Statistica Sinica Preprint No: SS-2023-0165	
Title	Minimum Aberration Factorial Designs Under A Mixed Parametrization
Manuscript ID	SS-2023-0165
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202023.0165
Complete List of Authors	Guanzhou Chen and Boxin Tang
Corresponding Authors	Guanzhou Chen
E-mails	guanzhou_chen@sfu.ca

Statistica Sinica

MINIMUM ABERRATION FACTORIAL DESIGNS UNDER A MIXED PARAMETRIZATION

Guanzhou Chen[1] and Boxin Tang[2]

[1]*Nankai University and* [2]*Simon Fraser University*

Abstract: The baseline parametrization for two-level factorial designs has been receiving increasing attention recently. While the orthogonal parametrization is appropriate for experiments where the two levels of each factor are symmetrical, the baseline parametrization is well suited for experiments where the two levels of each factor are asymmetrical and one level, called a baseline level, is the default level. This paper considers a general situation where some factors have a baseline level while others do not. A mixed parametrization of factorial effects is proposed and its connection with the existing parametrizations is established. Under this new parametrization, we show that orthogonal arrays continue to be optimal for estimating main effects, and then put forward two minimum aberration criteria for further design selection. Both theoretical and algorithmic constructions of minimum aberration designs are examined and useful designs are obtained.

Key words and phrases: Baseline parametrization, contamination, orthogonal array.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

1. Introduction

Two-level factorial designs are a class of experimental plans useful in scientific and technological investigations for studying the causal relationship between several input factors and a response variable. Factorial effects are utilized to attribute changes of the mean response due to various level combinations to the factors under study. The most commonly used factorial effects are those given by the orthogonal parametrization (Box and Hunter, 1961), which is termed so because those factorial effects form a set of orthogonal treatment contrasts. When it is too expensive to examine all level combinations, factorial effects cannot be all estimated and a fractional factorial design needs to be selected to entertain the estimation of the lower-order effects. One popular approach to design selection is to employ the minimum aberration criterion (Fries and Hunter, 1980; Tang and Deng, 1999). We refer to Mee (2009), Cheng (2014) and Wu and Hamada (2021) for comprehensive accounts on factorial designs under the orthogonal parametrization.

Under the orthogonal parametrization, the two levels of the factors are symmetrical and hence equally important. While this is true in most applications, there are situations, such as in microarray experiments (Yang and Speed, 2002; Glonek and Solomon, 2004; Banerjee and Mukerjee, 2008),

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

where one of the two levels represents a baseline or default setting and is thus more important than the other level. Investigators are interested in the impact on the mean response by changing the levels of a few factors while keeping other factors set at the baseline levels. This calls for a baseline parametrization in which factorial effects are defined in relation to the baseline levels. To select a fractional factorial design under this parametrization, Mukerjee and Tang (2012) put forward a minimum aberration criterion which aims at minimizing the bias caused by higher-order interactions on the estimation of main effects.

The blanket approach to defining factorial effects via either the orthogonal parametrization or the baseline parametrization can hardly represent all practical situations. Entirely conceivable are the scenarios that we know the importance of one of the two levels for some factors but are indifferent to the two levels for other factors. In an industrial experiment on quality improvement, besides studying the potential impact of changing the current settings of several machine components in a production line, we may also want to examine some additional factors along the way. Then the current settings may be regarded as the baseline levels for the machine components, but no importance can be attached to any of the two levels for the additional factors. To deal with such practical situations, we propose a mixed

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

parametrization of factorial effects in which some factors have baseline levels while the others do not. Our mixed parametrization includes as special cases both the orthogonal and baseline parametrizations.

The remainder of the paper is arranged as follows. Section 2 first reviews orthogonal and baseline parametrizations, and then introduces the mixed parametrization. A connection between the mixed parametrization and the existing parametrizations is established, through which we show that orthogonal arrays are optimal for estimating the main effects under the main-effects model. To protect the main effects from the contamination of nonnegligible higher-order interactions, two minimum aberration criteria are developed in Section 3, depending on whether or not the main effects of the factors with baseline levels need more protection than those of the other factors. Theoretical constructions are then provided to minimize the leading terms of these criteria. In Section 4, we present two algorithms to search for designs that are exactly optimal or nearly optimal under these criteria. All designs with 8, 12, 16 and 20 runs are found and made available online. The paper is concluded with a discussion in Section 5. All the proofs and some selected designs are provided in the Supplementary Material.

2. A mixed parametrization and optimality results

Consider a factorial experiment for m two-level factors $F_1, F_2, \ldots, F_m$ in which the two levels are denoted by -1 and $+1$. Let $S = \{1, 2, \ldots, m\}$ collect the indices of these factors. Then for any subset $u \subseteq S$, there corresponds a treatment combination $x_u = (x_{u1}, \ldots, x_{um})$ where $x_{uj} = +1$ if $j \in u$ and $x_{uj} = -1$ otherwise. We use τ_u to represent the treatment mean under the treatment combination x_u.

We first review the orthogonal parametrization of factorial effects. For any subset $w = \{j_1, \ldots, j_k\} \subseteq S$, let β_w be the factorial effect involving the k factors $F_{j_1}, \ldots, F_{j_k}$ under the orthogonal parametrization. Then we have

$$\tau_u = \sum_{w \subseteq S} \beta_w \prod_{j \in w} x_{uj}, \qquad \beta_w = \frac{1}{2^m} \sum_{u \subseteq S} \tau_u \prod_{j \in w} x_{uj}. \tag{2.1}$$

Mathematically, the treatment means τ_u's and the factorial effects β_w's are just a linear transformation of each other. However, the β_w's are statistically meaningful because they describe the change in treatment means due to the level changes of factors indexed by w. More concretely, the factorial effect β_w defines a treatment contrast by averaging over all possible level combinations of factors not contained in w. For example, the main effects are given by $\beta_j = (1/2^m) \sum_{u \subseteq S \setminus \{j\}} (\tau_{u \cup \{j\}} - \tau_u)$ for $j = 1, \ldots, m$.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

The orthogonal parametrization is well suited for situations where the two levels are symmetrical. For the opposite situations where one of the two levels corresponds to a baseline or default setting, the baseline parametrization may be more appropriate. We suppose the level -1 is the baseline level. For $w \subseteq S$, let θ_w be the factorial effect involving factors indexed by w under the baseline parametrization. Let $z_{uj} = x_{uj} + 1$ for $u \subseteq S$ and $j = 1, \ldots, m$. Then we have

$$\tau_u = \sum_{w \subseteq S} \theta_w \prod_{j \in w} z_{uj}, \qquad \theta_w = \frac{1}{2^{|w|}} \sum_{u \subseteq w} \tau_u \prod_{j \in w} x_{uj}, \tag{2.2}$$

where $|w|$ denotes the cardinality of w. In contrast to β_w's, the θ_w's characterize the factorial effect due to factors in w by fixing all other factors at the baseline level -1. For example, the main effects under the baseline parametrization are $\theta_j = (\tau_j - \tau_\phi)/2$ for $j = 1, \ldots, m$.

In the existing work on baseline designs, the two levels ± 1 are converted to 0 and 1 by $z_{uj} = (x_{uj} + 1)/2$. Our slightly different definition transforms ± 1 to 0 and 2, which is to ensure that β_w and θ_w have the same scale and are comparable. This modification gives rise to the extra $1/2^{|w|}$ in the expression of θ_w in (2.2).

We now consider a general situation in which the two levels are asymmetrical for some factors and symmetrical for the others. Without loss of generality, we assume the level -1 is the baseline level for the first m_1 factors $F_1, \ldots, F_{m_1}$, and for the remaining $m_2 = m - m_1$ factors $F_{m_1+1}, \ldots, F_m$, the two levels are symmetrical. For convenience, we call the first m_1 factors B-factors and the last m_2 factors O-factors. To define a mixed parametrization of factorial effects, we need to introduce some notation. Let $S_1 = \{1, \ldots, m_1\}$ and $S_2 = \{m_1 + 1, \ldots, m\}$, representing the index sets of B-factors and O-factors, respectively. For $w_1 \subseteq S_1$ and $w_2 \subseteq S_2$, let $\xi_{w_1 \cup w_2}$ be the factorial effect involving factors in $w_1 \cup w_2$ under the mixed parametrization. Then we have

$$\tau_u = \sum_{w_1 \subseteq S_1} \sum_{w_2 \subseteq S_2} \xi_{w_1 \cup w_2} \prod_{j \in w_1} z_{uj} \prod_{j \in w_2} x_{uj},$$

$$\xi_{w_1 \cup w_2} = \frac{1}{2^{|w_1| + m_2}} \sum_{u \subseteq w_1 \cup S_2} \tau_u \prod_{j \in w_1 \cup w_2} x_{uj}, \quad (2.3)$$

where $z_{uj} = x_{uj} + 1$. Clearly, (2.3) reduces to (2.1) if $S_1 = \phi$ and to (2.2) if $S_2 = \phi$. Therefore, our mixed parametrization includes as special cases the orthogonal and baseline parametrizations. The factorial effects under the mixed parametrization inherit features of the two parametrizations introduced above: The parameter $\xi_{w_1 \cup w_2}$ measures the effect of factors in $w_1 \cup w_2$ by averaging over all level combinations of O-factors in $S_2 \setminus w_2$ while fixing the B-factors in $S_1 \setminus w_1$ at the baseline level. For example, the main effects for B-factors are given by $\xi_j = (1/2^{m_2+1}) \sum_{u \subseteq S_2} (\tau_{u \cup \{j\}} - \tau_u)$ for $j = 1, \ldots, m_1$, and those for O-factors are defined as $\xi_j = (1/2^{m_2}) \sum_{u \subseteq S_2 \setminus \{j\}} (\tau_{u \cup \{j\}}$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

$-\tau_u$) for $j = m_1 + 1, \ldots, m$. The following example illustrates the three parametrizations by a 2^2 factorial.

Example 1. Suppose that $m = 2$ with $m_1 = m_2 = 1$ so the first factor is a B-factor and the second is an O-factor. There are 4 treatment combinations τ_ϕ, τ_1, τ_2 and τ_{12}. Under the three parametrizations discussed above, we obtain that

$$\beta_\phi = (\tau_\phi + \tau_1 + \tau_2 + \tau_{12})/4, \quad \theta_\phi = \tau_\phi, \quad \xi_\phi = (\tau_\phi + \tau_2)/2;$$

$$\beta_1 = \xi_1 = (-\tau_\phi + \tau_1 - \tau_2 + \tau_{12})/4, \quad \theta_1 = (\tau_1 - \tau_\phi)/2;$$

$$\beta_2 = (-\tau_\phi - \tau_1 + \tau_2 + \tau_{12})/4, \quad \theta_2 = \xi_2 = (\tau_2 - \tau_\phi)/2;$$

and $\beta_{12} = \theta_{12} = \xi_{12} = (\tau_\phi - \tau_1 - \tau_2 + \tau_{12})/4$.

As can be seen from (2.1), (2.2) and (2.3), the factorial effects under the three parametrizations are all linear transformations of the treatment means, and hence must be linearly related to each other. Sun and Tang (2022) established a linear relationship between the orthogonal and baseline parametrizations. Theorem 1 further reveals relationships between the mixed parametrization and the other two.

Theorem 1. *For any $w_1 \subseteq S_1$ and $w_2 \subseteq S_2$, we have that*

(i) $\xi_{w_1 \cup w_2} = \sum_{v_1 \supseteq w_1} (-1)^{|v_1| - |w_1|} \beta_{v_1 \cup w_2}$ *and* $\beta_{w_1 \cup w_2} = \sum_{v_1 \supseteq w_1} \xi_{v_1 \cup w_2}$*; and*

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

(ii) $\xi_{w_1 \cup w_2} = \sum_{v_2 \supseteq w_2} \theta_{w_1 \cup v_2}$ *and* $\theta_{w_1 \cup w_2} = \sum_{v_2 \supseteq w_2} (-1)^{|v_2|-|w_2|} \xi_{w_1 \cup v_2}$.

We note that the relationship between the orthogonal and baseline parametrizations can be obtained by taking $S_1 = S$ and $S_2 = \phi$ in part (i) of Theorem 1. More importantly, one can easily deduce from Theorem 1 the equivalency of the three conditions: (a) $\xi_w = 0$ for all $|w| \geq k$, (b) $\beta_w = 0$ for all $|w| \geq k$, and (c) $\theta_w = 0$ for all $|w| \geq k$, for any given positive integer k. This leads to the following result.

Corollary 1. *The factorial effects involving k or more factors are negligible under any one parametrization implies the same under the other two parametrizations. In particular, if all interactions are negligible under one parametrization, they must be negligible under the two parametrizations, in which case we have that $\xi_j = \beta_j = \theta_j$ for $j = 1, \ldots, m$.*

Now let's focus on the estimation of main effects ξ_j's under the mixed parametrization, using a design $D = (d_{ij})$ of N runs for m factors. Let X_1 be an $N \times m$ matrix with its (i, j)th element equal to $(d_{ij} + 1)$ if $j \leq m_1$ and d_{ij} otherwise. Consider the following main-effects model

$$Y = 1_N \xi_\phi + X_1 \xi^{(1)} + \epsilon, \tag{2.4}$$

where $Y = (Y_1, \ldots, Y_N)^{\mathrm{T}}$ is the vector of responses, 1_N is a column of N ones, $\xi^{(1)} = (\xi_1, \ldots, \xi_m)^{\mathrm{T}}$ and ϵ is the vector of uncorrelated random errors that have a zero mean and a constant variance σ^2. The results of Corollary 1 imply that such a model is equivalent to a main-effects model under the orthogonal parametrization. Then the following optimality results as stated in Corollary 2 can be established, where part (i) follows directly from Proposition 1 of Mukerjee and Tang (2012) and the fact that $\xi_j = \beta_j$ for $j = 1, \ldots, m$, and part (ii) is proved in the Supplementary Material. Recall that D is an orthogonal array of strength t if any t columns of D contain all possible level combinations of -1 and $+1$ the same number of times; we denote such an array by $\text{OA}(N, 2^m, t)$.

Corollary 2. *With reference to the model* (2.4)*, we have that*

(i) the best linear unbiased estimator $\hat{\xi}_j$ *of* ξ_j *satisfies* $\text{Var}(\hat{\xi}_j) \geq \sigma^2/N$ *for* $j = 1, \ldots, m$*, where the equality holds if and only if* D *is an* $\text{OA}(N, 2^m, 2)$*; and*

(ii) if design D *is an* $\text{OA}(N, 2^m, 2)$*, then* D *is universally optimal for estimating* $\xi^{(1)}$*.*

3. Two minimum aberration criteria

3.1 Bias caused by nonnegligible interactions

Corollary 2 shows that under the model (2.4) which ignores interactions, an orthogonal array is optimal for estimating the main effects $\xi^{(1)}$ in a very

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

broad sense. Let $\xi = (\xi_\phi, \xi^{(1)\mathrm{T}})^\mathrm{T}$. Then the best linear unbiased estimator for ξ is $\hat{\xi} = (X^\mathrm{T}X)^{-1}X^\mathrm{T}Y$, where $X = (1_N, X_1)$. However, this estimator is actually biased if interactions are not negligible. Suppose the true model is the full model

$$Y = 1_N\xi_\phi + X_1\xi^{(1)} + X_2\xi^{(2)} + \cdots + X_m\xi^{(m)} + \epsilon,$$

where $\xi^{(k)}$ collects all k-factor interactions ξ_w's with $|w| = k$, and X_k is the corresponding model matrix for $k = 1, \ldots, m$. Then the bias in the estimator $\hat{\xi}$ is given by

$$E(\hat{\xi}) - \xi = (X^\mathrm{T}X)^{-1}X^\mathrm{T}X_2\xi^{(2)} + \cdots + (X^\mathrm{T}X)^{-1}X^\mathrm{T}X_m\xi^{(m)}. \tag{3.5}$$

In this section, we concentrate on selecting an orthogonal array that minimizes the contamination of the potentially active interactions on the estimation of main effects. Two minimum aberration criteria are proposed to implement the idea, depending on whether or not the main effects of the B-factors need more protection than those of the O-factors.

3.2 Main effects of B-factors are more important

Under the mixed parametrization, there are two sets of main effects, one for the B-factors and the other for the O-factors. In practice, the two sets of main effects may not be of equal interest and thus ought to be treated

doi:10.5705/ss.202023.0165

differently. In this subsection, we consider the situation that the main effects of the B-factors are more important than those of the O-factors, and therefore need more protection from contamination by nonnegligible interactions. This is reasonable because the B-factors may well be those that have current default settings and the O-factors are some additional factors the investigator want to study. Default settings need to be protected; so do the B-factors that have default settings.

From the bias expression (3.5), one can see that for $k = 2, \ldots, m$, the k-factor interactions $\xi^{(k)}$ contribute a bias term of $B_k\xi^{(k)}$ to the estimation of main effects for B-factors, where B_k collects the rows $2, \ldots, m_1+1$ of the matrix $(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}X_k$. Similarly, the bias caused by $\xi^{(k)}$ on the estimation of main effects for O-factors is $O_k\xi^{(k)}$, where O_k collects the last m_2 rows of the matrix $(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}X_k$. If all components of $\xi^{(k)}$ are equally likely to be active with the same scale, then $\pi_k^B = \mathrm{tr}(B_k^{\mathrm{T}}B_k)$ and $\pi_k^O = \mathrm{tr}(O_k^{\mathrm{T}}O_k)$ provide reasonable measures of the amount of bias from $\xi^{(k)}$ on main-effects estimation for B-factors and O-factors, respectively.

Under the assumption that the main effects of B-factors are more important, it is a priority to protect these main effects from the contamination of interaction terms. On the other hand, the effect hierarchy principle says that lower-order interactions are more likely to be active than the higher-

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

order ones. Therefore, when only two-factor interactions are present, an orthogonal array that sequentially minimizes π_2^B and π_2^O is desirable. If, in addition, there are nonnegligible three-factor interactions, we then proceed to minimize π_3^B and π_3^O. Continuing this line of arguments, we obtain the following minimum π_B-aberration criterion for design selection.

Definition 1. An orthogonal array for m factors is said to have minimum π_B-aberration if it sequentially minimizes $\pi_2^B, \pi_2^O, \pi_3^B, \pi_3^O, \ldots, \pi_m^B, \pi_m^O$.

The idea of minimum π_B-aberration criterion is similar in spirit to those of the minimum G_2-aberration under the orthogonal parametrization (Tang and Deng, 1999) and the minimum K-aberration under the baseline parametrization (Mukerjee and Tang, 2012). To find a minimum aberration design is challenging, and our problem is further complicated by the presence of two types of factors. Nevertheless, good designs can still be obtained theoretically by concentrating on the leading terms in the criterion of minimum π_B-aberration.

Given k vectors $a_1, \ldots, a_k$ where $a_j = (a_{1j}, \ldots, a_{Nj})^{\mathrm{T}}$ for $j = 1, \ldots, k$, the J-characteristic of these vectors is defined as $J(a_1, \ldots, a_k) = \sum_{i=1}^{N} \prod_{j=1}^{k} a_{ij}$ (Tang, 2001). The next result expresses π_2^B and π_2^O in terms of the J-characteristics of columns of a design D. Note that the design matrix D has elements $+1$ and -1 in all columns.

Lemma 1. *Suppose that* $D = (b_1, \ldots, b_{m_1}, o_1, \ldots, o_{m_2})$ *is an orthogonal array of* N *runs for* m_1 *B-factors and* m_2 *O-factors. Then we have that*

$$\pi_2^B = \frac{3}{N^2}\sum_{i<j<k} J^2(b_i, b_j, b_k) + \frac{2}{N^2}\sum_{i<j}\sum_{k} J^2(b_i, b_j, o_k) + \frac{1}{N^2}\sum_{i}\sum_{j<k} J^2(b_i, o_j, o_k) + m_1(m_1-1)$$

and

$$\pi_2^O = \frac{1}{N^2}\sum_{i<j}\sum_{k} J^2(b_i, b_j, o_k) + \frac{2}{N^2}\sum_{i}\sum_{j<k} J^2(b_i, o_j, o_k) + \frac{3}{N^2}\sum_{i<j<k} J^2(o_i, o_j, o_k) + m_1 m_2.$$

The J-characteristics are 0 for any three columns of an $\text{OA}(N, 2^m, 3)$, which exists whenever $m \le N/2$ and a Hadamard matrix of order $N/2$ exists (Cheng, 2014). By Lemma 1, such a design minimizes the bias from two-factor interactions in estimating main effects of B-factors and O-factors. Another implication of Lemma 1 is that switching signs of columns of a design does not affect the values of π_2^B and π_2^O.

For $m > N/2$, we use regular designs to minimize π_2^B and π_2^O. Let the columns of D be selected from a saturated regular design $\text{OA}(2^h, 2^{2^h-1}, 2)$ for some integer h. Such an $\text{OA}(2^h, 2^{2^h-1}, 2)$ can be constructed by first writing down h independent columns $r_1, \ldots, r_h$ that form a full factorial and then adding all possible Hadamard products thereof. We assume that the columns of a regular $\text{OA}(2^h, 2^{2^h-1}, 2)$ are arranged in Yates order. For example, the 15 columns of an $\text{OA}(2^4, 2^{15}, 2)$ are given by

$$(r_1, r_2, r_1r_2, r_3, r_1r_3, r_2r_3, r_1r_2r_3,$$
$$r_4, r_1r_4, r_2r_4, r_1r_2r_4, r_3r_4, r_1r_3r_4, r_2r_3r_4, r_1r_2r_3r_4),$$

where, for example, r_1r_2 denotes the Hadamard product of r_1 and r_2. For experiments involving only O-factors, Chen and Hedayat (1996) showed that a design obtained by taking the last m columns of a regular $\text{OA}(2^h, 2^{2^h-1}, 2)$ minimizes π_2^O among all regular designs. Inspired by their construction, we establish Theorem 2.

Theorem 2. *Suppose R is a regular $\text{OA}(2^h, 2^{2^h-1}, 2)$. Let D_B select the last m_1 columns of R and D_O select the remaining m_2 columns from the last $m = m_1 + m_2$ columns of R that are not already in D_B. Then we have the following results for the design $D = (D_B, D_O)$.*

(i). If m_1 and m satisfy that $m_1 \le 2^h - 2^{h_1}$ and $m \ge 2^h - 2^{h_1}$ for some integer $h_1 \in \{0, 1, \ldots, h-1\}$, then design D minimizes π_2^B over all $\text{OA}(2^h, 2^m, 2)$s and sequentially minimizes π_2^B and π_2^O over all regular $\text{OA}(2^h, 2^m, 2)$s.

(ii). If m satisfies that $m = 2^h - 2^{h_1}$ for some integer $h_1 \in \{0, 1, \ldots, h-1\}$, then D sequentially minimizes π_2^B and π_2^O over all $\text{OA}(2^h, 2^m, 2)$s.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

It is worth remarking that although the constructed design D in Theorem 2 is regular, its optimality properties are established in the whole class of orthogonal arrays in two of the three optimality statements. Specifically, design D minimizes π_2^B over all $\text{OA}(2^h, 2^m, 2)$s in part (i) of Theorem 2, and sequentially minimizes π_2^B and π_2^O over all $\text{OA}(2^h, 2^m, 2)$s in part (ii) of Theorem 2.

The restriction on m_1 and m values in part (i) of Theorem 2 is fairly mild. Because $m \geq N/2 = 2^{h-1}$, we see that the condition is always satisfied so long as $m_1 \leq 2^{h-1}$. Example 2 further illustrates Theorem 2 with a case for $m_1 > 2^{h-1}$.

Example 2. Suppose we would like to study $m_1 = 18$ B-factors and $m_2 = 7$ O-factors with $2^5 = 32$ runs. Then for $h_1 = 3$, we have that $m_1 \leq 32 - 2^{h_1}$ and $m \geq 32 - 2^{h_1}$. Let $D_B = (r_2r_3r_4, r_1r_2r_3r_4, r_5, \ldots, r_1r_2r_3r_4r_5)$ and $D_O = (r_1r_2r_3, r_4, r_1r_4, r_2r_4, r_1r_2r_4, r_3r_4, r_1r_3r_4)$. By Theorem 2, the design $D = (D_B, D_O)$ minimizes π_2^B over all $\text{OA}(32, 2^{25}, 2)$s and sequentially minimizes π_2^B and π_2^O over all regular $\text{OA}(32, 2^{25}, 2)$s.

Remark 1. As careful readers may observe, the results of Theorem 2 hold no matter whether baseline or orthogonal parametrization is used for each factor of the design D. As long as the main effects are divided into two groups and more protection from two-factor interactions is needed for one

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

of the two groups, the results of Theorem 2 are applicable. The existence of two types of factors provides a natural application scenario for these results.

3.3 Main effects of all factors are equally important

If the main effects of the B-factors and the O-factors are of equal interest, then, naturally, one wishes to minimize $\pi_k = \pi_k^B + \pi_k^O$ for $k = 2, \ldots, m$, as π_k measures the contamination of k-factor interactions on the estimation of all main effects. Combined with the effect hierarchy principle, the idea can be formulated as the following minimum π-aberration criterion.

Definition 2. An orthogonal array for m factors is said to have minimum π-aberration if it sequentially minimizes $\pi_2, \pi_3, \ldots, \pi_m$.

Lemma 1 indicates that for a design $D = (d_1, \ldots, d_m)$ of N runs for m factors, we have $\pi_2 = 3A_3 + m_1(m-1)$ where $A_3 = \sum_{i<j<k} J^2(d_i, d_j, d_k)/N^2$ is the leading term in the minimum G_2-aberration criterion. However, for $\pi_3, \pi_4, \ldots, \pi_m$, such a simple connection with the minimum G_2-aberration criterion no longer exists. The expressions of $\pi_3, \pi_4, \ldots, \pi_m$ become more complex as sign-switching columns of D may affect their values.

In the following, we focus on sequential minimization of π_2 and π_3 through the use of regular designs. Consider a regular design D of 2^h runs for a total of $m = 2^h - 2^{h_1}$ factors where h_1 and h are integers. Chen and Hedayat (1996) and Tang and Wu (1996) proved that A_3, and thus π_2, are minimized if and only if columns of D are isomorphic to the last m columns of a saturated regular design. We show that π_3 of such a design D is determined by the J-characteristics of the B-factors alone.

Lemma 2. *Suppose that $D = (b_1, \ldots, b_{m_1}, o_1, \ldots, o_{m_2})$ is a regular* OA$(2^h, 2^m, 2)$ *that minimizes π_2, where $m = 2^h - 2^{h_1}$ for some integer h_1. Then we have that $\pi_3 = c_1 \sum_{i<j<k} J(b_i, b_j, b_k) + c_0$, where c_0 and $c_1 > 0$ are constants.*

Lemma 2 enables us to decide which columns should be assigned to the B-factors and how to switch their signs to minimize π_3. Note that among the last $m = 2^h - 2^{h_1}$ columns of a regular OA$(2^h, 2^{2^h-1}, 2)$, there are $h - h_1$ independent columns $r_{h_1+1}, \ldots, r_h$. Let's arrange these $h - h_1$ columns and all their possible Hadamard products in Yates order. Then let D_B collect the first m_1 columns with their signs all switched, where $m_1 \le 2^{h-h_1} - 1$. Let D_O include the remaining $m - m_1$ columns in the last m columns of the regular OA$(2^h, 2^{2^h-1}, 2)$. Finally, let $D = (D_B, D_O)$. We have the following result for this design D.

Theorem 3. *The design D sequentially minimizes π_2 and π_3 over all regular designs.*

The design D in Theorem 3 can be constructed as long as the total num-

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

ber m of factors satisfies $m = 2^h - 2^{h_1}$ for some integer h_1 and the number m_1 of B-factors satisfies $m_1 \leq 2^{h-h_1} - 1$. In the saturated case of $m = 2^h - 1$, such a design is obtainable for any choice of m_1 and m_2. In particular, if $m_1 = m = 2^h - 1$, then we have $D = (-r_1, -r_2, -r_1 r_2, -r_3, -r_1 r_3, \ldots, -r_1 r_2 r_3 \cdots r_h)$ which must have a row of -1's. Mukerjee and Tang (2012) showed that a saturated orthogonal array has minimum aberration under the baseline parametrization if it contains a run of all baseline levels. Therefore our result is consistent with theirs in this special case.

We illustrate Theorem 3 with an example.

Example 3. Suppose 64 experiments are allowed to examine the main effects of $m_1 = 6$ B-factors and $m_2 = 50$ O-factors. Let $D_B = (-r_4, -r_5, -r_4 r_5, -r_6, -r_4 r_6, -r_5 r_6)$ and $D_O = (r_4 r_5 r_6, r_1 r_4, \ldots, r_1 r_2 r_3 r_4 r_5 r_6)$ which consists of all columns that do not occur in D_B but do occur in the last 56 columns of the regular $\text{OA}(64, 2^{63}, 2)$. According to Theorem 3, the design $D = (D_B, D_O)$ sequentially minimizes π_2 and π_3 over all regular $\text{OA}(64, 2^{56}, 2)$s.

Theorems 2 and 3 provide two theoretical constructions for minimum π_B- and π-aberration designs. These methods have some restrictions on the run size as well as the numbers of B-factors and O-factors. In the next section, we develop efficient algorithms to search for minimum π_B- and

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

π-aberration designs for general cases.

4. Searching designs by algorithms

4.1 A complete search algorithm

Two orthogonal arrays are combinatorially isomorphic if one can be obtained from the other by permuting rows, permuting columns, switching signs of columns, or a combination of these operations (Hedayat et al., 1999). All orthogonal arrays can be generated by applying these operations to a complete set of non-isomorphic orthogonal arrays. Complete sets of non-isomorphic orthogonal arrays are available for small run sizes (Sun et al., 2008; Schoen et al., 2010), which allows us to find minimum π_B- and π-aberration designs over all orthogonal arrays.

When using an $\text{OA}(N, 2^m, 2)$ as a design for m_1 B-factors and m_2 O-factors, there is no need to inspect all isomorphic operations, as many of them lead to designs with the same π_B- or π-aberration. Clearly, permuting rows, permuting the first m_1 columns and permuting the last m_2 columns won't affect the π_B- or π-aberration. In addition, we have the following results on sign-switching columns.

Lemma 3. *Switching the signs of O-factors in an* $\text{OA}(N, 2^m, 2)$ *does not change* π_k^B, π_k^O *and thus* π_k *values for* $k = 2, \ldots, m$.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

Based on the above, we propose the following complete search algorithm for minimum aberration designs. The algorithm used by Mukerjee and Tang (2012) for the baseline parametrization can be seen as a special case where all factors are B-factors.

Step 1: Obtain a complete list of non-isomorphic OA$(N, 2^m, 2)$s.

Step 2: For each OA$(N, 2^m, 2)$ in the list, consider all $m!/(m_1!(m-m_1)!)$ possible ways to assign m_1 columns to the B-factors. The remaining $m_2 = m - m_1$ columns are used for the O-factors.

Step 3: For every possible assignment of B-factors and O-factors in Step 2, switch signs of the m_1 columns of the B-factors in all 2^{m_1} possible ways. Calculate the π_k^B, π_k^O and π_k values for all possible designs.

Note that for the minimum π-aberration criterion, only those OA$(N, 2^m, 2)$s with minimum π_2 values need to be considered in Step 1. We apply this complete search algorithm to obtain minimum π_B- and π-aberration designs of $N = 8$, 12 and 16 runs for all choices of m_1 and m_2, the numbers of B-factors and O-factors. For $N = 20$ runs, the complete search is done for $m \leq 13$. All the obtained designs are available online at `https://github.com/gz-chen/Mixed-Param`.

Suppose there are $q(N, m)$ non-isomorphic $\text{OA}(N, 2^m, 2)$s to be considered in Step 1. Then the total number of designs to be compared in a complete search is $q(N, m)2^{m_1}m!/(m_1!(m-m_1)!)$, which, as N, m and m_1 increases, soon becomes too large for computer to handle, not to mention that the computation of J-characteristics also grows rapidly and that complete sets of non-isomorphic orthogonal arrays are no longer available for large designs. Therefore, it is necessary to come up with an efficient algorithm for the cases where the complete search is impossible.

4.2 An algorithm based on minimum G_2-aberration designs

The aim of this subsection is to conduct an algorithmic search for large designs that perform well under the minimum π_B- or π-aberration criterion. To achieve this, several measures are taken to reduce the computation. The first is to focus on orthogonal arrays with minimum G_2-aberrations instead of all non-isomorphic ones in Step 1 of the complete search algorithm.

An $\text{OA}(N, 2^m, 2)$, say $D = (d_1, \ldots, d_m)$, is said to have minimum G_2-aberration if it sequentially minimizes $A_3, A_4, \ldots, A_m$, where $A_k = \sum_{j_1<\cdots<j_k} J^2(d_{j_1}, \ldots, d_{j_k})/N^2$ for $k = 3, \ldots, m$. As mentioned in Section 3.3, a minimum G_2-aberration design minimizes π_2 in the minimum π-aberration criterion. The next result shows that such a design is also promising in se-

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

quentially minimizing higher-order terms π_k for $k = 3, \ldots, m$ and entries in the minimum π_B-aberration criterion.

Theorem 4. *Suppose the B-factors of a design are generated by randomly selecting and sign-switching m_1 columns of an* $\text{OA}(N, 2^m, 2)$ *and the O-factors are given by the remaining columns. Let $\bar{\pi}_k$ be the average of π_k's over all possible designs generated in this way. Then, for $k = 2, \ldots, m$, we have*

$$\bar{\pi}_k = c_{k+1}^{(k)} A_{k+1} + c_k^{(k)} A_k + \cdots + c_3^{(k)} A_3 + c_0^{(k)},$$

where $c_0^{(k)}, c_3^{(k)}, \ldots, c_{k+1}^{(k)}$ are positive constants, $A_3, \ldots, A_m$ are determined by the $\text{OA}(N, 2^m, 2)$ *and we define $A_{m+1} = 0$. Similar results also hold for π_k^B and π_k^O.*

Theorem 4 provides a rationale for the use of minimum G_2-aberration designs in Step 1 of the complete search algorithm. Related to Theorem 4 is a result of Xiao and Xu (2018) who justified the use of generalized minimized aberration designs in generating space-filling designs.

Next, we improve the efficiency of Steps 2 and 3 of the complete search algorithm through a local search algorithm (Aarts and Lenstra, 2003). The idea is to iteratively replace a current design with the best one in a small neighbourhood of the current design, until no further improvement can be

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

made. A full description of our algorithm for minimum π-aberration designs is given below.

Step 1: Obtain a minimum G_2-aberration design from a list of $\text{OA}(N, 2^m, 2)$s. Randomly permute and sign-switch its columns. Denote this design by $D = (b_1, \ldots, b_{m_1}, o_1, \ldots, o_{m_2})$ and calculate $\pi = (\pi_2, \ldots, \pi_m)$ for D.

Step 2: Exchange a column b_j ($j = 1, \ldots, m_1$) and a column $\pm o_k$ ($k = 1, \ldots, m_2$). Among all $2m_1m_2$ designs generated this way, continue to the next step if none of them improves π; otherwise select one with the least π-aberration, denote it by D and update π. Then repeat this step.

Step 3: Exchange a column pair (b_{j_1}, b_{j_2}) ($1 \le j_1 < j_2 \le m_1$) and a column pair $(\pm o_{k_1}, \pm o_{k_2})$ ($1 \le k_1 < k_2 \le m_2$). Among all $m_1m_2(m_1 - 1)(m_2 - 1)$ designs generated this way, continue to the next step if none of them improves π; otherwise select one with the least π-aberration, denote it by D and update π. Then go back to Step 2.

Step 4: Replace a column b_j by $-b_j$ ($j = 1, \ldots, m_1$). Among all m_1 designs generated this way, continue to the next step if none of them improves π; otherwise select one with the least π-aberration, denote it by D and update π. Then repeat this step.

Step 5: Replace a column pair (b_{j_1}, b_{j_2}) by $(-b_{j_1}, -b_{j_2})$ $(1 \le j_1 < j_2 \le m_1)$. Among all $m_1(m_1-1)/2$ designs generated this way, continue to the next step if none of them improves π; otherwise select one with the least π-aberration, denote it by D and update π. Then go back to Step 4.

Step 6: Output the design D and the associated vector $\pi = (\pi_2, \ldots, \pi_m)$.

The algorithm above generalizes that for the baseline parametrization presented in Li et al. (2014). One can replace the vector $\pi = (\pi_2, \ldots, \pi_m)$ in the algorithm by $\pi_B = (\pi_2^B, \pi_2^O, \ldots, \pi_m^B, \pi_m^O)$ if a minimum π_B-aberration design is the goal. If there is more than one minimum G_2-aberration design in Step 1, then we can apply the algorithm to all those designs and then find the best output design.

To evaluate the performance of our algorithm, we apply it to 20-run designs for 13 factors. There are 730 non-isomorphic $\text{OA}(20, 2^{13}, 2)$s in total; five of them have weak minimum G_2-aberration with $A_3 = 15.92$; and three of them have minimum G_2-aberration with $A_4 = 43.64$ and $A_5 = 62.4$, while the other two weak minimum G_2-aberration designs have $A_4 = 43.64$ and $A_5 = 62.56$. Therefore in a complete search, we search 730 orthogonal arrays for minimum π_B-aberration designs and 5 orthogonal arrays for minimum π-aberration designs, whereas in the incomplete search we focus on the 3 minimum G_2-aberration designs. For each case of the number of B-factors $m_1 = 1, \ldots, 13$, we run the incomplete search algorithm 200 times for minimum π_B- and π-aberration designs separately and compare the results with those obtained from the complete search.

Under the minimum π_B-aberration criterion, we are surprised to find that all the designs obtained by the incomplete search algorithm sequentially minimize the leading terms π_2^B and π_2^O among all orthogonal arrays. So we move on to the next term and compare the 200 π_3^B values in the incomplete search with all the π_3^B values of orthogonal arrays that have sequentially minimized π_2^B and π_2^O. For each $m_1 = 1, \ldots, 13$, the distributions of these two sets of π_3^B values can be described by two boxplots, as shown in the left panel of Figure 1. It can be seen that the π_3^B values from the incomplete search are all centered near the minimum π_3^B values from the complete search. In Table 1, we provide the minimum and maximum π_3^B values found by our incomplete search algorithm, as well as proportions of π_3^B values in the complete search that are no less than these values. It can be seen that in many cases the best designs from the incomplete search algorithm attain the minimum π_3^B values. When the algorithm cannot find

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

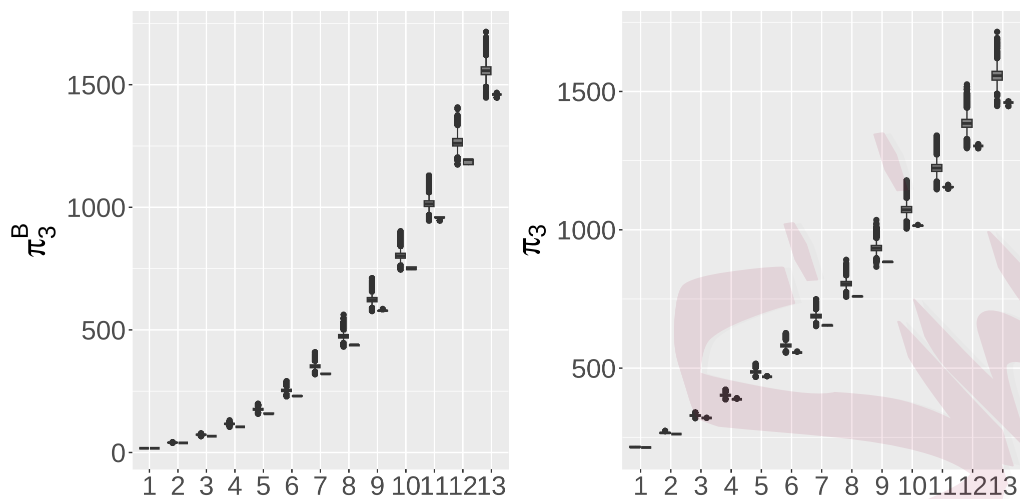

Figure 1: The π_3^B and π_3 values obtained by 200 incomplete searches and the complete search, where the x-axis represents the numbers of B-factors. For each $m_1 = 1, \ldots, 13$, the left and right boxplots show the values from the complete and incomplete searches, respectively.

the optimal designs, even the worst designs found by the algorithm have good performance in terms of the π_3^B values, as the proportions of designs beaten by them in the complete search are close to 100%. Similar observations on π_3 values can also be made from the searching results for minimum π-aberration designs, as presented in the right panel of Figure 1 and Table 2.

These empirical results demonstrate that our incomplete search algo-

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

Table 1: The range of π_3^B values obtained in the complete and incomplete search for minimum π_B-aberration designs. For each $m_1 = 1, \ldots, 13$, the two percentages are the proportions of OA$(20, 2^{13}, 2)$s that are no better than the best and worst designs found by the incomplete search.

	Complete search	Incomplete search			Complete search	Incomplete search	
m_1	π_3^B values	min π_3^B	max π_3^B	m_1	π_3^B values	min π_3^B	max π_3^B
1	[17.2, 17.2]	17.2 (100%)	17.2 (100%)	8	[431.36, 561.92]	432.64 (100%)	443.36 (99.953%)
2	[38.96, 42.16]	39.12 (99.697%)	39.12 (99.697%)	9	[577.04, 711.12]	577.04 (100%)	585.68 (99.919%)
3	[66.44, 77.96]	66.44 (100%)	66.44 (100%)	10	[746, 902.32]	746 (100%)	758.48 (99.828%)
4	[104.64, 131.52]	104.64 (100%)	104.64 (100%)	11	[946.12, 1129.32]	946.12 (100%)	961.8 (99.738%)
5	[157.76, 198.72]	157.76 (100%)	160.8 (99.989%)	12	[1174.08, 1407.68]	1174.08 (100%)	1197.12 (99.775%)
6	[228.72, 290.48]	228.72 (100%)	234.96 (99.934%)	13	[1447.52, 1715.52]	1447.52 (100%)	1467.04 (99.824%)
7	[318.84, 408.92]	319.8 (99.999%)	325.88 (99.967%)				

Table 2: The range of π_3 values obtained in the complete and incomplete search for minimum π-aberration designs. For each $m_1 = 1, \ldots, 13$, the two percentages are the proportions of OA$(20, 2^{13}, 2)$s that are no better than the best and worst designs found by the incomplete search.

	Complete search	Incomplete search			Complete search	Incomplete search	
m_1	π_3 values	min π_3	max π_3	m_1	π_3 values	min π_3	max π_3
1	[210.32, 217.04]	213.2 (98.462%)	213.2 (98.462%)	8	[758.4, 891.52]	758.4 (100%)	760.64 (99.996%)
2	[259.8, 272.92]	261.72 (98.462%)	261.72 (98.462%)	9	[866.48, 1035.44]	882.48 (100%)	886.32 (99.994%)
3	[319.08, 339.56]	319.88 (99.965%)	320.36 (99.528%)	10	[1004.6, 1178.36]	1014.2 (99.998%)	1017.56 (99.984%)
4	[387.28, 422]	387.28 (100%)	391.12 (99.633%)	11	[1146.92, 1340.68]	1148.52 (99.999%)	1162.6 (99.859%)
5	[468.64, 515.2]	468.8 (99.999%)	471.04 (99.863%)	12	[1295.12, 1525.52]	1295.12 (100%)	1309.68 (99.862%)
6	[556.36, 626.6]	556.36 (100%)	560.68 (99.93%)	13	[1447.52, 1715.52]	1447.52 (100%)	1464.16 (99.866%)
7	[652.28, 748.76]	652.28 (100%)	657.08 (99.981%)				

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

rithm can be used to obtain designs that perform well under the minimum π_B- or π-aberration criterion. We apply this algorithm to 20-run designs with more than 13 factors under the both criteria. All findings are available at `https://github.com/gz-chen/Mixed-Param`.

5. Concluding remarks

In this paper, we concern ourselves with the estimation of main effects. However, in some situations, we may also wish to estimate a few two-factor interactions in addition to the main effects. When it is uncertain which two-factor interactions are active, our incomplete search algorithm based on minimum G_2-aberration designs is still useful from the viewpoint of model efficiency, as it can be justified as follows. Let $\mathcal{F}$ collect certain f subsets of size two of $S = \{1, \ldots, m\}$, and $\xi_{\mathcal{F}}$ be the set of factorial effects ξ_ϕ, ξ_j's for $j = 1, \ldots, m$ and ξ_w's for $w \in \mathcal{F}$. Consider the model

$$Y = X_{\mathcal{F}}\xi_{\mathcal{F}} + \epsilon, \tag{5.6}$$

where $X_{\mathcal{F}}$ is the model matrix corresponding to $\xi_{\mathcal{F}}$ for the design D. Then the D-efficiency of design D under model (5.6) is given by $\det(X_{\mathcal{F}}^{\mathrm{T}} X_{\mathcal{F}})$. On the other hand, assume the orthogonal parametrization is used for all factors and consider the model $Y = Z_{\mathcal{F}}\beta_{\mathcal{F}} + \epsilon$, where $\beta_{\mathcal{F}}$ collects β_ϕ, β_j's for $j = 1, \ldots, m$ and β_w's for $w \in \mathcal{F}$ and $Z_{\mathcal{F}}$ is the corresponding model

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0165

matrix. Then we have the following result.

Proposition 1. *We have* $\det(X_{\mathcal{F}}^{\mathrm{T}} X_{\mathcal{F}}) = \det(Z_{\mathcal{F}}^{\mathrm{T}} Z_{\mathcal{F}})$.

Cheng et al. (2002) showed that when f is small, the minimum G_2-aberration criterion is a good surrogate for maximizing $E\{\det(Z_{\mathcal{F}}^{\mathrm{T}} Z_{\mathcal{F}})\}$, where the expectation $E\{\cdot\}$ is taken over all possible $\mathcal{F}$. The result of Proposition 1 implies that minimum G_2-aberration designs should also perform well in maximizing $E\{\det(X_{\mathcal{F}}^{\mathrm{T}} X_{\mathcal{F}})\}$. Therefore, the designs obtained by our incomplete search algorithms, which must be minimum G_2-aberration designs themselves, should allow efficient estimation of main effects and a few two-factor interactions, at least when averaging over all possible $\mathcal{F}$.

When the prior knowledge as to which two-factor interactions are active is available, it is preferable to use a design that entertains the estimation of these active effects. To address this problem under the baseline parametrization, Chen et al. (2021) carried out an algorithmic search for non-isomorphic models with up to 3 two-factor interactions. For the mixed parametrization, one needs to additionally take care of the type of two-factor interactions (say, B×B, B×O or O×O), which makes it more complicated to enumerate all possibilities. Nevertheless, the algorithm of Chen et al. (2021) can easily be easily modified for the mixed parametrization and used to search for an efficient design in practical applications.

There are several other possible directions for future research. First, all the designs considered in this paper are orthogonal arrays, because, as shown in Corollary 2, they are optimal under the main-effects model. On the other hand, under the baseline parametrization, Mukerjee and Tang (2012) showed that one-factor-at-a-time designs may be more desirable when the biases of the main effect estimators dominate their variances. It is interesting to investigate for the mixed parametrization how to obtain designs suitable for these situations.

Stallings and Morgan (2015) developed a weighted optimality theory which allows variable interests in different estimable functions. When the main effects of B-factors are more important, one possible approach is to apply their framework by placing greater weights on the estimation of main effects of B-factors under a model with interaction terms. This approach is different from the one adopted in this paper, which is to find an orthogonal array that protects the main effects of B-factors from the contamination of potential two-factor interactions. The problem as to how the resulting optimal designs are related to the designs studied in this paper is worthy of future research.

Supplementary Material

Supplementary material available online includes all the proofs of theoretical results in this paper and all minimum π_B- and π-aberration designs of 8 and 12 runs.

Acknowledgments

This research is supported by the Natural Sciences and Engineering Research Council of Canada. Most of the work was carried out while the first author was a PhD candidate at Simon Fraser University.

References

Aarts, E. and J. K. Lenstra (Eds.) (2003). *Local Search in Combinatorial Optimization.* Princeton, New Jersey: Princeton University Press.

Banerjee, T. and R. Mukerjee (2008). Optimal factorial designs for cDNA microarray experiments. *Ann. Appl. Stat. 2*(1), 366–385.

Box, G. E. P. and J. S. Hunter (1961). The 2^{k-p} fractional factorial designs. *Technometrics 3*, 311–351.

Chen, A., C.-Y. Sun, and B. Tang (2021). Selecting baseline designs using a minimum aberration criterion when some two-factor interactions are important. *Stat. Theory Relat. Fields 5*(2), 95–101.

Chen, H. and A. S. Hedayat (1996). 2^{n-l} designs with weak minimum aberration. *Ann. Statist. 24*(6), 2536–2548.

Cheng, C.-S. (2014). *Theory of Factorial Design: Single- and Multi-Stratum Experiments*. Boca Raton: CRC Press.

Cheng, C.-S., L.-Y. Deng, and B. Tang (2002). Generalized minimum aberration and design efficiency for nonregular fractional factorial designs. *Statist. Sinica 12*(4), 991–1000.

Fries, A. and W. G. Hunter (1980). Minimum aberration 2^{k-p} designs. *Technometrics 22*(4), 601–608.

Glonek, G. F. V. and P. J. Solomon (2004). Factorial and time course designs for cDNA microarray experiments. *Biostatistics 5*(1), 89–111.

Hedayat, A. S., N. J. A. Sloane, and J. Stufken (1999). *Orthogonal Arrays: Theory and Applications*. Springer Series in Statistics. New York: Springer-Verlag.

Li, P., A. Miller, and B. Tang (2014). Algorithmic search for baseline minimum aberration designs. *J. Statist. Plann. Inference 149*, 172–182.

Mee, R. (2009). *A Comprehensive Guide to Factorial Two-Level Experimentation*. New York: Springer Science & Business Media.

Mukerjee, R. and B. Tang (2012). Optimal fractions of two-level factorials under a baseline parameterization. *Biometrika 99*(1), 71–84.

Schoen, E. D., P. T. Eendebak, and M. V. M. Nguyen (2010). Complete enumeration of purelevel and mixed-level orthogonal arrays. *J. Combin. Des. 18*(2), 123–140.

Stallings, J. W. and J. P. Morgan (2015). General weighted optimality of designed experiments. *Biometrika 102*(4), 925–935.

Sun, C.-Y. and B. Tang (2022). Relationship between orthogonal and baseline parameterizations and its applications to design constructions. *Statist. Sinica 32*(1), 239–250.

Sun, D. X., W. Li, and K. Q. Ye (2008). Algorithmic construction of catalogs of non-isomorphic two-level orthogonal designs for economic run sizes. *Statist. Appl. 6*, 141–155.

Tang, B. (2001). Theory of J-characteristics for fractional factorial designs and projection justification of minimum G_2-aberration. *Biometrika 88*(2), 401–407.

Tang, B. and L.-Y. Deng (1999). Minimum G_2-aberration for nonregular fractional factorial designs. *Ann. Statist. 27*(6), 1914–1926.

Tang, B. and C. F. J. Wu (1996). Characterization of minimum aberration 2^{n-k} designs in terms of their complementary designs. *Ann. Statist. 24*(6), 2549–2559.

Wu, C. F. J. and M. S. Hamada (2021). *Experiments: Planning, Analysis, and Optimization, 3rd ed.* Hoboken, New Jersey: Wiley.

Xiao, Q. and H. Xu (2018). Construction of maximin distance designs via level permutation and expansion. *Statist. Sinica 28*(3), 1395–1414.

Yang, Y. H. and T. Speed (2002). Design issues for cDNA microarray experiments. *Nature Genet. 3*(8), 579–588.

doi:10.5705/ss.202023.0165

Guanzhou Chen

School of Statistics and Data Science, LPMC & KLMDASR, Nankai University

E-mail: guanzhou.chen@foxmail.com

Boxin Tang

Department of Statistics and Actuarial Science, Simon Fraser University

E-mail: boxint@sfu.ca