# Design for Computer Experiments with Qualitative and Quantitative Factors

Xinwei Deng, Ying Hung and C. Devon Lin

*Virginia Tech, Rutgers University and Queen's University*

*Abstract:* We introduce a new class of designs, called marginally coupled designs, for computer experiments with both qualitative and quantitative variables. The proposed design maintains an economic run size with attractive space-filling properties. The design points for quantitative factors forms a Latin hypercube design. In addition, for each level of any qualitative factor of a marginally coupled design, the corresponding design points for quantitative factors form a small Latin hypercube design. Existence of the proposed designs is studied. Constructions are provided for various types of designs with qualitative factors.

*Key words and phrases:* Difference scheme; Fold-over design; Latin square; Mixed level; Orthogonal array; Resolvable orthogonal array.

## 1. Introduction

Computer experiments refer to the study of real systems using complex simulation models. They have been widely used as alternatives to physical experiments, especially for studying complex systems. Computational expense of computer experiments often prohibits the naive approach of running the experiment over a dense grid of input configurations. Therefore, an efficient design of experiments is important in the study of computer experiments.

A widely used design in computer experiments is a Latin hypercube design (McKay et al., 1979; Santner et al., 2003; Fang et al., 2010). The popularity of this class of designs is due to their appealing feature that when projected onto any one dimension, the equally spaced design points ensure that each of the factors has all portions of its range represented. Different variants of Latin hypercube designs have been developed in the literature, including orthogonal Latin hypercube designs (Owen, 1994; Ye, 1998; Steinberg and Lin, 2006; Bingham et al., 2009; Lin et al., 2009; Pang et al., 2009; Sun et al., 2009; Lin et al., 2010), maximin Latin hypercube designs (Morris and Mitchell, 1995; Joseph and Hung, 2008; Moon et al., 2011), nested Latin hypercube designs (Qian, 2009; He and Qian, 2011), sliced Latin hypercube designs (Qian and Wu, 2009; Qian, 2012), among many others. These designs are primarily used for computer experiments with only quantitative factors. However, in many scientific problems, qualitative factors occur frequently and play important roles in the studies (Qian et al., 2008; Han et al., 2009; Hung et al., 2009). Particularly, in some applications, a large number or proportion of factors can be qualitative factors. How to efficiently design experiments with quantitative and qualitative factors is an important but unresolved issue in computer experiments. The objective of this research is to propose and construct a new class of designs, *marginally coupled designs*, for computer experiments involving both quantitative and qualitative factors.

Throughout, let $D = (D_1, D_2)$ be a design with $q$ qualitative factors and $p$ quantitative factors, where $D_1$ and $D_2$ are sub-designs for qualitative and quantitative factors, respectively. A design $D$ is called a marginally coupled design if $D_2$ is a Latin hypercube design and the rows in $D_2$ corresponding to each level of any factor in $D_1$ form a small Latin hypercube design. In this work, we focus on the situation of $D_1$ being orthogonal arrays (Hedayat et al., 1999).

For constructing designs of computer experiments with quantitative and qualitative factors, one alternative method can be using the sliced Latin hypercube designs (Qian, 2012) for $D_2$, where the design points in every slice of $D_2$ corresponds to one level combination of the qualitative factors. The run size of such designs increases dramatically with the number of qualitative factors in that there are a large number of level combinations. Thus, they are only suitable for experiments with a very small number of qualitative factors. In contrast, the proposed design can accommodate a large number of qualitative factors with an economic run size. In addition, it has the following attractive space-filling properties: (1) for each level of any qualitative factor, the corresponding design points of quantitative factors achieve maximum uniformity in any one-dimensional projection, and (2) the design points of quantitative factors possess maximum uniformity in any one-dimensional projection.

The remainder of the paper is organized as follows. Section 2 presents notation, definitions and an example of marginally coupled designs. Section 3 provides some existence results on the proposed designs. Several construction methods for such designs are given in Section 4. Section 5 concludes the paper with a discussion.

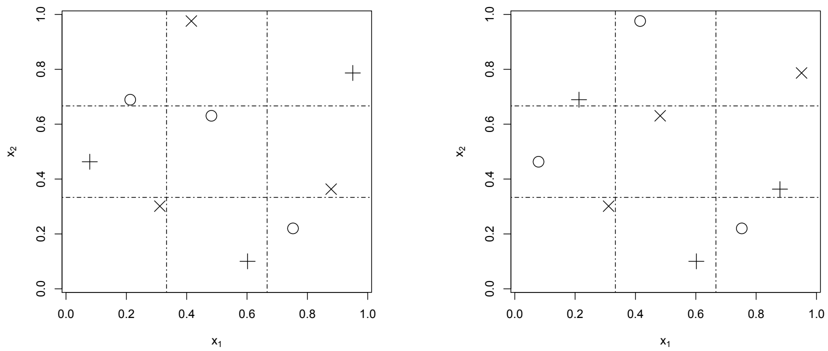
## 2. Notation, Definitions and an Example

An orthogonal array $A$ of strength $t$, denoted by $\mathrm{OA}(n, s_1 \cdots s_q, t)$, is an $n \times q$ matrix of which the $i$th column has $s_i$ levels $0, \ldots, s_i - 1$ and for every $n \times t$ submatrix of $A$, each of all possible level combinations appears equally often (Hedayat et al., 1999). If not all $s_i$'s are equal, an orthogonal array is mixed. We use $\mathrm{OA}(n, s_1^{q_1} \cdots s_k^{q_k}, t)$ to represent an orthogonal array in which the first $q_1$ columns have $s_1$ levels, the next $q_2$ columns have $s_2$ levels, and so on, and the last $q_k$ columns have $s_k$ levels. An $\mathrm{OA}(n, s_1^{q_1} \cdots s_k^{q_k}, 2)$ is said to be $(\alpha_1 \times \alpha_2 \times \cdots \times \alpha_k)$-resolvable if for $1 \leq j \leq k$, its rows can be partitioned into $n/(\alpha_j s_j)$ subarrays $A_1, \ldots, A_{n/(\alpha_j s_j)}$ of $\alpha_j s_j$ rows each such that each of $A_1, \ldots, A_{n/(\alpha_j s_j)}$ is an $\mathrm{OA}(\alpha_j s_j, s_1^{q_1} \cdots s_k^{q_k}, 1)$. Note that $\alpha_j s_j$'s are identical for all $j$'s. If all $s_j$'s are equal and $\alpha_1 = \cdots = \alpha_k = \alpha$ then an $(\alpha_1 \times \alpha_2 \times \cdots \times \alpha_k)$-resolvable orthogonal array reduces to an $\alpha$-resolvable orthogonal array. If $\alpha = 1$, the orthogonal array is called completely resolvable.

A Latin hypercube of $n$ runs for $p$ factors is represented by an $n \times p$ matrix of which each column is a random permutation of $n$ equally-spaced levels. For convenience, throughout we assume that the $n$ levels are taken to be $-(n-1)/2, -(n-3)/2, \ldots, (n-3)/2, (n-1)/2$. In Qian (2012), a Latin hypercube $L$ of $n = rm$ runs is called a sliced Latin hypercube of $r$ slices if $L$ can be expressed as $L = (L_1^{\mathrm{T}}, \ldots, L_r^{\mathrm{T}})^{\mathrm{T}}$ where $m$ levels in each column of $L_i$ have exactly one level from each of the $m$ equally-spaced intervals $\{[-n/2 + (j-1)r, -n/2 + jr] : 1 \leq j \leq m\}$. Given an $n \times p$ Latin hypercube $L = (l_{ij})$, a Latin hypercube design $X = (x_{ij})$ is generated via

$$x_{ij} = \frac{l_{ij} + (n-1)/2 + u_{ij}}{n}, \quad 1 \leq i \leq n, 1 \leq j \leq p, \tag{1}$$

where $u_{ij}$'s are independent random numbers from $[0, 1)$. We say $L$ is a Latin hypercube corresponding to $X$. Note that a $D_2$ in a marginally coupled design $D = (D_1, D_2)$ is a sliced Latin hypercube design with respect to each column of $D_1$. An example of a marginally coupled design of 9 runs for two quantitative factors and two qualitative factors is given in Example 1.




(a) (b)

Figure 1: Scatter plots of $x_1$ versus $x_2$ in Example 1, where rows of $D_2$ corresponding to levels 0,1,2 of $z_i$ are marked by $\times$, $\circ$, and $+$: (a) the levels of $z_1$; (b) the levels of $z_2$.

**Example 1.** *Design $D = (D_1, D_2)$ in Table 1 is a marginally coupled design of 9 runs for two quantitative variables ($x_1$, $x_2$) and two qualitative factors ($z_1$, $z_2$) each at three levels. Figure 1 displays the scatter plots of $x_1$ versus $x_2$. Rows of $D_2$ corresponding to levels 0,1,2 of $z_1$ or $z_2$ are represented by $\times$, $\circ$, and $+$. Projected onto $x_1$ or $x_2$, three points represented by $\times$ or $\circ$ or $+$ are exactly located in each of three intervals [0,1/3), [1/3,2/3), [2/3,1).*

Table 1. A marginally coupled design $D = (D_1, D_2)$

| $D_1$ | | $D_2$ | |
|---|---|---|---|
| 0 | 0 | 0.311 | 0.301 |
| 0 | 1 | 0.415 | 0.975 |
| 0 | 2 | 0.878 | 0.363 |
| 1 | 0 | 0.481 | 0.630 |
| 1 | 1 | 0.752 | 0.220 |
| 1 | 2 | 0.212 | 0.689 |
| 2 | 0 | 0.950 | 0.786 |
| 2 | 1 | 0.078 | 0.463 |
| 2 | 2 | 0.601 | 0.100 |

## 3. Existence of Marginally Coupled Designs

This section provides some results on the existence of marginally coupled designs. Recall the definition of marginally coupled designs. A marginally coupled design of $n$ runs has two sub-designs, $D_1$ and $D_2$, which are for $q$ qualitative factors and $p$ quantitative factors, respectively, where $q > 0$ and $p > 0$. For a given $n \times q$ design $D_1$, we say a marginally coupled design exists if there exists an $n \times p$ design $D_2$ with $p > 0$ such that $D = (D_1, D_2)$ is a marginally coupled design. Propositions 1 and 2 establish the necessary and sufficient condition of the existence of marginally coupled designs when $D_1$ is an $\mathrm{OA}(n, s^q, 2)$ and an $\mathrm{OA}(n, s_1^{q_1} s_2^{q_2}, 2)$ with $s_1 = \alpha_2 s_2$, respectively.

**Proposition 1.** *Given $D_1 = OA(n, s^q, 2)$, a marginally coupled design exists if and only if $D_1$ is a completely resolvable orthogonal array.*

*Proof.* We first show if $D_1$ is a completely resolvable orthogonal array, a marginally coupled design exists. A completely resolvable orthogonal array can be expressed as $D_1 = (A_1^T, \ldots, A_m^T)^T$ such that each $A_i$ is an $\mathrm{OA}(s, s^q, 1)$ for $1 \le i \le m = n/s$. Let $d = (d_j)$ be the column vector of length $n$ with $d_j = (j - 1 + u_j)/n$ where $u_j$'s are independent numbers from $[0, 1)$, $1 \le j \le n$. Now construct an $n \times p$ design $D_2$ in the following way. Write $d = (b_1^T, \ldots, b_m^T)^T$ where $b_i$ is the $\{(i-1)s + 1, \ldots, is\}$th entries of $d$, $1 \le i \le m = n/s$. Obtain each column of $D_2$ by randomly permuting $b_i$'s of $d$ and/or randomly permuting the entries in one or more of $b_i$'s. Then $(D_1, D_2)$ forms a marginally coupled design by definition. Next, we show that if a marginally coupled design exists, $D_1$ must be a completely resolvable orthogonal array. Suppose that $D_2$ has $p$ columns and the $j$th column is denoted by $D_2^{(j)}$. Then for each $D_2^{(j)}$, $1 \le j \le p$, and for $0 \le i \le m-1$, the rows of $D_1$ corresponding to the $s$ elements of $D_2^{(j)}$ in the interval $[i/m, (i+1)/m)$ must form an $\mathrm{OA}(s, s^q, 1)$. This is because, otherwise there will be a level, say $k$, in a column $w$ of $D_1$ such that the entries in $D_2^{(j)}$ corresponding to the level $k$ contain more than one element from $[i/m, (i+1)/m)$. In other words, for the column $w$ of $D_1$, $D_2$ is not a sliced Latin hypercube design. This completes the proof.

**Proposition 2.** *Given $D_1 = OA(n, s_1^{q_1} s_2^{q_2}, 2)$ with $s_1 = \alpha_2 s_2$, a marginally coupled design exists if and only if $D_1$ is a $(1 \times \alpha_2)$-resolvable orthogonal array that can be expressed as*

$$\begin{pmatrix} A_{11} & A_{12} \\ \vdots & \vdots \\ A_{m1} & A_{m2} \end{pmatrix} \tag{2}$$

*such that* $(A_{i1}, A_{i2})$ *is an* $OA(s_1, s_1^{q_1}s_2^{q_2}, 1)$*, where* $m = n/s_1$*, and for* $1 \le i \le m$*, the* $A_{i2}$ *is completely resolvable.*

*Proof.* We first show the necessary part, that is, for a $D_1$ in the proposition, a marginally coupled design exists. Let $d = (d_j)$ be the column vector of length $n$ with $d_j = (j-1+u_j)/n$ where $u_j$'s are independent numbers from $[0,1)$, $1 \le j \le n$. Now construct an $n \times p$ design $D_2$ in the following way. Write $d = (b_1^T, \ldots, b_m^T)^T$ where $b_i$ is the $\{(i-1)s+1, \ldots, is\}$th entries of $d$, $1 \le i \le m = n/s$. Set each column of $D_2$ to a column obtained by randomly permuting $b_i$'s of $d$ and/or randomly permuting the entries in one or more of $b_i$'s. Then $(D_1, D_2)$ forms a marginally coupled design by definition. Now we show the sufficient part. That is, we shall show that (a) the first $q_1$ columns in $D_1$ are a completely resolvable orthogonal array, and (b) the rows of each $A_{i2}$ in (2) can be partitioned into $\alpha_2$ subarrays $B_1, \ldots, B_{\alpha_2}$ of $s_2$ rows each such that each of $B_1, \ldots, B_{\alpha_2}$ is an $\mathrm{OA}(s_2, s_2^{q_2}, 1)$. Part (a) uses the arguments from the proof of Proposition 1 with $q = q_1$ and $s = s_1$. To complete the proof, we now validate part (b). Suppose $D_2$ has $p$ columns and the $j$th column is denoted by $D_2^{(j)}$. Then for each $D_2^{(j)}$, $1 \le j \le p$, and $0 \le i \le m-1$, the rows of $D_1$ corresponding to the $s_1$ elements of $D_2^{(j)}$ in the interval $[i/m, (i+1)/m)$ must be $A_{h1}$ for an $h$ in $\{1, \ldots, m\}$. Now for the given $i$ and $h$ and for $g = 0, \ldots, \alpha_2 - 1$, the rows of $A_{h2}$ corresponding to the $s_2$ elements of $D_2^{(j)}$ in the interval $[i/m + g/(m\alpha_2), i/m + (g+1)/(m\alpha_2))$ must form an $\mathrm{OA}(s_2, s_2^{q_2}, 1)$. This is because, otherwise there will be a level, say $k$, in a column $w$ of $(A_{12}^T, \ldots, A_{m2}^T)^T$ such that the entries in $D_2^{(j)}$ corresponding to the level $k$ contain more than one element from $[i/m + g/(m\alpha_2), i/m + (g+1)/(m\alpha_2))$. In other words, for the column $w$ in $(A_{12}^T, \ldots, A_{m2}^T)^T$, $D_2$ is not a sliced Latin hypercube design. We thus show part (b). This completes the proof.

The following lemma from Suen (1989) shows that the maximum number of columns in an $n/(sr)$-resolvable $s$-level orthogonal array of $n$ runs is $(n-r)/(s-1)$. For a completely resolvable $s$-level orthogonal array, we have $r = n/s$, and thus the maximum number of columns in a completely resolvable $s$-level orthogonal array is $n/s$, which leads to Corollary 1.

**Lemma 1.** *If a resolvable* $OA(n, s^q, 2)$ *can be partitioned into* $r$ $OA(n/r, s^q, 1)$*'s, then* $q \le (n-r)/(s-1)$.

**Corollary 1.** *Let $q*$ be the maximum value of $q$ such that a marginally coupled design $D = (D_1, D_2)$ with $D_1 = OA(n, s^q, 2)$ exists. We have $q^* \leq n/s$.*

Corollary 2 provides a result on the maximum number of columns in a two-level orthogonal array of $n$ runs where $n$ is a multiple of 4 for which a marginally coupled design exists.

**Corollary 2.** *Let $q*$ be the maximum value of $q$ such that a marginally coupled design $D = (D_1, D_2)$ with $D_1 = OA(4\lambda, 2^q, 2)$ exists, where $\lambda$ is an integer such that a Hadamard matrix of order $2\lambda$ exists. We have $q^* = 2\lambda$.*

*Proof.* By Proposition 1, $D_1 = \mathrm{OA}(4\lambda, 2^q, 2)$ is completely resolvable. Such $D_1$'s are fold-over designs. When the two levels are represented by 1 and -1, a fold-over design can be represented by $(A^T, -A^T)^T$ for some matrix $A$. Let $A$ be a Hadamard matrix of order $2\lambda$ (Hedayat et al., 1999). Then the maximum number of columns in a fold-over orthogonal array of $4\lambda$ runs is $2\lambda$. This completes the proof.

## 4. Construction of Marginally Coupled Designs

We develop several procedures to construct marginally coupled designs $D = (D_1, D_2)$ when $D_1$'s are $s$-level orthogonal arrays of $s^2$ and $\lambda s^2$ runs, mixed orthogonal arrays, and two-level orthogonal arrays, respectively.

### 4.1 Construction for $D_1$ being $s$-level orthogonal arrays of $s^2$ runs

Suppose an $\mathrm{OA}(s^2, s^k, 2)$, say $A$, is available and $D_1$ for qualitative factors is obtained by randomly taking $q$ columns from $A$. Let $A \backslash D_1$ be the complement of $D_1$ within $A$. Construction 1 below is based on the idea in Tang (1993), which is originally proposed for constructing orthogonal array-based Latin hypercubes.

*Construction 1.* Obtain a design, $B$, by randomly taking $p$ columns from $A \backslash D_1$, where $q + p \leq k$. For each column of $B$, replace the $s$ positions having level $i - 1$ by a random permutation of $\{(i-1)s+1\} - (s^2+1)/2, \ldots, \{(i-1)s+s\} - (s^2+1)/2$, for $1 \leq i \leq s$. Denote the resulting design by $L$ and obtain a Latin hypercube design $D_2$ based on $L$ via (1).

**Proposition 3.** *Let $D_1 = OA(s^2, s^q, 2)$. Design $D = (D_1, D_2)$ is a marginally coupled design, where $D_2$ is obtained by Construction 1.*

Proposition 3 can be readily verified by noting the two-dimensional projection property of an orthogonal array. This property implies that for each column in $D_2$, the $s$ points corresponding to each level of each column in $D_1$ are evenly distributed in the $s$ equally-spaced intervals.

**Example 2.** *Taking $s$ in Construction 1 to be $3, 4, 5, 7, 8, 9$, we have $OA(9, 3^4, 2)$, $OA(16, 4^5, 2)$, $OA(25, 5^6, 2)$, $OA(49, 7^8, 2)$, $OA(64, 8^9, 2)$ and $OA(81, 9^{10}, 2)$, which provide designs of computer experiments of $s^2$ runs for $q$ qualitative variables and $p$ quantitative variables, where $p + q \leq s + 1$.*

### 4.2 Constructions for $D_1$ being $s$-level orthogonal arrays of $\lambda s^2$ runs

This section introduces a method for constructing marginally coupled designs with $D_1$ being $s$-level orthogonal arrays of $n = \lambda s^2$ runs. The approach, Construction 2 below, uses mixed orthogonal arrays $\mathrm{OA}(\lambda s^2, s^k(\lambda s), 2)$. Suppose an $\mathrm{OA}(\lambda s^2, s^k(\lambda s), 2)$, denoted by $A$, is available and $D_1$ for qualitative factors is obtained by randomly taking $q$ columns from the first $k$ columns of $A$, where $\lambda$ is a positive integer and $q \leq k$.

*Construction 2.* Denote the last column of $A$ by $a$. For $1 \leq j \leq p$, let $\pi_j$ be a random permutation of $\{0, \ldots, \lambda s - 1\}$ and $\pi_j(i)$ be the $i$th entry of $\pi_j$. Replace the $s$ positions having level $\pi_j(i)$ in $a$ by a random permutation of $\{(i-1)s + 1\} - (\lambda s^2 + 1)/2, \ldots, \{(i-1)s + s\} - (\lambda s^2 + 1)/2$, for $1 \leq i \leq \lambda s$. Denote the resulting design by $L$ and obtain a Latin hypercube design $D_2$ based on $L$ via (1).

**Proposition 4.** *Let $D_1 = OA(\lambda s^2, s^q, 2)$. Design $D = (D_1, D_2)$ with $D_2$ in Construction 2 is a marginally coupled design.*

Analogous to Proposition 3, Proposition 4 is the consequence of the two-dimensional projection property of an orthogonal array. The website Sloane (2014) lists $\mathrm{OA}(\lambda s^2, s^k(\lambda s), 2)$ with $\lambda s^2 \leq 100$. For these orthogonal arrays, we have $k = \lambda s$ for $s$ being a prime or prime power except the cases in Table 2 and the cases having $(s = 2,\ \text{odd } \lambda, k = 2)$.

Table 2. The $s$, $\lambda$ and $k$ such that an $\mathrm{OA}(\lambda s^2, s^k(\lambda s), 2)$ exists and $k \neq \lambda s$, $s \geq 3$, $\lambda s^2 \leq 100$

| | | | | | | |
|---|---|---|---|---|---|---|
| $s$ | 3 | 3 | 5 | 4 | 4 | 3 |
| $\lambda$ | 5 | 7 | 3 | 5 | 6 | 11 |
| $k$ | 9 | 11 | 7 | 8 | 16 | 11 |

### 4.3 Construction for $D_1$ being mixed orthogonal arrays

Mixed orthogonal arrays were constructed via small mixed orthogonal arrays and difference schemes in Wang and Wu (1991), Hedayat et al. (1992), and Dey and Midha (1996), among others. A general formulation is provided in Theorem 9.15 in Hedayat et al. (1999). A slightly different version of this formulation is stated in Lemma 2.

**Lemma 2.** *Let $B = (B_1 \cdots B_v)$ be an $OA(n, s_1^{k_1} \cdots s_v^{k_v}, 2)$ where $B_j$ is the orthogonal array for $k_j$ factors with $s_j$ levels. If, for some $u$, there are difference schemes $D(u, c_j, s_j)$ (of strength 2), denoted by $D(j)$, for $1 \leq j \leq v$, then the design*

$$A = \Big( \; D(1) \odot B_1, \quad \cdots \quad , D(v) \odot B_v \; \Big), \tag{3}$$

*is an $OA(nu, s_1^{k_1 c_1} \cdots s_v^{k_v c_v}, 2)$, where $X \odot Y = (x_{ij} * Y)$ stands for the Kronecker product of a $u \times c$ matrix $X = (x_{ij})$ and an $n \times k$ matrix $Y = (y_{rs})$ with $x_{ij} * Y$ being the matrix with entries $x_{ij} * y_{rs}$ and the binary operation $*$ representing addition.*

Let $M$ be the Latin hypercube corresponding to $D_2$ in a marginally coupled design $D = (D_1, D_2)$. For convenience and notational simplicity, hereafter we call $M$ a *marginally sliced Latin hypercube* for $D_1$. The following method is proposed to construct marginally coupled designs when $D_1 = A$ in (3).

*Construction 3.* Let $C = (c_{ij})$ be a $u \times f$ matrix with $c_{ij} = \pm 1$, $H$ be a $u \times (pf)$ Latin hypercube and $M$ be a marginally sliced Latin hypercube for $B = \mathrm{OA}(n, s_1^{k_1} \cdots s_v^{k_v}, 2)$ in Lemma 2. Obtain an $(nu) \times (pf)$ matrix $L = C \otimes M + nH \otimes 1_n$, where $\otimes$ represents the Kronecker product and $1_n$ is a column of all 1's, and further obtain a Latin hypercube design $D_2$ based on $L$ via (1).

Construction 3 provides a way to construct marginally coupled designs for mixed orthogonal arrays of the form (3). A precise result is given in Proposition 5. Lemmas 3 and 4 are used to show Proposition 5. The proof of both lemmas are straightforward and thus omitted.

**Lemma 3.** *If $M$ is a marginally sliced Latin hypercube for an $OA(n, s_1^{k_1} \cdots s_v^{k_v}, 2)$, so is $-M$.*

**Lemma 4.** *If $M$ is a marginally sliced Latin hypercube for $B = (B_1, \cdots, B_v)$ in Lemma 2, $M$ is a marginally sliced Latin hypercube for $(a_1 \odot B_1, \cdots, a_v \odot B_v)$ for all $a_i \in \{0, \ldots, s_i - 1\}$ and $1 \leq i \leq v$.*

**Proposition 5.** *Let $D_1$ be $A$ in (3). Design $D = (D_1, D_2)$ with $D_2$ in Construction 3 is a marginally coupled design.*

*Proof.* To prove Proposition 5, we need to show that $L$ in Construction 3 is a marginally sliced Latin hypercube for $A$ in (3). First, $L$ is a Latin hypercube by verifying that each column of $L$ has levels $-(nu-1)/2, \ldots, (nu-1)/2$. Second, we show that for each column of $L$, the entries corresponding to a level in any column of $A$ with $s_i$ levels have exactly one value from each of the $nu/s_i$ intervals

$$\Phi_i = [\{-nu/2 + (j-1)s_i, -nu/2 + js_i\} : 1 \leq j \leq nu/s_i]. \tag{4}$$

Without loss of generality, consider the first column $l_1$ of $L$. Note that $l_1 = c_1 \otimes m_1 + nh_1 \otimes 1_n$, where $c_1$, $m_1$ and $h_1$ are the first column of $C$, $M$ and $H$, respectively. Now consider any column $a$ of $A$ and suppose $a = d \odot b$ where $d$ is a column from $D(i)$ in (3) and $b$ is a column from $B_i$ in (3). Let $c_{1j}$ and $d_j$ be the $j$th entry of $c_1$ and $d$, respectively. By Lemmas 3 and 4, for the column $c_{1j} \otimes m_1$, the entries corresponding to a level in $d_j \odot b$ with $s_i$ levels have exactly one value from each of the $n/s_i$ intervals $\phi_i = [\{-n/2 + (j-1)s_i, -n/2 + js_i\} : 1 \leq j \leq n/s_i]$. Let $\omega = \{-(u-1)/2, -(u-3)/2, \ldots, (u-3)/2, (u-1)/2\}$. Then for $l_1$, the entries corresponding to a level in $a$ with $s_i$ levels have exactly one value from each of the intervals

$$\{n\omega_j \phi_i : 1 \leq j \leq u\}, \tag{5}$$

where $n\omega_j\phi_i$ represents the intervals whose lower bounds and upper bounds are obtained by multiplying the lower bound and upper bound of each interval in $\phi_i$ by $n\omega_j$. It is straightforward to verify that the intervals in (5) are identical to $\Phi_i$ in (4) and thus we complete the proof.

**Example 3.** *Consider a design* $A = OA(32, 2^8 4^2, 2)$ *constructed as in Lemma 2 using*

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \\ 0 & 1 & 3 \\ 1 & 0 & 3 \end{pmatrix}, D(1) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \textit{ and } D(2) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 2 \\ 0 & 3 \end{pmatrix}.$$

*A marginally sliced Latin hypercube for* $B$ *is*

$$M = \frac{1}{2}\begin{pmatrix} -5 & 1 & -7 & -1 & 3 \\ 3 & -3 & 7 & 7 & -3 \\ 1 & -1 & 5 & 5 & -1 \\ -7 & 3 & -5 & -3 & 1 \\ -1 & 5 & -1 & -5 & 5 \\ 7 & -7 & 1 & 1 & -7 \\ 5 & -5 & 3 & 3 & -5 \\ -3 & 7 & -3 & -7 & 7 \end{pmatrix}. \textit{ By choosing } C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix},$$

*and any* $4 \times 20$ *Latin hypercube* $H$*, the proposed procedure above provides a* $32 \times 20$ *marginally sliced Latin hypercube for* $D_1 = A$*.*

### 4.4 Construction for $D_1$ being unreplicated or replicated $s$-level orthogonal arrays

We now introduce a construction for marginally coupled designs of $n$ runs when $D_1$'s are unreplicated or replicated $s$-level orthogonal arrays and $D_2$'s have $s$ slices with respect to each column of $D_1$, where $n$ is a multiple of $s^2$. The advantage of this construction method over Construction 1 in Section 4.1 is, Construction 1 works for $p \leq s + 1 - q$ while this method works for any value of $p$. The gain from Construction 1 is that columns in $D_2$ have two-dimensional stratification. This method is different from those in Section 4.2 in that $D_1$'s provided by the latter are not replicated.

Let $W_1, \ldots, W_q$ be mutually orthogonal Latin squares (Hedayat et al., 1999) of order $s$ with symbols $0, \ldots, s-1$. Two Latin squares are called orthogonal if the following holds: when one Latin square is superimposed upon the other, every ordered pair of variables occurs exactly once in the resulting square. Suppose we wish to have $p$ columns in $D_2$, the following procedure is proposed.

I. For $1 \leq i \leq q$, let $W_i(j, k)$ be the $(j, k)$th entry of $W_i$. For $1 \leq r \leq s$, let $(\zeta_{(r-1)s+1}, \ldots, \zeta_{rs})$ be the $(j, k)$'s such that $W_1(j, k) = r-1$, and let $\zeta = (\zeta_1, \ldots, \zeta_{s^2})$. Obtain an $s^2 \times q$ orthogonal array $H$ by letting its $(t,i)$th entry being $W_i(\zeta_t)$. If $n/s^2$ is greater than 1, for each row of $H$, add $n/s^2-1$ replications of that row and denote the resulting design by $D_1$. Otherwise, let $D_1 = H$.

II. For $1 \leq i,j \leq s$, define $\xi_{ij} = \{(t-1)s^2 + (i-1)s + j - (n+1)/2 : 1 \leq t \leq n/s^2\}$. Obtain an $n \times p$ array $L$ whose $k$th column is constructed in the following manner. Let $\alpha = (\alpha_1, \ldots, \alpha_s)$ and $\beta = (\beta_1, \ldots, \beta_s)$ be two independent random permutations of $\{1, \ldots, s\}$, and $\tilde{\xi}_{\alpha_i\beta_j}$ be a random permutation of the elements in $\xi_{\alpha_i\beta_j}$. For $1 \leq k \leq p$, the $k$th column of $L$ is obtained by stacking $\tilde{\xi}_{\alpha_i\beta_j}$'s row by row where $\alpha_i\beta_j$'s are in the following order $\alpha_1\beta_1, \ldots, \alpha_s\beta_1, \alpha_1\beta_2, \ldots, \alpha_s\beta_2, \ldots, \alpha_1\beta_s, \ldots, \alpha_s\beta_s$. Obtain $D_2$ based on $L$ via (1).

**Proposition 6.** *Let $q$ be the integer such that there exist $q$ mutually orthogonal Latin squares of order $s$. For $D_1$ and $D_2$ constructed above, we have that*

*(i) design $D_1$ is an unreplicated $OA(s^2, s^q, 2)$ when $n = s^2$ or a replicated $OA(s^2, s^q, 2)$ of $\lambda$ replicates when $n = \lambda s^2$ for an integer $\lambda > 1$, and*

*(ii) $D = (D_1, D_2)$ is a marginally coupled design.*

We sketch a proof. Part (i) of Proposition 6 follows from the definition of mutually orthogonal Latin squares. For part (ii), note that for $1 \leq j \leq s$, $\xi_{1j}, \ldots, \xi_{sj}$ forms a slice of a Latin hypercube of $n$ runs and $s$ slices. Part (ii) follows because for each column of $L$, the row entries corresponding to each level in each column of $D_1$ are $\xi_{1j}, \ldots, \xi_{sj}$ for certain $j$.

**Example 4.** *Consider $n = 16, s = 4, p = 9$. There are three mutually orthogonal Latin squares. A marginally coupled design given by the above procedure for $D_1$ being an unreplicated $OA(16, 4^3, 2)$ is given in Table 3. Note that we choose $p = 9$ in this example. However, the approach works for any value of $p$.*

### 4.5 Construction for $D_1$ being two-level orthogonal arrays

This section presents a method for constructing marginally coupled designs for $D_1$ being two-level orthogonal arrays. It extends the method in Lin et al. (2010) which introduced a general approach for constructing designs for computer experiments. For convenience, we use $-1, 1$ to represent two levels in an orthogonal array in this section. Let $A = (a_{ij})$ be an $n_1 \times m_1$ matrix with $a_{ij} = \pm 1$, $B = (b_{ij})$ be an $n_2 \times m_2$ Latin hypercube, $C = (c_{ij})$ be an

Table 3. Designs $D_1$ and $D_2$ in Example 4

| $D_1$ | | | $D_2$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0.684 | 0.422 | 0.854 | 0.940 | 0.483 | 0.822 | 0.123 | 0.379 | 0.035 |
| 0 | 3 | 2 | 0.913 | 0.682 | 0.316 | 0.226 | 0.698 | 0.624 | 0.851 | 0.671 | 0.274 |
| 0 | 1 | 3 | 0.384 | 0.914 | 0.602 | 0.710 | 0.224 | 0.371 | 0.576 | 0.876 | 0.768 |
| 0 | 2 | 1 | 0.164 | 0.125 | 0.109 | 0.466 | 0.977 | 0.099 | 0.363 | 0.161 | 0.527 |
| 1 | 1 | 1 | 0.617 | 0.349 | 0.906 | 0.903 | 0.397 | 0.807 | 0.149 | 0.316 | 0.113 |
| 1 | 2 | 3 | 0.831 | 0.620 | 0.433 | 0.183 | 0.687 | 0.519 | 0.910 | 0.608 | 0.341 |
| 1 | 0 | 2 | 0.327 | 0.824 | 0.658 | 0.641 | 0.136 | 0.310 | 0.670 | 0.823 | 0.851 |
| 1 | 3 | 0 | 0.083 | 0.079 | 0.158 | 0.425 | 0.892 | 0.019 | 0.418 | 0.116 | 0.573 |
| 2 | 2 | 2 | 0.550 | 0.481 | 0.990 | 0.796 | 0.290 | 0.889 | 0.026 | 0.310 | 0.167 |
| 2 | 1 | 0 | 0.784 | 0.701 | 0.470 | 0.006 | 0.503 | 0.663 | 0.796 | 0.503 | 0.399 |
| 2 | 3 | 1 | 0.307 | 0.962 | 0.700 | 0.556 | 0.041 | 0.403 | 0.557 | 0.793 | 0.896 |
| 2 | 0 | 3 | 0.004 | 0.191 | 0.189 | 0.270 | 0.765 | 0.166 | 0.267 | 0.032 | 0.646 |
| 3 | 3 | 3 | 0.703 | 0.294 | 0.767 | 0.863 | 0.322 | 0.997 | 0.210 | 0.454 | 0.238 |
| 3 | 0 | 1 | 0.999 | 0.539 | 0.304 | 0.101 | 0.620 | 0.741 | 0.997 | 0.699 | 0.465 |
| 3 | 2 | 0 | 0.445 | 0.791 | 0.556 | 0.616 | 0.104 | 0.473 | 0.705 | 0.996 | 0.976 |
| 3 | 1 | 2 | 0.196 | 0.040 | 0.035 | 0.325 | 0.854 | 0.244 | 0.473 | 0.204 | 0.713 |

$n_1 \times m_1$ Latin hypercube, and $H = (h_{ij})$ be an $n_2 \times m_2$ matrix with $h_{ij} = \pm 1$. Lin et al. (2010) consider the design

$$L = A \otimes B + n_2 C \otimes H. \tag{6}$$

Lemma 5 from Lin et al. (2010) provides the conditions for $L$ in (6) to be a Latin hypercube.

**Lemma 5.** *Design $L$ in (6) is a Latin hypercube if at least one of the following two conditions is true:*

*(a) $A$ and $C$ satisfy that for any $i$, if $p$ and $p'$ are such that $c_{pi} = -c_{p'i}$, then $a_{pi} = a_{p'i}$;*

*(b) $B$ and $H$ satisfy that for any $j$, if $q$ and $q'$ are such that $b_{qj} = -b_{q'j}$, then $h_{qj} = h_{q'j}$.*

**Proposition 7.** *Suppose that $D_0 = (E, F)$ is a marginally coupled design and $B$ is the corresponding Latin hypercube of $F$, where $E$ is an $n_2 \times q_0$ array and $F$ is an $n_2 \times m_2$ array. In addition, $A$, $B$, $C$ and $H$ are chosen to satisfy conditions in Lemma 5. Let $D_1 = A \otimes E$ and $D_2$ be the Latin hypercube design based on $L$ in (6) via (1). Then we have that design $D = (D_1, D_2)$ is a marginally coupled design, where $D_1$ is an $(n_1 n_2) \times (q_0 m_1)$ array and $D_2$ is an $(n_1 n_2) \times (m_1 m_2)$ array.*

*Proof.* Because $D_0 = (E, F)$ is a marginally coupled design, for each column of $E$, $B$ forms a sliced Latin hypercube of two slices. That is, let $\omega_{k1} = \{i : e_{ik} = 1\}$ and $\omega_{k2} = \{i : e_{ik} = -1\}$, where $e_{ik}$ is the $(i,k)$th element of $E$, $1 \le k \le q_0$. Then for $1 \le j \le m_2$, both $\{\lceil [b_{ij} + (n_2 + 1)/2]/2 \rceil : i \in \omega_{k1}\}$ and $\{\lceil [-b_{ij} + (n_2+1)/2]/2 \rceil : i \in \omega_{k1}\}$ are a permutation of $\{1, \ldots, n_2/2\}$, where $\lceil x \rceil$ is the smallest integer not less than $x$ . Similarly, both $\{\lceil [b_{ij} + (n_2+1)/2]/2 \rceil : i \in \omega_{k2}\}$ and $\{\lceil [-b_{ij} + (n_2+1)/2]/2 \rceil : i \in \omega_{k2}\}$ are a permutation of $\{1, \ldots, n_2/2\}$. It is easy to verify that for $1 \le j' \le m_1$ and $1 \le j \le m_2$, $\{\lceil (a_{i'j'} b_{ij} + n_2 c_{i'j'} h_{ij} + (n_1 n_2 + 1)/2)/2 \rceil : i \in \omega_{k1}, 1 \le i' \le n_1\}$ is a permutation of $\{[c_{i'j'} + (n_1+1)/2 - 1] n_2/2 + 1, \ldots, [c_{i'j'} + (n_1+1)/2] n_2/2\}$. Let $a_{j'}$ and $c_{j'}$ be the $j'$th column of $A$ and $C$, respectively. Because $c_{j'}$ is a permutation of $\{-(n_1-1)/2, \ldots, (n_1-1)/2\}$, $\{\lceil [a_{j'} b_{ij} + n_2 c_{j'} h_{ij} + (n_1 n_2 + 1) 1_{n_1}/2]/2 \rceil : i \in \omega_{k1}\}$ is a permutation of $\{1, \ldots, (n_1 n_2 + 1)/2\}$ where $1_{n_1}$ is a column of $n_1$ 1's. Likewise, $\{\lceil [a_{j'} b_{ij} + n_2 c_{j'} h_{ij} + (n_1 n_2 + 1) 1_{n_1}/2]/2 \rceil : i \in \omega_{k2}\}$ is a permutation of $\{1, \ldots, (n_1 n_2 + 1)/2\}$. Thus we show for each column of $D_1 = A \otimes E$, $L$ forms a sliced Latin hypercube of two slices. This completes the proof.

Analogous to Theorem 1 in Lin, Bingham, Sitter and Tang (2010), an orthogonal $D_2$ in Proposition 7 can be obtained by letting (1) $A$, $B$, $C$ and $H$ are orthogonal, (2) $A^T C = 0$ or $B^T H = 0$, and (3) $u_{ij}$'s in (1) are a constant between 0 and 1.

**Example 5.** *Let*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}, B = \frac{1}{2} \begin{pmatrix} -3 & -1 & 1 & 3 \\ 3 & 1 & -3 & -1 \\ -1 & -3 & 3 & 1 \\ 1 & 3 & -1 & -3 \end{pmatrix},$$

$$C = \frac{1}{2} \begin{pmatrix} 1 & -3 \\ 3 & 1 \\ -1 & 3 \\ -3 & -1 \end{pmatrix}, \text{ and } H = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

*Design $B$ is a marginally sliced Latin hypercube for*

$$E = \begin{pmatrix} 1 & 1 \\ -1 & 1 \\ -1 & -1 \\ 1 & -1 \end{pmatrix}.$$

*By Proposition 7, design $D = (D_1, D_2)$ with $D_1 = A \otimes E$ and $D_2$ based on $L$ in (6) via (1) is a marginally coupled design. If instead we choose both $B$ and $H$ to be orthogonal and let $u_{ij}$ in (1) be a constant between 0 and 1, then the resulting $D_2$ is orthogonal.*

Proposition 7 provides a way to construct marginally coupled designs when $D_1$ is a fold-over orthogonal array of $2^k$ runs and $2^{k-1}$ columns. To explain this, let $D^{(k)} = (D_1^{(k)}, D_2^{(k)})$ be such a marginally coupled design for a given $k$. For $k = 2$, the design with $D_1^{(2)}$ and $D_2^{(2)}$ being $E$ and $B$ in Example 5 is a marginally coupled design. For $k \geq 3$, the design with $D_1^{(k)} = A \otimes D_1^{(k-1)}$ and $D_2^{(k)} = A \otimes D_2^{(k-1)} + 2^{k-1} C \otimes H$, where $A = ((1,1)^T, (1,-1)^T)$, $C$ is a $2 \times 2$ Latin hypercube, and $H$ is a matrix of all 1's of the same size as $D_2^{(k-1)}$, is a marginally coupled design. The design $D_1^{(k)}$ is a fold-over orthogonal array of $2^k$ runs and $2^{k-1}$ factors.

## 5. Conclusions and Discussion

We introduce a new class of designs, marginally coupled designs, to accommodate a large number of qualitative factors in computer experiments with both qualitative and quantitative factors. Construction methods are given for various types of designs for qualitative factors. The existence of such designs is studied when design $D_1$ for qualitative factors are $s$-level orthogonal arrays and $\mathrm{OA}(n, s_1^{q_1}(\lambda s_1)^{q_2}, 2)$. Although completely solving the existence issue for general orthogonal arrays is likely to be quite nontrivial, it would be possible to obtain some useful general results. We will not dwell on this but conclude the paper with a couple of the future research. One important future research direction is the extension of marginally coupled designs with certain optimal criteria. This is because the proposed designs are space-filling in one-dimension, but it is not guaranteed that the designs are space-filling in higher dimensions. Therefore, additional criteria such as orthogonality or maximin distance ( Johnson et al., 1990; Yang et al., 2013; Huang et al., 2014; Ba et al., 2014) can be used to further enhance the space-filling properties. Another direction is the extension of marginally coupled designs to allow the design for quantitative factors to possess space-filling property with respect to any two columns of the design for qualitative factors. One possibility is that, for each level combination of any two columns of the design for qualitative factors, the corresponding design points for quantitative factors form a Latin hypercube design.

**Acknowledgement**

Deng and Hung are supported by the National Science Foundation. Lin is supported by the Natural Sciences and Engineering Research Council of Canada. We are grateful to the

Associate Editor and the referees for constructive comments that have helped improve the article significantly.

**References**

Ba. S., Brenneman, W.A. and Myers, W.R. (2014). Optimal sliced Latin hypercube designs. Revised for *Technometrics*.

Bingham, D., Sitter, R.R. and Tang, B. (2009). Orthogonal and nearly orthogonal designs for computer experiments. *Biometrika* **96**, 51–65.

Dey, A. and Midha, C.K. (1996). Construction of some asymmetrical orthogonal arrays. *Statist. Probab. Lett.* **28**, 211–217.

Fang, K.T., Li, R. and Sudjianto, A. (2010). *Design and Modeling for Computer Experiments*. New York: Chapman and Hall/CRC.

Han, G., Santner, T.J., Notz, W.I. and Bartel, D.L. (2009). Prediction for computer experiments having quantitative and qualitative input variables. *Technometrics* **51**, 278–288.

Hedayat, A.S., Pu, K. and Stufken, J. (1992). On the construction of asymmetrical orthogonal arrays. *Ann. Statist.* **20**, 2142–2152.

Hedayat, A.S., Sloane, N.J.A. and Stufken, J. (1999). *Orthogonal Arrays: Theory and Applications*. New York: Springer.

He, X. and Qian, P.Z.G. (2011). Nested orthogonal array based Latin hypercube designs. *Biometrika* **98**, 721–731.

Huang, H., Yang, J.F. and Liu, M.Q. (2014). Construction of sliced (nearly) orthogonal Latin hypercube designs. *J. of Complexity* 30(3), 355–365.

Hung, Y., Joseph, V.R. and Melkote, S.N. (2009). Design and analysis of computer experiments with branching and nested factors. *Technometrics* **51**, 354–365.

Johnson, M., Moore, L. and Ylvisaker, D. (1990). Minimax and maximin distance design. *J. Statist. Plann. Inference* **26**, 131–148.

Joseph, V.R. and Hung, Y. (2008). Orthogonal-maximin Latin hypercube designs. *Stat. Sinica* **18(1)**, 171–186.

Lin, C.D., Bingham, D., Sitter, R.R. and Tang, B. (2010). A new and flexible method for constructing designs for computer experiments. *Ann. Stat.* **38**, 1460–1477.

Lin, C.D., Mukerjee, R. and Tang, B. (2009). Construction of orthogonal and nearly orthogonal Latin hypercubes. *Biometrika* **96**, 243–247.

McKay, M.D., Beckman, R.J. and Conover, W.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245.

Moon, H., Dean, A. and Santner, T. (2011). Algorithms for generating maximin Latin hypercube and orthogonal designs. *J. Stat. Theory Pract.* **5**, 81-98.

Morris, M.D. and Mitchell, T.J. (1995). Exploratory designs for computer experiments. *J. Statist. Plann. Inference* **43**, 381–402.

Owen, A. (1994). Controlling correlations in Latin hypercube samples. *J. Amer. Statist. Assoc.* **89**, 1517–1522.

Qian, P.Z.G. (2009). Nested Latin Hypercube Designs. *Biometrika* **96**, 957–970.

Qian, P.Z.G. (2012). Sliced Latin hypercube designs. *J. Amer. Statist. Assoc.* **107**, 393–399.

Qian, P.Z.G., Wu, H. and Wu, C.F.J. (2008). Gaussian process models for computer experiments with qualitative and quantitative factors. *Technometrics* **50**, 383–396.

Qian, P.Z.G. and Wu, C.F.J. (2009). Sliced space-filling designs. *Biometrika* **96**, 733–739.

Pang, F., Liu, M.Q. and Lin, D.K.J. (2009). A construction method for orthogonal Latin hypercube Designs with prime power levels. *Stat. Sinica* **19**, 1721–1728.

Santner, T.J., Williams, B.J. and Notz, W.I. (2003). *The Design and Analysis of Computer Experiments.* New York: Springer.

Sloane, N.J.A. (2014). http://neilsloane.com/oadir/.

Steinberg, D.M. and Lin, D.K.J. (2006). A construction method for orthogonal Lain hypercube designs. *Biometrika* **93**, 279–288.

Sun, F., Liu, M. and Lin, D.K.J. (2009). Construction of orthogonal Latin hypercube designs. *Biometrika* **96**, 971–974.

Suen, C.Y. (1989). Some resolvable orthogonal arrays with two symbols. *Comm. Statist. Theory Methods* **18**, 3875–3881.

Tang, B. (1993). Orthogonal array-based Latin hypercubes. *J. Amer. Statist. Assoc.* **88**, 1392–1397.

Wang, J.C. and Wu, C.F.J. (1991). An approach to the construction of asymmetrical orthogonal arrays. *J. Amer. Statist. Assoc.* **86**, 450–456.

Yang, J., Lin, C.D., Qian, P.Z.G. and Lin, D.K.J. (2013). Construction of sliced orthogonal Latin hypercube designs. *Stat. Sinica* **23**, 1117-1130.

Ye, K.Q. (1998). Orthogonal column Latin hypercubes and their applications in computer experiments. *J. Amer. Statist. Assoc.* **93**, 1430–1439.