ELSEVIER

Journal of Statistical Planning and Inference 71 (1998) 331–348

journal of statistical planning and inference

Construction of response surface designs for qualitative and quantitative factors

C.F.J. Wu [a,*], Yuan Ding [b]

[a] Department of Statistics, University of Michigan, Ann Arbor, MI 48109-1027, USA

[b] Department of Mathematics and Statistics, Concordia University, Montreal, Quebec, Canada H3G 1M8

Received 26 June 1995; received in revised form 4 December 1997

Abstract

A general approach is proposed for constructing response surface designs of economical size for qualitative and quantitative factors. It starts with an efficient design (e.g. central composite design) for the quantitative factors and then partitions the design points into groups corresponding to different level combinations of the qualitative factors. Good designs are selected to ensure high estimation efficiency for models that reflect some stated objectives. The method is illustrated with the construction of designs for one qualitative factor. Some designs for two qualitative factors are constructed and extensions considered. Algorithmic construction of designs is briefly discussed.

Keywords: Central composite design; Determinant criterion; Fractional factorial design; Interchange algorithm; Plackett–Burman design

1. Introduction

An unresolved problem in response surface methodology is the construction of economical designs for both quantitative and qualitative factors. It may arise in situations in which some of the factors are qualitative by nature. Consider, for example, a machining process for slab caster rolls in a steel plant. To improve the machining time while keeping a low rating of surface roughness, four factors are identified as potentially important: (i) feed, which is the distance the tool advances in one revolution, (ii) speed at which the surface moves past the cutting tool, (iii) lead angle at which the tool meets the work piece, and (iv) insert, which is the replaceable part of the cutting tool that does the cutting. The first three factors are quantitative, while the fourth can only take two shapes: round and square. Without the availability of economical designs, the investigator may adopt one of the following two approaches.

* Corresponding author.

0378-3758/98/$19.00
PII: S0378-3758(98)00003-2

In the first approach, a second-order response surface design such as the central composite design is used for each level of the qualitative factor (i.e., for round and for square). It can take a large number of runs and may be impractical. Alternatively, one may ignore the difference between the two types of factors and use a standard design for both types. As Draper and John (1988) aptly pointed out, standard designs such as the central composite designs may not be suitable because they would require four to five levels which the qualitative factors may not have. Were it possible to find that many quantitative levels, they would still be meaningless for the qualitative factors.

The importance of the problem was recognized by Cox (1984), who posed it as one of eleven open problems in experimental design. Draper and John (1988) were the first to tackle it seriously. They discussed the relations between designs and models and gave some illustrative examples. In this paper we give a systematic method for constructing designs of economical size and discuss the underlying objectives and models. An obvious requirement for good designs is that when collapsed over the qualitative factors they should possess desirable properties of standard response surface designs for quantitative factors (Box and Draper, 1987; Khuri and Cornell, 1987; Myers and Montgomery, 1995), which is captured by Objectives C and D of Section 2. Some issues related to the presence of qualitative factors are addressed by objectives A, B, and C. In Section 3 we outline a method of constructing designs to meet these objectives. The main idea is to start with an efficient design for the quantitative factors and then partition its points into groups. Each group corresponds to a level combination of the qualitative factors. From these designs we then select those with high overall efficiencies as measured by the determinant criteria (3.1) and (3.3) for several models. In Sections 4 and 5 we further develop this idea by constructing designs with one qualitative factor that use central composite designs as the starting designs. A related class of designs for one qualitative and two quantitative factors is given in Section 6. Strategies for constructing designs with two qualitative and two quantitative factors are briefly discussed in Section 7. Section 8 outlines some extensions. In Section 9 we discuss algorithmic construction of designs and compare it with the approach taken in the paper.

2. Objectives and supporting models

The designs will be constructed to meet the following objectives. Denote the quantitative factors by $x_1,\ldots,x_k$ and the qualitative factors by $z_1,\ldots,z_r$.

A. The overall design is efficient for a model that is second order in $x_1,\ldots,x_k$, and has the main effects of $z_1,\ldots,z_r$ and the interactions between x_i and z_j.

B. At each combination of the qualitative factors or each level of a qualitative factor z_j, it is an efficient first-order design in $x_1,\ldots,x_k$.

C. The design in A consists of two parts: the first part is a first-order design for both $x_1,\ldots,x_k$ and $z_1,\ldots,z_r$, and the second part can be viewed as a sequential addition to the first part so that the expanded design is second-order.

D. When collapsed over the levels of $z_1,\ldots,z_r$, it is an efficient second-order design for $x_1,\ldots,x_k$.

Objective A is the most important. It ensures that the first and second order effects of x_i, the main effects of z_j, and the interactions between x_i and z_j can all be estimated with high overall efficiency. Because the effects of x_i may vary with the levels of z_j, it is also desirable to have the design at each combination of $z_1,\ldots,z_r$ (or at each level of z_j) that allows separate estimation of the first-order effects of x_i. This goal is met by implementing Objective B. Objective C enables the experiment to be conducted in two stages. The initial experiment allows the estimation of the main effects of x_i and z_j and some of their interactions. If warranted, the second experiment can be conducted to ensure the estimation of second-order effects in the combined experiment. Objective D ensures that, when there is no significant difference among the z_j's, the combined design is a second-order design with good overall properties. The issue of estimation efficiency will be addressed in the next section. If these objectives cannot be met simultaneously, priority should be given to A and B. Objective C is important only if a two-stage experiment is contemplated. Objective D is emphasized only if the effects of the qualitative factors are believed to be small.

These objectives can be stated more precisely with the aid of regression models. For simplicity we only consider $r=1$, the case of one qualitative factor. A second-order model for x_i and z is given by

$$E(y)=\sum_{z=1}^{m} W_z\left(\beta_{oz}+\sum_{i=1}^{k}\beta_{iz}x_i\right)+\sum_{i,j=1}^{k}\beta_{ij}x_ix_j, \tag{2.1}$$

where m is the number of levels of the qualitative factor, W_j is 1 when y is taken at level j of the variable z and 0 otherwise, β_{oz} is the constant term and β_{iz} is the slope of x_i, both depending on the choice of z. If the run size is small, we may only be able to fit the following submodels of Eq. (2.1):

$$E(y)=\sum_{z=1}^{m} W_z\left(\beta_{oz}+\sum_{i=1}^{k}\beta_{iz}x_i\right)+\sum_{i=1}^{k}\beta_{ii}x_i^2+\text{some of }\beta_{ij}x_ix_j(i<j), \tag{2.2a}$$

$$E(y)=\sum_{z=1}^{m} W_z\beta_{oz}+\sum_{i=1}^{k}(\beta_i x_i+\beta_{ii}x_i^2)+\text{ some of }W_z\beta_{iz}x_i\text{ and }\beta_{ij}x_ix_j(i<j). \tag{2.2b}$$

Model (2.2a) excludes some interaction terms $\beta_{ij}x_ix_j$, $i<j$, in model (2.1). Model (2.2b) further excludes some $\beta_{iz}x_i$ terms. Objective A stipulates that the overall design allows one of these models to be fitted with high efficiency.

Objective B requires that the coefficients in the model

$$E(y)=\beta+\sum_{i=1}^{k}\beta_i x_i+\text{some of }\beta_{ij}x_ix_j(i<j), \tag{2.3}$$

be estimated with high efficiency from the design at each level of the qualitative factor. By excluding the x_i^2 terms in Eqs. (2.2a) and (2.2b), we have the following submodels:

$$E(y)=\sum_{z=1}^{m} W_z\left(\beta_{oz}+\sum_{i=1}^{k}\beta_{iz}x_i\right)+\text{some of }\beta_{ij}x_ix_j(i<j), \tag{2.4a}$$

$$E(y)=\sum_{z=1}^{m} W_z\beta_{oz}+\sum_{i=1}^{k}\beta_i x_i+\text{some of }W_z\beta_{iz}x_i\text{ and }\beta_{ij}x_ix_j(i<j). \tag{2.4b}$$

Objective C requires that the coefficients in Eq. (2.4a) or Eq. (2.4b) be estimated with high efficiency from the first-order design.

3. Selection criteria and construction method

In this section we describe a method for constructing designs that can meet the objectives stated in Section 2. Because many designs for quantitative factors are available, we start with a good design for $x_1,\dots,x_k$, say, $\boldsymbol{d}$. If a large run size is affordable, we can repeat $\boldsymbol{d}$ for each level combination of the qualitative factors. In many practical situations we do not have this luxury. Instead we partition the points in $\boldsymbol{d}$ into several groups, each group being associated with a level combination of the qualitative factors. Among all possible partitionings, we choose those to satisfy Objectives A–C. There are several choices for $\boldsymbol{d}$. See Khuri and Cornell (1987). In order to explain the construction method in detail, we need to fix the choice of $\boldsymbol{d}$. Because of its ubiquitous role in response surface methodology, we choose central composite designs for $\boldsymbol{d}$. *The proposed approach works generally for other choices of* $\boldsymbol{d}$.

For $r=1$ and $m=2$ the constructed designs can be generically represented in Table 1 . For $x_1,\dots,x_k$, the first $t(=2^{k-p})$ runs are chosen according to a 2^{k-p} fractional factorial design with high resolution and $x_i=\pm 1$. The $(t+1)$st and $(t+2)$nd runs are at the center point. The last $2k$ runs are the 'star points' whose distance from the origin is α. The value of α may be chosen so that the designs for $x_1,\dots,x_k$ are rotatable. Such designs are particularly useful when the experimental region is naturally spherical. In Table 1 $\alpha=(2^{k-p})^{1/4}$ is chosen because it satisfies the rotatability property if the 2^{k-p} design for the corner points has resolution V or higher. This requirement on resolution is satisfied for the designs considered in the paper. The number of runs at the center point can be greater, which will be further discussed in Section 8.

For the first t runs, $z=1$ or -1 can be chosen according to an interaction column among the x_i's or by searching over different combinations of ±1's. The z values for the $(t+1)$st and the $(t+2)$nd runs are $(1,-1)$ or $(-1,1)$, which ensures that each of $z=1$ and $z=-1$ has one run at the center point. The z value for the last $2k$ runs are chosen by searching over different combinations of ±1's. The search in both cases is directed by objectives A–C.

Table 1
Second-order designs for quantitative factors ($x_1, \dots, x_k$) and one qualitative factor (z)

Run	x_1	x_2	$\cdots$	x_k	z
1					
2	± 1 According to a 2^{k-p} design				See Note 1
$\vdots$					
$t(=2^{k-p})$					
$t+1$	0	0	$\cdots$	0	1 or −1
$t+2$	0	0	$\cdots$	0	−1 or 1
$t+3$	α	0	$\cdots$	0	See Note 2
$t+4$	$-\alpha$	0	$\cdots$	0	
$t+5$	0	α	$\cdots$	0	
$t+6$	0	$-\alpha$	$\cdots$	0	
$\vdots$					
$t+2k+1$	0	$\cdots$		α	
$t+2k+2$	0	$\cdots$		$-\alpha$	

Notes:
1. $z=1$ or -1 by (i) equating column z to an interaction column among the x_i's, or (ii) search over different combinations of ± 1's according to some criteria.
2. See 1(ii).

We use the determinant criterion (D-criterion) for efficiency comparison,

$$|X^{\mathrm{T}}X|^{1/n}, \tag{3.1}$$

where the model is described by $Ey = X\beta$, y is the vector of observations, β is the vector of parameters, n is the dimension of β, i.e., the number of parameters in the model, and X is the *model matrix* consisting of all the important effects stipulated in the objectives and the intercept term.

According to A, the overall design with $t+2k+2$ runs should allow the parameters in models (2.1) or (2.2) to be estimated with high efficiency in terms of (3.1). (Throughout the paper the term 'efficiency' refers to comparison with other designs constructed by the same method as previously described.) When p in the 2^{k-p} design is small, we can usually find a design to accommodate the estimation of parameters in the complete model (2.1). Otherwise, we look for designs that can accommodate the estimation of the largest number of parameters in models (2.2). The same principle is used for models (2.3) and (2.4). According to B, the runs at each level of z should allow the parameters in model (2.3) to be estimated with high D-efficiency. According to C, the first $t+1$ runs, including one center point, should allow the parameters in models (2.4a) or (2.4b) to be estimated with high D-efficiency. If the first $t+1$ runs and the remaining $2k+1$ runs are conducted at different times, the D-criterion for the overall design should accommodate the possible difference between the mean levels of the two sets of runs, that is a 'block' effect. The model is

$$E(y)=[X,u]\begin{bmatrix}\beta\\ b\end{bmatrix}=X\beta+ub, \tag{3.2}$$

where y, X, β are the same as in Eq. (2.1), $u=(u_i)$ is the $(t+2k+2)\times 1$ vector with $u_i=1$ for $i\leqslant t+1$ and -1 for $i>t+1$, and b is the block effect. Then the determinant criterion for β is

$$|X^{\mathrm{T}}X-(X^{\mathrm{T}}u)(u^{\mathrm{T}}u)^{-1}(u^{\mathrm{T}}X)|^{1/n}, \tag{3.3}$$

where $u^{\mathrm{T}}u=t+2k+2$. In the optimal design literature, Eq. (3.3) is called the D_s-criterion (Silvey, 1980). We will use $\tilde{D}_1$ and, respectively, $\tilde{D}_{-1}$ to denote the value of Eq. (3.3) when the $(t+1)$th run (center point) is at $z=1$ and, respectively, $z=-1$.

To recapitulate, we should choose the column for z in Table 1 so that the design has high D-efficiency for the overall model (2.1) or (2.2), for the model (2.3) at $z=1$ and $z=-1$, for the first-order submodel (2.4) based on the first $t+1$ runs, and high D_s-efficiency for the overall model (2.1) or (2.2) plus the block effect. The computation for design search is manageable in most practical situations. An intelligent search over the $t+2k$ values of z is feasible for $t+2k\leqslant 16$. For larger $t+2k$, we define the first t values of z as an interaction column among the x_i's. Then search is done over the remaining $2k$ values of z. From combinatorial considerations, there is usually a small number of interactions that can be used for z. Among them we choose the ones with the best overall values of the D and D_s criteria. A more elaborate search algorithm will be discussed in Section 9.

4. Illustration of construction

We now illustrate the method by constructing designs for x_1, x_2 and z. There are ten runs whose design and model matrix are given in Table 2.

Since the levels for the x_i's are determined by the requirement that they form a rotatable central composite design, it remains to choose the levels of z for the ten runs. Therefore, the choice of designs amounts to the choice of the vector $\mathbf{z}=(z_1,\ldots,z_{10})$, where z_i is the z level of the ith run. According to the requirements stated above, first, the designs should allow the estimation of the parameters in Eq. (2.1), which can be rewritten as

$$E(y)=\beta_0+\beta_1x_1+\beta_2x_2+\beta_{12}x_1x_2+\beta_{11}x_1^2+\beta_{22}x_2^2+\gamma z+\gamma_1x_1z+\gamma_2x_2z$$

Where z takes the values 1 or -1. We denote this model by $(1,x_1,x_2,x_1x_2,x_1^2,x_2^2,z,x_1z,x_2z)$, where 1 is the constant term, and the corresponding value in Eq. (3.1) by D. Second, the design at each level of z should allow the estimation of the parameters in the model $(1,x_1,x_2,x_1x_2)$. The corresponding value of the determinant criterion is denoted by d_{-1} for $z=-1$ and d_1 for $z=1$. Third, the first-order design, consisting of the first five runs, should be able to fit models represented by $(1,x_1,x_2,z\|x_1x_2,x_1z,x_2z)$,

Table 2
Design and model matrix for x_1, x_2 and z

x_1	x_2	x_1x_2	x_1^2	x_2^2	z	x_1z	x_2z
1	1	1	1	1			
1	−1	−1	1	1			
−1	1	−1	1	1			
−1	−1	1	1	1			
0	0	0	0	0	1		
0	0	0	0	0	or		
$\sqrt{2}$	0	0	2	0			
$-\sqrt{2}$	0	0	2	0	−1		
0	$\sqrt{2}$	0	0	2			
0	$-\sqrt{2}$	0	0	2			

where the first four variables to the left of the double bar must be included in the model and one of x_1x_2, x_1z or x_2z is added as the fifth variable. We use $D_1(i)$ (and, respectively, $D_{-1}(i)$) to denote the value of the D-criterion for the first-order design whose center point takes $z=1$ (and, respectively, $z=-1$) and the model consisting of the first four variables and the ith variable to the right of the double bar. We denote the value of the D_s-criterion (3.3) by $\tilde{D}_1$ or $\tilde{D}_{-1}$, depending on the z value of the center point in the first-order design.

Because good designs must satisfy multi-objective criteria, we compare the designs in Table 3 according to their values of $D, d_{-1}, d_1, \tilde{D}_1, \tilde{D}_{-1}, D_1(i)$, and $D_{-1}(i)$. (The meaning of $d_1(x_1^2)$, $d_1(x_2^2)$ and so forth will be explained in Section 8,2.) A high D value is desired because of the primary importance of fitting the overall model to the data. Subject to this, we may choose designs with approximately equal d_1 and d_{-1} values if the information at each level of z is deemed equally important. Or, we may choose those with a large value of d_1 or d_{-1} if information at one level of z is more important for investigation (e.g., a more commonly used machine, or a major supplier). There are ten designs with positive values of D, d_1, and d_{-1}. If sequential design is contemplated, then the first-order design should be efficient for some selected models as reflected by high values of $D_1(i)$ or $D_{-1}(i)$ for some i, and the overall design should be efficient for fitting the model (3.2) with a block effect, which is reflected by high values of $\tilde{D}_1$ or $\tilde{D}_{-1}$. Overall designs 1, 2, 6 and 8 are the best. Note that designs 1 and 6 have four points with $z=1$ and six points with $z=-1$.

In Fig. 1 we represent these ten designs graphically. The geometry of these designs may provide some insight into their values of d_1, d_{-1} and $\tilde{D}$. For example, design 8 is the only one with $d_1 = d_{-1}$. Its designs at $z=1$ and at $z=-1$ are identical after a 90° rotation. In general, designs with equal numbers of points at $z=\pm 1$ have similar d_1 and d_{-1} values. The high values of $\tilde{D}$ in Table 3 can be explained partially by the orthogonality between the block effect and some second-order effects. This orthogonality is apparent from the geometry in some situations. Take, for example, design 6, which has the best $\tilde{D}$ value. From Fig. 2, the estimation of x_1^2 and x_2^2 is based solely on

Table 3
Ten designs with two center runs, $r=1$ and $k=2$

Design no.	$z=(z_1,\ldots,z_{10})$									
1	−1	−1	−1	1	1	−1	1	−1	1	−1
2	−1	−1	−1	1	1	−1	1	−1	1	1
3	−1	−1	−1	1	1	−1	−1	1	1	−1
4	−1	−1	−1	1	1	−1	−1	1	1	1
5	−1	−1	−1	1	1	−1	−1	1	−1	1
6	−1	1	1	1	1	−1	−1	−1	−1	−1
7	−1	−1	1	1	1	−1	−1	−1	−1	1
8	−1	1	1	−1	1	−1	−1	−1	1	1
9	−1	1	1	−1	1	−1	−1	−1	−1	1
10	−1	1	1	−1	1	−1	−1	1	−1	1

Design no.	D	d_1	d_{-1}	$d_1(x_1^2)$	$d_1(x_2^2)$	$d_{-1}(x_1^2)$	$d_{-1}(x_2^2)$	$d_{-1}(x_1^2,x_2^2)$
1	5.7	1.4	4.3	0.0	0.0	3.7	3.7	3.4
2	5.6	2.2	3.4	0.0	2.3	2.9	2.0	0.0
3	4.2	1.4	3.9	0.0	0.0	2.7	3.4	2.5
4	3.9	2.2	2.4	0.0	2.3	1.4	2.0	0.0
5	3.3	1.4	2.9	0.0	0.0	2.5	2.5	1.9
6	5.6	2.0	3.0	0.0	0.0	3.3	3.3	3.2
7	4.2	1.7	3.2	0.0	0.0	3.1	2.5	2.5
8	5.0	2.6	2.6	0.0	2.6	2.6	0.0	0.0
9	4.9	1.7	3.4	0.0	0.0	3.4	2.9	2.8
10	4.9	2.2	2.7	2.0	2.0	2.0	2.0	0.0

Design no.	$\tilde{D}_1$	$\tilde{D}_{-1}$	$D_1(1)$	$D_1(2)$	$D_1(3)$	$D_{-1}(1)$	$D_{-1}(2)$	$D_{-1}(3)$
1	4.6	4.6	3.6	3.0	3.0	2.3	0.0	0.0
2	3.9	4.8	3.6	3.0	3.0	2.3	0.0	0.0
3	3.1	3.1	3.6	3.0	3.0	2.3	0.0	0.0
4	3.9	3.2	3.6	3.0	3.0	2.3	0.0	0.0
5	2.0	2.1	3.6	3.0	3.0	2.3	0.0	0.0
6	4.2	5.4	2.3	0.0	0.0	3.6	3.0	3.0
7	3.1	3.1	3.0	0.0	3.0	3.0	0.0	3.0
8	4.9	4.9	3.0	0.0	0.0	3.0	0.0	0.0
9	3.6	4.6	3.0	0.0	0.0	3.0	0.0	0.0
10	3.6	4.6	3.0	0.0	0.0	3.0	0.0	0.0

the observations at $z=-1$ of the supplementary design. It is therefore orthogonal to the block effect since the latter measures the overall difference between the first-order design and the supplementary design.

5. Small designs for one qualitative factor

The same construction method can be used to obtain designs for $k\geqslant 3$. In this section we list some efficient designs in Table 4 for $x_1,\ldots,x_k$ with $k=3,4$ and 5 and one

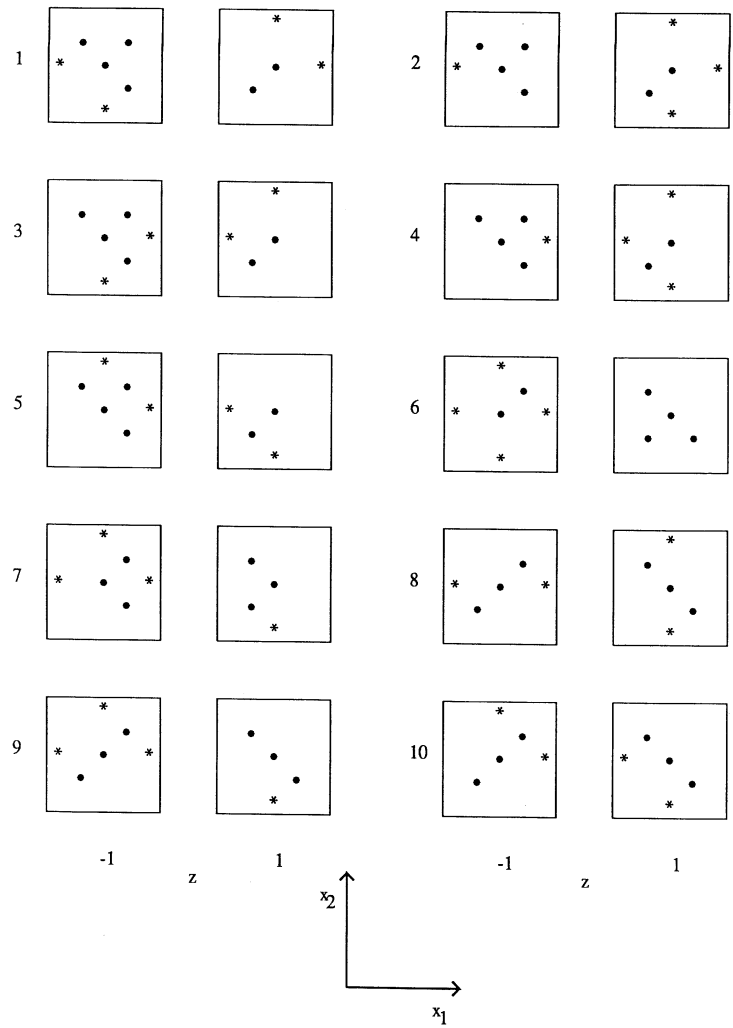

Fig. 1. Ten designs for x_1, x_2 and z with two center points. The four star points are denoted by $*$.

qualitative factor z. A more complete collection up to $k=7$ and discussion of their properties can be found in Wu and Ding (1991). Since z has two levels ± 1, models (2.1) and (2.2a) can be represented by the vector

$$(1, x_i, x_i^2, x_i x_j, z, x_i z), \tag{5.1}$$

and models (2.3) and (2.4a) can be represented, respectively, by the vectors

$$(1, x_i, x_i x_j), \tag{5.2}$$

and

$$(1, x_i, x_i x_j, z, x_i z), \tag{5.3}$$

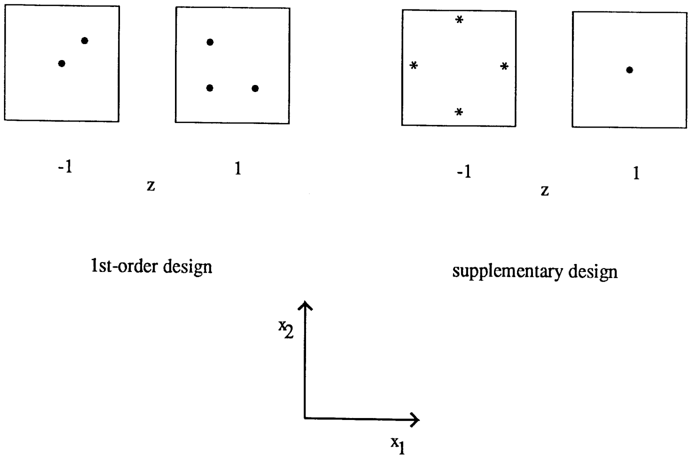

Fig. 2. Decomposition of design 6 in Fig. 1.

Table 4
Constructed designs for $x_1,\ldots,x_k$ with $k=3,4,5$ and one z

1. $k=3$[a]
   (1) 2^3 design for x_1, x_2 and x_3, $N=16$, $\alpha=1.682$.
   (2) Overall model: Eq. (5.1) with all possible i and j.
   (3) Model for $z=1$ and for $z=-1$: Eq. (5.2) with all possible i and j.
   (4) First-order model: $(1,x_1,x_2,x_3,z \| x_1x_2,x_1x_3,x_2x_3,x_1z,x_2z,x_3z)$.

Design no.	$z=(z_1,\ldots,z_{16})$																D	d_1	d_{-1}	$\tilde{D}_1$	$\tilde{D}_{-1}$
1	1	1	−1	1	1	−1	−1	−1	1	−1	−1	1	−1	1	−1	−1	9.5	2.4	3.8	8.0	8.0
2	1	1	−1	1	1	−1	−1	−1	1	−1	−1	−1	−1	1	1	1	9.2	3.1	3.1	8.4	7.9
3	1	1	1	−1	−1	−1	−1	−1	1	−1	−1	1	−1	1	1	1	8.7	3.0	4.3	7.4	7.4

2. $k=4$
   (1) 2^4 design for x_1, x_2, x_3 and x_4, $z=x_1x_2x_3$ for the first 16 runs, $N=26$, $\alpha=2$.
   (2) Overall model: Eq. (5.1) with all possible i and j.
   (3) Model for $z=1$ and for $z=-1$: $(1,x_1,x_2,x_3,x_4,x_1x_4,x_2x_4,x_3x_4)$.
   (4) First-order model: Eq. (5.3) with $(i,j)=(1,4),(2,4),(3,4)$.

Design no.	$z=(z_{19},\ldots,z_{26})$								D	d_1	d_{-1}
1	−1	−1	−1	−1	−1	−1	−1	−1	16.6	8.1	12.4
2	−1	−1	−1	−1	−1	−1	1	1	16.5	9.1	11.2
3	−1	−1	−1	−1	1	1	1	1	16.4	10.1	10.1
4	−1	−1	−1	−1	−1	−1	−1	1	16.4	8.6	11.9

3. $k=5$
   (1) 2^5 design for $x_1,\ldots,x_5$, $z=x_1x_2x_3x_4x_5$ for the first 32 runs, $N=44$, $\alpha=2.378$.
   (2) Overall model: Eq. (5.1) with all possible i and j.
   (3) Model for $z=1$ and for $z=-1$: Eq. (5.2) with all possible i and j.
   (4) First-order model: Eq. (5.3) with all possible i and j.

Table 4 (continued)

Design no.	$z=(z_{35},\ldots,z_{44})$										D	d_1	d_{-1}
1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	38.4	16.1	19.5
2	−1	−1	−1	−1	−1	−1	−1	−1	1	1	38.2	16.7	18.8
3	−1	−1	−1	−1	−1	−1	1	1	1	1	38.2	17.4	18.1
4	−1	−1	−1	−1	−1	1	1	1	1	1	37.9	17.8	17.8

4. $k=5$

(1) 2^{5-1} design with $x_5=x_2x_3x_4$ for $x_1,\ldots,x_5$, $z=x_1x_2x_3$ for the first 16 runs, $N=28$, $\alpha=2$.

(2) Overall model: Eq. (5.1) with $(i,j)=(1,2),(1,3),(1,4),(1,5),(2,3),(2,4)$ and $(2,5)$.

(3) Model for $z=1$ and for $z=-1$: $(1,x_i,x_2x_4,x_2x_5)$, $i=1,\ldots,5$.

(4) First-order model: Eq. (5.3) with $(i,j)=(2,4),(2,5)$.

Design no.	$z=(z_{19},\ldots,z_{28})$										D	d_1	d_{-1}
1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	15.8	8.1	13.8
2	−1	−1	−1	−1	−1	−1	−1	−1	1	1	15.7	9.1	12.4
3	−1	−1	−1	−1	−1	−1	1	1	1	1	15.6	10.1	11.2
4	−1	−1	−1	−1	−1	1	1	1	1	1	14.6	10.7	10.7

[a] Let the six variables to the right of the double bar in (4) of $k=3$ be denoted by 1, 2, 3, 4, 5, and 6, and $D_1(\boldsymbol{i})$ be the D value for the first-order design including the center point at $z=1$ and the model consisting of $1,x_1,x_2,x_3,z$ and four additional variables in the set $\boldsymbol{i}$, for example, x_1x_2,x_1x_3,x_2x_3 and x_1z for $\boldsymbol{i}=(1,2,3,4)$. There are 12 sets of $\boldsymbol{i}$ with positive values of $D_1(\boldsymbol{i})$ and $D_{-1}(\boldsymbol{i})$ for designs 1 and 2, and eight sets for design 3. For the same case Draper and John (1988) gave three designs in their Fig. 8. Using our definitions of D,d_1 and d_{-1}, their design 8a has $D=0,d_1=d_{-1}=3.56$; design 8b has $D=7.90$, $d_1=d_{-1}=0$; design 8c has $D=11.76$, $d_1=0$ and $d_{-1}=8.13$. Their designs 8a and 8b are inferior to the three designs given above. Design 8c has larger values of D and d_{-1} at the expense of having $d_1=0$. This can be seen from the distribution of points between $z=1$ and -1. All the star points and one center point are assigned to $z=1$. For the design at $z=1$, the columns of x_1x_2,x_1x_3 and x_2x_3 are zero. Therefore, the coefficients of these three variables are not estimable.

where $i=1,\ldots,k$ and $i<j$ take all possible pairs or are chosen from selected pairs as specified in each case. A similar representation was used in Section 4.

In each case, we give the information on the design and models in four parts.

(1) A 2^{k-p} design for $x_1,\ldots,x_k$ with its defining relations, or a 2^k full factorial design. (This is for the upper left part of Table 1.) If the first 2^{k-p} components of the z column in Table 1 are defined by an interaction among the x_i's, say $x_1x_2x_3$, we write $z=x_1x_2x_3$. Otherwise, this part is omitted. The total run size is $N=2^{k-p}+2k+1$. There is no general rule for choosing an interaction among the x_i's for z. In Table 4 we used a numerical search over different interactions to find one with the best overall performance.

(2) Overall model for the whole design.

(3) Model for $z=1$ and for $z=-1$.

(4) Model for the first-order design with optional terms to the right of the double bar.

The information on each design is completed by the values of z. If the first $t(=2^{k-p})$ components of z are defined by an interaction among the x_i's (see (1) above), we only give values of $\boldsymbol{z}=(z_{t+3},\ldots,z_{t+2k+2})$ for the star points. Otherwise, we give $\boldsymbol{z}=(z_1,\ldots,z_{t+2k+2})$. In each case only designs with high values of D,d_1,d_{-1}, are given. Values of the D_s-criterion (3.3) are only given for $k=2$ and 3. For $k \geqslant 4$, the rankings of designs according to $\tilde{D}$ are essentially the same as D. The tabulated designs are also efficient when the number of runs at the center point (with $z=1$ or $z=-1$) is greater than one. Further discussion is given in Section 8,3.

6. Nine-point designs with one center run

Draper and John (1988) gave a nine-run design for $r=1$ and $k=2$ in their Fig. 5, which consists of one center point, four corner points and four star points. By

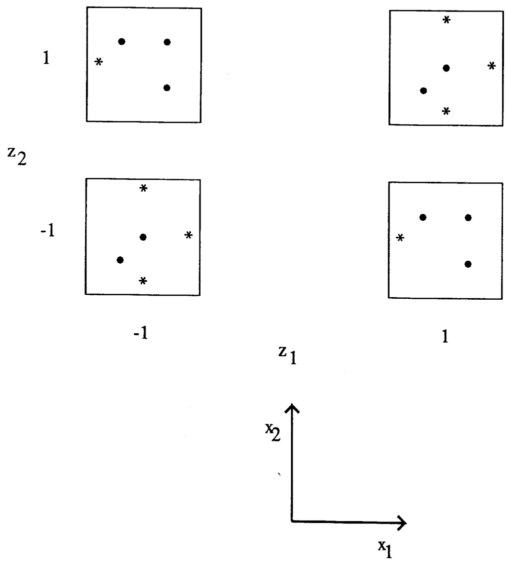

Fig. 3. A design for x_1, x_2, z_1 and z_2. Its top part ($z_2 = 1$) is the same as design 1 in Table 5. Its bottom part ($z_2 = -1$) is obtained from changing the sign of z_1 of the top part.

Table 5
Eleven designs with one center run, $r=1$, $k=2$

Design no.	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	D	d_1	d_{-1}	$D_1(1)$	$D_1(2)$	$D_1(3)$
1	−1	−1	−1	1	1	1	−1	1	1	4.9	2.2	3.1	3.6	3.0	3.0
2	−1	−1	−1	1	1	−1	1	1	−1	3.4	1.4	3.6	3.6	3.0	3.0
3	−1	−1	−1	1	1	−1	1	1	1	3.3	2.2	1.3	3.6	3.0	3.0
4	−1	−1	−1	1	1	1	−1	1	−1	3.3	2.2	1.3	3.6	3.0	3.0
5	−1	−1	−1	1	1	−1	1	−1	1	2.3	1.4	1.7	3.6	3.0	3.0
6	−1	−1	1	1	1	−1	−1	−1	1	3.4	1.7	3.0	3.0	0.0	3.0
7	−1	1	1	1	1	−1	−1	1	−1	4.9	3.4	2.0	2.3	0.0	0.0
8	−1	1	1	1	1	−1	−1	−1	−1	4.7	2.0	2.8	2.3	0.0	0.0
9	−1	1	1	−1	1	−1	1	−1	1	4.0	2.2	2.4	3.0	0.0	0.0
10	−1	1	1	−1	1	−1	−1	−1	1	4.0	1.7	3.1	3.0	0.0	0.0
11	−1	1	1	1	1	−1	−1	−1	1	3.3	2.4	2.0	2.3	0.0	0.0

using the search procedure in Section 4, we are able to find all possible nine-run designs with desirable properties in terms of D, $D_1(i)$, d_1, and d_{-1} whose definitions and models are the same as in Section 4. Without loss of generality, we can assume that the center run is assigned to $z=1$. There are eleven such designs in Table 5 with positive values of D, d_1 and d_{-1}. The x_i values of these designs are the same as before (see Table 2). The values of z_1 to z_4 are the z values for the corner points, z_5 for the center point, and z_6 to z_9 for the star points. From their values of D, d_1, d_{-1}, $D_1(1), D_1(2)$ and $D_1(3)$, design 1 is the best overall choice. If $D_1(i)$ are less important, design 7 is comparable to design 1. Both designs are quite unbalanced in the distribution of the star points over $z=\pm 1$. Note that the design given in Draper and John is equivalent to design 4 which has much smaller values of D and d_{-1}.

7. Designs for $r=2$ and $k=2$

When there are two or more qualitative factors, construction of designs becomes more complicated because Objective B in Section 2 can take several forms. We only consider the simplest case of $r=2$ and $k=2$. The following are some reasonable criteria for selecting designs. Since Objective C in Section 2 is essentially the same as in $r=1$, we will not discuss it here for brevity.

(1) The model $(1, x_i, x_i^2, x_1x_2, z, x_iz)$, where $i=1,2$ and $z=z_1, z_2, z_1z_2$, can be fitted with high D-efficiency from the overall design.

(2) For each level of z_1, the model $(1, x_i, x_i^2, x_1x_2, z_2, x_iz_2)$, $i=1,2$, can be fitted with high D-efficiency. For each level of z_2, the model $(1, x_i, x_i^2, x_1x_2, z_1, x_iz_1)$, $i=1,2$, can be fitted with high D-efficiency.

(3) For each of the four combinations of $z_1=\pm 1$ and $z_2=\pm 1$, the model $(1, x_1, x_2, x_1x_2)$ can be fitted with high D-efficiency.

For each level of z_1 and of z_2, the criteria (2) and (3) are exactly the same as for $r=1$ and $k=2$ (see Eqs. (5.1) and (5.2)). Therefore, we can use any of the eleven designs in Section 6 for, say, $z_2=1$. Once a design is chosen for $z_2=1$, the design at $z_2=-1$ is obtained by changing the sign of z_1 of the design at $z_2=1$. This is done so that the design at each of $z_2=\pm 1$ and of $z_1=\pm 1$ consists of four corner points, four star points and one center point. An example is given in Fig. 3. Altogether there are eleven such designs based on those in Table 5. For each design the D values for the criteria in (2) and (3) are the same as those in Table 5. For the overall design we consider the D-efficiency for three models. The first is the model given in (1) with 15 terms. The second and third models are obtained by adding $(x_1^2z_1, x_2^2z_1, x_1x_2z_1)$ and, respectively $(x_1^2z_2, x_2^2z_2, x_1x_2z_2)$ to the first model. The D values (3.1) for these three models are denoted by $D(1)$, $D(2)$ and $D(3)$. In terms of the $D(i)$, the best designs are 4, 1, 7 and then 8 and 9 (whose values are given in Table 6). When criteria (2) and (3) are also considered, we refer to the values of D, d_1 and d_{-1} in Table 5 for comparison purpose. On balance, we would recommend designs 1, 7, 8, 4 and 9 in descending order.

If four runs at the center point with one at each of $z_1 = \pm 1$ and $z_2 = \pm 1$ are required, we can use any of the 10 designs in Fig. 1 for $z_2 = 1$ and the rest of the construction is the same.

If the run size is much smaller than 18, we must be content with less ambitious objectives such as the following:

(1) The model $(1, z_1, z_2, z_1 z_2, x_1, x_2, x_1 x_2, x_1^2, x_2^2)$ can be fitted with high D-efficiency from the overall design.
(2) If the effect of z_1 (and resp. of z_2) is not significant, the model $(1, z_2, x_1, x_2, x_1 x_2, x_1^2, x_2^2, x_1 z_2, x_2 z_2)$ (and resp. the model $(1, z_1, x_1, x_2, x_1 x_2, x_1^2, x_2^2, x_1 z_1, x_2 z_1)$) can be fitted with high D-efficiency from the overall design.
(3) For each level of z_1 (and respectively of z_2), the model $(1, x_1, x_2, x_1 x_2)$ can be fitted with high D-efficiency.

Table 6
D values of five designs with $r = 2$ and $k = 2$. (The design number corresponds to the design number in Table 5.)

Design no.	$D(1)$	$D(2)$	$D(3)$
1	11.7	9.7	9.7
4	12.1	10.2	10.2
7	11.7	9.7	9.7
8	11.3	9.3	9.3
9	9.9	8.0	8.0

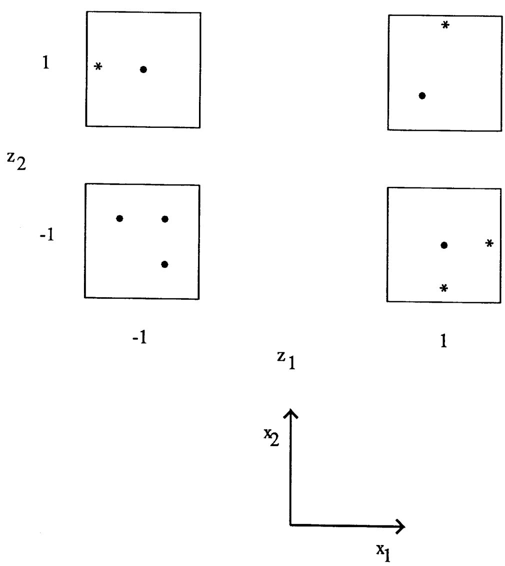

Fig. 4. When collapsed over $z_1 = \pm 1$ (and resp. $z_2 = \pm 1$), it becomes design 3 (and resp. design 2) in Fig. 1.

When the design is collapsed over the levels of z_1 (and, respectively, z_2), we call the resulting design a *marginal* design in z_2 (and, respectively, z_1), which has one qualitative factor and two quantitative factors. It is clear that, for the marginal design in z_1 or in z_2, the models in (2) and (3) are identical to the corresponding ones in Section 4 for $r=1$ and $k=2$. If the overall design consists of four corner points, four star points and two runs at the center point, then the two marginal designs must be chosen from the 10 designs given in Table 3 or Fig. 1. Once the marginal designs are specified, the allocation of the corner and star points to z_1 and z_2 is uniquely determined. There are two choices of the two center points: $(z_1,z_2)=\{(-1,-1),(1,1)\}$ or $\{(-1,1),(1,-1)\}$. One such example is given in Fig. 4. The design in Fig. 4 has the largest D value 6.16 for the model in (1) among all possible combinations of the marginal designs. In general, we can compare designs by the three D-efficiency values for (1)–(3). To save space, the details are omitted.

8. Some extensions

8.1. Run size reduction through smaller plans for the corner points

Sometimes we can achieve dramatic reduction of run size by this means without greatly sacrificing the capacity for effect estimation. Take, for example, the case of $r=1$ and $k=4$. The designs in Table 4.2 have 26 points. If we choose instead a 2^{4-1} design with $x_4=x_1x_2$ and define $z=x_2x_3$ for the first 8 runs, the total run size is 18. By further choosing $(z_{11},\ldots,z_{18})=(-1,1,-1,1,-1,1,-1,1)$ for the eight star points, the resulting design can estimate the 18 terms $(1,x_i,x_i^2,x_1x_2,x_1x_4,x_2x_3,x_2x_4,z,zx_i)$, $i=1,2,3,4$, with $D=5.7$. Obviously, it is not as efficient per observation as the designs in Table 4.2. It can estimate $(1,x_1,x_2,x_3,x_4,x_1x_2,x_1x_4,x_2x_3,x_2x_4)$ for $z=1$ and for $z=-1$, with $d_1=d_{-1}=2.5$. If the first-order model is $(1,x_1,x_2,x_3,x_4,x_1x_3,z,x_1z)$, the design has $D_1=D_{-1}=8.2$.

To reduce the number of corner points, we can employ the Plackett–Burman designs for the corner points. For example, designs represented in Table 1 with $r=1$, $k=6$, $x_6=x_1x_2x_3x_4x_5$ have 46 runs including 32 corner points and can entertain an overall model with 35 terms (see Wu and Ding, 1991 (Table 3.6)). By taking a Plackett–Burman design with 28, 24 or 20 runs for the corner points, the total run size is reduced by 4, 8, or 12. Note that the Plackett–Burman designs have been employed to reduce the size of central composite designs (see Draper and Lin, 1990, and its references).

8.2. Additional variables to be entertained

For most designs discussed so far, the models we have considered are not saturated. There is often the possibility of entertaining one or more variables. Here we only discuss two cases. Finding such variables for general designs is computationally straightforward. An illuminating example is the designs for $r=1$ and $k=2$. From Fig. 1, it is clear that for some designs x_1^2 or x_2^2 can be entertained at $z=1$ or $z=-1$. Let $d_{-1}(*)$ and $d_1(*)$ be, respectively, the D value (3.1) for the design at $z=-1$ and at $z=1$ and the model consisting of $1, x_1, x_2, x_1x_2$ and the additional variables indicated by $*$. For $z=-1$, $*$ can be x_1^2, x_2^2 or $\{x_1^2, x_2^2\}$. For $z=1, *$ cannot take both x_1^2 and x_2^2. The values are given in the second part of Table 3. The best two are designs 1 and 6 at $z=-1$. Design 6 having all the four star points at $z=-1$ clearly has a high value of $d_{-1}(x_1^2, x_2^2)$ for entertaining both x_1^2 and x_2^2 because the quadratic curvature of x_1 and of x_2 can be efficiently estimated by the two star points, one corner point and the center point. Another example concerns the 16-run designs for $r=1$ and $k=3$. The overall model (5.1) has 14 terms, leaving two degrees of freedom for additional terms. Consider the possibility of fitting one or two of the six cubic terms, $x_1x_2z, x_1x_3z, x_2x_3z, x_1^2z, x_2^2z$ and x_3^2z, which we denote by 1, 2, 3, 4, 5 and 6. Draper and John (1988) stated that their design 8b allows the pairs $(2,4)$, $(2,5)$ and $(2,6)$ to be fitted. In fact, $(4,5)$ and $(5,6)$ can also be fitted, but not others. In contrast, design 2 in Table 4.1 allows any of the $\binom{6}{2}=15$ pairs of cubic terms to be fitted. Design 1 allows 13 out of the 15 pairs to be fitted with the exception of $(1,6)$ and $(2,3)$. Design 3 still allows nine pairs to be fitted. Overall, our designs 1 and 2 are better.

8.3. Number of runs at the center point

The designs given in Sections 4 and 5 have one center run at each of $z=1$ and $z=-1$. Are they still good designs when the number of center runs is greater than one? First, it is easy to show that the nonsingularity of the D and D_s criteria in Section 3 is not affected by the number of center runs as long as it is at least one, i.e., the positiveness of the D, d or $\tilde{D}$ values in Sections 4 and 5 is not affected. Second, based on a complete search, these designs remain efficient when the number of center runs is increased to two or three. In practice, a smaller number of center runs is preferred if it is expensive to conduct the experiment. On the other hand, a larger number is required for testing lack of fit or achieving stable prediction variance. For details see the three texts cited in Section 1.

9. Comparison with algorithmic constructions

Our proposed approach is to fix a marginal design for the quantitative factors x and then search over the qualitative factor(s) z. Its main advantage is that we can choose a marginal design for x with known properties. In the construction we choose the central composite design as the marginal design because it is widely used and also possesses several desirable properties (details in Khuri and Cornell, 1987). Once the marginal design is chosen, computer search over z is relative easy. For larger problems, however, we may need to speed up the computations by using intelligent search over z. Two such techniques are available. First we can interchange any pair of $z=1$ and $z=-1$ in the z column. Then we select the pair that gives the best improvement in terms of the values of D, d_1, and d_{-1} and possibly other criteria. Because we must deal with several criteria, there are several ways to judge improvement. We may convert several criteria into one by using a weighted average of D, d_1, and d_{-1}, with more weight assigned to D. Alternatively we can use D as the primary criterion subject to the requirement that d_1, and d_{-1} cannot be decreased by a certain amount. Since interchange does not affect the run size at each level of the qualitative factor, we can repeat this procedure by starting with another design with different run sizes. The second technique is to switch $z=1$ to $z=-1$ or vice versa. Then we select the one switch that gives the best improvement in the sense as discussed above. This technique can change the run sizes, say from $(5,5)$ in the case of Table 3 to $(4,6)$ or $(6,4)$. We can start from a good design with $(5,5)$ and move to a better design with $(4,6)$ or $(6,4)$. Both techniques can be easily implemented. The *interchange* technique was used by Harville (1974), Wu (1981) and Cook and Nachtsheim (1989) among others. The *switch* technique was used, for example, by Wu (1981) for treatment allocation with covariates.

Another approach is to search over x and z simultaneously. Such an approach will be necessary if the design region for x has an irregular shape, e.g., when there are constraints on x. Problems with this approach may arise because it is often driven by a single criterion, say, the D criterion (3.1) for the most comprehensive model. Although the resulting design will have high efficiency for a particular criterion, it may not have other desirable properties. For example, the symmetry of central composite designs can be quite appealing to experimenters. The spacing of design points is fairly uniform over the design region, which can aid the search for optimum. These and other prescribed properties can be ensured by choosing a marginal design for x in advance. On the other hand, there is no guarantee that designs obtained through an optimal design algorithm possess such properties. A related problem with this approach concerns the multi-purpose nature of designs. In the present set-up there are several criteria to be satisfied. Designs that are optimal with respect to one criterion may be poor with respect to other criteria. To illustrate this point, we used, in the case of $r=1$ and $k=2$, the OPTEX procedure in SAS to search for D-optimal designs according to the D criterion in Eq. (3.1). We used the 18 points in Table 2 for x_1, x_2 and z as the candidate points for search. All the 100 searches (with different seeds) lead to a design with $D=7.029$, $d_1=4.229$ and $d_{-1}=0$. Its D and d_1 values are much higher than the best in Table 3 but it is obviously unacceptable because $d_{-1}=0$ means that the design at $z=-1$ is not estimable. Geometrically this design has all its corner points and one center point at $z=1$ and all its star points and one center point at $z=-1$. Since $x_1x_2=0$ at the star points, the design at $z=-1$ does not allow the x_1x_2 term to be estimated. This example clearly demonstrates the limitations of algorithms that are driven by a single criterion. Because computer-aided construction has the advantage of speed and can easily be adapted for different models and design regions, this promising direction of research needs to be further explored.

Acknowledgements

We would like to thank D.X. Sun for his meticulous assistance in computing, the associate editor and referees for helpful comments. The research was supported by the Natural Sciences and Engineering Research Council of Canada and the National Science Foundation DMS 9704649.

References

Box, G.E.P., Draper, N.R., 1987. Empirical Model-Building and Response surfaces. Wiley, New York.

Cox, D.R., 1984. Present position and potential developments: some personal views: design of experiments and regression. J. Roy. Statist. Soc. Ser. A. 147, 306–315.

Cook, R.D., Nachtsheim, C.J., 1989. Computer-aided blocking of factorial and response-surface designs. Technometrics 31, 339–346.

Draper, N.R., John, J.A., 1988. Response-surface designs for quantitative and qualitative variables. Technometrics 30, 423–428.

Draper, N.R., Lin, D.K.J., 1990. Small response-surface designs. Technometrics 32, 187–194.

Harville, D., 1974. Nearly optimal allocation of experimental units using observed covariate values. Technometrics 16, 589–599.

Khuri, A.I., Cornell, J.A., 1987. Response surface. Marcel Dekker, New York.

Myers, R.H., Montgomery, D.C., 1995. Response Surface Methodology: Process and Product Optimization Using Designed Experiments. Wiley, New York.

Silvey, S.D., 1980. Optimal Design. Chapman & Hall, London.

Wu, C.F., 1981. Iterative construction of nearly balanced assignment I: categorical covariates. Technometrics 23, 37–44.

Wu, C.F.J., Ding, Y., 1991. Construction of response surface designs for qualitative and quantitative factors. IIQP Research Report 91-03, University of Waterloo, Waterloo, Canada.