Statistica Sinica **26** (2016), 1001-1019
doi:http://dx.doi.org/10.5705/ss.202015.0214

OPTIMAL TWO-LEVEL REGULAR DESIGNS UNDER BASELINE PARAMETRIZATION VIA COSETS AND MINIMUM MOMENT ABERRATION

Rahul Mukerjee and Boxin Tang

Indian Institute of Management Calcutta and Simon Fraser University

Abstract: We consider two-level fractional factorial designs under a baseline parametrization that arises naturally when each factor has a control or baseline level. While the criterion of minimum aberration can be formulated as usual on the basis of the bias that interactions can cause in the estimation of main effects, its study is hindered by the fact that level permutation of any factor can impact such bias. This poses a serious challenge especially in the practically important highly fractionated situations where the number of factors is large. We address this problem for regular designs via explicit consideration of the principal fraction and its cosets, and obtain certain rank conditions which, in conjunction with the idea of minimum moment aberration, are seen to work well. The role of simple recursive sets is also examined with a view to achieving further simplification. Details on highly fractionated minimum aberration designs having up to 256 runs are provided.

Key words and phrases: Bias, level permutation, minimum aberration, orthogonal array, principal fraction, rank condition, simple recursive set, wordlength.

1. Introduction

Fractional factorial designs have been widely studied in the recent literature, with particular emphasis on their exploration under the minimum aberration (MA) and allied model robustness criteria; see Mukerjee and Wu (2006), Xu, Phoa, and Wong (2009) and Cheng (2014) for surveys and further references. While a vast majority of this work centers around the usual orthogonal parametrization (OP), a baseline parametrization (BP) for factorial designs has started gaining attention in recent years. It arises naturally in many situations where each factor has a control or a baseline level. An example, from Kerr (2006), is given by a toxicological study with binary factors, each representing the presence or absence of a toxin, the state of absence being a natural baseline level of each factor. The BP has found use in microarray experiments (Yang and Speed (2002)). It can also arise in agricultural or industrial experiments, with the currently used level of each factor constituting the baseline level.

Optimal paired comparison designs for full factorials under BP were investigated by several authors in the context of microarrays; see Banerjee and Mukerjee

(2008), Zhang and Mukerjee (2013), and the references there. The study of factorial fractions under BP was initiated by Mukerjee and Tang (2012). Focusing on the two-level case in view of its popularity among practitioners, they observed that orthogonal arrays (OAs) of strength two ensure optimal estimation of main effects when interactions are absent, and hence explored MA designs as OAs which sequentially minimize the bias that interactions of successively higher orders can cause in the estimation of main effects. Further results on two-level MA designs were reported by Li, Miller, and Tang (2014). Very recently, Miller and Tang (2015) obtained certain useful formulae for the bias terms under BP in the case of two-level regular designs.

As noted by these authors, BP has a special feature that significantly complicates the task of finding MA designs - level permutation of any factor can influence the bias terms which the MA criterion seeks to minimize. As a result, with m two-level factors, one needs to account for all the 2^m possible factor level permutations in any OA. This looks formidable, if not impossible, when m is large and, precisely because of this reason, existing tables of two-level MA designs under BP (Mukerjee and Tang (2012), Li, Miller, and Tang (2014)) cover only up to 19 factors. Even the formulae in Miller and Tang (2015), as they stand, are very hard to apply for large m.

In the present paper, we continue with two-level regular designs and build on the findings in Miller and Tang (2015). Through explicit consideration of the interplay between the principal fraction and its cosets, we obtain certain rank conditions which, jointly with the idea of minimum moment aberration (MMA; (Xu (2003)), are seen to work well especially for large m, i.e., in highly fractionated situations which are of practical importance due to their economy. It is also seen that simple recursive sets, introduced recently by Tang and Xu (2014) in a different context, play an effective role in achieving further simplification. We present the MMA formulation for BP in the next section. The main results appear in Section 3 preceded by a brief review of the relevant background material for regular designs. Design tables and other details are given in Section 4 and we conclude in Section 5 with some remarks on future work.

There are several reasons, in addition to their popularity among users, for considering regular designs as done here. First, they are very prospective, e.g., Mukerjee and Tang (2012) found that 16-run regular designs having MA under BP enjoy the same property also among all designs. Therefore, it is of natural interest to investigate how far the existing rich literature on regular de-signs under OP can be exploited under BP. An even more compelling reason is that our results on regular designs provide an important benchmark against which any future work on the nonregular case has to be compared. Unless the regular case is well understood, there is no way of assessing, through future research, whether nonregular designs are more advantageous or not. Indeed, a complete listing of nonisomorphic OAs for large m is neither available nor likely to emerge in the foreseeable future and our findings in the regular case will certainly provide an attractive option until such discovery takes place. Finally, as noted in Section 4, regular designs tend to compare very favorably with some nonregular designs that have been of recent interest.

2. Minimum Moment Aberration

For ease of reference, we first introduce BP and the MA criterion under this parametrization, following Mukerjee and Tang (2012). Then the MMA formulation is presented and its advantages discussed. The contents of this section apply to both regular and nonregular designs.

If there are two factors, each at levels 0 and 1 with 0 as the control or baseline level, then under BP, the effects of the four treatment combinations are expressed as

$$\tau_{00} = \theta_0, \quad \tau_{10} = \theta_0 + \theta_1, \quad \tau_{01} = \theta_0 + \theta_2, \quad \tau_{11} = \theta_0 + \theta_1 + \theta_2 + \theta_{12},$$

where θ_0 is the baseline effect, θ_1 and θ_2 are the two main effects, and θ_{12} represents the two-factor interaction. This can be readily extended to m two-level factors using heavier notation. With a 2^m factorial and BP as above, consider now an N-run design, where each treatment combination is obviously a binary m-tuple. Let $Z = (z_{uj})$, $1 \le u \le N$, $1 \le j \le m$, be the $N \times m$ binary design matrix with rows given by these N treatment combinations. As noted by Mukerjee and Tang (2012), in the absence of interactions, the design estimates each of the m main effects with the smallest possible variance if and only if Z forms an OA of strength (at least) two. This is just as in OP, and hence in the spirit of what is done under OP (Tang and Deng (1999)), one can discriminate among such OAs by taking cognizance of the bias that interactions of successive orders can cause in the estimation of main effects. From this perspective, in conformity with the effect hierarchy principle, Mukerjee and Tang (2012) proposed choosing Z as an OA which sequentially minimizes $K_2, \ldots, K_m$, where K_s is a measure of bias due to the s-factor interactions. In order to present the expression for K_s as given by them, let Ω_s be the set of s-tuples $g_1 \cdots g_s$, $1 \le g_1 < \cdots < g_s \le m$, and for any $g_1 \cdots g_s \in \Omega_s$, let $c(g_1 \cdots g_s)$ be the binary $N \times 1$ vector with the uth element $\prod_{l=1}^{s} z_{ug_l}$, $1 \le u \le N$. Then from their equations (4)$-$(6),

$$K_s = 4N^{-2} \sum_{\Omega_s} c(g_1 \cdots g_s)' W W' c(g_1 \cdots g_s), \; 2 \le s \le m, \tag{2.1}$$

where the sum $\sum_{\Omega_s}$ extends over $g_1 \cdots g_s \in \Omega_s$, the primes denote transposition, and

$$W = J_{Nm} - 2Z, \tag{2.2}$$

with J_{Nm} as the $N \times m$ matrix of ones. Here W is obtained from Z replacing 0 and 1 there by 1 and -1, respectively.

A major problem with (2.1) is that the sum $\sum_{\Omega_s}$ becomes unmanageable for large m, unless s is small or close to m. The following result alleviates this difficulty. Here $(ZZ')^{[s]}$ is the s-fold Schur product of ZZ', any element of $(ZZ')^{[s]}$ is the sth power of the corresponding element of ZZ'.

Lemma 1. *Sequential minimization of* $K_2, \ldots, K_m$ *is equivalent to that of* $M_2, \ldots,$ M_m*, where*

$$M_s = N^{-2} tr\{(ZZ')^{[s]}WW'\}, \quad 2 \le s \le m.$$

Proof. Denote the N rows of W by $w'_{(1)}, \ldots, w'_{(N)}$. Then from (2.1),

$$K_s = 4N^{-2}\text{tr}\{H_s WW'\} = 4N^{-2}\sum_{u=1}^{N}\sum_{v=1}^{N} H_s(u,v)\{w'_{(u)}w_{(v)}\}, \quad 2 \le s \le m, \tag{2.3}$$

where $H_s = \sum_{\Omega_s} c(g_1 \cdots g_s)c(g_1 \cdots g_s)'$ is a square matrix of order N and $H_s(u,v)$ is the (u,v)th element of H_s. By the definition of $c(g_1 \cdots g_s)$,

$$H_s(u,v) = \sum_{\Omega_s}\prod_{l=1}^{s}(z_{ug_l}z_{vg_l}), \quad 1 \le u, v \le N. \tag{2.4}$$

Let $T(u,v)$ be the set of indices j such that $z_{uj} = z_{vj} = 1$. Since Z is binary, the product in (2.4) equals 1 if $\{g_1, \ldots, g_s\} \subseteq T(u,v)$, and 0 otherwise. Hence, from (2.4), writing $t(u,v)$ for the cardinality of $T(u,v)$, we get $H_s(u,v) = \binom{t(u,v)}{s}$. Using this in (2.3),

$$K_s = 4N^{-2}\sum_{u=1}^{N}\sum_{v=1}^{N}\binom{t(u,v)}{s}\{w'_{(u)}w_{(v)}\}, \quad 2 \le s \le m. \tag{2.5}$$

Now, write $z'_{(1)}, \ldots, z'_{(N)}$ for the rows of Z, and observe that

$$t(u,v) = z'_{(u)}z_{(v)}, \quad 1 \le u, v \le N. \tag{2.6}$$

Hence, using (2.2) and the fact that Z is an OA of strength two, after some simplification,

$$\sum_{u=1}^{N}\sum_{v=1}^{N} t(u,v)\{w'_{(u)}w_{(v)}\} = \text{tr}(ZZ'WW') = \frac{mN^2}{4},$$

which does not depend on the design. So, by (2.5), sequential minimization of $K_2, \ldots, K_m$ is equivalent to that of $\sum_{u=1}^{N}\sum_{v=1}^{N}\{t(u,v)\}^s\{w'_{(u)}w_{(v)}\}$, $2 \le s \le m$. The result now follows from (2.6).

We call $M_2, \ldots, M_m$ the moment sequence due to their similarity with moments, and a design sequentially minimizing $M_2, \ldots, M_m$ is called an MMA design. While Lemma 1 shows the equivalence of the MA and MMA criteria, the M_s do not involve any sum over Ω_s, allow direct matrix calculation, and hence are much easier to compute than the K_s. Indeed, consideration of the M_s can also facilitate theoretical results. For instance, they allow a proof of Lemma 2 in the next section which, though not necessarily shorter than the original proof in Miller and Tang (2015), is more straightforward in the sense of eliminating the case enumeration in the original proof. We omit the details to save space.

These points are akin to those in Xu (2003) regarding MA vis-à-vis MMA under OP. But there is a major difference. MMA is dictated under OP by numbers of positions where pairs of rows of Z have the same level, whereas under BP it is dictated by numbers of positions where both rows in such pairs have 1. This is due to the asymmetry between the levels of any factor under BP.

3. Main Results

3.1. Background material

In what follows, all vector and matrix operations, including rank statements, are over the finite field $GF(2)$. A regular design $d(B, y)$, for a 2^m factorial in $N = 2^r$ $(2 \leq r < m)$ runs, is specified by (a) a set of m distinct nonnull $r \times 1$ binary vectors $b_1, \ldots, b_m$ such that the $r \times m$ matrix $B = [b_1 \cdots b_m]$ has rank r, and (b) a $1 \times m$ binary vector $y = (y_1, \ldots, y_m)$. The design consists of the N treatment combinations obtained by adding y to each of the N vectors in $R(B)$, the row space of B. Given B, there are 2^{m-r} distinct designs of this form, as $d(B, y)$ and $d(B, y^*)$ are identical if $y - y^* \in R(B)$ due to the subgroup structure of $R(B)$. We call $d(B, y)$ the principal fraction if it contains the treatment combination $(0, \ldots, 0)$, and a coset thereof otherwise. Clearly, the principal fraction is given by $R(B)$ itself and each coset is obtained by level permutation of one or more factors in the principal fraction. Hence, the principal fraction and the cosets are anticipated to play a crucial role under BP.

Let 0_r be the null column vector of order r. Then the wordlength pattern of the design $d(B, y)$ is given by the sequence $(A_3, \ldots, A_m)$, with

$$A_s = \sum_{\Omega_s} \phi(b_{g_1}, \ldots, b_{g_s}), \quad 3 \leq s \leq m, \tag{3.1}$$

where $\sum_{\Omega_s}$ is as in (2.1) and, for any $g_1 \cdots g_s \in \Omega_s$, $\phi(b_{g_1}, \ldots, b_{g_s})$ equals 1 or 0 according as whether $b_{g_1} + \cdots + b_{g_s}$ equals 0_r or not, respectively. The resolution of the design is the smallest s such that $A_s > 0$. With reference to $d(B, y)$, we also define

$$A_s^0 = \textstyle\sum_{\Omega_s}^0 \phi(b_{g_1}, \ldots, b_{g_s}), \quad A_s^1 = \textstyle\sum_{\Omega_s}^1 \phi(b_{g_1}, \ldots, b_{g_s}), \tag{3.2}$$

the sum $\sum_{\Omega_s}^{l}$ being over $g_1 \cdots g_s \in \Omega_s$ such that $y_{g_1} + \cdots + y_{g_s} = l \pmod 2$; $l = 0, 1$. Note that B alone determines A_s, whereas A_s^0 and A_s^1 depend on y as well. Thus a regular MA design under OP, which sequentially minimizes $A_3, \ldots, A_m$, is determined by B alone. We are now in a position to present a lemma from Miller and Tang (2015).

Lemma 2. *For any regular design,*
(a) $K_2 = m(m-1)/4 + (3/4)A_3$, *and*
(b) $K_3 = (1/16)\{3\binom{m}{3} + 4A_4 + 3(m-4)A_3^0 + 3mA_3^1\}$.
(c) *Furthermore, if* $A_3 = 0$, *then* $K_4 = (1/64)\{4\binom{m}{4} + 5A_5 + 4(m-1)A_4^0 + 4(m-5)A_4^1\}$.

Lemma 2(a), applicable to nonregular designs as well, is also implicit in Mukerjee and Tang (2012) while Miller and Tang (2015) gave a more general version of (c) without the condition $A_3 = 0$. However, the present form of (c) will suffice for our purpose.

3.2. Rank conditions and their application

As a first step towards finding the regular MA design under BP, we need to sequentially minimize K_2 and K_3. By Lemma 2(a), (b), this calls for

(i) characterizing B so as to sequentially minimize A_3 and A_4, and if the smallest possible A_3 is positive which happens for $m > N/2$, then
(ii) for every B as in (i), characterizing $y = (y_1, \ldots, y_m)$ so that

$$b_{g_1} + b_{g_2} + b_{g_3} = 0_r \;\Rightarrow\; y_{g_1} + y_{g_2} + y_{g_3} = 0 \pmod 2,\;\; \forall\, g_1g_2g_3 \in \Omega_3. \quad (3.3)$$

Condition (ii) is evident from (3.1) and (3.2), because $A_3 = A_3^0 + A_3^1$ and A_3^0 has a smaller coefficient than A_3^1 in K_3, by Lemma 2(b). While (3.3) is obviously met by any y in the principal fraction, we need to characterize all such y in order to assess their possible impact on $K_4, \ldots, K_m$.

To that end, suppose $m > N/2$. For any given B, define Q_3 as the $A_3 \times m$ matrix such that each $g_1g_2g_3 \in \Omega_3$ with $b_{g_1} + b_{g_2} + b_{g_3} = 0_r$ contributes a row to Q_3 having 1 in the g_1th, g_2th, g_3th positions, and 0 elsewhere. Clearly, $BQ_3' = 0$, so that $R(B) \subseteq \bar{R}(Q_3)$, where $\bar{R}(Q_3)$ is the ortho-complement of the row space of Q_3. Since $\text{rank}(B) = r$, this yields $r \le m - \rho$ or $\rho \le m - r$, where $\rho = \text{rank}(Q_3)$. If $\rho < m - r$, in which case $R(B)$ is a proper subspace of $\bar{R}(Q_3)$, let $\tilde{B}$ be an $(m - \rho - r) \times m$ matrix such that the rows of $[B' \; \tilde{B}']'$ form a basis of $\bar{R}(Q_3)$, and write $R(\tilde{B})$ for the row space of $\tilde{B}$. Using the standard softwares for matrix calculation, suitably adapted to $GF(2)$, one can obtain Q_3, ρ, and $\tilde{B}$ readily – in fact, up to $N = 128$ runs, almost instantaneously. We now have a result summarizing two useful rank conditions.

Proposition 1. *Suppose $m > N/2$ and consider any B.*
(a) *If $\rho = m - r$, then y meets* (3.3) *if and only if y is in the principal fraction.*
(b) *If $\rho < m - r$, then y meets* (3.3) *if and only if y is in $d(B, \tilde{y})$ for some $\tilde{y} \in R(\tilde{B})$.*

Proof. By the definition of Q_3, y meets (3.3) if and only if $y \in \bar{R}(Q_3)$. If $\rho = m - r$, then $\bar{R}(Q_3) = R(B)$, and (a) follows. Else, if $\rho < m - r$, then the rows of $[B' \ \tilde{B}']'$ span $\bar{R}(Q_3)$. Therefore, y meets (3.3) if and only if $y - \tilde{y} \in R(B)$ for some $\tilde{y} \in R(\tilde{B})$, and (b) follows.

Proposition 1 goes a long way in reducing the complexity due to factor level permutations which, for regular designs, is manifest in the principal fraction and its cosets. The gains are particularly significant in highly fractionated situations where even the smallest possible A_3 is large and hence ρ is often close, if not equal, to $m - r$. Given B, if $\rho = m - r$, then one needs to consider only the principal fraction. On the other hand, if $\rho < m - r$, then it suffices to take care of only as many as $2^{m-\rho-r}$ designs $d(B, \tilde{y})$, $\tilde{y} \in R(\tilde{B})$, which typically form a much smaller subclass of the totality of the 2^{m-r} distinct designs $d(B, y)$ arising from B. The next result is immediate from Proposition 1(a). Here $\underline{0}$ denotes the $1 \times m$ vector of zeros.

Theorem 1. *Let $m > N/2$. If up to isomorphism, there is a unique B, say B_0, which sequentially minimizes A_3 and A_4, and the condition $\rho = m - r$ holds for B_0, then the principal fraction $d(B_0, \underline{0})$ has MA among all regular designs under BP.*

This has wide-ranging applications. For instance, it applies to $19 \leq m \leq 31$ if $N = 32$, and $m = 36$ as well as $39 \leq m \leq 63$ if $N = 64$. Examples 1 and 2 below illustrate its use. More generally, for $m > N/2$, if any of the conditions in Theorem 1 fails, then the following procedure, illustrated in Examples 3 and 4 below, turns out to be quite convenient. The last step of the procedure involves calculation of the moment sequence $M_2, \ldots, M_m$ which, as discussed earlier, is much easier than computing $K_2, \ldots, K_m$.

Step I. List all nonisomorphic B which sequentially minimize A_3 and A_4. Existing catalogs of regular designs, such as the one in Xu (2009), together with complementary design theory as reviewed in Mukerjee and Wu (2006, Chap. 3), are helpful for this purpose.

Step II. For every such B, consider the principal fraction $d(B, \underline{0})$ if $\rho = m - r$, or the designs $d(B, \tilde{y})$, $\tilde{y} \in R(\tilde{B})$, if $\rho < m - r$. Let D be the class of all designs so obtained. By Proposition 1, the designs in D are the only ones that sequentially minimize K_2 and K_3 among regular designs.

Step III. Find an MMA design in D. By Lemma 1, this design has MA in D and hence among all regular designs under BP.

Here are a few examples. We follow the standard practice of representing any nonnull binary vector $b = (b(1), \ldots, b(r))'$ by the number $\sum_{l=1}^{r} b(l)2^{l-1}$, so $(1,0,0,0,1)'$ and $(0,1,1,0,1,1)'$ are denoted by 17 and 54, respectively.

Example 1. Let $N = 32$ and $m = 28$. From Xu (2009), up to isomorphism, there is a unique B, say

$$B_0 = [1\ 2\ 4\ 8\ 16\ 31\ 7\ 11\ 21\ 25\ 13\ 14\ 19\ 22\ 26\ 28\ 3\ 5\ 9\ 17\ 15\ 23\ 27\ 29\ 6\ 10\ 18\ 30],$$

that sequentially minimizes A_3 and A_4. Here $r = 5$ and, upon finding Q_3, one can check that $\rho = \text{rank}(Q_3) = 23 = m - r$. Hence by Theorem 1, the principal fraction $d(B_0, \underline{0})$ has MA among all regular designs under BP.

Example 2. Let $N = 256$ and $m = 245$. By complementary design theory, up to isomorphism, there is a unique B, say B_0, which sequentially minimizes A_3 and A_4. Following Tang and Wu (1996), B_0 has all nonnull binary 8×1 vectors, except 1, 2, 3, 4, 5, 6, 8, 9, 10 and 12, as columns. Here again, $\rho = 237 = m - r$, and the MA property of the principal fraction $d(B_0, \underline{0})$ holds as before by Theorem 1.

Example 3. Let $N = 32$ and $m = 18$. From Xu (2009), up to isomorphism, there is a unique B, say

$$B_0 = [1\ 2\ 4\ 8\ 16\ 31\ 7\ 11\ 21\ 25\ 13\ 14\ 19\ 22\ 26\ 28\ 3\ 5],$$

that sequentially minimizes A_3 and A_4. Here $\rho = 12 < m - r$, and $\tilde{B}$ consists of the single row

$$1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0.$$

It suffices to consider designs $d(B_0, \tilde{y})$, $\tilde{y} \in R(\tilde{B})$. There are only two such designs and, comparing their moment sequences, we find that the design $d(B_0, \tilde{y}_0)$, where $\tilde{y}_0$ is the row of $\tilde{B}$ as shown above, has MMA and hence MA among regular designs under BP.

Example 4. Let $N = 64$ and $m = 37$. Complementary design theory, used in conjunction with Xu's (2009) catalog, shows that, up to isomorphism, there are two choices of B, say B_1 and B_2, that sequentially minimize A_3 and A_4. The choice B_1 has all nonnull binary 6×1 vectors except
1, 2, 4, 8, 16, 31, 7, 11, 21, 13, 14, 26, 3, 17, 23, 9, 27, 29, 5, 19, 28, 6, 10, 18, 12, and 15
as columns, while B_2 has all such vectors except
1, 2, 4, 8, 16, 31, 7, 11, 21, 25, 13, 14, 19, 22, 26, 28, 3, 5, 9, 17, 15, 23, 10, 18, 6, and 24

as columns. For both B_1 and B_2, it turns out that $\rho = 31 = m - r$. Hence one needs to consider only the two principal fractions $d(B_1, \underline{0})$ and $d(B_2, \underline{0})$. Comparing their moment sequences, we find that $d(B_1, \underline{0})$ has MMA and hence MA among regular designs under BP. Incidentally, B_1 also has MA under OP and is isomorphic to the design 37-31.1 shown in Mee (2009, p.491).

3.3. Simple recursive sets

Some of the developments in the last subsection are closely linked with the idea of simple recursive sets considered recently by Tang and Xu (2014) for three-level regular designs with quantitative factors. We now examine how this idea helps in avoiding actual rank calculation in many situations, especially in the highly fractionated case. In our context, a set S of distinct nonnull $r \times 1$ binary vectors is called simple recursive, if there exist r linearly independent vectors, say $b_1, \ldots, b_r$, in S and a sequence $S_0 \subset S_1 \subset \cdots \subset S_q$ of sets of vectors, such that $S_0 = \{b_1, \ldots, b_r\}$ and

$$S_{l+1} = S_l \cup \{b : b \in S, b = a_1 + a_2 \text{ where } a_1, a_2 \in S_l\},\ 0 \leq l \leq q-1, \quad S_q = S. \tag{3.4}$$

With a view to illustrating the recursive process in (3.4) clearly, we present an example where any nonnull binary vector $(b(1), \ldots, b(r))'$ is represented by the string $1^{b(1)} \cdots r^{b(r)}$, with the convention that $j^{b(j)}$ is dropped if $b(j) = 0$, e.g., the vector $(1\ 0\ 1\ 0\ 1)'$, which was earlier denoted by 21, is now represented as 135. Thus the addition of two such vectors amounts to multiplication of the corresponding strings with squared symbols dropped.

Example 5. Let S consist of the columns of B_0 in Example 1. In our present notation,

$$\begin{aligned} S = \{&1, 2, 3, 4, 5, 12345, 123, 124, 135, 145, 134, 234, 125, 235, 245,\\ &345, 12, 13, 14, 15, 1234, 1235, 1245, 1345, 23, 24, 25, 2345\}. \end{aligned}$$

It is readily seen that S is simple recursive because it meets (3.4) with

$$\begin{aligned} &S_0 = \{1, 2, 3, 4, 5\}, \quad S_1 = S_0 \cup \{12, 13, 14, 15, 23, 24, 25\},\\ &S_2 = S_1 \cup \{123, 124, 125, 134, 135, 234, 235, 145, 245, 1234, 1235, 1245\},\\ &S_3 = S_2 \cup \{1345, 2345, 12345, 345\} = S. \end{aligned}$$

Example 5 shows the set of columns of B_0 to be simple recursive, and earlier in Example 1, the condition $\rho = m - r$ of Proposition 1(a) was seen to hold for B_0. We present a general result in this direction which links simple recursive sets with this rank condition.

Proposition 2. *Let S be a set of m distinct nonnull $r \times 1$ binary vectors and B be the $r \times m$ matrix with these vectors as columns. If S is simple recursive then B satisfies the rank condition $\rho = m - r$.*

Proof. If S is simple recursive then there exists a sequence of sets $S_0 \subset S_1 \subset \cdots \subset S_q$ such that S_0 contains r linearly independent vectors of S, and (3.4) holds. Arrange the columns of B such that the first r columns are the vectors in S_0, followed by columns given by the vectors in S_1 but not in S_0, and so on. With columns so arranged, if we write $B = [b_1 \cdots b_m]$, then by (3.4), for each $r+1 \le j \le m$, there exist $\delta(j)$ and $\epsilon(j)$ satisfying $1 \le \delta(j) < \epsilon(j) < j$, such that $b_j = b_{\delta(j)} + b_{\epsilon(j)}$. Let G be an $(m-r) \times m$ binary matrix with rows indexed by $r+1, \ldots, m$ and columns indexed by $1, \ldots, m$, such that row j of G has 1 in positions $\delta(j)$, $\epsilon(j)$, and j, and 0 elsewhere, $r+1 \le j \le m$. Then

(a) $BG' = 0$,

(b) G has three ones in each row, and

(c) $G = [G_1 \; G_2]$, where G_1 is $(m-r) \times r$ and G_2 is a lower triangular matrix of order $m - r$ with each diagonal element 1.

By (a) and (b), each row of G is also a row of the Q_3 introduced earlier, while by (c), $\text{rank}(G) = m - r$. So, $\rho = \text{rank}(Q_3) \ge m - r$, which completes the proof because as noted earlier, we also have $\rho \le m - r$.

Our next result shows a general structure of S that ensures the simple recursive property. Let F_r be the space of the 2^r binary vectors of order $r \times 1$, F_{r-1} be any $(r-1)$-dimensional subspace of F_r, and $\bar{F}$ be the complement of F_{r-1} in F_r. Consider

$$S = \bar{F} \cup F, \tag{3.5}$$

where F is any subset of nonnull vectors of F_{r-1}.

Proposition 3. *If F contains $r-1$ linearly independent vectors, then the set S in* (3.5) *is simple recursive.*

Proof. The case $r = 2$ is trivial. With $r \ge 3$, let $f_1, \ldots, f_{r-1}$ be linearly independent vectors in F and f_r be any vector in $\bar{F}$. For $1 \le l \le r-1$, write E_l for the set of the $\binom{r-1}{l}$ vectors of the form $f_r + f$, where f is the sum of any l of $f_1, \ldots, f_{r-1}$, e.g., $E_1 = \{f_r + f_1, \ldots, f_r + f_{r-1}\}$, etc. Clearly,

$$\bar{F} = \{f_r\} \cup E_1 \cup \cdots \cup E_{r-1}. \tag{3.6}$$

Consider now the sets

$$V_0 = \{f_1, \ldots, f_{r-1}, f_r\}, \qquad V_{l+1} = V_l \cup E_{l+1}, \;\; 0 \le l \le r-2, \quad V_r = S.$$

From (3.5), (3.6) and the definition of $E_1, \ldots, E_{r-1}$, observe that V_0 consists of r linearly independent members of S and that

$$V_{l+1} \subseteq V_l \cup \{b : b \in S, b = a_1 + a_2 \text{ where } a_1, a_2 \in V_l\}, \quad 0 \le l \le r-1.$$

This is similar to (3.4) with the only change that the equality in (3.4) connecting S_{l+1} with S_l is now replaced by the set inclusion ($\subseteq$) connecting V_{l+1} with V_l. Hence it is clear that if we take $S_0 = V_0$ and obtain $S_1, S_2, \ldots$, recursively as in (3.4), then $V_1 \subseteq S_1$, $V_2 \subseteq S_2$, and so on. As $V_r = S$, it follows that the process will end up with $S_q = S$, for some $q \le r$. This guarantees the existence of a sequence $S_0 \subset S_1 \subset \cdots \subset S_q$ of sets meeting (3.4), and completes the proof.

Propositions 2 and 3 lead to a result that significantly narrows the search for the regular MA design under BP, or even pinpoints it over a wide range of m, without rank calculation. Here m_j is the largest m such that a regular m-factor two-level design having resolution five or higher exists in 2^j runs. For instance, from Mukerjee and Wu (2006) and Xu (2009), $m_2 = 2$, $m_3 = 3$, $m_4 = 5$, $m_5 = 6$, $m_6 = 8$, $m_7 = 11$.

Theorem 2. *Let $m \ge N/2 + m_{r-2} + 1$.*

(a) *If $B_1, \ldots, B_p$ represent all nonisomorphic choices of B which sequentially minimize A_3 and A_4, then the MA design under BP among the principal fractions $d(B_1, \underline{0}), \ldots, d(B_p, \underline{0})$ also enjoys the same MA property among all regular designs.*

(b) *In particular, if up to isomorphism there is a unique B, say B_0, that sequentially minimizes A_3 and A_4, then the principal fraction $d(B_0, \underline{0})$ has MA among all regular designs under BP.*

Proof. With reference to any set F as in (3.5), let $A_3(F)$ and $A_4(F)$ denote, respectively, the numbers of triplets and quadruplets formed by the vectors in F that are linearly dependent, adding to 0_r; cf. (3.1). Consider any B that sequentially minimizes A_3 and A_4. By complementary design theory, the set, S, of columns of B must

(i) have the structure in (3.5), with

(ii) the set F sequentially minimizing $A_3(F)$ and $A_4(F)$ among all subsets of F_{r-1} that have the same cardinality as F and consist of nonnull vectors.

Since S consists of the m columns of B, by (i), $m = 2^{r-1} + (\#F)$, where $\#F$ is the cardinality of F. As $m \ge N/2 + m_{r-2} + 1$ and $N = 2^r$, $\#F > m_{r-2}$. As a result, if there are at most $r-2$ linearly independent vectors in F, then either $A_3(F) > 0$ or $A_3(F) = 0$, $A_4(F) > 0$. Clearly, in this situation there exists a nonnull vector, say f_0, in F_{r-1} which is not spanned by the vectors in

F. If $A_3(F) > 0$, then $\tilde{F}$ $(\subseteq F_{r-1})$ obtained from F by replacing any vector in F appearing in a linearly dependent triplet by f_0 has the same cardinality as F but entails $A_3(\tilde{F}) < A_3(F)$, contradicting (ii) above. A similar contradiction is reached if $A_3(F) = 0, A_4(F) > 0$. Thus F must contain $r-1$ linearly independent vectors. By Propositions 2 and 3, therefore, B satisfies the rank condition $\rho = m - r$, and the theorem follows from Proposition 1(a).

For $m \geq N/2 + m_{r-2} + 1$, Theorem 2 considerably simplifies Step II of the procedure described in the previous subsection and makes Examples 1 and 2 there more transparent. However, it does not cover Examples 3 and 4 where the need for rank calculation remains. We remark that Theorem 2 comes quite close to capturing all situations where $m > N/2$ and the regular MA design under BP is given by a principal fraction. For example, with 32, 64 and 128 runs, $r = 5, 6$ and 7, Theorem 2 tells that this should happen for $m \geq 20, 38$, and 71, respectively, while as reported in the next section, rank calculation shows that this actually happens for $m \geq 19, 36$ and 70, respectively. In addition to providing a neat theoretical result, Theorem 2 is practically useful for large N, such as $N = 512$, and correspondingly large m, where direct calculation of Q_3 and ρ can be slow. An illustrative example follows. To save space, we revert to the notation of the previous subsection for nonnull binary vectors, with any such vector denoted by a single number.

Example 6. Let $N = 512$. Then Theorem 2 applies to $m \geq 268$. Consider $m = 462$. By complementary design theory, together with Xu's (2009) catalog, up to isomorphism, there are three choices of B, say B_1, B_2 and B_3, that sequentially minimize A_3 and A_4. Of these, B_1 has all nonnull binary 9×1 vectors except those in the complement of

$$\{1, 2, 4, 8, 16, 32, 31, 39, 41, 51, 13, 21, 11, 52\}$$

in $\{1, 2, \ldots, 63\}$ as columns. Similarly, B_2 and B_3 have all such vectors except those in the complements of

$$\{1, 2, 4, 8, 16, 32, 31, 39, 41, 51, 42, 21, 22, 52\},$$

$$\text{and} \quad \{1, 2, 4, 8, 16, 32, 31, 39, 41, 51, 13, 21, 11, 46\},$$

respectively, in $\{1, 2, \ldots, 63\}$ as columns. By Theorem 2(a), it suffices to consider only the three principal fractions $d(B_1, \underline{0})$, $d(B_2, \underline{0})$ and $d(B_3, \underline{0})$. On the basis of $M_2, \ldots, M_5$ alone, we find that $d(B_1, \underline{0})$ has smaller MMA than the two other designs. Thus $d(B_1, \underline{0})$ has MMA and hence MA among regular designs under BP. We note that B_1 also entails MA under OP.

3.4. The case $m \le N/2$

If $m \le N/2$, then this approach does not work because the smallest possible A_3 is 0, and, for any B with $A_3 = 0$, (3.3) leading to Proposition 1 does not arise. By Lemma 2, as a first step towards finding the MA design, now one needs to
(i)′ characterize B with $A_3 = 0$ and, subject to this condition, minimize A_4; and if the minimum A_4 so obtained is positive, then
(ii)′ for every B as in (i)′, characterize y so that A_4^1 is the largest possible.
Condition (i)′ ensures sequential minimization of K_2 and K_3, and as $A_4 = A_4^0 + A_4^1$, then (ii)′ minimizes the contribution of $4(m-1)A_4^0 + 4(m-5)A_4^1$ to K_4 without affecting the term A_5 there. Because of (3.2) and in the hope of finding a counterpart of Proposition 1, one may wonder if, along the lines of (3.3), condition (ii)′ amounts to characterizing $y = (y_1, \ldots, y_m)$ so that

$$b_{g_1}+b_{g_2}+b_{g_3}+b_{g_4} = 0_r \quad \Rightarrow \quad y_{g_1}+y_{g_2}+y_{g_3}+y_{g_4} = 1 \pmod 2, \quad \forall\, g_1g_2g_3g_4 \in \Omega_4. \tag{3.7}$$

This turns out to be too ambitious because, unlike with (3.3), a choice of B meeting (i)′ may not admit any y that satisfies a condition as strong as (3.7). Thus if $N = 32$ and $m = 8$, then from Xu (2009), up to isomorphism, there is a unique $B = [1\ 2\ 4\ 8\ 16\ 15\ 19\ 21]$ meeting (i)′. This B has $A_3 = 0$ and $A_4 = 3$, i.e., three members of Ω_4 satisfy $b_{g_1} + b_{g_2} + b_{g_3} + b_{g_4} = 0_r$, and one can check that the relationship $y_{g_1} + y_{g_2} + y_{g_3} + y_{g_4} = 1 \pmod 2$ holds for at most two of these three, whatever be the choice of y.

In view of the above, unlike with $m > N/2$, a drastic reduction of the design problem does not seem to be possible for $m \le N/2$. Nevertheless, a matrix formulation and consideration of MMA allow us to make some progress and to suggest a procedure below on the basis of (i)′ and (ii)′. Given B, here $C(B)$ is a set of 2^{m-r} choices of y which account for the principal fraction and all its cosets, the designs $d(B, y)$, $y \in C(B)$, are distinct; for instance, if the first r columns of B are linearly independent, then $C(B)$ can be taken as the set of all y with 0 in first r positions.

<u>Step I</u>. List all nonisomorphic B that have $A_3 = 0$ and, subject to this condition, minimize A_4.

<u>Step II</u>. (a) If the minimum A_4 is positive, then for every B listed in Step I, find the subset $C_0(B)$ of $C(B)$ consisting of y which maximize A_4^1; by (3.2), this is facilitated by the fact that A_4^1 equals the number of ones in yQ_4' where, in the same manner as Q_3, the $A_4 \times m$ matrix Q_4 is constructed from $g_1g_2g_3g_4 \in \Omega_4$ satisfying $b_{g_1} + b_{g_2} + b_{g_3} + b_{g_4} = 0_r$.
(b) If the minimum A_4 is 0, then for every B listed in Step I, take $C_0(B) = C(B)$.

<u>Step III</u>. Find an MMA design over the class D of all $d(B, y)$ such that B is listed in Step I and $y \in C_0(B)$. By Lemma 1, this design also has MA in D and hence among all regular designs under BP.

In Step II, if (a) arises then $C_0(B)$ is often much smaller than $C(B)$, while if (b) arises then typically m is small and hence $C(B)$ itself is quite small; e.g., with $N = 32$ or 64, (b) arises only for $m = 6$ or $m = 7$ and 8, respectively. This simplifies the implementation of Step III where consideration of MMA also helps. Indeed, as illustrated in Example 7 below, this procedure works well for $N = 32$ and 64, where it yields regular MA designs under BP for all $m \leq N/2$, thus complementing our earlier results. However, Step II itself calls for maximization of A_4^1 over the 2^{m-r} choices of y in $C(B)$, and this becomes formidable for $N \geq 128$, unless m is relatively small.

Example 7. Let $N = 64$ and $m = 23$. From Xu's (2009) catalog, up to isomorphism, there are two choices of B which have $A_3 = 0$ and, subject to this condition, minimize A_4. These are

$$B_1 = [1\ 2\ 4\ 8\ 16\ 32\ 31\ 35\ 13\ 52\ 14\ 55\ 37\ 61\ 11\ 19\ 21\ 44\ 7\ 62\ 25\ 49\ 22],$$

$$\text{and} \quad B_2 = [1\ 2\ 4\ 8\ 16\ 32\ 31\ 35\ 13\ 52\ 14\ 55\ 37\ 61\ 11\ 19\ 21\ 44\ 7\ 62\ 25\ 22\ 41].$$

Step II yields $C_0(B_1)$ and $C_0(B_2)$ with respective sizes 6 and 96, both much smaller than the size 2^{m-r} of any $C(B)$. Thus the class D in Step III has 102 designs and, comparing their moment sequences, we find that the design $d(B_1, y)$, where

$$y = (0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1),$$

has MMA and hence MA among regular designs under BP. Note that B_1 also has MA under OP.

4. Design Tables and More Details

Along the lines of these examples, we now apply the techniques in Section 3 to describe and tabulate regular MA designs under BP for $N = 32$, 64, 128, and 256. For $N = 32$ and 64, all m are covered, while for $N = 128$ and 256, we cover large m where our main interest lies.

The following notation and conventions are used in this section:

(a) The B entailing MA under OP is denoted by B^*. Over the range of N and m considered here, this B^* is unique up to isomorphism and can be found either directly from Xu (2009) or Mee (2009), or by using complementary design theory in addition.

(b) As before, $\underline{0}$ stands for the $1 \times m$ vector of zeros.

(c) The design tables show both B and y to make the correspondence between the two clear.

(d) The B for a larger m is often conveniently expressed in terms of the B for a smaller m, e.g., the B for $m = 8$ in Table 1 is shown as $[B(m = 7)\ 21]$ to indicate that it is obtained by including the vector represented by 21 at the end of the B for $m = 7$ in the same table.

(e) The binary vector $y = (y_1, \ldots, y_m)$ is written simply as $y_1 \cdots y_m$. While exhibiting y in Table 3, we also write 0^u or 1^u to denote a string of u zeros or ones.

Our findings in this section on regular MA designs under BP are summarized below.

$\underline{N = 32}$: For $19 \leq m \leq 31$, the design $d(B^*, \underline{0})$ has MA. Table 1 shows MA designs for $6 \leq m \leq 18$. In this table, the B reported for each m has MA under OP.

$\underline{N = 64}$: For $m = 7, 8$ as well as $36 \leq m \leq 63$, the design $d(B^*, \underline{0})$ has MA. Table 2 shows MA designs for $9 \leq m \leq 35$. In this table, the B reported for each m has MA under OP, except for $m = 26$, where it is the second best under OP (Xu (2009)).

$\underline{N = 128}$: For $70 \leq m \leq 127$, the design $d(B^*, \underline{0})$ has MA. Table 3 shows MA designs for $65 \leq m \leq 69$. Thus every $m > N/2$ is covered. In Table 3, the B reported for each m has MA under OP, except for $m = 69$, where it is the second best under OP by complementary design theory.

$\underline{N = 256}$: For $192 \leq m \leq 255$, the design $d(B^*, \underline{0})$ has MA.

It is satisfying to observe that over the ranges of m considered here, BP and OP are in perfect agreement with regard to the choice of B under the MA criterion for $N = 32$ and 256, whereas their agreement is almost perfect for $N = 64$ and 128. From Mukerjee and Tang (2012), we also see that for $m = N - 1$ and $m = N - 2$, the saturated and nearly saturated cases, the designs reported above have MA under BP among all designs, regular or not.

We now briefly comment on how regular MA designs compare under BP with an important class of nonregular designs, namely quaternary code (QC) designs, which were introduced by Xu and Wong (2007) and have been of recent interest. The notion of wordlength pattern can be extended to these designs via the J-characteristics of Tang and Deng (1999). If $N = 64$ then, following Miller and Tang (2015), under BP the MA QC design dominates the MA regular design for $m = 13$ and 14; it is the other way round for $m = 15$ and 16. This is the same as under OP. For large m relative to N, which is the main thrust of this paper, there is not yet a single instance of the MA QC design having less aberration than the MA regular design under OP though there are quite a few situations where the

Table 1. Regular MA designs $d(B, y)$ under BP for $N = 32$ and $6 \le m \le 18$.

m	B	y
6	[1 2 4 8 16 31]	000001
7	[1 2 4 8 16 15 19]	0000001
8	$[B(m = 7)\quad 21]$	00000001
9	$[B(m = 8)\quad 25]$	000000011
10	$[B(m = 9)\quad 30]$	0000000011
11	[1 2 4 8 16 31 7 11 21 25 13]	00000001110
12	$[B(m = 11)\quad 14]$	000000010110
13	$[B(m = 12)\quad 19]$	0000000101101
14	$[B(m = 13)\quad 22]$	00000000001111
15	$[B(m = 14)\quad 26]$	000000000011111
16	$[B(m = 15)\quad 28]$	0000000000111111
17	$[B(m = 16)\quad 3]$	00011001000010000
18	$[B(m = 17)\quad 5]$	111000100000000000

reverse happens. Given the close conformity between BP and OP as seen above, we anticipate the same pattern also under BP. As a test case, let $N = 32$, where QC designs are well defined for $m \le 24$. Using the results in Mukerjee and Tang (2013) on minimization of A_3 for QC designs, together with Lemma 2(a) and a complete enumeration of all factor level permutations, we found MA QC designs under BP for $16 \le m \le 24$. In agreement with OP (Xu and Wong (2007)), it was seen that they are worse than their regular counterparts for $m = 20$ and 21, and make a tie for other m in this range. Thus, from available indications, regular designs tend to compare very favorably with QC designs under BP for large m.

5. Concluding Remarks

The present work leads to several open issues. The first of these concerns a comprehensive study of nonregular designs under BP. While this is likely to be very hard in general, it is of interest to explore QC designs in some detail, given their structured nature.

Even for regular designs, the case $m \le N/2$ turns out to be more difficult than $m > N/2$. Results that strengthen our findings in this case and further reduce the design search would be very useful.

The case of more general factorials including mixed factorials opens up new challenges. Under BP, Mukerjee and Tang (2012) found that OAs may not entail optimal estimation of the main effects beyond the two-level case even in the absence of interactions. Thus, in such general settings, formulation of the MA criterion itself becomes difficult. Recently, Mukerjee and Huda (2015) investigated model robust efficient designs under BP for general factorials under a minimaxity criterion. This was in the spirit of the corresponding work by

Table 2. Regular MA designs $d(B, y)$ under BP for $N = 64$ and $9 \le m \le 35$.

m	B	y
9	[1 2 4 8 16 32 31 39 41]	000000001
10	$[B(m = 9)$ 51]	0000000011
11	$[B(m = 10)$ 42]	00000000011
12	$[B(m = 11)$ 60]	000000000011
13	$[B(m = 11)$ 21 22]	0000000000111
14	$[B(m = 10)$ 13 21 11 52]	00000000001011
15	$[B(m = 14)$ 58]	000000000000111
16	$[B(m = 15)$ 22]	0000000000011111
17	$[B(m = 16)$ 25]	00000000000111110
18	$[B(m = 17)$ 28]	000000000011100111
19	$[B(m = 18)$ 46]	0000000000011011011
20	$[B(m = 19)$ 61]	00000000000011110111
21	[1 2 4 8 16 32 31 35 13 52 14 55 37 61 11 19 21 44 7 62 25]	000000001001000011111
22	$[B(m = 21)$ 49]	0000000010010000111111
23	$[B(m = 22)$ 22]	00000000100100001111111
24	$[B(m = 23)$ 41]	000000000011010011101111
25	$[B(m = 24)$ 38]	0000000000000111110011011
26	$[B(m = 25)$ 50]	00000000000001111100110111
27	$[B(m = 25)$ 26 28]	000000000000011111001101110
28	$[B(m = 27)$ 42]	0000000000000000111111111011
29	$[B(m = 28)$ 47]	00000000000000000111011111101
30	$[B(m = 29)$ 50]	000000000000000011110111111111
31	$[B(m = 30)$ 56]	0000000000000000000011111111111
32	$[B(m = 31)$ 59]	00000000000000000000111111111111
33	[1 2 4 8 16 32 7 11 13 14 19 21 25 31 35 37 44 52 55 61 62 49 22 41 38 26 28 42 47 50 56 59 63]	100100000000101100101011111000000
34	$[B(m = 33)$ 60]	0000010000000111100000000011100000000
35	$[B(m = 34)$ 43]	0001010100000011001100000000100000000

Table 3. Regular MA designs $d(B, y)$ under BP for $N = 128$ and $65 \le m \le 69$.

m	B	y
65	[1 64-127]	$0^{23}1^{8}0^{4}1^{4}0^{4}1^{4}001^{6}00110^{6}$
66	[1 2 64-127]	$0^{14}1^{8}0^{12}1^{12}0^{4}1^{4}0^{12}$
67	[1 2 4 64-127]	$0^{19}1^{8}0^{8}1^{8}0^{24}$
68	[1 2 4 8 64-127]	$0^{4}1^{16}0^{48}$
69	[1 2 4 8 15 64-127]	$0^{5}1^{16}0^{48}$

Yin and Zhou (2015) under OP. Any future result connecting this line of research with some version of MA would be illuminating.

Acknowledgement

We thank a referee for helpful comments. The work of RM was supported by the J.C. Bose National Fellowship of the Government of India and a grant from the Indian Institute of Management Calcutta. The work of BT was supported by the Natural Science and Engineering Research Council of Canada.

References

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

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

Kerr, K. F. (2006). Efficient 2^k factorial designs for blocks of size 2 with microarray applications. *J. Qual. Technol.* **38**, 309-318.

Li, P., Miller, A. and Tang, B. (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.* Springer, Berlin.

Miller, A. and Tang, B. (2015). Using regular fractions of two-level designs to find baseline designs. *Statist. Sinica* (to appear).

Mukerjee, R. and Huda, S. (2015). Approximate theory-aided robust efficient factorial fractions under baseline parametrization. *Ann. Inst. Statist. Math.* (to appear).

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

Mukerjee, R. and Tang, B. (2013). A complementary set theory for quaternary code designs. *Ann. Statist.* **41**, 2768-2785.

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

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

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**, 2549-2559.

Tang, Y. and Xu, H. (2014). Permuting regular fractional factorial designs for screening quantitative factors. *Biometrika* **101**, 333-350.

Xu, H. (2003). Minimum moment aberration for nonregular designs and supersaturated designs. *Statist. Sinica* **13**, 691-708.

Xu, H. (2009). http://www.stat.ucla.edu/∼hqxu/pub/ffd2r/

Xu, H., Phoa, F. K. H. and Wong, W. K. (2009). Recent developments in nonregular fractional factorial designs. *Statist. Surveys* **3**, 18-46.

Xu, H. and Wong, A. (2007). Two-level nonregular designs from quaternary linear codes. *Statist. Sinica* **17**, 1191-1213.

Yang, Y. H. and Speed, T. (2002). Design issues for cDNA microarray experiments. *Nature Genetics* (*Supplement*) **3**, 579-588.

Yin, Y. and Zhou, J. (2015). Minimax design criterion for fractional factorial designs. *Ann. Inst. Statist. Math.* **67**, 673-685.

Zhang, R. and Mukerjee, R. (2013). Highly efficient factorial designs for cDNA microarray experiments: use of approximate theory together with a step-up step-down procedure. *Stat. Appl. Genet. Mol. Biol.* **12**, 489-503.

Indian Institute of Management Calcutta, Joka, Diamond Harbour Road, Kolkata 700 104, India.

E-mail: rmuk0902@gmail.com

Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada.

E-mail: boxint@sfu.ca

(Received June 2015; accepted August 2015)