

Contents lists available at ScienceDirect

## Journal of Complexity

journal homepage: www.elsevier.com/locate/jco



# An algorithmic approach to finding factorial designs with generalized minimum aberration

### Fasheng Sun<sup>a</sup>, Min-Qian Liu<sup>a,\*</sup>, Wenrui Hao<sup>b</sup>

<sup>a</sup> Department of Statistics, School of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071, China <sup>b</sup> Department of Scientific Computing and Applied Software, School of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071, China

#### ARTICLE INFO

Article history: Received 11 January 2008 Accepted 22 August 2008 Available online 7 September 2008

Keywords: Discrepancy Factorial design Generalized minimum aberration Quadratic penalty function Uniformity

#### ABSTRACT

Factorial designs are arguably the most widely used designs in scientific investigations. Generalized minimum aberration (GMA) and uniformity are two important criteria for evaluating both regular and non-regular designs. The generation of GMA designs is a non-trivial problem due to the sequential optimization nature of the criterion. Based on an analytical expression between the generalized wordlength pattern and a uniformity measure, this paper converts the generation of GMA designs to a constrained optimization problem, and provides effective algorithms for solving this particular problem. Moreover, many new designs with GMA or near-GMA are reported, which are also (nearly) optimal under the uniformity measure.

© 2008 Published by Elsevier Inc.

#### 1. Introduction

Factorial designs are arguably the most widely used experimental designs in industrial and scientific investigations. Their practical success is due to the efficient use of experimental runs to study many factors simultaneously. From different viewpoints, various optimality criteria have been proposed for design construction and comparison. The maximum resolution criterion proposed by Box and Hunter [1] and minimum aberration criterion by Fries and Hunter [14] are two most successful optimality criteria. These two criteria show the rationality under the effect hierarchy principle. However, both of them are defined only for regular designs and they cannot be used to evaluate factorial designs in general. Recently, generalized minimum aberration (GMA) was proposed by Tang and Deng [24] for the two-level non-regular case, by Ma and Fang [21] for the multi-level case, and by Xu and Wu [27] for the mixed-level case.

\* Corresponding author.

E-mail address: mqliu@nankai.edu.cn (M.-Q. Liu).

<sup>0885-064</sup>X/\$ – see front matter 0 2008 Published by Elsevier Inc. doi:10.1016/j.jco.2008.08.002

For a design matrix  $P = (p_{ii})$ , which has n runs and s factors each having q levels, i.e.  $p_{ii} =$  $0, \ldots, q-1, i = 1, \ldots, n, j = 1, \ldots, s$ , let  $X_j$  be the contrast matrix of *j*-factor interactions of *P*. For j = 1, ..., s, if  $X_j = (x_{ik}^j)$ , let

$$A_{j}(P) = \frac{1}{n^{2}} \sum_{k} \left| \sum_{i=1}^{n} x_{ik}^{j} \right|^{2},$$
(1)

then  $W(P) = (A_1(P), \ldots, A_s(P))$  is called the generalized wordlength pattern (GWP) by Xu and Wu [27], and the index of the first non-zero element corresponds to the resolution. For a design of resolution III or higher, the GWP is usually simplified as  $W(P) = (A_3(P), \ldots, A_s(P))$ . The generalized minimum aberration (GMA) criterion proposed by them is to sequentially minimize  $A_i(P)$  for i =1,..., s. They showed that the definition of  $A_i(P)$  in (1) is invariant with respect to the choice of orthonormal contrasts and thus model free and also the GMA reduces to the minimum aberration for regular designs and the minimum  $G_2$ -aberration (Tang and Deng [24]) for two-level non-regular designs.

The GMA is also equivalent to the minimum generalized aberration proposed by Ma and Fang [21] for multi-level designs. Given the design matrix  $P = (p_{ij})$ , Ma and Fang [21] defined

$$A_j^g(P) = \frac{1}{n(q-1)} \sum_{k=0}^s P_j(k;s) E_k(P), \quad j = 1, \dots, s,$$
(2)

where  $P_j(k; s) = \sum_{r=0}^{j} (-1)^r (q-1)^{j-r} {k \choose r} {s-k \choose j-r}$  is the Krawtchouk polynomial  $\left( {x \choose y} = 0 \text{ for } x < y \right)$ ,  $E_k(P) = n^{-1} |\{(a, b)|a, b \in P, d_H(a, b) = k\}|, |A|$  denotes the cardinality of A, and  $d_H(a, b)$  is the Hamming distance between two runs a and b. Fang et al. [10] showed that

$$A_j(P) = (q-1)A_j^g(P) \tag{3}$$

for the multi-level case, in particular,  $A_j(P) = A_j^g(P)$  when q = 2. The GMA criterion is difficult and expensive to compute, because its definition involves a complicated coding of factorial effects that include all main effects and interactions (Xu [25]). In addition, it is very hard to operate with GMA because the GWP is a vector. There exist only a few approaches to the construction of GMA designs. Fang et al. [10] proposed the RBIBD method for constructing GMA multi-level supersaturated designs, which are of resolution II. For designs of resolution III or higher, Butler [2,3] developed alternative methods for constructing minimum  $G_2$ aberration two-level designs; Butler [4] obtained some GMA designs by projecting specific saturated orthogonal arrays; Xu [26] derived some GMA non-regular designs from the Nordstrom-Robinson code; Fang et al. [13] recently provided a formal optimization treatment on optimal designs with GMA, and proposed a general sub-design selection algorithm, which utilizes their newly developed lower bounds and optimality conditions. Note that, for two-level designs, the GMA criterion is a relaxed variant of the minimum G-aberration proposed by Deng and Tang [7], while Deng and Tang [8], Sun, Li and Ye [23] and Li, Deng and Tang [18] constructed many two-level orthogonal designs with minimum G-aberration.

The above drawbacks of GMA can be overcome if we can convert the vector problem to a scalar problem. The discrepancy measure of uniformity can play a key role for this. The discrepancy is another important measure used for evaluating factorial designs (Hickernell [15]; Fang et al. [11]). It measures how much the empirical distribution of the design points departs from the uniform distribution (Hickernell [16]). Recently Hickernell and Liu [17] defined a general discrepancy which has been proved to be a function of  $A_i(P)$ 's, i.e.

$$D^{2}(P;\gamma) = \frac{1}{n^{2}} \sum_{i,k=1}^{n} \prod_{j=1}^{s} \left\{ 1 + \gamma \left( -1 + q \delta_{p_{ij} p_{kj}} \right) \right\} - 1$$
(4)

$$=\sum_{j=1}^{s}\gamma^{j}A_{j}(P),$$
(5)

where  $\gamma$  is an arbitrary positive number. From a uniformity point of view, for a fixed number of points, n, a design with low discrepancy is preferred (Fang and Wang [12]).

This paper will find a surrogate for GWP based on the connection between GWP and the discrepancy in (5), which can reduce the computation and generate GMA designs more conveniently. The paper is organized as follows. Section 2 converts the problem of finding GMA designs to a constrained optimization problem, and Section 3 discusses the algorithms for solving this optimization problem. Many newly generated GMA or near-GMA designs are tabulated in Section 4, which are also (nearly) optimal under the discrepancy measure of uniformity. Section 5 contains some further discussions.

#### 2. Reformulation of the problem of finding GMA designs

To reformulate the problem of finding GMA designs, we need the following lemma.

**Lemma 1.** If  $a_i \ge 0$ ,  $b_i \ge 0$ , for i = 1, ..., k, and  $a_1 < b_1$ , then there exists an  $r_0 > 0$ , such that  $a_1r + \cdots + a_kr^k < b_1r + \cdots + b_kr^k$  for  $0 < r < r_0$ .

Proof. Since

$$\lim_{r \to 0} (a_1 + \dots + a_k r^{k-1}) = a_1 < b_1 = \lim_{r \to 0} (b_1 + \dots + b_k r^{k-1}),$$

there exists an  $r_0 > 0$  such that  $a_1 + \cdots + a_k r^{k-1} < b_1 + \cdots + b_k r^{k-1}$  for  $0 < r < r_0$ , namely,  $a_1r + \cdots + a_k r^k < b_1r + \cdots + b_k r^k$  for  $0 < r < r_0$ .  $\Box$ 

From Lemma 1, we obtain the following theorem.

**Theorem 1.** There exists a  $\gamma_0 > 0$ , such that minimizing  $D^2(P; \gamma)$  is equivalent to finding a GMA design when  $0 < \gamma < \gamma_0$ .

Let  $X = (x_{ij})$  be a full design matrix with *s* factors each having *q* levels,  $x_{ij} = 0, ..., q - 1$ ,  $i = 0, ..., q^s - 1, j = 1, ..., s$ . The level combinations of the full design matrix *X* are arranged lexicographically, e.g. the first level combination is (0, ..., 0, 0), the second one is (0, ..., 0, 1), and the last one is (q - 1, ..., q - 1, q - 1). Let **y** be a  $q^s \times 1$  vector, where the *i*th component  $y_i$  is *k* if the *i*th level combination of *X* repeats *k* times in design *P*,  $i = 0, ..., q^s - 1$ . Hence **y** satisfies the constraint

$$\mathbf{y}'\mathbf{1}_{q^{\mathrm{S}}}=n,\tag{6}$$

and  $y_i$  is a non-negative integer,  $i = 0, ..., q^s - 1$ , where  $\mathbf{1}_n$  denotes the  $n \times 1$  vector of ones. Then from (4) we obtain:

#### Lemma 2.

$$D^{2}(P;\gamma) = \frac{1}{n^{2}} \boldsymbol{y}' \boldsymbol{B}_{s} \boldsymbol{y} - 1,$$
(7)

where  $\mathbf{B}_s = \bigotimes^s \mathbf{B}_0$ ,  $\mathbf{B}_0 = (b_{ij})_{q \times q}$ ,  $b_{ij} = 1 + \gamma (q-1)$  for i = j and  $1 - \gamma$  otherwise, i, j = 1, ..., q, and  $\bigotimes$  denotes the Kronecker product.

From Theorem 1 and Lemma 2, it can be seen that given a sufficiently small positive value of  $\gamma$ , the problem of finding a GMA design can be transformed to the problem of finding a vector  $\mathbf{y}$  which minimizes (7) under the constraint (6) and  $y_i$  being a non-negative integer for  $i = 0, ..., q^s - 1$ .

Theorem 4.1 of Cheng and Ye [6] shows that the sum of GWP elements is larger for designs with higher degrees of replication, therefore they tend to have higher aberration than those with less replicates, so in this paper, for finding GMA designs via computer algorithms, we change the constraints on y to

$$\begin{cases} \mathbf{y}' \mathbf{1}_{q^{s}} = n, & \text{and} \\ y_{i} = 0 \text{ or } 1, & \text{for } i = 0, \dots, q^{s} - 1. \end{cases}$$
(8)

**Remark 1.** For some cases, if there does not exist any orthogonal array without replicates, our algorithm cannot find an orthogonal design under the constraints in (8), but can find a design with a smaller sum of GWP elements. For example, Cheng [5] showed that there is a unique  $OA(12, 2^4)$  which has 11 distinct runs. For this case, we can find a 12-run, 4-factor, two-level design with GWP (0, 1/9, 2/9, 0); the sum of the elements of this GWP is 1/3. However, the GWP of the unique  $OA(12, 2^4)$  is (0, 0, 4/9, 1/9), and the sum is 5/9 which is greater than 1/3. In addition, the constraints in (8) can accelerate the computer search.

#### 3. Optimization methods

The problem discussed in the previous section can be converted into the following optimization problem.

#### 3.1. A general algorithm

One fundamental approach to solving a constrained optimization problem is to replace the original problem by a penalty function that consists of (i) the original objective of the constrained optimization problem, and (ii) one additional term for each constraint, which is positive when the vector **y** violates that constraint and zero otherwise. There are many penalty functions available, among them a simple and commonly used one is the *quadratic penalty function*, in which the penalty terms are the squares of the constraint violations.

For the problem of generating GMA designs, we consider the optimization problem of finding a vector y to minimize (7) under the constraints in (8). The quadratic penalty function can be constructed as

$$\min_{\mathbf{y}} Q(\mathbf{y}; \mu) = \frac{1}{n^2} \mathbf{y}' \mathbf{B}_{s} \mathbf{y} - 1 + 1/(2\mu) \left( \sum_{i} y_i - n \right)^2 + 1/(2\mu) \sum_{i} \left( y_i (1 - y_i) \right)^2, \tag{9}$$

where  $\mu > 0$  is the *penalty parameter*. By driving  $\mu$  to zero, we penalize the constraint violations with increasing severity.

A general algorithm based on the penalty function can be specified as follows.

#### Algorithm 1.

Given  $\mu_0 > 0$ , a tolerance  $\tau_0 > 0$ , a  $\tau_{step} > 0$ , and a starting vector  $\mathbf{y}_0$ ; **for** k = 1, 2, ...**do** Use the Newton method (Nocedal and Wright [22]) to find an approximate

minimizer  $\mathbf{y}_k$  of  $Q(\cdot; \mu_{k-1})$ :

start with  $y_{k-1}$ , and terminate when  $\|\nabla Q(y; \mu_{k-1})\| \le \tau_{k-1}$ , where  $\nabla Q$  is the gradient of function Q;

if final convergence test is satisfied  $(||\mathbf{y}_k - \mathbf{y}_{k-1}|| \le \tau_{step})$  then **stop** with approximate solution  $\mathbf{y}_k$ ;

Choose a new tolerance  $\tau_k \in (0, \tau_{k-1})$ ;

Choose a new penalty parameter  $\mu_k \in (0, \mu_{k-1})$ ;

#### end for

**Remark 2.** We can choose any  $\tau_k$  and  $\mu_k$  as long as  $\tau_k \in (0, \tau_{k-1})$  and  $\mu_k \in (0, \mu_{k-1})$ . In the following, we will take  $\tau_k = \tau_{k-1}/2$  and  $\mu_k = \mu_{k-1}/2$ , which will be shown to perform very well in Section 4.

#### 3.2. Selection of $y_0$

To carry out Algorithm 1, we need a starting vector  $y_0$ . Although it can be randomly generated, the algorithm may converge slowly. Thus a suitable selected starting vector is called for. Generating a starting vector for Algorithm 1 can be regarded as finding the *shortest path of a graph* (Diestel [9]). We can regard  $y_i$  as a vertex and then define the weight of the vertex  $y_i$  as  $1/n^2(\mathbf{B}_s)_{ii}$  (the *ii*th element of  $\mathbf{B}_s$ )

and the weight of the edge connecting  $y_i$  and  $y_j$  as  $1/n^2(\mathbf{B}_s)_{ij}$  (the *ij*th element of  $\mathbf{B}_s$ ). Thus we modify Dijkstra's algorithm (Nocedal and Wright [22]) to solve this shortest path of the *n*-point problem.

#### Algorithm 2.

Denote the weighted graph by G = (V, E), where  $V = \{0, ..., q^s - 1\}$  is the index set of vertices, E is the set of the edges; Set  $\mathbf{y} = (y_0, ..., y_{q^s-1})' = (0, ..., 0)'$ ; Find a vertex  $y_{i_0}$  satisfying  $i_0 = \{i | \min_i(\mathbf{B}_s)_{ii}\}$ ; Set  $y_{i_0} = 1, W = \{i_0\}, dist[i_0] = +\infty$ ; for each  $w \in V \setminus W$  do  $dist[w] = (\mathbf{B}_s)_{ww} + 2(\mathbf{B}_s)_{w,i_0}$ ; end for for k = 1, ..., n - 1 do find  $i_0 = \min_w dist[w]$ ; set  $y_{i_0} = 1, W = W \cup \{i_0\}$ ; for each  $w \in V \setminus W$  do  $dist[w] + = 2(\mathbf{B}_s)_{w,i_0}$ ; end for end for

From Algorithm 2 we can obtain a  $q^s \times 1$  vector  $\mathbf{y} = (y_0, \dots, y_{q^s-1})'$ , where  $y_i = 1$ , for  $i \in W$ , and 0 otherwise. This vector can be used as a starting vector  $\mathbf{y}_0$  for Algorithm 1. Then we can obtain the approximate optimization solution. Note that this does not require a large computation.

#### 3.3. Selection of $\gamma$

Theorem 1 tells us that in order to obtain a GMA design through (9), we should have a sufficiently small  $\gamma > 0$ . Now let us discuss the selection of  $\gamma$ . First, we have:

**Lemma 3.** Suppose  $a_i$ ,  $b_i$  and m are all non-negative integers,  $a_i < m$ ,  $b_i < m$ , for i = 1, ..., k, and  $a_1 < b_1$ . If  $(a_1, ..., a_k)$  and  $(b_1, ..., b_k)$  are treated as the m-number system  $a_1 \cdots a_k$  and  $b_1 \cdots b_k$  respectively, then

$$a_1 \cdots a_k < b_1 \cdots b_k,$$

i.e.

 $a_1m^{k-1} + \cdots + a_k < b_1m^{k-1} + \cdots + b_k.$ 

Furthermore, from Theorem 4.1 of Cheng and Ye [6] we know that:

**Lemma 4.** For any  $n \times s$  factorial design P with no replicates, where each factor has q levels, we have

$$\sum_{j=1}^s A_j(P) = \frac{q^s}{n} - 1.$$

From these two lemmas, we obtain the following theorem.

**Theorem 2.** If the  $\gamma$  in (5) satisfies that

$$\frac{1}{\gamma}$$
 is a positive integer, and  $\frac{1}{\gamma} > nq^s - n^2$ ,

then minimizing  $D^2(P; \gamma)$  is equivalent to finding a GMA design.

**Proof.** From (2), we know that  $n^2(q-1)A_j^g(P)$  is a non-negative integer, so is  $n^2A_j(P)$ , since (3) holds. Thus from (5) and Lemmas 3 and 4, the conclusion can be proved easily.  $\Box$ 

| Table 1   |
|---|
| Some newly generated (near-) GMA designs for $q = 2$ and $n = 2^k < 64$ |

| n  | S  | $(A_3,\ldots,A_s)$ & Set of selected points   |
|----|----|---|
| 8  | 5  | (2, 1, 0)   |
|    |    | {0, 7, 9, 14, 18, 21, 27, 28}   |
| 8  | 6  | (4, 3, 0, 0)  |
|    |    | {0, 15, 19, 28, 37, 42, 54, 57}   |
| 16 | 5  | $(0, 0, 1)^a$   |
|    |    | {0, 3, 5, 6, 9, 10, 12, 15, 17, 18, 20, 23, 24, 27, 29, 30}   |
| 16 | 6  | (0, 3, 0, 0) ([8,23])   |
|    |    | {0, 7, 9, 14, 19, 20, 26, 29, 34, 37, 43, 44, 49, 54, 56, 63}   |
| 16 | 7  | (0, 7, 0, 0, 0) ([8,23])  |
|    |    | {0, 15, 19, 28, 37, 42, 54, 57, 70, 73, 85, 90, 99, 108, 112, 127}  |
| 32 | 6  | $(0, 0, 0, 1)^a$  |
|    |    | $S(16, 5) \cup \{33, 34, 36, 39, 40, 43, 45, 46, 48, 51, 53, 54, 57, 58, 60, 63\}$  |
| 32 | 7  | (0, 1, 2, 0, 0) ([26])  |
|    |    | $\{0, 7, 9, 14, 18, 21, 27, 28, 35, 36, 42, 45, 49, 54, 56, 63, 65, 70, 72, 79, 83, 84, 90, 93, 98, 101, 107,$  |
|    |    | 108, 112, 119, 121, 126}  |
| 32 | 8  | (0, 3, 4, 0, 0, 0) ([26])   |
|    |    | $S(16, 7) \cup \{129, 142, 146, 157, 164, 171, 183, 184, 199, 200, 212, 219, 226, 237, 241, 254\}$  |
| 32 | 9  | $(0, 6, 8, 0, 0, 1, 0)^b$   |
|    |    | {0, 31, 35, 60, 69, 90, 102, 121, 137, 150, 170, 181, 204, 211, 239, 240, 270, 273, 301, 306, 331, 340,   |
|    |    | 360, 375, 391, 408, 420, 443, 450, 477, 481, 510}   |
| 64 | 7  | $(0, 0, 0, 0, 1)^a$   |
|    |    | $S(32, 6) \cup \{65, 66, 68, 71, 72, 75, 77, 78, 80, 83, 85, 86, 89, 90, 92, 95, 96, 99, 101, 102, 105, 106, 108, 90, 90, 90, 90, 90, 90, 90, 90, 90, 90$   |
|    |    | 111, 113, 114, 116, 119, 120, 123, 125, 126}  |
| 64 | 8  | (0, 0, 2, 1, 0, 0) ([26])   |
|    |    | <i>S</i> (32, 7)∪{130, 133, 139, 140, 144, 151, 153, 158, 161, 166, 168, 175, 179, 180, 186, 189, 195, 196,   |
|    |    | 202, 205, 209, 214, 216, 223, 224, 231, 233, 238, 242, 245, 251, 252}   |
| 64 | 9  | (0, 1, 4, 2, 0, 0, 0) ([26])  |
|    |    | $S(32, 8) \cup \{258, 269, 273, 286, 295, 296, 308, 315, 324, 331, 343, 344, 353, 366, 370, 381, 387, 396, 370, 381, 387, 396, 381, 387, 396, 381, 387, 396, 381, 387, 396, 381, 387, 396, 381, 381, 381, 381, 381, 381, 381, 381$  |
|    |    | 405, 410, 422, 425, 432, 447, 448, 463, 470, 473, 485, 490, 499, 508}   |
| 64 | 10 | $(0, 2, 8, 4, 0, 1, 0, 0)^b$  |
|    |    | $S(32, 9) \cup \{518, 537, 549, 570, 579, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 604, 608, 639, 639, 639, 655, 656, 684, 691, 714, 725, 745, 758, 776, 791, 701, 701, 701, 701, 701, 701, 701, 70$ |
|    |    | 811, 820, 845, 850, 878, 881, 897, 926, 930, 957, 964, 987, 999, 1016}  |
|    |    |   |

S(n, s) represents the set of selected points for the design with *n* runs and *s* factors.

**Remark 3.** From this theorem, we should select a  $\gamma$  satisfying  $0 < \gamma < 1/(nq^s - n^2)$ , in particular, in this paper we will choose  $\gamma = 1/q^{2s}$ , which does not depend on *n*. For such a  $\gamma$ , the **B**<sub>s</sub> used frequently in the algorithms will remain unchanged for varying *n* and fixed *q* and *s*, thus can greatly save the computing time.

#### 4. Some newly generated designs

Tables 1–4 tabulate some newly generated designs from Algorithm 1, by using Algorithm 2 to find the starting vector  $\mathbf{y}_0$  for it. The values of  $\gamma$ ,  $\mu_0$ ,  $\tau_0$  and  $\tau_{step}$  are taken to be  $1/q^{2s}$ , 0.1,  $10^{-6}$  and  $10^{-10}$ respectively. We call all these designs (near-) GMA designs as the search in Algorithm 1 is not an exhaustive search. Tables 1–3 are for (near-) GMA designs with q = 2 and  $n = 2^k$ , q = 2 and  $n \neq 2^k$ , and q = 4 respectively; all designs have  $n \leq 64$  runs. For designs with larger numbers of runs, i.e. n > 64, we have only tabulated in Table 4 the GWPs of the newly generated designs, and omitted the sets of selected points for these designs for saving space. Interested readers can obtain them from the authors.

In these tables, the designs marked with <sup>*a*</sup> can be directly shown to be GMA designs based on their GWPs. Any design marked with a reference number can be checked to be a GMA design from the corresponding reference. For other designs, we are not sure whether they are GMA designs or not as there exist no conclusions to be used or designs to be compared with. In Table 1, the designs marked with <sup>*b*</sup> have the same aberration with the minimum aberration regular designs, and with the GMA non-regular designs derived from the Nordstrom–Robinson code by Xu [26]. These above discussions show

| - |
|---|
| ъ |
|   |
|   |

| Table 2              |             |               |               |                 |
|----------------------|-------------|---------------|---------------|-----------------|
| Some newly generated | (near-) GMA | designs for q | = 2 and $n =$ | $\neq 2^k < 64$ |

| n  | S  | $(A_3, \ldots, A_s)$ & Set of selected points   |
|----|----|---|
| 40 | 6  | (0.16, 0.44, 0, 0)  |
|    |    | $\{0, 1, 2, 3, 4, 11, 12, 13, 14, 15, 17, 18, 21, 22, 23, 24, 25, 26, 29, 30, 34, 36, 37, 38, 39, 40, 41, 42, 43, 45, 48, 49, 51, 52, 55, 56, 59, 60, 62, 63\}$   |
| 48 | 6  | (0, 0.3333, 0, 0)   |
|    |    | {0, 1, 2, 5, 6, 7, 8, 9, 11, 12, 14, 15, 16, 18, 19, 20, 21, 23, 25, 26, 27, 28, 29, 30, 33, 34, 35, 36, 37, 38, 40, 42, 43, 44, 45, 47, 48, 49, 51, 52, 54, 55, 56, 57, 58, 61, 62, 63}  |
| 48 | 7  | (0, 1.6667, 0, 0, 0)  |
|    |    | {0, 1, 6, 9, 14, 15, 18, 19, 21, 26, 28, 29, 35, 36, 37, 42, 43, 44, 48, 54, 55, 56, 57, 63, 64, 70, 71, 72, 73, 79, 83, 84, 85, 90, 91, 92, 98, 99, 101, 106, 108, 109, 112, 113, 118, 121, 126, 127}  |
| 48 | 8  | (0.3333, 1.6667, 2.2222, 0, 0.1111, 0)  |
|    |    | {0, 6, 14, 17, 25, 31, 35, 37, 45, 50, 58, 60, 67, 69, 75, 84, 90, 92, 96, 102, 104, 119, 121, 127, 135, 137, 143, 144, 150, 152, 164, 170, 172, 179, 181, 187, 194, 202, 204, 211, 213, 221, 225, 233, 239, 240, 246, 254}                   |
| 48 | 9  | (0.6667, 3.8889, 1, 4, 0, 0, 0.1111)  |
|    |    | {0, 3, 25, 38, 60, 63, 71, 77, 94, 97, 114, 120, 139, 146, 159, 160, 173, 180, 198, 204, 213, 234, 243, 249, 268, 277, 282, 293, 298, 307, 331, 337, 338, 365, 366, 372, 391, 404, 414, 417, 427, 440, 448, 456, 473, 486, 503, 511}          |
| 48 | 10 | (1, 7.1111, 1.6667, 8.4444, 0.7778, 1.2222, 0.1111, 0)  |
|    |    | {0, 11, 51, 76, 116, 127, 135, 145, 173, 210, 238, 248, 268, 281, 308, 331, 358, 371, 414, 423, 426, 469, 472, 481, 542, 554, 573, 578, 597, 609, 665, 692, 698, 709, 715, 742, 775, 800, 813, 850, 863, 888, 896, 918, 947, 972, 1001, 1023} |
| 56 | 6  | (0.0816, 0.0612, 0, 0)  |
|    |    | $\{0, 1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 32, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 47, 48, 49, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 62, 63\}$                   |

that our algorithms are effective for finding GMA or near-GMA designs. We should also be aware that those newly generated designs are also (nearly) optimal under the discrepancy measure of uniformity defined in (4).

Now let us show how to obtain the designs from the corresponding *sets of selected points* given in Tables 1–3. The *set of selected points* for any design with *n* runs and *s q*-level factors contains the positions of the *n* design points in the full design matrix with  $q^s$  runs which are arranged lexicographically and marked with  $0, \ldots, q^s - 1$ . To obtain such a design, we only need to change the *n* numbers in this set to *n s*-digit numbers in *q*-number system, and there is *absolutely no need* to enumerate all the  $q^s$  runs of the full design and then select the corresponding ones. This is an attractive advantage of our algorithms, and it is of course very convenient and useful to construct the design from this set. For example, the eight points of the first design in Table 1 are the 0th, 7th, 9th, 14th, 18th, 21th, 27th, and 28th runs of the full design with 2<sup>5</sup> runs. To write out this eight-run design, we only need to change 0, 7, 9, 14, 18, 21, 27, and 28 to eight *binary numbers* as shown below:

| 0  |               | 0 | 0 | 0 | 0 | 0  |
|----|---------------|---|---|---|---|----|
| 7  |               | 0 | 0 | 1 | 1 | 1  |
| 9  |               | 0 | 1 | 0 | 0 | 1  |
| 14 | $\rightarrow$ | 0 | 1 | 1 | 1 | 0  |
| 18 |               | 1 | 0 | 0 | 1 | 0  |
| 21 |               | 1 | 0 | 1 | 0 | 1  |
| 27 |               | 1 | 1 | 0 | 1 | 1  |
| 28 |               | 1 | 1 | 1 | 0 | 0. |

This is also true and very useful for the case of q = 4. As another illustration, to construct the design with q = 4, n = 16 and s = 4 given in Table 3, we only need to change the 16 selected points, say 0, 21, 42, 63, 70, 83, 108, 121, 139, 158, 161, 180, 205, 216, 231, 242 to numbers in quaternary system. In

| Table 3                |  |
|------------------------|--|
| Some newly generated ( | (near-) GMA designs for $q = 4$ and $n \le 64$ |

| n  | S | $(A_3, \ldots, A_s)$ & Set of selected points  |
|----|---|--|
| 16 | 4 | (12, 3) ([4])<br>{0, 21, 42, 63, 70, 83, 108, 121, 139, 158, 161, 180, 205, 216, 231, 242}   |
| 16 | 5 | (30, 15, 18) ([4])<br>{0, 85, 170, 255, 283, 334, 433, 484, 557, 632, 647, 722, 822, 867, 924, 969}  |
| 32 | 4 | (4, 3) ([4])<br>{0, 7, 18, 21, 42, 45, 56, 63, 65, 70, 83, 84, 107, 108, 121, 126, 139, 140, 153, 158, 161, 166, 179, 180, 202, 205, 216, 223, 224, 231, 242, 245}   |
| 32 | 5 | (10, 15, 6) ([4])<br>{0, 30, 75, 85, 170, 180, 225, 255, 261, 283, 334, 336, 431, 433, 484, 506, 557, 563, 614, 632, 647, 665, 716, 722, 808, 822, 867, 893, 898, 924, 969, 983}   |
| 48 | 4 | (1.3333, 3)<br>{0, 7, 9, 18, 21, 28, 35, 42, 45, 54, 56, 63, 65, 70, 79, 83, 84, 90, 101, 107, 108, 112, 121, 126, 130, 139, 140, 151, 153, 158, 161, 166, 168, 179, 180, 189, 196, 202, 205, 209, 216, 223, 224, 231, 238, 242, 245, 251}   |
| 48 | 5 | (3.3333, 15, 2)<br>{0, 30, 39, 75, 85, 114, 141, 170, 180, 216, 225, 255, 261, 283, 316, 334, 336, 361, 406, 431, 433, 451, 484, 506, 522, 557, 563, 607, 614, 632, 647, 665, 672, 716, 722, 757, 785, 808, 822, 836, 867, 893, 898, 924, 955, 969, 983, 1006}   |
| 64 | 4 | (0, 3) <sup><i>a</i></sup><br>{0, 5, 10, 15, 17, 20, 27, 30, 34, 39, 40, 45, 51, 54, 57, 60, 65, 68, 75, 78, 80, 85, 90, 95, 99, 102, 105, 108, 114, 119, 120, 125, 130, 135, 136, 141, 147, 150, 153, 156, 160, 165, 170, 175, 177, 180, 187, 190, 195, 198, 201, 204, 210, 215, 216, 221, 225, 228, 235, 238, 240, 245, 250, 255}  |
| 64 | 5 | (0, 15, 0) ([13])<br>{0, 21, 42, 63, 70, 83, 108, 121, 139, 158, 161, 180, 205, 216, 231, 242, 263, 274, 301, 312, 321, 340, 363, 382, 396, 409, 422, 435, 458, 479, 480, 501, 521, 540, 547, 566, 591, 602, 613, 624, 642, 663, 680, 701, 708, 721, 750, 763, 782, 795, 804, 817, 840, 861, 866, 887, 901, 912, 943, 954, 963, 982, 1001, 1020}   |
| 64 | 6 | (0, 45, 0, 18) ([13])<br>{0, 85, 170, 255, 283, 334, 433, 484, 557, 632, 647, 722, 822, 867, 924, 969, 1054, 1099, 1204, 1249, 1285, 1360, 1455, 1530, 1587, 1638, 1689, 1740, 1832, 1917, 1922, 2007, 2087, 2162, 2189, 2264, 2364, 2409, 2454, 2499, 2570, 2655, 2720, 2805, 2833, 2884, 3003, 3054, 3129, 3180, 3219, 3270, 3362, 3447, 3464, 3549, 3604, 3649, 3774, 3819, 3855, 3930, 4005, 4080} |

this way the transposed design matrix of this 16-run GMA design is:

| 0 | 21 | 42 | 63 | 70 | 83 | 108 | 121 | 139 | 158          | 161 | 180 | 205 | 216 | 231 | 242 |
|---|----|----|----|----|----|-----|-----|-----|--------------|-----|-----|-----|-----|-----|-----|
|   |    |    |    |    |    |     |     |     | $\downarrow$ |     |     |     |     |     |     |
| 0 | 0  | 0  | 0  | 1  | 1  | 1   | 1   | 2   | 2            | 2   | 2   | 3   | 3   | 3   | 3   |
| 0 | 1  | 2  | 3  | 0  | 1  | 2   | 3   | 0   | 1            | 2   | 3   | 0   | 1   | 2   | 3   |
| 0 | 1  | 2  | 3  | 1  | 0  | 3   | 2   | 2   | 3            | 0   | 1   | 3   | 2   | 1   | 0   |
| 0 | 1  | 2  | 3  | 2  | 3  | 0   | 1   | 3   | 2            | 1   | 0   | 1   | 0   | 3   | 2.  |
|   |    |    |    |    |    |     |     |     |              |     |     |     |     |     |     |

We can see that there is no need to enumerate all the  $4^4 = 256$  runs and then select the corresponding 16 ones.

#### 5. Further discussions

This paper has transformed the problem of finding GMA designs to an optimization problem, and provided effective algorithms for solving this problem. The newly generated designs can be easily obtained from the given sets of selected points, and they are both (nearly) optimal under the GMA criterion as well as the discrepancy measure of uniformity in (4). The algorithms can also be used to find optimal designs under any other criterion which can be expressed as a quadratic form of ylike the  $D(P; \gamma)$  in (7). For example, Ma and Fang ([19,20]) have expressed the squares of centered  $L_2$ -discrepancy ( $CL_2(D)$ ) and wrap-around  $L_2$ -discrepancy ( $WL_2(D)$ ) as quadratic forms of the vector y. Hence our algorithms can also be used to find uniform designs under the criteria of  $CL_2(D)$  and  $WL_2(D)$  respectively. But the algorithms will become very slow for large q and s and cannot even obtain a solution in practice. This is the main drawback of our algorithms.

|  | F. Sun et al. | ' Journal d | of Complexity | 25 (2009 | )75-84 |
|--|---------------|-------------|---------------|----------|--------|
|--|---------------|-------------|---------------|----------|--------|

| Table 4                                      |                                      |     |
|--|--------------------------------------|-----|
| The GWPs of more newly generated designs for | or $q = 2$ and $n \neq 2^k$ or $q =$ | = 4 |

| q | n   | S  | $(A_3,\ldots,A_s)$   |
|---|-----|----|--|
| 2 | 72  | 8  | (0.1481, 0.8642, 0.7901, 0.6914, 0.0494, 0.0123)                 |
|   | 80  | 8  | (0.12, 1.24, 0.8, 0, 0.04, 0)                                    |
|   | 80  | 9  | (0.24, 1.8, 0.36, 2.96, 0, 0, 0.04)                              |
|   | 80  | 10 | (0.36, 3.2, 0.6, 6.88, 0.28, 0.44, 0.04, 0)                      |
|   | 96  | 7  | (0, 0.1111, 0.2222, 0, 0)  |
|   | 96  | 8  | (0.0694, 0.3889, 1.0833, 0.1111, 0.0139, 0)                      |
|   | 96  | 9  | (0.2361, 1.3333, 1.7222, 0.5278, 0.1528, 0.3611, 0)              |
|   | 112 | 7  | (0, 0.1429, 0, 0, 0)   |
|   | 112 | 9  | (0.1224, 0.1837, 0.1837, 3.0612, 0, 0, 0.0204)                   |
|   | 112 | 10 | (0.2245, 3.6429, 0.7653, 2.0204, 0.4898, 0.8469, 0.1531, 0)      |
|   | 144 | 10 | (0.1111, 1.5802, 0.1852, 3.6049, 0.0864, 0.5309, 0.0123, 0)      |
|   | 160 | 8  | (0, 0.12, 0.48, 0, 0, 0)   |
|   | 160 | 10 | (0.0925, 1.6675, 0.2025, 3.2475, 0.0975, 0.0825, 0.0075, 0.0025) |
|   | 176 | 9  | (0.0496, 0.2893, 0.0744, 1.4876, 0, 0, 0.0083)                   |
|   | 176 | 10 | (0.0744, 0.9008, 0.1240, 3.2562, 0.0579, 0.3967, 0.0083, 0)      |
|   | 192 | 8  | (0, 0, 0.2222, 0.1111, 0, 0)                                     |
|   | 192 | 9  | (0, 0.1667, 0.5, 1, 0, 0, 0)                                     |
|   | 192 | 10 | (0, 0.2222, 0.8889, 2.2222, 0, 1, 0, 0)                          |
|   | 208 | 9  | (0.0355, 0.2781, 0.0533, 1.0769, 0, 0.0118, 0.0059)              |
|   | 208 | 10 | (0.0533, 0.8462, 0.0888, 2.5680, 0.0414, 0.3195, 0.0059, 0)      |
|   | 224 | 9  | (0, 0.1224, 0.1633, 0, 0, 1, 0)                                  |
|   | 224 | 10 | (0, 0.7296, 0, 2.6276, 0, 0.2092, 0, 0.0051)                     |
|   | 288 | 9  | (0, 0.0741, 0.0988, 0, 0, 0.6049, 0)                             |
| 4 | 80  | 5  | (1.2, 9.88, 0.72)  |
|   | 96  | 4  | (0.4444, 1.2222)   |
|   | 96  | 5  | (1.1111, 7.8889, 0.6667)   |
|   | 112 | 5  | (0.6122, 7.1633, 0.3673)   |
|   | 128 | 4  | $(0, 1)^a$   |
|   | 128 | 5  | (0, 7, 0)  |
|   | 128 | 6  | (0, 21, 0, 10)   |
|   | 176 | 5  | (0.2479, 4.4215, 0.1488)   |
|   | 192 | 4  | $(0, 0.3333)^a$  |
|   | 192 | 5  | (0, 4.3333, 0)   |
|   | 192 | 6  | (0, 14.7778, 0, 5.5556)  |
|   | 240 | 4  | (0.0533, 0.0133)   |
|   | 240 | 5  | (0.4356, 2.4133, 0.4178)   |
|   | 256 | 5  | $(0, 0, 3)^a$  |
|   | 256 | 6  | (0, 3, 12, 0)  |

#### Acknowledgments

The authors thank the Associate Editor and two anonymous referees for their valuable comments. This work was supported by the Program for New Century Excellent Talents in University (NCET-07-0454) of China and the NNSF of China grant 10671099.

#### References

- [1] G.E.P. Box, J.S. Hunter, The  $2^{k-p}$  fractional factorial designs, Technometrics 3 (1961) 311–352, 449–458.
- [2] N.A. Butler, Minimum aberration construction results for nonregular two-level fractional factorial designs, Biometrika 90 (2003) 891–898.
- [3] N.A. Butler, Minimum  $G_2$ -aberration properties of two-level foldover designs, Statist. Probab. Lett. 67 (2004) 121–132.
- [4] N.A. Butler, Generalised minimum aberration construction results for symmetrical orthogonal arrays, Biometrika 92 (2005) 485-491.
- [5] C.S. Cheng, Some projection properties of orthogonal arrays, Ann. Statist. 23 (1995) 1223-1233.
- [6] S.W. Cheng, K.Q. Ye, Geometric isomorphism and minimum aberration for factorial designs with quantitative factors, Ann. Statist. 32 (2004) 2168–2185.
- [7] L.Y. Deng, B. Tang, Generalized resolution and minimum aberration criteria for Plackett–Burman and other nonregular factorial designs, Statist. Sinica 9 (1999) 1071–1082.

- [8] L.Y. Deng, B. Tang, Design selection and classification for Hadamard matrices using generalized minimum aberration criteria, Technometrics 44 (2002) 173–184.
- [9] R. Diestel, Graph Theory, 3rd ed., Springer-Verlag, Heidelberg, New York, 2005.
- [10] K.T. Fang, G.N. Ge, M.Q. Liu, H. Qin, Construction of minimum generalized aberration designs, Metrika 57 (2003) 37–50.
- [11] K.T. Fang, R. Li, A. Sudjianto, Design and Modeling for Computer Experiments, Chapman & Hall, Boca Raton, 2006.
- [12] K.T. Fang, Y. Wang, Number-Theoretic Methods in Statistics, Chapman & Hall, London, 1994.
- [13] K.T. Fang, A. Zhang, R. Li, An effective algorithm for generation of factorial designs with generalized minimum aberration, J. Complexity 23 (2007) 740–751.
- [14] A. Fries, W.G. Hunter, Minimum aberration  $2^{k-p}$  designs, Technometrics 22 (1980) 601–608.
- [15] F.J. Hickernell, A generalized discrepancy and quadrature error bound, Math. Comp. 67 (1998) 299–322.
- [16] F.J. Hickernell, Goodness-of-fit statistics, discrepancies and robust designs, Statist. Probab. Lett. 44 (1999) 73–78.
- [17] F.J. Hickernell, M.Q. Liu, Uniform designs limit aliasing, Biometrika 89 (2002) 893–904.
- [18] Y. Li, L.Y. Deng, B. Tang, Design catalog based on minimum G-aberration, J. Statist. Plann. Inference 124 (2004) 219–230. [19] C.X. Ma, K.T. Fang, Applications of uniformity to orthogonal fractional factorial designs, Technical Report MATH-193, Hong
- Kong Bartist University, 1998.
- [20] C.X. Ma, K.T. Fang, Some connections between uniformity orthogonality and aberration in regular fractional factorial designs, Technical Report MATH-248, Hong Kong Baptist University, 1999.
- [21] C.X. Ma, K.T. Fang, A note on generalized aberration in factorial designs, Metrika 53 (2001) 85–93.
- [22] J. Nocedal, S.J. Wright, Numerical Optimization, Springer-Verlag, New York, 1999.
- [23] D.X. Sun, W. Li, K.Q. Ye, An algorithm for sequentially constructing nonisomorphic orthogonal designs and its applications, Technical report SUNYSB-AMS-02-13, Department of Applied Mathematics and Statistics, SUNY at Stony Brook, 2002.
- [24] B. Tang, L.Y. Deng, Minimum G<sub>2</sub>-aberration for nonregular fractional factorial designs, Ann. Statist. 27 (1999) 1914–1926.
- [25] H. Xu, Minimum moment aberration for nonregular designs and supersaturated designs, Statist. Sinica 13 (2003) 691–708.
- [26] H. Xu, Some nonregular designs from the Nordstrom-Robinson code and their statistical properties, Biometrika 92 (2005) 385–397.
- [27] H. Xu, C.F.J. Wu, Generalized minimum aberration for asymmetrical fractional factorial designs, Ann. Statist. 29 (2001) 1066–1077.