J Comb Optim (2009) 18: 51–63
DOI 10.1007/s10878-007-9137-6

Error-correcting pooling designs associated with the dual space of unitary space and ratio efficiency comparison

Geng-sheng Zhang · Xiao-lei Sun · Bo-li Li

Published online: 31 January 2008

Abstract In this paper, we construct a d^z-*disjunct* matrix with subspaces in a dual space of Unitary space $\mathbb{F}_{q^2}^{(n)}$, then give its several properties. As the smaller the ratio efficiency is, the better the pooling design is. We compare the ratio efficiency of this construction with others, such as the ratio efficiency of the construction of set, the general space and the dual space of symplectic space. In addition, we find it smaller under some conditions.

Keywords Group testing · d-disjunct · d^z-disjunct · Ratio efficiency

1 Introduction

Pooling designs have been widely used in many DNA-related applications, such as clone-library screening (including physical mapping), counting sequencing, determination of exon boundaries in eukaryotic genes, detecting gene complex, etc.

The basic problem of group testing is to identify the set of positives (defective) in a large population of items. As it is becoming more standard to use the term positive instead of defective, we shall use the former throughout the paper. We assume some testing mechanism exists which if applied to an arbitrary subset of the population gives a negative outcome if the subset contains no positive and positive outcome

Supported by NSF of the Education Department of Hebei Province (2007127) and NSF of Hebei Normal University (L2004B04).

G. Zhang (✉) · X. Sun
Department of Mathematics, Hebei Normal University, Shijiazhuang, 050016, People's Republic of China
e-mail: gshzhang@hebtu.edu.cn

B. Li
Department of Mathematics, The Branch of Hengshui College, Hengshui, 053000, People's Republic of China

otherwise. Objectives of group testing vary from minimizing the number of tests, limiting number of pools, limiting pool sizes to tolerating a few errors.

Group testing algorithms can roughly be divided into two categories: combinatorial group testing (CGT)and probabilistic group testing (PGT). In CGT, it is often assumed that the number of positives among n items is equal to or at most d for some given positive integer d. In PGT, we fix some probability p of having a positive. Group testing strategies can also be either adaptive or non-adaptive if all tests must be specified without knowing the outcomes of other tests. A Group testing algorithm is error-tolerant if it can detect or correct some errors in test outcomes. A mathematical model of error-tolerance designs is an d^z-disjunct matrix.

A $(0,1)$-matrix is said to be *d-disjunct* if and only if no column is contained in the union of d others. A *d-disjunct matrix* with t rows and n columns corresponds precisely to a pooling design which can identify at most d positives from n items with t tests.

A *d^z-disjunct matrix* is a matrix where given any $d+1$ column $C_0, C_1, \ldots, C_d$, the set $C_0 \backslash \bigcup_{i=1}^{d} C_i$ has at least z elements. A *d^z-disjunct* matrix can detect $z-1$ errors and correct $\lfloor (z-1)/2 \rfloor$ errors. The error-correcting capabilities is doubled by the addition of at most d confirmatory and guaranteed tests. A *d-disjunct* matrix is *d^1-disjunct* (D'yachkov et al. 2007).

A $(0,1)$-matrix has column (row) weight c if every column (row) has exactly c 1's.

Macula proposed a way of constructing *d-disjunct* matrix which uses the containment relation in a structure (Macula 1996). Ngo and Du extended the construction to some geometric structures, such as simplicial complexes, and some graph properties, such as matching (Ngo and Du 2002). Huang and Weng gave a comprehensive treatment of construction of d-disjunct matrices by using pooling spaces, which is a significant and important addition to the general theory (Huang and Weng 2004). Geng-sheng Zhang et al. constructed a *d^z-disjunct* matrix by using the dual space of symplectic space (Zhang et al. 2007).

In this paper, we construct a *d^z-disjunct* matrix with subspaces in a dual space of Unitary Space $\mathbb{F}_{q^2}^{(n)}$. In Sect. 3, we will give its several properties. Given some fixed items, our goal is detecting positive items. For a pooling design, the less the number of tests is, the better the pooling design is. In Sect. 4, we will give a new definition and compare the ratio t/n with others, such as in Macula (1996), Zhang et al. (2007) and D'yachkov et al. (2005). We find it smaller under some conditions.

2 Preliminary

In this section we shall briefly review some concepts of geometry of unitary groups over finite fields (Wan 2002). Let $\mathbb{F}_{q^2}$ is a finite field with q^2 elements, where q is a prime or a prime power. $\mathbb{F}_{q^2}$ has an involutive automorphism, i.e., an automorphism of order 2

$$a \mapsto \overline{a},$$

and the fixed filed of this automorphism is F_q. Let

$$A = (a_{ik})_{1\le i\le m, 1\le k\le n}$$

be an $m \times n$ matrix over $\mathbb{F}_{q^2}$. We use $\overline{A}$ to denote the matrix obtained from A by applying the automorphism $a \mapsto \overline{a}$ to all the mn elements of A, i.e.,

$$\overline{A} = (\overline{a}_{ik})_{1\le i\le m, 1\le k\le n}.$$

An $n \times n$ matrix H over $\mathbb{F}_{q^2}$ is said to be *Hermitian*, if $\overline{H}^t = H$.

We denote

$$H_0 = \begin{pmatrix} 0 & I^{(\nu)} \\ I^{(\nu)} & 0 \end{pmatrix}.$$

$$H_1 = \begin{pmatrix} 0 & I^{(\nu)} & \\ I^{(\nu)} & 0 & \\ & & 1 \end{pmatrix}.$$

In order to discuss these two cases simultaneously, we introduce the notation $n = 2\nu + \delta$ and H_δ, where $\delta = 0$ or 1.

Now let H be an $n \times n$ nonsingular Hermitian matrix over $\mathbb{F}_{q^2}$. An $n \times n$ matrix T over $\mathbb{F}_{q^2}$ is called a unitary matrix with respect to H if

$$TH\overline{T}^t = H.$$

Clearly, $n \times n$ unitary matrices with respect to a nonsingular Hermitian matrix H are nonsingular and they form a group with respect to matrix multiplication, called the unitary group of degree n with respect to H over $\mathbb{F}_{q^2}$ and denoted by $U_n(F_{q^2}, H)$. An *m-dimensional* subspace P is said to be of type (m, r) with respect to H, if $PH\overline{P}^t$ is of rank $2r$. In particular, subspaces of type $(m, 0)$ are called m-dimensional *totally isotropic subspaces* with respect to H. Subspaces of type (m, r) exist in the *n-dimensional* unitary space if and only if $2r \le 2m \le n + r$.

Two vectors x and y of $F_{q^2}^{(n)}$ are said to be *orthogonal* (with respect to H), if xHy^t=0. Let P be an *m-dimensional* subspace of $F_{q^2}^{(n)}$. Denote by $P^\perp$ the set of vectors which are orthogonal to every vector of P, i.e., $P^\perp = \{y \in F_{q^2}^{(n)} \mid yH\overline{x}^t = 0$ for all $x \in P\}$. Obviously, $P^\perp$ is a $(n-m)$-*dimensional* subspace of $F_{q^2}^{(n)}$ and is called the *dual subspace* of P with respect to H. A subspace P is *totally isotropic* if and only if $P \subseteq P^\perp$. The dual subspace of a subspace of type (m, r) is of type $(n-m, n-2m+r)$.

(m, s)-*space* denotes a subspace of type (m, s). Let $0 \le m \le [\frac{n}{2}]$, then the number of *m-dimensional totally isotropic* subspaces in the *n-dimensional* unitary space over F_{q^2} is

$$N(m, 0; n) = \frac{\prod_{i=n-2m+1}^{n}(q^i - (-1)^i)}{\prod_{i=1}^{m}(q^{2i} - 1)}.$$

Specially,

$$N(r-m_0,0;2(\nu-m_0)+\delta)=[r-m_0]_\delta=\frac{\prod_{i=2\nu-2r+1+\delta}^{2\nu-2m_0}(q^i-(-1)^i)}{\prod_{i=1}^{r-m_0}(q^{2i}-1)},$$

$$N(m-m_0,0;2(\nu-m_0)+\delta)=[m-m_0]_\delta=\frac{\prod_{i=2\nu-2m+1+\delta}^{2\nu-2m_0}(q^i-(-1)^i)}{\prod_{i=1}^{m-m_0}(q^{2i}-1)}.$$

$N(m_1,r_1;m,r;n)$ denote the number of (m_1,r_1)-*spaces* contained in a (m,r)-*space* in $\mathbb{F}_{q^2}^{(n)}$. $N'(m_1,r_1;m,r;n)$ denote the number of (m,r)-*spaces* containing a fixed (m_1,r_1)-*space* in $\mathbb{F}_{q^2}^{(n)}$.

Lemma 2.1 *Let*

$$2r\le 2m\le n+r$$

and

$$\max\left\{0,\frac{2m_1-r-r_1}{2}\right\}\le\min\{m-r,m-1-r_1\}.$$

Then

$$N(m_1,r_1;m,r;n)=\sum_{k=\max\{0,[\frac{2m_1-r-r_1+1}{2}]\}}^{\min\{m-r,m_1-r_1\}}q^{r_1(r+r_1-2m_1+2k)+2(m-1-k)(m-r-k)}$$
$$\times\frac{\prod_{i=r+r_1-2m_1+2k+1}^{r}(q^i-(-1)^i)\prod_{i=m-r-k+1}^{m-r}(q^{2i}-1)}{\prod_{i=1}^{r_1}(q^i-(-1)^i)\prod_{i=1}^{m_1-r_1-k}(q^{2i}-1)\prod_{i=1}^{k}(q^{2i}-1)}.$$

Lemma 2.2 *Assume that*

$$\left.\begin{array}{r}2r\le 2m\le n+r,\\ 2r_1\le 2m-1\le n+r_1,\\ 0\le r-r_1\le 2m-2r_1,\end{array}\right\}$$

holds.

$$N'(m_1,r_1;m,r;n)=\sum_{k=\max\{0,[\frac{2m_1-r-r_1+1}{2}]\}}^{\min\{m-r,m_1-r_1\}}q^{(n-2m+r)(r_1+r-2m_1+2k)+2(n-m-k)(m_1-r_1-k)}$$
$$\times\frac{\prod_{i=r+r_1-2m_1+2k+1}^{n-2m_1+r_1}(q^i-(-1)^i)\prod_{i=m_1-r_1-k+1}^{m_1-r_1}(q^{2i}-1)}{\prod_{i=1}^{n-2m+r}(q^i-(-1)^i)\prod_{i=1}^{m-r-k}(q^{2i}-1)\prod_{i=1}^{k}(q^{2i}-1)}.$$

Set $\left[{m-m_0 \atop r-m_0}\right]_\delta$ be the number of $(r,0)$-*spaces* contained in a $(m,0)$-*space* and $\left[{m-m_0 \atop r-m_0}\right]^*_\delta$ be the number of $(m,0)$-*spaces* containing a fixed $(r,0)$-*space* in $\mathbb{F}_{q^2}^{2(\nu-m_0)+\delta}$.

$$\begin{bmatrix} m-m_0 \\ r-m_0 \end{bmatrix}_\delta = \frac{\prod_{i=m-r-1}^{m-m_0}(q^{2i}-1)}{\prod_{i=1}^{r-m_0}(q^{2i}-1)},$$

$$\begin{bmatrix} m-m_0 \\ r-m_0 \end{bmatrix}^*_\delta = \frac{\prod_{i=2\nu-2m+1+\delta}^{2\nu-2r+1}(q^{i}-(-1)^i)}{\prod_{i=1}^{m-r}(q^{2i}-1)}.$$

We adopt the convention that

$$\prod_{i\in M} f(i) = 1,$$

where M denotes the empty set.

3 The construction

Definition 3.1 Select integers $0 \le m_0 < r < m \le \nu$. Assume P_0 is a fixed $(m_0,0)$-*space* of $\mathbb{F}_{q^2}^{(n)}$. Let M be the (0, 1)-matrix by taking all $(m,0)$-*spaces* which are contained in $P_0^\perp$ and contain P_0 as columns and all $(r,0)$-*spaces* which are contained in $P_0^\perp$ and contain P_0 as rows. M has a 1 in row i and column j if and only if i is contained in j.

We can show that M is a d^z-*disjunct* matrix with certain constant weights.

Theorem 3.2 *M is a $[r-m_0]_\delta \times [m-m_0]_\delta$ matrix, whose constant row weight (column) is* $\left[{m-m_0 \atop r-m_0}\right]_\delta$ $\left(\left[{m-m_0 \atop r-m_0}\right]_\delta^*\right)$.

Proof By the transitivity of $U_n(F_{q^2},H)$ on the set of subspaces of the same type we can assume that P_0 and $P_0^\perp$ have matrix representations of the forms

$$P_0 = \begin{pmatrix} I^{(m_0)} & 0 & 0 & 0 \end{pmatrix}, \quad \begin{matrix} m_0 & \nu-m_0 & m_0 & \nu-m_0+\delta \end{matrix}$$

$$P_0^\perp = \begin{pmatrix} I^{(m_0)} & 0 & 0 & 0 \\ 0 & I^{(\nu-m_0)} & 0 & 0 \\ 0 & 0 & 0 & I^{(\nu-m_0+\delta)} \end{pmatrix} \quad \begin{matrix} m_0 & \nu-m_0 & m_0 & \nu-m_0+\delta \end{matrix}$$

Let R be a $(r,0)$-*space* contained in $P_0^\perp$ and containing P_0. Then

$$R = \begin{pmatrix} I^{(m_0)} & 0 & 0 & 0 & 0 \\ 0 & R_1 & 0 & R_2 & R_3 \end{pmatrix} \begin{matrix} m_0 \\ r-m_0 \end{matrix}, \quad \begin{matrix} m_0 & \nu-m_0 & m_0 & \nu-m_0 & \delta \end{matrix}$$

where $R_2\overline{R_1}^t + R_1\overline{R_2}^t + R_3\overline{R_3}^t = 0$.

It can be easily verified that $(R_1 R_2)$ is a $(r-m_0, 0)$-*space* in $\mathbb{F}_q^{2(\nu-m_0)+\delta}$. Hence the number of rows in M is $N(r-m_0, 0; 2(\nu-m_0)+\delta) = [r-m_0]_\delta$.

Let Q be a $(m, 0)$-*space* contained in $P_0^\perp$ and containing P_0. So

$$Q = \begin{pmatrix} I^{(m_0)} & 0 & 0 & 0 & 0 \\ 0 & Q_1 & 0 & Q_2 & Q_3 \end{pmatrix} \begin{matrix} m_0 \\ r-m_0 \end{matrix}, \quad \begin{matrix} m_0 & \nu-m_0 & m_0 & \nu-m_0 & \delta \end{matrix}$$

where $Q_2\overline{Q_1}^t + Q_1\overline{Q_2}^t + Q_3\overline{Q_3}^t = 0$, where $(Q_1 Q_2 Q_3)$ is a $(m-m_0, 0)$-*space* in $\mathbb{F}_{q^2}^{2(\nu-m_0)+\delta}$. Hence the number of columns in M is $N(m-m_0, 0; 2(\nu-m_0)+\delta) = [m-m_0]_\delta$.

The row weight is the number of $(m-m_0, 0)$-*spaces* containing a fixed $(r-m_0, 0)$-*space* in $\mathbb{F}_{q^2}^{2(\nu-m_0)+\delta}$. Hence it is

$$\begin{aligned} N'(r-m_0, 0; m-m_0, 0; 2(\nu-m_0)+\delta) &= \frac{\prod_{i=2\nu-2m+1+\delta}^{2\nu-2r+1}(q^i-(-1)^i)}{\prod_{i=1}^{m-r}(q^{2i}-1)} \\ &= \begin{bmatrix} m-m_0 \\ r-m_0 \end{bmatrix}_\delta^* . \end{aligned}$$

The column weight is the number of $(r, 0)$-*spaces* containing P_0 and contained in a $(m, 0)$-*space*, namely the number of $(R_1 \quad R_2)$ contained in a $(Q_1 \quad Q_2)$. Hence it is

$$N(r-m_0, 0; m-m_0, 0; 2(\nu-m_0)+\delta) = \frac{\prod_{i=m-r-1}^{m-m_0}(q^{2i}-1)}{\prod_{i=1}^{r-m_0}(q^{2i}-1)} = \begin{bmatrix} m-m_0 \\ r-m_0 \end{bmatrix}_\delta . \qquad \square$$

Theorem 3.3 *Suppose $m-r \geq 2$ and set $b = \frac{(q^{2(m-m_0-1)}-1)(q^{2(m-m_0)}-q^{2(m-r-2)})}{(q^{2(m-r-2)}-1)(q^{2(m-m_0-1)}-1)-1}$. Then M is d^z-disjunct for $1 \leq d \leq b$ and*

$$z = \begin{bmatrix} m-m_0 \\ r-m_0 \end{bmatrix}_\delta - d\begin{bmatrix} m-m_0-1 \\ r-m_0 \end{bmatrix}_\delta + (d-1)\begin{bmatrix} m-m_0-2 \\ r-m_0 \end{bmatrix}_\delta .$$

Proof Let $C_0, C_1, \ldots, C_d$ be $d+1$ distinct columns ($(m, 0)$-*spaces* contained in $P_0^\perp$ and containing P_0) of M. There are $\left[\begin{smallmatrix} m-m_0 \\ r-m_0 \end{smallmatrix}\right]_\delta$ $(r, 0)$-*spaces* containing P_0 in C_0. Let $|C_0 \setminus \bigcup_{i=1}^d C_i|$ be the number of $(r, 0)$-*spaces* containing P_0 and contained in C_0 but not contained in C_i $(1 \leq i \leq d)$. Noticing that $|C_0 \setminus \bigcup_{i=1}^d C_i| = |C_0 \setminus \bigcup_{i=1}^d (C_0 \bigcap C_i)|$, to obtain the minimum of $|C_0 \setminus \bigcup_{i=1}^d C_i|$, we may assume that each C_i intersects C_0 at a $(m-1, 0)$-*space* containing P_0. Then each C_i covers $\left[\begin{smallmatrix} m-m_0-1 \\ r-m_0 \end{smallmatrix}\right]_\delta$ $(r, 0)$-*spaces* containing P_0 of C_0. However, the coverage of each pair of C_i and C_j overlaps at a $(m-2, 0)$-*space* containing P_0. Therefore only C_1 covers the full $\left[\begin{smallmatrix} m-m_0-1 \\ r-m_0 \end{smallmatrix}\right]_\delta$ $(r, 0)$-*spaces*, while each of $C_2, \ldots, C_d$ can cover a maximum of $\left(\left[\begin{smallmatrix} m-m_0-1 \\ r-m_0 \end{smallmatrix}\right]_\delta - \left[\begin{smallmatrix} m-m_0-2 \\ r-m_0 \end{smallmatrix}\right]_\delta\right)$ $(r, 0)$-*spaces* containing P_0 and not covered by C_1.

Consequently the number of $(r,0)$-*spaces* containing P_0 and contained in C_0 but not contained in C_i $(1 \le i \le d)$ is at least

$$z = \begin{bmatrix} m-m_0 \\ r-m_0 \end{bmatrix}_\delta - d \begin{bmatrix} m-m_0-1 \\ r-m_0 \end{bmatrix}_\delta + (d-1) \begin{bmatrix} m-m_0-2 \\ r-m_0 \end{bmatrix}_\delta .$$

For M to be d^z-*disjunct*, z must be positive, which implies

$$d < \frac{(q^{2(m-m_0-1)}-1)(q^{2(m-m_0)}-q^{2(m-r-2)})}{(q^{2(m-r-2)}-1)(q^{2(m-m_0-1)}-1)-1} + 1.$$

Set $b = \frac{(q^{2(m-m_0-1)}-1)(q^{2(m-m_0)}-q^{2(m-r-2)})}{(q^{2(m-r-2)}-1)(q^{2(m-m_0-1)}-1)-1}$. Then $1 \le d \le b$. □

Let C_0 be a column of M, namely a $(m,0)$-*space* contained in $P_0^\perp$ and containing P_0, and E be a fixed $(m-2,0)$-*space* containing P_0 and contained in C_0. Let D be a $(m-1,0)$-*space* containing E and contained in C_0. By the transitivity of $U_n(F_{q^2}, H)$ on the set of subspaces of the same type we can assume that

$$G = \begin{pmatrix} I^{(m_0)} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I^{(m-m_0-2)} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & D_1 & D_2 & D_3 & 0 & D_4 & D_5 & D_6 & D_7 \end{pmatrix} \begin{matrix} m_0 \\ m-m_0-2, \\ 1 \end{matrix}$$
$$\qquad\quad \begin{matrix} m_0 & m-m_0-2 & 1 & 1 & \nu-m & m-2 & 1 & 1 & \nu-m & \delta \end{matrix}$$

where $(D_1, D_2, D_3, D_4, D_5, D_6, D_7)$ is $(1,0)$-*space* in $F_{q^2}^{2(\nu-m_0)+\delta}$. Hence the number of $(m-1,0)$-*spaces* between E and C_0 is

$$b_\delta = [(m_0+1)-m_0]_\delta = \frac{\prod_{i=2(\nu-m_0)+\delta-1}^{2(\nu-m_0)+\delta}(q^i-(-1)^i)}{q^2-1}.$$

Corollary 3.4 *Suppose $m-r \ge 2$ and $1 \le d \le b'$, $b' = \min\{b, b_\delta\}$. Then M is not d^{z+1}-disjunct, where b and z are as in Theorem 3.3.*

Proof We will show that the minimum of $|C_0 \setminus \bigcup_{i=1}^d C_i|$ in the proof of Theorem 3.3 can be obtained. For $1 \le d \le b'$, we choose d distinct $(m-1,0)$-*spaces* between E and C_0, say $G_i (1 \le i \le d)$. For each G_i, we choose a $(m,0)$-*space* C_i such that $C_0 \cap C_i = G_i$. Hence each pair of C_i and C_j overlaps at the same $(m-2,0)$-*space* E. □

Corollary 3.5 *Suppose $r = m_0+1, m = r+3 = m_0+4 = \nu$ and $1 \le d \le q^6-1$. Then M is d^z-disjunct, but not d^{z+1}-disjunct, where*

$$z = (q^4-1)(q^6-1)[(q^4+1)(q^2+1)-d].$$

Proof Setting $r = m_0+1$ in the z formula of Theorem 3.3, we obtain

$$z = (q^4-1)(q^6-1)[(q^4+1)(q^2+1)-d].$$

The second statement follows directly from Corollary 3.4. □

In order to explain the pooling design plainly, we give an example to show that we can use less tests to detect more items.

Example 3.6 Choose $q = 3$, $m_0 = 3$, $r = 4$, $m = 7$ and $\nu = 9$. Then M is $728^{5358080}$*-disjunct*.

When $\delta = 0$ it has 11767941640 rows and 91278020175614232720832 columns. That is to say, approximately 10^{10} pools are necessary for identifying 728 positives from 9×10^{22} items. Moreover, 5358080 errors can be detected and 2679039 errors can be corrected.

When $\delta = 1$, it has 66430 rows and 374090246621369806528 columns. That is to say, approximately 66430 pools are necessary for identifying 728 positives from 3×10^{21} items. Moreover, 5358080 errors can be detected and 2679039 errors can be corrected. □

The following theorem tells us how to choose m so that the test to item is minimized.

Theorem 3.7 *For* $0 \le m \le \nu - m_0$, *the sequence* $N(m - m_0, 0; 2(\nu - m_0) + \delta)$ *is unimodal and gets its peak at* $m = [\frac{2\nu+2m_0}{4}]$ *or* $[\frac{2\nu+2m_0-1}{4}]$.

Proof

$$\frac{N(m_2 - m_0, 0; 2(\nu - m_0) + \delta)}{N(m_1 - m_0, 0; 2(\nu - m_0) + \delta)} = \frac{\prod_{i=2\nu-2m_2+1+\delta}^{2\nu-2m_1+\delta}(q^i - (-1)^i)}{\prod_{i=m_1-m_0+1}^{m_2-m_0}(q^{2i} - 1)}.$$

When $\delta = 0$,

$$\frac{\prod_{i=2\nu-2m_2+1}^{2\nu-2m_1}(q^i - (-1)^i)}{\prod_{i=m_1-m_0+1}^{m_2-m_0}(q^{2i} - 1)} = \frac{q^{2\nu-2m_2+1} + 1}{q^{2(m_1-m_0+1)} - 1} \frac{q^{2\nu-2m_2+2} - 1}{q^{2(m_1-m_0+2)} - 1} \cdots \frac{q^{2\nu-2m_1} - 1}{q^{2(m_2-m_0)} - 1}.$$

Note that

$$\frac{q^{2\nu-2m_2+1} + 1}{q^{2(m_1-m_0+1)} - 1} > \frac{q^{2\nu-2m_2+2} - 1}{q^{2(m_1-m_0+2)} - 1} > \cdots > \frac{q^{2\nu-2m_1} - 1}{q^{2(m_2-m_0)} - 1}.$$

If $[\frac{2\nu+2m_0}{4}] \le m_1 \le m_2 \le \nu - m_0$, then

$$\frac{q^{2\nu-2m_2+1} + 1}{q^{2(m_1-m_0+1)} - 1} < 1,$$

i.e.,

$$N(m_2 - m_0, 0; 2(\nu - m_0)) < N(m_1 - m_0, 0; 2(\nu - m_0)).$$

If $0 \le m_1 \le m_2 \le [\frac{2\nu+m_0}{4}]$, then

$$\frac{q^{2\nu-2m_1} - 1}{q^{2(m_2-m_0)} - 1} > 1,$$

i.e.,

$$N(m_2 - m_0, 0; 2(\nu - m_0)) > N(m_1 - m_0, 0; 2(\nu - m_0)).$$

When $\delta = 1$,

$$\frac{\prod_{i=2\nu-2m_2+2}^{2\nu-2m_1+1}(q^i - (-1)^i)}{\prod_{i=m_1-m_0+1}^{m_2-m_0}(q^{2i} - 1)} = \frac{q^{2\nu-2m_2+2} - 1}{q^{2(m_1-m_0+1)} - 1} \frac{q^{2\nu-2m_2+3} + 1}{q^{2(m_1-m_0+2)} - 1} \cdots \frac{q^{2\nu-2m_1+1} + 1}{q^{2(m_2-m_0)} - 1}.$$

Note that

$$\frac{q^{2\nu-2m_2+2} - 1}{q^{2(m_1-m_0+1)} - 1} > \frac{q^{2\nu-2m_2+3} + 1}{q^{2(m_1-m_0+2)} - 1} > \cdots > \frac{q^{2\nu-2m_1+1} + 1}{q^{2(m_2-m_0)} - 1}.$$

If $[\frac{2\nu+2m_0}{4}] \leq m_1 \leq m_2 \leq \nu - m_0$, then

$$\frac{q^{2\nu-2m_2+2} - 1}{q^{2(m_1-m_0+1)} - 1} < 1,$$

i.e.,

$$N(m_2 - m_0, 0; 2(\nu - m_0) + 1) < N(m_1 - m_0, 0; 2(\nu - m_0) + 1).$$

If $0 \leq m_1 \leq m_2 \leq [\frac{2\nu+m_0+1}{4}]$, then

$$\frac{q^{2\nu-2m_1} + 1}{q^{2(m_2-m_0)} - 1} > 1,$$

i.e.,

$$N(m_2 - m_0, 0; 2(\nu - m_0) + 1) > N(m_1 - m_0, 0; 2(\nu - m_0) + 1). \qquad \square$$

4 Comparison of test efficiency

In order to explain whether a pooling design is good or not, we give a measure of the constructions. As the less the number of tests is, the better the pooling design is. So we should compare t with n, where t denotes the number of tests, i.e. the number of rows of inclusion matrix, n denotes the number of detected items, i.e. the number of columns of inclusion matrix.

In Erdös et al. (1985), P. Erdös, P. Frankl and D. Füredi give a formula that $t(d,n) > d(1 + o(1)) \ln n$, where $t(d,n)$ denotes the minimum number of rows for a d-disjunct matrix with n columns. To take $t/\ln n$ as a measure of the construction of d^z-disjunct matrix is meaningful.

Definition 4.1 We call the ratio $t/\ln n$ test efficiency 1, denoted by $R(t/\ln n)$, where t denotes the number of tests, n denotes the number of detected items.

From the definition, we know that the smaller the value of $R(t/\ln n)$ is, the better the pooling design is. But in this paper, the number is very complex, it is very difficult to compare $R(t/\ln n)$ among the constructions.

One of our goals is to identify most positive items with least tests. So we can take t/n as a measure of the construction of the d^z-disjunct matrix also. The smaller the ratio of t/n is, the better the construction is. Another definition is given as follows.

Definition 4.2 We call the ratio t/n test efficiency 2, denoted by $R(t/n)$. Where t denotes the number of tests, n denotes the number of detected items.

In fact, $t/\ln n$ can be converted to t/n under some conditions. It is easier to compare t/n in this paper. We compare test efficiency 2 in the following.

Now we give the comparison of test efficiency 2.

In this paper

$$R(t_\delta/n_\delta) = \frac{\prod_{i=r-m_0+1}^{m-m_0}(q^{2i}-1)}{\prod_{i=2\nu-2m+1+\delta}^{2\nu-2r+\delta}(q^i-(-1)^i)}.$$

Zhang et al. (2007) construct a d^z-*disjunct* matrix with subspaces in a dual space of Symplectic Space $\mathbb{F}_q^{(2\nu)}$, where q is a prime or a prime power. Each of the columns (rows) is labelled by $(m,0)$-*spaces* ($(r,0)$-*spaces*) which are contained in $P_0^\perp$ and containing P_0. $m_{ij}=1$ if and only if i is contained in j. The test efficiency is

$$R(t/n) = \frac{\prod_{i=r-m_0+1}^{m-m_0}(q^i-1)}{\prod_{i=\nu-m+1}^{\nu-r}(q^{2i}-1)}.$$

A.G. D'yachkov et al. (2005) constructed with subspaces of $GF(q)$, where q is a prime or a prime power. Each of the columns (rows) is labelled by a $m(r)$-*dimensional* space, $m_{ij}=1$ if and only if the label of row i is contained in the label of column j. In order to compare with $R(t_\delta/n_\delta)$, let the dimension of the space be $2\nu-m_0+\delta$. Assume the test efficiency is $R(t_\delta{}^*/n_\delta{}^*)$, then

$$R(t_\delta{}^*/n_\delta{}^*) = \frac{\begin{bmatrix}2\nu-m_0+\delta\\ r\end{bmatrix}_q}{\begin{bmatrix}2\nu-m_0+\delta\\ m\end{bmatrix}_q} = \frac{\prod_{i=r+1}^{m}(q^i-1)}{\prod_{i=2\nu-m_0-m+\delta+1}^{2\nu-m_0-r+\delta}(q^i-1)}.$$

Macula (1996) proposed a way of constructing d-*disjunct* matrix which uses the containment relation in a structure. More specifically, let $S=\{1,2,\ldots,s\}$ be the base set, Then each of the columns (rows) is labelled by a $k(d)$ set of S, where $d<k<s$. $m_{ij}=1$ if and only if the label of row i is contained in the label of column j. In the same way, let $S=\{1,2,\ldots,2\nu-2m_0+\delta\}$ be the base set. Assume the test efficiency is $R(\overline{t_\delta}/\overline{n_\delta})$, then

$$R(\overline{t_\delta}/\overline{n_\delta}) = \frac{\binom{2\nu-2m_0+\delta}{r}}{\binom{2\nu-2m_0+\delta}{m}} = \frac{m!}{r!(2\nu-2m_0+\delta-r)\cdots(2\nu-2m_0-m+1)}.$$

Theorem 4.3 *If* $2\nu+m_0>2m+r-2\delta$, *then* $R(t_\delta/n_\delta)<R(t/n)$.

Proof

$$R\left(\frac{t_\delta}{n_\delta}\right)\Big/R\left(\frac{t}{n}\right)=\frac{\prod_{i=r-m_0+1}^{m-m_0}(q^{2i}-1)}{\prod_{i=r-m_0+1}^{m-m_0}(q^{i}-1)}\cdot\frac{\prod_{i=\nu-m+1}^{\nu-r}(q^{2i}-1)}{\prod_{i=2\nu-2m+1+\delta}^{2\nu-2r+\delta}(q^{i}-(-1)^{i})}$$

$$=\prod_{i=r-m_0+1}^{m-m_0}(q^{i}+1)\cdot\frac{\prod_{i=\nu-m+1}^{\nu-r}(q^{2i}-1)}{\prod_{i=2\nu-2m+1+\delta}^{2\nu-2r+\delta}(q^{i}-(-1)^{i})}.$$

When $\delta=0$,

$$R\left(\frac{t_0}{n_0}\right)\Big/R\left(\frac{t}{n}\right)$$

$$=\prod_{i=r-m_0+1}^{m-m_0}(q^{i}+1)\cdot\frac{\prod_{i=\nu-m+1}^{\nu-r}(q^{2i}-1)}{\prod_{i=2\nu-2m+1}^{2\nu-2r}(q^{i}-(-1)^{i})}$$

$$=(q^{r-m_0+1}+1)\cdots(q^{m-m_0}+1)$$

$$\times\frac{(q^{2(\nu-m+1)}-1)(q^{2(\nu-m+2)}-1)\cdots(q^{2(\nu-r)}-1)}{(q^{2(\nu-m+1)}+1)(q^{2(\nu-m+1)}-1)\cdots(q^{2(\nu-2r)}-1)}$$

$$=\prod_{i=1}^{m-r}\frac{q^{r-m_0+i}+1}{q^{2\nu-2m+2i-1}+1}.$$

When $r-m_0+i<2\nu-2m+2i-1$, $\frac{q^{r-m_0+i}+1}{q^{2\nu-2m+2i-1}+1}<1$.

That is, when $2\nu+m_0>2m+r$, $\prod_{i=1}^{m-r}\frac{q^{r-m_0+i}+1}{q^{2\nu-2m+2i-1}+1}<1$, i.e. $R(t_0/n_0)<R(t/n)$.

When $\delta=1$,

$$R\left(\frac{t_1}{n_1}\right)\Big/R\left(\frac{t}{n}\right)=\prod_{i=r-m_0+1}^{m-m_0}(q^{i}+1)\cdot\frac{\prod_{i=\nu-m+1}^{\nu-r}(q^{2i}-1)}{\prod_{i=2\nu-2m+2}^{2\nu-2r+1}(q^{i}-(-1)^{i})}$$

$$=\prod_{i=1}^{m-r}\frac{q^{r-m_0+i}+1}{q^{2\nu-2m+2i+1}+1}.$$

When $2\nu-2m+2i+1>r-m_0+i$, $\frac{q^{r-m_0+i}+1}{q^{2\nu-2m+2i+1}+1}<1$.

That is, when $2\nu+m_0>2m+r-2$, $\prod_{i=1}^{m-r}\frac{q^{r-m_0+i}+1}{q^{2\nu-2m+2i-1}+1}<1$, i.e. $R(t_1/n_1)<R(t/n)$.

Therefore when $2\nu+m_0>2m+r-2\delta$, $R(t_\delta/n_\delta)<R(t/n)$. □

Theorem 4.4 *If* $2\nu+m_0>3m-3$, *then* $R(t_\delta/n_\delta)<R(t^*_\delta/n^*_\delta)$.

Proof When $\delta = 0$,

$$R\left(\frac{t_0}{n_0}\right)\Big/R\left(\frac{t_0{}^*}{n_0{}^*}\right) = \frac{\prod_{i=r-m_0+1}^{m-m_0}(q^i-1)}{\prod_{i=r+1}^{m}(q^i-1)} \cdot \frac{\prod_{i=2\nu-m-m_0+1}^{2\nu-r-m_0}(q^i-1)}{\prod_{i=2\nu-2m+1}^{2\nu-2r}(q^i-(-1)^i)}$$

$$= \prod_{i=1}^{m-r}\frac{q^{r-m_0+i}-1}{q^{r+i}-1}\cdot\prod_{i=1}^{m-r}\frac{(q^{2\nu-m-m_0+i}-1)}{(q^{2\nu-2m+2i}+1)(q^{2\nu-2m+2i}-1)}$$

$$< \prod_{i=1}^{m-r}\frac{(q^{2\nu-m-m_0+i}-1)}{(q^{2\nu-2m+2i}+1)(q^{2\nu-2m+2i}-1)}.$$

When $m - m_0 - i - 1 < 2\nu - 2m + 2i - 1$,

$$\frac{(q^{2\nu-m-m_0+i}-1)}{(q^{2\nu-2m+2i}+1)(q^{2\nu-2m+2i}-1)} < 1.$$

Therefore when $2\nu + m_0 > 3m - 3$,

$$\prod_{i=1}^{m-r}\frac{(q^{2\nu-m-m_0+i}-1)}{(q^{2\nu-2m+2i}+1)(q^{2\nu-2m+2i}-1)} < 1,$$

i.e. $R(t_0/n_0) < R(t_0{}^*/n_0{}^*)$.

When $\delta = 1$,

$$R\left(\frac{t_1}{n_1}\right)\Big/R\left(\frac{t_1{}^*}{n_1{}^*}\right) = \frac{\prod_{i=r-m_0+1}^{m-m_0}(q^i-1)}{\prod_{i=r+1}^{m}(q^i-1)} \cdot \frac{\prod_{i=2\nu-m-m_0+2}^{2\nu-r-m_0+1}(q^i-1)}{\prod_{i=2\nu-2m+2}^{2\nu-2r+1}(q^i-(-1)^i)}$$

$$= \prod_{i=1}^{m-r}\frac{q^{r+m_0+i}-1}{q^{r+i}-1}\cdot\prod_{i=1}^{m-r}\frac{(q^{2\nu-m-m_0+1+i}-1)}{(q^{2\nu-2m+2i}-1)(q^{2\nu-2m+2i+1}+1)}$$

$$< \prod_{i=1}^{m-r}\frac{(q^{2\nu-m-m_0+1+i}-1)}{(q^{2\nu-2m+2i}-1)(q^{2\nu-2m+2i+1}+1)}.$$

When $m - m_0 - i < 2\nu - 2m + 2i$,

$$\frac{(q^{2\nu-m-m_0+1+i}-1)}{(q^{2\nu-2m+2i}-1)(q^{2\nu-2m+2i+1}+1)} < 1.$$

Therefore when $2\nu + m_0 > 3m - 3$,

$$\prod_{i=1}^{m-r}\frac{(q^{2\nu-m-m_0+1+i}-1)}{(q^{2\nu-2m+2i}-1)(q^{2\nu-2m+2i+1}+1)} < 1,$$

i.e. $R(t_1/n_1) < R(t_1{}^*/n_1{}^*)$. □

We do not find a general method to compare t_δ/n_δ with $\overline{t_\delta}/\overline{n_\delta}$. However We can show that $R(t_\delta/n_\delta)$ is smaller than $R(\overline{t_\delta}/\overline{n_\delta})$ sometimes through an example.

Example 4.5 When $q=3$, $m_0=3$, $r=4$, $m=7$ and $\nu=9$, $R(t_\delta/n_\delta) < R(\overline{t_\delta}/\overline{n_\delta})$.
$R(t_\delta/n_\delta) = \frac{\prod_{i=2}^{4}(q^{2i}-1)}{\prod_{i=5+\delta}^{10+\delta}(q^i-(-1)^i)}$.

When $\delta=0$, $R(\overline{t_0}/\overline{n_0}) = \frac{5}{8}$,
$R(t_0/n_0) = \frac{q^4-1}{(q^5+1)(q^7+1)(q^9+1)(q^{10}-1)} = \frac{80}{244\times 2188\times 19684\times 59048}$.
Clearly $R(t_0/n_0) < R(\overline{t_0}/\overline{n_0})$.

When $\delta=1$,
$R(\overline{t_1}/\overline{n_1}) = \frac{5}{12}$,
$R(t_0/n_0) = \frac{q^4-1}{(q^7+1)(q^9+1)(q^{10}-1)(q^{11}+1)} = \frac{80}{2188\times 19684\times 59048\times 177148}$.
Clearly $R(t_1/n_1) < R(\overline{t_1}/\overline{n_1})$.

References

D'yachkov AG, Hwang FK, Macula AJ, Vilenkin PA, Weng C (2005) A construction of pooling designs with some happy surprises. J Comput Biol 12:1129–1136

D'yachkov AG, Macula AJ, Vilenkin PA (2007) Nonadaptive and trivial two-stage group testing with error-correcting d^e-disjunct inclusion matrices. In: Boylai Society Mathematical Studies, vol 16. Springer, New York, pp 71–83

Erdös P, Frankl P, Füredi D (1985) Families of finite sets in which no set is covered by the union of r others. Isr J Math 51:79–89

Huang T, Weng C (2004) Pooling spaces and non-adaptive pooling designs. Discrete Math 282:163–169

Macula AJ (1996) A simple construction of d-disjunct matrices with certain constant weights. Discrete Math 162:311–312

Ngo HQ, Du D-Z (2002) New constructions of non-adaptive and error-tolerance pooling designs. Discrete Math 243:161–170

Wan Z (2002) Geometry of classical groups over finite fields, 2nd edn. Science, Beijing

Zhang G, Li B, Sun X (2007) A construction of d^z-disjunct matrices in a dual space of symplectic space. Discrete Appl Math (in press)