Statistics & Probability Letters 76 (2006) 1570–1577

Allocating factors to the columns of an orthogonal array when certain interactions are important

Ashish Das[a,*], Aloke Dey[a], Chand K. Midha[b]

[a] Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, New Delhi 110 016, India
[b] Department of Statistics, The University of Akron, Akron, OH 44325, USA

Received 21 September 2005; received in revised form 29 January 2006; accepted 1 March 2006
Available online 18 April 2006

Abstract

Fractional factorial plans represented by orthogonal arrays of strength two are known to be universally optimal under a model that includes the mean and all main effects, when all other factorial effects are assumed to be absent. However, if the number of factors in the experiment is smaller than the number of columns of a saturated or, tight orthogonal array and the experimenter is interested in estimating certain 2-factor interactions as well, one can possibly entertain these 2-factor interactions in the model apart from the mean and all main effects. The problem then is to allocate factors to the columns of the orthogonal array, so that the user-specified 2-factor interactions, in addition to the mean and the main effects, are optimally estimable. This problem is investigated in this paper with reference to the orthogonal array $OA(2^n, 2^n - 1, 2, 2)$, which exists for every integer $n \geqslant 2$. A method for the allocation of factors to factor representations is proposed that ensures the optimal estimation of the mean, all main effects and specified 2-factor interactions. The method is illustrated by considering in detail the cases $n = 3, 4$.

Keywords: Fractional factorial experiments; Universal optimality; Main effects; 2-factor interactions

1. Introduction

Two-level fractional factorial plans are used extensively in many diverse fields, notably in industrial experimentation and quality control work. Such fractional factorial plans when represented by orthogonal arrays have strong optimality properties. For instance, a fractional factorial plan represented by an orthogonal array of strength two is universally optimal in the sense of Kiefer (1975) and Sinha and Mukerjee (1982) for estimating the mean and all main effects when all interactions are assumed to be absent. Recall that a symmetric orthogonal array, denoted by $OA(N, n, m, g)$, with N rows (runs), n columns (factors), m symbols (levels) and strength $g (2 \leqslant g < n)$ is an $N \times n$ matrix, with $m \geqslant 2$ distinct symbols in each column such that in

*Corresponding author.
E-mail address: ashishdas.das@gmail.com (A. Das).

any $N \times g$ submatrix, each possible combination of the symbols appears equally often as a row. For comprehensive accounts on orthogonal arrays and their applications in the context of fractional factorials, see Hedayat et al. (1999) and Dey and Mukerjee (1999a).

For an integer $n \geq 2$, an orthogonal array $\mathrm{OA}(2^n, 2^n - 1, 2, 2)$ of strength two exists and can be constructed in the following manner: for $1 \leq i \leq n$, let x_i be a $2^n \times 1$ vector with entries 0 and 1, such that the rows of the $2^n \times n$ matrix formed by n columns $x_1, \ldots, x_n$ are the all possible combinations of a 2^n factorial. Next, form $2^n - 1$ columns $k_1x_1 + k_2x_2 + \cdots + k_nx_n$, where each k_i is either 0 or 1, not all k_i's are simultaneously zero and the elements in the sums are reduced modulo 2. Note that these $2^n - 1$ columns also include the (initial) columns $x_1, \ldots, x_n$. These $2^n - 1$ columns form an orthogonal array $\mathrm{OA}(2^n, 2^n - 1, 2, 2)$. It follows then that no two columns in the set $\mathscr{S} = \{k_1x_1 + k_2x_2 + \cdots + k_nx_n : k_i = 0, 1, (k_1, \ldots, k_n) \neq (0, \ldots, 0)\}$ are identical. We can represent the columns of the orthogonal array as $(\mathbf{1}, \mathbf{2}, \mathbf{12}, \mathbf{3}, \mathbf{13}, \mathbf{23}, \mathbf{123}, \mathbf{4}, \ldots, \mathbf{1234}, \ldots, \boldsymbol{n})$, where the representation $\boldsymbol{i}$ stands for the column x_i, $1 \leq i \leq n$, the representation $\boldsymbol{ij}$ stands for the column $x_i + x_j \pmod 2$, $i, j = 1, \ldots, n, i \neq j$, and so on. Such an array represents a 2^n-run fractional factorial plan for a 2-level experiment involving $2^n - 1$ factors. This plan is saturated and is universally optimal for estimating the mean and all main effects in the absence of 2-factor and higher order interactions. In view of this, fractional factorial plans represented by orthogonal arrays of strength two have traditionally been used for estimating main effects alone.

Restricting attention to the orthogonal array $\mathrm{OA}(2^n, 2^n - 1, 2, 2)$, as described above, suppose the number of factors in the experiment is k $(< 2^n - 1)$, and, furthermore, the experimenter is interested in estimating some 2-factor interactions as well. Can one then entertain these 2-factor interactions in the model, along with the mean and the main effects of the k factors involved? This leads to the problem of allocating the k factors to the above "factor representations", so that the user-specified 2-factor interactions, in addition to the mean and the main effects, are optimally estimable. This problem is investigated in this paper. A method for the allocation of factors to factor representations is proposed that ensures the optimal estimation of the mean, all main effects and specified 2-factor interactions. However, there are a few occasions where the proposed method is unable to suggest an allocation ensuring optimal estimation. The method is illustrated by considering in detail the cases $n = 3, 4$.

The problem of estimating main effects and specified 2-factor interactions via a fractional factorial plan has been studied earlier e.g., by Hedayat and Pesotan (1992, 1997), Wu and Chen (1992), Dey and Mukerjee (1999b) and Dey and Suen (2002). However, the problem addressed in this communication is slightly different from the ones considered hitherto in the literature.

For obtaining the optimal plans in this paper, we make use of a result of Dey and Mukerjee (1999b). For completeness, we state the result below in a form that is needed for this paper.

Theorem 1. *Let $\mathscr{D}$ be the class of all N-run fractional factorial plans for an arbitrary factorial experiment involving k factors, $F_1, \ldots, F_k$, such that each member of $\mathscr{D}$ allows the estimability of the mean, the main effects $F_1, \ldots, F_k$ and the t 2-factor interactions $F_{i_1}F_{j_1}, \ldots, F_{i_t}F_{j_t}$, where $1 \leq i_u \neq j_u \leq k$ for all $u = 1, \ldots, t$. A plan $d \in \mathscr{D}$ is universally optimal over $\mathscr{D}$ if all level combinations of the following sets of factors appear equally often in d:*

(a) $\{F_u, F_v\}$, $1 \leq u < v \leq k$;
(b) $\{F_u, F_{i_v}, F_{j_v}\}$, $1 \leq u \leq k$, $1 \leq v \leq t$;
(c) $\{F_{i_u}, F_{j_u}, F_{i_v}, F_{j_v}\}$, $1 \leq u < v \leq t$,

where a factor is counted only once if it is repeated in (b) *or* (c).

2. A method for allocation of factors

Throughout, we consider a fractional factorial plan represented by the orthogonal array $\mathrm{OA}(2^n, 2^n - 1, 2, 2)$, as described in the previous section. Clearly, the addition of any two distinct columns of this array, each of which is a member of $\mathscr{S}$, gives rise to a different column belonging to $\mathscr{S}$. To begin with, we have the following result.

Lemma 1. *Let $x_{i_1}, x_{i_2}, x_{i_3}$ be any three distinct columns of the orthogonal array $\mathrm{OA}(2^n, 2^n - 1, 2, 2)$, $n \geq 3$, such that $x_{i_1} + x_{i_2} + x_{i_3} \neq 0 \pmod 2$. Then, the columns $x_{i_1}, x_{i_2}, x_{i_3}$ form an orthogonal array of strength three.*

Proof. Let $f(u,v,w)$, $u,v,w = 0,1$, denote the frequency of the ordered triplet (u,v,w) under the columns $x_{i_1}, x_{i_2}, x_{i_3}$. Then, since $x_{i_1}, x_{i_2}, x_{i_3}$ are three distinct columns of the orthogonal array $\mathrm{OA}(2^n, 2^{n-1}, 2, 2)$ of strength two, we have

$$f(0,0,0)+f(0,0,1)=f(0,1,0)+f(0,1,1)=f(1,0,0)+f(1,0,1)=f(1,1,0)+f(1,1,1)=2^{n-2},$$

$$f(0,0,0)+f(1,0,0)=f(0,0,1)+f(1,0,1)=f(0,1,0)+f(1,1,0)=f(0,1,1)+f(1,1,1)=2^{n-2},$$

and

$$f(0,0,0)+f(0,1,0)=f(0,0,1)+f(0,1,1)=f(1,0,0)+f(1,1,0)=f(1,0,1)+f(1,1,1)=2^{n-2}.$$

These yield

$$f(0,0,1)=f(1,0,0)=f(0,1,0)=f(1,1,1)$$

and

$$f(0,0,0)=f(1,0,1)=f(0,1,1)=f(1,1,0).$$

Since $x_{i_1}+x_{i_2}\neq x_{i_3} \pmod 2$ and $x_{i_1}+x_{i_2}=x_{i_4}\in \mathscr{S}$ for some $x_{i_4}\neq x_{i_3}$, and since x_{i_3} and x_{i_4} form an orthogonal array of strength two, we have

$$f(0,0,0)+f(1,1,0)=f(0,0,1)+f(1,1,1)=f(0,1,0)+f(1,0,0)=f(0,1,1)+f(1,0,1)=2^{n-2}.$$

It follows then that $f(u,v,w)=2^{n-3}$ for all $u,v,w=0,1$, completing the proof. □

For convenience, we henceforth represent the columns of the orthogonal array $\mathrm{OA}(2^n, 2^n-1, 2, 2)$ by an n-tuple $(k_1,k_2,\ldots,k_n)$, where for $1\leqslant i\leqslant n$, $k_i=0$ or 1 and $(k_1,k_2,\ldots,k_n)\neq(0,0,\ldots,0)$. Thus, the representation **1** is now equivalently denoted by $(100\ldots0)$, **124** is equivalent to $(11010\ldots0)$, etc. We shall denote by Ω, the collection of all such 2^n-1 non-null binary vectors. It is not hard to see that the columns $\{\sum_{i=1}^n k_i x_i\}$ are completely characterized by their respective binary representations, $(k_1,k_2,\ldots,k_n)$, i.e., by the elements of Ω. Clearly, the result of Lemma 1 holds when x_{i_j} is replaced by its binary representation. Henceforth, unless otherwise mentioned, we work with the n-tuples, $\{z_i\}$ $1\leqslant i\leqslant 2^n-1$, each belonging to Ω, where $z_i=(k_1,k_2,\ldots,k_n)$ and $i=\sum_{j=1}^n k_j 2^{n-j}$.

Now, suppose the number of factors involved in the experiment is k $(<2^n-1)$, but the experimenter can afford to make 2^n runs. Additionally, the experimenter is interested in estimating certain 2-factor interactions, along with the mean and the k main effects. The model postulated is one that includes the mean, k main effects and the specified 2-factor interactions. It is assumed that all interactions, not included in the model are absent. How should one assign the factors to the columns of the orthogonal array so that the mean, all main effects and the specified 2-factor interactions are optimally estimable?

To begin with, suppose the experiment involves $k<2^n-1$ factors and it is desired to estimate the mean, the k main effects and only *one* 2-factor interaction. Let the representation of the specified 2-factor interaction be $z_{i_1}z_{i_2}$, where $z_{i_1}, z_{i_2}\in\Omega$ and let $z_p=z_{i_1}+z_{i_2} \pmod 2$. Then from Lemma 1, all triplets of columns of the orthogonal array represented by $\{z_{i_1}, z_{i_2}, z_{i_j}\}$, $z_{i_j}\in\Omega\backslash\{z_p\}$, form an orthogonal array of strength three. Using Theorem 1 then, the design involving the columns of the orthogonal array represented by $z_{i_1}, z_{i_2},\ldots,z_{i_k}$, such that for $1\leqslant u\leqslant k$, $z_{i_u}\neq z_p$, is universally optimal for estimating the mean, the k main effects and the interaction $z_{i_1}z_{i_2}$.

Next, suppose the experiment involves $k<2^n-2$ factors and one is interested in estimating the mean, the main effects of the k factors and *two* 2-factor interactions. There are only two possible types of interactions that need to be considered, namely,(a) $z_{i_1}z_{i_2}$ and $z_{i_3}z_{i_4}$, (b) $z_{i_1}z_{i_2}$ and $z_{i_1}z_{i_3}$.

Case a: The factorial effects involved here are the interactions of the type $z_{i_1}z_{i_2}$ and $z_{i_3}z_{i_4}$, in addition to the mean and main effects of k factors. Here, $z_{i_1}, z_{i_2}, z_{i_3}$ and z_{i_4} are all distinct. Then, all triplets of columns of the orthogonal array represented by $\{z_{i_1}, z_{i_2}, z_{i_j}\}$, $z_{i_j}\in\Omega\backslash\{z_p\}$, form an orthogonal array of strength three. Similarly, all triplets of columns of the orthogonal array represented by $\{z_{i_3}, z_{i_4}, z_{i_j}\}$, $z_{i_j}\in\Omega\backslash\{z_q\}$, where $z_q=z_{i_3}+z_{i_4} \pmod 2$, form an orthogonal array of strength three. Furthermore, arguing as in the proof of Lemma 1, it can be shown that the columns of the orthogonal array represented by $\{z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}\}$ form an orthogonal array of strength four, as long as $z_{i_1}+z_{i_2}+z_{i_3}+z_{i_4}\neq 0 \pmod 2$, i.e., as long as $z_p\neq z_q$. Now, invoking Theorem 1, we see that the design involving columns represented by $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4},\ldots,z_{i_k}$, such that none of these equal either z_p or z_q $(\neq z_p)$, is universally optimal for estimating the mean, the k main effects and the interactions $z_{i_1}z_{i_2}$ and $z_{i_3}z_{i_4}$.

Case b: In this case, we consider interactions of the type $z_{i_1}z_{i_2}$ and $z_{i_1}z_{i_3}$, apart from the mean and main effects of k factors, including $z_{i_1}, z_{i_2}, z_{i_3}$, $3 \leqslant k \leqslant 2^n - 3$. The vectors $z_{i_1}, z_{i_2}, z_{i_3}$ are distinct. Then, all triplets of columns of the orthogonal array represented by $\{z_{i_1}, z_{i_2}, z_{i_j}\}$, $z_{i_j} \in \Omega\backslash\{z_p\}$, form an orthogonal array of strength three. Similarly, if $z_s = z_{i_1} + z_{i_3}$ (mod 2), then all triplets of columns of the orthogonal array represented by $\{z_{i_1}, z_{i_3}, z_{i_j}\}$, $z_{i_j} \in \Omega\backslash\{z_s\}$, form an orthogonal array of strength three. From Theorem 1, the design involving columns represented by $z_{i_1}, z_{i_2}, z_{i_3}, \ldots, z_{i_k}$, such that none of these equal either z_p or z_s, is universally optimal for estimating the mean, the k main effects and the interactions $z_{i_1}z_{i_2}$ and $z_{i_1}z_{i_3}$.

Finally, consider a general setup where there are k factors and one is interested in estimating the k main effects and t interactions, say $z_{i_1}z_{j_1}, \ldots, z_{i_t}z_{j_t}$, $1 \leqslant t < 2^n - k$. For $t \geqslant 2$, any two of these t interactions may or may not have a common factor, i.e., i_l could be equal to i_m or, j_m, for $l \neq m$.

For $1 \leqslant l \leqslant t$, let $z_{i_l} + z_{j_l} = z_{u_l}$ (mod 2). Then, for $1 \leqslant l \leqslant t$, the triplets of columns of the orthogonal array represented by $\{z_{i_l}, z_{j_l}, z_v\}$, $z_v \neq z_{u_l}$, $z_v \in \Omega$, form an orthogonal array of strength three. Furthermore, consider four distinct vectors $z_{i_\alpha}, z_{j_\alpha}, z_{i_\beta}, z_{j_\beta}$, $1 \leqslant \alpha, \beta \leqslant t$. Then, the columns of the orthogonal array represented by these vectors form an orthogonal array of strength four, as long as $z_{i_\alpha} + z_{j_\alpha} + z_{i_\beta} + z_{j_\beta} \neq 0$ (mod 2), i.e., $z_{u_\alpha} \neq z_{u_\beta}$.

Then from Theorem 1, the design involving k columns of the orthogonal array represented by $z_{i_1}, z_{j_1}, \ldots, z_{i_t}, z_{j_t}$ in addition to other members of Ω , such that none of these equal the distinct vectors z_{u_1} or, $z_{u_2}, \ldots,$ or, z_{u_t}, is universally optimal for estimating the mean, the k main effects and the t interactions $z_{i_1}z_{j_1}, \ldots, z_{i_t}z_{j_t}$.

We illustrate these ideas by considering the 8- and 16-run plans in detail in the next two sections.

3. Eight-run plans

For $n = 3$, one gets the array OA(8, 7, 2, 2). The representations of the columns of this array are $z_1 = (001), z_2 = (010), z_3 = (011), z_4 = (100), z_5 = (101), z_6 = (110), z_7 = (111)$. These representations are equivalent to the representations **3, 2, 23, 1, 13, 12, 123**. Suppose there are $k \leqslant 6$ factors involved and one is interested in estimating the 2-factor interaction $z_{i_1}z_{j_1}$ in addition to the mean and the k main effects. Let $z_{i_1} + z_{j_1} = z_{u_1}$ (mod 2). Then, the design with k factors represented by z_{i_1}, z_{j_1} and an additional $k - 2$ members of $\Omega\backslash\{z_{i_1}, z_{j_1}, z_{u_1}\}$, is universally optimal. To elaborate further, let there be $k = 6$ factors, denoted by $A, B, \ldots, F$ and, suppose one is interested in estimating the mean, the 6 main effects and an interaction AB (without loss of generality). Then we can assign factors $A, B, \ldots, F$ to the columns **1, 2, 3, 13, 23, 123**, to arrive at the desired optimal design. An optimal design under a model that includes the mean, the interaction AB and main effects of fewer than 6 factors can be obtained by deleting an appropriate number of columns of the design with 6 factors.

Next, let there be $k = 4$ (or, $k = 5$) factors, denoted by A, B, C, D (or, $A, B, \ldots, E$, respectively) and, suppose one is interested in estimating the mean, the 4 (or, 5) main effects and two interactions, say AB and CD. Suppose the factors A, B, C, D have distinct representations $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}$, respectively. Furthermore, let $z_{i_1} + z_{i_2} = z_{i_5}$ (mod 2) and $z_{i_3} + z_{i_4} = z_{i_6}$ (mod 2). Then, $z_{i_1} + z_{i_2} + z_{i_5} = 0$ (mod 2) and $z_{i_3} + z_{i_4} + z_{i_6} = 0$ (mod 2). This implies that the remaining one factor representation z_{i_7} must be equal to zero, which is impossible. It follows then that with only 8 runs, it is impossible to estimate the interactions AB and CD (i.e., a pair of interactions with no common factor) simultaneously.

This fact can also be observed by first constructing all possible distinct sets of triplets $\{z_{i_1}, z_{i_2}, z_{i_3}\}$ such that $z_{i_1} + z_{i_2} + z_{i_3} = 0$ (mod 2). These triplets indicate the column representations of two factors whose interaction is present and the column which is not to be considered as a factor. The sets of triplets are

{**1 2 12**},

{**1 3 13**},

{**1 23 123**},

{**2 3 23**},

{**2 13 123**},

{**3 12 123**},

{**12 13 23**}.

Treating the column representations as the 7 treatment symbols and the sets as blocks one gets a balanced incomplete block (BIB) design with 7 treatments and 7 blocks each of size 3. From a well-known property of a symmetric BIB design, any two blocks has a common treatment. Thus, we cannot have 6 distinct symbols appearing in a pair of blocks. This leads us to the earlier observed fact that it is impossible to estimate simultaneously a pair of interactions with no common factor via the considered 8-run plan.

Next, we look at two or more 2-factor interactions such that any two interactions have a factor in common. Consider two interactions of the type AB and AC. Then we can assign factors $A, B, \ldots, E$ to the columns **1, 2, 3, 23, 123** to arrive at the desired optimal design. An optimal design with 4 factors can be obtained by deleting either of the columns **23** or **123**. When $t = 3$ interactions are of interest, the possible types of interactions are (i) AB, AC and AD, or (ii) AB, AC and BC. For each of these, we can assign factors A, B, C, D to the columns **1, 2, 3, 123** leading to the desired optimal design.

4. Sixteen-run plans

Consider the orthogonal array OA(16, 15, 2, 2) as a special case of the array OA(2^n, $2^n - 1$, 2, 2) with $n = 4$. The representations of the columns of this array are $z_1 = (0001), z_2 = (0010), z_3 = (0011), z_4 = (0100), z_5 = (0101), z_6 = (0110), z_7 = (0111), z_8 = (1000), z_9 = (1001), z_{10} = (1010), z_{11} = (1011), z_{12} = (1100), z_{13} = (1101), z_{14} = (1110), z_{15} = (1111)$. These representations are equivalent to the representations **4, 3, 34, 2, 24, 23, 234, 1, 14, 13, 134, 12, 124, 123, 1234**. Suppose there are $k \leqslant 14$ factors involved and one is interested in estimating the 2-factor interaction $z_{i_1} z_{j_1}$ in addition to the mean and the k main effects. Let $z_{i_1} + z_{j_1} = z_{u_1}$ (mod 2). Then, the design with k factors represented by z_{i_1}, z_{j_1} and an additional $k - 2$ members of $\Omega - \{z_{i_1}, z_{j_1}, z_{u_1}\}$ is universally optimal. To elaborate further, let there be $k = 14$ factors, denoted by $A, B, \ldots, N$ and, suppose one is interested in estimating the mean, the 14 main effects and an interaction AB (without loss of generality). Then we can assign factors $A, B, \ldots, N$ to the columns such that A and B are assigned to **1** and **2**, while the other factors are assigned to the remaining columns *except* column **12**. This gives us the desired optimal design. An optimal design with fewer than 14 factors can be obtained by deleting an appropriate number of columns of the design with 14 factors.

Next, let there be $k \leqslant 13$ factors, denoted by $A, B, \ldots, M$ and, suppose one is interested in estimating the mean, the k main effects and two interactions, say AB and CD. Suppose the factors A, B, C, D have distinct representations $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}$, respectively. Furthermore, let $z_{i_1} + z_{i_2} = z_{i_5}$ (mod 2) and $z_{i_3} + z_{i_4} = z_{i_6}$ (mod 2). Then, $z_{i_1} + z_{i_2} + z_{i_5} = 0$ (mod 2) and $z_{i_3} + z_{i_4} + z_{i_6} = 0$ (mod 2). This implies that we need to identify distinct values of $z_{i_1}, z_{i_2}, z_{i_5}, z_{i_3}, z_{i_4}, z_{i_6}$ which satisfy the above two equations.

As in the case of 8-run plans, we now construct, for $n = 4$, all possible distinct sets of triplets $\{z_{i_1}, z_{i_2}, z_{i_3}\}$ such that $z_{i_1} + z_{i_2} + z_{i_3} = 0$. These triplets indicate the column representations of two factors whose interaction is present and the column which is not to be considered as a factor. These sets of triplets are

{**1 2 12**},	{**3 4 34**},	{**13 24 1234**},	{**14 123 234**},	{**23 124 134**},
{**1 3 13**},	{**2 14 124**},	{**4 23 234**},	{**12 34 1234**},	{**24 123 134**},
{**1 4 14**},	{**2 134 1234**},	{**3 12 123**},	{**13 124 234**},	{**23 24 34**},
{**1 23 123**},	{**2 4 24**},	{**3 124 1234**},	{**12 134 234**},	{**13 14 34**},
{**1 24 124**},	{**2 34 234**},	{**3 14 134**},	{**4 123 1234**},	{**12 13 23**},
{**1 34 134**},	{**2 13 123**},	{**3 24 234**},	{**4 12 124**},	{**14 23 1234**},
{**1 234 1234**},	{**2 3 23**},	{**4 13 134**},	{**12 14 24**},	{**34 123 124**}.

As before, if we treat the column representations as the 15 treatment symbols, the above sets of triplets form a BIB design with 15 treatments and 35 blocks each of size 3. The above design is in fact a *resolvable* BIB design with 5 blocks in each row forming a complete replication. (Recall that a BIB design with v treatments, b blocks and replication number r is said to be resolvable if the b blocks can be partitioned into r sets of $m = b/r$ blocks such that each set contains every treatment exactly once.)

We can assign factors $A, B, \ldots, M$ to the columns such that A, B, C and D are assigned to **1**, **2**, **3** and **4**, respectively. The other factors are assigned to the remaining columns *except* columns **12** and **34**. This gives us the desired optimal design. An optimal design with fewer than 13 factors can be obtained by deleting an appropriate number of columns of the design with 13 factors.

Under the above setup, in case the two interactions are, say AB and AC, then let $z_{i_1} + z_{i_2} = z_{i_4}$ (mod 2) and $z_{i_1} + z_{i_3} = z_{i_5}$ (mod 2). Then, $z_{i_1} + z_{i_2} + z_{i_4} = 0$ (mod 2) and $z_{i_1} + z_{i_3} + z_{i_5} = 0$ (mod 2). This implies that we need to identify distinct values of $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}, z_{i_5}$ which satisfies the above two equations. We can assign factors $A, B, \ldots, M$ to the columns such that A, B and C are assigned to **1**, **2** and **3**, respectively. The other factors are assigned to the remaining columns *except* columns **12** and **13**. This gives us the desired optimal design. An optimal design with fewer than 13 factors can be obtained by deleting an appropriate number of columns of the design with 13 factors.

When $t = 3$, without loss of generality, there are four possible types of interactions: (i) AB, CD, EF; (ii) AB, AC, DE; (iii) AB, AC, AD; (iv) AB, AC, BC. For each of these cases, we can assign suitable columns to the factors, leading to the desired optimal design. The given resolvable BIB design is helpful in obtaining solutions to the equation sets that arise in each of the four cases. Suppose the factors A, B, C, D, E, F have distinct representations $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}, z_{i_5}, z_{i_6}$, respectively.

Case i: Interactions AB, CD, EF. Let $z_{i_1} + z_{i_2} = z_{i_7}$ (mod 2), $z_{i_3} + z_{i_4} = z_{i_8}$ (mod 2), $z_{i_5} + z_{i_6} = z_{i_9}$ (mod 2). Then, we have

$$z_{i_1} + z_{i_2} + z_{i_7} = 0 \pmod 2, \quad z_{i_3} + z_{i_4} + z_{i_8} = 0 \pmod 2, \quad z_{i_5} + z_{i_6} + z_{i_9} = 0 \pmod 2.$$

Now we need to identify distinct values of $z_{i_1}, z_{i_2}, z_{i_7}, z_{i_3}, z_{i_4}, z_{i_8}, z_{i_5}, z_{i_6}, z_{i_9}$ which satisfy the above three equations. The answer lies in the 7 sets of the resolvable BIB design. Considering any one set, say the first set,

{**1** **2** **12**}, {**3** **4** **34**}, {**13** **24** **1234**}, {**14** **123** **234**}, {**23** **124** **134**},

and after selecting any three blocks, for each block, assigning the two interacting factors to any two (of the three) elements of the block, we get our desired allocation. In each of the three blocks, the element not assigned to any factor would be the column in the original orthogonal array, which is *not* to be allocated to any factor. The non-interacting $k - 6$ factors are assigned the remaining columns. Thus, from the above set of 5 blocks, we can assign (for example) the k factors to the columns such that A, B, C, D, E and F are assigned to **1**, **2**, **3**, **4**, **13** and **24**, respectively. The other factors are assigned to the remaining columns *except* columns **12**, **34** and **1234**. This gives us the desired optimal design.

Case ii: Interactions AB, AC, DE. Let $z_{i_1} + z_{i_2} = z_{i_6}$ (mod 2), $z_{i_1} + z_{i_3} = z_{i_7}$ (mod 2), $z_{i_4} + z_{i_5} = z_{i_8}$ (mod 2). Then,

$$z_{i_1} + z_{i_2} + z_{i_6} = 0 \pmod 2, \quad z_{i_1} + z_{i_3} + z_{i_7} = 0 \pmod 2, \quad z_{i_4} + z_{i_5} + z_{i_8} = 0 \pmod 2.$$

Now we need to identify distinct values of $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}, z_{i_5}, z_{i_6}, z_{i_7}, z_{i_8}$ which satisfy the above three equations. The answer lies in any two of the 7 sets of the resolvable BIB design. We can assign the k factors to the columns such that A, B, C, D and E are assigned to **1**, **2**, **3**, **14** and **123**, respectively. The other factors are assigned to the remaining columns *except* columns **12**, **13** and **234**. This gives us the desired optimal design.

Case iii: Interactions AB, AC, AD. Let $z_{i_1} + z_{i_2} = z_{i_5}$ (mod 2), $z_{i_1} + z_{i_3} = z_{i_6}$ (mod 2), $z_{i_1} + z_{i_4} = z_{i_7}$ (mod 2). Then it follows that

$$z_{i_1} + z_{i_2} + z_{i_5} = 0 \pmod 2, \quad z_{i_1} + z_{i_3} + z_{i_6} = 0 \pmod 2, \quad z_{i_1} + z_{i_4} + z_{i_7} = 0 \pmod 2.$$

We need to identify distinct values of $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}, z_{i_5}, z_{i_6}, z_{i_7}$ which satisfy the above three equations This time, the answer lies in any three of the 7 sets of the resolvable BIB design. We can assign the k factors to the columns such that A, B, C and D are assigned to **1**, **2**, **3** and **4**, respectively. The other factors are assigned to the remaining columns *except* columns **12**, **13** and **14**. This gives us the desired optimal design.

Case iv: Interactions AB, AC, BC. Let $z_{i_1} + z_{i_2} = z_{i_4}$ (mod 2), $z_{i_1} + z_{i_3} = z_{i_5}$ (mod 2), $z_{i_2} + z_{i_3} = z_{i_6}$ (mod 2). It follows that

$$z_{i_1} + z_{i_2} + z_{i_4} = 0 \pmod 2, \quad z_{i_1} + z_{i_3} + z_{i_5} = 0 \pmod 2, \quad z_{i_2} + z_{i_3} + z_{i_6} = 0 \pmod 2.$$

Now we need to identify distinct values of $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}, z_{i_5}, z_{i_6}$ which satisfy the above three equations. Again, using the resolvable BIB design, we see that an allocation is to assign the k factors to the columns such that A, B and C are assigned to **1**, **2** and **3**, respectively. The other factors are assigned to the remaining columns *except* columns **12**, **13** and **23**. This gives us the desired optimal design.

On lines similar to the case of $t = 3$, we are able to solve for possible situations for $t = 4$. For illustration, we consider a few cases. In what follows, we represent any assignment of a factor say, A to a column, say 1, by $(A, 1)$.

Case i: Interactions AB, CD, EF, GH. The assignments are: $(A,1),(B,2),(C,3),(D,4),(E,13),(F,24),(G,14)$ and $(H,123)$. The columns that are not to be assigned any factors are 12, 34, 1234 and 234. The other factors are assigned to the remaining columns.

Case ii: Interactions AB, CD, EF, EG. The assignments are: $(A,1),(B,2),(C,3),(D,4),(E,13),(F,24)$ and $(G,124)$. The columns that are not to be assigned any factors are 12, 34, 1234 and 234. The other factors are assigned to the remaining columns.

Case iii: Interactions AB, CD, CE, CF. The assignments are: $(A,1),(B,2),(C,3),(D,4),(E,124)$ and $(F,14)$. The columns that are not to be assigned any factors are 12, 34, 1234 and 134. The other factors are assigned to the remaining columns.

Case iv: Interactions AB, CD, CE, DE. The assignments are: $(A,1),(B,2),(C,3),(D,4)$ and $(E,1234)$. The columns that are not to be assigned any factors are 12, 34, 124 and 123. The other factors are assigned to the remaining columns.

Case v: Interactions AB, AC, AD, AE. The assignments are: $(A,1),(B,2),(C,3),(D,4)$ and $(E,23)$. The columns that are not to be assigned any factors are 12, 13, 14 and 123. The other factors are assigned to the remaining columns.

Case vi: Interactions AB, BC, CD, DA. The assignments are: $(A,1),(B,2),(C,3)$ and $(D,4)$. The columns that are not to be assigned any factors are 12, 23, 34 and 14. The other factors are assigned to the remaining columns.

For $t = 5$, we come across some cases where inconsistent equations arise which do not allow one to solve for possible solutions. Again, for illustration, we consider a few cases.

Case i: Interactions AB, CD, EF, GH, IJ. The assignments are: $(A,1),(B,2),(C,3),(D,4),(E,13),(F,24),(G,14),(H,123),(I,23)$ and $(J,124)$. The columns that are not to be assigned any factors are 12, 34, 1234, 234 and 134.

Case ii: Interactions AB, CD, EF, GH, GI. Let $z_{i_1} + z_{i_2} = z_{i_{10}} \pmod 2$, $z_{i_3} + z_{i_4} = z_{i_{11}} \pmod 2$, $z_{i_5} + z_{i_6} = z_{i_{12}} \pmod 2$, $z_{i_7} + z_{i_8} = z_{i_{13}} \pmod 2$, $z_{i_7} + z_{i_9} = z_{i_{14}} \pmod 2$. Then it follows that

$$z_{i_1} + z_{i_2} + z_{i_{10}} = 0 \pmod 2, \quad z_{i_3} + z_{i_4} + z_{i_{11}} = 0 \pmod 2, \quad z_{i_5} + z_{i_6} + z_{i_{12}} = 0 \pmod 2,$$

$$z_{i_7} + z_{i_8} + z_{i_{13}} = 0 \pmod 2, \quad z_{i_7} + z_{i_9} + z_{i_{14}} = 0 \pmod 2.$$

Now we need to identify distinct values of $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}, z_{i_5}, z_{i_6}, z_{i_7}, z_{i_8}, z_{i_9}, z_{i_{10}}, z_{i_{11}}, z_{i_{12}}, z_{i_{13}}, z_{i_{14}}$ which satisfy the above five equations. However, note that the above equations imply that $z_{i_7} + z_{i_{15}} = 0$. This is impossible. Hence, there exists no allocations of the columns to the factors which would enable one to have such an interaction set in the model.

Case iii: Interactions AB, CD, EF, EG, EH. Let $z_{i_1} + z_{i_2} = z_{i_9} \pmod 2$, $z_{i_3} + z_{i_4} = z_{i_{10}} \pmod 2$, $z_{i_5} + z_{i_6} = z_{i_{11}} \pmod 2$, $z_{i_5} + z_{i_7} = z_{i_{12}} \pmod 2$, $z_{i_5} + z_{i_8} = z_{i_{13}} \pmod 2$. Then it follows that

$$z_{i_1} + z_{i_2} + z_{i_9} = 0 \pmod 2, \quad z_{i_3} + z_{i_4} + z_{i_{10}} = 0 \pmod 2, \quad z_{i_5} + z_{i_6} + z_{i_{11}} = 0 \pmod 2,$$

$$z_{i_5} + z_{i_7} + z_{i_{12}} = 0 \pmod 2, \quad z_{i_5} + z_{i_8} + z_{i_{13}} = 0 \pmod 2.$$

Now we need to identify distinct values of $z_{i_1}, z_{i_2}, z_{i_3}, z_{i_4}, z_{i_5}, z_{i_6}, z_{i_7}, z_{i_8}, z_{i_9}, z_{i_{10}}, z_{i_{11}}, z_{i_{12}}, z_{i_{13}}$ which satisfy the above five equations. However, note that the above equations imply that $z_{i_{14}} + z_{i_{15}} = 0$. This is also impossible. Hence, here too there exists no allocations of the columns to the factors which would enable one to have such an interaction set in the model.

Case iv: Interactions AB, CD, CE, CF, CG. The assignments are: $(A,2),(B,3),(C,1),(D,4),(E,24),(F,34)$ and $(G,234)$. The columns that are not to be assigned any factors are 23, 14, 124, 134 and 1234. The other $k - 7$ factors are assigned to the remaining columns.

The above ideas can, in principle, be extended to arrays with larger number of rows (runs), though the effort involved in larger arrays is obviously more. In general, an easy way to check the non-existence of an allocation of the factors to the columns (leading to universally optimal designs) is to add (mod 2) all the equations that arise and then use the fact that $\sum_{j=1}^{2^r-1} z_{i_j} = 0$. Non-existence is implied in case this leads to one z_{i_j} or sum of two z_{i_j}'s being equal to zero.

Acknowledgments

The authors thank a referee for the useful comments.

References

Dey, A., Mukerjee, R., 1999a. Fractional Factorial Plans. Wiley, New York.

Dey, A., Mukerjee, R., 1999b. Inter-effect orthogonality and optimality in hierarchical models. Sankhyā Ser. B 61, 460–468.

Dey, A., Suen, C., 2002. Optimal fractional factorial plans for main effects and specified two-factor interactions: a projective geometric approach. Ann. Statist. 30, 1512–1523.

Hedayat, A.S., Pesotan, H., 1992. Two-level factorial designs for main effects and selected two factor interactions. Statist. Sinica 2, 453–464.

Hedayat, A.S., Pesotan, H., 1997. Designs for two-level factorial experiments with linear models containing main effects and selected two-factor interactions. J. Statist. Plann. Inference 64, 109–124.

Hedayat, A.S., Sloane, N.J.A., Stufken, J., 1999. Orthogonal Arrays: Theory and Applications. Springer, New York.

Kiefer, J., 1975. Construction and optimality of generalized Youden designs. In: Srivastava, J.N. (Ed.), A Survey of Statistical Designs and Linear Models. North-Holland, Amsterdam, pp. 333–353.

Sinha, B.K., Mukerjee, R., 1982. A note on the universal optimality criterion for full rank models. J. Statist. Plann. Inference 7, 97–100.

Wu, C.F.J., Chen, Y., 1992. A graph-aided method for planning two-level experiments when certain interactions are important. Technometrics 34, 162–175.