On Construction of Optimal Two-Level Designs with Multi Block Variables[∗]

ZHAO Yuna *·* **ZHAO Shengli** *·* **LIU Minqian**

DOI: 10.1007/s11424-017-6144-2

Received: 1 July 2016 / Revised: 5 October 2016 -c The Editorial Office of JSSC & Springer-Verlag GmbH Germany 2018

Abstract When running an experiment, inhomogeneity of the experimental units may result in poor estimations of treatment effects. Thus, it is desirable to select a good blocked design before running the experiment. Mostly, a single block variable was used in the literature to treat the inhomogeneity for simplicity. However, in practice, the inhomogeneity often comes from multi block variables. Recently, a new criterion called B²-GMC was proposed for two-level regular designs with multi block variables. This paper proposes a systematic theory on constructing some B^2 -GMC designs for the first time. Experimenters can easily obtain the B^2 -GMC designs according to the construction method. Pros of B²-GMC designs are highlighted in Section 4, and the designs with small run sizes are tabulated in Appendix B for practical use.

Keywords Blocked design, general minimum lower order confounding, multi block variables, Yates order.

1 Introduction

Two-level fractional factorial designs are widely used in many areas of industry, science, and engineering. Under the assumption that the third- and higher-order effects are negligible, such designs can effectively identify important main effects and two-factor interactions in a linear model. In some experimental situations, however, especially when the size of experiment is relatively large, inhomogeneity of experimental units often exists and always causes bad

ZHAO Yuna

LPMC and Institute of Statistics, Nankai University, Tianjin 300071*, China.*

ZHAO Shengli

School of Statistics, Qufu Normal University, Qufu 273165*, China.*

LIU Minqian

LPMC and Institute of Statistics, Nankai University, Tianjin 300071*, China.* Email: mqliu@nankai.edu.cn.

[∗]This research was supported by the National Natural Science Foundation of China under Grant Nos. 11271205, 11371223, 11431006 and 11601244, the Specialized Research Fund for the Doctoral Program of Higher Education of China under Grant No. 20130031110002, the "131" Talents Program of Tianjin, and the Program for Scientific Research Innovation Team in Applied Probability and Statistics of Qufu Normal University under Grant No. 0230518.

This paper was recommended for publication by Editor SHI Jianjun.

influence on estimating treatment effects^[1, 2]. Blocking the experimental units into groups is an efficient way to solve this problem. Thus, selecting good blocked designs becomes an important issue.

As pointed out in [1], there are two kinds of blocking problems, one with a single block variable and the other with two or more block variables. We call the latter multi block variables problem. In the last decades, researchers have investigated the issue and proposed many optimality criteria for selecting a blocked design with a single block variable. The following four are the most popular ones among them. The first one is based on the minimum aberration (MA) criterion[3]. It extends the MA idea to the blocked case, see [4, 5] and the references therein. The second one is based on the clear effects criterion, see [6]. It aims at clearly estimating the maximum number of main effects and two-factor interactions of treatment factors in the blocked case^[7]. The third one is based on the maximum estimation capacity criterion^[8, 9] aiming at estimating as many models involving all the main effects and some two-factor interactions as possible. The fourth one is based on the general minimum lower order confounding (GMC) criterion proposed by Zhang, et al.^[10]. When the experimenters have some prior information on the importance ordering of treatment factors, the GMC designs are preferable.

Zhang, et al.^[10] proposed the GMC criterion for two-level regular designs. Since then, quite a few references studying the optimal GMC designs have appeared, such as [11–15]. Zhang and Mukerjee^[16] applied the GMC criterion to treat single block variable problem and established a blocked GMC (B-GMC) criterion to select s-level regular blocked designs. Zhao, et al.^[17] developed a theory on construction of B-GMC designs. Also for selecting optimal blocked designs with a single block variable, Wei, et al.^[18] extended the GMC idea in a different view from that in [16], and proposed another blocked GMC (B^1 -GMC) criterion. Zhao, et al.^[19] and Zhao, et al.^[20] discussed the construction of the B^1 -GMC designs.

However, the multi block variables case of blocked designs often happens in many practical experiments. As was mentioned in [1], in the agricultural context, when designs are laid out in rectangular schemes, both row and column inhomogeneity effects probably exist in the soil. We would like to give two more examples of multi block variables case. Consider an experiment to compare two gasoline additives by testing them on two cars with two drivers over two days. In this testing case, three variables, cars, drivers and days, have to be considered to partition the experimental units. Here is another one about testing the abrasion resistance of rubber-covered fabric in a Martindale wear tester. There are two types of materials, and two positions on the tester so that two samples can be tested at a time. Each time the tester is used, the setup could be a bit different; that is, there might be a systematic difference from application to application. So in this case, the "application" and position should be considered as two block variables.

For selecting two-level regular blocked designs with multi variables, Zhang, et al.^[21] proposed the blocked GMC (B^2 -GMC) criterion, and listed some 16-, 32-, and 64-run B^2 -GMC designs by computer search. However, when the run size of an experiment is large, computer search will be a time consuming work. In this paper, a systematic theory on constructing B^2 -GMC designs is established for the first time. Compared with the direct computer search, the construction methods in this paper are more efficient and easier to apply. The idea of the construction \mathcal{D} Springer

methods can be further extended to broader parameters setting. Comparison between B^2 - GMC designs and $B¹-GMC$ designs constructed based on single block variable is made under fair conditions. The comparison shows that in practical cases if a problem belongs to multi variables problem, we should not use the B^1 -GMC designs. Otherwise it will result in missing estimations of many important effects. So this work is valuable and necessary.

The rest of the paper is organized as follows. In Section 2, we introduce some related concepts and notation of GMC and B^2 -GMC criteria. Section 3 proposes the construction theory of B^2 -GMC designs along with two illustrative examples. In Section 4, we summarize the work of this paper and highlight the pros of B^2 -GMC designs by a fair comparison. A proof is deferred to Appendix A, and some small run-size B^2 -GMC designs are tabulated in Appendix B.

2 Preliminaries: GMC and B²**-GMC Criteria**

Here, we first introduce some notation which is useful in developing the coming up construction theory. Let

$$
H_q=(\boldsymbol{1},\boldsymbol{2},\boldsymbol{12},\boldsymbol{3},\boldsymbol{13},\boldsymbol{23},\boldsymbol{123},\cdots,\boldsymbol{q},\boldsymbol{1q},\cdots,\boldsymbol{1234\cdots q})_{2^q}
$$

denote the saturated design with Yates order, which is generated by q independent columns with 2^q components of entries 1 or -1. The q independent columns of H_q are in the form of

$$
1_{2^q} = (1, -1, 1, -1, \cdots, 1, -1, 1, -1)',
$$

\n
$$
2_{2^q} = (1, 1, -1, -1, \cdots, 1, 1, -1, -1)',
$$

\n
$$
3_{2^q} = (1, 1, 1, 1, -1, -1, -1, -1, \cdots, 1, 1, 1, 1, -1, -1, -1, -1)',
$$

\n
$$
\vdots
$$

\n
$$
q_{2^q} = (1, 1, \cdots, 1, -1, -1, \cdots, -1)'
$$

The other columns **12**, **13**, **23**, \cdots , **1234** \cdots *q* are generated by these *q* independent columns. Take **12** as an example. It is just the component-wise product of the independent columns $\mathbf{1}_{2^q}$ and **2**2*^q* , i.e.,

$$
({\bf 12})_{2^q}=(1,-1,-1,1,\cdots,1,-1,-1,1)'.
$$

In the following, for a vector a and a matrix $B = (b_1, b_2, \dots, b_l)$, let $aB = (ab_1, ab_2, \dots, ab_l)$ be a matrix obtained by taking the component-wise products of a and b_i for $i = 1, 2, \dots, l$. Furthermore, denote $H_1 = (1)_{2^q}$,

$$
H_r = (H_{r-1}, \mathbf{r}, \mathbf{r} H_{r-1}),
$$
\n(1)

$$
F_{j,r} = (\mathbf{j}, \mathbf{j}H_{r-1})
$$
\n⁽²⁾

for $j = r, r + 1, \dots, q$ and $r = 2, 3, \dots, q$, where H_{r-1} consists of the first $2^{r-1} - 1$ columns of H_q . Especially, when $j = r$, $(H_{r-1}, F_{j,r}) = H_r$. Without confusion, we will omit the subscript 2*^q* hereafter.

2 Springer

Let D_t denote a regular 2^{n-m} unblocked design consisting of n columns taken from H_q with $q = n - m$. The n columns of D_t comprise $n - m$ independent columns and the remaining m columns are determined by m independent defining relations. The design D_t is said to have resolution R if no c-factor effect is confounded with any other effect containing less than $R - c$ factors[22].

Now, we recall some concepts related to GMC given in [10]. Consider a regular 2*n*−*^m* unblocked design D_t , let ${}_{i}^{#C_j^{(k)}}(D_t)$ denote the number of *i*-th order treatment effects which are aliased with k j-th order treatment effects in D_t , where $k = 1, 2, \cdots, \frac{n!}{j!(n-j)!}$. Under the assumption that the third- and higher-order treatment effects are negligible, and the main effects are more important than two-factor interactions, put

$$
\begin{aligned}\n\#C(D_t) &= \left(\,^{\#}C_2(D_t), \,^{\#}C_2(D_t)\right), \text{ where} \\
\#C_2(D_t) &= \left(\,^{\#}C_2^{(0)}(D_t), \,^{\#}C_2^{(1)}(D_t), \cdots, \,^{\#}C_2^{(K_2)}(D_t)\right), \\
\#C_2(D_t) &= \left(\,^{\#}C_2^{(0)}(D_t), \,^{\#}C_2^{(1)}(D_t), \cdots, \,^{\#}C_2^{(K_2)}(D_t)\right),\n\end{aligned}\n\tag{3}
$$

and $K_2 = \frac{n(n-1)}{2}$. Pattern (3) is called the aliased-effect number pattern (AENP). A design sequentially maximizing the components of pattern (3) is called a general minimum lower order confounding (GMC) design^[10].

In many practical experiments, the inhomogeneity may come from different sources. See the examples given in Section 1. Suppose that there are s different inhomogeneity sources, namely s different block variables. For $j = 1, 2, \dots, s$, if the jth block variable partitions the experimental units into 2^{i_j} blocks, then i_j independent factors are needed. So, $\sum_{j=1}^{s} i_j$ factors are needed to indicate this blocking problem. Here, we would like to emphasize that it is not necessary to require the $\sum_{j=1}^{s} i_j$ factors to be independent, which is different from the single block variable case. The B^2 -GMC criterion for selecting optimal two-level regular blocked designs with multi block variables was proposed in [21] with $i_j = 1$. We use the notation 2^{n-m} : 2^{*s*} to denote a two-level regular blocked design $D = (D_t : D_b)$ with $N = 2^{n-m}$ runs, n treatment factors and s block factors, where D_t is the unblocked 2^{n-m} design, and D_b is a $2^q \times s$ blocking scheme matrix with each column representing a block factor. Conventionally, we suppose that the treatment factors and block factors have no interactions and assume that only main block effects and two-factor interactions of the block factors are significant block effects. Then a treatment effect confounded by main block effects or two-factor interactions of the block factors cannot be estimated. Under the assumption that all the interactions involving three or more treatment factors are negligible, we still consider only the main treatment effects and interactions of two treatment factors. For convenience, we present an equivalent form of the B²-GMC criterion. Let $^{#B}_{i}C_{j}^{(k)}(D)$ be the number of *i*-th order treatment effects which are aliased with k j-th order treatment effects but not with the grand mean or any significant block effects. Denote

$$
\begin{aligned} \#^B_1 C_2(D) &= (\#^B_1 C_2^{(0)}(D), \#^B_1 C_2^{(1)}(D), \cdots, \#^B_1 C_2^{(K_2)}(D)), \text{ and } \\ \#^B_2 C_2(D) &= (\#^B_2 C_2^{(0)}(D), \#^B_2 C_2^{(1)}(D), \cdots, \#^B_2 C_2^{(K_2)}(D)). \end{aligned}
$$

Let

$$
{}^{\#B}C(D) = ({}^{\#B}C_2(D), {}^{\#B}C_2(D)).
$$
\n(4)

Pattern (4) is called the blocked aliased-effect number pattern for multi block variables case, simply denoted by B^2 -AENP. A design that sequentially maximizes the components of (4) is called a B^2 -GMC design. The corresponding criterion is called the B^2 -GMC criterion. A resolution I or II 2^{n-m} design D_t will result in that some main effects cannot be estimated. So, only designs $D = (D_t : D_b)$ with D_t having resolution at least III are considered. Once a treatment effect is confounded by a significant block effect, it cannot be estimated. Thus, in the following, we consider only the designs $D = (D_t : D_b)$ that would not cause confounding of main treatment effects with significant block effects.

3 Construction of B²**-GMC Designs**

In this section, the construction theory of B²-GMC 2^{n-m} : 2^s designs with $\frac{5N}{16} + 1 \le n \le \frac{N}{2}$ are provided. For given n, suppose $2^{l} \leq \frac{N}{2} - n \leq 2^{l+1} - 1$ for some l. Let $2^{k} \leq s \leq 2^{k+1} - 1$, we first propose the construction of B²-GMC 2^{n-m} : 2^{*s*} designs with $k \leq l$ or $k \geq l+2$ in Theorem 3.1. The construction for the case of $k = l + 1$ is given in Algorithm 3.3.

For easy presentation of the following results, we first introduce some more pieces of notation here. Let I_{D_b} denote the matrix in which the columns are two-factor interactions of D_b . Hereafter, $a \in A$ means that a is a column of matrix A, $A \cup B$ denotes the matrix which consists of the columns of both matrices A and B but without duplication, $A \cap B$ denotes the matrix which consists of the common columns of A and B, $A \cap B = \emptyset$ means that matrices A and B have no common column, $A \setminus B$ denotes the matrix which consists of the columns of matrix A but not those of matrix B. The statement "A is an s-projection of B", denoted as $A \subset B$, implies that A is a matrix where the s columns come from matrix B .

Theorem 3.1 (B²-GMC 2^{n-m} : 2^{*s*} designs with $2^k \le s \le 2^{k+1} - 1$, $k \le l$ or $k \ge l + 2$) *Suppose* $D = (D_t : D_b)$ *is a* $2^{n-m} : 2^s$ *design with* $\frac{5N}{16} + 1 \le n \le \frac{N}{2}$, $2^k \le s \le 2^{k+1} - 1$ *for some* k, and $2^{l} \leq \frac{N}{2} - n \leq 2^{l+1} - 1$ *for some* l. Then $D = (D_t : D_b)$ *is a* B^2 -GMC design if D_t *consists of the last* ⁿ *columns of* ^F*q,q and*

- (a) when $1 \leq k \leq l$, D_b *is any s-projection of* $H_k \cup F_{a,(k+1)}$;
- (b) when $l + 2 \leq k \leq q 2$, D_b *is any s-projection of* H_{k+1} .

Theorem 3.1 provides the theoretical construction of B²-GMC designs for $k \leq l$ and $k \geq$ $l + 2$, which covers a wide range of s. Given parameters n, m and s, we can easily obtain the corresponding B^2 -GMC designs through some simple calculations. Now, we give an example as illustration of Theorem 3.1.

Example 3.2 Consider the construction of B²-GMC 2^{60-53} : 2^s designs for $s = 4$ and 17. Here $n = 60$, $m = 53$ and $N = 128$. Since $2^2 \le \frac{N}{2} - n \le 2^3 - 1$, thus $l = 2$. First, take the last 60 columns of $F_{7,7}$ to be D_t . Then $F_{7,7} \backslash D_t = F_{7,3}$.

When $s = 4$, there exists $k = 2$ such that $2^2 \le s \le 2^3 - 1$. Since $k = l$, this case belongs to Theorem 3.1 (a). We take any 4-projection of $H_2 \cup F_{7,3}$ as D_b . According to Lemma A.2

2 Springer

in Appendix A, $D_b \cup I_{D_b} = H_2 \cup F_{7,3}$. Note that each of the four alias sets represented by the columns of $F_{7,3}$ does not contain any two-factor interactions of D_t . Each of the three alias sets represented by the columns of H_2 , **1**, **2**, and **12**, contains 30 two-factor interactions of treatment factors. Except for the above 7 alias sets and the 60 main treatment effects alias sets, there are still 60 alias sets, each of which contains 28 two-factor interactions of treatment factors, which are orthogonal to significant block effects. Therefore, for B^2 -GMC 2^{60-53} : 2^4 design, we have ${}^{#}B_{C_2^{(0)}}(D) = 60, \, {}^{#}B_{C_2^{(27)}}(D) = 1680$ and thus the other components of ${}^{#}B_{C}(D)$ are all zeros.

When $s = 17$, there exists $k = 4$ such that $2^4 \leq s \leq 2^5 - 1$. Since $k = l + 2$, this case belongs to Theorem 3.1 (b). We take any 17-projection of H_5 as D_b . According to Lemma A.2, $D_b \cup I_{D_b} = H_5$. Except for the 31 alias sets represented by the columns of H_5 , the 60 main treatment effects alias sets, and the 4 alias sets represented by the columns of $F_{7,3}$, there are still 32 alias sets, each of which contains 28 two-factor interactions of treatment factors, which are orthogonal to significant block effects. For this B²-GMC 2^{60-53} : 2^{17} design, we have ${}^{#}B_{1}^{(0)}(D) = 60, \, {}^{#}B_{2}^{(27)}(D) = 896$ and the other components of ${}^{#}C(D)$ are all zeros.

Theorem 3.1 provides a theoretical method for constructing B^2 -GMC 2^{n-m} : 2^{*s*} designs with $2^k \leq s \leq 2^{k+1}-1$ for $k \leq l$ or $k \geq l+2$. However, Theorem 3.1 does not work on $k = l+1$. The following algorithm, as a complement of Theorem 3.1, can help us to construct the B^2 -GMC 2^{n-m} : 2^s designs with $2^{l+1} \leq s \leq 2^{l+2} - 1$. Throughout the algorithm, D_t is supposed to consist of the last *n* columns of $F_{q,q}$.

Algorithm 3.3 (B^2 **-GMC** 2^{n-m} : 2^{*s*} designs with $2^{l+1} \le s \le 2^{l+2} - 1$)

Step 1 Search all the blocking scheme matrix candidates B_i , $i = 1, 2, \dots, g$, where $B_i \cap D_t =$ \emptyset , and g denotes the possible number of blocking scheme matrix candidates.

Step 2 Calculate I_{B_i} , $i = 1, 2, \dots, g$.

Step 3 Check if $I_{B_i} \cap D_t = \emptyset$ for $i = 1, 2, \dots, g$. If yes, rank the columns of $B_i \cup I_{B_i}$ as they are in H_{q-1} , and denote $B_i \cup I_{B_i} = (b_1^{(i)}, b_2^{(i)}, \cdots, b_{j_i}^{(i)})$; otherwise, exclude it. Denote the remaining $B_i \cup I_{B_i}$ as $\Theta = \{B_1 \cup I_{B_1}, B_2 \cup I_{B_2}, \cdots, B_h \cup I_{B_h}\}\$, where $h \leq g$.

Step 4 Compare $B_i \cup I_{B_i}$ and $B_j \cup I_{B_j}$ of Θ for $i = 1, 2, \dots, h-1$ and $j = i+1, i+2, \dots, h$. Let $b_j^{(i)}$ and $b_j^{(j)}$ be the first different elements of $B_i \cup I_{B_i}$ and $B_j \cup I_{B_j}$, respectively. If $b_j^{(i)}$ is ranked ahead of $b_f^{(j)}$ in H_{q-1} , then remove $B_j \cup I_{B_j}$ from Θ . Repeat the process until there are only identical $B_i \cup I_{B_i}$ left in Θ and then take any B_i as D_b . Then, $D = (D_t : D_b)$ is the B²-GMC design.

Remark 3.4 In Step 1, we find all the possible blocking scheme matrices. With Step 3, we exclude the blocking scheme matrices which contradict $D_t \cap (D_b \cup I_{D_b}) = \emptyset$. The comparing process in Step 4 aims at finding out the blocking scheme matrices which satisfy the two conditions in Lemma A.7. Therefore, $D = (D_t : D_b)$ obtained in Step 4 is the B²-GMC design.

Algorithm 3.3 provides the construction of B²-GMC designs with $k = l + 1$. Since Algorithm 3.3 is based on the theory we have derived, it is easy to apply and can save more time than the direct computer search.

Here we would like to give an example to show the efficiency of Algorithm 3.3 and how it works.

Example 3.5 Consider the construction of B²-GMC 2^{61-54} : 2^4 design. First, D_t consists of the last 61 columns of $F_{7,7}$. In Step 1, search all the possible B_i from $H_7\backslash D_t$. There are $\binom{66}{4}$ = 720720 different choices for B_i . Here we list B_1 , B_2 and B_3 as partial illustrations:

$$
B_1 = \{1, 2, 12, 7\},
$$

\n
$$
B_2 = \{1, 2, 4, 5\},
$$

\n
$$
B_3 = \{1, 2, 12, 3\}.
$$

In Step 2, we obtain

$$
I_{B_1} = \{1, 2, 12, 17, 27, 127\},
$$

\n
$$
I_{B_2} = \{12, 14, 24, 15, 25, 45\},
$$

\n
$$
I_{B_3} = \{1, 2, 12, 13, 23, 123\}.
$$

In Step 3, all the possible I_{B_i} 's are checked whether they satisfy $I_{B_i} \cap D_t = \emptyset$. Because $I_{B_1} \cap D_t \neq \emptyset$, we exclude B_1 . After completely excluding, rank the columns of the remaining $B_i \cup I_{B_i}$'s and then put them in Θ . Here,

$$
B_2 \cup I_{B_2} = \{1, 2, 12, 4, 14, 24, 5, 15, 25, 45\},
$$

$$
B_3 \cup I_{B_3} = \{1, 2, 12, 3, 13, 23, 123\}.
$$

The columns in $B_2 \cup I_{B_2}$ and $B_3 \cup I_{B_3}$ are already ranked.

In Step 4, the $B_i \cup I_{B_i}$'s in Θ are compared. For example, compare $B_2 \cup I_{B_2}$ and $B_3 \cup I_{B_3}$, **4** and **3** are the first columns respectively such that $B_2 \cup I_{B_2}$ is different from $B_3 \cup I_{B_3}$. Since **3** is ranked ahead of **4**, we delete $B_2 \cup I_{B_2}$ from Θ . After a completely comparing and deleting process in Step 4, we obtain that $B_3 \cup I_{B_3}$ is the only identical one left in Θ . Thus, we take $B_3 = D_b$. Throughout the construction of B²-GMC 2⁶¹⁻⁵⁴ : 2⁴ design, we need not to calculate the pattern (4).

According to Lemma A.2, $D_b \cup I_{D_b} = H_3$. Except for the 7 alias sets represented by the columns of H_3 , the 61 main treatment effects alias sets, and the 3 alias sets represented by the 3 columns of $F_{7,7}\backslash D_t$, there are still 56 alias sets, each contains 29 two-factor interactions of treatment factors, which are orthogonal to significant block effects. For this B²-GMC 2^{61-54} : 2^4 design, we have ${}^{#B}_{1}C_{2}^{(0)}(D) = 61, \, {}^{#B}_{2}C_{2}^{(28)}(D) = 1624$ and the other components of ${}^{#B}C(D)$ are all zeros.

Zhang, et al.^[21] gave some 16-, 32-, and 64-run B^2 -GMC designs by the direct computer search. It took a few months to obtain them. We still take the construction of B²-GMC 2^{61-54} : 2⁴ design as an example. If one directly searches D_t from H_7 , D_t has $\binom{127}{61} = 1.090363E+37$ different choices. Corresponding to each D_t , there are $\binom{66}{4} = 720720$ different choices for D_b . For the $\binom{127}{61}$ $\binom{66}{4}$ possible pairs of D_t and D_b , calculating pattern (4) is undoubtedly a very hard task. Obviously, the construction methods we proposed are much more efficient than the direct computer search.

4 Concluding Remarks

How to arrange blocked designs with multi block variables is very important in practical experiments. Three practical examples of multi block variables have been included in Section 1. Zhang, et al.^[21] proposed the B^2 -GMC criterion for selecting optimal two-level regular designs with multi block variables. Although there are quite a few studies on blocked designs with a single block variable, the study of multi block variables blocking problem is valuable and necessary. We now take the comparison between the B^2 -GMC and B^1 -GMC designs to highlight this point.

For a fair comparison, we may as well suppose that, for the blocked designs with a single variable, only the main block effects and two-factor interactions of the block factors are significant block effects. When $\frac{5N}{16} + 1 \le n \le \frac{N}{2}$, the unblocked design of a B¹-GMC design consists of the last n columns of $F_{q,q}$. So does that of a B²-GMC design according to Theorem 3.1 and Algorithm 3.3. Therefore, for a given n, the B^2 -GMC and B^1 -GMC designs have the same unblocked design D_t . When $s = 2$, the B²-GMC and B¹-GMC designs are the same. When $s \geq 3$, the B²-GMC designs outperform the B¹-GMC designs in the way of estimating more two-factor interactions of D_t . Let D_b^1 and D_b^2 denote the blocking scheme matrices of the B¹-CMC and D_t^2 CMC and D_t^2 GMC and B²-GMC designs, respectively. According to the two conditions in Lemma A.7, we have $(D_b^2 \cup I_{D_b^2}) \cap H_{q-1} \subsetneq (D_b^1 \cup I_{D_b^1}) \cap H_{q-1}$. This means that the B²-GMC designs perform better than the B¹-GMC designs by noting that each column of H_{q-1} represents a two-factor interaction alias set of D_t . Now let us see an illustrative example.

Example 4.1 Consider 2^{11-6} : 2^4 designs. The unblocked designs for both the B²-GMC and B^1 -GMC designs consist of the last 11 columns of $F_{5,5}$. According to [19], the blocking scheme matrix D_b^1 of the B¹-GMC design are generated by the four independent columns **1**, **2**, **3**, **3** and **⁴**. For this design, there are 10 potentially significant block effects, **¹**, **²**, **³**, **⁴**, **¹²**, **¹³**, **¹⁴**, **²³**, **²⁴**, and **³⁴**, each is confounded with a two-factor interaction alias set of the unblocked design. According to Theorem 3.1 (a), the blocking scheme matrix D_b^2 of the B²-GMC design are generated by the four columns **¹**, **²**, **¹²**, and **⁵**. For this design, there are only 3 potentially significant block effects, **¹**, **²**, **¹²**, each is confounded with a two-factor interaction alias set of the unblocked design. As the two-factor interactions of treatment factors cannot be estimated once they are confounded with significant block effects, the B^2 -GMC design is better than the B^1 -GMC design.

Under the fair definition of significant block effects, the same conclusion can be reached when the B²-GMC deigns are compared with optimal blocked designs selected under the other criteria for the single block variable case. Zhang, et al.^[21] summarized that the number of clear two-factor interactions of a B^2 -GMC deign is larger than or equal to that of the B^1 -GMC deign and the optimal blocked design based on MA. For detailed discussion, please refer to [21]. In practical experiments, experimenters should first make clear which kind of blocking problem an experiment belongs to. If it belongs to the case of multi block variables, the B^2 -GMC design is a preferred choice, since more two-factor interactions are allowed to be estimated. On the other hand, if an experiment belongs to the case of single block variable, the experimenters cannot

<u>©</u> Springer

choose the B^2 -GMC design, since the single block variable case requires the s blocking factors to be completely independent which the B^2 -GMC designs cannot meet.

Zhang, et al.^[21] listed some 16-, 32-, and 64-run B^2 -GMC designs with the number of block variables at most 5 through computer search. In this paper we systematically establishes the construction methods of B²-GMC designs with $\frac{5N}{16} + 1 \le n \le \frac{N}{2}$ for all the possible number of block variables. In Appendix B, we list the B^2 -GMC designs with small run size for experimenter to use easily. In the future work, we will concentrate on providing the B^2 -GMC designs which cover a broader range of n.

References

- [1] Bisgaard S, A note on the definition for blocked 2*ⁿ*−*^p* designs, *Technometrics,* 1994, **36**: 308–311.
- [2] Wu C F J and Hamada M S, *Experiments: Planning, Analysis and Optimization*, 2nd Edition, Wiley, New Jersey, 2009.
- [3] Fries A and Hunter W G, Minimum aberration 2*^k*−*^p* designs, *Technometrics*, 1980, **22**: 601–608.
- [4] Xu H Q and Mee R W, Minimum aberration blocking schemes for 128-run designs, *J. Statist. Plann. Inference*, 2010, **140**: 3213–3229.
- [5] Zhao S L, Li P F, and Karunamuni R, Blocked two-level regular factorial designs with weak minimum aberration, *Biometrika*, 2013, **100**: 249–253.
- [6] Wu C F J and Chen Y, A graph-aided method for planning two-level experiments when certain interactions are important, *Technometrics*, 1992, **34**: 162–175.
- [7] Zhao S L, Li P F, and Liu M Q, On blocked resolution IV designs containing clear two-factor interactions, *J. Complexity*, 2013, **29**: 389–395.
- [8] Sun D X, Estimation capacity and related topics in experimental designs, Ph.D. dissertation, University of Waterloo, Waterloo, 1993.
- [9] Cheng C S and Mukerjee R, Blocked regular fractional factorial designs with maximum estimation capacity, *Ann. Statist.*, 2001, **29**: 530–548.
- [10] Zhang R C, Li P, Zhao S L, et al., A general minimum lower-order confounding criterion for two-level regular designs, *Statist. Sinica*, 2008, **18**: 1689–1705.
- [11] Zhang R C and Mukerjee R, Characterization of general minimum lower order confounding via complementary sets, *Statist. Sinica*, 2009, **19**: 363–375.
- [12] Chen J and Liu M Q, Some theory for constructing general minimum lower order confounding designs, *Statist. Sinica*, 2011, **21**: 1541–1555.
- [13] Li P F, Zhao S L, and Zhang R C, A theory on constructing 2*ⁿ*−*^m* designs with general minimum lower-order confounding, *Statist. Sinica*, 2011, **21**: 1571–1589.
- [14] Zhang R C and Cheng Y, General minimum lower order confounding designs: An overview and a construction theory, *J. Statist. Plann. Inference*, 2010, **140**: 1719–1730.
- [15] Cheng Y and Zhang R C, On construction of general minimum lower order confounding 2*ⁿ*−*^m* designs with $\frac{N}{4} + 1 \le n \le \frac{9N}{32}$, *J. Statist. Plann. Inference*, 2010, **140**: 2384–2394.
- [16] Zhang R C and Mukerjee R, General minimum lower order confounding in block designs using complementary sets, *Statist. Sinica*, 2009, **19**: 1787–1802.
- [17] Zhao S L, Lin D K J, and Li P F, A note on the construction of blocked two-level designs with general minimum lower order confounding, *J. Statist. Plann. Inference*, 2016, **172**: 16–22.
- [18] Wei J L , Li P, and Zhang R C, Blocked two-level regular designs with general minimum lower order confounding, *J. Statist. Theory Pract.*, 2014, **8**: 46–65.
- [19] Zhao S L, Li P F, Zhang R C, et al., Construction of blocked two-level regular designs with general minimum lower order confounding, *J. Statist. Plann. Inference*, 2013, **143**: 1082–1090.
- [20] Zhao Y N, Zhao S L, and Liu M Q, A theory on constructing blocked two-level designs with general minimum lower order confounding, *Front. Math. China*, 2016, **11**: 207–235.
- [21] Zhang R C, Li P, and Wei J L, Optimal two-level regular designs with multi block variables, *J. Statist. Theory Pract.*, 2011, **5**: 161–178.
- [22] Box G E P and Hunter J S, The 2*^k*−*^p* fractional factorial designs I and II, *Technometrics*, 1961, **3**: 311–351; 449–458.
- [23] Zhou Q, Balakrishnan N, and Zhang R C, The factor aliased effect number pattern and its application in experimental planning, *Canad. J. Statist.*, 2013, **41**: 540–555.

Appendix A Proof of Theorem 3.1

Several lemmas are developed to gradually introduce necessary conditions for building the construction theory.

Lemma A.1 *Suppose* $A = (a_1, a_2, \dots, a_{r_1}), B = (b_1, b_2, \dots, b_{r_2})$ and $r_1 > r_2$, where $a_i, b_j \in H_q$ are mutually different. Let $a_iB = (a_i b_1, a_i b_2, \cdots, a_i b_{r_2})$ for $i = 1, 2, \cdots, r_1$, then $\bigcap_{i=1}^{r_1} a_i B = \emptyset.$

Proof Suppose $\bigcap_{i=1}^{r_1} a_i B \neq \emptyset$. Then there exists $a_i b_{j_i} \in a_i B$, $i = 1, 2, \dots, r_1$, such that $a_1b_{j_1} = a_2b_{j_2} = \cdots = a_r, b_{j_r}$. Note that $r_1 > r_2$, then there is at least one pair, say j_1 and j_2 , such that $b_{j_1} = b_{j_2}$. This implies $a_1 = a_2$ which contradicts the assumption. This completes the proof the proof.

Using $\#\{A\}$ to denote the number of columns in matrix A, we give the following lemma.

Lemma A.2 *Let E be an s-projection of* $H_k \cup F_{j,(k+1)}$ *with* $2^k ≤ s ≤ 2^{k+1} - 1$ *. Then* $E \cup I_E = H_k \cup F_{i,(k+1)},$ where $j = k+1, k+2, \cdots, q$.

Proof Let

$$
E = (a_1, a_2, \cdots, a_s) \subset H_k \cup F_{j,(k+1)}, \text{ and}
$$

$$
\overline{E} = (H_k \cup F_{j,(k+1)}) \setminus E = (a_{s+1}, a_{s+2}, \cdots, a_{2^{k+1}-1}).
$$

Since $\#\{\overline{E}\} = 2^{k+1} - 1 - s < 2^k \leq \#\{E\}$, according to Lemma A.1, we have $\bigcap_{i=1}^s a_i \overline{E} =$ 0. For any $a_i, i = 1, 2, \dots, s$, we have $(a_i, a_i(E \setminus a_i), a_i \overline{E}) = H_k \cup F_{i,(k+1)}$, which implies $(a_i, a_i(E \backslash a_i)) = (H_k \cup F_{i,(k+1)}) \backslash a_i \overline{E}, i = 1, 2, \cdots, s$. Therefore,

$$
E \cup I_E = \bigcup_{i=1}^{s} (a_i, a_i(E \setminus a_i))
$$

=
$$
\bigcup_{i=1}^{s} ((H_k \cup F_{j,(k+1)}) \setminus a_i \overline{E})
$$

=
$$
(H_k \cup F_{j,(k+1)}) \setminus (\bigcap_{i=1}^{s} a_i \overline{E}).
$$
 (5)

The proof is completed following (5) and $\bigcap_{i=1}^s a_i \overline{E} = \emptyset$.

For easy reference, in the following lemma we introduce some results of [12].

Lemma A.3 *Let* E *be an s-projection of* $F_{q,q}$ *,* $s = 2^{k-1} + \delta \geq 3$ *and* $0 < \delta \leq 2^{k-1}$ *, then* $#{I_E} \geq 2^k - 1$, where the equality can be achieved when the number of independent columns *of* E *is* $k + 1$ *.*

When we select columns from H_q as the blocking scheme matrix D_b , there are two possibilities: (a) $D_b \cap F_{q,q} = \emptyset$ and (b) $D_b \cap F_{q,q} \neq \emptyset$. Lemmas A.4 and A.5 respectively reveal that no matter how to select D_b from H_q as blocking scheme matrix, there are at least a certain number of columns in H_{q-1} confounded by significant block effects.

Lemma A.4 *Let* $2^k \leq s \leq 2^{k+1} - 1$ *, when we select s columns from* H_q *as the blocking scheme matrix* D_b *, if* $D_b \cap F_{q,q} = \emptyset$ *, then*

$$
\inf_{D_b} \{ \# \{ (D_b \cup I_{D_b}) \cap H_{q-1} \} \} = 2^{k+1} - 1,
$$

where inf{·} *represents the infimum.*

Proof If $D_b \cap F_{q,q} = \emptyset$, then $D_b \subset H_{q-1}$ and $(D_b \cup I_{D_b}) \subset H_{q-1}$. Denote $S = (I, D_b)$, then *qS* ⊂ $F_{q,q}$, where *I* is the column with all elements unity. By the structures of H_{q-1} and $F_{q,q}$, it is easy to obtain $D_b \cup I_{D_b} = I_{\mathbf{q}S}$. As $2^k + 1 \leq \#\{\mathbf{q}S\} \leq 2^{k+1}$, according to Lemma A.3, we can obtain $#{I_qs} ≥ 2^{k+1} − 1$, i.e., $#{D_b ∪ I_{D_b} ≥ 2^{k+1} − 1}$. When take the first *s* columns of H_{k+1} as D_b , the equality can be achieved. The proof is completed.

With the help of Lemmas A.3 and A.4, we have the following lemma which considers the possibility (b), i.e., $D_b \cap F_{q,q} \neq \emptyset$.

Lemma A.5 *Let* $2^k \leq s \leq 2^{k+1} - 1$ *, when we select s columns from* H_q *as the blocking scheme matrix* D_b *, if* $D_b \cap F_{q,q} \neq \emptyset$ *, then*

$$
\inf_{D_b} \{ \# \{ (D_b \cup I_{D_b}) \cap H_{q-1} \} \} = 2^k - 1.
$$

Proof Note that

$$
\# \{ (D_b \cup I_{D_b}) \cap H_{q-1} \} = \# \{ (D_b \cap H_{q-1}) \cup I_{D_b \cap H_{q-1}} \cup I_{D_b \cap F_{q,q}} \}.
$$
 (6)

(i) Suppose $\#{D_b \cap F_{q,q}} \leq 2^{k-1}$, then $\#{D_b \cap H_{q-1}} \geq 2^{k-1}$. According to Lemma A.4, we have $\#\{(D_b \cap H_{q-1}) \cup I_{(D_b \cap H_{q-1})}\}\geq 2^k - 1$. Obviously, by (6),

$$
#\{(D_b \cup I_{D_b}) \cap H_{q-1}\} \geq #\{(D_b \cap H_{q-1}) \cup I_{(D_b \cap H_{q-1})}\} \geq 2^k - 1.
$$

(ii) Suppose $\#\{D_b \cap F_{q,q}\} \geq 2^{k-1} + 1$. By Lemma A.3, we obtain $\#\{I_{D_b \cap F_{q,q}}\} \geq 2^k - 1$. Then, by (6) ,

$$
#\{(D_b \cup I_{D_b}) \cap H_{q-1}\} \geq #\{I_{D_b \cap F_{q,q}}\} \geq 2^k - 1.
$$

In Lemma A.2, taking $j = q$, for any blocking scheme matrix $D_b^* \subset H_k \cup F_{q,(k+1)}$ with $\#\{D_b^*\} =$ s, we have $D_b^* \cup I_{D_b^*} = H_k \cup F_{q,(k+1)}$, and hence

$$
\# \{ (D_b^* \cup I_{D_b^*}) \cap H_{q-1} \} = 2^k - 1.
$$
\n⁽⁷⁾\n
$$
\textcircled{2} \text{ Springer}
$$

Г

This completes the proof following (i), (ii) and (7).

Lemma A.6 below is a result belonging to [23]. It plays an important role in the proof of Lemma A.7 below and Algorithm 3.3 in Section 3. Let $B_2(D_t, \gamma)$ denote the number of two-factor interactions of design D_t appearing in the alias set that contains γ .

Lemma A.6 *Suppose that* D_t *consists of the last* n *columns of* $F_{q,q}$ *,* γ_1 *and* γ_2 *are columns in* H_{q-1} *. If* γ_1 *is ranked ahead of* γ_2 *in* H_{q-1} *in Yates order, then*

$$
B_2(D_t, \gamma_1) \ge B_2(D_t, \gamma_2).
$$

As mentioned in Section 2, we consider only designs $D = (D_t : D_b)$ with $D_t \cap (D_b \cup I_{D_b}) = \emptyset$, which means ${}^{#B}C_2^{(k)}(D) = {}^{#C_2^{(k)}}(D_t)$ for any k. Thus, ${}^{#B}_{1}C_2(D)$ depends only on D_t . Therefore, we should consider the part ^D*^t* of ^D as an unblocked design and optimally choose it first to maximize ${}_{1}^{\#}C_2(D_t)$. Then, when $\frac{5N}{16} + 1 \le n \le \frac{N}{2}$, D_t must have resolution at least IV, and we can suppose $D_t \subset F_{q,q}$ ^[13].

Li, et al.^[13] investigated the construction of GMC 2^{n-m} designs with $n \ge \frac{5N}{16} + 1$. They showed that when $\frac{5N}{16} + 1 \le n \le \frac{N}{2}$, if D_t consists of the last n columns of $F_{q,q}$ then D_t is a GMC design, i.e., it sequentially maximizes pattern (3). Lemma A.7 below gives two sufficient conditions for a blocked design $D = (D_t : D_b)$ to be a B²-GMC design.

Lemma A.7 *Suppose that* D_t *consists of the last* n *columns of* $F_{q,q}$ *with* $\frac{5N}{16} + 1 \le n \le \frac{N}{2}$, *if* ^D*^b satisfies the following two conditions:*

 (C_1) *For any candidate blocking scheme matrix* D'_b ,

$$
\#\{(D_b \cup I_{D_b}) \cap H_{q-1}\} \leq \#\{(D'_b \cup I_{D'_b}) \cap H_{q-1}\};
$$

 (C_2) $(D_b \cup I_{D_b}) \cap H_{a-1}$ *consists of the first* $\#\{(D_b \cup I_{D_b}) \cap H_{a-1}\}$ *columns of* H_{a-1} *, then* $D = (D_t : D_b)$ *is a B*²*-GMC design.*

Proof By the definitions of GMC and B²-GMC criteria, to construct a B²-GMC 2^{n-m} : 2^{*s*} design, we can carry out the following two steps:

Step 1 Construct a 2^{n-m} GMC design D_t ;

Step 2 Construct a blocking scheme matrix D_b such that $D = (D_t : D_b)$ is a B²-GMC design.

Let D_t consist of the last *n* columns of $F_{q,q}$. Then according to [13], D_t is a 2^{n-m} GMC design. When selecting columns from H_q as the blocking scheme matrix D_b , we should first select the columns such that those in $D_b \cup I_{D_b}$ are neither aliased with the main effects of D_t nor with the two-factor interactions of D_t , then the columns such that those in $D_b \cup I_{D_b}$ are aliased with the two-factor interactions of D_t at the most serious degree.

Note that according to [13] each column $\gamma \in H_q$ corresponds to an alias set of D_t . For any $\gamma \in D_t$, the alias set contains a main effect of D_t . For any $\gamma \in F_{q,q} \backslash D_t$, the alias set contains only interactions involving three or more factors of D_t . For any $\gamma \in H_{q-1}$, there are at least $n - \frac{N}{4}$ two-factor interactions of D_t aliased with it. If D_b satisfies the conditions C_1 and C_2 , then by Lemma A.6, $D = (D_t : D_b)$ is a B²-GMC design. This completes the proof.

Proof of Theorem 3.1 When we select columns from H_q as D_b , there are two possibilities: (i) $D_b \cap F_{q,q} = \emptyset$ and (ii) $D_b \cap F_{q,q} \neq \emptyset$.

(a) If $D_b \cap F_{q,q} = \emptyset$, from Lemma A.4, $\#\{(D_b \cup I_{D_b}) \cap H_{q-1}\} \geq 2^{k+1} - 1$. On the other hand, if $D_b \cap F_{q,q} \neq \emptyset$, from the proof of Lemma A.5, there exists a $D_b \subset H_k \cup F_{q,(k+1)}$ such that $\#\{(D_b \cup I_{D_b}) \cap H_{q-1}\} = 2^k - 1$. As $2^{k+1} - 1 > 2^k - 1$, by C_1 of Lemma A.7, $D_b \cap F_{q,q} = \emptyset$ is not a good selection for D_b . Let D_b be any s-projection of $H_k \cup F_{q,(k+1)}$. By Lemma A.2, $D_b \cup I_{D_b} = H_k \cup F_{q,(k+1)}$. Obviously, we have $(D_b \cup I_{D_b}) \cap H_{q-1} = H_k$ and $\# \{ (D_b \cup I_{D_b}) \cap H_{q-1} \} = 2^k - 1$. According to Lemma A.7, $D = (D_t : D_b)$ is a B²-GMC design for $k \leq l$.

(b) When $k \geq l + 2$, we have $\#\{F_{q,q} \setminus D_t\} = \frac{N}{2} - n \leq 2^{k-1} - 1$. As $2^k \leq s \leq 2^{k+1} - 1$, then D_b has at least $2^{k-1} + 1$ columns from H_{q-1} , i.e., $\#\{D_b \cap H_{q-1}\} \geq 2^{k-1} + 1$. Suppose $D_b \cap F_{q,q} \neq \emptyset$. Note that for any $\gamma \in D_b \cap F_{q,q}$, we have $\gamma(D_b \cap H_{q-1}) \subset I_{D_b} \cap F_{q,q}$, which implies

$$
#{I_{D_b} \cap F_{q,q}} \ge #{\gamma(D_b \cap H_{q-1})} = #{D_b \cap H_{q-1}} \ge 2^{k-1} + 1.
$$

Recall that we consider only designs $D = (D_t : D_b)$ with $D_t \cap (D_b \cup I_{D_b}) = \emptyset$. Thus, we have $D_t \subset F_{q,q} \backslash I_{D_b}$, and hence

$$
n = #\{D_t\} \le #\{F_{q,q} \setminus I_{D_b}\} = #\{F_{q,q}\} - #\{(I_{D_b} \cap F_{q,q})\} \le \frac{N}{2} - (2^{k-1} + 1) \le n - 2.
$$

This contradiction shows that $D_b \cap F_{q,q} = \emptyset$ for $k \geq l+2$. Let D_b be any s-projection of H_{k+1} , then $D_b \cup I_{D_b} = H_{k+1}$ by Lemma A.2. Obviously, by Lemma A.4, D_b satisfies the two conditions in Lemma A.7. Thus, $D = (D_t : D_b)$ is a B²-GMC design.

Appendix B Some Small Run-Size B²**-GMC Designs**

The 16-, 32-, 64-run B²-GMC designs are listed in this section. For given $n = \frac{N}{2}$, the design $D = (D_t : D_b)$ with $D_t = F_{q,q}$ and D_b consisting of the first s columns of H_{q-1} is the B²-GMC design. In the following tables, $H_{\{\cdot\}}$ and $F_{\{\cdot\}}$ are defined as in (1) and (2), respectively.

$\,n$	S	D_{t}	D_h	Source
6 6	$2 - 3$ $4 - 7$	the last 6 columns of $F_{4,4}$	$\{4\}$ s-projection of $H_1 \cup F_{42}$ s-projection of H_3	Theorem 3.1 (a) Theorem 3.1 (a) Algorithm 3.3
	$2 - 3$ $4 - 7$	the last 7 columns of $F_{4,4}$	$\{4\}$ s-projection of H_2 s-projection of H_3	Theorem 3.1 (a) Algorithm 3.3 Theorem 3.1 (b)

Table B1 16-run B²-GMC 2^{n-m} : 2^s designs

$\,n$	$\mathcal{S}_{\mathcal{S}}$	D_t	D_b	Source			
$11 - 12$	1	the last <i>n</i> columns of $F_{5,5}$	${5}$	Theorem 3.1 (a)			
$11 - 12$	$2 - 3$		s-projection of $H_1 \cup F_{5,2}$	Theorem 3.1 (a)			
$11 - 12$	$4 - 7$		s-projection of $H_2 \cup F_{5,3}$	Theorem 3.1 (a)			
$11 - 12$	$8 - 15$		s-projection of H_4	Algorithm 3.3			
$13 - 14$	1	the last <i>n</i> columns of $F_{5,5}$	${5}$	Theorem 3.1 (a)			
$13 - 14$	$2 - 3$		s-projection of $H_1 \cup F_{5,2}$	Theorem 3.1 (a)			
$13 - 14$	$4 - 7$		s-projection of H_3	Algorithm 3.3			
$13 - 14$	$8 - 15$		s-projection of H_4	Theorem 3.1 (b)			
15	$\mathbf{1}$	the last 15 columns of $F_{5,5}$	$\{5\}$	Theorem 3.1 (a)			
15	$2 - 3$		s-projection of H_2	Algorithm 3.3			
15	$4 - 7$		s-projection of H_3	Theorem 3.1 (b)			
15	$8 - 15$		s-projection of H_4	Theorem 3.1 (b)			

Table B2 32-run B²-GMC 2^{n-m} : 2^{*s*} designs

