On the determination of optimal designs for an interference model

by J. Kunert[1] and R.J. Martin[2]

[1] Fachbereich Statistik, Universität Dortmund, D-44221 Dortmund, Germany.

[2] Department of Probability and Statistics, University of Sheffield, Sheffield S3 7RH, U.K.

Abstract

This paper generalizes Kushner's (1997) method for finding optimal repeated measurements designs to optimal designs under an interference model. The model we assume is for a one-dimensional layout without guard plots and with different left and right neighbour effects. The resulting optimal designs may need many blocks or may not even exist as a finite design. The results give lower bounds for optimality criteria on finite designs, and the design structure can be used to suggest efficient small designs.

Some key words: Interference model; Neighbour effect; Optimal design; Universal optimality.

AMS subject classification: Primary 62K05, 62K10, Secondary 62P10

Abbreviated form of the title: Optimal designs for an interference model

1. Introduction

Many agricultural and horticultural trials are susceptible to treatment interference, that is the treatment on one unit affecting the response on neighbouring units (see e.g. Besag and Kempton, 1986). There is increasing interest in the practical use of models to analyse data from such trials (e.g. David, Monod and Philippeau, 1998), and in the design of experiments in which treatment interference may occur (e.g. David and Kempton, 1996). A wide variety of possible models have been postulated (e.g. David and Kempton, 1996, David, Monod and Philippeau, 1998). There are only very limited results on optimal designs under interference models. Gill (1993) restricts the class of competing designs to those for which each treatment appears once in each block. Druilhet (1999) avoids this restriction but considers the case of very few blocks. Both papers assume a one-dimensional layout of plots within blocks, and that each block has a guard plot at each end, so that each interior plot has two neighbours. They concentrate on model (1) below, or its special case of equal left and right neighbour effects.

The present paper presents a general approach to determine optimal designs for contrasts among direct treatment effects that can be useful for many kinds of interference models. We consider experiments for comparing t treatments using b blocks of size k with a one-dimensional arrangement of plots in each block. We demonstrate the theory for the model with no guard plots, and the treatments having different left and right neighbour interference effects. Similar results to the ones given here will be possible for many other related models.

Let $d(i, j) \in \{1, ..., t\}$ be the treatment assigned to the plot (i, j) in the j-th position of the i-th block. In our model the response at plot (i, j) can be written

$$y_{ij} = \mu + \tau_{d(i,j)} + \lambda_{d(i,j-1)} + \rho_{d(i,j+1)} + \beta_i + e_{i,j}\,. \qquad (1)$$

Here

μ is the general mean,

$\tau_{d(i,\,j)}$ is the direct effect of treatment $d(i, j)$,

$\lambda_{d(i,\,j\text{-}1)}$ and $\rho_{d(i,\,j+1)}$ are, respectively, the left and right neighbour effects, that is the interference effect of the treatment assigned to, respectively, the left and right neighbour plots $(i, j\text{-}1)$ and $(i, j+1)$,

β_i is the effect of the i-th block, and $e_{i,\,j}$ is the random error, $1 \le i \le b$, $1 \le j \le k$.

We assume that the errors are i.i.d. with expectation 0. The generalization of the method to correlated errors and generalized least squares estimation is straightforward, cf. Kushner (1997). Since we assume there are no guard plots we have $\lambda_{d(i,\,0)} = \rho_{d(i,\,k+1)} = 0$.

We seek the optimal design among designs $d \in \Omega_{t,b,k}$, the set of all designs with b blocks of size k and with t treatments. Let T_{du} be the treatment design matrix of the direct effects in block u, $1 \le u \le b$. Further define $T_d = \left[T_{d1}^T, \ldots, T_{db}^T\right]^T$ as the design matrix of direct effects.

Let $Y = \left[y_{1,1}, \ldots, y_{1,k}, y_{2,1}, \ldots y_{b,k}\right]^T$ be the vector of the observations, 1_k be the k-vector of ones, I_b the b-dimensional identity matrix, and $\otimes$ denote the Kronecker product. Let V denote the $k \times k$ left neighbour incidence matrix with (i, j)-th element $v_{i,\,j}$ equal to 1 if $i - j = 1$, and 0 otherwise. For each u we define $L_{du} = VT_{du}$ and $R_{du} = V^T T_{du}$. Then $L_d = \left[L_{d1}^T, \ldots, L_{db}^T\right]^T$ = $\left(I_b \otimes V\right) T_d$ and $R_d = \left[R_{d1}^T, \ldots, R_{db}^T\right]^T$ = $\left(I_b \otimes V^T\right) T_d$ are, respectively, the design matrices of the left and right neighbour effects. Let e be the vector of the errors, and let τ, λ, ρ, and β be the vectors of direct effects, of left neighbour effects, of right neighbour effects, and of block effects, respectively. Then, we can write model (1) in vector notation as

$$Y = 1_{bk}\mu + T_d\tau + L_d\lambda + R_d\rho + (I_b \otimes 1_k)\beta + e\,.$$

For an $n \times p$ matrix M define $\omega^{\perp}(M) = I_n - M(M^T M)^- M^T$, where $(M^TM)^-$ is a generalized inverse (g-inverse) of M^TM. Then (see e.g. Kunert, 1983) the information matrix for the least squares estimate of τ, with zero row and column sums, is

$$C_d = T_d^T \omega^{\perp}([I_b \otimes 1_k, L_d, R_d])T_d\,.$$

A $t \times t$ matrix M is said to be completely symmetric, if all its diagonal elements are equal and all its off-diagonal elements are equal. A completely symmetric information matrix is a scalar multiple of the matrix $B_t = I_t - \frac{1}{t}1_t 1_t^T$. Assume we have a design $d^* \in \Omega_{t,b,k}$ such that C_{d^*} is completely symmetric and that tr C_{d^*} is maximal over $\Omega_{t,b,k}$. Then the design d^* is universally optimum, i.e. it is optimal under all the optimality criteria considered by Kiefer (1975).

2. Determination of an upper bound for tr C_d.

For a partitioned matrix $M = [S, U]$, we can write

$$\omega^{\perp}([S,U]) = \omega^{\perp}(S) - \omega^{\perp}(S)U\left\{U^T\omega^{\perp}(S)U\right\}^- U^T\omega^{\perp}(S)\,. \qquad (2)$$

Applying this formula twice and defining

$$C_{d11} = T_d^T\omega^{\perp}(I_b \otimes 1_k)T_d\,,\ C_{d12} = T_d^T\omega^{\perp}(I_b \otimes 1_k)L_d\,,\ C_{d13} = T_d^T\omega^{\perp}(I_b \otimes 1_k)R_d\,,$$

$$C_{d22} = L_d^T\omega^{\perp}(I_b \otimes 1_k)L_d\,,\ C_{d23} = L_d^T\omega^{\perp}(I_b \otimes 1_k)R_d\,,\ C_{d33} = R_d^T\omega^{\perp}(I_b \otimes 1_k)R_d$$

we get that

$$C_d = C_{d11} - C_{d12}C_{d22}^{-}C_{d12}^{T} - \left(C_{d13} - C_{d12}C_{d22}^{-}C_{d23}\right)\left(C_{d33} - C_{d23}^{T}C_{d22}^{-}C_{d23}\right)^{-}\left(C_{d13} - C_{d12}C_{d22}^{-}C_{d23}\right)^{T}. \tag{3}$$

Note that $\omega^{\perp}(I_b \otimes 1_k) = I_b \otimes B_k$. The formula for C_d contains g-inverses of C_{d22} and of $C_{d33} - C_{d23}^{T}C_{d22}^{-}C_{d23}$, both of which depend on the design d. This makes the determination of tr C_d for an arbitrary design d difficult. Hence, we try to find a simple upper bound for tr C_d.

The derivation of this bound is inspired by the convexity argument of Pukelsheim (1993, p. 75), see also Kushner (1997, Lemma 5.1). We give a slightly different proof, which is also valid if the matrices do not have full rank. We begin with a technical proposition.

Proposition 1:

<u>Assume</u> $A_1,\ldots,A_n,D_1,\ldots,D_n$ <u>are matrices,</u> $A_i \in IR^{m_i\times r}, D_i \in IR^{m_i\times s}, 1 \le i \le n$. <u>Then</u>

$$\sum A_i^T A_i - \left(\sum A_i^T D_i\right)\left(\sum D_i^T D_i\right)^{-}\left(\sum D_i^T A_i\right) \ge \sum\left\{A_i^T A_i - A_i^T D_i\left(D_i^T D_i\right)^{-} D_i^T A_i\right\}$$

<u>in the Loewner-ordering.</u>

Proof:

Consider the partitioned matrices

$$M_1 = \begin{bmatrix} D_1 \\ \vdots \\ D_n \end{bmatrix} \text{ and } M_2 = \begin{bmatrix} D_1 & & \\ & \ddots & \\ & & D_n \end{bmatrix}.$$

The column-space of M_1 is contained in the column-space of the block diagonal matrix M_2.

Hence,

$$\omega^{\perp}(M_1) \geq \omega^{\perp}(M_2) = \begin{bmatrix} \omega^{\perp}(D_1) & & \\ & \ddots & \\ & & \omega^{\perp}(D_n) \end{bmatrix}$$

and

$$\sum A_i^T A_i - \left(\sum A_i^T D_i\right)\left(\sum D_i^T D_i\right)^{-}\left(\sum D_i^T A_i\right) = \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}^T \omega^{\perp}(M_1) \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}$$

$$\geq \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}^T \begin{bmatrix} \omega^{\perp}(D_1) & & \\ & \ddots & \\ & & \omega^{\perp}(D_n) \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix} = \sum \left(A_i^T A_i - A_i^T D_i \left(D_i^T D_i \right)^{-} D_i^T A_i \right). \qquad \square$$

Note that $T_d 1_t$ is in the column-space of $I_b \otimes 1_k$, while $R_d 1_t$ and $L_d 1_t$ are not. This implies (see Kunert, 1983) that C_{d11} has row and column sums zero, that C_{d12} and C_{d13} have column sums zero, but not necessarily row sums zero, and that C_{d22}, C_{d23} and C_{d33} need not have zero row sums or column sums. For our bound, we use the traces of $B_t C_{dij} B_t$, and define $c_{dij} = \operatorname{tr} B_t C_{dij} B_t$ for $1 \leq i \leq j \leq 3$.

Since the matrix

$$\begin{bmatrix} B_t C_{d11} B_t & B_t C_{d12} B_t & B_t C_{d13} B_t \\ B_t C_{d12}^T B_t & B_t C_{d22} B_t & B_t C_{d23} B_t \\ B_t C_{d13}^T B_t & B_t C_{d23}^T B_t & B_t C_{d33} B_t \end{bmatrix} = \begin{bmatrix} B_t T_d^T \\ B_t L_d^T \\ B_t R_d^T \end{bmatrix} \omega^{\perp}(I_b \otimes 1_k) \begin{bmatrix} T_d B_t & L_d B_t & R_d B_t \end{bmatrix}$$

is nonnegative definite, this also holds for

$$\begin{bmatrix} c_{d11} & c_{d12} & c_{d13} \\ c_{d12} & c_{d22} & c_{d23} \\ c_{d23} & c_{d23} & c_{d33} \end{bmatrix}.$$

This implies directly that $c_{dii} \geq 0$, $1 \leq i \leq 3$ and that $c_{d22}c_{d33} - c_{d23}^2 \geq 0$. It also follows that (see, e.g. Rao and Toutenburg, 1995, Theorem A74)

$$Q = \begin{bmatrix} c_{d22} & c_{d23} \\ c_{d23} & c_{d33} \end{bmatrix}, \text{ satisfies } QQ^- \begin{bmatrix} c_{d12} \\ c_{d13} \end{bmatrix} = \begin{bmatrix} c_{d12} \\ c_{d13} \end{bmatrix}, \tag{4}$$

and, consequently, that

$$[c_{d12} \quad c_{d13}]Q^- \begin{bmatrix} c_{d12} \\ c_{d13} \end{bmatrix}$$

does not depend on the choice of the g-inverse Q^-.

We are therefore in a position to define

$$q^*_d = c_{d11} - [c_{d12} \quad c_{d13}]Q^- \begin{bmatrix} c_{d12} \\ c_{d13} \end{bmatrix}.$$

Then q^*_d depends on the following four cases (i) to (iv):

(i) If $c_{d22}c_{d33} - c_{d23}^2 > 0$, then Q is nonsingular and

$$q^*_d = c_{d11} - \frac{c_{d12}^2 c_{d33} - 2c_{d12}c_{d13}c_{d23} + c_{d13}^2 c_{d22}}{c_{d22}c_{d33} - c_{d23}^2}.$$

(ii) If $c_{d22}c_{d33} - c_{d23}^2 = 0$ and $c_{d22} > 0$, then $q^*_d = c_{d11} - c_{d12}^2 / c_{d22}$.

(iii) If $c_{d22} = 0$ and $c_{d33} > 0$, then $c_{d23} = 0$ and $q^*_d = c_{d11} - c_{d13}^2 / c_{d33}$.

(iv) If $c_{d22} = c_{d33} = 0$, then $c_{d23} = 0$ and $q^*_d = c_{d11}$.

With these definitions we can show

Proposition 2:

Every design $d \in \Omega_{t,b,k}$ has tr $C_d \le q^*_d$. If a design f has all C_{fij}, $1 \le i \le j \le 3$, completely symmetric, then tr $C_f = q^*_f$.

Proof:

Using formula (2), C_d can also be written as

$$C_d = \tilde{T}_d^T \omega^{\perp}([\tilde{L}_d, \tilde{R}_d])\tilde{T}_d\,, \tag{5}$$

where $\tilde{T}_d = \omega^{\perp}(I_b \otimes 1_k)T_d$, $\tilde{L}_d = \omega^{\perp}(I_b \otimes 1_k)L_d$, and $\tilde{R}_d = \omega^{\perp}(I_b \otimes 1_k)R_d$.

In Proposition 1 let $n = t!$ and consider $\{S_1 = I_t, S_2, ..., S_n\}$, the set of all $t \times t$ permutation matrices. Then define $A_i = \tilde{T}_d S_i$, $D_i = [\tilde{L}_d S_i, \tilde{R}_d S_i]$, $1 \le i \le n$. It can be shown with straightforward algebra, using (3) and (4), that $A_i^T \omega^{\perp}(D_i) A_i = S_i^T C_d S_i$ for all $1 \le i \le n$. On the other hand

$$\sum A_i^T A_i - \left(\sum A_i^T D_i\right)\left(\sum D_i^T D_i\right)^- \left(\sum D_i^T A_i\right)$$

$$= \sum S_i^T \tilde{T}_d^T \tilde{T}_d S_i -$$

$$\left[\sum S_i^T \tilde{T}_d^T \tilde{L}_d S_i, \quad \sum S_i^T \tilde{T}_d^T \tilde{R}_d S_i\right] \begin{bmatrix} \sum S_i^T \tilde{L}_d^T \tilde{L}_d S_i & \sum S_i^T \tilde{L}_d^T \tilde{R}_d S_i \\ \sum S_i^T \tilde{R}_d^T \tilde{L}_d S_i & \sum S_i^T \tilde{R}_d^T \tilde{R}_d S_i \end{bmatrix}^- \begin{bmatrix} \sum S_i^T \tilde{L}_d^T \tilde{T}_d S_i \\ \sum S_i^T \tilde{R}_d^T \tilde{T}_d S_i \end{bmatrix}$$

$$= \sum S_i^T C_{d11} S_i -$$

$$\left[\sum S_i^T C_{d12} S_i, \quad \sum S_i^T C_{d13} S_i\right] \begin{bmatrix} \sum S_i^T C_{d22} S_i & \sum S_i^T C_{d23} S_i \\ \sum S_i^T C_{d23}^T S_i & \sum S_i^T C_{d33} S_i \end{bmatrix}^- \begin{bmatrix} \sum S_i^T C_{d12}^T S_i \\ \sum S_i^T C_{d13}^T S_i \end{bmatrix}$$

Since the summations are over all permutations of the numbers $\{1, ..., t\}$, $\sum S_i^T C_{drs} S_i$ is completely symmetric for all $1 \le r \le s \le 3$. As C_{d11}, C_{d12}, and C_{d13} have column sums zero, we conclude that $\sum S_i^T C_{drs} S_i = \{c_{drs} n/(t-1)\}B_t + z_{rs} 1_t 1_t^T$, for some z_{rs}, with $z_{rs} = 0$ if $r = 1$.

To proceed, we need a g-inverse of

$$F = \begin{bmatrix} \{c_{d22} n/(t-1)\} B_t + z_{22} 1_t 1_t^T & \{c_{d23} n/(t-1)\} B_t + z_{23} 1_t 1_t^T \\ \{c_{d23} n/(t-1)\} B_t + z_{23} 1_t 1_t^T & \{c_{d33} n/(t-1)\} B_t + z_{33} 1_t 1_t^T \end{bmatrix} = \frac{t-1}{n} (Q \otimes B_t - \begin{bmatrix} z_{22} & z_{23} \\ z_{23} & z_{33} \end{bmatrix} \otimes 1_t 1_t^T).$$

One such g-inverse, for appropriate w_{ij}, is

$$F^- = \frac{n}{t-1} (Q^- \otimes B_t - \begin{bmatrix} w_{22} & w_{23} \\ w_{23} & w_{33} \end{bmatrix} \otimes 1_t 1_t^T).$$

Therefore

$$\sum A_i^T A_i - \left(\sum A_i^T D_i\right)\left(\sum D_i^T D_i\right)^- \left(\sum D_i^T A_i\right)$$

$$= \frac{n}{t-1} c_{d11} B_t - \frac{n^2}{(t-1)^2} ([c_{d12}, \quad c_{d13}] \otimes B_t) F^- \left(\begin{bmatrix} c_{d12} \\ c_{d13} \end{bmatrix} \otimes B_t \right) = \frac{n}{t-1} q*_d \, B_t.$$

Then Proposition 1 implies that $\operatorname{tr} C_d \le q*_d$.

Finally note that for design f we have $C_{frs} = \sum S_i^T C_{frs} S_i / n$ for every $1 \le r \le s \le 3$. ❐

3. Methods for determination of a maximal $q*_d$.

An optimal design $d*$ should have a completely symmetric C_{d*}, with tr $C_{d*} = q*_{d*}$, and it should have the right proportions of blocks assigned to the treatment sequences such that $q*_{d*}$ is maximal. Therefore, we need to maximize the bound $q*_d$. Define

$c_{d11}^{(u)} = \operatorname{tr}(T_{du}^T B_k T_{du})$, $c_{d12}^{(u)} = \operatorname{tr}(T_{du}^T B_k L_{du})$, $c_{d13}^{(u)} = \operatorname{tr}(T_{du}^T B_k R_{du})$, $c_{d22}^{(u)} = \operatorname{tr}(B_t L_{du}^T B_k L_{du} B_t)$,

$c_{d23}^{(u)} = \operatorname{tr}(B_t L_{du}^T B_k R_{du} B_t)$, and $c_{d33}^{(u)} = \operatorname{tr}(B_t R_{du}^T B_k R_{du} B_t)$.

We then get that

$$c_{drs} = \sum_{u=1}^{b} c_{drs}^{(u)}\,,\ 1 \le r \le s \le 3.$$

Note that each $c_{drs}^{(u)}$ remains unchanged if the treatments are relabelled, i.e. if T_{du}, L_{du} and R_{du} are replaced by $T_{du}S$, $L_{du}S$ and $R_{du}S$, respectively, where S is any $t \times t$ permutation matrix. We call two sequences of treatments equivalent if one can be transformed to the other by relabelling the treatments. Hence, two equivalent treatment sequences give the same $c_{drs}^{(u)}$. Therefore, for given t and k, we can divide the set of all possible treatment sequences into K equivalence classes $s_1, \ldots, s_K$. If, for example, $k = 3$ and $t \ge 3$, then there are the $K = 5$ equivalence classes given in Table 1.

Since $c_{drs}^{(u)}$ is the same for each u receiving a treatment sequence in a given equivalence class s_ℓ, $1 \le \ell \le K$, we can define $c_{rs}(\ell) = c_{drs}^{(u)}$ and get $c_{drs} = b\sum_{\ell=1}^{K} \pi_{d\ell} c_{rs}(\ell)$, where $\pi_{d\ell}$ is the proportion of blocks assigned to the class s_ℓ. This, however, implies that the bound q^*_d of any design $d \in \Omega_{t,b,k}$ is determined by the proportions $\pi_{d\ell}$. Note that the c_{dij} are linear in the $\pi_{d\ell}$, but that q^*_d is a quotient, where the $\pi_{d\ell}$ are third order in the numerator and second order in the denominator. This makes the maximization of q^*_d difficult.

The situation is similar to the models (with carryover effects) for repeated measurements designs. For these Kushner (1997) showed how to use the linearity of the c_{drs} to maximize q^*_d. This idea can be generalized to interference models.

Proposition 3:

<u>For any design $d \in \Omega_{t,b,k}$ define the function</u> $q_d : IR^2 \to IR$ as

$$q_d(x, y) = c_{d11} + 2c_{d12}x + 2c_{d13}y + 2c_{d23}xy + c_{d22}x^2 + c_{d33}y^2 .$$

Then for every x and y, we have $q_d(x, y) \geq q^*_d$. There is at least one point (x_d, y_d) such that $q_d(x_d, y_d) = q^*_d$.

Proof:

We can write

$$q_d(x,y) = c_{d11} + 2[c_{d12} \quad c_{d13}]\begin{bmatrix} x \\ y \end{bmatrix} + [x \quad y]Q\begin{bmatrix} x \\ y \end{bmatrix}$$

$$= c_{d11} + 2[c_{d12} \quad c_{d13}](u - Q^{-}\begin{bmatrix} c_{d12} \\ c_{d12} \end{bmatrix}) - (u^T - [c_{d12} \quad c_{d13}]Q^{-})Q(u - Q^{-}\begin{bmatrix} c_{d12} \\ c_{d12} \end{bmatrix})$$

where

$$u = \begin{bmatrix} x \\ y \end{bmatrix} + Q^{-}\begin{bmatrix} c_{d12} \\ c_{d12} \end{bmatrix}.$$

Then equation (4) implies that

$$q_d(x,y) = c_{d11} - [c_{d12} \quad c_{d13}]Q^{-}\begin{bmatrix} c_{d12} \\ c_{d13} \end{bmatrix} + u^T Qu \,.$$

Therefore, $q_d(x, y)$ is minimal iff $Q\,u = 0$, i.e. iff

$$Q\begin{bmatrix} x \\ y \end{bmatrix} = -\begin{bmatrix} c_{d12} \\ c_{d13} \end{bmatrix}.$$

This, however, holds if and only if the partial derivatives of q_d with respect to x and y are both 0. The minimum of q_d equals q^*_d. ❒

From the proof of Proposition 3, we immediately get

Corollary 1:

Consider a point (x_d, y_d) such that the partial derivatives $\partial q_d(x, y) / \partial x$ and $\partial q_d(x, y) / \partial y$ are both 0 for $(x, y) = (x_d, y_d)$. Then $q_d(x_d, y_d) = q^*_d$.

The elegance of $q_d(x, y)$ is that it can be written as a linear combination of functions $h_\ell(x, y)$, which depend on the equivalence classes of treatment sequences. Define

$$h_\ell(x,y) = c_{11}(\ell) + 2c_{12}(\ell)x + 2c_{13}(\ell)y + 2c_{23}(\ell)xy + c_{22}(\ell)x^2 + c_{33}(\ell)y^2 ,$$

for every $1 \le \ell \le K$. Then

$$q_d(x,y) = b\sum_{\ell=1}^{K} \pi_{d\ell} h_\ell(x,y) .$$

Proposition 4:

For a design $d \in \Omega_{t,b,k}$ consider a point (x_d, y_d) for which $q_d(x_d, y_d) = q^*_d$.

If $bh_\ell(x_d, y_d) \le q_d(x_d, y_d) = q^*_d$ for every $1 \le \ell \le K$, then for every $f \in \Omega_{t,b,k}$ we have

tr $C_f \le q^*_d = a^*_{t,b,k}$, say.

Proof:

For any f we have

$$q^*_f = b\sum_{\ell=1}^{K} \pi_{f\ell} h_\ell(x_f, y_f) \le \sum_{\ell=1}^{K} \pi_{f\ell} b h_\ell(x_d, y_d) \le \sum_{\ell=1}^{K} \pi_{f\ell} \; q^*_d = q^*_d.$$

The rest follows from Prop. 2. ❒

Note that the proportions $\pi_{d\ell}$ must be such that the partial derivatives of $\sum \pi_{d\ell} h_\ell(x,y)$ at (x_d, y_d) are both 0, and that only such classes ℓ of sequences are included for which $h_\ell(x_d, y_d) = \max_{1\le\ell\le K} h_\ell(x_d, y_d)$. Therefore (x_d, y_d) must be either at the minimum of an h_ℓ or at the intersection of two or more of the h_ℓ.

In many situations there is no design fulfilling both the conditions of Proposition 4 and of Proposition 2. In that case, however, one practical use of the $a^*_{t,b,k}$ is the lower bound which it provides for the optimality criteria.

As an example, consider the A-criterion $\varphi_A(C_f)$, which is the trace of the Moore-Penrose generalized inverse of C_f. From Prop. 2 we get

$$\varphi_A(C_f) \geq \varphi_A\left(\frac{q^*_f}{t-1} B_t\right) = \mathrm{tr}\left(\frac{t-1}{q^*_f} B_t\right) = \frac{(t-1)^2}{q^*_f}.$$

With Prop. 4 it follows that

$$\varphi_A(C_f) \geq \frac{(t-1)^2}{a^*_{t,b,k}}.$$

4. Some examples

In this section we demonstrate the methods derived in this paper by finding optimal or efficient designs for $k = 3$ and 4 for all $t \geq 2$. Note that, to save space, blocks are represented as columns in Examples 1 to 4.

4.1 The case of 3 plots per block

Table 1 lists the equivalence classes and the corresponding $c_{rs}(\ell)$ for the case that there are $k = 3$ plots per block. If t=2, then only the first four sequences are possible.

A design d^* which has half of its sequences from s_2 and half of its sequences from s_4 has

$$q_{d^*}(x,y) = b\,\{\tfrac{1}{2}h_2(x,y) + \tfrac{1}{2}h_4(x,y)\} = b\left(\frac{4}{3} - \frac{1}{3}x - \frac{1}{3}y - \frac{4t-2}{3t}xy + \frac{3t-2}{3t}x^2 + \frac{3t-2}{3t}y^2\right).$$

If $x = y = t/\{2(t\text{-}1)\}$, then the derivatives of $q_{d^*}(x, y)$ with respect to x and y are both 0. Therefore from Corollary 1, we have

$$x_{d^*} = y_{d^*} = \frac{t}{2(t-1)} \text{ and } q^*_{d^*} = q_{d^*}(x_{d^*}, y_{d^*}) = \left(\frac{7t-8}{6(t-1)}\right)b\,.$$

Table 1

The classes s_ℓ of sequences and adjusted $c_{rs}(\ell)$ for $k = 3$, $t \geq 2$

ℓ	Repre-sentative sequence	$3\,c_{11}(\ell)$	$3\,c_{12}(\ell)$	$3\,c_{13}(\ell)$	$3\,c_{22}(\ell)+\frac{2}{t}$	$3\,c_{23}(\ell)-\frac{1}{t}$	$3\,c_{33}(\ell)+\frac{2}{t}$
1	[1 1 1]	0	0	0	2	–1	2
2	[1 1 2]	4	–1	0	2	–2	4
3	[1 2 1]	4	–3	–3	4	1	4
4	[1 2 2]	4	0	–1	4	–2	2
5	[1 2 3]	6	–2	–2	4	–1	4

To prove the optimality of $q^*_{d^*}$ we have to calculate $h_\ell(x_{d^*}, y_{d^*})$ for every $1 \leq \ell \leq 5$, and to verify that $q^*_{d^*} / b$ - $h_\ell\,(x_{d^*}, y_{d^*})$ is nonnegative for every ℓ. Some algebra shows that $q^*_{d^*} / b$ - $h_\ell(x_{d^*}, y_{d^*})$ equals $(3t–4) / (4t–4) > 0$, 0, $(3t^2–5t) / (3t^2–6t+3) > 0$, 0, $(t–2) / (3t^2–6t+3) > 0$ (since $t > 2$) for $\ell = 1, ..., 5$, respectively.

Hence, we have shown

Theorem 1:

If $k = 3$ and $t \geq 2$, then for any design $d \in \Omega_{t,b,3}$ we have

$$\operatorname{tr} C_d \leq a^*_{t,b,3} = \left(\frac{7t-8}{6(t-1)}\right) b .$$

If a design d^* has half of its blocks with treatment sequences which are equivalent to [1 1 2] and half of its blocks with treatment sequences equivalent to [1 2 2], and if C_{d^*11}, C_{d^*12}, C_{d^*13}, C_{d^*22}, C_{d^*23}, and C_{d^*33} are completely symmetric, then d^* is universally optimal over $\Omega_{t,b,3}$.

Example 1:

If $t = 2$, then a 4 block example of a design fulfilling the conditions of Theorem 1 is

$$d^*_1 = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 2 & 2 & 1 \\ 2 & 1 & 2 & 1 \end{bmatrix} \in \Omega_{2,4,3} .$$

If $t = 3$ then a 12 block example of a design fulfilling the conditions of Theorem 1 is

$$d^*_2 = \begin{bmatrix} 1 & 1 & 2 & 2 & 3 & 3 & 1 & 1 & 2 & 2 & 3 & 3 \\ 1 & 1 & 2 & 2 & 3 & 3 & 2 & 3 & 1 & 3 & 1 & 2 \\ 2 & 3 & 1 & 3 & 1 & 2 & 2 & 3 & 1 & 3 & 1 & 2 \end{bmatrix} \in \Omega_{3,12,3} .$$

If $t = 4$, then a 24 block example is

$$d^*_3 = \begin{bmatrix} 111222333444111222333444 \\ 111222333444234134124123 \\ 234134124123234134124123 \end{bmatrix} \in \Omega_{4,24,3} .$$

4.2 The case of 4 plots per block

If k = 4 and $t \geq 4$, then we have 15 equivalence classes. The representative sequences and the $c_{rs}(\ell)$ for the 15 classes are given in Table 2. For t = 3, only the 14 classes s_1 to s_{14} are possible. For t = 2, only the 8 classes s_1, s_2, s_3, s_4, s_6, s_7, s_9 and s_{10} are possible.

We start with the case t = 2. Then consider a design d^* with half of its blocks from s_4 and half of its blocks from s_9. In that case

$$q_{d^*}(x,y) = b\,\{\tfrac{1}{2}h_4(x,y) + \tfrac{1}{2}h_9(x,y)\} \ = b\left(2 + \frac{11}{8}x^2 - 2xy + \frac{11}{8}y^2\right) \geq 2b\,,$$

with equality holding iff $x = y = 0$. Now, $h_\ell(0, 0) = c_{11}(\ell) \leq 2$ for all 8 possible classes s_ℓ of sequences, with equality for ℓ = 4, 7 and 9. Thus we have shown

Theorem 2:

<u>If</u> t = 2 <u>and</u> k = 4<u>, then for every design</u> $d \in \Omega_{2,b,4}$ <u>we have</u> tr $C_d \leq a^*_{2,b,4} = 2\,b$.

<u>If a design</u> d^* <u>has</u> $b/4$ <u>of its blocks with each of the sequences</u> [1 1 2 2], [2 2 1 1], [1 2 2 1] <u>and</u> [2 1 1 2]<u>, then</u> d^* <u>is universally optimal over</u> $\Omega_{2,b,4}$.

Note that sequences [1 1 2 2] and [2 2 1 1] are from s_4, while [1 2 2 1] and [2 1 1 2] are from s_9. Because $h_7(0, 0) = 2$, it is possible to show that there is another design f that has $q^*_f = a^*_{2,b,k}$. Design f has $3b/4$ of its blocks with sequences from s_4 and $b/4$ with sequences from s_7.

Example 2:

Theorem 2 requires that b is divisible by 4. Suppose b = 2 and consider the two designs

$$d = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 2 \\ 2 & 1 \end{bmatrix} \text{ and } f = \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 2 \\ 2 & 1 \end{bmatrix}.$$

While $q^*_d = a^*_{2,2,4} = 4$, for d we have tr $C_d = 16/7 < 4$, because C_{d33} is not completely symmetric. Design f, for which the C_{dij}, except for C_{d23}, are completely symmetric, has tr $C_d = 3$. Calculating the information matrix for all 256 possible designs, we find that f is universally optimal (since rank $C_d = 1$).

Table 2

The classes s_ℓ of sequences and adjusted $c_{rs}(\ell)$ for $k = 4$

ℓ	Representative sequence	$4\,c_{11}(\ell)$	$4\,c_{12}(\ell)$	$4\,c_{13}(\ell)$	$4\,c_{22}(\ell)+\frac{3}{t}$	$4\,c_{23}(\ell)-\frac{1}{t}$	$4\,c_{33}(\ell)+\frac{3}{t}$
1	[1 1 1 1]	0	0	0	3	–1	3
2	[1 1 1 2]	6	–1	1	3	–2	7
3	[1 1 2 1]	6	–3	–3	7	–1	7
4	[1 1 2 2]	8	2	2	7	–4	7
5	[1 1 2 3]	10	–1	0	7	–3	9
6	[1 2 1 1]	6	–3	–3	7	–1	7
7	[1 2 1 2]	8	–6	–6	7	4	7
8	[1 2 1 3]	10	–5	–4	7	1	9
9	[1 2 2 1]	8	–2	–2	7	–5	7
10	[1 2 2 2]	6	1	–1	7	–2	3
11	[1 2 2 3]	10	–1	–1	7	–4	7
12	[1 2 3 1]	10	–4	–4	9	–3	9
13	[1 2 3 2]	10	–4	–5	9	1	7
14	[1 2 3 3]	10	0	–1	9	–3	7
15	[1 2 3 4]	12	–3	–3	9	–2	9

As $h_{14}(0, 0) = c_{11}(14) = 10/4 > 2$, an optimal design for $t = 3$ must have other sequences than just s_4, s_7 and s_9. The case $k = 3$ suggests the candidate design d^* with $\pi_{d^*5} = \pi_{d^*14} = ½$. In fact, we find that $q_{d^*}(x,y) = b(\frac{5}{2} - \frac{1}{4}x - \frac{1}{4}y - \frac{4}{3}xy + \frac{7}{4}x^2 + \frac{7}{4}y^2)$, with a minimum at $x_{d^*} = y_{d^*} = 3/26$. Therefore $q^*_{d^*} = q_{d^*}(x_{d^*}, y_{d^*}) = (^{257}/_{104})\, b$. It is easy to check that for every ℓ, $1 \le \ell \le 14$, we have $257/104 - h_\ell(x_{d^*}, y_{d^*}) \ge 0$, with equality holding only for $\ell = 5$ and $\ell = 14$. Hence, we have shown

Theorem 3:

If $k = 4$ and $t = 3$, then for any design $d \in \Omega_{3,b,4}$ we have tr $C_d \leq a^*_{3,b,4} = (^{257}/_{104})\, b$.

If a design d^* has b/2 blocks with treatment sequences which are equivalent to each of [1 1 2 3] and [1 2 3 3], and if C_{d^*11}, C_{d^*12}, C_{d^*13}, C_{d^*22}, C_{d^*23}, and C_{d^*33} are completely symmetric, then d^* is universally optimal over $\Omega_{3,b,4}$.

Example 3:

The design

$$d^* = \begin{bmatrix} 1 & 2 & 3 & 3 & 1 & 2 & 1 & 2 & 3 & 2 & 3 & 1 \\ 1 & 2 & 3 & 2 & 3 & 1 & 1 & 2 & 3 & 3 & 1 & 2 \\ 2 & 3 & 1 & 1 & 2 & 3 & 3 & 1 & 2 & 1 & 2 & 3 \\ 3 & 1 & 2 & 1 & 2 & 3 & 2 & 3 & 1 & 1 & 2 & 3 \end{bmatrix} \in \Omega_{3,12,4}$$

fulfils the conditions of Theorem 3. The first 6 blocks of d^* form a design d which maximizes q^*_d in $\Omega_{3,6,4}$ but for which tr $C_d < q^*_d$, since C_{d12} and C_{d13} are not completely symmetric. However, when we calculate the A-criterion $\varphi_A(C_d)$ of d and compare it to the unattainable lower bound $\varphi^*_A = (t-1)^2 / a^*_{3,b,4}$, then we find that $\varphi^*_A / \varphi_A(C_d) = 0.996$, i.e. d has an efficiency of 99.6% and is likely to be A-optimal.

Finally, we consider the case $k = 4$ and $t \geq 4$. We try a design with a proportion π of sequences from the class s_{15} and proportions $(1-\pi)/2$ of classes s_5 and s_{14}, each. The three $h_\ell(x,y)$ intersect at $x = y = (5-\sqrt{17})/4 = x^*$, say. For $x = y = x^*$, we have

$$h_5(x, y) = h_{14}(x, y) = h_{15}(x, y) = \frac{(135-23\sqrt{17})t-(42-10\sqrt{17})}{16t}.$$

Note that $h_{15}(x, y) = h_{15}(y, x)$. Thus the derivative of $h_{15}(x+\delta, x-\delta)$ with respect to δ is zero if $\delta = 0$. The same holds for ½ $h_5(x, y)$ + ½ $h_{14}(x, y)$. It hence remains to find a π such that

$$\frac{\partial}{\partial x}\left(\pi\, h_{15}(x,x)+\frac{1-\pi}{2}h_5(x,x)+\frac{1-\pi}{2}h_{14}(x,x)\right)$$

is zero for $x = x^* \approx 0.219$. Therefore, set

$$\pi=\frac{(23-5\sqrt{17})t-(10-2\sqrt{17})}{2\sqrt{17}\,t}=\pi^*, \text{ say.}$$

It is easy to see that the differences $h_5(x^*, x^*) - h_\ell(x^*, x^*)$, for $\ell = 1,2, ..., 15$ are all positive, except for $\ell = 4, 5, 14$ and 15, when they are 0.

Hence we have an optimal design using sequence classes s_5, s_{14} and s_{15}. Since $h_5(x^*,x^*) - h_4(x^*,x^*) = 0$, we can construct an optimal design with some sequences from the class s_4. In fact, a second optimal design exists which consists of s_4 and s_{15} only having a proportion

$$\delta^*=\frac{(23-3\sqrt{17})t-(10-2\sqrt{17})}{4\sqrt{17}\,t}$$

of sequences from the class s_{15} and a proportion of $1 - \delta^*$ of sequences from s_4. Any convex combination of these two designs is also optimal. Hence, we have shown

Theorem 4

If $k = 4$ and $t \geq 4$, then for every design $d \in \Omega_{t,b,4}$ we have

$$\text{tr } C_d \leq a^*_{t,b,4} = b\ \frac{(135-23\sqrt{17})t-(42-10\sqrt{17})}{16t}.$$

To achieve this bound, we would need to construct a design d^* as follows:

Define

$$\pi^* = \frac{(23-5\sqrt{17})t-(10-2\sqrt{17})}{2\sqrt{17}\,t} \ \text{ and } \ \delta^* = \frac{(23-3\sqrt{17})t-(10-2\sqrt{17})}{4\sqrt{17}\,t}.$$

Choose $0 \leq \alpha \leq 1$. Let proportions $(1-\alpha)(1-\delta^*)$, $\alpha(1-\pi^*)/2$, $\alpha(1-\pi^*)/2$ and $\{\alpha\pi^* + (1-\alpha)\delta^*\}$ of the blocks of d^* have treatment sequences which are equivalent to [1 1 2 2], [1 1 2 3], [1 2 3 3] and [1 2 3 4], respectively, such that C_{d^*11}, C_{d^*12}, C_{d^*13}, C_{d^*22}, C_{d^*23}, and C_{d^*33} are completely symmetric.

Remark: The design d^* in Theorem 3 cannot exist for finite b. To see this, note that $1 - \delta^* = (1 - \pi^*) / 2$, which is irrational. Therefore, $(1 - \alpha)(1 - \delta^*) = (1 - \alpha)(1 - \pi^*) / 2$ and there is no α such that both $(1 - \alpha)(1 - \pi^*) / 2$ and $\alpha(1 - \pi^*) / 2$ are rational.

Despite the non-existence of d^*, Theorem 4 has two useful aspects. Firstly, it suggests the structure of an efficient design, and secondly $a^*_{v,b,4}$ gives a lower bound for the A-value. This is demonstrated in Example 4.

Example 4

It is possible to construct highly efficient designs if we can approximate reasonably well the fractions π^* or δ^* from Theorem 4. If $t = 4$, then the upper bound $a^*_{4,b,4}$ for tr C_d is approximately $b\times 2.49852$. To construct an efficient design, we select $\alpha = 0$. We would need a proportion of $\delta^* \approx 0.617995$ of blocks with a sequence from s_{15}. We use 2/3 instead and construct the 36 block design

$$f = \begin{bmatrix} 111222333444123412431432123412431432 \\ 111222333444241323144213241323144213 \\ 234134124123314241323124314241323124 \\ 234134124123432134212341432134212341 \end{bmatrix} \in \Omega_{4,36,4}\,.$$

It is easy to verify that C_{f11}, ..., C_{f33} are completely symmetric, and that tr $C_f \approx 89.8064...$. This is extremely close to the upper bound which is approximately $36\times 2.49852 = 89.94672$, so that f is highly efficient (efficiency ≈ 0.9984).

With 12 blocks, a design similarly constituted to f is

$$g_1 = \begin{bmatrix} 1 & 2 & 3 & 4 & 1 & 2 & 3 & 4 & 1 & 2 & 4 & 3 \\ 1 & 2 & 3 & 4 & 2 & 4 & 1 & 3 & 2 & 3 & 1 & 4 \\ 2 & 3 & 4 & 1 & 3 & 1 & 4 & 2 & 4 & 1 & 3 & 2 \\ 2 & 3 & 4 & 1 & 4 & 3 & 2 & 1 & 3 & 4 & 2 & 1 \end{bmatrix} \in \Omega_{4,12,4.}$$

Its C_g is not completely symmetric. However, its relative A-efficiency with respect to the bound $(t-1)^2/a^*_{4,12,4}$ is 0.968.

If we prefer not to repeat treatments, we have the universally optimal binary design using a type I orthogonal array with efficiency 0.924:

$$h = \begin{bmatrix} 1 & 2 & 3 & 4 & 1 & 2 & 3 & 4 & 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 & 2 & 4 & 1 & 3 & 2 & 3 & 4 & 1 \\ 2 & 4 & 1 & 3 & 3 & 1 & 4 & 2 & 4 & 1 & 2 & 3 \\ 4 & 3 & 2 & 1 & 4 & 3 & 2 & 1 & 3 & 4 & 1 & 2 \end{bmatrix}.$$

With 6 blocks, a design with relative A-efficiency 0.885, similarly constituted to f is

$$g_2 = \begin{bmatrix} 1 & 3 & 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 & 1 & 3 \\ 2 & 4 & 3 & 1 & 4 & 2 \\ 2 & 4 & 4 & 3 & 2 & 1 \end{bmatrix} \in \Omega_{4,6,4}.$$

For 8 treatments, and 24 blocks similar ideas lead to the design $g_3 \in \Omega_{8,24,4}$, where

$$g_3 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 1 & 1 & 3 & 6 & 2 & 2 & 6 & 5 & 4 & 7 & 8 & 8 & 3 & 5 & 4 & 7 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 2 & 4 & 1 & 7 & 5 & 8 & 3 & 6 & 3 & 2 & 6 & 1 & 8 & 7 & 5 & 4 \\ 5 & 6 & 7 & 8 & 1 & 2 & 3 & 4 & 3 & 2 & 6 & 1 & 8 & 7 & 5 & 4 & 2 & 4 & 1 & 7 & 5 & 8 & 3 & 6 \\ 5 & 6 & 7 & 8 & 1 & 2 & 3 & 4 & 4 & 7 & 8 & 8 & 3 & 5 & 4 & 7 & 1 & 1 & 3 & 6 & 2 & 2 & 6 & 5 \end{bmatrix}.$$

This design has a relative A-efficiency of 0.910. Note that each treatment is replicated 12 times in the design g_3, as in g_1.

The methods of the present paper can be used for blocks with $k > 4$ as well. However, with larger k the number K of equivalence classes increases rapidly. For $k = 5$ and $t = 5$ there are 52 classes of sequences. It is possible, though, to show that a design d with a proportion

$$\pi^* = \frac{3}{4\sqrt{41}\,t}\left(25t - 3\sqrt{41}\,t - 7 + \sqrt{41}\right)$$

of blocks with a sequence equivalent to [1 2 3 4 5] and the other blocks with a sequence equivalent to [1 1 2 3 3] has a maximal $q^*_d = a^*_{t,b,5}$. In the special instance $t = 5$ (with $\pi^* \approx 0.6643$), then a binary type I orthogonal array h, which uses only sequences equivalent to [1 2 3 4 5], has an efficiency of $q^*_h \,/\, a^*_{5,b,5} = 0.959$. This is slightly higher than for $k = 4$.

Further work is aimed at obtaining bounds on the $c_{ij}(\ell)$ to get results for a general k. We conjecture that $a^*_{t,b,k}$ is achieved by a design with a majority of sequences from the class containing [1 2 3 ... k–2 k–1 k] and the rest of the sequences equivalent to [1 1 2 ... k–3 k–2 k–2]. We also conjecture that for $t \geq k > 5$, a binary type I orthogonal array will have an efficiency of more than 0.95.

Acknowledgement

The authors wish to thank H.B. Kushner for prior notice of his work and the referees for their helpful comments. Financial support from the Volkswagen Foundation and the Deutsche Forschungsgemeinschaft (SFB 475, "Reduction of complexity in multivariate data structures") is gratefully acknowledged.

References:

David, O. and Kempton, R.A. (1996). Designs for interference. *Biometrics* **52**, 224 – 233.

David, O., Monod, H. and Philippeau, G. (1998): Protocoles expérimentaux pour contrôler les effets de voisinage dans les essais variétaux. In: *Actes des XXXèmes Journées de Statistique, Volume II: Communications libres*, (ENSAI, Rennes, 25 - 29 mai 1998), 162 - 164.

Druilhet, P. (1999): Optimality of circular neighbour balanced designs. *Journal of Statistical Planning and Inference* **81**, 141 - 152.

Gill, P.S. (1993): Design and analysis of field experiments incorporating local and remote effects of treatments. *Biometrical Journal* **35**, 343 – 354.

Kiefer, J. (1975): Construction and optimality of generalized Youden designs. In '*A survey of statistical design and linear models*', Ed. J. N. Srivastava, pp. 333 - 353. North Holland: Amsterdam.

Kunert, J. (1983): Optimal design and refinement of the linear model with applications to repeated measurements designs. *Annals of Statistics* **11**, 247 - 257.

Kushner, H. B. (1997): Optimal repeated measurements designs: the linear optimality equations. *Annals of Statistics* **25**, 2328 - 2344.

Pukelsheim, F. (1993): *Optimal Design of Experiments*. Wiley, New York.

Rao, C. R. and Toutenburg, H. (1995): *Linear models. Least squares and alternatives*. Springer, New York.