Discrete Applied Mathematics 157 (2009) 387–390

Contents lists available at ScienceDirect

Discrete Applied Mathematics

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

Note

Transforming an error-tolerant separable matrix to an error-tolerant disjunct matrix☆

Hong-Bin Chen [a], Yongxi Cheng [b], Qian He [c,*], Chongchong Zhong [c]

[a] *Department of Applied Mathematics, National Chiao Tung University, Hsinchu, 30050, Taiwan*

[b] *Institute for Theoretical Computer Science, Tsinghua University, Beijing, 100084, China*

[c] *Department of Mathematics, Shanghai Jiao Tong University, Shanghai, 200240, China*

ARTICLE INFO

Article history:
Received 7 September 2007
Received in revised form 29 May 2008
Accepted 3 June 2008
Available online 9 July 2008

Keywords:
Error-tolerant
Separable matrices
Disjunct matrices

ABSTRACT

Recently, Chen and Hwang [H.B. Chen, F.K. Hwang, Exploring the missing link among d-separable, $\bar{d}$-separable and d-disjunct matrices, Discrete Applied Mathematics 133 (2007) 662–664] provided a method for transforming a separable matrix to a disjunct matrix. In [D.Z. Du, F.K. Hwang, Pooling Designs and Nonadaptive Group Testing — Important Tools for DNA Sequencing, World Scientific, 2006], Du and Hwang attempted to extend this result to its error-tolerant version; unfortunately, they gave an incorrect extension. This note gives a solution to this problem.

1. Introduction

Let M be a $(0, 1)$ matrix. For any set S of columns of M, $U(S)$ will denote the union of the row indices of 1-entries of all columns in S. When S is the singleton set $\{C\}$, we abuse the notation by writing $U(S)$ simply as C. M is called *d-separable* if for any two distinct d-sets S and S' of columns, $U(S) \neq U(S')$. M is called *$\bar{d}$-separable* if the restrictions $|S| = d$ and $|S'| = d$ above are changed to $|S| \leq d$ and $|S'| \leq d$, respectively. Finally, M is called *d-disjunct* if for any d-set S of columns and any column C not in S, C is not contained in $U(S)$. These three properties of $(0, 1)$ matrices have been widely studied in the literature of nonadaptive group testing designs (pooling designs), which have applications in DNA screening [2–7].

It has long been known that d-disjunctness implies $\bar{d}$-separability which in turn implies d-separability [3, Chapter 2]. Recently, Chen and Hwang [1] found a way to construct a disjunct matrix from a separable matrix to complete the cycle of implications.

Theorem 1.1 (*Chen and Hwang [1]*). *Suppose M is a $2d$-separable matrix. Then one can construct a d-disjunct matrix by adding at most one row to M.*

The notions of d-separability, $\bar{d}$-separability and d-disjunctness have error-tolerant versions. A $(0, 1)$ matrix M is called *$(d; z)$-separable* if $|U(S) \triangle U(S')| \geq z$ for any two d-sets of columns of M. It is *$(\bar{d}; z)$-separable* if the restriction of d-sets is changed to two sets each with at most d elements. Finally, M is *$(d; z)$-disjunct* if for any d-set S of columns and any column

☆ He and Zhong are supported by Science Technology Commission of Shanghai Municipality (Grant 06ZR14048) and Chinese Ministry of Education (Grant 108056).

* Corresponding author. Fax: +86 21 54743152.
E-mail addresses: andan.am92g@nctu.edu.tw (H.-B. Chen), cyx@mails.tsinghua.edu.cn (Y. Cheng), heqian@sjtu.edu.cn, nd4ren@163.com (Q. He), zhongchongchong@sjtu.edu.cn (C. Zhong).

0166-218X/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2008.06.004

C not in S, $|C \setminus U(S)| \geq z$. Note that the variable z represents some redundancy for tolerating errors [3]. For $z = 1$, the error-tolerant version is reduced to the original version.

Du and Hwang attempted to extend Theorem 1.1 to its error-tolerant version.

Theorem 1.2 (*[3, Theorem 2.7.6]*). *Suppose M is a $(2d; z)$-separable matrix. Then one can obtain a $(d; z)$-disjunct matrix by adding at most z rows to M.*

By Theorem 1.2, Du and Hwang obtained the following corollary.

Corollary 1.3 (*[3, Theorem 2.7.7]*). *A $(d; 2z)$-separable matrix can be obtained from a $(2d; z)$-separable matrix by adding at most z rows.*

Unfortunately, Theorem 1.2 is incorrect; thus Corollary 1.3 is incorrect as seen from the following counter-example. Let

$$M_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

It is easily verified that M_1 is $(2; 2)$-separable. We now show that adding two rows to M_1 cannot produce a $(1; 2)$-disjunct matrix.

Let C_1, C_2, C_3, C_4 denote the four columns of M_1. Suppose we set $C = C_i$ and $S = \{C_j\}$, $i \neq j$. Then we need two rows each containing C_i but not C_j. One such row is already provided by M_1. So we need one $(1, 0)$-pair in a new row. Since this is required for each pair of (i, j) with $i \neq j$, there are $4 \times 3 = 12$ choices of (i, j) pairs and each such pair needs a $(1, 0)$-pair in a new row; or equivalently, we need the new rows to provide twelve such $(1, 0)$-pairs. But one new row can provide at most four $(1, 0)$-pairs (achieved by a row with two 1-entries and two 0-entries). So two new rows are not sufficient for providing the twelve $(1, 0)$-pairs required by the $(1; 2)$-disjunctness property.

In this note we give a correct version of Theorem 1.2, and obtain a more rigorous statement of Theorem 1.1.

2. Main results

Lemma 2.1 (*[3, Lemma 2.1.1]*). *Suppose M is a d-separable matrix with n columns where $d < n$; then it is k-separable for every positive integer $k \leq d$.*

Note that the condition $d < n$ in Lemma 2.1 is necessary as seen from the following example: Let

$$M_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

M_2 is trivially 3-separable. But it is not 2-separable, as the union of any pair of its columns is identical.

We now generalize Lemma 2.1 to an error-tolerant version.

Lemma 2.2. *If a matrix M with n columns is $(d; z)$-separable for $d < n$, then it is $(k; z)$-separable for every positive integer $k \leq d$.*

Proof. It suffices to prove that M is $(d-1; z)$-separable. Assume that M is not $(d-1; z)$-separable. Then there exist two distinct sets S and S' each consisting of $d-1$ columns of M such that $|U(S) \triangle U(S')| < z$.

If $|S \setminus S'| = |S' \setminus S| \geq 2$, then there must exist a pair of columns (C_x, C_y) such that $C_x \in S \setminus S'$ and $C_y \in S' \setminus S$. It is easy to see that

$$|U(S \cup \{C_y\}) \triangle U(S' \cup \{C_x\})| \leq |U(S \cup \{C_y\}) \triangle U(S')| \leq |U(S) \triangle U(S')|.$$

This violates the $(d; z)$-separability of M, as desired.

Now consider the case of $|S \setminus S'| = |S' \setminus S| = 1$. It is obvious that $|S \cup S'| = d$. Thanks to $d < n$, we can take a column C of M which is in neither S nor S'. It is easily seen that $|U(S \cup \{C\}) \triangle U(S' \cup \{C\})| \leq |U(S) \triangle U(S')| < z$. This contradicts the $(d; z)$-separability of M, completing the proof. ■

We are ready to give a correct version of Theorem 1.2.

Theorem 2.3. *Suppose M is a $(2d; z)$-separable matrix with n columns where $n \geq 2d + 1$. Then one can obtain a $(d; \lceil z/2 \rceil)$-disjunct matrix by adding at most $\lceil z/2 \rceil$ rows to M.*

Proof. Suppose M is not $(d; \lceil z/2\rceil)$-disjunct. Then there exist a column C and a set S of d other columns such that $|C \setminus U(S)| < \lceil z/2\rceil$. By adding at most $\lceil z/2\rceil$ rows to M such that each row has a 1-entry at column C and 0-entries at all columns in S, we can obtain $|C \setminus U(S)| \geq \lceil z/2\rceil$. Of course, there may exist another pair (C', S') where C' is a column and S' is a set of d columns other than C', such that $|C' \setminus U(S')| < \lceil z/2\rceil$ in M. Then we break it up by using those $\lceil z/2\rceil$ rows in the same fashion. What we need to show is that this procedure is not self-conflicting, i.e., there do not exist two pairs (C, S) and (C', S') such that $|C \setminus U(S)| < \lceil z/2\rceil$, yet on the other hand $C \in S'$ while $|C' \setminus U(S')| < \lceil z/2\rceil$.

Suppose to the contrary that there exist two pairs (C, S) and (C', S') in M as described above with $|S| = |S'| = d$. Define $S_0 = \{C'\} \cup S \cup S'$, $S_1 = S_0 \setminus \{C\}$, and $S_2 = S_0 \setminus \{C'\}$. Let $s = |S_0|$; then $s \leq 2d + 1$ and $|S_1| = |S_2| = s - 1 \leq 2d$.

Note that $S_1 \neq S_2$, but they have the same cardinality which is less than $2d + 1$. We now show the symmetric difference of $U(S_1)$ and $U(S_2)$ is less than z, thus violating the assumption of $(2d; z)$-separability.

Since the only column in S_1 but not in S_2 is C' and $|C' \setminus U(S')| < \lceil z/2\rceil$, we have

$$|U(S_1) \setminus U(S_2)| < \lceil z/2\rceil. \tag{1}$$

Similarly, we can obtain

$$|U(S_2) \setminus U(S_1)| < \lceil z/2\rceil. \tag{2}$$

Eq. (1) along with Eq. (2) gives $|U(S_1) \triangle U(S_2)| < z$, implying that M is not $(s - 1; z)$-separable. This contradicts Lemma 2.2 and so we have completed the proof. ■

Corollary 2.4. *Suppose M is a 2d-separable matrix with n columns where $n \geq 2d + 1$. Then one can obtain a d-disjunct matrix by adding at most one row to M.*

Proof. It follows from Theorem 2.3 on setting $z = 1$. ■

Corollary 2.4 is a more rigorous version of Theorem 1.1. The following example shows the necessity of the extra condition $n \geq 2d + 1$ in Corollary 2.4. Let

$$M_3 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then M_3 is trivially 4-separable; but it can be easily verified that no row can be added to M_3 to make it 2-disjunct. Similarly, any matrix with $2d$ columns is trivially $(2d; z)$-separable and one does not expect that adding $\lceil z/2\rceil$ rows to an arbitrary matrix with $2d$ columns would make it $(d; \lceil z/2\rceil)$-disjunct. To see a specific counter-example, note that M_1 is trivially a $(4; 4)$-separable matrix; but adding two rows does not make it a $(2; 2)$-disjunct matrix – it is even not $(1; 2)$-disjunct as indicated at the end of Section 1.

Corollary 2.5. *Suppose M is a $(2d; z)$-separable matrix with n columns where $n \geq 2d + 1$. Then, for any positive integer $k \leq \lceil z/2\rceil$, one can obtain a $(d; k)$-disjunct matrix by adding at most k rows to M.*

Proof. The proof of Theorem 2.3 shows that there do not exist two pairs (C, S) and (C', S') such that $|C \setminus U(S)| < \lceil z/2\rceil$, yet on the other hand $C \in S'$ while $|C' \setminus U(S')| < \lceil z/2\rceil$. In fact, the term $\lceil z/2\rceil$ can be replaced by any positive integer k which satisfies the symmetric difference of $U(S_1)$ and $U(S_2)$ is less than z. Therefore, for any $k \leq \lceil z/2\rceil$, we can obtain a $(d; k)$-disjunct matrix by adding at most k rows to M in the same fashion. ■

The following equivalence relation is given in [3] without giving a proof. We now give a proof and use the equivalence relation to obtain a stronger result.

Lemma 2.6 *([3, Lemma 2.7.5]). A matrix M is $(\bar{d}; z)$-separable if and only if it is $(d; z)$-separable and $(d - 1; z)$-disjunct.*

Proof. Suppose M is $(\bar{d}; z)$-separable but not $(d - 1; z)$-disjunct, in other words, there exists a set S of $d - 1$ columns other than a column C such that $|C \setminus U(S)| \leq z$. Then it is easy to see that $|U(S \cup \{C\}) \triangle U(S)| = |U(S \cup \{C\}) \setminus U(S)| \leq z$, a contradiction to $(\bar{d}; z)$-separability. Thus, M is $(d - 1; z)$-disjunct and $(d; z)$-separable trivially.

Let M be $(d; z)$-separable and $(d - 1; z)$-disjunct. It suffices to show that $|U(X) \triangle U(Y)| \geq z$ for any two sets X, Y of at most d columns. If $|X| = |Y| \leq d$, then $|U(X) \triangle U(Y)| \geq z$ by $(d; z)$-separability and Lemma 2.2. Assume $|X| < |Y| \leq d$; then there exists a column $C_y \in Y$ but not in X. By $(d - 1; z)$-disjunctness, we obtain $|C_y \setminus U(X)| \geq z$; hence $|U(X) \triangle U(Y)| \geq z$. This completes the proof. ■

By Lemmas 2.6 and 2.2, we extend Corollary 2.5 to a stronger version.

Corollary 2.7. *Suppose M is a $\overline{(2d}; z)$-separable matrix with n columns where $n \geq 2d + 1$. Then, for any positive integer $k \leq \lceil z/2\rceil$, one can obtain a $(\overline{d + 1}; k)$-separable matrix by adding at most k rows to M.*

3. Concluding remarks

The following remarks demonstrate the optimality of our results.

Remark 1. The constraint $k \leq \lceil z/2 \rceil$ in Corollary 2.5 is necessary if we want the number of rows added to be independent of n and d. To see a specific example, consider that M is an $(n\lceil z/2 \rceil) \times n$ matrix such that each column has $\lceil z/2 \rceil$ 1-entries and any two columns have no intersection. Then, M is $(2d; z)$-separable. Since every column has only $\lceil z/2 \rceil$ 1-entries, to make $M(d; k)$-disjunct by adding rows, the rows added must form a $(d; k - \lceil z/2 \rceil)$-disjunct submatrix when $k > \lceil z/2 \rceil$. In this case, the minimum number of rows required would depend on n, d and $k - \lceil z/2 \rceil$.

Remark 2. Let N be a $(0, 1)$ matrix of constant row sum 1 and constant column sum z and let M be obtained from N by adding one zero column. It is easy to verify that M is $(2d; z)$-separable. Since there is a zero column in M, we cannot obtain from M a $(d; k)$-disjunct matrix by adding less than k rows. This shows that the bound on the number of additional rows given in Corollary 2.5 is optimal in this sense.

References

[1] H.B. Chen, F.K. Hwang, Exploring the missing link among d-separable, $\bar{d}$-separable and d-disjunct matrices, Discrete Applied Mathematics 133 (2007) 662–664.
[2] D.Z. Du, F.K. Hwang, Combinatorical Group Testing and its Applications, 2nd edition, World Scientific, 1999.
[3] D.Z. Du, F.K. Hwang, Pooling Designs and Nonadaptive Group Testing — Important Tools for DNA Sequencing, World Scientific, 2006.
[4] P. Erdős, P. Frankl, Z. Füredi, Families of finite sets in which no set is covered by the union of r others, Israel Journal of Mathematics 51 (1985) 79–89.
[5] F.K. Hwang, V.T. Sós, Non-adaptive hypergeometric group testing, Studia Scientiarum Mathematicarum Hungarica 22 (1987) 257–263.
[6] A.J. Macula, A simple construction of d-disjunct matrices with certain constant weights, Discrete Mathematics 162 (1996) 311–312.
[7] A.J. Macula, Error-correcting nonadaptive group testing with d^e-disjunct matrices, Discrete Applied Mathematics 80 (1997) 217–222.