The Annals of Statistics
1995, Vol. 23, No. 4, 1223–1233

SOME PROJECTION PROPERTIES OF ORTHOGONAL ARRAYS[1]

BY CHING-SHUI CHENG

University of California, Berkeley

The definition of an orthogonal array imposes an important geometric property: the projection of an OA($\lambda 2^t, 2^k, t$), a $\lambda 2^t$-run orthogonal array with k two-level factors and strength t, onto any t factors consists of λ copies of the complete 2^t factorial. In this article, projections of an OA($N, 2^k, t$) onto $t+1$ and $t+2$ factors are considered. The projection onto any $t+1$ factors must be one of three types: one or more copies of the complete 2^{t+1} factorial, one or more copies of a half-replicate of 2^{t+1} or a combination of both. It is also shown that for $k \geq t+2$, only when N is a multiple of 2^{t+1} can the projection onto some $t+1$ factors be copies of a half-replicate of 2^{t+1}. Therefore, if N is not a multiple of 2^{t+1}, then the projection of an OA($N, 2^k, t$) with $k \geq t+2$ onto any $t+1$ factors must contain at least one complete 2^{t+1} factorial. Some properties of projections onto $t+2$ factors are established and are applied to show that if N is not a multiple of 8, then for any OA($N, 2^k, 2$) with $k \geq 4$, the projection onto any four factors has the property that all the main effects and two-factor interactions of these four factors are estimable when the higher-order interactions are negligible.

1. Introduction. Orthogonal arrays, first introduced by Rao (1946, 1947), have been used extensively in factorial designs. Specifically, an orthogonal array of size N, k constraints, s levels and strength t, denoted OA(N, s^k, t), is a $k \times N$ matrix $\mathbf{X}$ of s symbols such that all the ordered t-tuples of the symbols occur equally often as column vectors of any $t \times N$ submatrix of $\mathbf{X}$. It is clear that N must be of the form λs^t, where λ is usually called the index of the orthogonal array. In applications to factorial designs, each row corresponds to a factor, the symbols are factor levels and each column represents a combination of the factor levels. Thus every OA(N, s^k, t) defines an N-run factorial design for k factors each having s levels.

This definition imposes an important projection property of an orthogonal array: when projected onto any t factors, it yields λ copies of a complete factorial. Statistically this geometric property implies that all the main effects and interactions of any t factors are estimable when the other factors are ignored. It also relates to the concept of *resolution* introduced by Box and Hunter (1961): *regular* fractional factorial designs (those constructed by using defining relations) of resolution $t+1$ are orthogonal arrays with strength t. In general, an orthogonal array with strength $t = 2s$ (respec-

Received September 1993; revised June 1994.
[1]Research supported by NSF Grant DMS-91-00938.
AMS 1991 *subject classification*. Primary 62K15.
Key words and phrases. Foldover, orthogonal array, Plackett–Burman design, resolution.

tively, $2s - 1$) can be used to estimate all the main effects and interactions involving at most s (respectively, $s - 1$) factors under the assumption that all the interactions involving more than s factors are negligible. For example, an orthogonal array with strength two defines a design in which all the main effects are estimable if the interactions are negligible. Important examples of such arrays are the Plackett–Burman designs (1946), which are constructed from Hadamard matrices. Recall that a Hadamard matrix of order n is an $n \times n$ $(1, -1)$-matrix in which any two columns (and hence any two rows) are orthogonal. Such a matrix can be normalized so that all the entries in the first row are equal to 1. An $\mathrm{OA}(n, 2^{n-1}, 2)$ can be obtained by deleting the first row. Therefore, if an $n \times n$ Hadamard matrix exists, then it can be used to conduct an experiment with $n - 1$ two-level factors in only n runs. Such a design allows the estimation of all main effects under the assumption that the interactions are not present. The economy in run sizes provided by such designs makes them suitable for screening experiments in which the primary purpose is to identify important factors.

The purpose of this paper is to investigate the projections of orthogonal arrays onto more than t factors. Studying projections onto more than t factors can shed light on other statistical properties of orthogonal arrays. Even though many potential factors may be considered in the initial stage of a study, the number of active factors which have significant effects is often small. This is called *effect sparsity*. An $\mathrm{OA}(n, 2^{n-1}, 2)$ can be used to estimate all the main effects when the interactions are negligible. However, if there are just a few active factors, then it may be possible to study their interactions.

Projections of Plackett–Burman designs were studied by Lin and Draper (1991, 1992). Their computer searches found all the projections of 12-, 16-, 20-, 24-, 28-, 32-, and 36-run Plackett–Burman designs onto three factors. For example, they found that for each of these designs, the projection onto any three factors must be one of the following three types: one or more copies of the complete 2^3 factorial, one or more copies of a half-replicate of 2^3 or a combination of both. In particular, the projection of a 12-run Plackett–Burman design onto *any* three factors is always a 2^3 complete factorial plus a half-replicate. This was also observed by Box and Bisgaard (1993), who commented that the interesting projective properties of Plackett–Burman designs, which the experimenters have sometimes been reluctant to use for industrial experimentation due to their complicated alias structures, provide a compelling rationale for their use.

Lin and Draper further considered projections of 12-, 16-, 20- and 24-run Plackett–Burman designs onto four and five factors. Projections onto four factors begin to look messier. However, projecting the 12-run Plackett–Burman design onto any four factors always yields a design with the property that all the main effects and two-factor interactions of the four factors are estimable when the higher-order interactions are negligible [Lin and Draper (1993), Wang and Wu (1995)]. Wang and Wu (1995) also observed this important property for 20-run Plackett–Burman designs, and coined the term *hidden projection*. They attributed the success of Hamada and Wu's (1992) strategy for entertaining and estimating two-factor interactions from Plackett–Burman type designs to the hidden projection property. Wang and Wu (1995) further examined projections onto five factors. In this case, even though the resulting projections may not allow the estimation of all the two-factor interactions, many of them can be estimated when the others are assumed negligible.

These are mostly computer works. This paper is an attempt to derive some general results on the projections of two-level orthogonal arrays. Some of the structures evidenced in earlier computation results can be derived analytically and shown to hold more generally than the scope limited by the computer power. Such a study also provides more insight.

Section 2 is devoted to properties of projections of an $\mathrm{OA}(N, 2^k, t)$ onto $t + 1$ factors. A result by Seiden and Zemach (1966) implies that for any $\mathrm{OA}(N, 2^k, t)$, the projection onto any $t + 1$ factors must be one of three types: one or more copies of the complete 2^{t+1} factorial, one or more copies of a half-replicate of 2^{t+1} or a combination of both. So the pattern observed by Lin and Draper (1992) for Plackett–Burman designs is a special case. We also show that for $k \geq t + 2$, only when N is a multiple of 2^{t+1} can the projection onto some $t + 1$ factors be copies of a half-replicate of 2^{t+1}. This result has important statistical consequences. Section 3 discusses projections onto $t + 2$ factors for the case in which t is even. These results are applied in Section 4 to show that if N is *not a multiple of* 8, then for an $\mathrm{OA}(N, 2^k, 2)$ with $k \geq 4$, the projection onto any four factors allows the estimation of all the main effects and two-factor interactions.

Throughout the rest of the paper, only two-level orthogonal arrays will be considered. The levels of each factor will be denoted 1 and -1. For any two $1 \times N$ vectors $\mathbf{x} = (x_1, \dots, x_N)$ and $\mathbf{y} = (y_1, \dots, y_N)$, we define their *Hadamard product* $\mathbf{x} \circ \mathbf{y}$ to be the vector $(x_1 y_1, \dots, x_N y_N)$.

2. Projections onto $t + 1$ factors. Proposition 2.2 of Seiden and Zemach (1966) showed that in an $\mathrm{OA}(\lambda 2^t, 2^{t+1}, t)$, any two columns differing in an even number of components appear the same number of times, while any two columns differing in an odd number of components appear together λ times. This result can be rephrased as the following theorem.

THEOREM 2.1. *Suppose* $\mathbf{X}$ *is an* $\mathrm{OA}(N, 2^k, t)$ *with* $k \geq t + 1$. *Let* $\mathbf{Y}$ *be a* $(t + 1) \times N$ *submatrix of* $\mathbf{X}$. *Then there exist two nonnegative integers* α *and* β *such that each* $(t + 1) \times 1$ *vector* $\mathbf{x} = (x_1, x_2, \dots, x_{t+1})^T$ *with* $x_1 x_2 \cdots x_{t+1} = 1$, *where* $x_i = 1$ *or* -1, *appears* α *times as a column vector of* $\mathbf{Y}$, *and each of those with* $x_1 x_2 \cdots x_{t+1} = -1$ *appears* β *times*.

In Theorem 2.1, if $\alpha = 0$ (or $\beta = 0$), then the projection of $\mathbf{X}$ onto these $t + 1$ factors is $2^{-t}N$ copies of the half-replicate of 2^{t+1} consisting of the combinations satisfying $x_1 x_2 \cdots x_{t+1} = -1$ (or $x_1 x_2 \cdots x_{t+1} = 1$) where $x_1, x_2, \dots, x_{t+1}$ are the levels of the $t + 1$ factors onto which the orthogonal array is projected. We call this kind of projections type I. If $\alpha = \beta$, then the projection is $2^{-(t+1)}N$ copies of the complete 2^{t+1} factorial and is called type II. Otherwise the projection contains copies of a complete 2^{t+1} factorial plus copies of a half replicate and is called type III.

Both types II and III projections contain at least one copy of the complete 2^{t+1} factorial. Type II projections are the best, while type I projections are the least desirable. From the statistical efficiency point of view, α and β should be as close to each other as possible. Unfortunately, type I projections always arise in projections of the regular fractional factorial designs, which are popular in practice because of their simple alias structures. For a regular fractional factorial design with strength t, there must exist certain $t+1$ factors such that all the combinations in the design satisfy $x_{i_1}x_{i_2}\cdots x_{i_{t+1}} = 1$ (or -1), where $x_{i_1}, x_{i_2}, \dots$ and $x_{i_{t+1}}$ are the levels of these $t+1$ factors. It follows that the projection onto these factors is of type I.

Even nonregular designs can produce type I projections. For any $k \times N$ $(1, -1)$-matrix $\mathbf{X}$, let $\bar{\mathbf{X}}$ be obtained from $\mathbf{X}$ by interchanging 1 and -1 and let $\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X} & \bar{\mathbf{X}} \\ -\mathbf{1} & \mathbf{1} \end{bmatrix}$ where $\mathbf{1}$ is the $1 \times N$ vector of 1's. This array is called the *foldover* of $\mathbf{X}$. Seiden and Zemach (1966) proved that if $\mathbf{X}$ is an OA$(N, 2^k, t)$ and t is *even*, then $\tilde{\mathbf{X}}$ is an orthogonal array with strength $t+1$. Let $\mathbf{X}$ be an OA$(n, 2^{n-1}, 2)$ obtained from an $n \times n$ Hadamard matrix, as described in the Introduction. Then $\tilde{\mathbf{X}}$ is an OA$(2n, 2^n, 3)$. Pick any row of $\tilde{\mathbf{X}}$ and form its Hadamard product with each of the remaining $n-1$ rows. Supplementing $\tilde{\mathbf{X}}$ with these $t-1$ products as row vectors, we obtain an OA$(2n, 2^{2n-1}, 2)$. Clearly the projections of this array onto certain sets of three factors are of type I. If the conjecture that an $n \times n$ Hadamard matrix exists for every n which is a multiple of 4 is true, then we would have shown that for every N that is a multiple of 8, there exists a two-level orthogonal array of size N and strength 2 whose projections onto certain sets of three factors are of type I. In the following, we shall prove that for $k \geq 4$, this can never happen when N is *not* a multiple of 8. In general, if N is *not* a multiple of 2^{t+1}, then the projection of an OA$(N, 2^k, t)$ with $k \geq t+2$ onto any $t+1$ factors must be of type III. We first prove a preliminary result.

LEMMA 2.2. *Let $\mathbf{X}$ be an OA$(N, 2^k, t)$ with $k \geq t+2$. If the projection of $\mathbf{X}$ onto certain $t+1$ factors, say factors $i_1, i_2, \dots, i_{t+1}$, is of type I, then its projection onto any other $t+1$ factors which have exactly t factors in common with $i_1, i_2, \dots, i_{t+1}$ must be of type II.*

PROOF. Without loss of generality, we may assume that the projection of $\mathbf{X}$ onto the first $t+1$ rows is of type I, and we shall show that its projection onto rows $2, \dots,$ and $t+2$ is of type II. Since the projection of $\mathbf{X}$ onto the first $t+1$ rows is of type I, by changing, if necessary, the signs of all the entries in the same row, we may assume that every column vector $\mathbf{x} = (x_1, x_2, \dots, x_k)^T$ of $\mathbf{X}$ satisfies $x_1 x_2 \cdots x_{t+1} = 1$; that is,

$$x_1 = x_2 \cdots x_{t+1}. \tag{2.1}$$

Suppose the projection of $\mathbf{X}$ onto rows $2, \ldots, t+2$ is not of type II. Then by Theorem 2.1, it must be of either type I or type III. We shall show that this would lead to a contradiction. Let $\mathbf{Z}$ be the submatrix of $\mathbf{X}$ consisting of rows $2, \ldots,$ and $t+2$. Then without loss of generality, we may assume that there is a nonnegative integer s (which can be zero) such that the first $s2^{t+1}$ columns of $\mathbf{Z}$ are s copies of all the 2^{t+1} vectors of ± 1's each of length $t+1$, and all the last $N - s2^{t+1}$ columns of $\mathbf{X}$ satisfy

$$x_{t+2} = x_2 \cdots x_{t+1}. \tag{2.2}$$

Now consider the first and the $(t+2)$nd rows of $\mathbf{X}$. It is clear that all the four pairs $(1,1)^T$, $(1,-1)^T$, $(-1,1)^T$ and $(-1,-1)^T$ appear equally often in the first $s2^{t+1}$ columns (if $s > 0$). However, it follows from (2.1) and (2.2) that only $(1,1)^T$ and $(-1,-1)^T$ can appear in the last $N - s2^{t+1}$ columns. This contradicts the assumption that $\mathbf{X}$ is an orthogonal array. □

Since N must be a multiple of 2^{t+1} if the projection onto some $t+1$ factors is of type II, the following theorem follows immediately from Lemma 2.2:

THEOREM 2.3. *Let $\mathbf{X}$ be an* OA$(N, 2^k, t)$ *with $k \geq t+2$. If there are $t+1$ factors onto which the projection of $\mathbf{X}$ is of type I, then N must be a multiple of 2^{t+1}.*

From Theorems 2.1 and 2.3, we have the following corollary.

COROLLARY 2.4. *If N is not a multiple of 2^{t+1}, then the projection of an* OA$(N, 2^k, t)$ *with $k \geq t+2$ onto any $t+1$ factors must be of type III, and therefore contains at least one copy of the complete 2^{t+1} factorial.*

Corollary 2.4 has important statistical implications. Lemma 2.2 is also interesting in its own right. We expect it to be useful for studying projections onto more than $t+1$ factors.

3. Projections onto $t+2$ factors. Results on projections onto $t+2$ factors can be obtained by considering foldovers. For convenience, two $(1,-1)$-vectors which can be obtained from each other by interchanging 1 and -1 are called *mirror images*. We shall denote the mirror image of a vector $\mathbf{x}$ by $\bar{\mathbf{x}}$.

Suppose $\mathbf{X}$ is an OA$(N, 2^k, t)$, in which t is even. Since its foldover $\tilde{\mathbf{X}}$ has strength $t+1$, the projection of $\tilde{\mathbf{X}}$ onto any $t+2$ rows must be one of the three types given in the last section. The following can easily be obtained by applying the results in Section 2 to $\tilde{\mathbf{X}}$.

COROLLARY 3.1. *Suppose $\mathbf{X}$ is an* OA$(N, 2^k, t)$ *in which t is even and $k \geq t+2$. Let $\mathbf{Y}$ be a $(t+2) \times N$ submatrix of $\mathbf{X}$. For each $\mathbf{x} = (x_1, x_2, \ldots, x_{t+2})^T$ with $x_i = 1$ or -1, let $f(\mathbf{x})$ be the number of times $\mathbf{x}$*

appears as a column vector of $\mathbf{Y}$. *Then there exist two nonnegative integers* α *and* β *such that* $f(\mathbf{x}) + f(\bar{\mathbf{x}}) = \alpha$ *for all* $\mathbf{x}$ *such that* $x_1 x_2 \cdots x_{t+2} = 1$, *and* $f(\mathbf{x}) + f(\bar{\mathbf{x}}) = \beta$ *for all* $\mathbf{x}$ *such that* $x_1 x_2 \cdots x_{t+2} = -1$.

COROLLARY 3.2. *Suppose* t *is even*, N *is not a multiple of* 2^{t+1} *and* $k \geq t + 2$. *Then for any* $\mathbf{x} = (x_1, x_2, \dots, x_{t+2})^T$ *with* $x_i = 1$ *or* -1, *at least one of* $\mathbf{x}$ *and its mirror image must appear in the projection of an* $\mathrm{OA}(N, 2^k, t)$ *onto any* $t + 2$ *factors.*

Lin and Draper (1992) found that except for permutations of columns and/or sign changes, there is only one possible projection of a 12-run Plackett–Burman design onto any four factors. As an example to demonstrate how Corollaries 3.1 and 3.2 can be applied, we shall derive this result analytically. In fact, the same result holds for all $\mathrm{OA}(12, 2^k, 2)$ with $4 \leq k \leq 11$, not just for the Plackett–Burman design. What happens is that there is only one $\mathrm{OA}(12, 2^4, 2)$.

Let $\mathbf{X}$ be an $\mathrm{OA}(12, 2^4, 2)$. We first observe two simple facts which follow from Theorems 2.1 and 2.3 and Corollaries 3.1 and 3.2.

FACT 1. In any three rows of $\mathbf{X}$, either each $\mathbf{x} = (x_1, x_2, x_3)^T$ with $x_1 x_2 x_3 = 1$ appears twice as a column vector and each of those with $x_1 x_2 x_3 = -1$ appears exactly once, or each $\mathbf{x}$ with $x_1 x_2 x_3 = 1$ appears once and each of those with $x_1 x_2 x_3 = -1$ appears exactly twice.

FACT 2. For each $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$, let $f(\mathbf{x})$ be the number of times $\mathbf{x}$ appears as a column vector of $\mathbf{X}$. Then either $f(\mathbf{x}) + f(\bar{\mathbf{x}}) = 2$ for all $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = 1$ and $f(\mathbf{x}) + f(\bar{\mathbf{x}}) = 1$ for all $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = -1$, or $f(\mathbf{x}) + f(\bar{\mathbf{x}}) = 1$ for all $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = 1$ and $f(\mathbf{x}) + f(\bar{\mathbf{x}}) = 2$ for all $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = -1$.

Now we shall use these two facts to prove the following theorem.

THEOREM 3.3. *Let* $\mathbf{X}$ *be an* $\mathrm{OA}(12, 2^4, 2)$. *Then* $\mathbf{X}$ *must have exactly two identical columns. Furthermore, all* $\mathrm{OA}(12, 2^4, 2)$ *can be obtained from one another by permuting columns and/or changing the signs of all the entries in the same row.*

PROOF. From Fact 1, by appropriately changing signs and/or permuting columns, we may assume that the first three rows of $\mathbf{X}$ are

$$\begin{matrix} + & + & + & + & - & - & - & - & + & + & - & - \\ + & + & - & - & + & + & - & - & + & - & + & - \\ + & + & - & - & - & - & + & + & - & + & + & - \end{matrix},$$

where $+$ and $-$ represent 1 and -1, respectively. We first prove that $\mathbf{X}$ must have at least two identical columns. If not, then all the columns of $\mathbf{X}$ are distinct and $\mathbf{X}$ can be written as

$$\begin{array}{cccccccccccc}
+ & + & + & + & - & - & - & - & + & + & - & - \\
+ & + & - & - & + & + & - & - & + & - & + & - \\
+ & + & - & - & - & - & + & + & - & + & + & - \\
+ & - & + & - & + & - & + & - & \ast & \ast & \ast & \ast
\end{array},$$

where $\ast$ is to be determined. Without loss of generality, suppose the first of the four $\ast$'s is $+$. Applying Fact 1 to the last three rows of $\mathbf{X}$, we conclude that $\mathbf{X}$ must be

$$\begin{array}{cccccccccccc}
+ & + & + & + & - & - & - & - & + & + & - & - \\
+ & + & - & - & + & + & - & - & + & - & + & - \\
+ & + & - & - & - & - & + & + & - & + & + & - \\
+ & - & + & - & + & - & + & - & + & + & - & -
\end{array}.$$

However, then $+\,+$ would appear four times in the first and last rows of $\mathbf{X}$. This is not possible because $\mathbf{X}$ is an orthogonal array with strength 2. Therefore, $\mathbf{X}$ has at least two identical columns. Without loss of generality, suppose they are $(1,1,1,1)^T$. Then $\mathbf{X}$ is

$$\begin{array}{cccccccccccc}
+ & + & + & + & - & - & - & - & + & + & - & - \\
+ & + & - & - & + & + & - & - & + & - & + & - \\
+ & + & - & - & - & - & + & + & - & + & + & - \\
+ & + & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast
\end{array}.$$

By Fact 2, the last column must be $(-1,-1,-1,1)^T$; otherwise $(1,1,1,1)^T$ together with its mirror image would appear three times. Furthermore, $f(\mathbf{x}) + f(\bar{\mathbf{x}}) = 2$ for all $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = 1$ and $f(\mathbf{x}) + f(\bar{\mathbf{x}}) = 1$ for all $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = -1$. Then we apply Fact 1 to the last three rows to show that the third and fourth columns of $\mathbf{X}$ must be $(1,-1,-1,1)^T$ and $(1,-1,-1,-1)^T$. Considering $(1,-1,-1,1)^T$ and its mirror image, we see that the next to last column of $\mathbf{X}$ must be $(-1,1,1,-1)^T$. Continuing this kind of argument, we conclude that $\mathbf{X}$ must be

$$\begin{array}{cccccccccccc}
+ & + & + & + & - & - & - & - & + & + & - & - \\
+ & + & - & - & + & + & - & - & + & - & + & - \\
+ & + & - & - & - & - & + & + & - & + & + & - \\
+ & + & + & - & + & - & + & - & - & - & - & +
\end{array}. \tag{3.1}$$

Therefore, array (3.1) is the unique $\mathrm{OA}(12, 2^4, 2)$ (up to possible permutations of columns and/or changes of signs). This completes the proof. □

The following corollary is an immediate consequence of Theorem 3.3:

COROLLARY 3.4. *Let* $\mathbf{X}$ *be an* $\mathrm{OA}(12, 2^k, 2)$ *with* $4 \le k \le 11$. *Then any* 4×12 *submatrix of* $\mathbf{X}$ *can be obtained from* (3.1) *by permuting columns and/or changing the signs of all the entries in the same row.*

4. Projections of an OA($N, 2^k, 2$) onto 4 factors. Corollary 3.4 shows that the projection of an OA($12, 2^k, 2$) with $4 \le k \le 11$ onto any four factors is unique up to permutations of columns and/or changes of signs. The computer work of Lin and Draper (1992) shows that projections of larger Plackett–Burman designs are more complex. For instance, uniqueness no longer holds. Although more laborious, one can still work out the possible projections analytically. However, we shall not pursue it here.

Notice that array (3.1) contains 11 distinct columns. It can easily be verified that these 11 columns constitute a design under which all the main effects and two-factor interactions can be estimated when the higher-order interactions are negligible. Therefore, as observed by Lin and Draper (1993) and Wang and Wu (1995), all the projections of the 12-run Plackett–Burman design onto four factors have this important statistical property. The main purpose of this section is to prove that it holds generally for all OA($N, 2^k, 2$) with $k \ge 4$ as long as N is not a multiple of 8. By contrast, unless a regular fractional factorial design already has resolution at least 5, not all its projections onto four factors enjoy this property.

For any orthogonal array $\mathbf{X}$ with k factors and N runs, we say that a defining relation exists among factors $i_1, \dots, i_s$, where $s \le k$, if the corresponding rows $\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_s}$ of $\mathbf{X}$ satisfy $\mathbf{x}_{i_1} \odot \cdots \odot \mathbf{x}_{i_s} = \mathbf{1}$ or $-\mathbf{1}$, where $\mathbf{1}$ is the $1 \times N$ vector of 1's. In this case, the main effect of any of these s factors is totally aliased with the interaction of the other $s - 1$ factors, and the interaction of any subset of the s factors is totally aliased with the interaction of the remaining factors. It is clear that if a defining relation exists among a subset of four or fewer factors, then not all the two-factor interactions can be estimated. For regular fractional factorial designs, the converse is also true (although it does not hold in general). This is the well known rule of using the length of the shortest word in the defining relation to determine the resolution of a regular fractional factorial design. Therefore, in a regular fractional factorial design, if no defining relation exists among any four or fewer factors, then all the main effects and two-factor interactions are estimable. We shall first show that this is true for *all* OA($N, 2^k, 2$) with $k = 4$, even though in general, effects which are not totally aliased are not necessarily orthogonal.

THEOREM 4.1. *Let* $\mathbf{X}$ *be an* OA($N, 2^4, 2$). *Then all the main effects and two-factor interactions are estimable under the assumption that the higher-order interactions are negligible if and only if no defining relation exists among all the four factors as well as any three of them.*

PROOF. It is sufficient to prove the "if" part. Since no defining relation exists among all the four factors and any three of them, similar to the two facts mentioned in Section 3, we have the following statements:

(i) In any three rows of $\mathbf{X}$, each $\mathbf{x} = (x_1, x_2, x_3)^T$ with $x_i = 1$ or -1 appears at least once.

(ii) For each $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$ with $x_i = 1$ or -1, at least one of $\mathbf{x}$ and its mirror image appears as a column vector of $\mathbf{X}$.

Now consider the 16 4×1 vectors of 1's and -1's. If all of them appear at least once as column vectors of $\mathbf{X}$, then there is nothing to prove. So it is enough to consider the case in which at least 1 of these 16 vectors does not appear. Without loss of generality (by changing signs if necessary), we may assume that at least one vector with $x_1 x_2 x_3 x_4 = 1$ is absent. By (ii), at most four such vectors can be absent, and no two of them can be mirror images. We shall divide the proof into four cases:

Case 1. There is exactly one $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = 1$ which does not appear as a column vector of $\mathbf{X}$. We may assume that this vector is $(-1, -1, -1, -1)^T$. Then by (i), all the four vectors $(1, -1, -1, -1,)^T$, $(-1, 1, -1, -1)^T$, $(-1, -1, 1, -1)^T$ and $(-1, -1, -1, 1)^T$ appear as column vectors of $\mathbf{X}$. Together with the seven vectors other than $(-1, -1, -1, -1)^T$ which satisfy $x_1 x_2 x_3 x_4 = 1$. they give 11 distinct columns of $\mathbf{X}$. It is easy to verify that these 11 columns define a design which allows the estimation of all the main effects and two-factor interactions: let $\mathbf{D}$ be the 4×11 matrix whose column vectors are the 11 vectors that we have just shown to be column vectors of $\mathbf{X}$, and let $\mathbf{F}$ be the 11×11 matrix whose row vectors are the four row vectors of $\mathbf{D}$ (corresponding to main effects), the 1×11 vector of ones (corresponding to the general mean) and the $\binom{4}{2} = 6$ Hadamard products of pairs of row vectors of $\mathbf{D}$ (corresponding to the two-factor interactions). Then we show that $\mathbf{FF}'$ is nonsingular. This can be verified, for example, by a computer.

Case 2. There are exactly two $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = 1$ which do not appear as column vectors of $\mathbf{X}$. We may assume that they are $(-1, -1, -1, -1)^T$ and $(-1, -1, 1, 1)^T$. Then by (i), all the six vectors $(1, -1, -1, -1)^T$, $(-1, 1, -1, -1)^T$, $(-1, -1, 1, -1)^T$, $(-1, -1, -1, 1)^T$, $(1, -1, 1, 1)^T$ and $(-1, 1, 1, 1)^T$ appear as column vectors of $\mathbf{X}$. Together with the six vectors other than $(-1, -1, -1, -1)^T$ and $(-1, -1, 1, 1)^T$ which satisfy $x_1 x_2 x_3 x_4 = 1$, they give 12 distinct columns of $\mathbf{X}$. Again, it can be verified that these 12 columns define a design with the desired property.

Case 3. There are exactly three $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = 1$ which do not appear as column vectors of $\mathbf{X}$. We may assume that these three vectors are $(-1, -1, -1, -1)^T$, $(-1, -1, 1, 1)^T$, and $(-1, 1, -1, 1)^T$. Then the same argument shows that $\mathbf{X}$ has at least 12 distinct columns: $(1, -1, -1, -1)^T$, $(-1, 1, -1, -1)^T$, $(-1, -1, 1, -1)^T$, $(-1, -1, -1, 1)^T$, $(1, -1, 1, 1)^T$, $(1, 1, -1, 1)^T$, $(-1, 1, 1, 1)^T$, and the five vectors other than $(-1, -1, -1, -1)^T$, $(-1, -1, 1, 1)^T$ and $(-1, 1, -1, 1)^T$ which satisfy $x_1 x_2 x_3 x_4 = 1$. This also gives a design under which all the main effects and two-factor interactions are estimable.

Case 4. There are exactly four $\mathbf{x}$ with $x_1 x_2 x_3 x_4 = 1$ which do not appear as column vectors of $\mathbf{X}$. We may assume that these four vectors are $(-1, -1, -1, -1)^T$, $(-1, -1, 1, 1)^T$, $(-1, 1, -1, 1)^T$, $(-1, 1, 1, -1)^T$, or $(-1, -1, -1, -1)^T$, $(-1, -1, 1, 1)^T$, $(-1, 1, -1, 1)^T$, $(1, -1, -1, 1)^T$. Proceeding in the same way, we can handle these two subcases separately. In the first subcase, all the eight $\mathbf{x}$ with $x_1x_2x_3x_4 = -1$ must appear, giving a total of at least 12 distinct columns in $\mathbf{X}$, while in the second subcase, all the $\mathbf{x}$ with $x_1x_2x_3x_4 = -1$, except possibly $(1,1,1,-1)^T$ must appear, giving a total of at least 11 distinct columns. □

For a resolution 3 or 4 regular fractional factorial design, there are always certain sets of three or four factors such that a defining relation exists among these factors. On the other hand, suppose N is not a multiple of 8 and $\mathbf{X}$ is an $\mathrm{OA}(N, 2^k, 2)$ with $k \geq 4$. Then by the results in Sections 2 and 3, the projection of $\mathbf{X}$ onto any three factors and the projection of $\tilde{\mathbf{X}}$ onto any four factors are of type III. It follows that defining relations cannot exist among *any* three or four factors of $\mathbf{X}$. Therefore we have the following theorem.

THEOREM 4.2. *Suppose N is not a multiple of 8. Let $\mathbf{X}$ be an $\mathrm{OA}(N, 2^k, 2)$ with $k \geq 4$. Then the projection of $\mathbf{X}$ onto any four factors has the property that all the main effects and two-factor interactions of these four factors are estimable when the higher-order interactions are negligible.*

In a regular fractional factorial design, any two factorial effects (main effects or interactions) are either totally aliased or orthogonal. As mentioned in Section 2, such a simple alias structure is one reason why these designs are popular. It is interesting to note that the important projective property obtained in Theorem 4.2 can be attributed to the complex alias structures of the nonregular designs.

We conclude this article with the following theorem.

THEOREM 4.3. *Let $\mathbf{X}$ be an $\mathrm{OA}(N, 2^k, 2)$ with $k \geq 4$. Suppose the projection of $\mathbf{X}$ onto certain three factors is of type III. Then the projection onto these three and any other factor has the property that all the main effects and two-factor interactions are estimable when the higher-order interactions are negligible.*

PROOF. Without loss of generality, assume that the projection of $\mathbf{X}$ onto the first three factors is of type III. We shall show that its projection onto the first four factors has the property described in this theorem. By Lemma 2.2, the projection onto any three of the first four factors cannot be of type I. Therefore, there is no defining relation among any three of the first four factors. By Theorem 4.1, it remains to show that there is also no defining relation among all four factors. Let $\mathbf{Y}$ be the submatrix of $\mathbf{X}$ consisting of the first four rows. Suppose there is a defining relation among these four factors. We shall show that this would lead to a contradiction. We may assume that all column vectors $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$ of $\mathbf{Y}$ satisfy

$$x_4 = x_1x_2x_3, \tag{4.1}$$

and there is a postive integer s such that in the first $8s$ columns of $\mathbf{Y}$, each of the eight 3×1 vectors of 1's and -1's appears s times in the first three rows, and each of the last $N - 8s$ columns of $\mathbf{Y}$ satisfies

$$x_3 = x_1 x_2. \tag{4.2}$$

From (4.1) and (4.2), we have $x_4 = (x_1 x_2)(x_1 x_2) = 1$ for each of the last $N - 8s$ columns of $\mathbf{Y}$, that is, all the last $N - 8s$ entries of the fourth row of $\mathbf{Y}$ are equal to 1. On the other hand, 1 and -1 are equally represented in the first $8s$ entries of the fourth row of $\mathbf{Y}$. This contradicts the fact that $\mathbf{Y}$ is an orthogonal array. □

REFERENCES

BOX, G. E. P. and BISGAARD, S. (1993). What can you find out from 12 experimental runs? *Quality Engineering* **5** 663–668.

BOX, G. E. P. and HUNTER, J. S. (1961). The 2^{k-p} fractional factorial designs. *Technometrics* **3** 311–351; 449–458.

HAMADA, M. and WU, C. F. J. (1992). Analysis of designed experiments with complex aliasing. *Journal of Quality Technology* **24** 130–137.

LIN, D. K. J. and DRAPER, N. R. (1991). Projection properties of Plackett and Burman designs. Technical Report 885, Dept. Statistics, Univ. Wisconsin.

LIN, D. K. J. and DRAPER, N. R. (1992). Projection properties of Plackett and Burman designs. *Technometrics* **34** 423–428.

LIN, D. K. J. and DRAPER, N. R. (1993). Generating alias relationships for two-level Plackett and Burman designs. *Comput. Statist. Data Anal.* **15** 147–157.

PLACKETT, R. L. and BURMAN, J. P. (1946). The design of optimum multi-factorial experiments. *Biometrika* **33** 305–325.

RAO, C. R. (1946). Hypercubes of strength d leading to confounded designs in factorial experiments. *Bull. Calcutta Math. Soc.* **38** 67–78.

RAO, C. R. (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *J. Roy. Statist. Soc. Suppl.* **9** 128–139.

SEIDEN, E. and ZEMACH, R. (1966). On orthogonal arrays. *Ann. Math. Statist.* **37** 1355–1370.

WANG, J. C. and WU, C. F. J. (1995). A hidden projection property of Plackett–Burman and related designs. *Statist. Sinica* **5** 235–250.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720