Sankhyā : The Indian Journal of Statistics
1982, Volume 44, Series B, Pt. 2, pp. 154–166.

CONSTRUCTION OF FACTORIAL DESIGNS WITH ALL MAIN EFFECTS BALANCED

By RAHUL MUKERJEE

University of Calcutta

SUMMARY. **The concept of difference array has been introduced to construct connected two-factor designs, symmetric and asymmetric, ensuring inter-effect-orthogonality, balancing main effects and retaining full information on at least one main effect. Two methods for constructing two-factor designs retaining full information on both main effects have been presented. Extensions to multifactor designs have been considered. The methods described cover almost all cases of factorial designs and require, in most cases, a smaller number of replications than any of the existing methods.**

1. Introduction

Consider a factorial experiment involving m factors $F_1, F_2, \ldots, F_m$, the j-th factor being at $s_j(\geqslant 2)$ levels, $1 \leqslant j \leqslant m$. A particular selection of levels $i = (i_1, i_2, \ldots, i_m)$ will be termed the i-th level combination. Throughout this paper the $v = \prod_{j=1}^{m} s_j$ level combinations will be lexicographically ordered (cf. Kurkjian and Zelen, 1963). Let the v level combinations be arranged in a block design with b blocks and incidence matrix $\boldsymbol{N}^{(v\times b)} = (n_{ih})$. The fixed effects intrablock model with no block-treatment interaction and with a constant error variance is assumed.

In such a design *inter-effect orthogonality* holds if best linear estimates of estimable treatment contrasts belonging to different factorial effects are orthogonal (i.e. uncorrelated). Any factorial effect is called *balanced* if all normalised contrasts belonging to that effect are estimated with the same variance. In the equireplicate case *full information* is retained on an effect if the effect be balanced and there is no loss of information on any contrast belonging to that effect (relative to the comparable complete block design).

Before presenting the main results, we introduce some notations, definitions and preliminary results. Let for $1 \leqslant j \leqslant m$, $\mathbf{1}_j^{(s_j\times 1)} = (1, 1, \ldots, 1)'$, $\boldsymbol{E}_j = \mathbf{1}_j\mathbf{1}_j'$, $\boldsymbol{I}_j = \boldsymbol{I}^{(s_j\times s_j)}$. $\boldsymbol{P}_j =$ an $(s_j-1)\times s_j$ matrix such that $(s_j^{-1/2}\,\mathbf{1}_j, \boldsymbol{P}_j')$ is orthogonal. For any $\boldsymbol{x} = (x_1, \ldots, x_m)$, $x_j = 0, 1, \forall j$, $\boldsymbol{x} \neq \mathbf{0}$, let $\boldsymbol{P}^{\boldsymbol{x}} = \boldsymbol{P}_1^{x_1}\times\ldots\times\boldsymbol{P}_m^{x_m}$, $\boldsymbol{\varepsilon}^{\boldsymbol{x}} = \boldsymbol{\varepsilon}_1^{x_1}\times\ldots\times\boldsymbol{\varepsilon}_m^{x_m}$, where

$$\boldsymbol{P}_j^{x_j} \begin{cases} s_j^{-1/2}\,\mathbf{1}_j' & \text{if } x_j = 0, \\ \boldsymbol{P}_j & \text{if } x_j = 1 \end{cases} \qquad \boldsymbol{\varepsilon}_j^{x_j} \begin{cases} \mathbf{1}_j & \text{if } x_j = 0 \\ \boldsymbol{I}_j & \text{if } x_j = 1. \end{cases}$$

2C

Definition 1.1 : A proper matrix is a square matrix with all row sums and all column sums equal.

Definition 1.2 : A $v \times v$ matrix $\boldsymbol{A}$, where $v = \prod_{j=1}^{m} s_j (s_j \geqslant 2, \forall j)$ is said to have structure K if it be expressible as a linear combinatiom of Kronecker products of proper matrices of orders $s_1, \ldots, s_m$ (taken in that order), i.e., if

$$\boldsymbol{A} = \sum_{g=1}^{w} \xi_g (\boldsymbol{V}_{g1} \times \ldots \times \boldsymbol{V}_{gm}), \qquad \ldots \quad (1.1)$$

where w is a positive integer, $\xi_1, \ldots, \xi_w$ are some real numbers and for each g, $\boldsymbol{V}_{gj}$ is some proper matrix of order s_j, $1 \leqslant j \leqslant m$. Structure K will always be with respect to a particular factorisation $v = \Pi s_j$ with the factors occurring in a particular order.

For an equireplicate factorial experiment in a block design with common replication number r, constant block size k and incidence matrix $\boldsymbol{N}$, the following theorems were proved by Mukerjee (1979, 1980a, 1980b) :

Theorem 1.1 : *A sufficient condition for inter-effect-orthogonality to hold is that the matrix* $\boldsymbol{NN}'$ *has structure* K. *In the connected case this is also a necessary condition for inter-effect-orthogonality.*

Theorem 1.2 : *Given that* $\boldsymbol{NN}'$ *has structure* K, *any main effect* $F_j (1 \leqslant j \leqslant m)$ *is balanced if and only if* $\varepsilon^{\boldsymbol{x}'} \boldsymbol{NN}' \varepsilon^{\boldsymbol{x}} = u_1 \boldsymbol{I}_j + u_2 \boldsymbol{E}_j$, *where* u_1, u_2 *are real numbers and* $\boldsymbol{x} = (x_1, \ldots, x_m)$, $x_j = 1$, $x_{j'} = 0 \ \forall j' \neq j$. *In this case the loss of information on* F_j *is given by* $L(F_j) = (rkv)^{-1} s_j u_1$.

Theorem 1.3 : *Full information is retained on any main effect if and only if in each block the levels of the corresponding factor occur equal number of times.*

Theorem 1.4 : *Given that* $\boldsymbol{NN}'$ *has structure* K *and the design is connected, the average loss of information on a complete set of orthonormal contrasts belonging to any factorial effect*

$$F_1^{x_1} F_2^{x_2} \ldots F_m^{x_m} \ (x_j = 0, 1, \forall j, \ \boldsymbol{x} = (x_1, \ldots, x_m) \neq \boldsymbol{0})$$

is given by

$$1 - r^{-1} (\Pi (s_j - 1)^{x_j}) \, [\mathit{Trace} \ (\boldsymbol{P}^{\boldsymbol{x}} \boldsymbol{C} \boldsymbol{P}^{\boldsymbol{x}'})^{-1}]^{-1},$$

where $\boldsymbol{C}$ *is the* $\boldsymbol{C}$ *matrix of the design.*

Earlier Kurkjian and Zelen (1963) and Kshirsagar (1966) considered a property of $\boldsymbol{NN}'$, which they called property A. $\boldsymbol{NN}'$ has property A if it be of the form (1.1) with $\boldsymbol{V}_{gj} = \boldsymbol{I}_j$ or $\boldsymbol{E}_j \ \forall g, j$.

Obviously property A of $\boldsymbol{NN'}$ is a special case of structure K. For equireplicate factorial designs with constant block size, property A of $\boldsymbol{NN'}$ is necessary and sufficient for inter-effect-orthogonality with all factorial effects balanced while structure K of $\boldsymbol{NN'}$ is necessary and sufficient for inter-effect-orthogonality alone.

In the present paper Theorems 1.1–1.3 will be utilised to construct connected factorial designs for which inter-effect-orthogonality holds with all main effects balanced, retaining full information on at least one main effcet. The methods described are applicable to a very wide variety of cases and require, in most cases, a smaller number of replicates than any of the existing methods. In many cases the smaller size of the design has been achieved at the cost of balancing of interactions. This, however, poses no problem since the analysis of the proposed designs can be done using formulae given by Mukerjee (1979). Also the properties of the proposed designs with respect to interactions may be explored using Theorem 1.4.

2. Construction of two-factor designs from varietal designs

Consider the problem of constructing an $s_1 \times s_2$ design in two factors F_1, F_2. For $j = 1, 2$, the levels of F_j will be denoted by $0, 1, \ldots, s_j - 1$. Restricting ourselves to equireplicate block designs if we want to retain full information on main effect F_1, by Theorem 1.3, block size must be a multiple of s_1 so that minimum block size will be s_1. In the following subsections the actual method of construction will be described.

2.1. *Difference array.* Let $\gamma^{(\alpha)}_{i_2 i_2'}$, $1 \leqslant \alpha \leqslant s_1 - 1$, $0 \leqslant i_2, i_2' \leqslant s_2 - 1$, be given nonnegative integers such that for each α, $\boldsymbol{\Gamma}^{(\alpha)} = (\gamma^{(\alpha)}_{i_2 i_2'})$ is a proper matrix with each row sum and each column sum equal to the same positive integer r. Let $\boldsymbol{M}$ be an $s_1 \times (rs_2)$ matrix whose rows are serially numbered $0, 1, \ldots, s_1 - 1$ and whose entries are chosen from the set $\{0, 1, \ldots, s_2 - 1\}$ such that for any i_1, i_1' $(i_1 \neq i_1',\ 0 \leqslant i_1, i_1' \leqslant s_1 - 1)$ if $i_1' - i_1 = \alpha (\text{mod } s_1)$ then in the $2 \times (rs_2)$ matrix $\boldsymbol{M}_{i_1 i_1'}$ formed by taking the i_1-th and i_1'-th rows of $\boldsymbol{M}$ as first and second rows respectively the ordered pair $\begin{pmatrix} i_2 \\ i_2' \end{pmatrix}$ occurs as a column vector $\gamma^{(\alpha)}_{i_2 i_2'}$ times, $1 \leqslant \alpha \leqslant s_1 - 1$, $0 \leqslant i_2, i_2' \leqslant s_2 - 1$. Then $\boldsymbol{M}$ is defined as a difference array $[rs_2, s_1, s_2, 2, \boldsymbol{\Gamma}^{(\alpha)}, 1 \leqslant \alpha \leqslant s_1 - 1]$ with rs_2 assemblies, s_1 constraints, s_2 levels, strength 2 and index parameters $\boldsymbol{\Gamma}^{(\alpha)}$, $1 \leqslant \alpha \leqslant s_1 - 1$. Obviously $\gamma^{(\alpha)}_{i_2 i_2'} = \gamma^{(-\alpha)}_{i_2' i_2}$, where $(-\alpha)$ is reduced mod s_1 and in each row of $\boldsymbol{M}$ each symbol is repeated r times.

If $\gamma^{(\alpha)}_{i_2 i'_2} = \gamma^{(\alpha)}_{i'_2 i_2} = \psi_{i_2 i'_2}$ (say), $\forall\ \alpha$, $\forall\ i_2, i'_2$, then the difference array reduces to a balanced array of strength 2. If $\gamma^{(\alpha)}_{i_2 i'_2} =$ constant $= \gamma$ (say), $\forall\ \alpha$, $\forall\ i_2, i'_2$, then $r = \gamma s_2$ and the difference array reduces to an orthogonal array $[\gamma s_2^2, s_1, s_2, 2]$.

If a difference array $\boldsymbol{M}$ as described above exists, taking its entries as levels of F_2, associating its rows with the levels of F_1 and taking its columns as blocks, one gets an $s_1 \times s_2$ design in rs_2 s_1-plot blocks with common replication number r for which the following theorem holds :

Theorem 2.1.1 : *In an $s_1 \times s_2$ factorial design constructed as above from a difference array $[rs_2, s_1, s_2, 2, \boldsymbol{\Gamma}^{(\alpha)}, 1 \leqslant \alpha \leqslant s_1-1]$, (i) inter-effect-orthogonality holds and (ii) full information is retained on main effect F_1. Further, main effect F_2 is also balanced if and only if $\sum_{\alpha=1}^{s_1-1} \boldsymbol{\Gamma}^{(\alpha)} = g_1\boldsymbol{I}_2 + g_2\boldsymbol{E}_2$, for some numbers g_1, g_2, and in that case, $L(F_2) = (rs_1)^{-1}(r+g_1)$.*

Proof : (i) For $0 \leqslant \alpha \leqslant s_1-1$, let $\boldsymbol{R}_\alpha$ be $(s_1 \times s_1)$ matrices such that the (β, β')-th cell of $\boldsymbol{R}_\alpha$ is filled by 1 if $\beta' = \beta + \alpha \pmod{s_1}$ and by 0 otherwise, $0 \leqslant \beta, \beta' \leqslant s_1-1$. Obviously $\boldsymbol{R}_0 = \boldsymbol{I}_1$ and $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_{s_1-1}$ are permutation matrices. Observing that two distinct level combinations (i_1, i_2), (i'_1, i'_2) occur together in no block if $i_1 = i'_1$ and in $\gamma^{(\alpha)}_{i_2 i'_2}$ blocks if $i'_1 - i_1 = \alpha \pmod{s_1}$, $1 \leqslant \alpha \leqslant s_1-1$, for the factorial design under consideration

$$\boldsymbol{NN}' = r(\boldsymbol{I}_1 \times \boldsymbol{I}_2) + \sum_{\alpha=1}^{s_1-1} \boldsymbol{R}_\alpha \times \boldsymbol{\Gamma}^{(\alpha)}. \qquad \ldots \quad (2.1.1)$$

Since $\boldsymbol{\Gamma}^{(\alpha)}$' s are proper matrices, $\boldsymbol{NN}'$ has structure K and inter-effect-orthogonality holds by Theorem 1.1.

(ii) Obvious by Theorem 1.3.

From (2.1.1) it is easily seen that

$$\boldsymbol{\varepsilon}^{01\prime}\boldsymbol{NN}'\boldsymbol{\varepsilon}^{01} = rs_1\boldsymbol{I}_2 + s_1 \sum_{\alpha=1}^{s_1-1} \boldsymbol{\Gamma}^{(\alpha)}.$$

Hence by Theorem 1.2, main effect F_2 is balanced if and only if $\sum_{\alpha=1}^{s_1-1} \boldsymbol{\Gamma}^{(\alpha)} = g_1\boldsymbol{I}_2 + g_2\boldsymbol{E}_2$, for some numbers g_1, g_2, in which event (with $u_1 = s_1(r+g_1)$, $u_2 = s_1 g_2$ in Theorem 1.2) $L(F_2) = (rs_1)^{-1}(r+g_1)$.

Q.E.D.

2.2. *The method of cyclic rotation.* A method for generating difference arrays from varietal designs will be described now. Consider an $s_1 \times d$ matrix $\boldsymbol{A}_0 = (a_{i_1 h})$, $a_{i_1 h} \in \{0, 1, \ldots, s_2-1\}$, $0 \leqslant i_1 \leqslant s_1-1$, $0 \leqslant h \leqslant d-1$. This can be looked upon as a varietal (possibly non-binary) design in s_2 varieties laid out in d blocks (columns) and s_1 rows. Let $\boldsymbol{A}_0^*$ be an $s_1 \times (ds_1)$ matrix such that

$$\boldsymbol{A}_0^* = [\boldsymbol{A}_0, \boldsymbol{R}_1\boldsymbol{A}_0, \ldots, \boldsymbol{R}_{s_1-1}\boldsymbol{A}_0], \qquad \ldots \quad (2.2.1)$$

where $\boldsymbol{R}_t$'s are as defined in the proof of Theorem 2.1.1.

Theorem 2.2.1 : *If $\boldsymbol{A}_0$ represents an equireplicate varietal design then $\boldsymbol{A}_0^*$ given by* (2.2.1) *forms a difference array.*

Proof : Let r be the common replication number in $\boldsymbol{A}_0$. Since $rs_2 = ds_1$, $\boldsymbol{A}_0^*$ is evidently an $s_1 \times (rs_2)$ matrix in symbols $0, 1, \ldots, s_2-1$. In any two-rowed submatrix of $\boldsymbol{A}_0^*$ consisting of, say, the i_1-th and i_1'-th rows $(0 \leqslant i_1 \neq i_1' \leqslant s_1-1, i_1'-i = \alpha (mod\, s_1))$ number of times the ordered pair $\begin{pmatrix} i_2 \\ i_2' \end{pmatrix}$ occurs as a column vector is equal to the number of times in $\boldsymbol{A}_0$ the symbols i_2 and i_2' are at some β-th and β'-th positions in a column where β, β' satisfy $\beta'-\beta = \alpha \pmod{s_1}$ (i.e., this number depends on i_1, i_1' only through α). Let this number be $\gamma^{(\alpha)}_{i_2 i_2'}$. For any i_2, α, $0 \leqslant i_2 \leqslant s_2-1$, $1 \leqslant \alpha \leqslant s_1-1$, $\sum_{i_2'=0}^{s_2-1} \gamma^{(\alpha)}_{i_2 i_2'}$ = number of times the symbol i_2 occurs in $\boldsymbol{A}_0 = r$. Similarly for any i_2', α, $\sum_{i_2=0}^{s_2-1} \gamma^{(\alpha)}_{i_2 i_2'} = r$. Hence for each α, $\boldsymbol{\Gamma}^{(\alpha)} = (\gamma^{(\alpha)}_{i_2 i_2'})$ is a proper matrix and $\boldsymbol{A}_0^*$ is a difference array $[rs_2, s_1, s_2, 2, \boldsymbol{\Gamma}^{(\alpha)}, 1 \leqslant \alpha \leqslant s_1-1]$. Q.E.D.

The above method of generating difference arrays starting from an $\boldsymbol{A}_0$ will be called the *method of cyclic rotation.* The two-factor design that can be constructed from such a difference array as described in subsection 2.1 will be called a *derived factorial design.*

John (1973) describes a method of construction of cyclic two-factor designs starting from some initial blocks or subsets of level combinations. Further blocks are developed (possibly discarding repeated blocks) by adding to each member of each initial block the $s_1 s_2$ level combinations (i_1, i_2) $(0 \leqslant i_j \leqslant s_j-1;\ j = 1, 2)$ addition being mod s_j for the levels of the j-th factor, $j = 1, 2$. The present method of construction of difference array and the $s_1 \times s_2$ factorial amounts to starting from the d initial subsets $\{(0, a_{0h}), (1, a_{1h}), \ldots, (s_1-1, a_{s_1-1,h})\}$, $0 \leqslant h \leqslant d-1$ and developing blocks by the addition of only the s_1 level combinations $(i_1, 0)$, $0 \leqslant i_1 \leqslant s_1-1$. Thus the resulting design is not a cyclic design in John's sense. Actually we are starting from a larger number of initial blocks and using a sort of curtailed cycle while John's intention seems to be to start from a smaller number of initial blocks and develop the full cycle.

2.3. *Use of balanced block designs* : *Definition* 2.3.1 : An equireplicate varietal design (possibly non-binary) in s_2 varieties laid out in d blocks each of size s_1, with incidence matrix $\boldsymbol{N}^*$ will be called a balanced block design (BBD) if $\boldsymbol{N}^*\boldsymbol{N}^{*\prime} = (r^*-\lambda^*)\boldsymbol{I}_2+\lambda^* \boldsymbol{E}_2$, for some integers r^*, λ^*.

Theorem 2.3.1 : (i) *For the difference array* (2.2.1) *based on the equireplicate varietal design* $\boldsymbol{A}_0$ *with common replication number* r, $\sum_{\alpha=1}^{s_1-1} \boldsymbol{\Gamma}^{(\alpha)} = g_1\boldsymbol{I}_2+g_2\boldsymbol{E}_2$ *for some numbers* g_1, g_2 *if and only if* $\boldsymbol{A}_0$, *when looked upon as a block design with columns as blocks, forms a BBD.*

(ii) *If* $\boldsymbol{A}_0$, *when looked upon as a block design, forms a BBD with incidence matrix* $\boldsymbol{N}^*$ *and* $\boldsymbol{N}^*\boldsymbol{N}^{*\prime} = (r^*-\lambda^*)\boldsymbol{I}_2+\lambda^* \boldsymbol{E}_2$, *then in the derived factorial design main effect* F_2 *is balanced with* $L(F_2) = (rs_1)^{-1}(r^*-\lambda^*)$.

Proof : (i) Let the design represented by the columns of $\boldsymbol{A}_0$ have incidence matrix $\boldsymbol{N}^* = (n^*_{i_2 h})$, $0 \leqslant i_2 \leqslant s_2-1$, $0 \leqslant h \leqslant d-1$. For the difference array (2.2.1) based on $\boldsymbol{A}_0$, for any i_2, i_2', $0 \leqslant i_2, i_2' \leqslant s_2-1$,

$$\sum_{\alpha=1}^{s_1-1} \gamma^{(\alpha)}_{i_2 i_2'} = \sum_{\alpha=1}^{s_1-1} \{\text{number of times the symbols } i_2, i_2' \text{ occur together at some } \beta\text{-th and } \beta'\text{-th positions in a column of } A_0 \text{ where } \beta, \beta' \text{ satisfy } \beta'-\beta = \alpha \pmod{s_1}\}$$

$$= \begin{cases} \sum_{h=0}^{d-1} n^*_{i_2 h}(n^*_{i_2 h}-1) & \text{if } i_2 = i_2' \\ \sum_{h=0}^{d-1} n^*_{i_2 h} n^*_{i_2' h} & \text{if } i_2 \neq i_2' \end{cases} \qquad \ldots \quad (2.3.1)$$

Since $r = \sum_{h=0}^{d-1} n^*_{i_2 h} \; \forall \; i_2$, by (2.3.1),

$$\sum_{\alpha=1}^{s_1-1} \boldsymbol{\Gamma}^{(\alpha)} = \boldsymbol{N}^*\boldsymbol{N}^{*\prime}-r\boldsymbol{I}_2 \qquad \ldots \quad (2.3.2)$$

whence (i) follows at once by the definition of BBD.

(ii) In this case by (2.3.2), for the difference array (2.2.1), $\sum_{\alpha=1}^{s_1-1} \boldsymbol{\Gamma}^{(\alpha)} = g_1\boldsymbol{I}_2+g_2\boldsymbol{E}_2$, where $g_1 = r^*-\lambda^*-r$, $g_2 = \lambda^*$, whence by Theorem 2.1.1 the result follows. Q.E.D,

The above theorem, together with Theorems 2.1.1, 2.2.1, gives a systematic method for constructing two-factor designs ensuring inter-effect-orthogonality, balancing main effects and retaining full information on at least one main effect (viz., main effect F_1). Since for every integral s_1, s_2 ($\geqslant 2$) a BBD for s_2 varieties in blocks of size s_1 exists, the method has very wide applicability.

(I) In particular, if $s_1 < s_2$ one may use a balanced incomplete block design (BIBD), with usual parameters r, λ. In this case for the derived factorial design (DFD), $L(F_2) = (rs_1)^{-1}(r-\lambda)$.

(II) If $s_1 = s_2$ one may use a randomised block design (RBD). In this case by Theorem 1.3 it is easy to check that in the DFD full information is retained on main effect F_2 as well.

(III) If $s_1 > s_2$ one may start with a BBD with incidence matrix $\boldsymbol{N}^{*(s_2\times s_2)} = (n^*_{i_2h})$, where $n^*_{i_2h} = x^*$ whenever $i_2 = h$ and $n^*_{i_2h} = y^*$ whenever $i_2 \neq h$. where x^*, y^* are positive integers such that $x^*+(s_2-1)y^* = s_1$. Writing $\boldsymbol{N}^*\boldsymbol{N}^{*\prime}$ explicitly and applying Theorem 2.3.1, it can be seen that in this case for the DFD, $L(F_2) = s_1^{-2}(x^*-y^*)^2$. $L(F_2) = 0$ if and only if $x^* = y^*$. But $x^* = y^* \Longrightarrow s_1 = x^*+(s_2-1)x^* = s_2x^*$. Hence unless $s_1(> s_2)$ is an integral multiple of s_2 full information cannot be retained on main effect F_2. As a working rule, in order to minimise $L(F_2)$, $|x^*-y^*|$ should be as small as possible.

In cases (I) and (III) above it can be shown that there always exists an arrangement of varieties within the blocks of the BIBD or BBD under use such that the method of cyclic rotation applied to the corresponding matrix $\boldsymbol{A}_0$ leads to a connected DFD. The same can be said also for case (II) provided $s_1 = s_2 > 2$ and the RBD used involves at least two blocks (cf. Mukerjee, 1980b).

Before concluding this section we compare the method described with that due to Nair and Rao (1948) who used orthogonal arrays $[s_2^2, s_1, s_2, 2]$ for constructing $s_1 \times s_2$ balanced confounded designs in blocks of size s_1 with $L(F_1) = 0$. In many cases our method will give designs with smaller number of replicates e.g. if $s_1 = 3$, $s_2 = 7$, starting with the symmetric BIBD $s_2 = d = 7$, $r = s_1 = 3$, $= 1$, our method will give a design in only three replicates while that due to Nair and Rao will require at least as many as ($s_2-1 =$) 6 replicates. Again in some cases (e.g., if $s_1 = 5$, $s_2 = 6$) the method by Nair and Rao fails as the required orthogonal arrays are nonexistent. However, in our method a 5×6 design can be constructed from the symmetric BIBD $s_2 = d = 6$, $r = s_1 = 5$, $\lambda = 4$. For $s_1 = 2$, $s_2 \geqslant 2$, it can be checked that our method gives a balanced confounded design and is equivalent to that due to Nair and Rao.

The following example serves as an illustration :

Example 2.3.1 : If $s_1 = 5, s_2 = 3$, starting from a BBD (blocks are written as columns) :

$$\begin{matrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{matrix}$$

we can get a 5×3 design in five replicates with the desired properties and with $L(F_2) = 1/25$. Incidentally, this 5×3 design is a balanced confounded design in the sense of Nair and Rao and interaction F_1F_2 is also balanced in it. Such a design cannot, however, be obtained by the method of Nair and Rao since an orthogonal array $[3^2, 5, 3, 2]$ does not exist. Muller (1966) gave a 5×3 balanced confounded design in blocks of size 5 with $L(F_1) = 0$, $L(F_2) = 1/25$. But his design required as many as ten replicates.

3. Retaining full information on both main effects

For an $s_1 \times s_2$ design with $s_1 = s_2$ or $s_1 > s_2$ and an integral multiple of s_2, this problem has already been considered in subsection 2.3. Here we present a general method applicable for $s_1 \neq s_2$. If $s_1 \neq s_2$ to retain full information on both the main effects in an equireplicate block design, block size must be a multiple of both s_1 and s_2 by Theorem 1.3. If s_1 and s_2 be prime to each other it is, therefore, impossible to retain full information on both main effects in an incomplete block design. Hence we consider only the case when s_1, s_2 are not prime to each other. In this case let $f (> 1)$ be their highest common factor and let $s_1 = f_1f$, $s_2 = f_2f$, $f_1, f_2 \geqslant 1, f_1 \neq f_2$. Minimum possible block size is f_1f_2f. Without loss of generality let $s_1 > s_2$.

Let $\boldsymbol{\theta}$ be an f_1s_2 component vector such that $\boldsymbol{\theta}' = (0, 0, \ldots, 0, 1, 1, \ldots, 1, \ldots, s_2-1, s_2-1, \ldots, s_2-1)$, where in $\boldsymbol{\theta}$ each symbol $0, 1, \ldots, s_2-1$ is repeated f_1 times. Let $\boldsymbol{R}^*_\alpha$ $(0 \leqslant \alpha \leqslant f_1s_2-1)$ be permutation matrices of order $f_1 s_2$ similarly defined as before. Let $\boldsymbol{B} = [\boldsymbol{\theta}, \boldsymbol{R}_1^* \boldsymbol{\theta}, \ldots, \boldsymbol{R}_{s_1 - 1}^* \boldsymbol{\theta}]$. $\boldsymbol{B}$ has $f_1 s_2 = f_2 s_1$ rows. Associate the first s_1 rows of $\boldsymbol{B}$ with the levels of F_1 in the order $0, 1, \ldots, s_1 - 1$, the next s_1 rows with the levels of F_1 in the same order and so on till all the rows are exhausted. Identify the elements of $\boldsymbol{B}$ with the levels of F_2 and form one block from each column.

The above method of constructing an $s_1 \times s_2$ design in s_1 blocks each of $f_1 f_2 f$ plots actually amounts to starting from an initial block $\{(gf_1 + g', g + g''f), 0 \leqslant g \leqslant f-1, 0 \leqslant g' \leqslant f_1 - 1, 0 \leqslant g'' \leqslant f_2 - 1\}$ and developing $(s_1 - 1)$ further blocks by adding to each member of the initial block the level combinations $(i_1, 0)$, $s_1 \geqslant i_1 \geqslant 1$, addition being mod s_1 for the levels of F_1.

For the above $s_1 \times s_2$ design which is easily seen to be connected and equireplicate (with common replication number f_1) the following theorem holds :

Theorem 3.1 : (i) *Inter-effect-orthogonality holds*, (*ii*) $L(F_1) = L(F_2) = 0$.

Proof : (i) For $0 \leqslant i_2 \leqslant s_2 - 1 = f_2 f - 1$, let $\boldsymbol{e}_{i_2}$ be an s_2-component vector having 1 at the i_2-th position and 0 at every other position. For each g, $0 \leqslant g \leqslant f-1$, let

$$\boldsymbol{e}_g^* = \boldsymbol{e}_g + \boldsymbol{e}_{f+g} + \ldots + \boldsymbol{e}_{(f_2 - 1)f + g}, \quad \ldots (3.1)$$

$$\boldsymbol{T}_g^{(f_1 s_2 \times f_1)} = \begin{bmatrix} \boldsymbol{e}_g^* \boldsymbol{e}_g^* & \ldots & \boldsymbol{e}_g^* \boldsymbol{e}_g^* \\ \boldsymbol{e}_g^* \boldsymbol{e}_g^* & \ldots & \boldsymbol{e}_g^* \boldsymbol{e}_{g+1}^* \\ \ldots & \ldots & \ldots \\ \boldsymbol{e}_g^* \boldsymbol{e}_g^* & \ldots & \boldsymbol{e}_{g+1}^* \boldsymbol{e}_{g+1}^* \\ \boldsymbol{e}_g^* \boldsymbol{e}_{g+1}^* & \ldots & \boldsymbol{e}_{g+1}^* \boldsymbol{e}_{g+1}^* \end{bmatrix}$$

where $(g+1)$ in $\boldsymbol{T}_g$ is reduced mod f. Then it can be seen that the incidence matrix $\boldsymbol{N}$ of the $s_1 \times s_2$ design is given by

$$\boldsymbol{N}^{(s_1 s_2 \times s_1)} = \begin{bmatrix} \boldsymbol{T}_0 & \boldsymbol{T}_1 & \ldots & \boldsymbol{T}_{f-1} \\ \boldsymbol{T}_1 & \boldsymbol{T}_2 & \ldots & \boldsymbol{T}_0 \\ \ldots & & \ldots & \ldots \\ \boldsymbol{T}_{f-1} & \boldsymbol{T}_0 & \ldots & \boldsymbol{T}_{f-2} \end{bmatrix},$$

whence, writing $\boldsymbol{N}$ in full, after some simplification

$$\boldsymbol{NN}' = \sum_{g=0}^{f-1} \sum_{g'=0}^{f_1-1} \boldsymbol{R}_{gf_1+g'} \times \boldsymbol{Z}_{gg'}, \qquad \ldots \quad (3.2)$$

where $\boldsymbol{R}_\alpha$ $(0 \leqslant \alpha \leqslant s_1-1 = f_1f-1)$ are as defined in subsection 2.1 and for $0 \leqslant g \leqslant f-1$, $0 \leqslant g' \leqslant f_1-1$,

$$\boldsymbol{Z}_{gg'} = (f_1-g') \sum_{a=0}^{f-1} \boldsymbol{e}_a^* \boldsymbol{e}_{a+g}^{*\prime} + g' \sum_{a=0}^{f-1} \boldsymbol{e}_a^* \boldsymbol{e}_{a+g+1}^{*\prime}, \qquad \ldots \quad (3.3)$$

where $(a+g)$ and $(a+g+1)$ are reduced mod f. From (3.1) for $0 \leqslant a \leqslant f-1$, $\boldsymbol{e}_a^{*\prime} \boldsymbol{1}_2 = f_2$. Hence from (3.3)

$$\boldsymbol{Z}_{gg'}\boldsymbol{1}_2 = f_2(f_1-g') \sum_{a=0}^{f-1} \boldsymbol{e}_a^* + f_2 g' \sum_{a=0}^{f-1} \boldsymbol{e}_a^* = f_1 f_2 \boldsymbol{1}_2.$$

Similarly $\boldsymbol{1}_2'\boldsymbol{Z}_{gg'} = f_1 f_2 \boldsymbol{1}_2'$. Hence $\boldsymbol{Z}_{gg'}$'s are proper matrices and from (3.2) applying Theorem 1.1, it follows that inter-effect-orthogonality holds.

(ii) Since in each block each level of F_1 occurs f_2 times and each level of F_2 occurs f_1 times, this is obvious by Theorem 1.3. Q.E.D.

A formula for the average loss of information on the two-factor interaction in such designs will be derived in the Appendix.

Example 3.1 : A 6×4 design in three replications constructed by the above method is shown below. Blocks are written as rows. Here $f = 2$, $f_1 = 3$, $f_2 = 2$, $\boldsymbol{\theta}' = (0, 0, 0, 1, 1, 1, 2, 2, 2, 3, 3, 3)$

level of F_1	0	1	2	3	4	5	0	1	2	3	4	5
	0	0	0	1	1	1	2	2	2	3	3	3
	0	0	1	1	1	2	2	2	3	3	3	0
level of F_2	0	1	1	1	2	2	2	3	3	3	0	0
	1	1	1	2	2	2	3	3	3	0	0	0
	1	1	2	2	2	3	3	3	0	0	0	1
	1	2	2	2	3	3	3	0	0	0	1	1

Another method for constructing two-factor designs with $L(F_1) = L(F_2) = 0$ will be described in Section 4.

4. Extension to multifactor designs

Multifactor designs with desirable properties can be constructed as Kronecker products of simpler designs. This method was used to a limited extent by Tyagi (1971) for constructing a $2\times5\times6$ balanced confounded design.

For $1 \leqslant t \leqslant n$, let D_t be a factorial design in m_t ($\geqslant 1$, if $m_t = 1$, D_t is a varietal design) factors F_{tj}, $1 \leqslant j \leqslant m_t$, such that (a) D_t is equireplicate with common replication number r_t and constant block size k_t, (b) D_t is connected, (c) inter-effect-orthogonality holds for D_t, and (d) each main effect in D_t is balanced with $L(F_{tj}) = q_{tj}$ ($\geqslant 0$), $1 \leqslant j \leqslant m_t$. Then D, the Kronecker product of $D_1, \ldots, D_n$ will be a design in $m (= \Sigma m_t)$ factors which, by known results, will be connected, will have a common replication number Πr_t and a constant block size Πk_t. We assume that in $D, D_1, \ldots, D_n$ the rows of the corresponding incidence matrices are associated with the relevant level combinations in lexicographic order. The following theorem can be proved easily using Theorems 1.1, 1.2 :

Theorem 4.1 : (i) *For D inter-effect-orthogonality holds.* (ii) *In D each main effect is balanced with $L(F_{tj}) = q_{tj}$, $1 \leqslant j \leqslant m_t$, $1 \leqslant t \leqslant n$.*

As an interesting application of the above result we present an alternative method for constructing $s_1\times s_2$ designs retaining full information on both main effects.

Using the notations of Section 3, let $f > 2$. Introducing pseudofactors G_1, G_2, G_3, G_4 with f, f, f_1, f_2 levels respectively the f_1f (f_2f) level combinations of G_3 and G_1 (G_4 and G_2) are identified with the f_1f (f_2f) levels of $F_1(F_2)$. Let D_1 = connected $f\times f$ design in G_1, G_2 in two replicates retaining full information on both main effects obtained as described in subsection 2.3; D_2 = single replicate randomised block design (RBD) in levels of G_3; D_3 = single replicate RBD in levels of G_4. Here $r_1 = 2$, $r_2 = r_3 = 1$, $k_1 = f$, $k_2 = {}_1$, $k_3 = f_2$. Let $D = D_1\times D_2\times D_3$. In D interpreting the level combinations of G_1, G_2, G_3, G_4 in terms of the level combinations of F_1, F_2, by Theorem 4.1, it easily follows that the resulting two-factor design (i) is equireplicate with only two replications, (ii) has constant block size ff_1f_2, (iii) is connected, (iv) satisfies inter-effect-orthogonality, and (v) ensures $L(F_1) = L(F_2) = 0$. If $f > 2$, $f_1 > 2$ (e.g., when $s_1 = 9, s_2 = 6$) this alterntative method will be more economic (in terms of the number of replicates) than that described in Section 3.

Appendix

A general formula for the average loss of information on different factorial effects for the designs considered in this paper was presented in Theorem 1.4. For the designs considered in Section 3 which retain full information on main effects, we shall derive here a still simpler formula for the average loss of information on the two factor interaction.

By (3.2),

$$\boldsymbol{C} = f_1(\boldsymbol{I}_1\times\boldsymbol{I}_2)-(f_1f_2f)^{-1}\sum_{g=0}^{f-1}\sum_{g'=0}^{f_1-1}\boldsymbol{R}_{gf_1+g'}\times\boldsymbol{Z}_{gg'}, \qquad \ldots \quad (A.1)$$

where for $0\leqslant g\leqslant f-1$, $0\leqslant g'\leqslant f_1-1$, $\boldsymbol{R}_{gf_1+g'}$ and $\boldsymbol{Z}_{gg'}$ as defined earlier are circulant matrices. Denoting by $w_0=1, w_1, \ldots, w_{s_1-1}$ the distinct s_1-th roots and by $z_0=1, z_1, \ldots, z_{s_2-1}$ the distunct s_2-th roots of unity, by (A.1) the eigenvalues of $\boldsymbol{C}$ are

$$\lambda_{i_1i_2}=f_1-(f_1f_2f)^{-1}\sum_{g=0}^{f-1}\sum_{g'=0}^{f_1-1}w_{i_1}^{gf_1+g'}\,z_{i_2}^{g}(f_1-g'+g'z_{i_2})$$

$$\times\left\{1+(z_{i_2}^{f})+\ldots+(z_{i_2}^{f})^{f_2-1}\right\}, \qquad \ldots \quad (A.2)$$

with respective eigenvectors $\boldsymbol{W}_{i_1}\times\boldsymbol{Z}_{i_2}$, where $\boldsymbol{W}_{i_1}=(1, w_{i_1}, \ldots, w_{i_1}^{s_1-1})'$, $\boldsymbol{Z}_{i_2}=(1, z_{i_2}, \ldots, z_{i_2}^{s_2-1})'$, $0\leqslant i_1\leqslant s_1-1$, $0\leqslant i_2\leqslant s_2-1$.

Let $\boldsymbol{M}_1^{(s_1-1\times s_1)}$, $\boldsymbol{M}_2^{(s_2-1\times s_2)}$ be such that $\boldsymbol{M}_1'=(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_{s_1-1})$, $\boldsymbol{M}_2'=(\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{s_2-1})$. Since

$$\text{row space }(\boldsymbol{P}^{11})\equiv\text{row space }(\boldsymbol{M}_1\times\boldsymbol{M}_2),$$

it follows that Trace $(\boldsymbol{P}^{11}\boldsymbol{C}\boldsymbol{P}^{11'})^{-1}=\sum_{i_1=1}^{s_1-1}\sum_{i_2=1}^{s_2-1}\lambda^{-1}_{i_1i_2}$. Hence by Theorem 1.4 and (A.2), after some simplification,

average loss of information on F_1F_2

$$=1-f_1^{-1}(s_1-1)(s_2-1)[f_1^{-1}\{(s_1-1)(s_2-1)-f_1(f-1)\}+\Delta]^{-1}, \qquad \ldots \quad (A.3)$$

where

$$\Delta = \sum_{g=1}^{f-1} \sum_{l=1}^{f_1} [f_1 - f_1^{-1}\{1-\cos 2\pi g f^{-1}\}\{1-\cos 2\pi(lf-g)s_1^{-1}\}^{-1}]^{-1}.$$

Using the above, the average loss of information on F_1F_2 in the design in Example 3.1 is ·103.

The above treatment is possible for the designs considered in Section 3 since the matrices involved in the right-hand side of (3.2) are not only proper but also circulant. This is not, however, necessarily true for the designs considered in other sections. Hence for such designs it is difficult to obtain an algebraic reduction like (A.3) of Theorem 1.4, though for any particular design with a numerically known $\boldsymbol{C}$ matrix, it is straightforward to apply Theorem 1.4 directly.

Acknowledgement. The author is grateful to Professor S. K. Chatterjee, Department of Statistics, University of Calcutta, for his kind guidance throughout the work. Thanks are also due to the referee for his constructive suggestions.

References

JOHN, J. A. (1973): Generalised cyclic designs in factorial experiments. *Biometrika,* **60**, 55-63.

KSHIRSAGAR, A. M. (1966): Balanced factorial designs. *J. Roy. Statist. Soc.*, Ser. B, **28**, 559-569.

KURKJIAN, B. and ZELEN, M. (1963): Applications of the calculus of factorial arrangements I. Block and direct product designs. *Biometrika,* **50**, 63-73.

MUKERJEE, R. (1979): Inter-effect-orthogonality in factorial experiments. *Calcutta Statist. Assoc. Bull.*, **28**, 83-108.

——— (1980a): Further results on the anaysis of factorial experiments. *Calcutta Statist. Assoc. Bull.*, **29**, 1-26.

——— (1980b): Asymmetric factorial designs and allied problems. Unpublished Ph.D. thesis submitted to the University of Calcutta.

MULLER, E. R. (1966): Balanced confounding in factorial experiments. *Biometrika,* **53**, 507-524.

NAIR, K. R. and RAO, C. R. (1948): Confounding in asymmetric factorial experiments. *J. Roy. Statist. Soc.*, Ser. B, **10**, 109-131.

TYAGI, B. N. (1971): Confounded asymmetric factorial designs. *Biometrics,* **27**, 229-232.

Paper received : May, 1980.

Revised : June, 1981.