

Construction of column-orthogonal designs for computer experiments

SUN FaSheng^{1,2}, PANG Fang² & LIU MinQian^{2,*}

¹*Department of Statistics, KLAS and School of Mathematics and Statistics,
Northeast Normal University, Changchun 130024, China;*

²*Department of Statistics, School of Mathematical Sciences and LPMC,
Nankai University, Tianjin 300071, China*

Email: sfxsfx2001@gmail.com, pangfang@mail.nankai.edu.cn, mqliu@nankai.edu.cn

Received January 27, 2010; accepted April 8, 2011; published online October 18, 2011

Abstract Latin hypercube design and uniform design are two kinds of most popular space-filling designs for computer experiments. The fact that the run size equals the number of factor levels in a Latin hypercube design makes it difficult to be orthogonal. While for a uniform design, it usually has good space-filling properties, but does not necessarily have small or zero correlations between factors. In this paper, we construct a class of column-orthogonal and nearly column-orthogonal designs for computer experiments by rotating groups of factors of orthogonal arrays, which supplement the designs for computer experiments in terms of various run sizes and numbers of factor levels and are flexible in accommodating various combinations of factors with different numbers of levels. The resulting column-orthogonal designs not only have uniformly spaced levels for each factor but also have uncorrelated estimates of the linear effects in first order models. Further, they are 3-orthogonal if the corresponding orthogonal arrays have strength equal to or greater than three. Along with a large factor-to-run ratio, these newly constructed designs are economical and suitable for screening factors for physical experiments.

Keywords computer experiment, Latin hypercube design, orthogonal array, rotation, uniform design

MSC(2000): 62K15, 62K20

Citation: Sun F S, Pang F, Liu M Q. Construction of column-orthogonal designs for computer experiments. *Sci China Math*, 2011, 54(12): 2683–2692, doi: 10.1007/s11425-011-4284-8

1 Introduction

Many physical phenomena encountered in science and engineering are governed by a set of complicated equations. These equations often have only numerical solutions that are carried out by computer programs. Latin hypercube design (LHD) and uniform design are two kinds of most popular space-filling designs for computer experiments (see [9, 27]). LHDs were introduced by [25]. The fact that each factor in an LHD has as many uniformly spaced levels as its run size makes it attractive in that the design achieves the maximum stratification when projected into any univariate dimension. Efforts have been made to find orthogonal or nearly orthogonal LHDs, see e.g., [1–7, 18, 19, 26, 28, 31–34, 36, 37]. However, the factors in an LHD have as many levels as the run size, which makes it very difficult for an LHD to be orthogonal (see [4]). Uniform designs were proposed by [8] and [35], and have received great attention in recent decades, see e.g., [9, 10, 13, 29] and the references therein. A uniform design seeks design points that are uniformly scattered on the design domain, it is robust against the model specification (see [12]) and

*Corresponding author

limits the effects of aliasing to yield reasonable efficiency and robustness together (see [16]). However, a uniform design does not necessarily have small or zero correlations between factors.

For computer experiments, practical experiments have revealed that designs with many levels are desirable, but it is not essential that the run size equals the number of levels at which each factor is observed (see [4]), as in an LHD. As we know, screening important factors and then estimating the effects accurately are the main objectives of experimental designs. Therefore, lower correlations among effect estimates are preferred, which will achieve the lowest correlation when the model matrix is orthogonal. By relaxing the condition that the number of levels for each factor must be identical to the run size, we, in this paper, propose some methods to construct column-orthogonal designs and nearly column-orthogonal designs, which not only have uniformly spaced levels for each factor but also have some other attractive properties, as will be discussed later.

The paper is organized as follows. Section 2 provides some notations and related work on rotation designs. The construction methods for column-orthogonal designs are proposed in Section 3. Section 4 constructs a new class of nearly column-orthogonal designs. Some discussions and concluding remarks are provided in Section 5.

2 Some notations and related work on rotation designs

A design with n runs and m factors, each having q_1, \dots, q_m levels, respectively, is denoted by $D(n, q_1 \cdots q_m)$. A $D(n, q_1 \cdots q_m)$ design is an $n \times m$ matrix with entries of the j th column from a set of q_j symbols, which are assumed here to be $\{(2i - q_j - 1)/2, i = 1, \dots, q_j\}$ for odd q_j and $\{2i - q_j - 1, i = 1, \dots, q_j\}$ for even q_j . If in each column the symbols occur equally often, the design is called a U-type design. The q_j 's are not necessarily distinct, for example, a $D(n, q_1^{m_1} q_2^{m_2})$ is a design that has m_1 factors of q_1 levels and m_2 factors of q_2 levels. In particular, when all the q_j 's are equal, the design is said to be symmetrical, otherwise, asymmetrical. A $D(n, n^m)$ is called an LHD and denoted by $\text{LHD}(n, m)$. A U-type design $D(n, q_1 \cdots q_m)$ is called a column-orthogonal design, denoted by $\text{COD}(n, q_1 \cdots q_m)$, if the inner product of any two columns is zero; and is called an orthogonal array of strength t , denoted by $\text{OA}(n, q_1 \cdots q_m, t)$, if all possible level-combinations for any t columns appear equally often. We shall call the latter orthogonality *combinatorial orthogonality* to distinguish it from the *column-orthogonality*. Clearly, the combinatorial orthogonality implies the column-orthogonality, but the inverse is not necessarily true. Furthermore, a column-orthogonal design is called 3-orthogonal (see [4]) if the sum of elementwise products of any three columns (whether they are distinct or not) is zero.

Let X denote the regression matrix for the first-order model of a column-orthogonal design with m factors, including a column of ones and the m factors in the design. Let X_{int} denote the $n \times m(m-1)/2$ matrix with all the possible bilinear interactions, and let X_{quad} denote the $n \times m$ matrix with all the pure quadratic terms. The alias matrices for the first-order model associated with the bilinear interactions and the pure quadratic terms are then given by $(X'X)^{-1}X'X_{\text{int}}$ and $(X'X)^{-1}X'X_{\text{quad}}$, respectively. A good design for factor screening should maintain relatively small terms in these alias matrices (see [28]). It is easy to see that if a column-orthogonal design is 3-orthogonal, then these two alias matrices are both zero matrices.

[1–3] showed that a class of LHDs can be constructed by rotating the points in d -factor, q -level standard full factorial designs, where d is a power of 2, and defined a sequence of rotation matrices by a recursive scheme. [5] proposed the idea of independently rotating groups of factors in two-level designs. Recently, [26, 28] combined the above two ideas with the knowledge of Galois field to produce the orthogonal LHD matrix with $n = q^d$ runs, where q is a prime and d is a power of 2. This severe run size constraint is the primary limitation to their rotation methods.

Let $R_0^q = 1$, and

$$R_c^q = \begin{pmatrix} q^{2^{c-1}} R_{c-1}^q & -R_{c-1}^q \\ R_{c-1}^q & q^{2^{c-1}} R_{c-1}^q \end{pmatrix} \quad \text{for } c = 1, 2, \dots, \quad (1)$$

then we have

Lemma 1 (cf. [26]). *The matrix R_c^q in (1) is a rotation of the d -factor ($d = 2^c$), q -level standard full factorial design which yields unique and equally-spaced projections to each dimension.*

Remark 1. Here, we relax the definition of rotation to be a matrix R satisfying $R'R = kI$ for some scalar k , instead of $R'R = I$. It can be easily checked that the matrix R_c^q in (1) consists of columns (and rows) of permutations of $\{1, q, \dots, q^{2^c-1}\}$ up to sign changes, which guarantees that, for a 2^c -factor q -level standard full factorial design A , AR_c^q yields unique and equally-spaced projections to each dimension.

This paper will extend the rotation method to orthogonal arrays for accommodating various run sizes. Though the obtained designs are not always LHDs, they are column-orthogonal designs or nearly column-orthogonal designs, and the factors have enough levels to be employed in computer experiments.

3 Construction of column-orthogonal designs

In this section, we present the construction methods for column-orthogonal designs by rotating symmetrical as well as asymmetrical orthogonal arrays.

3.1 Construction from symmetrical orthogonal arrays

For convenience, we denote

$$R_{(c_1, \dots, c_v)}^q = \text{diag}\{R_{c_1}^q, \dots, R_{c_v}^q\}, \tag{2}$$

and

$$R_{m,q} = \text{diag}\{R_1^q, \dots, R_1^q\}, \tag{3}$$

where R_c^q is given in (1) and R_1^q occurs $m/2$ times in the diagonal of $R_{m,q}$.

Theorem 1. *Suppose A is an $OA(n, q^m, t)$ with $m = 2k$ and $t \geq 2$, $D = AR_{m,q}$, then*

- (1) D is a $COD(n, (q^2)^m)$;
- (2) if $t \geq 3$, D is a 3-orthogonal $COD(n, (q^2)^m)$.

The proof of Theorem 1 is given in the Appendix. Now, let us see some illustrative examples.

Example 1. Suppose A is an $OA(12, 2^{11}, 2)$, A_1 consists of the first 10 columns of A and $D = A_1R_{10,2}$, then D is a $COD(12, 4^{10})$. The $OA(12, 2^{10}, 2)$ and $COD(12, 4^{10})$ are listed in Table 1.

Example 2. Suppose A is an $OA(18, 3^7, 2)$, A_1 consists of the first 6 columns of A and $D = A_1R_{6,3}$, then D is a $COD(18, 9^6)$. A_1 and D are shown in Table 2.

Example 3. Suppose A is an $OA(24, 2^{12}, 3)$, and $D = AR_{12,2}$, then D is a 3-orthogonal $COD(24, 4^{12})$. A and D are shown in Table 3.

Table 1 $OA(12, 2^{10}, 2)$ and $COD(12, 4^{10})$

OA(12, 2 ¹⁰ , 2)										COD(12, 4 ¹⁰)									
1	1	1	1	1	1	1	1	1	1	3	1	3	1	3	1	3	1	3	1
-1	1	-1	1	1	1	-1	-1	-1	1	-1	3	-1	3	3	1	-3	-1	-1	3
-1	-1	1	-1	1	1	1	-1	-1	-1	-3	-1	1	-3	3	1	1	-3	-3	-1
1	-1	-1	1	-1	1	1	1	-1	-1	1	-3	-1	3	-1	3	3	1	-3	-1
-1	1	-1	-1	1	-1	1	1	1	-1	-1	3	-3	-1	1	-3	3	1	1	-3
-1	-1	1	-1	-1	1	-1	1	1	1	-3	-1	1	-3	-1	3	-1	3	3	1
-1	-1	-1	1	-1	-1	1	-1	1	1	-3	-1	-1	3	-3	-1	1	-3	3	1
1	-1	-1	-1	1	-1	-1	1	-1	1	1	-3	-3	-1	1	-3	-1	3	-1	3
1	1	-1	-1	-1	1	-1	-1	1	-1	3	1	-3	-1	-1	3	-3	-1	1	-3
1	1	1	-1	-1	-1	1	-1	-1	1	3	1	1	-3	-3	-1	1	-3	-1	3
-1	1	1	1	-1	-1	-1	1	-1	-1	-1	3	3	1	-3	-1	-1	3	-3	-1
1	-1	1	1	1	-1	-1	-1	1	-1	1	-3	3	1	1	-3	-3	-1	1	-3

Table 2 OA(18, 3⁶, 2) and COD(18, 9⁶)

OA(18, 3 ⁶ , 2)						COD(18, 9 ⁶)					
1	1	1	1	1	1	4	2	4	2	4	2
-1	-1	-1	-1	-1	-1	-4	-2	-4	-2	-4	-2
0	0	0	0	0	0	0	0	0	0	0	0
1	1	-1	0	-1	0	4	2	-3	1	-3	1
-1	-1	0	1	0	1	-4	-2	1	3	1	3
0	0	1	-1	1	-1	0	0	2	-4	2	-4
1	-1	1	0	0	-1	2	-4	3	-1	-1	-3
-1	0	-1	1	1	0	-3	1	-2	4	3	-1
0	1	0	-1	-1	1	1	3	-1	-3	-2	4
1	0	0	1	-1	-1	3	-1	1	3	-4	-2
-1	1	1	-1	0	0	-2	4	2	-4	0	0
0	-1	-1	0	1	1	-1	-3	-3	1	4	2
1	-1	0	-1	1	0	2	-4	-1	-3	3	-1
-1	0	1	0	-1	1	-3	1	3	-1	-2	4
0	1	-1	1	0	-1	1	3	-2	4	-1	-3
1	0	-1	-1	0	1	3	-1	-4	-2	1	3
-1	1	0	0	1	-1	-2	4	0	0	2	-4
0	-1	1	1	-1	0	-1	-3	4	2	-3	1

Table 3 OA(24, 2¹², 3) and COD(24, 4¹²)

OA(24, 2 ¹² , 3)												COD(24, 4 ¹²)											
-1	1	1	1	1	1	1	1	1	1	1	1	-1	3	3	1	3	1	3	1	3	1	3	1
-1	-1	1	-1	1	1	1	-1	-1	-1	1	-1	-3	-1	1	-3	3	1	-3	-3	-1	1	-3	-1
-1	-1	-1	1	-1	1	1	1	-1	-1	-1	1	-3	-1	-1	3	-1	3	3	1	-3	-1	-1	3
-1	1	-1	-1	1	-1	1	1	1	-1	-1	-1	-1	3	-3	-1	1	-3	3	1	1	-3	-3	-1
-1	-1	1	-1	-1	1	-1	1	1	1	-1	-1	-3	-1	1	-3	-1	3	3	1	-3	-1	-1	-1
-1	-1	-1	1	-1	-1	1	-1	1	1	1	-1	-3	-1	-1	3	-3	-1	1	-3	3	1	1	-3
-1	-1	-1	-1	1	-1	-1	1	-1	1	1	1	-3	-1	-3	-1	1	-3	-1	3	-1	3	3	1
-1	1	-1	-1	-1	1	-1	-1	1	-1	1	1	-1	3	-3	-1	-1	3	-3	-1	1	-3	3	1
-1	1	1	-1	-1	-1	1	-1	-1	1	-1	1	-1	3	1	-3	-3	-1	1	-3	-1	3	-1	3
-1	1	1	1	-1	-1	-1	1	-1	-1	1	-1	-3	-1	3	1	-3	-3	-1	1	-3	-1	1	-3
-1	-1	1	1	1	-1	-1	-1	1	-1	-1	1	-3	-1	3	3	1	-3	-1	1	-3	-1	3	-1
-1	1	-1	1	1	1	-1	-1	-1	1	-1	-1	-1	3	-3	-1	-3	-1	-3	-1	3	-3	-1	-1
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	-3	-3	-1	-3	-1	-3	-1	-3	-1	-3	-1
1	1	-1	1	-1	-1	-1	1	1	1	-1	1	3	1	-1	3	-3	-1	3	3	1	-1	3	3
1	1	1	-1	1	-1	-1	-1	1	1	1	-1	3	1	1	-3	-3	-1	3	1	1	-3	1	-3
1	-1	1	1	-1	1	-1	-1	-1	1	1	1	1	-3	3	1	-1	3	-3	-1	-1	3	3	1
1	1	-1	1	1	-1	1	-1	-1	-1	1	1	3	1	-1	3	1	-3	-3	-1	3	1	-3	1
1	1	1	-1	1	1	-1	1	-1	-1	-1	1	3	1	1	-3	3	1	-1	3	-3	-1	-1	3
1	1	1	1	-1	1	1	-1	1	-1	-1	-1	3	1	3	1	-3	1	-3	1	-3	-3	-1	-1
1	-1	1	1	1	-1	1	1	-1	1	-1	-1	1	-3	3	1	1	-3	3	1	-1	3	-3	-1
1	-1	-1	1	1	1	-1	1	1	-1	1	-1	1	-3	-1	3	3	1	-1	3	1	-3	1	-3
1	-1	-1	-1	1	1	1	-1	1	1	-1	1	1	-3	-3	-1	3	1	1	-3	3	1	-1	3
1	1	-1	-1	-1	1	1	1	-1	1	1	-1	3	1	-3	-1	-1	3	3	1	-1	3	1	-3
1	-1	1	-1	-1	-1	1	1	1	-1	1	1	1	-3	1	-3	-3	-1	3	1	1	-3	3	1

Corollary 1. Suppose $q \geq 2$ is a prime power, $l \geq 2$ and $m = (q^l - 1)/(q - 1)$, then a $COD(q^l, (q^2)^k)$ exists, where k is the largest even integer not greater than m .

Proof. It is seen from Theorem 3.20 of [14] that for any prime power $q \geq 2$, an $OA(q^l, q^m, 2)$ exists whenever $l \geq 2$, where $m = (q^l - 1)/(q - 1)$. Thus from Theorem 1, the COD can be obtained by rotating the OA consisting of some k columns of this OA by $R_{k,q}$.

Next, we discuss the construction of asymmetrical column-orthogonal designs by rotating symmetrical orthogonal arrays.

Remark 2. In Example 1, if we take

$$R = \begin{pmatrix} R_{10,2} & 0 \\ 0 & R_0^q \end{pmatrix},$$

then AR is an asymmetrical COD(12, $4^{10}2^1$). From the OA in Example 2, a COD(18, 9^63^1) can also be obtained similarly.

Now, we propose a general method to construct asymmetrical column-orthogonal designs. Suppose A is an OA(n, q^m, t) with $t \geq 2$, in which the strength of the first 2^{c_1} factors is 2^{c_1} , the strength of the next 2^{c_2} factors is $2^{c_2}, \dots$, the strength of the last 2^{c_v} factors is 2^{c_v} , and $2^{c_1} + \dots + 2^{c_v} = m$. Let $D = AR_{(c_1, \dots, c_v)}^q$, where $R_{(c_1, \dots, c_v)}^q$ is defined in (2). Then

- Theorem 2.** (1) D is a COD($n, (q^{2^{c_1}})^{2^{c_1}} \dots (q^{2^{c_v}})^{2^{c_v}}$).
 (2) If $t \geq 3$, then D is a 3-orthogonal COD($n, (q^{2^{c_1}})^{2^{c_1}} \dots (q^{2^{c_v}})^{2^{c_v}}$).

The proof of Theorem 2 can be easily obtained from Lemma 1 along the lines of the proof of Theorem 1. Now, let us see two examples for illustration.

Example 4. Suppose A is an OA(16, $2^{15}, 2$), columns 1–4, 5–8 and 9–12 form three full 2^4 factorial sets, respectively, A_1 consists of the first 14 columns of A , and A_2 consists of the first 12 columns of A . Then, $D = AR_{(2,2,2,1,0)}^2$, $D_1 = A_1R_{(2,2,2,1)}^2$ and $D_2 = A_2R_{(2,2,2)}^2$ are COD(16, $16^{12}4^22^1$), COD(16, $16^{12}4^2$) and COD(16, 16^{12}), respectively. Moreover, the COD(16, 16^{12}) is in fact an orthogonal LHD(16, 12), and each 4-column part is a 3-orthogonal column-orthogonal design. The OA(16, $2^{15}, 2$) and COD(16, $16^{12}4^22^1$) are shown in Table 4.

Methods of partitioning the saturated factorial designs to the maximal number of full factorial sets using the Galois field are provided in [26, 28].

Example 5. Suppose D is a 3_{IV}^{8-4} design with the defining relation $5 = 1^2234, 6 = 12^234, 7 = 123^24, 8 = 1234^2$, then 1, 2, 3, 4 and 5, 6, 7, 8 constitute two full factorial sets, respectively. Let $D_1 = DR_{(2,2)}^3$, then D_1 is a 3-orthogonal COD(81, 81^8) which is in fact an orthogonal LHD(81, 8) with the property that the sum of elementwise products of any three columns is zero.

Table 4 OA(16, $2^{15}, 2$) and COD(16, $16^{12}4^22^1$)

OA(16, $2^{15}, 2$)	COD(16, $16^{12}4^22^1$)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	15 5 9 3 15 5 9 3 15 5 9 3 3 1 1
1 1 1-1 1 1 1-1 1-1-1-1-1-1-1-1	13 1 1-13 13 1 1-13 1-13-13 -1-3-1-1
1 1-1 1 1-1-1 1-1 1 1-1-1-1-1-1	11 7 -7 11 3 -9 -5 15 -3 9 5-15-3-1-1
1 1-1-1 1-1-1-1-1-1-1 1 1 1 1	9 3-15 -5 1-13-13 -1-13 -1 -1 13 3 1 1
1-1 1 1-1 1-1 1-1-1-1-1 1 1-1	7-11 11 7 -5 15 -3 9-15 -5 -9 -3 3 1 -1
1-1 1-1-1 1-1-1-1 1 1 1-1-1 1	5-15 3 -9 -7 11-11 -7 -1 13 13 1-3-1 1
1-1-1 1-1-1 1 1 1-1-1 1-1-1 1	3 -9 -5 15 -9 -3 15 5 3 -9 -5 15-3-1 1
1-1-1-1-1-1 1-1 1 1 1-1 1 1-1	1-13-13 -1-11 -7 7-11 13 1 1-13 3 1 -1
-1 1 1 1-1-1-1-1 1 1-1 1 1-1-1	-1 13 13 1-15 -5 -9 -3 11 7 -7 11 1-3 -1
-1 1 1-1-1-1-1 1 1-1 1-1-1 1 1	-3 9 5-15-13 -1 -1 13 5-15 3 -9-1 3 1
-1 1-1 1-1 1 1-1-1 1-1-1-1 1 1	-5 15 -3 9 -3 9 5-15 -7 11-11 -7-1 3 1
-1 1-1-1-1 1 1 1-1-1 1 1 1-1-1	-7 11-11 -7 -1 13 13 1 -9 -3 15 5 1-3 -1
-1-1 1 1 1-1 1-1-1-1 1-1 1-1 1	-9 -3 15 5 5-15 3 -9-11 -7 7-11 1-3 1
-1-1 1-1 1-1 1 1-1 1-1 1-1 1-1	-11 -7 7-11 7-11 11 7 -5 15 -3 9-1 3 -1
-1-1-1 1 1 1-1-1 1-1 1 1-1 1-1	-13 -1 -1 13 9 3-15 -5 7-11 11 7-1 3 -1
-1-1-1-1 1 1-1 1 1 1-1-1 1-1 1	-15 -5 -9 -3 11 7 -7 11 9 3-15 -5 1-3 1

Table 5 OA(36, 2⁸3⁴6², 2) and COD(36, 4⁸9⁴36²)

OA(36, 2 ⁸ 3 ⁴ 6 ² , 2)												COD(36, 4 ⁸ 9 ⁴ 36 ²)													
-5	-5	0	0	1	1	-1	-1	-1	-1	-1	-1	-35	-15	0	0	4	2	-3	-1	-3	-1	-3	-1	-3	-1
-3	-3	0	0	1	1	1	-1	-1	-1	1	-1	-11	-15	0	0	4	2	1	-3	-3	-1	1	-3	1	-3
-5	-3	1	1	0	0	-1	1	1	1	1	-1	-33	-13	4	2	0	0	-1	3	3	1	1	-3	-3	-1
-3	-5	1	1	0	0	1	1	1	-1	1	-1	-13	-17	4	2	0	0	3	1	1	-3	1	-3	-1	3
-1	3	0	1	0	1	-1	1	-1	-1	1	1	1	1	1	3	1	3	-1	3	-3	-1	3	1	3	1
1	5	0	1	0	1	-1	-1	1	-1	-1	1	11	19	1	3	1	3	-3	-1	1	-3	-1	3	-1	3
-1	5	1	0	1	0	1	-1	-1	1	1	1	-1	31	3	-1	3	-1	1	-3	-1	3	3	1	-1	3
1	3	1	0	1	0	1	1	-1	1	-1	1	9	17	3	-1	3	-1	3	1	-1	3	-1	3	-3	-1
3	-1	0	1	1	0	1	-1	1	1	-1	-1	17	-9	1	3	3	-1	1	-3	3	1	-3	-1	3	1
5	1	0	1	1	0	-1	1	-1	1	-1	-1	31	1	1	3	3	-1	-1	3	-1	3	-3	-1	3	1
3	1	1	0	0	1	1	1	1	-1	-1	1	19	3	3	-1	1	3	3	1	1	-3	-1	3	1	-3
5	-1	1	0	0	1	-1	-1	1	1	1	1	19	-11	3	-1	1	3	-3	-1	3	1	3	1	1	-3
-1	-1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-7	-5	4	2	-4	-2	-3	-1	-3	-1	-3	-1	-3	-1
1	1	1	1	-1	-1	1	-1	-1	-1	1	-1	7	5	4	2	-4	-2	1	-3	-3	-1	1	-3	1	-3
-1	1	-1	-1	1	1	-1	1	1	1	1	-1	-5	7	-4	-2	4	2	-1	3	3	1	1	-3	-3	-1
1	-1	-1	-1	1	1	1	1	1	-1	1	-1	5	-7	-4	-2	4	2	3	1	1	-3	1	-3	-1	3
3	-5	1	-1	1	-1	-1	1	-1	-1	1	1	13	-33	2	-4	2	-4	-1	3	-3	-1	3	1	3	1
5	-3	1	-1	1	-1	-1	-1	1	-1	-1	1	17	-13	2	-4	2	-4	-3	-1	1	-3	-1	3	-1	3
3	-3	-1	1	-1	1	1	-1	-1	1	1	1	15	-11	-2	4	-2	4	1	-3	-1	3	3	1	-1	3
5	-5	-1	1	-1	1	1	1	-1	1	-1	1	15	-35	-2	4	-2	4	3	1	-1	3	-1	3	-3	-1
-5	3	1	-1	-1	1	1	-1	1	1	-1	1	-17	13	2	-4	-2	4	1	-3	3	1	-3	-1	3	1
-3	5	1	-1	-1	1	-1	1	-1	1	-1	1	-13	33	2	-4	-2	4	-1	3	-1	3	-3	-1	3	1
-5	5	-1	1	1	-1	1	1	1	-1	-1	1	-15	35	-2	4	2	-4	3	1	1	-3	-1	3	1	-3
-3	3	-1	1	1	-1	-1	-1	1	1	1	1	-15	11	-2	4	2	-4	-3	-1	3	1	3	1	1	-3
3	3	-1	-1	0	0	-1	-1	-1	-1	-1	-1	11	15	-4	-2	0	0	-3	-1	-3	-1	-3	-1	-3	-1
5	5	-1	-1	0	0	1	-1	-1	-1	1	-1	35	15	-4	-2	0	0	1	-3	-3	-1	1	-3	1	-3
3	5	0	0	-1	-1	1	1	1	1	1	-1	13	17	0	0	-4	-2	-1	3	3	1	1	-3	-3	-1
5	3	0	0	-1	-1	1	1	1	-1	1	-1	33	13	0	0	-4	-2	3	1	1	-3	1	-3	-1	3
-5	-1	-1	0	-1	0	-1	1	-1	-1	1	1	-31	-1	-3	1	-3	1	-1	3	-3	-1	3	1	3	1
-3	1	-1	0	-1	0	-1	-1	1	-1	-1	1	-17	9	-3	1	-3	1	-3	-1	1	-3	-1	3	-1	3
-5	1	0	-1	0	-1	1	-1	-1	1	1	1	-19	11	-1	-3	-1	-3	1	-3	-1	3	3	1	-1	3
-3	-1	0	-1	0	-1	1	1	-1	1	-1	1	-19	-3	-1	-3	-1	-3	3	1	-1	3	-1	3	-3	-1
-1	-5	-1	0	0	-1	1	-1	1	1	-1	-1	-11	-19	-3	1	-1	-3	1	-3	3	1	-3	-1	3	1
1	-3	-1	0	0	-1	-1	1	-1	1	-1	1	3	-19	-3	1	-1	-3	-1	3	-1	3	-3	-1	3	1
-1	-3	0	-1	-1	0	1	1	1	-1	-1	1	-9	-17	-1	-3	-3	1	3	1	1	-3	-1	3	1	-3
1	-5	0	-1	-1	0	-1	-1	1	1	1	1	1	-31	-1	-3	-3	1	-3	-1	3	1	3	1	1	-3

3.2 Construction from asymmetrical orthogonal arrays

We can also obtain column-orthogonal designs by rotating asymmetrical orthogonal arrays. Suppose A is an $OA(n, q_1^{m_1} \dots q_v^{m_v}, t)$ with $t \geq 2$, and $R(q_1^{m_1} \dots q_v^{m_v}) = \text{diag}(R_{m_1, q_1}, \dots, R_{m_v, q_v})$, where R_{m_i, q_i} is defined in (3) and m_i is even, $i = 1, \dots, v$. Let $D = AR(q_1^{m_1} \dots q_v^{m_v})$, then we have the following theorem, which can be proved easily along the lines of the proof of Theorem 1.

Theorem 3. (1) D is a $COD(n, (q_1^2)^{m_1} \dots (q_v^2)^{m_v})$.
 (2) If $t \geq 3$, D is a 3-orthogonal $COD(n, (q_1^2)^{m_1} \dots (q_v^2)^{m_v})$.

Remark 3. Let $A = (A_1, \dots, A_v)$, where A_i is an $OA(n, q_i^{m_i}, t)$ with $t \geq 2$, a more flexible partition of A_i can be obtained according to the discussion that precedes Theorem 2 when we rotate A .

Example 6. Suppose A is an $OA(36, 2^8 3^4 6^2, 2)$, then from Theorem 3 we can construct a $COD(36, 4^8 9^4 36^2)$. The two designs are shown in Table 5.

4 Construction of nearly column-orthogonal designs

In this section, Theorem 1 is modified to construct nearly column-orthogonal designs, which have flexible run sizes and numbers of factors. Two measures defined in [4] are used to assess the near orthogonality of a design D with m columns: the maximum correlation $\rho_M(D) = \max_{i,j} |\rho_{ij}(D)|$ and the average squared correlation $\rho^2(D) = \sum_{i < j} \rho_{ij}^2(D) / [(m(m-1)/2)]$, where $\rho_{ij}(D)$ denotes the correlation coefficient between the i th and j th columns. Firstly, we can easily have

Lemma 2. Suppose A is an $n \times m$ matrix with $A'A = cI_m$, and $D = AT$, where c is a constant and T is a matrix with m rows, then

(1) $\rho_M(D) = \rho_M(T)$ and $\rho^2(D) = \rho^2(T)$;

(2) if A is a 3-orthogonal column-orthogonal design, then the estimates of the linear effects of all factors of D are uncorrelated with the estimates of all quadratic effects and bilinear interactions.

From Lemma 2, the following theorem can be obtained.

Theorem 4. Suppose A is an $OA(n, q^m, t)$ with $t \geq 2$, in which the strength of the first m_1 factors is m_1 , the strength of the next m_2 factors is m_2, \dots , the strength of the last m_v factors is m_v , $\sum_{i=1}^v m_i = m$, and $T = \text{diag}\{T_{m_1, k_1}^q, \dots, T_{m_v, k_v}^q\}$ is a matrix with order $m \times k$, where T_{m_i, k_i}^q is an $m_i \times k_i$ matrix comprised of columns of permutations of $\{1, q, \dots, q^{m_i-1}\}$ (up to sign changes), $k_i \leq m_i$ and $\sum_{i=1}^v k_i = k$. Let $D = AT$, then

(1) $\rho_M(D) = \rho_M(T)$ and $\rho^2(D) = \rho^2(T)$;

(2) if $t \geq 3$, then the estimates of the linear effects of all factors of D are uncorrelated with the estimates of all quadratic effects and bilinear interactions.

Remark 4. (1) Since the order of $T_{m_i, k_i}^q, i = 1, \dots, v$, is usually far smaller than D , finding a T_{m_i, k_i}^q with low correlations is easier than finding a D with low correlations.

(2) $T = \text{diag}\{T_{m_1, k_1}^q, \dots, T_{m_v, k_v}^q\}$ is a partitioned diagonal matrix, thus any two columns not in the same part are orthogonal.

Corollary 2. Suppose $A = (A_1, \dots, A_v)$ is an $OA(n, q^m, t)$ with $m = vm_0$ and $t \geq 2$, A_i is an $OA(n, q^{m_0}, m_0)$, and $T = \text{diag}\{T_{m_0, k}^q, \dots, T_{m_0, k}^q\}$ with $T_{m_0, k}^q$ repeating v times in the diagonal. Let $D = AT$, then

(1) $\rho_M(D) = \rho_M(T_{m_0, k}^q)$ and $\rho^2(D) = \frac{k-1}{kv-1} \rho^2(T_{m_0, k}^q)$;

(2) if $t \geq 3$, then the estimates of the linear effects of all factors of D are uncorrelated with the estimates of all quadratic effects and bilinear interactions.

Example 7. Suppose $A = (A_1, \dots, A_6)$ is an $OA(32, 2^{30}, 2)$ with A_i 's being full 2^5 factorial sets. Let $T_1 = \text{diag}\{T_{5,4}^2, \dots, T_{5,4}^2\}$ with $T_{5,4}^2$ repeating 6 times in the diagonal, and $T_2 = \text{diag}\{T_{5,5}^2, \dots, T_{5,5}^2\}$ with $T_{5,5}^2$ repeating 6 times in the diagonal, where

$$T_{5,4}^2 = \begin{pmatrix} 2 & 4 & 8 & 16 \\ 4 & -2 & -16 & 8 \\ 8 & 16 & -2 & -4 \\ 16 & -8 & 4 & -2 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad T_{5,5}^2 = \begin{pmatrix} 1 & 4 & 16 & -4 & 8 \\ 2 & 8 & 4 & -1 & -16 \\ 4 & -16 & 8 & 2 & -4 \\ 8 & -1 & -2 & -16 & 1 \\ 16 & 2 & -1 & 8 & 2 \end{pmatrix}.$$

It can be calculated that $\rho_M(T_{5,4}^2) = 0.0029$, $\rho^2(T_{5,4}^2) = 8.5999 \times 10^{-6}$, $\rho_M(T_{5,5}^2) = 0.0938$ and $\rho^2(T_{5,5}^2) = 0.0048$.

(1) Let $D_1 = AT_1$, then from Corollary 2, D_1 is an LHD(32, 24) and $\rho_M(D_1) = \rho_M(T_{5,4}^2) = 0.0029$, $\rho^2(D_1) = \frac{k-1}{kv-1} \rho^2(T_{5,4}^2) = \frac{4-1}{4 \times 6-1} \rho^2(T_{5,4}^2) = 1.1217 \times 10^{-6}$. Comparing to the nearly orthogonal LHD(32, 15) with $\rho_M = 0.0191$ and $\rho^2 = 2 \times 10^{-5}$ in [18], D_1 can accommodate more factors and has lower values of ρ_M and ρ^2 .

(2) Similarly, let $D_2 = AT_2$, then D_2 is an LHD(32, 30) with $\rho_M(D_2) = \rho_M(T_{5,5}^2) = 0.0938$ and $\rho^2(D_2) = 6.6207 \times 10^{-4}$.

Corollary 3. Suppose $n = q^d$, where q is a prime and $d \geq 2$ is an integer, then an orthogonal LHD(n, k) exists, where $k = \lfloor \frac{q^d-1}{d(q-1)} \rfloor$.

Proof. From [26, 28], a saturated $D(q^d, q^m)$ design A with $m = (q^d - 1)/(q - 1)$ can be partitioned into $k = \lfloor \frac{q^d-1}{d(q-1)} \rfloor$ full factorial sets and a remainder part, without loss of generality, assume $A = (A_1, \dots, A_k, S)$, where A_i for $i = 1, \dots, k$ are full factorial sets. Let $A_0 = (A_1, \dots, A_k)$, $T = \text{diag}\{T_{d,c}^q, \dots, T_{d,c}^q\}$ with $T_{d,c}^q$ repeating k times in the diagonal, and $D_0 = A_0T = (A_1T_{d,c}^q, \dots, A_kT_{d,c}^q)$.

We take one column from every part $A_i T_{d,c}^q$ to form a design D , then from Theorem 4 and Lemma 1, D is an orthogonal LHD(n, k).

5 Discussions and concluding remarks

In this paper, we propose some methods to construct column-orthogonal designs and nearly column-orthogonal designs by rotating orthogonal arrays. The methods are easy to implement, and the resulting column-orthogonal designs keep the estimates of the linear effects of all factors uncorrelated with each other, sometimes even uncorrelated with the estimates of all quadratic effects and bilinear interactions, along with flexible and economical run sizes. In addition, in each rotation part, the resulting designs also preserve the geometric configuration of orthogonal arrays, thus have good space-filling properties. It is seen from our methods in the previous sections that if the orthogonal arrays have no repeated runs, so do the constructed designs. Therefore, such designs can be used for computer experiments. In addition, our asymmetrical column-orthogonal designs and nearly column-orthogonal designs are useful if one feels the need of studying some factors in more detail than others (cf. [4]).

Note that the column-orthogonality and uniformity do not necessarily agree with each other, i.e., the uniformity does not guarantee that the design possesses low correlations among its effects, and vice versa. The proposed designs can guarantee the nice column-orthogonality properties, and thus are optimal in terms of the column-orthogonality criteria as we have discussed, but they may be not optimal under the uniformity criteria. Now let us see an illustrative example.

Example 8. (Column-orthogonal design vs uniform design) Table 6 shows two designs with 12 runs and 5 factors, each having 4 levels, where the COD(12, 4⁵) is obtained by taking the 1st, 3rd, 5th, 7th and 9th columns from the COD(12, 4¹⁰) in Table 1, and the uniform design $D(12, 4^5)$ is taken from the Appendix of [13, p. 237], which was obtained by minimizing the uniformity measure of CL_2 -discrepancy (cf. [15]). We can see that the COD(12, 4⁵) is slightly worse than the uniform design $D(12, 4^5)$ under the CL_2 -discrepancy, but is better in terms of both ρ_M and ρ^2 , which are zero.

Further, with a large factor-to-run ratio, these new designs are economical and suitable for factor screening.

Example 9. Consider a screening experiment with 12 treatment factors each having 16 levels, and two block factors each having 4 levels. The scientist wants to reduce the cost of this experiment and plans to use a 16-run design. How to design a 16-run experiment and make sure that the estimates of the linear effects of all factors (including treatment factors and block factors) are uncorrelated with each other? The smallest run size of an orthogonal array for this problem is at least 256, which does not satisfy the run size constraint. Here, the scientist can use the COD(16, 16¹²4²) constructed in Example 4 for conducting the experiment, and the active factors can then be easily identified under the first-order model, since the estimates of the linear effects of all factors are uncorrelated with each other under this model.

Table 6 COD(12, 4⁵) and uniform design $D(12, 4^5)$

	COD(12, 4 ⁵)					Uniform design $D(12, 4^5)$				
Optimal design matrix	3	3	3	3	3	-3	-1	1	-3	-3
	-1	-1	3	-3	-1	-3	1	-3	3	-1
	-3	1	3	1	-3	-3	3	-1	-1	3
	1	-1	-1	3	-3	-1	-3	-3	-3	1
	-1	-3	1	3	1	-1	-1	3	3	3
ρ_M	-3	1	-1	-1	3	-1	3	3	1	-3
	-3	-1	-3	1	3	1	-3	-1	3	-3
	1	-3	1	-1	-1	1	1	-1	-1	-1
	3	-3	-1	-3	1	1	1	1	1	1
	3	1	-3	1	-1	3	-3	3	-1	-1
ρ^2	-1	3	-3	-1	-3	3	-1	-3	1	3
	1	3	1	-3	1	3	3	1	-3	1
	0					0.0400				
	0					0.0160				
	0.2320					0.2277				

As we know, supersaturated designs are popularly used for factor screening. Such designs are mainly evaluated under the $E(s^2)$ criterion for the two-level case, and the $E(f_{\text{NOD}})$ and χ^2 criteria for the multi-level and mixed-level cases. There are also several other criteria for evaluating supersaturated designs. Please refer to [11, 17, 21, 22] for the definitions and some reviews on these criteria. Some most recent developments on such designs can be found in, e.g., [20, 23, 24, 30]. But we should note that the $E(s^2)$ is only defined for two-level designs, which is equivalent to the ρ^2 criterion in this case, while other criteria like $E(f_{\text{NOD}})$ and χ^2 are defined for measuring the near orthogonality of a design combinatorially. The factors evaluated under these criteria are usually considered to be qualitative ones, while those in a column-orthogonal design or nearly column-orthogonal design are treated as quantitative ones. If the factors here are treated as qualitative ones, then almost all the resulting designs are supersaturated.

Acknowledgements This work was supported by the Program for New Century Excellent Talents in University of China (Grant No. NCET-07-0454), National Natural Science Foundation of China (Grant No. 10971107), and the Fundamental Research Funds for the Central Universities (Grant No. 10QNJJ003). The authors thank the referees for their valuable comments and suggestions.

References

- 1 Beattie S D, Lin D K J. Rotated factorial designs for computer experiments. Technical Report TR#98-02, Department of Statistics, The Pennsylvania State University, University Park, PA, 1998
- 2 Beattie S D, Lin D K J. Rotated factorial designs for computer experiments. *J Chin Statist Assoc*, 2004, 42: 289–308
- 3 Beattie S D, Lin D K J. A new class of Latin hypercube for computer experiments. In: Fan J, Li G, eds. *Contemporary Multivariate Analysis and Experimental Designs in Celebration of Professor Kai-Tai Fang's 65th Birthday*. Singapore: World Scientific, 2005, 206–226
- 4 Bingham D, Sitter R R, Tang B. Orthogonal and nearly orthogonal designs for computer experiments. *Biometrika*, 2009, 96: 51–65
- 5 Bursztyn D, Steinberg D M. Rotation designs for experiments in high bias situations. *J Statist Plann Inference*, 2002, 97: 399–414
- 6 Butler N A. Optimal and orthogonal Latin hypercube designs for computer experiments. *Biometrika*, 2001, 88: 847–857
- 7 Cioppa T M, Lucas T W. Efficient nearly orthogonal and space-filling Latin hypercubes. *Technometrics*, 2007, 49: 45–55
- 8 Fang K T. The uniform design: application of number-theoretic methods in experimental design. *Acta Math Appl Sinica*, 1980, 3: 363–372
- 9 Fang K T, Li R, Sudjianto A. *Design and Modeling for Computer Experiments*. Boca Raton: Chapman & Hall, 2006
- 10 Fang K T, Lin D K J. Uniform designs and their application in industry. In: Khattree R, Rao C R, eds. *Handbook on Statistics in Industry*. Amsterdam: Elsevier, 2003, 131–170
- 11 Fang K T, Lin D K J, Liu M Q. Optimal mixed-level supersaturated design. *Metrika*, 2003, 58: 279–291
- 12 Fang K T, Lin, D K J, Winker P, et al. Uniform design: theory and application. *Technometrics*, 2000, 42: 237–248
- 13 Fang K T, Ma C X. *Orthogonal and Uniform Experimental Designs (in Chinese)*. Beijing: Science Press, 2001
- 14 Hedayat A S, Sloane N J A, Stufken J. *Orthogonal Arrays: Theory and Applications*. New York: Springer-Verlag, 1999
- 15 Hickernell F J. A generalized discrepancy and quadrature error bound. *Math Comp*, 1998, 67: 29–322
- 16 Hickernell F J, Liu M Q. Uniform designs limit aliasing. *Biometrika*, 2002, 89: 893–904
- 17 Li P F, Liu M Q, Zhang R C. Some theory and the construction of mixed-level supersaturated designs. *Statist Probab Lett*, 2004, 69: 105–116
- 18 Lin C D, Bingham D, Sitter R R, et al. A new and flexible method for constructing designs for computer experiments. *Ann Statist*, 2010, 38: 1460–1477
- 19 Lin C D, Mukerjee R, Tang B. Construction of orthogonal and nearly orthogonal Latin hypercubes. *Biometrika*, 2009, 96: 243–247
- 20 Liu M Q, Cai Z Y. Construction of mixed-level supersaturated designs by the substitution method. *Statist Sinica*, 2009, 19: 1705–1719
- 21 Liu M Q, Fang K T, Hickernell F J. Connections among different criteria for asymmetrical fractional factorial designs. *Statist Sinica*, 2006, 16: 1285–1297
- 22 Liu M Q, Lin D K J. Construction of optimal mixed-level supersaturated designs. *Statist Sinica*, 2009, 19: 197–211
- 23 Liu M Q, Zhang L. An algorithm for constructing mixed-level k -circulant supersaturated designs. *Comput Statist Data Anal*, 2009, 53: 2465–2470

- 24 Liu Y, Liu M Q. Construction of optimal supersaturated design with large number of levels. *J Statist Plann Inference*, 2011, 141: 2035–2043
- 25 McKay M D, Beckman R J, Conover W J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 1979, 21: 239–245
- 26 Pang F, Liu M Q, Lin D K J. A construction method for orthogonal Latin hypercube designs with prime power levels. *Statist Sinica*, 2009, 19: 1721–1728
- 27 Santner T J, Williams B J, Notz W I. *The Design and Analysis of Computer Experiments*. New York: Springer, 2003
- 28 Steinberg D M, Lin D K J. A construction method for orthogonal Latin hypercube designs. *Biometrika*, 2006, 93: 279–288
- 29 Sun F S, Chen J, Liu M Q. Connections between uniformity and aberration in general multi-level factorials. *Metrika*, 2011, 73: 305–315
- 30 Sun F S, Lin D K J, Liu M Q. On construction of optimal mixed-level supersaturated designs. *Ann Statist*, 2011, 39: 1310–1333
- 31 Sun F S, Liu M Q, Lin D K J. Construction of orthogonal Latin hypercube designs. *Biometrika*, 2009, 96: 971–974
- 32 Sun F S, Liu M Q, Lin D K J. Construction of orthogonal Latin hypercube designs with flexible run sizes. *J Statist Plann Inference*, 2010, 140: 3236–3242
- 33 Tang B. Orthogonal array-based Latin hypercubes. *J Amer Statist Assoc*, 1993, 88: 1392–1397
- 34 Tang B. Selecting Latin hypercubes using correlation criteria. *Statist Sinica*, 1998, 8: 965–977
- 35 Wang Y, Fang K T. A note on uniform distribution and experimental design. *KeXue TongBao*, 1981, 26: 485–489
- 36 Yang J Y, Liu M Q. Construction of orthogonal and nearly orthogonal Latin hypercube designs from orthogonal designs. *Statist Sinica*, 2012, doi: 10.5705/ss.2010.021
- 37 Ye K Q. Orthogonal column Latin hypercubes and their application in computer experiments. *J Amer Statist Assoc*, 1998, 93: 1430–1439

Appendix. Proof of Theorem 1

(1) Since A is an orthogonal array with centered levels, we have $A'A = aI_k$, where a is the square norm of any column of A . Then from the column-orthogonality of $R_{m,q}$, we know the conclusion is true.

(2) Let $D = (d_{ij})$, $A = (x_{ij})$, $R_{m,q} = (r_{ij})$. Then from $D = AR_{m,q}$, we have

$$\begin{aligned}
 \sum_{i=1}^n d_{ij}d_{ik}d_{il} &= \sum_{i=1}^n \sum_{s=1}^m x_{is}r_{sj} \sum_{u=1}^m x_{iu}r_{uk} \sum_{v=1}^m x_{iv}r_{vl} \\
 &= \sum_{i=1}^n \sum_{s=1}^m \sum_{u=1}^m \sum_{v=1}^m x_{is}r_{sj}x_{iu}r_{uk}x_{iv}r_{vl} \\
 &= \sum_{s=1}^m \sum_{u=1}^m \sum_{v=1}^m \sum_{i=1}^n x_{is}x_{iu}x_{iv}r_{sj}r_{uk}r_{vl} \\
 &= \sum_{s=1}^m \sum_{u=1}^m \sum_{v=1}^m r_{sj}r_{uk}r_{vl} \sum_{i=1}^n x_{is}x_{iu}x_{iv}.
 \end{aligned}$$

For the three numbers s, u, v , there are the following three cases:

- (i) $s = u = v$, then $\sum_{i=1}^n x_{is}x_{iu}x_{iv} = \sum_{i=1}^n x_{is}^3$;
- (ii) only two of s, u, v are equal, without loss generality, suppose $s = u \neq v$ then $\sum_{i=1}^n x_{is}x_{iu}x_{iv} = \sum_{i=1}^n x_{is}^2x_{iv}$;
- (iii) s, u, v are all different from each other.

Since A is an orthogonal array with strength $t \geq 3$ and centered levels $(2i - q - 1)/2$ for odd q or $(2i - q - 1)$ for even q , $i = 1, \dots, q$, we can easily see that $\sum_{i=1}^n x_{is}x_{iu}x_{iv} = 0$ for any of these three cases. Thus the conclusion is true.