Available online at www.sciencedirect.com

ScienceDirect

Journal of Complexity 23 (2007) 740–751

Journal of COMPLEXITY

www.elsevier.com/locate/jco

An effective algorithm for generation of factorial designs with generalized minimum aberration

Kai-Tai Fang[a,*], Aijun Zhang[b], Runze Li[c]

[a] *BNU - HKBU United International College, Zhuhai Campus of Beijing Normal University, Jinfeng Road, Zhuhai, 519085, China*

[b] *Department of Statistics, University of Michigan, Ann Arbor, MI 48105, USA*

[c] *Department Statistics, Penn State University, University Park, PA 16802, USA*

Received 28 September 2006; accepted 16 March 2007
Available online 4 May 2007

Abstract

Fractional factorial designs are popular and widely used for industrial experiments. Generalized minimum aberration is an important criterion recently proposed for both regular and non-regular designs. This paper provides a formal optimization treatment on optimal designs with generalized minimum aberration. New lower bounds and optimality results are developed for resolution-III designs. Based on these results, an effective computer search algorithm is provided for sub-design selection, and new optimal designs are reported.

Keywords: Fractional factorial design; Generalized minimum aberration; Lagrange analysis; Sub-design selection

1. Introduction

Fractional factorial designs (FFDs) are popular choices of designs of experiments in industry. Extensive research has been done on factorial designs in recent decades, with main focus on optimality theory and design construction. The two most successful optimality criteria are *maximum resolution* by Box and Hunter [2] and *minimum aberration* by Fries and Hunter [7]. However, these criteria are defined for *regular* designs only; they cannot be used to assess a factorial design in general. Recently, *generalized minimum aberration* (GMA) was proposed for both regular and non-regular designs, with the two-level case by Tang and Deng [14], and multi-level case by Ma

* Corresponding author.
E-mail addresses: ktfang@uic.edu.hk (K.-T. Fang), ajzhang@umich.edu (A. Zhang), rli@stat.psu.edu (R. Li).

0885-064X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jco.2007.03.010

and Fang [10] and Xu and Wu [17]. More background on FFD and GMA will be presented in the next section.

This paper is to study the optimal conditions for GMA designs, which is a non-trivial problem due to the sequential optimization nature of the criterion. Unlike conventional combinatorial approaches, we provide a formal treatment for resolution-III designs from the optimization perspective. It will be shown that our new optimality results can be viewed as a natural extension of the weak-equidistance optimality for resolution-II designs by Zhang et al. [18]. Here we restrict ourselves to the symmetrical designs, i.e. those in which all the factors take the same number of levels, while the methodology is readily generalized to mixed-level designs in a straightforward manner.

There exist several approaches to the construction of GMA designs. Among others, Lin [8] proposed to use half fractions of Hadamard matrices for constructing two-level supersaturated designs (SSDs), and Fang et al. [5] proposed the RBIBD method for constructing multi-level SSDs. However, these construction methods are restricted to GMA designs of resolution II. For designs of resolution III or higher, it is most natural to consider the subset design approach based on existing classes of orthogonal arrays. Butler [3] obtained some GMA designs by projecting specific saturated orthogonal arrays. In this paper, we propose a general sub-design selection algorithm, which utilizes the newly developed lower bounds and optimality conditions.

The paper is organized as follows. Some background material is presented in Section 2. In Section 3, new lower bounds and optimality results are developed for orthogonal FFDs of resolution III, by Lagrange analysis for the nonlinear programming problem and a strengthening technique to take into account the integer-valued condition. These results are applied in Section 4 for sub-design selection, where we provide an effective computer search algorithm and report some new optimal designs with GMA. In the final section, we discuss some possible routes for future works.

Throughout the paper, we use the symbol $\lfloor x \rfloor$ for representing the largest integer not exceeding x, $\lceil x \rceil$ for the smallest integer not less than x and $\langle x \rangle = x - \lfloor x \rfloor$ for the fractional part of x. The Kronecker delta function is defined as $\delta(x, y) = 1$ if $x = y$ and 0 otherwise. For non-negative integers j and k, $S(j, k)$ denotes the Stirling number of the second kind, i.e. the number of ways of partitioning a set of j elements into k non-empty sets. Furthermore, we extend the definition of binomial coefficient function $\binom{x}{j}$ to cover any non-negative argument $x \in \mathbb{R}_+ : \binom{x}{j} = 1$ if $j = 0$, $\binom{x}{j} = 0$ if $x < j$ and $\binom{x}{j} = \frac{x(x-1)\cdots(x-j+1)}{j!}$ otherwise.

2. Background

A factorial design of n runs and s factors for which each factor takes q levels is denoted by $D(n, q^s)$. The full factorial design with $n = q^s$ runs comprises all possible level combinations. An FFD with $n \ll q^s$ runs takes only some fraction of the runs required for the full factorial; see Wu and Hamada [15] for details. A particular FFD is often chosen to satisfy some constraint or optimize some condition or set of conditions. Two common conditions of interest are balance and orthogonality. Balance means that for each factor each level appears in the same number of runs. Orthogonality means that for each pair of factors, all the q^2 possible level-combinations appear equally often. Two designs are said to be isomorphic if one can be obtained from the other by re-ordering runs, permuting factors or switching levels of one or more factors. For given parameters (n, s, q), we use $\mathcal{D}(n, q^s), \mathcal{U}(n, q^s)$ and $\mathcal{L}(n, q^s)$ to denote the sets of non-isomorphic designs that have no constraint, balanced constraint and orthogonality constraint, respectively. A design in the set of $\mathcal{U}(n, q^s)$ is also called a U-type design in the uniform design literature; see

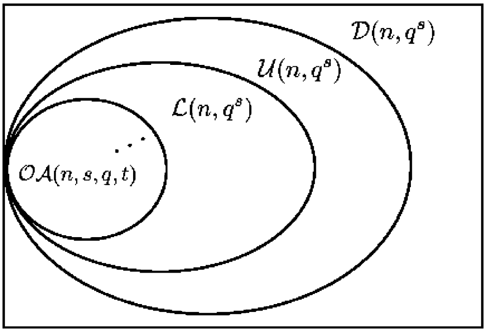

Fig. 1. Nested classes of fractional factorial designs.

the recent monograph by Fang et al. [6]. Note that these sets of designs are nested by $\mathcal{D}(n, q^s) \supset \mathcal{U}(n, q^s) \supset \mathcal{L}(n, q^s)$.

The column-wise study of fractional factorials is tightly connected with the notion of orthogonal arrays. An FFD $D(n, q^s)$ can be viewed as an orthogonal array of strength t, often denoted by $\mathrm{OA}(n, s, q, t)$, if for each t-tuple of factors each level combination appears equally often. Similarly, let us use $\mathcal{OA}(n, s, q, t)$ to denote the set of non-isomorphic orthogonal arrays, where $\mathcal{OA}(n, s, q, 1) \equiv \mathcal{U}(n, q^s)$ and $\mathcal{OA}(n, s, q, 2) \equiv \mathcal{L}(n, q^s)$. For given (n, s, q), an illustration of the nested structure is given in Fig. 1. Rao [13] presented the following well-known conditions for the existence of $\mathrm{OA}(n, s, q, t)$:

$$n \geqslant \begin{cases} \sum_{i=0}^{u} \binom{s}{i}(q-1)^i & \text{if } t = 2u \\ \sum_{i=0}^{u} \binom{s}{i}(q-1)^i + \binom{s-1}{u}(q-1)^{u+1} & \text{if } t = 2u+1. \end{cases} \tag{1}$$

These general lower bounds of n for given (s, q, t) are called Rao's bounds. They have been improved for many specific parameter settings, see e.g. Bose and Bush [1] and Mukerjee and Wu [12].

The row-wise study of factorial designs is tightly connected with the notion of error-correcting codes in MacWilliams and Sloane [11]. It leads to the definition of the generalized minimum aberration (GMA) criterion. Let δ_{ik}^{j} be the coincidence indicator between the ith and kth runs at the jth factor. For any $1 \leqslant i, k \leqslant n$, the (i, k)-coincidence of the design is defined as $\beta_{ik} = \sum_{j=1}^{s} \delta_{ik}^{j}$. For an $\mathrm{OA}(n, s, q, t)$, Bose and Bush [1] derived the following necessary conditions on the coincidences:

$$\sum_{k=1}^{n} \binom{\beta_{ik}}{j} = \frac{n}{q^j}\binom{s}{j} \quad \text{for } 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant t. \tag{2}$$

On the other hand, Zhang et al. [18] proved that for any $D(n, q^s)$,

$$\sum_{i=1}^{n} \left(\sum_{k=1}^{n} \binom{\beta_{ik}}{j} - \frac{n}{q^j}\binom{s}{j} \right) = \sum_{u}\sum_{v} \left(N_v^{(u)} - \frac{n}{q^j} \right)^2 \quad \text{for } 1 \leqslant j \leqslant s, \tag{3}$$

where $\sum_u$ denotes the summation over all j-element subset of $\{1, \ldots, s\}$, $\sum_v$ denotes the summation over q^j j-tuple level-combinations, and $N_v^{(u)}$ is the frequency of level-combination v appearing in a u-factor sub-design. It is implied that D is an $\mathrm{OA}(n, s, q, t)$ if the right-hand side of (3) vanishes for all $j \leqslant t$. Thus, it is clear that the Bose–Bush identities (2) are also sufficient conditions for $D(n, q^s)$ to have orthogonal strength t. Besides, for a design $D(n, q^s)$ that is *saturated* in the sense that $n = 1 + s(q-1)$, [12] derived that

$$\beta_{ik} = s - n/q \quad \text{for any } 1 \leqslant i < k \leqslant n, \tag{4}$$

a useful property in many ways, with a typical example in the study of complementary designs.

Maximum resolution and minimum aberration are well-known criteria for regular designs. They are extended to non-regular designs via row-wise coincidences and MacWilliams identities in coding theory, as noted by Ma and Fang [10] and Xu and Wu [17]. For any FFD $D(n, q^s)$, define

$$A_j(D) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} \sum_{w=0}^{j} (-1)^w (q-1)^{j-w} \binom{s-\beta_{ik}}{w} \binom{\beta_{ik}}{j-w} \tag{5}$$

for $1 \leqslant j \leqslant s$. The vector $\mathbf{w}(D) = (A_1(D), \ldots, A_s(D))$ is called the generalized word-length pattern (GWP), and the index of the first non-zero element corresponds to the resolution. For two designs, D_1 is said to have less generalized aberration than D_2 if the first non-zero element of $\mathbf{w}(D_1) - \mathbf{w}(D_2)$ is negative. A design D_* is said to have GMA if no other design has less generalized aberration than it.

3. Lower bounds and optimality results

For given parameters (n, s, q), the GMA criterion tends to sequentially minimize the GWP from low to high orders, such that the selected designs not only have the maximum resolution (say r), but also have the smallest A_r-value. Furthermore, if there are multiple resolution-r designs with the same smallest A_r, the GMA criterion will sequentially reduce the set of candidates by minimizing $A_{r+1}(D), A_{r+2}(D), \ldots,$ until the resulting candidates all have the same optimal GWP. The lower bounds of $A_j(D)$ for $j = 1, 2, 3, \ldots$ are of crucial importance in the search of GMA designs.

Delsarte [4] derived that $A_j(D) \geqslant 0$ for $1 \leqslant j \leqslant s$, where the equality holds for all $j \leqslant t$ if and only if $D(n, q^s)$ is an orthogonal array of strength t. Consider the set of candidate designs

$$D \in \mathcal{S}_t \equiv \mathcal{OA}(n, s, q, t) \setminus \mathcal{OA}(n, s, q, t+1),$$

for which $A_1(D) = \cdots = A_t(D) = 0$ while $A_{t+1}(D) > 0$. Schematically in Fig. 1, $\mathcal{S}_t$ for $t = 1, 2, \ldots$ represent the resolution-$(t+1)$ rings from outer to inner areas. For $D \in \mathcal{S}_t$, there is lack of a tight lower bound for $A_{t+1}(D)$.

This section presents lower bounds and optimality results for $\mathcal{S}_t$ with $t = 1$ and 2. We begin with a brief review of the weak-equidistant optimality for $\mathcal{S}_1$, as obtained by [18] through majorization inequality. Then, we develop a general treatment for $\mathcal{S}_2$ from a formal optimization perspective. It will be shown that the optimality results for $\mathcal{S}_2$ can be viewed as a natural extension of those for $\mathcal{S}_1$.

3.1. Balanced designs of resolution II

Zhang et al. [18] provides the optimality results for U-type designs $D \in \mathcal{S}_1(n, q^s)$, namely, the weak-equidistance lower bounds:

$$A_2(D) \geqslant \frac{q^2}{2n}((n-1)\theta(\theta + 2\gamma - 1) + s(s-1)(1 - n/q^2)), \tag{6}$$

where $\theta = \lfloor \mu_0 \rfloor$ and $\gamma = \langle \mu_0 \rangle$, based on the average of pairwise coincidences

$$\mu_0 = \frac{1}{\binom{n}{2}} \sum_{1 \leqslant i<k \leqslant n} \beta_{ik} = \frac{s(n-q)}{q(n-1)}, \tag{7}$$

as a consequence of (2) when $j = 1$. Recall that a design is called *equidistant* if any two distinct runs of the design have the same coincidence μ_0, where μ_0 must be a positive integer. The lower bounds (6) are achieved if there exist weak-equidistant designs, i.e. designs whose pairwise coincidences satisfy

$$\beta_{ik} = \begin{cases} \theta & \text{with proportion } (1-\gamma), \\ \theta + 1 & \text{with proportion } \gamma \text{ for } 1 \leqslant i < k \leqslant n. \end{cases} \tag{8}$$

For $\gamma = 0$, this reduces to the equidistant case. Such weak-equidistant designs have been shown to have GMA in $\mathcal{S}_1(n, q^s)$. More details can be referred to [9,16,18] in the context of optimal SSDs.

3.2. Orthogonal designs of resolution III

For orthogonal FFDs in $\mathcal{S}_2(n, q^s)$ that have resolution III, the necessary conditions (2) imply both the mean type of constraint (7) and the following variance type of constraint on the pairwise coincidences:

$$\frac{1}{\binom{n}{2}} \sum_{1 \leqslant i<k \leqslant n} \left(\beta_{ik} - \mu_0\right)^2 = \frac{ns(q-1)(n-1-s(q-1))}{q^2(n-1)^2} \equiv \sigma_0^2. \tag{9}$$

Our goal is to minimize $A_3(D)$ for resolution-III designs $D \in \mathcal{S}_2(n, q^s)$. Since the original formulas (5) for $A_j(D)$ involve Krawtchouk polynomials which are difficult to analyze, we employ a reformulation through power moments. A similar consideration appears in [16], who, however, employs an overly complicated formulation and proof. Here we need only a basic property of Stirling numbers $S(j, l)$ of the second kind defined by

$$x^j = \sum_{l=1}^{j} S(j, l) x(x-1)\cdots(x-l+1) \quad \text{(for any real } x\text{)},$$

as well as a recent result by [18]

$$\sum_{1 \leqslant i,k \leqslant n} \binom{\beta_{ik}}{l} = \frac{n^2}{q^l} \sum_{w=1}^{l} \binom{s-w}{l-w} A_w(D) + \binom{s}{l} \left(\frac{n^2}{q^l} - n\right)$$

for $l = 1, \ldots, s$. Then, it is straightforward to obtain the following:

$$\sum_{1 \leqslant i<k \leqslant n} \beta_{ik}^j = \alpha_j + \sum_{l=1}^{j} \eta_{jl} \sum_{w=1}^{l} \binom{s-w}{l-w} A_w(D) \quad \text{for } j = 1, \ldots, s, \tag{10}$$

where $\eta_{jl} = \frac{n^2 S(j,l) l!}{2q^l}$ and $\alpha_j = \sum_{l=1}^{j} \frac{S(j,l) s!}{2(s-l)!} \left(\frac{n^2}{q^l} - n\right)$. Since for each j the leading coefficient for $A_j(D)$ is evaluated as $\frac{n^2 j!}{2q^j} > 0$, it is clear that sequentially minimizing $\sum_{i<k} \beta_{ik}^j$ for $j = 1, 2, \ldots$ is equivalent to the GMA criterion in assessing and selecting designs.

Now consider the optimization problem

$$\min_{D\in\mathcal{S}_2(n,q^s)} \sum_{i<k} \beta_{ik}^3(D) \tag{11}$$

over pairwise coincidences $\boldsymbol{\beta}(D)=\{\beta_{ik}(D), 1\leqslant i<k\leqslant n\}$. Besides the mean and variance types of constraints (7) and (9), we have a boundary constraint for each β_{ik} based on the existence of OA$(n,s,q,2)$:

$$\max\{0, s-n/q\}\leqslant \beta_{ik}\leqslant s \quad \text{for } 1\leqslant i<k\leqslant n, \tag{12}$$

where the second inequality is obvious. The first inequality can be verified from (4) if $D(n,q^s)$ can be embedded in a saturated design; otherwise one can refer to [3] for an alternative proof.

In what follows we concentrate on the derivation of the lower bounds for (11) subject to the conditions (7), (9) and (12), as divided into two steps.

3.2.1. Step 1: Lagrange analysis

We use the classical Lagrange analysis by assuming that each argument β_{ik} is a *continuous* variable falling in the interval given by (12). The Lagrangian function takes the form

$$L(\boldsymbol{\beta};\lambda_1,\lambda_2)=\sum_{i<k}\beta_{ik}^3-\lambda_1\sum_{i<k}(\beta_{ik}-\mu_0)-\lambda_2\sum_{i<k}(\beta_{ik}^2-\mu_0^2-\sigma_0^2),$$

where λ_1 and λ_2 are undetermined Lagrange multipliers. It is necessary for $\hat{\boldsymbol{\beta}}$ to be an optimal solution that there exist $\hat{\lambda}_1$ and $\hat{\lambda}_2$ such that the gradient

$$\nabla L(\hat{\boldsymbol{\beta}};\hat{\lambda}_1,\hat{\lambda}_2)=\mathbf{0}. \tag{13}$$

Any root $\hat{\boldsymbol{\beta}}$ with zero gradient is said to be stationary, which corresponds to a minimum, maximum or saddle point.

The equation (13) leads to

$$\hat{\beta}_{ik}=\frac{1}{3}\left(-\hat{\lambda}_2\pm\sqrt{\hat{\lambda}_2^2+3\hat{\lambda}_1}\right) \quad \text{for } 1\leqslant i<k\leqslant n.$$

Denote by $\hat{\beta}_a$ and $\hat{\beta}_b$ the two undetermined roots and assume $\hat{\beta}_a\leqslant\hat{\beta}_b$. Let $p\in(0,1)$ be the proportion of $\hat{\beta}_b$ that appears in $\hat{\boldsymbol{\beta}}$. Based on (7) and (9), we get the alternative expressions

$$\hat{\beta}_a=\mu_0-\sigma_0\sqrt{\frac{p}{1-p}}, \quad \hat{\beta}_b=\mu_0+\sigma_0\sqrt{\frac{1-p}{p}}.$$

The objective function (11) is evaluated to be

$$\sum_{1\leqslant i<k\leqslant n}\hat{\beta}_{ik}^3=\binom{n}{2}\left(\mu_0^3+3\mu_0\sigma_0^2+\frac{1-2p}{\sqrt{p(1-p)}}\sigma_0^3\right), \quad p\in(0,1). \tag{14}$$

Now using the boundary condition that $\max\{0, s-n/q\} \leqslant \beta_{ik} \leqslant s$, we get the feasible range of p:

$$\frac{\sigma_0^2}{(s-\mu_0)^2+\sigma_0^2} \equiv p_{\min} \leqslant p \leqslant p_{\max} \equiv \begin{cases} \dfrac{\mu_0^2}{\sigma_0^2+\mu_0^2} & \text{if } n \geqslant qs, \\ \dfrac{\left(\mu_0 - s + \dfrac{n}{q}\right)^2}{\sigma_0^2 + \left(\mu_0 - s + \dfrac{n}{q}\right)^2} & \text{otherwise.} \end{cases} \tag{15}$$

Substituting $p_{\max}$ into (14) leads to the lower bound for $\sum_{i<k} \beta_{ik}^3$, since (14) is strictly decreasing in p (So, one can also obtain the upper bound by substituting $p_{\min}$). Denote this lower bound by LB_1.

3.2.2. Step 2: strengthening

This step is to strengthen the lower bound obtained above by taking into account the constraint that both $\hat{\beta}_a$ and $\hat{\beta}_b$ must be integer-valued.

In Step 1 analysis, when $p = p_{\max}$,

$$\hat{\beta}_a = \max\{0, s-n/q\}, \quad \hat{\beta}_b = \mu_0 + \sigma_0^2/(\mu_0 - \hat{\beta}_a). \tag{16}$$

It is straightforward that

$$\sum_{i<k} \beta_{ik}^3 = \sum_{i<k} (\beta_{ik} - \hat{\beta}_a)(\beta_{ik} - \hat{\beta}_b)^2 + \mathrm{LB}_1 \geqslant \mathrm{LB}_1,$$

where the equality holds if β_{ik} takes either $\hat{\beta}_a$ or $\hat{\beta}_b$ for $1 \leqslant i < k \leqslant n$.

When $\hat{\beta}_b$ is not integer-valued, we have that

$$(\beta_{ik} - \hat{\beta}_a)(\beta_{ik} - \lfloor \hat{\beta}_b \rfloor)(\beta_{ik} - \lceil \hat{\beta}_b \rceil) \geqslant 0$$

for any integer-valued $\beta_{ik} \geqslant \hat{\beta}_a$. Thus, we can strengthen the lower bound by taking the values of β_{ik} to be either $\hat{\beta}_a$, $\lfloor \hat{\beta}_b \rfloor$ or $\lceil \hat{\beta}_b \rceil$. By (7) and (9), we have the optimal distribution of the pairwise coincidences

$$\beta_{ik} = \begin{cases} \hat{\beta}_a & \text{with proportion } 1 - p_1 - p_2, \\ \lfloor \hat{\beta}_b \rfloor & \text{with proportion } p_1, \\ \lceil \hat{\beta}_b \rceil & \text{with proportion } p_2 \qquad \text{for } 1 \leqslant i < k \leqslant n, \end{cases} \tag{17}$$

where $\hat{\beta}_a, \hat{\beta}_b$ are as in (16), and the proportions are given by

$$p_1 = (\mu_0 - \hat{\beta}_a - \alpha)/(\lfloor \hat{\beta}_b \rfloor - \hat{\beta}_a), \quad p_2 = \alpha/(\lceil \hat{\beta}_b \rceil - \hat{\beta}_a)$$

and $\alpha = \mu_0^2 + \sigma_0^2 + \lfloor \hat{\beta}_b \rfloor \hat{\beta}_a - \mu_0(\lfloor \hat{\beta}_b \rfloor + \hat{\beta}_a)$. For a non-empty $\mathcal{L}(n, q^s)$ of orthogonal FFDs, it can be verified that there uniquely exists such an optimal distribution of pairwise coincidences. Obviously, the optimality conditions (17) for $\mathcal{S}_2(n, q^s)$ can be viewed as a natural extension of (8) for $\mathcal{S}_1(n, q^s)$.

Based on (17), the refined lower bound for (11) is given by

$$\sum_{i<k} \beta_{ik}^3 \geqslant \mathrm{LB}_2 \tag{18}$$

$$= \binom{n}{2} \left[(\lfloor \hat{\beta}_b \rfloor + \lceil \hat{\beta}_b \rceil)(\mu_0^2 + \sigma_0^2 - \mu_0 \hat{\beta}_a) - \lfloor \hat{\beta}_b \rfloor \lceil \hat{\beta}_b \rceil (\mu_0 - \hat{\beta}_a) + \hat{\beta}_a (\mu_0^2 + \sigma_0^2) \right].$$

LB_2 reduces to LB_1 if $\hat{\beta}_b$ is an integer. Back to the generalized word-length pattern, by (5), the lower bound of $A_3(D)$ is given by

$$A_3(D) \geqslant \frac{n-1}{n} \Big[(1 - p_1 - p_2) K_3(s - \hat{\beta}_a; s, q) + p_1 K_3(s - \lfloor \hat{\beta}_b \rfloor; s, q) + p_2 K_3(s - \lceil \hat{\beta}_b \rceil; s, q) \Big] + \frac{(q-1)^j}{n} \binom{s}{j}, \tag{19}$$

where $K_3(x; s, q)$ denotes the Krawtchouk polynomial

$$K_3(x; s, q) = \sum_{w=0}^{3} (-1)^w (q-1)^{3-w} \binom{x}{w} \binom{s-x}{3-w}.$$

Finally, we claim that the optimal pairwise distance distribution (17) leads to not only the smallest $\sum_{i<k} \beta_{ik}^3$ and $A_3(D)$, but also the generalized minimum aberration. This is true since (17) is the unique minimizer of $\sum_{i<k} \beta_{ik}^3$, and under (17) all the higher-order power moments $\sum_{i<k} \beta_{ik}^j$ for $j > 3$ are uniquely determined, too.

4. Orthogonal sub-design selection

In the algorithmic approach to construct optimal SSDs, there exist construction methods based on randomly generated balanced designs, e.g. the columnwise–pairwise algorithm and its stochastic versions. However, we can hardly construct orthogonal FFDs from scratch, since every candidate needs to be an orthogonal array, while the construction of orthogonal arrays itself is a non-trivial problem. Instead, the sub-design selection approach is most popular for constructing orthogonal FFDs.

Given a specified criterion, a common way of design construction is to select sub-designs from a superset design that contains a larger number of factors. For example, Hadamard matrices are often employed as superset designs for two-level sub-design selection. However, it is left open, especially for the GMA criterion, how close the best selected designs are to the global optimality across the whole class of non-isomorphic designs. We will resolve this issue by utilizing the lower bounds developed in the previous section. Our concentration is on the generation of GMA designs of resolution III or higher.

Suppose there exists a superset design $D_0 \in \mathcal{S}_2(n, q^{s_0})$ with $s_0 > s$. Our objective is to select the sub-designs $D(n, q^s)$ from a sequence of $\binom{s_0}{s}$ possible s-factor projections under the GMA criterion. According to the optimal condition (17), it is sufficient to claim that a candidate sub-design has GMA if one of the following rules is satisfied.

Stopping rules for orthogonal sub-design selection:

(1) The pairwise distances β_{ik} for $i < k$ are all the same. This corresponds to the saturated design that is equidistant.

(2) The pairwise distances β_{ik} for $i < k$ take two different integers, with the smaller one given by $\max\{0, s - n/q\}$. This corresponds to the optimal condition (16) when $\hat{\beta}_b$ is integer-valued.
(3) The pairwise distances β_{ik} for $i < k$ take three different integers, with the smallest given by $\max\{0, s - n/q\}$ and the rest two are consecutive integers. This corresponds to the strengthened optimal condition (17).

Based on these stopping rules, we present a computer search algorithm for orthogonal sub-design selection. In the algorithm, the power moments $\sum_{i<k} \beta_{ik}^j$ for $j = 3, \ldots, s$ are evaluated to serve as the GMA criterion.

Algorithm (*Orthogonal Sub-Design Selection*). Given a superset design $D_0(n, q^{s_0})$ that is orthogonal, do the following to find the s-factor sub-design:

(1) Compute the coincidence matrix for each factor (only need to evaluate the upper triangular part, i.e. δ_{ik}^j for $1 \leqslant i < k \leqslant n$):

$$C_j = \left(\delta_{ik}^j\right)_{n\times n}, \quad j = 1, \ldots, s_0.$$

(2) Collect all s-factor combinations among $\{1, 2, \ldots, s_0\}$:

$$G_l = \left\{j_1^{(l)}, \ldots, j_s^{(l)}\right\}, \quad l = 1, \ldots, \binom{s_0}{s}.$$

(3) For each sub-design $l = 1, 2, \ldots, \binom{s_0}{s}$,
   (a) compute the coincidence matrix by $[\beta_{ik}]_{n\times n} = \sum_{j\in G_l} C_j$. Note that one can also update the coincidence matrix from $l - 1$ to l by adding

$$\sum_{j\in G_l\setminus G_{l-1}} C_j - \sum_{j\in G_{l-1}\setminus G_l} C_j$$

   when $|G_l \setminus G_{l-1}| = 1$ or 2.
   (b) Check β_{ik}'s to see if they satisfy any of the stopping rules. If so, terminate the loop and output the sub-design with G_l-factors. Otherwise, compute the power moments

$$B_l = \left(\sum\nolimits_{i<k} \beta_{ik}^3, \ldots, \sum\nolimits_{i<k} \beta_{ik}^s\right);$$

   update the best record $B_{l^*} = B_l$ if the first non-zero element of $B_l - B_{l^*}$ is negative.

If any stopping rule is satisfied, the algorithm outputs a GMA design $D(n, q^s)$; otherwise, it reports the best sub-design among $\binom{s_0}{s}$ candidates that has the smallest $(\sum_{i<k} \beta_{ik}^3, \ldots, \sum_{i<k} \beta_{ik}^s)$ in the sense of sequential comparison.

There are many resources on superset designs on Sloane's orthogonal arrays website (http://www.research.att.com/~njas/oadir/). As an example, we choose oa.27.13.3.2 from Sloane's website as the superset design $D_0(27, 3^{13})$ for selecting 3-level sub-designs with 27 runs. Computational results are shown in Table 1. The resulting sub-designs have GMA for $k = 4, 6, 8$–13. However, for $k = 5$ and 7, the selected best sub-designs cannot achieve the lower bounds.

This algorithm works fast for small s_0, say, no greater than 16. For example, running through $s = 4, \ldots, 12$ in the oa.27.13.3.2 example costs only about 5 s of CPU time under our MATLAB 7 (The MathWorks, Inc.) environment, on a PC with 1.60 GHz Pentium M processor and 512 MB

Table 1
Sub-design selection results from superset design oa.27.13.3.2

$D(27, 3^s)$	Factor combination	GMA	$\sum_{i<k} \beta_{ik}^3$	Lower bound
$s = 4$	{9, 11, 12, 13}	Yes	1404	1404
5	{7, 9, 11, 12, 13}	NS	2322	2160
6	{6, 7, 8, 9, 12, 13}	Yes	3402	3402
7	{6, 7, 8, 9, 11, 12, 13}	NS	4968	4914
8	{5, 6, 7, 8, 9, 11, 12, 13}	Yes	6696	6696
9	{3, 5, 6, 7, 8, 9, 11, 12, 13}	Yes	8748	8748
10	{3, 5, 6, 7, 8, 9, 10, 11, 12, 13}	Yes	11 772	11 772
11	{3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13}	Yes	14958	14 958
12	{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13}	Yes	18 468	18 468

"NS" under "GMA" means "not sure". The selected design has resolution IV for $s = 4$ and resolution III for $s > 4$. The GMA design at $s = 12$ is weak-equidistant.

Table 2
CPU time in seconds resulting from the early stopping rule (A) and the exhaustive search (B) for the oa.27.13.3.2 example

s	4	5	6	7	8	9	10	11	12	Total
A	0.0501	1.8426	0.0601	2.6638	0.0601	0.0601	0.0501	0.0501	0.0401	4.8770
B	3.5852	1.9728	3.5651	2.5537	3.7354	1.3119	1.2518	0.4006	0.2203	18.5967

RAM. Compared to the exhaustive search in traditional sub-design selection procedures, the computational complexity of our algorithm is greatly reduced, since it terminates immediately upon hitting any stopping rule. To illustrate this gain in computing time, Table 2 lists CPU time cost for the early stopping rule and the exhaustive search, respectively.

For large s_0, the algorithm becomes intractable since the total number of combinations, $\binom{s_0}{s}$, grows exponentially. Because many of the sub-designs are actually isomorphic and yield the same coincidence distribution, we can employ a simple trick of randomization to avoid the useless duplicates and overcome the intractability problem. Let us call $D(n, q^s)$ a random sub-design of $D_0(n, q^{s_0})$ if it selects s out of s_0 factors at random. Let M be a pre-defined maximal number of trials. Then, the second step in the algorithm above can be modified to M random sub-designs instead of using the complete $\binom{s_0}{s}$ candidates. Our experience suggests that $M = 200$ would succeed in finding the GMA sub-design, whenever there exists a candidate sub-design that satisfies a stopping rule. Otherwise, a more conservative choice of M is recommended.

By inputting different supersets from Sloane's orthogonal arrays website (in particular, those oa.n.s_0.q.2 of strength 2), scanning through the sub-designs for $s = 3, \ldots, s_0$, a large number of GMA designs can be found by our orthogonal sub-design selection algorithm. These optimal designs have resolution III or higher, including both two-level and multi-level cases. It is also interesting to note that the construction results of [3] can be all reproduced by our algorithm, including the GMA designs $D(27, 3^s)$ for $s = 8, \ldots, 12$ (cf. Table 1), $D(81, 4^s)$ for $s = 15, \ldots, 20$ and $D(50, 5^s)$ for $s = 3, \ldots, 11$.

Finally, we report some GMA designs that are new to the FFD literature, as summarized in Table 3. These GMA designs have not appeared elsewhere, to the best of our knowledge. Furthermore, there is no restriction to choose other superset designs in the search of new GMA designs.

Table 3
Some new GMA designs to the FFD literature

Superset	GMA Design	Factor combination
oa.27.13.3.2	$D(27, 3^4)$	cf. Table 1
	$D(27, 3^6)$	
oa.81.40.3.2	$D(81, 3^{36})$	{2, 8, 16, 37} deleted
	$D(81, 3^{37})$	{16, 26, 36} deleted
oa.64.21.4.2	$D(64, 4^5)$	{9, 16, 17, 18, 19}
	$D(64, 4^6)$	{2, 8, 13, 16, 18, 19}
	$D(64, 4^9)$	{3, 4, 10, 12, 14, 15, 16, 17, 20}
	$D(64, 4^{12})$	{1, 3, 4, 5, 6, 7, 8, 9, 11, 17, 18, 20}

5. Conclusion

This paper demonstrates how to apply the classical optimization approach to study the GMA criterion for FFDs, by employing Lagrange analysis and a strengthening trick to sharpen the lower bounds. We have shown that the new optimality conditions (17) for orthogonal designs can be viewed as an extension of weak-equidistance optimality (8) for balanced designs. These conditions serve as the early stopping rules in our orthogonal sub-design selection algorithm.

Before ending the paper, we discuss several open problems for possible routes of future research. First, it is interesting to extend the GMA optimality to mixed resolution-III designs $D(n, q_1^{s_1} q_2^{s_2})$. One can refer to [9] for this kind of effort in extending the weak-equidistance optimality of resolution-II designs. Second, optimal conditions for GMA designs with higher resolution ($r \geqslant 4$) can be also analyzed via pairwise coincidences. Similar to the optimization setting (11), one could focus on the minimization of $\sum_{i<k} \beta_{ik}^r$, subject to the necessary conditions (2) and possible boundary conditions like (12). However in this case, Lagrange analysis requires analytically finding the roots of higher-order polynomials. The optimal conditions seem not as obvious as those for the resolution-II and III designs. Third, the sub-design selection algorithm in Section 4 enumerates a large number of sub-designs, $\binom{s_0}{s}$ in total. It becomes not efficient when s_0 is large, in which case we suggested a simple random sampling scheme to reduce the computational complexity. More advanced algorithms can be developed by incorporating some heuristics. All these problems are worthy of our future investigation.

Acknowledgments

The authors thank the Associate Editor and two anonymous referees for their valuable comments. Fang and Zhang were partially supported by the Hong Kong RGC-Grant RGC/HKBU 200804. Zhang was also supported by University of Michigan and thanks Professor Vijay Nair for valuable suggestions. Li was supported by a National Institute on Drug Abuse (NIDA) Grant P50 DA10075, and partially supported by NSF Grants DMS 0348869 and CCF 0430349.

References

[1] R.C. Bose, K.A. Bush, Orthogonal arrays of strength two and three, Ann. Math. Statist. 23 (1952) 508–524.
[2] G.E.P. Box, J.S. Hunter, The 2^{k-p} fractional factorial designs, Technometrics 3 (1961) 311–352, 449–458.

[3] N.A. Butler, Generalised minimum aberration construction results for symmetrical orthogonal arrays, Biometrika 92 (2005) 485–491.
[4] P. Delsarte, An algebraic approach to the association schemes of coding theory, Philips Res. Rep. Suppl. 10 (1973).
[5] K.-T. Fang, G.N. Ge, M.-Q. Liu, H. Qin, Construction of minimum generalized aberration designs, Metrika 57 (2003) 37–50.
[6] K.-T. Fang, R. Li, A. Sudjianto, Design and Modeling for Computer Experiments, Chapman & Hall, Boca Raton, 2006.
[7] A. Fries, W.G. Hunter, Minimum aberration 2^{k-p} designs, Technometrics 22 (1980) 601–608.
[8] D.K.J. Lin, A new class of supersaturated designs, Technometrics 35 (1993) 28–31.
[9] M.Q. Liu, K.T. Fang, F.J. Hickernell, Connections among different criteria for assymetrical fractional factorial designs, Statist. Sinica 16 (2006) 1285–1297.
[10] C.X. Ma, K.T. Fang, A note on generalized aberration in factorial designs, Metrika 53 (2001) 85–93.
[11] F.J. MacWilliams, N.J.A. Sloane, The Theory of Error-Correcting Codes, North-Holland, Amsterdam, 1977.
[12] R. Mukerjee, C.F.J. Wu, On the existence of saturated and nearly saturated asymmetrical orthogonal arrays, Ann. Statist. 23 (1995) 2102–2115.
[13] C.R. Rao, Factorial experiments derivable from combinatorial arrangements of arrays, J. Roy. Statist. Soc. B 9 (1947) 128–139.
[14] B. Tang, L.Y. Deng, Minimum G_2-aberration for nonregular fractional factorial designs, Ann. Statist. 27 (1999) 1914–1926.
[15] C.F.J. Wu, M. Hamada, Experiments: Planning, Analysis, and Parameter Design Optimization, Wiley, New York, 2000.
[16] H. Xu, Minimum moment aberration for nonregular designs and supersaturated designs, Statist. Sinica 13 (2003) 691–708.
[17] H. Xu, C.F.J. Wu, Generalized minimum aberration for asymmetrical fractional factorial designs, Ann. Statist. 29 (2001) 1066–1077.
[18] A. Zhang, K.T. Fang, R. Li, A. Sudjianto, Majorization framework for balanced lattice designs, Ann. Statist. 33 (2005) 2837–2853.