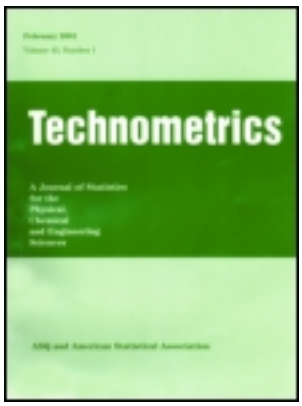

Technometrics
Publication details, including instructions for authors and subscription information:
http://amstat.tandfonline.com/loi/utch20

Computer-Aided Blocking of Factorial and Response-Surface Designs

R. Dennis Cook [a] & Christopher J. Nachtsheim [b]

[a] Department of Applied Statistics, University of Minnesota, St. Paul, MN, 55108

[b] Curtis L. Carlson School of Management, University of Minnesota, Minneapolis, MN, 55455

Version of record first published: 23 Mar 2012

To cite this article: R. Dennis Cook & Christopher J. Nachtsheim (1989): Computer-Aided Blocking of Factorial and Response-Surface Designs, Technometrics, 31:3, 339-346

To link to this article: http://dx.doi.org/10.1080/00401706.1989.10488556

TECHNOMETRICS, AUGUST 1989, VOL. 31, NO. 3

Computer-Aided Blocking of Factorial and Response-Surface Designs

R. Dennis Cook

Department of Applied Statistics
University of Minnesota
St. Paul, MN 55108

Christopher J. Nachtsheim

Curtis L. Carlson School of Management
University of Minnesota
Minneapolis, MN 55455

We consider the problem of blocking response-surface and factorial designs when block sizes are prespecified, often rendering standard approaches inapplicable. A generalization of Harville's (1974) algorithm is given for generating nearly D-optimal block designs and is applied in a number of representative settings. The algorithm is time and space efficient and can be used in conjunction with qualitative treatments for generating many of the classical designs—balanced and partially balanced incomplete block designs, for example.

KEY WORDS: Block designs; D optimality; Exchange algorithms; Incomplete blocks; k-exchange algorithm.

1. INTRODUCTION

Designing experiments can be a frustrating task when experimental units are scarce and not exchangeable. For example, in a recent process-control experiment, the objective was to study the main effects and two-factor interactions of four process variables on the texture of a finished food product in a plant scale-up. We became involved when the manufacturer realized that the process seemed hopelessly out of control because an unacceptably high proportion of the product was being scrapped. Only three days were available for experimentation and, since each change in the process required approximately 1.25 to 1.5 hours to establish a new equilibrium, only 6 runs per day, or 18 runs in total, were available. Further, since day-to-day variation in plant conditions could not be ruled out, it was hoped that a suitable design incorporating days as blocks could be devised. The constraints on the number of blocks and their sizes made application of standard design techniques somewhat elusive, however; a sensible distribution of a basic 2^4 factorial into three blocks was, at the very least, not obvious.

If four instead of three days had been available, it may have been possible to adapt standard approaches in a fairly straightforward manner. For example, a 2^4 factorial in four blocks of size 4 might have been appropriate, with the remaining two runs in each block used to provide replication. Demanding additional resources or tailoring the inquiry to fit those available may occasionally be effective, but often such tactics fail to provide a satisfactory solution.

In another instance, a team of chemical engineers and food scientists was concerned with the optimization of stability in a whipped topping. Here, stability relates to the amount of "melting" that occurs after the aerosol topping is dispensed. Previous experimentation had indicated that the most important factors were the amounts of two emulsifiers and the amount of fat. Curvature of the response in each of the three factors was also a strong possibility. These results, together with the objectives of the study, suggested the use of a three-factor central composite design to estimate the associated full quadratic response-surface model. The experiment was to be carried out in a pilot plant, which was available for five days. Since four batches per day could be produced in the pilot plant, at most 20 experimental units were available. If day-to-day effects are of concern, the experiment should be run in five blocks of size (at most) 4. The question again arises: How should the experimental units (runs) be partitioned into blocks (days)?

The literature is of limited help for the types of design problems illustrated by these examples. Some useful approaches to blocking response-surface designs were described by Box and Draper (1987, sec. 15.3). For example, a three-factor central composite design can be implemented usefully in three blocks of sizes 5, 5, and 7. Each of the first two blocks is composed of a half replication of the 2^3 factorial plus one center point. The third block consists of the six star points and an additional center point. If, in addition, the star points are carefully located (Box and Draper 1987, sec. 15.3), the block effects will be orthogonal to the treatment (regression) effects. Unfortunately, the block size requirements are not met

in our illustration, and only 17 of the 20 available units have been used.

In the simple examples discussed previously, application of classical design is difficult, if not impossible, without some modification of the sample or block size constraints. More complicated situations often arise, particularly in industrial settings. Experimental units often differ on a set of relevant attributes that force blocks of varying size. In such situations, an examination of "optimal" designs, along with a comparison of standard design schemes on various design criteria, can provide valuable insight regarding sensible blocking schemes.

1.1 Structure of Optimal Block-Design Problems

The structure of an optimal block design consists of three broad components—the design constraints, the model, and the optimality criterion—that depend on the particular application.

Design Constraints. For the purposes of this report, the design constraints consist of three components: (a) the sample size n, (b) the feasible region χ from which the $t \times 1$ vectors x_i $(i = 1, \ldots, n)$ of factor settings or treatment combinations can be obtained, and (c) the set E composed of the $N \geq n$ available experimental units. We assume that E can be partitioned into b available blocks such that the kth block consists of $N_k > 0$ homogeneous experimental units. Further, let $n_k \leq N_k$ denote the number of experimental units from the kth available block actually used in the experiment. If $N = n$, then $N_k = n_k$ for all k. Otherwise, the kth block in the actual design will be a subset of the kth available block, including the possibility of the empty set, so some of the available blocks may not appear as blocks in the design. Finally, associated with each experimental unit in the kth available block is a $b \times 1$ indicator vector of the form $u^T = (0, \ldots, 0, 1, 0, \ldots, 0)$ with a 1 in the kth position and zeros elsewhere. The collection of these indicator vectors is denoted by $U_n = \{u_j, j = 1, \ldots, N\}$.

Model. We are concerned only with designs under the additive blocking model

$$y_i = f_1^T(x_i)\tau + u_i^T\beta + \varepsilon_i, \qquad i = 1, \ldots, n, \quad (1)$$

where τ $(p \times 1)$ and β $(b \times 1)$ are unknown parameters and the errors ε_i are iid with mean 0 and constant variance σ^2. The component f_1 assumes various forms depending on the problem. In the response-surface case, f_1 will typically consist of polynomial terms in the factor levels. Qualitative treatments can be represented by choosing f_1 from the appropriate class of indicator vectors.

Model (1) can be represented in matrix form as

$$\begin{aligned} Y &= X_1\tau + X_2\beta + \varepsilon \\ &= X\alpha + \varepsilon, \end{aligned} \quad (2)$$

where $Y = (y_i)$ is the $n \times 1$ vector of responses, $\varepsilon = (\varepsilon_i)$ is the $n \times 1$ vector of errors, X_1 and X_2 are $n \times p$ and $n \times b$ matrices with rows $f_1^T(x_i)$ and u_i^T, respectively, $X = (X_1, X_2)$, and $\alpha^T = (\tau^T, \beta^T)$.

Design Criteria. Although several design criteria have been proposed, we deal with only two, D and D_s optimality. D-optimal designs are those that minimize the determinant of the covariance matrix $(X^TX)^{-1}\sigma^2$ of the least squares estimate of α.

Similarly, D_s-optimal designs minimize the covariance matrix of the least squares estimates of a selected subset of parameters. We investigate D_s-optimal designs that minimize the determinant of the covariance matrix of the least squares estimate of τ, in the context of Model (2), the block effects β being regarded as nuisance parameters. Specifically, partition $(X^TX)^{-1} = (D_{ij})$ $(i, j = 1, 2)$ where $D_{11}\sigma^2$ $(p \times p)$ is the covariance matrix of the least squares estimate of τ, and let $Q_2 = I - X_2(X_2^TX_2)^{-1}X_2$ denote the projection operator for the orthogonal complement of the column space of X_2. Then a D_s-optimal design for τ minimizes the determinant of $D_{11} = (X_1^TQ_2X_1)^{-1}$. We assume that there is at least one design for which the determinant of $X_1^TQ_2X_1$ is nonzero and that overparameterization is avoided by choice of f_1. In part, this implies that, because of our choice os u_i, the column space of X_1 may not contain a constant vector.

Under Model (2), the D-optimal design for α and the D_s-optimal design for τ are equivalent when $N = n$:

$$|X^TX| = |X_2^TX_2| \times |X_1^TQ_2X_1|, \quad (3)$$

and maximizing $|X_1^TQ_2X_1|$ is equivalent to maximizing $|X^TX|$, since $|X_2^TX_2|$ is a constant (Harville 1974, 1975). For further background on optimality criteria and recent developments, see Atkinson (1982) and Fedorov (1972).

Following these specifications, a design algorithm is typically employed to determine the feasible set of points that is best, according to the chosen criterion, for inference based on the specified regression model. Very little has been published, however, regarding such algorithms for blocked response-surface models as those described previously. Existing algorithms like DETMAX (Mitchell 1974) are not immediately applicable when $n = N$ because they do not explicitly consider block size constraints. For the most part, previous work has focused on the qualitative treatment case. Harville (1974) was the first

to develop an algorithm that explicitly allows for blocking. Harville's algorithm searches for D-optimal designs and is applicable to Model (2) with f_1 restricted to the class of indicator vectors. Since our approach to blocking starts with Harville's framework, we give a relatively detailed overview.

1.2 Harville's Algorithm

Beginning with a nonsingular starting design, Harville's algorithm searches for improvements by sequentially attempting three classes of changes in three corresponding phases. In the first phase, improvements are accomplished through *exchanges*. In an exchange, the treatment assigned to a given experimental unit is replaced by some other treatment. In the *interchange* phase, improvements are accomplished by switching the allocation of a treatment from one experimental unit to a different experimental unit and simultaneously switching the allocation of the treatment assigned to the second experimental unit back to the first experimental unit. Clearly, different treatments and blocks must be involved if improvements are to result. If there are N experimental units available, of which only $n < N$ are to be used, an attempt to identify an optimal set of n experimental units is made during the *replacement* phase. This is accomplished by sequentially replacing assigned units $U_n = \{u_i, I = 1, \ldots, n\}$ with units in $U_N - U_n$, provided that improvements in the optimality criterion result.

To speed convergence, Harville exploited both the assumed qualitative treatment structure and the space-saving concomitant with D_s optimality [see Eq. (3)]. For changes during phases 1 and 2 of Harville's algorithm, Q_2 is constant, and it is, therefore, only necessary to evaluate the $p \times p$ matrix D_{11}. Further, because of the simple treatment structure, exchanges and interchanges affect just two rows and two columns of $X_1^T Q_2 X_1$, and consequently the impact of any interchange (after the first) can be assessed by computing the determinant of a 2×2 matrix. Harville made use of these ideas in developing a relatively efficient search procedure in phases 1 and 2. Complete reevaluation of $|X_1^T Q_2 X_1|$ is required in any phase 3 replacement or when a new pair of treatments is encountered during phases 1 and 2.

1.3 Other Approaches to Blocking

Freedman (1976), Jones (1976), Jones and Eccleston (1980), and Eccleston and Jones (1980) reported algorithms in conjunction with a weighted trace criterion. These approaches do not take advantage of Harville's algorithm and consequently are not quite as versatile, nor do they allow f_1 to represent a general response-surface model.

If sample sizes are very large, approximate or large-sample designs may be used effectively. Relevant theory and algorithms were provided by Cook and Thibodeau (1980) and were recently extended by Nachtsheim (in press). Unfortunately, as shown by Cook and Nachtsheim (1980), large-sample designs are not particularly effective when n or some of the blocks are small. Finally, Draper and John (1988) investigated response-surface designs for quantitative and qualitative variables.

In Section 2, we present an extension of the Harville algorithm that can be used in conjunction with Model (2). In Section 3, a number of examples are discussed. Conclusions are presented in Section 4.

2. CONSTRUCTING BLOCK DESIGNS FOR RESPONSE-SURFACE MODELS

A number of difficulties preclude direct application of Harville's algorithm for Model (2). First, exchanges typically impact every element of $X_1^T Q_2 X_1$. Second, response-surface design spaces tend to be much larger than design spaces for qualitative treatments. Third, changes in experimental units during phase 3 require recomputation of $|X_1^T Q_2 X_1|$ from scratch, and this leads to further inefficiencies. Finally, it is fairly easy for phase 3 to get stuck at a local optimum, particularly when an entire block must be deleted from the starting design to obtain the optimal design.

In the remainder of this article, our general approach follows the first two phases of Harville's algorithm. This means in particular that $N = n$, the block sizes are predetermined ($N_k = n_k$), and all available experimental units are to be used in the design. Algorithms for the case $n < N$ are under investigation.

2.1 Phase 1: Exchange Algorithm

The k-exchange algorithm (Johnson and Nachtsheim 1983) is used in phase 1. The user-specified constant k can range from 1 to n. For $k = 1$, the algorithm is essentially the same as Wynn's (1972) algorithm. When $k = n$, the algorithm reduces to the modified Fedorov algorithm (Cook and Nachtsheim 1980). Typically, $k < n/4$ is sufficient (Cook and Nachtsheim 1982).

The k-exchange procedure was originally developed for producing D-optimal designs and is based on the fractional increase in the determinant of $X^T X$ when x_{ij}, the jth treatment combination in block i in the present design, is exchanged for an arbitrary treatment combination x in χ, in effect replacing the corresponding row of X_1, $f_1^T(x_{ij})$, with $f_1^T(x)$. Letting

"up" denote the updated form of an expression after an exchange has taken place,

$$|(X^TX)_{up}| = |X^TX|\,\Delta_E(x_{ij}, x), \tag{4}$$

where

$$\Delta_E(x_{ij}, x) = 1 + v(x) - v(x_{ij}) - v(x)v(x_{ij}) + v^2(x_{ij}, x)$$

and

$$v(x_{ij}, x) = (f_1(x_{ij})^T, u_i^T)(X^TX)^{-1}(f(x)^T, u_i^T)^T,$$

where $v(x) = v(x, x)$ is proportional to the variance of a fitted value in the present design prior to updating and u_i is the indicator vector for block i. Each iteration of the k-exchange algorithm begins by identifying the points in the present design with the k smallest values of the variance function $v(x_{ij})$. Starting with the design point $x_{i_1j_1}$ minimizing v, these k points are then exchanged sequentially for the respective points in χ that maximize the Δ_E function. For example $x_{i_1j_1}$ is exchanged for the point that achieves $\max_x \Delta_E(x_{i_1j_1}, x)$. The design $|X^TX|$ and $(X^TX)^{-1}$ are updated after each exchange.

During the review process, we became aware of a generalization of the k-exchange algorithm, the kl-exchange algorithm of Atkinson and Donev (in press). In the kl-exchange algorithm, only l points in χ are considered for inclusion. Atkinson and Donev applied the algorithm in blocking situations, although their approach is markedly different from that given here.

Since the D and D_s criteria are equivalent when $N = n$, the k-exchange algorithm can be used without modification or with the modifications suggested by Atkinson and Donev (in press). A considerable increase in efficiency can be achieved, however, by exploiting the block structure when updating and maximizing the Δ_E function. First, recalling that n_i is the size of the ith block, partition $X_1^T = (F_1^T, \ldots, F_b^T)$, where F_i is the $(n_i \times p)$ submatrix consisting of the rows of X_1 assigned to block i and, for notational convenience, let $M = X_1^TQ_2X_1$. Then straightforward algebra will verify that

$$v(x_{ij}, x) = 1/n_i + (f_1(x_{ij}) - \bar{f}_{1i})^TM^{-1}(f_1(x) - \bar{f}_{1i}), \tag{5}$$

where $\bar{f}_{1i}$ is the row average of F_i, $\bar{f}_{1i} = F_i^T1_i/n_i$. Consequently, when using the Δ_E function, we can work in terms of the $p \times p$ matrix M rather than the full $(p + b) \times (p + b)$ matrix X^TX, which may be quite large. For this to be an effective strategy, however, we need to have the ability to update M directly after each exchange.

Direct updating of M is accomplished by first noting that the $n \times n$ projection operator Q_2 is block diagonal, $Q_2 = \text{block} - \text{diag}(Q_{ii})$, where

$$Q_{ii} = I_i - 1_i1_i^T/n_i, \tag{6}$$

I_i is the $n_i \times n_i$ identity, and 1_i is an $n_i \times 1$ vector of 1's. Then

$$M = \sum_{k=1}^{b} F_k^TQ_{kk}F_k, \tag{7}$$

and, since replacing x_{ij} with x affects only F_i, M can be updated by replacing F_i in (7) with

$$F_i^{up} = F_i + c_{ij}(f_1^T(x) - f_1^T(x_{ij})) = F_i + c_{ij}d_{ij}^T, \tag{8}$$

where "up" again denotes an updated form, c_{ij} is the $n_i \times 1$ vector with a 1 in the jth position and 0's elsewhere, and $d_{ij} = f_1(x) - f_1(x_{ij})$. Combining (7) and (8),

$$M_{up} = \sum F_k^TQ_{kk}F_k + d_{ij}c_{ij}^TQ_{ii}c_{ij}d_{ij}^T + d_{ij}c_{ij}^TQ_{ii}F_i + F_i^TQ_{ii}c_{ij}d_{ij}. \tag{9}$$

Using (6), this updated form can be expressed more conveniently as

$$M_{up} = M + A_{ij}B_{ij}^T, \tag{10}$$

where $A_{ij} = [a_{ij}, d_{ij}]$ and $B_{ij} = [d_{ij}, b_{ij}]$ are matrices of dimension $p \times 2$, $a_{ij} = (f_1(x) - \bar{f}_{1i}) - d_{ij}/n_i$, $b_{ij} = f_1(x_{ij}) - \bar{f}_{1i}$, and d_{ij} is as defined near (9).

As a consequence of (10), standard formulas for rank-2 updating of M^{-1} apply (e.g., see Fedorov 1972, pp. 99–105):

$$M_{up}^{-1} = M^{-1} - M^{-1}A_{ij}[I_2 + B_{ij}^TM^{-1}A_{ij}]^{-1}B_{ij}^TM^{-1}. \tag{11}$$

In short, updating the $p \times p$ matrix M^{-1} requires computation of the inverse of the 2×2 matrix $[I_2 + B_{ij}^TM^{-1}A_{ij}]$, the determinant of which is equal to Δ_E:

$$|(X^TX)_{up}| = |X_2^TX_2| \times |M_{up}| = |X_2^TX_2| \times |M| \times |I_2 + B_{ij}^TM^{-1}A_{ij}| = |X^TX|\,\Delta_E(x_{ij}, x). \tag{12}$$

As a consequence of these results, $(X_1^TQ_2X_1)^{-1}$ is computed from scratch only once, and computation of the full information matrix X^TX and its inverse is avoided.

2.2 Phase 2: Interchanges

In a phase 2 interchange, χ is ignored and treatment combinations are "traded" between blocks. Specifically, this involves systematically replacing x_{ij},

the jth treatment in block i, with x_{kl}, the lth treatment in block k, while simultaneously replacing x_{kl} with x_{ij} ($i \neq k$) as long as improvement results. To understand how to implement this procedure, it is necessary to investigate relationships between a starting design and the design obtained after a single interchange. The required updating could be accomplished by repeated application of the rank-2 updating procedures associated with (10) and (11), since a single interchange corresponds to two exchanges. Surprisingly, an interchange can be accomplished with a single rank-2 update. To see this, note that the interchange between blocks i and k affects only F_i and F_k. Thus $F_i^{up} = F_i - c_{ij}d_{ij,kl}^T$ and $F_k^{up} = F_k - c_{kl}d_{ij,kl}^T$, where $d_{ij,kl} = f_1(x_{ij}) - f_1(x_{kl})$. Following the steps that led to (10), it can be shown that a double update of M resulting from an interchange is of the form

$$M_{up} = M + A_{ij,kl}B_{ij,kl}^T, \tag{13}$$

where $A_{ij,kl} = [a_{ij,kl}, d_{ij,kl}]$ and $B_{ij,kl} = [d_{ij,kl}, b_{ij,kl}]$ are again $p \times 2$ matrices, $a_{ij,kl} = f_1(x_{ij}) - \bar{f}_{1i} - f_1(x_{kl}) + \bar{f}_{1k}$, $b_{ij,kl} = a_{ij,kl} + vd_{ij,kl}$, $v = (n_i - 1)/n_i + (n_k - 1)/n_k$, and $d_{ij,kl}$ is as defined previously. Updating M_{up}^{-1} is accomplished by using (11), replacing A_{ij} and B_{ij} with $A_{ij,kl}$ and $B_{ij,kl}$, respectively. Similarly, the fractional change in the determinant of M is

$$\begin{aligned} |M_{up}| &= |M| \times |I_2 + B_{ij,kl}^T M^{-1} A_{ij,kl}| \\ &= |M|\, \Delta_I(x_{ij}, x_{kl}). \end{aligned} \tag{14}$$

Thus in this phase of the algorithm the operational criterion is to select x_{ij} and x_{kl} to maximize Δ_I, each evaluation requiring the determinant of a 2×2 matrix.

2.3 Starting Designs

To begin, the algorithm described in Sections 2.1 and 2.2 requires a starting design with $|M| \neq 0$. We select such a starting design by first determining the treatment combinations and then distributing these across the experimental material.

Treatment Combinations. The treatment combinations for the starting design are chosen by ignoring the block structure and using the algorithm of Galil and Kiefer (1980) to generate a $(p + 1)$-point starting design for the model $y_i = \beta + f_1^T(x_i)\tau + \varepsilon_i$ ($i = 1, \ldots, n$). Since the constant term is included, $p + 1$ points are required for a nonsingular design. Briefly, the Galil–Kiefer algorithm begins with a 1-point design and then adds points sequentially until a $(p + 1)$-point design is obtained. Each point is added to maximize $|X_1 X_1^T|$, where X_1 corresponds to the previous design in the sequence.

Allocation to Experimental Units. Let $x_1, x_2, \ldots, x_{p+1}$ denote the treatment combinations determined from the Galil–Kiefer algorithm. To insure that M is nonsingular, we use the following simple scheme for constructing a connected starting design: Beginning with x_1, assign the treatment combinations sequentially to the units in block 1, starting over if $n_1 > p + 1$. Let x_{k_1} denote the last treatment assigned to block 1. The assignment to block 2 begins with x_{k_1} and proceeds sequentially, x_{k_1}, $x_{k_1+1}, \ldots, x_{k_2}$. Similarly, the assignment to block 3 begins with x_{k_2} and again proceeds sequentially, $x_{k_2}, x_{k_2+1}, \ldots, x_{k_3}$. Generally, the treatments are assigned sequentially to blocks, subject to the constraint that the last treatment assigned to block i is the first treatment assigned to block $i + 1$.

3. ILLUSTRATIONS

3.1 Constructing Classical Block Designs

Classical block designs involving qualitative treatments are typically D optimal, and when permitted by the design constraints, the algorithm of Section 2 will often yield such familiar designs. For example, suppose that $f_1^T(x) = (x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3)$, $N = 8$, $t = 3$, $\chi = [-1, 1]^3$, and it is necessary to run the experiment in $b = 2$ blocks of size 4. Using traditional notation to indicate treatments consisting of high and low combinations of the factor levels, the algorithm assigns the treatment combinations (1), ab, ac, and bc to one block and a, b, c, and abc to the other. This usual fractional factorial design is, of course, obtained from a 2^3 by applying the defining relation $I = ABC$ and assigning the treatments in each half-fraction to the respective blocks.

Similarly, the algorithm can be used to construct balanced and partially balanced incomplete block designs for a suitable choice of sample sizes, numbers of treatments, and block sizes. In our experience, the algorithm has not failed to find such designs when they are known to exist. Notably, the proposed algorithm is significantly faster than the Harville algorithm for such problems. For example, when $t = 3$, $b = 7$, and $N = 21$, balanced incomplete block designs are known to exist. The Chen–Harville FORTRAN code (Chen 1977), implemented on a Macintosh Plus computer, constructed one such design in 509 seconds. From the same starting design and with $k = 3$, our algorithm converged to the same design in 66 seconds. We found this substantial difference in time to be a little surprising, since Harville's procedure is tailored to a qualitative treatment structure. We have observed similar percentage changes in time in other problems.

3.2 Constructing Blocked Factorial and Response-Surface Designs

Let us return to the first example discussed in Section 1, in which 18 experimental units in three blocks each of size 6 were available for estimation of the model

$$f_1^T(x) = (x_1, x_2, x_3, x_4, x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4, x_3x_4),$$

with $\chi = [-1, 1]^4$. This model allows estimation of main effects and two-factor interactions when each factor is to be run at two levels. Our algorithm produced the following design: Block 1—*b*, *d*, *c*, *ab*, *ad*, *abcd*; Block 2—(1), *ac*, *abc*, *abd*, *acd*, *bcd*; and Block 3—*a*, *ac*, *bc*, *bd*, *cd*, *abcd*. This design consists of the complete 2^4 factorial with treatment combination *abcd* repeated in blocks 1 and 3 and treatment combination *ac* repeated in blocks 2 and 3.

We now turn to blocking in standard response-surface experiments. Suppose that $t = 2$, $N = 14$, and $b = 2$. We used the proposed algorithm to construct *D*-optimal designs for block sizes (7, 7), (8, 6), and (9, 5) in conjunction with the second-order response-surface model $f_1^T(x) = (x_1, x_2, x_1^2, x_2^2, x_1x_2)$ and $\chi = [-1, 0, 1]^2$. Resulting designs are pictured in Figure 1. Of interest in the (7, 7) case is the similarity of the layouts in each block. The treatment combinations in one block can be obtained by a simple reflection about the line $x_1 = x_2$ of the coordinates in the other. For the (9, 5) case, familiar designs result in each block. Moreover, in every case, if block effects are not significant, the design is simply a standard 3^2 factorial with corner and center points having two replicates. This represents an extremely efficient use of the five repeat observations; not only are 5 df obtained for pure error, but the placement at the corners leads to more precise inference regarding main effects and interactions.

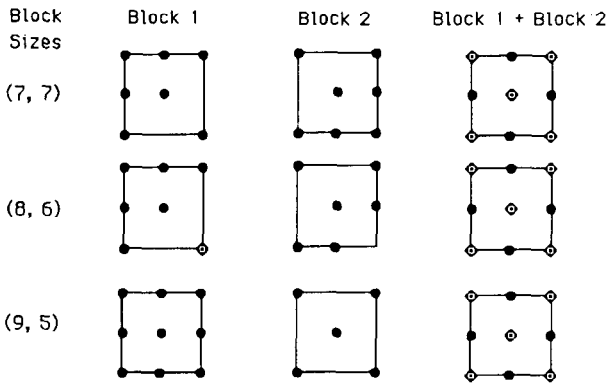

Figure 1. Three Block Designs for a Second-Order Response-Surface Model With Two Factors, 14 Observations, Two Blocks, and the Indicated Block Sizes. Solid dots indicate one replicate; hollow dots indicate two replicates.

As a final example, we consider the second design problem discussed in the introduction. In this case, $n = 20$, $t = 3$, $b = 5$, and the block sizes are equal. Again, a second-order response-surface model is of interest, this time in three factors. The best design found is pictured in Figure 2. In an interesting near-symmetry, four out of five blocks consist of two corner points, one star point, and one edge center. The fifth block consists of three corners and one edge center.

3.3 Blocking Existing Designs

It is often desirable to run a *prespecified* collection of treatment combinations in a blocked experiment. In this section, we discuss the use of our algorithm in several such situations.

In the four-factor, 18-run experiment discussed previously, the fact that all 16 treatment combinations in the standard 2^4 were included is reassuring; if block effects are not significant, a fairly standard analysis ensues. This will not always occur, however. For example, if a different model had been specified (e.g., one having fewer interaction terms), some corner points might not be included in the *D*-optimal design. In such situations, the experimenter can simply prespecify the set of treatments desired in the final design, and omit phase 1 of the algorithm. Phase 2 will then find the optimal assignment of required treatments to blocks. In the preceding example, the experimenter may wish that the 18 treatment combinations consist of the 2^4 corner points plus two center points. Omitting phase 1, the algorithm produced the following design: Block 1—*a*, *d*, *ab*, *bc*, *bd*, *abcd*; Block 2—(1), *c*, *ad*, *abc*, *acd*, *bcd*; and Block 3—*b*, *ac*, *cd*, *abd*, (0), (0), where (0) indicates a center point. In contrast with the previous design, the replicated points are placed in the same block.

As noted in Section 1, for a second-order response-surface model in three factors, Box and Draper (1987, p. 512) suggested a three-factor central composite design in three blocks of sizes 5, 5, and 7 for a total of 17 observations. Various advantages accrue from the use of central composite designs. In this example, if the star points are located at distance $\alpha = 2.8^{1/2}$ from the center, block effects will be orthogonal to treatment effects. Further, as long as $\alpha > 1$, each factor is observed at five levels, allowing the adequacy of the second-order model to be studied. Of course, the suggested block sizes (5, 5, and 7) may not be possible. When this is so, a compromise approach involves using the central composite design as a starting design and then applying only phase 2 as described previously. In the current example, if $\alpha = 2.8^{1/2}$ and block sizes of 5, 5, and 7 are specified, the phase 2 interchange procedure results in the Box–

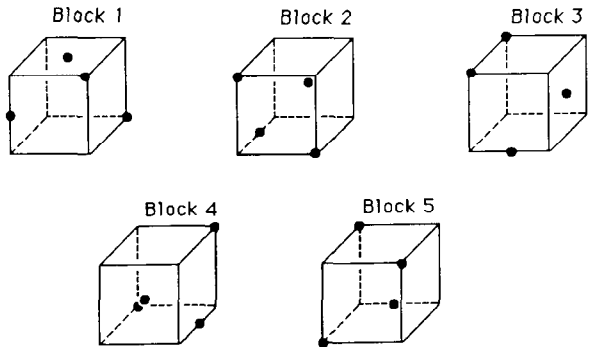

Figure 2. Block Design for a Second-Order Response-Surface Model With Three Factors, 20 Observations, and Five Blocks Each of Size 4.

Draper orthogonal blocking design. This procedure was tried for various other values of t in a full second-order response-surface model with similar results.

Finally, experiments are sometimes blocked simply as protection against possible effects that may not be real. A reasonable approach in such instances would be to construct an optimal design ignoring blocks and then to optimally allocate the chosen treatment combinations to blocks via phase 2. In this way, if the block effects do not materialize, the experiment is as efficient as possible. As in all of the cases discussed in this section, the experimenter's choice between designs produced by the full algorithm and those produced from phase 2 alone should be guided by an inspection of the correlation structure, the nature of the replication schemes, the relative values of the criterion, and other relevant design considerations.

4. DISCUSSION: OPTIMAL VERSUS CLASSICAL DESIGN

Optimal design is occasionally held to have little practical relevance. The reasons for such attitudes vary, but the dependence of the final design on the model is usually toward the top of the list. Further, most studies rely on a relatively narrow view that comes with adopting a single optimality criterion that by itself cannot adequately reflect the real value of a design. On the other hand, one clear advantage of the optimal-design approach is its ability to produce sensible solutions in irregular problems in which an appropriate application of more standard designs is elusive at best. In our experience, the frequency with which optimality considerations result in relevant information is clearly high enough to justify routine consideration of the designs, particularly in nonstandard problems.

Classical design has its own advantages and disadvantages. Like optimal designs, classical designs depend on a model, although to a somewhat lesser degree. When confounding a 2^n factorial, for example, prior knowledge of which interactions are likely to be negligible is essential. Further, we find that the situations in which classical designs work well are rather confining, occasionally leading to unfortunate results. On the other hand, classical designs often do well on several criteria, and the analysis associated with a classical design tends to be relatively easy to conduct and interpret.

Judging from experience, experimenters often have an overwhelming urge to tailor the scientific question or trim the experimental material to allow application of a particular classical design. For example, this may imply discarding treatments and experimental units so that the largest possible randomized complete block (RCB) design can be implemented. We do not universally condemn such manipulations, but the associated consequences cannot be adequately assessed without knowledge of the alternatives. Discarding experimental units will of course produce a drop in $|M|$ relative to a D-optimal design. Clearly, we cannot tell if the advantages of an RCB design override this drop if we have not computed the D-optimal design in the first place.

Our basic contention is simply that optimality considerations allow the experimenter to broaden the base of available designs that must then be judged on a range of relevant criteria. Optimal designs often, but not always, fare well on such judgments. The algorithm presented here can be used to generate D-optimal blocking designs for general response-surface models. A technical summary of the algorithm, which was recently implemented in RS/Discover (BBN Software Products Corporation) is given in the Appendix.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation Grant DMS-8801544. We thank David Harville, Brad Jones, the referees, and the editors for useful comments on an earlier version of this manuscript.

APPENDIX: OPTIMAL BLOCKING ALGORITHM

1. Obtain a nonsingular starting design by using the algorithm described in Section 2.3 and form X_1 and X_2. Select k for use in the k-exchange algorithm. Select ε for use in the termination criterion. Set $N_I = -1$.
2. *Exchanges*. Let $d_0 = |X_{10}^T Q_2 X_{10}|$, where $X_0 = [X_{10}, X_2]$ is the design matrix for current best design. Set $N_E = 0$.
   2.1. Set $s = 0$. Compute the variances of prediction for points in the design via (5). Identify and rank the k points $x^1, \ldots, x^k$ having smallest variances.

2.2. Set $s = s + 1$. Consider the exchange of the point x^s. Find the treatment combination x^* in χ at which the multiplicative increase $\Delta_E(x^s, x)$ in the determinant criterion is maximal: $\Delta_E(x^s, x^*) = \max_x \Delta_E(x^s, x)$.

2.3. If $\Delta_E(x^s, x^*) > 1$, exchange x^s for x^*, update $(X_1^T Q_2 X_1)^{-1}$ and $|X_1^T Q_2 X_1|$ by using (11) and (12), and set $N_E = 1$.

2.4. If $s < k$, go to step 2.2; otherwise, go to step 2.5.

2.5. Let $d_1 = |X_1^T Q_2 X_1|$. If $(d_1 - d_0)/d_0 > \varepsilon$, set $d_0 = d_1$ and go to step 2.1. If $(d_1 - d_0)/d_0 < \varepsilon$, $N_E = 0$, and $N_I > 0$, go to step 4. Otherwise, go to step 3.

3. *Interchanges.*

3.1. Set $N_I = 0$.

3.2. Set $s = 0$. Sequentially consider all nonidentical pairs, (x_{ij}, x_{kl}) $i \neq k$, $j \neq l$, of treatment combinations in different blocks for interchange. In particular, if $\Delta_I(x_{ij}, x_{kl}) > 1$, interchange x_{ij} and x_{kl}, update $(X_1^T Q_2 X_1)^{-1}$ and $|X_1^T Q_2 X_1|$ by using (11) and (14), respectively, and set $s = s + 1$. The next pair in the sequence is considered after the necessary updating. After the last pair, go to step 3.3.

3.3. Set $N_I = N_I + s$. If $N_I = 0$, go to step 4. If $s > 0$, go to step 3.2. If $s = 0$ and $N_I > 0$, go to step 2.

4. Stop.

[*Received May 1988. Revised January 1989.*]

REFERENCES

Atkinson, A. C. (1982), "Developments in the Design of Experiments," *International Statistical Review*, 50, 161–177.

Atkinson, A. C., and Donev, A. N. (in press), "The Construction of Exact *D*-Optimum Experimental Designs With Application to Blocking Response Surface Designs," *Biometrika*, 76.

Box, G. E. P., and Draper, N. (1987), *Empirical Model-Building and Response Surfaces*, New York: John Wiley.

Chen, Mon-Gy (1977), "Computer Construction of *D*-Optimal Balanced Incomplete Block Designs," unpublished manuscript, Iowa State University, Dept. of Statistics.

Cook, R. D., and Nachtsheim, C. J. (1980), "A Comparison of Algorithms for Constructing Exact *D*-Optimal Designs," *Technometrics*, 22, 315–324.

——— (1982), "An Analysis of the *k*-Exchange Algorithm," technical report, University of Minnesota, Dept. of Management Sciences.

Cook, R. D., and Thibodeau, L. A. (1980), "Marginally Restricted *D*-Optimal Designs," *Journal of the American Statistical Association*, 75, 336–371.

Draper, N. A., and John, J. A. (1988), "Response-Surface Designs for Quantitative and Qualitative Variables," *Technometrics*, 30, 423–428.

Eccleston, J. A., and Jones, B. (1980), "Exchange and Interchange Procedures to Search for Optimal Row-and-Column Designs," *Journal of the Royal Statistical Society*, Ser. B, 42, 372–376.

Fedorov, V. V. (1972), *Theory of Optimal Experiments*, New York: Academic Press.

Freedman, G. H. (1976), "On the Selection of Designs for Comparative Experiments," *Biometrics*, 32, 195–199.

Galil, Z., and Kiefer, J. (1980), "Time- and Space-Saving Computer Methods, Related to Mitchell's DETMAX, for Finding *D*-Optimum Designs," *Technometrics*, 21, 301–313.

Harville, D. (1974), "Nearly Optimal Allocation of Experimental Units Using Observed Covariate Values," *Technometrics*, 16, 589–599.

——— (1975), "Computing Optimum Designs for Covariance Models," in *A Survey of Statistical Design and Linear Models*, ed. J. N. Srivastava, Amsterdam: North-Holland, pp. 209–228.

Johnson, M. E., and Nachtsheim, C. J. (1983), "Some Guidelines for Constructing Exact *D*-Optimal Designs on Convex Design Spaces," *Technometrics*, 25, 271–277.

Jones, B. (1976), "An Algorithm for Deriving Optimal Block Designs," *Technometrics*, 18, 451–458.

Jones, B., and Eccleston, J. A. (1980), "Exchange and Interchange Procedures to Search for Optimal Designs," *Journal of the Royal Statistical Society*, Ser. B, 42, 238–243.

Mitchell, T. J. (1974), "An Algorithm for the Construction of *D*-Optimal Experimental Designs," *Technometrics*, 16, 203–211.

Nachtsheim, C. J. (in press), "On the Design of Experiments in the Presence of Fixed Covariates," *Journal of Statistical Planning and Inference*.

Wynn, H. P. (1972), "Results in the Theory and Construction of *D*-Optimum Experimental Designs," *Journal of the Royal Statistical Society*, Ser. B, 34, 133–147.