \min\{1/s\,;\,\overline{R}(s-1,1)\},\qquad s=2,3,\ldots,\tag{13}$$

where $\overline{R}(z,1)$, $z=1,2,\ldots$, is the recurrent upper bound from Theorem 1. Hence, the asymptotic upper bound

$$R_1\,(\le s)\le\overline{R}(s-1,1)=\frac{2\cdot\log_2 s}{s^2}\cdot(1+o(1)),\quad s\to\infty,$$

holds.

In [9]-[10] (see, also [20]), we obtained the best known asymptotic random coding lower bounds on $R_1\,(\le s)$ and $R_1\,(=s)$ along with the best known upper bound on $R_1\,(=s)$. These bounds have the form:

$$R_1\,(\le s)\;\ge\;\underline{R}(s,1)\;=\;\frac{1}{s^2\cdot\log_2 e}\cdot(1+o(1))=\frac{1.386}{s^2}\cdot(1+o(1)),\quad s\to\infty, \tag{14}$$

$$R_1\,(=s)\ge\underline{R}_1\,(=s)=\frac{2}{s^2\cdot\log_2 e}\cdot(1+o(1))=\frac{1.386}{s^2}\cdot(1+o(1)),\quad s\to\infty, \tag{15}$$

$$R_1\,(=s)\le\overline{R}_1\,(=s)=\frac{4\cdot\log_2 s}{s^2}\cdot(1+o(1)),\quad s\to\infty. \tag{16}$$

Lower bound (14), i.e., function $\underline{R}(s,1)$ is defined in Theorem 6. For the particular case $\ell=1$, bound (14) is better than the lower bound of Proposition 7. The numerical values of lower bound (15), i.e., numbers $\underline{R}_1\,(=s)$, $s=2,3,\dots,8$, are given in Table 2.

In addition, applying the corresponding non-asymptotic results [20], one can calculate numerical values of upper bound (16), i.e., numbers $\overline{R}_1\,(=s)$, $s\ge 1$, which lead to inequalities: $R_1\,(=s)<1/s$ if $s\ge 11$. For the case $s=2$, the nontrivial inequality $R_1\,(=2)<0.4998<1/2$ was proved in [24]. For $3\le s\le 10$, the inequality $R_1\,(=s)<1/s$ can be considered as our conjecture.

5.3 Upper Bound on the Rate of Threshold $(\ell,\le s)$–Designs

For threshold $(\ell,\le s)$–designs, $\ell\ge 1$, the universal upper bound of Proposition 5 can be improved [30]. An improvement is obtained with the help of the following auxiliary concept.

Definition 5. [30] [2]. Let ℓ, $1\le\ell<s<t/2$ be integers. A binary $(N\times t)$-matrix X is called a *superimposed* $\mathcal{M}_s^\ell$*-code* (briefly, $\mathcal{M}_s^\ell$*-code*) if for any two non-intersecting subsets $Z,\,U\in\mathcal{P}(t,\le s)$, $Z\cap U=\emptyset$, such that $\ell\le|U|\le s$ and for any element $j\in U$, the matrix X contains a row $\mathbf{x}_i=(\,x_i(1),x_i(2)\dots,x_i(t)\,)$, $i\in[N]$, for which

$$x_i(j)=1,\quad \sum_{k\in U}x_i(k)\;=\;\ell\quad\text{and}\quad x_i(k)=0\quad\text{for all}\quad k\in Z.$$

Let $t\left(N,\mathcal{M}_s^\ell\right)$ denote the maximal size of $\mathcal{M}_s^\ell$–codes. For fixed $1\le\ell<s$, introduce

$$R\left(\mathcal{M}_s^\ell\right)\triangleq\overline{\lim_{N\to\infty}}\,\frac{\log_2 t\left(N,\mathcal{M}_s^\ell\right)}{N},\qquad 1\le\ell<s.$$

called a *rate* of $\mathcal{M}_s^\ell$–codes. The evident connection between $\mathcal{M}_s^\ell$–codes and superimposed $(2s-\ell,1)$-codes is given by

Proposition 8. [30]. **1.** *Let $2\le s<t/2$. If $\ell=1$, then any $\mathcal{M}_s^1$–code X of size t is a superimposed $(2s-1,1)$-code and, vice versa, any superimposed $(2s-1,1)$-code X of size t is a $\mathcal{M}_s^1$–code, i.e., the rate $R\left(\mathcal{M}_s^1\right)=R(2s-1,1)$.* **2.** *If $2\le\ell<s<t/2$, then any $\mathcal{M}_s^\ell$-code X of size t is a superimposed $(2s-\ell,1)$-code, i.e., the rate $R\left(\mathcal{M}_s^\ell\right)\,\le\,R(2s-\ell,1)$.*

In Sect. 5.4, the following connections between threshold $(\ell,\le s)$–designs and $\mathcal{M}_s^\ell$-codes will be proved.

[2] From an original definition of paper [30] we canceled the extra condition $|Z|\le|U|$ because for the original definition of paper [30], the author's proofs of Propositions 9 which is given below will be incorrect.

Proposition 9. [30]. *If $1 \le \ell < s < t/2$, then any threshold $(\ell, \le 2s-1)$–design X of size t is a $\mathcal{M}_s^\ell$-code, i.e., the rate $R_\ell\,(\le 2s-1) \le R\left(\mathcal{M}_s^\ell\right)$ if $1 \le \ell < s$.*

Proposition 10. [30]. *If $1 \le \ell < s < t/2$, then any $\mathcal{M}_s^\ell$-code X of size t is a threshold $(\ell, \le s)$–design, i.e. the rate $R\left(\mathcal{M}_s^\ell\right) \le R_\ell\,(\le s)$.*

Hence, the trivial upper bound (12), Propositions 8-9 and Theorems 1-2 lead to

Theorem 8. *Let $s = 2, 3, \dots$ be fixed. If $1 \le \ell < s$, then any threshold $(\ell, \le 2s-1)$–design is a superimposed $(2s-\ell, 1)$-code, i.e., the rate*

$$R_\ell\,(\le 2s-1) \;\le\; \overline{R}_\ell\,(\le 2s-1) \triangleq \min\left\{\frac{1}{2s-1}\,;\,\overline{R}(2s-\ell,1)\right\}, \quad 1 \le \ell < s, \tag{17}$$

and the rate

$$R_\ell\,(\le 2s) \;\le\; \overline{R}_\ell\,(\le 2s) \triangleq \min\left\{\frac{1}{2s}\,;\,\overline{R}(2s-\ell,1)\right\}, \quad 1 \le \ell < s, \tag{18}$$

where $\overline{R}(z,1)$, $z \ge 1$, is the recurrent upper bound from Theorem 1. In addition, for any fixed $\ell \ge 1$ and $s \to \infty$, the rate

$$R_\ell\,(\le s) \;\le\; \frac{2\log_2 s}{s^2}\,(1+o(1)). \quad \ell = 1, 2, \dots, \quad s \to \infty.$$

If $\ell = 1$ and $s = 3, 5, 7, \dots$ is odd, then upper bound (17) improves upper bound (13). For $\ell = 1, 2, 3$ and $s = \ell+1, \ell+2, \dots, 8$, numerical values of upper bounds (17)-(18) are presented in Table 2.

The reciprocal values of $\overline{R}(z,1)$, i.e., numbers $1/\overline{R}(z,1) > z$, $z = 2, 3, \dots, 17$, which were presented in Table 1, yield

Proposition 11. *If $1 \le \ell \le 5$ and $s \ge 2\ell+1$, then $\overline{R}_\ell\,(\le s) < 1/s$. If $\ell \ge 6$ and $s \ge 2\ell$, then $\overline{R}_\ell\,(\le s) < 1/s$.*

Remark 4. Using the inequality of Proposition 10 and a random coding method [30], one can asymptotically improve (?) the lower bound of Proposition 7.

5.4 Proofs of Propositions 9-10

Proof of Proposition 9. Let $X = \|x_i(k)\|$, $k \in [t]$, $i \in [N]$, be an arbitrary threshold $(\ell, \le 2s-1)$–design. Fix arbitrary subsets:

$$U \in \mathcal{P}(t, \le s), \quad \ell \le |U| \le s, \quad j \in U, \quad Z \in \mathcal{P}(t, \le s), \quad U \cap Z = \emptyset.$$

Introduce subsets $P, P' \in \mathcal{P}(t, \le 2s-1) \setminus \mathcal{P}(t, \le \ell-1)$ as follows:

$$P \triangleq U, \quad P' \triangleq (U \setminus j) \cup Z, \quad P \setminus P' \ne \emptyset, \quad \ell \le |P| \le s, \quad \ell \le |P'| \le 2s-1.$$

Definition 4 of threshold $(\ell, \le 2s-1)$–design means that there exists an index $i \in [N]$ such that

$$\big(\, y_i(P,X) = 1,\; y_i(P',X) = 0 \,\big) \;\Rightarrow\; \left(\sum_{k\in P} x_i(k) \ge \ell,\; \sum_{k\in P'} x_i(k) \le \ell-1\right) \;\Rightarrow$$

$$\Rightarrow \left(\sum_{k\in U} x_i(k) \ge \ell,\; \sum_{k\in U\setminus j} x_i(k) + \sum_{k\in Z} x_i(k) \le \ell-1\right) \;\Rightarrow$$

$$\Rightarrow \left(\sum_{k \in U \setminus j} x_i(k) = \ell - 1,\; x_i(j) = 1,\; \sum_{k \in Z} x_i(k) = 0 \right),$$

i.e., code X is a $\mathcal{M}_s^\ell$-code. Proposition 9 is proved.

Proof of Proposition 10. Let $X = \|x_i(k)\|$, $k \in [t]$, $i \in [N]$, be an arbitrary $\mathcal{M}_s^\ell$-code. Consider arbitrary subsets: $P, P' \in \mathcal{P}(t, \le s)$, $P \ne P'$, and such that

$$P \setminus P' \ne \emptyset, \quad P, P' \in \mathcal{P}(t, \le s) \setminus \mathcal{P}(t, \le \ell - 1), \quad \ell \le |P| \le s,\; \ell \le |P'| \le s.$$

Fix an arbitrary $j \in P \setminus P'$, $j \notin P'$ and define non-intersecting subsets $U \triangleq P$ and $Z \triangleq P' \setminus P$. We have

$$\ell \le |U| \le s, \quad j \in U, \quad U \cap Z = \emptyset, \quad Z \subset P', \quad P' \setminus Z \subset U, \quad |Z| \le |P'| \le s.$$

Definition 5 of $\mathcal{M}_s^\ell$-code implies that there exists an index $i \in [N]$ such that

$$\left(\sum_{k \in U} x_i(k) = \ell,\; \sum_{k \in Z} x_i(k) = 0,\; x_i(j) = 1,\; \sum_{k \in P' \setminus Z} x_i(k) \le \ell - 1 \right) \Rightarrow$$

$$\Rightarrow \left(\sum_{k \in P} x_i(k) = \ell,\; \sum_{k \in P'} x_i(k) \le \ell - 1 \right) \Rightarrow \left(y_i(P, X) = 1,\; y_i(P', X) = 0 \right),$$

i.e., code X is a threshold $(\ell, \le s)$–design. Proposition 10 is proved.

5.5 Random Coding Bound on the Rate $R\left(\mathcal{M}_s^\ell\right)$

If $\beta \triangleq \Pr\{x_i(k) = 0\}$ and $1 - \beta \triangleq \Pr\{x_i(k) = 0\}$, then for any $j \in [t]$, the probability

$$\Pr\left\{ \mathbf{x}(j) \text{ is } \mathcal{M}_s^\ell - \text{bad} \right\} \le \sum_{u=\ell}^{s} \sum_{z=0}^{s} \binom{t-1}{u+z-1} \binom{u+z-1}{u-1} \times$$

$$\times \left[1 - \binom{u-1}{\ell-1} \beta^\ell (1-\beta)^{u+z-\ell} \right]^N.$$

The given inequality leads to the following lower bound on the rate of $\mathcal{M}_s^\ell$-codes: for any β, $0 < \beta < 1$, the rate $R\left(\mathcal{M}_s^\ell\right)$ satisfies inequality

$$R\left(\mathcal{M}_s^\ell\right) \ge \min_{\ell \le u \le s;\, 0 \le z \le s} \left\{ \frac{-\log_2 \left[1 - \binom{u-1}{\ell-1} \beta^\ell (1-\beta)^{u+z-\ell} \right]}{u+z-1} \right\}.$$

The minimax is attained when $u = \ell$ and $z = s$. Therefore, the lower bound takes the standard form of Proposition 7:

$$R\left(\mathcal{M}_s^\ell\right) \ge \frac{-\log_2 \left[1 - \max\limits_{0<\beta<1} \left\{ \beta^\ell (1-\beta)^s \right\} \right]}{s + \ell - 1} = \frac{-\log_2 \left[1 - \frac{s^s \, \ell^\ell}{(s+l)^{s+\ell}} \right]}{s + \ell - 1}$$

where the maximum is attained at $\beta = \ell/(s+\ell)$. Thus, if $\ell \ge 1$ is fixed and $s \to \infty$, then the asymptotic form of the random coding bound is

$$R\left(\mathcal{M}_s^\ell\right) \ge \frac{\ell^\ell \exp\{-\ell\} \log_2 e}{s^{\ell+1}} (1 + o(1)).$$

6 Comments on Definitions 4-5

We deem that instead of Definitions 4-5, Mahdi Cheraghchi [30] actually introduced the following definitions.

Definition 4'. A code X is called a *threshold* $\widetilde{(\ell, \leq s)}$*–design*, of length N and size t if for any $P, P' \in \mathcal{P}(t, \leq s)$, $P \neq P'$, and such that

$$P \setminus P' \neq \emptyset, \quad P, P' \in \mathcal{P}(t, \leq s) \setminus \mathcal{P}(t, \leq \ell - 1), \qquad |P'| \leq |P|,$$

there exists an index $i \in [N]$, where the i-th outcome of non-adaptive threshold group testing is

$$y_i^{\ell}(P, X) = 1 \quad \text{and} \quad y_i^{\ell}(P', X) = 0.$$

Let $\widetilde{t}_{\ell}(N, \leq s)$, be the maximal size of threshold $\widetilde{(\ell, \leq s)}$–designs. For fixed $1 \leq \ell < s$, define the corresponding *rate*

$$\widetilde{R}_{\ell}(\leq s) \triangleq \overline{\lim_{N \to \infty}} \frac{\log_2 \widetilde{t}_{\ell}(N, \leq s)}{N}.$$

Definition 5'. [30] Let ℓ, $1 \leq \ell < s < t/2$ be integers. A binary $(N \times t)$-matrix X is called a *superimposed* $\widetilde{\mathcal{M}}_s^{\ell}$*-code* (briefly, $\widetilde{\mathcal{M}}_s^{\ell}$*-code*) if for any two non-intersecting subsets $Z, U \in \mathcal{P}(t, \leq s)$, $Z \cap U = \emptyset$, such that $\ell \leq |U| \leq s$, $|Z| \leq |U|$ and for any element $j \in U$, the matrix X contains a row $\mathbf{x}_i = (x_i(1), x_i(2) \ldots, x_i(t))$, $i \in [N]$, for which

$$x_i(j) = 1, \quad \sum_{k \in U} x_i(k) = \ell \quad \text{and} \quad x_i(k) = 0 \quad \text{for all} \quad k \in Z.$$

Let $t\left(N, \widetilde{\mathcal{M}}_s^{\ell}\right)$ denote the maximal size of $\widetilde{\mathcal{M}}_s^{\ell}$–codes. For fixed $1 \leq \ell < s$, introduce

$$R\left(\widetilde{\mathcal{M}}_s^{\ell}\right) \triangleq \overline{\lim_{N \to \infty}} \frac{\log_2 t\left(N, \widetilde{\mathcal{M}}_s^{\ell}\right)}{N}, \qquad 1 \leq \ell < s.$$

called a *rate* of $\widetilde{\mathcal{M}}_s^{\ell}$–codes.

For Definitions 4' and 5', Propositions 8 and 10 will be true and, hence, the inequality

$$\widetilde{R}_{\ell}(\leq s) \geq R\left(\widetilde{\mathcal{M}}_s^{\ell}\right)$$

takes place, but we cannot state that Proposition 9 and the corresponding upper bound of Theorem 8 are correct.

One can easily see that the random coding arguments of Sect.5.5 leads to

Proposition 12. *For any* β, $0 < \beta < 1$, *the rate* $R\left(\widetilde{\mathcal{M}}_s^{\ell}\right)$ *satisfies inequality*

$$R\left(\widetilde{\mathcal{M}}_s^{\ell}\right) \geq \min_{\ell \leq u \leq s;\ 0 \leq z \leq u} \left\{ \frac{-\log_2 \left[1 - \binom{u-1}{\ell-1} \beta^{\ell} (1-\beta)^{u+z-\ell}\right]}{u + z - 1} \right\}.$$

Therefore,

$$R\left(\widetilde{\mathcal{M}}_s^{\ell}\right) \geq \underline{R}\left(\widetilde{\mathcal{M}}_s^{\ell}\right) \triangleq \max_{0 < \beta < 1} \min_{\ell \leq u \leq s} \left\{ \frac{-\log_2 \left[1 - \binom{u-1}{\ell-1} \beta^{\ell} (1-\beta)^{2u-\ell}\right]}{2u - 1} \right\}. \tag{17}$$

Problem. Find the asymptotic form of lower bound (17) if $\ell \geq 1$ is fixed and $s \to \infty$.

7 Concluding Remarks

In this Sect., we would like to distinguish the principal achievements for the theory of non-adaptive group testing models and superimposed codes obtained in the last decade.

1. In 2003, Vladimir Lebedev [21] proved Theorem 3 which established a recurrent inequality for the rate $R(z,u)$ of superimposed (z,u)-codes. This inequality and the best known numerical values [4, 17] of upper bound on the rate $R(z,1)$ gave the best known numerical values of upper bound on the rate $R(z,u)$, $z \geq u \geq 2$.

2. In 2004, Vladimir Lebedev and Hyun Kim [22] presented the best known and optimal constructions of superimposed (z,u)-codes, $z \geq u \geq 2$.

3. In 2004, Annalisa De Bonis and Ugo Vaccaro [25] proved Theorem 7 which established an upper bound on the rate of superimposed $\mathcal{D}_s^\ell$-codes via the rate $R(z,1)$ of superimposed $(z,1)$-codes. The result leads to the universal upper bound (Proposition 4) on the rate of group testing designs motivated by compressed genotyping models in molecular biology.

4. In 2010, Mahdi Cheraghchi [30] introduced the concept of superimposed $\mathcal{M}_s^\ell$-codes and proved Propositions 8-10 which actually established an upper bound (Theorem 8) on the rate of threshold $(\ell, \leq s)$–designs for non-adaptive threshold group testing model via the recurrent upper bound $\overline{R}(z,1)$ on the rate $R(z,1)$ of superimposed $(z,1)$-codes.

References

[1] W.H. Kautz, R.C. Singleton, "Nonrandom Binary Superimposed Codes," *IEEE Trans. Inform. Theory*, vol. 10, no. 4, pp. 363-377, 1964.

[2] F.J. MacWilliams, N.J.A. Sloane, "The Theory of Error-Correcting Codes", North Holland, 1977.

[3] I. Csizar, J. Korner, "Information Theory. Coding Theorems for Discrete Memoryless Systems", Academiai Kiado, Budapest, 1981.

[4] A.G. D'yachkov, V.V. Rykov, "Bounds on the Length of Disjunct Codes," *Problemy Peredachi Informatsii*, vol. 18, no. 3, pp. 7-13, 1982 (in Russian).

[5] A.G. D'yachkov, V.V. Rykov, "A Survey of Superimposed Code Theory," *Problems of Control and Inform. Theory*, vol. 12, no. 4, pp. 229-242, 1983.

[6] A.G. D'yachkov, V.V. Rykov, "Generalized Superimposed Codes and their Application to Random Multiple Access", *Proc. of the 6-th International Symposium on Information Theory*, part 1, Taschkent, 1984.

[7] A.G. D'yachkov, V.V. Rykov, "On a Model of Associative Memory," *Problemy Peredachi Inform.*, 1988, v.24, no.3, 107-110 (in Russian).

[8] Nguyen Quang A., T. Zeisel, "Bounds on constant weight binary syperimposed codes," *Probl. of Control and Inform. Theory*, 1988, v.17, no.4, pp.223-230.

[9] A.G. D'yachkov, V.V. Rykov, A.M. Rashad, "Superimposed Distance Codes," *Problems of Control and Inform. Theory*, vol. 18, no. 4, pp. 237-250, 1989.

[10] A.G. D'yachkov, A.M. Rashad, "Universal Decoding for Random Design of Screening Experiments", *Microelectronics and Reliability*, vol. 29, no. 6, pp. 965-971, 1989.

[11] A.M. Rashad, "Random coding bounds on the rate for list-decoding superimposed codes," *Problems of Control and Inform. Theory*, v.19, No 2, 1990, pp. 141-149.

[12] A.G. D'yachkov, V.V. Rykov, "The Capacity of the Boolean Associative Memory". *Proc. of the 5-th International Conference on Artificial Neural Networks*, Churchill Colledge, Cambridge, UK, pp. 158-160, 1997.

[13] A.G. D'yachkov, A.J. Macula, V.V. Rykov, "New Constructions of Superimposed Codes," *IEEE Trans. Inform. Theory*, vol. 46, no. 1, pp. 284-290, 2000.

[14] A.G. D'yachkov, A.J. Macula, V.V. Rykov, "New Applications and Results of Superimposed Code Theory Arising from the Potentialities of Molecular Biology". *In the book "Numbers, Information and Complexity"*, pp. 265-282, Kluwer Academic Publishers, 2000.

[15] A.G. D'yachkov, A.J. Macula, D.C. Torney, P.A. Vilenkin, "Two Models of Non-Adaptive Group Testing for Designing Screening Experiments". *Advances in model oriented design and analysis: proceedings of the 6th International Workshop on Model Oriented Design and Analysis*, Puchberg/Schneeberg, Austria, June 25-29, 2001, pp. 63-75, Physica-Verlag Heidelberg, 2001.

[16] A.G. D'yachkov, V.V. Rykov, "Optimal Superimposed Codes and Designs for Renyi's Search Model". *Journal of Statistical Planning and Inference*, vol. 100, pp. 281-302, 2002.

[17] A. D'yachkov, A. Macula, D. Torney, P. Vilenkin, "Families of Finite Sets in which No Intersection of ℓ Sets Is Covered by the Union of s Others", *Journal of Combinatorial Theory, Series A*, vol. 99, pp. 195-218, 2002.

[18] A.G. D'yachkov, P.A. Vilenkin P.A., S.M. Yekhanin, "Upper bounds on the rate of superimposed (s,ℓ)-codes, based on Engel's inequality". *Proceedings of the 8th International Workshop "Algebraic and Combinatorial Coding Theory"*, Tsarskoe Selo, Russia, September 8-14, pp. 95-99, 2002.

[19] A.G. D'yachkov, "Superimposed Designs and Codes for Non-Adaptive Search of Mutually Obscuring Defectives". *2003 IEEE International Symposium on Information Theory*, Yokohama, Japan, June 29 - July 4, 2003, p. 134.

[20] A.G. D'yachkov, "Lectures on Designing Screening Experiments", Lecture Note Series 10, Feb. 2004, Combinatorial and Computational Mathematics Center, Pohang University of Science and Technology (POSTECH), Korea Republic, (monograph, pp. 112).

[21] V.S. Lebedev, "An Asymptotic Upper Bound on the Rate of (w,r)-Cover-Free Codes", *Problemy Peredachi Informatsii*, 2003, v.39, no.4, pp.3-9, 2003, (in Russian).

[22] H. Kim, V. Lebedev, "On Optimal Superimposed Codes", *Journal of Combinatorial Designs*, v. 12, no.2, pp. 79-91, 2004.

[23] D.-Z. Du, F.K. Hwang, "Combinatorial Group Testing and Its Applications," World Scientific, Singapore-New Jersey-London-Hong Kong, 1993.

[24] D. Copperersmith, J. Shearer, "New Bounds for Union-free Families of Sets", *The Electronic Journal of Combinatorics*, vol. 5, no.1, R39, 1998.

[25] A. De Bonis, U. Vaccaro, "Optimal Algorithms for Group Testing Problems, and New Bounds on Generalized Superimposed Codes", *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4673-4680, 2006.

[26] P. Damaschke, "Threshold Group Testing", *In the book "General Theory of Information Transfer and Combinatorics"*, pp. 707-718, Kluwer Academic Publishers, 2006.

[27] Chen and Fu, "Nonadaptive algorithms for threshold group testing", *Discrete Applied Mathematics*, vol. 157, pp. 1581–1585, 2009.

[28] V.S. Lebedev, "Separating Codes and a New Combinatorial Search Model", *Problemy Peredachi Informatsii*, vol. 46, no. 1, pp. 1-6, 2010 (in Russian).

[29] Y. Erlich, A. Gordon, M. Brand, G. Hannon, P. Mitra, "Compressed Genotyping", *IEEE Trans. Inform. Theory*, vol. 56, no. 2, pp. 706-723, 2010.

[30] M. Cheraghchi, "Improved Constructions for Non-adaptive Threshold Group Testing", *In Proceedings of the 37th International Colloquium on Automata, Languages and Programming (ICALP)*, 2010. arXiv:1002.2244.

[31] M.B. Malyutov, "The separating Property of Random Matrices". *Mat. Zametki*, vol. 23, pp. 155-160, 1978.

[32] M.B. Malyutov, "Maximal Rates of Screening Designs", *Th. Probab. and Appl.*, vol. 24, pp. 655-657, 1979.

[33] M.B.Malyutov, P.S. Mateev, "Screening Designs for Non-Symmetric Response Function", *Mat. Zametki*, vol. 27, pp. 109-127, 1980.

[34] M.B. Malyutov, A.G. Dyachkov, "On Weakly Separating Designs", *In the book: Methods of transmission and processing information*, Nauka, Moscow, pp. 87-104, 1980.

[35] M. Malyutov, H. Sadaka, "Jaynes Principle in Testing Significan Variables of Linear Model", *Random Operators and Stochastic Equations*, vol. 6, pp. 311-330, 1998.

[36] M.B. Malyutov, "Recovery of Sparse Active Inputs in general systems: a review", *In Proceedings, IEEE region 8, Conference SIBIRCON-2010*, pp. 15-22, 2010, Proceedings are available via IEEEXplore search.