A Coset Pattern Identity between a $2^{n-p}$ Design and its Complement

Peng Zeng, Hong Wan, and Yu Zhu
Auburn University, Purdue University, and Purdue University
February 8, 2010

Abstract: The coset pattern matrix contains more detailed information about effect aliasing in a factorial design than the commonly used wordlength pattern. More flexible and elaborate design criteria can be proposed using the coset pattern matrix. In this article, we establish an identity that relates the coset pattern matrix of a design to that of its complement. As an application, the identity is used to characterize minimum $M$-aberration designs through their complements.

Key words and phrases: Complementary design, coset pattern matrix, fractional factorial design, minimum $M$-aberration, wordlength pattern.

1 Introduction

The $2^{n-p}$ fractional factorial designs are among the most popularly used experimental plans in practice. For given $n$ and $p$, a $2^{n-p}$ design is determined by $p$ independent defining relations or words. These defining words generate the so-called defining contrasts subgroup, commonly denoted by $G$, which contains the $p$ defining words, other relations induced by the $p$ defining words, and the identity. Factorial effects in a $2^{n-p}$ design are either orthogonal to each other or completely aliased with each other. The aliasing property of a $2^{n-p}$ design is
completely determined by its defining contrasts subgroup $G$. In the literature, the wordlength pattern of $G$ is used to quantify the aliasing property of a design, which is defined to be $W_0 = (A_{01}, A_{02}, \ldots, A_{0n})$, where $A_{0i}$ ($1 \leq i \leq n$) is the number of defining words of length $i$ in $G$. The resolution of a design, denoted by $R$, is defined to be the smallest $i$ such that $A_{0i} > 0$. Designs that sequentially minimize $A_{0i}$’s are said to have minimum aberration (MA). MA designs are considered optimal and recommended for use in practice when no additional information is available regarding the experimental factors and their effects (Wu and Hamada, 2000).

Although $W_0$ is able to capture the aliasing property of a design to a large degree, it does have some limitations. First, $W_0$ is far from giving the whole picture of effect aliasing in a design. For example, the number of clear effects (Wu and Chen, 1992) is an important quantity, but in general it does not have a clear relationship with $W_0$. Second, $W_0$ is not able to take into account possible prior information regarding a design’s structural properties and factorial effects. For example, in a robust parameter design, experimental factors are divided into two groups which are the group of control factors and the group of noise factors. The wordlength pattern $W_0$ fails to distinguish these two groups of factors (Wu and Zhu, 2003). It is not difficult to see the cause for the limitations of $W_0$. The defining contrasts subgroup $G$ contains factorial effects aliased with the grand mean, and $W_0$ reports only the frequencies of effects of various lengths in $G$. When two effects are of the same length in $G$, they are counted as the same in $W_0$. If an effect is not included in $G$, $W_0$ does not provide information about how it is aliased with other effects. As a matter of fact, the complete picture of effect aliasing is given by the so-called alias structure that consists of all the cosets of $G$. Each coset is a collection of the effects that are aliased with each other. Note that $G$ itself is a coset. If two effects belong to two different cosets, then they are orthogonal to each other. For each coset, a frequency vector similar to $W_0$ can be defined, which records the frequencies of effects of various lengths. These frequency vectors can be stacked to form a matrix that is referred to as the coset pattern matrix (Zhu and Zeng 2005). The coset pattern matrix contains more comprehensive information about effect aliasing than $W_0$ and
can be used to develop more sensitive and flexible criteria for constructing and selecting optimal designs. One such example is the minimum $M$-aberration criterion proposed by Zhu and Zeng (2005).

When the number of factors under consideration is relatively large, choosing and comparing designs become difficult. To overcome the difficulty, Chen and Hedayat (1996) and Tang and Wu (1996) proposed a complementary approach, which tries to characterize designs through their complements. In particular, the wordlength pattern $W_0$ of a design can be related to that of its complement through a certain identity. The complementary approach has been further generalized to $q^{n-p}$ designs by Suen et al. (1997), to $2^{n-p}$ designs with multiple groups of factors by Zhu (2003), and to nonregular designs by Xu and Wu (2001). Because of the one-to-one correspondence between a design and its complement, it is expected that the coset pattern matrices of a design and its complement should also be related to each other. The purpose of this article is to establish this type of relationship explicitly. Furthermore, we will show how the relationship can be used to identify minimum $M$-aberration designs through the complementary approach.

The rest of the article is organized as follows. Section 2 introduces notations and some concepts. Section 3 explores the correspondence between the cosets of a design and its complement and further derives the explicit relationship between their coset pattern matrices. Section 4 demonstrates how these relations can be used to select minimum $M$-aberration designs through the complementary approach when the number of factors is large. The proofs of the theorems, propositions and corollaries are included in the Appendix. All designs discussed in this article are regular two-level fractional factorial designs with resolution at least III.

2 Notations and Definitions

We use letters $1, 2, \ldots, n$ to denote $n$ factors in an experiment. Factorial effects are represented by words that juxtapose the involved letters (or factors) from the smallest to the
largest. For example, 1 represents the main effect of factor 1, and 12 represents the interaction between factor 1 and factor 2. Following the convention, we use $I$ to denote the grand mean. Including $I$, there are in total $2^n$ factorial effects, which form an Abelian group denoted by $S$. We introduce an order between these factorial effects. One effect $i_1i_2\cdots i_k$ is said to be smaller than another effect $j_1j_2\cdots j_l$, written as $i_1i_2\cdots i_k \prec j_1j_2\cdots j_l$, if $k < l$ or if $k = l$ and $i_1i_2\cdots i_k$ should be listed ahead of $j_1j_2\cdots j_l$ lexicographically. The defining contrast subgroup $G$ of a $2^{n-p}$ design is a subgroup of $S$, and it can generate $2^{n-p}$ cosets that form a partition of $S$. Each coset contains $2^p$ effects that are aliased with each other. The smallest effect (under $\prec$) in a coset is defined to be the coset leader. If a coset has $i_1\cdots i_k$ as its coset leader, it is represented by $i_1i_2\cdots i_kG$. It is clear that $\prec$ can also be applied to the coset leaders, so the cosets can be rank-ordered, according to their coset leaders, from the coset of the lowest rank (i.e., 0) to the coset of the highest rank (i.e., $2^{n-p} - 1$). The rank of the coset $i_1i_2\cdots i_kG$ is denoted by $r(i_1i_2\cdots i_kG)$. The coset of rank 0 is $IG$, which is the defining contrast subgroup itself. The next $n$ cosets are $1G$, $2G$, $\ldots$, $nG$ with rank 1, 2, $\ldots$, $n$, respectively, followed by cosets with higher ranks. A coset is said to be of order $d$ if its coset leader is an effect of order $d$ (i.e., a $d$-factor interaction). For example, $i_1G$ is a coset of order 1 and $i_1i_2G$ is a coset of order 2. We use $\mathcal{F}$ to denote the collection of all the cosets of a design.

Suppose $i_1\cdots i_lG$ is the coset of rank $r$, that is, $r(i_1\cdots i_lG) = r$, with $0 \leq r \leq 2^{n-p} - 1$. Let $A_{rj}$ be the number of words of length $j$ in $i_1\cdots i_lG$. The vector $W_r = (A_{r1}, A_{r2}, \ldots A_{rn})$ is defined to be the coset pattern of $i_1\cdots i_lG$. Note that when $k = 0$, $W_0$ is exactly the wordlength pattern of $G$. The $2^{n-p} \times n$ matrix $A = (W_0^T, W_1^T, \ldots, W_{2^{n-p}-1}^T)^T = (A_{rj})$ is called the coset pattern matrix. Compared to the wordlength pattern $W_0$, the coset pattern $A$ contains more comprehensive information about the effect aliasing of a design. For example, the number of clear two-factor interactions can be directly calculated from $A$, which is equal to $\sum_{i=1}^{2^{n-p}-1} I(A_{i1} = 0)I(A_{i2} = 1)$, where $I(\cdot)$ is the indicator function. As mentioned in the introduction, using coset pattern matrices, more elaborate criteria can be proposed to discriminate and select useful designs. The minimum $M$-aberration criterion proposed by...
Zhu and Zeng (2005) is such an example and will be discussed in Section 4.

Example 1. Consider a $2^{8-4}$ design $D$ with independent defining relations $\{125, 136, 147, 2348\}$. There are in total 256 factorial effects in $D$, which are partitioned into 16 distinct cosets $IG, 1G, 2G, 3G, 4G, 5G, 6G, 7G, 8G, 18G, 23G, 24G, 26G, 27G, 28G, \text{ and } 37G,$ arranged from rank 0 to rank 15. Because some cosets share the same coset pattern, $D$ has only six distinct coset patterns; see Table 1. The coset pattern of $IG$ is the wordlength pattern of $D$. The coset pattern matrix $A$ is a $16 \times 8$ matrix formed by the 16 coset patterns.

<table>
<thead>
<tr>
<th>coset</th>
<th>coset pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>$IG$</td>
<td>0 0 3 7 4 0 1 0</td>
</tr>
<tr>
<td>$1G$</td>
<td>1 3 0 4 7 1 0 0</td>
</tr>
<tr>
<td>$2G, 3G, 4G, 5G, 6G, 7G$</td>
<td>1 1 4 4 3 3 0 0</td>
</tr>
<tr>
<td>$8G$</td>
<td>1 0 4 7 3 0 0 1</td>
</tr>
<tr>
<td>$18G$</td>
<td>0 1 7 4 0 3 1 0</td>
</tr>
<tr>
<td>$23G, 24G, 26G, 27G, 28G, 37G$</td>
<td>0 3 3 4 4 1 1 0</td>
</tr>
</tbody>
</table>

3 Main Result

The main purpose of this section is to derive the relationship between the coset pattern matrix of a design and that of its complement. To facilitate the derivation, some commonly used concepts and tools for fractional factorial designs are needed. Let $F_2$ be the Galois field $\{0, 1\}$, Let $PG(k - 1, 2) = \{x = (x_1, x_2, \ldots, x_k)^T : x_i \in F_2 \text{ for } 1 \leq i \leq k \text{ and at least one } x_i \neq 0\}$. Note that $T$ means transpose. In the literature on finite geometries, $PG(k - 1, 2)$ is known as the $(k - 1)$-dimensional projective geometry over $F_2$. There are $2^k - 1$ distinct points in $PG(k - 1, 2)$. Letting $m = 2^k - 1$, we denote the points in $PG(k - 1, 2)$ as $p_1, p_2,$
\[ \mathbf{0}^k = (0, 0, \ldots, 0)^T \] and \( \text{EG}(k, 2) = \{ \mathbf{0}^k \} \cup \text{PG}(k - 1, 2) \). \text{EG}(k, 2) is known as the \( k \)-dimensional finite Euclidean geometry on \( F_2 \), and it is also a linear space over \( F_2 \). There exists a natural connection between \( \text{PG}(k - 1, 2) \), the Sylvester-type Hadamard matrix \( H_k(2) \) and \( 2^{n-p} \) fractional factorial designs. Note that the entries of the Sylvester-type Hadamard matrix are commonly represented by 0 and 1 instead of -1 and 1 as used in general Hadamard matrices, and furthermore, we remove the all 0 column of the usual Sylvester-type Hadamard matrix. Therefore, in this article, \( H_k(2) \) is a \( 2^k \times (2^k - 1) \) matrix with 0 and 1 as its entries. Let \( P = (p_1, p_2, \ldots, p_m) \), a \( k \times m \) matrix whose columns are the \( m \) points from \( \text{PG}(k - 1, 2) \). The \( k \) rows of \( P \) are \( m \)-dimensional vectors over \( F_2 \) and can generate \( 2^k \) distinct linear combinations, which are also \( m \)-dimensional row vectors. These linear combinations or vectors form exactly the Sylvester-type Hadamard matrix \( H_k(2) \). Note that the columns of \( H_k(2) \) correspond to the columns of \( P \) and further to the points of \( \text{PG}(k - 1, 2) \). Therefore, the relationship between the columns of \( H_k(2) \) is exactly the same as the relationship between the points of \( \text{PG}(k - 1, 2) \). In the rest of the article, we will use the columns of \( H_k(2) \) and the points of \( \text{PG}(k - 1, 2) \) interchangeably.

A \( 2^{n-p} \) design is a collection of \( n \) columns of \( H_k(2) \) with rank \( k \), or equivalently, a collection of \( n \) points of \( \text{PG}(k - 1, 2) \) with rank \( k \), where \( k = n - p \) and the ranks are the maximum numbers of linearly independent columns of \( H_k(2) \) or points of \( \text{PG}(k - 1, 2) \) over \( F_2 \), respectively. Again in the rest of the article, a \( 2^{n-p} \) design and a collection of \( n \) points of rank \( k \) in \( \text{PG}(k - 1, 2) \) are used interchangeably. For more discussions on the connection between finite projective geometries and factorial designs, readers are referred to Bose (1947), Cameron and van Lint (1991), Chen and Hedayat (1996), and Mukerjee and Wu (2006).

Let \( D \) be a \( 2^{n-p} \) design. As discussed previously, \( D \) is a collection of \( n \) points from \( \text{PG}(k - 1, 2) \). Because \( \text{PG}(k - 1, 2) \) consists of \( 2^k - 1 \) points, there are \( 2^k - 1 - n \) remaining points in \( \text{PG}(k - 1, 2) \) that are not chosen by \( D \). These remaining points form another fractional factorial design, denoted by \( \hat{D} \), which is referred to as the complementary design of \( D \) in the literature. Recall that \( G, \mathcal{F} \) and \( A \) are the defining contrasts subgroup, the collection of cosets, and the coset pattern matrix of \( D \), respectively. Let \( G, \mathcal{F} \), and \( \hat{A} \) be
those of \( \bar{D} \), respectively. Because \( D \) and \( \bar{D} \) are complementary to each other, an intrinsic correspondence exists between their cosets, i.e., between \( \mathcal{F} \) and \( \bar{\mathcal{F}} \). In what follows, we use two examples to demonstrate this correspondence first.

**Example 2.** Let \( k = 4 \). Then \( PG(4 - 1, 2) \) consists of 15 points, which are denoted by

\[
\{a, b, c, d, ab, ac, ad, bc, bd, cd, abc, abd, acd, bcd, abcd\},
\]

where \( a = (1, 0, 0, 0)^T \), \( b = (0, 1, 0, 0)^T \), \( c = (0, 0, 1, 0)^T \) and \( d = (0, 0, 0, 1)^T \) are linearly independent and the remaining points are their linear combinations. The design \( D \) in Example 1 can be obtained by associating the factors with the points of \( PG(4 - 1, 2) \) as follows.

\[
1 = a, 2 = b, 3 = c, 4 = d, 5 = ab, 6 = ac, 7 = ad, 8 = bcd.
\]

The defining contrasts subgroup \( G \) is

\[
IG = \{I, 125, 136, 147, 2348, 2356, 2457, 2678, 3467, 3578, 4568, 12378, 12468, 13458, 15678, 1234567\},
\]

and the coset including the main effect 1 is

\[
1G = \{1, 25, 36, 47, 2378, 2468, 3458, 5678, 12348, 12356, 12457, 12678, 13467, 13578, 14568, 234567\}.
\]

Any effect in \( IG \) corresponds to \( 0^4 \) in \( EG(4, 2) \) because the sum of the involved factors is equal to \( 0^4 \). For example, consider the effect \( 2348 \in IG \). \( 2 + 3 + 4 + 8 = b + c + d + bcd = 0^4 \). Therefore, as a coset, \( IG \) corresponds to one point (i.e., \( 0^4 \)) in \( EG(4, 2) \). Similarly, it can be verified that all the effects in \( 1G \) correspond to \( a \) in \( EG(4, 2) \); therefore, the coset \( 1G \) corresponds to \( a \) of \( EG(4, 2) \). In general, two effects are aliased with each other if and only they correspond to the same point in \( EG(4, 2) \). This leads to a correspondence between the cosets of \( D \) and the points of \( EG(4, 2) \), as demonstrated in Table 2.

The remaining seven points of \( PG(4 - 1, 2) \) form the complementary design of \( D \), denoted by \( \bar{D} \). Let

\[
\bar{1} = bc, \bar{2} = bd, \bar{3} = cd, \bar{4} = abc, \bar{5} = abd, \bar{6} = acd, \bar{7} = abcd.
\]
Then $\bar{D}$ is a $2^{7-3}$ design with independent defining relations $\{123, 156, 345\}$. Similar to $D$, $\bar{D}$ also has 16 different cosets, which are $1\bar{G}, 1\bar{G}, 2\bar{G}, 3\bar{G}, 4\bar{G}, 5\bar{G}, 6\bar{G}, 7\bar{G}, 14\bar{G}, 17\bar{G}, 27\bar{G}, 37\bar{G}, 47\bar{G}, 57\bar{G}, 67\bar{G}, 147\bar{G}$, arranged from rank 0 to rank 15. Following similar arguments, the correspondence between the cosets of $\bar{D}$ and the points in $E\Gamma(4, 2)$ can be established as shown in Table 2.

Through the points in $E\Gamma(4, 2)$, the cosets of $D$ and $\bar{D}$ match with each other. We introduce a mapping $\tau^*$ from $\mathcal{F}$ to $\bar{\mathcal{F}}$ to pair the matching cosets. Let $i_1 \cdots i_7G$ be an arbitrary coset of $D$, $\tau^*(i_1 \cdots i_7G) = \bar{j}_1 \cdots \bar{j}_7\bar{G}$ if and only if the cosets $i_1 \cdots i_7G$ and $\bar{j}_1 \cdots \bar{j}_7\bar{G}$ correspond to the same point in $E\Gamma(4, 2)$. For example, $\tau^*(1G) = 14\bar{G}$, because $1G$ and $14\bar{G}$ match each other through $a$. See Table 2 for the other matching pairs.

<table>
<thead>
<tr>
<th>$E\Gamma(4, 2)$</th>
<th>$0^4$</th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$d$</th>
<th>$ab$</th>
<th>$ac$</th>
<th>$ad$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i_1 \cdots i_7G$</td>
<td>$1G$</td>
<td>$1G$</td>
<td>$2G$</td>
<td>$3G$</td>
<td>$4G$</td>
<td>$5G$</td>
<td>$6G$</td>
<td>$7G$</td>
</tr>
<tr>
<td>$\tau^*(i_1 \cdots i_7G)$</td>
<td>$1G$</td>
<td>$14\bar{G}$</td>
<td>$67\bar{G}$</td>
<td>$57\bar{G}$</td>
<td>$47\bar{G}$</td>
<td>$37\bar{G}$</td>
<td>$27\bar{G}$</td>
<td>$17\bar{G}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$E\Gamma(4, 2)$</th>
<th>$bcd$</th>
<th>$abc$</th>
<th>$bc$</th>
<th>$bd$</th>
<th>$abc$</th>
<th>$abd$</th>
<th>$cd$</th>
<th>$acd$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i_1 \cdots i_7G$</td>
<td>$8G$</td>
<td>$18G$</td>
<td>$23G$</td>
<td>$24G$</td>
<td>$26G$</td>
<td>$27G$</td>
<td>$28G$</td>
<td>$37G$</td>
</tr>
<tr>
<td>$\tau^*(i_1 \cdots i_7G)$</td>
<td>$147\bar{G}$</td>
<td>$7\bar{G}$</td>
<td>$1\bar{G}$</td>
<td>$2\bar{G}$</td>
<td>$4\bar{G}$</td>
<td>$5\bar{G}$</td>
<td>$3\bar{G}$</td>
<td>$6\bar{G}$</td>
</tr>
</tbody>
</table>

**Example 3.** Consider a $2^{13-9}$ design $D$ with independent defining relations

$$\{125, 136, 237, 1238, 149, 24t_0, 124t_1, 34t_2, 134t_3\},$$

where $t_0, \ldots, t_3$ represent factors 10, \ldots, 13, respectively. This design can be obtained by associating the factors to the points of $P\Gamma(4 - 1, 2)$ as follows.

$$1 = a, 2 = b, 3 = c, 4 = d, 5 = ab, 6 = ac, 7 = bc, 8 = abc, 9 = ad,$$

$$t_0 = bd, t_1 = abd, t_2 = cd, t_3 = acd.$$ 

The correspondence between the cosets of $D$ and the points of $E\Gamma(4, 2)$ is given in Table 3.
There are two remaining points \{bcd, abcd\} in PG(4−1, 2), which form a full 2^2 factorial design. In terms of \(H_4(2)\), the remaining columns bcd and abcd consist of 4 replicated 2^2 designs. In this example, the complementary design \(\tilde{D}\) of \(D\) is not of rank 4; instead, it is of rank 2 and is a replicated 2^2 full factorial design. We refer to \(\tilde{D}\) as a degenerate complementary design. The defining contrasts subgroup of \(\tilde{D}\) is \(G = \{I\}\), and \(\tilde{D}\) has only four cosets \(IG, \tilde{1}G, \tilde{2}G\) and \(\tilde{1}\tilde{2}G\), corresponding to the points \(0^4, bcd, abcd, \) and \(a\) of \(EG(4, 2)\), respectively. Therefore, only four cosets of \(D\) have corresponding cosets of \(\tilde{D}\), whereas the remaining 12 cosets of \(D\) do not have corresponding cosets of \(\tilde{D}\). This phenomenon occurs because the complementary design \(\tilde{D}\) is degenerate. For convenience, we expand the collection of cosets of \(\tilde{D}\) to include the empty set \(\varnothing\). Let \(\mathcal{F}^* = \{\varnothing\} \cup \mathcal{F}\). We define a mapping \(\tau^*\) from \(\mathcal{F}\) to \(\mathcal{F}^*\) as follows. A coset of \(D\) is mapped to a coset of \(\tilde{D}\) by \(\tau^*\) if the two cosets correspond to the same point in \(EG(4, 2)\). If a coset of \(D\) does not have a corresponding coset of \(\tilde{D}\), it is mapped to \(\varnothing\). The mapping \(\tau^*\) is shown in Table 3.

<table>
<thead>
<tr>
<th>(EG(4, 2))</th>
<th>(0^4)</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(ab)</th>
<th>(ac)</th>
<th>(bc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i_1 \cdots i_l G)</td>
<td>(IG)</td>
<td>(1G)</td>
<td>(2G)</td>
<td>(3G)</td>
<td>(4G)</td>
<td>(5G)</td>
<td>(6G)</td>
<td>(7G)</td>
</tr>
<tr>
<td>(\tau^*(i_1 \cdots i_l G))</td>
<td>(I\tilde{G})</td>
<td>(\tilde{1}\tilde{G})</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(EG(4, 2))</th>
<th>(abc)</th>
<th>(ad)</th>
<th>(bd)</th>
<th>(abd)</th>
<th>(cd)</th>
<th>(acd)</th>
<th>(bcd)</th>
<th>(abcd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i_1 \cdots i_l G)</td>
<td>(8G)</td>
<td>(9G)</td>
<td>(t_0 G)</td>
<td>(t_1 G)</td>
<td>(t_2 G)</td>
<td>(t_3 G)</td>
<td>(2t_2 G)</td>
<td>(2t_3 G)</td>
</tr>
<tr>
<td>(\tau^*(i_1 \cdots i_l G))</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\tilde{1}G)</td>
<td>(\tilde{2}G)</td>
</tr>
</tbody>
</table>

The mapping between \(\mathcal{F}\) and \(\mathcal{F}\) (or \(\mathcal{F}^*\)) illustrated in the previous two examples hold in general. We state it in the following proposition, and give a proof in the Appendix.

**Proposition 1.** Suppose \(D\) is a \(2^{n−p}\) fractional factorial design with rank \(k\) (= \(n − p\)) and \(\tilde{D}\) is its complementary design with rank \(\tilde{k}\). Let \(\mathcal{F}\) and \(\mathcal{F}\) be the collections of cosets of \(D\) and \(\tilde{D}\), respectively. Let \(\varnothing\) be the empty set.
(i) If $k = \bar{k}$, then both $\mathcal{F}$ and $\tilde{\mathcal{F}}$ contain $2^k$ cosets, and there exists a one-to-one mapping $\tau^*$ from $\mathcal{F}$ to $\tilde{\mathcal{F}}$ such that for an arbitrary coset $C \in \mathcal{F}$, $\tau^*(C)$ correspond to the same point in $EG(k, 2)$. In particular, $\tau^*(IG) = I\tilde{G}$, cosets of $D$ of order 1 are mapped to cosets of $D$ of order 2 or higher, and cosets of $D$ of order 2 or higher are mapped to cosets of $\tilde{D}$ of order 1.

(ii) If $k > \bar{k}$, then $\mathcal{F}$ contains $2^k$ cosets and $\tilde{\mathcal{F}}$ contains $2^k$ cosets, and there exists a mapping $\tau^*$ from $\mathcal{F}$ to $\tilde{\mathcal{F}} \cup \{\emptyset\}$ such that for an arbitrary coset $C \in \mathcal{F}$, $\tau^*(C)$ either is $\emptyset$ or corresponds to the same point in $EG(k, 2)$ as $C$. In particular, $\tau^*(IG) = I\tilde{G}$, $2^k - 2^k$ cosets of $D$ of order 1 are mapped to $\emptyset$, the remaining $n - (2^k - 2^k)$ cosets of $D$ of order 1 are mapped to cosets of $\tilde{D}$ of order 2 or higher, and cosets of $D$ of order 2 or higher are mapped to cosets of $\tilde{D}$ of order 1.

Under Scenario (ii) in Proposition 1, the coset pattern matrix $\tilde{A}$ of $\tilde{D}$ has $2^k$ rows, less than the number of rows of the coset pattern matrix $A$ of $D$. In order to treat the two scenarios uniformly in the rest of the article, we append $2^k - 2^k$ rows of zeros to $\tilde{A}$ to make it have the same number of rows as $A$. From Proposition 1, the first row of $A$ and the first row of $\tilde{A}$ correspond to each other, because they are patterns of $IG$ and $I\tilde{G}$ satisfying $\tau^*(IG) = I\tilde{G}$. The next $n$ rows of $A$, which are the patterns of the cosets of $D$ of order 1 (i.e., $A_{r1} = 1$ for $1 \leq r \leq n$) correspond to the last $n$ rows of $\tilde{A}$, which are either the rows of zeros or the patterns of the cosets of $\tilde{D}$ of order 2 or higher, due to the mapping $\tau^*$ between the cosets of $D$ of order 1 and $\emptyset$ or the cosets of $\tilde{D}$ of order 2 or higher. The last $2^{n-p} - n - 1$ rows of $A$, which are the patterns of the cosets of $D$ of order 2 or higher (i.e. $A_{r1} = 0$ for $n + 1 \leq r \leq 2^{n-p} - 1$), correspond to the rows of $\tilde{A}$, which are the patterns of the cosets of $D$ of order 1, again due to the mapping $\tau^*$ between the involved cosets. Therefore the rows of $A$ and $\tilde{A}$ are paired with each other through $\tau^*$. For ease of presentation, we introduce a mapping $\tau$ that maps the rows of $A$ to the rows of $\tilde{A}$. For row $i$ of $A$ ($0 \leq i \leq 2^{n-p} - 1$), $\tau(i)$ is the row of $\tilde{A}$ such that the coset corresponding to row $i$ of $A$ is mapped by $\tau^*$ to $\emptyset$ or the coset corresponding to row $\tau(i)$ of $\tilde{A}$. Therefore, $\tau$ is essentially a mapping from
\{0, 1, \ldots, 2^{n-p} - 1\} to itself and we refer to it as the \textit{complementary mapping}. The properties of \(\tau\) is summarized as a proposition given below.

**Proposition 2.** Let \(\tau\) be the mapping from \(\{0, 1, \ldots, 2^{n-p} - 1\}\) to itself induced by the pairing of rows of \(A\) and \(\bar{A}\). It has the following properties,

\(i\) \(\tau(0) = 0\);

\(ii\) \(2^{n-p} - n \leq \tau(r) \leq 2^{n-p} - 1\) for \(1 \leq r \leq n\);

\(iii\) \(1 \leq \tau(r) \leq 2^{n-p} - 1 - n\) for \(n + 1 \leq r \leq 2^{n-p} - 1\).

When proposing the complementary approach, Tang and Wu (1996) established the relationship between \(W_0 = (A_{0j})\) and \(\bar{W}_0 = (\bar{A}_{0j})\), both of which are the worldlength patterns of \(D\) and \(\bar{D}\), iteratively. As a matter of fact, a similar relationship holds for the coset patterns of any two paired cosets of \(D\) and \(\bar{D}\), that is, between \(W_r = (A_{rj})\) and \(W_{\tau(r)} = (A_{\tau(r)j})\). The next theorem states an explicit identity that relates \(A\) and \(\bar{A}\). We call it the \textit{coset pattern identity} between \(D\) and \(\bar{D}\). The proof of the theorem uses the identity for any three-way partition of \(PG(k - 1, 2)\) derived in Zhu (2003) and is given in the appendix.

**Theorem 1.** Suppose that \(D\) is a \(2^{n-p}\) design with coset pattern matrix \(A = (A_{ij})\) and its complementary design \(\bar{D}\) has coset pattern matrix \(\bar{A} = (\bar{A}_{ij})\). Let \(\tau\) be the complementary mapping between \(A\) and \(\bar{A}\). Then

\[
A_{i,j} = \frac{1}{m} \binom{n}{j} - \frac{1}{m} \sum_{j_1, j_2 = j} (-1)^{j_2} \binom{n - m/2}{j_1} \binom{m/2}{j_2} + (-1)^j \sum_{t_1 + t_2 = j} (-1)^{[t_2/2]} \binom{n - m/2}{[t_2/2]} \bar{A}_{\tau(i), t_1},
\]

where \(m = 2^{n-p}\) and \([x]\) denotes the largest integer that is less than or equal to \(x\).

The notation \(\binom{n}{k}\) in (1) is defined by

\[
\binom{n}{k} = \begin{cases} 
0, & \text{if } k < 0 \text{ or } k \text{ is not an integer}; \\
1, & \text{if } k = 0; \\
\frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots2\cdot1}, & \text{otherwise}.
\end{cases}
\]
It coincides with its usual definition \( n! / ((n - k)!k!) \) when both \( n \) and \( k \) are nonnegative integers such that \( n \geq k \).

When \( k < k \), there may not exist a coset of \( \bar{D} \) corresponding to a coset of \( D \) with a main effect as its leader. According to Theorem 1, the coset pattern of such coset can be explicitly calculated as

\[
A_{i,j} = \frac{1}{m} \binom{n}{j} - \frac{1}{m} \sum_{j_1 + j_2 = j} (-1)^{j_2} \binom{n-m/2}{j_1} \binom{m/2}{j_2},
\]

because this coset pattern is mapped to a row of zeroes.

In practice, effects of lower orders such as 1, 2, 3, and 4 are of particular interest, because effects of order higher than 4 are unlikely to be important. In other words, coset patterns \( A_{ij} \) with \( j \leq 4 \) are usually sufficient in practice. Working out (1) for \( 1 \leq j \leq 4 \), we derive the explicit relations between \( (A_{ij}) \) and \( (\bar{A}_{\tau(i)j}) \) \( (1 \leq j \leq 4) \), which are reported in the following corollary. In the rest of the article, we assume \( m = 2^{n-p} \).

**Corollary 1.** Let \( i' = \tau(i) \). For any \( 0 \leq i \leq 2^{n-p} - 1 \),

\[
\begin{align*}
A_{i,1} &= 1 - \bar{A}_{i',1} - \bar{A}_{i',0} \\
A_{i,2} &= b + \bar{A}_{i',2} + \bar{A}_{i',1} - b\bar{A}_{i',0} \\
A_{i,3} &= \binom{n}{2} - \frac{nm}{2} + \frac{m^2}{6} + \frac{1}{3} - \bar{A}_{i',3} - \bar{A}_{i',2} + b\bar{A}_{i',1} + b\bar{A}_{i,0} \\
A_{i,4} &= \binom{n}{3} - \frac{m}{2} \binom{n}{2} + \frac{nm^2}{6} - \frac{m^3}{24} - \frac{m}{3} + \frac{n}{3} \\
&\quad + \bar{A}_{i',4} + \bar{A}_{i',3} - b\bar{A}_{i',2} - b\bar{A}_{i',1} + \binom{b}{2} \bar{A}_{i',0}
\end{align*}
\]

where \( b = n - m/2 \), \( \bar{A}_{i',0} = 1 \) when \( i' = 0 \); and \( = 0 \), otherwise.

The equations for \( A_{0,3} \) and \( A_{0,4} \) were also reported in Chen and Hedayat (1996). If a coset of rank \( i \) contains a clear two-factor interaction, then \( A_{i,1} = 0 \) and \( A_{i,2} = 1 \). Applying the equations in Corollary 1, we have \( \bar{A}_{i',1} = 1 \) and \( \bar{A}_{i',2} = m/2 - n \). Because \( \bar{A}_{i',2} \geq 0 \), we have \( n \leq m/2 \), which implies that clear two-factor interactions exist only when \( n \leq m/2 \).

**Example 4.** The design \( D \) in Example 1 has six distinct coset patterns and so does its complement \( \bar{D} \). The cosets and coset patterns of \( \bar{D} \) are listed in Table 4. We place the
Table 4: The coset patterns of $\bar{D}$.

<table>
<thead>
<tr>
<th>Coset of $\bar{D}$</th>
<th>rows of $\bar{A}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{G}$</td>
<td>0 0 4 3 0 0 0</td>
</tr>
<tr>
<td>$1G, 2G, 3G, 4G, 5G, 6G$</td>
<td>1 2 2 2 1 0 0</td>
</tr>
<tr>
<td>$7\bar{G}$</td>
<td>1 0 0 4 3 0 0</td>
</tr>
<tr>
<td>$14\bar{G}$</td>
<td>0 3 4 0 0 1 0</td>
</tr>
<tr>
<td>$17\bar{G}, 27\bar{G}, 37\bar{G}, 47\bar{G}, 57\bar{G}, 67\bar{G}$</td>
<td>0 1 2 2 2 1 0</td>
</tr>
<tr>
<td>$147\bar{G}$</td>
<td>0 0 3 4 0 0 1</td>
</tr>
</tbody>
</table>

Table 5: The paired distinct rows of the coset pattern matrices.

<table>
<thead>
<tr>
<th>rows of $A$</th>
<th>rows of $\bar{A}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0 3 7 4 0 1 0</td>
<td>0 0 4 3 0 0 0 0</td>
</tr>
<tr>
<td>1 3 0 4 7 1 0 0</td>
<td>0 3 4 0 0 1 0 0</td>
</tr>
<tr>
<td>1 1 4 4 3 3 0 0</td>
<td>0 1 2 2 2 1 0 0</td>
</tr>
<tr>
<td>1 0 4 7 3 0 0 1</td>
<td>0 0 3 4 0 0 1 0</td>
</tr>
<tr>
<td>0 1 7 4 0 3 1 0</td>
<td>1 0 0 4 3 0 0 0</td>
</tr>
<tr>
<td>0 3 3 4 4 1 1 0</td>
<td>1 2 2 2 1 0 0 0</td>
</tr>
</tbody>
</table>

paired coset patterns of $D$ and $\bar{D}$ in the same row in Table 5. As an illustration of Theorem 1, we apply (1) to obtain the explicit expression of $A$ in terms of $\bar{A}$ given below.

$$A_{i,j} = \frac{1 - (-1)^j}{16} \binom{8}{j} + (-1)^j \bar{A}_{i',j} + (-1)^j \bar{A}_{i',j-1},$$

where $i' = \tau(i)$. The above equation can be further simplified by specifying $j$. For example, when $j = 4$, it becomes $A_{i,4} = \bar{A}_{i',4} + \bar{A}_{i',3}$. We can also expresses $\bar{A}$ in terms of $A$ as follows.

$$\bar{A}_{i',j} = \frac{1}{16} \binom{7}{j} - \frac{1}{16} \sum_{j_1+j_2=j} (-1)^{j_2} \binom{-1}{j_1} \binom{8}{j_2} + (-1)^j \sum_{t_1+t_2=j} (-1)^{[t_2/2]_j} \binom{-1}{t_2/2} \binom{A_{i,t_1}}{t_2},$$

Similarly, when $j = 4$, the equation becomes $\bar{A}_{i',4} = -8 + A_{i,4} + A_{i,3} + A_{i,2} + A_{i,1} + A_{i,0}$. 

13
Readers can verify that the paired coset patterns in Table 5 satisfy the simple equations derived above.

**Example 5.** The $2^{13-9}$ design discussed in Example 3 has 16 cosets and four distinct coset patterns. Its complementary design has 4 cosets and 3 distinct coset patterns. In order to match the number of distinct coset patterns for the designs, we include an additional row of zeros for the complementary design, so that the coset patterns of the two designs have a one-to-one correspondence, as listed in Table 6. Although the first and third rows in Table 6 for $\bar{A}$ appear to be the same, they are in fact different from each other. In fact, the first row is essentially the wordlength pattern, while the third row is the added row of zeros.

Table 6: The paired distinct rows of the coset pattern matrices.

<table>
<thead>
<tr>
<th>rows of $A$</th>
<th>rows of $\bar{A}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0 22 55 72 96 116 87 40 16 6 1 0</td>
<td>0 0</td>
</tr>
<tr>
<td>1 6 16 40 87 116 96 72 55 22 0 0 1</td>
<td>0 1</td>
</tr>
<tr>
<td>1 5 17 45 82 106 106 82 45 17 5 1 0</td>
<td>0 0</td>
</tr>
<tr>
<td>0 6 22 40 72 116 116 72 40 22 6 0 0</td>
<td>1 0</td>
</tr>
</tbody>
</table>

Applying Theorem 1, we obtain the explicit expression of $A$ in terms of $\bar{A}$ given below.

$$A_{i,j} = \frac{1}{16} \binom{13}{j} - \frac{1}{16} \sum_{j_1+j_2=j} (-1)^{j_2} \binom{5}{j_1} \binom{8}{j_2} + (-1)^j \sum_{t_1+t_2=j} (-1)^{[t_2/2]} \binom{5}{[t_2/2]} \bar{A}_{t_1,t_1},$$

where $i' = \tau(i)$. For example, when $j = 4$, we have $A_{i,4} = 45 - 5\bar{A}_{4,2} - 5\bar{A}_{4,1} + 10\bar{A}_{4,0}$. We can also express $\bar{A}$ in terms of $A$ as follows, which can be obtained directly from Corollary 1.

$$\bar{A}_{i,1} = 1 - A_{i,1} - A_{i,0} \quad \bar{A}_{i,2} = -6 + A_{i,2} + A_{i,1} + 6A_{i,0}.$$  

Readers can verify that the paired coset patterns in Table 6 satisfy the equations given above.
4 Application to Minimum $M$-Aberration Design

As briefly discussed in the introduction section, the coset pattern matrix can be used to define more elaborate criteria for discriminating and selecting designs. The minimum $M$-aberration criterion is such an example. In this section, we will first introduce the minimum $M$-aberration and then apply the coset pattern identity to construct minimum $M$-aberration designs via the complementary approach.

If an effect of order $i$ is aliased with another effect of order $j$ and both effects belong to a coset of order $k$, then we say that these two effects form a pair of aliased effects of type $(i, j)_k$. Let $M(i, j)_k$ be the number of pairs of aliased effects of type $(i, j)_k$. It is clear that $k \leq \min\{i, j\}$. For convenience, we always assume $i \leq j$. $M(i, j)_k$ can be calculated from the coset pattern matrix $A$ as follows. When $i = j$,

$$M(i, i)_k = \sum_{h \in R_k} A_{hi}(A_{hi} - 1)/2 = \sum_{h=1}^{m-1} I(A_{h1} = \cdots = A_{h,k-1} = 0) A_{hi}(A_{hi} - 1)/2,$$

and when $i \neq j$,

$$M(i, j)_k = \sum_{h \in R_k} A_{hi} A_{hj} = \sum_{h=1}^{m-1} I(A_{h1} = \cdots = A_{h,k-1} = 0) A_{hi} A_{hj},$$

where $R_k$ is the collection of the ranks of the cosets of order $k$. We arrange the $M(i, j)_k$’s into a sequence, denoted by $M$, according to the following order: $M(i, j)_k$ is placed ahead of $M(i', j')_k$ if (i) $i + j < i' + j'$; or (ii) $i + j = i' + j'$ and $|i - j| < |i' - j'|$; or (iii) $i = i'$, $j = j'$ and $k > k'$. We refer to $M$ as the aliasing type pattern of a design. The first 10 entries of $M$ are given below.

$$M = (M_{(1,2)}_1, M_{(2,2)}_2, M_{(2,2)}_1, M_{(1,3)}_1, M_{(2,3)}_2, M_{(2,3)}_1, M_{(1,4)}_1, M_{(3,3)}_3, M_{(3,3)}_2, M_{(3,3)}_1, \ldots).$$

Although the aliasing type pattern $M$ cannot be determined by $W_0$ alone, there does exist a relation between $M$ and $W_0$, which is

$$\sum_{k=1}^{i} M(i, j)_k = \sum_{k=0}^{i} \binom{n - i - j + 2k}{k} \binom{i + j - 2k}{i - k} A_{0,i+j-2k}. \quad (2)$$
A proof of (2) is given in the Appendix. Applying (2) for the first few entries of \( M \), we have

\[
M_{(1,2)_1} = 3A_{0,3}, \quad M_{(2,2)_2} + M_{(2,2)_1} = 3A_{0,4}, \quad M_{(1,3)_1} = 4A_{0,4}, \\
M_{(2,3)_2} + M_{(2,3)_1} = 3(n - 3)A_{0,3} + 10A_{0,5}, \quad M_{(1,4)_1} = (n - 3)A_{0,3} + 5A_{0,5}.
\]

Clearly the aliasing type pattern is more elaborate than the wordlength pattern when used to discriminate designs. Designs that sequentially minimize the entries of \( M \) are said to have minimum \( M \)-aberration. For further discussion about \( M \)-aberration, readers are referred to Zhu and Zeng (2005).

When \( n \geq 2^{n-p-1} \), the number of factors in \( D \) is larger than the number of factors in \( \bar{D} \) and the complementary approach becomes appealing. The following proposition asserts that the order of a coset of \( D \) cannot be higher than two when \( n \geq 2^{n-p-1} \).

**Proposition 3.** When \( n \geq 2^{n-p-1} \), then every coset of \( D \) except the coset with rank 0 has either a main effect or a two-factor interaction as its coset leader. In other words, the order of a coset of \( D \) can only be zero, one or two, when \( n \geq 2^{n-p-1} \).

According to Proposition 3, when \( n \geq 2^{n-p-1} \), \( M_{(i,j)_k} \)'s with \( k \geq 3 \) are all equal to zero. Therefore, the aliasing type pattern \( M \) can be simplified by removing \( M_{(i,j)_k} \)'s with \( k \geq 3 \). For convenience, we still use \( M \) to denote the simplified aliasing type pattern. Similarly, \( \bar{M}_{(i,j)_k} \) can be defined for the complementary design \( \bar{D} \). Because \( M_{(i,j)_k} \) and \( \bar{M}_{(i,j)_k} \) are respectively functions of \( A_{i,j} \) and \( \bar{A}_{i,j} \), which are related to each other by the coset pattern identity (1), we expect that \( M_{(i,j)_k} \) can be expressed in terms of \( \bar{M}_{(i,j)_k} \). This turns out to be true as stated in the following theorem.

**Theorem 2.** When \( n \geq 2^{n-p-1} \), \( M_{(i,j)_k} \) can be expressed in terms of \( \bar{M}_{(i,j)_k} \),

\[
M_{(i,j)_k} = c_{0,0,0} + \sum c_{i,j,k} \bar{M}_{(i,j)_k}
\]

where \( c_{i,j,k} \)'s are constants depending on \( n, p, i, j, \) and \( k \) only.

General expressions of \( c_{i,j,k} \)'s are fairly involved and thus are omitted in this article. The first seven entries of \( M \) are of particular interest, because they involve at least a main effect
or a two-factor interaction and both effects are of order lower than four. Therefore, we work out (3) explicitly for these seven terms and include the results in the following corollary.

**Corollary 2.**

\[
\begin{align*}
M_{(1,2)_1} &= \text{constant} - \bar{M}_{(1,2)_1} \\
M_{(2,2)_2} &= \text{constant} + \bar{M}_{(2,2)_1} + (b + 1)\bar{M}_{(1,2)_1} \\
M_{(2,2)_1} &= \text{constant} + \bar{M}_{(2,2)_2} - b\bar{M}_{(1,2)_1} \\
M_{(1,3)_1} &= \text{constant} + \bar{M}_{(1,3)_1} + (4/3)\bar{M}_{(1,2)_1} \\
M_{(2,3)_2} &= \text{constant} - \bar{M}_{(2,3)_1} - 2\bar{M}_{(2,2)_1} - (b + 1)\bar{M}_{(1,3)_1} + (a - 2)\bar{M}_{(1,2)_1} \\
M_{(2,3)_1} &= \text{constant} - \bar{M}_{(2,3)_2} - 2\bar{M}_{(2,2)_2} + b\bar{M}_{(1,3)_1} + (4b/3 - a + 1)\bar{M}_{(1,2)_1} \\
M_{(1,4)_1} &= \text{constant} - \bar{M}_{(1,4)_1} - (5/4)\bar{M}_{(1,3)_1} + (b - 1/3)\bar{M}_{(1,2)_1},
\end{align*}
\]

where “constant” means a constant only depending on \(n\), \(m = 2^n - p\), \(a = n(n - 1)/2 - nm/2 + m^2/6 + 1/3\), and \(b = n - m/2\).

Corollary 2 implies that sequentially minimizing

\[
M_{(1,2)_1}, \ M_{(2,2)_2}, \ M_{(2,2)_1}, \ M_{(1,3)_1}, \ M_{(2,3)_2}, \ M_{(2,3)_1}, \text{ and } M_{(1,4)_1}
\]

is equivalent to sequentially minimizing

\[
(-1)\bar{M}_{(1,2)_1}, \ \bar{M}_{(2,2)_1}, \ \bar{M}_{(2,2)_2}, \ \bar{M}_{(1,3)_1}, \ (-1)\bar{M}_{(2,3)_1}, \ (-1)\bar{M}_{(2,3)_2}, \text{ and } (-1)\bar{M}_{(1,4)_1}.
\]

Based on this fact, we can establish some general rules to identify minimum \(M\)-aberration designs with \(n\) factors and \(m = 2^n - p\) runs via the complementary approach.

**Rule 1.** Find \(\mathcal{D}_1\), the collection of designs with \(m - 1 - n\) factors that maximize \(\bar{M}_{(1,2)_1}\). If \(\mathcal{D}_1\) contains exactly one design, then the complement of the design has minimum \(M\)-aberration.

**Rule 2.** If \(\mathcal{D}_1\) contains more than one design, find \(\mathcal{D}_2 \subset \mathcal{D}_1\), the collection of designs that minimizes \(\bar{M}_{(2,2)_1}\) in \(\mathcal{D}_1\). If \(\mathcal{D}_2\) contains exactly one design, the complement of the design has minimum \(M\)-aberration.
Rule 3. If $\mathcal{D}_2$ contains more than one design, find $\mathcal{D}_3 \subset \mathcal{D}_2$, the collection of designs that minimizes $\bar{M}_{(2,2)}$ in $\mathcal{D}_2$. If $\mathcal{D}_3$ contains exactly one design, the complement of the design has minimum $M$-aberration.

Similar rules involving $\bar{M}_{(1,3)}$, $(-1)\bar{M}_{(2,3)}$, $(-1)\bar{M}_{(2,3)}$, and $(-1)\bar{M}_{(1,4)}$, respectively can be stated as Rules 1-3. When $m - 1 - n$ is not large, Rules 1-3 are usually sufficient to identify the minimum $M$-aberration designs.

**Appendix**

*Proof of Proposition 1.* Denote the $2^k - 1$ points in $PG(k - 1, 2)$ as $p_1, p_2, \ldots, p_m$ where $m = 2^k - 1$. Then $EG(k, 2) = \{0^k\} \cup PG(k - 1, 2)$. Because $D$ is a $2^{n-p}$ fractional factorial design with rank $k$, it is equivalent to a collection of $n$ points from $PG(k - 1, 2)$ of rank $k$. Without loss of generality, assume $D = \{p_1, p_2, \ldots, p_n\}$. Then the remaining $m - n$ points of $PG(k - 1, 2)$ form the complementary design, that is, $\tilde{D} = \{p_{n+1}, \ldots, p_m\}$.

$D$ involves $n$ factors (or equivalently $n$ points) and has in total $2^n$ effects including the grand mean. Because the rank of $D$ is $k$, the points or factors $p_1, p_2, \ldots, p_n$ are linearly dependent. Without loss of generality, we assume that $p_1, p_2, \ldots, p_k$ are linearly independent. Then, the remaining points $p_{k+1}, \ldots, p_n$ can be generated from $p_1, p_2, \ldots, p_k$ via linear combination. These generating relations are the so-called defining relations for $D$, which further generate the defining contrasts subgroup $G$. $G$ consists of $2^n$ effects which are aliased with the grand mean $I$. Using $G$, the $2^n$ effects of $D$ are partitioned into $2^k$ cosets, each of which contains $2^{n-k}$ effects aliased with each other.

Similarly, $\tilde{D}$ involves $m - n$ factors (or points) and has in total $2^{m-n}$ effects. Because the rank of $\tilde{D}$ is $\bar{k}$, only $\bar{k}$ points are linearly independent. The defining contrasts subgroup $\tilde{G}$ of $\tilde{D}$ consists of $2^{m-n-k}$ effects aliased with the grand mean, and the $2^{m-n}$ effects are partitioned into $2^k$ cosets, each of which contains $2^{m-n-k}$ effects aliased with each other.

Let $C$ be an arbitrary coset of $D$, and let $i_1i_2\cdots i_h$ and $j_1j_2\cdots j_l$ be two arbitrary effects in $C$. In terms of the points in $PG(k - 1, 2)$, the two effects are $p_{i_1}p_{i_2}\cdots p_{i_h}$ and $p_{j_1}p_{j_2}\cdots p_{j_l}$.
and they are aliased with each other if and only if
\[ p_{i_1} + p_{i_2} + \cdots + p_{i_h} = p_{j_1} + p_{j_2} + \cdots + p_{j_l}. \]

Because all the points involved in the sums above are in \( PG(k - 1, 2) \), the sums must be the same point in \( EG(k, 2) \). We denote the point by \( s \). It is not difficult to see that every effect in \( C \) sums to \( s \). Thus the coset \( C \) corresponds to \( s \). Furthermore, if two cosets \( C_1 \) and \( C_2 \) correspond to a same point in \( PG(k - 1, 2) \), then \( C_1 \) and \( C_2 \) must be identical. Therefore, every coset of \( D \) corresponds to a point of \( EG(k, 2) \) uniquely. Similarly, we can show that every coset of \( \tilde{D} \) corresponds to a point of \( EG(k, 2) \) uniquely. Now we are ready to prove (i) and (ii) of the proposition.

We first prove (i). When \( k = \bar{k} \), the number of cosets of \( D \), the number of cosets of \( \tilde{D} \), and the number of points in \( EG(k, 2) \) are all equal to \( 2^k \). Since the correspondence from the cosets of \( D \) or \( \tilde{D} \) to \( EG(k, 2) \) is unique, a one-to-one mapping \( \tau^* \) from the cosets of \( D \) to the cosets of \( \tilde{D} \) follows. For any coset \( C \in \mathcal{F} \), \( \tau^*(C) \) is the coset in \( \mathcal{F} \) such that \( C \) and \( \tau^*(C) \) correspond to the same point in \( EG(k, 2) \). Because \( IG \) and \( I\tilde{G} \) correspond to \( 0^k \) in \( EG(k, 2) \), so \( \tau^*(IG) = IG \). For any coset of \( D \) of order 2 or higher (e.g., \( C = i_1 i_2 \cdots i_h G \), where \( i_1 i_2 \cdots i_h \) is the coset leader with \( h \geq 2 \)), consider \( s = p_{i_1} + p_{i_2} + \cdots + p_{i_h} \). We claim that \( s \) cannot be a point in \( D = \{ p_1, p_2, \ldots, p_n \} \), because if \( s = p_i \) for some \( i \) between 1 and \( n \), then the main effect \( i \) must be in the coset and the coset’s order is 1 instead of 2 or higher, which is a contradiction. Moreover, \( s \) cannot be \( 0^k \), because otherwise, the coset is \( IG \) and of order zero. Since \( s \) must be a point in \( EG(k, 2) \), it must be in \( \tilde{D} = \{ p_{n+1}, \ldots, p_m \} \). Assume \( s = p_j \) where \( n + 1 \leq j \leq m \). Then \( \tau^*(i_1 i_2 \cdots i_h G) = j\tilde{G} \), which is a coset of order 1. Similarly, for any coset of \( \tilde{D} \) of order 2 or higher, its inverse under \( \tau^* \) is a coset of \( D \) of order 1. Because \( \tau^* \) is one to one, (i) is proved.

Next we prove (ii). When \( k > \bar{k} \), the number of cosets of \( D \) is \( 2^k \) and the number of cosets of \( \tilde{D} \) is \( 2^k \) (\(< 2^k \)). The cosets of \( D \) have a one to one correspondence with the points of \( EG(k, 2) \), but the cosets of \( \tilde{D} \) now only have a one to one correspondence with a subset of points of \( EG(k, 2) \) with cardinality \( 2^k \). For a coset \( C \) of \( D \) and a coset \( \tilde{C} \) of \( \tilde{D} \) that correspond to the same point, \( C \) is mapped to \( \tilde{C} \), that is, \( \tau^*(C) = \tilde{C} \). Under this mapping,
$2^k - 2^k$ cosets of $D$ do not have corresponding cosets of $D$. We simply map them to the empty set $\emptyset$. For cosets of $D$ that are not mapped to $\emptyset$, similar to the proof of (i), we can show that cosets of $D$ of order 1 are mapped to cosets of $\bar{D}$ of order 2 or higher and cosets of $D$ of order 2 or higher are mapped to cosets of $\bar{D}$ of order 1. What remains to be shown is that the cosets that are mapped to $\emptyset$ must be cosets of order 1. This immediately follows from the fact that the cosets of $D$ of order 2 or higher must correspond to a point in $\bar{D}$.

The proof of Theorem 1 uses a result in Zhu (2003), which is restated below as Lemma 1.

A general word of $PG(k - 1, 2)$ is defined to be a collection of points that sum to 0 in $F_2$.

**Lemma 1 (Equation (35) in Zhu (2003)).** Suppose that $S_1$, $S_2$ and $S_3$ form a three-way partition of $PG(k - 1, 2)$. Let the number of points in $S_1$, $S_2$, and $S_3$ be $l_1$, $l_2$, $l_3$, respectively. Let $N_{i,j,k}$ be the number of general words that contain $i$ points of $S_1$, $j$ points of $S_2$, and $k$ points of $S_3$. Then

$$N_{i,j,0} = \frac{1}{m} \binom{l_1}{i} \binom{l_2}{j} - \frac{1}{m} \sum_{i_1 + i_2 = i} (-1)^{i_2} \binom{l_1 - m/2}{i_1} \binom{m/2}{i_2} \binom{l_2}{j}$$

$$+ \sum_{t_1 + t_2 = i} \sum_{s_2 + s_3 = t_1} \binom{l_1 - m/2}{[t_2/2]} (-1)^{s_3} N_{0,u,s_3} Q_{l_2,u} (s_2, j),$$

where $m = 2^k$ and

$$Q_{n,k}(s,t) = (-1)^t \binom{k}{(t-s+k)/2} \binom{n-k}{(t+s-k)/2}.$$

**Remark 1.** According to the definition of $N_{i,j,k}$, the subscripts $i$, $j$, $k$ are all nonnegative and $N_{i,j,k} = 0$ if $i > l_1$, or $j > l_2$, or $k > l_3$.

**Proof of Theorem 1.** We consider the following three scenarios separately.

(i) When $i = 0$, then $\tau(i) = 0$. Let $S_1 = D$, $S_2 = \{\bar{a}\}$, and $S_3 = \bar{D}\setminus\{\bar{a}\}$, where $\bar{a}$ is an arbitrary point in $\bar{D}$. The equation (1) is obtained by applying Lemma 1 with $l_1 = n$ and $l_2 = 1$, noticing that $A_{0,j} = N_{j,0,0}$ and $\bar{A}_{0,j} = N_{0,0,j} + N_{0,1,j-1}$.
(ii) When \(1 \leq i \leq n\), then \(2^{n-p} - n \leq \tau(i) \leq 2^{n-p} - 1\). Suppose the coset leader of the corresponding coset for \(D\) is \(a \in D\). Let \(S_1 = D \setminus \{a\}\), \(S_2 = \{a\}\), and \(S_3 = \bar{D}\). The equation (1) is obtained by applying Lemma 1 with \(l_1 = n - 1\) and \(l_2 = 1\), noticing that \(A_{i,j} = N_{j-1,0,0} + N_{j,1,0}\) and \(\bar{A}_{\tau(i),j} = N_{0,1,j}\).

(iii) When \(n + 1 \leq i \leq 2^{n-p} - 1\), then \(1 \leq \tau(i) \leq 2^{n-p} - 1 - n\). Suppose the coset leader of the corresponding coset for \(\bar{D}\) is \(\bar{a}\). Let \(S_1 = D\), \(S_2 = \{\bar{a}\}\), and \(S_3 = \bar{D} \setminus \{\bar{a}\}\). The equation (1) is obtained by applying Lemma 1 with \(l_1 = n\) and \(l_2 = 1\), noticing that \(A_{i,j} = N_{j,1,0}\) and \(\bar{A}_{\tau(i),j} = N_{0,0,j-1} + N_{0,1,j}\).

When \(D\) is degenerate, the same proof still applies.

Proof of Equation (2). The two sides of the equation provide two different ways to count the number of pairs of an \(i\)-factor interaction and a \(j\)-factor interaction that are aliased with each other. When the pairs are classified according to which coset they belong to, the total number of pairs is \(\sum_{k=1}^{i} M_{(i,j)_k}\), the left-hand side of the equation. When the pairs are classified according to whether two effects share common factors, it yields the right-hand side of the equation. In fact, when they do not share any common factors, this type of aliasing pair can be derived from a word of length \(i + j\), and the total count is \((\binom{i+j}{i})A_{0,i+j}\). When they share exactly one common factor, this type of aliasing pair can be derived from a word of length \(i + j - 2\), and the total count is \((\binom{n-i-j+2}{1})\binom{i+j-2}{i-1}A_{0,i+j-2}\). Following the similar arguments, we can obtain other items on the right-hand side of the equation.

Proof of Proposition 3. Suppose that one coset of rank \(i\) has order larger than two. It implies \(A_{i,1} = A_{i,2} = 0\). Therefore, let \(i' = \tau(i)\) and, noticing that \(\bar{A}_{i',0} = 0\), we have

\[
0 = A_{i,1} = 1 - \bar{A}_{i',1} - \bar{A}_{i',0},
\]
\[
0 = A_{i,2} = (n - 2^{n-p-1}) + \bar{A}_{i',2} + \bar{A}_{i',1} - (n - 2^{n-p-1})\bar{A}_{i',0},
\]

which yields \(\bar{A}_{i',1} = 1\) and \(\bar{A}_{i',2} = -(n - 2^{n-p-1}) - 1\). When \(n \geq 2^{n-p-1}\), \(\bar{A}_{i',2} < 0\), which is impossible. Hence this proposition holds.
Alternative Proof of Proposition 3. This proof is provided to the authors by a private communication. For any point in $\bar{D}$, there are exactly $2^{n-p-1} - 1$ lines passing through it. Those lines can be categorized into three types according to whether the other two points are (i) both in $D$, (ii) one in $D$ and one in $\bar{D}$, (iii) both in $\bar{D}$. Let $r$, $s$, and $t$ denote the number of lines of these three types. Then $r + s + t = 2^{n-p-1} - 1$ and $2r + s = n$. Hence $t = 2^{n-p-1} - 1 - n + r$. If $n \geq 2^{n-p-1}$ then $r > 0$ because $t$ cannot be negative. So every alias set given by a point of $D$ has a main effect and every alias set given by a point of $\bar{D}$ has a 2fi.

Proof of Theorem 2. When $n \geq 2^{n-p-1}$, there are only four different types of patterns in $M$, which are $M_{(i,j)_1}$, $M_{(i,i)}$, $M_{(i,j)_2}$, and $M_{(i,e)}$. In what follows, We will prove the theorem for the pattern $M_{(i,j)_1}$ only. The proof for the other patterns are similar and thus omitted. Let

$$c_k = \frac{1}{m} \binom{n}{k} - \frac{1}{m} \sum_{j_1+j_2=k} (-1)^{j_2} \binom{n-m/2}{j_1} \binom{m/2}{j_2}.$$ 

Then we can write $M_{(i,j)_1}$ as

$$M_{(i,j)_1} = \sum_{h \in R_1} A_{hi} A_{hj}$$

$$= \sum_{h \in \cup_{k \geq 2} R_k} \left\{ c_i + (-1)^i \sum_{t_1+t_2=i} (-1)^{t_2/2} \binom{n-m/2}{t_2/2} A_{h,t_1} \right\}$$

$$\left\{ c_j + (-1)^j \sum_{t_1+t_2=j} (-1)^{t_2/2} \binom{n-m/2}{t_2/2} A_{h,t_1} \right\}$$

$$= c_i c_j n + c_j (-1)^i \sum_{t_1+t_2=i} (-1)^{t_2/2} \binom{n-m/2}{t_2/2} \sum_{h \in \cup_{k \geq 2} R_k} A_{h,t_1}$$

$$+ c_i (-1)^j \sum_{t_1+t_2=j} (-1)^{t_2/2} \binom{n-m/2}{t_2/2} \sum_{h \in \cup_{k \geq 2} R_k} A_{h,t_1}$$

$$+ (-1)^{i+j} \sum_{t_1+t_2=i} \sum_{t_3+t_4=j} (-1)^{t_2/2+t_4/2} \binom{n-m/2}{t_2/2} \binom{n-m/2}{t_4/2} \sum_{h \in \cup_{k \geq 2} R_k} A_{h,t_1} A_{h,t_3}$$

Note that when $t_1 \neq t_3$

$$\sum_{h \in \cup_{k \geq 2} R_k} A_{h,t_1} A_{h,t_3} = \sum_{k \geq 2} M_{(t_1,t_3)_k}$$

22
and when \( t_1 = t_3 \)

\[
\sum_{\bar{h} \in \mathcal{U}_{k \geq 2}} \bar{A}_{\bar{h},t_1} \bar{A}_{\bar{h},t_3} = 2 \sum_{k \geq 2} M_{(t_1,t_3)k} + \sum_{\bar{h} \in \mathcal{U}_{k \geq 2}} \bar{A}_{\bar{h},t_1}.
\]

We also have

\[
\sum_{\bar{h} \in \mathcal{U}_{k \geq 2}} \bar{A}_{\bar{h},t_1} = \binom{m-1-n}{t_1} - \bar{A}_{0,t_1} - \sum_{\bar{h} \in \mathcal{R}_1} \bar{A}_{\bar{h},t_1} = \binom{m-1-n}{t_1} - \bar{A}_{0,t_1} - M_{(1,t_1)1}.
\]

Applying equation (2) to \( \bar{D} \) yields

\[
\bar{M}_{(1,j)1} = (j+1)\bar{A}_{0,j+1} + (m-1-n-j+1)\bar{A}_{0,j-1},
\]

which implies

\[
\bar{A}_{0,j+1} = \bar{M}_{(1,j)1}/(j+1) - (m-n-j)\bar{A}_{0,j-1}/(j+1).
\]

Since \( \bar{A}_{0,3} = (1/3)\bar{M}_{(1,2)1}, \bar{A}_{0,4} = (1/4)\bar{M}_{(1,3)1} \), we can express all the remaining \( \bar{A}_{0,j} \) in terms of \( \bar{M}_{(1,j)1} \). Therefore, all terms involving \( \bar{A}_{i,j} \) on the right-hand side of the expression of \( M_{(i,j)1} \) can be expressed in terms of \( \bar{M}_{(i,j)1} \), and so can \( M_{(i,j)1} \).

**Proof of Corollary 2.** The proof of this corollary follows the general strategy discussed in the proof of Theorem 2. We only verify one equation as a demonstration. Reader can verify the remaining equations similarly.

\[
M_{(2,2)2} = \sum_{h \in \mathcal{R}_2} A_{h,2}(A_{h,2} - 1)/2 = \sum_{h \in \mathcal{R}_1} (b + \bar{A}_{h,2} + \bar{A}_{\bar{h},1})(b + \bar{A}_{h,2} + \bar{A}_{\bar{h},1} - 1)/2
\]

\[
= \text{constant} + (b + 1) \sum_{h \in \mathcal{R}_1} \bar{A}_{h,2} + \sum_{h \in \mathcal{R}_1} \bar{A}_{h,2}(\bar{A}_{h,2} - 1)/2
\]

\[
= \text{constant} + (b + 1)\bar{M}_{(1,2)1} + \bar{M}_{(2,2)1}.
\]

**References**


23


Peng Zeng
Department of Mathematics and Statistics, Auburn University, Auburn, AL 38849, USA
Email: zengpen@auburn.edu
Hong Wan
School of Industrial Engineering, Purdue University, West Lafayette, IN 47907, USA
Email: hwan@purdue.edu

Yu Zhu
Department of Statistics, Purdue University, West Lafayette, IN 47907, USA
Email: yuzhu@stat.purdue.edu