Journal of Statistical Planning and Inference 185 (2017) 29–40

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Nearly column-orthogonal designs based on leave-one-out good lattice point sets

Ru Yuan [a], Dennis K.J. Lin [b], Min-Qian Liu [a,*]

[a] LPMC and Institute of Statistics, Nankai University, Tianjin 300071, China

[b] Department of Statistics, Pennsylvania State University, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 26 July 2016
Received in revised form 28 December 2016
Accepted 10 January 2017
Available online 21 January 2017

Keywords:
Column-orthogonal
Fold-over
Good lattice point set
L_1-distance
Space-filling

ABSTRACT

Good lattice point sets have desirable space-filling properties, and many designs with large L_1-distance can be obtained by the leave-one-out good lattice point method (Zhou and Xu, 2015). However, there are negatively fully correlated columns in such a design. This is undesirable in the modeling of computer experiments. To overcome such a deficiency, we propose a class of designs based on the leave-one-out good lattice point method, whose columns can be divided into two groups, such that any two columns are column-orthogonal when they are from different groups and nearly column-orthogonal when they are in the same group. The new designs can also estimate the linear effects without being correlated with the second-order effects. Moreover, they have good stratification properties and their L_1-distances are comparable with the corresponding designs in Zhou and Xu (2015).

1. Introduction

Computer experiments are becoming increasingly prevalent as a useful tool for study of uncertainty quantification. In designing a computer experiment, it is crucial to spread the design points throughout the experimental region as evenly as possible. Such designs are referred to as space-filling designs, including Latin hypercube designs (LHDs) (McKay et al., 1979), and their modifications, such as maximin distance designs (Johnson et al., 1990), and uniform designs (Fang and Wang, 1994). Due to the curse of dimensionality, it is rather difficult for design points to cover a large portion of a high dimensional design region. Thus, it is more reasonable to consider designs that are space-filling in lower dimensional projections of the input space. One fruitful approach to constructing such designs is to make use of orthogonal arrays. Owen (1992) and Tang (1993) considered randomized orthogonal arrays and orthogonal array based LHDs, representing an important development in this area. Recently, He and Tang (2013) introduced strong orthogonal arrays and Mukerjee et al. (2014) studied mappable nearly orthogonal arrays, both with the focus on constructing better space-filling designs than those based on ordinary orthogonal arrays.

Besides space-filling, orthogonality is another useful criterion for designing computer experiments. Orthogonal LHDs have received a great deal of attention in the literature. Ye (1998) was the first to construct LHDs that are exactly orthogonal. Further research along this direction results in significant enhancements in the methodology and results; see for example, Steinberg and Lin (2006), Lin et al. (2009), Pang et al. (2009), Sun et al. (2009, 2010), Lin et al. (2010), Yang and Liu (2012), and Wang et al. (2015), among others. It was argued (Bingham et al., 2009) that orthogonal designs with many levels are as suitable as orthogonal LHDs for computer experiments. Construction of such orthogonal designs was further studied by

* Corresponding author.
E-mail address: mqliu@nankai.edu.cn (M.-Q. Liu).

http://dx.doi.org/10.1016/j.jspi.2017.01.002

Sun et al. (2011), Georgiou (2011), and Georgiou et al. (2014). Orthogonality is of importance when polynomial modeling is considered. For the first-order model, orthogonal designs ensure the orthogonality among all main effects. When second-order effects are present, however, it is desirable to have orthogonal designs that are able to estimate the linear effects without being correlated with the second-order effects, i.e., the quadratic effects and bilinear interactions. Thus designs with the following two properties are desirable:

(a) all the columns are orthogonal to each other; and
(b) the sum of the elementwise product of any three columns is zero.

Property (b) is an important property that ensures the estimates of the linear effects being uncorrelated with that of the second-order effects. The designs by Ye (1998), Sun et al. (2009, 2010), Yang and Liu (2012), Georgiou (2011), and Georgiou et al. (2014) have such properties.

Korobov (1959) proposed the good lattice point method for numerical evaluation of multivariate integrals. Fang and Wang (1981) proposed the leave-one-out good lattice point method, which can be used to obtain a design with a better L_1-distance than the good lattice point sets. Recently, Zhou and Xu (2015) studied the space-filling properties of good lattice point sets, and constructed some maximin distance designs by the leave-one-out good lattice point method. However, such designs will result in negatively fully correlated columns, which is an undesirable property in computer experiments.

Motivated by the desirable space-filling properties of good lattice point sets, in this paper, we use leave-one-out good lattice point sets to construct designs that satisfy property (b), with part of their column pairs being column-orthogonal and other column pairs being nearly column-orthogonal. Moreover, the new designs retain the desirable space-filling properties of the original good lattice point sets, and have a better stratification property. Some theoretical results regarding the orthogonality and space-filling property of the new designs are established. Some newly constructed designs are tabulated for practical use. Note that Lin et al. (2010) developed a general method for constructing "good" designs for computer experiments, in which four initial small designs and a numerical parameter are needed to build large "good" designs. There is some similarity between the newly proposed methods and that of Lin et al. (2010). However, our methods need only one initial design, and focus on eliminating the deficiency (i.e., negatively fully correlations) in the initial design, instead of constructing designs that inherit the good properties of the initial designs as in Lin et al. (2010).

The rest of this paper is organized as follows. Section 2 provides relevant definitions and notation. The construction methods and properties of the newly constructed designs are given in Section 3. Some further discussion and concluding remarks are given in Section 4. All proofs are deferred to Appendix.

2. Definitions and notation

Let $D(N, s^m)$ denote a design with N runs, m factors, and s equally-spaced levels. When $s = N$, the design is called a Latin hypercube design (LHD), denoted by LHD(N, m). If all levels are equally replicated in each column of a $D(N, s^m)$, the design is called a U-type design (Fang et al., 2006). The s equally-spaced levels can be denoted as $-(s-1)/2, -(s-3)/2, \ldots, (s-1)/2$. When each column is orthogonal to the mean effect, the design is called a mean orthogonal design; and when the inner product of any two distinct columns is zero, the design is called a column-orthogonal design, denoted by COD(N, s^m). We use $D(N, z, s^m)$ to denote a design with N runs and m factors, with z (> 0) zeros and $s-1$ nonzero equally replicated levels in each factor. When $z = N/s$, $D(N, z, s^m)$ is also a U-type design and will be simply denoted as $D(N, s^m)$. When $z \neq N/s$, then the designs are mean orthogonal but not of U-type and such designs will be suitable only for experiments with quantitative factors (Georgiou et al., 2014). Finally, design F is said to be full fold-over (or fold-over) of design D if $F = (D^T, -D^T)^T$, where T is the notation for transpose.

For an integer $p \geq 1$, define $d_p(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{m} |x_i - y_i|^p$ as the L_p-distance of any two rows $\boldsymbol{x} = (x_1, \ldots, x_m)$ and $\boldsymbol{y} = (y_1, \ldots, y_m)$ in a $D(N, s^m)$ design D. Define the L_p-distance of the design D to be

$$d_p(D) = \min\{d_p(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{x} \neq \boldsymbol{y}, \boldsymbol{x}, \boldsymbol{y} \in D\}.$$

The correlation between two vectors $\boldsymbol{a} = (a_1, \ldots, a_N)^T$ and $\boldsymbol{b} = (b_1, \ldots, b_N)^T$ is

$$\rho(\boldsymbol{a}, \boldsymbol{b}) = \frac{\sum_{i=1}^{N}(a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^{N}(a_i - \bar{a})^2 \sum_{i=1}^{N}(b_i - \bar{b})^2}},$$

where $\bar{a} = \sum_{i=1}^{N} a_i/N$ and $\bar{b} = \sum_{i=1}^{N} b_i/N$. The correlation matrix of a design D with m columns is $\rho(D) = (\rho_{ij}(D))_{m\times m}$, where $\rho_{ij}(D)$ is the correlation between the ith and jth columns of D. Two commonly used measures for evaluating the orthogonality of D are

$$\rho_M(D) = \max_{i<j} |\rho_{ij}(D)| \quad \text{and} \quad \rho^2(D) = 2\sum_{i<j} \rho_{ij}^2(D)/(m(m-1)). \tag{1}$$

If $\bar{a} = \bar{b} = 0$, $\rho(\boldsymbol{a}, \boldsymbol{b})$ reduces to $\tilde{\rho}(\boldsymbol{a}, \boldsymbol{b}) = \sum_{i=1}^{N} a_i b_i / \left(\sqrt{\sum_{i=1}^{N} a_i^2 \sum_{i=1}^{N} b_i^2}\right) = \boldsymbol{a}^T\boldsymbol{b}/(\sqrt{\boldsymbol{a}^T\boldsymbol{a}}\sqrt{\boldsymbol{b}^T\boldsymbol{b}})$.

Table 1
The 7- and 14-run LHDs generated by the good lattice point method.

$D_1: N=7$						$D_2: N=14$					
1	2	3	4	5	6	1	3	5	9	11	13
2	4	6	1	3	5	2	6	10	4	8	12
3	6	2	5	1	4	3	9	1	13	5	11
4	1	5	2	6	3	4	12	6	8	2	10
5	3	1	6	4	2	5	1	11	3	13	9
6	5	4	3	2	1	6	4	2	12	10	8
7	7	7	7	7	7	7	7	7	7	7	7
						8	10	12	2	4	6
						9	13	3	11	1	5
						10	2	8	6	12	4
						11	5	13	1	9	3
						12	8	4	10	6	2
						13	11	9	5	3	1
						14	14	14	14	14	14

3. Construction based on the leave-one-out good lattice point method

3.1. The leave-one-out good lattice point method

The key idea behind the good lattice point method is to use a generator vector $\boldsymbol{h} = (h_1, \ldots, h_m)$ with $1 = h_1 < \cdots < h_m < N$, where each h_i is a positive integer less than and coprime to N. The number of entries in $\boldsymbol{h}$ satisfies $m = \phi(N)$, where $\phi(N)$ is the Euler function, i.e., the number of positive integers that are less than and coprime to N. For $k = 1, \ldots, N$, the kth row of the good lattice point set can be obtained by $\boldsymbol{x}_k = k * \boldsymbol{h} = (kh_1, \ldots, kh_m) \pmod N$, where the multiplication operation modulo N is modified so that the result falls into $[1, N]$, and the last row of a good lattice point set is always $(N, \ldots, N)$. The N-run good lattice point set is an LHD(N, m).

Zhou and Xu (2015) gave some theoretical results on the space-filling properties (in terms of L_1-distance) of good lattice point sets. They showed several classes of good lattice point sets that are nearly maximin L_1-distance designs.

Lemma 1 (*Zhou and Xu, 2015*)**.** *Let m and $\boldsymbol{h}$ be defined as above, D be the good lattice point set generated by $\boldsymbol{h}$.*

(i) *If $N = p_1^t$ for an odd prime p_1, then $d_1(D) = (N^2 + p_1)(1 - 1/p_1)/4$.*
(ii) *If $N = 2p_1$ for an odd prime p_1, then $d_1(D) = (p_1 - 1)^2/2$.*
(iii) *If $N = 2^t$, then $d_1(D) = N^2/8$.*

In some situations, one can obtain a design with greater L_1-distance than the good lattice point set by using the leave-one-out good lattice point method (Fang and Wang, 1981). Deleting the last row of an $(N+1)$-run good lattice point set, we can obtain an N-run point set that forms an LHD$(N, \phi(N+1))$. Zhou and Xu (2015) showed that the N-run design generated by the leave-one-out good lattice point method has the same L_1-distance as the good lattice point set with $N+1$ runs.

Example 1. The 7- and 14-run LHDs generated by the good lattice point method are shown in Table 1. Their L_1-distances can be easily calculated from Lemma 1, i.e., $d_1(D_1) = (N^2 + p_1)(1 - 1/p_1)/4 = 12$ for $N = 7$; and $d_1(D_2) = (p_1 - 1)^2/2 = 18$ for $N = 14$.

Using the leave-one-out good lattice point method, we can obtain another two LHDs, i.e. LHD(6, 6) and LHD(13, 6), from the LHDs in Table 1 by deleting the 7th row of D_1 and the 14th row of D_2, respectively. After their levels being centered, the two LHDs can be re-expressed as in Table 2.

For either design in Table 2, it is easy to see that the 1st and 6th, 2nd and 5th, 3rd and 4th columns are all negatively fully correlated, this can also be seen in Fig. 1 which shows the two-dimensional projections of both designs. Furthermore, each of them has a fold-over structure. In fact, a general result can be established as follows.

Lemma 2. *For the* LHD(N, m) $(m = \phi(N+1))$ *generated by the leave-one-out good lattice point method, let $\boldsymbol{x}_k$ be the kth row and $\boldsymbol{x}^j$ be the jth column of the design, $k = 1, \ldots, N, j = 1, \ldots, m$. Then we have*

$$\boldsymbol{x}_k + \boldsymbol{x}_{N+1-k} = (N+1)\boldsymbol{1}_m^T, \quad k = 1, \ldots, \lfloor N/2 \rfloor, \quad \text{and} \tag{2}$$

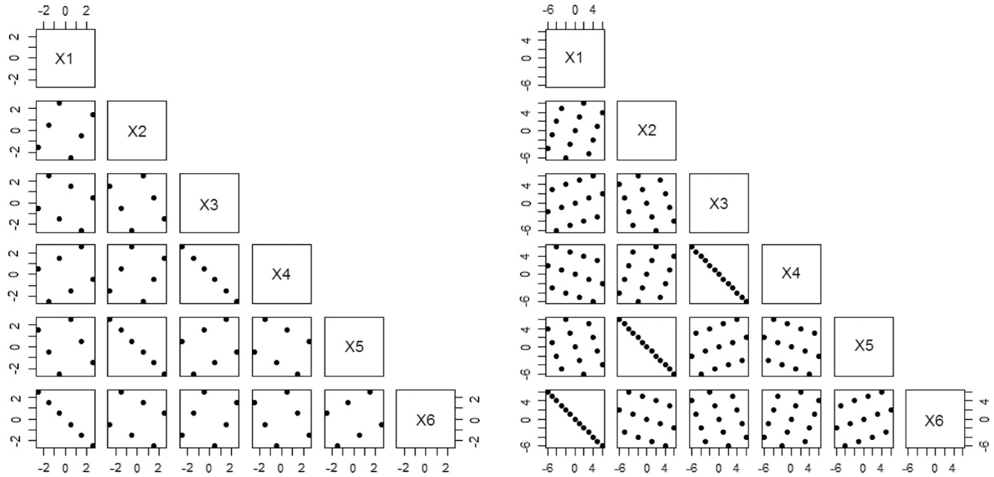
$$\boldsymbol{x}^j + \boldsymbol{x}^{m+1-j} = (N+1)\boldsymbol{1}_N, \quad j = 1, \ldots, m/2, \tag{3}$$

where $\lfloor t \rfloor$ denotes the largest integer not exceeding t, and $\boldsymbol{1}_m$ is an $m \times 1$ vector with all elements unity.

Note that Ye et al. (2000) constructed a kind of designs, called symmetric LHDs, which also satisfy the property in Eq. (2) after a properly reordering of the rows. In fact, this property implies the fold-over structure as shown in Eq. (4).

Table 2
Two LHDs with centered levels obtained from Table 1 by the leave-one-out good lattice point method.

LHD(6, 6)						LHD(13, 6)					
−2.5	−1.5	−0.5	0.5	1.5	2.5	−6	−4	−2	2	4	6
−1.5	0.5	2.5	−2.5	−0.5	1.5	−5	−1	3	−3	1	5
−0.5	2.5	−1.5	1.5	−2.5	0.5	−4	2	−6	6	−2	4
0.5	−2.5	1.5	−1.5	2.5	−0.5	−3	5	−1	1	−5	3
1.5	−0.5	−2.5	2.5	0.5	−1.5	−2	−6	4	−4	6	2
2.5	1.5	0.5	−0.5	−1.5	−2.5	−1	−3	−5	5	3	1
						0	0	0	0	0	0
						1	3	5	−5	−3	−1
						2	6	−4	4	−6	−2
						3	−5	1	−1	5	−3
						4	−2	6	−6	2	−4
						5	1	−3	3	−1	−5
						6	4	2	−2	−4	−6



(a) Two-dimensional projections of the LHD(6, 6). (b) Two-dimensional projections of the LHD(13, 6).

Fig. 1. Two-dimensional projections of the LHDs in Table 2.

Remark 1. For the design generated by the leave-one-out good lattice point method, Eq. (3) implies that for each of its columns, there is another column that is negatively fully correlated with it. Furthermore, if all the levels in each column are centered, we have

$$\boldsymbol{x}_k + \boldsymbol{x}_{N+1-k} = \boldsymbol{0}_m^T, \quad \text{i.e. } \boldsymbol{x}_k = -\boldsymbol{x}_{N+1-k},\ k = 1, \ldots, \lfloor N/2 \rfloor, \quad \text{and} \tag{4}$$

$$\boldsymbol{x}^j + \boldsymbol{x}^{m+1-j} = \boldsymbol{0}_N, \quad \text{i.e. } \boldsymbol{x}^j = -\boldsymbol{x}^{m+1-j},\ j = 1, \ldots, m/2, \tag{5}$$

where $\boldsymbol{0}_m$ is an $m \times 1$ vector with all elements zero. Eq. (4) implies that the design with centered levels has a fold-over structure, and Eq. (5) shows that $\rho(\boldsymbol{x}^j, \boldsymbol{x}^{m+1-j}) = \tilde{\rho}(\boldsymbol{x}^j, \boldsymbol{x}^{m+1-j}) = -1$, for $j = 1, \ldots, m/2$.

Lemma 2 and Remark 1 demonstrate that there are negatively fully correlated columns in the designs generated by the leave-one-out good lattice point method. However, if the design has columns less than $\phi(N+1)$, this is not always true and negatively fully correlated columns may not exist. The details are summarized in the following lemma.

Lemma 3. *For the design obtained by the leave-one-out good lattice point method, assume it has N runs and u ($u \le \phi(N+1)$) factors, its u factors correspond to u entries of the generator vector* $\boldsymbol{h}$ *for* $(N+1)$*-run good lattice point set. Without loss of generality, denote these u entries as* $h_1 < \cdots < h_u$.

(i) *If there exist no two distinct positive integers* $i_1,\ i_2$ ($i_1,\ i_2 \le u$) *such that* $h_{i_1} + h_{i_2} = N+1$*, then there are no fully correlated columns in the design; otherwise, there is at least one pair of columns that are negatively fully correlated in the design.*
(ii) *In particular, if* $u \le \phi(N+1)/2$*, and* $1 \le h_1 < \cdots < h_u \le \lfloor N/2 \rfloor$ *or* $\lfloor N/2 \rfloor < h_1 < \cdots < h_u \le N$*, then there are no fully correlated columns in the design; if* $u > \phi(N+1)/2$*, then there is at least one pair of columns that are negatively fully correlated in the design.*

For this lemma, it is obvious that case (ii) is a special case of case (i). The proof can be easily derived from that of Lemma 2, so we omit it here.

3.2. The proposed construction methods

Let D_0 be the LHD(N_0, m) $(m = \phi(N_0+1))$ obtained by the leave-one-out good lattice point method with centered levels, its columns can be divided into two groups

$$D_0 = (D_0^{(1)} \vdots D_0^{(2)}), \tag{6}$$

where $D_0^{(1)}$ consists of the first $m/2$ columns, and $D_0^{(2)}$ consists of the last $m/2$ columns. From Remark 1, for each column in $D_0^{(1)}$, there is one column in $D_0^{(2)}$ that is negatively fully correlated with it; and D_0, $D_0^{(1)}$ and $D_0^{(2)}$ all have a fold-over structure, which imply the following result.

Theorem 1. *For the designs D_0, $D_0^{(1)}$ and $D_0^{(2)}$ in (6), any of them satisfies property* (b).

From the discussion in Section 3.1, there are negatively fully correlated columns in the designs obtained by the leave-one-out good lattice point method, which is undesirable. In regression analysis, when a polynomial model is used, complete correlations among the design columns will cause the corresponding effects unidentifiable. Here, we propose methods to eliminate the complete correlations in the designs.

Next, we give the construction methods for odd N_0 and even N_0, respectively, due to different properties of the new designs for both cases. When N_0 is an odd integer, from Remark 1, D_0, $D_0^{(1)}$ and $D_0^{(2)}$ in (6) have a fold-over structure, and they can be re-expressed as:

$$D_0 = (D_0^{(1)} \vdots D_0^{(2)}) = \begin{pmatrix} E & \vdots & -E \\ \mathbf{0}_{m/2}^T & \vdots & \mathbf{0}_{m/2}^T \\ -E & \vdots & E \end{pmatrix}, \tag{7}$$

where E is an $(N_0 - 1)/2 \times m/2$ matrix.

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

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

Algorithm 1 (*Construction of Designs with $2N_0 - 1$ or $2N_0 - 2$ Runs and $\phi(N_0 + 1)$ Columns, for **Odd** N_0*).

Step 1. Given an odd integer N_0, generate an $(N_0 + 1)$-run good lattice point set.

Step 2. Delete the last row of the $(N_0 + 1)$-run good lattice point set, and center the levels, denote the obtained design by D_0. Divide the columns of D_0 into two groups as in (7), where $m = \phi(N_0 + 1)$.

Step 3. Construct two new designs based on D_0:

$$D_1 = (D_1^{(1)} \vdots D_1^{(2)}) = \begin{pmatrix} E - \frac{1}{2}S_E & \vdots & -\left(E - \frac{1}{2}S_E\right) \\ -\left(E - \frac{1}{2}S_E\right) & \vdots & E - \frac{1}{2}S_E \\ E - \frac{1}{2}S_E & \vdots & E - \frac{1}{2}S_E \\ -\left(E - \frac{1}{2}S_E\right) & \vdots & -\left(E - \frac{1}{2}S_E\right) \end{pmatrix}, \quad \text{and} \tag{8}$$

$$D_2 = (D_2^{(1)} \vdots D_2^{(2)}) = \begin{pmatrix} D_0^{(1)} & \vdots & D_0^{(2)} \\ D_0^{(1)} & \vdots & -D_0^{(2)} \end{pmatrix} \setminus \mathbf{0}_m^T \tag{9}$$

where $D \setminus \mathbf{0}_m^T$ means deleting a row with all elements zero from D.

Since for any two successive positive integers, the Euler function value $\phi(\cdot)$ (i.e., the number of columns of the design generated by the leave-one-out good lattice point method) of the odd number is larger than the value of the other one in most cases, i.e., the corresponding good lattice point set can evaluate more factors. Next, we give an algorithmic construction based on the N_0-run design D_0 with $\phi(N_0 + 1)$ columns in (6), where $N_0 + 1$ is odd, i.e., N_0 is even.

Algorithm 2 (*Construction of Designs with $2N_0$ or $2N_0 + 1$ Runs and $\phi(N_0 + 1)$ Columns, for **Even** N_0*).

Step 1. Given an even integer N_0, generate an $(N_0 + 1)$-run good lattice point set.

Step 2. Delete the last row of the $(N_0 + 1)$-run good lattice point set, and center the levels, denote the obtained design by D_0. Divide the columns of D_0 into two groups as in (6).

Step 3. Construct two new designs based on D_0:

$$D_3 = (D_3^{(1)} \vdots D_3^{(2)}) = \begin{pmatrix} D_0^{(1)} & \vdots & D_0^{(2)} \\ D_0^{(1)} & \vdots & -D_0^{(2)} \end{pmatrix}, \quad \text{and} \tag{10}$$

$$D_4 = (D_4^{(1)} \vdots D_4^{(2)}) = \begin{pmatrix} D_0^{(1)} + \frac{1}{2}S_{D_0^{(1)}} & \vdots & D_0^{(2)} + \frac{1}{2}S_{D_0^{(2)}} \\ \mathbf{0}_{m/2}^T & \vdots & \mathbf{0}_{m/2}^T \\ D_0^{(1)} + \frac{1}{2}S_{D_0^{(1)}} & \vdots & -\left(D_0^{(2)} + \frac{1}{2}S_{D_0^{(2)}}\right) \end{pmatrix}. \tag{11}$$

Theorem 2. *For the four designs constructed in Algorithms 1 and 2,*

(i) *for an odd integer N_0, D_1 in (8) is a $D(2N_0 - 2, (N_0 - 1)^m)$, and D_2 in (9) is a $D(2N_0 - 1, 1, N_0^m)$; for an even integer N_0, D_3 in (10) is a $D(2N_0, N_0^m)$, and D_4 in (11) is a $D(2N_0 + 1, 1, (N_0 + 1)^m)$, where $m = \phi(N_0 + 1)$;*
(ii) *they all satisfy property* (b)*, and each column in $D_i^{(1)}$ is column-orthogonal to the columns in $D_i^{(2)}$, for $i = 1, 2, 3, 4$;*
(iii) *D_2 has the same L_1-distance as the good lattice point set with $N_0 + 1$ runs. In particular, if $N_0 + 1 = 2p_1$ for an odd prime p_1, then $d_1(D_2) = (p_1 - 1)^2/2$; if $N_0 + 1 = 2^t$, then $d_1(D_2) = (N_0 + 1)^2/8$; and*
(iv) *for an even integer N_0, if $\phi(N_0 + 1) \geq \phi(N_0 + 2)$, then D_3 or D_4 contains more columns than the corresponding D_1 or D_2 with the same run size.*

Note that for an even integer N_0, $\phi(N_0 + 1) \geq \phi(N_0 + 2)$ may not always be true, and one may occasionally have $\phi(N_0^* + 1) < \phi(N_0^* + 2)$ for some even N_0^*. When this happens, for given run size $2N_0^*$ or $2N_0^* + 1$, we can increase the number of columns of the corresponding design by taking $N_0 = N_0^* + 1$ in Algorithm 1, finally, which design will be used in the experiments depends on the actual circumstances.

For a design $D(N, s^m)$, if its any two-dimensional projection has equal number of points in each grid of the $s_1 \times s_2$ ($s_1, s_2 \leq s$) grids, we say it achieves stratification on $s_1 \times s_2$ grids.

Theorem 3. *For an odd integer N_0, D_1 in (8) achieves stratification on $(N_0 - 1)/2 \times 2$ or $2 \times (N_0 - 1)/2$ grids. For an even integer N_0, D_3 in (10) achieves stratification on $N_0/2 \times 2$ or $2 \times N_0/2$ grids.*

The proof of Theorem 3 also implies the following result.

Corollary 1. *For an odd integer N_0, the projection of D_1 in (8) onto any two columns from different groups also achieves stratification on $(N_0 - 1) \times 2$ or $2 \times (N_0 - 1)$ grids. For an even integer N_0, the projection of D_3 in (10) onto any two columns from different groups also achieves stratification on $N_0 \times 2$ or $2 \times N_0$ grids.*

Corollary 2. *If we ignore the zero point, D_2 in (9) achieves stratification on $(N_0 - 1)/2 \times 2$ or $2 \times (N_0 - 1)/2$ grids, and D_4 in (11) achieves stratification on $N_0/2 \times 2$ or $2 \times N_0/2$ grids.*

It is noted that the newly constructed design has a better stratification property and a comparable space-filling property under L_1-distance.

3.3. An example and some comparisons

Example 2. For $N_0 = 7$, the 12- and 13-run designs generated by Algorithm 1 are given in Table 3. For $N_0 = 6$, the 12- and 13-run designs generated by Algorithm 2 are given in Table 4; the corresponding design D_0 is the LHD(6, 6) in Table 2.

Fig. 2(a) shows the two-dimensional projections of the 12-run design in Table 3; Fig. 2(b) shows the two-dimensional projections of the 12-run design in Table 4, where the stratification properties given in Theorem 3 can be easily verified. Fig. 3 shows illustrations of Corollary 1 for the two designs $D(12, 6^4)$ and $D(12, 6^6)$, respectively. Fig. 4 shows illustrations of Corollary 2 for the two designs $D(13, 1, 7^4)$ and $D(13, 1, 7^6)$, respectively.

For comparison, take a look at the two-dimensional projections of the original LHD(6, 6), denoted by D_0, in Fig. 1(a), it can be easily seen that D_0 does not have the property in Theorem 3, as neither (X_1, X_6), (X_2, X_5) nor (X_3, X_4) has equal number of points in the 3×2 or 2×3 grids. Also, the stratification property in Corollary 2 does not hold for the design in Fig. 1(b).

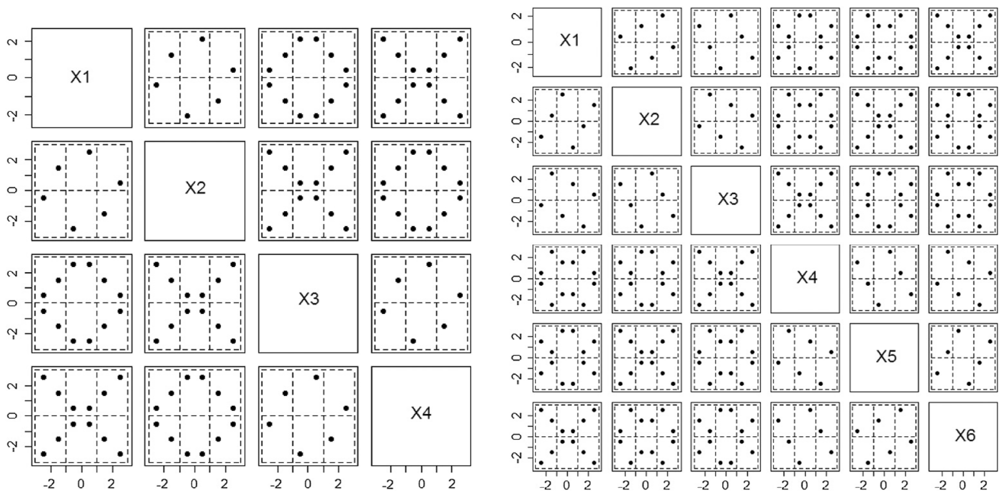
Some useful designs constructed from Algorithms 1 and 2 for run sizes less than 50 are tabulated in Table 5. It also lists the correlations among columns for each case, where N is the number of rows of the new design D, m is the number of columns, and s is the number of levels in each column of D, $\rho^2(D)$ and $\rho_M(D)$ are the two measures in (1) for evaluating the orthogonality of the new design D. From Table 5, it is easy to see that the new designs perform well under the criterion of

Table 3
Designs generated by Algorithm 1 for $N_0 = 7$.

$D(12, 6^4)$				$D(13, 1, 7^4)$			
−2.5	−0.5	0.5	2.5	−3	−1	1	3
−1.5	1.5	−1.5	1.5	−2	2	−2	2
−0.5	−2.5	2.5	0.5	−1	−3	3	1
0.5	2.5	−2.5	−0.5	0	0	0	0
1.5	−1.5	1.5	−1.5	1	3	−3	−1
2.5	0.5	−0.5	−2.5	2	−2	2	−2
−2.5	−0.5	−0.5	−2.5	3	1	−1	−3
−1.5	1.5	1.5	−1.5	−3	−1	−1	−3
−0.5	−2.5	−2.5	−0.5	−2	2	2	−2
0.5	2.5	2.5	0.5	−1	−3	−3	−1
1.5	−1.5	−1.5	1.5	1	3	3	1
2.5	0.5	0.5	2.5	2	−2	−2	2
				3	1	1	3

Table 4
Designs generated by Algorithm 2 for $N_0 = 6$.

$D(12, 6^6)$						$D(13, 1, 7^6)$					
−2.5	−1.5	−0.5	0.5	1.5	2.5	−3	−2	−1	1	2	3
−1.5	0.5	2.5	−2.5	−0.5	1.5	−2	1	3	−3	−1	2
−0.5	2.5	−1.5	1.5	−2.5	0.5	−1	3	−2	2	−3	1
0.5	−2.5	1.5	−1.5	2.5	−0.5	1	−3	2	−2	3	−1
1.5	−0.5	−2.5	2.5	0.5	−1.5	2	−1	−3	3	1	−2
2.5	1.5	0.5	−0.5	−1.5	−2.5	3	2	1	−1	−2	−3
−2.5	−1.5	−0.5	−0.5	−1.5	−2.5	0	0	0	0	0	0
−1.5	0.5	2.5	2.5	0.5	−1.5	−3	−2	−1	−1	−2	−3
−0.5	2.5	−1.5	−1.5	2.5	−0.5	−2	1	3	3	1	−2
0.5	−2.5	1.5	1.5	−2.5	0.5	−1	3	−2	−2	3	−1
1.5	−0.5	−2.5	−2.5	−0.5	1.5	1	−3	2	2	−3	1
2.5	1.5	0.5	0.5	1.5	2.5	2	−1	−3	−3	−1	2
						3	2	1	1	2	3



(a) $D(12, 6^4)$ in Table 3.

(b) $D(12, 6^6)$ in Table 4.

Fig. 2. Two-dimensional projections of the 12-run designs in Tables 3 and 4, respectively.

$\rho^2(D)$, since the values of $\rho^2(D)$ are all less than 0.05. From the last column of Table 5, it can be seen that the percentage of the absolute correlations among columns that do not exceed 0.1 is larger than 60% in most cases. In almost all cases but $N = 28, 29$, designs constructed in Algorithm 2 contain more columns than the corresponding designs constructed in Algorithm 1, which is ensured by Theorem 2(iv).

Comparisons of correlations between columns of the newly constructed designs and the corresponding initial leave-one-out good lattice point sets are shown in Figs. 5 and 6. Fig. 5 plots the percentages of the low absolute correlations (not larger than 0.1) among all absolute correlations for the new designs constructed in Algorithm 1 and the corresponding initial

Table 5
Newly constructed designs by Algorithms 1 and 2 for $N < 50$.

Design	N	m	s	$\rho^2(D)$	$\rho_M(D)$	Percentage of $\lvert\rho(i,j)\rvert \le 0.1$*
D_1 in (8)	12	4	6	0.0002	0.0285	1.0000
	16	4	8	0.0000	0.0000	1.0000
	20	4	10	0.0027	0.0909	1.0000
	24	6	12	0.0164	0.2027	0.6000
	28	8	14	0.0085	0.1428	0.5714
	32	6	16	0.0019	0.0705	1.0000
	36	8	18	0.0218	0.2507	0.5714
	40	10	20	0.0099	0.2000	0.7777
	44	8	22	0.0111	0.2004	0.7143
	48	12	24	0.0150	0.2730	0.8182
D_3 in (10)	12	6	6	0.0160	0.2000	0.6000
	16	6	8	0.0326	0.2857	0.6000
	20	10	10	0.0335	0.3333	0.5555
	24	12	12	0.0255	0.3636	0.8182
	28	8	14	0.0484	0.3846	0.5714
	32	16	16	0.0300	0.4000	0.7333
	36	18	18	0.0294	0.4117	0.6470
	40	12	20	0.0328	0.4210	0.8182
	44	22	22	0.0300	0.4285	0.7143
	48	20	24	0.0264	0.4347	0.7895
D_2 in (9)	13	4	7	0.0068	0.1428	0.6666
	17	4	9	0.0000	0.0000	1.0000
	21	4	11	0.0001	0.0181	1.0000
	25	6	13	0.0213	0.2307	0.6000
	29	8	15	0.0068	0.1428	0.7143
	33	6	17	0.0013	0.0588	1.0000
	37	8	19	0.0219	0.2631	0.5714
	41	10	21	0.0105	0.2000	0.7778
	45	8	23	0.0101	0.2094	0.7143
	49	12	25	0.0154	0.2800	0.8182
D_4 in (11)	13	6	7	0.0020	0.0714	1.0000
	17	6	9	0.0217	0.2333	0.6000
	21	10	11	0.0232	0.2545	0.5555
	25	12	13	0.0211	0.3186	0.6364
	29	8	15	0.0350	0.3285	0.5714
	33	16	17	0.0265	0.3627	0.7333
	37	18	19	0.0257	0.3684	0.6470
	41	12	21	0.0279	0.3896	0.8182
	45	22	23	0.0263	0.3932	0.7143
	49	20	25	0.0242	0.4076	0.7895

* Percentage of the absolute correlations not larger than 0.1 among all correlations.

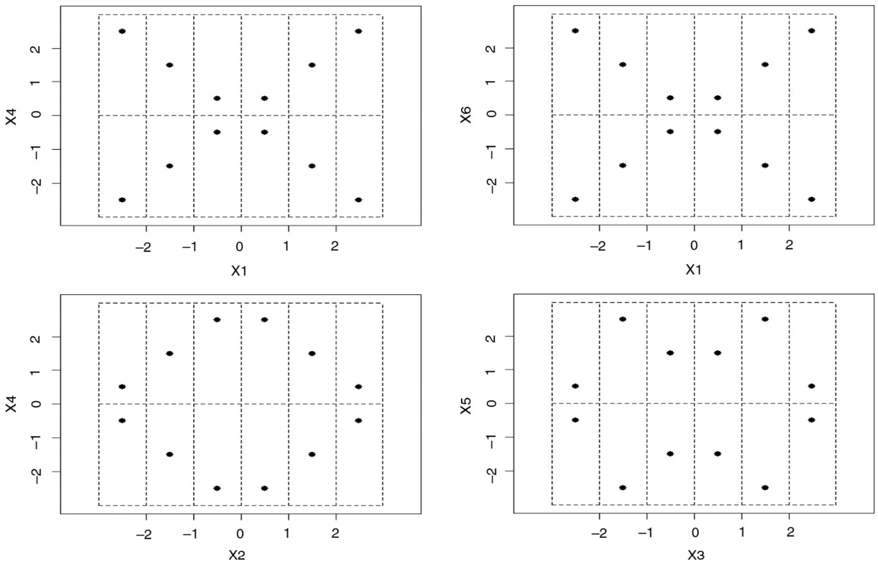
designs. Fig. 6 plots the percentages of the low absolute correlations (not larger than 0.1) among all absolute correlations for the new designs constructed in Algorithm 2 and the corresponding initial designs. The run sizes of the new designs are ranging from 12 to 1000 (there are about 250 such designs), and the run sizes of corresponding initial designs are ranging from 6 to 500. It is easy to see that the newly constructed designs have more column pairs with low absolute correlations (not larger than 0.1).

In practice, when the design has N runs, u ($u < \phi(N + 1)$) factors, and s levels, there exist two possible designs. For example, when $(N, u, s) = (12, 4, 6)$, there are two ways to generate the design: (1) using Algorithm 1 to obtain the design directly, and (2) using Algorithm 2 to generate the base design, then choose four columns from the base design (under a criterion like $\rho^2(D)$ or $\rho_M(D)$). It is found that there are only a few cases that designs obtained from Algorithm 2 are strictly superior to the corresponding designs from Algorithm 1; in most cases, designs from Algorithm 1 are better than the best one obtained from Algorithm 2.

4. Concluding remarks

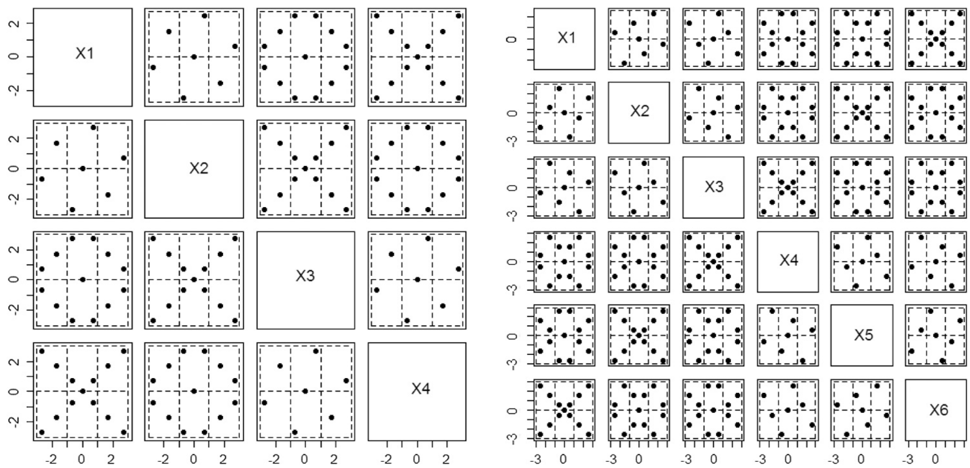
In this paper, we propose algorithms to construct designs from the leave-one-out good lattice point method. The new designs have a better stratification property than the original good lattice point sets, meanwhile they retain the space-filling property of good lattice point sets. Zhou and Xu (2015) studied the space-filling properties of good lattice point sets and showed that designs generated by the leave-one-out good lattice point method have large L_1-distances. The newly constructed design D_2 in (9) also has a large L_1-distance, as proved in Theorem 2.

The designs generated by the leave-one-out good lattice point method by Zhou and Xu (2015) will result in negatively fully correlated columns. It is shown that the newly constructed designs are able to convert such negatively fully correlated



(a) $D(12, 6^4)$ in Table 3. (b) $D(12, 6^6)$ in Table 4.

Fig. 3. Two-dimensional projections of the 12-run designs for columns from different groups.



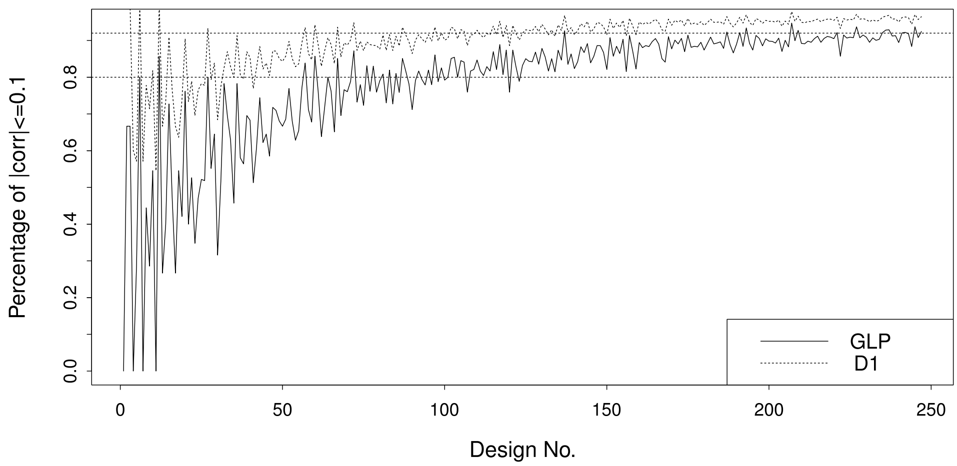
(a) $D(13, 1, 7^4)$ in Table 3. (b) $D(13, 1, 7^6)$ in Table 4.

Fig. 4. Two-dimensional projections of the 13-run designs in Tables 3 and 4, respectively.

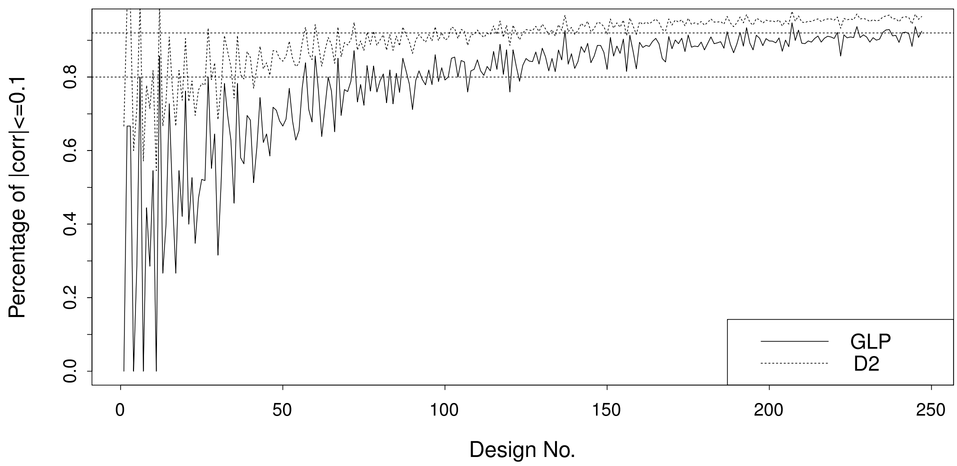
columns into column-orthogonal ones. The columns of the new designs can be divided into two groups such that any two columns are column-orthogonal when they are from different groups and nearly column-orthogonal when they are in the same group. Moreover, the proposed designs have a fold-over structure and thus satisfy property (b), which makes the linear effects estimable when second-order effects are present in the polynomial model. Some general theoretical results on the newly constructed designs are established.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 11271205, 11431006, 11401321 and 11501305), the Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. 20130031110002), the "131" Talents Program of Tianjin, and National Security Agent via Grant H98230-15-1-0253. The authors thank the Executive-Editor, an Associate-Editor and three referees for their valuable comments.



(a) The designs from Algorithm 1 vs. the corresponding initial designs.



(b) The designs from Algorithm 1 vs. the corresponding initial designs.

Fig. 5. Comparisons of the percentages of low correlation column pairs: the newly constructed designs vs. the corresponding initial designs.

Appendix. Proofs

A.1. Proof of Lemma 2

For the LHD(N, m) generated by the leave-one-out good lattice point method, it is easy to know that: if h_j is coprime to $(N+1)$, then $N+1-h_j$ is also coprime to $(N+1)$, and $h_{m+1-j} = N+1-h_j$. As

$$kh_j \pmod{(N+1)} + (N+1-k)h_j \pmod{(N+1)} = N+1, \quad \text{and}$$
$$kh_j \pmod{(N+1)} + kh_{m+1-j} \pmod{(N+1)} = N+1,$$

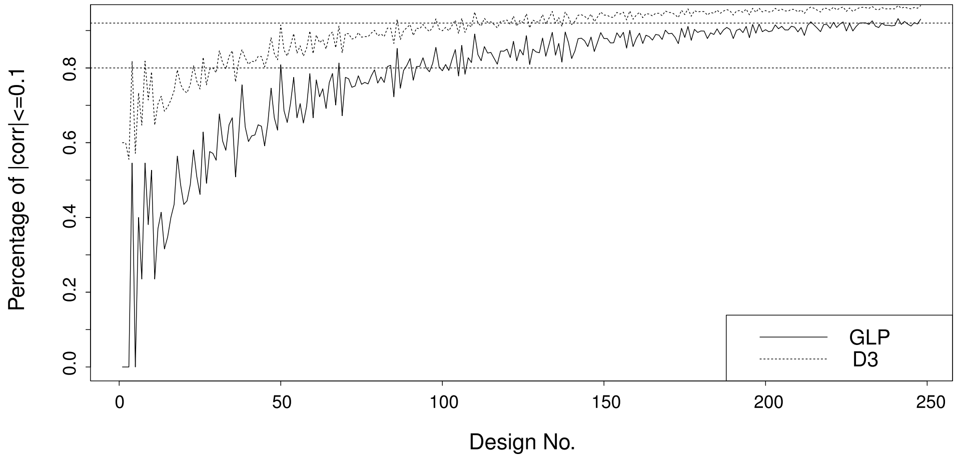
for $k = 1, \ldots, N,\ j = 1, \ldots, m$, we then have (2) and (3) directly, this completes the proof.

A.2. Proof of Theorem 2

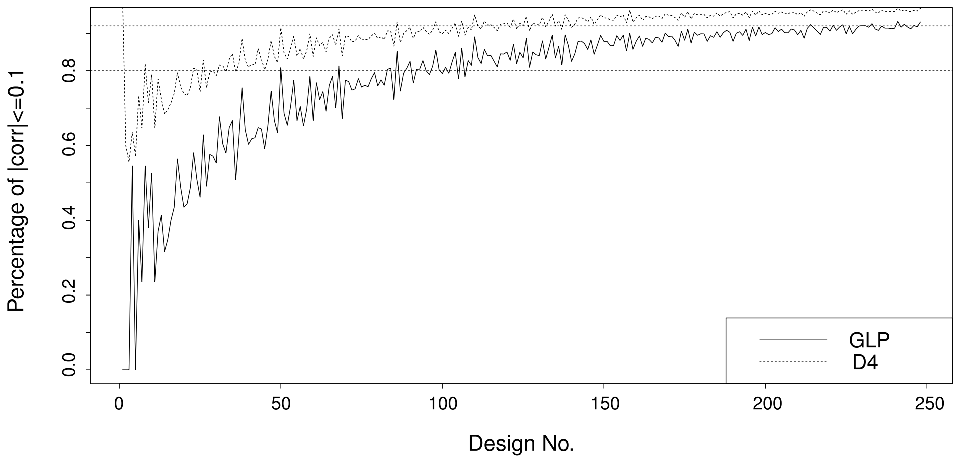
(i) and (iv) are obvious, so we only prove (ii) and (iii).

(ii) From (6), (7), (8), (9), (10), (11), we can easily verify that all of D_1, D_2, D_3, D_4 have a fold-over structure and thus satisfy property (b). For any one column from $D_i^{(1)}$ and another one from $D_i^{(2)}$, $i = 1, 2, 3, 4$, from Algorithms 1 and 2, they can be expressed as

$$\begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{a} & -\boldsymbol{b} \end{pmatrix} \quad \text{for } i = 1, 3, \quad \text{or} \quad \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \\ 0 & 0 \\ \boldsymbol{a} & -\boldsymbol{b} \end{pmatrix} \quad \text{for } i = 2, 4,$$



(a) The designs from Algorithm 2 vs. the corresponding initial designs.



(b) The designs from Algorithm 2 vs. the corresponding initial designs.

Fig. 6. Comparisons of the percentages of low correlation column pairs: the newly constructed designs vs. the corresponding initial designs.

where for an odd integer N_0, $\boldsymbol{a}$ and $\boldsymbol{b}$ are two $(N_0 - 1) \times 1$ vectors with all elements nonzero; for an even integer N_0, $\boldsymbol{a}$ and $\boldsymbol{b}$ are two $N_0 \times 1$ vectors with all elements nonzero. This implies that each column in $D_i^{(1)}$ is column-orthogonal to the columns in $D_i^{(2)}$, $i = 1, 2, 3, 4$.

(iii) From (7), we have that, up to the orders of the rows and columns, D_2 in (9) can be written as:

$$D_2 = (D_2^{(1)} \vdots D_2^{(2)}) = \begin{pmatrix} D_0^{(1)} & \vdots & D_0^{(2)} \\ D_0^{(1)} & \vdots & -D_0^{(2)} \end{pmatrix} \setminus \mathbf{0}_m^T = \begin{pmatrix} E & \vdots & -E \\ -E & \vdots & E \\ \mathbf{0}_{m/2}^T & \vdots & \mathbf{0}_{m/2}^T \\ E & \vdots & E \\ -E & \vdots & -E \end{pmatrix}. \quad (12)$$

Let $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{i,m/2})$ and $\boldsymbol{e}_j = (e_{j1}, \ldots, e_{j,m/2})$ be any two rows of E, $i, j = 1, \ldots, (N_0 - 1)/2$. Then, for any two rows of D_0, their L_1-distance can be expressed as

$$2(|e_{i1} - e_{j1}| + \cdots + |e_{i,m/2} - e_{j,m/2}|) \quad (i \neq j), \quad \text{or} \quad (13)$$

$$2(|e_{i1} + e_{j1}| + \cdots + |e_{i,m/2} + e_{j,m/2}|), \quad \text{or} \quad (14)$$

$$2(|e_{i1}| + \cdots + |e_{i,m/2}|), \quad (15)$$

and for any two rows of D_2, their L_1-distance can be expressed as (13), (14), (15) or [(13)+(14)]/2. Since D_0 is an LHD(N_0, m), both L_1-distances in (13) and (14) are greater than zero, thus $[(13) + (14)]/2 > \min\{(13), (14)\}$, which implies that D_2 has the same L_1-distance as D_0, and hence D_2 has the same L_1-distance as the good lattice point set with N_0+1 runs. Moreover, the L_1-distances of D_2 for the two cases can be directly obtained from Lemma 1. This completes the proof.

A.3. Proof of Theorem 3

For an even integer N_0, from Lemma 2 and Remark 1, up to the orders of the rows and columns, D_3 can be re-expressed as

$$D_3 = (D_3^{(1)} \vdots D_3^{(2)}) = \begin{pmatrix} D_0^{(1)} & \vdots & D_0^{(2)} \\ D_0^{(1)} & \vdots & -D_0^{(2)} \end{pmatrix} = \begin{pmatrix} E & \vdots & -E \\ -E & \vdots & E \\ E & \vdots & E \\ -E & \vdots & -E \end{pmatrix}. \tag{16}$$

For any two columns of D_3, without loss of generality, assume they correspond to the ith and jth columns of E, $i,\ j = 1, \ldots, m/2$, then all possible level combinations in these two columns can be expressed as

(i) $(e_{ki}, e_{kj}), (-e_{ki}, -e_{kj})$ for $k = 1, \ldots, N_0/2$, when they are in the same group, and
(ii) $(e_{ki}, -e_{kj}), (-e_{ki}, e_{kj}), (e_{ki}, e_{kj}), (-e_{ki}, -e_{kj})$ for $k = 1, \ldots, N_0/2$, when they are from different groups.

Then, according to properties of the good lattice point set, it is easy to verify that D_3 achieves stratification on $N_0/2 \times 2$ or $2 \times N_0/2$ grids for both cases.

Similarly, we can prove that for an odd integer N_0, D_1 in (8) achieves stratification on $(N_0-1)/2 \times 2$ or $2 \times (N_0-1)/2$ grids. This completes the proof.

References

Bingham, D., Sitter, R.R., Tang, B., 2009. Orthogonal and nearly orthogonal designs for computer experiments. Biometrika 96, 51–65.
Fang, K.T., Li, R., Sudjianto, A., 2006. Design and Modeling for Computer Experiments. Chapman & Hall/CRC, New York.
Fang, K.T., Wang, Y., 1981. A note on uniform distribution and experimental design. Chinese Sci. Bull. 26, 485–489.
Fang, K.T., Wang, Y., 1994. Number-Theoretic Methods in Statistics. Chapman & Hall, London.
Georgiou, S.D., 2011. Orthogonal designs for computer experiments. J. Statist. Plann. Inference 141, 1519–1525.
Georgiou, S.D., Koukouvinos, C., Liu, M.Q., 2014. U-type and column-orthogonal designs for computer experiments. Metrika 77, 1057–1073.
He, Y., Tang, B., 2013. Strong orthogonal arrays and associated Latin hypercubes for computer experiments. Biometrika 100, 254–260.
Johnson, M.E., Moore, L.M., Ylvisaker, D., 1990. Minimax and maximin distance designs. J. Statist. Plann. Inference 26, 131–148.
Korobov, N., 1959. The approximate computation of multiple integrals. Dokl. Akad. Nauk SSSR 124, 1207–1210.
Lin, C.D., Bingham, D., Sitter, R.R., Tang, B., 2010. A new and flexible method for constructing designs for computer experiments. Ann. Statist. 38, 1460–1477.
Lin, C.D., Mukerjee, R., Tang, B., 2009. Construction of orthogonal and nearly orthogonal Latin hypercubes. Biometrika 96, 243–247.
McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21, 239–245.
Mukerjee, R., Sun, F.S., Tang, B., 2014. Nearly orthogonal arrays mappable into fully orthogonal arrays. Biometrika 101, 957–963.
Owen, A.B., 1992. Orthogonal arrays for computer experiments, integration and visualization. Statist. Sinica 2, 439–452.
Pang, F., Liu, M.Q., 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., Lin, D.K.J., 2006. A construction method for orthogonal Latin hypercube designs. Biometrika 93, 279–288.
Sun, F.S., Liu, M.Q., Lin, D.K.J., 2009. Construction of orthogonal Latin hypercube designs. Biometrika 96, 971–974.
Sun, F.S., Liu, M.Q., 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., Liu, M.Q., 2011. Construction of column-orthogonal designs for computer experiments. Sci. China Math. 54, 2683–2692.
Tang, B., 1993. Orthogonal array-based Latin hypercubes. J. Amer. Statist. Assoc. 88, 1392–1397.
Wang, L., Yang, J.F., Lin, D.K.J., Liu, M.Q., 2015. Nearly orthogonal Latin hypercube designs for many design columns. Statist. Sinica 25, 1599–1612.
Yang, J.Y., 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 application in computer experiments. J. Amer. Statist. Assoc. 93, 1430–1439.
Ye, K.Q., Li, W., Sudjianto, A., 2000. Algorithmic construction of optimal symmetric Latin hypercube designs. J. Statist. Plann. Inference 90, 145–159.
Zhou, Y.D., Xu, H., 2015. Space-filling properties of good lattice point sets. Biometrika 102, 959–966.