New constructions of error-tolerance pooling designs

Haixia Guo[1,2,*] Jizhu Nan[2,†]

1. College of Science, Tianjin University of Technology and Education,Tianjin,300222, P. R. China

2. School of Mathematical Sciences, Dalian University of Technology, Dalian,116024, P. R. China

Abstract

It's well known that all of the pooling designs constructed in a finite set or a finite vector space. In this paper, we construct two families of pooling designs not only based on finite set (resp. finite vector space) but also partial mappings (resp. partial linear mappings), and discuss their error-tolerance properties.

Key words: Pooling design; Partial mapping; Partial linear mapping

1 Introduction

The basic problem of non-adaptive group testing is to identify the defective parts as the subset of objects being tested. Pooling design is a mathematical tool to find the defective items using the minimum number of tests. A pooling design is usually represented by a binary matrix whose columns are indexed with items and rows are indexed with pools. An entry at cell (i, j) is 1 if and only if the jth pool is contained by the ith item, and 0, otherwise. A mathematical model with error-correcting presented in [1] is an s^e-disjunct matrix. A binary matrix M is said to be s^e-disjunct if given

*E-mail address: ghx626@126.com

†E-mail address: jznan@163.com

any $s+1$ columns of M with one designated, there are $s+1$ rows with a 1 in the designate column and 0 in each of the other s columns (see [2]). An s^e-disjunct matrix can be employed to discern s defectives, detect e errors and correct $\lfloor e/2 \rfloor$ errors (see [3]). The s^e-disjunct matrix has become an important tool for determining pooling design. Most of pooling designs were constructed by the containment relation of subsets (resp. subspaces) in a finite set (resp. vector space)(see [4]-[8]). Inspired by these studies, we will restrict our attention to construct pooling designs based on partial mappings and partial linear mappings, and the error-tolerant properties are the same with Macula's designs but our pooling can contain more items.

2 Construction I

Given integers m, n and prime power q. Then Gaussian coefficient denoted by

$$\begin{bmatrix} n \\ m \end{bmatrix}_q = \begin{cases} \frac{n(n-1)\cdots(n-m+1)}{m!}, & \text{if } q=1; \\ \frac{\prod_{i=n-m+1}^{n}(q^i-1)}{\prod_{i=1}^{m}(q^i-1)}, & \text{if } q \neq 1. \end{cases}$$

For convenience, we write $\binom{n}{m}$ to substitute $\begin{bmatrix} n \\ m \end{bmatrix}_1$. And we let $\begin{bmatrix} n \\ 0 \end{bmatrix}_q = 1$ and $\begin{bmatrix} n \\ m \end{bmatrix}_q = 0$ whenever $m < 0$ or $n < m$.

Let $[n] = \{1, 2, \cdots, n\}$. Denote by $\mathcal{M}(k;n)$ the set of all k-pairs (C, f) where C is a k-subset of $[n]$ and $f : C \to [n]$ is a mapping. The pairs are called *partial mappings*. For $(C, f) \in \mathcal{M}(k;n)$ and $(D, g) \in \mathcal{M}(d;n)$, the pair (D, g) is called a *d-pair* of (C, f) if $D \subseteq C$ and $f|_D = g$, where $f|_D$ is the restriction of f on D. If (D, g) is a d-pair of (C, f), we also say that (C, f) contains (D, g).

Definition 2.1. *Given integers* $1 \leq d < k < n$. *Let* $M(d, k; n)$ *be the binary matrix with rows indexed with* $\mathcal{M}(d;n)$ *and columns indexed with* $\mathcal{M}(k;n)$, *such that* $M((A, f), (B, g)) = 1$ *if and only if* (B, g) *contains* (A, f).

Note that $M(d, k; n)$ is an $n^d\binom{n}{d} \times n^k\binom{n}{k}$ matrix, whose constant row (resp. column) weight is $n^{k-d}\binom{n-d}{k-d}$ (resp. $\binom{k}{d}$). And $\{(n^k\binom{n}{k})/(n^d\binom{n}{d})\}$:

$(\binom{n}{k}/\binom{n}{d}) = n^{(k-d)} > 1$, so this pooling can contain more items than Macula's in [2].

Theorem 2.2. *Given integers $1 \le d < k < n$, and $1 \le s < d$. Then $M(d,k;n)$ is an s^{e_1}-disjunct matrix, where $e_1 = \binom{k-s}{d-s} - 1$.*

Proof. Let $(C_0, f_0), (C_1, f_1), \dots, (C_s, f_s)$ be any $s+1$ distinct columns of $M(d,k;n)$. For each $j \in [s]$, let C_{0j} be the largest element in the set $\{C \subseteq C_0 \cap C_j \mid f_0|_C = f_j|_C\}$. Then $|C_{0j}| \le k-1$. Therefore, there exists an $a_j \in C_0 \setminus C_j$ such that $f_0(a_j) \ne f_j(a_j)$. Suppose $A_0 := \{a_1, a_2, \dots, a_s\}$. Then (C_0, f_0) contains $(A_0, f_0|_{A_0})$ but (C_j, f_j) does not contain $(A_0, f_0|_{A_0})$ for each $j \in [s]$. Note that the d-pair of (C_0, f_0) containing $(A_0, f_0|_{A_0})$ is at least $\binom{k-s}{d-s}$. Therefore $e_1 = \binom{k-s}{d-s} - 1$. □

3 Construction II

In this section, we consider the q-analog of Construction I. We begin with some useful results.

Let F_q be the finite field with q elements, where q is a prime power, and n a positive integer. Denote by $F_q^{(n)}$ the n-dimensional vector space over F_q.

Theorem 3.1. *([9, Theorem 1.7, Corollary 1.8, 1.9]) Let $0 \le k \le m \le n$. Then the number of m-dimensional vector subspaces of $F_q^{(n)}$ is $\begin{bmatrix} n \\ m \end{bmatrix}_q$, the number of k-dimensional vector subspaces contained in a given m-dimensional vector subspace of $F_q^{(n)}$ is $\begin{bmatrix} m \\ k \end{bmatrix}_q$, and the number of m-dimensional vector subspaces containing a given k-dimensional vector subspace of $F_q^{(n)}$ is $\begin{bmatrix} m-k \\ n-k \end{bmatrix}_q$.*

Let $\mathcal{M}_q(k;n)$ denote the set of all k-pairs (P, f^*) where P is a k-dimensional vector subspace of $F_q^{(n)}$ and $f^* : P \to F_q^{(n)}$ is a linear mapping. The pairs are called *partial linear mappings*. For $(P, f^*) \in \mathcal{M}_q(k;n)$ and $(Q, g^*) \in \mathcal{M}_q(d;n)$, the pair (Q, g^*) is called a *d-pair* of (P, f^*) if $Q \subseteq P$ and $f^*|_Q = g^*$, where $f^*|_Q$ is the restriction of f^* on Q. If (Q, g^*) is a d-pair of (P, f^*), we also say that (P, f^*) contains (Q, g^*).

Definition 3.2. *Given integers $1 \leq d < k < n$. Let $M_q(d,k;n)$ be the binary matrix with rows indexed with $\mathcal{M}_q(d;n)$ and columns indexed with $\mathcal{M}_q(k;n)$ of vector space $F_q^{(n)}$, such that $M_q((U,f^*),(V,g^*)) = 1$ if and only if (V,g^*) contains (U,f^*).*

Note that $M_q(d,k;n)$ is a $q^{d\times n}\begin{bmatrix} n \\ d \end{bmatrix}_q \times q^{k\times n}\begin{bmatrix} n \\ k \end{bmatrix}_q$ matrix, whose constant row (resp. column) weight is $q^{n(k-d)}\begin{bmatrix} n-d \\ k-d \end{bmatrix}_q$ (resp. $\begin{bmatrix} k \\ d \end{bmatrix}_q$). And

$$\{(q^{k\times n}\begin{bmatrix} n \\ k \end{bmatrix}_q)/(q^{d\times n}\begin{bmatrix} n \\ d \end{bmatrix}_q)\} : (\begin{bmatrix} n \\ k \end{bmatrix}_q / \begin{bmatrix} n \\ d \end{bmatrix}_q) = q^{(k-d)n} > 1,$$

so this pooling can contain more items than Macula et al's in [3].

Theorem 3.3. *Let $1 \leq d < k < n$, and set $p := \frac{q(q^{k-1}-1)}{q^{k-d}-1}$. Then $M_q(d,k;n)$ is an s^{e_2}-disjunct matrix for $1 \leq s \leq p$, where $e_2 = \begin{bmatrix} k \\ d \end{bmatrix}_q - s\begin{bmatrix} k-1 \\ d \end{bmatrix}_q + (s-1)\begin{bmatrix} k-2 \\ d \end{bmatrix}_q - 1$.*

Proof. Let $(P_0,f_0^*),(P_1,f_1^*),\ldots,(P_s,f_s^*)$ be any $s+1$ distinct columns of $M_q(d,k;n)$. To obtain the maximum numbers of d-pairs of (P_0,f_0^*) not covered by $(P_1,f_1^*),\ldots,(P_s,f_s^*)$, we may assume that for $1 \leq i \leq j$, $(P_{0i}, f_0^*|_{P_{0i}})$, is $(k-1)$-pair of (P_0,f_0^*), where P_{0i} be the largest subspace in $\{P \subseteq P_0 \cap P_i \mid f_0^*|_P = f_i^*|_P\}$ and $(P_{ij}, f_0^*|_{P_{ij}})$ is $(k-2)$-pair of $(P_{0i}, f_0^*|_{P_{0i}})$ where P_{ij} be the largest subspace in $\{P \subseteq P_0 \cap P_i \cap P_j \mid f_0^*|_P = f_i^*|_P = f_j^*|_P\}$. So, there are $\begin{bmatrix} k \\ d \end{bmatrix}_q$ many d-pairs of (P_0,f_0^*), $\begin{bmatrix} k-1 \\ d \end{bmatrix}_q$ many d-pairs of $(P_{0i}, f_0^*|_{P_{0i}})$, and $\begin{bmatrix} k-2 \\ d \end{bmatrix}_q$ many d-pairs of $(P_{ij}, f_0^*|_{P_{ij}})$. Therefore all the d-pairs of (P_0,f_0^*) not covered by $(P_1,f_1^*),\ldots,(P_s,f_s^*)$ is at least

$$\begin{bmatrix} k \\ d \end{bmatrix}_q - s\begin{bmatrix} k-1 \\ d \end{bmatrix}_q + (s-1)\begin{bmatrix} k-2 \\ d \end{bmatrix}_q.$$

Since $e_2 = \begin{bmatrix} k \\ d \end{bmatrix}_q - s\begin{bmatrix} k-1 \\ d \end{bmatrix}_q + (s-1)\begin{bmatrix} k-2 \\ d \end{bmatrix}_q - 1 \geq 0$, we obtain

$$s \leq \frac{\begin{bmatrix} k \\ d \end{bmatrix}_q - \begin{bmatrix} k-2 \\ d \end{bmatrix}_q}{\begin{bmatrix} k-1 \\ d \end{bmatrix}_q - \begin{bmatrix} k-2 \\ d \end{bmatrix}_q} = \frac{q(q^{k-1}-1)}{q^{k-d}-1}.$$

This proves Theorem 3.3. □

Acknowledgements

This research is supported by the NSF of Tianjin Municipal of China (No. 11JCYBJC00500) and Research Fund for the Doctoral Program of Higher Education of China (No. 201101647).

References

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

[2] Macula AJ (1997) Error-correcting non-adaptive group testing with d^e-disjunct matrices. Discrete Appl Math 80: 217–222

[3] 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

[4] Du D, Hwang F, Wei W, et al (2006) New construction for transversal design. J Comput Biol 13: 990–995

[5] Guo J (2010) Pooling designs associated with unitary space and ratio efficiency comparison. J Comb Optim 19: 492–500

[6] Guo J, Wang Y, Gao S (2010) Constructing error-correcting pooling designs with sympelectic space. J Comb Optim 20: 413–421

[7] Gao S, Li Z, Du H, et al (2010) Approaching pooling design with smaller efficient ratio. J Global Optim 49: 125–135

[8] Gao S, Li Z, Wang Y, et al (2011) DNA library screening pooling design and unitary spaces. Theor Comput Sci 412(3): 217–224

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