Discrete Applied Mathematics 200 (2016) 18–22

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

# Improvements on some error-tolerance pooling designs[☆]

Huilan Chang[*], Yi-Tsz Tsai

*Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung 811, Taiwan, ROC*

ARTICLE INFO

*Article history:*
Received 26 July 2014
Received in revised form 18 May 2015
Accepted 6 July 2015
Available online 29 July 2015

*Keywords:*
Group testing
Pooling design
Error-tolerance
Johnson graph

ABSTRACT

Pooling designs are fundamental tools in many applications such as biotechnology and network security. Many famous pooling designs have been constructed from mathematical structures by "containing relations". Recently, pooling designs constructed by "intersecting relation" have been proposed by Guo and Wang (2011). Constructing by intersecting relation provides much better error-tolerance capabilities. We study the error-tolerance capabilities of pooling designs constructed by intersecting relation from combinatorial structures proposed by D'yachkov et al. (2007) and Lv et al. (2014).



## 1. Introduction

Group testing was originally introduced by Dorfman [3] as a potential approach to economically screen soldiers with syphilis during World War II and the combinatorial group testing was first studied by Li [12]. There are many applications and variations of group testing such as blood testing, chemical leak testing, codes and DNA screening. We refer to the general references [4,5] by Du and Hwang. The idea of group testing is to group samples and then test the content of the groups. In classical group testing, there are $n$ items, each of which is either positive or negative and among them are up to $d$ positive ones. A *group test*, also called a *pool*, is a subset of items and it yields a positive outcome if and only if it contains at least one positive item.

Two types of group testing algorithms are often investigated: A *sequential algorithm* conducts tests one by one and the outcomes of previous tests are known at the time of preparing the current test. In a *nonadaptive algorithm*, also called a *pooling design*, all tests are specified beforehand and are conducted simultaneously. A renewed interest of group testing has developed due to its applications in computational molecular biology (Balding et al. [2]; Farach et al. [8]). Since biological experiments are more time-consuming, non-adaptive algorithms are more applicable. A nonadaptive algorithm is usually represented by a binary matrix with columns indexed by items and rows indexed by tests, where a cell $(i, j)$ contains a 1-entry if and only if the $i$th test contains the $j$th item.

A binary matrix $M$ is called *$s$-disjunct* if for any $s+1$ columns of $M$ with one designated, there is a row intersecting the designated column and none of the other $s$ columns. Disjunct matrices have a pivotal role in designing nonadaptive group testing algorithms [5, Chapter 2].

Biological experiments are known for producing erroneous outcomes. Therefore, it would be beneficial for pooling designs to have error-tolerance capability. A binary matrix $M$ is called *$s^e$-disjunct* if given any $s+1$ columns of $M$ with one designated, there are $e+1$ rows intersecting the designated column and none of the other $s$ columns. An $s^0$-disjunct matrix is simply $s$-disjunct. An $s^e$-disjunct matrix is called *fully* $s^e$-disjunct if it is not $s_1^{e_1}$-disjunct whenever $s_1 > s$ or $e_1 > e$.

☆ Partially supported by National Science Council, Taiwan under grant NSC 102-2115-M-390-004.
* Corresponding author.
*E-mail addresses:* huilan0102@gmail.com (H. Chang), gfedc0606@gmail.com (Y.-T. Tsai).

http://dx.doi.org/10.1016/j.dam.2015.07.003
0166-218X/

An $s^e$-disjunct matrix is $\lfloor e/2 \rfloor$-error-correcting [6]. Many famous pooling designs have been constructed from mathematical structures by "containing relation" (see [1,6,7,11] and [14]). Subsequently, some improvements on error-tolerance capability have been proposed by considering the "intersecting relation" (see [9–11,13] and [15]).

We study the error-tolerance capabilities of pooling designs constructed by intersecting relation from combinatorial structures proposed by D'yachkov et al. [7] and Lv et al. [13].

## 2. Pooling designs constructed from vectors

D'yachkov et al. [6] discussed three types of inclusion matrices (subset, subspace and vector) and exhibited their disjunct properties and error-tolerance capabilities. One of these types of inclusion matrices is constructed by vector as follows: A *q-ary vector* is a vector whose entries are from $\{0, 1, 2, \ldots, q\}$. The *weight* of a vector $\alpha$ is the number of its non-zero entries and is denoted by $|\alpha|$. Let $[n] := \{1, 2, \ldots, n\}$. For $k \le n$, let $\binom{[n]}{k}[q]^k$ denote the set of all $q$-ary vectors of length $n$ and weight $k$. For $\alpha \in \binom{[n]}{d}[q]^d$ and $\beta \in \binom{[n]}{k}[q]^k$, let $\alpha(i)$ and $\beta(i)$ denote the $i$th entry of $\alpha$ and $\beta$, respectively. D'yachkov et al. [6] defined that $\alpha \prec \beta$ if for $1 \le i \le n$, $\alpha(i) = \beta(i)$ whenever $\alpha(i) \ne 0$. For example, $(1, 2, 0, 0) \prec (1, 2, 0, 1)$ and $(1, 2, 0, 0) \not\prec (1, 3, 0, 1)$.

**Definition 1.** Let $1 \le d \le k \le n, q \ge 1$, and $\pi_q(d, k, n)$ (or $\pi$ in shorthand) be the binary matrix by taking all members of $\binom{[n]}{k}[q]^k$ and all members of $\binom{[n]}{d}[q]^d$ as columns and rows, respectively. For $\alpha \in \binom{[n]}{d}[q]^d$ and $\beta \in \binom{[n]}{k}[q]^k$, $\pi(\alpha, \beta) = 1$ if and only if $\alpha \prec \beta$.

The error-tolerance capability of $\pi_q(d, k, n)$ has been studied by D'yachkov et al. [7].

**Theorem 2** *([7, Proposition 2]). Let $1 \le s \le d$. Then $\pi_q(d, k, n)$ is fully $s^e$-disjunct where $e = \binom{k-s}{d-s} - 1$.*

For $\alpha \in \binom{[n]}{d}[q]^d$ and $\beta \in \binom{[n]}{k}[q]^k$, let $\alpha \cap \beta := \gamma$, where $\gamma$ is a $q$-ary vector of length $n$ with $\gamma(i) = \alpha(i)$ if $\alpha(i) = \beta(i) \ne 0$ and $\gamma(i) = 0$ for otherwise. For example, $(0, 1, 0, 2) \cap (0, 1, 2, 3) = (0, 1, 0, 0)$. Note that $|\alpha \cap \beta|$ is the number of $i$ such that $\alpha(i) = \beta(i) \ne 0$. We study the error-tolerance capabilities of matrices defined by intersecting relation as follows.

**Definition 3.** Let $1 \le r \le d \le k \le n$ and $q \ge 1$. Let $\pi_q(r; d, k, n)$ be the binary matrix by taking all members of $\binom{[n]}{k}[q]^k$ and all members of $\binom{[n]}{d}[q]^d$ as columns and rows, respectively. For $\alpha \in \binom{[n]}{d}[q]^d$ and $\beta \in \binom{[n]}{k}[q]^k$, $\pi(\alpha, \beta) = 1$ if and only if $|\alpha \cap \beta| = r$.

Note that $\pi_q(d, k, n)$ and $\pi_q(r; d, k, n)$ have the same number of rows and columns, and $\pi_q(d; d, k, n) = \pi_q(d, k, n)$. We have the following basic property of such matrices.

**Lemma 4.** *The row weight of $\pi_q(r; d, k, n)$ is*

$$\binom{d}{r}\sum_{j=0}^{d-r}\binom{d-r}{j}\binom{n-d}{k-r-j}(q-1)^j q^{k-r-j}$$

*and the column weight of $\pi_q(r; d, k, n)$ is*

$$\binom{k}{r}\sum_{j=0}^{d-r}\binom{k-r}{j}\binom{n-k}{d-r-j}(q-1)^j q^{d-r-j}.$$

**Proof.** Let $\alpha$ be any $q$-ary vector in $\binom{[n]}{d}[q]^d$. Then the row weight of $\pi_q(r; d, k, n)$ is the number of all $\beta \in \binom{[n]}{k}[q]^k$ satisfying $|\alpha \cap \beta| = r$. The number of $q$-ary vectors $\gamma$ of weight $r$ satisfying $\gamma \prec \alpha$ is $\binom{d}{r}$. Fix $\gamma$ and let $Y := \{y : \gamma(y) \ne 0\}$. Then the number of all $\beta \in \binom{[n]}{k}[q]^k$ satisfying $\alpha \cap \beta = \gamma$ is $\sum_{j=0}^{d-r}\binom{d-r}{j}(q-1)^j\binom{n-d}{k-r-j}q^{k-r-j}$ by considering $\beta(i) \ne \alpha(i)$ whenever $\beta(i) \ne 0$ and $i \in [n] \setminus Y$. Then the row weight is obtained. Similarly, we have the column weight as stated. □

In the following, we discuss the error-tolerance capability of the matrix $\pi_q(r; d, k, n)$.

**Theorem 5.** *Let $1 \le s \le r$, $\lfloor (d+1)/2 \rfloor \le r \le d \le k$ and $n - k - s(k + d - 2r) \ge d - r$. Then $\pi_q(r; d, k, n)$ is $s^e$-disjunct, where $e = q^{d-r}\binom{k-s}{r-s}\binom{n-k-s(k+d-2r)}{d-r} - 1$.*

**Proof.** Let $\beta_0, \beta_1, \ldots, \beta_s$ be any distinct $s + 1$ columns. Then for each $t \in [s]$, there exists $x_t \in [n]$ such that $\beta_0(x_t) \ne 0$ and either $\beta_t(x_t) = 0$ or $\beta_t(x_t) \ne 0$ and $\beta_0(x_t) \ne \beta_t(x_t)$. Let $X = \{x_1, x_2, \ldots, x_s\}$. Let $N$ be the number of $q$-ary vectors $\gamma$ of weight $r$ satisfying $\gamma(x) = \beta_0(x)$ for all $x \in X$ and $\gamma \prec \beta_0$. Then the number of $q$-ary vectors $\gamma$ of weight $r$ satisfying