Statistica Sinica 2(1992), 453-464

# TWO-LEVEL FACTORIAL DESIGNS FOR MAIN EFFECTS AND SELECTED TWO-FACTOR INTERACTIONS

A. S. Hedayat and H. Pesotan

*University of Illinois at Chicago and University of Guelph*

*Abstract:* From a practical viewpoint the first decision to be made in the construction of a design of a two-level factorial experiment is the choice of the parameters of interest. It is convenient to represent such a choice by considering an undirected graph $g$ with $n$ vertices and $e$ edges. The vertices and edges of $g$ are used respectively to identify the main effects of $n$ two-level factors and the $e$ two-factor interactions of interest. The parameters identified by $g$ together with the general mean are taken to be the parameters of interest. A design $d$ of the $2^n$ factorial will be called a $g$-design if and only if $d$ is saturated and is capable of providing an unbiased estimator of the parameters of interest relative to the orthogonal polynomial model. In this paper (i) a $g$-design is constructed for each graph $g$ and certain features of $g$-designs are noted, (ii) some $D$-optimality results for $g$-designs within the class of all $g$-designs are obtained.

*Key words and phrases:* Experimental design, fractional factorials, resolution III and V designs, $D$-optimal, orthogonal polynomial model, undirected graphs, isomorphic graphs, $g$-designs.

## 1. Introduction

Factorial designs in which each factor is studied at two-levels (high and low; present and absent) are very important in factor screening experiments and many scientific investigations. As a first approximation for the model of response controlled by these factors it is customary to include the main effects and some specified set of two-factor interactions. Thus, for example, if there are five factors to be studied, then we may want to explore and study a model involving the general mean, the five main effects and some three two-factor interaction effects. In this case, out of $\binom{5}{2} = 10$ two-factor interactions we have selected some three to be included in the model. Aside from the main effects and the mean, from a practical viewpoint, the selection of the other parameters of interest, mainly two factor interactions, is a crucial and weighty one and is often suggested by previous knowledge or other background information concerning the experiment. A discussion of how such interactions arise, and how one selects those which may be deemed important for the purposes of estimation, along with examples is given in Wu and Chen (1992).

A general design question to be studied is this: suppose there are $n$ factors under study and we intend to include the general mean, the main effects and $m \leq \binom{n}{2}$ two-factor interactions in the model. We would like to specify a design with at least $N = 1 + n + m$ level combinations so that the $N$ parameters in the exploratory model can be estimated unbiasedly. The literature offers some methods to tackle this problem. We review these briefly below.

If $N$ is small the desired design may be constructed on an ad hoc search basis. To save time in such a search, Greenfield (1976) suggests a more systematic tree search procedure to select an appropriate fraction. However, such procedures ultimately involve trial and error and rapidly become unwieldy as $N$ becomes large.

Taguchi (1959, 1960) proposes a graph aided method to tackle this problem. He identifies each of the $k$ vertices of an undirected graph $g(k, e)$ with a factor and each of the $e$ edges with a corresponding two factor interaction. With $g(k, e)$ he associates a 2-level fractional factorial design, which is capable of providing an unbiased estimator of the $(1 + k + e)$ parameters including the mean. For various values of $k$ and $e$ Taguchi, op. cit., has catalogued such graph aided designs. Thus, in the case of a $2^n$ experiment involving $N = 1 + n + m$ parameters consisting of the mean, $n$ main effects and a specified set of $m$ two factor interactions, Taguchi's method is as follows: (1) draw a graph $g^*(n, m)$ whose $n$ vertices are labelled by the factors and whose $m$ edges connect the vertices specified by the two factor interactions, (2) search the catalogued graphs to find a graph $g(k, e)$ such that $g^*$ is a subgraph of $g$ with the values $k \geq n$, $e \geq m$ as close to the values $n$ and $m$ respectively as obtainable in the catalogue. The method then indicates how the vertices of $g^*$ may be relabelled in accordance with the labelling of $g$. Then the design associated with $g$ solves the design selection problem for the unbiased estimation of the $N$ parameters. For further elucidation on Taguchi's procedure, see Wu and Chen (1992).

Wu and Chen, op. cit., have also considered the above design construction problem. They suggest that a satisfactory solution should meet three goals: (i) the design should be capable of estimating the specified parameters assuming the rest are negligible, (ii) the estimation in (i) should be possible even if the assumption that the remaining interactions are negligible is somewhat relaxed, (iii) there should be some built in flexibility in the design to allow for the estimation of other interactions not in the set of specified parameters under (ii). Wu and Chen (1992) discuss graph aided methods designed to meet these objectives. In this connection we mention that in Hedayat (1990) this design construction problem has been considered within the framework of meeting the objectives (i) to (iii) via a different approach, namely the introduction and study of the concept of orthogonal arrays of strength $t+$.

In this paper we follow the graphical approach of Taguchi in studying the design construction problem stated above, but there is a big difference between our method and that of Taguchi. As discussed above, Taguchi translates the problem in terms of a graph and uses this graph to find a design from among a set of catalogued graph aided designs. By contrast, in our approach we specify an undirected graph and use it to construct a saturated design directly. Thus no comparison with any catalogued graph aided design is involved.

Specifically, we do the following: for each $n$ and each selected set of two-factor interactions we specify a design with the minimum number of observations (saturated) so that all the parameters of the model can be estimated unbiasedly. If further observations are needed for variance estimation, then while our design can be augmented arbitrarily with one or more level combinations, in practice such an augmentation should be done carefully to attain statistical and other efficiencies.

As in Taguchi (1959) and Wu and Chen (1992), we find it useful to present each model under study with an undirected graph whose vertices are labelled with the main effects and edges with two-factor interactions which are of interest. From this perspective, specifying all nonisomorphic graphs on $n$ vertices and $e$ edges gives us the set of all possible models for the $2^n$ factorial in which the parameters of interest are identified by the labelled graph. This saves time and space in cataloging, storing and retrieving.

Section 2 outlines the linear model considered here and the use of a graph in representing a specific model. From a practical viewpoint, once the parameters of interest have been specified by a graph, the problem is to select a design so as not to alias these parameters. To tackle this problem, the concept of a $g$-design is introduced in Section 3 and a specific $g$-design is constructed for each graph $g$. In Section 4, some $D$-optimality results for the class of $g$-designs are obtained. Finally, in Section 5, further results on the bounds of the determinants of the information matrices of $g(n,e)$-designs are obtained.

## 2. Preliminaries

We assume that the response under study is influenced by $n$ quantitative factors labelled by $1,2,\ldots,n$. Each factor will be studied at two different levels. These levels will be coded by 0 and 1. A *design* $d$ of the $2^n$ factorial is a set of level combinations $\boldsymbol{t}$, where $\boldsymbol{t}=(i_1\, i_2 \ldots i_n)$ is a $(0,1)$-vector, namely $i_j \in \{0,1\}$. The response at the level combination $\boldsymbol{t}$ will be denoted by $Y_t$ and will be assumed to follow the standard orthogonal polynomial model (see Chapter 4, in Raktoe, Hedayat and Federer (1981)),

$$Y_t = \mu + \sum_{j=1}^{n} f(i_j)\alpha_j + \sum_{(j,k)\in I} g(i_j, i_k)\beta_{jk} + \varepsilon_t. \tag{2.1}$$

The various terms in the model (2.1) are defined as follows: $\mu$ is the general mean, $\alpha_j$ is the main effect of the factor $j$, $\beta_{jk}$ is the interaction effect between factors $j$ and $k$. The set $I$ consists of those pairs $(j,k)$ where interaction between factors $j$ and $k$ are assumed to be present in the model. Further, the known coefficients in the model are defined as

$$f(i_j) = \begin{cases} 1, & \text{if } i_j = 1, \\ -1, & \text{if } i_j = 0, \end{cases} \quad \text{and} \quad g(i_j, i_k) = \begin{cases} 1, & \text{if } i_j = i_k, \\ -1, & \text{if } i_j \neq i_k. \end{cases}$$

If we make $N$ observations at $N$ level combinations then we use an $N \times 1$ vector $Y$ to represent these observations. Thus under model (2.1), we can summarize our data with its associated model in the form $Y = X\boldsymbol{\theta}+\boldsymbol{\varepsilon}$, where each row of $X$ represents the coefficients in the model related to the corresponding observation in $Y$ and $\boldsymbol{\theta}$ is a column vector consisting of $\mu$, the main effects and the two-factor interactions in the model. The related error vector is $\boldsymbol{\varepsilon}$.

With each model we associate an undirected labelled graph $g$ on $n$ vertices. The vertices of $g$ represent the main effects and the edges of $g$ are used to represent the interactions of interest. Two seemingly different models may produce isomorphic graphs. Two undirected labelled graphs $g$ and $h$ are said to be *isomorphic* if one graph can be obtained from relabelling the vertices of the other graph. From the design viewpoint we need not identify designs for all the different models, but rather for nonisomorphic graphs. This saves us time and space. The number of nonisomorphic graphs on $n$ vertices grows very fast with $n$. For example, there are 156 and 1044 nonisomorphic graphs with $n = 6$ and $n = 7$ vertices, respectively. For further details on graphs and graph enumeration we refer to Harary and Palmer (1973).

## 3. The Class of $g$-Designs

Let $g(n,e)$ be a graph with $n$ vertices and $e$ edges. Let $V = \{1, 2, \ldots, n\}$ denote the set of $n$ vertices of $g$ and let $J(g)$ be a set of $e$ distinct pairs of the form $(ij)$, $i \neq j$, of the elements of $V$. The set $J(g)$ will be called the *edge* set of $g$, and a pair $(i\,j)$ in $J(g)$ indicates that there is an edge in $g$ joining the vertices $i$ and $j$.

A design of the $2^n$ factorial will be called a *$g(n,e)$-design* if and only if (1) it is capable of providing an unbiased estimator for the parameters specified by $g$, namely the mean, the main effects of the $n$ factors (identified by $V$), and the two-factor interactions specified by the edge set $J(g)$ relative to the orthogonal polynomial model (2.1), and, (2) it is *saturated*, namely, it contains precisely $n+e+1$ level combinations, which is the number of parameters in the model specified by $g$. A $g(n,e)$-design is called an orthogonal design if its corresponding information matrix is diagonal.

The class of all $g(n,e)$-designs will be denoted by $D(g,n,e)$. To shorten both the notation and the language we will sometimes refer to a $g(n,e)$-design simply as a $g$-design, especially when the number of vertices $n$ and the number of edges $e$ of the graph $g$ need no emphasis. Similarly, we will write $D(g)$ as an abbreviation for the class $D(g,n,e)$ when no confusion is likely. Note that by definition each $g$-design is nonsingular, that is, the determinant of its design matrix is nonzero.

A vector or a matrix all of whose entries are $+1$ or $-1$ will be called a $(-1,1)$-vector or a $(-1,1)$-matrix respectively. A column vector all of whose entries is $+1$ will be denoted by $\mathbf{1}$ and its dimension will be apparent from the context. A $(-1,1)$-matrix will be called *normalized* if its first column is the vector $\mathbf{1}$. Two $(-1,1)$-matrices will be called *equivalent* if either one can be obtained from the other by a finite sequence of row interchanges, column interchanges, or multiplication of a row or a column by $-1$. Let $\boldsymbol{u}'=(u_1,u_2,\ldots,u_n)$ and $\boldsymbol{v}'=(v_1,v_2,\ldots,v_n)$ be $(-1,1)$-vectors. The *Schur product* of $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined to be the vector $\boldsymbol{w}$ where $\boldsymbol{w}'=(u_1v_1,u_2v_2,\ldots,u_nv_n)$.

A square $(-1,1)$-matrix $T$ of order $n+e+1$ will be called a $g(n,e)$-*matrix* if and only if $T$ is equivalent to a $(-1,1)$-matrix of the form

$$[\,\mathbf{1}\,|X_1|\,X_2\,], \tag{3.1}$$

where
(i) $X_1$ has dimension $(n+e+1)\times n$ and whose columns are labelled by the elements of $V$ in order,
(ii) $X_2$ has dimension $(n+e+1)\times e$, and whose columns are indexed by $J(g)$, and, for $(ij)$ in $J(g)$, the $(ij)$th column of $X_2$ is the Schur product of the $i$th and $j$th columns of $X_1$.

The form (3.1) will be called the *standard form* for a $g(n,e)$-matrix and the submatrix $X_1$ will be called the *core*. Hereafter, we will assume that all $g(n,e)$-matrices under consideration are in standard form. The class of all nonsingular $g(n,e)$-matrices in standard form will be denoted by $M(g,n,e)$.

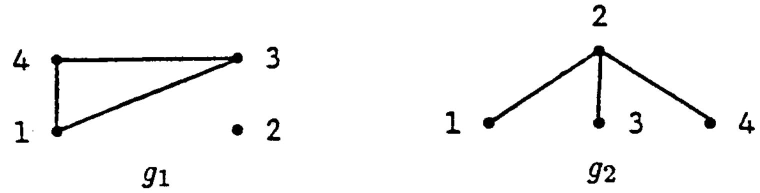
If $d$ is a $g(n,e)$-design, then clearly its design matrix is a nonsingular $g(n,e)$-matrix in standard form. Conversely, by replacing $-1$ by $0$ in the core of a nonsingular $g(n,e)$-matrix we obtain a $g(n,e)$-design whose design matrix is the given $g(n,e)$-matrix. Hence, we have

**Proposition 3.1.** *There is a one-to-one correspondence between the classes of all $g(n,e)$-designs and nonsingular $g(n,e)$-matrices in standard form.*

Proposition 3.1, in effect, means that the study of the class $D(g,n,e)$ is equivalent to the study of the class $M(g,n,e)$. Note that as a class $D(g,n,e)$ is intermediate to the classes of saturated resolution V and resolution III designs of the $2^n$ factorial. At one extreme when $g$ is the complete graph and $e = \binom{n}{2}$ then $D(g,n,e)$ equals the class of saturated resolution V designs of the $2^n$ factorial. At the other extreme, when $g$ is the trivial graph, namely $J(g) = \emptyset$, then the corresponding class of $g$-designs is the class of saturated resolution III designs.

Clearly if $g_1$ is a graph isomorphic to $g$ then, knowing the class $D(g)$, we also know the class $D(g_1)$, for, given $d$ in $D(g)$ by suitably interchanging the columns of $d$, we obtain a design in $D(g_1)$. However, it is possible for a design to be a $g_1$- and a $g_2$-design for two nonisomorphic graphs $g_i$ with $n$ vertices and $e$ edges. This is illustrated in the following example.

*Example 3.1.* Consider the two nonisomorphic graphs $g_1$ and $g_2$, used by Taguchi (1959,1960), with $n = 4$ and $e = 3$ whose labelled diagrams are given below:




Let $d = \{(1111),(1100),(1010),(1001),(0110),(0000),(0101),(0011)\}$. Then $d$ is a $g_1$- and a $g_2$-design whose design matrices are normalized Hadamard matrices of order 8. Thus, $d$ is a $D$-optimal $g_i$-design in the class $D(g_i,4,3)$ for $i = 1,2$, namely, that it has maximum determinant for its information matrix or, equivalently, since we are considering saturated designs it has maximum determinant for its design matrix in the class $D(g_i, 4, 3)$.

We now proceed to establish that the class $D(g)$ is nonempty for each graph $g$, by constructing a specific $g$-design for each $g$.

Let $g(n,e)$ be a graph with vertex set $V = \{1,2,\ldots,n\}$ and edge set $J(g)$. We define the design $d(g)$ of the $2^n$ factorial whose level combinations are the rows of the $(n+e+1)\times n$ matrix below:

$$\begin{pmatrix} 0 \\ I_n \\ D_e \end{pmatrix}, \qquad (3.2)$$

where

(1) the lead row is the zero vector $\mathbf{0}$,

(2) the next $n$ rows are the rows of $I_n$, the identity $n \times n$ matrix,

(3) The last $e$ rows are based on the set $J(g)$: for an edge $(ij)$ in $J(g)$, the $(ij)$th row of $D_e$ is a (0,1)-vector with exactly the $i$th and $j$th entries being $+1$.

For the present we show that for each graph $g$, the design $d(g)$ is indeed a $g(n,e)$-design. Some statistical properties of these designs will be addressed in the ensuing sections. From the perspective of balance in a design (an attractive feature possessed by regular designs), the designs $d(g)$ in general can be very unbalanced for estimating main effects or interactions. For instance, for any factor that does not appear in any of the interactions, the column in the design matrix corresponding to this factor has one $+1$ and $(n+e-1)$ $-1$'s. However, for some small values of $n$ and $e$, it will be shown in the next section that the corresponding designs $d(g)$ are, indeed, $D$-optimal.

We note that the design matrix associated with $d(g)$ and the given graph $g(n,e)$ whose first column corresponds to the general mean, whose next $n$ columns correspond to the main effects and whose final $e$ columns correspond to the $e$ interactions specified by $g(n,e)$ is as given in (3.3):

$$X = \begin{pmatrix} 1 & -\mathbf{1}' & \mathbf{1}' \\ \mathbf{1} & M & E \\ \mathbf{1} & N & F \end{pmatrix}. \tag{3.3}$$

The entries of the matrices $M$, $N$, $E$ and $F$ in (3.3) may be described as follows: (i) the $(i,j)$th entry of the $n \times n$ matrix $M$ is $+1$, if $i = j$ and $-1$, if $i \neq j$, (ii) the $((ij),t)$th entry of the $e \times n$ matrix $N$ is $+1$, if $t \in \{i,j\}$, and $-1$, otherwise, (iii) the $(i,(k\ell))$th entry of the $n \times e$ matrix $E$ is $-1$, if $i \in \{k,\ell\}$, and $+1$, otherwise, and (iv) the $((ij),(k\ell))$th entry of the $e \times e$ matrix $F$ is $+1$, if $|\{i,j\} \cap \{k,\ell\}| = 0$ or 2, and $-1$, if $|\{i,j\} \cap \{k,\ell\}| = 1$. Thus $d(g)$ will be a $g(n,e)$-design if we can show that its design matrix $X$ is of full rank. We establish this by explicitly computing the determinant of $X$. We now convert $X$ to a $(0,1)$-matrix by performing the following elementary row and column operations on $X$ in sequence: replace the $i$th column $C_i$ of $X$ by $C_i + \mathbf{1}$, for each $i \in V$, replace the $(ij)$th column $C_{(ij)}$ of $X$ by $(-1)C_{(ij)} + \mathbf{1}$ for each $(ij)$ in $J(g)$. Call the resulting matrix $K$ after performing these column operations. Replace each row $R$ (not the lead row) of the resulting matrix $K$ by $R - R_0$, where $R_0$ is the lead row of $K$. Let $W$ be the matrix that results from $X$ after these column and row operations. The matrix $W$ has the form

$$W = \begin{pmatrix} 1 & \mathbf{0}' & \mathbf{0}' \\ \mathbf{0} & 2I_n & P \\ \mathbf{0} & Q & S \end{pmatrix}.$$

We claim that (a) $P' = 2D_e$ and $Q = 2D_e$, and, (b) $S = 2D_eD_e' - 4I_e$, where $D_e$ is defined in (3.2). We need to recall how the matrices $P$, $Q$ and $S$ are obtained. Now the $(i,\ (k\ell))$th entry of $P$ is 2 if $i \in \{k,\ell\}$ and is 0, otherwise, and this coincides with the $((k\ell),i)$th entry of $Q$. Since $Q = 2D_e$ is straightforward, (a) is established. Next, the $((ij),(k\ell))$th entry of the matrix $S$ is 0, if $|\{i,j\}\cap\{k,\ell\}| = 0$ or 2, and is 2, if $|\{i,j\}\cap\{k,\ell\}| = 1$. Notice that the entries of $D_eD_e'$ are double indexed by the elements of $J(g)$. Now the $((ij),(k\ell))$th entry of $D_eD_e'$ is 0, 1, 2 respectively according as $|\{i,j\}\cap\{k,\ell\}|$ is 0, 1, 2. From this (b) follows. Thus from (a) and (b), $|\det(X)| = |\det(W)| = 2^{n+2e}$. Hence, we have

**Theorem 3.1.** *For each graph $g(n,e)$, the design $d(g)$ of the $2^n$ factorial defined in (3.2) is a $g(n,e)$-design and the absolute value of the determinant of its design matrix is $2^{n+2e}$.*

## 4. $D$-optimal $g$-designs

In the previous section it is noted that the study of the classes $D(g,n,e)$ and $M(g,n,e)$ are equivalent. Thus, from the viewpoint of $D$-optimality, we find it convenient to confine our attention to the construction of $D$-optimal $g(n,e)$-matrices in $M(g,n,e)$. The core of such matrices then provide the $D$-optimal $g(n,e)$-designs in $D(g,n,e)$.

Any two graphs with $n$ vertices and one edge are isomorphic. Thus it is enough to construct $D$-optimal $g(n,1)$-matrices for one specific $g(n,1)$. We call $m$ an $H$-number if a Hadamard matrix of order $m$ exists. Let $n = 2m-2$, where $m$ is an $H$-number. Let $L$ be a normalized Hadamard matrix of order $m$ and let $M$ be obtained from $L$ by deleting its first column. Let

$$K = \begin{pmatrix} \mathbf{1} & M & M & \mathbf{1} \\ \mathbf{1} & M & -M & -\mathbf{1} \end{pmatrix}. \tag{4.1}$$

Then $K$ is a Hadamard matrix of order $2m$ whose final column is the Schur product of an $i$th and $j$th column of $K$ for any $i$, $2 \le i \le m-1$, and $j = i+m-1$. If the given interaction is between factors $F$ and $G$, then we assign $F$ to any one of the first $m-1$ factors and $G$ to its counterpart factor $i+m-1$ in the core of $K$. Hence, we have

**Theorem 4.1.** *For each graph $g(n,1)$, $n = 2m-2$, with $m$ being an $H$-number, the design defined by the core of $K$ in (4.1) is an orthogonal $D$-optimal $g(n,1)$-design in the class $D(g,n,1)$.*

Let $\boldsymbol{v}_1' = (\mathbf{1}',\mathbf{1}',\mathbf{1}',\mathbf{1}')$, $\boldsymbol{v}_2' = (\mathbf{1}',\mathbf{1}',-\mathbf{1}',-\mathbf{1}')$, $\boldsymbol{v}_3' = (\mathbf{1}',-\mathbf{1}',\mathbf{1}',-\mathbf{1}')$ and $\boldsymbol{v}_4' = (\mathbf{1}',-\mathbf{1}',-\mathbf{1}',\mathbf{1}')$, where each component $\mathbf{1}'$ and $-\mathbf{1}'$ of these vectors is a row vector of order $2m+1$, $m \ge 1$. Note that $\boldsymbol{v}_4$ is the Schur product of $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$. Suppose that there is a $(-1,1)$-vector $\boldsymbol{x}' = (\boldsymbol{x}_1',\boldsymbol{x}_2',\boldsymbol{x}_3',\boldsymbol{x}_4')$, where the dimension of $\boldsymbol{x}_i'$ is $(2m+1)\times 1$ , such that $\boldsymbol{x}$ is orthogonal to each of $\boldsymbol{v}_i$. Then, if there are $k_i$ plus ones in $\boldsymbol{x}_i$, the orthogonality assumption leads to the equations

$k_1 + k_2 + k_3 + k_4 = 4m + 2$ and $k_1 = k_2 = k_3 = k_4$, clearly a contradiction. Hence no such vector $\boldsymbol{x}$ exists and we have , in contrast to Theorem 4.1, the following:

**Theorem 4.2.** *There is no orthogonal $g(n, 1)$-design with $n = 8m + 2$ factors as long as $m \geq 1$ and $n + 2$ is an $H$-number.*

Let $X$ be any $g(n, 1)$-matrix. Up to a row and column interchange we can always present $X$ in the form

$$X = \begin{pmatrix} 1 & \boldsymbol{x} & \boldsymbol{x} & A & \mathbf{1} \\ 1 & \boldsymbol{y} & -\boldsymbol{y} & B & -\mathbf{1} \end{pmatrix}. \tag{4.2}$$

Form the matrix $C = X + J$, where $J = \mathbf{1}\mathbf{1}'$. Then replacing the second column of $C$ by the sum of its second and third columns we obtain the matrix

$$G = 2Z = 2\begin{pmatrix} 1 & 2\boldsymbol{x}_1 & \boldsymbol{x}_1 & A_1 & \mathbf{1} \\ 1 & 0 & -\boldsymbol{y}_1 & B_1 & \mathbf{0} \end{pmatrix},$$

where $\boldsymbol{x}_1$, $\boldsymbol{y}_1$, $A_1$ and $B_1$ are obtained from their counterparts in $X$ by replacing $-1$ by 0. This leads us to the following result:

**Theorem 4.3.** *If $X$ is the design matrix of any $g(n, 1)$-design then* $\det(X)$ *is an integer multiple of* $2^{n+2}$, *and*

$$2^{n+2} \leq |\det(X)| \leq 2^{n+2}\left[\frac{(n+2)^{(n+2)/2}}{2^{n+2}}\right], \tag{4.3}$$

*where* [ ] *denotes the greatest integer less than or equal to the quantity in the brackets.*

**Proof.** It is easy to verify that $|\det(X)| = 2^{n+1}|\det(Z)|$. Expanding the determinant of $Z$ by its second column, we conclude that $\det(X)$ is an integer multiple of $2^{n+2}$. This, in conjunction with the Hadamard determinant inequality, establishes the theorem.

When $n = 3$, the two bounds in (4.3) coincide. Hence, $|\det(X)| = 32$ for all $g(3, 1)$-matrices $X$. When $n = 4$, by Theorem 4.3, for any $g(4, 1)$-matrix we have $|\det(X)| = 64k$ for some integer $k$. By Williamson (1946), the $D$-optimal value for the class of all $(-1, 1)$-matrices of order 6 is 160. Hence, we conclude that $k \leq 2$ and $|\det(X)| \leq 128$ for all $g(4, 1)$-matrices $X$. Consider the matrix $L$, where for simplicity we write + and − for +1 and −1 respectively:

$$L = \begin{pmatrix} + & + & + & + & + & + \\ + & + & - & - & + & - \\ + & - & + & - & + & - \\ + & - & - & + & - & - \\ + & + & + & - & - & + \\ + & - & - & - & - & + \end{pmatrix}. \tag{4.4}$$

Now $L$ is a $g(4,1)$-matrix with $J(g) = \{(34)\}$ and $|\det(L)| = 128$. Thus, we have

**Corollary 4.1.** (a) *Any $g(3,1)$-matrix in the class $M(g,3,1)$ is $D$-optimal. In particular, the design matrix $X$ of the design $d(g)$ defined in (3.2) is $D$-optimal in $M(g,3,1)$ and $|\det(X)| = 32$,* (b) *The matrix $L$ defined in (4.4) is a $g(4,1)$-matrix, where $J(g) = \{(34)\}$, which is $D$-optimal in $M(g,4,1)$ and $|\det(L)| = 128$.*

**Remark.** We would like to point out that in general when constructing $g(n,e)$-matrices, the requirement of normalization and the Schur product condition for $D$-optimal $g(n,1)$-matrices are independent. For example, consider the matrices

$$P = \begin{pmatrix} + & + & + & - & - \\ + & + & - & + & - \\ + & - & + & + & + \\ - & + & + & + & + \\ + & + & - & - & + \end{pmatrix} \qquad Q = \begin{pmatrix} + & + & + & + & + & + \\ + & + & - & - & + & - \\ + & - & + & - & + & - \\ + & - & - & + & - & - \\ + & + & + & - & - & + \\ - & + & + & + & - & - \end{pmatrix}.$$

Now $|\det(P)| = 48$ and $|\det(Q)| = 160$. It is known that these are $D$-optimal values for the classes of all $(-1,1)$-matrices of order 5 and 6 respectively (see for example, Williamson (1946)). Note that the final column in each of $P$ and $Q$ is respectively the Schur product of the two immediately preceding columns. However, neither can be converted into a $g(3,1)$ nor a $g(4,1)$-matrix, respectively, according to Corollary 4.1.

In the remaining part of the paper we shall analyze further the lower bound on the determinant of $g(n,e)$-matrices.

## 5. More on $g(n,e)$-Designs

We noted in Section 3 that the designs $d(g)$ while being $g(n,e)$-designs, suffer from imbalance for large $n$ and $e$. The purpose of this section is to show how these designs may be converted into more efficient ones.

For $n$ factors and $e$ two-factor interactions we gave a method for constructing a $g(n,e)$-design in (3.2). We would like to point out that any $g(m,t)$-design can be looked upon as a $g(n,e)$-design with $t \geq e$ so long as $m+t = n+e$, by simply taking $t-e$ arbitrary columns of $X_2$ into $X_1$ in the design matrix representation as given in (3.1). Therefore, by a direct method and the above conversion technique, for a given $n$ and $e$ we can construct precisely $N$, $g(n,e)$-designs using the $g(n-i, e+i)$-design, $i = 0,1,\ldots,k$ in (3.2), where $N = k+1$, $k$ being the largest integer less than $n$ and $\binom{n-k}{2} - k \geq e$, i.e., $k = \left[\frac{1}{2}[(2n+1) - \sqrt{8n+8e+1}]\right]$. By Theorem 3.1, the determinant of the design matrix for the $g(n,e)$-design obtained by converting the $g(n-i, e+i)$-design is, $2^{n+2e+i}, i = 0, 1, \ldots, k$. Therefore, the $D$-optimal $g(n,e)$-design among these $N$ designs is the one which is obtained from the conversion of the $g(n-k, e+k)$-design. Thus, we have

**Theorem 5.1.** *Among all $g(n,e)$-designs, the determinant of the design matrix of a D-optimal design is at least $2^{n+2e+k}$, where*

$$k = \left\lfloor \frac{(2n+1) - \sqrt{8n+8e+1}}{2} \right\rfloor.$$

Note that the lower bound given in Theorem 5.1 is sharper than the one given in (4.3) when $e = 1$.

We close this section by proving the following result.

**Corollary 5.1.** *Among all $g(5,1)$-designs, the design obtained by converting the $g(3,3)$-design in (3.2) is D-optimal.*

**Proof.** The determinant of the design matrix of this converted design is $2^{5+2+2} = 2^9$. On the other hand, Williamson (1946) has shown that the maximum determinant of a $(-1,1)$-matrix of order 7 is $2^6(9)$. This fact and the upper bound given in (4.3) establish the result.

## Acknowledgement

This research is supported by NSF Grant DMS-9113038 and by NSERC Grant A8776. The authors wish to thank the referees for their helpful comments on an earlier version of the paper.

## References

Greenfield, A. A. (1976). Selection of defining contrasts in two-level experiments. *Appl. Statist.* **25**, 64–67.

Harary, F. and Palmer, E. M. (1973). *Graphical Enumeration.* Academic press, New York.

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, Marcel Dekker, New York.

Raktoe, B. L., Hedayat, A. S. and Federer, W. T. (1981). *Factorial Designs.* John Wiley, New York.

Taguchi, G. (1959). Linear graphs for orthogonal arrays and their applications to experimental design, with the aid of various techniques. *Rep. Statist. Appl. Res., JUSE* **6**, 1–43.

Taguchi, G. (1960). Table of orthogonal arrays and linear graphs. *Rep. Statist. Appl. Res., JUSE* **7**, 1–52.

Williamson, J. (1946). Determinants whose elements are 0 and 1. *Amer. Math. Monthly* **53**, 427–434.

Wu, C. F. J. and Chen, Y. (1992). Graph-aided method for planning two-level experiments when certain interactions are important. in To appear in *Technometrics.*

Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, IL 60680, U. S. A.

Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario N1G 2W1, Canada.

(Received December 1990; accepted January 1992)