Study of two classes of composite designs

Si Qiu Jianhui Ning

Central China Normal University

Abstract: In this paper, we study the efficiencies of two classes of composite designs. The fisrt class is the central composite designs (CCDs) which are the most popular used second-order designs in response surface studies. The second class is the composite designs based on orthogonal array (OACDs). Given the assuming second-order polynomial model, the $D-$efficiency of two classes of designs for general α were derived. Moreover, the determination of the α value for star points was also discussed from the perspective of space-filling and $C-$efficiency.

Key words: Central composite design; orthogonal array composite design; α value; $D-$efficiency; $C-$efficiency; centered discrepancy.

1 Introduction

In many case studies, the goal of experiment is to find a level-combination of the factors that can optimize the response or achieve a desirable value of the response. Suppose that the process or system involves a response y that depends on the input factors $x_1,...,x_k$. Their relationship can be modeled by

$$y = f(x_1,...,x_k) + \epsilon \tag{1}$$

where the form of the true function f is unknown and ϵ is an error term. Because f is unknown and y is observed with random error, we need to run experiments to obtain data about f. Success of the investigation depends on how well f can be approximated. Response surface methodology, proposed by Box and Wilson (1951), is a strategy to achieve this goal that involves experimentation, modeling, data analysis and optimization. The first stage in many experiments is to perform a screening process by fitting a first-order linear model to the data. If the first-order model suffers from lack of fit due to the existence of some surface curvature, a second-order polynomial model is employed. For quantitative factors denoted by $x_1,...,x_k$, a second-order polynomial model is

$$y = \beta_0 + \sum_{i=1}^{k}\beta_i x_i + \sum_{i=1}^{k}\beta_{ii}x_i^2 + \sum_{i=1<j\leq k}\beta_{ij}x_i x_j + \epsilon \tag{2}$$

where $\beta_0,\beta_i,\beta_{ii},\beta_{ij}$ are the intercept, linear, quadratic and bilinear terms respectively, and ϵ is the error term. For a comprehensive account of response surface methodology, see Box and Draper (2007), Khuri and Cornell (1996), and Myers, Montgomery and Anderson-Cook (2009).The common way that researchers address response surface methodology in the literature is to use central composite designs (CCDs) introduced by Box and Wilson (1951). There are also other variations such as small composite designs (SCDs) proposed by Draper and Lin (1990) and augmented-pair designs (APDs) proposed by Morris (2000). Composite designs are often used to fit the second-order model. A designs is called a second-order design if it allows all parameters in the second-order model (2) to be estimated. The above mentioned composite designs are second-order designs.

We give a brief introduction to composite designs. For k factors, a composite design consists of 3 parts: (i)n_1 cube points (or corner points) with $x_i = 1$ or -1; (ii)n_2 star points (or additional points) with $x_i = \alpha$ or $-\alpha$; (iii)n_0 center points with all $x_i = 0$. If $\alpha = 1$, the composite design has 3 different levels, if not, the design has 5 different levels. Two-level orthogonal arrays such as full or fractional factorial designs are often used as the cube points. Xu, Jaynes and Ding (2014) introduced a new class of composite designs, called orthogonal array composite designs (OACDs), which use a 3-level orthogonal array as the star points. Zhou and Xu (2016) developed some theoretical results on CCDs and OACDs for $\alpha=1$. If the α value is not chosen to be 1, what will the results be? So in this paper, we derived bounds of $D-$efficiency of the two classes of designs for estimating all and part of the parameters in a second-order model for general α. We further introduced the $C-$efficiency to compare the two classes of composite designs. In choosing a composite design, the determination of the α value should be considered. For central composite designs, Box and Hunter (1957) provides one criterion, that is rotatability. We proceed from the perspective of space-filling and efficiency to disscuss the determination of the α value for both two classes of composite designs.

This paper is organized as follows. Central composite designs and orthogonal array composite designs are discussed in section 2 and 3 respectively. The determination of the α value is discussed in section 4. Conclusion remarks are given in section 5 and we give proofs in the Appendix.

2 Central composite designs

Suppose that there are k input factors, a central composite design uses the axial points as star points, then it consists of the following three parts:

- n_1 cube points with $x_i=1$ or -1 for $i=1,...,k$;
- n_2 axial points of the form $(0,...,x_i,..,0)$ with $x_i=\alpha$ or $-\alpha$ for $i=1,...,k$;
- n_0 center points with all $x_i=0$ for $i=1,...,k$.

2.1 The definition of $D-$efficiency

Let d be the k-factor composite design, $X=(\mathbf{1},Q,L,B)$ be the model matrix of the second-order model (2), where $\mathbf{1}$ is a column of ones, Q,L and B are quadratic, linear and bilinear terms, respectively. Let d_i be the part i of the design for $i=1,2$. The total number of runs of d is $N=n_1+n_2+n_0$. The $D-$efficiency of d is

$$D(d)=N^{-1}|X'X|^{1/p} \tag{3}$$

where $p=(k+1)(k+2)/2$ is the number of parameters in the second-order model (2). Sometimes the partial efficiency also need to considered, we use the D_s-efficiency describes the precision for estimating a subset of the model parameters. Define D_s-efficiency as

$$D_s(d)=N^{-1}|X_s'X_s-X_s'X_{(s)}(X_{(s)}'X_{(s)})^{-1}X_{(s)}'X_s|^{1/t} \tag{4}$$

where s is a subset of factors, X_s and $X_{(s)}$ are the sub-matrices of X corresponding to the parameters in s or not in s, respectively, and t is the number of parameters in s. Since

$$|X'X|=|X_{(s)}'X_{(s)}||X_s'X_s-X_s'X_{(s)}(X_{(s)}'X_{(s)})^{-1}X_{(s)}'X_s| \tag{5}$$

then the D_s-efficiency can also be calculated by

$$D_s(d)=N^{-1}(\frac{|X'X|}{|X_{(s)}'X_{(s)}|})^{1/t} \tag{6}$$

Let D_L, D_B and D_Q be the D_s-efficiency for the linear, bilinear and quadratic terms, respectively.

2.2 The $D-$efficiencies of central composite designs

From equation (3) and (6), we obtain the follow Theorem 1.

Theorem 1. For a CCD with k factors and n_0 center points, if the 2-level portion is an $OA(n_1, 2^k, 4)$, its $D-$, D_L-, D_B- and D_Q- efficiencies are, respectively,

$$D(CCD) = \frac{1}{N}\{n_1^q(2\alpha^4 n_1 + 4\alpha^6)^k[(1+\frac{kn_1}{2\alpha^4})n_0 + (1-\frac{k}{\alpha^2})^2 n_1]\}^{1/p} \tag{7}$$

$$D_L(CCD) = \frac{1}{N}(n_1 + 2\alpha^2) \tag{8}$$

$$D_B(CCD) = \frac{n_1}{N} \tag{9}$$

$$D_Q(CCD) = \frac{2\alpha^4}{N^{\frac{k+1}{k}}}[(1+\frac{kn_1}{2\alpha^4})n_0 + (1-\frac{k}{\alpha^2})^2 n_1]^{1/k} \tag{10}$$

where $N = n_1 + 2k + n_0$, $q = k(k-1)/2$, $p = (k+1)(k+2)/2$.

We can see that the α value affects the efficiencies of CCDs except the D_B-efficiency. So it is important to choose a proper α.

3 Orthogonal array composite designs

Xu, Jaynes and Ding (2014) introduce a new class of composite designs that combine a two-level factorial or fractional factorial design and a three-level orthogonal array, and refer to them as orthogonal array composite designs (OACDs). An orthogonal array of n runs, k columns, s levels, and strength t, denoted by $OA(n, s^k, t)$, is an $n \times k$ matrix. In general, the strength t is omitted when $t = 2$. Detailed account of the theory and application of orthogonal array can see Hedayat, Sloane and Stufken (1999).

In an orthogonal array composite design, its star points form a 3-level orthogonal array. So an orthogonal array composite design with k factors has three parts as follows:

- n_1 cube points with $x_i = 1$ or -1 for $i = 1, ..., k$;
- a 3-level orthogonal array with n_2 runs;
- n_0 center points with all $x_i = 0$ for $i = 1, ..., k$.

Orthogonal array composite designs are constructed with careful consideration of experimental cost, time and statistical efficiency. Xu et al.(2014) present several advantages for considering an OACD. First, they present a good trade-off between model parameter estimation efficiency and run size economy. Second, an OACD uses a better initial design than existing composite designs, enabling all linear effects to be estimated clearly from any interaction effects. Third, they offer more in-depth analyses than traditional designs using either a two-level or three-level design, that is, an OACD has the built-in ability to perform cross-validation on the data quality and analysis results. Separate analyses for the two-level and three-level parts, as well as the combined two-level and three-level OACD, can be performed by building three models: a model with linear and bilinear terms based on the two-level design, a model with linear and quadratic terms based on the three-level orthogonal array, and a full second-order model based on the entire OACD. Because each linear effect is estimated three times, and each bilinear and quadratic effect are estimated twice, we can check the consistency of their estimates, providing model confirmation. Fourth, an OACD can be used in either a single or sequential experiment, that is, starting with a two-level design, followed by a three-level design or starting with a three-level design, followed by a two-level design; this feature is particularly appealing in many industrial and engineering experiments. Because of its desirable features, many researchers use OACDs to investigate biological systems. Jaynes et al.(2015) present a novel application of OACDs to optimize lipid production of a cell-free system for algae.

Because of the information matrix and efficiencies for OACDs depend on the specific 3-level orthogonal array used, we cannot get general theoretical results for the $D-$efficiency and D_s-efficiencies. We now consider lower bounds of the efficiencies and have the following Theorem 2 and Theorem 3. For the D_L-efficiency, we also have upper bound.

Theorem 2. Let an $OA(n_1, 2^k, 4)$ be the first part and an $OA(n_2, 3^k)$ be the second part of the OACD. The determinant of its information matrix and $D-$efficiency have the following lower bounds, respectively,

$$|X'X| \geq \eta \tag{11}$$

$$\eta = n_1^q [\frac{(4n_2\alpha^2 + 6n_1)\alpha^2 n_2}{27}]^k [(1+2k\alpha^2)n_0 + n_2 + n_1(1+2k\alpha^2 + \frac{9kn_0}{2n_2\alpha^2} + \frac{9k}{2\alpha^2} - 6k)] \tag{12}$$

$$D(OACD) \geq LB(OACD) = N^{-1}(\eta)^{1/p} \tag{13}$$

where $N = n_1 + n_2 + n_0$, $q = k(k-1)/2$, $p = (k+1)(k+2)/2$.

Theorem 3. Suppose that an OACD satisfies the conditions in Theorem 2. Its D_L-, D_B-, D_Q-efficiencies have the following lower bounds, respectively,

$$D_L(OACD) \geq \frac{1}{N}(\frac{9n_1}{9n_1 + 4n_2\alpha^4})^{q/k}(\frac{2}{3}n_2\alpha^2 + n_1) \tag{14}$$

$$D_B(OACD) \geq \frac{n_1}{N} \tag{15}$$

$$D_Q(OACD) \geq \frac{2n_2\alpha^2}{9N^{\frac{k+1}{k}}}(\frac{9n_1}{9n_1 + 4n_2\alpha^4})^{q/k}[(1+2k\alpha^2)n_0 + n_2 + n_1(1+2k\alpha^2 + \frac{9kn_0}{2n_2\alpha^2} + \frac{9k}{2\alpha^2} - 6k)]^{1/k} \tag{16}$$

where $N = n_1 + n_2 + n_0$, $q = k(k-1)/2$. Furthermore, its D_L-efficiency has an upper bound

$$D_L(OACD) \leq \frac{1}{N}(\frac{2}{3}n_2\alpha^2 + n_1) \tag{17}$$

the equality holds when the linear terms of second part of the design are orthogonal to the bilinear terms of second part of the design.

In order to compare the two classes of composite designs intuitively, next we give an example to compare the efficiency of the composite designs.

Example 1. We compare OACDs with CCDs consisting of the same 2-level portion for $k = 4, ..., 12$. We choose a full factorial 2^k for $k = 4$ or a regular 2^{k-m} design with resolution at least V for $k = 5, ..., 11$. For $k = 12$, we use an $OA(128, 2^{15}, 4)$ from Xu (2005) as the 2-level portion. For the 3-level OA, we choose the first k columns of $OA(9, 3^4, 2)$, $OA(18, 3^7, 2)$ and $OA(27, 3^{13}, 2)$ from Sloane's website. Table 1 and Table 2 shows the $D-$efficiency of OACDs and CCDs as well as the lower bound for $\alpha = 1$ and $\alpha = 1.5$, respectively, with $n_0 = 5$ center points. For every $k \geq 4$, an OACD has larger $D-$efficiency than a CCD for every α. And the lower bound of OACD is also larger than CCD's $D-$efficiency when $\alpha = 1$. When $\alpha = 1.5$, the lower bound of OACD no longer larger than CCD's $D-$efficiency for $k = 4, ..., 7$, but the $D-$efficiency of OACDs still larger than CCD's. From Figure 1 and Figure 2 we can get this result intuitively.

Table 1. Comparison of $D-$efficiency between OACDs and CCDs for $\alpha=1$ with $n_0=5$ center points

k	d_1	d_2	$D(OACD)$	$LB(OACD)$	$D(CCD)$
4	2^4	$OA(9,3^4)$	0.42108	0.40179	0.39835
5	2_V^{5-1}	$OA(18,3^5)$	0.44523	0.38248	0.37968
6	2_{VI}^{5-1}	$OA(18,3^6)$	0.48160	0.43847	0.41672
7	2_{VII}^{7-1}	$OA(18,3^7)$	0.50102	0.47804	0.44163
8	2_V^{8-2}	$OA(27,3^8)$	0.52416	0.48455	0.44919
9	2_V^{9-2}	$OA(27,3^9)$	0.54165	0.51972	0.46666
10	2_V^{10-3}	$OA(27,3^{10})$	0.55634	0.53500	0.47925
11	2_V^{11-4}	$OA(27,3^{11})$	0.56969	0.54881	0.48990
12	$OA(128,2^{12},4)$	$OA(27,3^{12})$	0.58118	0.56132	0.49885

Table 2. Comparison of $D-$efficiency between OACDs and CCDs for $\alpha=1.5$ with $n_0=5$ center points

k	d_1	d_2	$D(OACD)$	$LB(OACD)$	$D(CCD)$
4	2^4	$OA(9,3^4)$	0.72977	0.53363	0.56000
5	2_V^{5-1}	$OA(18,3^5)$	0.93602	0.51456	0.52616
6	2_{VI}^{5-1}	$OA(18,3^6)$	0.86086	0.55279	0.55928
7	2_{VII}^{7-1}	$OA(18,3^7)$	0.77859	0.57529	0.57837
8	2_V^{8-2}	$OA(27,3^8)$	0.86378	0.58210	0.57862
9	2_V^{9-2}	$OA(27,3^9)$	0.79590	0.60460	0.59011
10	2_V^{10-3}	$OA(27,3^{10})$	0.79343	0.61569	0.59763
11	2_V^{11-4}	$OA(27,3^{11})$	0.79161	0.62566	0.60319
12	$OA(128,2^{12},4)$	$OA(27,3^{12})$	0.78745	0.63466	0.60717

In fact, for the 3-level orthogonal array, permuting its level lead to OACD with different $D-$effficiency. We can search a high efficiency design by permuting level for several factors.

4 The determination of the α value

4.1 $C-$efficiency

The most well-known and popularly used design criterion is the $D-$efficiency. An implicit assumption for $D-$efficiency is that all parameters are equally important, this may not be appropriate for the two-stage design. To overcome its disadvantage, Lu, Lin and Zhou (2009) propose a new criterion, called $C-$efficiency. The $C-$efficiency is defined as

$$C=D_{\mathbf{1}}^{w_{\mathbf{1}}}D_L^{w_L}D_B^{w_B}D_Q^{w_Q} \tag{18}$$

where $w_{\mathbf{1}}$, w_L, w_B, w_Q, are nonnegative weights and $w_{\mathbf{1}}+w_L+w_B+w_Q=1$. The weights represent the importance of the parameter sets, more important parameter sets carry heavier weights. For different first-stage points, one can specify different weights to $C-$efficiency. Lu, Lin and Zhou (2009) suggests five specifications of the weights.

We now use the $C-$efficiency to compare CCDs and OACDs for $k=4$ factors.

Example 2. Consider the same CCDs and OACDs as in example 1 for $k=4$ and compare their $C-$efficiency for $\alpha=1,...,2$. Because the 2-level portion of the design is a 2^4 full factorial design, following Lu, Lin and Zhou (2009) we set $w_{\mathbf{1}}=w_L=w_B=0$ and $w_Q=1$. Figure 3 shows that the $C-$efficiency of CCDs larger than OACDs, but in fact this phenomena just appear in $k=4$. Another results is that the $C-$efficiency increases as the α increases. So the $C-$efficiency is maximized in selecting larger α.

4.2 New criterion for choosing α value

For CCDs, Wu and Hamada(2009) gives some suggestions on the determination of the α value. In general, α should be chosen between 1 and $\sqrt{k}$ and rarely outside this range. For $\alpha = 1$, the star points are placed at the center of the faces of the cube, $\alpha = \sqrt{k}$ makes the star points and cube points lie on the same sphere. The efficiency of the parameter estimates is increased by pushing the star points toward the extreme, however, for large k, this choice should be taken with caution because the star points are too far from the center point and no information is available for the response surface in the intermediate range of the factors, especially along the axes. From example 1 and example 2, we can find the effect of the α value on efficiency, the larger α is, the greater efficiency is. In general, if it is desired to collect information closer to the faces of the cube, a smaller α value should be chosen. If estimation efficiency is a concern, the star points should be pushed toward the extremes of the region, namely choosing a larger α. Box and Hunter (1957) provides a criterion for CCDs to determine the α value. The criterion is rotatability. If the central composite design is rotatable, then $\alpha = 2^{(k-m)/4}$.

Yet there is no general criterion for choosing α for composite designs, and it is indeed needed. So one purpose of this paper is to propose a new criterion for choosing α and applicable to both classes composite designs. What we need to pay attention to is that, above discussion is under the second-order model. When the true model is unknown and the second-order model does not fit the true model, it is meaningless to choose larger α to get higher efficiency design, and sometimes the bound of the α value cannot achieve. We all know this, if we lack the information about the true model, we prefer to distribute points uniformly. That is to say, we would like to get a design as uniformly as possible to conduct experiments. Uniform design has robustness for different models. Take these into consideration, we propose to determine the α value from the perspective of space-filling, using the uniformity to choose α. So the criterion to determine α is the discrepancy. We give a simple example to show our idea.

Example 3. Suppose that the response y depends on four input factors x_1, x_2, x_3, x_4. $x_1 \in [11.20, 13.60]$, $x_2 \in [1.98, 3.04]$, $x_3 \in [0.75, 1.75]$, $x_4 \in [1.00, 3.00]$. For the sake of later calculation easier and eliminate the influence of variable dimension, we convert x_i to coded variables firstly. We transfer the lower bound of the actual level into -2 and the upper bound into 2, then α can be arbitrary value from 1 to 2. In this example, we use the centered discrepancy to measure the uniformity of designs. We calculate centered discrepancy of OACDs and CCDs for different α. Figure 4 and Figure 5 show the tendency of centered discrepancy changing with different α. From the two figures, we can find that the OACD is the most uniformly design for $\alpha = 1.4$; when $\alpha = 1.6$, the CCD is the most uniformly design. If we determine the α value from the perspective of rotatability, then $\alpha = 2$.

5 Conclusion remarks

We study the estimation efficiencies of CCDs and OACDs under a second-order polynomial model for general α value. We find that OACDs are more effective in estimating the parameters than CCDs especially the number of factors is large. The OACDs provide a good trade-off between estimation efficiency and run size economy, so it can be used as an alternative to the popular CCDs. We also provided some criterions to choose the α value. Different criterions will make different results. In practice, the determination of the α value also depends on the objectives of each experiment and the geometric nature of and the practical constraints on the design region. We proposed an idea to determine α value from the the perspective of space-filling, find a α to make the design most uniformly, but we haven't give the theoretical results. We hope to address this issue in future work. In addition, the number of center points will affect the composite designs, further discussion on the number of runs at the center point can be found in Box and Draper (2007), it is an important problem for future investigation.

Appendix

Lemma 1. Let $a \neq 0$, $b \neq 0$,

$$\begin{pmatrix} c_0 & c\mathbf{1}'_k \\ c\mathbf{1}_k & a\mathbf{J}_k + b\mathbf{I}_k \end{pmatrix} = b^{k-1}(bc_0 + k(ac_0 - c^2)).$$

Lemma 2. Let E and F be two $n \times n$ nonnegative definite matrices with partions

$$E = \begin{pmatrix} E_1 & \mathbf{0} \\ \mathbf{0} & E_2 \end{pmatrix} \geq 0, \ F = \begin{pmatrix} F_1 & F_3 \\ F'_3 & F_2 \end{pmatrix} \geq 0,$$

where E_1 and F_1 are $m \times m$ matrices. Then

$$|E+F| \geq |E_2| \cdot |E_1 + F_1|.$$

Proof of Theorem 1. The information matrix of the central composite design is block diagonal as follows:

$$X'X = \begin{pmatrix} N & (n_1+2\alpha^2)\mathbf{1}'_k & \mathbf{0} & \mathbf{0} \\ (n_1+2\alpha^2)\mathbf{1}_k & n_1 J_k + 2\alpha^4 I_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (n_1+2\alpha^2)I_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & n_1 I_q \end{pmatrix}, \tag{19}$$

where $\mathbf{1}_k$ is a column of k ones, I_k is $k \times k$ identity matrix, J_k is the $k \times k$ matrix of ones. So it is easy to obtain that

$$|X'X| = n_1^q(2\alpha^4 n_1 + 4\alpha^6)^k[(1+\frac{kn_1}{2\alpha^4})n_0 + (1-\frac{k}{\alpha^2})^2 n_1], \tag{20}$$

$$|X'_{(L)}X_{(L)}| = n_1^q(2\alpha^4)^k[(1+\frac{kn_1}{2\alpha^4})n_0 + (1-\frac{k}{\alpha^2})^2 n_1], \tag{21}$$

$$|X'_{(B)}X_{(B)}| = (2\alpha^4 n_1 + 4\alpha^6)^k[(1+\frac{kn_1}{2\alpha^4})n_0 + (1-\frac{k}{\alpha^2})^2 n_1], \tag{22}$$

$$|X'_{(Q)}X_{(Q)}| = N(n_1+2\alpha^2)^k n_1^q, \tag{23}$$

then use the definition of $D-$efficiency we can get formula (7), (8), (9), (10).

Proof of Theorem 2. Denote $X_0 = (\mathbf{1}_{n_0}, \mathbf{0}, \mathbf{0}, \mathbf{0})$ and $X_i = (\mathbf{1}_{n_i}, Q_i, L_i, B_i)$, where Q_i, L_i, B_i respectively are the quadratic, linear and bilinear terms of d_i in the second-order model, $i = 1, 2$.

$$X'_1X_1 = \begin{pmatrix} n_1 & n_1\mathbf{1}'_k & \mathbf{0} & \mathbf{0} \\ n_1\mathbf{1}_k & n_1 J_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & n_1 I_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & n_1 I_q \end{pmatrix} = \begin{pmatrix} n_1 J_{k+1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & n_1 I_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & n_1 I_q \end{pmatrix}, \tag{24}$$

$$X_2'X_2 = \begin{pmatrix} n_2 & \frac{2}{3}n_2\alpha^2\mathbf{1}_k' & \mathbf{0} & \mathbf{0} \\ \frac{2}{3}n_2\alpha^2\mathbf{1}_k & \frac{4}{9}n_2\alpha^4 J_k + \frac{2}{9}n_2\alpha^4 I_k & \mathbf{0} & Q_2'B_2 \\ \mathbf{0} & \mathbf{0} & \frac{2}{3}n_2\alpha^2 I_k & L_2'B_2 \\ \mathbf{0} & B_2'Q_2 & B_2'L_2 & B_2'B_2 \end{pmatrix}, \tag{25}$$

let $Y = X_2'X_2 + X_0'X_0$, then

$$Y = \begin{pmatrix} n_2+n_0 & \frac{2}{3}n_2\alpha^2\mathbf{1}_k' & \mathbf{0} & \mathbf{0} \\ \frac{2}{3}n_2\alpha^2\mathbf{1}_k & \frac{4}{9}n_2\alpha^4 J_k + \frac{2}{9}n_2\alpha^4 I_k & \mathbf{0} & Q_2'B_2 \\ \mathbf{0} & \mathbf{0} & \frac{2}{3}n_2\alpha^2 I_k & L_2'B_2 \\ \mathbf{0} & B_2'Q_2 & B_2'L_2 & B_2'B_2 \end{pmatrix}, \tag{26}$$

denote

$$B_{11} = \begin{pmatrix} n_2+n_0 & \frac{2}{3}n_2\alpha^2\mathbf{1}_k' \\ \frac{2}{3}n_2\alpha^2\mathbf{1}_k & \frac{4}{9}n_2\alpha^4 J_k + \frac{2}{9}n_2\alpha^4 I_k \end{pmatrix}, B_{13} = \begin{pmatrix} \mathbf{0} \\ Q_2'B_2 \end{pmatrix}, \tag{27}$$

then

$$X'X = X_1'X_1 + Y = \begin{pmatrix} B_{11}+n_1 J_{k+1} & \mathbf{0} & B_{13} \\ \mathbf{0} & (\frac{2}{3}n_2\alpha^2+n_1)I_k & L_2'B_2 \\ B_{13}' & B_2'L_2 & n_1 I_q + B_2'B_2 \end{pmatrix}, \tag{28}$$

from Lemma 2, we get

$$|X'X| = |X_1'X_1+Y| \geq |n_1 I_q| \cdot \begin{vmatrix} B_{11}+n_1 J_{k+1} & \mathbf{0} \\ \mathbf{0} & (\frac{2}{3}n_2\alpha^2+n_1)I_k \end{vmatrix} = n_1^q(\frac{2}{3}n_2\alpha^2+n_1)^k |B_{11}+n_1 J_{k+1}|, \tag{29}$$

from Lemma 1, we have

$$|B_{11}+n_1 J_{k+1}| = (\frac{2}{9}n_2\alpha^2)^k[(1+2k\alpha^2)n_0+n_2+n_1(1+2k\alpha^2+\frac{9kn_0}{2n_2\alpha^2}+\frac{9k}{2\alpha^2}-6k)], \tag{30}$$

therefore

$$|X'X| \geq n_1^q[\frac{(4n_2\alpha^2+6n_1)\alpha^2 n_2}{27}]^k[(1+2k\alpha^2)n_0+n_2+n_1(1+2k\alpha^2+\frac{9kn_0}{2n_2\alpha^2}+\frac{9k}{2\alpha^2}-6k)], \tag{31}$$

then we can obtain formula (13).

Proof of Theorem 3. When $s = L$, from formula (28) and Fischer inequality, we have

$$|X_{(L)}'X_{(L)}| \leq |B_{11}+n_1 J_{k+1}| \cdot |n_1 I_q + B_2'B_2|, \tag{32}$$

because all of the diagonal elements of $B_2'B_2$ are $\frac{4}{9}n_2\alpha^4$, we have

$$|n_1 I_q + B_2'B_2| \leq (n_1+\frac{4}{9}n_2\alpha^4)^q, \tag{33}$$

so

$$|X_{(L)}'X_{(L)}| \leq |B_{11}+n_1 J_{k+1}| \cdot (n_1+\frac{4}{9}n_2\alpha^4)^q, \tag{34}$$

then using formula (6) and (31), we obtain formula (14). Moreover, from Fischer inequality, we have

$$|X'X| \leq |X'_{(L)}X_{(L)}| \cdot |X'_L X_L|, \tag{35}$$

so we have

$$\frac{|X'X|}{|X'_{(L)}X_{(L)}|} \leq |X'_L X_L| = |(\frac{2}{3}n_2\alpha^2 + n_1)I_k| = (\frac{2}{3}n_2\alpha^2 + n_1)^k, \tag{36}$$

then get formula (17), if the linear terms of d_2 are orthogonal to the bilinear terms of d_2, then

$$\frac{|X'X|}{|X'_{(L)}X_{(L)}|} = |X'_L X_L|, \tag{37}$$

and the upper bound of D_L-efficiency is achieved.

When $s = B$, from formula (28),

$$|X'_{(B)}X_{(B)}| = \begin{vmatrix} B_{11} + n_1 J_{k+1} & \mathbf{0} \\ \mathbf{0} & (\frac{2}{3}n_2\alpha^2 + n_1)I_k \end{vmatrix} = |B_{11} + n_1 J_{k+1}| \cdot (\frac{2}{3}n_2\alpha^2 + n_1)^k, \tag{38}$$

then follows from formula (6) and (31),we get the lower bound of D_B-efficiency.

When $s = Q$, from formula (28) and Fischer inequality,

$$\begin{aligned} |X'_{(Q)}X_{(Q)}| &= \begin{vmatrix} N & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\frac{2}{3}n_2\alpha^2 + n_1)I_k & L'_2 B_2 \\ \mathbf{0} & B'_2 L_2 & n_1 I_q + B'_2 B_2 \end{vmatrix} \leq N \cdot (\frac{2}{3}n_2\alpha^2 + n_1)^k \cdot |n_1 I_q + B'_2 B_2| \\ &\leq N \cdot (\frac{2}{3}n_2\alpha^2 + n_1)^k \cdot (n_1 + \frac{4}{9}n_2\alpha^4)^q, \end{aligned} \tag{39}$$

then follows from formula (6), (11) and (33),we get the lower bound of D_Q-efficiency.

References

1 Box, G. E. P. and Hunter, J. S. (1957). Multifactor Experimental Designs for Exploring Response Surfaces. Annals of Mathematical Statistics, 28(1), 195-241.

2 Box, G. E. P. and Wilson, K. B. (1951). On the experimental attainment of optimum conditions. Journal of Royal Statistics Society, 13, 1-45.

3 Box, G. E. P. and Draper, N. R. (2007). Response Surfaces, Mixtures, and Ridge Analyses. 2nd edition. Wiley, New York.

4 Draper, N. R. and Lin, D. K. (1990). Small response-surface designs. Technometrics, 32(2), 187-194.

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

6 Khuri, A. and Cornell, J. A. (1996). Response Surfaces: Designs and Analyses. 2nd edition. Marcel Dekker, New York.

7 Lu, X., Lin, D. K. and Zhou, D. (2009). On Construction of Two-stage Response Surface Designs. Quality Technology and Quantitative Management, 6(4), 493-502.

8 Morris, M. D. (2000). A class of three-level experimental designs for response surface modeling. Technometrics, 42(2), 111-121.

9 Myers, R. H., Montgomery, D. C. and Anderson-Cook, C. M. (2009). Response Surface Methodology: Process and Product Optimization Using Designed Experiments. 3rd edition. Wiley, New York.

10 Wu, C. F. J. and Hamada, M. (2009). Experiments: Planning, Analysis and Parameter Design Optimization. 2nd edition. Wiley, New York.

11 Xu, H. (2005). Some nonregular designs from the Nordstrom – Robinson code and their statistical properties. Biometrika, 92(2), 385-397.

12 Xu, H., Jaynes, J. and Ding, X. (2014). Combining two-level and three-level orthogonal arrays for factor screening and response

surface exploration. Statistica Sinica, 24, 269-289.

13 Zhou, Y. D. and Xu, H. Q. (2016). Composite Designs Based on Orthogonal Arrays and Definitive Screening Designs. Journal of the American Statistical Association,, 0-0.

14 Jaynes, J., Zhao, Y., Xu, H. and Ho, C. (2015). Use of Orthogonal Array Composite Designs to Study Lipid Accumulation in a Cell-Free System. Quality and Reliability Engineering International, 32(5), 1965-1974.

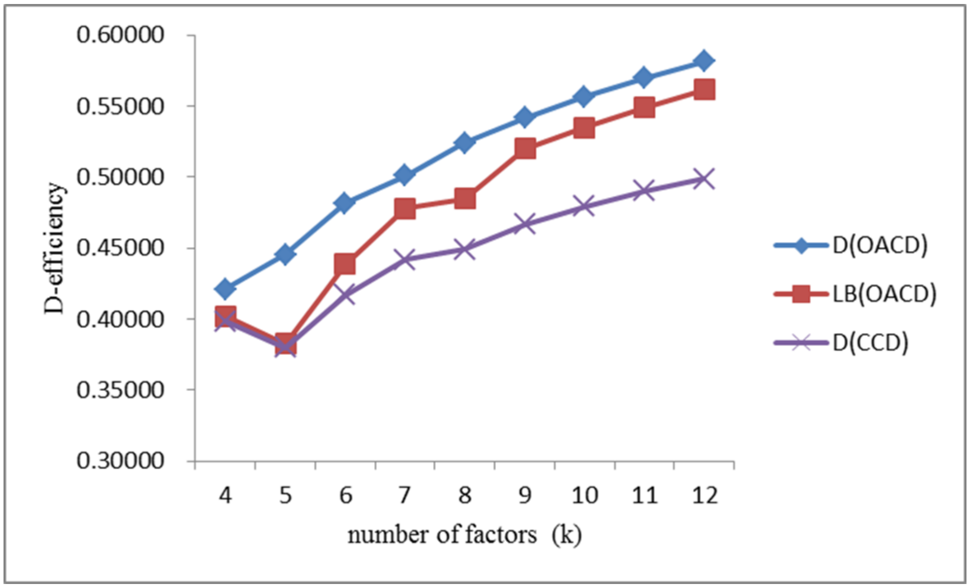

Figure. 1. Comparison of D-efficiency between OACDs and CCDs for $\alpha = 1$ with $n_0 = 5$ center points

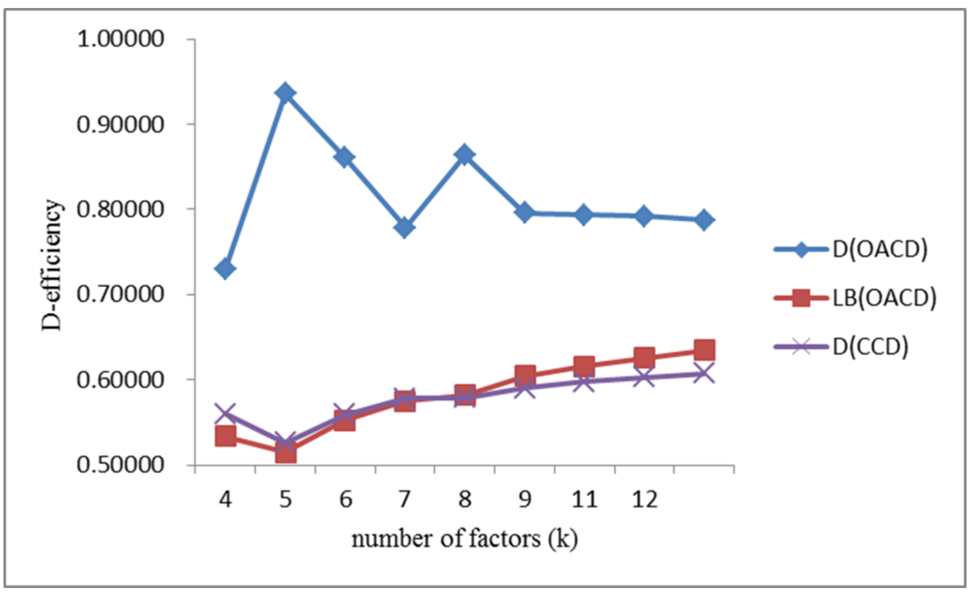

Figure. 2. Comparison of D-efficiency between OACDs and CCDs for $\alpha = 1.5$ with $n_0 = 5$ center points

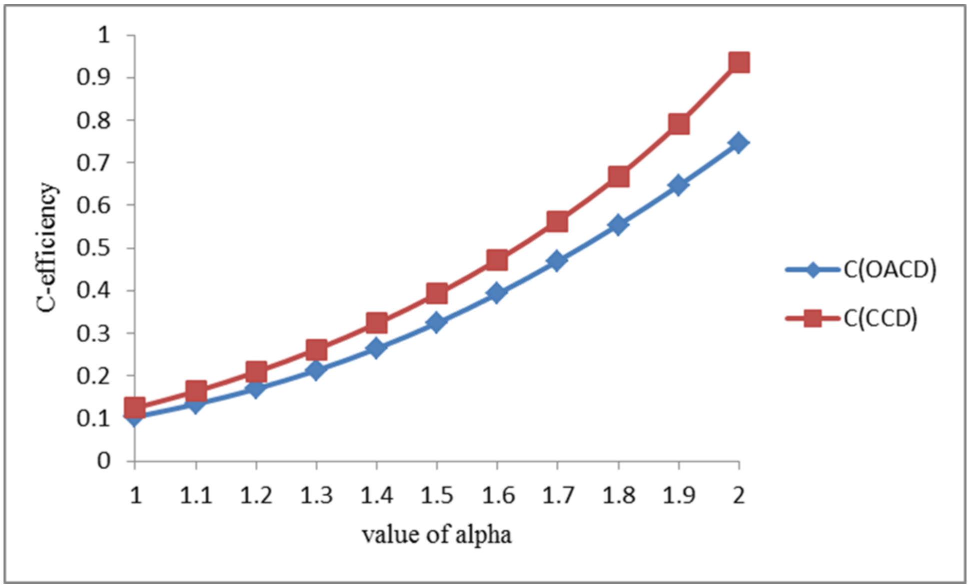

Figure. 3. C-efficiency of OACDs and CCDs for $k = 4$ factors

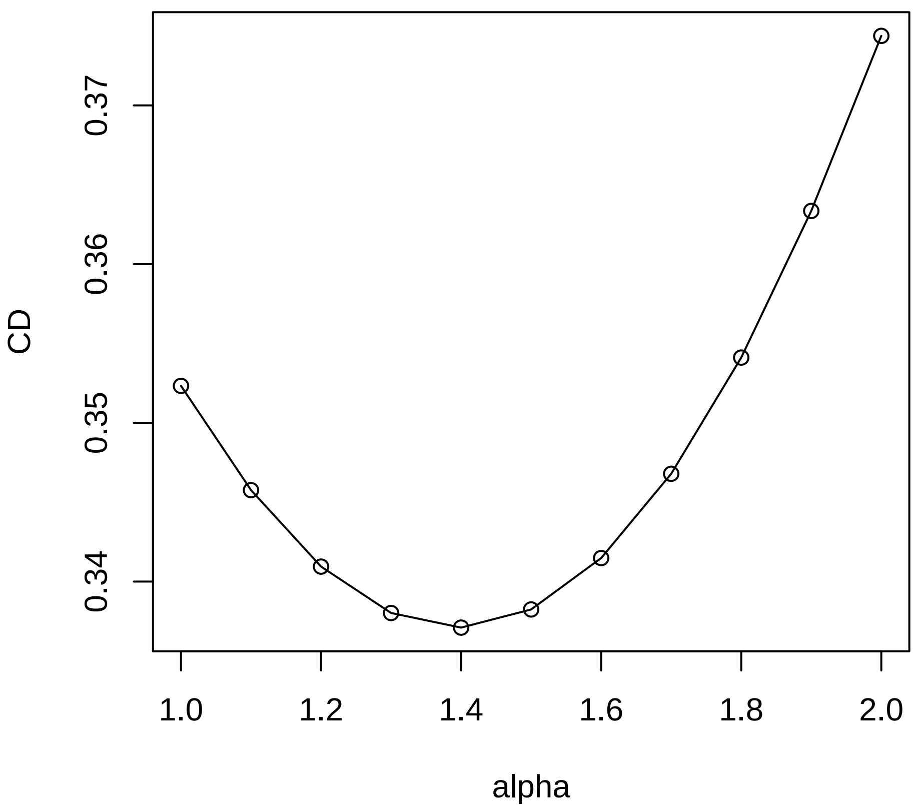

Figure. 4. Centered discrepancy of OACDs for $k = 4$ factors

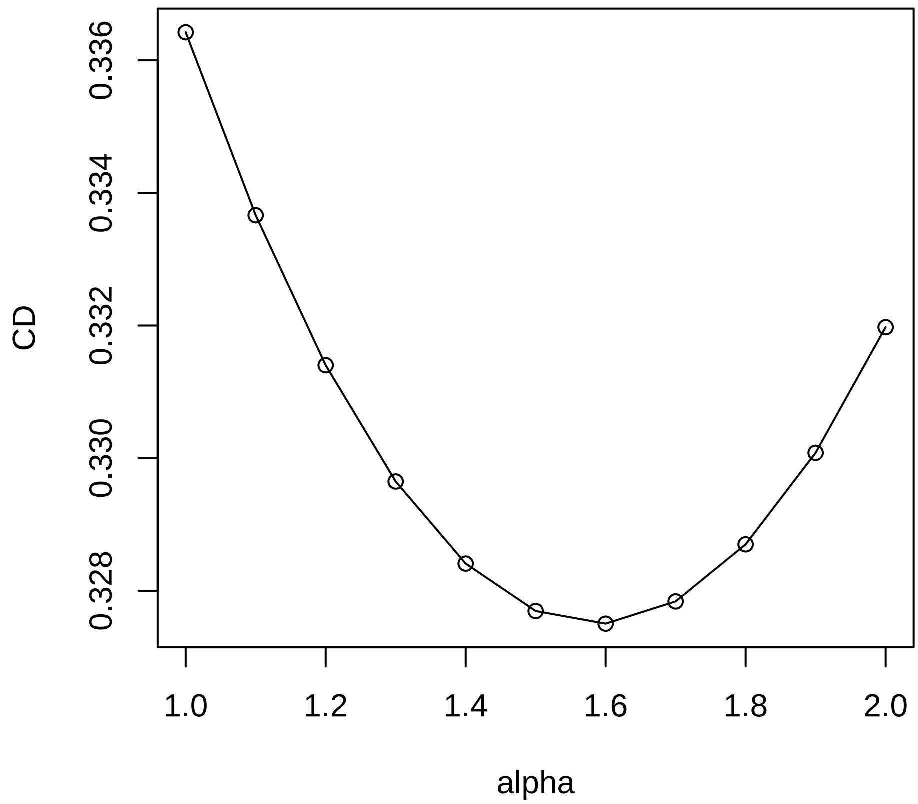

Figure. 5. Centered discrepancy of CCDs for $k = 4$ factors