Threshold and majority group testing

R. Ahlswede[a], C. Deppe[a], V. Lebedev[b]

[a] *Department of Mathematics, University of Bielefeld*
[b] *IPPI (Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute))*

Abstract

We consider two generalizations of group testing: threshold group testing (introduced by Damaschke [7]) and majority group testing (a further generalization, including threshold group testing and a model introduced by Lebedev [14]).
We show that each separating code gives a nonadaptive strategy for threshold group testing for some parameters. This is a generalization of a result in [2] on "guessing secrets", introduced in [5].
We introduce threshold codes and show that each threshold code gives a nonadaptive strategy for threshold group testing. We show that there exist threshold codes such that we can improve the lower bound of [3] for the rate of threshold group testing.
We consider majority group testing if the number of defective elements is unknown (otherwise it reduces to threshold group testing). We show that cover-free codes and separating codes give strategies for majority group testing. We give a lower bound for the rate of majority group testing.

Keywords: group testing, pooling, threshold group testing, separating codes, cover-free codes

1

Email addresses: ahlswede@math.uni-bielfeld.de (R. Ahlswede), cdeppe@math.uni-bielfeld.de (C. Deppe), lebedev37@mail.ru (V. Lebedev)

Supported by the DFG in the project "Advances in Search and Sorting"
Supported in part by the Russian Foundation for Basic Research, project no 09-01-00536.

1. Introduction

Group testing is of interest for many applications like in molecular biology. For an overview of results and applications we refer to the books [8] and [9].

The classical group testing problem is to find the unknown subset $\mathcal{D}$ of all defective elements in the set $[N] = \{1, 2, \ldots, N\}$.

For a subset $\mathcal{S} \subset [N]$ a test $t_{\mathcal{S}}$ is the function $t_{\mathcal{S}} : 2^{[N]} \to \{0, 1\}$ defined by

$$t_S(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| = 0 \\ 1 & , \text{ otherwise.} \end{cases}$$

We define search strategies as in [1]. A strategy is called successful, if we can uniquely determine $\mathcal{D}$. We remind the reader of the concepts of adaptive and nonadaptive strategies.

Strategies are called adaptive if the results of the first $k-1$ tests determine the kth test. Strategies in which we choose all tests independently are called nonadaptive.

In the present paper we study two generalizations of group testing which are quite natural.

In **threshold group testing** the integers $0 \leq l < u$ are given and a test $t_{\mathcal{S}}$ is the function $t_{\mathcal{S}} : 2^{[N]} \to \{0, 1, \{0, 1\}\}$, defined by

$$t_S(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \leq l \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \geq u \\ \{0, 1\} & , \text{ otherwise} \\ & \text{(meaning that the result can be arbitrary 0 or 1).} \end{cases}$$

In threshold group testing it is not possible to find the set $\mathcal{D}$ of all defective elements exactly if the gap $g = u - l - 1 > 0$ (see [7]). It is only possible to find a set $\mathbb{F}$ of subsets of $[N]$ which includes $\mathcal{D}$ and $\forall \mathcal{P}, \mathcal{P}' \in \mathbb{F} : |\mathcal{P}' \backslash \mathcal{P}| \leq g$ and $|\mathcal{P} \backslash \mathcal{P}'| \leq g$.

In **majority group testing** there are two functions $f_1, f_2 : \{0, 1, \ldots, N\} \to \mathbb{R}^+$ which put weights on the number $D = |\mathcal{D}| \in \{0, 1, \ldots, N\}$ of defective elements and $f_1(D) < f_2(D)\ \forall D \in [0, 1, \ldots, N]$.

They describe the structure of tests $t_{\mathcal{S}} : 2^{[N]} \to \{0, 1, \{0, 1\}\}$ as follows

$$t_S(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \leq f_1(D) \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \geq f_2(D) \\ \{0, 1\} & , \text{ otherwise} \\ & \text{(meaning that the result can be arbitrary 0 or 1).} \end{cases}$$

Clearly majority group testing is a generalization of threshold group testing. We get threshold group testing as a special case by setting $f_1(D) = l$ and $f_2(D) = u$. Furthermore the models are equivalent if the number $\mathcal{D}$ of defectives is known. In majority group testing, in particular also for threshold group testing, it is not possible to find the set $\mathcal{D}$ of all defective elements. We can find a set of subsets $\mathbb{F} \subset 2^{[N]}$, which contains $\mathcal{D}$. This set depends on f_1 and f_2, on $\mathcal{D}$, and on the strategy used. In this case we call a strategy successful, if we can find an $\mathbb{F}$ with the smallest possible size in the worst case.

A special case of majority group testing was introduced by Lebedev [14] as follows

$$t_{\mathcal{S}}(\mathcal{D}) = \begin{cases} 0 & \text{, if } |\mathcal{S} \cap \mathcal{D}| < \frac{D}{2} \\ 1 & \text{, if } |\mathcal{S} \cap \mathcal{D}| > \frac{D}{2} \\ \{0, 1\} & \text{, if } |\mathcal{S} \cap \mathcal{D}| = \frac{D}{2}. \end{cases}$$

It was shown in [14] that a (w, w) separating code gives a successful non-adaptive strategy if it is assumed that D is odd and that $D < 2w$. (See Section 5 for other special cases studied in [14]).

In [2] it was shown that for **guessing secrets** (that means $l = 0$ and $u = D$ for threshold group testing) a (D, D) separating code gives a successful nonadaptive strategy. We prove in **Section 2** that for threshold group testing a $(u, D - l)$ separating code gives a successful nonadaptive strategy if $D = u + l$. This improves the result of [3] for this special case, because the authors use a $(u, D - l)$ cover-free code for the strategy, which has a smaller rate than a separating code.

In **Section 3** we **introduce threshold codes** and show that these codes give nonadaptive strategies for threshold group testing, if the number of defectives are known. We improve the lower bound for the rate of threshold group testing of [3]. In **Section 4** we give a list of values for the bound of Section 3 by computer calculation.

Finally, in **Section 5** we consider **majority group** testing for $f_1(D) = \lceil \frac{D}{k} \rceil - 1$ and $f_2(D) = \lfloor \frac{D}{k} \rfloor + 1$ where $2 \leq k \in \mathbb{N}$. We first give conditions for a successful nonadaptive strategy. Then we give a lower bound for its rate. Again we find relations to separating codes and cover-free codes.

We call a set $\mathcal{S} \subset [N]$ a test set if it is used in defining the test $t_{\mathcal{S}}$. We assume that $D \geq u$, because otherwise it is not possible to find any defective element.

2. Nonadaptive threshold group testing using separating codes

It is obvious that it is not possible to find defective elements if $D \leq u-1$. In [7] it is shown that if $D \geq u$ we can find a set $\mathcal{P}$ such that

$$|\mathcal{D}\backslash\mathcal{P}| \leq g \text{ and } |\mathcal{P}\backslash\mathcal{D}| \leq g \tag{1}$$

or we can even find a set $\mathbb{F}$ of subsets of $[N]$ with

$$\mathcal{D} \in \mathbb{F} \quad \text{and} \quad \forall \mathcal{P}, \mathcal{P}' \in \mathbb{F}: \ |\mathcal{P}'\backslash\mathcal{P}| \leq g \text{ and } |\mathcal{P}\backslash\mathcal{P}'| \leq g. \tag{2}$$

We say that a strategy for threshold group testing is successful if we found a set $\mathbb{F}$ satisfying condition (2). The size of $\mathbb{F}$ cannot be reduced. It is shown in [7] that all answers given for a strategy can be the same for all sets in the set $\mathbb{F}$ as for the set $\mathcal{D}$ of defective elements. Thus we cannot distinguish these sets. First we consider the case where the number D of defectives is known.

Definition 1. *$n_{Th}(N, l, u, D)$ is the minimal number of tests of a nonadaptive strategy for threshold group testing with lower bound l and upper bound u (see the definition in the introduction) to find a set $\mathbb{F}$ which fulfills (2), if there are D defective elements. $R_{Th} = R_{Th}(l, u, D) = \sup_N \frac{\log N}{n_{Th}(N,l,u,D)}$ denotes the maximal achievable rate of a nonadaptive strategy for threshold group testing for given D, u, l.*

Definition 2. *An $n \times N$ matrix $(m_{ij})_{1\leq i\leq n, 1\leq j\leq N}$ is called a (w, r) separating code of size $n \times N$, if for any pair of subsets $I, J \subset [N]$ such that $|I| = w$, $|J| = r$, and $I \cap J = \emptyset$, there exists a row index $k \in [n]$ such that $m_{ki} = 1$ $\forall i \in I$ and $m_{kj} = 0$ $\forall j \in J$ or vice versa.*

By $n_S(N, w, r)$ we denote the minimal number of rows of a (w, r) separating code with N columns and by R_S the corresponding maximal achievable rate.

Theorem 1. *Let $D = u + l$, then $n_{Th}(N, l, u, D) \leq n_S(N, u, D - l)$.*

Proof. Let $(m_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,N}}$ be a $(u, D-l)$ separating code of size $n \times N$.

We use the n rows as test sets (written in binary representation) for our strategy and show that we can find a set $\mathbb{F}$ of sets such that (2) is fulfilled.

Let $\mathbb{F}_0 = \{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_{\binom{N}{D}}\}$ be the set of all D element subsets of $\{1, 2, \ldots, N\}$.

First consider $\mathcal{A}_1$ and search for the set $\mathcal{A}_i$ with the smallest index $i > 1$, such that $|\mathcal{A}_1 \backslash \mathcal{A}_i| > g = u - l - 1$.
Now we compare these two sets.

Case 1: $u - l \leq |\mathcal{A}_1 \backslash \mathcal{A}_i| < u$.
Set $\mathcal{I} = (\mathcal{A}_1 \backslash \mathcal{A}_i) \cup \mathcal{B}$, where $\mathcal{B} \subset \mathcal{A}_i \cap A_1$, such that $|\mathcal{I}| = u$
and set $\mathcal{J} \subset \mathcal{A}_i \backslash \mathcal{B}$, such that $|\mathcal{J}| = D - l$. This is possible because $|\mathcal{B}| \leq l$.

Case 2: $|\mathcal{A}_1 \backslash \mathcal{A}_i| \geq u$.
Set $\mathcal{I} \subset \mathcal{A}_1 \backslash \mathcal{A}_i$ such that $|\mathcal{I}| = u$ and set $\mathcal{J} \subset \mathcal{A}_i$, such that $|\mathcal{J}| = D - l$.

There exists a row (a test set $\mathcal{S}$), because of the properties of separating codes, such that $\mathcal{I} \subset \mathcal{S}$ and $\mathcal{J} \not\subset \mathcal{S}$ or vice versa, where $\mathcal{S}$ is the subset which corresponds to the row.

If the result is 1 and ($\mathcal{I} \subset \mathcal{S}$ and $\mathcal{J} \not\subset \mathcal{S}$) then we continue our strategy with the set $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{A}_i\}$. Otherwise we continue with the set $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{A}_1\}$. If the result is 1 and ($\mathcal{I} \not\subset \mathcal{S}$ and $\mathcal{J} \subset \mathcal{S}$) then we continue our strategy with the set $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{A}_1\}$. Otherwise we continue with the set $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{A}_i\}$.

If $\mathcal{A}_1 \in \mathbb{F}_1$ we search again for the set with the smallest index $i > 1$, such that $|\mathcal{A}_1 \backslash \mathcal{A}_i| > g = u - l - 1$, if such a set exists.
Otherwise we continue with $\mathcal{A}_2$. We stop at step s if there are no sets $\mathcal{A}$ and $\mathcal{B}$ in the set such that $|\mathcal{A} \backslash \mathcal{B}| > g$ and $|\mathcal{B} \backslash \mathcal{A}| > g$. The remaining set $\mathbb{F}_s$ has the claimed properties:

We did not exclude the set $\mathcal{D}$ which contains all defective elements from our set for the following reason. If we compare $\mathcal{D}$ and $\mathcal{A}$ and the result of our test is 1, we remove $\mathcal{A}$, because more than u elements of $\mathcal{D}$ are in the test set. If the result is 0 we also remove $\mathcal{A}$, because then less than l elements are in the test set $\mathcal{S}$. Therefore our remaining set $\mathbb{F}_s$ contains the set with all defective elements and for $\mathbb{F}_s$ (2) holds. □

The following is an upper bound for n_S (the authors use the terminology (N, u)-universal sets, which are (u, u)-separating codes).

Theorem [16] $n_S(N, u, u) \leq 2u2^{2u} \log N$.

In conjunction with our Theorem 1 this implies (this was shown in [2] for $l = 0$ only)

Corollary 1. *If $D = u + l$, then $n_{Th}(N, l, u, D) \leq 2u2^u \log N$ and $R_{Th} \geq \frac{1}{2u2^{2u}}$.*

By random choice of a separating code (see [6]) we get a lower bound on R_S and thus we get from Theorem 1

Corollary 2.
$$R_{Th} \geq R_S \geq \frac{-\log(1 - 2^{-(2u)})}{2u - 1}. \tag{3}$$

3. A general lower bound for nonadaptive threshold group testing

In the previous section we got a lower bound for threshold group testing if $D = u + l$. The best known bound for general D was given in [3] using cover-free codes.

Definition 3. *An $n \times N$ matrix $(m_{ij})_{1 \leq i \leq n, 1 \leq j \leq N}$ is called a (w, r) cover-free code of size $n \times N$, if for any pair of subsets $\mathcal{I}, \mathcal{J} \subset [N]$ such that $|\mathcal{I}| = w$, $|\mathcal{J}| = r$, and $\mathcal{I} \cap \mathcal{J} = \emptyset$, there exists a row index $k \in [n]$ such that $m_{ki} = 1$ $\forall i \in \mathcal{I}$ and $m_{kj} = 0$ $\forall j \in \mathcal{J}$.*

$n_c(N, w, r)$ denotes the minimal number of rows among all (w, r) cover-free codes with N columns.

Threshold group testing without gap is a special case of the complex group testing model, which was introduced in [20]. In complex group testing we have a set of N elements and a family $\mathcal{P}$ of defective subsets of this set. The test gives a positive result, if it includes all elements of a defecive subset. The goal is to find all defective subsets.

Let $\mathcal{D}$ the set of defective elements in threshold group testing with the upper bound u and the lower bound l. If we choose $\mathcal{P} = \binom{\mathcal{D}}{u}$ then threshold group testing and complex group testing are the same. Therefore the bounds for complex group testing in [11] can be used for threshold group testing without gap. For $u = 3$ it is the same bound like in [3].

In [3] it is shown that every $(u, D' - l)$ cover-free code is a nonadaptive strategy for threshold group testing, if D is unknown but bounded by D'. This implies

Theorem [3] $n'_{Th}(N,l,u,D') \leq n_c(N,u,D'-l)$, where n'_{Th} denotes the minimal number of tests of a nonadaptive strategy for threshold group testing with lower bound l, upper bound u, and D bounded by D'.

Applying a bound for cover-free codes it is shown in [3]

Theorem [3]

$$\begin{aligned} n'_{Th}(N,l,u,D') \;\leq\; & \left(\frac{u+D'-l}{D'-l}\right)^{D'-l}\left(\frac{u+D'-l}{u}\right)^{u} \cdot \\ & \left(1+(u+D'-l)\log\left(\frac{N}{u+D'-l}+1\right)\right). \end{aligned}$$

For the rate this gives

$$R \geq \frac{\left(\left(\frac{D'-l}{D'-l+u}\right)^{D'-l}\left(\frac{u}{D'-l+u}\right)^{u}\right)}{D'-l+u}. \tag{4}$$

The best known lower bound using cover-free codes for the rate of threshold group testing is given in [10] (see also [18], [19] for constructions of cover-free codes) by

$$R \geq \frac{-\log\left(1-\left(\frac{D'-l}{D'-l+u}\right)^{D'-l}\left(\frac{u}{D'-l+u}\right)^{u}\right)}{D'-l+u-1}. \tag{5}$$

One gets this lower bound by random choice of a $(u, D'-l)$ cover-free code.

We first consider the case when D is given and derive another lower bound for threshold group testing. Let $(m_{ij})_{1\leq i\leq n, 1\leq j\leq N}$ be an $n\times N$ matrix. We denote by $r_i = (m_{i1},\dots,m_{iN})$ the ith row and by $c_j = (m_{1j},\dots m_{nj})$ the jth column.

Definition 4. *We call an $n\times N$ matrix a (D,u,l)-threshold code, if for all $\mathcal{A},\mathcal{B}\subset\{1,2,\dots,N\}$, $|\mathcal{A}| = |\mathcal{B}| = D$, and $|\mathcal{A}\backslash\mathcal{B}| \geq u-l$ there exists an $i\in\{1,2,\dots,n\}$ such that*

$$\begin{array}{c} \left(\sum_{a\in\mathcal{A}} m_{ia} \geq u \quad \text{and} \quad \sum_{b\in\mathcal{B}} m_{ib} \leq l\right) \\ \text{or} \\ \left(\sum_{a\in\mathcal{A}} m_{ia} \leq l \quad \text{and} \quad \sum_{b\in\mathcal{B}} m_{ib} \geq u\right). \end{array} \tag{6}$$

We call the rows of the matrix tests and the columns codewords.

In the previous section we have shown how to get a nonadaptive group testing strategy in case $D = u+l$ by an $(u, D-l)$ separating code. A (D, u, l) threshold code is defined in such a way that it gives a nonadaptive strategy for threshold group testing for every u, l, D. Therefore we get the following

Lemma 1. *Every (D, u, l) threshold code gives a nonadaptive strategy for threshold group testing if the number D of defectives is known.*

Now we want to find a lower bound for the rate $R = \frac{\log N}{n}$ of a (D, u, l) threshold code by random choice. First we calculate the rate for codes with a weaker condition (6'), that is if (6) holds only for all $\mathcal{A}$ and $\mathcal{B}$ with $|\mathcal{A} \cap \mathcal{B}| = z$ for some z fixed.

Given an integer N, what is the minimal number (of rows) n such that a threshold code of size $n \times N$ fulfills this weaker condition?

We say that c_j is bad if there exists a pair of sets $\mathcal{A}, \mathcal{B} \subset \{1, 2, \ldots, N\}$ with $|\mathcal{A}| = |\mathcal{B}| = D$ and for which (6') is not true for any row. Otherwise we call c_j good. Consider a random matrix $(X_{ij})_{1 \le i \le n, 1 \le j \le N}$ where the X_{ij}'s are independent identical distributed random variables. We choose $P(X_{ij} = 1) = p$ and $P(X_{ij} = 0) = q$.

Let $\mathcal{A}, \mathcal{B} \subset [N]$ with $|\mathcal{A} \cap \mathcal{B}| = D - u + l$. Then every test (row) of a (D, u, l)-threshold code contains exactly l 1s inside of the positions corresponding to $|\mathcal{A} \cap \mathcal{B}|$. If there are less, then in the first set we cannot have more than u 1s, and if there are more then in the second set we will have more than l 1s. Therefore in this case

$$\begin{aligned} &P(c_j \text{ is bad} \wedge |\mathcal{A} \cap \mathcal{B}| = D - u + l) = \\ &\binom{N-1}{D+u-l-1} \cdot \binom{D+u-l-1}{D} \cdot \left(1 - 2\binom{D-u+l}{l} p^l q^{D-u} p^{u-l} q^{u-l}\right)^n . \end{aligned} \tag{7}$$

If we assign

$$n = n_* = -\frac{\log\left(\binom{N-1}{D+u-l-1}\binom{D+u-l-1}{D}\right)}{\log\left(1 - 2\binom{D-u+l}{l} p^u q^{D-l}\right)} + 1$$

then the right-hand side of (7) does not exceed $\frac{1}{2}$ and the average number of bad columns does not exceed $\frac{N}{2}$. Thus there exists a matrix which has at least $\frac{N}{2}$ good columns. By using $R \ge \lim_{N\to\infty} \frac{\log \frac{N}{2}}{n_*}$ we get

$$R \ge \frac{-\log\left(1 - 2\binom{D-u+l}{l} p^u q^{D-l}\right)}{D + u - l - 1}. \tag{8}$$

We want to consider the general case. We say that c_j is bad if there exists a pair of sets $\mathcal{A}, \mathcal{B} \subset \{1, 2, \ldots, N\}$ with $|\mathcal{A}| = |\mathcal{B}| = D$ and for which (6) is not true for any row. Clearly

$$P(c_j \text{ is bad}) = \sum_{k=0}^{D-u+l} P(c_j \text{ is bad} \wedge |\mathcal{A} \cap \mathcal{B}| = k). \tag{9}$$

If $|\mathcal{A} \cap \mathcal{B}| = k$ we get

$$P(c_j \text{ is bad} \wedge |\mathcal{A} \cap \mathcal{B}| = k) = \tbinom{N-1}{2D-k-1} \cdot \tbinom{2D-k-1}{D} \cdot$$
$$\left(1 - 2(\textstyle\sum_{j=0}^{\min\{k,l\}} \binom{k}{j} p^j q^{k-j} (\sum_{i=u-j}^{D-k} \binom{D-k}{i} p^i q^{D-k-i})(\sum_{t=0}^{l-j} \binom{D-k}{h} p^t q^{D-k-t}))\right)^n. \tag{10}$$

Now we need an upper bound

$$P(c_j \text{ is bad}) \le (D-u+l+1) \max_{k \in \{0,1,\ldots,D-u+l\}} P(c_j \text{ is bad} \wedge |\mathcal{A} \cap \mathcal{B}| = k). \tag{11}$$

We calculate for each k the rate like for $k = D - u + l$. The factor $(D - u + l + 1)$ in (11) does not change the rate. Therefore the minimal of these rates gives a bound for the rate in the general case.

Hence we get

Theorem 2. *Let $0 \le l < u \le D$ be given, then*

$$R_{Th} \ge R_T = \max_{0 \le p \le 1} \min_{0 \le k \le D-u+l} \tag{12}$$
$$\frac{-\log(1 - 2(\sum_{j=0}^{\min\{k,l\}} \binom{k}{j} p^j q^{k-j} (\sum_{i=u-j}^{D-k} \binom{D-k}{i} p^k q^{D-k-i})(\sum_{t=0}^{l-j} \binom{D-k}{t} p^t q^{D-k-t})))}{2D-k-1}.$$

This formula is better than the known one, because every $(u, D-l)$ cover-free code is a (D, u, l) threshold code. The bound for the rate of cover-free codes is derived in the same way as we did for threshold codes. Recall that the best known bound for the rate, using cover-free codes is

$$R_{Th} \ge R_C = \frac{-\log\left(1 - \left(\frac{D-l}{D-l+u}\right)^{D-l} \left(\frac{u}{D-l+u}\right)^u\right)}{D-l+u-1}. \tag{13}$$

It holds $-\log(1-x) \sim x$ if x is small. Thus we want to compare

$$R_T = \max_{0\le p\le 1} \min_{0\le k\le D-u+l} \tag{14}$$
$$\frac{2(\sum_{j=0}^{\min\{k,l\}} \binom{k}{j} p^j q^{k-j} (\sum_{i=u-j}^{D-k} \binom{D-k}{i} p^i q^{D-k-i})(\sum_{t=0}^{l-j} \binom{D-k}{t} p^t q^{D-k-t})))}{2D-k-1}$$

and

$$R_C = \frac{\left(\frac{D-l}{D-l+u}\right)^{D-l} \left(\frac{u}{D-l+u}\right)^{u}}{D-l+u-1}. \tag{15}$$

Note first that, since $u \ge l+1$, for all $0 \le k \le D-u+l$

$$F_T = \frac{2}{2D-k-1} \ge \frac{2}{2D-1} \ge F_C = \frac{1}{D-l+u-1} \ge \frac{1}{D}. \tag{16}$$

We have $\frac{F_T}{F_C} > 1$, for instance for $u = l+1$ and $k = D-1$ $\frac{F_T}{F_C} = 2$, or $\lambda = 0$, $\mu = 1$, and therefore $\kappa = 0$, as it occurs in the result of [2].

We start with our analysis for $u = l+1$, that is in relative quantities

$$\lambda = \frac{l}{D},\ \mu = \frac{u}{D},\ \kappa = \frac{k}{D}, \tag{17}$$

and for the probabilities in (15)

$$\frac{u}{D-l+u} = \frac{\mu}{1-\lambda+\mu} = \mu = p, \quad 1-\mu = q. \tag{18}$$

We begin first with the **entropy description** of the lower bounds for $u = l+1$

$$R_C = F_C \mu^{\mu D}(1-\mu)^{1-\mu D} = F_C 2^{-h(\mu)D}, \tag{19}$$

$$R_T \ \ge\ F_T \min_{0\le k\le D-u+l} \sum_{j=0}^{\min\{k,l\}} \binom{k}{j} p^j q^{k-j} \cdot \tag{20}$$
$$(\sum_{i=u-j}^{D-k} \binom{D-k}{i} p^i q^{D-k-i}) \cdot$$
$$(\sum_{t=0}^{l-j} \binom{D-k}{t} p^t q^{D-k-t}).$$

For $u = l + 1$ we have

$$\begin{aligned} R_T(l) &\geq F_T \min_{0 \leq k \leq l} \sum_{j=0}^{k} \binom{k}{j} \mu^j (1-\mu)^{k-j} \\ & (1-*) \sum_{t=0}^{l-j} \binom{D-k}{t} \mu^t (1-\mu)^{k-j}, \end{aligned}$$

where $(1-*)SUM$ stands for $(1-SUM)SUM$.

Argument We choose the optimal j for all sums (max. entropy principle). If $j^* = \mu k = \lambda \kappa D$, then

$$\binom{k}{j^*} \mu^{j^*} (1-\mu)^{k-j^*} \geq \frac{1}{k+1} \geq \frac{1}{D} \tag{21}$$

$$\binom{D-k}{u-j^*} \mu^{u-j^*} (1-\mu)^{D-k-u+j^*} \geq \frac{1}{D-k+1} \geq \frac{1}{D}$$

$$\binom{D-k}{l-j^*} \mu^{l-j^*} (1-\mu)^{D-k-l+j^*} \geq \frac{1}{D-k+1} \geq \frac{1}{D}.$$

Consequently

$$\frac{R_T(l)}{R_C} \geq \frac{F_T}{F_C} 2^{h(\mu)D - o(D)}. \tag{22}$$

Therefore we have

Theorem 3. *In the case without gap, that is $u = l + 1$, it holds*

$$\frac{R_T}{R_C} \geq 2^{h(\mu)D - o(D)}. \tag{23}$$

We write now

$$R_T = F_T \min_{0 \leq k \leq D-u+l} E_1 E_2 E_3$$

with $E_1 = \sum_{j=0}^{\min\{k,l\}} \binom{k}{j} p^j q^{k-j}$, $E_2 = \sum_{i=u-j}^{D-k} \binom{D-k}{i} p^i q^{D-k-i}$,

and $E_3 = \sum_{t=0}^{l-j} \binom{D-k}{t} p^t q^{D-k-t}$.

To approximate E_1 for all k observe first that we need

$$j(\kappa) \leq \kappa p D \text{ and } j(\kappa) \leq \lambda D. \tag{24}$$

Let $\kappa \leq 1 - \mu + \lambda$ and

$$p = p(\mu, \lambda) = \frac{\lambda}{1 - \mu + \lambda} \tag{25}$$

then for $j = \frac{\lambda}{1-\mu+\lambda}\kappa D$ the inequalities in (24) hold.

Lemma 2. *1. $\lambda \leq p(\mu, \lambda) \leq \mu$.*
2. $p(\mu, \lambda) \leq \frac{\mu+\lambda}{2}$ for $\mu \leq \frac{1}{2}$.
3. The inequality $\sqrt{\lambda\mu} \leq p(\mu, \lambda)$ does not hold.

Proof.

1. Since $\mu \geq \lambda$, $1 - \mu + \lambda \leq 1$ and $\lambda \leq p(\mu, \lambda)$ then

$$\lambda = (1 - \mu + \lambda)p(\mu, \lambda) \leq (1 + (\lambda - \mu))\mu$$

or equivalently

$$\lambda - \mu \leq (\lambda - \mu)\mu,$$

which holds, because $0 \leq \mu \leq 1$ and $\lambda - \mu$ is negative.

2. The inequality is equivalent to

$$\lambda \leq \frac{\mu + \lambda}{2} - \frac{\mu^2 - \lambda^2}{2} \tag{26}$$

or to

$$\lambda - \lambda^2 \leq \mu - \mu^2,$$

which holds, because $\lambda \leq \mu \leq \frac{1}{2}$ and thus

$$\lambda(1 - \lambda) \leq \mu(1 - \mu).$$

3. Counterexample: $\mu = \frac{1}{2}$ and $\lambda = \frac{1}{4}$.

It remains to estimate E_2 and E_3 from below

$$E_2 = \sum_{i=u-j(\kappa)}^{D-k} \binom{D-k}{i} p^i q^{D-k-i},$$

where $u - j(\kappa) = (\mu - \frac{\lambda}{1-\mu+\lambda}\kappa)D \leq p(1 - \kappa)D$.

$$E_2 \geq \binom{(1-\kappa)D}{(\mu - p\kappa)D} p^{(\mu - p\kappa)D}(1-p)^{(1-\mu-(1-p)\kappa)D}$$

$$E_2 \geq 2^{D(1-\kappa)[h(\frac{\mu-p\kappa}{1-\kappa})+\frac{\mu-p\kappa}{1-\kappa}\log p+\frac{1-\mu-(1-p\kappa)}{1-\kappa}\log(1-p)]}.$$

We set

$$f(\kappa,\mu,\lambda,p) = h(\frac{\mu-p\kappa}{1-\kappa}) + \frac{\mu-p\kappa}{1-\kappa}\log p + \frac{1-\mu-(1-p\kappa)}{1-\kappa}\log(1-p)]$$

and want to find $\min_\kappa f(\kappa,\mu,\lambda,p)$. Let $r = \frac{\mu-p\kappa}{1-\kappa}$, then we have

$$\begin{aligned}
& -r\log r-(1-r)\log(1-r)+r\log p+(1-r)\log\frac{1-p}{1-r} \\
= \;& r\log\frac{p}{r}+(1-r)\log\frac{1-p}{1-r} \\
= \;& -D((r,1-r)||(p,1-p)),
\end{aligned}$$

where $D((r,1-r)||(p,1-p))$ is called information divergence. For E_3 we have

$$E_3 = \sum_{t=0}^{l-j(\kappa)} \binom{D-k}{t} p^t q^{D-k-t},$$

where $l-j(\kappa) = (\lambda-p\kappa)D$. Therefore

$$E_3 \geq 2^{D(1-\kappa)[h(\frac{\lambda-p\kappa}{1-\kappa})+\frac{\lambda-p\kappa}{1-\kappa}\log p+\frac{1-\kappa-lambda+p\kappa)}{1-\kappa}\log(1-p)]}.$$

We set

$$g(\kappa,\mu,\lambda,p) = h(\frac{\lambda-p\kappa}{1-\kappa}) + \frac{\lambda-p\kappa}{1-\kappa}\log p + \frac{1-\kappa-\lambda+p\kappa}{1-\kappa}\log(1-p)]$$

and want to find $\min_\kappa f(\kappa,\mu,\lambda,p)$. Let $s = \frac{\lambda-p\kappa}{1-\kappa}$, then we have

$$\begin{aligned}
& -s\log s-(1-s)\log(1-s)+s\log p+(1-s)\log\frac{1-p}{1-s} \\
= \;& s\log\frac{p}{s}+(1-s)\log\frac{1-p}{1-s} \\
= \;& -D((s,1-s)||(p,1-p)).
\end{aligned}$$

It follows

Proposition 1.

$$R_T \sim F_T 2^{(-D((r,1-r)||(p,1-p))-D((s,1-s)||(p,(1-p)))(1-\kappa)}.$$

Therefore we compare

$-D((r,1-r)||(p,1-p))-D((s,1-s)||(p,(1-p)))(1-\kappa)$ with $-\max(h(\lambda),h(\mu))$.

We need some basic calculation for the bound on R_C.

Lemma 3. *Denote $P = P(\lambda,\mu) = \frac{\mu}{1-\lambda+\mu}$, then*

$$R_C/F_C = \left(\frac{u}{D-l+u}\right)^u \left(\frac{D-l}{D-l+u}\right)^{D-l} = 2^{-h(P)(1-\lambda+\mu)D}.$$

Proof. We have

$$\begin{aligned} & \left(\frac{\mu}{1-\lambda+\mu}\right)^{\mu D} \left(\frac{1-\lambda}{1-\lambda+\mu}\right)^{(1-\lambda)D} \\ = \;& P^{PD(1-\lambda+\mu)}(1-P)^{(1-P)D(1-\lambda+\mu)} \\ = \;& 2^{-h(P)(1-\lambda+\mu)D}. \end{aligned}$$

We note that $1-\lambda+\mu > 1$ and $P > \mu$.
So far $\frac{1}{2} \geq P = \frac{\mu}{1-\lambda+\mu}$ and we have

$$h(P)(1-\lambda+\mu) > h(\mu) > h(\lambda).$$

Now we want to show that

$$h(P)(1-\lambda+\mu) > \max_{\kappa}(1-\kappa)(D((r,1-r)||(p,1-p)+D((s,1-s)||(p,1-p))$$

with $P = \frac{\mu}{1-\lambda+\mu}$, $r = \frac{\mu-p\kappa}{1-\kappa}$, $s = \frac{\lambda-p\kappa}{1-\kappa}$, and $p = \frac{\lambda}{1-\mu+\lambda}$.
We consider the special case $D = u+l$. Then we have $1 = \mu+\lambda$, $0 \leq \kappa \leq 2\lambda$, $P = \frac{1}{2}$, and $p = \frac{1}{2}$.

For $\mu = \lambda$ we have

$$h(P)(1-\lambda+\mu) = 2\mu > 0 = (1-\kappa)2D((\frac{1}{2},\frac{1}{2})||(\frac{1}{2},\frac{1}{2})).$$

For $\lambda < \mu$ we have $r = \frac{2\mu-\kappa}{2-2\kappa}$ and $s = \frac{2\lambda-\kappa}{2-2\kappa}$.

We have to show that for $0 \leq \kappa \leq \lambda$

$$2\mu > \max_{\kappa} -(1-\kappa)(D((\frac{2\mu-\kappa}{2-2\kappa}, 1-\frac{2\mu-\kappa}{2-2\kappa})||(\frac{1}{2},\frac{1}{2})+D((\frac{2\lambda-\kappa}{2-2\kappa}, 1-\frac{2\lambda-\kappa}{2-2\kappa})||(\frac{1}{2},\frac{1}{2})).$$

This is equivalent to (because $\lambda = 1-\mu$)

$$\begin{aligned}
2 &\leq h(\frac{2\mu-\kappa}{2-2\kappa}) + h(\frac{2(1-\mu)-\kappa}{2-2\kappa}) + \frac{2\mu}{1-\kappa} \qquad (27)\\
2 &\leq h(\frac{2\mu-\kappa}{2-2\kappa}) + h(1-\frac{2\mu-\kappa}{2-2\kappa}) + \frac{2\mu}{1-\kappa}\\
2 &\leq 2h(1-\frac{2\mu-\kappa}{2-2\kappa}) + \frac{2\mu}{1-\kappa}\\
1-\frac{2\mu}{2-2\kappa} &\leq h(1-\frac{2\mu-\kappa}{2-2\kappa}). \qquad (28)
\end{aligned}$$

(27) is true for $\mu > 1-\kappa$. For $\mu \leq 1-\kappa$ it holds $\frac{1}{2} \leq \frac{2\mu-\kappa}{2-2\kappa} \leq 1$. Therefore

$$1-\frac{2\mu-\kappa}{2-2\kappa} \leq h(1-\frac{2\mu-\kappa}{2-2\kappa})$$

and (28) holds.

For $\kappa = 0$ and $\mu = 1$, that means $\lambda = 0$, the two terms are the same. In general we get

Theorem 4. *Let $1 = \mu + \lambda$ and $\lambda \leq \mu$, then*

$$R_T \geq 2^{2\lambda)} R_C.$$

4. Computer Results

We set $R_T = F_T \cdot E_1 \cdot E_2 \cdot E_3$ and $R_C = F_C E_0$ with
$E_0 = \left(\frac{D-l}{D-l+u}\right)^{D-l}\left(\frac{u}{D-l+u}\right)^{u}$
$E_1 = \sum_{j=0}^{\min\{k,l\}} \binom{k}{j} p^j q^{k-j}$
$E_2 = \sum_{i=u-j}^{D-k} \binom{D-k}{i} p^i q^{D-k-i}$

$E_3 = \sum_{t=0}^{l-j} \binom{D-k}{t} p^t q^{D-k-t}$

We set $\mu D = u$, $\lambda D = l$, $\kappa D = k$, $\alpha D = j$
$E_0 = 2^{De_0}$ with $e_0 = -h(\frac{\mu}{(1-\lambda+\mu)})(1-\lambda+\mu)$

$E_1 = 2^{De_1}$ with $e_1 = \kappa h(\frac{\alpha}{\kappa}) + \alpha \log(p) + (\kappa - \alpha)\log(1-p)$

$E_2 = 2^{De_2}$ with $e_2 = (1-\kappa)h(\frac{f_2}{1-\kappa}) + (1-\kappa-f_2)\log(1-p) + f_2 \log(p)$ where $f_2 = (\mu - \alpha)\beta + (1-\kappa)(1-\beta)$.

$E_3 = 2^{De_3}$ with $e_3 = (1-\kappa)h(\frac{f_3}{1-\kappa}) + f_3 \log(p) + (1-\kappa-f_3)\log(1-p)$ where $f_3 = (\lambda - \alpha)\gamma$.

The computer calculates now for given μ and λ:

$$\max_{0\le p\le 1} \min_{0\le\kappa\le 1-\mu+\lambda} \max_{0\le\gamma,\beta\le 1, \max\{0,\mu+\kappa-1\}\le\alpha\le\min\{\kappa,\lambda\}} e_1 + e_2 + e_3 - e_0.$$

We have $\frac{R_T}{R_C} = 2^{D(e_1+e_2+e_3-e_0)}$.

Observations:

- $0 \le e_1 + e_2 + e_3 - e_0 \le 1$
- The biggest value for $e_1 + e_2 + e_3 - e_0$ we have for $\mu = \lambda = \frac{1}{2}$
- The smallest value for $e_1 + e_2 + e_3 - e_0$ we have for ($\lambda = 0$ and μ arbitrary) and for ($\mu = 1$ and λ arbitrary).
- For $\mu = \lambda$ (13) gives the correct value.
- $p = \frac{\mu}{1-\lambda+\mu}$ gives always the biggest value (therefore it is not listed in the table below).

λ	μ	$e_1 + e_2 + e_3 - e_0$	λ	μ	$e_1 + e_2 + e_3 - e_0$
0	x	0.00	0.4	0.4	0.97
0.1	0.1	0.47	0.4	0.5	0.89
0.1	0.2	0.45	0.4	0.6	0.80
0.1	0.3	0.43	0.4	0.7	0.69
0.1	0.4	0.41	0.4	0.8	0.55
0.1	0.5	0.39	0.4	0.9	0.36
0.1	0.6	0.36	0.4	1.0	0.00
0.1	0.7	0.32	0.5	0.5	1.00
0.1	0.8	0.27	0.5	0.6	0.89
0.1	0.9	0.21	0.5	0.7	0.76
0.1	1.0	0.00	0.5	0.8	0.60
0.2	0.2	0.72	0.5	0.9	0.39
0.2	0.3	0.69	0.5	1.0	0.00
0.2	0.4	0.65	0.6	0.6	0.97
0.2	0.5	0.60	0.6	0.7	0.83
0.2	0.6	0.55	0.6	0.8	0.65
0.2	0.7	0.49	0.6	0.9	0.41
0.2	0.8	0.40	0.6	1.0	0.00
0.2	0.9	0.28	0.7	0.7	0.88
0.2	1.0	0.00	0.7	0.8	0.69
0.3	0.3	0.88	0.7	0.9	0.43
0.3	0.4	0.83	0.7	1.0	0.00
0.3	0.5	0.76	0.8	0.8	0.72
0.3	0.6	0.69	0.8	0.9	0.45
0.3	0.7	0.60	0.8	1.0	0.00
0.3	0.8	0.49	0.9	0.9	0.47
0.3	0.9	0.32	0.9	1.0	0.00
0.3	1.0	0.00	1.0	1.0	0.00

5. Majority group testing

We remind the reader of the definition of majority group testing in Section 1.

This is a generalization of the model considered in [14].

Furthermore the author introduced and analyzed the following test:

1.
$$t_{\mathcal{S}}(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < \frac{D}{2} \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| \geq \frac{D}{2}. \end{cases}$$

He mentioned as other tests:

2.
$$t_{\mathcal{S}}(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < \frac{D}{2} \\ & , \text{ or } |\mathcal{N} \backslash \mathcal{S}| > |\mathcal{S}| \text{ and } |\mathcal{S} \cap \mathcal{D}| = \frac{D}{2} \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| > \frac{D}{2} \\ & , \text{ or } |\mathcal{N} \backslash \mathcal{S}| \leq |\mathcal{S}| \text{ and } |\mathcal{S} \cap \mathcal{D}| = \frac{D}{2}. \end{cases}$$

3.
$$t_{\mathcal{S}}(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < \frac{D}{2} \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| > \frac{D}{2} \\ * & , \text{ otherwise.} \end{cases}$$

It is clear that the test 1. is included in our definition of majority group testing. The other two are not included.

We will now consider the case $f_1(D) = \lceil \frac{D}{k} \rceil - 1$ and $f_2(D) = \lfloor \frac{D}{k} \rfloor + 1$ and write $f(D) = \frac{D}{k}$.

$$t_{\mathcal{S}}(\mathcal{D}) = \begin{cases} 0 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| < \frac{D}{k} = f(D) \\ 1 & , \text{ if } |\mathcal{S} \cap \mathcal{D}| > \frac{D}{k} = f(D) \\ \{0, 1\} & , \text{ otherwise.} \end{cases}$$

If D is known this problem can be reduced to threshold group testing:

1. For $D \mod k \equiv 0$ we set $l = \frac{D}{k} - 1$ and $u = \frac{D}{k} + 1$. Therefore we get a strategy by a $(\frac{D}{k} + 1, \frac{k-1}{k} D + 1)$ cover-free code, by a $(\frac{D}{2} + 1, \frac{D}{2} + 1)$ separating code for $k = 2$, or by a $(D, \frac{D}{k} + 1, \frac{k-1}{k} D + 1)$ threshold code.
2. For $D \mod k \equiv s$ with $0 < s < k$ we set $l = \frac{D-s}{k}$ and $u = \frac{D+k-s}{k}$. Therefore we get strategy by a $(\frac{D+k-s}{k}, \frac{(k-1)D+s}{k})$ cover-free code, by a $(\frac{D+1}{2}, \frac{D+1}{2})$ separating code for $k = 2$, or by a $(D, \frac{D+k-s}{k}, \frac{(k-1)D+s}{k})$ threshold code.

Now we will consider the case when D is bounded by some $D' < N$. The number of tests depends on D'.

First we consider the case $k = 2$.

It is clear that as in threshold group testing it is not always possible to determine the set of defectives.

Definition 5. *We say that two sets $\mathcal{A}, \mathcal{B}$ of possible defective elements are* **indistinguishable** *if for all strategies the results can be the same for both $\mathcal{A}$ and $\mathcal{B}$. We call a set $\mathbb{F}$ of subsets of $[N]$ a* **solution**, *if the set of all defective elements $\mathcal{D} \in \mathbb{F}$ and the sets in $\mathbb{F}$ are pairwise indistinguishable. The next theorem gives conditions for a solution.*

Theorem 5. *Let $\mathcal{D} \subset [N]$ be the set of defectives and $f(D) = \frac{D}{2}$. We can determine a solution $\mathbb{F}$ such that for every sets $\mathcal{P}_1, \mathcal{P} \in \mathbb{F}$ with $|\mathcal{P}_1| \geq |\mathcal{P}|$ the following holds*

1. *If $\mathcal{P} \subset \mathcal{P}_1$ and $\mathcal{P}$ is even then*

$$|\mathcal{P}_1 \backslash \mathcal{P}| \leq 2. \tag{29}$$

2. *If $\mathcal{P} \subset \mathcal{P}_1$ and $\mathcal{P}$ is odd then*

$$|\mathcal{P}_1 \backslash \mathcal{P}| \leq 1. \tag{30}$$

3. *If $|\mathcal{P} \backslash \mathcal{P}_1| = 1$ and $\mathcal{P}$ is even then*

$$|\mathcal{P}_1 \backslash \mathcal{P}| \leq 1. \tag{31}$$

4. *In all other cases we can distinguish two sets $\mathcal{A}, \mathcal{B}$ with $\mathcal{A} \neq \mathcal{B}$.*

Proof. First we show that there exists a strategy, such that we get a set $\mathbb{F}$ which satisfies (29), (30), and (31). We consider the set $\mathbb{F}_0 = \cup_{j=0}^{D'} \binom{[N]}{j}$ of all possible sets of defectives. If there are two sets in $\mathbb{F}_0$ which do not fulfill (29), (30), and (31) we show that there exists a test such that one will be removed and we show that the set of defectives will not be removed.

1. Let $\mathcal{P} \subset \mathcal{P}_1$ and $|\mathcal{P}| = 2a$.
   It is enough to have a test set $\mathcal{S}$ which includes exactly $a-1$ elements of $\mathcal{P}$ and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \mathcal{P}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
   If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 < \frac{D}{2} = a$ and the result is "0".
   If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 + w > \frac{D}{2} = a + \frac{w}{2}$, because $|\mathcal{P}_1 \backslash \mathcal{P}| = w \geq 3$ by assumption and the result is "1".

2. Let $\mathcal{P} \subset \mathcal{P}_1$ and $|\mathcal{P}| = 2a + 1$.
   It is enough to have a test set $\mathcal{S}$ which includes exactly a elements of $\mathcal{P}$ and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
   If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a < \frac{D}{2} = a + \frac{1}{2}$ and the result is "0".
   If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a + w > \frac{D}{2} = a + \frac{w+1}{2}$, because $|\mathcal{P}_1 \backslash \mathcal{P}| = w \geq 2$ by assumption and the result is "1".
3. Let $\mathcal{P} \not\subset \mathcal{P}_1$, $|\mathcal{P} \backslash \mathcal{P}_1| = 1$, and $|\mathcal{P}| = 2a$. It is enough to have a test set $\mathcal{S}$ which includes exactly $a - 1$ elements of $\mathcal{P}$ and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
   If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 < \frac{D}{2} = a$ and the result is "0".
   If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 + w > \frac{D}{2} = a + \frac{w-1}{2}$, because $|\mathcal{P}_1 \backslash \mathcal{P}| = w \geq 2$ by assumption and the result is "1".
4. Let $\mathcal{P} \not\subset \mathcal{P}_1$, $|\mathcal{P} \backslash \mathcal{P}_1| > 1$, and $|\mathcal{P}| = 2a$.
   (a) $|\mathcal{P} \cap \mathcal{P}_1| = l \geq a - 1$.
   It is enough to have a test set $\mathcal{S}$ which includes exactly $a - 1$ elements of $\mathcal{P} \cap \mathcal{P}_1$, no other elements of $\mathcal{P}$ and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
   If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 < \frac{D}{2} = a$ and the result is "0".
   If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 + w > \frac{D}{2} = a + \frac{w - |\mathcal{P} \backslash \mathcal{P}_1|}{2}$, because $|\mathcal{P} \backslash \mathcal{P}_1| > 1$ by assumption and the result is "1".
   (b) $|\mathcal{P} \cap \mathcal{P}_1| = l < a - 1$.
   It is enough to have a test set $\mathcal{S}$ which includes all elements of $\mathcal{P} \cap \mathcal{P}_1$, $a - 1 - |\mathcal{P} \cap \mathcal{P}_1|$ elements of $\mathcal{P}$, and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
   If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 < \frac{D}{2} = a$ and the result is "0".
   If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = D > \frac{D}{2}$ and the result is "1".
5. Let $\mathcal{P} \not\subset \mathcal{P}_1$, $|\mathcal{P} \backslash \mathcal{P}_1| > 1$, and $|\mathcal{P}| = 2a + 1$.
   (a) $|\mathcal{P} \cap \mathcal{P}_1| = l \geq a$.
   It is enough to have a test set $\mathcal{S}$ which includes exactly a elements of $\mathcal{P} \cap \mathcal{P}_1$, no other elements of $\mathcal{P}$ and the remaining w elements

of $\mathcal{P}\backslash\mathcal{P}_1$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0\backslash\{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0\backslash\{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a < \frac{D}{2} = a + \frac{1}{2}$ and the result is "0".
If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a + w > \frac{D}{2} = a + \frac{w - |\mathcal{P}\backslash\mathcal{P}_1|}{2}$, because $|\mathcal{P}\backslash\mathcal{P}_1| > 0$ by assumption and the result is "1".

(b) $|\mathcal{P} \cap \mathcal{P}_1| = l < a$.
It is enough to have a test set $\mathcal{S}$ which includes all elements of $\mathcal{P} \cap \mathcal{P}_1$, $a - |\mathcal{P} \cap \mathcal{P}_1|$ elements of $\mathcal{P}$, and the remaining w elements of $\mathcal{P}_1\backslash\mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0\backslash\{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0\backslash\{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a < \frac{D}{2} = a + \frac{1}{2}$ and the result is "0".
If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = D > \frac{D}{2}$ and the result is "1".

In all cases it is not possible to distinguish the sets of possible solutions. $\square$

Remarks:

1. In the corresponding matrix of a strategy for majority group testing with $f(D) = \frac{D}{2}$ it is possible to exchange the zeros and the ones.
2. The proof of Theorem 5 shows that for a successful strategy we have to have for all disjoint pairs $\mathcal{J}, \mathcal{I} \subset [N]$ with the size $\lfloor \frac{D'}{2} \rfloor + 1$ two test sets such that the elements of $\mathcal{J}$ are contained in one test set and no element of $\mathcal{I}$ is contained in the other test set or vice versa. This is exactly a separating code and therefore we have the following

Theorem 6. *Let $f(D) = \frac{D}{2}$ and D be bounded by D', which is known, then a $(\lfloor \frac{D'}{2} \rfloor + 1, \lfloor \frac{D'}{2} \rfloor + 1)$ separating code gives a nonadaptive strategy for majority group testing.*

Now let us consider the case $k > 2$.

As before we have conditions for a solution. They are given by the following

Theorem 7. *Let $\mathcal{D} \subset [N]$ be the set of defectives and $f(D) = \frac{D}{k}$. We can determine a solution $\mathbb{F}$ such that for every sets $\mathcal{P}_1, \mathcal{P} \in \mathbb{F}$ with $|\mathcal{P}_1| \geq |\mathcal{P}| = ak + s$ $(0 \leq s \leq k-1)$ the following holds*

1. *If* $\mathcal{P} \subset \mathcal{P}_1$ *and* $s = 0$ *or* $s = k - 1$ *then*

$$|\mathcal{P}_1 \backslash \mathcal{P}| \leq 1. \tag{32}$$

2. *If* $|\mathcal{P} \backslash \mathcal{P}_1| = 1$ *and* $s = 0$ *then*

$$|\mathcal{P}_1 \backslash \mathcal{P}| \leq 1. \tag{33}$$

3. *In all other cases we can distinguish two sets* $\mathcal{A}, \mathcal{B}$ *with* $\mathcal{A} \neq \mathcal{B}$.

Proof.

The proof follows the same ideas as the proof of Theorem 5. First we show that there exists a strategy, such that we get a set $\mathbb{F}$ which fulfills (32) and (33). We consider the set $\mathbb{F}_0 = \cup_{j=0}^{D'} \binom{[N]}{j}$ of all possible sets of defectives. If there are two sets in $\mathbb{F}_0$ which do not fulfill (32) and (33) we show that there exists a test such that one will be removed and we show that the set of defectives will not be removed.

1. Let $\mathcal{P} \subset \mathcal{P}_1$ and $|\mathcal{P}| = ak + s$ with $0 \leq s \leq k - 1$.
   (a) $s = 0$:
   It is enough to have a test set $\mathcal{S}$ which includes exactly $a - 1$ elements of $\mathcal{P}$ and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
   If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 < \frac{D}{2} = a$ and the result is "0".
   If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 + w > \frac{D}{2} = a + \frac{w}{k}$, because $|\mathcal{P}_1 \backslash \mathcal{P}| = w \geq 2$ by assumption and $k > 2$. Thus the result is "1".
   (b) $s = k - 1$:
   It is enough to have a test set $\mathcal{S}$ which includes exactly a elements of $\mathcal{P}$ and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
   If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a < \frac{D}{2} = a + \frac{k-1}{k}$ and the result is "0".
   If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a + w > \frac{D}{2} = a + \frac{w+k-1}{k}$, because $|\mathcal{P}_1 \backslash \mathcal{P}| = w \geq 2$ by assumption and the result is "1".
2. Let $\mathcal{P} \not\subset \mathcal{P}_1$, $|\mathcal{P} \backslash \mathcal{P}_1| = 1$, and $s = 0$.
   It is enough to have a test set $\mathcal{S}$ which includes exactly $a - 1$ elements of $\mathcal{P} \cap \mathcal{P}_1$ and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this

test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 < \frac{D}{k} = a$ and the result is "0".
If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a - 1 + w > \frac{D}{2} = a + \frac{w-1}{k}$, because $|\mathcal{P}_1 \backslash \mathcal{P}| = w \geq 2$ by assumption and the result is "1".

3. Let $\mathcal{P} \not\subset \mathcal{P}_1$, $|\mathcal{P} \backslash \mathcal{P}_1| > 1$, and $s \neq 0$.
   (a) $|\mathcal{P} \cap \mathcal{P}_1| = l \geq a$.
   It is enough to have a test set $\mathcal{S}$ which includes exactly a elements of $\mathcal{P} \cap \mathcal{P}_1$, no other elements of $\mathcal{P}$ and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
   If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a < \frac{D}{k} = a + \frac{s}{k}$ and the result is "0".
   If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a + w > \frac{D}{2} = a + \frac{w - |\mathcal{P} \backslash \mathcal{P}_1| + s}{w}$, because $|\mathcal{P} \backslash \mathcal{P}_1| \geq 2$ and $s \leq k - 1$ by assumption and the result is "1".
   (b) $|\mathcal{P} \cap \mathcal{P}_1| = l < a$.
   It is enough to have a test set $\mathcal{S}$ which includes all elements of $\mathcal{P} \cap \mathcal{P}_1$, $a - |\mathcal{P} \cap \mathcal{P}_1|$ elements of $\mathcal{P}$, and the remaining w elements of $\mathcal{P}_1 \backslash \mathcal{P}$. If the result of this test is "1" we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}\}$, otherwise we continue with $\mathbb{F}_1 = \mathbb{F}_0 \backslash \{\mathcal{P}_1\}$. Now we have to show that we will not remove $\mathcal{D}$.
   If $\mathcal{P} = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = a < \frac{D}{k} = a + \frac{s}{k}$ and the result is "0".
   If $\mathcal{P}_1 = \mathcal{D}$ then $|\mathcal{S} \cap \mathcal{D}| = D > \frac{D}{2} = \frac{D}{2}$ and the result is "1".

If the sets fulfill (32) and (33) it is not possible to distinguish them, because it is possible to get the same test results if $\mathcal{P} = \mathcal{D}$ or if $\mathcal{P}_1 = \mathcal{D}$. $\square$

The proof of Theorem 7 shows that every $(\lfloor \frac{D'}{k} \rfloor + 1, D' - \lceil \frac{D'}{k} \rceil + 1)$ cover-free code gives a strategy and thus the following

Theorem 8. *Let $f(D) = \frac{D}{k}$ and D be bounded by D', which is known, then a $(\lfloor \frac{D'}{k} \rfloor + 1, D' - \lceil \frac{D'}{k} \rceil + 1)$ cover-free code is a nonadaptive strategy for majority group testing.*

Acknowledgment

The authors would like to thank Christian Kleinewächter for improvements of the computer program.

References

[1] R. Ahlswede and I. Wegener, Suchprobleme, Teubner Verlag, Stuttgart, 1979, Russian Edition: Zadatsi Poiska, MIR, 1982, English Edition: Search Problems, Wiley-Interscience Series in Discrete Mathematics and Optimization, 1987.

[2] N. Alon, V. Guruswami, T. Kaufman, and M. Sudan, Guessing secrets efficiently via list decoding, 13th SODA, 254-262, 2002.

[3] H.-B. Chen and H.-L. Fu, Nonadaptive algorithms for threshold group testing, Discrete Appl. Math. 157, No. 7, 1581-1585, 2009.

[4] H.-B. Chen, H.-L. Fu, and F.K. Hwang, An upper bound of the number of tests in pooling designs for the error-tolerant complex model, Opt. Lett. 2, 425-431, 2008.

[5] F. Chung, R. Graham, and F.T. Leighton, Guessing secrets, Electronic J. on Combinatorics, 8, 1-25, 2001.

[6] G. Cohen and H.G. Schaathun, Asymptotic overview on separating codes, Technical report no. 248 from Department of Informatics, University of Bergen, 2003.

[7] P. Damaschke, Threshold group testing, General Theory of Information Transfer and Combinatorics, R. Ahlswede et al. editors, Lecture Notes in Computer Science, Vol. 4123, Springer Verlag, 707-718, 2006.

[8] D.Z. Du and F.K. Hwang, Combinatorial Group Testing and its Applications, 2nd edition, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, Series on Applied Mathematics, 12, 2000.

[9] D.Z. Du and F.K. Hwang, Pooling Designs and Nonadaptive Group Testing. Important Tools for DNA Sequencing, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, Series on Applied Mathematics, 18, 2006.

[10] A. D'yachkov, A. Macula, P. Vilenkin, and D. Torney, Families of finite sets in which no intersection of l sets is covered by the union of s others, J. Combin. Theory Ser. A 99, No. 2, 195–218, 2002.

[11] A. D'yachkov, A. Macula, P. Vilenkin, and D. Torney, Two models of nonadaptive group testing for designing screening experiments, Proc. 6th Int.Workshop on Model-Orented Designs and Analysis, 63-75, 2001.

[12] W. Kautz and R. Singleton, Nonrandom binary superimposed codes, IEEE Trans. Information Theory, Vol. 10, No. 4, 363-377, 1964.

[13] V.S. Lebedev, An asymptotic upper bound for the rate of (w, r)-cover-free codes, Probl. Inf. Transm. 39, No. 4, 317-323, 2003.

[14] V.S. Lebedev, Separating codes and a new combinatorial search model, Probl. Inf. Transm. 46, No. 1, 1-6, 2010.

[15] C.J. Mitchell and F.C. Piper, Key storage in secure networks, Discrete Appl. Math. 21, No. 3, 215-228, 1988.

[16] M. Naor, L.J. Schulman, and A. Srinivasan, Splitters and near-optimal derandomization, Proceedings. of th 36th Annual Symposium on Foundations of Computer Science, 182-191, 1995.

[17] D.R. Stinson, Generalized cover-free families, In honour of Zhu Lie, Discrete Math. 279, No. 1-3, 463-477, 2004.

[18] D.R. Stinson, T. van Trung, and R. Wei, Secure frameproof codes, key distribution patterns, group testing algorithms and related structures, Journal of Statistical Planning and Inference 86, 595-617, 2000.

[19] D.R. Stinson, R. Wei, and L. Zhu, Some new bounds for cover-free families, J. Combin. Theory A. 90, 224-234, 2000.

[20] D.C. Torney, Sets pooling designs, Ann. Combin. 3, 95101, 1999.