

## ***D*-Optimal Designs for Second-Order Response Surface Models with Qualitative Factors**

Chuan-Pin Lee and Mong-Na Lo Huang  
*National Sun Yat-sen University*

*Abstract:* Central composite design (CCD) is widely applied in many fields to construct a second-order response surface model with quantitative factors to help to increase the precision of the estimated model. When an experiment also includes qualitative factors, the effects between the quantitative and qualitative factors should be taken into consideration. In the present paper, *D*-optimal designs are investigated for models where the qualitative factors interact with, respectively, the linear effects, or the linear effects and 2-factor interactions or quadratic effects of the quantitative factors. It is shown that, at each qualitative level, the corresponding *D*-optimal design also consists of three portions as CCD, i.e. the cube design, the axial design and center points, but with different weights. An example about a chemical study is used to demonstrate how the *D*-optimal design obtained here may help to design an experiment with both quantitative and qualitative factors more efficiently.

*Key words:* Central composite design, approximate design, dispersion function, equivalence theorem, flue gas desulfurization.

### **1. Introduction**

The second-order response surface model is applied widely in many research fields through the central composite design (CCD). A chemical study of flue gas desulfurization published by Zainudin et al. (2005), for illustration, deals with an adsorbent for removing the sulfur dioxide (SO<sub>2</sub>) in the flue gas. The experiment including a total of 40 runs with three quantitative factors and one 2-level qualitative factor was performed to fit a second-order response surface model. At each qualitative level, the settings of the quantitative factors form a 20-run CCD including a 2<sup>3</sup> full factorial design, six axial points and six center points for three quantitative factors. Detailed description of the experiment including the factors, design settings and data are given in the Appendix. Their purpose is to find out the optimal treatment combinations yielding the maximal response

according to the estimated model. However, if the qualitative factor is treated as quantitative as a coded variable  $x_{k+1}$  to fit the model

$$E [y (\mathbf{x}^T, x_{k+1})] = \alpha_0 + \sum_{i=1}^{k+1} \alpha_i x_i + \sum_{i < j}^{k+1} \alpha_{ij} x_i x_j + \sum_{i=1}^k \alpha_{ii} x_i^2, \quad (1)$$

some estimatable effects are naturally absent in model (1), such as the two-factor interactions of the quantitative factors interacting with the qualitative factor, even if those are statistically significant.

A general model of the mean response at the  $j$ -th level of a  $J$ -level qualitative factor with  $k$  quantitative factors described as follows is more adequate to be considered,

$$E [y (j, \mathbf{x}^T)] = f_1^T (\mathbf{x}) \boldsymbol{\beta}_j + f_2^T (\mathbf{x}) \boldsymbol{\gamma}, \quad \mathbf{x} \in \mathcal{X}, \quad (2)$$

where  $f_1 (\mathbf{x})$  and  $f_2 (\mathbf{x})$  are  $p_1$ - and  $p_2$ -dimension vectors containing the quantitative effects, respectively, such that  $f_1 (\mathbf{x})$  is the part of the regression functions having interaction with the qualitative factor but  $f_2 (\mathbf{x})$  is the part invariant at each qualitative level,  $\{\boldsymbol{\beta}_j, j = 1, \dots, J\}$  and  $\boldsymbol{\gamma}$  are the corresponding vectors of unknown parameters, and  $\mathcal{X} \subseteq \mathbf{R}^k$  is the design region of the quantitative factors. The responses are assumed to be independent of each other as well as with constant variance. If the interactions between qualitative and quantitative factors are assumed to be negligible, one can express the quantitative effects under second-order regression model as

$$E [y (j, \mathbf{x}^T)] = \beta_j + [f_L^T (\mathbf{x}), f_I^T (\mathbf{x}), f_Q^T (\mathbf{x})] \boldsymbol{\gamma}^{(S)}, \quad (3)$$

where

$$f_L (\mathbf{x}) = (x_1, \dots, x_k)^T, \\ f_I (\mathbf{x}) = (x_1 x_2, x_1 x_3, \dots, x_{k-1} x_k)^T,$$

and

$$f_Q (\mathbf{x}) = (x_1^2, \dots, x_k^2)^T,$$

which are associated with the linear effects, 2-factor interactions and quadratic effects of the quantitative factors. Model (1) with the linear effects of the quantitative factors interacting with the qualitative factor  $x_{k+1}$  can be rewritten as

$$E [y (j, \mathbf{x}^T)] = [1, f_L^T (\mathbf{x})] \boldsymbol{\beta}_j^{(L)} + [f_I^T (\mathbf{x}), f_Q^T (\mathbf{x})] \boldsymbol{\gamma}^{(IQ)}. \quad (4a)$$

From the analysis of variance table shown in the Appendix, it presents that not only the linear effects but also the interactions between the quantitative factors

are influenced by the qualitative factor. Therefore, an appropriate model for the example of SO<sub>2</sub> removal is described by

$$E[y(j, \mathbf{x}^T)] = [1, f_L^T(\mathbf{x}), f_I^T(\mathbf{x})] \boldsymbol{\beta}_j^{(LI)} + f_Q^T(\mathbf{x}) \boldsymbol{\gamma}^{(Q)}. \quad (4b)$$

When the quadratic effects of the quantitative factors are of interest at each qualitative level, one can consider the following model,

$$E[y(j, \mathbf{x}^T)] = [1, f_L^T(\mathbf{x}), f_Q^T(\mathbf{x})] \boldsymbol{\beta}_j^{(LQ)} + f_I^T(\mathbf{x}) \boldsymbol{\gamma}^{(I)}. \quad (4c)$$

The saturated model where  $f_1$  includes all effects such as constant, linear, interactions and quadratic effects and  $f_2$  being zero can be taken into account as well and expressed as

$$E[y(j, \mathbf{x}^T)] = [1, f_L^T(\mathbf{x}), f_I^T(\mathbf{x}), f_Q^T(\mathbf{x})] \boldsymbol{\beta}_j^{(S)}.$$

However, there is no difference between fitting such a model and fitting the second-order response surface model for quantitative factors at individual level of the qualitative factor. In this work, design problems for estimation of the unknown parameters will be considered where it is assumed to have one qualitative factor with  $J$  levels, which is corresponding to the total number of level combinations if there are more than one qualitative factor.

Qualitative and quantitative factors considered simultaneously in one experiment perplex the response surface designs if some quantitative effects of interest are varied at the qualitative levels. An approximate design is a probability measure on the design region  $\Omega$  having finite support points with positive weights. More specifically, an  $N$ -run exact design should maintain that each weight times  $N$  to be an integer. The quality of a design  $\tau$  associated with the assumed model that  $E[y(\mathbf{z})] = h^T(\mathbf{z}) \boldsymbol{\theta}$  can be measured by the corresponding information matrix

$$\mathbf{M}_h(\tau) := \int_{\Omega} h(\mathbf{z}) h^T(\mathbf{z}) \tau(d\mathbf{z}).$$

A design is defined to be  $D$ -optimal if it maximizes the determinant of the information matrix. When there is no interaction between qualitative and quantitative factors,  $D$ -optimal design can be constructed as a product of those designs which are  $D$ -optimal in the corresponding single-factor models, see Schwabe and Wierich (1995). This product design is also  $D$ -optimal for the model in which all quantitative effects interact with the qualitative factor.

Draper and John (1988) presented the original idea to construct exact designs with certain objectives. A systematic approach of design constructions was proposed by Wu and Ding (1998) with objectives such as (i) the design can be partitioned into two parts, one for fitting the main effects of all factors and the

other for estimating the interactions and quadratic effects; (ii) the design collapsed over the qualitative levels forms an efficient design for the second-order response surface design for quantitative factors. The resulting designs in both articles are based on the support points of CCD. Atkinson and Donev (1989) proposed the BLKL-exchange algorithm to search for exact optimal designs with specified block sizes according to the  $D$ -criterion. Theoretical verifications of exact  $D$ -optimal designs as well as the construction method are provided by Huang et al. (2008) for model (3) and the settings of the quantitative factors are selected within the unit circle. In the present work, approximate  $D$ -optimal designs for models (4a), (4b) and (4c) are investigated.

In Section 2, some properties are derived about the relationships of information matrices between model (2) and the model with only quantitative factors. For models (4a), (4b) and (4c), the  $D$ -optimal designs constructed through CCD are shown in Section 3. Section 4 concludes the paper, and the proofs of  $D$ -optimality are left in Appendix.

## 2. Preliminaries

Let the general model (2) be rewritten as

$$\begin{aligned} E[y(j, \mathbf{x}^T)] &= [\mathbf{e}_j^T \otimes f_1^T(\mathbf{x}), f_2^T(\mathbf{x})] (\boldsymbol{\beta}_1^T, \dots, \boldsymbol{\beta}_J^T, \boldsymbol{\gamma}^T)^T \\ &= g^T(j, \mathbf{x}) \boldsymbol{\theta}, \end{aligned}$$

where  $\mathbf{e}_j \in \mathbf{R}^J$  is the unit vector whose  $j$ -th component is equal to 1 and all others are zero, and  $\otimes$  is used to denote the Kronecker product of two matrices. Let  $\mathcal{X}_J = \{1, \dots, J\}$  be the index set of the qualitative levels and  $\Omega$  be the product set  $\mathcal{X}_J \times \mathcal{X}$ . An arbitrary design  $\tau$  on  $\Omega$  can be expressed as

$$\tau(j, \mathbf{x}^T) = \eta(j) \xi_j(\mathbf{x}),$$

where  $\eta$  and  $\xi_j$  are the marginal and the conditional designs on  $\mathcal{X}_J$  and  $\mathcal{X}$ , respectively. Then the information matrix of  $\tau$  is presented as

$$\mathbf{M}_g(\tau) = \int_{\Omega} g(j, \mathbf{x}^T) g^T(j, \mathbf{x}^T) d\tau(j, \mathbf{x}^T) = \sum_{j=1}^J \eta(j) \mathbf{M}_g^{(j)}, \quad (5)$$

where, for  $j = 1, \dots, J$ ,

$$\mathbf{M}_g^{(j)} = \begin{bmatrix} \mathbf{e}_j \mathbf{e}_j^T \otimes \mathbf{M}_{11}(\xi_j) & \mathbf{e}_j \otimes \mathbf{M}_{12}(\xi_j) \\ \mathbf{e}_j^T \otimes \mathbf{M}_{21}(\xi_j) & \mathbf{M}_{22}(\xi_j) \end{bmatrix},$$

and

$$\mathbf{M}_{ab}(\xi_j) := \int_{\mathcal{X}} f_a(\mathbf{x}) f_b^T(\mathbf{x}) d\xi_j(\mathbf{x}), \quad a, b \in \{1, 2\}.$$

If  $\tau$  is considered as a product design and denoted by  $\eta \times \xi$ , which indicates that  $\xi_j = \xi$  for all  $j$ , then (5) is simplified to

$$\mathbf{M}_g(\tau) = \begin{bmatrix} \left(\sum_j \eta(j) \mathbf{e}_j \mathbf{e}_j^T\right) \otimes \mathbf{M}_{11}(\xi) & \left(\sum_j \eta(j) \mathbf{e}_j\right) \otimes \mathbf{M}_{12}(\xi) \\ \left(\sum_j \eta(j) \mathbf{e}_j^T\right) \otimes \mathbf{M}_{21}(\xi) & \mathbf{M}_{22}(\xi) \end{bmatrix}.$$

Note that the information matrix of the design  $\xi$  is associated with the quantitative model

$$E[y(\mathbf{x})] = [f_1^T(\mathbf{x}), f_2^T(\mathbf{x})] \left(\tilde{\boldsymbol{\beta}}^T, \tilde{\boldsymbol{\gamma}}^T\right)^T = f^T(\mathbf{x}) \tilde{\boldsymbol{\theta}}, \tag{6}$$

in

$$\mathbf{M}_f(\xi) = \begin{bmatrix} \mathbf{M}_{11}(\xi) & \mathbf{M}_{12}(\xi) \\ \mathbf{M}_{21}(\xi) & \mathbf{M}_{22}(\xi) \end{bmatrix}.$$

Then Lemma 1 can be obtained easily as follows.

**Lemma 1.** If  $\tau$  is a product design with the marginal designs  $\eta$  and  $\xi$  on  $\mathcal{X}_J$  and  $\mathcal{X}$ , respectively, then the following equation of determinants holds for model (2) and (6):

$$\det(\mathbf{M}_g(\tau)) = \left(\prod_{j=1}^J \eta(j)\right)^{p_1} [\det(\mathbf{M}_{11}(\xi))]^{J-1} \det(\mathbf{M}_f(\xi)),$$

where  $p_1$  is the dimension of  $\mathbf{M}_{11}$ .

Under  $D$ -criterion the above lemma implies the marginal design  $\eta$  should be considered as an uniform design on  $\mathcal{X}_J$ , i.e.  $\eta(j) = 1/J$  for all  $j$ , since the maximization of  $\det(\mathbf{M}_g(\tau))$  can be separated into two parts corresponding to the marginal designs.

To prove the  $D$ -optimality by the equivalence theorem in next section, the dispersion function is introduced here which is proportional to the variance of the predicted response and defined by

$$d_h(\mathbf{z}; \tau) := h^T(\mathbf{z}) \mathbf{M}_h^{-1}(\tau) h(\mathbf{z}), \text{ for } \mathbf{z} \in \Omega.$$

In the following, a connection of dispersion functions between the model (2) and (6) is derived for product designs and the proof is given in the Appendix. The determinant and inverse of a partitioned matrix can be obtained according to the formulas in Khuri (2003, pp. 35-36).

**Lemma 2.** Under the same assumptions of Lemma 1, we have

$$d_g(j, \mathbf{x}^T; \tau) = d_f(\mathbf{x}; \xi) + \left(\frac{1}{\eta(j)} - 1\right) \Delta_{f_1}(\mathbf{x}; \xi), \text{ for } (j, \mathbf{x}) \in \mathcal{X}_J \times \mathcal{X}, \tag{7}$$

where  $d_g(j, \mathbf{x}^T; \tau)$  and  $d_f(\mathbf{x}; \xi)$  are the dispersion functions associated with the models (2) and (6), respectively, and

$$\Delta_{f_1}(\mathbf{x}; \xi) = f_1^T(\mathbf{x}) \mathbf{M}_{11}^{-1}(\xi) f_1(\mathbf{x}).$$

### 3. $D$ -optimal designs

The optimization problem is reduced to search for the optimal design  $\xi^*$  on  $\mathcal{X}$  through implementing result in Lemma 1, and the optimal designs can be constructed by the support points of central composite designs. Let  $\Xi_{CCD}$  be a class collecting designs with different weights on three portions of CCD, namely the cube, axial and center points, which can be expressed as

$$\Xi_{CCD} = \{ \xi : \xi = w_c \xi_c + w_s \xi_s + w_0 \xi_0, w_c + w_s + w_0 = 1 \},$$

where  $\xi_c$  is a  $2^k$  factorial design,  $\xi_s$  is equally supported on the standard basis of  $\mathbf{R}^k$  multiplied by  $\pm r$ , and  $\xi_0$  is one-point design at the origin. Then the non vanishing moments of any design in  $\Xi_{CCD}$  are

$$\mu_u = \int_{\mathcal{X}} x_i^u \xi(d\mathbf{x}) = w_c + \left( \frac{r^u}{k} \right) w_s, \quad u \in \{2, 4\} \text{ and } i \in \{1, \dots, k\},$$

and

$$\mu_{22} = \int_{\mathcal{X}} x_i^2 x_j^2 \xi(d\mathbf{x}) = w_c, \text{ for all } i \neq j.$$

In the following  $\mathcal{X}$  is assumed to be the sphere with radius  $r = \sqrt{k}$ , that is,  $\mathcal{X} = \{ \mathbf{x} \in \mathbf{R}^k : \mathbf{x}^T \mathbf{x} \leq k \}$ . With this choice the support points of designs  $\xi_c$  and  $\xi_s$  lie on the boundary of  $\mathcal{X}$ . On this region the design supporting on the origin with weight  $2/[(k+1)(k+2)]$  and on the boundary uniformly with remaining weight is  $D$ -optimal for the second-order response surface model (Kiefer, 1960). For  $k = 2$  and  $3$ , the vertices of regular polygons and polyhedrons equally spaced on the boundary can be used to constructed the  $D$ -optimal designs. Instead of these,  $D$ -optimal designs made by CCD have the advantages of the sequential examination of model adequacy and structure invariance for all  $k$ . Kiefer and Wolfowitz (1960) gave the celebrated results on the equivalence between  $D$ -optimality and minimaxity of the dispersion function, i.e. the  $D$ -optimality of a design can be verified via the properties of the corresponding dispersion function. In order to find the  $D$ -optimal designs, we first present some results concerning the dispersion function of a design in  $\Xi_{CCD}$ . It is obvious that the dispersion functions for the quantitative model (6) with  $f_1$  and  $f_2$  given in (4a), (4b) and (4c) are identical. The verification of Lemma 3 is shown in the Appendix.

**Lemma 3.** For a design  $\xi$  in  $\Xi_{CCD}$ , the function  $d_f(\mathbf{x};\xi)$  with  $f^T = (f_1^T, f_2^T)$  in one of the (4a), (4b) and (4c) models can be expressed as

$$d_f(\mathbf{x};\xi) = \frac{1}{w_0} + a_1 \sum_{l=1}^k x_l^2 + a_2 \left( \sum_{l=1}^k x_l^2 \right)^2 + a_3 \sum_{l < m}^k x_l^2 x_m^2, \quad (8)$$

where

$$a_1 = \frac{(k+2)w_0 - 2}{kw_0(1-w_0)}, \quad a_2 = \frac{k-1}{k^2w_s} + \frac{1}{k^2w_0(1-w_0)} \quad \text{and} \quad a_3 = \frac{1}{w_c} - \frac{2}{kw_s}.$$

Additionally, the function  $\Delta_{f_1}(\mathbf{x};\xi)$  is

- (i)  $1 + \left(\frac{1}{1-w_0}\right) \sum_{l=1}^k x_l^2$ , for (4a);
- (ii)  $1 + \left(\frac{1}{1-w_0}\right) \sum_{l=1}^k x_l^2 + \frac{1}{w_c} \sum_{l < m}^k x_l^2 x_m^2$ , for (4b);
- (iii)  $d_f(\mathbf{x};\xi) - \frac{1}{w_c} \sum_{l < m}^k x_l^2 x_m^2$ , for (4c).

By symbolizing the support points of designs  $\xi_c$ ,  $\xi_s$  and  $\xi_0$  as  $\mathbf{x}_c$ ,  $\mathbf{x}_s$  and  $\mathbf{x}_0$ , respectively, we solve the following equations with respect to the weights  $w_c$ ,  $w_s$  and  $w_0$  under the constraint that  $w_c + w_s + w_0 = 1$  for each  $j \in \mathcal{X}_J$  and each of the models (4a), (4b) and (4c),

$$d_g(j, \mathbf{x}_c^T; \tau) - d_g(j, \mathbf{x}_s^T; \tau) = 0 \quad \text{and} \quad d_g(j, \mathbf{x}_s^T; \tau) - d_g(j, \mathbf{x}_0^T; \tau) = 0, \quad (9)$$

where  $\tau = \eta \times \xi$  and  $\xi \in \Xi_{CCD}$ . From the former equation, we obtain that the weight  $w_c$  is equal to  $kw_s/2$  for (4a),  $kw_s/[2\eta(j)]$  for (4b) and  $\eta(j)kw_s/2$  for (4c), respectively. The results of the *D*-optimal designs are presented in the following theorem and the proofs based on the equivalence theorem of Kiefer and Wolfowitz (1960) are presented in the Appendix.

**Theorem 4.** The product design  $\tau^*$  with the marginal designs  $\eta$  and  $\xi$  is *D*-optimal if  $\eta$  is a uniform design on  $\mathcal{X}_J$  and  $\xi$  belongs to  $\Xi_{CCD}$  with the optimal weights given as follows:

1. for model (4a),

$$w_s^* = \frac{2k(k+2J+1)}{(k+2)(k^2+2Jk+k+2)} \quad \text{and} \quad w_c^* = \frac{kw_s^*}{2};$$

2. for model (4b),

$$w_s^* = \frac{2k(Jk + J + 2)}{(k + 1)(Jk + 2)^2} \text{ and } w_c^* = \frac{Jkw_s^*}{2};$$

3. for model (4c),

$$w_s^* = \frac{2Jk(k + 4J - 1)}{(k + 2J)[2J(2k + 1) + k(k - 1)]} \text{ and } w_c^* = \frac{kw_s^*}{2J};$$

and for each model  $w_0^* = 1 - w_s^* - w_c^*$ .

Note that the  $D$ -optimal design for each model is rotatable, that is, the associating dispersion function depends only on  $J$ . When  $J = 1$ , these three models are simplified to the second-order response surface model for  $k$ -quantitative factors and the  $D$ -optimal design can be obtained from one of the results in Theorem 4, that is,

$$\xi_{2nd}^* = \frac{k^2(k + 3)}{(k + 1)(k + 2)^2}\xi_c + \frac{2k(k + 3)}{(k + 1)(k + 2)^2}\xi_s + \frac{2}{(k + 1)(k + 2)}\xi_0.$$

This theorem shows that the  $D$ -optimal design puts more weights on the cube portion than the axial portion for models (4a) and (4b) with  $k > 2$ . This agrees with the intuition that the estimations of the linear effects and interactions of the quantitative factors are efficient through the full  $2^k$  factorial design at each level. Furthermore, the comparisons between  $w_s^*$ ,  $w_c^*$  and  $w_0^*$  for each of models (4a) and (4b) show that, for  $J \geq 1$  and  $k \geq 2$ ,

$$w_c^* \geq w_s^* \geq w_0^*.$$

For model (4c), we have that  $w_s^* \geq w_0^*$  for  $k \geq 2$  and

$$w_c^* \geq w_s^*, \text{ if } J \leq \frac{k}{2},$$

and

$$w_c^* \leq w_0^*, \text{ if } J \geq \frac{k}{4} \left( 2k - 1 + \sqrt{4k^2 - 3} \right).$$

Table 1 lists the optimal weights for these models with  $k$  and  $J \in \{2, \dots, 5\}$ . By comparing the optimal weights between the three models, for given  $k$  and  $J$ , both of  $w_s^*$  and  $w_0^*$  for model (4c) are greater than those for models (4a) and (4b), and those for model (4b) are the smallest. These results are given in Figure 1 for  $J = 3$  and  $2 \leq k \leq 20$ .



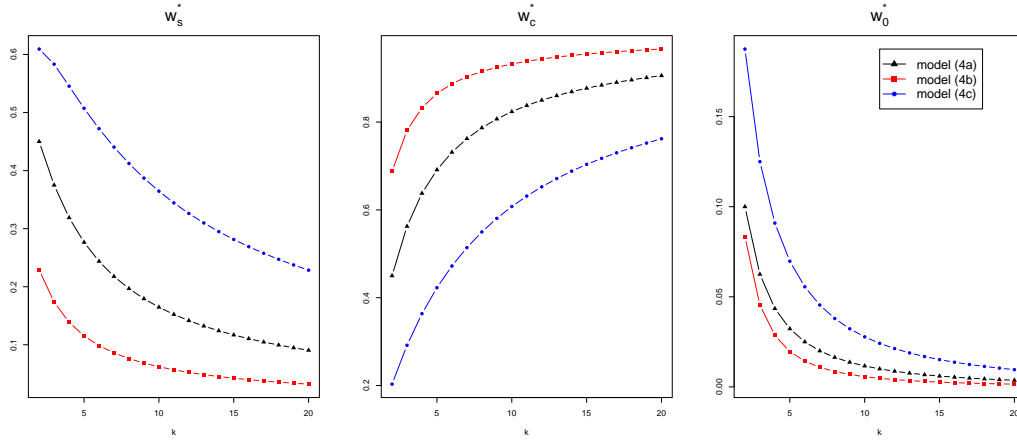


Figure 1: Plots for comparing the optimal weights between models (4a), (4b) and (4c) with  $J = 3$  and  $2 \leq k \leq 20$

Table 1: The weights of  $D$ -optimal design for  $2 \leq k \leq 5$  and  $2 \leq J \leq 5$

k	J	Model (4a)			Model (4b)			Model (4c)		
		$w_s^*$	$w_c^*$	$w_0^*$	$w_s^*$	$w_c^*$	$w_0^*$	$w_s^*$	$w_c^*$	$w_0^*$
2	2	$\frac{7}{16}$	$\frac{7}{16}$	$\frac{1}{8}$	$\frac{8}{27}$	$\frac{16}{27}$	$\frac{1}{9}$	$\frac{6}{11}$	$\frac{3}{11}$	$\frac{2}{11}$
		$\frac{9}{20}$	$\frac{9}{20}$	$\frac{1}{10}$	$\frac{11}{48}$	$\frac{11}{16}$	$\frac{1}{12}$	$\frac{39}{64}$	$\frac{13}{64}$	$\frac{3}{16}$
2	3	$\frac{11}{24}$	$\frac{11}{24}$	$\frac{1}{12}$	$\frac{14}{75}$	$\frac{56}{75}$	$\frac{1}{15}$	$\frac{68}{105}$	$\frac{17}{105}$	$\frac{4}{21}$
		$\frac{13}{28}$	$\frac{13}{28}$	$\frac{1}{14}$	$\frac{17}{108}$	$\frac{85}{108}$	$\frac{1}{18}$	$\frac{35}{52}$	$\frac{7}{52}$	$\frac{5}{26}$
2	4	$\frac{24}{65}$	$\frac{24}{65}$	$\frac{1}{13}$	$\frac{15}{64}$	$\frac{45}{64}$	$\frac{1}{16}$	$\frac{60}{119}$	$\frac{45}{119}$	$\frac{2}{17}$
		$\frac{24}{65}$	$\frac{36}{65}$	$\frac{1}{13}$	$\frac{15}{64}$	$\frac{45}{64}$	$\frac{1}{16}$	$\frac{60}{119}$	$\frac{45}{119}$	$\frac{2}{17}$
2	5	$\frac{28}{65}$	$\frac{28}{65}$	$\frac{1}{13}$	$\frac{17}{108}$	$\frac{85}{108}$	$\frac{1}{18}$	$\frac{35}{52}$	$\frac{7}{52}$	$\frac{5}{26}$
		$\frac{28}{65}$	$\frac{28}{65}$	$\frac{1}{14}$	$\frac{17}{108}$	$\frac{85}{108}$	$\frac{1}{18}$	$\frac{35}{52}$	$\frac{7}{52}$	$\frac{5}{26}$
3	2	$\frac{3}{8}$	$\frac{9}{16}$	$\frac{1}{16}$	$\frac{21}{121}$	$\frac{189}{242}$	$\frac{1}{22}$	$\frac{7}{12}$	$\frac{7}{24}$	$\frac{1}{8}$
		$\frac{3}{8}$	$\frac{9}{16}$	$\frac{1}{16}$	$\frac{21}{121}$	$\frac{189}{242}$	$\frac{1}{22}$	$\frac{7}{12}$	$\frac{7}{24}$	$\frac{1}{8}$
3	3	$\frac{36}{95}$	$\frac{54}{95}$	$\frac{1}{19}$	$\frac{27}{196}$	$\frac{81}{98}$	$\frac{1}{28}$	$\frac{216}{341}$	$\frac{81}{341}$	$\frac{4}{31}$
		$\frac{36}{95}$	$\frac{54}{95}$	$\frac{1}{19}$	$\frac{27}{196}$	$\frac{81}{98}$	$\frac{1}{28}$	$\frac{216}{341}$	$\frac{81}{341}$	$\frac{4}{31}$
3	4	$\frac{21}{55}$	$\frac{63}{110}$	$\frac{1}{22}$	$\frac{33}{289}$	$\frac{495}{578}$	$\frac{1}{34}$	$\frac{165}{247}$	$\frac{99}{494}$	$\frac{5}{38}$
		$\frac{21}{55}$	$\frac{63}{110}$	$\frac{1}{22}$	$\frac{33}{289}$	$\frac{495}{578}$	$\frac{1}{34}$	$\frac{165}{247}$	$\frac{99}{494}$	$\frac{5}{38}$
3	5	$\frac{6}{19}$	$\frac{12}{19}$	$\frac{1}{19}$	$\frac{24}{125}$	$\frac{96}{125}$	$\frac{1}{25}$	$\frac{11}{24}$	$\frac{11}{24}$	$\frac{1}{12}$
		$\frac{6}{19}$	$\frac{12}{19}$	$\frac{1}{19}$	$\frac{24}{125}$	$\frac{96}{125}$	$\frac{1}{25}$	$\frac{11}{24}$	$\frac{11}{24}$	$\frac{1}{12}$
4	2	$\frac{22}{69}$	$\frac{44}{69}$	$\frac{1}{23}$	$\frac{34}{245}$	$\frac{204}{245}$	$\frac{1}{35}$	$\frac{6}{11}$	$\frac{4}{11}$	$\frac{1}{11}$
		$\frac{22}{69}$	$\frac{44}{69}$	$\frac{1}{23}$	$\frac{34}{245}$	$\frac{204}{245}$	$\frac{1}{35}$	$\frac{6}{11}$	$\frac{4}{11}$	$\frac{1}{11}$
4	3	$\frac{26}{81}$	$\frac{52}{81}$	$\frac{1}{27}$	$\frac{44}{405}$	$\frac{352}{405}$	$\frac{1}{45}$	$\frac{38}{63}$	$\frac{19}{63}$	$\frac{2}{21}$
		$\frac{26}{81}$	$\frac{52}{81}$	$\frac{1}{27}$	$\frac{44}{405}$	$\frac{352}{405}$	$\frac{1}{45}$	$\frac{38}{63}$	$\frac{19}{63}$	$\frac{2}{21}$
4	4	$\frac{10}{31}$	$\frac{20}{31}$	$\frac{1}{31}$	$\frac{54}{605}$	$\frac{108}{121}$	$\frac{1}{55}$	$\frac{230}{357}$	$\frac{92}{357}$	$\frac{5}{51}$
		$\frac{10}{31}$	$\frac{20}{31}$	$\frac{1}{31}$	$\frac{54}{605}$	$\frac{108}{121}$	$\frac{1}{55}$	$\frac{230}{357}$	$\frac{92}{357}$	$\frac{5}{51}$
4	5	$\frac{25}{91}$	$\frac{125}{182}$	$\frac{1}{26}$	$\frac{35}{216}$	$\frac{175}{216}$	$\frac{1}{36}$	$\frac{5}{12}$	$\frac{25}{48}$	$\frac{1}{16}$
		$\frac{25}{91}$	$\frac{125}{182}$	$\frac{1}{26}$	$\frac{35}{216}$	$\frac{175}{216}$	$\frac{1}{36}$	$\frac{5}{12}$	$\frac{25}{48}$	$\frac{1}{16}$
5	2	$\frac{60}{217}$	$\frac{150}{217}$	$\frac{1}{31}$	$\frac{100}{867}$	$\frac{250}{289}$	$\frac{1}{51}$	$\frac{240}{473}$	$\frac{200}{473}$	$\frac{3}{43}$
		$\frac{60}{217}$	$\frac{150}{217}$	$\frac{1}{31}$	$\frac{100}{867}$	$\frac{250}{289}$	$\frac{1}{51}$	$\frac{240}{473}$	$\frac{200}{473}$	$\frac{3}{43}$
5	3	$\frac{5}{18}$	$\frac{25}{36}$	$\frac{1}{36}$	$\frac{65}{726}$	$\frac{325}{363}$	$\frac{1}{66}$	$\frac{200}{351}$	$\frac{125}{351}$	$\frac{2}{27}$
		$\frac{5}{18}$	$\frac{25}{36}$	$\frac{1}{36}$	$\frac{65}{726}$	$\frac{325}{363}$	$\frac{1}{66}$	$\frac{200}{351}$	$\frac{125}{351}$	$\frac{2}{27}$
5	4	$\frac{80}{287}$	$\frac{200}{287}$	$\frac{1}{41}$	$\frac{160}{2187}$	$\frac{2000}{2187}$	$\frac{1}{81}$	$\frac{8}{13}$	$\frac{4}{13}$	$\frac{1}{13}$
		$\frac{80}{287}$	$\frac{200}{287}$	$\frac{1}{41}$	$\frac{160}{2187}$	$\frac{2000}{2187}$	$\frac{1}{81}$	$\frac{8}{13}$	$\frac{4}{13}$	$\frac{1}{13}$

#### 4. Discussion

In this study the approximate  $D$ -optimal designs for models (4a), (4b) and (4c) have been provided when the quantitative factors are considered within a sphere design region. The framework for investigating  $D$ -optimal designs can be applied to higher order models for quantitative factors with qualitative factors.

The performance of designs comparing to the  $D$ -optimal design for model  $g(\mathbf{x})$  are measured by the  $D$ -efficiency which is defined by

$$D_{\text{eff}}(\tau) = \frac{|\mathbf{M}_g(\tau)|}{|\mathbf{M}_g(\tau_g^*)|}.$$

The design with 40 trials in the example of  $\text{SO}_2$  removal, denoted by  $\tau_{ex}$ , has poor  $D$ -efficiencies (0.140, 0.034 and 0.151) for estimating the model parameters of (4a), (4b) and (4c). Furthermore, by denoting the  $D$ -optimal designs for models (3), (4a), (4b) and (4c) as  $\tau_3^*$  ( $= \bar{\eta} \times \xi_{2nd}^*$ ),  $\tau_{4a}^*$ ,  $\tau_{4b}^*$  and  $\tau_{4c}^*$ , these designs are compared mutually with each other for  $k = 3$  and  $J = 2$  and the  $D$ -efficiencies are presented in the following table. For models (3) and (4a), the designs  $\tau_3^*$  and  $\tau_{4a}^*$  provide the excellent information for the estimations, and  $\tau_{4b}^*$  and  $\tau_{4c}^*$  have acceptable efficiencies. However, for estimating models (4b) and (4c), the corresponding  $D$ -optimal designs provide the information about only 13.8% for each other models. Thus it can be concluded that  $\tau_3^*$  and  $\tau_{4a}^*$  have robust performance on fitting these models. Further,  $\tau_3^*$  is most robust for these models according to the maximin criterion, that is, the minimum  $D$ -efficiency under other models is the greatest among all 5 candidate designs.

Table 2: Comparisons of  $D$ -efficiency for  $k = 3$  and  $J = 2$

	Model (3)	Model (4a)	Model (4b)	Model (4c)
$\tau_3^*$	1.000	0.960	0.582	0.641
$\tau_{4a}^*$	0.966	1.000	0.655	0.554
$\tau_{4b}^*$	0.689	0.747	1.000	0.138
$\tau_{4c}^*$	0.732	0.662	0.138	1.000
$\tau_{ex}$	0.310	0.140	0.034	0.151

## Appendix

### A.1 Experiment of flue gas desulfurization

The original data described in Zainudin et al. (2005) is summarized in the following tables. Totally four factors as code variables are shown in Table A.1

with the original levels, and Table A.2 presents the results of the experiment with a 20-run CCD at each qualitative level. The analysis of variance is summarized in Table A.3 in which the total variation is explained about 98.6 percent by fitting model (4b).

Table A.1: Factors and levels of flue gas desulfurization experiment

Quantitative factors :	-1.682	-1	0	1	1.682
$x_1$ : Hydration period (h)	3.18	10	20	30	36.82
$x_2$ : Ratio of OPA to CaO or Ca(OH) <sub>2</sub> (g/g)	0.32 : 1	1 : 1	2 : 1	3 : 1	3.68 : 1
$x_3$ : Amount of CaSO <sub>4</sub> (g)	0.32	1	2	3	3.68
Qualitative factor:	$j = 1$	$j = 2$			
CaO or Ca(OH) <sub>2</sub>	CaO	Ca(OH) <sub>2</sub>			

Table: A.2 The design matrix of the 20-run CCD for each qualitative level and the response data

Factor			response( $y$ )		Factor			response( $y$ )	
$x_1$	$x_2$	$x_3$	$j = 1$	$j = 2$	$x_1$	$x_2$	$x_3$	$j = 1$	$j = 2$
-1	-1	-1	31.85	36.25	-1	-1	1	24.45	35.82
1	-1	-1	89.22	71.74	1	-1	1	45.65	68.77
-1	1	-1	24.36	34.05	-1	1	1	28.65	30.10
1	1	-1	120.5	126.00	1	1	1	99.11	119.30
-1.682	0	0	15.48	18.74	0	0	0	48.93	59.37
1.682	0	0	78.67	134.20	0	0	0	55.13	55.11
0	-1.682	0	27.34	26.01	0	0	0	55.17	40.34
0	1.682	0	79.24	105.90	0	0	0	42.68	54.35
0	0	-1.682	56.91	41.28	0	0	0	47.97	52.73
0	0	1.682	33.97	50.60	0	0	0	47.77	48.47

Table: A.3 Analysis of variance for the experiment of flue gas desulfurization

Source of variation	Degrees of freedom	Sum of squares	Mean square	$F$	$p$ -value
Model	17	163289.856	9605.286	95.833	< 0.0001
Qualitative factor	2	128564.439	64282.220	641.351	< 0.0001
Within Qualitative					
Linear effects ( $j = 1$ )	3	11961.379	3987.126	39.780	< 0.0001
Linear effects ( $j = 2$ )	3	18335.423	6111.808	60.978	< 0.0001
Interaction ( $j = 1$ )	3	1590.235	530.0783	5.289	0.006
Interaction ( $j = 2$ )	3	1598.011	532.6704	5.315	0.006
Between Qualitative					
Quadratic effects	3	1240.369	413.456	4.125	0.018
Residual	23	2305.274	100.229		
Total	40	165595.131			

## A.2 Proof of Lemma 2

By calculating the inverse matrices of  $\mathbf{M}_g(\tau)$  according to the formula of a partitioned matrix in Khuri (2003, pp. 36), we have

$$\mathbf{M}_g^{-1}(\tau) = \begin{pmatrix} \mathbf{D}_{11}(\tau) & \mathbf{D}_{12}(\tau) \\ \mathbf{D}_{21}(\tau) & \mathbf{D}_{22}(\tau) \end{pmatrix}$$

where  $\tau = \eta \times \xi$ ,

$$\begin{aligned} \mathbf{D}_{11}(\tau) &= \text{Diag}\left(\frac{1}{\eta(1)}, \dots, \frac{1}{\eta(J)}\right) \otimes \mathbf{M}_{11}^{-1}(\xi) \\ &\quad + (\mathbf{1}_J \mathbf{1}_J^T) \otimes [\mathbf{M}_{11}^{-1}(\xi) \mathbf{M}_{12}(\xi) \mathbf{D}_{22}(\tau) \mathbf{M}_{21}(\xi) \mathbf{M}_{11}^{-1}(\xi)], \\ \mathbf{D}_{12}(\tau) &= -\mathbf{1}_J \otimes [\mathbf{M}_{11}^{-1}(\xi) \mathbf{M}_{12}(\xi) \mathbf{D}_{22}(\tau)] = \mathbf{D}_{21}^T(\tau), \\ \mathbf{D}_{22}(\tau) &= [\mathbf{M}_{22}(\xi) - \mathbf{M}_{21}(\xi) \mathbf{M}_{11}^{-1}(\xi) \mathbf{M}_{12}(\xi)]^{-1}, \end{aligned}$$

and  $\mathbf{1}_J$  is the  $k \times 1$  vector  $[1, \dots, 1]^T$ . It is obvious that these submatrices are connected with the submatrices of  $\mathbf{M}_f^{-1}(\xi)$ . From the definition of dispersion function in Section 2, equation (7) can be obtained after the algebraic simplification with properties of the Kronecker product.

## Appendix A.3 Proof of Lemma 3

Note that model (6) is identical to the second-order response surface model with the  $f_1$  and  $f_2$  under all three models (4a), (4b) and (4c). For any  $\xi$  in  $\Xi_{CCD}$ , the moment matrices associated with the linear effects, 2-factor interactions and quadratic effects by denoting as  $\mathbf{M}_L$ ,  $\mathbf{M}_I$  and  $\mathbf{M}_Q$  are, respectively,

$$\mathbf{M}_L(\xi) = \int_{\mathcal{X}} f_L(\mathbf{x}) f_L^T(\mathbf{x}) d\xi(\mathbf{x}) = \mu_2 \mathbf{I}_k,$$

$$\mathbf{M}_I(\xi) = \int_{\mathcal{X}} f_I(\mathbf{x}) f_I^T(\mathbf{x}) d\xi(\mathbf{x}) = \mu_{22} \mathbf{I}_{\frac{k(k-1)}{2}},$$

and

$$\mathbf{M}_Q(\xi) = \int_{\mathcal{X}} f_Q(\mathbf{x}) f_Q^T(\mathbf{x}) d\xi(\mathbf{x}) = (\mu_4 - \mu_{22}) \mathbf{I}_k + \mu_{22} \mathbf{J}_k,$$

where  $\mathbf{I}_k$  is the  $k \times k$  identity matrix,  $\mathbf{J}_k = \mathbf{1}_k \mathbf{1}_k^T$ . As a result, we have

$$\begin{aligned} d_f(\mathbf{x}; \xi) &= f^T(\mathbf{x}) \mathbf{M}_f^{-1}(\xi) f(\mathbf{x}) \\ &= f_L^T(\mathbf{x}) \mathbf{M}_L^{-1}(\xi) f_L(\mathbf{x}) + f_I^T(\mathbf{x}) \mathbf{M}_I^{-1}(\xi) f_I(\mathbf{x}) \\ &\quad + [1, f_Q^T(\mathbf{x})]^T \begin{pmatrix} 1 & \mu_2 \mathbf{1}_k^T \\ \mu_2 \mathbf{1}_k & \mathbf{M}_Q(\xi) \end{pmatrix}^{-1} \begin{bmatrix} 1 \\ f_Q(\mathbf{x}) \end{bmatrix}, \end{aligned}$$

where  $f^T(\mathbf{x}) = [1, f_L^T(\mathbf{x}), f_I^T(\mathbf{x}), f_Q^T(\mathbf{x})]$ ,

$$\mathbf{M}_f(\xi) = \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} & \mu_2 \mathbf{1}_k^T \\ \mathbf{0} & \mathbf{M}_L(\xi) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_I(\xi) & \mathbf{0} \\ \mu_2 \mathbf{1}_k & \mathbf{0} & \mathbf{0} & \mathbf{M}_Q(\xi) \end{pmatrix}.$$

It can be simplified by using the results that  $\mu_2 = w_c + w_s$ ,  $\mu_{22} = w_c$ ,  $\mu_4 - \mu_{22} = kw_s$  and

$$\begin{pmatrix} 1 & \mu_2 \mathbf{1}_k^T \\ \mu_2 \mathbf{1}_k & \mathbf{M}_Q(\xi) \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{w_0} & -\frac{1}{kw_0} \mathbf{1}_k^T \\ -\frac{1}{kw_0} \mathbf{1}_k & \frac{1}{kw_s} \mathbf{I}_k + \left( \frac{1}{k^2 w_0(1-w_0)} - \frac{1}{k^2 w_s} \right) \mathbf{J}_k \end{pmatrix}.$$

Then (8) can be obtained easily. The function  $\Delta_{f_1}(\mathbf{x}; \xi)$  can be obtained that, for model (4a),

$$\Delta_{f_1}(\mathbf{x}; \xi) = 1 + f_L^T(\mathbf{x}) \mathbf{M}_L^{-1}(\xi) f_L(\mathbf{x}),$$

and for model (4b),

$$\Delta_{f_1}(\mathbf{x}; \xi) = 1 + f_L^T(\mathbf{x}) \mathbf{M}_L^{-1}(\xi) f_L(\mathbf{x}) + f_I^T(\mathbf{x}) \mathbf{M}_I^{-1}(\xi) f_I(\mathbf{x}),$$

and for model (4c),

$$\begin{aligned} \Delta_{f_1}(\mathbf{x}; \xi) &= f_L^T(\mathbf{x}) \mathbf{M}_L^{-1}(\xi) f_L(\mathbf{x}) \\ &+ [1, f_Q^T(\mathbf{x})]^T \begin{pmatrix} 1 & \mu_2 \mathbf{1}_k^T \\ \mu_2 \mathbf{1}_k & \mathbf{M}_Q(\xi) \end{pmatrix}^{-1} \begin{bmatrix} 1 \\ f_Q(\mathbf{x}) \end{bmatrix}. \end{aligned}$$

Lemma 3 is proved.

#### A.4 Proof of Theorem 4

According to the equivalence theorem, to verify the *D*-optimality is equivalent to show that the maximum of the dispersion function  $d_g(j, \mathbf{x}^T; \tau^*)$  w.r.t  $(j, \mathbf{x}^T)$  is no greater than the number of model parameters and the maximum will be attained at the support points of  $\tau^*$ .

First, the coefficient of the term  $\sum_{i < j}^k x_i^2 x_j^2$  in  $d_g(j, \mathbf{x}; \tau^*)$  can be computed easily for each model according to Lemma 2 and 3, then the relationship between  $w_c^*$  and  $w_s^*$  described in the theorem makes it vanish for all  $j$ . This implies that the dispersion function depends only on  $\mathbf{x}^T \mathbf{x}$ , in other words,  $d_g(j, \mathbf{x}; \tau^*)$  is a second-degree polynomial in  $\mathbf{x}^T \mathbf{x}$ . Further, the positive coefficient for the terms of the highest degree indicates that the maximum of the dispersion function for

each  $j$  occurs at  $\mathbf{x}^T \mathbf{x} = 0$  or  $k$ . Since the optimal weights are the solutions of the equations (9), then we have

$$d_g(j, \mathbf{x}_0; \tau^*) = d_g(j, \mathbf{x}_c; \tau^*) = d_g(j, \mathbf{x}_s; \tau^*), \quad \forall j.$$

Next, for each model, the maximum can be easily evaluated from  $d_g(j, \mathbf{x}_0; \tau^*)$ , i.e. the intercept of the dispersion function. Then applying the results of Lemma 3 to Lemma 2, we have

$$d_g(j, \mathbf{x}_0; \tau^*) = \begin{cases} \frac{1}{w_0^*} + (J - 1), & \text{for (4a) and (4b),} \\ \frac{J}{w_0^*}, & \text{for (4c),} \end{cases}$$

and

$$w_0^* = \begin{cases} 1 - \frac{k+2}{2} w_s^*, & \text{for (4a),} \\ 1 - \frac{Jk+2}{2} w_s^*, & \text{for (4b),} \\ 1 - \frac{k+2J}{2J} w_s^*, & \text{for (4c).} \end{cases}$$

A straightforward algebraic computation yields that the maximum is equal to  $(k+1)(k+2J)/2$  for model (4a),  $[J+k+Jk(k+1)/2]$  for model (4b) and  $[J(2k+1)+k(k-1)/2]$  for model (4c), respectively, which are the number of the model parameters. Then the proof is complete.

## References

- Atkinson, A. C. and Donev, A. N. (1989). The construction of exact  $D$ -optimum experimental designs with application to blocking response surface designs. *Biometrika* **76**, 515-526.
- Draper, N. R. and John, J. A. (1988). Response-surface designs for quantitative and qualitative variables. *Technometrics* **30**, 423-428.
- Huang, M.-N. L., Lee, C. P., Chen, R. B., and Klein, T. (2010). Exact  $D$ -optimal designs for a second-order response surface model on a circle with qualitative factors. *Comput. Statist. Data Anal.* **54**, 516-530.
- Khuri, A. I., 2003. *Advanced Calculus with Applications in Statistics*, 2nd edition. Wiley.
- Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* **12**, 363-366.
- Schwabe, R. and Wierich, W. (1995).  $D$ -optimal designs of experiments with non-interacting factors. *J. Statist. Plann. Inference* **44**, 371-384.
- Wu, C. F. J. and Ding, Y. (1998). Construction of response surface designs for qualitative and quantitative factors. *J. Statist. Plann. Inference* **71**, 331-384.

---

Zainudin, N. F., Lee, K. T., Kamaruddin, A. H., Bhatia, S., and Mohamed, A. R. (2005).  
Study of adsorbent prepared from oil palm ash (OPA) for flue gas desulfurization.  
*Sep. Purif. Technol.* **45**, 50-60.

Received August 29, 2009; accepted January 25, 2010.

Chuan-Pin Lee

Department of Applied Mathematics

National Sun Yat-sen University

No. 70, Lienhai Rd., Kaohsiung 80424, Taiwan, R.O.C.

leecb@math.nsysu.edu.tw

Mong-Na Lo Huang

Department of Applied Mathematics

National Sun Yat-sen University

No. 70, Lienhai Rd., Kaohsiung 80424, Taiwan, R.O.C.

lomn@math.nsysu.edu.tw