The Annals of Statistics
1981, Vol. 9, No. 4, 737–757

OPTIMUM BALANCED BLOCK AND LATIN SQUARE DESIGNS FOR CORRELATED OBSERVATIONS

BY J. KIEFER[1] AND H. P. WYNN[1,2]

Cornell University, University of California, Berkeley, and Imperial College, London

In this paper designs are found which are optimum for various models that include some autocorrelation in the covariance structure V. First it is noted that the ordinary least squares estimator is quite robust against small perturbations in V from the uncorrelated case $V_0 = \sigma^2 I$. This "local" argument justifies our use of such estimators and restriction to the class of designs $\mathscr{X}^*$ (balanced incomplete block or Latin squares) optimum under V_0. Within $\mathscr{X}^*$ we search for designs for which the least squares estimator minimizes appropriate functionals of the dispersion matrix under various correlation models V. In particular, we consider "nearest neighbor" correlation models in detail. The solutions lead to interesting combinatorial conditions somewhat similar to those encountered in "repeated measurement" designs. Typically, however, the latter need not be BIBD's and require twice as many blocks. For Latin squares, and hypercubes, the conditions are less restrictive than those giving "completeness."

1. Introduction; literature. The relevant literature concentrates on the estimation of the fixed effects of independent variables, or factors, rather than the estimation of the underlying error process. In other words, the emphasis is on correcting the usual estimators and finding designs for which estimators do well under given error assumptions.

The two papers by Papadakis (1937) and Bartlett (1938) are the first in the field. Papadakis produced estimators which are close to the weighted least squares estimators for certain designs. A thorough analysis of the method was carried out by Atkinson (1969). R. M. Williams (1952) studied designs with treatments laid out in a one-dimensional array, under similar models, with the first-order autocorrelation in that dimension. Recent work on the analysis of spatial patterns by Bartlett (1975), Besag (1974), Ripley (1977), and others has a bearing on the design and analysis problems. The recent paper by Bartlett (1978) stimulated a lively and useful discussion of the issues.

There have been other approaches to guarding against effects from neighboring plots in experimental design. The classical methods of randomization, due primarily to Fisher, should certainly be mentioned. For example, selecting a Latin square at random from all, or a subclass, of Latin squares has been advocated.

There is a considerable literature on designs for so-called "change-over" or "residual" effects where fixed treatment effects are carried over to neighboring plots. A good presentation and source of references for this subject is Hedayat and Afsarinejad (1975). Important early papers in the area are by E. J. Williams (1949, 1950) and Patterson (1950, 1951, 1952).

In the field of optimum design Kiefer (1960) extends the work of R. M. Williams (1952). The papers by Sacks and Ylvisaker (1966, 1968, 1969) and the extensions by Wahba (1971,

Received April 1979; revised June 1980.
[1] Research sponsored by the National Science Foundation.
[2] Research also sponsored by the Grants Council.
AMS 1970 *subject classifications*. Primary 62K05; secondary 05B20.
Key words and phrases. Optimum designs, correlated observations, difference sets, balanced incomplete block designs, Latin squares and hypercubes.

1974) study the estimation of fixed effects under very general continuous time error processes. Bickel and Herzberg (1979) develop some asymptotic theory for similar processes with simple linear regression models for the fixed effects. O'Hagan (1978) gives a Bayesian approach to the same kind of model.

Closest to the approach presented in the present paper are the papers by Berenblut and Webb (1974) and by Duby et al. (1977), and the recent more comprehensive thesis by Martin (1977). These authors investigate the behavior of different classical designs under various correlation assumptions. The stationary correlation model considered by them turns out to be less tractable than the "nearest neighbor" structure we treat in detail; thus, Duby et al. give numerical illustrations of design comparisons rather than general combinatorial conclusions of the type we reach. The present paper may be seen as giving some theoretical backing to (1) the choice of classical designs as a class for investigation, (2) the use of the ordinary least squares estimators, and (3) the use of various "equineighbored" conditions on combinatorial designs.

Further references to optimum design theory and combinatorial analysis will be given as they are required.

2. Least squares versus BLU estimators. We assume the linear regression model

$$E(Y) = X\theta,$$

where $Y = (Y_1, \ldots, Y_n)^T$ is a vector of observations and X is an $n \times p$ design matrix belonging to a class $\mathscr{X}$ of such matrices. The parameter vector $\theta = (\theta_1, \ldots, \theta_p)^T$ is unknown. The covariance matrix of the observations $\mathrm{Cov}(Y) = V$ belongs to a class $\mathscr{V}$ of covariance matrices which contains that of the standard case, $V_0 = \sigma^2 I$, where I is the $n \times n$ identity matrix. We assume every V in $\mathscr{V}$ is positive definite.

Let $t = B\theta$, with B $v \times p$, be a v-vector of estimable functions of θ. A well-known necessary and sufficient condition for this estimability is that $B^T = X^T XC$ for some $p \times v$ matrix C. Let $\hat{t}_0$ be the unique minimum variance linear unbiased estimator (BLUE) of t under $V = V_0$, which we shall refer to as the least squares (LS) estimator. Thus $\hat{t}_0 = B\hat{\theta}_0$, where $\hat{\theta}_0$ is any solution of the normal equations. Let M^- be the Moore-Penrose g-inverse of M. The covariance matrix of $\hat{t}_0$ under $V = V_0$ is

$$\mathrm{Cov}(\hat{t}_0 \mid V_0) = \sigma^2 B(X^T X)^- B^T.$$

The estimability of t is unaffected by the choice of V in $\mathscr{V}$. The BLUE for t under a general V in $\mathscr{V}$ is the weighted least squares (WLS) estimator

$$\hat{t}_V = B(X^T V^{-1} X)^- X^T V^{-1} Y,$$

whose covariance matrix under V is

$$\mathrm{Cov}(\hat{t}_V \mid V) = B(X^T V^{-1} X)^- B^T.$$

(It should cause no confusion that $\hat{t}_{V_0}$ has been abbreviated $\hat{t}_0$.) We may also calculate the covariance matrix of the LS estimator $\hat{t}_0$ under V, which is

$$\mathrm{Cov}(\hat{t}_0 \mid V) = B(X^T X)^- X^T V X (X^T X)^- B^T.$$

(Note that $\hat{t}_0$ is an unbiased estimator of t under V.) Writing $A_0 = X^T X$ and $A_V = X^T V^{-1} X$ and $V = \sigma^2(I + \Gamma)$, we obtain (as proved by Strand (1973) in the case $B = I$)

$$\begin{aligned} &\sigma^{-2}[\mathrm{Cov}(\hat{t}_0 \mid V) - \mathrm{Cov}(\hat{t}_V \mid V)] \\ &\quad = \sigma^{-2}(A_0^- X^T V - A_V^- X^T) V^{-1} (A_0^- X^T V - A_V^- X^T)^T B^T \qquad (2.1) \\ &\quad = B A_0^- X^T \Gamma [I - X A_0^- X^T] \Gamma X A_0^- B^T + O(\Gamma^3) = O(\Gamma^2), \end{aligned}$$

the last as $\Gamma \to 0$.

If we take, for example, the case of $\Gamma = \rho L$ where L is a fixed symmetric matrix and ρ is sufficiently small, then $O(\Gamma^2)$ becomes $O(\rho^2)$.

A well-known necessary and sufficient condition that $\mathrm{Cov}(\hat{t}_0 \mid V) = \mathrm{Cov}(\hat{t}_V \mid V)$, or, equivalently, that $\hat{t}_0 = \hat{t}_V$ (for all Y), is that

$$VXA_0^- B^T \subset R(X) \tag{2.2}$$

where $R(U)$ is the column space (range) of a matrix U. This condition, and similar ones, can be found in the work of a number of authors, such as Zyskind (1967), Rao (1967), Watson (1967, Theorem 1), McElroy (1967), Kruskal (1968), and others. It is interesting to note that condition (2.2) in turn is a necessary and sufficient condition for the first term in the third form of (2.1) to be zero, which is equivalent to

$$[I - XA_0^- X^T]\Gamma XA_0^- B^T = 0.$$

Thus we cannot choose X, Γ and B to get a better approximation than $O(\Gamma^2)$ in (2.1) without forcing $\mathrm{Cov}(\hat{t}_0 \mid V) = \mathrm{Cov}(\hat{t}_V \mid V)$ and hence $\hat{t}_0 = \hat{t}_V$.

The expansion (2.1) says that in an approximate sense we are justified in using the ordinary least squares estimate if we feel that any autocorrelation present is small. Furthermore it supplies our basic motivation for the approach adopted in this paper. Thus assuming V is unknown, but that we are primarily concerned with V's whose perturbation from $V = V_0$ is small, we suggest the following robustness argument. It is somewhat similar in style to the approach taken by Box, Draper, and others in the search for designs robust against the possible presence of higher order polynomial terms in regression models, that is, the so-called variance-bias methods. Our motivation is also two-stage:

(1) Find the class of designs in $\mathscr{X}$, say $\mathscr{X}^*$, which are optimum (in some specified sense) under $V = V_0$, using $\hat{t}_0$ as our estimator of t.

(2) Among all the designs $\mathscr{X}^*$ from stage (1), find the class of designs $\mathscr{X}^{**}$ which are optimum, again using $\hat{t}_0$ but under a specific structure V or class of V that may be present. (These V can be quite far from V_0.)

Thus at stage (1) we seek to minimize specified functional(s) Φ of $\mathrm{Cov}(\hat{t}_0 \mid V_0)$ (or of its g-inverse, and at stage (2) we minimize functionals Ψ of $\mathrm{Cov}(\hat{t}_0 \mid V)$. We now consider this process in more detail.

3. Optimum designs. In the settings considered in this paper we take the components $t_1, \ldots, t_v$ of t to be certain estimable contrasts $t_i = \sum_j b_{ij}\theta_j$; that is, $\sum_j b_{ij} = 0\ \forall i$. (It may help the reader to think of these in terms of (4.2) in the BIBD setting.) To stress the dependence of the covariances on the design, we define, for X in $\mathscr{X}$,

$$D(X, V) = \mathrm{Cov}(\hat{t}_0 \mid V).$$

We will always be dealing with choices of the t_i for which

$$D(X, V) \text{ has zero row and column sums } \forall X, \tag{3.1}$$

and hereafter *assume* that (3.1) holds. Define

$$C(X, V) = D^-(X, V). \tag{3.2}$$

This, too, has zero row and column sums. When $V = V_0$ and the $t_1, \ldots, t_v$ are appropriately defined (as in settings considered later herein), $C(X, V_0)$ is the so-called "C-matrix" of experimental design theory. We emphasize that, when $V \neq V_0$, $C(X, V)$ is the inverse of the covariance matrix of the LS estimator $\hat{t}_0$, *not* of some $\hat{t}_{V_1}$.

Kiefer (1975) pointed out that the class of convex decreasing functionals Φ on the set of C is more general than the class of convex increasing functionals Ψ on the matrices $D = C^-$, since $\Phi(C) = \Psi(C^-)$ is convex if Ψ is. The strict inclusion of one class in the other is illustrated by the D-optimality criterion (as discussed below). Thus it is more satisfactory to know that a design is optimum relative to all suitably symmetric convex criteria on the class of C-matrices than to know such a result on the class of D-matrices. This is the spirit of the definition employed in Kiefer (1975) in the settings of

(i) two factors, blocks and treatments with a fixed number b of blocks and fixed block size k (with v treatments and the t_i appropriate treatment contrasts),

(ii) three factors, rows, columns, and treatments in a $v \times v$ array with one observation per cell (with the t_i contrasts of effects of any one of the factors).

A design is called *universally optimum relative to* $\mathscr{X}$ if it minimizes $\Phi(C(X, V_0))$ over $\mathscr{X}$ for every Φ which is convex and invariant under permutations of coordinates (rows and columns of C) and has the property that

$$\Phi(bC) \leq \Phi(C) \qquad \forall b > 1$$

(or for every Φ which is an increasing function of such a function).

In the settings (i) and (ii) above the BIBD's (if they exist) and the Latin squares are the only universally optimum designs (and in fact are the only optimum designs for any single strictly convex Φ). They will therefore exactly comprise the class $\mathscr{X}^*$ referred to at the end of Section 2. Note that when $k = v$ in (i), $\mathscr{X}^*$ is exactly a set of b complete blocks.

We now turn to the second stage of optimization referred to in Section 2. The computation of $C(X, V)$ is often complicated, even for $X = X^* \subset \mathscr{X}^*$. In essence the procedure involves first computing $D(X^*, V)$ and then obtaining $C(X^*, V)$ as its Moore-Penrose inverse, which may be a formidable task if it is to be exhibited as a specific function of X, Γ and B. We try to avoid this inversion by use of an analogue of the method used to prove universal optimality, but cannot achieve quite such a strong optimality result. Call X^{**} *weakly universally optimum relative to* $\mathscr{X}^*$ for a covariance matrix V if it minimizes $\Psi(D(X^*, V))$ over $\mathscr{X}^*$ for every convex Ψ invariant under permutation of coordinates and such that

$$\Psi(bD) \geq \Psi(D) \qquad \forall b > 1$$

(or for every Ψ which is an increasing function of such a function). An even simpler argument than that used to prove universal optimality then shows that X^{**} is weakly universally optimum relative to $\mathscr{X}^*$ for covariance matrix V if

(3.3) (a′) $D(X^{**}, V)$ is completely symmetric (CS),

(b′) $\operatorname{trace} D(X^{**}, V) = \min X^* \in \mathscr{X}^* \operatorname{trace} D(X^*, V)$

(If Ψ is strictly convex and X^{**} satisfies (3.3), no other D can be Ψ-optimum.)

The criteria Ψ covered by this definition include (with λ_i = eigenvalues)

$$\Psi_p(D) = \textstyle\sum_1^{v-1} \lambda_i^p(D) \qquad \text{for } p \geq 1$$

(which for $p = 1$ is the A-optimality criterion, equivalent to $\sum_{i,j} \operatorname{Var}(\hat{t}_{0i} - \hat{t}_{0j} \mid V)$),

$$\Psi_\infty(D) = \max_i \lambda_i(D) \qquad (E\text{-optimality}),$$

but not, unfortunately,

$$\Psi_0(D) = \textstyle\sum_1^{v-1} \log \lambda_i(D) \qquad (D\text{-optimality}).$$

(The criteria Ψ_p, which depend only on the λ_i, are not merely permutation invariant, but are also orthogonally invariant. Two other criteria, which are of the former but not the latter form, are mentioned at the end of this section.)

Conditions (3.3) (a′) and (b′) also imply, by a similar averaging argument to that used to prove (weak) universal optimality, the optimality of d^* for every Ψ which is a nondecreasing Schur-convex function of the ordered eigenvalues of D. See Giovagnoli and Wynn (1980) or Constantine (1980) for a full discussion in the more usual context of $\Phi(C)$. The use of Schur convexity in design optimality was suggested to one of the present authors by I. Olkin about 1970, and design majorization arguments have appeared in Cheng (1979) and Kiefer (1975, page 339). D-optimality is also excluded in these terms: a competitor of d^* satisfying (3.3) could in general have larger trace(D) but smaller det(D).

In some settings, however, D-optimality and more can be concluded for V near V_0 (and V in $\mathscr{V}$); the following two conditions will be shown to suffice: (1) Suppose (as is the case for $\mathscr{X}^*$ consisting of all BIBD's or all Latin squares) that $D(X, V_0) = D_0$ (say) is the same for all X in $\mathscr{X}^*$ and that $\mathscr{X}^*$ is finite. (2) Suppose furthermore, that for fixed σ there is a constant $c_1 > 0$ such that as $D \to D_0$

$$\Psi(D) = \Psi(D_0) + c_1 \operatorname{trace}(D - D_0) + o(D - D_0)$$

for D satisfying (3.1) (as is the case for Ψ_0 and other suitably regular and symmetric Ψ when D_0 is CS, which is true for D_0 in the examples we have mentioned). Under these conditions since (Section 2) $\sigma^{-2}\mathrm{Cov}(\hat{t}_0 \mid V) = B(A_0^- + A_0^- X^T \Gamma X A_0^-)B^T$, we have

$$\sigma^{-2} \operatorname{trace} D(X, V) = \operatorname{trace} BA_0^- B^T + \operatorname{trace}(XA_0^- B^T B A_0^- X^T)\Gamma,$$

where, as above, $V = \sigma^2(I + \Gamma)$. It then follows from the finiteness of $\mathscr{X}^*$ and the fact that $D(X, V) - D_0 = O(\Gamma)$ as $\Gamma \to 0$ that only an A-optimum X in $\mathscr{X}^*$ can possibly minimize $\Psi(D(X, V))$ for V sufficiently near V_0. Moreover, if all A-optimum X^* yield the same $D(X^*, V)$ for V sufficiently near V_0, we conclude that all these X^*, and only these, are also Ψ-optimum for V in a neighborhood of V_0 (and in $\mathscr{V}$). This last is true in some of the examples of interest.

The results obtained in Sections 4 and 5 are of the following kind. At stage (1) the Latin squares, BIBD's, etc., are universally optimum under V_0 for the $\mathscr{X}$ under consideration, and thus constitute the class $\mathscr{X}^*$. At stage (2) within $\mathscr{X}^*$ the weakly universally optimum designs for given V (or Γ) are found (and the optimality can be extended locally as indicated in the previous paragraph). The form of $\hat{t}_0$, whose components are the LS estimators of some standard treatment contrasts, are simple because of the balanced or orthogonal nature of the designs in $\mathscr{X}^*$. This makes calculation of $D(X^*, V)$ a relatively easy exercise. This means, in turn, that minimization of a $\Psi(D(X^*, V))$ is *easier* than minimization of a $\Psi(\mathrm{Cov}(\hat{t}_V \mid V))$ notwithstanding the simple form of $\mathrm{Cov}(\hat{t}_V \mid V)$ given in Section 2.

It will be a feature of the examples discussed in Sections 4 and 5 that trace$D(X^*, V)$ is a constant for all X^* in $\mathscr{X}^*$. Thus with respect to minimizing the A-optimality criterion trace(D), all such X^* are equally good. Furthermore if we can find an X^* for which all the diagonal elements of $D(X^*, V)$ are equal and hence all $\mathrm{Var}(\hat{t}_{0i} \mid V)$ are equal, then obviously this X^* minimizes $\max_i \mathrm{Var}(\hat{t}_{0i} \mid V)$. (This certainly holds for any X^* with the CS property (3.3)(a′), but this last condition is not necessary, just as it was not for A-optimality.) Although this last optimality criterion is not of as much interest as the Ψ_p because of the meaning of the $\hat{t}_{0i}$ (see (4.2)), it has been considered in the literature, e.g., in Duby et al. (1977). (A-optimality is more meaningful because of its equivalent meaning, given shortly below (3.3).) Of greater interest is the criterion $\max_{i,j} \mathrm{Var}(\hat{t}_{0i} - \hat{t}_{0j} \mid V)$, the maximum of variances of estimators of "principal contrasts" $\alpha_i - \alpha_j$; when all X^* in $\mathscr{X}^*$ have the same value of $\operatorname{tr} D(X^*, V)$ and (3.1) holds, any X^* satisfying (3.3)(a′) is optimum in this sense. The last two criteria mentioned are additional to the orthogonally invariant Ψ_p, but both also fall within weakly universal optimality.

4. M-way layouts. In this paper the examples will be confined to special M-way layouts, for models with M factors. In the present section we consider settings where balanced orthogonal designs (complete blocks, Latin squares or hypercubes, etc.) exist.

Consider a model with M factors labelled $1, \ldots, M$. Let

$$N(i_1, \ldots, i_M)$$

be the number of observations at the i_jth level of factor j ($j = 1, \ldots, M$). We shall assume for simplicity when $M \geq 3$ that $N(i_1, \ldots, i_M) = 0$ or 1 for all $i_1, \ldots, i_M$. Factor j is assumed to have n_j levels ($j = 1, \ldots, M$).

Factor 1 will be called the treatment factor and its levels called *treatments*. We put $n_1 = v$ in accordance with standard notation. The usual linear additive model is

$$E(Y_{(i_1,\ldots,i_M)}) = \alpha_{i_1}^{(1)} + \cdots \alpha_{i_M}^{(M)},$$

where $Y_{(i_1,\ldots,i_M)}$ is the observation (if there is one) at level i_j of factor j ($j = 1, \ldots, M$). If Y is the vector of $Y_{(i_1,\ldots,i_M)}$ in some order (e.g., lexicographic), then let $V = \mathrm{Cov}(Y)$. In this section for $M \geq 3$ we shall take V to be such that

$$\begin{aligned} \mathrm{Cov}(Y_{(i_1,\ldots,i_M)}, Y_{(i_1',\ldots,i_M')}) &= \sigma^2 && \text{if } \textstyle\sum_{j=2}^M |i_j - i_j'| = 0 \\ &= \rho\sigma^2 && \text{if } \textstyle\sum_{j=2}^M |i_j - i_j'| = 1 \qquad (4.1) \\ &= 0 && \text{otherwise.} \end{aligned}$$

This is a "nearest neighbor" (NN) correlation structure in that observations are considered to be positioned by the levels of their last $M - 1$ factors, and an observation has correlation ρ with its nearest neighbors but not with observations "further away". The NN structure is described slightly differently, below and in Section 5, for $M = 2$; this is because we are considering (for example) only one Latin square for $M = 3$, but several blocks for $M = 2$.

The dot notation will be used for summation (rather than averaging) here, for convenience in considering N as well as Y:

$$N(i_1, \cdot, \ldots, \cdot) = \textstyle\sum_{i_2=1}^{n_2} \cdots \sum_{i_M=1}^{n_M} N(i_1, \ldots, i_M),$$

etc. Similarly,

$$Y_{(i_1,\cdot,\ldots,\cdot)} = \textstyle\sum_{i_2=1}^{n_2} \cdots \sum_{i_M=1}^{n_M} Y_{(i_1,\ldots,i_M)},$$

etc.

Define the standard treatment contrasts (chosen to make $t_i - t_j = \alpha_i^{(1)} - \alpha_j^{(1)}$),

$$t_i = \alpha_i^{(1)} - v^{-1} \textstyle\sum_1^v \alpha_j^{(1)} \qquad i = 1, \ldots, v. \qquad (4.2)$$

We shall suppress the superscript so that $\alpha^{(1)} \equiv \alpha$ in all that follows. Thus $t = (t_1, \ldots, t_v)^T$ is the vector t of the previous sections.

When $M = 2$ we consider $\mathscr{X}$ to be the class of designs for which

(i) $n_1 = v$, $n_2 = b$,

(ii) $N(\cdot, j) \leq v$.

(Often (ii) is strengthened to assume equality.) That is, we have b "blocks", each permitted to be of size $\leq v$. A treatment may appear more than once in a block. It is well known that the complete block designs, with b blocks of size v and each treatment repeated once per block, are uniquely universally optimum (Kiefer, 1958, 1975), and we take the class of such designs to be $\mathscr{X}^*$. Of course $\hat{t}_{0i} = b^{-1}Y_{(i,\cdot)} - (bv)^{-1}Y_{(\cdot,\cdot)}$. To save space, we postpone detailed calculations of $D(X^*, V)$ until the block design considerations of Section 5 for general block size k. The NN covariance structure (as in Section 5) is that observations in different blocks are uncorrelated, and observations in the same block are correlated only if they are neighbors. (The position in a block can refer to time or to position in a linear array of plots, etc.) Theorem 5.1, specialized to the present case, then shows that a design X^* in $\mathscr{X}^*$ is weakly universally optimum (for every possible value $\rho \neq 0$) if and only if, for all pairs i, i' with $1 \leq i < i' \leq v$, the quantities

$$v\#\{j \mid i, i' \text{ are NN's in block } j\} + \#\{j \mid i \text{ is at an end of block } j\} + \#\{j \mid i' \text{ is at an end of block } j\}$$

are the same. (If such an X^* exists, no X not satisfying this condition can be optimum for strictly convex Ψ.) A simple counting argument (special case of (5.5)) shows that a necessary condition for this is that $v \mid b$ when v is odd and $v \mid 4b$ when v is even; however, when $4 \mid v$ the value $b = v/4$ is too small for the last displayed expression to be independent of i, i', since at most half the treatment pairs can occur as NN's. We thus find that $b = v$ is the smallest number of blocks for v odd and that $b = v/2$ is the smallest number when v is even. Designs satisfying these conditions, and which are thus complete block designs of minimal size to be weakly universally optimum for the NN model, can be obtained by taking the blocks to be the columns of the Latin square design of Theorem 4.2 below when v is odd, and to be the first $v/2$ columns of that Latin square design when v is even.

When $M = 3$ define the class of designs $\mathscr{X}$ with

(i) $n_1 = n_2 = n_3 = v$,

(ii) $N(i, j, k) = 0$ or 1 and $N(\cdot, j, k) = 0$ or 1 $(1 \le i, j, k \le v)$.

(Often one assumes all $N(\cdot, j, k) = 1$, but this is not necessary.) Kiefer (1958, 1975) proves that Latin squares, namely those designs for which

$$N(i, j, \cdot) = N(i, \cdot, k) = 1$$

$(1 \le i, j, k \le v)$ are the class of universally optimum designs, $\mathscr{X}^*$, within this $\mathscr{X}$. (Of course, such designs do have all $N(\cdot, j, k) = 1$.)

Because of the orthogonal and balanced nature of Latin squares the LS estimators of the t_i have a simple form, namely,

$$\hat{t}_{0i} = v^{-1}Y_{(i,\cdot,\cdot)} - v^{-2}Y_{(\cdot,\cdot,\cdot)}.$$

From this we can calculate for a Latin square X^*

$$D(X^*, V_0) = v^{-1}\sigma^2(I_v - v^{-1}J_{v,v}).$$

Let $N_{ii'}$ denote the number of times that treatments i and i' are adjacent in the square X^* in any direction (vertically, horizontally but *not* diagonally, and counting each adjacency just once). For any Latin square X^* we have

$$\sum_{i'(\ne i)} N_{ii'} = 4(v - 1), \tag{4.3}$$

$$\sum_{i<i'} N_{ii'} = 2v(v - 1).$$

Moreover if all the $N_{ii'}$ are equal for $i \ne i'$, their common value is 4.

Making use of (4.3), after some manipulation we find that with V of (NN) form (4.1),

$$D(X^*, V) = v^{-1}\sigma^2(I - v^{-1}J) - (v - 1)v^{-3}4\rho\sigma^2J + v^{-2}\rho\sigma^2N, \tag{4.4}$$

where N is the matrix whose i, i'th entry is $N_{ii'}$ and whose diagonal entries are 0.

Notice that since the diagonal entries of $D(X^*, V)$ are the same for all Latin squares X^*, all such X^* are equally good with respect to the A-optimality and $\min_{X^*}\max_i \mathrm{Var}(\hat{t}_{0i} \mid V)$ criteria, as described in Section 3. Also there are no "edge effects" which typically arise with some other types of V matrices and $\hat{t}$. (See Martin, 1977.)

We can immediately see that (b') of (3.3) is automatically satisfied and that (a') is satisfied if and only if all $N_{ii'}$ are equal. We shall soon see that such designs always exist. Thus we have

THEOREM 4.1. *Among the class $\mathscr{X}^*$ of Latin square designs and under the nearest neighbor covariance V (for each $\rho \ne 0$), the class of weakly universally optimum designs is the class of designs with all $N_{ii'}$ equal* $(1 \le i, i' \le v;\ i \ne i')$.

The condition "all $N_{ii'}$ equal" (= 4) is much weaker than the condition that a square be "complete", that is, row- and column-complete. See Dénes and Keedwell (1974) for a discussion of the latter. However the method, first proposed by E. J. Williams (1949) and rediscovered by a number of authors, which is used to construct complete Latin squares when v is even, can be used to construct squares with all $N_{ii'}$ equal for all values of v. The construction is given by

THEOREM 4.2. *Form a* $v \times v$ *square whose* (j, k) *cell* $(1 \le j, k \le v)$ *contains the treatment number*

$$\sum_{r=1}^{j} (-1)^r(r-1) + \sum_{r=1}^{k} (-1)^r(r-1),$$

reduced (mod v). *This has all* $N_{ii'}$ *equal.*

EXAMPLE 4.1. For $v = 5$, identifying $0 \equiv 5$, we obtain

5	1	4	2	3
1	2	5	3	4
4	5	3	1	2
2	3	1	4	5
3	4	2	5	1

A common alternate description of this construction places successive integers in alternate rows of the first column on the way down and continues this process on the way back, completes the design cyclically in rows, and then reorders columns so that the first row and first column are the same. Notice that these squares have the property that in any row (column) the differences between *successive* treatment numbers produce every nonzero residue (mod v) exactly twice. This property is discussed extensively in Section 5 for BIBD's.

The method of Theorem 4.2 when v is even can be used to give designs with columns of size v but rows of size only $v/2$ which still have the property that each i, i' pair of adjacent treatments occurs once in every column. The reason that we may use a number of columns equal to only half of the number of rows, compared to the full number used in repeated measurement designs, is that we do not distinguish the order of adjacent pairs. As an example consider the design on page 232 of Hedayat and Afsarinejad (1975) for $v = 6$. The first three columns are $(1, 6, 2, 5, 3, 4)^T$, $(2, 1, 3, 6, 4, 5)^T$ and $(3, 2, 4, 1, 5, 6)^T$. Each column gives successive differences equivalent to (5, 2, 3, 4, 1) (mod 6), that is, all nonzero residues (mod 6). Each adjacent i, i' pair occurs just once. Notice the somewhat different order to that in Theorem 4.2. Also, the order of columns is immaterial in the Hedayat-Afsarinejad design, so that the latter setting is closer to that of block designs discussed in Section 5.

The results for Latin squares can be extended to higher way layouts and in particular to Latin cubes and hypercubes.

Suppose, then, for $M \ge 3$ and for a given positive integer $p < M - 1$, that $\mathscr{X}$ is the class of designs with

(i) $n_1 = v = k^p$, $n_j = k (j = 2, \ldots, M)$,

(ii) $N(i_1, \ldots, i_M) = 0$ or 1 and $N(\cdot, i_2, \ldots, i_M) = 0$ or 1 $(1 \le i_1 \le v, 1 \le i_2, \ldots, i_M \le k)$.

The previous optimality considerations can be extended to M-way layouts (Kiefer, 1958, page 690), and in fact a similar argument to that in Kiefer (1975) shows that the fully balanced and orthogonal designs are universally optimum; these are the designs X^* for which

$$N(i_1, i_2, \cdot, \ldots, \cdot) = N(i_1, \cdot, i_3, \ldots, \cdot) = \cdots = N(i_1, \cdot, \cdot, \ldots, i_M) = k^{M-p-2}.$$

(This implies that all $N(\cdot, i_2, \ldots, i_M) = 1$.) That is, every $k^p \times k$ two-way layout formed by taking treatments and one other factor (and ignoring the others) is a complete two-way layout with k^{M-p-2} observations in each cell.

We hereafter let $\mathscr{X}^*$ be the class of such designs X^*.

It is convenient to think of the k^{M-1} cells of the $(M-1)$-cube, with k levels in each direction, as specifying the last $M-1$ factors. Then a design in $\mathcal{X}$ is an assignment of treatment labels, one to each cell, or with none to a cell for which $N(\cdot, i_2, \ldots, i_M) = 0$. The designs X^* in $\mathcal{X}^*$ were called Latin hypercubes by Kishen (1949) when $M > 3$. (For discussion of other definitions see Kerr et al., 1973.) When $M = 3$, $p = 1$, the formulation will be seen to reduce to the earlier Latin square setup.

By a calculation similar to that for Latin squares, again using the NN structure V, one can show that, for $i \neq i'$, a Latin hypercube with $n = k^{M-1}$, $v = k^p$, and $r = n/v$ replications of each treatment, has

$$\mathrm{Cov}(\hat{t}_{0i}, \hat{t}_{0i'}) = -n^{-1}\sigma^2 + \rho\sigma^2 r^{-2}[N_{ii'} - v^{-1}(N_{i.} + N_{i'.}) + v^{-2}N_{..}],$$

where $N_{ii'}$ counts the number of times treatments i and i' are neighbors in any of the $M-1$ coordinate directions (not diagonally) and where $N_{i.} = \sum_{i'} N_{ii'}$ and $N_{..} = \sum_i N_{i.}$; we note that the sum over i' includes N_{ii}, which need not be 0 for a Latin hypercube (as it was automatically for a Latin square) unless $p = M-2$. A simple counting argument yields

$$N_{..} = \textstyle\sum_{i,i'} N_{ii'} = 2(M-1)k^{M-2}(k-1).$$

Thus, equality of all $N_{i.}$ and of all $N_{ii'}$ $(i \neq i')$ implies (3.3)(a′), and since (3.1) entails $\mathrm{tr}D = -\sum_{i\neq i'} \mathrm{Cov}(\hat{t}_{0i}, \hat{t}_{0i'})$, we see that (3.3)(b′) will be satisfied if the design minimizes

$$\sigma^{-2}\,\mathrm{tr}\,D = n^{-1}v(v-1) - \rho r^{-2}(v^{-1}N_{..} - \textstyle\sum_i N_{ii}), \tag{4.5}$$

which reduces to what one obtains from (4.4) in the Latin square case.

Our $\mathcal{X}^*$ consists of the Kishen hypercubes, for which we have mentioned that it is not automatic that $N_{ii} = 0$ unless $p = M-2$. (Greater "strength" of the array, restricting designs to a small subset of the designs universally optimum under V_0, would be required.) Hence we can no longer minimize (4.5) without knowledge of sgn ρ. There are three options, which parallel those considered in Kiefer (1960) in a simpler setting:

(α) Assume it known that $\rho \geq 0$, so that we do want all $N_{ii} = 0$;

(β) assume it known that $\rho \leq 0$, so we want $\sum_i N_{ii}$ maximized;

(γ) not knowing sgn ρ, take a minimax approach, by minimizing $|N_{..} - v\sum_i N_{ii}|$.

The most interesting and perhaps the most natural of these is (α), which is the only formulation we consider in the remainder of this section.

Under assumption (α), then, the weakly universally optimum designs in $\mathcal{X}^*$ are the Latin hypercubes with all $N_{ii} = 0$ and with all $N_{ii'}$ equal for $i \neq i'$. This requirement places rather rigid constraints on the parameter values, since we must equate the total number of adjacencies with the common value of $N_{ii'}$ times the total number of i, i' pairs. That is, $N_{..}/2$ must be a multiple of the total number of treatment pairs, $k^p(k^p-1)/2$. When $p = 1$ this is always achievable. When $p > 1$ we obtain that $2(M-1)$ is a multiple of $k^{p+2-M}(k^p - 1)/(k-1)$, and since k^{M-p-2} is relatively prime to $(k^p-1)/(k-1)$ this means that

$$2(M-1) = h(k^p - 1)/(k-1) \tag{4.6}$$

where h is a positive integer; and each treatment pair then occurs hk^{M-p-2} times as a neighboring pair, if the desired design is achieved. When $p = 2$, we obtain that $M-1$ is a multiple of $(k+1)/2$ for k odd $(k = 1, 3, 5, \ldots)$ and of $k+1$ for k even $(k = 2, 4, 6, \ldots)$. (For $k = 3$, $M-1$ must not only be a multiple of 2 but must also be ≥ 3 to satisfy $p < M-1$; that is, $h \geq 2$.) For larger p, except for the case $p = 3$, $k = 2$, when $M-1$ must be at least 7, we find that the number of observations $n = k^{M-1}$ is always at least 2^{15} (and $M - 1 \geq 13$); see Table 1. This removes such designs, if they exist, from the realm of practicality for almost all applications, and suggests the usefulness of future investigation of smaller designs which are not balanced but are close to it.

When $p = 1$ we can obtain a design of the required form by an extension of the method of Theorem 4.2 to more than 2 dimensions.

TABLE 1
Minimum value of $M - 1$ satisfying (4.6)

		p			
		2	3	4	5
k	2	3	7	15	31
	3	2	13	20	121
	4	5	21	85	341
	5	3	31	78	781

THEOREM 4.3. *The following design with v treatments, each replicated v^{M-2} times, with one observation in each cell of an $(M-1)$-dimensional hypercube of side v, has all $N_{ii'}$ $(i \neq i')$ equal and all $N_{ii} = 0$. In the $(i_2, \ldots, i_M)$ cell of the hypercube place the treatment number*

$$\sum_{j=2}^{M} \sum_{r=1}^{i_j} (-1)^r (r-1),$$

reduced (mod v).

When $p > 1$ it is difficult to find designs with the right structure. Here are two of the simplest examples. Both constructions can be described in the geometric terms of Kishen (1949), and we represent them in the resulting notation without giving detailed description of the geometric configurations.

EXAMPLE 4.2. For $k = 2$, $p = 2$, $M - 1 = 3$, the levels of the last 3 factors are represented as $x_1, x_2, x_3 = 0$ or 1. The treatment in cell (x_1, x_2, x_3) is $(x_1 + x_3, x_2 + x_3)$ mod (2, 2). Rewriting the resulting treatments mod (2, 2) as 0, 1, 2, 3 (by considering them as binary numbers), we obtain

```
0 1   3 2
2 3   1 0
```

for the two "layers" for $x_3 = 0, 1$.

EXAMPLE 4.3. For $k = 3$, $p = 2$, $M - 1 = 4$, letting the levels of the last 4 factors be represented as $x_1, x_2, x_3, x_4 = 0$, 1 or 2, and letting the treatment in cell (x_1, x_2, x_3, x_4) be $(x_1 + 2x_3 + 2x_4, x_2 + x_3 + 2x_4)$ mod (3, 3), we obtain an equineighbored design in which each treatment pair is adjacent six times. Rewriting the treatments 0, 1, $\cdots$, 8 by considering them as ternary numbers, and letting x_1, x_2 be the rows and columns of "little" squares below and x_3, x_4 be the "large" rows and columns, we obtain

```
012   867   453
345   201   786
678   534   120

786   345   201
120   678   534
453   012   867

534   120   678
867   453   012
201   786   345.
```

Not all choices of parallel pencils of 2-flats for treatments in Kishen's construction will yield an equineighbored design. For example, the assignment $(x_1 + x_2 + x_4, x_2 + x_3 + x_4)$ of treatments will yield a hypercube, but it is not equineighbored.

Despite considerable trial and error, the authors have been unable to find a solution to the case $k = 5$, $p = 2$, $M - 1 = 3$. Simple trial and error methods break down after a few "layers" have been obtained.

A direct analogue for $M > 2$ of the consideration of several blocks when $M = 2$ is the consideration of several uncorrelated Latin squares or hypercubes. The use of several hypercubes extends the parameter values k, M, p for which designs with the right structure for weak universal optimality can be obtained. We shall not pursue this here. Similarly, the considerations can be extended to the larger but less practical design settings in which block size (for $M = 2$) or possibly unequal Latin rectangle or hyperrectangle sides are multiples of v, so that $\mathscr{X}^*$ still consists of designs with classical orthogonality and balance properties under V_0. The considerations of R.M. Williams (1952) and of Kiefer (1960) can be viewed as treating, for $M = 2$ and a different covariance structure, a single block ($b = 1$) of length k much greater than v. For the approach of the present paper, k would be a multiple of v and $\mathscr{X}^*$ would consist of designs for which each treatment appears k/v times. No design in $\mathscr{X}^*$ satisfies (3.3)(a′), but (for example) designs with all $N_{ii} = 0$ and all $N_{ii'}$, as nearly equal as possible will be close to weak universal optimality when k is large, if $\rho \geq 0$ is assumed.

5. Non-orthogonal settings.

5.1. *Optimality.* Consider the designs with $M = 2$ factors. Take $\mathscr{X}$ to be the class of block designs with b blocks, v treatments and block size k. Thus, as in the case $M = 2$ of Section 4, we assume

(i) $n_1 = v, \qquad n_2 = b$
(ii) $N(\cdot, j) \leq k \qquad (1 \leq i \leq v, 1 \leq j \leq b)$,

where (ii) is strengthened to equality if we impose complete use of all plots in each block, and where $N(i, j)$ is often restricted to be 0 or 1 if $k \leq v$. With or without such restrictions, the universally optimum designs, $\mathscr{X}^*$, are the balanced block designs with parameters b, v, k, when such designs exist; see Kiefer (1958, 1975). We shall only consider the case of $k < v$ here, the case of BBD's when $k > v$ requiring only slight modification in our development and being of less practical interest; the case $k = v$ was treated in Section 4. As usual, we use r and λ to denote the replication and treatment intersection numbers.

For BIBD let T_i be the sum of the observations on treatment i, i.e., $T_i = Y_{(i,\cdot)}$; and let B_j be the sum of observations on block j, $B_j = Y_{(\cdot,j)}$. LET A_i be the set of blocks in which treatment i occurs: $A_i = \{j \mid N(i, j) = 1\}$. Defining t_i as in (4.2), we have by the standard analysis of the BIBD

$$\hat{t}_{0i} = Q_i/\lambda v,$$

where

$$Q_i = kT_i - \textstyle\sum_{j\in A_i} B_j.$$

As in Section 4, $\hat{t}_{0i}$ is then the LS estimator of $\alpha_i^{(1)} - v^{-1} \sum_1^v \alpha_j^{(1)}$.

It is slightly more convenient to work with $Q = (Q_1, \cdots, Q_V)^T$ rather than $\hat{t}_0$, so we define

$$\bar{D}(X^*, V) = \mathrm{Cov}(Q \mid V).$$

We distinguish the position of an observation in a block. So let $g(j, r)$ be the treatment number of the rth observation in the jth block. We now list the observations $Y_{(i,j)}$ by lexicographic order of (j, r), not (i, j). Assuming a fixed covariance structure determined only by position in a block, but assuming zero covariance between blocks, we may write

$$\begin{aligned}\mathrm{Cov}(Y_{(i,j)}, Y_{i',j'}) &= \sigma^2 && \text{if } |i-i'|+|j-j'|=0,\\ &= \rho_{rs}\sigma^2 && \text{if } j=j' \text{ and } i=g(j,r),\ i'=g(j,s),\\ &= 0 && \text{otherwise.}\end{aligned}$$

Define the $k \times k$ matrix $V^* = \{V^*_{rs}\}$ by $V^*_{rs} = \sigma^2\rho_{rs}$, where $\rho_{rr} = 1$. We assume this V^* is positive definite. Then the overall covariance structure is given by

$$V = I_b \otimes V^*.$$

Define $k \times v$ matrices P_j ($j = 1, \cdots, b$) by

$$\begin{aligned}\{P_j\}_{ri} &= 1 && \text{if } g(j,r) = i,\\ &= 0 && \text{otherwise.}\end{aligned}$$

Then we obtain

$$\bar{D}(X^*, V) = k^2 \textstyle\sum_{j=1}^{b} P_j^T W P_j, \tag{5.1}$$

where

$$W = (I_k - k^{-1}J_{k,k})V^*(I_k - k^{-1}J_{k,k}).$$

For $j \in A_i$ define the position of treatment i by $h(i, j) = r$ if $g(j, r) = i$. Then the (i, i')th entry of (5.1) is

$$\mathrm{Cov}(Q_i, Q_{i'}) = k^2 \textstyle\sum_{j\in A_i\cap A_{i'}} W_{h(i,j),h(i',j)}. \tag{5.2}$$

Thus

$$\mathrm{trace}\ \bar{D}(X^*, V) = bk^2\ \mathrm{trace}\ W.$$

This means trace $\bar{D}(X^*, V)$ is independent of the choice of BIBD X^* so that every BIBD is equally good in terms of A-optimality for every V, and condition (b′) of (3.3) is automatically satisfied.

Now specialize to the NN covariance structure within a block. Put

$$\rho_{rs} = \begin{cases} 1 & \text{if } r = s,\\ r & \text{if } |r-s| = 1,\\ 0 & \text{otherwise.}\end{cases}$$

From (5.2) we then have

$$\begin{aligned}\sigma^{-2}\,\mathrm{Var}(Q_i) &= r[k(k-1) - 2\rho(k+1)] + 2\rho k e_i,\\ \sigma^{-2}\,\mathrm{Cov}(Q_i, Q_{i'}) &= -\lambda[k - 2\rho(k+1)] + k\rho[e_{ii'} + kN_{ii'}], \qquad i \neq i',\end{aligned} \tag{5.3}$$

where

(i) e_i is the number of blocks with treatment i at an end, that is,

$$e_i = \#\{j \mid g(j,1) = i \ \text{ or } \ g(j,k) = i\};$$

(ii) $e_{ii'}$ is the number of blocks in $A_i \cap A_{i'}$ in which i occurs at an end plus the number where i' occurs at an end, that is,

$$\begin{aligned}e_{ii'} = {}&\#\{j \mid j \in A_i \cap A_{i'}, g(j,1) = i \ \text{ or } \ g(j,k) = i\}\\ &+ \#\{j \mid j \in A_i \cap A_{i'}, g(j,1) = i' \ \text{ or } \ g(j,k) = i'\},\end{aligned}$$

a block counting twice if i and i' are at the two ends of it; and

(iii) $N_{ii'}$ is the number of times i and i' are adjacent in a block:

$$N_{ii'} = \#\{j \mid g(j,r) = i, g(j,s) = i', |r-s| = 1\}.$$

Simple enumeration gives

$$\textstyle\sum_{i'(\neq i)} e_{ii'} = 2r + (k-2)e_i \tag{5.4}$$

and

(5.5) $$\sum_{i'(\neq i)} N_{ii'} = 2r - e_i.$$

From (5.2), (5.4), (5.5), and the fact that $\sum_i e_i = 2b$ for every design in $\mathscr{X}^*$, we thus have

THEOREM 5.1. *A* BIBD *is weakly universally optimum for the* NN *model if all the quantities*

$$e_{ii'} + kN_{ii'} \qquad (i \neq i')$$

are equal.

(See just below (3.3), regarding uniqueness.)

A sufficient condition is that all the $N_{ii'}$ are equal and all the $e_{ii'}$ are equal. To make all $\mathrm{Var}(\hat{t}_{0i})$ equal we only need all the e_i equal; this is relevant for handling criteria of the form $\sum_i f(\mathrm{Var}(\hat{t}_{0i}))$ with f convex.

The condition of Theorem 5.1 imposes an extra condition on the parameters of a BIBD for the theorem to be usable to show it is weakly universally optimum. Further enumeration gives $\sum_{i<i'} e_{ii'} = 2b(k-1)$ and $\sum_{i<i'} N_{ii'} = rv - b$. Thus equality of all $e_{ii'} + kN_{ii'}$ implies

(5.6) $$v(v-1) \mid 2b(k-1)(k+2).$$

Since $2b(k-1)(k+2) = v(v-1)[2\lambda(k+2)/k]$, (5.6) is equivalent to $k \mid 4\lambda$. If all $N_{ii'}$ are equal, condition (5.6) certainly holds since equality of the $N_{ii'}$ implies

(5.7) $$k \mid 2\lambda$$

(in which case all $N_{ii'} = 2\lambda/k$); in fact, (5.7) is equivalent to (5.6) if k is odd, and, more simply then, to $k \mid \lambda$.

Our main interest in the remainder of this section is in the construction, for given k and v, of BIBD's which satisfy the condition of Theorem 5.1 and for which b *is as small as possible*; the obvious design with $b = v(v-1) \cdots (v-k+1)/2$ (half the permutations of length k) is of little practical interest. We hereafter assume $k \geq 3$ to avoid trivialities.

It is clear that not all (v, b, k) for which a BIBD exists will satisfy (5.6). The familiar (7, 7, 3) is an example. Even when a BIBD satisfies (5.6) there need not exist a BIBD satisfying the condition of Theorem 5.1, as the following example shows.

EXAMPLE 5.1. For $v = b = 7$ and $k = 4$ the condition of Theorem 5.1 requires that all $e_{ii'} + 4N_{ii'}$ equal 6. Thus $N_{ii'} = 1$ and $e_{ii'} = 2$. Since all BIBD's with $v = b = 7$ and $k = 4$ are isomorphic if order within blocks (and of blocks) is ignored, we may assume the first ordered block is (1, 2, 3, 4). Since $\lambda = 2$, treatments 2 and 3 occur together in one other block. To make $e_{23} = 2$ that block must be $(2, x, y, 3)$ or its reverse. But by the known structure of this BIBD, neither x nor y can be 1 or 4. So we may take the block to be (2, 5, 6, 3). Repeating the argument, to make $e_{56} = 2$ we require a block $(5, x', y', 6)$, or its reverse, and the known structure of the BIBD forces $\{x', y'\} = \{1, 4\}$. Picking any fourth block of the BIBD and repeating the argument, one finds that one must be led either to the repetition of one of the first three blocks, which is not allowed, or else to another cycle of three blocks, with one block hence left over. (The latter actually never occurs.) In either case, the design cannot be completed with the desired structure.

An argument that sometimes works to construct designs with minimum b for which all $e_{ii'}$ and $N_{ii'}$ are equal, when $k = 4$, will now be illustrated for the above example. Take (1, 2, 5, 3) as the initial block. This is a difference set with the additional property that the difference ± 1, ± 3, ± 2 (mod 7) between pairs of *successive* treatments give all the nonzero residues (mod 7) exactly once. It is easy to see that all the $N_{ii'}$ are equal in the BIBD with $b = 7$ obtained by developing this initial block in the usual fashion, the jth block being obtained by adding $j - 1$ (mod 7) to each element of (1, 2, 5, 3). This method is considered more generally later. Moreover for any pair $\{i, i'\}$ with $i - i' \equiv \pm 3$ (mod 7) the pair occurs once as the middle pair of a block, contributing 0 to $e_{ii'}$, and once where one of i, i' is at an end. Thus for the 7 such $\{i, i'\}$ we have $e_{ii'} = 1$. Similarly for the 7 pairs with $i - i' \equiv \pm 2 \pmod 7$ we have $e_{ii'} = 3$ and for the remaining 7 pairs with $i - i' \equiv \pm 1 \pmod 7$ we have $e_{ii'} = 2$. Now consider the 7 blocks obtained from developing (5, 1, 3, 2). All the $N_{ii'}$ are still equal but the roles of $i - i' \equiv \pm 2$ and $i - i' \equiv \pm 3 \pmod 7$ have been reversed. Hence in the BIBD with all 14 blocks we have all $N_{ii'} = 2$ and all $e_{ii'} = 4$. Moreover the previous paragraph implies that $b = 14$ is the smallest value of b for which the condition of Theorem 5.1 can be satisfied when $v = 7$, $k = 4$.

The case $k = 3$ always yields a solution. The result of Hanani (1961) asserts that the usual necessary conditions for existence of a BIBD with given (v, b, k, r, λ) are sufficient when $k = 3$. Thus a design exists if and only if

$$\lambda(v-1) \equiv 0 \pmod 2, \qquad \lambda v(v-1) \equiv 0 \pmod 6.$$

We also recall that, as noted just below (5.7), the latter is equivalent to (5.6) for k odd as in the present case.

Thus Hanani's conditions and (5.6) reduce to the consideration of a family of designs depending on two nonnegative integers m and v with

$$k = 3, \quad \lambda = 3m, \quad b = mv(v-1)/2, \quad r = 3m(v-1)/2, \tag{5.8}$$

with m even or m and v odd.

THEOREM 5.2. *For $k = 3$ condition* (5.8) *is necessary and sufficient for the existence of a design satisfying the condition of Theorem* 5.1.

PROOF. When $k = 3$, each adjacent pair of varieties i, i' in the same block contributes 1 to each of $N_{ii'}$ and $e_{ii'}$; thus it suffices to order the blocks of a BIBD satisfying (5.8) in such a way that all $N_{ii'}$ are equal. Since $\pi_{ii'} = \lambda - N_{ii'}$ is the number of blocks in which both i and i' occur at the ends, it is sufficient to make all $\pi_{ii'}$ equal. Given a BIBD satisfying (5.8), write each block as the triple $\{\tau_1, \tau_2, \tau_3\}$ of subsets of size two contained in it. In what follows, two blocks of the design that contain identical elements are considered different. Let S_τ be the set of all blocks containing a fixed subset τ of size 2. For any p different subsets $\tau_1, \cdots, \tau_p$ of size 2 $\left(1 \le p \le \binom{v}{2}\right)$ the number of distinct blocks in $\cup_{i=1}^{p} S_{\tau_i}$ is at least $p\lambda/3 = mp$, since each τ occurs in λ blocks of the design and there are three τ's per block. Using a theorem of Agrawal (1966) on m-ple systems of distinct representatives, we conclude that we can select a collection H_i of m blocks in each S_{τ_i}, $1 \le i \le \binom{v}{2}$, with the H_i disjoint. Each of the m blocks in H_i is then ordered so that the pairs of treatments in τ_i occur at the ends. This makes all the $\pi_{ii'} = m$. □

When $k = 4$, although Hanani's work again shows that the usual necessary conditions for existence of a BIBD are sufficient, those conditions together with (5.5) no longer guarantee the existence of a design satisfying the condition of Theorem 5.1, as Example 5.1 shows. Thus although we have a comprehensive picture of the combinatorial considerations associated with our optimality criteria for the NN model when $k = 2$, 3 or v (Section 4), the combinatorics seem much more difficult for other values of k.

5.2. *Equineighbored* BIBD *construction.* For large k the contribution of $N_{ii'}$ in (5.3) is much larger than that of $e_{ii'}$, since $kN_{ii'}$ is, on average, $k/2$ times as large as $e_{ii'}$. Also, the equality of the $N_{ii'}$ implies equality of the e_i by (5.5). We shall therefore devote most of the rest of this section to the construction of BIBD's with all the $N_{ii'}$ equal, which we call *equineighbored* BIBD's or EBIBD's. Such designs, besides being A-optimum, make all the $\mathrm{Var}(\hat{t}_i)$ equal and can be expected to perform fairly efficiently for criteria other than A-optimality.

The NN model leads us to seek the equal appearance of unordered pairs of neighbors, not of ordered ones. This distinguishes the present work from the problems treated by such authors as E.J. Williams (1949), Hedayat and Afsarinejad (1975, 1978), and Hedayat and Magda (1979) in the construction of repeated measurement designs, who distinguish the order of pairs. In seeking to minimize b for given v and k we may typically construct designs with half the value of b required by those authors. However, it must be emphasized that we restrict our designs to be BIBD's, which Hedayat and Afsarinejad (1975) do not. For example, their design with $v = 5$, $k = 3$, $b = 10$ is not a BIBD, and no BIBD with those parameters can have equal appearance of all ordered pairs, although by our Theorem 5.2 a BIBD of our type exists. Similarly for $v > k$ their design has $b \geq 2v$, which is not necessarily the case for our EBIBD's, which often exist with $b = v$. On the other hand, they construct designs with $k = (v + 1)/2$ and $b = 2v$ for all odd v, but these are not always BIBD's except in special cases, for example, if v is a prime power $\equiv 3 \pmod 4$. In this last case their design will of course automatically be an EBIBD of our type. The main aim of Hedayat and Afsarinejad (1975) is the construction, for given v, of designs with the minimum conceivable value $2v$ of b, which forces $k = (v + 1)/2$. On the other hand, for fixed v and k, we want to construct equineighbored designs with minimum b. Obviously any equineighbored design can be replicated with blocks in reversed order to yield a design of the type Hedayat and Afsarinejad consider. If $N_{ii'} = 1$ for the equineighbored design, the resulting Hedayat and Afsarinejad design has minimum b for given v, k and minimum k for given v and b.

The designs of Hwang (e.g., 1973) and others have NN structure for *circular* (cyclic) block designs with $k = (v - 1)/2$ because of the extra pair of neighbors per block. Those designs are easily seen to be optimum in their setting, by our argument. Powers of a primitive element can be used for prime power v construction there, but not in our problem. C.S. Cheng has recently shown how Hwang's designs can sometimes be modified to yield designs for our problem, but they do not always attain the minimum b we seek.

Assuming (v, b, k) are such that a BIBD exists and that (5.7) is satisfied, we want to construct an EBIBD, and, for given v and k, to do this for smallest possible b. One may try to make all the $N_{ii'}$ equal by reordering treatments within each block of a given BIBD, but we have no general algorithm for doing this even for the case $k = 3$ (which is treated in alternate fashion in Theorem 5.2). We shall give methods of generating EBIBD's based on the difference set technique of developing cyclic designs.

Let $\{a_1, \cdots, a_k\}$ be a difference set (mod v), i.e., a set of integers such that amongst all the difference $\pm(a_i - a_j)$ $(1 \leq i, j \leq k, i \neq j)$ each nonzero residue (mod v) $1, \cdots, v - 1$ appears the same number of times. When v is odd and $k = (v + 1)/2$ define an *equineighbored difference set* (mod v) as a difference set $\{a_1, \cdots, a_k\}$ (mod v) with the additional property that amongst the $2(k - 1)$ *successive differences*

$$\pm(a_i - a_{i+1}), \qquad 1 \leq i \leq k - 1,$$

each nonzero residue appears exactly once. We recall that for the analogous case $v = k$ ($M = 2$ in Section 4), the property is similar to that held by the first row and column of the squares generated by Theorem 4.1. We give the following without proof:

THEOREM 5.3. *If* $\boldsymbol{a} = \{a_1, \cdots, a_k\}$ *is an equineighbored difference set* (mod v) *where* $k = (v + 1)/2$, *then the symmetric* BIBD *formed by developing* $\boldsymbol{a}$ *as the initial block so that the* $(1 + j)$th *block is*

$$(a_1 + j, a_2 + j, \cdots, a_k + j) \pmod v$$

$(0 \leq j \leq v - 1)$ *is an* EBIBD *with all* $N_{ii'} = 1$.

Difference sets are known for many cases where $k = (v + 1)/2$. (See Hall, 1967; Baumert, 1971.) Since the complement of a difference set is a (smaller) difference set of $k - 1$ elements (mod v), it is usually listed in these references. For all $v \leq 35$ of the form $4\lambda - 1$, such sets my be constructed by one of the following methods:

(1) If $v = p^m$, p prime, $v \equiv 3 \pmod 4$, then all the nonzero quadratic residues (mod v) yield a difference set with $k = (v - 1)/2$ elements (cyclic if $m = 1$).

(2) If $v = pq$ where p and $q = p + 2$ are twin primes, a slightly more complex recipe (Hall, page 141) yields a difference set of $k = (v - 1)/2$ elements.

These are the only two methods we discuss here.

In case (1) with $m = 1$, the element 0 may be adjoined to the nonzero quadratic residues to yield a difference set (mod v) of k elements which, in their natural order, yield an equineighbored difference set.

THEOREM 5.4. *If v is prime and $v \equiv 3 \pmod 4$, then*

$$\left\{0^2, 1^2, 2^2, \cdots, \left(\frac{v-1}{2}\right)^2\right\} \pmod v$$

is an equineighbored difference set.

PROOF. It is well known that this is a difference set. Moreover,

$$\pm(a_{i+1} - a_i) \equiv \pm(2i - 1),$$

$1 \leq i \leq (v - 1)/2$, and these are exactly the $(v - 1)$ nonzero residues (mod v). □

We remark that it is in general necessary to establish that $\boldsymbol{a}$ is a difference set, as well as that $\pm(a_{i+1} - a_i)$ are distinct. For example (0, 1, 6, 2) (mod 7) has distinct successive differences (±1, ±5, ±4), but is not a difference set because of the twelve differences $a_i - a_j$ $(i \neq j)$ there are three appearances of ±1, two of ±2, and one of ±3. When developed, this block thus does not yield a BIBD.

In case (1) with $m > 1$ the quadratic residues of $GF(p^m)$ with 0 included again form a difference set with $k = (v + 1)/2$ elements. However, there is no obvious ordering as there was in the case of Theorem 5.4. It is tempting to let x be a primitive element and consider the successive differences of the set $x^2, x^4, \cdots, x^{(v-1)}$; however, the difficulty is that adjoining the remaining element 0 anywhere in the sequence need not work, as is shown by the case $v = 11$, $x = 2$, $(x^2, x^4, x^6, x^8, x^{10}) = (4, 5, 9, 3, 1) \pmod{11}$. The case (2), above, also yields no simple mechanism. We construct a difference set of $(v - 1)/2$ elements (mod v) by the cited method, and its complement is then a difference set of the required size, $(v + 1)/2$. But again there is no obvious ordering scheme that produces an equineighbored difference set. This inability to find a scheme analogous to that of Theorem 5.4 in the cases of the previous two paragraphs has led us to the following simple routine.

Start with a given unordered difference set $c_1, \cdots, c_{(v+1)/2}$. We shall try to arrange the elements of this set in a special sequence $a_1, \cdots, a_{(v+1)/2}$ by a simple iterative method which adjoins one element at a time to the ends of the sequence already constructed. First arrange all the pairs $\{c_i, c_j\}$ $(i \neq j)$ in a list consisting of $(v - 1)/2$ groups in such a way that group h $(1 \leq h \leq (v - 1)/2)$ consists of all those pairs $\{c_i, c_j\}$ with $c_i - c_j = \pm h$. Begin with some pair $\{c_i, c_j\} = \{a_{2,1}, a_{2,2}\}$. In general when there are s out of the c_i in the sequence label them

$$a_{s,1}, \cdots, a_{s,s}. \tag{5.9}$$

At this stage we must be sure that the following are deleted from the list of pairs (since none is available to yield a difference on a subsequent step):

(1) any group h for which

$$a_{s,r} - a_{s,r+1} = \pm h \pmod v$$

for some r with $1 \leq r \leq s - 1$;

(2) any pair $\{c_i, c_j\}$ for which c_i or $c_j = a_{s,r}$ for some r with $2 \leq r \leq s - 2$;

(3) the pair $\{a_{s,1}, a_{s,s}\}$ for $s > 2$.

To add the next, $(s + 1)$st, element to the sequence (5.9), we proceed as follows. Choose an undeleted group that contains the *smallest* number of undeleted pairs $\{c_i, c_j\}$ among undeleted groups. If the chosen group contains an undeleted $\{c_i, c_j\}$ one member of which, say c_i, is either $a_{s,1}$ or $a_{s,s}$, then adjoin the other member c_j to the appropriate end to form $a_{s+1,1}, \cdots, a_{s+1,s+1}$. If no such $\{c_i, c_j\}$ exists then we are "blocked" and we choose instead a group with the next smallest number of undeleted $\{c_i, c_j\}$. Special rules may be adopted if a previously "blocked" group is still undeleted at some stage, if there are several "tied' groups, and so on.

It is easy to construct examples where this simple routine can fail. However, in every case of the form $v = b = 4\lambda - 1$ and $k = r = 2\lambda$, that we have been discussing for $\lambda \leq 9$, the method has worked on the first try after less systematic attempts had failed. For larger values of λ, when there are only four sparse groups left, the various combinations of remaining additions at both ends were considered, to avoid failure. The routine is also simple to program on a computer and can be extended to the case of more than one initial block and to generalized difference sets, discussed later. The equineighbored difference sets we obtained in the cases $v = 15$, 27 and 35 for $\lambda \leq 9$, not covered by Theorem 5.4, are as follows.

(5.10)
v = 15: (11, 14, 4, 12, 8, 2, 3, 1) (mod 15),
v = 27: (020, 102, 111, 202, 221, 121, 022, 021, 001, 120, 000, 110, 211, 100) (mod (3, 3, 3)),
v = 35: (19, 6, 30, 5, 8, 31, 23, 24, 26, 20, 2, 32, 18, 22, 15, 34, 25, 10) (mod 35).

In the case $v = 27$ the differencing is carried out component-wise and the difference set is not cyclic. In all cases we started with an unordered difference set. The only case in this family of parameter values with $v \leq 50$ not covered by Theorem 5.4 or (5.10) is $v = 39$ for which it is a consequence of a theorem of Hall and Ryser (see Baumert, 1971, page 25) that no cyclic difference set exists. There is a symmetric BIBD in that case but we have not attempted to reorder it (by necessarily different methods), and thus do not know whether there is a symmetric EBIBD for $\lambda = 10$.

Since the classification of all difference sets is itself incomplete we cannot hope for a simple listing of all equineighbored sets. We now make some brief remarks on the classification of equineighbored difference sets, for the simple case when v is prime for the family considered above.

We consider equineighbored difference sets to be equivalent if they can be transformed into each other by one or more of the following transformations: (1) addition to each element $a_1, \cdots, a_{(v+1)/2}$ of an arbitrary integer $\boldsymbol{a}$ (and reduction (mod v)); (2) multiplication of each element by an arbitrary integer $c \not\equiv 0 \pmod v$, and reduction (mod v) (with c not restricted to be a "multiplier" in the technical sense of the difference set literature, as defined by Hall or Baumert); (3) reversal of the order $a_1, \cdots, a_{(v+1)/2}$ to $a_{(v+1)/2}, \cdots, a_1$. Note that under (1) the successive differences $d_1, \cdots d_{(v-1)/2}$ of an equineighbored difference set are preserved in the same order, where $d_i = a_{i+1} - a_i$ $(i = 1, \cdots, (v - 1)/2)$. Since we distinguish between d_i and $-d_i$, (2) and (3) do not preserve the order of the d_i's; nevertheless, it seems natural to regard difference sets which can be obtained from each other through these transformations, as equivalent.

EXAMPLE 5.2. In the case $\lambda = 2$, $v = 7$, the two sets of successive differences (1, 3, 5) and (1, 5, 4) identify two equivalence classes of equineighbored difference sets. Typical members of each are respectively (0, 1, 4, 2) and (1, 2, 0, 4). Thus neither of these can be transformed into the other by tranformations (1), (2), or (3), or by combinations of them. Under (2) and (3), (1, 3, 5) and (1, 5, 4) together generate half of the $2^3 3!$ possible difference triples $(\pm\alpha, \pm\beta, \pm\gamma)$ where (α, β, γ) is a permutation of (1, 2, 3); the other half cannot occur. For example, (1, 2, 3) occurs but (1, 2, 4) does not, nor indeed does any permutation of $\pm(1, 2, 4)$. For large v the picture is more complex.

We now give a few additional examples to illustrate the extension to settings where more than one initial block must be developed. There is often more than one possible set of initial blocks, but we give only one in each example. For brevity we include among these examples only one of the common infinite "families" of BIBD's (Example 5.9), but it is easy to extend the method to others. All the examples yield the smallest b for which (5.7) is satisfied, except Example 5.5.

All of these examples will be treated by the use of an elementary development we next set forth.

Suppose, for a given v, that we have a set of $\bar{b}$ initial blocks of size k, that together constitute a difference set. We now describe a simple device for obtaining a set $\mathscr{B}$ (say) of $\bar{b}k$ or $\bar{b}k/2$ initial blocks when k is odd or even, respectively, so that $\mathscr{B}$ is an equineighbored difference set which can be used as an initial block set to generate an EBIBD with $b = v\bar{b}k$ or $v\bar{b}k/2$. To this end, we recall from the discussion of the case $M = 2$ of complete blocks in Section 4, but with block size now k, that the k rows of a $k \times k$ Latin square given by Theorem 4.2 for k odd, or the first $k/2$ rows for k even, yield an equineighbored complete block design $\mathscr{D}$ (say). For each of the $\bar{b}$ initial blocks of size k, substitute its k symbols into $\mathscr{D}$ to yield k or $k/2$ blocks of size k; we call this the $\mathscr{D}$-development of the initial block. Doing this for all $\bar{b}$ initial blocks yields the set $\mathscr{B}$. From the nature of $\mathscr{D}$ it follows that, although any of the $\bar{b}$ initial blocks need not be a difference set, every difference occurring in the initial block occurs proportionally often as a NN difference in the $\mathscr{D}$-development of that block. Since the $\bar{b}$ initial blocks are a difference set, we have

THEOREM 5.5. *The set $\mathscr{B}$ of blocks just described is an equineighbored difference set.*

EXAMPLE 5.3. Even when $k = 3$, it is convenient to have a simple explicit formula for the construction of EBIBD's. Consider the case $v = 7$, $k = 3$. It follows from (5.7) that the minimum b is 21. The initial blocks (1, 2, 4), (4, 1, 2), (2, 4, 1) may be constructed by using Theorem 5.5, and contain, among them, each pair of successive differences (± 1, ± 2, ± 3) twice. Thus these blocks developed by adding all residues (mod 7) yield an EBIBD. Note that for this small value of k it was possible to choose the initial blocks to be cyclic permutations of (1, 2, 4).

EXAMPLE 5.4. Consider the case $k = 4$, $v = 13$. The familiar design with $\lambda = 1$ is obtained by developing (0, 1, 3, 9) (mod 13). A second initial block is needed to achieve (5.7). From Theorem 5.5 we take it to be (1, 9, 0, 3), and then each successive difference pair $\pm i$ ($1 \le i \le 6$) appears once in the set of two initial blocks. Clearly since $k = 4$ the second initial block could not be obtained by a cyclic permutation of the first.

EXAMPLE 5.5. Here is an example based on an acyclic difference set. For $v = 16$, $k = 6$, the usual symmetric design with $\lambda = 2$ is developed from an initial block consisting of a difference set of six suitably ordered elements mod (2, 2, 2, 2), listed as the first line of (5.11), below. By Theorem 5.5 we obtain 3 initial blocks which can be developed into an EBIBD with $b = 48$. (We do not know whether a BIBD with $b = 24$, which exists, can be ordered to be equineighbored.)

$$\begin{gathered}((0010), (0001), (0011), (1100), (1000), (0100)),\\ ((0001), (1100), (0010), (0100), (0011), (1000)),\\ ((0011), (0010), (1000), (0001), (0100), (1100)).\end{gathered} \tag{5.11}$$

EXAMPLE 5.6. Here are two illustrations in which the basic design (non-EBIBD with minimum b) which we expand, is based on more than one initial block. For $v = 13$, $k = 3$, $b = 26$, a BIBD with $\lambda = 1$ is often constructed by developing (mod 13) the two initial blocks (1, 3, 9) and (2, 5, 6). It is easily seen that the six initial blocks obtained from the cyclic permutation of these two blocks yields an equineighbored difference set (as in Example 5.3). This amounts to using Theorem 5.5 on the original two initial blocks.

For $v = 9$, $k = 4$, $b = 18$, a BIBD with $\lambda = 3$ is often obtained by developing the two initial blocks (1, 4, 0, 2) and (1, 0, 4, 6) (mod 9). As in Example 5.4, a set consisting only of cyclic permutations of these will not work, but we may use Theorem 5.5 to adjoin the two blocks (4, 2, 1, 0) and (0, 6, 1, 4), to obtain an equineighbored difference set and thus an EBIBD with the minimum b value of 72. Interestingly, each of the last two initial blocks contains a repeated neighbored difference.

EXAMPLE 5.7. For an example involving generalized differences including an "∞" treatment, we consider the case $v = 12$, $k = 4$, $b = 33$ where a design with $\lambda = 3$ is often obtained by developing (mod 11) the three initial blocks (0, 3, 7, 1), (0, 1, 3, 9), (∞, 0, 1, 5). If we use Theorem 5.5 to adjoin (3, 1, 0, 7), (1, 9, 0, 3) and (0, 5, ∞, 1), we obtain a set of six initial blocks with equineighbored successive differences, which can be developed to yield an EBIBD with the minimum b of 66. Once more, each of the last two initial blocks contains a repeated neighbored difference.

EXAMPLE 5.8. We conclude our examples with an illustration involving mixed difference sets. The T_1 system of BIBD's consists of designs with (for t a positive integer) $v = 6t + 3$, $k = 3$, $b = (2t + 1)(3t + 1)$, and $\lambda = 1$. These are often obtained by developing, mod $(2t + 1)$, the initial set of $3t + 1$ blocks

$$(5.12)\qquad \begin{array}{ll} (i_1, (2t+1-i)_1, 0_2) & 1 \le i \le t, \\ (i_2, (2t+1-i)_2, 0_3) & 1 \le i \le t, \\ (i_3, (2t+1-i)_3, 0_1) & 1 \le i \le t, \\ (0_1, 0_2, 0_3). & \end{array}$$

It is easily verified that the blocks of (5.12) and their cyclic permutations (which arise from using Theorem 5.5) yield a set of $3(3t + 1)$ equineighbored initial blocks which can be developed into an EBIBD with the minimum $b = (6t + 3)(3t + 1)$.

We close with brief mention of nonorthogonal layouts for $M \ge 3$. For $M = 3$, optimality of Youden designs (YD's) or generalized Youden designs under V_0 is considered by Kiefer (1958, 1975), and for $M > 3$ results for Youden hyperrectangles are obtained by Cheng (1978). We limit discussion here to the case $M = 3$. If treatments are assigned to a $k \times mv$ array with $k < v$ and m an integer, there exists a "regular" YD if there is a BIBD with $b = mv$. The LS estimators for such a YD are in fact those of the BIBD consisting of the columns of the YD. The YD's are now our $\mathscr{X}^*$. If $m > 1$, it is possible that $N_{ii} > 0$ in a YD. The expression for $D(X^* \mid V)$ is slightly more complicated than it is for a Latin square, and will not be given here. As was the case for hypercubes, if $\rho > 0$ one is led to make all $N_{ii} = 0$ and to look for designs for which all $N_{ii'}$ are equal for $i \ne i'$. Unfortunately, simple number-theoretic considerations show that *no regular* YD *can be equineighbored.* In the nonregular setting in which there are $k_1 \times k_2$ generalized YD's with both $k_i > v$ and neither k_i divisible by v, this simple nonexistence argument no longer applies; we do not know the extent to which equineighbored generalized YD's may exist.

REFERENCES

AGRAWAL, H. (1966). Some generalisations of distinct representatives with applications to statistical designs. *Ann. Math. Statist.* **37** 525–526.

ATKINSON, A. C. (1969). The use of residuals as a concomitant variable. *Biometrika* **56** 33–41.

BARTLETT, M. S. (1938). The approximate recovery of information from replicated field experiments with large blocks. *J. Agric. Sci.* **28** 418–427.

BARTLETT, M. S. (1975). *The Statistical Analysis of Spatial Pattern.* Chapman and Hall, London.

BARTLETT, M. S. (1978). Nearest neighbour models in the analysis of field experiments (with discussion). *J. R. Statist. Soc. B* **40** 147–174.

BAUMERT, L. (1971). *Cyclic difference sets. Lecture Notes in Mathematics* No. 182. Springer, Berlin.
BERENBLUT, I. I. AND WEBB, G. I. (1974). Experimental design in the presence of autocorrelated errors. *Biometrika* **61** 427–437.
BESAG, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. R. Statist. Soc. B* **36** 192–236.
BICKEL, P. J. AND HERZBERG, A. M. (1979). Robustness of designs against autocorrelation in time. I. Asymptotic theory for location and linear regression. *Ann. Statist.* **7** 77–95.
CHENG, C.-S. (1978). Optimal designs for the elimination of multi-way heterogeneity. *Ann. Statist.* **6** 1262–1272.
CHENG, C.-S. (1979). Optimal incomplete block designs with four varieties. *Sankhya, Ser. B* **41** 1–14.
CONSTANTINE, G. M. (1980). On Schur-optimality. Preprint, Univ. of Indiana.
DÉNES, J. AND KEEDWELL, A. D. (1974). *Latin Squares and Their Applications.* English Universities, London.
DUBY, C., GUYON, X. AND PRUM, B. (1977). The precision of different experimental designs for a random field. *Biometrika* **64** 59–66.
GIOVAGNOLI, A. AND WYNN, H. (1980). A majorization theorem for the C-matrices of binary designs. *J. Statist. Planning and Inf.* **4** 145–154.
HALL, M. JR. (1967). *Combinatorial Theory.* Blaisdell, Waltham, Mass.
HANANI, H. (1961). The existence and construction of balanced incomplete block designs. *Ann. Math. Statist.* **32** 361–386.
HEDAYAT, A. AND AFSARINEJAD, K. (1975). Repeated measurement designs. I. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastara, ed.), 229–242. North-Holland, Amsterdam.
HEDAYAT, A., AND AFSARINEJAD, K. (1978). Repeated measurement designs. II. Characterizations, construction and optimality. *Ann. Statist.* **6** 61–62.
KERR, J. R., PEARCE, S. C. AND PREECE, D. A. (1973). Orthogonal designs for three-dimensional experiments. *Biometrika* **60** 349–358.
KIEFER, J. (1958). On the nonrandomized optimality and randomized nonoptimality of symmetric designs. *Ann. Math. Statist.* **29** 675–699.
KIEFER, J. (1960). Optimum experimental designs. V, with applications to systematic and rotatable designs. *Proc. Fourth Berk. Symp.* **2** 381–405.
KIEFER, J. (1975). Construction and optimality of generalised Youden designs. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastara, ed.), 333–353. North-Holland, Amsterdam.
KISHEN, K. (1949). On the construction of Latin and hyper-graeco-latin cubes and hypercubes. *J. Ind. Soc. Ag. Statist.* **2** 2–48.
KRUSKAL, W. (1968). When are Gauss-Markov and least squares estimators identical? A coordinate free approach. *Ann. Math. Statist.* **39** 70–75.
MARTIN, R. J. (1977). Spatial models with applications in sampling and experimental design. Ph.D. thesis, London School of Economics, University of London.
O'HAGAN, A. (1978). Curve fitting and optimal design for prediction (with discussion). *J. R. Statist. Soc. B* **40** 1–42.
PAPADAKIS, J. S. (1937). Méthode statistique pour des expériences sur champ. *Bull. Inst. Amél. Plantes à Salonique*, No. 23.
PATTERSON, H. D. (1950). The analysis of change-over trials. *J. Agric. Sci.* **40** 375–380.
PATTERSON, H. D. (1951). Change-over trials. *J. R. Statist. Soc. B* **13** 256–271.
PATTERSON, H. D. (1952). The construction of balanced designs for experiments involving sequences of treatments. *Biometrika* **39** 32–48.
RAO, C. R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. *Proc. V Berk. Symp.* **1** 355–372.
RIPLEY, B. D. (1977). Modelling spatial patterns (with discussion). *J. R. Statist. Soc. B* **39** 172, 212.
SACKS, J. AND YLVISAKER, D. (1966). Designs for regression problems with correlated errors. *Ann. Math. Statist.* **37** 66–89.
SACKS, J. AND YLVISAKER, D. (1968). Designs for regression problems with correlated errors; many parameters. *Ann. Math. Statist.* **39** 49–69.
SACKS, J. AND YLVISAKER, D. (1969). Designs for regression problems with correlated errors. III. *Ann. Math. Statist.* **41** 2057–2074.
WAHBA, G. (1971). On the regression design problem of Sacks and Ylvisaker. *Ann. Math. Statist.* **42** 1035–1053.
WAHBA, G. (1974). Regression design for some equivalence classes of kernels. *Ann. Statist.* **5** 925–934.
WILLIAMS, E. J. (1949). Experimental designs balanced for the estimation of pairs of residual effects of treatments. *Aust. J. Sci. Res.* **2** 149–164.
WILLIAMS, E. J. (1950). Experimental designs balanced for pairs of residual effects. *Aust. J. Sci. Res.* **3** 351–363.

WILLIAMS, R. M. (1952). Experimental designs for serially correlated observations. *Biometrika* **39** 151–167.

WATSON, G. S. (1967). Linear least squares regression. *Ann. Math. Statist.* **38** 1679–1699.

ZYSKIND, G. (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *Ann. Math. Statist.* **38** 1092–1109.

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE OF LONDON
LONDON SW7 2B7, ENGLAND

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720