行政院國家科學委員會專題研究計畫成果報告

多面體相關的圖形和組合設計的研究 A Study of Graphs and Designs Related to Polyhedra

計畫編號: NSC90-2115-M-009-021

執行期間:90年8月1日 至 91年7月31日

主持人: 黃大原

執行機構:國立交通大學應用數學系

E-mail: thuang@math.nctu.edu.tw

本成果報告包括以下應繳交之附件:

□赴國外出差或研習心得報告一份

□赴大陸地區出差或研習心得報告一份

□出席國際學術會議心得報告及發表之論文各一份

□國際合作研究計畫國外研究報告書一份

執行單位:國立交通大學應用數學系

中華民國 91 年 10 月 25 日

Pooling Spaces and Non-Adaptive Pooling Designs

Tayuan Huang* Chih-wen Weng[†]

June 5, 2002

Abstract

A pooling space is defined to be a ranked (meet) semi-lattice with atomic intervals. We show how to construct non-adaptive pooling designs from a pooling space. All these pooling designs are e—error detecting for some e which can be chosen to be very large compared to d, the maximal number of defective items(or positives). Eight new classes of non-adaptive pooling designs are given, which are related to the Hamming matroid, the attenuated space, and six classical polar spaces. The general constructions of pooling spaces from known ones are discussed.

Keywords: pooling spaces, pooling designs, ranked, atomic, semilattice

1 Introduction

The basic problem of group testing is to identify the set of defective ones in a large population of items. A group testing algorithm is *non-adaptive* if all tests must be specified without knowing the outcomes of other tests. A

^{*}Department of Applied Mathematics, National Chiao Tung University, Taiwan R.O.C.

[†]Department of Applied Mathematics, National Chiao Tung University, Taiwan R.O.C.

non-adaptive group testing algorithm is useful in many areas. One of the examples is the problem of DNA library screening. Suppose we have n items to be tested. Assume there are at most d defective items and we want to determine them in simultaneous t tests. Each test (or pool) contains a subset of items. The output of a pool is positive if an item in the pool is infected. A mathematical model of this problem is a d-disjunct matrix (see Section ?? for formal definitions). In this paper, we define a pooling space to be a ranked (meet) semi-lattice with atomic intervals and show how to construct d-disjoint matrices from a pooling space. These d-disjunct matrices have a special property described below. If we view these d-disjunct matrices as (d-1)-disjunct matrices, then they detect e errors for some positive integer e. As our examples show the number e is very large compared to the maximal number d of defective items. A. Macula [4], [5] gave a construction of ddisjunct matrices from the poset consisting of the subsets of a finite set. H. Ngo and D. Du [7] gave a construction of d-disjunct matrices from the poset consisting of the subspaces of a vector space. Our construction is a generalization of their results. This generalization problems were initially proposed by H. Ngo and D. Du[6, p177].

2 Preliminaries

Let M be a $t \times n$ 01-matrix. We view each column i (resp. row j) as a set R_i (resp. R_j) that contains all row indices j (resp. column indices i) such that $M_{ji} = 1$. M is said to be d-disjunct if the union of any d columns does not contain another. A d-disjunct $t \times n$ matrix M can be used to design a non-adaptive group testing algorithm on n items by associating the column indices with the items and the row indices with the pools. If $M_{ij} = 1$ then item j is contained in pool i. Let M be a d-disjunct matrix. The weight wt(u) of a column vector or a row vector u of M is the number of 1's in u. For $S \subseteq \{1, 2, \cdots, n\}$ with $|S| \leq d$, the test result of S in M is the union of those columns indexed by S. A d-test result of M is a test result of M in M for some M are all distinct. The design of a M-disjunct matrix is also called non-adaptive pooling design.

Let M be a d-disjunct matrix. Then we say M detects e errors and

corrects $\lfloor \frac{e}{2} \rfloor$ errors if any two distinct d-test results of M have Hamming distance at least e+1; in other words, any two distinct d-test results have at least e+1 rows with different entries. We emphasize that the concept of e-error-detecting and $\lfloor \frac{e}{2} \rfloor$ -error correcting does not only depend on the matrix M, but also depends on the integer d. The meaning of the variables t, n, d, e is fixed throughout the paper. e is also called the e-rror-tolerance number of M.

We now give the basic definitions and properties of a partial ordered set. The expert may want to skip the remaining of this section and go to the next section.

Let P denote a finite set. By a partial order on P, we mean a binary relation \leq on P such that

- (i) $x \le x$ $(\forall x \in P)$,
- (ii) $x \le y$ and $y \le z \longrightarrow x \le z$ $(\forall x, y, z \in P)$,
- (iii) $x \le y$ and $y \le x \longrightarrow x = y$ $(\forall x, y \in P)$.

By a partially ordered set (or poset, for short), we mean a pair (P, \leq) , where P is a finite set, and where \leq is a partial order on P. Abusing notation, we will suppress reference to \leq , and just write P instead of (P, \leq) .

Let P denote a poset, with partial order \leq , and let x and y denote any elements in P. As usual, we write x < y whenever $x \leq y$ and $x \neq y$. We say y covers x whenever x < y, and there is no $z \in P$ such that x < z < y. An element $x \in P$ is said to be *minimal* whenever there is no $y \in P$ such that y < x. Let $\min(P)$ denote the set of all minimal elements in P. Whenever $\min(P)$ consists of a single element, we denote it by 0, and we say P has a 0.

Suppose P has a 0. By an atom in P, we mean an element in P that covers 0. We let A_P denote the set of atoms in P.

Suppose P has a 0. By a rank function on P, we mean a function

$$\operatorname{rank}: P \longrightarrow \mathbb{Z}$$

such that rank(0) = 0, and such that for all $x, y \in P$,

$$y \text{ covers } x \longrightarrow \operatorname{rank}(y) - \operatorname{rank}(x) = 1.$$

Observe the rank function is unique if it exists. P is said to be ranked whenever P has a rank function. In this case, we set

$$rank(P) := \max\{rank(x)|x \in P\},\$$

$$P_i := \{x | x \in P, \operatorname{rank}(x) = i\} \quad (i \in \mathbb{Z}),$$

and observe $P_0 = \{0\}, P_1 = A_P$.

Let P denote any poset, and let S denote any subset of P. Then there is a unique partial order on S such that for all $x, y \in S$,

$$x \le y \text{ (in } S) \longleftrightarrow x \le y \text{ in } P.$$

This partial order is said to be *induced* from P. By a *subposet* of P, we mean a subset of P, together with the partial order induced from P. Pick any $x, y \in P$ such that $x \leq y$. By the *interval* [x, y], we mean the subposet

$$[x,y] := \{z | z \in P, x \le z \le y\}$$

of P.

Let P denote any poset, and pick any $x,y \in P$. By a lower bound for x,y, we mean an element $z \in P$ such that $z \leq x$ and $z \leq y$. Suppose the subposet of lower bounds for x,y has a unique maximal element. In this case we denote this maximal element by $x \wedge y$, and say $x \wedge y$ exists. The element $x \wedge y$ is known as the meet of x and y. P is said to be (meet) semi-lattice whenever P is nonempty, and $x \wedge y$ exists for all $x,y \in P$. A semi-lattice has a 0. Suppose P is a semi-lattice, and pick $x,y \in P$. By an upper bound for x and y, we mean an element $z \in P$ such that $z \geq x$ and $z \geq y$. Observe that the subset of upper bounds for x and y is closed under $x \in P$ in particular, it has a unique minimal element iff it is non-empty. In this case we denote this minimal element by $x \vee y$, and say that $x \vee y$ exists. The element $x \vee y$ is known as the join of x and y.

Suppose P is a semi-lattice. Then P is said to be atomic whenever each element of P is a join of atoms. Observe if P is a ranked atomic semi-lattice, then

$$|[0,x] \cap P_1| \ge \operatorname{rank}(x) \tag{2.1}$$

for all $x \in P$.

Let q be a positive integer. The Gaussian binomial coefficients with basis q is defined by

$$\begin{bmatrix} N \\ i \end{bmatrix}_q = \begin{cases} \prod_{j=0}^{i-1} \frac{N-j}{i-j} & \text{if } q = 1, \\ \prod_{j=0}^{i-1} \frac{q^N - q^j}{q^i - q^j} & \text{if } q \neq 1. \end{cases}$$

In the case q = 1, for convenience, we write $\binom{N}{i}$ instead of $\binom{N}{i}_1$. Now assume q = 1, or a prime power. Set

$$L_q(N) = \begin{cases} \text{all subsets of } \{1, 2, \cdots, N\} & \text{if } q = 1, \\ \text{subspaces of } GF(q)^N & \text{if } q \text{ is a prime power,} \end{cases}$$

where GF(q) is the finite field of q elements. Let $P = L_q(N)$ be a poset with the usual set inclusion order. Note that

$$\begin{bmatrix} N \\ i \end{bmatrix}_{a} = |P_i|.$$

3 Construct Error Tolerable d-disjunct Matrices

Let P be a poset. For any $w \in P$, define

$$w^+ = \{ y \ge w | y \in P \}.$$

A pooling space is a ranked semi-lattice P such that w^+ is atomic for all $w \in P$. Note that this definition is equivalent to that a pooling space is a ranked semi-lattice with each interval being atomic. It is also clear that a pooling space is atomic. If P is a pooling space, then so is w^+ for any $w \in P$. We show how to construct d-disjunct matrices from a pooling space in this section.

Theorem 3.1. Let P be a pooling space with rank $D \geq 2$. Fix an element $x \in P_D$ and fix an integer ℓ $(1 \leq \ell < D)$. Let $T \subseteq P_D$ be a subset such that $|T| \leq \ell$ and $x \notin T$. Then the following (i)-(iii) hold.

- (i) There exists an element $y \in [0, x] \cap P_{\ell}$ such that $y \not\leq z$ for all $z \in T$.
- (ii) For any integer ℓ' $(\ell \leq \ell' < D)$, there are at least $f_{xT\ell'} + 1$ elements $w \in [0, x] \cap P_{\ell'}$ such that $w \not\leq z$ for all $z \in T$, where

$$f_{xT\ell'} = |\bigcup[y, x] \cap P_{\ell'}| - 1 \tag{3.1}$$

with the union taking from all y in (i).

(iii) For any integer ℓ' ($\ell \leq \ell' < D$), there are at least $e_{xT\ell'} + 1$ elements $w \in [0, x] \cap P_{\ell'}$ such that $w \not\leq z$ for all $z \in T$, where

$$e_{xT\ell'} = \max |[y, x] \cap P_{\ell'}| - 1 \tag{3.2}$$

with the maximum taking from all y in (i).

Proof. (i) We prove (i) by induction on D. When D=2, we can assume $\ell=1$ and |T|=1, otherwise the theorem clearly holds. Suppose $T=\{z\}$. Then the meet $x \wedge z \in P_0 \cup P_1$. Hence $|[0,x] \cap [0,z] \cap P_1| = |[0,x \wedge z] \cap P_1| \leq 1$. Observe that $|[0,x] \cap P_1| \geq 2$ by (??). Thus we can pick an element $y \in ([0,x] \cap P_1) \setminus [0,z]$. This proves the case D=2.

In general, pick an element $z \in T$. Then $x \neq z$. Since [0, x] is atomic and since $[0, x \wedge z]$ is proper contained in [0, x], we can pick an atom $w \in [0, x] \setminus [0, x \wedge z]$. Observe that $w \not\leq z$. Hence $T \cap w^+$ has $\leq \ell - 1$ elements. In the pooling space w^+ , the element x and the elements of $T \cap w^+$ have rank D-1, and the elements of $w^+ \cap P_\ell$ have rank $\ell - 1$. Hence by induction, we can choose $y \in [w, x] \cap P_\ell$ such that $y \not\leq u$ for all $u \in T \cap w^+$. Note that clearly $y \not\leq u$ for all $u \in T \setminus w^+$. This proves (i).

(ii) This is immediate from (i).

Note that the properties of a pooling space is preserved by truncation. That is if P is a pooling space with rank D, then

$$P_0 \cup P_1 \cup \cdots \cup P_k$$

is a pooling space of rank k for each k $(0 \le k \le D)$. The following corollary is immediate from above Theorem.

Corollary 3.2. Let P be a pooling space with rank D. Fix integers d, k $(1 \le d < k \le D)$. Let M = M(D, k, d) be a 01- matrix whose rows (resp. columns) are indexed by P_d (resp. P_k) such that $M_{xy} = 1$ iff $x \le y$. Then

- (i) M = M(D, k, d) is a d-disjunct matrix.
- (ii) For each integer d' $(1 \le d' \le d)$, M is a d'-disjunct matrix which detects f errors and corrects $\lfloor \frac{f}{2} \rfloor$ errors, where referring to (??),

$$f = \min f_{xTd} - 1$$

with the minimum taking from all $x \in P_D$, $T \subseteq P_D$ and $x \notin T$.

(iii) For each integer d' $(1 \le d' \le d)$, M is a d'-disjunct matrix which detects e errors and corrects $\lfloor \frac{e}{2} \rfloor$ errors, where referring to (??),

$$e = \min e_{xTd} - 1$$

with the minimum taking from all $x \in P_D$, $T \subseteq P_D$ and $x \notin T$.

Proof. Since the truncation of a pooling space is a pooling space, we can assume k = D.

- (i) This is clear from the definition and Theorem ??(i) with $d = \ell$.
- (ii) Suppose $S, T \subseteq P_k$ are distinct subsets with |S|, $|T| \leq d$. Then at least one of them is nonempty, so assume $S \neq \emptyset$. Pick $x \in S$. Applying Theorem ??(ii) with $d = \ell'$ and $d' = \ell$, we find there are at least $f_{xTd} + 1$ rows of M which have value 1 in the column x and the value 0 in all columns indexed by T. Hence the test result of S and the test result of S have Hamming distance at least $f_{xTd} + 1$. This proves (ii).
- (iii) This is immediate from (ii) with the observation $e \leq f$.

4 Examples

In this section, we give some examples of pooling spaces P with rank D. All of these examples are called the *quantum matroids* with the base q [9], where q is 1 or a prime power. The number $|P_i|$ can be computed from results given in [9]. We suppress the details of the computing. For integers $1 \le d < k \le D$,

the examples produce the d-disjunct matrices M = M(D, k, d) have size $t \times n$, where $t = |P_d|$ and $n = |P_k|$. For each d' $(1 \le d' \le d)$, M is a d'-disjunct matrix which detects e errors and corrects $\lfloor \frac{e}{2} \rfloor$ errors, where

$$e = \begin{bmatrix} k - d' \\ d - d' \end{bmatrix}_q - 1.$$

The weight of each column of M is

$$\begin{bmatrix} k \\ d \end{bmatrix}_q$$

and the weight of each row of M is

$$\frac{|P_k|}{|P_d|} \begin{bmatrix} k \\ d \end{bmatrix}_a.$$

1. The Hamming matroid H(D, N) $(2 \le N)$ [2], [8].

Set

$$A = A_1 \cup A_2 \cup \cdots \cup A_D$$
 (disjoint union),

where

$$|A_i| = N \qquad (1 \le i \le D).$$

$$P = \{x \mid x \subseteq A, |x \cap A_i| \le 1 \text{ for all } i \quad (1 \le i \le D)\},\$$

 $x \le y$ whenever x is a subset of $y \quad (x, y \in P)$,

$$rank(x) = |x| \quad (x \in P),$$

$$|P_i| = \binom{D}{i} N^i.$$

2. The attenuated space $A_q(D, N)$ $(D \le N)$ [2], [3].

Let V denote a vector space of dimension N over the field GF(q), and fix a subspace $w \subseteq V$ of dimension N - D.

$$P = \{x \mid x \text{ is a subspace of } V, x \cap w = 0\},\$$

 $x \leq y$ whenever x is a subspace of y $(x, y \in P)$, $\operatorname{rank}(x) = \dim(x) \quad (x \in P),$ $|P_i| = \begin{bmatrix} D \\ i \end{bmatrix}_q q^{i(N-D)}.$

3. The classical polar spaces of rank D over GF(q) [1].

Let V denote a vector space over the field GF(q), and assume V possesses a given non-degenerate form. We call a subspace of V isotropic whenever the form vanishes completely on that subspace. The maximal isotropic subspaces have the same dimension, denoted by D.

 $P = \{x \mid x \text{ is an isotropic subspace of } V\},$ $x \leq y \text{ whenever } x \text{ is a subspace of } y \quad (x,y \in P),$ $\operatorname{rank}(x) = \dim(x) \quad (x \in P),$

name	$\mathbf{dim}\ V$	${f form}$	$ P_i $
$B_D(q)$	2D + 1	quadratic	$\begin{bmatrix} D \\ i \end{bmatrix}_q (1+q^D)(1+q^{D-1})\cdots(1+q^{D-i+1})$
$C_D(q)$	2D	alternating	$\begin{bmatrix} D \\ i \end{bmatrix}_q (1+q^D)(1+q^{D-1})\cdots(1+q^{D-i+1})$
$D_D(q)$	2D	$\begin{array}{c} \text{quadratic} \\ \text{(Witt index } D) \end{array}$	$\begin{bmatrix} D \\ i \end{bmatrix}_q (1 + q^{D-1})(1 + q^{D-2}) \cdots (1 + q^{D-i})$
$^2D_{D+1}(q)$	2D + 2	$\begin{array}{c} \text{quadratic} \\ \text{(Witt index } D) \end{array}$	$\begin{bmatrix} D \\ i \end{bmatrix}_q (1 + q^{D+1})(1 + q^D) \cdots (1 + q^{D-i+2})$
$^2A_{2D}(r)$	2D + 1	Hermitian	$\begin{bmatrix} D \\ i \end{bmatrix}_q (1 + q^{D + \frac{1}{2}}) (1 + q^{D - \frac{1}{2}}) \cdots (1 + q^{D - i + \frac{3}{2}})$
$^2A_{2D-1}(r)$	2D	Hermitian	$\begin{bmatrix} D \\ i \end{bmatrix}_q (1 + q^{D - \frac{1}{2}}) (1 + q^{D - \frac{3}{2}}) \cdots (1 + q^{D - i + \frac{1}{2}})$

5 Pooling Polynomials

Let P be a pooling space with rank D. The ratio $\frac{|P_d|}{|P_k|}$ is the main concerned of the construction of pooling designs, and the structure of P is less important. With this motivation, we give the following definition.

Definition 5.1. Let P be a pooling space with rank D. The pooling polynomial of P is

$$f(P) := \sum_{i=0}^{D} |P_i| x^i.$$

Note that the constant term of a pooling polynomial is always 1. With lexicographical order, 1 and 1 + x are the first two pooling polynomials.

Let P', P'' be pooling spaces with rank D', D'' respectively. We define the direct sum P'+P'' of P' and P'' in the following. The element set of P'+P'' is the disjoint union of elemnets of P' and P'' except the 0 of P' and the 0 of P'' are identical. Hence P'+P'' has |P'|+|P''|-1 elements. The partial order of P'+P'' is naturally inherited from P' and P''. It is easy to see P'+P'' is a pooling space with rank $\max\{D',D''\}$. We define the product $P' \otimes P''$ of P' and P'' in the following. The element set of $P' \otimes P''$ is

$$\{(a,b) \mid a \in P', b \in P''\}.$$

The partial order in $P' \bigotimes P''$ is defined by

$$(a,b) \le (c,d)$$
 iff $a \le c$ and $b \le d$,

for any $a, c \in P'$ and any $b, d \in P''$. It is easy to see that for any $a, c \in P'$ and $b, d \in P''$, the following (i)-(iii) hold.

- (i) rank((a, b)) = rank(a) + rank(b);
- (ii) $(a,b) \wedge (c,d) = (a \wedge c, b \wedge d);$

(iii) (a,b) is the join of $(a_1,0), \dots, (a_r,0), (0,b_1), \dots, (0,b_s)$, where a_i, b_j are atoms that satisfy $a = a_1 \vee \dots \vee a_r$ and $b = b_1 \vee \dots \vee b_s$;

(iv)
$$[(a, b), (c, d)] = (a, c) \otimes (b, d)$$
 in $P' \otimes P''$.

We conclude from (i)-(iv) above that $P' \otimes P''$ is a pooling space with rank D' + D''.

Note that if P is a pooling space then so is $P \setminus w^+$ for any $w \in P$. Let f be a pooling polynomial. By a *reduction* of f, we mean a polynomial obtained by replacing the leading coefficient of f with a smaller nonnegative integer. We immediately have the following theorem.

Theorem 5.2. Let \mathcal{F} be the set of pooling polynomials. Suppose $f_1(x), f_2(x) \in \mathcal{F}$. Then the following (i)-(iii) hold.

- (i) A reduction of $f_1(x)$ is in \mathcal{F} ;
- (ii) $f_1(x) + f_2(x) 1 \in \mathcal{F}$;
- (iii) $f_1(x)f_2(x) \in \mathcal{F}$.

Example. $(1+3x+2x^2)^m$ is a pooling polynomial, since it can be obtained from the pooling polynomial 1+x by using productions and reductions as shown in the equation

$$(1+3x+2x^2)^m = (((1+x)^3 - x^3) - x^2)^m.$$

6 Concluding Remarks

We construct d-disjunct matrices from a pooling space in Section $\ref{eq:construct}$. If we view these d-disjunct matrices as (d-1)-disjunct matrices, they are proved to be e-error-detecting for some integer e. Some examples of pooling spaces are given in Section $\ref{eq:construction}$. By checking these examples, the ratio $\frac{t}{n} = \frac{|P_d|}{|P_k|}$ is small and the error-tolerance number e is large if d, k are well chosen.

However it seems d is too small compared to n in all these examples. We show how to construct a new pooling space from given ones in Section ??. This can be used to obtain a pooling space with a desired $|P_i|$ ranges.

Of course, our list of pooling spaces is not exhausted. It can be expected that there are a lot of unknown ones and a complete list of them is unlikely to be completed. We give another class to show this line of study might have number theory involved. Fix a positive integer m, and set

 $P = \{i \mid 2 \le i \le m, \text{ and } i \text{ is an integer which contains no square factors}\}.$

The partial order in P is defined by

$$i \le j$$
 iff i divides j .

By identifying an element in P with a subset of primes, the poset P can be obtained from the infinite poset consisting all the subsets of primes and then deleting each subposet w^+ for each integer w > m (in natural integers ordering). It can be easily checked directly that P is a pooling space. However the computing of $|P_i|$ is not likely to be written as a nice formula of i and m.

A class of pooling space related to the Hermitian forms graphs are constructed in [10]. All of the examples we mentioned in this paper are obtained by some other mathematical objects. To close the paper, we ask if there is an algorithm to construct the pooling spaces directly. More precisely, try to construct a pooling space according to a given pooling polynomial if it exists.

Acknowledgements The authors would like to thank Frank Hwang for giving this manuscript a careful reading and offering valuable suggestions.

References

- [1] P. Cameron, Projective and polar spaces, QMW Maths Notes 13.
- [2] P. Delsarte, Association schemes and t-designs in regular semi-lattice, J. Combin. Theory Ser. A, 20:230-243, 1976.
- [3] T. Huang, A characterization of the association schemes of bilinear forms, European J. Combin., 8:159-173, 1987.

- [4] A. J. Macula, A simple construction of d-disjunct matrices with certain constant weights, Discrete Math., 162:311-312, 1996.
- [5] A. J. Macula, Probabilistic nonadaptive group testing in the presence of errors and dna library screening, Annals of Combinatorics, 3:61-69, 1999.
- [6] H. Ngo and D. Zu, A Survey on Combinatorial Group Testing Algorithms with Applications to DNA Library Screening, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 55:171-182, 2000.
- [7] H. Ngo and D. Du, New Constructions of Non-Adaptive and Error-Tolerance Pooling Designs, preprint.
- [8] P. Terwilliger, The incidence algebra of a uniform poset, Coding Theory and Design Theory part I: Coding Theory, IMA volumes in Mathematics and its applications Vol. 20, Springer, New York, 193-212, 1990.
- [9] P. Terwilliger, Quantum matroids, Progress in algebraic combinatorics (Fukuoka, 1993), 323–441, Adv. Stud. Pure Math., 24, Math. Soc. Japan, Tokyo, 1996.
- [10] C. Weng, *D*-bounded Distance-regular Graphs, Europ. J. Combinatorics, 18:211-229, 1997.