Statistica Sinica **25** (2015), 1599-1612
doi:http://dx.doi.org/10.5705/ss.2014.160

NEARLY ORTHOGONAL LATIN HYPERCUBE DESIGNS FOR MANY DESIGN COLUMNS

Lin Wang[1], Jian-Feng Yang[1], Dennis K. J. Lin[2] and Min-Qian Liu[1]

[1]*Nankai University and* [2]*The Pennsylvania State University*

Abstract: Latin hypercube designs (LHDs) have found wide applications in computer experiments. Some methods have been proposed to construct orthogonal (or nearly orthogonal) LHDs. This paper proposes methods for expanding a fold-over orthogonal (or nearly orthogonal) LHD to a nearly orthogonal LHD which is able to accommodate many factors. The number of factors is flexible and can be almost as twice large as the number of factors of the original LHD, while the run size remains unchanged. It is shown that the upper bound of the maximum correlation between any two distinct columns of the resulting design is very small (smaller than 0.10 for most cases). The proposed methods can be applied to any fold-over LHDs.

Key words and phrases: Computer experiment, correlation, Latin hypercube design, orthogonality.

1. Introduction

Designs of computer experiments have received a great deal of attention in the past decades. Scientists are increasingly using experiments on computer simulators to help understand complicated physical phenomena (Fang, Li, and Sudjianto (2006)). Latin hypercube designs (LHDs), introduced by McKay, Beckman, and Conover (1979), have been popularly used for computer experiments because of their uniform coverage property. An $n \times k$ LHD, denoted by LHD(n, k), is a matrix of k columns each being a permutation of n equally-spaced levels. In this paper, the n levels of an LHD(n, k) are taken to be $\{-(n-1)/2, -(n-3)/2, \ldots, (n-1)/2\}$. Orthogonality is an important criterion for evaluating LHDs. An LHD is called orthogonal if the correlation coefficient between any two distinct columns in the design is zero. For any design $L = (l_1, \ldots, l_k)$, where l_i is the ith column of L, we define $\rho_{ij}(L) = l_i' l_j / (l_i' l_i l_j' l_j)^{1/2}$. If the sum of the elements in l_i for all $i = 1, \ldots, k$ is zero (i.e., centerized), then $\rho_{ij}(L)$ is simply the correlation coefficient between l_i and l_j. A design L is called column-orthogonal if $\rho_{ij}(L) = 0$ for all $i \neq j$. Otherwise, take $\rho_M(L) = \max_{i<j} |\rho_{ij}(L)|$ to be the maximum correlation of L. For an LHD in this paper, column-orthogonality is equivalent to orthogonality since it has centered levels.

For first-order polynomial models, orthogonal LHDs are useful because they ensure the estimates of linear effects are uncorrelated. Construction of orthogonal LHDs has been widely studied, see e.g., Ye (1998), Steinberg and Lin (2006), Cioppa and Lucas (2007), Bingham, Sitter, and Tang (2009), Pang, Liu, and Lin (2009), Georgiou (2009), Lin, Mukerjee, and Tang (2009), Sun, Liu, and Lin (2009, 2010), Lin et al. (2010), Sun, Pang, and Liu (2011), Georgiou and Stylianou (2011) (and its corrigendum Georgiou and Stylianou (2012)), Yang and Liu (2012), and Georgiou and Efthimiou (2014), among others. Note that some of them considered LHDs with fold-over structures, which makes sure that the sum of elementwise product of any three columns is zero. However, the resulting LHDs of these methods usually have severe restrictions on the design dimensions. Georgiou and Stylianou (2011), Yang and Liu (2012), and Efthimiou, Georgiou, and Liu (2014) constructed nearly orthogonal LHDs by adding runs to existing LHDs. Inevitably, these methods increase experimental costs. By contrast, our studies show how to accommodate more design columns while keeping (near) orthogonality and the same run size. The resulting designs have larger factor-to-run ratios and are more economical.

We propose new methods to construct nearly orthogonal LHDs with flexible run sizes. The number of possible factors is flexible, and can be almost as large as the run size. Bingham, Sitter, and Tang (2009) Lin, Mukerjee, and Tang (2009), Lin et al. (2010), Sun, Pang, and Liu (2011), and Gu and Yang (2013) also constructed nearly orthogonal LHDs. In particular, the method of adding columns to the fold-over LHDs of Sun, Liu, and Lin (2009), proposed by Gu and Yang (2013), can be regarded as a special case of ours. As will be seen in Section 3, the resulting LHDs in this paper can study more factors than the existing designs with the same run sizes (under the restriction of small correlations). In addition, though there are some algorithmic methods intended for searching nearly orthogonal LHDs of any sizes, they are not always able to find designs with good properties and are too cumbersome for finding designs even with moderate sizes. The proposed methods are shown to have high efficiency even for LHDs with large sizes, and the resulting designs have small correlations between distinct columns.

This paper is organized as follows. Section 2 proposes the new construction methods. The upper bound of the maximum correlation of the resulting LHDs is also provided. Section 3 shows some results and comparisons among the existing methods and the proposed methods. Concluding remarks are provided in Section 4. All proofs are deferred to the Appendix.

2. The Construction Methods

A design is fold-over if when d is one of its rows, $-d$ is also one of its rows. In this section, we consider adding columns (H) to a fold-over orthogonal or nearly orthogonal LHD (L), such that the combined design (L, H) is an LHD with a small correlation between distinct columns. The run size is unchanged.

2.1. Construction by adding columns to a $2n$-run LHD

Consider the case that L is an LHD$(2n, m)$ with $L = (D', -D')'$. Let h be any column of H. Here, to reduce the correlation between h and any column of L, it is natural to assign the two closest levels of h to the ith and $(n+i)$th rows, $i = 1, \ldots, n$ (Gu and Yang (2013)). It is then desirable that the $2n \times 1$ column h satisfies

(a) h is a permutation of $\{-(2n-1)/2, -(2n-3)/2, \ldots, (2n-1)/2\}$; and

(b) for $i = 1, \ldots, n$, $|h_i - h_{n+i}| = 1$.

Theorem 1. *For $n \geq 2$, if ρ_{hj} is the correlation coefficient between a column h satisfying* (a) *and* (b) *and the jth column of L, $j = 1, \ldots, m$, then*

$$|\rho_{hj}| \leq \frac{3n}{4n^2 - 1}. \tag{2.1}$$

The upper bound in (2.1) is less than 0.1 when $n \geq 8$. We now provide an algorithm to generate the design H such that any column h of H satisfies (a) and (b).

Algorithm 1 (Construction of H when L has a run size of $2n$)**.**

Step 1. Take X to be an orthogonal or nearly orthogonal LHD(n, k).

Step 2. Take $E = 2X - J_{nk}/2$ and $F = 2X + J_{nk}/2$, where J_{nk} is an $n \times k$ matrix with all elements unity.

Step 3. Take $H = (E', F')'$.

Theorem 2. *For an H constructed via Algorithm* 1*,*

(i) *each column of H satisfies* (a) *and* (b)*;*

(ii) *H is a nearly orthogonal LHD$(2n, k)$ with*

$$\rho_{ij}(H) = \frac{4(n^2-1)\rho_{ij}(X) + 3}{4n^2 - 1}, \textit{for any } i \neq j; \textit{ and} \tag{2.2}$$

(iii) *if X has a fold-over structure, so does H.*

Corollary 1. *For $n \geq 2$, suppose L is a fold-over orthogonal or nearly orthogonal LHD$(2n, m)$. The design (L, H) formed by combining L and an H in Algorithm 1 is a nearly orthogonal LHD$(2n, m + k)$ with*

$$\rho_M((L, H)) \leq \max\left\{\rho_M(L), \frac{3n}{4n^2 - 1}, \frac{4(n^2 - 1)\rho_M(X) + 3}{4n^2 - 1}\right\}.$$

Example 1. Suppose L is the fold-over orthogonal LHD$(24, 12)$ in Georgiou and Efthimiou (2014), and X is the nearly orthogonal LHD$(12, 11)$ with $\rho_M(X) = 0.056$ constructed by the algorithm of Lin (2008) (L and X are listed in Tables S1 and S2, respectively, in the supplementary materials). Algorithm 1 gives a nearly orthogonal LHD$(24, 23)$ of H (Table S3 in the supplementary materials) with $\rho_M(H) = 0.0609$. According to Theorem 1, the maximum correlation between any column of H and any column of L is less than $3n/(4n^2 - 1) = 0.0626$ (in fact, it is 0.0365). Thus (L, H) is a nearly orthogonal LHD$(24, 23)$ with $\rho_M((L, H)) = 0.0609$. This design is apparently new and not a product of any existing method.

Algorithm 1 is easy to implement. Though H is generally not orthogonal, the correlation between any two distinct columns of H is quite small and is smaller than the upper bound in (2.1) if $\rho_M(X) < 3/(4(n + 1))$, thus will not increase the maximum correlation of the whole matrix (L, H). However, sometimes it is desirable to have a much smaller maximum correlation $\rho_M(H)$. To accomplish this, we offer a modified algorithm.

Definition 1. The sign matrix of an $n \times m$ matrix $A = (a_{ij})$ is an $n \times m$ matrix $S_A = (s_{ij})$ with

$$s_{ij} = \begin{cases} 1, & \text{if } a_{ij} \geq 0; \\ -1, & \text{if } a_{ij} < 0. \end{cases}$$

Algorithm 2 (Modified construction of H when L has a run size of $2n$)**.**

Step 1. Take $X = (x_{ij})$ to be an orthogonal or nearly orthogonal LHD(n, k) with $S_X = (s_{ij})$ as its sign matrix.

Step 2. Take $E = (e_{ij})$ with

$$e_{ij} = \begin{cases} s_{ij}(2|x_{ij}| - \frac{1}{2}), & \text{for } 1 \leq i \leq \lceil \frac{n}{2} \rceil; \\ s_{ij}(2|x_{ij}| + \frac{1}{2}), & \text{for } \lceil \frac{n}{2} \rceil + 1 \leq i \leq n, \end{cases}$$

where $\lceil c \rceil$ is the smallest integer not less than c.

Step 3. Take $F = (f_{ij})$ with

$$f_{ij} = \begin{cases} s_{ij}(2|x_{ij}| + \frac{1}{2}), & \text{for } 1 \leq i \leq \lceil \frac{n}{2} \rceil; \\ s_{ij}(2|x_{ij}| - \frac{1}{2}), & \text{for } \lceil \frac{n}{2} \rceil + 1 \leq i \leq n. \end{cases}$$

Step 4. Take $H = (E', F')'$.

Theorem 3. *For an H constructed via Algorithm* 2*,*

(i) *each column of H satisfies* (a) *and* (b)*;*

(ii) *H is a nearly orthogonal LHD$(2n, k)$ with*

$$\rho_{ij}(H) = \frac{4(n^2-1)\rho_{ij}(X) + 3\rho_{ij}(S_X)}{4n^2-1}, \text{for any } i \neq j; \text{and} \tag{2.3}$$

(iii) *if X has a fold-over structure, so does H.*

Corollary 2. *For $n \geq 2$, suppose L is a fold-over orthogonal or nearly orthogonal LHD$(2n, m)$. The design (L, H) formed by combining L and an H in Algorithm 2 is a nearly orthogonal LHD$(2n, m + k)$ with*

$$\rho_M((L, H)) \leq \max\left\{\rho_M(L), \frac{3n}{4n^2-1}, \frac{4(n^2-1)\rho_M(X) + 3\rho_M(S_X)}{4n^2-1}\right\}.$$

Corollary 3. *For the H constructed via Algorithm* 2*, if S_X is column-orthogonal, then*

$$\rho_{ij}(H) = \frac{4(n^2-1)\rho_{ij}(X)}{4n^2-1}, \text{ for any } i \neq j.$$

Furthermore if X is orthogonal, then so is H.

Remark 1. In Algorithm 2, let the matrices E and F be specified by

$$e_{ij} = \begin{cases} s_{ij}(2|x_{ij}| - \frac{1}{2}), & \text{for } i \leq \lceil \frac{n}{2} \rceil \text{ and } j \leq \lceil \frac{k}{2} \rceil, \\ s_{ij}(2|x_{ij}| + \frac{1}{2}), & \text{for } i \geq \lceil \frac{n}{2} \rceil + 1 \text{ and } j \leq \lceil \frac{k}{2} \rceil, \\ s_{ij}(2|x_{ij}| + \frac{1}{2}), & \text{for } i \leq \lceil \frac{n}{2} \rceil \text{ and } j \geq \lceil \frac{k}{2} \rceil + 1, \\ s_{ij}(2|x_{ij}| - \frac{1}{2}), & \text{for } i \geq \lceil \frac{n}{2} \rceil + 1 \text{ and } j \geq \lceil \frac{k}{2} \rceil + 1; \end{cases}$$
$$f_{ij} = \begin{cases} s_{ij}(2|x_{ij}| + \frac{1}{2}), & \text{for } i \leq \lceil \frac{n}{2} \rceil \text{ and } j \leq \lceil \frac{k}{2} \rceil, \\ s_{ij}(2|x_{ij}| - \frac{1}{2}), & \text{for } i \geq \lceil \frac{n}{2} \rceil + 1 \text{ and } j \leq \lceil \frac{k}{2} \rceil, \\ s_{ij}(2|x_{ij}| - \frac{1}{2}), & \text{for } i \leq \lceil \frac{n}{2} \rceil \text{ and } j \geq \lceil \frac{k}{2} \rceil + 1, \\ s_{ij}(2|x_{ij}| + \frac{1}{2}), & \text{for } i \geq \lceil \frac{n}{2} \rceil + 1 \text{ and } j \geq \lceil \frac{k}{2} \rceil + 1. \end{cases}$$

Theorem 3 still holds. Algorithm 4 in Gu and Yang (2013) is a special case of Algorithm 2.

Example 2. Let L and X be the fold-over orthogonal LHD$(64, 32)$ and LHD$(32, 16)$ constructed in Sun, Liu, and Lin (2009) (see Tables S4–S6 in the supplementary materials). From Lemma 1 in Sun, Liu, and Lin (2009), $\rho_M(S_X) = 0$.

Then Algorithm 2 gives an orthogonal LHD$(64, 16)$ of H. In addition, H has a fold-over structure. The maximum correlation between any column of H and any column of L is 0.0234. Thus (L, H) is a nearly orthogonal LHD$(64, 48)$ with $\rho_M((L, H)) = 0.0234$; this is smaller than the 0.0238 of the LHD$(64, 48)$ constructed by Lin, Mukerjee, and Tang (2009).

The condition that S_X is column-orthogonal is rather common for an orthogonal or nearly orthogonal LHD X. Most of the fold-over orthogonal LHDs with even run sizes constructed by, for example, Ye (1998), Sun, Liu, and Lin (2009, 2010), Georgiou (2009), Georgiou and Stylianou (2011), Yang and Liu (2012), and Georgiou and Efthimiou (2014) satisfy this condition. Even if S_X is not column-orthogonal, the value $\rho_{ij}(S_X)$ in (2.3) is usually quite small. Thus the value in (2.3) is usually much smaller than that in (2.2). Hence, Algorithm 2 provides an LHD with a lower maximum correlation than Algorithm 1 does. If the bound given in Theorem 2 is acceptable, we recommend Algorithm 1, as it is a much straightforward approach.

2.2. Construction by adding columns to a $(2n + 1)$-run LHD

Consider the case that L is an LHD$(2n + 1, m)$ with $L = (D', 0_m, -D')'$, where 0_m denotes an $m \times 1$ vector with all entries zero. Let h be any column of H. To reduce the correlation between h and any column of L, we require the $(2n + 1) \times 1$ column h to satisfy (see, Gu and Yang (2013))

(c) the column h is a permutation of $\{-n, -(n - 1), \ldots, n\}$; and

(d) for $i = 1, \ldots, n$, $|h_i - h_{n+1+i}| = 1$ and $h_{n+1} = 0$.

Theorem 4. *For $n \geq 2$, if ρ_{hj} is the correlation coefficient between a column h satisfying* (c) *and* (d) *and the jth column of L, $j = 1, \ldots, m$, then*

$$|\rho_{hj}| \leq \frac{3}{4n + 2}. \tag{2.4}$$

The upper bound in (2.4) is less than 0.1 when $n \geq 7$.

Algorithm 3 (Construction of H when L has $2n + 1$ runs)**.**

Step 1. Take $X = (x_{ij})$ to be an orthogonal or nearly orthogonal LHD(n, k) with $S_X = (s_{ij})$ as its sign matrix (note that the sum of the elements in each column of S_X is zero if n is even).

Step 2. Take $E = (e_{ij})$ with

$$e_{ij} = \begin{cases} 2x_{ij}, & \text{for } 1 \leq i \leq \lceil \frac{n}{2} \rceil; \\ s_{ij}(2|x_{ij}| + 1), & \text{for } \lceil \frac{n}{2} \rceil + 1 \leq i \leq n. \end{cases}$$

Step 3. Take $F = (f_{ij})$ with

$$f_{ij} = \begin{cases} s_{ij}(2|x_{ij}|+1), & \text{for } 1 \leq i \leq \lceil \frac{n}{2} \rceil; \\ 2x_{ij}, & \text{for } \lceil \frac{n}{2} \rceil + 1 \leq i \leq n. \end{cases}$$

Step 4. If n is even, let $H = (E', 0_k, F')'$; otherwise, let $H = (E', -1_k, F')'$, where 1_k denotes a $k \times 1$ vector with all entries unity.

Theorem 5. *For an H constructed in Algorithm* 3*,*

(i) *each column of H satisfies* (c) *and* (d)*;*

(ii) *H is a nearly orthogonal LHD*$(2n+1,k)$ *with*

$$|\rho_{ij}(H)| \leq \frac{2(n-1)|\rho_{ij}(X)|}{2n+1} + \frac{3n+3|\rho_{ij}(S_X)|}{2n^2+3n+1}; \tag{2.5}$$

(iii) *for an even n, if both X and S_X are orthogonal and $S_X'X + X'S_X = n^2 I_k/2$, where I_k denotes the identity matrix of order k, then H is orthogonal;*

(iv) *for an odd n, if X is orthogonal* (*note that S_X cannot be column-orthogonal in this case*) *and $S_X'X + X'S_X = (n^2-1)I_k/2$, then*

$$|\rho_{ij}(H)| \leq \frac{3|\rho_{ij}(S_X)|}{2n^2+3n+1} + \frac{3}{n(2n^2+3n+1)}; \textit{ and}$$

(v) *if X has a fold-over structure, so does H.*

Corollary 4. *For any $n \geq 2$, suppose L is a fold-over orthogonal or nearly orthogonal LHD*$(2n+1,m)$*. The design* (L,H)*, formed by combining L and an H constructed in Algorithm* 3*, is a nearly orthogonal LHD*$(2n+1,m+k)$ *with*

$$\rho_M((L,H)) \leq \max\left\{\rho_M(L), \frac{3}{4n+2}, \frac{2(n-1)\rho_M(X)}{2n+1} + \frac{3n+3\rho_M(S_X)}{2n^2+3n+1}\right\}.$$

Remark 2. In Algorithm 3, if we exchange the right $\lceil k/2 \rceil$ columns of E and F as in Remark 1, Theorem 5 still holds. In this case, Algorithm 3 in Gu and Yang (2013) is a special case of Algorithm 3.

Remark 3. There are many fold-over LHDs satisfying the conditions in Theorem 5(iii) and (iv); see, for example, the orthogonal LHDs constructed by Ye (1998), Sun, Liu, and Lin (2009, 2010), Georgiou (2009), Georgiou and Stylianou (2011), Yang and Liu (2012), and Georgiou and Efthimiou (2014).

Example 3. Suppose L is the fold-over orthogonal LHD$(49,24)$ (Table S7 in the supplementary materials) constructed in Georgiou and Stylianou (2011), and X is the nearly orthogonal LHD$(24,23)$ constructed in Example 1. It can be checked

Table 1. Some existing methods for constructing the L.

Method	Run size N	Number of factors
Ye	2^{r+1} or $2^{r+1}+1$	$2r$
CL	2^{r+1} or $2^{r+1}+1$	$1+r+\binom{r}{2}$
Ge	4, 5, 8, 9, 16, 17	2, 2, 4, 4, 8, 8, respectively
SLL	$c2^{r+1}$ or $c2^{r+1}+1$	2^r
YL	$c2^{r+1}+i$, $i=0,1,2,3$	2^r
GS	7, $8k+3$	3, $4k(k=1,\ldots,6)$, respectively
	$8k+i$, $i=0,1,2$	$4k(k=1,\ldots,6,8)$
GE	$2mk$ or $2mk+1$	$m=12,16,20,24$
EGL	$2mk+2$ or $2mk+3$	$m=2,4,8,12,16,20,24$

that $\rho_M(S_X)=0.3333$. Algorithm 3 gives a nearly orthogonal LHD$(49,23)$ of H with $\rho_M(H)=0.0580$. The maximum correlation between any column of H and any column of L is 0.0306. Thus (L,H) is a nearly orthogonal LHD$(49,47)$ with $\rho_M((L,H))=0.0580$. This design is new.

3. Some Results and Comparisons

The methods in the previous section are able to accommodate more columns to nearly orthogonal LHDs. To apply them, we need fold-over orthogonal (or nearly orthogonal) LHDs L with $N=2n$ or $2n+1$ runs, and orthogonal (or nearly orthogonal) LHDs X with n runs. The existences of these two classes of designs are thus critical to the proposed methods. Table 1 provides some existing methods for constructing L. For the simplicity, denote the methods of Ye (1998), Cioppa and Lucas (2007), Georgiou (2009), Sun, Liu, and Lin (2009, 2010), Georgiou and Stylianou (2011), Yang and Liu (2012), Georgiou and Efthimiou (2014), and Efthimiou, Georgiou, and Liu (2014) by Ye, CL, Ge, SLL, GS, YL, GE, and EGL, respectively. They can all be used to construct the design L. For the design X, any orthogonal or nearly orthogonal LHDs with n runs can be used, for example the designs systematically constructed by the methods mentioned in Table 1, and by Steinberg and Lin (2006), Lin, Mukerjee, and Tang (2009), Lin et al. (2010), and Gu and Yang (2013). There are also some algorithmic methods for searching such designs, including particle swarm optimization (Chen et al. (2013); Leatherman, Dean, and Santner (2014)), simulated annealing (Morris and Mitchell (1995)), some versions of genetic algorithms (e.g., Bates, Sienz, and Toropov (2004); Liefvendahl and Stocki (2006)), the search algorithm proposed by Lin (2008), and so on. These algorithmic methods could be used to search for a nearly orthogonal LHD of any size. However, the resulting designs often have relatively large correlation coefficients. In addition, even for moderate n ($n>30$, say), the algorithmic methods may fail to produce LHDs with small correlations and relatively large column sizes because of the cost of computation. By contrast,

Table 2. Some nearly orthogonal LHDs with $N < 100$ runs.

Generated design	L [source]	X [source]	Method
LHD(16, 15)	LHD(16, 8) [YL]	LHD(8, 7) [Lin]	Algorithm 1 or 2
LHD(17, 15)	LHD(17, 8) [YL]	LHD(8, 7) [Lin]	Algorithm 3
LHD(18, 16)	LHD(18, 8) [YL]	LHD(9, 8) [Lin]	Algorithm 1 or 2
LHD(19, 16)	LHD(19, 8) [YL]	LHD(9, 8) [Lin]	Algorithm 3
LHD(24, 23)	LHD(24, 12) [GE]	LHD(12, 11) [Lin]	Algorithm 1 or 2
LHD(25, 23)	LHD(25, 12) [GE]	LHD(12, 11) [Lin]	Algorithm 3
LHD(26, 24)	LHD(26, 12) [GS]	LHD(13, 12) [Lin]	Algorithm 1 or 2
LHD(27, 24)	LHD(27, 12) [GS]	LHD(13, 12) [Lin]	Algorithm 3
LHD(32, 31)	LHD(32, 16) [YL]	LHD(16, 15) [Lin]	Algorithm 1 or 2
LHD(33, 31)	LHD(33, 16) [YL]	LHD(16, 15) [Lin]	Algorithm 3
LHD(34, 32)	LHD(34, 16) [YL]	LHD(17, 16) [Lin]	Algorithm 1 or 2
LHD(35, 32)	LHD(35, 16) [YL]	LHD(17, 16) [Lin]	Algorithm 3
LHD(48, 47)	LHD(48, 24) [GE]	LHD(24, 23)	Algorithm 1 or 2
LHD(49, 47)	LHD(49, 24) [GE]	LHD(24, 23)	Algorithm 3
LHD(50, 47)	LHD(50, 24) [GS]	LHD(25, 23)	Algorithm 1 or 2
LHD(51, 47)	LHD(51, 24) [GS]	LHD(25, 23)	Algorithm 3
LHD(64, 63)	LHD(64, 32) [SLL]	LHD(32, 31)	Algorithm 1 or 2
LHD(65, 63)	LHD(65, 32) [SLL]	LHD(32, 31)	Algorithm 3
LHD(66, 63)	LHD(66, 32) [YL]	LHD(33, 31)	Algorithm 1 or 2
LHD(67, 63)	LHD(67, 32) [YL]	LHD(33, 31)	Algorithm 3
LHD(96, 71)	LHD(96, 24) [GE]	LHD(48, 47)	Algorithm 1 or 2
LHD(97, 71)	LHD(97, 24) [GE]	LHD(48, 47)	Algorithm 3

Note: YL=Yang and Liu (2012); Lin=Lin (2008); GE=Georgiou and Efthimiou (2014); GS=Georgiou and Stylianou (2011); SLL=Sun, Liu, and Lin (2009).

the proposed methods in Section 2 offer a systematic procedure for constructing nearly orthogonal LHDs with smaller correlations as well as the ability to accommodate more columns. The procedure is easy to implement and outperforms other algorithmic methods (especially for large run sizes). Nevertheless, other algorithmic methods can be straightforwardly incorporated into our proposed algorithms for generating the design X as well. For example, Lin (2008) provided the smallest maximum correlations of some orthogonal and nearly orthogonal LHDs according to her algorithm. Based on these designs and our proposed methods, we can obtain other nearly orthogonal LHDs.

Table 2 presents some nearly orthogonal LHDs constructed via the proposed methods. The first column displays the newly generated nearly orthogonal LHDs. The second and third columns list the designs L and X used for the construction of the corresponding generated LHDs with the contents in the brackets being their sources. Note that the designs X without sources are the generated designs with the same sizes from the first column of the same table. The last column lists the corresponding construction algorithms. Table 2 reveals that the proposed methods can produce designs which are able to accommodate a large number of factors. The numbers of factors of the generated LHDs are almost as large as the corresponding run sizes, which indicates that the LHDs we obtain have more flexible choices for the numbers of factors.

4. Concluding Remarks

We have proposed methods to expand a fold-over orthogonal (or nearly orthogonal) LHD to a nearly orthogonal LHD of a larger column size. The method is easy to implement. The number of possible factors of the resulting design is flexible—it can be made as large as nearly the run size while maintaining the run size. The upper bound of the maximum correlation of any of the resulting LHDs is quite small. Since only the fold-over structure is required for the original designs, the proposed methods work for expanding any fold-over LHDs, not restricted to the ones referred in this paper.

Efthimiou, Georgiou, and Liu (2014) and some other references constructed nearly orthogonal LHDs by adding runs to the existing LHDs, while we accommodate more design columns while keeping the same run size and preserving near orthogonality. This is important as computer experiments often look at more factors than existing orthogonal LHDs can afford (Butler (2005)).

This paper focuses on expanding fold-over LHDs while still keeping their near orthogonality. For LHDs without fold-over structure, the proposed methods do not work. How to extend our work to enlarge LHDs without fold-over structure is an issue worth further study.

Acknowledgements

The authors thank Editor Naisyin Wang, an associate editor, and two referees for their valuable comments and suggestions that have resulted in improvements in the paper. This work was supported by the National Natural Science Foundation of China (Grant Nos. 11271205, 11431006, 11101024, 11471172 and 11271355), the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20130031110002), and the "131" Talents Program of Tianjin. The first two authors contributed equally to this work.

Appendix: Proofs

A.1. Proof of Theorem 1

The proof is similar to that of Theorems 1 and 2 in Gu and Yang (2013), and is omitted. Note that $n = 2^c$ is required in Gu and Yang (2013), while the result here works for any positive integer greater than 1.

A.2. Proof of Theorem 2

Each column of X is a permutation of $\{-(n-1)/2, -(n-3)/2, \ldots, (n-3)/2, (n-1)/2\}$. Thus each column of E is a permutation of $\{-(2n-1)/2, -(2n-5)/2, \ldots, (2n-7)/2, (2n-3)/2\}$, and each column of F is a permutation of $\{-(2n-3)/2, -(2n-7)/2, \ldots, (2n-5)/2, (2n-1)/2\}$. Parts (i) and (iii) follow from the structures of E and F in Step 2 of Algorithm 1. Hence, we need only to prove Part (ii). It is clear that H is an LHD. From the structure of H, we have

$$\begin{aligned}
H'H &= E'E + F'F \\
&= (2X - \frac{1}{2}J_{nk})'(2X - \frac{1}{2}J_{nk}) + (2X + \frac{1}{2}J_{nk})'(2X + \frac{1}{2}J_{nk}) \\
&= 8X'X + \frac{1}{2}J_{nk}'J_{nk} \\
&= 8X'X + \frac{n}{2}J_{kk}.
\end{aligned}$$

This indicates that

$$\begin{aligned}
\rho_{ij}(H) &= \frac{8n(n+1)(n-1)\rho_{ij}(X)/12 + n/2}{n(2n+1)(2n-1)/6} \\
&= \frac{4(n^2-1)\rho_{ij}(X) + 3}{4n^2 - 1}
\end{aligned}$$

by noting that $x'x = n(n+1)(n-1)/12$ for any column x of X. This completes the proof.

A.3. Proof of Theorem 3

For $j = 1, \ldots, k$, the jth column of H is a permutation of $\{s_{1j}(2|x_{1j}| - 1/2), \ldots, s_{nj}(2|x_{nj}| - 1/2), s_{1j}(2|x_{1j}| + 1/2), \ldots, s_{nj}(2|x_{nj}| + 1/2)\}$, which is in fact a permutation of $\{-(2n-1)/2, -(2n-3)/2, \ldots, (2n-3)/2, (2n-1)/2\}$. Part (i) follows from the structures of E and F in Steps 2 and 3 of Algorithm 2. For Part (ii), it is clear that H is an LHD. For $i, j = 1, \ldots, k$ and $i \neq j$, the inner product between the ith and jth columns of H is

$$\begin{aligned}
&\sum_{t=1}^{n} \Big(s_{ti}(2|x_{ti}| - \frac{1}{2})s_{tj}(2|x_{tj}| - \frac{1}{2})\Big) + \sum_{t=1}^{n} \Big(s_{ti}(2|x_{ti}| + \frac{1}{2})s_{tj}(2|x_{tj}| + \frac{1}{2})\Big) \\
&\quad = \sum_{t=1}^{n} \Big(8s_{ti}s_{tj}|x_{ti}||x_{tj}| + \frac{1}{2}s_{ti}s_{tj}\Big) \\
&\quad = \sum_{t=1}^{n} \Big(8x_{ti}x_{tj} + \frac{1}{2}s_{ti}s_{tj}\Big).
\end{aligned}$$

Thus

$$\rho_{ij}(H) = \frac{8n(n+1)(n-1)\rho_{ij}(X)/12 + n\rho_{ij}(S_X)/2}{n(2n+1)(2n-1)/6}$$
$$= \frac{4(n^2-1)\rho_{ij}(X) + 3\rho_{ij}(S_X)}{4n^2-1}.$$

For Part (iii), if X is a fold-over design, without loss of generality, suppose $X = (X_1', -X_1')'$ if n is even and $X = (X_1', 0_k, -X_1')'$ if n is odd. Then for the tth row, say e_t, of E, $e_t = (s_{t1}(2|x_{t1}| - 1/2), \ldots, s_{tk}(2|x_{tk}| - 1/2))$ for $t \le \lceil (n-1)/2 \rceil$, the $(t + \lceil n/2 \rceil)$th row of F is $(s_{(t+\lceil n/2\rceil)1}(2|x_{(t+\lceil n/2\rceil)1}| - 1/2), \ldots, s_{(t+\lceil n/2\rceil)k}(2|x_{(t+\lceil n/2\rceil)k}| - 1/2)) = -e_t$. Similarly, for $t > \lceil n/2 \rceil$, the $(t - \lceil n/2 \rceil)$th row of F is $-e_t$. When n is odd, the $\lceil n/2 \rceil$th row of E is $(-1/2, \ldots, -1/2)$ and the $\lceil n/2 \rceil$th row of F is $(1/2, \ldots, 1/2)$. Thus H has a fold-over structure.

A.4. Proof of Theorem 4

The proof is similar to that of Theorem 1 of Gu and Yang (2013), and is omitted. Note that $n = 2^c$ is required in Gu and Yang (2013), while the result here works for any positive integer greater than 1.

A.5. Proof of Theorem 5

The proofs of Parts (i) and (v) are similar to that of Theorem 3 and are omitted. For Part (ii), we need only prove (2.5). For $i, j = 1, \ldots, k$ and $i \neq j$, the inner product between the ith and jth columns of H is

$$\sum_{t=1}^{n} (s_{ti}(2|x_{ti}| + 1)s_{tj}(2|x_{tj}| + 1)) + \sum_{t=1}^{n} 4s_{ti}|x_{ti}|s_{tj}|x_{tj}| + \delta$$
$$= 8\sum_{t=1}^{n} x_{ti}x_{tj} + 2\left(\sum_{t=1}^{n} x_{ti}s_{tj} + \sum_{t=1}^{n} s_{ti}x_{tj}\right) + \sum_{t=1}^{n} s_{ti}s_{tj} + \delta,$$

where

$$\delta = \begin{cases} 0, & \text{if } n \text{ is even;} \\ 1, & \text{if } n \text{ is odd.} \end{cases}$$

By noting that $\sum_{t=1}^{n} x_{ti}s_{tj} + \sum_{t=1}^{n} s_{ti}x_{tj} \le 2\sum_{t=1}^{n} |x_{ti}| = (n^2 - \delta)/2$, we have

$$|\rho_{ij}(H)| \le \frac{8n(n+1)(n-1)|\rho_{ij}(X)|/12 + n^2 + n|\rho_{ij}(S_X)|}{n(n+1)(2n+1)/3}$$
$$= \frac{2(n-1)|\rho_{ij}(X)|}{2n+1} + \frac{3n + 3|\rho_{ij}(S_X)|}{2n^2+3n+1}.$$

Parts (iii) and (iv) are now obvious.

References

Bates, S. J., Sienz, J. and Toropov, V. V. (2004). Formulation of the optimal Latin hypercube design of experiments using a permutation genetic algorithm. *AIAA* 2004-2011, 1-7.

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

Butler, N. A. (2005). Supersaturated Latin hypercube designs. *Commun. Stat. Theory Methods* **34**, 417-428.

Chen, R. B., Hsieh, D. N., Hung, Y. and Wang, W. C. (2013). Optimizing Latin hypercube designs by particle swarm. *Stat. Comput.* **23**, 663-676.

Cioppa, T. M. and Lucas, T. W. (2007). Efficient nearly orthogonal and space-filling Latin hypercubes. *Technometrics* **49**, 45-55.

Efthimiou, I., Georgiou, S. D. and Liu, M. Q. (2014). Construction of nearly orthogonal Latin hypercube designs. *Metrika*, `http://dx.doi.org/10.1007/s00184-014-0489-5`.

Fang, K. T., Li, R. and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments*. Chapman and Hall/CRC, New York.

Georgiou, S. D. (2009). Orthogonal Latin hypercube designs from generalized orthogonal designs. *J. Statist. Plann. Inference* **129**, 1530-1540.

Georgiou, S. D. and Efthimiou, I. (2014). Some classes of orthogonal Latin hypercube designs. *Statist. Sinica* **24**, 101-120.

Georgiou, S. D. and Stylianou, S. (2011). Block-circulant matrices for constructing optimal Latin hypercube designs. *J. Statist. Plann. Inference* **141**, 1933-1943.

Georgiou, S. D. and Stylianou, S. (2012). Corrigendum to: Block-circulant matrices for constructing optimal Latin hypercube designs. *J. Statist. Plann. Inference* **142**, 1623-1623.

Gu, L. and Yang, J. F. (2013). Construction of nearly orthogonal Latin hypercube designs. *Metrika* **76**, 819-830.

Leatherman, E., Dean, A. and Santner, T. (2014). Computer experiment designs via particle swarm optimization. *Preprint NI*14015-*DAE*, Isaac Newton Institute for Mathematical Sciences.

Liefvendahl, M. and Stocki, R. (2006). A study on algorithms for optimization of Latin hypercubes. *J. Statist. Plann. Inference* **136**, 3231-3247.

Lin, C. D. (2008). New developments in designs for computer experiments and physical experiments. Ph.D. thesis, Simon Fraser Univ.

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

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

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

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

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

Sun, F. S., Liu, M. Q. and Lin, D. K. J. (2010). Construction of orthogonal Latin hypercube designs with flexible run sizes. *J. Statist. Plann. Inference* **140**, 3236-3242.

Sun, F. S., Pang, F. and Liu, M. Q. (2011). Construction of column-orthogonal designs for computer experiments. *Sci. China Math.* **54**, 2683-2692.

Yang, J. Y. and Liu, M. Q. (2012). Construction of orthogonal and nearly orthogonal Latin hypercube designs from orthogonal designs. *Statist. Sinica* **22**, 433-442.

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

LPMC and Institute of Statistics, Nankai University, Tianjin 300071, China.

E-mail: lin.happyforever@163.com

LPMC and Institute of Statistics, Nankai University, Tianjin 300071, China.

E-mail: jfyang@nankai.edu.cn

Department of Statistics, The Pennsylvania State University, University Park, PA 16802, USA.

E-mail: DKL5@psu.edu

LPMC and Institute of Statistics, Nankai University, Tianjin 300071, China.

E-mail: mqliu@nankai.edu.cn

(Received May 2014; accepted October 2014)