

# SECOND ORDER SATURATED ORTHOGONAL ARRAYS OF STRENGTH THREE

Ching-Shui Cheng<sup>1</sup>, Robert W. Mee<sup>2</sup>, and Oksoun Yee<sup>3</sup>

<sup>1</sup>*Academia Sinica and University of California, Berkeley*, <sup>2</sup>*University of Tennessee*, and  
<sup>3</sup>*Schering-Plough Corporation*

*Abstract:* Strength three two-level orthogonal arrays with the number of runs ( $N$ ) equaling twice the number of factors ( $k$ ) are second-order saturated (SOS) designs. That is, for such designs one can construct a saturated model with an intercept,  $k$  ( $= N/2$ ) main effects, and  $N/2 - 1$  two-factor interactions. Projections of this design onto subsets of factors provide no more degrees of freedom for two-factor interactions. This article explores the construction of other second-order saturated strength three arrays which allocate more than  $N/2$  degrees of freedom for two-factor interactions. These new orthogonal arrays are constructed using two methods, one based on a foldover technique that reverses the signs of a subset of the columns of the strength three orthogonal array with  $k = N/2$ , and the second based on the Kronecker product of an SOS design and a Hadamard matrix. We compare these new designs with respect to their generalized word length and alias length patterns.

*Key words and phrases:* Alias length pattern, complex aliasing, confounding frequency vector, doubling, foldover, nonregular design, resolution IV, two-factor interaction, word length pattern.

01/31/2006 – Revised for *Statistica Sinica*

## 1. Introduction

It is well known that strength three two-level orthogonal arrays exist for  $k$  factors in  $2k$  runs for  $k$  any multiple of 4 - assuming the existence of Hadamard matrices of all orders that are multiples of 4. We denote these arrays as  $OA(N, 2^k, 3)$  with  $N = 2k$ .

These orthogonal arrays, whether the regular fractional factorials such as  $2_{III}^{4-1}$  and  $2_{III}^{8-4}$  or the nonregular designs such as discussed by Miller and Sitter (2001) are all obtained by folding over a  $k \times k$  Hadamard matrix  $\mathbf{H}_k$  (Hedayat, Sloane, and Stufken 1999, p. 148).

That is, these  $k$ -factor designs are given by

$$\begin{bmatrix} \mathbf{H}_k \\ -\mathbf{H}_k \end{bmatrix}. \quad (1)$$

If the objective is to maximize the number of factors in a strength three orthogonal array with  $N$  runs, the foldover design (1) is optimal. For this reason, these designs are known as minimal strength three designs (Margolin 1969) since they minimize the number of runs for strength three orthogonal arrays.

For a two-level design  $\mathbf{D}$  written as an  $N \times k$  matrix of 1's and -1's, define  $\mathbf{X}_1$  to be the first-order model matrix where a column of 1's is appended to  $\mathbf{D}$ , and let  $\mathbf{X}_2$  denote the  $N \times [k(k-1)/2]$  matrix of two-factor interaction contrasts, i.e.,  $\mathbf{X}_2$  consists of all columns of the form  $\mathbf{a} \odot \mathbf{b}$ , where  $\mathbf{a}$  and  $\mathbf{b}$  are columns of  $\mathbf{D}$  and  $\mathbf{a} \odot \mathbf{b}$  is their component-wise product. The design  $\mathbf{D}$  is "second-order saturated" (SOS) if  $\text{rank}[\mathbf{X}_1 \mathbf{X}_2] = N$  (see Block and Mee 2003). For strength-three orthogonal arrays,  $\text{rank}[\mathbf{X}_1 \mathbf{X}_2] = 1 + k + \text{rank}[\mathbf{X}_2]$ , since  $\mathbf{X}_1' \mathbf{X}_2 = \mathbf{0}$ . Thus a strength three  $\mathbf{D}$  is SOS if and only if  $\text{rank}[\mathbf{X}_2] = N - k - 1$ . This is the case for all regular and nonregular designs mentioned above with  $N = 2k$ .

Regular foldover designs of the form (1) are called “even” designs since all the words of the defining contrast subgroup have even length. The concept of even designs can be extended to the nonregular case. Suppose  $(A_1^s, \dots, A_k^s)$  is the generalized wordlength pattern (gwlp) of a design as defined in Xu and Wu (2001) or Ma and Fang (2001). Then the design is said to be even if  $A_i^s = 0$  for all odd  $i$ . We show in Appendix 1 that a two-level design is even if and only if it is a “foldover design” (or “mirror-image pair design”), and for all even designs,  $\text{rank}[\mathbf{X}_2] \leq N/2 - 1$ , i.e., they allocate fewer than half of the available degrees of freedom for two-factor interactions.

For  $N \geq 16$ , there exist SOS resolution IV designs with  $\text{rank}[\mathbf{X}_2] > N/2$ . These SOS designs provide more information for two-factor interactions than foldover designs of the same size and number of factors. In Section 2, we summarize results for regular resolution IV SOS designs and discuss two constructions of SOS designs. Then in Section 3, we apply the same methods of construction to strength three orthogonal arrays. New strength three orthogonal arrays are obtained, many of which are SOS. The method of foldover is by far the most popular method of constructing (even) orthogonal arrays of strength three. One contribution of the present paper is the construction of strength three two-level orthogonal arrays that are not even.

For regular fraction SOS designs, there necessarily exist an orthogonal set of  $k$  main effect and  $N - 1 - k$  two-factor interactions contrasts. For non-regular SOS designs, a set of such orthogonal contrasts is not guaranteed. In Section 3, we identify certain situations where such orthogonal contrasts can be found.

In addition to word length pattern, resolution IV designs are usefully characterized by their “alias length pattern” (Block and Mee 2003), which summarizes

the lengths of the chains of aliased two-factor interactions; also see Cheng, Steinberg and Sun (1999). For orthogonal arrays with partial aliasing, the generalized word length pattern (Ma and Fang 2001; Xu and Wu 2001) is used to summarize the aberration of the design. However, for  $OA(N, 2^k, 3)$  a more direct characterization of the aliasing among two-factor interactions is needed. We propose such a measure of aliasing for nonregular designs of strength three in Section 4. This measure, which we refer to as generalized alias length pattern, reflects how different the extent of aliasing is among the two-factor interactions. The final section discusses briefly the analysis of data obtained using these designs, as well as the need for additional work comparing these new SOS designs and their projections.

## 2. Regular Resolution IV SOS Designs

The smallest examples of SOS designs of resolution IV (or higher) for which  $\text{rank}[\mathbf{X}_2] > N/2$  are:

- $2_{IV}^{5-1}$ , with  $\text{rank}[\mathbf{X}_2] = 10$
- the minimum aberration  $2_{IV}^{10-5}$  design, with  $\text{rank}[\mathbf{X}_2] = 21$
- The  $2_{IV}^{9-4}$  design with generators  $6 = 1 \cdot 2 \cdot 3$ ,  $7 = 1 \cdot 2 \cdot 4$ ,  $8 = 1 \cdot 3 \cdot 4$ ,  $9 = 2 \cdot 3 \cdot 4 \cdot 5$ , where the symbols 1-9 denote the nine factors. This design has  $\text{rank}[\mathbf{X}_2] = 22$ , but it has seven length-four words, one more than the minimum aberration design.

As mentioned above, regular  $2_{IV}^{k-p}$  SOS designs with  $k = N/2$  are even designs consisting of  $N/2$  mirror image pairs of runs. Regular  $2_{IV}^{k-p}$  SOS designs with  $k < N/2$  are even/odd designs, meaning the defining contrast subgroup consists of  $2^p$  words, half of which are even and half odd in length. Even/odd designs do not contain any mirror image pairs.

Chen and Cheng (2006) pointed out that SOS designs of resolution IV or higher are mathematically equivalent to maximal caps in a finite projective geometry, and reviewed some results on the construction of maximal caps (Davydov and Tombak, 1990; Bruen, Haddad and Wehlau, 1998; Bruen and Wehlau, 1999). We mention two important types of regular SOS designs:

- Type 1: SOS designs obtained by partial foldover, with  $k = N/4 + 1$
- Type 2: SOS designs obtained by doubling (to be defined later)

The SOS designs already listed represent both types. We state the following results from the above-mentioned papers without proof.

### 2.1. Construction of Type 1 resolution IV designs by partial foldover: $k = N/4 + 1$

Let  $\mathbf{D}$  denote a regular resolution IV design with  $N/2$  runs and  $N/4$  factors.

Partition  $\mathbf{D}$  so that  $\mathbf{D} = [\mathbf{B} \ \mathbf{C}]$ , and create the  $N$ -run design

$$\mathbf{S}_1 = \begin{bmatrix} \mathbf{1} & \mathbf{B} & \mathbf{C} \\ -\mathbf{1} & -\mathbf{B} & \mathbf{C} \end{bmatrix}; \quad (2)$$

that is,  $\mathbf{S}_1$  is obtained by foldover of  $\mathbf{D}$ , reversing the columns of  $\mathbf{B}$ , but not those of  $\mathbf{C}$ , and appending an initial column to represent one more factor. (Note that the usual foldover method reverses all columns.) Then  $\mathbf{S}_1$  is of resolution IV or higher. Such designs contain at least  $N/4$  clear two-factor interactions. If  $\mathbf{B}$  contains an odd number of columns, then  $\mathbf{S}_1$  is SOS. Otherwise, it may or may not be SOS. We now provide three examples:

- Let  $\mathbf{D}$  denote the  $2_{IV}^{4-1}$  fraction and let  $\mathbf{B}$  be any single column (or any three of the four columns). Then  $\mathbf{S}_1$  is the  $2_V^{5-1}$  design.

- Let  $\mathbf{D}$  denote the  $2_{IV}^{8-4}$  fraction and let  $\mathbf{B}$  be any single column of  $\mathbf{D}$ . Then  $\mathcal{S}_1$  is the  $2_{IV}^{9-4}$  SOS design mentioned earlier.
- Let  $\mathbf{D}$  denote the  $2_{IV}^{16-11}$  fraction. Depending on which columns are placed in  $\mathbf{B}$ , one may construct any one of the five non-isomorphic SOS  $2_{IV}^{17-11}$  designs. For details, see the beginning of Appendix 2.

## 2.2. Construction of Type 2 SOS designs: doubled designs

Suppose  $\mathcal{S}$  is a resolution four-or-higher SOS design with  $N/2$  runs and  $k/2$  factors. Then the design

$$\mathcal{S}_2 = \begin{bmatrix} \mathcal{S} & \mathcal{S} \\ \mathcal{S} & -\mathcal{S} \end{bmatrix}$$

is a resolution IV SOS design with  $k$  factors in  $N$  runs. For instance, if  $\mathcal{S}$  is the  $2_V^{5-1}$  design, then  $\mathcal{S}_2$  is the minimum aberration  $2_{IV}^{10-5}$  SOS design mentioned earlier; if  $\mathcal{S}$  is the  $2_{IV}^{9-4}$  SOS design, then  $\mathcal{S}_2$  is the unique  $2_{IV}^{18-12}$  SOS design. Type 2 SOS designs have no clear two-factor interactions.

## 2.3. Additional SOS designs with $k \leq N/4 + 1$

There exist regular SOS designs that cannot be constructed by partial foldover or doubling. The smallest example is the  $2_{IV}^{13-7}$  design with 36 clear two-factor interactions (see Block and Mee 2003). For  $N = 128$ , SOS designs not of Type 1 or 2 exist for  $k = 21, 22, 24-29, 31$  and  $33$  (see Block 2003). Most but not all of these have clear two-factor interactions.

### 3. SOS Strength Three Orthogonal Arrays

In this section we construct  $OA(N, 2^k, 3)$  that do not include any mirror image pairs, and yet are SOS. We construct these orthogonal arrays using two methods – by partial foldover of SOS designs, and by a generalization of doubling. For regular SOS designs, there exist  $N-1$  main effects and two-factor interactions that are pairwise orthogonal. This is not guaranteed for SOS designs with partial aliasing. Therefore, we find it helpful to distinguish the following two cases:

- If for an  $OA(N, 2^k, 3)$ ,  $\text{rank}[\mathbf{X}_2] = N-k-1$ , then the array is second-order saturated.
- If in addition, there exist a set of  $N-k-1$  two-factor interactions that are pairwise orthogonal, then the array is orthogonal second-order saturated (OSOS).

#### 3.1. Type 1 SOS $OA(N, 2^k, 3)$ with $k = N/4 + 1$

Let  $\mathbf{H}_m$  denote a Hadamard matrix of order  $m = N/4$  and let  $\mathbf{D}$  denote the even  $OA(2m, 2^m, 3)$  obtained by foldover. Partition the design into two sets of columns, which

we denote as  $\mathbf{D} = [\mathbf{B} \ \mathbf{C}]$ , and construct  $\mathbf{S}_1 = \begin{bmatrix} \mathbf{I} & \mathbf{B} & \mathbf{C} \\ -\mathbf{I} & -\mathbf{B} & \mathbf{C} \end{bmatrix}$ . This extends to nonregular

designs a construction described in Section 2.1, which is translated from the geometric literature into design language. Usually the geometric approach only works for regular designs, but once the construction is expressed in the form of (2), it can readily be extended to nonregular designs, and produces many new orthogonal arrays of strength three. Tang (2005), in a study of clear two-factor interactions under nonregular designs,

presented a construction which is a special case of our method with  $\mathbf{B}$  consisting of one single column.

Lemma 1.  $\mathcal{S}_1$  is an orthogonal array of strength three.

This lemma is quickly verified by noting that since  $\mathbf{D}$  is strength three, so is  $[-\mathbf{B} \ \mathbf{C}]$ , and hence so is  $\begin{bmatrix} \mathbf{B} & \mathbf{C} \\ -\mathbf{B} & \mathbf{C} \end{bmatrix}$ . Projecting  $\begin{bmatrix} \mathbf{B} & \mathbf{C} \\ -\mathbf{B} & \mathbf{C} \end{bmatrix}$  into two or three columns, and appending the final column also results in an equally replicated full factorial.

In the regular case,  $\mathcal{S}_1$  is SOS as long as  $\mathbf{B}$  contains an odd number of columns. We believe that this is also true for nonregular designs. The proof for the regular case utilizes the fact that a regular SOS design of resolution four or higher is equivalent to a maximal cap. There is no such equivalence result for nonregular designs and so the proof does not carry over. But we do have the following partial result:

Theorem 1: If  $\mathbf{B}$  contains a single column of  $\mathbf{D}$ , then  $\mathcal{S}_1$  is SOS.

Proof: Since  $\mathcal{S}_1$  is strength three, it may be verified that  $\mathcal{S}_1$  is SOS by showing that  $\text{rank}[\mathbf{X}_2] = N - k - 1 = 3k - 5$ . It is easily verified that all  $2k-3$  two-factor interactions involving at least one of the first two columns of  $\mathcal{S}_1$  are orthogonal to one another and the remaining  $(k-2)(k-3)/2$  interactions of pairs of columns of  $\begin{bmatrix} \mathbf{C} \\ \mathbf{C} \end{bmatrix}$ . Thus, that  $\mathcal{S}_1$  is SOS

follows from observing that the matrix of  $(k-2)(k-3)/2$  aliased two-factor interactions among the columns of  $\mathbf{C}$  has rank  $k-2$ . This can be proved as follows. Without loss of

generality, we may assume that  $\mathbf{B} = \begin{bmatrix} \mathbf{I}_m \\ -\mathbf{I}_m \end{bmatrix}$ , where  $\mathbf{I}_m$  is the  $m \times 1$  vector of 1's. Then

$\mathbf{C} = \begin{bmatrix} \mathbf{Y} \\ -\mathbf{Y} \end{bmatrix}$ , where  $\mathbf{Y}$  is a saturated strength two array of size  $m = k - 1$ . It suffices to show

that the matrix of two-factor interactions among the columns of  $Y$  has rank  $k-2$ . Let  $y_1, \dots, y_{k-2}$  be the columns of  $Y$  and, for  $i = 2, \dots, k-2$ , let  $e_i$  be the two-factor interaction contrast  $y_1 \odot y_i$ . Then  $y_1, e_2, \dots, e_{k-2}$  are orthogonal to one another and thus form an orthogonal basis of the linear space of contrasts. Let  $f = y_2 \odot y_3$ . Then  $f$  can be expressed as a linear combination of  $y_1, e_2, \dots, e_{k-2}$ , say  $f = \lambda y_1 + \sum_{i=2}^{k-2} \mu_i e_i$ . Since  $Y$  is a saturated orthogonal array of strength two, without loss of generality, we may assume that the subarray  $[y_1 \ y_2 \ y_3]$  is not of strength three. Then the sum of the components of  $y_1 \odot y_2 \odot y_3$  is not equal to zero. Therefore  $\lambda \neq 0$ . It follows that the  $k-2$  columns  $f, e_2, \dots, e_{k-2}$ , all of which are two-factor interactions among the columns of  $Y$ , are linearly independent. This proves Theorem 1.

From the above proof, it is clear that  $S_1$  is orthogonal SOS if there exist  $k-2$  two-factor interactions among the columns of  $Y$  that are mutually orthogonal. Now suppose  $Y$  has at least one word of length three, i.e., there are three columns, say  $y_1, y_2$ , and  $y_3$ , such that their component-wise product is the column of all 1's or all -1's. Then it can easily be verified that the  $k-2$  two-factor interaction contrasts  $f, e_2, \dots, e_{k-2}$  in the proof of Theorem 1 are mutually orthogonal. Thus we have proved the following result.

**Theorem 2:** In Theorem 1, if  $D$  is the foldover of a Hadamard matrix that has at least one word of length three, then  $S_1$  is orthogonal SOS.

We now present an example. Construct the 12-run Plackett-Burman design by cycling the row +++---+--+ and then appending the row ++++++. Adding the first column of +1's, we obtain  $H_{12}$ ; folding over this 12x12 matrix we obtain the 24-run

strength three array  $\mathbf{D}$ . We obtained different  $OA(48, 2^{13}, 3)$  designs by different choices for the columns of  $\mathbf{B}$ . Let  $b$  denote the number of columns of  $\mathbf{B}$ . Various  $OA(48, 2^{13}, 3)$  designs and their partial generalized word length patterns are listed below:

$b$	$\text{rank}(\mathbf{X}_2)$	SOS?	$A_4^g$	$A_5^g$
1	34	yes	36.667	18.333
2	33	no	28.333	26.667
3	34	yes	26	29
4	34	yes	26.556	28.444
5	34	yes	27.778	27.222
6	32 or 34	some	28.333	26.667

We need not consider  $b > 6$  since the results would be the same as for reversing  $12-b$  columns. As conjectured, we do get SOS designs for all odd values of  $b$ . Note that  $A_4^g + A_5^g = 55$  for each design, which is the  $A_4^g$  value for the even design  $\mathbf{D}$ . For a given value of  $b$ , we obtained the same  $A_4^g$  and  $A_5^g$  values for all sets of  $b$  columns from  $\mathbf{D}$ . Also, note how  $A_4^g$  varies in a predictable manner as a function of  $b$ . Two designs having the same gwlp does not guarantee that the designs are isomorphic. Obviously, for  $b = 6$ , there are at least two non-isomorphic designs, since some choices of six columns produce an SOS design, and others do not.

The best Type 1 SOS  $OA(N, 2^{N/4+1}, 3)$  designs that may be generated in this manner are summarized in Table 1 for  $N = 48, 64, 80$  and  $96$ . Appendix 3 gives additional details regarding the best choice of columns to reverse. The case  $N = 64$  is interesting in that  $A_4^g = 60$  matches the minimum number of length four words for a

regular  $2^{17-11}$  SOS design. Note however, that of the four nonregular  $OA(16, 2^{15}, 2)$  used to construct these  $OA(64, 2^{17}, 3)$ , two did not achieve the minimum  $A_4^g$  of 60. For the  $OA(96, 2^{25}, 3)$ , it appears that the optimum is achieved at  $b = 8$ ;  $A_4^g = 245.56$  is the best result to date. Finally, since there are Hadamard matrices of order 16 that have words of length three, new nonregular OSOS  $OA(64, 2^{17}, 3)$ 's can be constructed.

### 3.2. Type 2 SOS $OA(N, 2^k, 3)$

As an extension of the doubling construction for regular SOS designs discussed in Chen and Cheng (2006), we present the following general theorem regarding the construction of Type 2 SOS designs.

*Theorem 3:* Suppose  $\mathbf{S}$  is an  $OA(N, 2^k, 3)$  and  $\mathbf{H}_m$  is a Hadamard matrix of order  $m$ . Then  $\mathbf{H}_m \otimes \mathbf{S}$  is an  $OA(mN, 2^{km}, 3)$ . Furthermore, if  $\mathbf{S}$  is SOS, then  $\mathbf{H}_m \otimes \mathbf{S}$  is SOS, and if  $\mathbf{S}$  is OSOS, then  $\mathbf{H}_m \otimes \mathbf{S}$  is OSOS.

*Proof:* Let  $\mathbf{D} = \mathbf{H}_m \otimes \mathbf{S}$ . That  $\mathbf{D}$  is an orthogonal array of strength three can easily be checked and so the details are omitted. Now suppose  $\mathbf{S}$  is SOS. To show that  $\mathbf{D}$  is SOS, we need to show that  $\text{rank}[\mathbf{X}_2] = mN - mk - 1$ , where  $\mathbf{X}_2$  is as defined in the Introduction. Let the  $m$  columns of  $\mathbf{H}_m$  be  $\mathbf{b}_1, \dots, \mathbf{b}_m$ . Fix an  $i \in \{1, \dots, m\}$ , and for each  $j \in \{1, \dots, m\}$  let  $\mathbf{A}_j$  be the  $mN \times [k(k-1)/2]$  matrix consisting of columns of the form  $(\mathbf{b}_i \otimes \mathbf{x}) \odot (\mathbf{b}_j \otimes \mathbf{y})$ , where  $\mathbf{x}$  and  $\mathbf{y}$  are different columns of  $\mathbf{S}$ . Recall that  $(\mathbf{b}_i \otimes \mathbf{x}) \odot (\mathbf{b}_j \otimes \mathbf{y})$  is a two-factor interaction contrast of the array  $\mathbf{D}$ . Furthermore, fix a column  $\mathbf{z}$  of  $\mathbf{S}$  and let  $\mathbf{Z}$  be the  $mN \times (m-1)$  matrix consisting of the two-factor interaction contrasts  $(\mathbf{b}_i \otimes \mathbf{z}) \odot (\mathbf{b}_l \otimes \mathbf{z})$ , where  $l \in \{1, \dots, m\}$  and  $l \neq i$ . Then it can be verified that  $\mathbf{A}_j' \mathbf{Z} = \mathbf{0}$  for all  $j$ ,  $\mathbf{A}_{j_1}' \mathbf{A}_{j_2} = \mathbf{0}$

for all  $j_1 \neq j_2$ , and the  $m - 1$  columns of  $\mathbf{Z}$  are also pairwise orthogonal . Thus

$\text{rank}[\mathbf{X}_2] \geq \text{rank}[\mathbf{Z}] + \sum_{j=1}^m \text{rank}[A_j]$  and  $\text{rank}[\mathbf{Z}] = m - 1$ . Since  $\mathbf{S}$  is SOS,  $\text{rank}[A_j] = N - k - 1$ . Therefore  $\text{rank}[\mathbf{X}_2] \geq m - 1 + m(N - k - 1) = mN - mk - 1$ . It follows that  $\mathbf{X}$  is SOS.

Now suppose  $\mathbf{S}$  is OSOS. Then there exist a set of  $N - k - 1$  two-factor interactions of columns of  $\mathbf{S}$  that are pairwise orthogonal. For each of these interaction contrasts, say  $\mathbf{x} \odot \mathbf{y}$ , where  $\mathbf{x}$  and  $\mathbf{y}$  are columns of  $\mathbf{S}$ , let  $(\mathbf{b}_i \otimes \mathbf{x}) \odot (\mathbf{b}_j \otimes \mathbf{y})$  be the corresponding columns of  $A_j$ . Then these  $N - k - 1$  columns of  $A_j$  are pairwise orthogonal. Collecting the columns so constructed in each of the sets  $A_1, \dots, A_m$  and the  $m - 1$  pairwise orthogonal columns of  $\mathbf{Z}$  mentioned earlier, we have identified a set of pairwise orthogonal two-factor interaction contrasts that make the array  $\mathbf{D}$  OSOS. This completes the proof of Theorem 3.

OA(192,  $2^k$ , 3) Examples:

Suppose  $m = 12$  and  $\mathbf{S}$  is the regular  $2_v^{5-1}$  design. Then the Kronecker product is an OA(192,  $2^{60}$ , 3). SOS designs with the same number of runs can be obtained by doubling a Type 1 SOS OA(96,  $2^{25}$ , 3) or by twice doubling an OA(48,  $2^{13}$ , 3). The resulting designs accommodate 50, 52 or 60 factors. The SOS options with the minimum generalized aberration we have found are summarized in Table 2. These designs and their projections will be considered further in the next section.

#### 4. Generalized Alias Length Pattern

Block and Mee (2003) introduced alias length pattern  $(alp) = (a_1, a_2, \dots, a_L)$  to characterize the aliasing among two factor interactions for regular  $2_{IV}^{k-p}$  designs, where  $L$  denotes the longest chain of aliased two-factor interactions, and  $a_j$  denotes the number of alias chains of length  $j$  ( $j = 1, \dots, L$ ). Note that  $a_1$  denotes the number of clear two-factor interactions. Let  $df_2$  denote the rank of  $\mathbf{X}_2$  and partition  $\mathbf{X}_2$  as  $\mathbf{X}_2 = [\mathbf{X}_{21}, \mathbf{X}_{22}]$ , where  $\mathbf{X}_{21}$  has  $df_2$  columns with full column rank. If the columns of  $\mathbf{X}_{21}$  are orthogonal to one another, the alias matrix involving two-factor interactions for the fitted model defined by  $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_{21}]$  is

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}_{22} = \begin{bmatrix} \frac{1}{N}\mathbf{X}'_1\mathbf{X}_{22} \\ \frac{1}{N}\mathbf{X}'_{21}\mathbf{X}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{A} \end{bmatrix}.$$

Here we use  $\mathbf{A}$  to denote only the last  $df_2$  rows of the alias matrix, since for strength three designs, the other rows are null. We note further that, using the same partitioning of  $\mathbf{X}_2$ ,

$$\frac{1}{N}\mathbf{X}'_2\mathbf{X}_2 = \begin{bmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{A}' & \frac{1}{N}\mathbf{X}'_{22}\mathbf{X}_{22} \end{bmatrix}. \quad (3)$$

For designs with no partial aliasing, this matrix contains 0's, 1's, and -1's. Let  $m_i$  denote the number of non-zero elements of the  $i^{\text{th}}$  row of  $[\mathbf{I}, \mathbf{A}]$  ( $i = 1, \dots, df_2$ ). That is,  $m_i$  represents the length of the alias chain that includes the interaction in the  $i^{\text{th}}$  column of  $\mathbf{X}_{21}$ . These quantities were used by Cheng, Steinberg and Sun (1999) to study the model robustness of minimum aberration designs. Note that the  $i^{\text{th}}$  row of  $[\mathbf{I}, \mathbf{A}]$  (or its negative) is repeated  $m_i - 1$  times in the remaining  $k(k-1)/2 - df_2$  rows of (3). Computationally, alp may be obtained from (3) as well as from  $[\mathbf{I}, \mathbf{A}]$ . Define

$$\mathbf{d} = \text{diag}(\mathbf{X}'_2\mathbf{X}_2\mathbf{X}'_2\mathbf{X}_2)/N^2$$

For regular fractions,  $a_j = f_j / j$  ( $j = 1, \dots, L$ ), where  $f_j$  denotes the frequency of  $j$ 's in  $\mathbf{d}$ .

Note also that  $a_j$  is the number of  $m_i$ 's ( $i = 1, \dots, df_2$ ) equal to  $j$ .

We propose using  $\mathbf{d}$  as a measure of aliasing among two-factor interactions for nonregular strength three designs. For nonregular designs,  $\mathbf{X}'_2\mathbf{X}_2 / N$  contains elements other than 0's and  $\pm 1$ 's, reflecting partial aliasing among two-factor interactions. The elements of  $\mathbf{d}$  correspond to the sum of squares of each row of  $\mathbf{X}'_2\mathbf{X}_2 / N$ , and these are not necessarily integers. A referee noted that  $\mathbf{X}'_2\mathbf{X}_2 / N$  consists of elements of the confounding frequency vector (see Deng and Tang 1999), and that the average for  $\mathbf{d}$  equals  $\bar{d} = 1 + 12A_4^g / [k(k-1)]$ .

For both regular and nonregular designs, we define the *generalized alias length pattern* (galp) as the vector of  $f_j$ 's corresponding to  $\mathbf{d}$ . For designs proposed earlier:

- The OA(192,  $2^{60}$ , 3) has galp:  $f_{12}=1440; f_{30}=330$ . That is, no two-factor interactions are clear, but 1440 of the 1770 interactions have considerably less aliasing ( $j = 12$ ) than the other 330.
- The OA(192,  $2^{52}$ , 3) defined by the Kronecker product of  $\mathbf{H}_4$  and the OA(48,  $2^{13}$ , 3) with  $A_4^g = 2613$  has galp:  $f_4=192; f_{11.1111}=432; f_{14.666667}=576; f_{20}=48; f_{26}=78$ , so that 192 interactions are involved in relatively little aliasing.
- Results for a doubled OA(96,  $2^{25}$ , 3) with  $A_4^g = 245.56$  are reported in Table 2.

The initial OA(96,  $2^{25}$ , 3) has 24 clear two factor interactions. After doubling, the OA(192,  $2^{50}$ , 3) has 48 aliased pairs of two factor interactions, as reflected in  $f_2 = 4*24 = 96$ . If one begins with the OA(96,  $2^{25}$ , 3) obtained by reversing a single column (which has 47 clear interactions), the doubled array has  $f_2 = 4*47 = 188$ .

We know from Cheng, Steinberg, and Sun (1999) that minimum aberration designs tend to have uniform or nearly uniform length alias chains. This fact suggests that deleting 8-10 columns from the  $OA(192, 2^{60}, 3)$  will likely produce orthogonal arrays with less aberration than for the other two SOS designs in Table 2. We sequentially deleted columns from the  $OA(192, 2^{60}, 3)$ , in each case dropping a column that reduces  $A_4^g$  the most. The results are reported in Table 3 for 12 projections. Obviously, projections of the  $OA(192, 2^{60}, 3)$  design dominate (with respect to  $A_4^g$ ) the other SOS arrays with 192 runs in Table 2. This is to be expected in that the regular SOS designs with  $k = 5N/16$  have many projections that are minimum aberration.

## 5. Discussion

This article has presented the construction of new strength three arrays that increase the degrees of freedom for two-factor interactions. In so doing we have greatly enlarged the number of available designs that avoid aliasing of main effects and two-factor interactions. We have only distinguished the designs constructed based on generalized word length pattern and  $\text{rank}(\mathbf{X}_2)$ . However, non-isomorphic designs may have the same gwlp and identical aliasing lengths as reflected by the vector  $\mathbf{d} = \text{diag}(\mathbf{X}_2' \mathbf{X}_2 \mathbf{X}_2' \mathbf{X}_2) / N^2$ . To confirm that two designs are isomorphic requires examination of the projections, as described by Clark and Dean (2001). It could certainly be the case that two SOS designs may have identical gwlp and  $\mathbf{d}$ , but one has better projection designs for  $k' (< k)$  factors. Thus, further study of these designs is needed.

Wu and Hamada (2000, Section 8.4) propose methods for considering interactions in the modeling of data from nonregular designs. Even more closely connected to our designs, Miller and Sitter (2001) proposed a method for considering interactions in the analysis of 24-, 40-, and 48-run strength three designs based on the two-step procedure of first identifying significant main effects and then performing an all-subsets regression analysis involving two-factor interactions that contain one or more of the significant main effects. Miller and Sitter state that all-subsets regression is believed to be preferred to stepwise regression procedures. When the number of significant main effects is large, all-subsets regression will be computationally infeasible, even with the weak heredity constraint. For such cases, they propose doing all-subsets regression up to the largest computationally feasible size model, select the best subset, and then consider additional terms in a third stage that forces in the terms already identified as significant. This approach seems reasonable, although the question of whether one can discriminate between alternative models must be considered. Given these new  $OA(N, 2^k, 3)$  with complex aliasing, the need has increased for further study of model selection approaches. For some designs proposed here, there exist a full set of  $df_2$  orthogonal two-factor interaction columns. Whether this feature is useful for the analysis has not yet been investigated.

## **Appendix 1. Some Properties of Even Designs**

It is known that all even regular designs are foldover designs. In an unpublished work, one of the authors (C. S. Cheng) showed that the result holds more generally for nonregular designs. It was also obtained independently by Balakrishnan and Yang (2003)

at about the same time. That foldover designs are even is trivial. We include the converse here and use it to show that for even designs,  $\text{rank}[X_2] \leq N/2 - 1$ .

Lemma A1: A two-level even design must be a foldover design, i.e., it is of the form  $\begin{bmatrix} \mathbf{Y} \\ -\mathbf{Y} \end{bmatrix}$ .

Proof: For each factor-level combination  $\mathbf{x} = (x_1, \dots, x_k)$ , where  $x_i = \pm 1$ , let  $f(\mathbf{x})$  denote the number of times  $\mathbf{x}$  appears in the design. For each subset  $\mathbf{S}$  of  $\{1, \dots, k\}$ , let

$x_{\mathbf{S}} = \prod_{i \in \mathbf{S}} x_i$ . As shown in Ye (2003) as well as Tang (2001) in a slightly different form,

$f(\mathbf{x})$  can be expressed as  $f(\mathbf{x}) = b_0 + \sum_{\mathbf{S}} b_{\mathbf{S}} x_{\mathbf{S}}$ , where the sum is over all nonempty

subsets of  $\{1, \dots, k\}$ ; explicit forms of the coefficients  $b_{\mathbf{S}}$  are given in their papers. It

suffices to say that  $A_i^s$  is proportional to the sum of squares of the coefficients  $b_{\mathbf{S}}$  over all  $\mathbf{S}$

of size  $i$ . Thus, for even designs,  $b_{\mathbf{S}} = 0$  for all  $\mathbf{S}$  containing an odd number of elements.

Since  $x_{\mathbf{S}} = (-x)_{\mathbf{S}}$  for all  $\mathbf{S}$  containing an even number of elements, we have  $f(\mathbf{x}) = f(-\mathbf{x})$ ,

i.e., each factor-level combination and its mirror image appears the same number of times in the design.

According to Lemma A1, each even design is a foldover design, i.e., it can be partitioned as  $\mathbf{C}$ , followed by the foldover  $-\mathbf{C}$ , where  $\mathbf{C}$  is  $(N/2) \times k$ . Let  $\mathbf{x}_2$  denote the  $(N/2) \times k(k-1)/2$  matrix of all two-factor interactions from the columns of  $\mathbf{C}$ . Then the corresponding matrix  $\mathbf{X}_2$  for the  $N$ -run even design satisfies

$$\text{rank}[\mathbf{X}_2] = \text{rank} \begin{bmatrix} \mathbf{x}_2 \\ \mathbf{x}_2 \end{bmatrix} = \text{rank}[\mathbf{x}_2] \leq N/2 - 1.$$

In fact, this can be strengthened to show that the column space of all even effects

(excluding the intercept) has  $\text{rank} \leq N/2 - 1$ .

## Appendix 2. Construction of Alternative SOS OA(64, 2<sup>17</sup>, 3)

Let  $\mathbf{D}$  denote the regular  $2_{IV}^{16-11}$  fraction. The five non-isomorphic SOS  $2_{IV}^{17-11}$  designs may be obtained as follows:

- Let  $\mathbf{B}$  contain any single column. Then  $\mathcal{S}_1$  is Chen, Sun and Wu's (1993) (CSW) design 17-11.6 with the number of length four words ( $A_4$ ) = 105 and alp = (31,0,0,0,0,15), where alp (alias length pattern) is defined in the beginning of Section 4.
- Let  $\mathbf{B}$  contain any three columns. Then  $\mathcal{S}_1$  is CSW's design 17-11.5 with  $A_4 = 73$  and alp = (19,0,12,0,12,0,3).
- Let  $\mathbf{B}$  contain any four columns that do not form a length-four word for  $\mathbf{D}$ . Then  $\mathcal{S}_1$  is CSW's design 17-11.4 with  $A_4 = 68$  and alp = (16,6,0,18,0,6).
- Let  $\mathbf{B}$  contain any five columns such that no subset forms a length-four word for  $\mathbf{D}$ . Then  $\mathcal{S}_1$  is CSW's design 17-11.3 with  $A_4 = 65$  and alp = (16,0,15,0,15).
- Let  $\mathbf{B}$  contain any six columns such that no subset forms a length-four word for  $\mathbf{D}$ . Then  $\mathcal{S}_1$  is CSW's design 17-11.2 with  $A_4 = 60$  and alp = (16, 0,0,30).

SOS designs 17-11.2 and 17-11.4 are obtained by reversing an even number of columns. While reversing an even number of columns does not guarantee that  $\mathcal{S}_1$  is SOS, in this case and others in Table 1, the best SOS design cannot be constructed only by considering  $\mathbf{B}$  with an odd number of columns.

There exist five non-isomorphic OA(16,2<sup>15</sup>,2) designs – only one of which is regular (Hedayat, Sloane, and Stufkin 1999, p. 156). Taking  $\mathbf{D}$  as the foldover of one of these nonregular OAs, one can obtain many nonregular Type 1 SOS OA(64,2<sup>17</sup>,3) that have  $A_4^g$  and galp identical to the five regular SOS above, but with partial aliasing of

effects. Hall types II and III (Wu and Hamada 2000) both resulted in  $OA(64, 2^{17}, 3)$  designs with  $A_4^g = 60$  for  $b = 6$ ; no  $b$  achieved this small an  $A_4^g$  for Hall types IV and V.

### **Appendix 3. Construction of the Best Type 1 SOS $OA(N, 2^k, 3)$**

Here we provide additional detail regarding the best choice of columns for the arrays described in Table 1. To construct the  $OA(48, 2^{13}, 3)$ , any partitioning of the  $OA(24, 2^{12}, 3)$  into three columns for  $\mathbf{B}$  produces a Type 1 SOS design with  $A_4^g = 26$ . For a nonregular  $OA(64, 2^{17}, 3)$  with  $A_4^g = 60$ , one may begin with Hadamard matrix 16.1 from Sloane (1999), foldover to create  $\mathbf{D}$ , and then select columns 1-3, 5, 11, 14 for  $\mathbf{B}$ . For an  $OA(80, 2^{21}, 3)$  with  $A_4^g = 136.52$ , begin with Hadamard matrix 20.1 from Sloane (1999), foldover over to create  $\mathbf{D}$ , and then select columns 11, 13, 14, 16-20 for  $\mathbf{B}$ . Finally, for an  $OA(96, 2^{25}, 3)$  with  $A_4^g = 245.56$ , one may begin with Hadamard matrix 24.1 from Sloane (1999), foldover to create  $\mathbf{D}$ , and then select columns 12, 14, 15, 18, 20, and 22-24 for  $\mathbf{B}$ . These choices are just one of many yielding the same  $A_4^g$ . They are provided here for the convenience of the reader.

### **Acknowledgement**

We would like to thank Professor Boxin Tang for helpful discussions in the proofs of Theorems 1 and 2. Ching-Shui Cheng's research was partially supported by grant DMS-0505556 from the National Science Foundation.

## References

- Balakrishnan, N. and Yang, P. (2003). Connections between the resolutions of general two-level factorial designs. Preprint.
- Block, R.M. (2003). *Theory and Construction Methods for Large Regular Resolution IV Designs*. University of Tennessee Ph.D. dissertation: Knoxville. [Available at <http://etd.utk.edu/2003/BlockRobert.pdf>]
- Block, R.M. and Mee, R.W. (2003). Second order saturated resolution IV designs. *J. of Statistical Theory and Applications* **2**, 96-112.
- Bruen, A., Haddad, L., and Wehlau, L. (1998). Binary codes and caps. *J. of Combinatorial Designs* **6**, 275-284.
- Bruen, A. and Wehlau, D. (1999). Long binary linear codes and large caps in projective space. *Designs, Codes and Cryptography* **17**, 37-60.
- Chen, H. and Cheng, C.S. (2006). Doubling and projection: a method of constructing two-level designs of resolution IV, *Ann. Statist.* **34**, to appear.
- Chen, J., Sun, D.X., and Wu, C.F.J. (1993). A catalogue of two-level and three-level fractional factorial designs with small runs. *Int. Statist. Rev.* **61**, 131-145.
- Cheng, C. S., Steinberg, D. M., and Sun, D. X. (1999). Minimum aberration and model robustness for two-level fractional factorial designs. *J. of Roy. Statist. Soc. Ser. B* **61**, 85-93.
- Clark, J.B. and Dean, A.M. (2001). Equivalence of fractional factorial designs. *Statist. Sinica* **11**, 537-547.

- Davydov, A. A. and Tombak, L. M. (1990). Quasiperfect linear binary codes with distance 4 and complete caps in projective geometry. *Problems of Information Transmission* **25**, 265-275.
- Deng, L.Y. and Tang, B. (1999). Generalized resolution and minimum aberration criteria for Plackett-Burman and other nonregular factorial designs. *Statist. Sinica* **9**, 1071-1082.
- Hedayat, A.S., Sloane, N.J.A., and Stufken, J. (1999). *Orthogonal Arrays*. Springer: New York.
- Ma, C.X. and Fang, K.T. (2001). A note on generalized aberration in factorial designs. *Metrika* **53**, 85-93.
- Margolin, B.H. (1969). Results on factorial designs of resolution IV for the  $2^n$  and  $2^n 3^m$  series. *Technometrics* **11**, 431-444.
- Miller, A. and Sitter, R.R. (2001). Using the folded-over 12-run Plackett-Burman design to consider interactions. *Technometrics* **43**, 44-55.
- Sloane, N.J.A. (1999). A Library of Hadamard Matrices,  
<http://www.research.att.com/~njas/hadamard/index.html>.
- Tang, B. (2001). Theory of J-characteristics for fractional factorial designs and projection justification of minimum  $G_2$ -aberration. *Biometrika*, **88**, 401-407.
- Tang, B. (2005). Orthogonal arrays robust to nonnegligible two-factor interactions, preprint.
- Wu, C.F.J. and Hamada, M. (2000). *Experiments: Planning, Analysis, and Parameter Design Optimization*. John Wiley & Sons: New York.

Xu, H. and Wu, C.F.J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *Ann. Statist.* **29**, 1066-1077.

Ye, K.Q. (2003). Indicator functions and its application in two level factorial design. *Ann. Statist.* **31**, 984-994.

Table 1. Best Type 1 nonregular SOS designs

$N$	$k$	$k(k-1)/2$	$\text{rank}(\mathbf{X}_2)$	Best $A_4^g$ Found	Best $b$
48	13	78	34	26	3
64	17	136	46	60	6
80	21	210	58	136.52	8
96	25	300	70	245.56*	$\geq 8$

\* Search was not exhaustive.

Table 2: Alternative SOS OA(192,2<sup>k</sup>,3)

$N$	$k$	$k(k-1)/2$	$\text{rank}(\mathbf{X}_2)$	$A_4^g$	galp
192	50	1225	141	2264.4	$f_2 = 96; f_{7.56} = 4; f_{9.33} = 40;$ $f_{10.22} = 40; f_{11.11} = 100; f_{11.56} = 80;$ $f_{12} = 256; f_{12.89} = 260; f_{13.78} = 120;$ $f_{14.67} = 40; f_{15.11} = 160; f_{16.44} = 4;$ $f_{25} = 25$
192	52	1326	139	2613	$f_4 = 192; f_{11.111} = 432; f_{14.667} = 576;$ $f_{20} = 48; f_{26} = 78$
192	60	1770	131	4235	$f_{12} = 1440; f_{30} = 330$

Table 3: Twelve projections from OA(192,2<sup>60</sup>,3)

$k$	Column removed	Best $A_4^g$	galp and $\bar{d}$
60		4235	$f_{12}=1440; f_{30}=330$ $\bar{d}=15.356$
59	60	3952.7	$f_{11}=528; f_{12}=864; f_{29}=319$ $\bar{d}=14.861$
58	1	3681.3	$f_{10}=30; f_{10.11}=90; f_{11}=793; f_{12}=432; f_{28}=308$ $\bar{d}=14.362$
57	17	3421	$f_{10}=90; f_{10.11}=270; f_{11}=795; f_{12}=144; f_{27}=297$ $\bar{d}=13.861$
56	23	3171.7	$f_{10}=180; f_{10.11}=540; f_{11}=534; f_{26}=286$ $\bar{d}=13.357$
55	29	2933.3	$f_{10}=300; f_{10.11}=900; f_{11}=10; f_{25}=275$ $\bar{d}=12.852$
54	31	2720	$f_9=72; f_{9.11}=72; f_{9.22}=288; f_{10}=188; f_{10.11}=540;$ $f_{11}=6; f_{24}=72; f_{24.11}=189; f_{25}=4$ $\bar{d}=12.405$
53	4	2515.7	$f_{8.11}=18; f_{8.33}=72; f_9=118; f_{9.11}=108; f_{9.22}=432;$ $f_{10}=102; f_{10.11}=270; f_{11}=3; f_{23}=27; f_{23.11}=66;$ $f_{23.22}=156; f_{24}=6$ $\bar{d}=11.954$
52	2	2320.3	$f_{8.11}=54; f_{8.33}=216; f_9=138; f_{9.11}=108; f_{9.22}=432;$ $f_{10}=42; f_{10.11}=90; f_{11}=1; f_{22}=14; f_{22.11}=9;$ $f_{22.22}=90; f_{22.33}=126; f_{23}=6$ $\bar{d}=11.499$
51	3	2134	$f_{8.11}=108; f_{8.33}=432; f_9=132; f_{9.11}=72; f_{9.22}=288;$ $f_{10}=8; f_{21}=6; f_{21.22}=18; f_{21.33}=108; f_{21.44}=99;$ $f_{22}=4$ $\bar{d}=11.042$
50	5	1956.7	$f_{8.11}=180; f_{8.33}=720; f_9=100; f_{20.33}=30;$ $f_{20.44}=120; f_{20.56}=75$ $\bar{d}=10.584$
49	55	1799.8	$f_{7.22}=32; f_{7.33}=64; f_{7.56}=224; f_8=32; f_{8.33}=432; f_9$ $=60; f_{19.44}=24; f_{19.56}=58; f_{19.67}=80; f_{19.78}=50;$ $f_{20.44}=4$ $\bar{d}=10.183$
48	11	1650.2	$f_{6.56}=28; f_{6.89}=42; f_7=7; f_{7.22}=48; f_{7.33}=100;$ $f_{7.56}=336; f_8=48; f_{8.11}=66; f_{8.33}=216; f_9=30;$ $f_{18.56}=12; f_{18.67}=36; f_{18.78}=73; f_{18.89}=48; f_{19}=30;$ $f_{19.67}=8$ $\bar{d}=9.778$