Statistica Sinica **12**(2002), 905-916

QUASI-ORTHOGONAL ARRAYS AND OPTIMAL FRACTIONAL FACTORIAL PLANS

Kashinath Chatterjee*, Ashish Das and Aloke Dey

**Asutosh College, Calcutta and Indian Statistical Institute, New Delhi*

Abstract: Generalizing orthogonal arrays, a new class of arrays called quasi-orthogonal arrays, are introduced and it is shown that fractional factorial plans represented by these arrays are universally optimal under a wide class of models. Some methods of construction of quasi-orthogonal arrays are also described.

Key words and phrases: Fractional factorial plan, quasi-orthogonal array, universal optimality.

1. Introduction

The study of optimal fractional factorial plans is of considerable recent interest; see Dey and Mukerjee (1999a, Chapters 2 and 6) for a review. Most of these results however, relate to situations where all factorial effects involving the same number of factors are considered equally important and, as such, the underlying model involves the general mean and all effects involving up to a specified number of factors. The presumption of equality in the importance of all factorial effects involving the same number of factors may not be tenable in many practical situations. For example, one might be interested in estimating the general mean, all main effects and only a subset of two-factor interactions. The issue of estimability and optimality in situations of this kind in the context of two-level factorials has been addressed by Hedayat and Pesotan (1992, 1997) and Chiu and John (1998). For some other related work, see Hedayat (1990), Wu and Chen (1992), Sun and Wu (1994) and, Dey and Mukerjee (1999b).

In this paper, further results on the optimality of fractional factorial plans for arbitrary factorials are obtained. In Section 2, some preliminaries are introduced. Generalizing orthogonal arrays, a new class of arrays called *quasi-orthogonal arrays* are introduced in Section 3. It is shown that fractional factorial plans represented by quasi-orthogonal arrays are universally optimal (and hence, in particular A-, D- and E-optimal) under several alternative models. As such, fractional factorial plans represented by quasi-orthogonal arrays exhibit a kind of model robustness in the sense that the same plan remains optimal under two or more rival models.

There is also a practical motivation for studying quasi-orthogonal arrays, as these can lead to useful plans for designing experiments for quality improvement. In a production line, the quality of a product depends on two types of factors, called *control* and *noise* factors. The control factors are those that can be set at specified levels during the production process, while the noise factors can be fixed at selected levels during the experiment but *not* during the production or, later use of the product. It is often desired to plan an experiment to study the main effects of the control and noise factors plus the control versus noise two-factor interactions (see e.g., Shoemaker, Tsui and Wu (1991)). In such a situation, one desires an optimal plan under a model that includes the mean, all main effects and a chosen subset of two-factor interactions, all other factorial effects being assumed absent. Quasi-orthogonal arrays can provide such optimal plans. Some methods of construction of quasi-orthogonal arrays are given in Section 4.

2. Preliminaries

Throughout this paper, we closely follow the notations and terminology used in Dey and Mukerjee (1999a). Consider the set up of an $m_1 \times \cdots \times m_n$ factorial experiment involving n factors $F_1, \ldots, F_n$ appearing at $m_1, \ldots, m_n$ (≥ 2) levels respectively. The $v = \prod_{i=1}^{n} m_i$ treatment combinations are represented by ordered n-tuples $j_1 \ldots j_n$, $0 \leq j_i \leq m_i - 1$, $1 \leq i \leq n$. Let $\boldsymbol{\tau}$ denote the $v \times 1$ vector with elements $\tau(j_1 \ldots j_n)$ arranged in the lexicographic order, where $\tau(j_i \ldots j_n)$ is the fixed effect of the treatment combination $j_1 \ldots j_n$. Also, let Ω denote the set of all binary $n-$tuples. For each $\boldsymbol{x} = x_1 \ldots x_n \in \Omega$, define $\gamma(\boldsymbol{x}) = \{i \ : x_i = 1\}$ and $\alpha(\boldsymbol{x}) = \prod_{i=1}^{n} (m_i - 1)^{x_i}$. Furthermore, for any subset Ω_s of Ω, define $\alpha_s = \sum_{\boldsymbol{x} \in \Omega_s} \alpha(\boldsymbol{x})$. Note that $\sum_{\boldsymbol{x} \in \Omega} \alpha(\boldsymbol{x}) = v$. We denote the $a \times 1$ vector of all ones by $\mathbf{1}_a$, the identity matrix of order a by I_a and a generalized inverse of a matrix A by A^-. For $1 \leq i \leq n$, let P_i be an $(m_i - 1) \times m_i$ matrix such that the $m_i \times m_i$ matrix ${m_i}^{-\frac{1}{2}} \mathbf{1}_{m_i}, P_i'$ is orthogonal. For each $\boldsymbol{x} = x_1 \ldots x_n \in \Omega$, let the $\alpha(\boldsymbol{x}) \times v$ matrix $P^{\boldsymbol{x}}$ be defined as

$$P^{\boldsymbol{x}} = {P_1}^{x_1} \otimes \cdots \otimes {P_n}^{x_n}, \tag{2.1}$$

where for $1 \leq i \leq n$,

$$P_i^{x_i} = \begin{cases} {m_i}^{-\frac{1}{2}} \mathbf{1'}_{m_i} & \text{if } x_i = 0 \\ P_i & \text{if } x_i = 1, \end{cases} \tag{2.2}$$

and $\otimes$ denotes the Kronecker product. Then for each $\boldsymbol{x} = x_1 \ldots x_n \in \Omega$, $\boldsymbol{x} \neq 00 \ldots 0$, the elements of $P^{\boldsymbol{x}} \boldsymbol{\tau}$ represent a complete set of orthonormal contrasts belonging to the factorial effect ${F_1}^{x_1} \cdots {F_n}^{x_n} \equiv F^{\boldsymbol{x}}$, say; *cf.* Gupta and Mukerjee (1989). Also $P^{00\ldots0} \boldsymbol{\tau} = v^{\frac{1}{2}} \bar{\tau}$, where $\bar{\tau}$ is the general mean, and in this sense the general mean will be represented by $F^{00\ldots0}$.

Let Ω_t be a subset of Ω, containing $00\ldots0$. We consider a model which is such that $F^{\boldsymbol{x}}$ is included in the model if and only if $\boldsymbol{x} \in \Omega_t$, the effects not included in the model being assumed negligible. Then,

$$P^{\boldsymbol{x}}\boldsymbol{\tau} = \mathbf{0}, \text{ for each } \boldsymbol{x} \in \bar{\Omega}_t, \tag{2.3}$$

where $\bar{\Omega}_t = \Omega - \Omega_t$ and $\mathbf{0}$ is a null vector. Under (2.3), let

$$\boldsymbol{\beta}_{\boldsymbol{x}} = P^{\boldsymbol{x}}\boldsymbol{\tau}, \text{ for each } \boldsymbol{x} \in \Omega_t, \tag{2.4}$$

which provides an interpretation for $\boldsymbol{\beta}_{\boldsymbol{x}}$ $(\boldsymbol{x} \in \Omega_t,\ \boldsymbol{x} \neq 00\ldots0)$ in terms of a complete set of orthonormal contrasts belonging to a possibly non-negligible factorial effect $F^{\boldsymbol{x}}$. Also, $\beta_{00\ldots0}$ can be interpreted in terms of the general mean.

Under the absence of the factorial effects $F^{\boldsymbol{x}}$, $\boldsymbol{x} \in \bar{\Omega}_t$, suppose that interest lies in estimating the factorial effects $F^{\boldsymbol{x}}$, $\boldsymbol{x} \in \Omega_f$, where Ω_f is a subset of Ω_t containing $00\ldots0$. If Ω_f is a proper subset of Ω_t, then the factorial effects included in $\Omega_t - \Omega_f$ are treated as nuisance parameters. By (2.4), the objects of interest are then $\boldsymbol{\beta}_{\boldsymbol{x}}$ for $\boldsymbol{x} \in \Omega_f(\subset \Omega_t)$, namely interest lies in

$$\boldsymbol{\beta}^{(1)} = (\cdots, \boldsymbol{\beta}_{\boldsymbol{x}}', \cdots)'_{\boldsymbol{x}\in\Omega_f}. \tag{2.5}$$

Define

$$P^{(1)} = (\cdots, (P^{\boldsymbol{x}})', \cdots)'_{\boldsymbol{x}\in\Omega_f} \tag{2.6}$$

and, if Ω_f is a proper subset of Ω_t, let

$$\boldsymbol{\beta}^{(2)} = (\cdots, \boldsymbol{\beta}_{\boldsymbol{x}}', \cdots)'_{\boldsymbol{x}\in\Omega_t-\Omega_f},\ \ P^{(2)} = (\cdots, (P^{\boldsymbol{x}})', \cdots)'_{\boldsymbol{x}\in\Omega_t-\Omega_f}. \tag{2.7}$$

Clearly, $\boldsymbol{\beta}^{(1)}$ and $\boldsymbol{\beta}^{(2)}$ are column vectors of orders α_f and $\alpha_t - \alpha_f$ respectively; similarly $P^{(1)}$ and $P^{(2)}$ are matrices of orders $\alpha_f \times v$ and $(\alpha_t - \alpha_f) \times v$ respectively.

Consider an N-run fractional factorial plan for an $m_1 \times \cdots \times m_n$ factorial and, as before, let the effects included in the model be $F^{\boldsymbol{x}}$, $\boldsymbol{x} \in \Omega_t$, where Ω_t is a fixed subset of Ω containing $00\ldots0$. For given N, $0 < N < v$, let $\mathcal{D}_N$ be the class of all N-run plans $\{d\}$ such that, under any $d \in \mathcal{D}_N$, each of the factorial effects represented by $F^{\boldsymbol{x}}$, $\boldsymbol{x} \in \Omega_f(\subset \Omega_t)$ is estimable. Let R_d be a $v \times v$ diagonal matrix with diagonal elements representing, in the lexicographic order, the replication numbers of the v treatment combinations in d. The parametric functions of interest are $P^{(1)}\boldsymbol{\tau}$, where $P^{(1)}$ is as defined in (2.6). Assuming that the observations are homoscedastic and uncorrelated, the information matrix for $P^{(1)}\boldsymbol{\tau}$, under d, following Dey and Mukerjee (1999a), is given by

$$\mathcal{I}_d = P^{(1)}R_d(P^{(1)})' - P^{(1)}R_d(P^{(2)})'(P^{(2)}R_d(P^{(2)})')^{-}P^{(2)}R_d(P^{(1)})', \tag{2.8}$$

where the second term does not arise if $\Omega_f = \Omega_t$.

Also, for any $d \in \mathcal{D}_N$, it can be shown that

$$\mathrm{tr}(\mathcal{I}_d) \leq (N/v)\alpha_f, \tag{2.9}$$

where $\mathrm{tr}(\cdot)$ denotes trace. Furthermore, if $\Omega_f = \Omega_t$, then for any $d \in \mathcal{D}_N$,

$$\mathrm{tr}(\mathcal{I}_d) = (N/v)\alpha_f. \tag{2.10}$$

3. Quasi-Orthogonal Arrays and Optimality

For an arbitrary subset T of Ω, define $T^* = \{\boldsymbol{x} \ : \boldsymbol{x} \in T$, there does not exist $\boldsymbol{y} \in T,\ \boldsymbol{y} \neq \boldsymbol{x}$, such that $\boldsymbol{x} \leq \boldsymbol{y}\}$, where $\boldsymbol{x} \leq \boldsymbol{y}$ means $x_i \leq y_i,\ 1 \leq i \leq n$. Henceforth, for a $T \subset \Omega$, we call T^* the *reduced set* of T. Note that for a given T^*, there might exist more than one T giving rise to it. For example, let $n = 3$, $T_1 = \{000, 100, 010, 001, 110, 101\}$ and $T_2 = \{000, 010, 110, 101\}$. For both T_1 and T_2, $T^* = \{110, 101\}$.

Definition. Let T^* be the reduced set of a given subset of Ω. A quasi-orthogonal array $QOA(N, n, m_1 \times \cdots \times m_n, T^*)$, having N rows and $n(\geq 2)$ columns, say $A_1, A_2, \ldots, A_n$, is an $N \times n$ array with elements in the ith column A_i having $m_i(\geq 2)$ distinct symbols for $i = 1, \ldots, n$, such that for every $\boldsymbol{x} \in T^*$, all combinations of symbols corresponding to the columns $\{A_i \ : i \in \gamma(\boldsymbol{x})\}$ appear equally often as a row.

Note that an orthogonal array $OA(N, n, m_1 \times \cdots \times m_n, g)$ is a quasi-orthogonal array $QOA(N, n, m_1 \times \cdots \times m_n, T^*)$ with $T^* = \{\boldsymbol{x} : |\gamma(\boldsymbol{x})| = g\}$ where, for a set W, $|W|$ denotes its cardinality. In this sense, one may regard quasi-orthogonal arrays as a generalization of orthogonal arrays. For a definition of an $OA(N, n, m_1 \times \cdots \times m_n, g)$, see Hedayat, Sloane and Stufken (1999). Following standard terminology, we denote a *symmetric* orthogonal array (i.e., when $m_1 = \cdots = m_n = m$, say) by $OA(N, n, m, g)$.

Another generalization of orthogonal arrays are the compound orthogonal arrays considered by Rosenbaum (1994) and Hedayat et al. (1999); for a definition of compound orthogonal arrays, see e.g., Hedayat et al. (1999, p.230). While all compound orthogonal arrays are necessarily quasi-orthogonal arrays, every quasi-orthogonal array need not be a compound orthogonal array as seen in Example 1 below.

Example 1. Let $N = 16,\ n = 7,\ m_1 = 4,\ m_2 = m_3 = \cdots = m_7 = 2$. Further, let $T^* = \{A_1 \cup A_2\}$, where $A_1 = \{1000100, 1000010, 1000001\}$ and $A_2 = \{\boldsymbol{x} : |\gamma(\boldsymbol{x})| = 3\} - \{1000110, 1000101, 1000011, 1100100, 1100010, 1100001, 1010100, 1010010, 1010001, 1001100, 1001010, 1001001\}$, $\boldsymbol{x}$ denoting a binary 7-tuple.

The following array (in transposed form) is a $QOA(16, 7, 4 \times 2^6, T^*)$, where T^* is as given above:

$$\begin{bmatrix} 0000 & 1111 & 2222 & 3333 \\ 0101 & 1010 & 1010 & 0101 \\ 0011 & 1100 & 1100 & 0011 \\ 0110 & 1001 & 1001 & 0110 \\ 1010 & 0101 & 1010 & 0101 \\ 1100 & 0011 & 1100 & 0011 \\ 1001 & 0110 & 1001 & 0110 \end{bmatrix}'.$$

It may be noted that this array is not a compound orthogonal array.

Clearly, the rows of a quasi-orthogonal array $QOA(N, n, m_1 \times \cdots \times m_n, T^*)$ can be identified with the treatment combinations of an $m_1 \times \cdots \times m_n$ factorial set up, and the array itself can be regarded as an N-run fractional factorial plan for such a factorial. For instance, the array in Example 1 represents a 16-run plan for a 4×2^6 factorial.

As before, let $\Omega_f \subset \Omega_t \subset \Omega$ such that $00\ldots0 \in \Omega_f$. Furthermore, for given Ω_f, Ω_t, define $S = \{\boldsymbol{x}_i \vee \boldsymbol{x}_j \ : \ \boldsymbol{x}_i \in \Omega_f,\ \boldsymbol{x}_j \in \Omega_t\}$ where, for $\boldsymbol{u} = u_1 \ldots u_n \in \Omega$ and $\boldsymbol{w} = w_1 \ldots w_n \in \Omega$, $\boldsymbol{u} \vee \boldsymbol{w} = z_1 \ldots z_n$ with $z_t = \max(u_t, w_t)$, $1 \le t \le n$. Let S^* be the reduced set of S.

Let $\mathcal{M}(\Omega_f, \Omega_t)$ denote a linear model in which a factorial effect $F^{\boldsymbol{x}}$ is included if and only if $\boldsymbol{x} \in \Omega_t$, and suppose interest lies in estimating all the factorial effects in Ω_f, where $\Omega_f \subset \Omega_t \subset \Omega$. With reference to the chosen Ω_f, Ω_t, let S be the set defined above and S^* the reduced set of S.

Theorem 1. *Under the model $\mathcal{M}(\Omega_f, \Omega_t)$, let $d_0 \in \mathcal{D}_N$ be represented by a $QOA(N, n, m_1 \times \cdots \times m_n, S^*)$. Then d_0 is a universally optimal plan for estimating complete sets of orthonormal contrasts belonging to the factorial effects $F^{\boldsymbol{x}}$, $\boldsymbol{x} \in \Omega_f$.*

Proof. Let $\Omega_t (\subset \Omega)$ and $\Omega_f (\subset \Omega_t)$ be as specified by the model $\mathcal{M}(\Omega_f, \Omega_t)$. Then clearly, $\Omega_f \subset \Omega_t \subset S$ and $\boldsymbol{x} \vee \boldsymbol{y} \in S$ for every $\boldsymbol{x} \in \Omega_f$ and $\boldsymbol{y} \in \Omega_t$. Following the line of proof of Lemma 2.6.1 in Dey and Mukerjee (1999a, p.25) it can be shown that (a) for each $\boldsymbol{x} \in S$, $P^{\boldsymbol{x}} R_{d_0} (P^{\boldsymbol{x}})' = (N/v) I_{\alpha(\boldsymbol{x})}$, and (b) for each $\boldsymbol{x}$, $\boldsymbol{y} \in S$, such that $\boldsymbol{x} \vee \boldsymbol{y} \in S$ and $\boldsymbol{x} \neq \boldsymbol{y}$, $P^{\boldsymbol{x}} R_{d_0} (P^{\boldsymbol{y}})' = O$, where O is a null matrix. We thus have $P^{\boldsymbol{x}} R_{d_0} (P^{\boldsymbol{x}})' = (N/v) I_{\alpha(\boldsymbol{x})}$ for each $\boldsymbol{x} \in \Omega_f$, and $P^{\boldsymbol{x}} R_{d_0} (P^{\boldsymbol{y}})' = O$ for each $\boldsymbol{x} \in \Omega_f$, $\boldsymbol{y} \in \Omega_t, \boldsymbol{x} \neq \boldsymbol{y}$. It follows that $P^{(1)} R_{d_0} (P^{(1)})' = (N/v) I_{\alpha_f}$, $P^{(1)} R_{d_0} (P^{(2)})' = O$, so that by (2.8),

$$\mathcal{I}_{d_0} = (N/v) I_{\alpha_f}. \tag{3.1}$$

Now from (2.9) and (3.1), following Kiefer (1975) and Sinha and Mukerjee (1982), the claimed universal optimality of d_0 is established.

Fractional factorial plans represented by quasi-orthogonal arrays can be optimal under two or more rival models, thus exhibiting a kind of model robustness. The following example illustrates this fact.

Example 2. Consider a 3×2^4 experiment and let the factors be F_1 at three levels and F_i, $2 \leq i \leq 5$, each at two levels. Suppose we are interested in estimating the mean, all main effects and all two-factor interactions except the interactions F_2F_5, F_3F_5 and F_4F_5. The interactions F_2F_5, F_3F_5, F_4F_5 and all interactions involving three or more factors are assumed to be absent. Then $\Omega_f = \{00000,\ 10000,\ 01000,\ 00100,\ 00010,\ 00001,\ 11000,\ 10100,\ 10010,\ 10001,\ 01100,\ 01010,\ 00110\}$, $\Omega_t = \Omega_f$, $S = \Omega - \{01111, 11111\}$, $S^* = \{10111,\ 11011,\ 11101,\ 11110\}$. For this S^*, we have a quasi-orthogonal array $QOA(24, 5, 3 \times 2^4, S^*)$ as shown (in transposed form) below .

$$\begin{bmatrix} 00000000 & 11111111 & 22222222 \\ 00001111 & 00001111 & 00001111 \\ 00111001 & 00111001 & 00111001 \\ 01010101 & 01010101 & 01010101 \\ 01100011 & 01100011 & 01100011 \end{bmatrix}' .$$

A 24-run plan represented by this quasi-orthogonal array is therefore universally optimal for the estimation of all the factorial effects in the considered model $\mathcal{M}(\Omega_f, \Omega_f)$.

Now suppose that all two and three factor interactions *not* involving F_1 are absent, along with all interactions involving four or more factors. The three-factor interactions involving F_1 are included in the model, though we are not interested in estimating them. The factorial effects of interest (that are to be estimated) are the mean, all main effects and all two-factor interactions among F_1 and F_j, $2 \leq j \leq 5$. Then $\Omega_f = \{00000, 10000, 01000, 00100, 00010, 00001, 11000, 10100, 10010, 10001\}$, $\Omega_t - \Omega_f = \{11100, 11010, 11001, 10110, 10101, 10011\}$, $S = \Omega - \{00111, 01011, 01101, 01110, 01111, 11111\}$, $S^* = \{10111, 11011, 11101, 11110\}$. This S^* is the same as the one given above. Thus a quasi-orthogonal array $QOA(24, 5, 3 \times 2^4, S^*)$ as given above represents a 24-run universally optimal plan for the estimation of all the factorial effects in the considered model $\mathcal{M}(\Omega_f, \Omega_t)$.

Consider the model $\mathcal{M}(\Omega_f, \Omega_f)$ and, with reference to this model, let S and S^* be the sets defined earlier. Then, in analogy with the Rao's bound for the number of rows of an orthogonal array (see e.g., Dey and Mukerjee (1999a, p.28), one can show that a necessary condition for the existence of a $QOA(N, n, m_1 \times \cdots \times m_n, S^*)$ is that

$$N \geq \alpha_f. \tag{3.2}$$

Any N-run plan represented by a quasi-orthogonal array for which the equality in (3.2) holds is saturated. Recall that orthogonal arrays for which the number of rows attains the Rao's bounds are called *tight*.

4. Construction of Quasi-Orthogonal Arrays

In this section, some methods of construction of quasi-orthogonal arrays are discussed. Theorems 2 and 3 give methods of constructing 2-symbol symmetric quasi-orthogonal arrays, while Theorem 4 gives a method of construction of asymmetric quasi-orthogonal arrays. Some comments about the use of such arrays in obtaining universally optimal fractional factorial plans under different models are also made.

We need the following preliminaries. A positive integer u is called a *Hadamard number* if a Hadamard matrix of order u, H_u, exists. Throughout, unless stated otherwise, the trivial Hadamard number $u = 1$ will be left out of consideration. Without loss of generality, the first column of H_u will be taken as $\mathbf{1}_u$. A set of three distinct columns of H_u, $u \geq 4$, will be said to have the *Hadamard property* if the Hadamard product of any two columns in the set equals the third, where the Hadamard product of two vectors $\boldsymbol{a} = (a_1, \ldots, a_n)$ and $\boldsymbol{b} = (b_1, \ldots, b_n)$ is defined as $\boldsymbol{a} * \boldsymbol{b} = (a_1 b_1, \ldots, a_n b_n)$.

Theorem 2. *If u, w are Hadamard numbers such that $4 \leq w \leq u$, then there exists a quasi-orthogonal array $QOA(uw, w(u-1), 2^{w(u-1)}, T^*)$ with T^* a set whose elements are the rows of the matrix*

$$\begin{pmatrix} \mathbf{1}_2' \otimes A_{(w-1)(2)} & O \\ \mathbf{1}_2' \otimes I_{w-1} \otimes \mathbf{1}_m & \mathbf{1}_{w-1} \otimes I_m \\ O & A_{(m)(2)} \end{pmatrix},$$

where $A_{(s)(t)}$ is an $\binom{s}{t} \times s$ matrix whose rows are all possible binary s-tuples with exactly t unities, O is a null matrix of appropriate order, and $m = w(u-3)+2$.

Proof. Let $H_u = [\mathbf{1}_u \vdots \boldsymbol{a}_1 \vdots \cdots \vdots \boldsymbol{a}_{u-1}]$, $H_w = [\mathbf{1}_w \vdots \boldsymbol{b}_1 \vdots \cdots \vdots \boldsymbol{b}_{w-1}]$. Consider $H_{uw} = H_u \otimes H_w$. It is not hard to see that H_{uw} contains $w-1$ disjoint sets of columns given by $\{\mathbf{1}_u \otimes \boldsymbol{b}_i, \boldsymbol{a}_i \otimes \mathbf{1}_w, \boldsymbol{a}_i \otimes \boldsymbol{b}_i\}$ $1 \leq i \leq w-1$, each set having the Hadamard property. For each $1 \leq i \leq w-1$, identify the column $\mathbf{1}_u \otimes \boldsymbol{b}_i$ by a two-level factor F_i, the column $\boldsymbol{a}_i \otimes \mathbf{1}_w$ by a two-level factor F_{w-1+i} and delete the columns $\boldsymbol{a}_i \otimes \boldsymbol{b}_i$ and the column $\mathbf{1}_u \otimes \mathbf{1}_w$ in H_{uw}. Further, identify the remaining columns of H_{uw} by the two-level factors $F_{2w-1}, \ldots, F_{w(u-1)}$. Then the $uw \times w(u-1)$ matrix represents the required quasi-orthogonal array.

Consider a two-level factorial experiment involving $w(u-1)$ factors which can be grouped into three sets of factors : the first group having factors $F_1, \ldots, F_{w-1}$;

the second group having factors $F_w, \ldots, F_{2(w-1)}$; the third group having the factors $F_j,\ 2w-1 \le j \le w(u-1)$. Suppose it is desired to estimate the mean, all main effects and all two-factor interactions of type $F_i F_{w-1+i},\ 1 \le i \le w-1$. All other factorial effects are assumed negligible. Then the quasi-orthogonal array of Theorem 2 provides a universally optimal saturated plan for the above experiment under the stated model.

Theorem 3. *If u, w are Hadamard numbers such that $4 \le w \le u$, then there exists a quasi-orthogonal array $QOA(uw, w(u-w+2)-2, 2^{w(u-w+2)-2}, T^*)$ with T^* a set whose elements are the rows of the matrix*

$$\begin{pmatrix} A_{(w-1)(2)} \otimes \mathbf{1}_{\binom{w-1}{2}} & \mathbf{1}_{\binom{w-1}{2}} \otimes A_{(w-1)(2)} & O \\ I_{w-1} \otimes \mathbf{1}_{w-1} \otimes \mathbf{1}_m & \mathbf{1}_{w-1} \otimes I_{w-1} \otimes \mathbf{1}_m & \mathbf{1}_{w-1} \otimes \mathbf{1}_{w-1} \otimes I_m \\ O & O & A_{(m)(2)} \end{pmatrix},$$

where $m = w(u-w)$.

Proof. Following the proof of Theorem 2, for each $1 \le i \le w-1$ and $1 \le j \le w-1$, identify the column $\mathbf{1}_u \otimes \boldsymbol{b}_i$ by a two-level factor F_i, the column $\boldsymbol{a}_j \otimes \mathbf{1}_w$ by a two-level factor F_{w-1+j} and delete the columns $\boldsymbol{a}_j \otimes \boldsymbol{b}_i$ in H_{uw}. Also, delete the column $\mathbf{1}_u \otimes \mathbf{1}_w$. Further, identify the remaining columns of H_{uw} by the two-level factors $F_{2w-1}, \ldots, F_{w(u-w+2)-2}$. Then the $uw \times w(u-w+2)-2$ matrix represents the required quasi-orthogonal array.

Consider a two-level factorial experiment involving $w(u-w+2)-2$ factors which can be grouped into three sets of factors : the first group having factors $F_1, \ldots, F_{w-1}$; the second group having factors $F_w, \ldots, F_{2(w-1)}$; the third group having the factors $F_j,\ 2w-1 \le j \le w(u-w+2)-2$. Suppose it is desired to estimate the mean, all main effects and all two-factor interactions of the type $F_i F_j,\ 1 \le i \le w-1,\ w \le j \le 2(w-1)$. All other factorial effects are assumed negligible. Then the quasi-orthogonal array of Theorem 3 provides a universally optimal saturated plan for the above experiment under the stated model.

Theorem 4. *Let there exist orthogonal arrays $OA(N, n-r+1, m \times m_{r+1} \times \cdots \times m_n, g_1)$ and $OA(m, r, m_1 \times \cdots \times m_r, g_2)$, where $g_1 = 2s_1 + i$ and $g_2 = 2s_2 + j$, s_1, s_2 being positive integers, $i, j = 0, 1$. Then there exists a quasi-orthogonal array $QOA(N, n, m_1 \times \cdots \times m_r \times m_{r+1} \times \cdots \times m_n, T^*)$ with T^* a set whose elements are the rows of the matrix*

$$\begin{pmatrix} \mathbf{1}_{\binom{n-r}{g_1-1}} \otimes A_{(r)(g_2)} & A_{(n-r)(g_1-1)} \otimes \mathbf{1}_{\binom{r}{g_2}} \\ O & A_{(n-r)(g_1)} \end{pmatrix}.$$

Proof. Replacing the m symbols in the first column of the $OA(N, n-r+1, m \times m_{r+1} \times \cdots \times m_n, g_1)$ by the rows of the array $OA(m, r, m_1 \times \cdots \times m_r, g_2)$, we get the required quasi-orthogonal array.

Fractional factorial plans represented by the quasi-orthogonal arrays of Theorem 4 involve n factors, say $G_1, \ldots, G_r$ and $F_1, \ldots, F_{n-r}$, where for $1 \le k \le r$, G_k appears at m_k levels, while for $1 \le l \le n-r$, F_l appears at m_{r+l} levels. These plans are universally optimal under a model $\mathcal{M}_1$, where the model $\mathcal{M}_1$ is such that Ω_f contains binary n-tuples corresponding to (i) the mean; (ii) all effects involving at most s_2 (r) factors among the first r factors, if $g_2 < r$ $(g_2 = r)$; (iii) all effects involving at most s_1 factors among the last $(n-r)$ factors; (iv) all interactions involving at most s_2 factors among the first r factors and at most (s_1+i-1) factors among the last $(n-r)$ factors for $i = 0, 1$, if $g_1 > 2$ and $g_2 < r$; (iv$'$) all interactions involving at most r factors among the first r factors and at most (s_1+i-1) factors among the last $(n-r)$ factors for $i = 0, 1$, if $g_1 > 2$ and $g_2 = r$.

Also, if $j = 1$, the set $\Omega_t - \Omega_f$ contains binary n-tuples corresponding to (v) (s_2+1)-factor interactions among the first r factors; (vi) all interactions involving (s_2+1) factors among first r factors and at most (s_1+i-1) factors among the last $(n-r)$ factors, for $i = 0, 1$, if $g_1 > 2$. If $j \neq 1$, then $\mathcal{M}_1$ is such that $\Omega_t - \Omega_f$ is null i.e., $\mathcal{M}_1 \equiv \mathcal{M}_1(\Omega_f, \Omega_f)$.

Under the model $\mathcal{M}_1$, the plan represented by the quasi-orthogonal array of Theorem 4 is universally optimal. In contrast to the model $\mathcal{M}_1$, consider another model $\mathcal{M}_2$ which is such that Ω_f contains binary n-tuples corresponding to (i) the mean; (ii) all effects involving at most s_2 factors among the first r factors; (iii) all effects involving at most s_1 factors among the last $(n-r)$ factors; (iv) all interactions involving at most s_2 factors among the first r factors and at most (s_1-1) factors among the last $(n-r)$ factors, if $g_1 > 3$.

Also, if $(i, j) \neq (0, 0)$, the set $\Omega_t - \Omega_f$ contains binary n-tuples corresponding to (v) (s_2+1)-factor interactions among the first r factors, if $j = 1$; (vi) (s_1+1)-factor interactions among the last $(n-r)$ factors, if $i = 1$; (vii) all interactions involving (s_2+1) factors among first r factors and at most (s_1+i-1) factors among the last $(n-r)$ factors $(i = 0, 1)$, if $j = 1$ and $g_1 > 2$; (viii) all interactions involving at most s_2 factors of the first r factors and s_1 factors among the last $(n-r)$ factors if $i = 1$.

Then a fractional factorial plan represented by the quasi-orthogonal array of Theorem 4 is universally optimal under $\mathcal{M}_2$ as well.

Example 3. Consider the orthogonal arrays $OA(32, 4, 8 \times 2^3, 3)$ and $OA(8, 5, 4 \times 2^4, 2)$. Then we have $s_1 = 1, i = 1$ and $s_2 = 1, j = 0$. Replacing the 8 symbols in the first column of the first array by the rows of the second array, one gets an array with 32 rows and 8 columns, the first column having four symbols and the remaining columns having two symbols each. This array is a quasi-orthogonal array $QOA(32, 8, 4 \times 2^7, T^*)$, where the elements of T^* are the rows of the following matrix:

$$\begin{pmatrix} \mathbf{1}_3 \otimes A_{(5)(2)} & A_{(3)(2)} \otimes \mathbf{1}_{10} \\ O & \mathbf{1}'_3 \end{pmatrix}.$$

The 32-run plan represented by this quasi-orthogonal array is universally optimal under a model $\mathcal{M}_1$ that includes the mean, all main effects and all two-factor interactions involving any one of the first five factors and any one of the last three factors, assuming that all other factorial effects not included in the model are absent. Note that this plan is saturated under $\mathcal{M}_1$. The plan is also universally optimal under a model $\mathcal{M}_2$ which is such that Ω_f contains binary 8-tuples corresponding to the mean and all main effects, and $\Omega_t - \Omega_f$ contains binary 8-tuples corresponding to all two-factor interactions among the last three factors and all two-factor interactions involving any one of the first five factors and any one of the last three factors.

As stated earlier, plans represented by quasi-orthogonal arrays can be used for planning experiments for quality improvement. Thus, from Theorem 4 with $g_1 = 3$ and $g_2 = 2$, one can obtain optimal plans for an experiment having r control factors and $n-r$ noise factors. For instance, if there are 5 control factors and 3 noise factors, then the 32-run saturated plan of Example 3 is optimal for the estimation of the mean, main effects of the control and noise factors and the 15 control versus noise two-factor interactions.

Remark. Under a model $\mathcal{M}(\Omega_f, \Omega_f)$, fractional factorial plans based on quasi-orthogonal arrays can be saturated in some cases. Such a saturated plan does not provide an internal estimate of the error variance and thus precludes the use of a standard F-test for testing the significance of the relevant factorial effects. In such a situtation, one might like to add one or more runs to the quasi-orthogonal array to get an estimate of the error variance. The question that arises then is how to add the run(s) so that the resulting plan is also optimal in some sense for estimating the relevant parameters in the model. The optimality of orthogonal array plus one run plans, when the original orthogonal array is of even strength ($= 2\alpha$) and the model includes the mean and all factorial effects involving α factors or less, has been tackled recently by Mukerjee (1999). The result of Mukerjee (1999) in the context of quasi-orthogonal arrays, under a model $\mathcal{M}(\Omega_f, \Omega_f)$ can be extended, following essentially the same arguments as in Mukerjee (1999). With reference to a model $\mathcal{M}(\Omega_f, \Omega_f)$, giving rise to the set S, suppose there exists a quasi-orthogonal array $QOA(N-1, n, m_1 \times \cdots \times m_n, S^*)$, and let $d_0 \in \mathcal{D}_N$ be obtained by adding any one run to the $N-1$ runs given by the array. Then, following Mukerjee (1999), it can be shown that d_0 is optimal in $\mathcal{D}_N$ with respect to every generalized criterion of type 1 (*cf.* Cheng (1980)) if $\mathrm{HCF}({m_i}^{x_i},\ 1 \le i \le n,\ x_i \ne 0) \ge 2$ for each $\boldsymbol{x} = x_1 \ldots x_n \in S^*$, where HCF stands for the highest common factor.

Note that the plan obtained by adding just one run to a saturated plan represented by a quasi-orthogonal array is the smallest plan providing an internal estimate of the error. This estimate, however, may not be very precise as it is based on only one degree of freedom. The issue of optimality of plans obtained by adding two or more runs to a plan represented by quasi-orthogonal arrays remains open.

Acknowledgements

The authors thank the referees and an Associate Editor for useful comments on earlier versions.

References

Cheng, C.-S. (1980). Optimality of some weighing and 2^n fractional factorial designs. *Ann. Statist.* **8**, 436-446.

Chiu, W. Y. and John, P. W. M. (1998). *D*-optimal fractional factorial designs. *Statist. Probab. Lett.* **37**, 367-373.

Dey, A. and Mukerjee, R. (1999a). *Fractional Factorial Plans.* Wiley, New York.

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

Gupta, S. and Mukerjee, R. (1989). *A Calculus for Factorial Arrangements*. Springer-Verlag, Berlin.

Hedayat, A. S. (1990). New properties of orthogonal arrays and their statistical applications. In *Statistical Design and Analysis of Industrial Experiments* (Edited by S. Ghosh), 407-422. Dekker, New York.

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

Hedayat, A. S. and 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. and Stufken, J. (1999). *Orthogonal Arrays : Theory and Applications*. Springer-Verlag, New York.

Kiefer, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (Edited by J. N. Srivastava) 333-353. North-Holland, Amsterdam.

Mukerjee, R. (1999). On the optimality of orthogonal array plus one run plans. *Ann. Statist.* **27**, 82-93.

Rosenbaum, P. R. (1994). Dispersion effects from fractional factorials in Taguchi's method of quality design. *J. Roy. Statist. Soc. Ser. B* **56**, 641-652.

Shoemaker, A., Tsui, K. and Wu, C. F. J. (1991). Economical experimentation methods for robust design. *Technometrics* **33**, 415-428.

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

Sun, D. X. and Wu, C. F. J. (1994). Interaction graphs for three-level fractional factorial designs. *J. Quality Tech.* **26**, 297-307.

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

Department of Statistics, Asutosh College, 92, S. P. Mukherjee Road, Calcutta 700 026, India.

Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, 7, SJS Sansanwal Marg, New Delhi 110 016, India.

E-mail: ashish@isid.ac.in

Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, 7, SJS Sansanwal Marg, New Delhi 110 016, India.

E-mail: adey@isid.ac.in

(Received June 2000; accepted January 2002)