

# On the construction of multi-level supersaturated designs

Kai-Tai Fang<sup>a</sup>, Dennis K.J. Lin<sup>b, \*</sup>, Chang-Xing Ma<sup>c</sup>

<sup>a</sup>*Department of Mathematics, Hong Kong Baptist University and Chinese Academy of Sciences, Beijing, People's Republic of China*

<sup>b</sup>*Department of Management Science, Penn State University, University Park, PA 16803, USA*

<sup>c</sup>*Department of Statistics, Nankai University, People's Republic of China*

Received 8 January 1999; received in revised form 15 July 1999; accepted 18 August 1999

---

## Abstract

New criteria of comparing multi-level supersaturated designs are proposed and their properties are studied. A new class of multi-level supersaturated designs are obtained by collapsing a *U*-type uniform design to an orthogonal array. A global optimization algorithm, the threshold accepting algorithm, is then applied to search for the best supersaturated designs under any prespecified criterion. Examples show that these newly constructed supersaturated designs have good modeling properties. © 2000 Elsevier Science B.V. All rights reserved.

*MSC:* 62K15; 62K10

*Keywords:* Collapsing method; Effect sparsity; Near orthogonal design; Orthogonal array; Supersaturated designs; Threshold accepting; *U*-type designs; Uniform designs

---

## 1. Introduction

Many preliminary studies in industrial and scientific experimentation contain a large number of potentially relevant factors, but often only a few are believed to have significant effects. The goal is to identify these few active factors with a relatively small number of runs. One approach is to use a supersaturated design where the number of runs is smaller than the number of unknown parameters. Developing such screening designs has received a great deal of attention: Satterthwaite (1959) proposed the idea of supersaturated design in random balanced designs, Booth and Cox (1962) examined these designs systematically. Lin (1993) provided a new class of supersaturated designs based on half-fractions of Hadamard matrices. Nguyen (1996) described a method of constructing supersaturated designs from balanced incomplete block (BIB) designs that

---

\* Correspondence address: Department of Management Science and Information Systems, 314 Beam Building, Penn State University, University Park, PA 16802-1913, USA. Tel.: (814)-865-6377; fax: (814)-863-2381.

*E-mail address:* dkl5@psu.edu (D.K.J. Lin)

is a generalization of Lin (1993). Wu (1993) augmented Hadamard matrices by adding interaction columns. Some algorithms for constructing supersaturated designs have been studied by many authors, for example, Lin (1991,1995), Li and Wu (1997), Yamada and Lin (1997) and Nguyen (1996). However, all these authors considered only 2-level designs. Designs with multi-levels are often requested in industrial and scientific experimentation for exploring nonlinear effects of the factors. Especially, designs with 3 levels have been widely used. Recently, Liu and Zhang (1997) proposed a method of constructing 3-level supersaturated designs. But their method is far too complicated to implement when the number of runs is large.

In this paper we study the problem of constructing multi-level supersaturated designs. Several criteria related to orthogonality are defined for comparing supersaturated designs. The definition of these criteria and their properties are studied in Section 2. In Section 3 we propose a unified way to construct multi-level supersaturated designs. This method collapses a  $U$ -type design to an orthogonal array. A number of supersaturated designs are obtained by this method and compared with other existing designs. Throughout the paper, we will use the popular notations;  $-1$  and  $1$  for the elements of a two-level factor and  $0, \dots, s-1$  for the elements of an  $s$ -level ( $s \geq 3$ ) factor.

## 2. Some new criteria for comparing supersaturated designs

A design matrix  $X$  is called *column-orthogonal* if  $X'X$  is a diagonal matrix. In a column-orthogonal design, the parameters corresponding to each column of the design can be estimated independently of other columns. The column orthogonality has been used in constructing 2-level supersaturated designs. Suppose  $X$  is an  $n \times m$  design matrix of a design with  $n$  runs and  $m$  2-level factors each with  $n/2$  of  $+1$  and  $-1$  (for even  $n$ ). Let  $s_{ij}$  be the  $(i, j)$ -element of  $X'X$ . Column-orthogonality is equivalent to all  $s_{ij} = 0$  for all  $i \neq j$ . Many criteria for generating saturated and supersaturated designs based on  $s_{ij}$  have been proposed. Booth and Cox (1962) suggested a criterion to minimize

$$\text{Ave}(s^2) = \sum_{i < j} s_{ij}^2 / \binom{m}{2}.$$

The criteria

$$\text{Ave}|s| = \sum_{i < j} |s_{ij}| / \binom{m}{2},$$

$s_{\max} = \max_{i < j} |s_{ij}|$  and the frequency of  $\{s_{ij} = \pm s_{\max}\}$  are also widely used (e.g., Lin, 1993,1995; Wu, 1993; Deng et al., 1999).

Two design columns  $\mathbf{x}_u$  and  $\mathbf{x}_v$  are called *fully aliased* if one column can be obtained from another by permuting levels. We cannot use two fully aliased columns to accommodate two different factors. For comparing multi-level designs, column-orthogonality is not sufficient. For example, two fully aliased 3-level columns  $\mathbf{x}_1 = (0, 1, 2, 2, 0,$

$1, 1, 2, 0)'$  and  $\mathbf{x}_2 = (1, 2, 0, 0, 1, 2, 2, 0, 1)'$ , where  $\mathbf{x}_2$  is obtained from  $\mathbf{x}_1$  by permuting levels  $(0, 1, 2) \rightarrow (1, 2, 0)$ , have correlation coefficient  $-0.5$ , not  $\pm 1$ . An  $n \times m$  design matrix with  $n$  runs and  $m$   $s$ -level factors is called a  $U$ -type design, denoted by  $U(n, s^m)$ , if for each column all of its entries appear equally often. When level  $s = n$ , we denote the  $U$ -type design by  $U(n, n^m)$ . A  $U$ -type design is called an orthogonal design, denoted by  $L_n(s^m)$ , if every pair of design columns has all of their level combinations appear equally often. In this case  $n$  is a multiple of  $s^2$ . We shall call such orthogonality *combination-orthogonality* to distinguish it from column-orthogonality. Clearly, combination-orthogonality implies column-orthogonality (column-orthogonality means that correlation coefficients of linear effects of factors are 0), but the inverse is not necessarily true. For example, two 3-level factors after relabeling,  $\mathbf{x}_1 = (-1, -1, -1, 0, 0, 0, 1, 1, 1)'$  and  $\mathbf{x}_2 = (-1, -1, 1, 0, 1, 1, -1, 0, 0)'$ , have column-orthogonality (i.e.,  $\mathbf{x}'_1 \mathbf{x}_2 = 0$ ) but not combination-orthogonality. Note that the orthogonal design defined here can be regarded as a special case of orthogonal arrays. An orthogonal array of strength  $r$  and size  $n$  with  $m$  constraints is given by an  $n \times m$  matrix with entries from a set of  $s \geq 2$  symbols, where each  $n \times r$  submatrix contains all possible  $1 \times r$  row vectors equally often. Orthogonal arrays of strength two are the orthogonal designs we have defined here. We next define some new criteria based on combination-orthogonality for comparing supersaturated designs.

**Definition 2.1.** Let  $\mathbf{x}_u$  and  $\mathbf{x}_v$  be two columns of an  $n \times m$  design matrix with  $s_u$  and  $s_v$  levels, respectively. Define

$$f(\mathbf{x}_u, \mathbf{x}_v) = \sum_{i=0}^{s_u-1} \sum_{j=0}^{s_v-1} \left| N_{u,v}(i, j) - \frac{n}{s_u s_v} \right|,$$

where  $N_{u,v}(i, j)$  is the number of  $(i, j)$ -pairs in  $(\mathbf{x}_u, \mathbf{x}_v)$ .

The value of  $f(\mathbf{x}_u, \mathbf{x}_v)$  gives a nonbalance measure among all pairs in  $(\mathbf{x}_u, \mathbf{x}_v)$ . Columns  $\mathbf{x}_u$  and  $\mathbf{x}_v$  are combination-orthogonal if and only if  $f(\mathbf{x}_u, \mathbf{x}_v) = 0$ . In this case  $n$  is a multiple of  $s^2$  when  $s_u = s_v = s$ .

**Definition 2.2.** Let  $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)$  be an  $n \times m$   $U$ -design matrix with  $n = ks^2$  runs (rows) and  $m$  factors (columns) each having  $s$  levels:  $0, \dots, s - 1$ . The following criteria are defined for measuring combination-nonorthogonality of the design  $\mathbf{X}$ .

- (1)  $\text{Ave}|f| = \sum_{1 \leq u < v \leq m} f(\mathbf{x}_u, \mathbf{x}_v) / \binom{m}{2}$ ;
- (2)  $\text{Ave}(f^2) = \sum_{1 \leq u < v \leq m} f(\mathbf{x}_u, \mathbf{x}_v)^2 / \binom{m}{2}$ ;
- (3)  $f_{\max} = \max_{1 \leq u < v \leq m} f(\mathbf{x}_u, \mathbf{x}_v)$ ;
- (4)  $N_{f_{\max}}$  = the frequency of  $\{f(\mathbf{x}_u, \mathbf{x}_v) = f_{\max}\}$  if  $f_{\max} > 0$ ; otherwise  $N_{f_{\max}} = 0$ ;
- (5)  $N_{\text{non-od}}$  = the number of  $\{f(\mathbf{x}_u, \mathbf{x}_v) \neq 0\}$ .

The criteria  $\text{Ave}|f|$  and  $\text{Ave}(f^2)$  give the combination-nonorthogonality of all column pairs of  $\mathbf{X}$  in the average sense. The criterion  $f_{\max}$  shows the worst

nonorthogonality in all pairs of columns of  $\mathbf{X}$ . When  $f_{\max} > 0$ ,  $N_{\max}$  shows the number of the worst pairs in  $\mathbf{X}$ . The  $N_{\text{non-od}}$  criterion gives the number of nonorthogonal column-pairs of the design. The design  $\mathbf{X}$  is combination-orthogonal if and only if one of the above criteria is zero (in fact, this also implies that all of these five criteria are zero). Some properties of these criteria are given in the following theorem.

**Theorem 2.1.** *For any design matrix  $\mathbf{X}$  defined in Definition 2.2, we have the following assertions:*

- (i) *For a two-level design ( $s = 2$ ),  $f(\mathbf{x}_i, \mathbf{x}_j) = |s_{ij}|$ ,  $\text{Ave}|f| = \text{Ave}|s|$ , and  $\text{Ave}(f^2) = E(s^2)$ , where  $s_{ij}$  is the  $(i, j)$ -element of  $\mathbf{X}'\mathbf{X}$ .*
- (ii)  *$f(\mathbf{x}_i, \mathbf{x}_j) \leq 2ks(s - 1)$  for  $0 \leq i, j \leq s - 1$ , and the equality holds if and only if  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are fully aliased.*
- (iii) *If  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are not combination-orthogonal, then  $f(\mathbf{x}_i, \mathbf{x}_j) \geq 4$ .*
- (iv) *All five criteria  $\text{Ave}|f|$ ,  $\text{Ave}(f^2)$ ,  $f_{\max}$ ,  $N_{f_{\max}}$  and  $N_{\text{non-od}}$  are invariant under exchanging rows and columns of  $\mathbf{X}$  and permuting levels of each column of  $\mathbf{X}$ .*
- (v) *For any given orthogonal design  $L_n(s^m)$  and positive integer  $r$ , there exists a  $U$ -type design  $U(n, s^{rm})$  satisfying*

$$\text{Ave}|f| = ks(2s - 2)m \binom{r}{2} \binom{rm}{2}^{-1} \tag{2.1}$$

and

$$\text{Ave}(f^2) = (ks(2s - 2))^2 m \binom{r}{2} \binom{rm}{2}^{-1}, \tag{2.2}$$

where  $n = ks^2$ . In particular, when  $s = 2$  and  $m = n - 1$ , (2.2) gives the lower bound of  $E(s^2)$  as given in Nguyen (1996) and Tang and Wu (1997).

**Proof.** (i) Let  $N(s, t)$  be the number of pairs  $(s, t)$  in  $(\mathbf{x}_i, \mathbf{x}_j)$ , for  $s, t = -1$  or  $1$ . Then

$$\begin{aligned} s_{ij} &= N(-1, -1) - N(-1, 1) - N(1, -1) + N(1, 1) \\ &= (N(-1, -1) - k) + (k - N(-1, 1)) + (k - N(1, -1)) + (N(1, 1) - k). \end{aligned}$$

Since  $\mathbf{x}_i$  and  $\mathbf{x}_j$  have the same number of levels of  $+1$  or  $-1$ ,

$$N(-1, -1) + N(-1, 1) = N(1, -1) + N(1, 1) = N(-1, 1) + N(1, 1) = 2k,$$

$$N(-1, -1) - k = k - N(-1, 1) = k - N(1, -1) = N(1, 1) - k.$$

Therefore,

$$\begin{aligned} |s_{ij}| &= |(N(-1, -1) - k) + (k - N(-1, 1)) + (k - N(1, -1)) + (N(1, 1) - k)| \\ &= 4|N(-1, -1) - k| = f(\mathbf{x}_i, \mathbf{x}_j). \end{aligned}$$

Assertion (i) thus follows.

(ii) Because  $\mathbf{x}_i$  and  $\mathbf{x}_j$  have the same number of levels  $0, \dots, s - 1$ , we have

$$\sum_{r=0}^{s-1} N(r, t) = sk, \quad t = 0, \dots, s - 1,$$

$$\sum_{t=0}^{s-1} N(r, t) = sk, \quad r = 0, \dots, s - 1.$$

In order to maximize  $f(\mathbf{x}_i, \mathbf{x}_j)$ , we only need to maximize each  $\sum_{r=0}^{s-1} |N(r, t) - k|$  under condition  $\sum_{r=0}^{s-1} N(r, t) = sk$ , for  $t = 0, \dots, s - 1$ . Obviously,  $f(\mathbf{x}_i, \mathbf{x}_j)$  arrives at its maximum when  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are fully aliased. In this case, for any fixed  $t$ , there exists  $r_1$  such that  $N(r_1, t) = ks$  and  $N(r, t) = 0, r \neq r_1$ . Therefore,  $\sum_{r=0}^{s-1} |N(r, t) - k| = 2(s - 1)k$  and  $\sum_{i=0}^{s-1} \sum_{i=0}^{s-1} f(\mathbf{x}_i, \mathbf{x}_j) = 2ks(s - 1)$ .

(iii) If  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are not orthogonal, there exists at least one pair,  $(i_1, j_1)$ , such that the number of its appearances is not  $k$ . Since  $\sum_{i=0}^{s-1} N(i, j_1) = ks$ , there exists another pair  $(i_2, j_1), i_2 \neq i_1$  such that the number of its appearances is not  $k$ . Similarly, there exist at least four different pairs  $(i_1, j_1), (i_2, j_1), (i_2, j_2), (i_3, j_2)$  such that the number of their appearances is not  $k$ . Obviously,  $f(\mathbf{x}_i, \mathbf{x}_j) \geq 4$ , the equality holds when  $i_3 = i_1$ , the number of appearances of these four pairs is  $k \pm 1$ , and the frequency of other pairs are all equal to  $k$ .

(iv) Obvious.

(v) Denote this  $L_n(s^m)$  by  $\mathbf{L}$ , then the design  $\mathbf{S} = (\mathbf{L}, \mathbf{L}, \dots, \mathbf{L})$  satisfies (2.1), (2.2). In fact, any two different columns  $\mathbf{x}_i$  and  $\mathbf{x}_j$  of  $\mathbf{S}$  are orthogonal to each other and so  $f(\mathbf{x}_i, \mathbf{x}_j) = 0$  while any two same columns  $\mathbf{x}_i$  and  $\mathbf{x}_j$  of  $\mathbf{S}$  are fully aliased and so  $f(\mathbf{x}_i, \mathbf{x}_j) = s|ks - k| + (s^2 - s)|0 - k| = ks(2s - 2)$ . When  $s = 2, m = n - 1$ , Tang and Wu (1997) proved the lower bound.  $\square$

Assertion (i) shows that the criteria defined in Definition 2.2 are extensions of  $\text{Ave}|s|, \text{Ave}(s^2), s_{\max}$  and the frequency of  $s_{ij} = s_{\max}$  in the case  $s = 2$ . In the case  $s = 2$ , the absolute value of correlation coefficient between  $\mathbf{x}_i$  and  $\mathbf{x}_j$  is  $|s_{ij}|/n = f(\mathbf{x}_i, \mathbf{x}_j)/n$ . For the multi-level case, we define  $f(\mathbf{x}_i, \mathbf{x}_j)/2ks(s - 1)$  as a generalized correlation coefficient from property (ii). Assertion (ii) gives the upper bound of  $f(\mathbf{x}_i, \mathbf{x}_j)$  while assertion (iii) gives the lower bound of  $f(\mathbf{x}_i, \mathbf{x}_j)$  if  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are not orthogonal. Property (iv) ensures that all proposed criteria are invariant under exchanging rows or columns of the design, or under permuting level values.

### 3. Constructing multi-level supersaturated designs

In the last decade, a number of methods and algorithms for generating 2-level supersaturated designs has been proposed. Unfortunately, it is not straightforward to extend these methods to multi-level supersaturated designs. The only exception may be the columnwise-pairwise algorithm. Li and Wu (1997) have stated the possibility of application of the columnwise-pairwise algorithm for constructing any level of supersaturated designs, although no specific multi-level designs were given. In this section

we adopt the *collapsing method* from Addelman (1962). The basic idea of the construction method is to collapse a multi-level factor into several low-level factors. Here we collapse the popular  $U$ -type uniform design (see, for example, Fang and Wang, 1994; Fang and Hickernell, 1995) to an orthogonal design.

Denote an orthogonal design  $(L_n(s^d))$ ,  $U$ -type design  $(U(n, n^r))$  and the resulting  $n \times rd$  design by  $L = (l_{ij})$ ,  $U = (u_{uv})$ , and  $X$ , respectively. Without loss of generality, let the entries of  $U$  be  $1, 2, \dots, n$  and the first column of  $U$  be  $(1, 2, \dots, n)'$ . Denote  $L$  in terms of its rows by  $L = (l_1, \dots, l_n)'$ , where each  $l_k$  is a column vector. Define a mapping:  $k \rightarrow l'_k$ , for  $k = 1, \dots, n$ . Applying this mapping to the first column of  $U$  produces the first block,  $X_1$ , of  $X$ . Clearly,  $X_1 = L$ . Applying the mapping to the second column of  $U$  produces to the second block  $X_2$  of  $X$ , and so on. As a result, a supersaturated design  $X = (X_1, \dots, X_r)$  of size  $n \times rd$  is obtained. Specifically, we have

$$X = \begin{pmatrix} l'_{u_{11}}, l'_{u_{12}}, \dots, l'_{u_{1r}} \\ l'_{u_{21}}, l'_{u_{22}}, \dots, l'_{u_{2r}} \\ \dots \dots \dots \dots \\ l'_{u_{n1}}, l'_{u_{n2}}, \dots, l'_{u_{nr}} \end{pmatrix}. \tag{3.1}$$

To be more specific, denote  $X$  as  $S_n(s^{rd})$ , called the  $S$ -design, and denote the above operation by  $S_n(s^{rd}) = U \oplus L$ . The  $U$  and  $L$  matrices are called the *generating designs*.

**Example 1.** Suppose that we extend the orthogonal design  $L = L_9(3^4)$  and a generating  $U$ -type design  $U = U(9, 9^2)$  to an  $S$ -design  $S_9(3^8)$  by the collapsing method. By the collapsing method, we have

$$U \oplus L = \begin{bmatrix} 1 & 1 \\ 2 & 7 \\ 3 & 3 \\ 4 & 9 \\ 5 & 5 \\ 6 & 6 \\ 7 & 2 \\ 8 & 8 \\ 9 & 4 \end{bmatrix} \oplus \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 1 & 2 & 0 & 1 \\ 2 & 0 & 2 & 1 \\ 2 & 1 & 0 & 2 \\ 2 & 2 & 1 & 0 \end{bmatrix} = X = \begin{bmatrix} 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & | & 2 & 0 & 2 & 1 \\ 0 & 2 & 2 & 2 & | & 0 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 & | & 2 & 2 & 1 & 0 \\ 1 & 1 & 2 & 0 & | & 1 & 1 & 2 & 0 \\ 1 & 2 & 0 & 1 & | & 1 & 2 & 0 & 1 \\ 2 & 0 & 2 & 1 & | & 0 & 1 & 1 & 1 \\ 2 & 1 & 0 & 2 & | & 2 & 1 & 0 & 2 \\ 2 & 2 & 1 & 0 & | & 1 & 0 & 1 & 2 \end{bmatrix}.$$

The first block (the first four columns) of  $X$  is the original  $L$  and the second block (the last four columns) are formed by rearranging row vectors of  $L$  according to the order defined by the second column of  $U$ . The values of  $Ave|f|$  and  $Ave(f^2)$  are 2.428 and 12.857, respectively. The crucial step of this collapsing method is how to choose the best  $U$ -type design  $U(n, n^r)$  in the sense of minimizing any specific criterion (see Definition 2.2). Denote  $\mathcal{U}(n; s^r)$  as the set of all  $U$ -type  $U(n, s^r)$ . We search for a supersaturated design  $X = S_n(s^{rd}) = U \oplus L$  by minimizing any specific criterion with respect to  $U \in \mathcal{U}(n, n^r)$ .

Table 1  
The nonorthogonalities of  $S$ -designs obtained by TA

$n, r, s, d$	$c$	$m$	$T$	Ave $ f $	Ave( $f^2$ )	$f_{\max}$	$N_{f_{\max}}$	$E( t )$	$E(t^2)$
9,2,3,4	1	4	3	2.5714	12.0000	6	4	0.1250	0.0446
9,3,3,4	1	4	3	3.2727	15.2727	6	12	0.1843	0.0720
9,4,3,4	2	4	4	3.6000	16.8000	6	24	0.2236	0.0854
16,2,4,5	1	4	3	4.3556	38.4000	12	2	0.0833	0.0211
16,3,4,5	1	4	3	5.6000	49.3714	12	6	0.1424	0.0468
16,4,4,5	2	4	4	6.2526	56.9684	16	1	0.1582	0.0491
18,2,3,7	1	4	3	3.1868	21.3187	12	1	0.1071	0.0300
18,3,3,7	1	4	3	4.2000	29.2381	12	5	0.1258	0.0370
18,4,3,7	2	4	4	4.8254	34.2645	12	9	0.1481	0.0424
25,2,5,6	1	4	3	6.4242	84.9697	22	1	0.0906	0.0224
25,3,5,6	1	4	3	8.4052	115.9477	24	2	0.1072	0.0272
25,4,5,6	2	4	4	10.1957	146.6812	30	1	0.1167	0.0285
27,2,3,13	1	4	3	3.7846	34.4738	16	2	0.0814	0.0207
27,3,3,13	1	4	3	5.2659	47.5735	18	1	0.1031	0.0256
27,4,3,13	2	4	4	5.9774	54.6094	18	2	0.1204	0.0299

Taking  $L_9(3^4)$  as an example, 2000  $U$ -type designs in  $\mathcal{U}(n, n^t)$  were randomly generated to obtain the corresponding  $S$ -designs. We then evaluated all their values for the following criteria: Ave $|f|$ , Ave( $f^2$ ),  $f_{\max}$ ,  $N_{f_{\max}}$ ,  $N_{\text{non-od}}$ , the discrepancy and  $E(t^2)$ , where

$$E(t^2) = \frac{\sum_{1 \leq i < j \leq m} r_{ij}^2}{\binom{m}{2}},$$

$m$  is the number of columns of  $X$ ,  $r_{ij}$  is the correlation coefficient between  $x_i$  and  $x_j$ . As pointed out by one referee, the correlation coefficient  $r_{ij}$  is a useful measure for quasi-Monte Carlo methods but not useful for multi-level supersaturated designs. The reason is that a correlation coefficient is meaningful only for quantitative factors. The two columns in Table 1 list  $E(|t|)$  and  $E(t^2)$  for reference purpose only. The discrepancy is a popular measure of uniformity in quasi-Monte Carlo methods (Niederreiter, 1992). Some empirical observations are:

- An  $S$ -design with a low Ave( $f^2$ ) value has a low  $E(t^2)$  value, but the inverse is not necessarily true. We find that in the 2000  $S$ -designs, most of designs with low  $E(t^2)$  value have many fully aliased column pairs. For example, if we choose  $U$ -type design

$$\begin{pmatrix} 123456789 \\ 294537861 \end{pmatrix}'$$

as the generating design, the corresponding  $S$ -design has the smallest  $E(t^2)=0.0357$ , but it includes four fully aliased column pairs. This is the reason why we did not use  $E(t^2)$  as a criterion for comparing multi-level designs.

- An  $S$ -design with a large Ave( $f^2$ ) value, in general, has a large Ave $|f|$  value, and vice versa. An  $S$ -design with small Ave( $f^2$ ) value has a small  $f_{\max}$  value, but it

does not necessarily have a small  $\text{Ave}|f|$  value. In many cases, it has a large  $\text{Ave}|f|$  value. We also find that many  $S$ -designs with small  $\text{Ave}|f|$  value may have a larger  $f_{\max}$ . Thus, the criterion  $\text{Ave}|f|$  is not recommended, and we recommend using  $\text{Ave}(f^2)$  or using the two criteria  $\text{Ave}|f|$  and  $f_{\max}$  together.

- If an  $S$ -design has a low  $\text{Ave}(f^2)$  value, its generating design  $U$  generally has a low discrepancy. However, the inverse is not necessarily true. Since computing  $\text{Ave}(f^2)$  is much easier than computing the discrepancy, this suggests a way to find a nearly uniform design by minimizing  $\text{Ave}(f^2)$  with respect to  $U \in \mathcal{U}(n, s^r)$ . The minimum design  $U^*$  is a nearly uniform design (cf. Fang and Hickernell, 1995).

Our simulation recommends the use of  $\text{Ave}(f^2)$  among the five criteria defined in Definition 2.2. When two  $S$ -designs have the same  $\text{Ave}(f^2)$  value, we choose one with smaller  $f_{\max}$  value. On the other hand, if we choose  $f_{\max}$  as the first criterion, the procedure will tend to be conservative, namely, many good designs with small  $\text{Ave}(f^2)$  values will be screened out. It is unrealistic to evaluate all possible  $S$ -designs generated by  $U \oplus L$ . Therefore, we need an optimization algorithm that can work on NP or NP hard problems. One such optimization algorithm is the *threshold accepting* (TA) algorithm proposed by Dueck and Scheuer (1990). TA has been successfully applied to many NP hard problems. By the use of TA, Winker and Fang (1996) found many  $U$ -type designs with lower discrepancy than that generated by other methods. The following gives the detailed algorithm to search for  $S$ -designs with low  $\text{Ave}(f^2)$ :

*Step 1:* Randomly choose  $U^c = (u_1^c, \dots, u_r^c) \in \mathcal{U}(n, n^r)$ , where “c” means “current”. Set control parameters  $T, t, m, JJ$  and  $II$  and compute  $D_c = \text{Ave}(f^2)$  of  $U_c \oplus L$ , where  $L$  is the pre-decided orthogonal design.

*Step 2:* Randomly choose the  $t$  columns of  $U_c$  and randomly exchange  $m$  pairs of elements in each of these  $t$  columns. This results in a new  $U$ -type design, denoted by  $U^u = (u_1^u, \dots, u_r^u)$ , where “u” stands for “update”. Compute  $D_u = \text{Ave}(f^2)$  of  $U^u \oplus L$ . Let  $J := J + 1$ .

*Step 3:* If  $D_u - D_c < T$ , let  $U^c := U^u$  and go to Step 2; otherwise go to Step 4.

*Step 4:* If  $J < JJ$ , go to Step 2; otherwise go to Step 5.

*Step 5:* Reduce the  $T$ -value according to a pre-assigned rule as follows. Let  $I := I + 1$ . If  $I < II$ , go to Step 2, otherwise terminate the process. The current matrix  $U^c$  can be considered a local minimum solution and  $U_c \oplus L$  is the desired  $S$ -design.

There are several control parameters  $T, t, m, JJ$  and  $II$  in this procedure. The threshold  $T > 0$  can help avoid converging to a local minimum. If  $T$  is too large, the procedure closes to a purely random search and the convergence rate is very slow. If  $T$  is too small, it is difficult to “jump out” from a local minimum. The integers  $t$  and  $m$  control the number of exchanging pairs. If  $t$  and  $m$  are too large, the procedure closes to a purely random search. Otherwise we may converge to a local minimum. The integer  $II$  controls the number of iterations and  $JJ$  furthermore controls the number of searches in the neighborhood of  $U_c$  and also controls how often to reduce the  $T$ -value. Generally, the larger  $II$  and  $JJ$  are, the better the output  $S$ -design is. However, larger  $II$  and  $JJ$  require more CPU time.



In this work, we have chosen  $L_9(3^4)$ ,  $L_{18}(3^7)$ ,  $L_{27}(3^{13})$ ,  $L_{16}(4^5)$  and  $L_{25}(5^6)$  as the generating designs. These popular orthogonal arrays are available in many design books, see, for example, Dey (1985). The corresponding generating  $U$ -type designs chosen have  $r = 2-4$ . Table 1 presents the control parameters  $t, m$ , and  $T$  values in the use of TA and values of some criteria for the output  $S$ -designs. These output designs were tabulated in the appendix. Note that the  $S_9(3^8)$  design obtained in Example 1 can be further improved by using the TA.

There are some advantages of the collapsing method proposed in this paper. First, this method is simple and can be easily implemented. Second, the output  $S$ -design is block-orthogonal. Let  $\mathbf{X} = \mathbf{U} \oplus \mathbf{L}$  be the  $S$ -design, where  $\mathbf{U} \in \mathcal{U}(n, s^r)$  and  $\mathbf{L} = L_n(s^d)$ . Then  $\mathbf{X} = (\mathbf{X}_1: \mathbf{X}_2: \cdots: \mathbf{X}_r)$ , where each  $\mathbf{X}_j$  has size  $n \times d$  and the  $\mathbf{X}_j$  are orthogonal. This block-orthogonal structure of the design has been considered by many authors, for example, Yamada and Lin (1997). If we put more important factors in the same orthogonal block, their main effects can be easily estimated.

#### 4. Subdesigns

The number of columns of the output  $S$ -design by the collapsing method is a multiple of  $d$ . When the number of factors is not a multiple of  $d$ , a subdesign can be obtained as follows. Let  $\mathbf{X}$  be the output  $S$ -design of size  $n \times dr$  by the collapsing method and let the number of factors  $m \in [d(r-1) + 1, dr]$ . We then delete  $m - rd$  columns from  $\mathbf{X}$  based on the criterion given below.

**Definition 4.1.** Let  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_m)$  be an  $S$ -design. Define

$$c(i) = \sum_{j \neq i, j=1}^m f(\mathbf{x}_i, \mathbf{x}_j)^2$$

as a measure for nonorthogonality of the  $i$ th column  $\mathbf{x}_i$  against the remaining  $m - 1$  columns.

For a given  $S$ -design  $S_n(s^{rd})$ , we delete the column that has the largest  $c$ -value. Then we calculate the  $c$ -value for the remaining  $rd - 1$  columns and delete the column that has the largest  $c$ -value, and so on, until there are  $m$  columns left. Table 2 gives the order of columns being deleted from the corresponding  $S$ -design. The corresponding values of  $\text{Ave}(f^2)$ ,  $\text{Ave}|f|$  for each subdesign were evaluated. They are satisfactory for practical use (see Fang et al., 1998, Table 3 for the lengthy display).

There are very few multi-level supersaturated designs in the literature. A thorough computation on the  $\text{Ave}(f^2)$  values, for example, indicates the superiority of the newly constructed designs. As compared to the supersaturated designs obtained by Liu and Zhang (1998), for example. The  $\text{Ave}(f^2)$  value for the cases  $(S_9(3^8), S_{16}(4^{10}), S_{18}(3^{14}))$ ,

Table 2  
Columns to be deleted

<i>n, r</i>	The Order of Columns to be deleted
9,2	8 7 6
9,3	12 11 10 9 8 7 6
9,4	16 15 14 13 12 11 10 9 8 7 6
18,2	13 11 9 8 12 10
18,3	10 8 11 16 13 9 14 12 6 18 20 21 17
18,4	1 7 4 6 3 2 5 23 22 28 27 24 26 25 8 12 18 10 9 11
27,2	2 9 5 3 12 8 1 4 13 10 6 7
27,3	25 22 15 17 19 26 16 21 24 18 23 20 14 1 10 3 7 9 12 5 4 6 2 11 8
27,4	22 40 52 45 44 41 49 43 46 50 42 48 51 47 23 24 16 21 20 14 26 25 17 19 18 15 33 34 30 29 32 31 37 39 28 27 38 36
16,2	10 8 7 9
16,3	14 15 12 13 11 8 6 10 9
16,4	15 13 12 14 11 16 18 19 20 17 6 10 8 7
25,2	7 12 8 11 10
25,3	15 14 18 17 16 13 12 9 11 8 7
25,4	12 8 7 10 9 11 15 14 16 13 17 18 22 23 24 21 20

Table 3  
For  $n = 9, s = 3$ , U-type design and orthogonal array

U-type design									Orthogonal array
$r = 2$		$r = 3$			$r = 4$				
1	5	1	7	6	1	6	3	9	0 0 0 0
2	4	2	2	8	2	3	7	7	0 1 1 1
3	9	3	5	2	3	2	2	3	0 2 2 2
4	3	4	4	7	4	5	8	1	1 0 1 2
5	1	5	6	9	5	7	5	5	1 1 2 0
6	6	6	3	5	6	1	4	8	1 2 0 1
7	8	7	9	3	7	8	6	4	2 0 2 1
8	2	8	8	1	8	4	9	6	2 1 0 2
9	7	9	1	4	9	9	1	2	2 2 1 0

$S_{27}(3^{26})$ ) are (12.00, 40.00, 23.34, 37.13) for the new designs. The corresponding values for the designs in Liu and Zhang are (19.29, NA, 28.30, NA), where NA stands for “Not Available”.

**Concluding Remarks.** One referee points out that in two-level supersaturated designs, some works have considered construction methods based on permutation of rows. These include Taguchi (1987), Tang and Wu (1997) and Deng et al. (1994). Apart from the fact that they are all for two-level designs, the main difference from our work is

Table 4  
For  $n = 18, s = 3, U$ -type design and orthogonal array

U-type design									Orthogonal array
$r = 2$			$r = 3$			$r = 4$			
1	16	1	12	14	1	15	17	17	0 0 0 0 0 0 0
2	1	2	6	6	2	3	8	8	0 1 1 1 1 1 1
3	8	3	1	8	3	12	2	10	0 2 2 2 2 2 2
4	15	4	17	5	4	4	16	7	1 0 0 1 1 2 2
5	6	5	8	18	5	18	11	4	1 1 1 2 2 0 0
6	10	6	5	15	6	5	1	18	1 2 2 0 0 1 1
7	4	7	18	16	7	16	4	5	2 0 1 0 2 1 2
8	18	8	15	1	8	8	3	9	2 1 2 1 0 2 0
9	12	9	2	10	9	14	14	14	2 2 0 2 1 0 1
10	9	10	16	7	10	13	7	16	0 0 2 2 1 1 0
11	13	11	14	11	11	10	10	15	0 1 0 0 2 2 1
12	2	12	11	9	12	1	18	13	0 2 1 1 0 0 2
13	14	13	7	3	13	9	6	11	1 0 1 2 0 2 1
14	17	14	9	12	14	6	9	3	1 1 2 0 1 0 2
15	7	15	4	17	15	11	13	2	1 2 0 1 2 1 0
16	11	16	10	2	16	7	12	6	2 0 2 1 2 0 1
17	3	17	13	4	17	2	5	1	2 1 0 2 0 1 2
18	5	18	3	13	18	17	15	12	2 2 1 0 1 2 0

Table 5  
For  $n = 27, s = 3, U$ -type design and orthogonal array

U-type design									Orthogonal array
$r = 2$			$r = 3$			$r = 4$			
1	22	1	2	7	1	19	13	9	0 0 1 0 1 2 1 1 2 0 1 1 1
2	14	2	14	22	2	2	15	7	0 1 0 1 2 1 1 2 0 1 1 1 0
3	21	3	19	14	3	5	3	21	1 0 1 2 1 1 2 0 1 1 1 0 0
4	20	4	21	17	4	12	10	27	0 1 2 1 1 2 0 1 1 1 0 0 2
5	23	5	5	26	5	25	22	22	1 2 1 1 2 0 1 1 1 0 0 2 0
6	15	6	20	8	6	13	26	25	2 1 1 2 0 1 1 1 0 0 2 0 2
7	6	7	18	25	7	3	8	17	1 1 2 0 1 1 1 0 0 2 0 2 1
8	12	8	15	1	8	14	23	23	1 2 0 1 1 1 0 0 2 0 2 1 2
9	18	9	4	3	9	16	7	4	2 0 1 1 1 0 0 2 0 2 1 2 2
10	13	10	27	11	10	1	9	8	0 1 1 1 0 0 2 0 2 1 2 2 1
11	4	11	10	16	11	22	21	2	1 1 1 0 0 2 0 2 1 2 2 1 0
12	1	12	26	21	12	27	20	13	1 1 0 0 2 0 2 1 2 2 1 0 2
13	24	13	24	2	13	17	17	12	1 0 0 2 0 2 1 2 2 1 0 2 2
14	9	14	1	6	14	9	6	19	0 0 2 0 2 1 2 2 1 0 2 2 2
15	25	15	8	15	15	4	12	5	0 2 0 2 1 2 2 1 0 2 2 2 0
16	16	16	12	12	16	26	1	11	2 0 2 1 2 2 1 0 2 2 2 0 0
17	8	17	16	19	17	10	18	16	0 2 1 2 2 1 0 2 2 2 0 0 1
18	7	18	7	4	18	6	19	18	2 1 2 2 1 0 2 2 2 0 0 1 0
19	10	19	11	20	19	15	16	20	1 2 2 1 0 2 2 2 0 0 1 0 1
20	17	20	23	13	20	24	25	24	2 2 1 0 2 2 2 0 0 1 0 1 2
21	26	21	9	23	21	18	14	14	2 1 0 2 2 2 2 0 0 1 0 1 2 1
22	19	22	3	10	22	20	5	3	1 0 2 2 2 0 0 1 0 1 2 1 1
23	3	23	6	5	23	21	4	6	0 2 2 2 0 0 1 0 1 2 1 1 2
24	27	24	13	24	24	8	24	26	2 2 2 0 0 1 0 1 2 1 1 2 0
25	11	25	17	9	25	23	27	15	2 2 0 0 1 0 1 2 1 1 2 0 1
26	5	26	22	27	26	7	2	10	2 0 0 1 0 1 2 1 1 2 0 1 1
27	2	27	25	18	27	11	11	1	0 0 0 0 0 0 0 0 0 0 0 0 0

Table 6  
For  $n = 16, s = 4, U$ -type design and orthogonal array

U-type Design									Orthogonal array
$r = 2$			$r = 3$			$r = 4$			
1	3	1	8	10	1	16	2	11	
2	13	2	6	4	2	7	3	1	0 1 1 1 1
3	4	3	5	13	3	8	12	16	0 2 2 2 2
4	2	4	2	7	4	10	8	9	0 3 3 3 3
5	16	5	10	1	5	2	9	12	1 0 1 2 3
6	9	6	4	3	6	1	15	6	1 1 0 3 2
7	5	7	9	5	7	5	16	13	1 2 3 0 1
8	8	8	13	9	8	12	13	3	1 3 2 1 0
9	1	9	15	16	9	4	6	15	2 0 2 3 1
10	12	10	3	15	10	13	10	4	2 1 3 2 0
11	7	11	12	14	11	6	5	5	2 2 0 1 3
12	15	12	11	12	12	9	14	7	2 3 1 0 2
13	6	13	7	6	13	3	1	10	3 0 3 1 2
14	10	14	14	2	14	14	7	2	3 1 2 0 3
15	14	15	1	11	15	15	4	14	3 2 1 3 0
16	11	16	16	8	16	11	11	8	3 3 0 2 1

Table 7  
For  $n = 25, s = 5, U$ -type design and orthogonal array

U-type design									Orthogonal array
$r = 2$			$r = 3$			$r = 4$			
1	5	1	2	4	1	14	12	19	
2	2	2	11	21	2	1	7	9	0 1 1 2 3 4
3	20	3	18	18	3	12	20	23	0 2 2 4 1 3
4	10	4	25	15	4	13	5	6	0 3 3 1 4 2
5	13	5	9	11	5	11	24	2	0 4 4 3 2 1
6	6	6	1	2	6	5	2	7	1 0 1 1 1 1
7	21	7	3	10	7	22	9	20	1 1 2 3 4 0
8	7	8	8	22	8	3	23	1	1 2 3 0 2 4
9	11	9	4	20	9	19	1	17	1 3 4 2 0 3
10	14	10	5	24	10	23	22	13	1 4 0 4 3 2
11	1	11	6	5	11	9	19	5	2 0 2 2 2 2
12	8	12	20	3	12	8	4	22	2 1 3 4 0 1
13	24	13	19	23	13	24	16	18	2 2 4 1 3 0
14	17	14	23	14	14	10	10	24	2 3 0 3 1 4
15	15	15	15	17	15	17	11	10	2 4 1 0 4 3
16	9	16	21	16	16	6	13	21	3 0 3 3 3 3
17	18	17	24	12	17	25	15	3	3 1 4 0 1 2
18	3	18	16	9	18	16	8	14	3 2 0 2 4 1
19	23	19	10	8	19	2	18	15	3 3 1 4 2 0
20	22	20	13	6	20	20	21	11	3 4 2 1 0 4
21	12	21	14	1	21	7	17	25	4 0 4 4 4 4
22	25	22	7	19	22	18	6	4	4 1 0 1 2 3
23	19	23	22	13	23	21	3	8	4 2 1 3 0 2
24	4	24	17	25	24	15	14	12	4 3 2 0 3 1
25	16	25	12	7	25	4	25	16	4 4 3 2 1 0

that Taguchi's randomly combined design is permuted in a random fashion; the other two are permuted to minimize the specific criterion  $E(s^2)$ . Our construction method is systematic and deterministic.

## Acknowledgements

We thank the reviewers for useful comments that helped improve the paper. This work was partially supported by a Hong Kong UGC-grant and Hong Kong Baptist University. The authors thank Professor R. C. Zhang for sending us related materials and Dr. Fred J. Hickernell for his support and help. Dennis Lin is partially supported by *National Science Foundation* via Grant DMS-9704711 and *National Science Council of ROC* via Contract NSC 87-2119-M-001-007.

## Appendix

The output designs were tabulated in Tables 3–7.

## References

- Addelman, S., 1962. Orthogonal main-effect plans for analytical factorial experiments. *Technometrics* 4, 21–46.
- Booth, K.H.V., Cox, D.R., 1962. Some systematic supersaturated designs. *Technometrics* 4, 489–495.
- Deng, L.Y., Lin, D.K.J., Wang, J.N., 1994. Supersaturated designs using hadamard matrices. IBM Research Report No. RC-19470, IBM Watson Research Center.
- Deng, L.Y., Lin, D.K.J., Wang, J.N., 1999. On resolution rank criterion for supersaturated designs. *Statist. Sinica* 9, 605–610.
- Dey, A., 1985. *Orthogonal Fractional Factorial Designs*. Wiley, New York.
- Fang, K.-T., Hickernell, F.J., 1995. The uniform design and its recent development. *Bulletin of The International Statistical Institute 50th Session*, Vol. 1, pp. 339–349.
- Fang, K.T., Lin, D.K.J., Ma, C.X., 1998. Construction of multi-level supersaturated designs by the collapsing method. Technical Report MATH-191, Hong Kong Baptist University.
- Fang, K.T., Wang, Y., 1994. *Number-Theoretic Methods in Statistics*. Chapman & Hall, London.
- Li, W.W., Wu, C.F.J., 1997. Columnwise-pairwise algorithms with applications to the construction of supersaturated designs. *Technometrics* 39, 171–179.
- Lin, D.K.J., 1991. Systematic supersaturated designs. Technical Report, No. 276, College of Business Administration, University of Tennessee.
- Lin, D.K.J., 1993. A new class of supersaturated designs. *Technometrics* 35, 28–31.
- Lin, D.K.J., 1995. Generating systematic supersaturated designs. *Technometrics* 37, 213–225.
- Liu, M.Q., Zhang, R.C., 1998. A criterion and an algorithm on constructing 3-level factorial supersaturated designs. Technical Report, Nankai University.
- Niederreiter, H., 1992. *Random Number Generation and Quasi-Monte Carlo Methods*, CBMC-NSF. SIAM, Philadelphia.
- Nguyen, N.-K., 1996. An algorithmic approach to constructing supersaturated designs. *Technometrics* 38, 69–73.
- Satterthwaite, F., 1959. Random balance experimentation (with discussion). *Technometrics* 1, 111–137.

- Tang, B., Wu, C.F.J., 1997. A method for constructing super-saturated designs and its  $E_s^2$  optimality. *Canadian J. Statist.* 25, 191–201.
- Taguchi, G., 1987. *System of Experimental Designs*. Kraus International Publications, New York.
- Winker, P., Fang, K.-T., 1996. Optimal U-type design. In: H. Niederreiter et al. (Eds.), *Proceedings of MC&QMC 96, Lecture Notes in Statistics*. Springer, New York, 1997.
- Wu, C.F.J., 1993. Construction of supersaturated designs through partially aliased interactions. *Biometrika* 80, 661–669.
- Yamada, S., Lin, D.K.J., 1997. Supersaturated designs including an orthogonal base. *Canadian J. Statist.* 25, 203–213.