Optimal Saturated Block Designs when Observations are Correlated

Bo Jin
Merck Research Laboratories
BL3-2, 10 Sentry Parkway
Blue Bell, PA 19422, USA

J. P. Morgan
Department of Statistics
Virginia Tech
Blacksburg, Virginia 24061, USA

15th June 2006

Summary

A- and MV-optimal block designs are identified in the class of minimally connected designs when the observations within blocks are spatially correlated. All connected designs are shown to be D-equal regardless of the correlation structure, and a sufficient condition for E-optimality is presented. Earlier results for the uncorrelated case are strengthened.

NOTE: *See pages 21-22 for additional results found after this paper was accepted for publication.*

1 Introduction

Consider comparing the relative effectiveness of v treatments employing $n = bk$ experimental units arranged into b blocks of size k. The standard linear model for the observation y_{ju} on unit u in block j when using design d is

$$y_{ju} = \mu + \tau_{d[j,u]} + \beta_j + \varepsilon_{ju} \tag{1}$$

for $u = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, b$. The components of (1) are an overall mean μ, the effect $\tau_{d[j,u]}$ of treatment $d[j, u]$ assigned to unit u in block j by design d, a block effect β_j, and a random error ε_{ju} with zero mean. Writing $\tau = (\tau_1, \ldots, \tau_v)'$ and $\beta = (\beta_1, \ldots, \beta_b)'$ for the vectors of treatment and block effects, and $y_{n\times 1}$ for the vector of yields arranged in lexicographic order, then (1) says the mean vector is $\mathrm{E}(y) = \mu 1_n + M_d\tau + L\beta$ for $L = I_b \otimes 1_k$ the $n \times b$ unit/block incidence matrix, and M_d the $n \times v$ unit/treatment incidence matrix defined by the $d[j, u]$. Here and elsewhere, 1_q denotes a column of q ones (likewise 0_q a column of q zeros). Choice of design is choice of M_d.

It is well known that all treatment contrasts are estimable under d if and only if the information matrix for estimation of treatment effects, $C_d = M_d'(I - \frac{1}{k}LL')M_d$, is of rank $v - 1$; equivalently, the rank of $X_d = (1, M_d, L)$ is $b + v - 1$. A design with this property is said to be *connected*. A necessary condition for d to be connected is that the number of rows of X_d is at least its required rank, that is, $n \geq b + v - 1$.

Several papers have appeared in the literature which discuss optimality of connected block designs when the number of experimental units is minimal, $n = b + v - 1$. Let $\mathcal{D}(v, b, k)$ denote the class of all connected block designs having v treatments, b blocks and constant block size $k \geq 2$ satisfying

$$bk = b + v - 1. \tag{2}$$

Then $\mathcal{D}(v, b, k)$ is the class of *minimally connected designs*. Alternatively, since estimation of block and treatment effects takes all $n - 1$ degrees of freedom (there are no degrees of freedom remaining for error), this is the class of *saturated designs*. When observations are equivariable and uncorrelated, the A-, MV-, and E-optimal designs in $\mathcal{D}$ are known, and all connected designs are D-equal: see Bapat and Dey (1991), Mandal, Shah and Sinha (1991), and Dey, Shah, and Das (1995). Here optimality of designs in $\mathcal{D}$ is studied when observations are equivariable and *correlated*. For the D-optimality problem, an arbitrary correlation structure is considered. For other optimality problems this spatial correlation structure is assumed:

$$\mathrm{cov}(y_{ju}, y_{j'u'}) = \begin{cases} 0 & \text{if } j \neq j' \\ \sigma^2 & \text{if } j = j', u = u' \\ \rho_{|u-u'|}\sigma^2 & \text{if } j = j', u \neq u' \end{cases} \quad (3)$$

where

$$1 > \rho_1 \geq \rho_2 \geq \rho_3 \geq \ldots \geq \rho_{k-1} \geq 0. \quad (4)$$

In addition it is everywhere required that the variance-covariance matrix $\Sigma_{n\times n}$ for the entire $n \times 1$ observations vector y be positive definite, i.e.

$$a^{'}\Sigma a > 0 \text{ for any } a \neq 0. \quad (5)$$

The common diagonal element of Σ is denoted by σ^2. The earlier results mentioned above are for $\Sigma = \sigma^2 I_n$.

The paper proceeds as follows. Section 2 presents the basic properties of minimally connected designs in a fashion suited to the current endeavor. Section 3 identifies M-optimal (this includes A- and MV-optimal) designs for estimation of elementary treatment contrasts. That all connected designs are D-equal is established in section 4. Section 5 studies the E-optimality problem, providing a sufficient condition and determining optimal designs in some special cases. Concluding remarks comprise section 6.

2 Properties of Minimally Connected Designs

This section presents a lemma from which several useful properties of minimally connected designs follow.

Lemma 2.1 *For each $d \in \mathcal{D}$, there is only one unbiased estimator for any treatment contrast $m'\tau$. Consequently, the ordinary least squares estimate (OLSE) and the general least squares estimate (GLSE) for $m'\tau$ are the same.*

Proof Suppose there are two unbiased estimators $p'y$ and $q'y$ for $m'\tau$. Let $a = p-q \neq 0$ and write $\theta' = (\mu, \tau', \beta')$. Then $0 = \mathrm{E}(a'y) = a'X_d\theta$ for all $\theta \Rightarrow a'X_d = 0 \Rightarrow n > \mathrm{rank}(X_d) = b + v - 1$, which contradicts (2). □

The result of lemma 2.1 also holds for block contrasts. The importance of the lemma is in the corollaries that follow. Corollary 2.2 first appeared in Bapat and Dey (1991); here the proof is much simpler.

Corollary 2.2 *Any $d \in \mathcal{D}$ is necessarily binary.*

Corollary 2.3 *For any $d \in \mathcal{D}$, no pair of blocks has more than one treatment in common, that is, no pair of treatments occurs in more than one block.*

Definition *A **chain** in a block design is an even-length sequence of experimental units (or the corresponding observations) such that two consecutive units share either the same treatment or the same block, but not both, and such that no treatment or block is common to more than two of the units.*

If $(y_1, y_2, \ldots, y_c)$ is a chain of observations connecting treatments i and i' (blocks j and j'), then the linear combination $y_1 - y_2 + y_3 - \cdots - y_c$ is an unbiased estimator of the elementary treatment contrast $\tau_i - \tau_{i'}$ (respectively, the elementary block contrast $\beta_j - \beta_{j'}$).

Corollary 2.4 *For any $d \in \mathcal{D}$, any two blocks are connected by exactly one chain, as are any two treatments.*

The development of connectedness in terms of rank can be seen in the text by Chakrabarti (1962, chapter 2). The equivalent formulation via chains is handled nicely by Searle (1971, section 7.4).

For later use two types of elementary treatment contrasts will be distinguished for any particular design. If two treatments appear in the same block then their elementary contrast estimate is the difference between the corresponding observations for those two units. This is called a *within-block elementary contrast* and its corresponding estimate a *within-block elementary contrast estimate.* If two treatments never appear in the same block then their elementary contrast is estimated by the unique chain of observations connecting them. This is called a *between-blocks elementary contrast* and its corresponding estimate a *between-blocks elementary contrast estimate.*

EXAMPLE 1 Consider the design in $\mathcal{D}(7,3,3)$ having blocks $B_1 = [1,2,3], B_2 = [4,1,5]$, and $B_3 = [5,6,7]$. For this design, $\widehat{\tau}_2 - \widehat{\tau}_3 = y_{12} - y_{13}$ is a within-block elementary treatment contrast estimate for B_1, while $\widehat{\tau}_4 - \widehat{\tau}_7 = (y_{21} - y_{23}) - (y_{33} - y_{31})$ is a between-blocks elementary treatment contrast estimate for blocks B_2 and B_3.

3 M-Optimal Designs for Elementary Treatment Contrasts

The results established here will be for elementary treatment contrasts, but the concept of Majorization optimality, also called M-optimality (Bagchi and Bagchi, 2001), can be more generally construed as follows (the reader is referred to Bhatia, 1997, for a complete discussion of majorization). Let $e_{1d}, e_{2d}, \ldots, e_{md}$ be variances of a set of m contrasts of interest when estimated using design d, and let e_d be the vector of the e_{gd} for $g = 1, 2, \ldots, m$. Let f be a monotonically increasing convex function. Design d^* is M-optimal for estimation of the m specified contrasts if d^* minimizes $\sum_{g=1}^{m} f(e_{gd})$ for every such f. A necessary and sufficient condition for M-optimality is that e_{d^*} is *(weakly) submajorized* by e_d (in notation, $e_{d^*} \prec_w e_d$) for every d (Bhatia, 1997, page 40). Authors like Bagchi and Bagchi (2001) have pursued M-optimality for the canonical variances, that is, the e_{gd} are the inverses of the nonzero eigenvalues of the information matrix C_d. For experiments in which estimation will focus on elementary treatment contrasts $\tau_i - \tau_{i'}$, a better choice is to let e_d contain the $\binom{v}{2}$ pairwise variances $\text{var}(\hat{\tau}_i - \hat{\tau}_{i'})$. Implications of M-optimality in this sense will be discussed following the main result of this section, Theorem 3.1, and its proof.

The design $d^* \in \mathcal{D}$ to be studied here is

$$\begin{array}{lcccccccccc} B_1: & 1 & 2 & \ldots & l-1 & l & l+1 & \ldots & k-1 & k \\ B_2: & k+1 & k+2 & \ldots & k+l-1 & l & k+l & \ldots & 2k-2 & 2k-1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ B_b: & v-k+2 & v-k+3 & \ldots & v-k+l & l & v-k+l+1 & \ldots & v-1 & v \end{array} \tag{6}$$

where treatment $l = \lfloor \frac{k+1}{2} \rfloor$ is placed on experiment unit $u = \lfloor \frac{k+1}{2} \rfloor$ in every block. If the block size k is odd then treatment l is at the middle unit in each block; if k is even then in what follows treatment l could be at either of the middle two units in each block, so for clarity unit $u = \frac{k}{2}$ is chosen.

THEOREM 3.1 *Under conditions (3)-(5) the design d^* displayed in (6) is M-optimal for estimation of elementary treatment contrasts.*

PROOF The factor σ^2 will be ignored when expressing treatment contrast variances. Let e_d be the $1 \times \binom{v}{2}$ vector of elementary treatment contrast variances for an arbitrary design $d \in \mathcal{D}(v, b, k)$. Decompose e_d as $e_d = (e_d^{(1)}, e_d^{(2)})$ where $e_d^{(1)}$ consists of the within-block elementary treatment contrast variances, and $e_d^{(2)}$ contains the between-blocks elementary treatment contrast variances. Corollaries 2.2 and 2.3 say that $e_d^{(1)}$ is $1 \times b\binom{k}{2}$, and $e_d^{(2)}$ is thus $1 \times [\binom{v}{2} - b\binom{k}{2}] = 1 \times \binom{b}{2}(k-1)^2$, the latter being $(k-1)^2$ variances for each of the $\binom{b}{2}$ pairs of blocks.

It is easy to see that $e_d^{(1)} = e_{d^*}^{(1)}$ for some permutation of their elements. The theorem will be shown if $e_d^{(2)} \succ_w e_{d^*}^{(2)}$, which can be established through consideration of two distinct scenarios.

SCENARIO 1: there is one treatment l common to all b blocks (these designs are exactly those for which various optimalities have been proven in the uncorrelated case). It is sufficient to compare the variances of between-blocks elementary treatment contrasts for any two blocks (say B_1 and B_2), since the collection of $(k-1)^2$ such variances has identical structure for any pair of blocks. Denote the replicates of treatment l in blocks B_1 and B_2 as $l_{[s]}$ and $l_{[t]}$ respectively, where the bracketed subscript denotes the unit on which this treatment appears, and with no loss of generality $s \le t \le \lfloor \frac{k+1}{2} \rfloor$. Then the layout of the two blocks for design d can be displayed as

$$\begin{array}{lcccccccccc} B_1: & 1 & 2 & \ldots & s-1 & l_{[s]} & s+1 & \ldots & \ldots & k-1 & k \\ B_2: & k+1 & k+2 & \ldots & \ldots & k+t-1 & l_{[t]} & k+t & \ldots & 2k-2 & 2k-1 \end{array}$$

By Lemma 2.1 there is only one unbiased estimate for any between-blocks elementary treatment contrast, which must be of the form $(y_{1u_1} - y_{1s}) - (y_{2u_2} - y_{2t})$ where $u_1 \in$

$\{1, 2, \ldots, k\}/\{s\}$ and $u_2 \in \{1, 2, \ldots, k\}/\{t\}$. Thus the elements of $e_d^{(2)}$ are

$$4 - 2\rho_{|u_1-s|} - 2\rho_{|u_2-t|} \tag{7}$$

again for $u_1 \in \{1, 2, \ldots, k\}/\{s\}$ and $u_2 \in \{1, 2, \ldots, k\}/\{t\}$.

Choice of design d under this scenario is simply choice of s and t. Choice of s is choice of $k-1$ values from $(\rho_1, \rho_2, \ldots, \rho_{k-1})$ for $\rho_{|u_1-s|}$ in (7), and choice of t is likewise choice of $k-1$ values for $\rho_{|u_2-t|}$. While there are other restrictions, this fact is true: among the $k-1$ values for $\rho_{|u_1-s|}$ induced by choice of s, no ρ_g can appear more than twice (likewise for t). If $s = \lfloor \frac{k+1}{2} \rfloor = s^*$ (say) is selected, then the $k-1$ induced values are $(\rho_1, \rho_1, \rho_2, \rho_2, \ldots, \rho_{\frac{k-1}{2}}, \rho_{\frac{k-1}{2}})$ if k is odd, or $(\rho_1, \rho_1, \rho_2, \rho_2, \ldots, \rho_{\frac{k-2}{2}}, \rho_{\frac{k}{2}})$ if k is even. Let a_s and a_t be the $(k-1)$-vectors of the $\rho_{|u_1-s|}$ and $\rho_{|u_2-t|}$ corresponding to choice of s and t, with their elements (a_{sh} and a_{th}, respectively) arranged in nonincreasing order. Clearly $a_{s^*h} \geq a_{sh}$ for all h and any s. Thus for all h, h' and any (s,t), taking $t^* = s^*$, $a_{s^*h} + a_{t^*h'} \geq a_{sh} + a_{th'} \Rightarrow 4 - 2a_{s^*h} - 2a_{t^*h'} \leq 4 - 2a_{sh} - 2a_{th'}$ and it immediately follows that $e_d^{(2)} \succ_w e_{d^*}^{(2)}$.

Scenario 2: there are at least two blocks with no common treatment. For any two blocks with a common treatment, the $(k-1)^2$ elementary between-blocks treatment contrasts still obey the argument in scenario 1. For two blocks (say B_1 and B_2) having no common treatment, corollary 2.4 says there must be a unique chain linking the two. Thus there is a sequence of blocks of the following form:

$$\begin{array}{llllll}
B_1: & \ldots & l_1 & \ldots & \ldots & \ldots \\
B_{j_1}: & \ldots & l_1 & \ldots & l_2 & \ldots \\
B_{j_2}: & \ldots & l_2 & \ldots & l_3 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
B_{j_w}: & \ldots & l_w & \ldots & l_{w+1} & \ldots \\
B_2: & \ldots & l_{w+1} & \ldots & \ldots & \ldots
\end{array} \tag{8}$$

B_{j_1}, B_{j_2}, …, B_{j_w} are called the *linking blocks* and treatments l_1, l_2, …, l_{w+1} the *linking treatments*. The subscripts for the linking treatments are not their positions: they simply indicate different treatments.

An elementary treatment contrast for any two linking treatments in (8) is not a between-blocks treatment contrast for B_1 and B_2, because it is either a within-block treatment contrast, or a between-blocks treatment contrast for some other pair of blocks. For example, the elementary contrast between l_1 and l_2 is a within-block treatment contrast for B_{j_1} and the elementary contrast between l_1 and l_{w+1} is a between-blocks treatment contrast for B_{j_1} and B_{j_w}. Thus the $(k-1)^2$ between-blocks elementary treatment variances for B_1 and B_2 are those for comparing the treatments in B_1 other than l_1 with the treatments in B_2 other than l_{w+1}. Suppose treatment l_1 is at the s^{th} unit in B_1 and l_{w+1} is at the t^{th} unit in B_2, where with no loss of generality $s \leq t \leq \lfloor \frac{k+1}{2} \rfloor$. Consider comparing two treatments, one at unit u_1 ($\neq s$) in B_1, the other at unit u_2 ($\neq t$) in B_2. The estimate of this between-blocks treatment contrast is

$$y_{1u_1} - y_{1s} - (y_{2u_2} - y_{2t}) + \sum_{m=1}^{w} (y_{j_m[l_m, B_{j_m}]} - y_{j_m[l_{m+1}, B_{j_m}]})$$

where $[l_m, B_{j_m}]$ is the position of the linking treatment l_m in the linking block B_{j_m}. The variance of this contrast is

$$4 - 2\rho_{|u_1-s|} - 2\rho_{|u_2-t|} + \Delta \tag{9}$$

where $\Delta = 2\sum_{m=1}^{w}(1 - \rho_{|[l_m,B_{j_m}]-[l_{m+1},B_{j_m}]|}) > 0$. Comparing (9) to (7), the submajorization $e_d^{(2)} \succ_w e_{d^*}^{(2)}$ is obvious.

The proof is complete if no other scenario is possible, that is, if Scenario 1 is implied by $B_j \cap B_{j'} \neq \emptyset$ for all $j \neq j'$. This follows easily from corollaries 2.3 and 2.4. □

Depending on the ρ_h's, d^* may or may not be uniquely M-optimal. Uniqueness (up to treatment relabelling) is guaranteed by $\rho_1 > \rho_2 > \ldots > \rho_{k-1} > 0$. If all ρ_h's are zero (the uncorrelated case), then the structure is the same except that the position of the common treatment l is irrelevant. This can be extended as:

COROLLARY 3.2 *If correlation decays rapidly relative to block size in the sense that ρ_h is zero for all $h > \tilde{h}$ and some $\tilde{h} \leq \lfloor \frac{k-1}{2} \rfloor$, then M-optimality is attained if the common treatment l in d^* appears in any position having at least $\tilde{h}$ plots to either side.*

The *A-value* of a design is the sum of all elementary treatment contrast variances (this is $f(x) = x$ in the first paragraph of this section). A-optimal designs minimize the average variance over elementary treatment contrasts. This can also be expressed as the average of the canonical variances, or the average of the variances of any $v - 1$ orthonormal contrasts (Kempthorne, 1956). A direct consequence of Theorem 3.1 is:

COROLLARY 3.3 *Under conditions (3)-(5) the design d^* displayed in (6) is A-optimal.*

The *MV-value* of a design d is $\text{MV}_d = \max_{i \neq i'} \text{var}_d(\hat{\tau}_i - \hat{\tau}_{i'})$. MV-optimal designs minimize the largest variance for elementary treatment contrasts. Theorem 3.1 and the definition of submajorization give:

COROLLARY 3.4 *Under conditions (3)-(5) the design d^* displayed in (6) is MV-optimal.*

The uniqueness property mentioned just prior to corollary 3.2 need not hold for MV-optimality. In fact, any d for which $\text{MV}_d = \max(e_d^{(1)})$ is MV-optimal.

Theorem 3.1 holds for the special case $\rho_h = 0$ for all h, showing that even in the uncorrelated case, the optimality of d^* is much stronger than has been previously established. Indeed, whether or not the errors are correlated, the optimality is even stronger than stated in Theorem 3.1, for it is clear from the proof that d^* minimizes $\sum_i \sum_{i'>i} f(\text{var}(\hat{\tau}_i - \hat{\tau}_{i'}))$ even if the convexity condition is removed from f.

4 D-OPTIMAL DESIGN

The *D-value* of a design d is the product of the positive eigenvalues of the Moore-Penrose inverse $C_d^\dagger$ of the information matrix C_d; D-optimal designs minimize the D-value. The first step in attacking the D-optimality problem is to establish a useful expression for $C_d^\dagger$.

LEMMA 4.1 *For any $d \in \mathcal{D}$,*

$$C_d^\dagger = \frac{1}{\sigma^2}(I_v - \frac{1}{v}J_v)\,Cov(\hat{\tau}_{(d)})(I_v - \frac{1}{v}J_v) \tag{10}$$

where $Cov(\hat{\tau}_{(d)})$ is the variance-covariance matrix for any solution $\hat{\tau}_{(d)}$ to the reduced normal equations for estimating treatment effects under design d, and J_v is the $v \times v$ all-ones matrix.

Proof Let $C_d = \sum_{i=1}^{v-1} z_{di} s_i s_i'$ be the spectral decomposition of C_d where the eigenvalues are $z_{d0} = 0 < z_{d1} \leq z_{d2} \leq \ldots \leq z_{d,v-1}$ and a corresponding set of orthonormal eigenvectors is $\{s_0 = \frac{1}{\sqrt{v}} 1_v, s_1, s_2, \ldots, s_{v-1}\}$. Then the Moore-Penrose inverse of C_d is $C_d^{\dagger} = \sum_{i=1}^{v-1} \frac{1}{z_{di}} s_i s_i^{'}$.

For given $\hat{\tau}_{(d)}$ let C_d^P denote the r.h.s. of (10). Denote an arbitrary generalized inverse of C_d by C_d^-. Then for any estimable contrasts, say $s^{'}\tau$ and $m^{'}\tau$ where $s'1 = m'1 = 0$,

$$s'\mathrm{Cov}(\hat{\tau}_{(d)})m = \mathrm{Cov}(s'\hat{\tau}_{(d)}, m'\hat{\tau}_{(d)}) = \sigma^2 s' C_d^- m \tag{11}$$

is invariant to the choice of C_d^- and so equals $\sigma^2 s' C_d^{\dagger} m$. Now $s_i^{'}(I_v - \frac{1}{v} J_v)$ is zero if $i = 0$ and otherwise is $s_i^{'}$. This fact with (11) gives that for any i, $s_0^{'} C_d^P s_i = 0$; for any $i > 0$, $s_i^{'} C_d^P s_i = \frac{1}{z_{di}}$; and for any $i \neq i' \neq 0$, $s_i^{'} C_d^P s_{i'} = 0$. Write $L = (s_0, s_1, \ldots, s_{v-1})$ and let $D^{\dagger}$ be a $v \times v$ diagonal matrix whose diagonal elements are 0 and $\frac{1}{z_{di}}$ $(i = 1, \ldots, v-1)$. Noting that L is orthogonal, it has just been shown that $L' C_d^P L = D^{\dagger}$ and so

$$C_d^P = LL^{'} C_d^P LL^{'} = LD^{\dagger} L' = \sum_{i=1}^{v-1} \frac{1}{z_{di}} s_i s_i^{'} = C_d^{\dagger}. \qquad \square$$

For any $d \in \mathcal{D}$, a solution to the reduced normal equations can be found by setting the estimator $\hat{\tau}_1$ for treatment 1 to zero. For this choice $\mathrm{Cov}(\hat{\tau}_{(d)})$ is

$$\mathrm{Cov}(\hat{\tau}_{(d)}) = \sigma^2 \begin{pmatrix} 0 & 0_{v-1}' \\ 0_{v-1} & \varphi_d \end{pmatrix} \tag{12}$$

where φ_d is the $(v-1) \times (v-1)$ variance-covariance matrix for $\hat{\tau}_2, \ldots, \hat{\tau}_v$.

The following lemma in terms of C_d was first given by Chakrabarti (1963).

Lemma 4.2 *All cofactors of C_d (or $C_d^{\dagger}$) have the same value. Thus for $d \in \mathcal{D}$*

$$\prod_{i=1}^{v-1} \frac{1}{z_{di}} = \frac{1}{v\, Co(C_d)} = v\, Co(C_d^{\dagger}),$$

where $Co(.)$ denotes a matrix cofactor.

Corollary 4.3 *For any $d \in \mathcal{D}$, the D-value is*

$$\prod_{i=1}^{v-1} \frac{1}{z_{di}} = \frac{1}{v} |\varphi_d| \tag{13}$$

where $|.|$ denotes determinant and φ_d is as defined in (12).

PROOF The cofactor of the (1,1) element of $C_d^\dagger$ is $|(I_{v-1} - \frac{1}{v}J_{v-1})\varphi_d(I_{v-1} - \frac{1}{v}J_{v-1})|$, which is easily seen from (10) and (12). This determinant is $|I_{v-1} - \frac{1}{v}J_{v-1}|^2|\varphi_d| = \frac{1}{v^2}|\varphi_d|$ and now apply lemma 4.2. □

LEMMA 4.4 *$|\varphi_d|$, and thus the D-value, is invariant to the choice of $d \in \mathcal{D}$.*

PROOF Label so that the treatment common to all blocks in d^* is treatment 1. Then the solution producing φ_{d^*} in (12) is (after the initial 0) just the $b(k-1)$ simple within-block differences relative to unit $\lfloor\frac{k+1}{2}\rfloor$:

$$\hat{\tau}_{(d^*)} = (0, y_{11} - y_{1\lfloor\frac{k+1}{2}\rfloor}, y_{12} - y_{1\lfloor\frac{k+1}{2}\rfloor}, \ldots, y_{1k} - y_{1\lfloor\frac{k+1}{2}\rfloor}, y_{21} - y_{2\lfloor\frac{k+1}{2}\rfloor}, \ldots, y_{bk} - y_{b\lfloor\frac{k+1}{2}\rfloor})' \tag{14}$$

The $b(k-1)$ nonzero elements of (14) are clearly a basis for the space of y projected orthogonally to blocks (that is, for the estimation space of τ) regardless of the design. It follows immediately that for any $d \in \mathcal{D}$ the solution $\hat{\tau}_{(d)}$ found by setting $\hat{\tau}_1 = 0$ is

$$\hat{\tau}_{(d)} = \begin{pmatrix} 1 & 0'_{v-1} \\ 0_{v-1} & R_d \end{pmatrix} \hat{\tau}_{(d^\star)}$$

where R_d is nonsingular. Then from (12), $\varphi_d = R_d\varphi_{d^*}R'_d$ and the proof is done if $|R_dR'_d| = 1$. Lemma 2.1 says that R_d does not depend on Σ, so is the same as in the case $\Sigma = \sigma^2 I$. The proof (Bapat and Dey, 1991) that the D-value is invariant to design d in that case gives the result.

However one need not lean on the earlier graph-theoretic results. Here is a transparent alternative that works directly with the estimators to establish $|R_d| = \pm 1$. The solution vector $\hat{\tau}_{(d)}$ with $\hat{\tau}_1 = 0$ is comprised exactly of the contrasts in y corresponding to the unique chains from each treatment to the first treatment. For treatments in any block containing 1, these are the same as in $\hat{\tau}_{(d^*)}$. For any block not containing 1, say B_1, there is a unique linking treatment through which the chain to treatment 1 passes for every treatment (other than the linking treatment) in B_1; see (8). Indeed, it is easily seen that these chains differ only in their first members, and that for any block (say B_j) these chains pass through prior to reaching a block with treatment 1, the chains to

1 for treatments in B_j are, aside from their first members, the segments of the chains from B_1 starting from the linking treatment in B_j. Consequently, the chains to blocks containing treatment 1 induce a *partial order* on the set of blocks (the last element in a chain being the block containing treatment 1). This will be a crucial fact in the induction on b to follow.

For $b = 1$ and any k, suppose with no loss of generality that treatment 1 in d appears on plot $s > \lfloor \frac{k+1}{2} \rfloor$, and that otherwise treatments $2, 3, \ldots$ appear in order of the plots. Then it can be checked that

$$R_d = \left(\begin{array}{c|c} I_t & (-1_t | 0_{t\times(t-1)})P \\ \hline 0_{t\times t} & \left(\begin{array}{c|c} -1_t & \begin{array}{c} 0'_{t-1} \\ \hline I_{t-1} \end{array} \end{array} \right) P \end{array} \right) \quad \text{or} \quad \left(\begin{array}{c|c} I_t & (-1_t | 0_{t\times t})P \\ \hline 0_{(t+1)\times t} & \left(\begin{array}{c|c} -1_{t+1} & \begin{array}{c} 0'_{t} \\ \hline I_{t} \end{array} \end{array} \right) P \end{array} \right) \qquad (15)$$

as k is odd or even, where $t = \lfloor \frac{k+1}{2} \rfloor - 1$ and P is the permutation matrix that moves the first column to position $s - t - 1$ ($P = (p_{ij})$ with nonzero p_{ij} specified by $p_{1,s-t-1} = 1$, $p_{i,i-1} = 1$ for $2 \leq i \leq s-t-1$, and $p_{i,i} = 1$ for $i \geq s-t$). Clearly R_d has determinant ± 1.

Now assume the result holds for all saturated designs with b blocks and any k. Let d be a saturated design with $b + 1$ blocks of size k, and let $B_1, B_2, \ldots, B_{b+1}$ be any ordering of the blocks of d that respects the partial order. Relabel treatments other than treatment 1 so that B_1 contains either $1, \ldots, k$ or $2, \ldots, k + 1$ as B_1 does or does not contain 1. Then B_1 is the first block in a chain of blocks, the last of which contains 1 (if B_1 contains treatment 1, the chain consists only of B_1). Consequently, removing B_1 from d leaves a saturated design d' with b blocks.

If B_1 contains treatment 1, then obviously

$$R_d = \left(\begin{array}{c|c} R_1 & 0 \\ \hline 0 & R_{d'} \end{array} \right)$$

where R_1 is of form (15) with treatment 1 on plot s, implying $|R_d| = |R_1||R_{d'}|$ which is ± 1 by the induction hypothesis.

If B_1 does not contain treatment 1, let s be the position of the linking treatment in B_1, and let the linking treatment be labeled $k + 1$. Without loss of generality and aside from plot s, treatments $2, \ldots, k$ appear in plot order. For treatments in B_1, solutions

$\hat\tau_2, \ldots, \hat\tau_k$ in $\hat\tau_{(d)}$ involve measurements in B_1, but $\hat\tau_{k+1}$ does not. By the partial order, no chain to treatment 1 beginning with a treatment not in B_1 involves the differences arising from B_1. These statements say

$$R_d = \left(\begin{array}{c|c} R_1 & H \\ \hline 0 & R_{d'} \end{array}\right)$$

where again R_1 is of form (15), and here H is some nonzero matrix. Again $|R_d| = |R_1||R_{d'}| = \pm 1$, completing the proof. □

Corollary 4.3 and lemma 4.4 together produce the main result of this section.

THEOREM 4.5 *All designs in $\mathcal{D}(v,b,k)$ are D-equal, regardless of Σ.*

5 E-optimal Design

A design is E-optimal if it minimizes (in d) the maximum eigenvalue of the Moore-Penrose inverse $C_d^\dagger$ of the information matrix C_d. For given d, this eigenvalue is the maximum over all normalized contrasts $m'\tau$ of $\mathrm{var}(\widehat{m'\tau_{(d)}})/\sigma^2$. The study of E-optimality to follow requires the Moore-Penrose inverse $C_{d^*}^\dagger$ for the design d^* displayed in (6).

COROLLARY 5.1 *$C_{d^*}^\dagger$ for design d^* is*

$$C_{d^*}^\dagger = \begin{pmatrix} \frac{b}{v^2} 1_{k-1}' V 1_{k-1} & -\frac{1}{v} 1_b' \otimes [1_{k-1}' V (I_{k-1} - \frac{b}{v} J_{k-1})] \\ -\frac{1}{v} 1_b \otimes [(I_{k-1} - \frac{b}{v} J_{k-1}) V 1_{k-1}] & I_b \otimes V - \frac{1}{v} J_b \otimes (V J_{k-1} + J_{k-1} V) + \frac{b}{v^2} J_b \otimes (J_{k-1} V J_{k-1}) \end{pmatrix}$$

where V is the $(k-1) \times (k-1)$ correlation matrix for the $k-1$ simple within-block differences relative to $y_{1\lfloor\frac{k+1}{2}\rfloor}$ for the first block:

$$V = \frac{1}{\sigma^2} \mathrm{Cov}(y_{11} - y_{1\lfloor\frac{k+1}{2}\rfloor}, y_{12} - y_{1\lfloor\frac{k+1}{2}\rfloor}, \cdots, y_{1k} - y_{1\lfloor\frac{k+1}{2}\rfloor}). \tag{16}$$

PROOF This follows from (10) in lemma 4.1, using the solution specified in (14). □

THEOREM 5.2 *The largest eigenvalue of $C_{d^*}^\dagger$ is the largest eigenvalue of the correlation matrix V in (16).*

PROOF After factoring, $0 = \lambda|V - \lambda I_{k-1}|^{b-1}|V(I_{k-1} - \frac{b}{v}J_{k-1}) - \lambda I_{k-1}|$ is the characteristic equation for $C^{\dagger}_{d^*}$. Thus the non-zero eigenvalues of $C^{\dagger}_{d^*}$ are the eigenvalues of V with frequency $b-1$ each, and the eigenvalues of $V(I_{k-1} - \frac{b}{v}J_{k-1})$. Since the largest eigenvalue λ_{max} obeys $\lambda_{max}(AB) \le \lambda_{max}(A)\lambda_{max}(B)$ for any two positive definite matrices A and B (e.g. Bhatia, 1997, page 94), and the largest eigenvalue of $I_{k-1} - \frac{b}{v}J_{k-1}$ is 1,

$$\lambda_{max}(V(I_{k-1} - \frac{b}{v}J_{k-1})) \le \lambda_{max}(V)\lambda_{max}(I_{k-1} - \frac{b}{v}J_{k-1}) \le \lambda_{max}(V). \tag{17}$$

□

Theorem 5.2 provides the standard for evaluating any design d relative to d^* in terms of the E-criterion. If there is a normalized treatment contrast $m'\tau$ such that $\text{var}(\widehat{m'\tau_{(d)}})/\sigma^2$ is no less than the largest eigenvalue of V, then d cannot be E-superior to d^*. The question is how to pick such a contrast for an arbitrary d.

Corollary 2.4 says that for any d one can always find two blocks, call them B_1 and B_2, sharing a common treatment, call it treatment 1. Setting the estimator for treatment 1 to zero, and labelling the other treatments in these two blocks appropriately, one *partial* solution to the reduced normal equations is

$$\begin{aligned}\hat{\tau}_{(d)} = (0, &y_{11} - y_{1s}, y_{12} - y_{1s}, \cdots, y_{1,s-1} - y_{1s}, y_{1,s+1} - y_{1s}, \cdots, y_{1k} - y_{1s},\\ &y_{21} - y_{2t}, y_{22} - y_{2t}, \cdots, y_{2,t-1} - y_{2t}, y_{2,t+1} - y_{2t}, \cdots, y_{2k} - y_{2t}, \cdots)',\end{aligned} \tag{18}$$

where s and t are the positions of treatment 1 in the two blocks. This is a partial solution in the sense that only the estimators for treatments in B_1 and B_2 have been displayed. It will be used to calculate variances for contrasts of treatments in these two blocks. Let V_1 and V_2 be the *(k-1)×(k-1)* correlation matrices

$$V_1 = \frac{1}{\sigma^2}\text{Cov}(y_{11} - y_{1s}, y_{12} - y_{1s}, \cdots, y_{1,s-1} - y_{1s}, y_{1,s+1} - y_{1s}, \cdots, y_{1k} - y_{1s})$$
$$V_2 = \frac{1}{\sigma^2}\text{Cov}(y_{21} - y_{2t}, y_{22} - y_{2t}, \cdots, y_{2,t-1} - y_{2t}, y_{2,t+1} - y_{2t}, \cdots, y_{2k} - y_{2t})$$

Then the partial expression of $\text{Cov}(\hat{\tau}_{(d)})$ corresponding to (18) is:

$$\frac{1}{\sigma^2}\text{Cov}(\hat{\tau}_{(d)}) = \begin{pmatrix} 0 & 0'_{k-1} & 0'_{k-1} & 0'_{(b-2)(k-1)} \\ 0_{k-1} & V_1 & 0_{k-1} & \cdots \\ 0_{k-1} & 0_{k-1} & V_2 & \cdots \\ 0_{(b-2)(k-1)} & \cdots & \cdots & \cdots \end{pmatrix} \tag{19}$$

Lemma 5.3 *If $\lambda_{max}(\frac{V_1+V_2}{2}) \geq \lambda_{max}(V)$, then design d cannot be E-superior to d^*.*

Proof Suppose the normalized eigenvector corresponding to the largest eigenvalue of $\frac{V_1+V_2}{2}$ is $x_{(k-1)\times 1}$. Then from (19), $\mathrm{var}(\widehat{m'\tau_{(d)}})/\sigma^2 = m'\mathrm{Cov}(\hat{\tau}_{(d)})m/\sigma^2 = \lambda_{max}(\frac{V_1+V_2}{2})$ where m is the $v \times 1$ vector specified by $m' = \frac{1}{\sqrt{2}}(0, x', -x', 0_{1\times(v-2k+1)})$. □

It also follows that no design with two blocks sharing a treatment at the center position $\lfloor \frac{k+1}{2} \rfloor$ can be E-superior to d^* (this is just $V_1 = V_2 = V$).

The variance-covariance structure Σ for the entire layout under (3)-(5) can be written as $\Sigma = I_b \otimes \Sigma_k$ where Σ_k is the $k \times k$ within-block covariance matrix. For $u = 1, 2, \cdots, k$ define H_u as the $k \times k$ matrix with u^{th} column all 1's, and all other elements 0's. Compute $(I - H_u)\Sigma_k(I - H_u^{'})/\sigma^2$, remove the u^{th} row and the u^{th} column, and name the resulting matrix Γ_u. Obviously Γ_u is a positive definite $(k-1) \times (k-1)$ matrix. In lemma 5.3, $V = \Gamma_{\lfloor \frac{k+1}{2} \rfloor}$, $V_1 = \Gamma_s$, and $V_2 = \Gamma_t$. This leads to

Theorem 5.4 *Design d^* is E-optimal if*

$$\min(\lambda_{max}\left(\frac{\Gamma_{u_1} + \Gamma_{u_2}}{2}\right)) = \lambda_{max}(\Gamma_{\lfloor \frac{k+1}{2} \rfloor})$$

where the minimum is over $1 \leq u_1 \leq u_2 \leq \lfloor \frac{k+1}{2} \rfloor$.

Proof This evaluates the eigenvalue bound employed in lemma 5.3 for every pair of linking positions. Since any $d \in \mathcal{D}$ must possess two blocks sharing a linking treatment, the minimum of these quantities is a lower bound for the largest eigenvalue of $C_d^{\dagger}$. □

The sufficient condition as stated in Theorem 5.4 is actually one of a family of sufficient conditions for d^* to be E-optimal. In the set of indices $\{1, 2, \ldots, \lfloor \frac{k+1}{2} \rfloor\}$ where u_1 and u_2 take their values, any member, say u, can be replaced by $k+1-u$. Consequently, there are $2^{\lfloor \frac{k+1}{2} \rfloor}$ sets of sufficient conditions, although some of these are identical. One of these will be employed in the proof for Corollary 5.6.

Theorem 5.4 provides a method for establishing E-optimality of d^*, but it requires comparing largest eigenvalues of $\lfloor \frac{k+1}{2} \rfloor - 1 + \binom{\lfloor \frac{k+1}{2} \rfloor}{2}$ matrices to that of C_{d^*}, that is, enumerating all possibilities for u_1 and u_2 except $u_1 = u_2 = \lfloor \frac{k+1}{2} \rfloor$. This is an analytically impossible task when k is large. That the condition of Theorem 5.4 may be too strong can been seen if enough ρ_h are zero, for then $\Gamma_u = P\Gamma_{\lfloor \frac{k+1}{2} \rfloor}P'$ for some $u \neq \lfloor \frac{k+1}{2} \rfloor$ and permutation matrix P, implying the vector of eigenvalues for $\Gamma_{\lfloor \frac{k+1}{2} \rfloor}$ majorizes that of $\frac{1}{2}(\Gamma_u + \Gamma_{\lfloor \frac{k+1}{2} \rfloor})$. A consequence is that d^* cannot be shown E-superior by this method even though this Γ_u is equivalent to $\Gamma_{\lfloor \frac{k+1}{2} \rfloor}$ in terms of its contribution to contrast variances. Theorem 5.4 can nonetheless be used to find E-optimal designs in particular cases. Below are two examples from Jin (2004). Proofs are in appendix A.

Corollary 5.5 *Design d^* is E-optimal when $k = 3$.*

Corollary 5.6 *Design d^* is E-optimal when $k = 4$ and the covariance structure is defined as $\rho_s = \rho^s$ for $0 < \rho < 1$.*

6 Discussion

Spatial correlation of observations has been found to impose positional conditions on optimal designs. This is not surprising, for much stronger positional balancing is found in the optimality conditions determined for non-saturated block designs with correlated errors in papers such as Kunert (1987), Morgan and Chakravarti (1988), Martin and Eccleston (1991), Bhaumik (1995), and Benchekroun and Chakravarti (1999). What may be surprising is that regardless of the strength of positive correlation, the unequal replication found for optimal designs when $\Sigma = \sigma^2 I$ is maintained. The results here hold uniformly for all correlations, known or unknown, that respect the spatial non-increasing property (4).

The general prescription of these results is to use the same design known to be optimal for uncorrelated measurements, but to place the common treatment at the center position in all blocks. Among other results, this strategy is shown to be A-optimal, MV-optimal, and D-optimal. We conjecture that the same strategy is E-optimal, but have been unable to obtain a proof other than in special cases. Theorem 5.4 captures the essence of the problem: if the ρ_h in (4) are all distinct and positive, one needs to get hold of the maximum eigenvalue for a pair of blocks with arbitrary linking positions.

This seems to be a complex task for any general correlation structure allowed by (3) and (4). Even for the 1-dimensional parameterization offered by AR(1) correlations, we have done no better than proving optimality of d^* for $k \leq 4$. Theorem 5.4 is certainly adequate, however, from an applied perspective: for given k and feasible correlations, one can computationally check that the condition for E-optimality of d^* holds.

Is there a substantial advantage to using d^* instead of any of the designs that are optimal when the errors are uncorrelated, that is, instead of an OLS-optimal design? The first scenario in the proof of Theorem 3.1 makes it clear that d^* dominates any OLS-optimal design, other than d^* itself, whenever the correlations (4) are not all equal. The strength of this dominance depends on the exact values of the ρ_h, and on the competitor, for the performance of an OLS-optimal design monotonically degrades as the position of the linking treatment in any block is moved further from the center. Since within-block contrasts are estimated with the same variance for any design, designs can be compared through their between-blocks variances (7). For d^* with odd k (even k is similar), the mean of these $(k-1)^2$ variances is $4[(k-1) - 2\sum_{h=1}^{(k-1)/2} \rho_h]/(k-1)$, while for two blocks linked at their end plots, the mean is $4[(k-1) - \sum_{h=1}^{k-1} \rho_h]/(k-1)$. The ratio of these two figures can be as large as 1 (when all ρ_h are equal), but can also be quite small (e.g. 0.5 for $k = 3, \rho_1 = 0.7, \rho_2 = 0.1$).

ACKNOWLEDGEMENT

This paper has evolved during the review process. We thank the referees for their patience and careful reading. J. P. Morgan was supported by National Science Foundation grant DMS01-04195.

A Proofs for Corollaries 5.5 and 5.6

Proof for corollary 5.5

Set $k = 3$ and let the within-block covariance matrix Σ_k be any matrix allowed by (3) and (4). Computing $(I - H_2)\Sigma_k(I - H_2^{'})/\sigma^2$ and deleting its second row and column produces Γ_2:

$$\Gamma_2 = \begin{pmatrix} 2-2\rho_1 & 1-2\rho_1+\rho_2 \\ 1-2\rho_1+\rho_2 & 2-2\rho_1 \end{pmatrix}$$

The eigenvalues of Γ_2 are $1-\rho_2$ and $3-4\rho_1+\rho_2$, which both must be positive; the larger is the E-value for d^*. Similarly Γ_1 is $(I-H_1)\Sigma_k(I-H_1^{'})/\sigma^2$ with the first row and column deleted:

$$\Gamma_1 = \begin{pmatrix} 2-2\rho_1 & 1-\rho_2 \\ 1-\rho_2 & 2-2\rho_2 \end{pmatrix}.$$

The largest eigenvalue γ_1 of Γ_1 is $2-\rho_1-\rho_2+\sqrt{1+\rho_1^2-2\rho_2-2\rho_1\rho_2+2\rho_2^2}$. Also needed is $\frac{\Gamma_1+\Gamma_2}{2}$, with largest eigenvalue $\gamma_{1,2}=\frac{1}{2}[4-3\rho_1-\rho_2+\sqrt{4-8\rho_1+5\rho_1^2-2\rho_1\rho_2+\rho_2^2}\ ]$.

With these quantities in hand, it remains to compare the various eigenvalues. For $\frac{\Gamma_1+\Gamma_2}{2}$ versus Γ_1, compute

$$\gamma_{1,2}-\gamma_1 = \frac{1}{2}[\sqrt{4(1-\rho_1)^2+(\rho_1-\rho_2)^2}-\sqrt{4(1-\rho_2)^2+4(\rho_1-\rho_2)^2}-(\rho_1-\rho_2)] \leq 0.$$

So it must be shown that Γ_2 has largest eigenvalue smaller than $\gamma_{1,2}$. Similarly straightforward manipulations give $1-\rho_2-\gamma_{1,2} \leq -(1-\rho_1) < 0$, and $3-4\rho_1+\rho_2-\gamma_{1,2} \leq -\frac{3(\rho_1-\rho_2)}{2} < 0$ for $\rho_1 \neq \rho_2$. $\square$

Proof for corollary 5.6

Take $k=4$ and let the within-block covariance matrix Σ_k be specified by (3) with $\rho_{|u-u'|}=\rho^{|u-u'|}$. The relevant matrix for d^* is either Γ_2 or Γ_3. Here Γ_3 will be compared to Γ_1 and $\Gamma_{1,3} \equiv \frac{\Gamma_1+\Gamma_3}{2}$. Explicit expressions for Γ_1, Γ_3, and $\Gamma_{1,3}$, as well as a more detailed version of this proof, can be found in Jin (2004). The characteristic equations for these three matrices are

$$\begin{aligned} G_1(\lambda) &= -2(2-\rho)(1-\rho)^3(1+\rho)^2-(1-\rho)^2(1+\rho)(-9-\rho-2\rho^2+2\rho^3)\lambda \\ &\quad +2(-1+\rho)(3+2\rho+\rho^2)\lambda^2+\lambda^3=0, \\ G_3(\lambda) &= -2(2-\rho)(1-\rho)^3(1+\rho)^2-(1-\rho)^2(1+\rho)(-9+\rho-\rho^2+\rho^3)\lambda \\ &\quad -2(1-\rho)(3+\rho)\lambda^2+\lambda^3=0, \quad \text{and} \\ G_{1,3}(\lambda) &= (1-\rho)^3(1+\rho)(-16-8\rho-6\rho^2+3\rho^3+\rho^4)-(1-\rho)^2(-36-36\rho-19\rho^2 \\ &\quad -6\rho^3+\rho^4)\lambda-4(1-\rho)(6+3\rho+\rho^2)\lambda^2+4\lambda^3=0. \end{aligned}$$

Evaluation of G_1 shows that $G_1(0) < 0$, $G_1(1-\rho) > 0$, $G_1(2-2\rho^2) < 0$, and $G_1(4) > 0$. So the three roots are within $(0, 1-\rho)$, $(1-\rho, 2-2\rho^2)$ and $(2-2\rho^2, 4)$ respectively. The largest eigenvalue of Γ_1, say γ_1, is within $(2-2\rho^2, 4)$. Evaluating G_3 and $G_{1,3}$ at the same four points yields the same strict inequalities, so that the eigenvalues for Γ_3 and $\Gamma_{1,3}$ fall in the same intervals. Denoting the largest eigenvalues for these two matrices by γ_3 and $\gamma_{1,3}$, then all of γ_1, γ_3 and $\gamma_{1,3}$ are in $(2-2\rho^2, 4)$. Each of $G_1(\gamma)$, $G_3(\gamma)$ and $G_{1,3}(\gamma)$ is negative at $\gamma = 2-2\rho^2$ and has a single root in $\gamma > 2-2\rho^2$.

The three characteristic functions obey

$$\begin{aligned} G_3(\lambda) &= G_1(\lambda) + \lambda\rho(1-\rho^2)[2\lambda - (2-\rho)(1-\rho)(1+\rho)] \quad (20) \\ &= \frac{G_{1,3}(\lambda)}{4} + \frac{\rho(1-\rho)}{4}[4\lambda^2(1+\rho) - (1-\rho)^2(1+\rho)\rho(-14+3\rho+\rho^2) \\ &\quad -\lambda(1-\rho)(4+19\rho+6\rho^2+3\rho^3)]. \quad (21) \end{aligned}$$

Now for $\gamma > 2-2\rho^2$, $G_3(\gamma) < 0 \Rightarrow \gamma < \gamma_3$ and $G_3(\gamma) > 0 \Rightarrow \gamma > \gamma_3$. Using (20) to evaluate G_3 at $\gamma_1 \in (2-2\rho^2, 4)$ gives

$$\begin{aligned} G_3(\gamma_1) &= G_1(\gamma_1) + \gamma_1\rho(1-\rho^2)[2\gamma_1 - (2-\rho)(1-\rho)(1+\rho)] \\ &> \gamma_1\rho(1-\rho^2)[2(2-2\rho^2) - (2-\rho)(1-\rho)(1+\rho)] \\ &= \gamma_1\rho(1-\rho^2)(1-\rho)(1+\rho)(2+\rho) > 0, \end{aligned}$$

showing that $\gamma_1 > \gamma_3$.

Finally, for $\gamma_{1,3} \in (2-2\rho^2, 4)$, (21) says that $G_3(\gamma_{1,3}) = H(\gamma_{1,3})$ where

$$\begin{aligned} H(\gamma) &= \frac{\rho(1-\rho)}{4}[4\gamma^2(1+\rho) - (1-\rho)^2(1+\rho)\rho(-14+3\rho+\rho^2) \\ &\quad -\gamma(1-\rho)(4+19\rho+6\rho^2+3\rho^3)]. \end{aligned}$$

Differentiating $H(\gamma)$ with respect to γ gives

$$\begin{aligned} \frac{\partial H(\gamma)}{\partial \gamma} &= \frac{\rho}{4}[8\gamma(1-\rho^2) - (1-\rho)^2(4+19\rho+6\rho^2+3\rho^3)] \\ &> \frac{\rho}{4}[8(2-2\rho^2)(1-\rho^2) - (1-\rho)^2(4+19\rho+6\rho^2+3\rho^3)] \\ &= \frac{1}{4}(1-\rho)^2\rho(12+13\rho+10\rho^2-3\rho^3) > 0 \end{aligned}$$

so that $H(\gamma)$ is increasing in $\gamma > 2-2\rho^2$. But $H(2-2\rho^2) = \frac{1}{4}(1-\rho)^3\rho(1+\rho)(8+8\rho+\rho^2-7\rho^3) > 0$. Thus $G_3(\gamma_{1,3}) > 0$ and consequently $\gamma_{1,3} > \gamma_3$. □

REFERENCES

Bagchi, B. and Bagchi, S. (2001). Optimality of partial geometric designs. *Annals of Statistics* **29**, 577–594.

Bapat, R. B. and Dey, A. (1991). Optimal block designs with minimal number of observations. *Statistics and Probability Letters* **11**, 399–402.

Benchekroun, K. and Chakravarti, I. M. (1999). Correlation uniform incomplete block designs. *Journal of Statistical Planning and Inference* **76**, 263–272.

Bhatia, R. (1997). *Matrix Analysis.* Springer-Verlag, New York.

Bhaumik, D. (1995). Majorization and D-optimality for complete block designs under correlations. *Journal of the Royal Statistical Society Series B* **57**, 139–143.

Chakrabarti, M. C. (1962). *Mathematics of Design and Analysis of Experiments.* Asia Publishing House, Bombay.

Chakrabarti, M. C. (1963). On the C-matrix in design of experiments. *Journal of the Indian Statistical Society* **1**, 8–23.

Dey, A., Shah, K. R., and Das, A. (1995). Optimal block designs with minimal and nearly minimal number of units. *Statistica Sinica* **5**, 547–558.

Jin, B. (2004). *Optimal Block Designs with Limited Resources.* Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Blacksburg.

Kempthorne, O. (1956). The efficiency factor of an incomplete block design. *Annals of Mathematical Statistics* **27**, 846–849.

Kunert, J. (1987). Neighbour balanced block designs for correlated errors. *Biometrika* **74**, 717–724.

Mandal, N. K., Shah, K. R., and Sinha, B. K. (1991). Uncertain resources and optimal designs: problems and perspectives. *Calcutta Statistical Association Bulletin* **40**, 267–282.

Martin, R. J. and Eccleston, B. K. (1991). Optimal incomplete block designs for general dependence structures. *Journal of Statistical Planning and Inference* **28**, 67–81.

Morgan, J. P. and Chakravarti, I. M. (1988). Block designs for first and second order neighbor correlations. *Annals of Statistics* **16**, 1206–1224.

Searle, S. (1971). *Linear Models.* John Wiley & Sons, New York.

Additional Results

A few additional results, found after the paper was accepted

Theorem 3.1 can be generalized to include blocks of fixed, but unequal, sizes. Let the block sizes be $k_1 \geq \ldots \geq k_b$ for a saturated setting, that is, $n = \sum_{j=1}^{b} k_j = b + v - 1$. Lemma 2.1 and so corollaries 2.2-2.4 still hold. The proof of Theorem 3.1 depends only on the blocks having fixed sizes and not on those sizes being equal. Thus M-optimality for estimation of elementary treatment contrasts holds for d^* defined to have one treatment common to every block, appearing on unit $\lfloor \frac{k_j+1}{2} \rfloor$ in block j. This result for A-optimality in the uncorrelated case (when the one common treatment can appear in any position in each block), here extended to all optimality criteria of the form $\sum_i \sum_{i'>i} f(\text{var}(\hat{\tau}_i - \hat{\tau}_{i'}))$ for nondecreasing f, was established by Das, Dean, and Notz (1998, *JSPI* 72, 133-147).

If all $\rho_h = 0$ (the uncorrelated case), then an optimal design as in the preceding paragraph for fixed $k_1 \geq \ldots \geq k_b$ has only two distinct variances for pairwise comparisons, the smaller of which is that for a within-blocks elementary treatment contrast. So consider fixing v, b, and thus $n = b + v - 1$, but otherwise allowing the block sizes to be arbitrary. The best design in this wider class can be easily determined: it will maximize the number of pairwise comparisons estimated by within-blocks contrasts. That is, it will maximize $\sum_{j=1}^{b} \binom{k_j}{2}$. Now a block with only one experimental unit is disconnected from the other observations and so does not contribute to treatment contrasts estimation. Consequently, a plausible but not necessary restriction is $k_j \geq 2$ for all j. The only necessary requirement imposed by fixed b is that $k_j \geq 1$ for all j. Alternatively, b and thus n could be allowed to vary subject only to $n = b + v - 1$; decreasing b is equivalent to setting some k_j to zero, in which case the k_j's are totally unrestricted. Next listed are the best designs in each case:

restriction	M-best block sizes
fixed b, all $k_j \geq 2$	$k_1 = v - b - 1,\ k_2 = \ldots k_b = 2$
fixed b, all $k_j \geq 1$	$k_1 = v,\ k_2 = \ldots k_b = 1$
b not fixed	$b = 1,\ k_1 = v$

The best of all saturated block designs is a single complete block, which is the choice that minimizes the number of experimental units employed! Of course, one is not always able to get sufficiently uniform experimental units to comprise a single block of size v. The bottom line here is that, within whatever limitations are in force in planning a particular experiment (which may well fix b and preclude a block of size v), one should choose blocks so as to maximize $\sum_{j=1}^{b} \binom{k_j}{2}$. Note that the second line in the above table with $k_2 = \ldots, k_b = 1$, while mathematically correct, is experimentally silly.

We are now exploring the issues for variable block sizes in the correlated case.

It is not hard to see that our results for D-optimality also hold for any fixed $k_1 \geq \ldots \geq k_b$.