Constructing Error-Correcting Pooling Designs with Singular Symplectic Space

Xuemei Liu *, Xing Gao

College of Science, Civil Aviation University of China, Tianjin, 300300, P.R.China

Abstract Pooling designs are standard experimental tools in many biotechnical applications. In this paper, we construct a family of error-correcting pooling designs with the incidence matrix of two types of subspaces of singular symplectic spaces over finite fields.

Keywords Pooling designs · d^e−disjunct matrix · Singular symplectic space

AMS classification 20G40 51D25

1. Introduction

Given a set of n items with some defectives, the group testing problem is asking to identify all defectives with the minimum number of tests each of which is on a subset of items, called a pool, and the test-outcome is negative if the pool does not contain any defective and positive if the pool contains a defective.

A pooling design is a group testing algorithm of special type, also called nonadaptive group testing, in which all pools are given at the beginning of the algorithm so that no test-outcome of one pool can effect the determination of another pool. The pooling design has many applications in molecular biology, such as DNA library screening, nonunique probe selection, gene detection, etc. (Du and Hwang [1]; Du et al. [2]; D'yachkov et al. [3]).

A pooling design can be 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 the ith pool contains the jth item, and 0, otherwise. Such a binary matrix is said to be d−disjunct. With d−disjunct pooling design, decoding is very simple. Remove all items in negative pools, the remaining items are all defectives. In practice, test-outcomes may contain errors. To make pooling design error tolerant, one introduced the

*Correspondence : College of Science, Civil Aviation University of China, Tianjin,300300, P.R.China; E-mail: xm-liu771216@163.com.

concept of d^e-disjunct matrix (Macula [4]). A binary matrix M is said to be d^e-disjunct if given any $d+1$ columns of M with one designated, there are $e+1$ rows with a 1 in the designated column and 0 in each of the other d columns. The d^0-disjunctness is actually the d-disjunctness. D'yachkov et al. [5] proposed the concept of fully d^e-disjunct matrices. A d^e-disjunct matrix is fully d^e-disjunct if it is neither $(d+1)^e$- disjunct nor d^{e+1}-disjunct. There are several constructions of d^e-disjunct matrices in the literature (Balding and Torney [6]; Erdös et al. [7]; Guo et al. [8]; Guo [9]; Huang and Weng [10]; Li et al. [11]; Macula [12]; Nan and Guo [13]; Ngo and Du [14]; Zhang et al. [15], [16]). In this paper we present a new construction associated with subspaces in $F_q^{(2\nu+l)}$.

2.Singular symplectic spaces

In this section we shall introduce the concepts of subspaces of type (m,s,k) in singular symplectic spaces. Notations and terminologies will be adopted from Wan [17].

We always assume that

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

Let F_q be a finite field with q elements, where q is a prime power. Let E denote the subspace of $F_q^{(2\nu+l)}$ generated by $e_{2\nu+1}, e_{2\nu+2}, \cdots, e_{2\nu+l}$, where e_i is the row vector in $F_q^{(2\nu+l)}$ whose ith coordinate is 1 and all other coordinates are 0.

The singular symplectic group of degree $2\nu+l$ over F_q, denoted by $Sp_{2\nu+l,2\nu}(F_q)$, consists of all $(2\nu+l)\times(2\nu+l)$ nonsingular matrices T over F_q satisfying $TK_lT^t = K_l$. The row vector space $F_q^{(2\nu+l)}$ together with the right multiplication action of $Sp_{2\nu+l,2\nu}(F_q)$ is called the $(2\nu+l)$-dimensional singular symplectic space over F_q. An m-dimensional subspace P in the $(2\nu+l)$-dimensional singular symplectic space is said to be of type (m,s,k), if PK_lP^t is of rank $2s$ and $\dim(P\cap E)=k$. In particular, subspaces of type $(m,0,0)$ are called m-dimensional totally isotropic subspaces.Clearly, singular symplectic group $Sp_{2\nu+l,2\nu}(F_q)$ is transitive on the set of all subspaces of the same type in $F_q^{(2\nu+l)}$.

We begin with some useful propositions.

Denote by $\mathcal{M}(m,s,k;2\nu+l,\nu)$ the set of all subspaces of $F_q^{(2\nu+l)}$ of type (m,s,k). Then we have

Proposition 2.1(Wan [17], Theorem 3.22) *$\mathcal{M}(m,s,k;2\nu+l,\nu)$ is non-empty if and only if $k\le l$ and $2s\le m-k\le \nu+s$, and if and only if* max $\{0, m-\nu-s\}\le k\le \min\{l, m-2s\}$.

Proposition 2.2 (Wan [17], Theorem 3.24) *Let* $\max\{0, m-\nu-s\} \leq k \leq \min\{l, m-2s\}$.*Denote by* $N(m,s,k;2\nu+l,\nu) = \mid \mathcal{M}(m,s,k;2\nu+l,\nu) \mid$. *Then*

$$N(m,s,k;2\nu+l,\nu) = q^{2s(\nu+s-m+k)+(m-k)(l-k)} \times \frac{\prod_{i=\nu+s-m+k+1}^{\nu}(q^{2i}-1)\prod_{i=l-k+1}^{l}(q^{i}-1)}{\prod_{i=1}^{s}(q^{2i}-1)\prod_{i=1}^{m-k-2s}(q^{i}-1)\prod_{i=1}^{k}(q^{i}-1)}.$$

Denote by $\mathcal{M}(m_1,s_1,k_1;m,s,k;2\nu+l,\nu)$ the set of all subspaces of type (m_1,s_1,k_1) contained in a given subspace of type (m,s,k) and denote by $N(m_1,s_1,k_1;m,s,k;2\nu+l,\nu) = \mid \mathcal{M}(m_1,s_1,k_1;m,s,k;2\nu+l,\nu) \mid$. Then we have

Proposition 2.3 (Wan [17], Theorem 3.28) $\mathcal{M}(m_1,s_1,k_1;m,s,k;2\nu+l,\nu)$ *is non-empty if and only if* $k_1 \leq k \leq l$, $2s \leq m-k \leq \nu+s$, $2s_1 \leq m_1-k_1 \leq \nu+s_1$, *and* $0 \leq s-s_1 \leq (m-k)-(m_1-k_1)$, *and these four conditions are equivalent to* $k_1 \leq k \leq l$, $2s \leq m-k \leq \nu+s$, *and* $\max\{0, m_1-k_1-s-s_1\} \leq \min\{m-k-2s, m_1-k_1-2s_1\}$.

Proposition 2.4 *Assume that* $k_1 \leq k \leq l$, $2s \leq m-k \leq \nu+s$, *and* $\max\{0, m_1-k_1-s-s_1\} \leq \min\{m-k-2s, m_1-k_1-2s_1\}$. *Then*

$$N(m_1,s_1,k_1;m,s,k;2\nu+l,\nu) = \left(\sum_{t=\max\{0,m_1-k_1-s-s_1\}}^{\min\{m-k-2s,m_1-k_1-2s_1\}} q^{2s_1(s+s_1-m_1+k_1+t)+(m_1-k_1-t)(m-k-2s-t)} \times \frac{\prod_{i=s+s_1-(m_1-k_1)+t+1}^{s}(q^{2i}-1)\prod_{i=m-k-2s-t+1}^{m-k-2s}(q^{i}-1)}{\prod_{i=1}^{s_1}(q^{2i}-1)\prod_{i=1}^{m_1-k_1-t-2s_1}(q^{i}-1)\prod_{i=1}^{t}(q^{i}-1)}\right) \times \frac{\prod_{i=k-k_1+1}^{k}(q^{i}-1)}{\prod_{i=1}^{k_1}(q^{i}-1)} q^{(m_1-k_1)(k-k_1)}.$$

Proof By Wan [17], Theorem 3.29, Corollary 3.25 and Theorem 1.7, the desired result follows.

Denote by $\mathcal{M}'(m_1,s_1,k_1;m,s,k;2\nu+l,\nu)$ the set of all subspaces of type (m,s,k) containing a given subspace of type (m_1,s_1,k_1) and denote by $N'(m_1,s_1,k_1;m,s,k;2\nu+l,\nu) = \mid \mathcal{M}'(m_1,s_1,k_1;m,s,k;2\nu+l,\nu) \mid$. Then we have

Proposition 2.5 *Assume that* $k_1 \leq k \leq l$, $2s \leq m-k \leq \nu+s$, *and* $\max\{0, m_1-k_1-s-s_1\} \leq \min\{m-k-2s, m_1-k_1-2s_1\}$. *Then*

$$N'(m_1,s_1,k_1;m,s,k;2\nu+l,\nu) = \frac{N(m_1,s_1,k_1;m,s,k;2\nu+l,\nu)N(m,s,k;2\nu+l,\nu)}{N(m_1,s_1,k_1;2\nu+l,\nu)}.$$

Proof Let

$$M = \{(P,Q) | P \in \mathcal{M}(m_1,s_1,k_1;2\nu+l,\nu), Q \in \mathcal{M}(m,s,k;2\nu+l,\nu), P \subseteq Q\}.$$

We compute the size of M in the following two ways.

For a fixed subspace P of type (m_1, s_1, k_1), there are $N'(m_1, s_1, k_1; m, s, k; 2\nu + l, \nu)$ subspaces of type (m, s, k) containing P. By Proposition 2.2

$$|M| = N'(m_1, s_1, k_1; m, s, k; 2\nu + l, \nu)N(m_1, s_1, k_1; 2\nu + l, \nu).$$

For a fixed subspace Q of type (m, s, k), there are $N(m_1, s_1, k_1; m, s, k; 2\nu + l, \nu)$ subspaces of type (m_1, s_1, k_1) containing Q. By Proposition 2.2 and Proposition 2.4

$$|M| = N(m_1, s_1, k_1; m, s, k; 2\nu + l, \nu)N(m, s, k; 2\nu + l, \nu).$$

Hence the desired result follows.

Proposition 2.6 (Wan [17], Corollary 1.9) *Let $0 \leq k \leq m \leq n$. Then the number $N'(k, m, n)$ of $m-$dimensional vector subspaces containing a given $k-$dimensional vector subspace $F_q^{(n)}$ is equal to* $\begin{bmatrix} n-k \\ m-k \end{bmatrix}_q$.

3. The construction

In this section, we construct a family of inclusion matrices associated with subspaces of $F_q^{(2\nu+l)}$, and exhibit its disjunct property.

Definition 3.1 Assume $k_1 \leq k \leq l$, $2s \leq m - k \leq \nu + s$, $0 \leq m_1 - k_1 \leq \nu$, $1 \leq m_1 \leq m$, $m - k_1 \leq 2s$, $s \geq 2$. Let $A(m_1, 0, k_1; m, s, k; 2\nu + l, \nu)$ be the binary matrix whose rows (resp. columns) are indexed by $\mathcal{M}(m_1, 0, k_1; 2\nu + l, \nu)$ (resp. $\mathcal{M}(m, s, k; 2\nu + l, \nu)$). We also order elements of these sets lexicographically. $A(m_1, 0, k_1; m, s, k; 2\nu + l, \nu)$ has a 1 in row i and column j if and only if the $i-$th subspace of $\mathcal{M}(m_1, 0, k_1; 2\nu + l, \nu)$ is a subspace of the $j-$th subspace of $\mathcal{M}(m, s, k; 2\nu + l, \nu)$.

By Propositions 2.2, 2.4 and 2.5, $A(m_1, 0, k_1; m, s, k; 2\nu + l, \nu)$ is $N(m_1, 0, k_1; 2\nu + l, \nu) \times N(m, s, k; 2\nu + l, \nu)$ matrix, whose constant row (resp. column) weight is $N'(m_1, 0, k_1; m, s, k; 2\nu + l, \nu)$ (resp. $N(m_1, 0, k_1; m, s, k; 2\nu + l, \nu)$).

Theorem 3.2 *Assume $k_1 \leq k \leq l$, $2s \leq m - k \leq \nu + s$, $0 \leq m_1 - k_1 \leq \nu$, $1 \leq m_1 \leq m, m - k_1 \leq 2s$, $s \geq 2$, and let $y = N(m_1, 0, k_1; m, s, k; 2\nu + l, \nu)$, $z = N(m_1, 0, k_1; m - 1, s - 1, k; 2\nu + l, \nu)$, $u = N(m_1, 0, k_1; m - 2, m - 1 - k_1 - s, k_1; 2\nu + l, \nu)$, $v = N(m_1, 0, k_1; m - 2, m - 2 - k_1 - s, k_1; 2\nu + l, \nu)$ and $x = \max\{z - u, z - v\}$, if $1 \leq d \leq \lfloor \frac{y-z-1}{x} \rfloor + 1$, then $A(m_1, 0, k_1; m, s, k; 2\nu + l, \nu)$ is d^e-disjunct, where $e = y - z - (d-1)x - 1$. In particular, if $1 \leq d \leq \min\{\lfloor \frac{y-z-1}{x} \rfloor + 1, q + 1\}$, then $A(m_1, 0, k_1; m, s, k; 2\nu + l, \nu)$ is fully d^e-disjunct.*

Proof Let $P, P_1, P_2, \cdots, P_d$ be $d+1$ distinct columns of $A(m_1, 0, k_1; m, s, k; 2\nu+l, \nu)$. To obtain the maximum number of subspaces of $\mathcal{M}(m_1, 0, k_1; 2\nu+l, \nu)$ in

$$P \cap \bigcup_{i=1}^{d} P_i = \bigcup_{i=1}^{d} (P \cap P_i),$$

we may assume that $\dim(P \cap P_i) = m-1$ and $\dim(P \cap P_i \cap P_j) = \dim((P \cap P_i) \cap (P \cap P_j)) = m-2$, for any two distinct i and j, where $1 \leq i, j \leq d$. Since $P \in \mathcal{M}(m, s, k; 2\nu+l, \nu)$, $P \cap P_i$ (resp. $P \cap P_i \cap P_j$) is a subspace of type $(m-1, s-1, k)$ (resp. type $(m-2, m-1-k_1-s, k_1)$) or type $(m-2, m-2-k_1-s, k_1)$ of $F_q^{(2\nu+l)}$. By Proposition 2.3, $u > 0$ and $v > 0$. By Proposition 2.4, the number of subspaces of P not covered by $P_1, P_2, \cdots, P_d$ is at least

$$y-z-(d-1)[z-\min\{u,v\}] = y-dz+(d-1)\times\min\{u,v\} = y-z-(d-1)x.$$

Hence, we may take $e = y-z-(d-1)x-1$ under the given d. Since $e \geq 0$, we obtain

$$d \leq \lfloor \frac{y-z-1}{x} \rfloor + 1.$$

Now we show that the maximal dimension of $P \cap \bigcup_{i=1}^{d} P_i$ is achieved by an explicit construction. For $P \cap P_1$, by Proposition 2.4, $N(m_1, 0, k_1; m-2, m-1-k_1-s, k_1; 2\nu+l, \nu) \geq 1$ and $N(m_1, 0, k_1; m-2, m-2-k_1-s, k_1; 2\nu+l, \nu) \geq 1$. Hence there exists an $(m-2)$–dimensional subspace contained in $P \cap P_1$, denoted by Q, such that the number of subspace of type $(m_1, 0, k_1)$ contained in Q is equal to $\min\{u, v\}$. By Proposition 2.6, the number of $(m-1)$–dimensional subspaces containing Q and contained in P is equal to $q+1$, and each of these subspaces is a subspace of type $(m-1, s-1, k)$. For1 $\leq d \leq \min\{\lfloor \frac{y-z-1}{x} \rfloor + 1, q+1\}$, we choose d distinct $(m-1)$-dimensional subspaces between Q and P, say $Q_i, (1 \leq i \leq d)$. Since $N'(m-1, s-1, k; m, s, k; 2\nu+l, \nu) \geq 2$, by Proposition 2.5, for each Q_i , we can choose a subspace of type (m, s, k) denoted by P_i, such that $P \cap P_i = Q_i$.Hence, each pair of P_i and P_j overlap at the same subspace Q. This completes the proof.

Acknowledgements This work is supported by the National Natural Science Foundation of China under Grant No. 61179026 and the Fundamental Research Funds for the Central Universities(3122013k001).

References

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

[2] Du D, Hwang F (2006) Pooling designs and non–adaptive group testing: important tools for DNA sequencing. World Scientific,Singapore.

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

[5] D′yachkov AG, Macula AJ, Vilenkin PA (2007) Nonadaptive group and trivial two-stage group testing with error-correction d^e–disjunct inclusion matrices. In: Csiszár I, Katona GOH, Tardos G (eds) Entropy,search, complexity,1st edn.Springer, Berlin, pp 71-84.ISBN-10:3540325735;ISBN-13:978-3540325734.

[6] Balding DJ, Torney DC (1996) Optimal pooling designs with error detection. J Comb Theory Ser A 74 : 131 – 140.

[7] 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.

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

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

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

[11] Li Z, Gao S, Du H, Zou F, Wu W (2010) Two constructions of new error–correcting pooling designs from orthogonal spaces over a finite field of characteristic 2. J Comb Optim 20 : 325 – 334.

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

[13] Nan J, Guo J (2010) New error-correcting pooling designs associated with finite vector spaces. J Comb Optim 20 : 96 – 100.

[14] Ngo H, Du D (2002) New constructions of non−adaptive and error-tolerance pooling designs. Discrete Math 243 : 167 − 170.

[15] Zhang G, Li B, Sun X, Li F (2008) A construction of d^z-disjunct matrices in a dual space of symplectic space. Discrete Appl Math 28 : 2400 − 2406.

[16] Zhang X, Guo J, Gao S (2009) Two new error-correcting pooling designs from d−bounded distance−regular graphs. J Comb Optim 17 : 339 − 345.

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