Discrete Mathematics, Algorithms and Applications
Vol. 1, No. 2 (2009) 235–251

World Scientific
www.worldscientific.com

ERROR-TOLERANT TRIVIAL TWO-STAGE GROUP TESTING FOR COMPLEXES USING ALMOST SEPARABLE AND ALMOST DISJUNCT MATRICES

WEIWEI LANG

Institute of Mathematics and Information Science
Hebei Normal University, Shijiazhuang
050016, P. R. China

YUEXUAN WANG*

Institute for Theoretical Computer Science
Tsinghua University, Beijing
100084, P. R. China

JAMES YU

Department of Computer Science
University of Texas at Dallas
Richardson, TX 75083, USA
jiyu@utdallas.edu

SUOGANG GAO

Institute of Mathematics and Information Science
Hebei Normal University, Shijiazhuang
050016, P. R. China
sggao@mail.hebtu.edu.cn

WEILI WU†

Department of Computer Science
University of Texas at Dallas
Richardson, TX 75083, USA
weiliwu@utdallas.edu

Accepted 12 February 2009

*This work was supported in part by the National Basic Research Program of China Grants 2007CB807900 and 2007CB807901, the National Natural Science Foundation of China Grant 60604033, and the Hi-Tech Research & Development Program of China Grant 2006AA10Z216.
†The author was supported in part by National Science Foundation under grant CCF 0621829 and 0514796.

In this paper, we define an *α-almost $(k;2e+1)$-separable matrix* and an *α-almost k^e-disjunct matrix*. Using their complements, we devise algorithms for fault-tolerant trivial two-stage group tests (pooling designs) for k-complexes. We derive the expected values for the given algorithms to identify all such positive complexes.

Keywords: Group testing; pooling design; complex; separable matrix; disjunct matrix.

Mathematical Subject Classification 2000: 05B30, 62K05

1. Introduction

Group testing is a well-known technique for identifying positive items efficiently from a given large population of items by conducting tests on subsets of items. An item in the population is said to be either positive or negative. The test outcome for any subset or group of items can also be either positive or negative as is determined in certain ways by its constituent items. The test outcomes for these groups of items are then used to determine which items or items sets are positive.

The classical group testing model consists of a set of n items, d of which are positive. A test on a group of items has negative outcome if all the items within the group are negative and has positive outcome otherwise. The complex group testing model, on the other hand, assumes that the test outcome of a group is positive only when the group completely contains a set of items known as a *positive complex*. We can think of it as if we replace the set of positive items in the first model with a set of positive complexes $D = \{D_1, D_2, \ldots, D_d\}$, where D_i is a positive complex. In complex group testing, if the test outcome is positive it means the group contains at least one positive complex in it. We normally assume that $D_i \not\subseteq D_j$ for $i \neq j$ [1].

Depending on how the tests are specified, group testing algorithms can be *sequential*, *non-adaptive*, or *multi-stage*. In sequential group testing, tests are specified with the knowledge of test outcomes from earlier tests. In non-adaptive testing, all tests are specified in parallel without the knowledge of any other test outcomes. In multi-stage algorithms, however, all the tests within each stage are specified in parallel but different stages are generally specified sequentially with the knowledge of test outcomes from earlier stages.

Due to their wide applications in biology experiments, non-adaptive group tests are now often referred to as *pooling designs*. For a multi-stage algorithm, if there are exactly s stages we often refer to it as an *s-stage algorithm*. For a two-stage algorithm, when each test in the second stage consists of only a single item or a single complex, we refer to it as a *trivial* two-stage algorithm.

It is sometimes unavoidable to have errors in test outcomes. Biology experiments, for example, are known for their unreliabilities [9]. It is therefore important to construct error-tolerant group tests or pooling designs to cope with these errors. Non-adaptive group tests or pooling designs are typically represented by incidence matrices, where columns corresponds items and rows corresponds to tests or pools.

In this paper, we give trivial two-phase pooling designs for complexes for situations where testing errors may exist. We first list some definitions and known results to provide a basis for subsequent discussions. We then provide the details for the error-tolerant matrices and the corresponding error-tolerant pooling design algorithms for complexes.

2. Related Work

Torney provided the first example of complexes on eukaryotic DNA transcription and RNA translation [10]. For error-tolerant models, Kautz and Singleton first presented a way to construct codes that can both detect and correct errors in 1964 [6]. Huang and Weng described a d^e-*disjunct* matrix that can be used to detect e errors and correct $\lfloor e/2 \rfloor$ errors [4, 5]. Later, Dyachkov, Macula, and Vilenkin proved that when applied to two-stage group testing algorithms, d^e-disjunct matrices can be used to correct e errors [2]. Du and Hwang proved that a $(d; z)$-*separable* matrix can be used to correct $\lfloor (z-1)/2 \rfloor$ errors [1].

For complex group testing, Macula, Rykov, and Yekhanin constructed a k-complex pooling design using the complement of an α-*almost* k-*disjunct* matrix and calculated the expected values for detecting all positive k-complexes under error-free conditions [8]. Macula and Popyack constructed a pooling design to detect k_1-complexes for $k_1 \leq k$ and calculated the expected values for detecting all k_1-complexes under error-free testing conditions [7]. None of the above tests, however, were designed to be error-tolerant.

In this paper, we extend the results found in [8] and [7] for error-free test outcomes to cases where there may be errors in test outcomes. We introduce the concept of a $(k; 2e+1)$-*separable matrix* based on an α-*almost* k-*disjunct matrix* and construct an error-tolerant trivial two-stage pooling design by intersecting the complement of such a matrix with a set of random rows. The expected value is given for detecting all positive k-complexes. We further introduce the concept of an α-*almost* k^e-*disjunct matrix* and similarly construct an error-tolerant trivial two-stage pooling design for complexes using its complement intersecting with a set of random rows. The expected value is given for detecting all positive k_1-complexes for $k_1 \leq k$.

3. Preliminaries

First, we fix up some definitions. For further details, the reader can refer to the corresponding references.

Definition 3.1 ([8]). Given a positive integer n, let $[n]$ denote the set $\{1, 2, \ldots, n\}$. A subset of $[n]$ with k elements is called a k-set. Given a set S, let $|S|$ denote the number of elements in S.

Definition 3.2 ([8]). Let $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$ be a set of positive complexes. If $|S_l| = k$, then S_l is called a k-complex. If $|S_l| = k$ for all $1 \leq l \leq d$, Γ a set of positive k-complexes.

Definition 3.3 ([8]). Let A be a $(0,1)$-matrix. The complement of A is the matrix obtained by interchanging the 0s and 1s in A.

Definition 3.4 ([8]). By n-vector, we mean a binary vector with n elements. Let X and Y be two n-vectors

$$X = (x_1, \ldots, x_i, \ldots, x_n)^t \quad \text{and} \quad Y = (y_1, \ldots, y_i, \ldots, y_n)^t.$$

The union or Boolean sum of X and Y is $X \vee Y = (x_1 \vee y_1, \ldots, x_i \vee y_i, \ldots, x_n \vee y_n)^t$, where

$$x_i \vee y_i = \begin{cases} 0, & \text{if } x_i = y_i = 0 \quad \text{for } i = 1, 2, \ldots, n; \\ 1, & \text{otherwise.} \end{cases}$$

The intersecion of X and Y is $X \wedge Y = (x_1 \wedge y_1, \ldots, x_i \wedge y_i, \ldots, x_n \wedge y_n)^t$, where

$$x_i \wedge y_i = \begin{cases} 1, & \text{if } x_i = y_i = 1 \quad \text{for } i = 1, 2, \ldots, n; \\ 0, & \text{otherwise.} \end{cases}$$

Definition 3.5 ([3]). Given a $(0, 1)$-matrix A, if the union of any d columns does not include any other columns, we call A a d-disjunct matrix.

For example,

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

is a 2-disjunct matrix.

Definition 3.6 ([5]). Given a $(0, 1)$-matrix A, for for any $d + 1$ columns C_0, $C_1, \ldots, C_d$ of A, if there are at least $e + 1$ 1s in C_0 but not in $\bigcup_{i=1}^{d} C_i$, we call A a d^e-disjunct matrix.

For example,

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

is a 2^1-disjunct matrix.

Remark 3.7 ([5]). A d^e-disjunct matrix must also be a $d_1^{e_1}$-disjunct matrix for all $d_1 \leq d$ and $e_1 \leq e$. For instance, the matrix in the above example is also a 1^1-disjunct matrix.

Definition 3.8 ([11]). Given two n-vectors X and Y where

$$X = (x_1, \ldots, x_i, \ldots, x_n)^t \quad \text{and} \quad Y = (y_1, \ldots, y_i, \ldots, y_n)^t,$$

the number of different elements $|\{i|x_i \neq y_i\}|$, denoted by $H(X,Y)$, is called the Hamming distance between X and Y.

Definition 3.9 ([1]). Given a $(0,1)$-matrix A, if the Hamming distance between any two d-column unions is at least z, we call A a $(d;z)$-separable matrix.

For example,

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is a $(3;2)$-separable matrix.

Definition 3.10 ([5]). For any particular column in a (0, 1)-matrix A, we consider the row indices of all 1 entries as a set. For any two column vectors

$$X = (x_1, \ldots, x_i, \ldots, x_n)^t \quad \text{and} \quad Y = (y_1, \ldots, y_i, \ldots, y_n)^t,$$

we use $|X - Y|$ to denote the number of 1s found in X but not Y. If the set of indices of all 1 entries in X is included in the set of similar indices in Y, we denote it by $X \subseteq Y$.

Definition 3.11 ([8]). Let p be a real number with $0 < p < 1$. Let r_i be a random row vector on (0,1) with t elements, where each element in r_i is 1 with a probability p. Given an $n \times t$ $(0,1)$-matrix Ω, define a $(m+n) \times t$ matrix $\Omega(m,p,t)$ by adding to Ω m random row vectors r_i with $1 \leq i \leq m$. We use w_j, where $1 \leq j \leq n$, to denote the jth row of Ω and use $u_1(j), \ldots, u_v(j), \ldots, u_t(j)$, where $1 \leq j \leq n$, to denote the columns of Ω. We use $u_1(i), \ldots, u_v(i), \ldots, u_t(i)$, where $1 \leq i \leq n+m$, to denote the columns of $\Omega(m,p,t)$.

Definition 3.12 ([8]). Given an $n \times t$ $(0,1)$-matrix Ω, define an $mn \times t$ $(0,1)$-matrix $\Omega^*(m,p,t)$ whose rows are the coordinate-wise intersections of random row r_i with row w_j from $\Omega(m,p,t)$, where $1 \leq i \leq m$ and $1 \leq j \leq n$. The rows $r_i \wedge w_j$ are ordered lexigraphically according to (i,j). We use $u_1(i,j), \ldots, u_v(i,j), \ldots, u_t(i,j)$ to denote its column vectors with $1 \leq i \leq m$ and $1 \leq j \leq n$.

Definition 3.13 ([7]). Suppose $S_1, \ldots, S_l, \ldots, S_d$ is a collection of sets of columns from Ω with $|S_l| \leq k$ for all $1 \leq l \leq d$. For any $l_0 \in [d]$, we say that the random part of $\Omega(m,p,t)$ separates S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$ with $l_0 \neq l$ if there exists a random row r_i in $\Omega(m,p,t)$ with $i \in [m]$ such that every column of S_{l_0} in $\Omega(m,p,t)$ has a 1 in row r_i but for each S_l with $l_0 \neq l$ there is at least one column of S_l in $\Omega(m,p,t)$ with a 0 in row r_i.

Definition 3.14 ([8]). Suppose A is an $n \times t$ $(0,1)$-matrix. Let $\{a_v(i)|i = 1, 2, \ldots, n; v = 1, 2, \ldots, t\}$ be the column vectors of A. Define E as the event that for any k-set of columns $\{a_{v_s}(i)\}_{s=1}^k$ we have $a_v(i) \leq \bigvee_{s=1}^k a_{v_s}(i)$" for all $a_v(i) \notin \{a_{v_s}(i)\}_{s=1}^k$. Let $0 < \alpha \leq 1$ be a real number and assume uniform probability distribution when choosing the k-set columns from A. We call A an α-almost k-disjunct matrix if the probability of E occurring satisfies the condition $prob(E) \leq 1 - \alpha$.

Definition 3.15 ([8]). Let $o(i,j)$, where $1 \leq i \leq m$ and $1 \leq j \leq n$, denote the test outcome vectors for when the test outcomes are error-free. For fixed i, let $o_i(j)$ denote the sub vector of $o(i,j)$.

Next, we list some of the known results from the literature to be used in later discussions.

Theorem 3.16 ([2]). *Let $1 \leq d < t$ be an integer. And let M be an $n \times t$ d^e-disjunct matrix. If S and T are column subsets of M with cardinalities of at most d, then*

(1) *If $S \subset T$, then $H(\vee S, \vee T) \geq e + 1$;*
(2) *If $S \not\subset T$ and $T \not\subset S$, then $H(\vee S, \vee T) \geq 2e + 2$.*

Theorem 3.17 ([8]). *Consider the set of columns from $\Omega^*(m,p,t)$: $u_1(i,j), \ldots, u_v(i,j), \ldots, u_t(i,j)$, where $1 \leq i \leq m$ and $1 \leq j \leq n$. Suppose there is a set of d positive k-complexes $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$, where $S_l = \{u_{v_s}(i,j)\}_{s=1}^k$. Each row of $\Omega^*(m,p,t)$, $r_i \wedge w_j$, forms one pool. Then $u_v(i,j)$ is in the test as determined by $r_i \wedge w_j$ if and only if the element at row $r_i \wedge w_j$ and column $u_v(i,j)$ is 1. Pool $r_i \wedge w_j$ is positive if and only if this pool contains a certain k-complex S_l, i.e. both r_i and w_j include S_l in $\Omega(m,p,t)$.*

Theorem 3.18 ([8]). *Let $i \in [m]$ and $l_0 \in [d]$. If there exists a row r_i in $\Omega(m,p,t)$ that separates S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$, then $\wedge S_{l_0} = o_i(j)$ in Ω.*

Theorem 3.19 ([8]). *Let $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$ be a collection of k-complexes of columns from $\Omega^*(m,p,t)$. Suppose $S_{l_0} \in \Gamma$. Define $\Phi(l_0,d,p)$ to be the probability that the random part of $\Omega(m,p,t)$ separates S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$. Then*

$$\Phi(l_0,d,p) \geq 1 - \left(1 - p^k \left(\sum_{y=1}^{k} \binom{t-k}{y} \binom{k}{k-y} \binom{t}{k}^{-1} (1-p^y)\right)^{d-1}\right)^m. \quad (3.1)$$

Theorem 3.20 ([7]). *Suppose $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$ is the set of positive complexes with $|S_l| \leq k$ for $1 \leq l \leq d$. Let $S_{l_0} \in \Gamma$ and $|S_{l_0}| = k_{l_0}$ for a fixed $l_0 \in d$. Set $h(l, l_0, \Gamma) = |S_l \backslash S_{l_0}|$. When l_0 and Γ are both fixed, we use $h(l)$ to replace $h(l, l_0, \Gamma)$. Define $\Phi(l_0,d,p)$ to be the probability that the random part of $\Omega(m,p,t)$ separates*

S_{l_0} *from* $S_1, \ldots, S_l, \ldots, S_d$. *Then*

$$\Phi(l_0, d, p) \geq 1 - \left(1 - p^{k_{l_0}} \prod_{l=1, l \neq l_0}^{d} (1 - p^{h(l)})\right)^m, \tag{3.2}$$

where $h(l) = |S_l \backslash S_{l_0}|$.

4. An α-Almost $(k; 2e+1)$-Separable Matrix

To construct error-tolerant pooling designs, we define an "α-almost $(k; 2e+1)$-separable matrix" as follows.

Definition 4.1. Suppose A is an $n \times t$ $(0,1)$-matrix. Let $\{a_v(i) \,|\, i = 1, 2, \ldots, n;\ v = 1, 2, \ldots, t\}$ denote all the columns of A. Define E as the event that given any two k-sets of columns from A, the Hamming distance between their respective column unions is at least $2e + 1$. Given that the k-sets of columns are chosen from A with uniform probability distribution, if the probability of event E satisfies the condition $prob(E) \geq \alpha$, where $0 < \alpha \leq 1$, we call A an α-almost $(k; 2e+1)$-separable matrix.

Property 4.2. *Suppose* Ω *is the complement of an* α*-almost* $(k; 2e+1)$*-separable matrix. Let* $\{u_v(i) \,|\, i = 1, 2, \ldots, n;\ v = 1, 2, \ldots, t\}$ *be the columns of* Ω. *Then for any two* k*-sets of columns in* Ω *the probability for the Hamming distance between their respective column intersections being* $\geq 2e + 1$ *is at least* α.

Proof. This follows from Definitions 3.3 and 4.1. □

Next, we construct a pooling design similar to Theorem 3.17. Let $o(i, j)$, where $1 \leq i \leq m$ and $1 \leq j \leq n$, be the error-free test outcomes. For a fixed i, we denote the sub-vector $o(i, j)$ by $o_i(j)$. Let $p(i, j)$, where $1 \leq i \leq m$ and $1 \leq j \leq n$, be the test outcomes with errors. For a fixed i, we denote the sub-vector of $p(i, j)$ by $p_i(j)$. We use the columns $u_v(j)$ of Ω to identify the target items in subsequent discussions.

Algorithm 4.3. Let $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$ be the collection of positive k-complexes. Suppose at most e errors can be found in the outcome of every m tests. Then $H(o_i(j), p_i(j)) \leq e$ for every $i \in [m]$. This trivial two-stage algorithm finds the positive k-complexes.

(1) First Stage

For each i, find the set of k-sets $T_{i_1}, T_{i_2}, \ldots, T_{i_q}$ of columns in Ω, where $1 \leq q \leq \binom{t}{k}$, that satisfies the condition $H(\wedge T_{i_x}, p_i(j)) \leq e$ for each $1 \leq x \leq q$.

(2) Second Stage

For each T_{i_x}, form $\lceil \frac{m}{e} \rceil$ redundant pools each of which consists of just items in T_{i_x}. T_{i_x} is confirmed to be a positive complex if and only if there is at most 1 negative tests.

Proof. First, consider the special case where $\alpha = 1$ and there exists a random row r_i in $\Omega(m, p, t)$ that separates S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$ for $l_0 \in [d]$.

From Theorem 3.18, we know that $\wedge S_{l_0} = o_i(j)$. We claim that in this case the first stage can be used to find the only column k-set T_i of Ω that satisfies the condition $H(\wedge T_i, p_i(j)) \leq e$. Furthermore, we have $T_i = S_{l_0}$ inside Ω.

In fact, since

$$\begin{aligned} H(\wedge S_{l_0}, p_i(j)) &= H(o_i(j), p_i(j)) \\ &\leq e, \end{aligned} \tag{4.1}$$

the k-sets that satisfy the specified condition in the first stage must exist since S_{l_0} is one example. Suppose there are two k-sets T_{i_1} and T_{i_2} satisfying the given condition, i.e.

$$H(\wedge T_{i_1}, p_i(j)) \leq e \tag{4.2}$$

and

$$H(\wedge T_{i_2}, p_i(j)) \leq e. \tag{4.3}$$

Then

$$H(\wedge T_{i_1}, p_i(j)) + H(\wedge T_{i_2}, p_i(j)) \leq 2e. \tag{4.4}$$

On the other hand, we have

$$\begin{aligned} H(\wedge T_{i_1}, p_i(j)) + H(\wedge T_{i_2}, p_i(j)) &\geq H(\wedge T_{i_1}, \wedge T_{i_2}) \\ &\geq 2e + 1, \end{aligned} \tag{4.5}$$

which is a contradiction. So there must exist just one k-set that satisfies the given condition. Let's denote this k-set by T_i,

Suppose that this T_i is not S_{l_0}. Then

$$\begin{aligned} H(\wedge T_i, p_i(j)) + H(o_i(j), p_i(j)) &\geq H(\wedge T_i, o_i(j)) \\ &= H(\wedge T_i, \wedge S_{l_0}) \\ &\geq 2e + 1. \end{aligned} \tag{4.6}$$

But $H(o_i(j), p_i(j)) \leq e$, so we must have $H(\wedge T_i, p_i(j)) \geq e + 1$, which contradicts our choice of T_i. So inside Ω $T_i = S_{l_0}$. Our claim is thus true and the algorithm can find the positive k-complex S_{l_0} in the first stage in this special case.

For the more general case where $0 < \alpha \leq 1$, the above happens only with probability α. On the other hand, when the m rows of $\Omega(m, p, t)$ are random, a row r_i in $\Omega(m, p, t)$ separates S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$ for $l_0 \in [d]$ with only a certain probability (see theorem below for the exact value). As a result, the k-sets found in the first stage are positive k complexes with a certain probability. Therefore, we need to use the second stage to confirm the k-sets and eliminate the negative k-sets. Since there are at most e errors in every m tests, if at most 1 out of $\lceil \frac{m}{e} \rceil$ tests are negative for a tested k-set, we conclude that it is positive. Otherwise, k-set is negative and can be discarded from the result. □

Example 4.4. We use this example to explain the above algorithm. Matrix Ω in Table 1 is the complement of a $(3;3)$-separable matrix. Hence $k = 3$ and $e = 1$. Matrix $\Omega(3, 0.6, 4)$ in Table 2 was obtained by adding 3 random rows to Ω according to Definition 3.11, And matrix $\Omega^*(3, 0.6, 4)$ in Table 3 was constructed according to Definition 3.12.

Assume the target set is $\{u_1(j), u_2(j), u_3(j), u_4(j)\}$ and there is a set of two positive 3-complexes $\Gamma = \{S_1, S_2\}$, where $S_1 = \{u_1(j), u_2(j), u_3(j)\}$ and $S_2 = \{u_1(j), u_3(j), u_4(j)\}$.

The error-free test outcomes would be

$$o(i,j) = (0,0,0,0,0,0,0,0,0,1,1,1,\\ 0,0,0,1,1,1,0,0,0,0,0,0,\\ 0,0,0,0,0,0,0,0,0,0,0,0)^t. \tag{4.7}$$

Table 1. Matrix Ω.

Ω	$u_1(i)$	$u_2(i)$	$u_3(i)$	$u_4(i)$
ω_1	0	1	1	1
ω_2	0	1	1	1
ω_3	0	1	1	1
ω_4	1	0	1	1
ω_5	1	0	1	1
ω_6	1	0	1	1
ω_7	1	1	0	1
ω_8	1	1	0	1
ω_9	1	1	0	1
ω_{10}	1	1	1	0
ω_{11}	1	1	1	0
ω_{12}	1	1	1	0

Table 2. Matrix $\Omega(3, 0.6, 4)$.

$\Omega(3, 0.6, 4)$	$u_1(i)$	$u_2(i)$	$u_3(i)$	$u_4(i)$
ω_1	0	1	1	1
ω_2	0	1	1	1
ω_3	0	1	1	1
ω_4	1	0	1	1
ω_5	1	0	1	1
ω_6	1	0	1	1
ω_7	1	1	0	1
ω_8	1	1	0	1
ω_9	1	1	0	1
ω_{10}	1	1	1	0
ω_{11}	1	1	1	0
ω_{12}	1	1	1	0
r_1	1	1	1	0
r_2	1	0	1	1
r_3	0	1	1	1

Table 3. Matrix $\Omega^*(3, 0.6, 4)$.

$\Omega^*(3, 0.6, 4)$	$u_1(i,j)$	$u_2(i,j)$	$u_3(i,j)$	$u_4(i,j)$
$r_1 \wedge \omega_1$	0	1	1	0
$r_1 \wedge \omega_2$	0	1	1	0
$r_1 \wedge \omega_3$	0	1	1	0
$r_1 \wedge \omega_4$	1	0	1	0
$r_1 \wedge \omega_5$	1	0	1	0
$r_1 \wedge \omega_6$	1	0	1	0
$r_1 \wedge \omega_7$	1	1	0	0
$r_1 \wedge \omega_8$	1	1	0	0
$r_1 \wedge \omega_9$	1	1	0	0
$r_1 \wedge \omega_{10}$	1	1	1	0
$r_1 \wedge \omega_{11}$	1	1	1	0
$r_1 \wedge \omega_{12}$	1	1	1	0
$r_2 \wedge \omega_1$	0	0	1	1
$r_2 \wedge \omega_2$	0	0	1	1
$r_2 \wedge \omega_3$	0	0	1	1
$r_2 \wedge \omega_4$	1	0	1	1
$r_2 \wedge \omega_5$	1	0	1	1
$r_2 \wedge \omega_6$	1	0	1	1
$r_2 \wedge \omega_7$	1	0	0	1
$r_2 \wedge \omega_8$	1	0	0	1
$r_2 \wedge \omega_9$	1	0	0	1
$r_2 \wedge \omega_{10}$	1	0	1	0
$r_2 \wedge \omega_{11}$	1	0	1	0
$r_2 \wedge \omega_{12}$	1	0	1	0
$r_3 \wedge \omega_1$	0	1	1	1
$r_3 \wedge \omega_2$	0	1	1	1
$r_3 \wedge \omega_3$	0	1	1	1
$r_3 \wedge \omega_4$	0	0	1	1
$r_3 \wedge \omega_5$	0	0	1	1
$r_3 \wedge \omega_6$	0	0	1	1
$r_3 \wedge \omega_7$	0	1	0	1
$r_3 \wedge \omega_8$	0	1	0	1
$r_3 \wedge \omega_9$	0	1	0	1
$r_3 \wedge \omega_{10}$	0	1	1	0
$r_3 \wedge \omega_{11}$	0	1	1	0
$r_3 \wedge \omega_{12}$	0	1	1	0

If the actual test outcomes with $e = 1$ are

$$\begin{aligned} p(i,j) = (&1,0,0,0,0,0,0,0,0,1,1,1,\\ &0,0,0,0,1,1,0,0,0,0,0,0,\\ &1,0,0,0,0,0,0,0,0,0,0,0)^t, \end{aligned} \tag{4.8}$$

then from Algorithm 4.3 we know that

$$T_1 = \{u_1(j), u_2(j), u_3(j)\}, \tag{4.9}$$

$$T_2 = \{u_1(j), u_3(j), u_4(j)\}, \tag{4.10}$$

$$T_3 = \emptyset. \tag{4.11}$$

Further tests in stage two would confirm T_1 and T_2 are the two positive 3-complexes.

Corollary 4.5. *Using the notations from Algorithm* 4.3, *let* Ω *be the complement of an* α*-almost* $(k; 2e+1)$*-separable matrix. Suppose row* r_i *of* $\Omega(m, p, t)$*, where* $i \in [m]$*, separates* S_{l_0} *from* $S_1, \ldots, S_l, \ldots, S_d$ *for* $l_0 \in [d]$*. If* $H(o_i(j), p_i(j)) \leq e$*, then probability for* $S_{l_0} = T_i$ *in* Ω *is at least* α*.*

Proof. This can be seen from the proof of 4.3 and Definition 4.1. □

Theorem 4.6. *Suppose* $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$ *is a set of positive* k*-complexes and* Ω *is the complement of an* α*-almost* $(k; 2e+1)$*-separable matrix. If* $H(o_i(j), p_i(j)) \leq e$*, the expected value for detecting all positive complexes using Algorithm* 4.3 *is at least* $\alpha \cdot d \cdot \Phi(l_0, d, p)$*.*

Proof. From Corollary 4.5, if the random part of matrix $\Omega(m, p, t)$ separates S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$ where $l_0 \in [d]$, then the probability to detect S_{l_0} using Algorithm 4.3 is at least α. From Theorem 3.19, we know the probability for the random part in $\Omega(m, p, t)$ to separate S_{l_0} with $l_0 \in [d]$ from $S_1, \ldots, S_l, \ldots, S_d$ is $\Phi(l_0, d, p)$. So the probability to detect S_{l_0} using Algorithm 4.3 is at least $\alpha \cdot \Phi(l_0, d, p)$. But S_{l_0} and the other $d-1$ complexes are in equal positions. Due to the additive property of expected values, we know that the expected value for finding out all positive complexes is at least $\alpha \cdot d \cdot \Phi(l_0, d, p)$. □

5. An α-Almost k^e-Disjunct Matrix

We first define the "α-almost k^e-disjunct matrix." We construct our pooling designs using the complement of such matrix. We discuss two separate cases based on the number of test errors that could occur.

5.1. *The case with at most* $\lfloor \frac{e}{2} \rfloor$ *errors in every* m *tests*

The "α-almost k-disjunct matrix" was given by Definition 3.14. For error-tolerant pooling design, we define the following.

Definition 5.1. Suppose A is an $n \times t$ $(0,1)$-matrix. Let $\{a_v(i) \mid i = 1, 2, \ldots, n;\ v = 1, 2, \ldots, t\}$ denote the columns of A. Define E as the event that for any k columns $\{a_{v_s}(i)\}_{s=1}^{k}$ chosen from the t columns of A and any column $a_v(i)$ of A with $a_v(i) \notin \{a_{v_s}(i)\}_{s=1}^{k}$, there exist at least $(e+1)$ 1s in $a_v(i)$ but not in $\bigvee_{s=1}^{k} a_{v_s}(i)$. Given that the k-set columns from A are chosen with uniform probability distribution, if the probability of event E satisfies the condition $prob(E) \geq \alpha$, then A is an α-almost k^e-disjunct matrix.

Property 5.2. *Suppose* Ω *is the complement of an* α*-almost* k^e*-disjunct matrix. Let* $\{u_v(i) \mid i = 1, 2, \ldots, n;\ v = 1, 2, \ldots, t\}$ *be columns of* Ω*. For any* k *columns from*

Ω $\{u_{v_s}(i)\}_{s=1}^{k}$ and any column $u_v(i)$ of Ω with $u_v(i) \notin \{u_{v_s}(i)\}_{s=1}^{k}$, the probability that there exist at least $(e+1)$ 1s in $\bigwedge_{s=1}^{k} u_{v_s}(i)$ but not in $u_v(i)$ is α.

Proof. This is a direct result of Definitions 3.12 and 5.1. □

We construct the pooling design similar to Theorem 3.17.

Algorithm 5.3. Suppose $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$ is the set of positive k_1-complexes with $k_1 \leq k$. Assume there are at most $\lfloor \frac{e}{2} \rfloor$ errors in every m tests. Then $H(o_i(j), p_i(j)) \leq \lfloor \frac{e}{2} \rfloor$ for all $i \in [m]$. This two-stage algorithm finds the positive k_1-complexes as follows.

(1) First Stage

For each i, find the set of columns in Ω, $T_i = \{u_v(j) : |p_i(j) - u_v(j)| \leq \lfloor \frac{e}{2} \rfloor$, where i is fixed and $u_v(j)$ is a column of Ω.$\}$.

(2) Second Stage

For all the sets T_i with $|T_i| = k_1$, confirm that they are k_1-complexes. For each such confirmation, form $\lceil \frac{m}{\lfloor \frac{e}{2} \rfloor} \rceil$ identical pools consisting of just items in T_i. T_i is confirmed to be positive if and only if there is at most 1 negative test among these pools. It is confirmed to be negative otherwise and can be discard from the result.

Proof. Since Ω is the complement of an α-almost k^e-disjunct matrix, we know that Ω is the complement of an α-almost k_1^e-disjunct matrix for $k_1 \leq k$ from Remark 3.7. Given any set of k_1 columns from Ω $\{u_{v_s}(i)\}_{s=1}^{k_1}$ and any column $u_v(i) \notin \{u_{v_s}(i)\}_{s=1}^{k_1}$ from Ω, the probability for there to be at least $(e+1)$ 1s in $\bigwedge_{s=1}^{k_1} u_{v_s}(i)$ but not in $u_v(i)$ is at least α.

First, consider the special case where $\alpha = 1$ and there exists a random row r_i of $\Omega(m, p, t)$ that separates S_{l_0} with $l_0 \in [d]$ from $S_1, \ldots, S_l, \ldots, S_d$ in Ω.

From Theorem 3.18, we know that $\wedge S_{l_0} = o_i(j)$. We claim that the $T_i = S_{l_0}$ when $H(o_i(j), p_i(j)) \leq \lfloor \frac{e}{2} \rfloor$. It can be shown as follows.

(i) $S_{l_0} \subseteq T_i$.

Since $\wedge S_{l_0} = o_i(j)$, we have $o_i(j) \subseteq u_v(j)$. Consequently,

$$\begin{aligned} |p_i(j) - u_v(j)| &\leq |p_i(j) - o_i(j)| \\ &\leq H(p_i(j), o_i(j)) \\ &\leq \left\lfloor \frac{e}{2} \right\rfloor. \end{aligned} \tag{5.1}$$

Hence, $u_v(j) \in T_i$.

(ii) $T_i \subseteq S_{l_0}$.

Suppose there exists a column $u_v(j)$ in Ω such that $u_v(j) \in T_i$ but $u_v(j) \notin S_{l_0}$. For $\alpha = 1$, Ω becomes the complement of a $k^e - disjunct$ matrix. So

$$| \wedge S_{l_0} - u_v(j)| \geq e + 1. \tag{5.2}$$

Since $H(o_i(j), p_i(j)) \le \lfloor \frac{e}{2} \rfloor$, then

$$\begin{aligned}
|p_i(j) - u_v(j)| &\ge | \wedge S_{l_0} - u_v(j)| - |p_i(j) - \wedge S_{l_0}| \\
&\ge | \wedge S_{l_0} - u_v(j)| - H(o_i(j), p_i(j)) \\
&\ge e + 1 - \left\lfloor \frac{e}{2} \right\rfloor \\
&= \left\lceil \frac{e}{2} \right\rceil + 1 \\
&\ge \left\lfloor \frac{e}{2} \right\rfloor + 1, \qquad (5.3)
\end{aligned}$$

which contradicts the assumption that $u_v(j) \in T_i$. Hence, $T_i \subseteq S_{l_0}$. Our claim is thus true.

This means that if $\alpha = 1$ and there exists a random row r_i in $\Omega(m, p, t)$ that separates S_{l_0}, where $l_0 \in [d]$, from $S_1, \ldots, S_l, \ldots, S_d$, the first stage of Algorithm 5.3 can identify the positive complex S_{l_0}.

Next, consider the general case where $0 < \alpha \le 1$ and the m rows of $\Omega(m, p, t)$ are random. The same argument is true as in Algorithm 4.3. We need to confirm the k_1-sets in the second stage to eliminate the negative k_1-sets. The second stage verifies that T_i is a positive k_1-complex with $k_1 = |T_i|$. The redundant pools help eliminate the effect from test errors. □

We still use the matrices from Example 4.4 to explain.

Suppose that Ω is the complement of a $3^2 - disjunct$ matrix (see Table 1) with $k = 3$ and $e = 2$, the target set is $\{u_1(j), u_2(j), u_3(j), u_4(j)\}$, and there is a set of two positive complexes $\Gamma = \{S_1, S_2\}$, where $S_1 = \{u_1(j), u_2(j)\}$ and $S_2 = \{u_2(j), u_3(j), u_4(j)\}$.

The error-free test outcome would be

$$\begin{aligned}
o(i,j) = (&0,0,0,0,0,0,1,1,1,1,1,1, \\
&0,0,0,0,0,0,0,0,0,0,0,0, \\
&1,1,1,0,0,0,0,0,0,0,0,0)^t. \qquad (5.4)
\end{aligned}$$

If the actual test result with $e = 3$ is

$$\begin{aligned}
p(i,j) = (&1,0,0,0,0,0,1,1,1,1,1,1, \\
&1,0,0,0,0,0,0,0,0,0,0,0, \\
&0,1,1,0,0,0,0,0,0,0,0,0)^t, \qquad (5.5)
\end{aligned}$$

then from Algorithm 5.3 we know that

$$T_1 = \left\{ u_v(j) : |p_1(j) - u_v(j)| \le \left\lfloor \frac{e}{2} \right\rfloor \right\} = \{u_1(j), u_2(j)\}, \qquad (5.6)$$

$$T_2 = \left\{ u_v(j) : |p_2(j) - u_v(j)| \le \left\lfloor \frac{e}{2} \right\rfloor \right\} = \{u_1(j), u_2(j), u_3(j), u_4(j)\}, \qquad (5.7)$$

$$T_3 = \left\{ u_v(j) : |p_3(j) - u_v(j)| \le \left\lfloor \frac{e}{2} \right\rfloor \right\} = \{u_2(j), u_3(j), u_4(j)\}. \qquad (5.8)$$

Because $|T_1| \leq 3$ and $|T_3| \leq 3$, by testing against T_1 and T_3 we would know that T_1 and T_3 are positive complexes.

Corollary 5.4. *Using the notation from Algorithm 5.3, let Ω be the complement of an α-almost k^e-disjunct matrix. Suppose row r_i in $\Omega(m,p,t)$ separates S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$, where $l_0 \in [d]$ and $|S_l| \leq k$. If $H(o_i(j), p_i(j)) \leq \lfloor \frac{e}{2} \rfloor$, then the probability that $S_{l_0} = T_i$ in Ω is at least α.*

Proof. This can been seen from the proof of Algorithm 5.3. □

Theorem 5.5. *Suppose $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$ is the set of k_1-complexes and Ω is the complement of an α-almost k^e-disjunct matrix. If $H(o_i(j), p_i(j)) \leq \lfloor \frac{e}{2} \rfloor$, then the expected value for finding out all the positive complexes using Algorithm 5.3 is at least $\alpha \cdot \sum_{l=1}^{d} \Phi(l,d,p)$.*

Proof. From Corollary 5.4, we know that if the random part of $\Omega(m,p,t)$ separates S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$ with $l_0 \in [d]$, then the probability of finding out S_{l_0} using Algorithm 5.3 is at least α. From Theorem 3.20 we know that the probability for the random part of $\Omega(m,p,t)$ to separate S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$ with $l_0 \in [d]$ is $\Phi(l_0, d, p)$. Thus the probability to identify S_{l_0} using Algorithm 5.3 is at least $\alpha \cdot \Phi(l_0, d, p)$. The probability to identify a positive k_1-complex S_l thus is at least $\alpha \cdot \Phi(l, d, p)$. From the additive property of expected values, we have that the expected value to find all the positive complexes is at least $\alpha \cdot \sum_{l=1}^{d} \Phi(l,d,p)$. □

5.2. *The case with at most e errors in every m tests*

Lemma 5.6. *Suppose $1 \leq d < t$ and M is the complement of an $n \times t$ d^e-disjunct matrix. Let S and T be two different column subsets of M, each with a cardinality no greater than d. Then*

(1) *If $S \subset T$, then $H(\wedge S, \wedge T) \geq e + 1$.*
(2) *If $S \not\subset T$ and $T \not\subset S$, then $H(\wedge S, \wedge T) \geq 2e + 2$.*

Proof. This is a direct result of Definition 3.3 and Theorem 3.16. □

When there are no more than e errors in m tests, and $\alpha = 1$, we have the following Theorem and Algorithm.

Theorem 5.7. *Suppose $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$ is the set of positive k_1-complexes with $k_1 \leq k$. For each $i \in [m]$, assume $H(o_i(j), p_i(j)) \leq e$. Consider set of columns in Ω,*

$$T_i = \left\{ u_v(j) : |p_i(j) - u_v(j)| \leq \left\lfloor \frac{e}{2} \right\rfloor \right\}, \tag{5.9}$$

where i is fixed and $u_v(j)$ is a column of Ω. If there exists a random row r_i in $\Omega(m,p,t)$ that separates S_{l_0} from $S_1, \ldots, S_l, \ldots, S_d$ in Ω for some $l_0 \in [d]$, then for

$|T_i| \le k$ we have

$$o_i(j) = p_i(j) \Leftrightarrow \wedge T_i = p_i(j), \tag{5.10}$$

which means we can test for e errors.

Proof. From the proof of Algorithm 5.3, we know $\wedge S_{l_0} = o_i(j)$.

($\Rightarrow$) Assume that $o_i(j) = p_i(j)$. Then

$$H(o_i(j), p_i(j)) = 0 \le \left\lfloor \frac{e}{2} \right\rfloor. \tag{5.11}$$

From the proof of Algorithm 5.3, we know that $T_i = S_{l_0}$. Hence

$$\begin{aligned} \wedge T_i &= \wedge S_{l_0} \\ &= o_i(j) \\ &= p_i(j). \end{aligned} \tag{5.12}$$

($\Leftarrow$) Assume that $\wedge T_i = p_i(j)$. If $S_{l_0} = T_i$, then obviously $o_i(j) = p_i(j)$. If $S_{l_0} \neq T_i$, then by the proof of Algorithm 5.3 we know

$$H(\wedge o_i(j), p_i(j)) > \left\lfloor \frac{e}{2} \right\rfloor, \tag{5.13}$$

i.e.

$$o_i(j) \neq p_i(j). \tag{5.14}$$

By triangle inequality, Lemma 5.6, and the fact that $H(o_i(j), p_i(j)) \le e$ we get

$$\begin{aligned} H(\wedge T_i^*, p_i(j)) &\ge H(\wedge T_i^*, \wedge S_{l_0}) - H(\wedge S_{l_0}, p_i(j)) \\ &= H(\wedge T_i^*, \wedge S_{l_0}) - H(o_i(j), p_i(j)) \\ &\ge e + 1 - e \\ &= 1, \end{aligned} \tag{5.15}$$

which contradicts the assumption that $\wedge T_i^* = p_i(j)$. Therefore, we have $o_i(j) = p_i(j)$. □

Algorithm 5.8. Suppose $\Gamma = \{S_1, \ldots, S_l, \ldots, S_d\}$ is the set of positive k_1 complexes with $k_1 \le k$. Suppose further that $H(o_i(j), p_i(j)) \le e$ for each $i \in [m]$. This two-stage algorithm can find positive k_1-complexes.

(1) First Stage

Find the set of k_1-sets $T_{i_1}, T_{i_2}, \ldots, T_{i_q}$ where $1 \le q \le \binom{t}{k_1}$ from Ω's columns that satisfy the condition $H(\wedge T_{i_x}, p_i(j)) \le e$ for $1 \le x \le q$.

(2) Second Stage

For all the sets T_i with $|T_i| = k_1$, confirm that they are k_1-complexes. For each such confirmation on T_i, form $\lceil \frac{m}{e} \rceil$ identical pools using items in T_i. The verification is positive if and only if there is at most 1 negative tests. It is negative otherwise and T_i can be discard from the result.

Proof. If there exist a random row r_i from $\Omega(m,p,t)$ that separates S_{l_0} from $S_1,\ldots,S_l,\ldots,S_d$ for $l_0 \in [d]$, then $\wedge S_{l_0} = o_i(j)$ from the proof of Algorithm 5.3.

Since

$$H(\wedge S_{l_0}, p_i(j)) = H(o_i(j), p_i(j)) \le e, \tag{5.16}$$

there must exist a subset that satisfies the condition specified in the first stage as S_{l_0} is one example. Suppose that T_{i_1} and T_{i_2} are two subsets that satisfy the condition, namely,

$$H(\wedge T_{i_1}(j), p_i(j)) \le e \tag{5.17}$$

and

$$H(\wedge T_{i_2}(j), p_i(j)) \le e. \tag{5.18}$$

We then have either $T_{i_1} \subset T_{i_2}$ or $T_{i_2} \subset T_{i_1}$. Otherwise, from Lemma 5.6, we have

$$H(\wedge T_{i_1}, \wedge T_{i_2}) \ge 2e+2. \tag{5.19}$$

Furthermore, since

$$\begin{aligned} H(\wedge T_{i_1}(j), p_i(j)) + H(\wedge T_{i_2}(j), p_i(j)) &\ge H(\wedge T_{i_1}, \wedge T_{i_2}) \\ &\ge 2e+2, \end{aligned} \tag{5.20}$$

then either $H(\wedge T_{i_1}, p_i(j)) > e$ or $H(\wedge T_{i_2}, p_i(j)) > e$, which contradicts the assumption that T_{i_1} and T_{i_2} both satisfy the given condition.

Therefore, the sets that satisfy the condition form a chain, where S_{l_0} is a part of it. We then just need to find the corresponding chain of sets in Ω and test the sets one by one. This allows us to find S_{l_0}. Since the m rows of $\Omega(m,p,t)$ are random, we must further test the sets to verify if they are the sets that satisfy the given condition. □

Again, we use the matrices from Example 4.4 to illustrate.

Suppose Ω is the complement of a $3^2 - \mathit{disjunct}$ matrix (see Table 2) with $k = 3$ and $e = 2$. Assume the target set is $\{u_1(j), u_2(j), u_3(j), u_4(j)\}$ and there is a set of two positive complexes $\Gamma = \{S_1, S_2\}$ with $S_1 = \{u_1(j), u_2(j)\}$ and $S_2 = \{u_2(j), u_3(j), u_4(j)\}$.

The error-free test result would be

$$\begin{aligned} o(i,j) = (&0,0,0,0,0,0,1,1,1,1,1,1,\\ &0,0,0,0,0,0,0,0,0,0,0,0,\\ &1,1,1,0,0,0,0,0,0,0,0,0)^t. \end{aligned} \tag{5.21}$$

If the test result is

$$\begin{aligned} p(i,j) = (&0,0,0,0,0,0,0,0,1,1,1,1,\\ &0,0,0,0,0,0,0,0,0,0,1,1,\\ &1,1,1,1,1,0,0,0,0,0,0,0)^t, \end{aligned} \tag{5.22}$$

then from Algorithm 5.8, we have

$$
\begin{aligned}
&T^*_{1_1} = \{u_1(j), u_2(j)\}, T^*_{1_2} = \{u_1(j), u_2(j), u_3(j)\},\\
&T^*_{2_1} = \{u_1(j), u_2(j), u_3(j)\},\\
&T^*_{3_1} = \{u_3(j), u_4(j)\}, T^*_{3_2} = \{u_2(j), u_3(j), u_4(j)\}.
\end{aligned} \tag{5.23}
$$

Upon further testing of these sets, we can identify that $T^*_{1_1}, T^*_{3_2}$ are the positive complexes.

References

[1] D.-Z. Du and F. K. Hwang, *Pooling Designs and Nonadaptive Group Testing*, Vol. 18, (World Scientific, Singapore, 2006).

[2] A. G. D'Yachkov, A. J. Macula and P. A. Villenkin, Nonadaptive and trivial two-stage group testing with errror-correcting d^e-disjunct inclusion matrices, in *Entropy, Search, Complexity*, (Springer, Berlin, Heidelberg, 2007), pp. 71–83.

[3] A. G. D'Yachkov, A. J. Macula, D. C. Torney and P. A. Villenkin, Two models of nonadaptive group testing for designing screening experiments, in *Proc. 6th Int. Workshop on Model-Oriented Design and Analysis*, eds. A. Atkinson, P. Hackl and W. Muller, (Physica-Verlag, 2001), pp. 63–75.

[4] T. Huang and C.-W. Weng, A note on decoding of superimposed codes, *J. Comb. Optim.* **7** (2003) 381–384.

[5] T. Huang and C.-W. Weng, Pooling spaces and non-adaptive pooling designs, *Discrete Math.* **282** (2004) 163–169.

[6] W. H. Kautz and R. C. Singleton, Nonrandom binary superimposed codes, *IEEE Trans. Info. Theor.* **10** (1964) 363–377.

[7] A. J. Macula and L. J. Popyack, A group testing method for finding patterns in data, *Discrete Appl. Math.* **144** (2004) 149–157.

[8] A. J. Macula, V. V. Rykov and S. Yekhanin, Trivial two-stage group testing for complexes using almost disjunct matrices, *Discrete Appl. Math.* **137** (2004) 97–107.

[9] H. Q. Ngo and D.-Z. Du, New constructions of non-adaptive and error-tolerance pooling designs, *Discrete Math.* **243** (2002) 161–170.

[10] D. C. Torney, Sets Pooling Designs, *Ann. Comb.* **3** (1999) 95–101.

[11] W. Wu, Y. Huang, X. Huang and Y. Li, On error-tolerant DNA screening, *Discrete Appl. Math.* **154** (2006) 1753–1758.