CHARACTERIZATION OF GENERAL MINIMUM LOWER ORDER CONFOUNDING VIA COMPLEMENTARY SETS

Runchu Zhang[1] and Rahul Mukerjee[2]

[1]*Nankai University,* [2]*Indian Institute of Management Calcutta*

Abstract: With reference to regular fractions of general s-level factorials, we consider the design criterion of general minimum lower order confounding (GMC) that aims, in an elaborate manner, at keeping the lower order factorial effects unaliased with one another to the extent possible. Using a finite projective geometric formulation, that involves identification of the alias sets with the points of the geometry, we derive explicit formulae connecting the key terms for this criterion with the complementary set. These results are then applied to find optimal designs under the GMC criterion.

Key words and phrases: Alias set, effect hierarchy principle, finite projective geometry, minimum aberration.

1. Introduction

The problem of optimal selection of regular fractional factorial plans, under model uncertainty, has received significant attention in the literature. The criterion of minimum aberration (MA) has gained much popularity in this context. In addition, criteria such as maximum estimation capacity (MEC) and clear effects have been proposed and studied. We refer to Mukerjee and Wu (2006) for a review. All these criteria are based on the effect hierarchy principle (Wu and Hamada, 2000, p.112) which treats factorial effects of the same order as equally important and lower order effects as more important than higher order effects. They all are motivated, in various senses, by the objective of keeping the lower order factorial effects unaliased with one another to the extent possible. With reference to two-level factorials, Zhang, Li, Zhao and Ai (2007, hereafter called ZLZA) recently introduced and discussed at length a new criterion of *general minimum lower order confounding* (GMLOC or GMC for short) that aims at achieving the same objective in a more elaborate and explicit manner.

The purpose of the present work is to develop a theory for the GMC criterion in terms of complementary sets. The results should be particularly useful in the practically important nearly saturated situation where the complementary set is relatively small in size and hence easy to handle. This approach has found much applicability in the study of optimal designs under the MA or MEC criteria. Considerable additional work is, however, required for the present problem because we now need to obtain explicit relations connecting the aliasing pattern of a design for the lower order factorial effects with the complementary set. In addition to being directly useful in the study of the GMC criterion, these identities facilitate a more transparent understanding of such aliasing pattern. Identification of the alias sets with the points of a finite projective geometry helps in their derivation. In contrast with ZLZA who studied the GMC criterion in the two-level case, we consider general s-level factorials where s is a prime or prime power. This enables us to include, for example, three-level factorials as well in the present study.

2. Background and preliminary results

Consider an s^n factorial involving n factors $F_1, \ldots, F_n$, each at s levels, where $s\,(\geq 2)$ is a prime or prime power. A typical pencil $b = (b_1, \ldots, b_n)'$ is a nonnull n-vector over the finite field $GF(s)$, and pencils with proportional elements are considered identical. A pencil with i nonzero elements is called an ith order pencil ($1 \leq i \leq n$). If these nonzero elements occur at the j_1th, ..., j_ith positions, then the pencil represents the ith order factorial effect $F_{j_1} \cdots F_{j_i}$. Since pencils with proportional elements are identical, any ith order effect is represented by $(s-1)^{i-1}$ pencils. The case $i = 1$ gives a *main effect* while the case $i > 1$ corresponds to an *interaction*.

We will be interested in regular $1/s^m$ fractions of an s^n factorial (hereafter, simply referred to as s^{n-m} designs), where $1 \leq m < n$ and, to avoid trivialities, $n \geq 3$. For $1 \leq i \leq n$, let A_i be the number of ith order pencils appearing in the defining relation of such a design. The *resolution* of the design is the smallest integer i for which $A_i > 0$. Since the main effects are of utmost importance in a factorial setting, in what follows only designs of resolution three or higher are considered. These designs do not assign any main effect pencil to the defining relation or entail aliasing of any two such pencils. Then $A_1 = A_2 = 0$ and the sequence $(A_3, \ldots, A_n)$ represents the *wordlength pattern* (WLP) of the design. The criterion of *minimum aberration* (MA) aims at sequential minimization of $A_3, A_4, \ldots$.

Following ZLZA, we now introduce the GMC criterion, as applicable to an s^{n-m} design of resolution three or higher. An ith order pencil is said to be aliased with jth order pencils at degree k if it does not appear in the defining relation and is aliased with k jth order pencils (excluding itself, in case $j = i$). Noting that there are K_j $[= \binom{n}{j}(s-1)^{j-1}]$ jth order pencils altogether, for $1 \leq i, j \leq n$ and $0 \leq k \leq K_j$, define ${}^{\#}_{i}C^{(k)}_{j}$ as the number of ith order pencils that are aliased with jth order pencils at degree k, and write ${}^{\#}_{i}C_j$ for the vector $({}^{\#}_{i}C^{(0)}_{j}, {}^{\#}_{i}C^{(1)}_{j}, \ldots, {}^{\#}_{i}C^{(K_j)}_{j})$. Also, for $1 \leq j \leq n$, let ${}^{\#}_{0}C_j$ be a row vector of order $K_j + 1$, with kth element 1 if $A_j = k$, and 0 otherwise $(0 \leq k \leq K_j)$.

The sequence

$$ {}^{\#}C = ({}_{1}^{\#}C_2, {}_{2}^{\#}C_1, {}_{2}^{\#}C_2, {}_{0}^{\#}C_3, {}_{1}^{\#}C_3, {}_{2}^{\#}C_3, {}_{3}^{\#}C_1, {}_{3}^{\#}C_2, {}_{3}^{\#}C_3, \ldots) \tag{2.1} $$

is called the *aliased effect-number pattern* (AENP) of the design. The AENP incorporates the WLP via the terms ${}_{0}^{\#}C_j$ in (2.1) and is, in fact, much more informative than the WLP because of the terms ${}_{i}^{\#}C_j^{(k)}(i, j \geq 1)$ which explicitly reflect the nature of aliasing. Note that (2.1) places ${}_{i}^{\#}C_j$ ahead of ${}_{i^*}^{\#}C_{j^*}$ if and only if either (i) $\max(i, j) < \max(i^*, j^*)$, or (ii) $\max(i, j) = \max(i^*, j^*)$ and $i < i^*$, or (iii) $\max(i, j) = \max(i^*, j^*)$, $i = i^*$ and $j < j^*$. As discussed by ZLZA in detail, this entails an arrangement of the terms ${}_{i}^{\#}C_j$ in (2.1) in decreasing order of importance from left to right, under the effect hierarchy principle. In addition, for any fixed i and j, the first element ${}_{i}^{\#}C_j^{(0)}$ of ${}_{i}^{\#}C_j$ signifies no aliasing, while the subsequent elements ${}_{i}^{\#}C_j^{(k)}$ signify progressively more severe aliasing as k increases from 1 to K_j. From these perspectives, the GMC criterion proposed by ZLZA aims at sequential maximization of the elements of ${}^{\#}C$, from left to right. A more precise definition appears below.

At this stage, we note that some of the terms in (2.1) are uniquely determined by others that precede them and hence are redundant under the GMC criterion. In a design of resolution three or higher, any jth order pencil ($j \geq 2$) is aliased with at most one first order pencil, and the number of jth order pencils that are aliased with one first order pencil equals $\sum_{k\geq 1} k\, {}_{1}^{\#}C_j^{(k)}$. Hence

$$ {}_{j}^{\#}C_1^{(1)} = \sum_{k\geq 1} k\, {}_{1}^{\#}C_j^{(k)},\ {}_{j}^{\#}C_1^{(0)} = K_j - {}_{j}^{\#}C_1^{(1)} - A_j,\ {}_{j}^{\#}C_1^{(k)} = 0\ (k \geq 2). $$

Furthermore, an inspection of the manners in which a defining pencil can entail aliasing of a first order pencil with a jth order one shows that

$$ \sum_{k\geq 1} k\, {}_{1}^{\#}C_j^{(k)} = (n - j + 1)(s - 1)A_{j-1} + j(s - 2)A_j + (j + 1)A_{j+1}, \tag{2.2} $$

where $A_{n+1} = 0$. Thus the ${}_{0}^{\#}C_j\ (j \geq 3)$ in (2.1) can be dropped as they are uniquely determined by the preceding terms ${}_{1}^{\#}C_u, 2 \leq u \leq j - 1$. Similarly, the ${}_{j}^{\#}C_1\ (j \geq 2)$ are redundant because of the terms ${}_{1}^{\#}C_u, 2 \leq u \leq j$. As a result, the GMC criterion can be defined on the basis of a simpler version of (2.1) as given by

$$ {}^{\#}C = ({}_{1}^{\#}C_2, {}_{2}^{\#}C_2, {}_{1}^{\#}C_3, {}_{2}^{\#}C_3, {}_{3}^{\#}C_2, {}_{3}^{\#}C_3, \ldots). \tag{2.3} $$

Definition 1. *Let ${}^{\#}C_l$ be the lth element of ${}^{\#}C$ in (2.3), and ${}^{\#}C(d_1)$ and ${}^{\#}C(d_2)$ be the AENPs of designs d_1 and d_2 respectively. Suppose t is the smallest integer such that*

${}^{\#}C_t(d_1) \neq {}^{\#}C_t(d_2)$. If ${}^{\#}C_t(d_1) > {}^{\#}C_t(d_2)$ then d_1 is said to have less general lower order confounding (GLOC) than d_2. A design d is said to have general minimum lower order confounding (GMC) if no other design has less GLOC than d and, in this case, d is called a GMC design.

A geometric formulation helps in characterizing GMC designs. Write $L_w = (s^w - 1)/(s-1)$. Let P denote the set of the L_{n-m} points of the finite projective geometry $PG(n-m-1, s)$. For any nonempty subset Q of P, define $V(Q)$ as the matrix given by the points of Q as columns. Then the following well-known lemma holds (see e.g., Mukerjee and Wu, 2006, p.40).

Lemma 1. *Any s^{n-m} design d of resolution three or higher is represented by an n-subset T of P such that $V(T)$ has full row rank and*

(a) the rows of $V(T)$ span the treatment combinations in d,

(b) any pencil b appears in the defining relation of d if and only if $V(T)b = 0$,

(c) any two pencils $b^{(1)}$ and $b^{(2)}$, neither of which is a defining pencil, are aliased with each other in d if and only if $V(T)b^{(1)}$ and $V(T)b^{(2)}$ are proportional to the same point of P.

In view of Lemma 1, hereafter, an s^{n-m} design of resolution three or higher is denoted simply by the corresponding set T. Lemma 1(c), exhibiting a one to one correspondence between the L_{n-m} points of P and the L_{n-m} alias sets, will be very useful. We now introduce some more notation. Consider any point γ of P and any nonempty subset Q of P. Let q denote the cardinality of Q and Ω_{iq} be the set of q-vectors over $GF(s)$ having i nonzero elements. For $i \geq 1$, define

$$A_i(Q) = (s-1)^{-1}\#\{\lambda :\ \lambda \in \Omega_{iq}, V(Q)\lambda = 0\}, \tag{2.4}$$

$$B_i(Q,\gamma) = (s-1)^{-1}\#\{\lambda :\ \lambda \in \Omega_{iq}, V(Q)\lambda \text{ is nonnull and proportional to } \gamma\}, \tag{2.5}$$

where # denotes the cardinality of a set. In particular, if $V(Q)$ has full row rank then Q represents an $s^{q-(q-n+m)}$ design; in this case, by Lemma 1(b), (c), $A_i(Q)$ is the same as the A_i in the WLP of Q and $B_i(Q,\gamma)$ is the number of ith order pencils appearing in the alias set, corresponding to γ, of Q. Even if $V(Q)$ is not of full row rank, the expressions in (2.4) and (2.5) are well-defined, albeit without the above interpretation. We are now in a position to present Lemma 2 below. Here $\overline{Q}$, of cardinality $\overline{q} = L_{n-m} - q$, is the complement of Q in P, and

$$G_3(q,\overline{q}) = \frac{1}{6}(s-1)\{q(q-1) + \overline{q}(\overline{q}-1) - q\overline{q}\}, \tag{2.6}$$

$$\begin{aligned} G_4(q,\overline{q}) = \ & \frac{1}{24}(s-1)[(s-1)\{q(q-1)(q-\overline{q}-2)-\overline{q}(\overline{q}-1)(\overline{q}-q-2)\} \\ & -(3s-5)\{q(q-1)+3\overline{q}(\overline{q}-1)-2q\overline{q}\}]. \end{aligned} \tag{2.7}$$

Lemma 2.

(a) $A_3(Q) = G_3(q,\overline{q}) - A_3(\overline{Q})$,

(b) $A_4(Q) = G_4(q,\overline{q}) + (3s-5)A_3(\overline{Q}) + A_4(\overline{Q})$,

(c) if $\gamma \notin Q$, *then* $A_i(Q \cup \{\gamma\}) = A_i(Q) + B_{i-1}(Q,\gamma)$, *for* $i = 3, 4$,

(d) if $\gamma \in Q$, *then* $A_3(Q) = A_3(Q\backslash\{\gamma\}) + B_2(Q,\gamma)$ *and* $A_4(Q) = A_4(Q\backslash\{\gamma\}) + B_3(Q,\gamma) - (s-2)B_2(Q,\gamma)$.

Parts (a) and (b) of Lemma 2 are due to Suen, Chen and Wu (1997), who gave expressions for G_3 and G_4 involving Krawtchouk polynomials. Additional algebra yields the more explicit forms shown in (2.6) and (2.7) above. These can also be obtained generalizing the approach in Tang and Wu (1996) to the s-level case. Unlike what happens with the MA criterion, we now require the details on G_3 and G_4 in order to obtain certain relationships in explicit forms which can actually be needed in discriminating among designs; see Theorem 1 and Example 2 below. Lemma 2(c) follows from Lemma 2 in Mukerjee and Wu (1999). Finally, Lemma 2(d) follows replacing Q by $Q\backslash\{\gamma\}$ in Lemma 2(c) and then invoking the following lemma.

Lemma 3. *If* $\gamma \in Q$, *then*

(i) $B_2(Q\backslash\{\gamma\},\gamma) = B_2(Q,\gamma)$ *and*

(ii) $B_3(Q\backslash\{\gamma\},\gamma) = B_3(Q,\gamma) - (s-2)B_2(Q,\gamma)$.

Proof. Since the points of P are nonnull and no two of them are proportional to each other, no nonnull linear combination of γ and a point from $Q\backslash\{\gamma\}$, with both combining coefficients nonzero, can be proportional to γ. Hence (i) follows from (2.5). Next suppose the relation $\lambda_1\gamma^{(1)} + \lambda_2\gamma^{(2)} + \lambda_3\gamma^{(3)} = \lambda_0\gamma$ holds for some nonzero elements $\lambda_1, \lambda_2, \lambda_3, \lambda_0$ of $GF(s)$ and some three points $\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}$ of Q. Then either each of $\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}$ belongs to $Q\backslash\{\gamma\}$, or one of them, say $\gamma^{(1)}$, equals γ and the other two belong to $Q\backslash\{\gamma\}$. In the latter case $\lambda_1 \neq 0, \lambda_0$, because no two points of P are proportional to each other, so that there are $s-2$ possibilities for λ_1. Recalling (2.5), these considerations yield $B_3(Q,\gamma) = B_3(Q\backslash\{\gamma\},\gamma) + (s-2)B_2(Q\backslash\{\gamma\},\gamma)$, which, in conjunction with (i), yields (ii). $\square$

3. Expressions in terms of the complementary set

Consider now an s^{n-m} design of resolution three or higher, represented by an n-

subset T of P as envisaged in Lemma 1. Let $\overline{T}$ be the complement of T in P. The cardinality of $\overline{T}$ is $f = L_{n-m} - n$. As before, γ is any point of P. Then the following result holds.

Theorem 1.

(a) $B_1(T,\gamma)$ *equals 1 if* $\gamma \in T$, *and 0 if* $\gamma \notin T$.

(b) If $\gamma \in T$ *then* $B_2(T,\gamma) = \frac{1}{2}(s-1)(n-f-1) + B_2(\overline{T},\gamma)$.

(c) If $\gamma \notin T$ *then* $B_2(T,\gamma) = \frac{1}{2}(s-1)(n-f+1) + B_2(\overline{T},\gamma)$.

(d) If $\gamma \in T$ *then* $B_3(T,\gamma) = H_{31}(n,f) - (2s-3)B_2(\overline{T},\gamma) - B_3(\overline{T},\gamma)$, *where*

$$H_{31}(n,f) = \frac{1}{6}(s-1)[(s-1)\{(n-1)(n-2) + f(f+3) - nf\} - (n+f-1)].$$

(e) If $\gamma \notin T$ *then* $B_3(T,\gamma) = H_{32}(n,f) - (2s-3)B_2(\overline{T},\gamma) - B_3(\overline{T},\gamma)$, *where*

$$H_{32}(n,f) = \frac{1}{6}(s-1)[(s-1)\{(n-1)(n-2) + f(f+3) - nf - 6\} + 2(n-2f+2)].$$

Proof. (a) This is obvious from (2.5).

(b) By Lemma 2(a),(c),(d), if $\gamma \in T$ then

$$A_3(T\backslash\{\gamma\}) + B_2(T,\gamma) = A_3(T) = G_3(n,f) - A_3(\overline{T}),$$
$$A_3(T\backslash\{\gamma\}) = G_3(n-1,f+1) - A_3(\overline{T}\cup\{\gamma\}) = G_3(n-1,f+1) - A_3(\overline{T}) - B_2(\overline{T},\gamma).$$

The truth of (b) now follows because $G_3(n,f) - G_3(n-1,f+1) = \frac{1}{2}(s-1)(n-f-1)$ by (2.6).

(c) Follows from (b) interchanging the roles of T and $\overline{T}$ and hence those of n and f.

(d) By Lemma 2(b),(c),(d), if $\gamma \in T$ then

$$A_4(T\backslash\{\gamma\}) + B_3(T,\gamma) - (s-2)B_2(T,\gamma) = A_4(T) = G_4(n,f) + (3s-5)A_3(\overline{T}) + A_4(\overline{T}),$$
$$\begin{aligned} A_4(T\backslash\{\gamma\}) &= G_4(n-1,f+1) + (3s-5)A_3(\overline{T}\cup\{\gamma\}) + A_4(\overline{T}\cup\{\gamma\}) \\ &= G_4(n-1,f+1) + (3s-5)\{A_3(\overline{T}) + B_2(\overline{T},\gamma)\} + A_4(\overline{T}) + B_3(\overline{T},\gamma). \end{aligned}$$

Hence

$$B_3(T,\gamma) = G_4(n,f) - G_4(n-1,f+1) + (s-2)B_2(T,\gamma) - (3s-5)B_2(\overline{T},\gamma) - B_3(\overline{T},\gamma).$$

Using part (b), the result now follows noting that by (2.7),

$$G_4(n,f) - G_4(n-1,f+1) + \frac{1}{2}(s-2)(s-1)(n-f-1) = H_{31}(n,f).$$

(e) If $\gamma \notin T$ then interchanging the roles of T and $\overline{T}$ and hence those of n and f in (d), $B_3(\overline{T},\gamma) = H_{31}(f,n) - (2s-3)B_2(T,\gamma) - B_3(T,\gamma)$, whence the result follows using part (c) and the fact that $H_{31}(f,n) - \frac{1}{2}(2s-3)(s-1)(n-f+1) = H_{32}(n,f)$. □

With reference to the design T, recall that $B_i(T,\gamma)$ is the number of ith order pencils appearing in the alias set that corresponds to γ. Hence, for this design

$$ {}_i^{\#}C_i^{(k)} = (k+1)\#\{\gamma : \gamma \in P,\ B_i(T,\gamma) = k+1\},\ 0 \le k \le K_i,\ 1 \le i \le n, \tag{3.1}$$

$$ {}_i^{\#}C_j^{(k)} = \sum\nolimits_{jk} B_i(T,\gamma),\ 0 \le k \le K_j,\ 1 \le i \ne j \le n, \tag{3.2}$$

where $\sum_{jk}$ is sum over γ such that $\gamma \in P$ and $B_j(T,\gamma) = k$. Theorem 1 can now be readily applied to (3.1) and (3.2) to yield expressions, in terms of the complementary set $\overline{T}$, for the leading terms ${}_1^{\#}C_2, {}_2^{\#}C_2, {}_1^{\#}C_3, {}_2^{\#}C_3, {}_3^{\#}C_2, {}_3^{\#}C_3$ in the AENP given by (2.3). Thus

$$ {}_1^{\#}C_2^{(k)} = \#\Big\{\gamma : \gamma \in T,\ \frac{1}{2}(s-1)(n-f-1) + B_2(\overline{T},\gamma) = k\Big\}, \tag{3.3}$$

$$ \begin{aligned} {}_2^{\#}C_2^{(k)} = \ & (k+1)\,[\#\{\gamma : \gamma \in T,\ \tfrac{1}{2}(s-1)(n-f-1) + B_2(\overline{T},\gamma) = k+1\} \\ & +\#\{\gamma : \gamma \in \overline{T},\ \tfrac{1}{2}(s-1)(n-f+1) + B_2(\overline{T},\gamma) = k+1\}\,], \end{aligned} \tag{3.4}$$

$$ {}_1^{\#}C_3^{(k)} = \#\{\gamma : \gamma \in T,\ H_{31}(n,f) - (2s-3)B_2(\overline{T},\gamma) - B_3(\overline{T},\gamma) = k\}, \tag{3.5}$$

$$ {}_2^{\#}C_3^{(k)} = \sum\nolimits_{3k}^{(1)} \Big\{\frac{1}{2}(s-1)(n-f-1) + B_2(\overline{T},\gamma)\Big\} + \sum\nolimits_{3k}^{(2)} \Big\{\frac{1}{2}(s-1)(n-f+1) + B_2(\overline{T},\gamma)\Big\},$$

and so on, where $\sum_{3k}^{(1)}$ is sum over γ such that $\gamma \in T$ and $H_{31}(n,f) - (2s-3)B_2(\overline{T},\gamma) - B_3(\overline{T},\gamma) = k$, while $\sum_{3k}^{(2)}$ is sum over γ such that $\gamma \in \overline{T}$ and $H_{32}(n,f) - (2s-3)B_2(\overline{T},\gamma) - B_3(\overline{T},\gamma) = k$.

The aforesaid expressions in terms of the complementary set $\overline{T}$ depend on $\overline{T}$ only through $B_2(\overline{T},\gamma)$ and $B_3(\overline{T},\gamma)$, $\gamma \in P$. The calculation of these quantities is facilitated if one writes $\overline{T} = \{\pi_1, \ldots, \pi_f\}$, considers the collections of vectors

$$ M_2 = \{\pi_i + \alpha\pi_j : 1 \le i < j \le f;\ \alpha(\ne 0) \in GF(s)\}, \tag{3.6}$$

$$ M_3 = \{\pi_i + \alpha_1\pi_j + \alpha_2\pi_u : 1 \le i < j < u \le f;\ \alpha_1, \alpha_2(\ne 0) \in GF(s)\}, \tag{3.7}$$

and observes from (2.5) that $B_2(\overline{T},\gamma)$ and $B_3(\overline{T},\gamma)$ are nothing but the numbers of vectors, respectively in M_2 and M_3, that are nonnull and proportional to γ. Thus one only needs to prepare frequency distributions, separately for M_2 and M_3, on the basis of the proportionality of the vectors therein to the various $\gamma\,(\in P)$. This is quite straightforward, even by hand calculation, for relatively small f, in which case M_2 and M_3 are also small in size. Here is an illustrative example.

Example 1. To simplify notation, denote any point $(x_1, x_2, \ldots, x_{n-m})'$ of P by $1^{x_1}2^{x_2}\cdots(n-m)^{x_{n-m}}$, with i^{x_i} dropped if $x_i = 0$. Consider a 3^{n-m} design represented by a set T of P such that $\overline{T} = \{1, 2, 12, 12^2, 3\}$. Then $f = 5$, and since $f < L_{n-m}$, we get $n-m \ge 3$, i.e., $n = L_{n-m} - f \ge 8$. Write $\overline{T}^* = \{1, 2, 12, 12^2\}$, $T^* =$

$\{13, 13^2, 23, 23^2, 123, 123^2, 12^23, 12^23^2\}$, and note that $\overline{T}^* \subset \overline{T}$, $T^* \subset T$. From (3.6), it can be seen that there are 20 vectors in M_2, of which three are proportional to each point of $\overline{T}^*$ and one is proportional to each point of T^*. Similarly, by (3.7), out of the 40 vectors in M_3, four are null and three are proportional to each point of $\overline{T}^*$ and T^*. Thus among the five points of $\overline{T}$, the four from $\overline{T}^*$ have $B_2(\overline{T},\gamma) = B_3(\overline{T},\gamma) = 3$ and the remaining one has $B_2(\overline{T},\gamma) = B_3(\overline{T},\gamma) = 0$. Similarly, among the n points of T, the eight from T^* have $B_2(\overline{T},\gamma) = 1$, $B_3(\overline{T},\gamma) = 3$, and the remaining $n-8$ have $B_2(\overline{T},\gamma) = B_3(\overline{T},\gamma) = 0$. Since $H_{31}(n,f) = \frac{1}{3}(2n^2 - 17n + 80)$ in this example, from (3.3)-(3.5) it now follows that for the design T,

$$
\begin{aligned}
&{}^{\#}_{1}C_2^{(n-6)} = n-8,\ {}^{\#}_{1}C_2^{(n-5)} = 8,\ {}^{\#}_{1}C_2^{(k)} = 0 \text{ for every other } k;\\
&{}^{\#}_{2}C_2^{(n-7)} = (n-6)(n-8),\ {}^{\#}_{2}C_2^{(n-6)} = 8(n-5),\ {}^{\#}_{2}C_2^{(n-5)} = n-4,\\
&\qquad {}^{\#}_{2}C_2^{(n-2)} = 4(n-1),\ {}^{\#}_{2}C_2^{(k)} = 0 \text{ for every other } k;\\
&{}^{\#}_{1}C_3^{(k)} = 8 \text{ if } k = \tfrac{1}{3}(2n^2-17n+80)-6,\ {}^{\#}_{1}C_3^{(k)} = n-8 \text{ if } k = \tfrac{1}{3}(2n^2-17n+80),\\
&\qquad {}^{\#}_{1}C_3^{(k)} = 0 \text{ for every other } k. \qquad \square
\end{aligned}
$$

While the results in this section cover the six leading terms in the AENP given by (2.3), similar techniques will work, at the expense of heavier algebra, if we wish to find their counterparts for the subsequent terms. However, this will not be needed because, as seen in the next section, the results reported above are quite comprehensive for discrimination among designs under the GMC criterion. Moreover, as ZLZA point out, in many applications it is reasonable to assume the absence of factorial effects beyond the third order, in which situation the six leading terms listed in (2.3) completely determine the AENP.

4. GMC designs via complementary sets

The results of the last section are now applied to characterize GMC designs for relatively small values of f which, as indicated earlier, are practically important. Special attention is given to two- and three-level factorials. We consider $f \geq 3$, since all designs are isomorphic for $f = 1, 2$.

For a design T, let $\delta(T) = (\delta_1(T), \ldots, \delta_n(T))$ be the vector with elements $B_2(\overline{T},\gamma)$, $\gamma \in T$, arranged in the nondecreasing order. If

$$g(T) = \#\{\gamma : \gamma \in T, B_2(\overline{T},\gamma) > 0\}, \tag{4.1}$$

then the first $n - g(T)$ elements of $\delta(T)$ are zeros, and the rest are positive. The next two propositions will be useful. The first of these emerges from (2.3), (3.3) and Definition 1, and the second, giving a necessary condition for a design to have GMC, is evident from the first.

Proposition 1. *Consider two designs T_1 and T_2. Suppose $\delta(T_1) \neq \delta(T_2)$ and let t be the smallest integer such that $\delta_t(T_1) \neq \delta_t(T_2)$. If $\delta_t(T_1) < \delta_t(T_2)$ then T_1 has less GLOC than T_2 and hence dominates the latter under the GMC criterion.*

Proposition 2. *A design T can have GMC only if it minimizes $g(T)$.*

As an immediate application of Proposition 2, suppose $f = L_w, 2 \leq w \leq n - m - 1$. Then by (2.5) and (4.1), $g(T)$ vanishes if and only if $\overline{T}$ is a $(w-1)$-flat of P, i.e., an L_w-subset of P that is closed (up to proportionality) under the formation of nonnull linear combinations. Since all $(w-1)$-flats are isomorphic, it is clear that in this case a design T has GMC if and only if $\overline{T}$ is a $(w-1)$-flat. Such a design is also known to have MA and MEC (Suen, Chen and Wu, 1997; Cheng and Mukerjee, 1998).

Turning to two-level factorials, we now obtain GMC designs for $3 \leq f \leq 15$. This will facilitate comparison with Tang and Wu (1996) who gave MA designs for $3 \leq f \leq 11$. The cases $f = 3, 7$ and 15 are settled from the discussion in the last paragraph. The following lemma helps in the remaining cases.

Lemma 4. *Let $s = 2$.*

(a) If $f = 4, 5$ or 6, then a design T can have GMC only if $\overline{T}$ is contained in a 2-flat.

(b) If $8 \leq f \leq 14$, then a design T can have GMC only if $\overline{T}$ is contained in a 3-flat.

Proof. Only (b) is proved. The proof of (a) is similar and simpler. Let $8 \leq f \leq 14$. Then $n - m \geq 4$. If $n - m = 4$, then $\overline{T}$ is contained in a 3-flat for every design and there is nothing to prove. Suppose $n - m \geq 5$. For $8 \leq f \leq 14$, we can always choose $\overline{T}$ as a subset of a 3-flat. For such a choice, say $\overline{T}_0$, all sums involving a pair of points of $\overline{T}_0$ belong to the same 3-flat. Since this 3-flat has 15 points of which f are in $\overline{T}_0$, there are at most $15 - f$ points outside $\overline{T}_0$ which equal one of these sums. By (2.5) and (4.1), therefore, $g(T_0) \leq 15 - f$. In view of Proposition 2, it now suffices to show that

$$g(T) > 15 - f, \tag{4.2}$$

whenever $\overline{T}$ is not contained in a 3-flat, i.e., $\text{rank}[V(\overline{T})] \geq 5$. If $\text{rank}[V(\overline{T})] = \rho$, then $\overline{T}$ has ρ linearly independent points which span $\binom{\rho}{2}$ additional points as pairwise sums. These $\binom{\rho}{2}$ points are distinct and, among them, at most $f - \rho$ are in $\overline{T}$, i.e., at least $\binom{\rho}{2} - (f - \rho)$ $[= \binom{\rho+1}{2} - f]$ are outside $\overline{T}$. Hence by (2.5) and (4.1), $g(T) \geq \binom{\rho+1}{2} - f$, and the truth of (4.2) follows for $\rho \geq 6$. For $\rho = 5$, we get $g(T) \geq 15 - f$. If equality holds here then $\overline{T}$ consists of the five independent points, say $\pi_1, \ldots, \pi_5$, and $f - 5\,(> 0)$ of the

10 pairwise sums arising out of $\pi_1, \ldots, \pi_5$. The remaining $15 - f$ of these 10 pairwise sums are outside $\overline{T}$. In addition, if say, $\pi_1 + \pi_2$ is one of the $f - 5$ pairwise sums in $\overline{T}$, then both $\pi_1 + \pi_2$ and π_3 are in $\overline{T}$ but not their sum. Thus, outside $\overline{T}$, there are at least $16 - f$ points each of which equals the sum of two points of $\overline{T}$. Hence by (2.5) and (4.1), the truth of (4.2) follows again. $\square$

For two-level factorials, Chen and Hedayat (1996) showed that if $2^{w-1} \le f \le 2^w - 1$, then the leading term A_3 in the WLP is minimized only if $\overline{T}$ is contained in a $(w-1)$-flat. One may wonder if this result can lead to a more general version of Lemma 4. However, as $A_3 = \frac{1}{3}\sum_{k\ge 1} k\,{}_1^{\#}C_2^{(k)}$ by (2.2), no exact connection between minimization of A_3 and the GMC criterion, that incorporates sequential maximization of ${}_1^{\#}C_2^{(k)}$, $0 \le k \le K_2$, emerges in an obvious manner. At any rate, Lemma 4 itself will suffice for the present purpose of characterizing GMC designs for $3 \le f \le 15$. The next example illustrates how it reduces the search for $\overline{T}$ through the use of the catalog of 16-run designs given by Chen, Sun and Wu (1993).

Example 2. Let $s = 2$ and $f = 11$. By Lemma 4(b), a design T can have GMC only if $\overline{T}$ is contained in a 3-flat. But then $\overline{T}$ itself represents a 2^{11-7} design, and following Chen, Sun and Wu (1993), one needs to consider only three nonisomorphic choices of $\overline{T}$, namely,

$$\overline{T_1} = \{1,2,3,4,12,13,23,14,24,134,234\},\ \overline{T_2} = \{1,2,3,4,12,13,23,123,14,24,34\},$$
$$\overline{T_3} = \{1,2,3,4,12,13,23,123,14,24,124\}.$$

The same notation as in Example 1 is used here for the points of P. Considering the pairwise sums arising out of $\overline{T_1}$, $\overline{T_2}$ and $\overline{T_3}$, we get $\delta(T_1) = (0^{n-4}, 4, 5, 5, 5)$, $\delta(T_2) = \delta(T_3) = (0^{n-4}, 4, 4, 4, 4)$, where 0^u is the null row vector of order u. Hence by Proposition 1, both T_2 and T_3 dominate T_1 under the GMC criterion. Moreover, as $\delta(T_2) = \delta(T_3)$, by (3.3), the designs T_2 and T_3 have the same ${}_1^{\#}C_2$, and one has to consider the next term ${}_2^{\#}C_2$ in the AENP (2.3) in order to compare them. To that effect, one can proceed as in Example 1 and employ (3.4) to show that for T_2,

$$\begin{aligned}
&{}_2^{\#}C_2^{(\frac{1}{2}n-7)} = (\tfrac{1}{2}n-6)(n-4),\ {}_2^{\#}C_2^{(\frac{1}{2}n-3)} = 9(\tfrac{1}{2}n-2),\\
&{}_2^{\#}C_2^{(\frac{1}{2}n-2)} = 3(n-2),\ {}_2^{\#}C_2^{(k)} = 0 \text{ for every other } k,
\end{aligned}$$

while for T_3,

$$\begin{aligned}
&{}_2^{\#}C_2^{(\frac{1}{2}n-7)} = (\tfrac{1}{2}n-6)(n-4),\ {}_2^{\#}C_2^{(\frac{1}{2}n-3)} = 6(n-4),\\
&{}_2^{\#}C_2^{(\frac{1}{2}n-1)} = \tfrac{3}{2}n,\ {}_2^{\#}C_2^{(k)} = 0 \text{ for every other } k.
\end{aligned}$$

Since $n \ge 20$ for $f = 11$, it follows that T_3 yields a larger ${}_2^{\#}C_2^{(\frac{1}{2}n-3)}$ than T_2. Thus

T_3 represents the GMC design. Interestingly, here T_2 is the MA design (Tang and Wu, 1996) and hence the MA and GMC criteria differ. Observe that the comparison of T_2 and T_3 under the GMC criterion involves the use of (3.4) which, in turn, requires explicit knowledge of G_3 as shown in (2.6). □

For $s = 2$ and $3 \le f \le 15$, Table 1 shows the sets $\overline{T}$ for GMC designs. In particular, this table yields 64-run $2^{n-(n-6)}$ GMC designs for $n \ge 48$ and 128-run $2^{n-(n-7)}$ designs for $n \ge 112$, and hence supplements the tables in ZLZA. A comparison with Tang and Wu (1996) shows that the designs shown in Table 1 also have MA except when $f = 10\,(n - m \ge 5)$ and $f = 11$.

We next consider three-level factorials and obtain GMC designs for $3 \le f \le 13$. This will facilitate comparison with Suen, Chen and Wu (1997) who gave MA designs over the same range of f. The case $f = 3$ is straightforward and the cases $f = 4$ and 13 are settled from the discussion below Proposition 2. The following lemma, with a proof analogous to that of Lemma 4, helps in the other cases.

Lemma 5. *Let $s = 3$. If $5 \le f \le 12$, then a design T can have GMC only if $\overline{T}$ is contained in a 2-flat.*

Example 3. Let $s = 3$ and $f = 5$. In view of Lemma 5, following Chen, Sun and Wu (1993) as in the previous example, one needs to consider only three nonisomorphic choices of $\overline{T}$, namely, $\overline{T_1} = \{1, 2, 3, 12, 12^2 3\}$, $\overline{T_2} = \{1, 2, 3, 12, 13\}$ and $\overline{T_3} = \{1, 2, 3, 12, 12^2\}$. Note that $\overline{T_3}$ is the same as $\overline{T}$ of Example 1. Hence proceeding as in that example, $\delta(T_1) = (0^{n-8}, 1, 2, 2, 2, 2, 2, 2, 4)$, $\delta(T_2) = (0^{n-8}, 1, 1, 1, 1, 2, 2, 3, 3)$, $\delta(T_3) = (0^{n-8}, 1, 1, 1, 1, 1, 1, 1, 1)$, and it is clear from Proposition 1 that T_3 is the GMC design. □

For $s = 3$ and $3 \le f \le 13$, Table 2 shows the sets $\overline{T}$ for GMC designs. In particular, this table yields 27-run $3^{n-(n-3)}$ GMC designs for $4 \le n \le 13$. A comparison with Suen, Chen and Wu (1997) shows that the designs shown in Table 2 also have MA. This, however, does not imply that the two criteria lead to the same ranking of all designs over this range of f. For instance with $f = 8$, even though the best design remains the same for both criteria, the rankings of other designs differ.

Acknowledgment

The authors are thankful to the editors and referees for their careful examination and encouraging comments to the paper. This work was supported by the NNSF of China grant No. 10571093 and the SRFDP of China grant No. 20050055038. Mukerjee's research was also supported by the Visiting Scholar Program at Chern Institute of Mathematics and a grant from CMDS, Indian Institute of Management Calcutta.

Appendix

Table 1. The sets $\overline{T}$ for 2^{n-m} designs

f	$\overline{T}$
3	$\{1, 2, 12\}$
4	$\{1, 2, 12, 3\}$
5	$\{1, 2, 12, 3, 13\}$
6	$\{1, 2, 12, 3, 13, 23\}$
7	$\{1, 2, 12, 3, 13, 23, 123\}$
8	$\{1, 2, 12, 3, 13, 23, 123, 4\}$
9	$\{1, 2, 12, 3, 13, 23, 123, 4, 14\}$
$10\,(n-m=4)$	$\{1, 2, 12, 3, 13, 23, 4, 14, 24, 34\}$
$10\,(n-m\geq 5)$	$\{1, 2, 12, 3, 13, 23, 123, 4, 14, 24\}$
11	$\{1, 2, 12, 3, 13, 23, 123, 4, 14, 24, 124\}$
12	$\{1, 2, 12, 3, 13, 23, 123, 4, 14, 24, 124, 34\}$
13	$\{1, 2, 12, 3, 13, 23, 123, 4, 14, 24, 124, 34, 134\}$
14	$\{1, 2, 12, 3, 13, 23, 123, 4, 14, 24, 124, 34, 134, 234\}$
15	$\{1, 2, 12, 3, 13, 23, 123, 4, 14, 24, 124, 34, 134, 234, 1234\}$

Table 2. The sets $\overline{T}$ for 3^{n-m} designs

f	$\overline{T}$
3	$\{1, 2, 12\}$
4	$\{1, 2, 12, 12^2\}$
5	$\{1, 2, 12, 12^2, 3\}$
6	$\{1, 2, 12, 12^2, 3, 13\}$
7	$\{1, 2, 12, 12^2, 3, 12^23, 12^23^2\}$
8	$\{1, 2, 12, 12^2, 3, 23^2, 12^23, 12^23^2\}$
9	$\{1, 2, 12^2, 3, 13^2, 23^2, 123^2, 12^23, 12^23^2\}$
10	$\{1, 2, 12, 12^2, 3, 13, 13^2, 23, 23^2, 123\}$
11	$\{1, 2, 12, 12^2, 3, 13, 13^2, 23, 23^2, 123, 123^2\}$
12	$\{1, 2, 12, 12^2, 3, 13, 13^2, 23, 23^2, 123, 123^2, 12^23^2\}$
13	$\{1, 2, 12, 12^2, 3, 13, 13^2, 23, 23^2, 123, 123^2, 12^23, 12^23^2\}$

References

Chen, J., Sun, D.X. and Wu, C.F.J. (1993). A catalogue of two-level and three-level fractional factorial designs with small runs. *Internat. Statist. Rev.* **61**, 131-145.

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

Cheng, C.S. and Mukerjee, R. (1998). Regular fractional factorial designs with minimum aberration and maximum estimation capacity. *Ann. Statist.* **26**, 2289-2300.

Mukerjee, R. and Wu, C.F.J. (1999). Blocking in regular fractional factorials: a projective geometric approach. *Ann. Statist.* **27**, 1256-1271.

Mukerjee, R. and Wu, C.F.J. (2006). *A Modern Theory of Factorial Designs.* Springer, New York.

Suen, C.Y., Chen, H. and Wu, C.F.J. (1997). Some identities on q^{n-m} designs with application to minimum aberration designs. *Ann. Statist.* **25**, 1176-1188.

Tang, B. and Wu, C.F.J. (1996). Characterization of minimum aberration 2^{n-k} designs in terms of their complementary designs. *Ann. Statist.* **24**, 1914-1926.

Wu, C.F.J. and Hamada, M. (2000). *Experiments: Planning, Analysis and Parameter Design Optimization.* Wiley, New York.

Zhang, R.C., Li, P., Zhao, S.L. and Ai, M.Y. (2007). A general minimum lower-order confounding criterion for two-level regular designs. *Statist. Sinica,* to appear.

LPMC and School of Mathematical Sciences, Nankai University, Tianjin 300071, China

E-mail: zhrch@nankai.edu.cn.

Indian Institute of Management Calcutta, Kolkata 700104, India

E-mail: rmuk1@hotmail.com.