JSPI: 5608 Model 3G pp. 1–10 (col. fig: NIL)

ARTICLE IN PRESS

Journal of Statistical Planning and Inference xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Robustness of orthogonal-array based composite designs to missing data

Xue-Ping Chen [a,b], Bing Guo [a], Min-Qian Liu [a,*], Xiao-Lei Wang [a]

[a] *LPMC and Institute of Statistics, Nankai University, Tianjin 300071, China*
[b] *Department of Mathematics, Jiangsu University of Technology, Changzhou, 213001, China*

ARTICLE INFO

Article history:
Received 28 December 2016
Received in revised form 10 October 2017
Accepted 11 October 2017
Available online xxxx

MSC:
62K20
62K15

Keywords:
Composite design
D-efficiency
Generalized scaled standard deviation
Minimax loss criterion

ABSTRACT

Missing observations can hardly be avoided even by a well-planned experiment. Based on the orthogonal-array based composite designs proposed by Xu et al. (2014), new orthogonal-array based composite minimax loss designs are constructed. Comparisons between the proposed designs and other composite designs, including orthogonal-array based composite designs, augmented pairs designs, augmented pairs minimax loss designs, central composite designs, and small composite designs are made in detail, which show that the new composite designs are more robust to one missing design point in terms of the *D*-efficiency and generalized scaled standard deviation. Moreover, it is demonstrated that the *D*-efficiency remains unchanged for both level permutation and column permutation in some special cases.

1. Introduction

In statistics, missing data is a common occurrence and can have a significant effect on the conclusions drawn from the data. The effect of missing observations can be of particular concern when the design is nearly saturated, saturated, or supersaturated. Many authors have discussed the robustness of statistical designs against missing data. For example, Akhtar and Prescott (1986) studied the robustness of central composite designs to one or two design points missing and constructed minimax loss central composite designs. Morris (2000) constructed augmented pairs designs — a class of nearly saturated experimental designs, with three equally spaced levels, for use in response surface modeling. The purpose was to construct some efficient designs that can estimate a second-order model by using fewer runs as compared with other second-order designs. Ahmad and Gilmour (2010) studied the robustness of subset response surface designs to missing observations. Later, by the minimax loss criterion, Ahmad et al. (2012) constructed augmented pairs minimax loss designs, which are more robust to one missing observation than the original augmented pairs designs.

Recently, motivated by an antiviral drug experiment, Ding et al. (2013), and Xu et al. (2014) suggested a new class of composite design based on a two-level factorial design and a three-level orthogonal array. Xu et al. (2014) illustrated that the new orthogonal-array based composite designs (OACDs) have many desirable features and are effective for factor screening and response surface modeling. More recently, Zhou and Xu (2016) derived bounds of their efficiencies for estimating parameters in a second-order model. In this paper, based on OACDs, we develop a new class of second-order design, called orthogonal-array based composite minimax loss design (OACM), which minimizes the maximum loss of a missing design point and is more robust to one missing observation than other composite designs for estimating parameters in a second

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

https://doi.org/10.1016/j.jspi.2017.10.004

Please cite this article in press as: Chen X., et al., Robustness of orthogonal-array based composite designs to missing data. J. Statist. Plann. Inference (2017), https://doi.org/10.1016/j.jspi.2017.10.004.

order model. On a more convenient way, the proposed designs can be seen as to choose the best value of α under the minimax loss criterion for OACD when $\alpha \neq 1$.

The paper proceeds as follows. In Section 2, we first give a brief introduction of all composite designs used in this paper. A construction procedure of OACMs and the choice of α (the distance of a nonzero coordinate in an additional design point from the center) for the designs are also discussed in this section. In Section 3, the new designs are compared with the other composite designs under the D-efficiency and the precision of regression coefficient estimates by calculating the generalized scaled standard deviations for full model, linear terms, quadratic terms, and bilinear terms, respectively. The robustness in terms of losses for various composite designs with $\alpha = 1$ is also provided in this section. Concluding remarks are provided in Section 4. All the proofs are deferred to Appendix.

2. Orthogonal-array based composite minimax loss designs

2.1. Composite designs

We first give a brief introduction of composite designs (Box and Draper, 2007). For k factors, denoted by $x_1, \ldots, x_k$, a composite design consists of three parts: (i) n_f cube points with all $x_i = -1$ or 1; (ii) n_α additional points with all $x_i = -\alpha, 0, \alpha$; and (iii) n_0 center points with all $x_i = 0$. A composite design has a total of $n = n_f + n_\alpha + n_0$ points and has 3 or 5 different levels depending on whether $\alpha = 1$ or not.

Different structures of additional points contribute to different composite designs. Box and Hunter (1957) originally proposed to use a full factorial or a fractional factorial design of resolution V in a central composite design. Thus the run size of the standard central composite design becomes very large especially for $k > 5$. To reduce the run size, Draper and Lin (1990) proposed the small composite designs by using Plackett–Burman designs as the factorial portion. In both central composite design and small composite design, $n_\alpha = 2k$ axial points (with one $x_i = \alpha$ or $-\alpha$ and $x_j = 0$ for $j \neq i$) are chosen as the additional points. Morris (2000) introduced the augmented pairs design by adding one point for each pair of the cube points. An augmented pairs design has $n_\alpha = n_f(n_f - 1)/2$ additional points. An augmented pairs minimax loss design (Ahmad et al., 2012) carefully chooses a value of α based on the minimax loss criterion. Hence, augmented pairs designs and augmented pairs minimax loss designs possess the same design structure but may have different α values. An orthogonal-array based composite design (OACD, Xu et al., 2014) is a composite design such that its n_α additional points form a 3-level orthogonal array. 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 which all s^t level-combinations appear equally often in every $N \times t$ submatrix.

2.2. Model description and minimax loss criterion

In this paper, an orthogonal-array based composite minimax loss design (OACM) is constructed by finding a new value of α by employing the minimax loss criterion presented by Akhtar and Prescott (1986). This criterion is popular for constructing designs that are more robust to missing observations. The value of α will be chosen so that the maximum loss of a missing design point is minimized. In order to use the minimax loss criterion, we proceed as follows.

Composite designs are often used to fit a second-order model. For k quantitative factors, the second-order model is

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{j=1}^{k} \beta_{jj} x_{ij}^2 + \sum_{j_1=1}^{k-1} \sum_{j_2=j_1+1}^{k} \beta_{j_1 j_2} x_{ij_1} x_{ij_2} + \epsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

where $\beta_0, \beta_j, \beta_{jj}$ and $\beta_{j_1 j_2}$ are the intercept, linear, quadratic and bilinear (or interaction) terms, respectively, and ϵ_i is a random error, with mean zero, variance σ^2 and independence between any pair of runs. Let $p = (k+1)(k+2)/2$, and $\mathbf{X}$ be the $n \times p$ model matrix in which the ith row $\mathbf{x}_i^T = (1, x_{i1}, \ldots, x_{ik}, x_{i1}^2, \ldots, x_{ik}^2, x_{i1}x_{i2}, \ldots, x_{i(k-1)}x_{ik})$. Then, the above model (1) can be rewritten in matrix notation as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\mathbf{Y} = (y_1, \ldots, y_n)^T$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k, \beta_{11}, \ldots, \beta_{kk}, \beta_{12}, \ldots, \beta_{k-1,k})^T$, and error term $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^T$. Here notation "^{T}" stands for the transpose.

In general, if m observations are missing from an experiment, then m rows in matrix $\mathbf{X}$ will be missing. Particularly, let the ith row corresponding to the ith missing observation be represented by $\mathbf{x}_i^T$ for $1 \le i \le n$, and $\mathbf{X}_{-i}$ be the $n-1$ remaining rows of $\mathbf{X}$ for the reduced design, excluding the missing data. Then $\mathbf{X}^T\mathbf{X}$ and $\mathbf{X}_{-i}^T\mathbf{X}_{-i}$ are the information matrices for the complete design and the reduced design, respectively. Now we may partition the whole model matrix as $\mathbf{X} = \begin{pmatrix} \mathbf{x}_i^T \\ \mathbf{X}_{-i} \end{pmatrix}$. Obviously,

$$\mathbf{X}^T\mathbf{X} = \mathbf{x}_i\mathbf{x}_i^T + \mathbf{X}_{-i}^T\mathbf{X}_{-i}.$$

The relative reduction in $d = |\mathbf{X}^T\mathbf{X}|$ due to a missing observation is called loss. The loss due to the ith observation being missing (Akhtar and Prescott, 1986; Andrews and Herzberg, 1979) is defined by

$$l_i = \frac{|\mathbf{X}^T\mathbf{X}| - |\mathbf{X}_{-i}^T\mathbf{X}_{-i}|}{|\mathbf{X}^T\mathbf{X}|}.$$

Please cite this article in press as: Chen X., et al., Robustness of orthogonal-array based composite designs to missing data. J. Statist. Plann. Inference (2017), https://doi.org/10.1016/j.jspi.2017.10.004.

Table 1
OACDs used in this paper for $k = 3, \ldots, 9$.

k	Two-level portion			Three-level portion	
	Design	n_f	Columns	Design	Columns
3	2^3	8	–	$OA(9)$	(1–3)
4	2^4	16	–	$OA(9)$	(1–4)
5	2_{V}^{5-1}	16	$E = ABCD$	$OA(18)$	(2–6)
6	2_{VI}^{6-1}	32	$F = ABCDE$	$OA(18)$	(1–6)
7	2_{IV}^{7-2}	32	$F = ABCD, G = ABE$	$OA(18)$	(3,1,5,7,4,2,6)
8	2_{IV}^{8-3}	32	$F = ABCD, G = ABE,$ $H = ACE$	$OA(27)$	(1,3,4,5,2,7,8,6)
9	2_{IV}^{9-4}	32	$F = ABCD, G = ABE,$ $H = ACE, J = ADE$	$OA(27)$	(5,6,1,7,2,4,9,3,8)

It can be easily shown that $l_i = \mathbf{x}_i^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_i$. Then,

$$\sum_{i=1}^{n} l_i = \sum_{i=1}^{n} \mathbf{x}_i^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_i = \mathrm{tr}(\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T). \tag{2}$$

Remark 1. Eq. (2) implies that the sum of all losses from the first run to the last run is a constant for a specific design with given run size n and number of parameters p, which implies that a particular loss can be reduced at the cost of increases in other losses. Thus, the most useful criterion for reducing the loss of a single missing observation is to minimize the maximum loss due to the ith observation being missing, which is called the minimax loss criterion. Now the loss due to the ith observation being missing for a given n-run design D can be written as $l_i(D)$, $1 \le i \le n$. Following the minimax loss criterion, the minimax loss design D^* for a set of designs $\mathcal{D}$ satisfies

$$\max_{1\le i\le n}\{l_i(D^*)\} = \min_{D\in\mathcal{D}} \max_{1\le i\le n}\{l_i(D)\}. \tag{3}$$

For augmented pairs minimax loss designs, Ahmad et al. (2012) showed that all specific types of losses are invariant to which one design point is missing. For instance, let $k = 5$, all the losses in an augmented pairs design can be classified into four types: L_f loss for a factorial point missing; L_{α_3} loss for an α_3-type [$(\alpha, \alpha, \alpha, 0, 0)$, $(\alpha, \alpha, 0, \alpha, 0)$, etc.] point missing; L_{α_2} loss for an α_2-type [$(\alpha, \alpha, 0, 0, 0)$, $(\alpha, 0, \alpha, 0, 0)$, etc.] point missing; L_0 loss for a center point (0,0,0,0,0) missing. More details can be found in Ahmad et al. (2012).

For OACDs with $k \le 5$, we apply the same method in Ahmad et al. (2012) to classify the losses. However, the number of types of missing will increase as the number of factors k increases, which leads the classification to be not proper in practice. Thus, due to the composite structure of an OACD, we classify all losses into three types for $k > 5$: L_f loss for a factorial point missing; L_α loss for a point of three-level portion missing, and L_0 loss for a center point missing. In each type, the losses are averaged to one value, since empirical studies show that there was little difference among the losses in each type.

2.3. A step-by-step procedure for constructing OACMs

Based on the previous discussion, we now present a step-by-step procedure for constructing the proposed OACMs:

Algorithm 1.

Step 1. Choose an OACD $\mathbf{H} = \begin{pmatrix}\mathbf{H}_2\\ \mathbf{H}_3\end{pmatrix}$ with k columns, where $\mathbf{H}_2$ is the two-level portion with n_f runs and $\mathbf{H}_3$ is the three-level (i.e., $-1, 0, 1$) portion with n_α runs. Let $\mathbf{H}' = \begin{pmatrix}\mathbf{H}_2\\ \alpha\mathbf{H}_3\end{pmatrix}$. Note that we may add n_0 center points in an OACD;
Step 2. Calculate the lose functions L_f, L_0, and L_{α_i} for $k \le 5$ (or L_α for $k > 5$) based on $\mathbf{H}'$ for the second-order polynomial model (1);
Step 3. Find α_0 such that the maximum loss of $\{L_f, L_{\alpha_i}, L_0\}$ for $k \le 5$ (or $\{L_f, L_\alpha, L_0\}$ for $k > 5$) is minimized;
Step 4. Substitute $\alpha = \alpha_0$ in $\mathbf{H}'$, an OACM with $n = n_f + n_\alpha + n_0$ runs and k columns is constructed.

Remark 2. Note that there are many OACDs for given parameters k, n_f, n_α, n_0 in Step 1. First, there may exist nonisomorphic fractional factorial designs for us to choose. Second, the properties of the resulting design may depend on which two-level column is aligned with which three-level column. In this paper, we use the D-optimal OACDs under the second-order model from Xu et al. (2014), which are listed in Table 1. Note that the D-optimal OACDs are only referred to different column alignments. In this table, $OA(9)$, $OA(18)$ and $OA(27)$ represent orthogonal arrays $OA(9, 3^4, 2)$, $OA(18, 3^7, 2)$ and $OA(27, 3^{13}, 2)$, respectively; 2^k means a full factorial design for k 2-level factors, 2_{R}^{k-p} is a 2^{k-p} fractional factorial design of resolution R. As an example, consider the OACD listed for $k = 5$. For the 2-level factorial portion we use a 2_{V}^{5-1} (i.e., a 16-run fractional factorial design of resolution V) with generator $E = ABCD$, and for the 3-level design we use an $OA(18)$ with columns (2–6), which results in the 34-run OACD in Table 2 with $n_f = 16$, $n_\alpha = 18$ and $n_0 = 0$.

Please cite this article in press as: Chen X., et al., Robustness of orthogonal-array based composite designs to missing data. J. Statist. Plann. Inference (2017), https://doi.org/10.1016/j.jspi.2017.10.004.

Table 2
A 34-runs OACD with five factors and $n_0 = 0$.

Runs	x_1	x_2	x_3	x_4	x_5
1	1	−1	−1	−1	−1
2	−1	1	−1	−1	−1
3	−1	−1	1	−1	−1
4	−1	−1	−1	1	−1
5	−1	−1	−1	−1	1
6	1	1	1	−1	−1
7	1	1	−1	1	−1
8	1	1	−1	−1	1
9	1	−1	1	1	−1
10	1	−1	1	−1	1
11	1	−1	−1	1	1
12	−1	1	1	1	−1
13	−1	1	−1	1	1
14	−1	1	1	−1	1
15	−1	−1	1	1	1
16	1	1	1	1	1
17	$-\alpha$	$-\alpha$	$-\alpha$	$-\alpha$	$-\alpha$
18	0	0	0	0	0
19	α	α	α	α	α
20	$-\alpha$	$-\alpha$	0	0	α
21	0	0	α	α	$-\alpha$
22	α	α	$-\alpha$	$-\alpha$	0
23	$-\alpha$	0	$-\alpha$	α	0
24	0	α	0	$-\alpha$	α
25	α	$-\alpha$	α	0	$-\alpha$
26	$-\alpha$	α	α	0	0
27	0	$-\alpha$	$-\alpha$	α	α
28	α	0	0	$-\alpha$	$-\alpha$
29	$-\alpha$	0	α	$-\alpha$	α
30	0	α	$-\alpha$	0	$-\alpha$
31	α	$-\alpha$	0	α	0
32	$-\alpha$	α	0	α	$-\alpha$
33	0	$-\alpha$	α	$-\alpha$	0
34	α	0	$-\alpha$	0	α

Remark 3. In Step 3 of Algorithm 1, we may use Gauss–Newton iterative method, bisection method or other classical optimal algorithm (cf., Pillo and Giannessi, 1996) to search α_0 from 0.50 to 2.00 (cf., Ahmad and Gilmour, 2010; Ahmad et al., 2012).

Example 1. Consider an example to illustrate the construction of an OACM for five factors using five center runs based on the OACD in Table 2.

Fig. 1 shows a curve for each kind of the loss values. It can be seen from Fig. 1 that when α increases from 0.50 to 2.00, L_f decreases while $L_{\alpha_5}, L_{\alpha_4}, L_{\alpha_3}$ have an increasing sharp, and L_0 keeps more robust and smaller comparing to others. It is interesting to note that losses $L_{\alpha_3}, L_{\alpha_4}, L_{\alpha_5}$ tend to coincide as α increases. In order to compute the value of α at which this maximum loss is minimized, L_f loss is equated with $L_{\alpha_4}, L_{\alpha_3}, L_{\alpha_5}$ turn by turn, and it is observed that the values of α at which L_f is equal to the other three losses $(L_{\alpha_5}, L_{\alpha_4}, L_{\alpha_3})$ are quite similar. Fig. 1 shows that L_f and L_{α_4} are equal for some $1.15 \le \alpha \le 1.20$. By bisection method, the maximum loss of missing one design point will be minimized at $\alpha_0 = 1.1648$ for the OACD with five factors and five center runs.

Remark 4. For other cases, the values of α can be obtained similarly following the previous steps. Note that for comparing the values of $\{L_f, L_\alpha, L_0\}$, we only need to choose the larger value between L_f and L_α, since L_0 are smaller than others in all cases (see Section 3.3). All these α values can be found in the square brackets of Table 3.

3. Comparison results

In this section, the proposed OACMs are compared with other composite designs based on the *D*-efficiency and D_s-optimality.

3.1. Comparison based on D-efficiency

Now we consider the relative (overall) *D*-efficiency of an OACM as compared with the corresponding OACD for $3 \le k \le 9$ and $1 \le n_0 \le 5$. The results are shown in Table 3. Here the relative *D*-efficiency is defined to be

$$D_{\text{eff}} = \left(\frac{|\mathbf{X}^T\mathbf{X}|_{\text{OACM}}}{|\mathbf{X}^T\mathbf{X}|_{\text{OACD}}} \right)^{1/p},$$

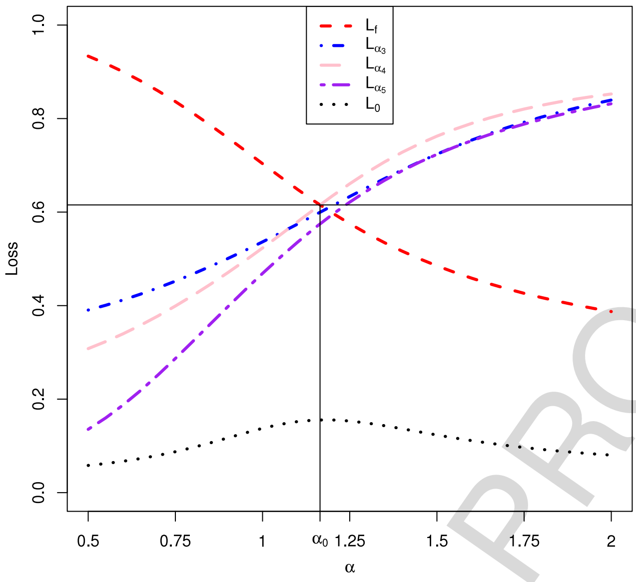

Fig. 1. Choice of α for the OACD with $k = 5$, $n_0 = 5$.

Table 3
D-efficiencies of OACMs relative to OACDs.

n_0	Number of factors k						
	3	4	5	6	7	8	9
1	1.0497	0.9879	1.2851	1.1126	1.0856	1.1821	1.1978
	[1.0486]	[0.9877]	[1.1775]	[1.1056]	[1.0727]	[1.1139]	[1.1155]
2	1.0757	1.0155	1.2946	1.1307	1.1068	1.1967	1.2297
	[1.0652]	[1.0148]	[1.1799]	[1.1178]	[1.0876]	[1.1203]	[1.1305]
3	1.0840	1.0272	1.2998	1.1408	1.1178	1.2061	1.2436
	[1.0754]	[1.0251]	[1.1811]	[1.1241]	[1.0948]	[1.1244]	[1.1365]
4	1.0910	1.0387	1.3032	1.1442	1.1243	1.2096	1.2511
	[1.0804]	[1.0350]	[1.1819]	[1.1256]	[1.0989]	[1.1255]	[1.1395]
5	1.0967	1.0400	1.2717	1.1476	1.1291	1.2135	1.2563
	[1.0845]	[1.0358]	[1.1648]	[1.1274]	[1.1019]	[1.1271]	[1.1416]

Note: The α values used to construct the corresponding OACMs are presented in the square brackets.

where $p = (k + 1)(k + 2)/2$. Designs with high relative *D*-efficiency are often preferred, especially among the designs with the same number of runs. From Table 3, it can be seen that OACMs perform relatively better than the corresponding OACDs for all cases except for the case when $k = 4$, $n_0 = 1$. For this special case, the OACM with $\alpha = 0.9877$ has only less than 2% *D*-efficiency loss compared with the corresponding OACD.

For two-level fractional factorial designs, level permutations will not change the determinant of the full information matrix $|\mathbf{X}^T\mathbf{X}|$. However, for OACDs, it can be easily verified that level permutations for different portions may change the value of $|\mathbf{X}^T\mathbf{X}|$. Here, level permutations are limited within each portion of an OACD to keep the orthogonal structure unchanged. Moreover, for a given OACD, Xu et al. (2014) noted that the *D*-efficiency may also depend on which two-level column is aligned with which three-level column. Hence, for a given OACD or OACM, the determinant of the information matrix may be changed due to both level permutations and column permutations, which makes the situation for investigating *D*-optimal OACMs more complex. However, the determinant remains unchanged for some special cases as shown below.

Theorem 1. *Under the condition that the two-level portion of an OACD or OACM has resolution V, the value of* $|\mathbf{X}^T\mathbf{X}|_{\text{OACD}}$ *or* $|\mathbf{X}^T\mathbf{X}|_{\text{OACM}}$ *will be a constant regardless of any column permutation or level permutation for the two-level portion.*

Theorem 2. *Under the same condition as in Theorem 1, the value of* $|\mathbf{X}^T\mathbf{X}|_{\text{OACD}}$ *or* $|\mathbf{X}^T\mathbf{X}|_{\text{OACM}}$ *will be a constant regardless of any column permutation or under some restricted level permutations for the three-level portion. Here the level permutations are restricted to two cases: mapping* $(-1, 0, 1)$ *to* $(1, 0, -1)$*, and mapping* $(1, 0, -1)$ *to* $(-1, 0, 1)$.

From Theorems 1 and 2, the determinant of information matrix remains unchanged under level permutations and column permutations of both two-level and three-level portions in some special cases.

Remark 5. The trace of information matrix remains unchanged for any level permutation of the two-level portion. Actually, under the same notations in the proof of Theorem 1, let $\mathbf{X}_2$ be the original two-level portion of model matrix, $\mathbf{X}_2^*$ be the new two-level portion after level permutation, we have

$$\mathrm{tr}(\mathbf{X}_2^{*T}\mathbf{X}_2^*) = \mathrm{tr}(\mathrm{diag}^T(\mathbf{e})\mathbf{X}_2^T\mathbf{X}_2\mathrm{diag}(\mathbf{e})) = \mathrm{tr}(\mathbf{X}_2^T\mathbf{X}_2\mathrm{diag}^T(\mathbf{e})\mathrm{diag}(\mathbf{e})) = \mathrm{tr}(\mathbf{X}_2^T\mathbf{X}_2)$$

for any level permutation of the two-level portion, which does not require the two-level portion to have resolution at least V. More details on "diag($\mathbf{e}$)″ can be found in the Appendix. A similar conclusion can be obtained for the column permutations.

3.2. Comparison based on D_s-optimality

In this subsection, we will compare the design efficiencies in estimating a subset of the model parameters. First we divide the model parameters into three groups: the k linear parameters ($\beta_j, j = 1, \ldots, k$), the k pure quadratic parameters ($\beta_{jj}, j = 1, \ldots, k$), and the $k(k-1)/2$ bilinear parameters (β_{ij}, $1 \le i < j \le k$). Corresponding to the parameter partition, the $n \times p$ model matrix can be partitioned as

$$\mathbf{X} = (\mathbf{1}, \mathbf{X}_l, \mathbf{X}_q, \mathbf{X}_b),$$

where $\mathbf{1}$ is an $n \times 1$ vector with all elements unity, representing the constant term in the model, and $\mathbf{X}_l$, $\mathbf{X}_q$, $\mathbf{X}_b$ are the portions of model matrix representing the linear, quadratic, and bilinear (linear × linear) terms with orders $n \times k$, $n \times k$, $n \times (k-1)k/2$, respectively. Let $\overline{\mathbf{X}}_s$ be the remaining portion of $\mathbf{X}$ by deleting $\mathbf{X}_s$. Here, the OACMs are compared with the central composite designs, small composite designs, augmented pairs designs, augmented pairs minimax loss designs, and OACDs in terms of generalized scaled deviations (Morris, 2000; Ahmad et al., 2012) for the full model ($\widetilde{D}$), linear terms ($\widetilde{D}_l$), quadratic terms ($\widetilde{D}_q$), and bilinear terms ($\widetilde{D}_b$), respectively. These measures are defined by

$$\widetilde{D} = \sqrt{(n|\mathbf{X}^T\mathbf{X}|^{-1/p})},$$

and

$$\widetilde{D}_s = \sqrt{(n|\mathbf{X}_s^T\mathbf{X}_s - \mathbf{X}_s^T\overline{\mathbf{X}}_s(\overline{\mathbf{X}}_s^T\overline{\mathbf{X}}_s)^{-1}\overline{\mathbf{X}}_s^T\mathbf{X}_s|^{-1/p_s})}, \ \text{for } s = l, q, b,$$

where $p_l = k$, $p_q = k$, $p_b = k(k-1)/2$. It is clear that the generalized scaled deviations used here are equivalent to the square roots of reciprocals of the overall D-efficiency and D_s-efficiency in Xu et al. (2014). Hence, for a fixed value of α, designs which minimize $\widetilde{D}$ are D-optimal for the second-order polynomial model, whereas those that minimize $\widetilde{D}_l$, $\widetilde{D}_q$, and $\widetilde{D}_b$ are D_s-optimal for estimating the corresponding subsets of coefficients.

Table 4 provides details of the generalized scaled deviations $\widetilde{D}$, $\widetilde{D}_l$, $\widetilde{D}_q$, and $\widetilde{D}_b$ for all composite designs with $3 \le k \le 9$ and $n_0 = 5$. The values of α obtained by minimax loss criterion for the corresponding designs are listed in the first column. Note that the values of α for APMs were derived by Ahmad et al. (2012). The small composite designs employed here are constructed by taking columns from Plackett–Burman designs of Draper and Lin (1990). In general, central composite designs have a well performance on generalized scaled deviations except for $\widetilde{D}_q$ but suffer larger run sizes especially for $k > 7$. To compare the results of designs with nearly the same run sizes, we drop central composite designs in Fig. 2, which shows a graphical representation of the various generalized scaled deviations of the designs under consideration for $k = 3, \ldots, 9$. Only values for OACMs are represented in solid line, and others are in dotted lines.

Fig. 2(a) compares the overall D-efficiencies for all designs. It is clear that in general the OACMs have the lowest $\widetilde{D}$ values, followed by the OACDs, augmented pairs minimax loss designs, augmented pairs designs, and small composite designs. For $k = 9$, the $\widetilde{D}$ value of the OACM is only 0.1323 (about 8.9%) higher than that of the corresponding augmented pairs minimax loss design. However, the OACM has a smaller run size, i.e. $n = 64$, than the corresponding augmented pairs minimax loss design whose run size is $n = 83$.

Fig. 2(b) compares the $\widetilde{D}_l$ values for estimating the linear parameters. It is clear that the general pattern is similar to the generalized scaled deviation for the full model considered in Fig. 2(a): the OACMs have the best $\widetilde{D}_l$ values, followed by the OACDs, augmented pairs minimax loss designs, augmented pairs designs, small composite designs, except for the case of $k = 9$. For $k = 9$, the $\widetilde{D}_l$ values of the OACM and augmented pairs minimax loss design are 1.3617 and 1.2435, respectively. However, the OACM has a smaller run size as discussed previously.

Fig. 2(c) shows the $\widetilde{D}_q$ values for the pure quadratic terms. Unlike other generalized scaled deviations, the OACMs and OACDs perform not so well. As noted in Xu et al. (2014), more design points located at the corners will lead to higher D, D_l, and D_b-efficiency while more design points located at the mid-sides and center will increase the quadratic efficiency. The OACDs and OACMs have relatively more corner points and less center point replicates than the corresponding augmented pairs designs, augmented pairs minimax loss designs, and small composite designs. This is the reason why the OACDs and OACMs have larger $\widetilde{D}_q$'s (equivalently smaller D_q-efficiency in Xu et al. (2014)).

Fig. 2(d) shows the performance of the $\widetilde{D}_b$ values for estimating the bilinear coefficients. Both OACDs and OACMs perform better than others. Moreover, the OACMs have the best $\widetilde{D}_b$ values among almost all cases for $3 \le k \le 9$.

Please cite this article in press as: Chen X., et al., Robustness of orthogonal-array based composite designs to missing data. J. Statist. Plann. Inference (2017), https://doi.org/10.1016/j.jspi.2017.10.004.

Table 4
Generalized scaled deviations of different designs with $n_0 = 5$.

	Design	n	$\widetilde{D}$	$\widetilde{D}_l$	$\widetilde{D}_q$	$\widetilde{D}_b$
$k = 3$	CCD	19	1.6196	1.3784	2.3489	1.5411
	SCD	15	2.0992	2.7386	2.2333	3.3541
	APD	15	2.0992	2.7386	2.2333	3.3541
$\alpha = 1.4142$	APM	15	1.6423	1.9365	1.3936	2.7386
	OACD	22	1.5411	1.1941	3.1086	1.2913
$\alpha = 1.0845$	OACM	22	1.4715	1.1779	2.7901	1.2707
$k = 4$	CCD	29	1.5844	1.2693	2.8329	1.3463
	SCD	21	2.0124	2.6499	2.5199	2.4227
	APD	41	1.7180	1.4693	2.0606	1.8806
$\alpha = 1.0353$	APM	41	1.6648	1.4281	1.9541	1.8266
	OACD	30	1.5411	1.1941	3.1087	1.2913
$\alpha = 1.0358$	OACM	30	1.5111	1.1870	2.9669	1.2828
$k = 5$	CCD	31	1.6229	1.3123	3.0056	1.3919
	SCD	27	2.1451	3.0715	2.8462	2.4980
	APD	41	1.8261	1.4842	2.0739	2.0554
$\alpha = 1.0997$	APM	41	1.6474	1.3662	1.7819	1.8398
	OACD	39	1.4987	1.2331	2.6651	1.3310
$\alpha = 1.1824$	OACM	39	1.3117	1.1866	2.0316	1.2077
$k = 6$	CCD	49	1.5491	1.2005	3.7701	1.2374
	SCD	33	2.0992	4.0620	3.1707	2.2287
	APD	41	1.9338	1.4927	2.0851	2.2113
$\alpha = 1.0692$	APM	41	1.7819	1.4038	1.8701	2.0154
	OACD	55	1.4450	1.1489	3.2660	1.2151
$\alpha = 1.1274$	OACM	55	1.3535	1.1302	2.7454	1.1737
$k = 7$	CCD	83	1.5048	1.1214	4.9384	1.1388
	SCD	43	2.2091	3.7007	3.6369	2.3268
	APD	41	2.0378	1.4974	2.0948	2.3496
$\alpha = 0.8165$	APM	41	2.6731	1.8442	2.9282	3.2239
	OACD	55	1.5402	1.2050	3.6393	1.3767
$\alpha = 1.1019$	OACM	55	1.4494	1.1971	3.1728	1.3228
$k = 8$	CCD	85	1.4921	1.1348	5.0926	1.1524
	SCD	57	1.9580	2.6464	4.2161	1.8419
	APD	83	1.9701	1.4058	2.0429	2.1998
$\alpha = 1.2659$	APM	83	1.4163	1.2288	1.3502	1.5332
	OACD	64	1.5715	1.2533	3.2149	1.4859
$\alpha = 1.1271$	OACM	64	1.4266	1.2341	2.6924	1.3585
$k = 9$	CCD	151	1.4639	1.0777	6.8444	1.0861
	SCD	63	2.1524	4.3787	4.4929	2.1062
	APD	83	2.0424	1.4058	2.0467	2.2861
$\alpha = 1.2429$	APM	83	1.4868	1.2435	1.3964	1.6135
	OACD	64	1.8148	1.3692	4.3663	1.8465
$\alpha = 1.1416$	OACM	64	1.6191	1.3617	3.6401	1.6660

Note: APD: augmented design; APM: pairs augmented pairs minimax loss design;
CCD: central composite design; SCD: small composite design.

Remark 6. In conclusion, OACMs can be seen as special cases of OACDs by carefully choosing the α values. Hence among the small composite designs, augmented pairs designs, augmented pairs minimax loss designs, OACDs, and OACMs with nearly the same run sizes, the OACMs usually possess best generalized scaled deviation values $\widetilde{D}$, $\widetilde{D}_l$, $\widetilde{D}_b$ for estimating the constant, linear, and bilinear terms.

3.3. *Comparison results when* $\alpha = 1$

Based on the empirical experience from Tables 3 and 4, the determinants of the information matrices of APDs, APMs, OACDs and OACMs depend on the values of α. Now we consider the robustness in terms of losses (i.e., L_f, L_α, L_0) for various composite designs (CCD, SCD, APD and OACD) with $\alpha = 1$.

The losses L_f, L_α, L_0 of various composite designs with $n_0 = 5$ and $3 \le k \le 9$ have different patterns as displayed in panels (a), (b) and (c) in Fig. 3, respectively. Firstly, losses L_f and L_α are larger than loss L_0 for all kinds of composite designs. That is, cube points and additional points provide more information for estimating parameters in model (1). Secondly, from Fig. 3(a) and 3(b), for most of the composite designs, the L_f as well as L_α values differ greatly for different values of k. The L_0 values of all kinds of composite designs in Fig. 3(c) have a stable pattern, especially for the L_0 of OACDs. And the L_0 values of CCDs are almost the same as that of SCDs. Moreover, the L_0 values of OACDs are larger than those of other composite designs for $3 \le k \le 9$.

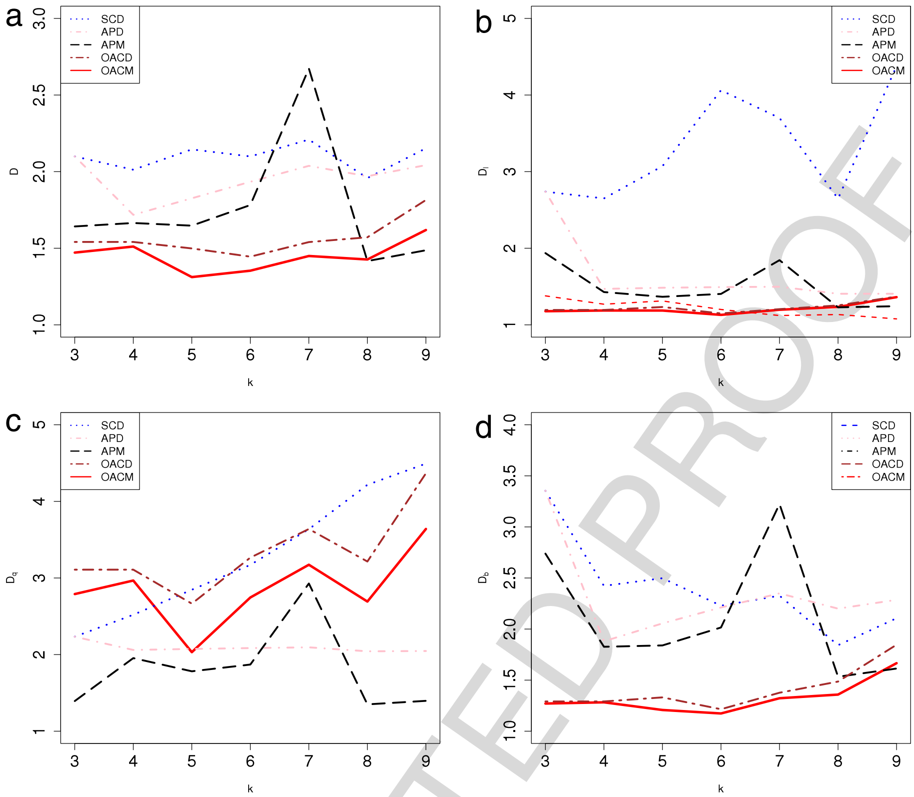

Fig. 2. Comparisons on $\widetilde{D}$ (a), $\widetilde{D}_l$ (b), $\widetilde{D}_q$ (c), and $\widetilde{D}_b$ (d) of some composite designs with $3 \le k \le 9$ and $n_0 = 5$.

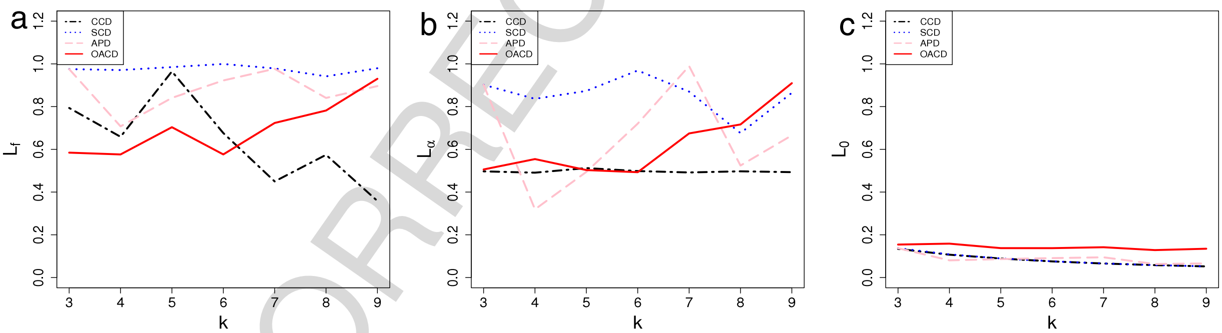

Fig. 3. Comparisons on L_f (a), L_α (b), and L_0 (c) of composite designs with $3 \le k \le 9$ and $n_0 = 5$.

In fact, from Remark 1, the sum of all losses from the first run to the last run equals the number of parameters p. So a certain loss will generally decrease as n increases. Fig. 4 shows the performances of losses L_f, L_α, L_0 of various composite designs with $k = 5$ for $1 \le n_0 \le 9$. Now the run size n increases as the number of center point n_0 increases, which leads to a decreasing pattern of L_0 as shown in Fig. 4(c). However, this influence becomes quite small for cube points and additional points. As shown in Fig. 4(a) and 4(b), all losses are quite consistent for different values of n_0. Note that OACDs have relatively smaller values of L_f and L_α, but larger values of L_0.

4. Discussion and further work

In this paper, the issue of orthogonal-array based composite design (OACD) with missing observations is considered. Orthogonal-array based composite minimax loss designs (OACMs) are constructed, which are more robust to missing observations. The construction uses the minimax loss criterion to choose the α value for a whole set of OACD points in the experimental region. The OACMs in general, perform better than the OACDs, small composite designs, augmented pairs

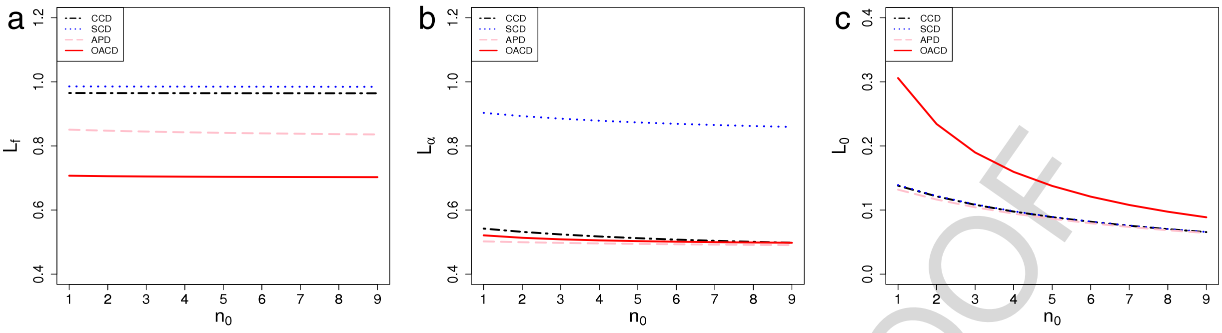

Fig. 4. Comparisons on L_f (a), L_α (b), and L_0 (c) of composite designs with $1 \le n_0 \le 9$ and $k = 5$.

designs, augmented pairs minimax loss designs, and are more efficient for estimating linear and bilinear coefficients of the second-order model. For estimating the linear and bilinear terms, OACMs perform the best in almost all cases as shown in Section 3. For estimating the quadratic terms, the augmented pairs designs and augmented pairs minimax loss designs have the best $\widetilde{D}_q$ values for all cases as in Morris (2000) and Xu et al. (2014).

The optimal α values of OACMs are carefully chosen by comparing losses L_f, L_α and L_0 in this paper. The reason is that the losses are always equal in each type and the comparison for various composite designs can be more convenient in this way. Actually, from Section 3.3, since loss l_0 is quite smaller than L_f and L_α, we only need to compare the losses L_f and L_α based on the minimax loss criterion. As an anonymous reviewer suggested, we may generalize the optimization procedure in various ways. For instance, the minimax criterion (3) is equivalent to

$$\min_{D \in \mathcal{D}} \sum_{1 \le i \le n} l_i^t(D) \tag{4}$$

as t tends to infinity. Thus the minimax lose criterion can also be done via (4). Moreover, we may combine the three types of losses into one objective function by some weight methods as follows.

$$\min_{D \in \mathcal{D}} w_1 L_f + w_2 L_\alpha + w_3 L_0, \tag{5}$$

where w_1, w_2, w_3 are some weight functions. However, this alternative way is not necessarily equivalent to the minimax loss criterion.

Note that the constructions of OACDs and OACMs usually are not unique for given parameters n_f, n_α, n_0, k due to their column alignments or level permutations. One issue needs to consider is to construct OACMs from OAs with other optimal properties, such as uniform OAs (Tang et al., 2012; Tang and Xu, 2014), and space-filling OAs (Zhou and Xu, 2014). This is under progress.

Acknowledgments

The authors thank the Executive-Editor, an Associate-Editor and two referees for their valuable comments. Xue-Ping Chen would like to thank Dr. Jinyu Yang for a useful discussion on OACDs. This work was supported by the National Natural Science Foundation of China (Nos. 11771220, 11431006, 11501305, and 11601195), Natural Science Foundation of Jiangsu Province of China (No. BK20160289), Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No. 16 kJB110005), China Postdoctoral Science Foundation (No. 2016M591378), Tianjin Development Program for Innovation and Entrepreneurship, Tianjin "131" Talents Program, and Project 613319. The authorship is listed in alphabetic order.

Appendix

Proof of Theorem 1. Without loss of generality, let $n_0 = 0$. We now divide the model matrix $\mathbf{X}$ as

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_2 \\ \mathbf{X}_3 \end{pmatrix},$$

where $\mathbf{X}_2$ denotes the first n_f runs corresponding to the two-level portion, and $\mathbf{X}_3$ denotes the remaining n_α runs corresponding to the three-level portion. Hence, we have

$$\mathbf{X}^T\mathbf{X} = \mathbf{X}_2^T\mathbf{X}_2 + \mathbf{X}_3^T\mathbf{X}_3.$$

Please cite this article in press as: Chen X., et al., Robustness of orthogonal-array based composite designs to missing data. J. Statist. Plann. Inference (2017), https://doi.org/10.1016/j.jspi.2017.10.004.

Since the two-level fractional factorial design has resolution at least V, $\mathbf{X}_2^T\mathbf{X}_2 = n_f I_p$. For any column permutation, let $\mathbf{X}_2^*$ be the new two-level portion of model matrix, $\mathbf{P}$ be the corresponding permutation matrix. Then, $\mathbf{X}_2^* = \mathbf{X}_2\mathbf{P}$, and

$$\mathbf{X}_2^{*T}\mathbf{X}_2^* = \mathbf{P}^T\mathbf{X}_2^T\mathbf{X}_2\mathbf{P} = n_f\mathbf{P}^T\mathbf{I}_p\mathbf{P} = n_f\mathbf{I}_p.$$

For any level permutation of two-level columns, we have

$$\mathbf{X}_2^* = \mathbf{X}_2\text{diag}(\mathbf{e}),$$

where $\mathbf{e}$ is a vector with entries 1 or -1, diag($\mathbf{e}$) is a diagonal matrix with the elements from vector $\mathbf{e}$. So

$$\mathbf{X}_2^{*T}\mathbf{X}_2^* = \text{diag}^T(\mathbf{e})\mathbf{X}_2^T\mathbf{X}_2\text{diag}(\mathbf{e}) = n_f\text{diag}^T(\mathbf{e})\mathbf{I}_p\text{diag}(\mathbf{e}) = n_f\mathbf{I}_p.$$

That is, due to this orthogonal structure between any pair of columns, $\mathbf{X}_2^T\mathbf{X}_2$ remains unchanged for any level permutation or column permutation of the two-level columns. ■

Proof of Theorem 2. Following the same notations in the proof of Theorem 1, let $\mathbf{X}_3^*$ be the new three-level portion of model matrix after a certain column permutation of the three-level portion, and $\mathbf{P}$ be the corresponding permutation matrix. Then, we have $\mathbf{X}_3^* = \mathbf{X}_3\mathbf{P}$. Let $\mathbf{X}^* = \begin{pmatrix}\mathbf{X}_2\\ \mathbf{X}_3^*\end{pmatrix}$. It can be shown that the determinant of information matrix $|\mathbf{X}^{*T}\mathbf{X}^*|$ will not be changed after the column permutation. In fact,

$$\begin{aligned}|\mathbf{X}^{*T}\mathbf{X}^*| &= |\mathbf{X}_2^T\mathbf{X}_2 + \mathbf{X}_3^{*T}\mathbf{X}_3^*|\\ &= |n_f\mathbf{I}_p + \mathbf{P}^T\mathbf{X}_3^T\mathbf{X}_3\mathbf{P}|\\ &= |\mathbf{P}^T(n_f\mathbf{I}_p + \mathbf{X}_3^T\mathbf{X}_3)\mathbf{P}|\\ &= |n_f\mathbf{I}_p + \mathbf{X}_3^T\mathbf{X}_3||\mathbf{P}^T\mathbf{P}|\\ &= |\mathbf{X}^T\mathbf{X}|.\end{aligned}$$

Level permutations are restricted to two cases: $(1, 0, -1)$ to $(-1, 0, 1)$ or vice versa. In both cases, we still let $\mathbf{X}_3^*$ be the new three-level portion of model matrix. Hence, we have $\mathbf{X}_3^* = \mathbf{X}_3\text{diag}(\mathbf{e})$, With the similar steps, we have

$$\begin{aligned}|\mathbf{X}^{*T}\mathbf{X}^*| &= |\mathbf{X}_2^T\mathbf{X}_2 + X_3^{*T}\mathbf{X}_3^*|\\ &= |n_f\mathbf{I}_p + \text{diag}^T(\mathbf{e})\mathbf{X}_3^T\mathbf{X}_3\text{diag}(\mathbf{e})|\\ &= |\text{diag}^T(\mathbf{e})(n_f\mathbf{I}_p + \mathbf{X}_3^T\mathbf{X}_3)\text{diag}(\mathbf{e})|\\ &= |n_f\mathbf{I}_p + \mathbf{X}_3^T\mathbf{X}_3||\text{diag}^T(\mathbf{e})\text{diag}(\mathbf{e})|\\ &= |\mathbf{X}^T\mathbf{X}|.\end{aligned}$$

This ends the proof. ■

References

Ahmad, T., Akhtar, M., Gilmour, S.G., 2012. Multilevel augmented pairs second-order response surface designs and their robustness to missing data. Comm. Statist. Theory Methods 41, 437–452.
Ahmad, T., Gilmour, S.G., 2010. Robustness of subset response surface designs to missing observations. J. Statist. Plann. Inference 140, 92–103.
Akhtar, M., Prescott, P., 1986. Response surface designs robust to missing observations. Comm. Statist. Simulation Comput. 15, 345–363.
Andrews, D.F., Herzberg, A.M., 1979. The robustness and optimality of response surface designs. J. Statist. Plann. Inference 3, 249–257.
Box, G.E.P., Draper, N.R., 2007. Response Surfaces, Mixtures, and Ridge Analyses, second ed.. Wiley, New York.
Box, G.E.P., Hunter, J.S., 1957. Multi-factor experimental designs for exploring response surfaces. Ann. Math. Statist. 28, 195–241.
Ding, X., Xu, H., Hopper, C., Yang, J., Ho, C.-M., 2013. Use of fractional factorial designs in antiviral drug studies. Qual. Reliab. Eng. Int. 29, 299–304.
Draper, N.R., Lin, D.K.J., 1990. Small response-surface designs. Technometrics 32, 187–194.
Morris, M.D., 2000. A class of three-level experimental designs for response surface modeling. Technometrics 42, 111–121.
Pillo, G.D., Giannessi, F., 1996. Nonlinear Optimization and Applications. Plenum Publishing, New York.
Tang, Y., Xu, H., 2014. Permuting regular fractional factorial designs for screening quantitative factors. Biometrika 101, 333–350.
Tang, Y., Xu, H., Lin, D.K.J., 2012. Uniform fractional factorial designs. Ann. Statist. 40, 891–907.
Xu, H., Jaynes, J., Ding, X., 2014. Combining two-level and three-level orthogonal arrays for factor screening and response surface exploration. Statist. Sinica 24, 269–289.
Zhou, Y.D., Xu, H., 2014. Space-filling fractional factorial designs. J. Amer. Statist. Assoc. 109, 1134–1144.
Zhou, Y.D., Xu, H., 2016. Composite designs based on orthogonal arrays and definitive screening designs. J. Amer. Statist. Assoc.. http://dx.doi.org/10.1080/01621459.2016.1228535.

Please cite this article in press as: Chen X., et al., Robustness of orthogonal-array based composite designs to missing data. J. Statist. Plann. Inference (2017), https://doi.org/10.1016/j.jspi.2017.10.004.