SFB 823

Optimal crossover designs in a model with self and mixed carryover effects with correlated errors

Adrian Wilk, Joachim Kunert

Nr. 51/2013

Discussion Paper

Optimal crossover designs in a model with self and mixed carryover effects with correlated errors.

Adrian Wilk and Joachim Kunert
Fakultät Statistik, Technische Universität Dortmund, 44221 Dortmund, Germany

Abstract

We determine optimal crossover designs for the estimation of direct treatment effects in a model with mixed and self carryover effects. The model also assumes that the errors within each experimental unit are correlated following a stationary first-order autoregressive process. The paper considers situations where the number of periods for each experimental unit is at least four and the number of treatments is greater or equal to the number of periods.

Key words: Carryover effects, Optimal design, Correlated errors

1 Introduction

In crossover designs experimental units are exposed to a number of treatments, one after the other. We consider the situation that treatments are liable to have a carryover effect on the measurement in the next period. [Afsarinejad and Hedayat(2002)] suggested a model with partial interaction between direct effects and carryover effects. In this model, each treatment has two types of carryover effects. If the treatment in the present period is the same as in the preceding period, the measurement is influenced by the self carryover effect, if the treatment is different, the mixed carryover effect appears. These authors considered designs with two periods for each experimental unit and determined optimal designs for direct treatment effects. [Kunert and Stufken(2002)] and [Kunert and Stufken(2008)] considered designs with more than two periods. In all three papers, the model assumed that the errors are independent. An extension of the model, assuming with correlated errors, was considered by [Hedayat and Yan(2008)]. They investigated designs where the number p of periods equals three and the number t of treatments is at least three. They also gave a numerical investigation of the case $t \geq p = 4$.

In what follows, we consider the same model. We extend the work by [Hedayat and Yan(2008)] to the case $t \geq p \geq 4$ and give a closed solution for that case. The proof uses the approach of [Kushner(1997)], see also [Kunert and Martin(2000)] and [Kunert and Stufken(2002)].

2 Notations and a tool for finding good designs

The model which we consider is exactly the model studied in [Hedayat and Yan(2008)]. The response on period j of unit i can be written as

$$y_{ij} = \begin{cases} \alpha_i + \beta_j + \tau_{d(i,j)} + \rho_{d(i,j-1)} + \varepsilon_{ij}, & \text{if } d(i,j) \neq d(i,j-1), \\ \alpha_i + \beta_j + \tau_{d(i,j)} + \gamma_{d(i,j-1)} + \varepsilon_{ij}, & \text{if } d(i,j) = d(i,j-1), \end{cases} \tag{1}$$

where $d(i,j) \in \{1,...,t\}$ is the treatment applied to unit i in period j. The unknown fixed parameters α_i, β_j, $\tau_{d(i,j)}$, $\rho_{d(i,j-1)}$ and $\gamma_{d(i,j-1)}$ are the effect of unit i, the effect of period j, the direct effect of treatment $d(i,j)$, the mixed carryover effect of treatment $d(i,j-1)$ and the self carryover effect of treatment $d(i,j-1)$, respectively. Since the first observation of any unit is not affected by a carryover effect, we define $\rho_{d(i,0)} = 0$ and $\gamma_{d(i,0)} = 0$. Further, we assume that the errors within any unit are correlated following a stationary first-order autoregressive process with known correlation parameter λ. Errors from different units are independent from each other. Thus $Cov(\varepsilon_{ij_1}, \varepsilon_{ij_2}) = \sigma^2\lambda^{|j_1-j_2|}/(1-\lambda^2)$ with variance $\sigma^2 > 0$ and correlation parameter $\lambda \in (-1,1)$ and $Cov(\varepsilon_{i_1j_1}, \varepsilon_{i_2j_2}) = 0$ for $i_1 \neq i_2$.

The class of all designs with t treatments, n experimental units and p periods for each unit is denoted by $\Omega_{t,n,p}$. We restrict attention to the case $t \geq p \geq 4$.

Using matrix notation the model becomes

$$Y = U\alpha + P\beta + T_d\tau + M_d\rho + S_d\gamma + \varepsilon \tag{2}$$

where $Y = [y_{11}, y_{12}, \ldots, y_{np}]^T$. The covariance matrix of ε is given by $Cov(\varepsilon)$

$= \sigma^2(I_n \otimes \Lambda)$, where $\otimes$ denotes the Kronecker product and

$$\Lambda = \frac{1}{1-\lambda^2} \begin{pmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^{p-1} \\ \lambda & 1 & \lambda & \dots & \lambda^{p-2} \\ \lambda^2 & \lambda & 1 & \dots & \lambda^{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda^{p-1} & \lambda^{p-2} & \lambda^{p-3} & \dots & 1 \end{pmatrix}.$$

The matrices $U = I_n \otimes 1_p$, $P = 1_n \otimes I_p$, T_d, M_d and S_d are the design matrices of the unit, period, direct treatment, mixed carryover and self carryover effects, respectively.

We assume that the parameter $\lambda \in (-1, 1)$ is known. As in [Kunert(1985)], we can determine a matrix V_λ such that $V_\lambda \Lambda_\lambda V_\lambda^T = I_p$ and transform model (2) to

$$\begin{aligned} (I_n \otimes V_\lambda)Y &= (I_n \otimes V_\lambda)U\alpha + (I_n \otimes V_\lambda)P\beta + (I_n \otimes V_\lambda)T_d\tau \\ &\quad +(I_n \otimes V_\lambda)M_d\rho + (I_n \otimes V_\lambda)S_d\gamma + \tilde{\varepsilon}, \end{aligned}$$

where $Cov(\tilde{\varepsilon}) = \sigma^2 I_{np}$. Then the information matrix for the estimation of direct effects becomes

$$C_d = T_d^T(I_n \otimes V_\lambda)\omega^\perp((I_n \otimes V_\lambda)[P, U, M_d, S_d])(I_n \otimes V_\lambda)T_d,$$

where $\omega^\perp(A) = I - A(A^TA)^- A^T$ for a matrix A and $(A^TA)^-$ is the generalized inverse of A^TA.

As in [Kunert and Stufken(2002)] we can see that the information matrix C_d of any design d must have row and column sums zero. It was shown by [Kiefer(1975)] that then a design is universally optimal if the corresponding information matrix is completely symmetric and has maximal trace among the information matrices of all designs in $\Omega_{t,n,p}$. Complete symmetry means that the matrix can be written as $aI_t + b1_t1_t^T$ where a and b are real numbers.

An upper bound of C_d in the Loewner ordering is

$$C_d \leq T_d^T(I_n \otimes V_\lambda)\omega^\perp((I_n \otimes V_\lambda)[U, M_d, S_d])(I_n \otimes V_\lambda)T_d$$

with equality if and only if

$$T_d^T(I_n \otimes V_\lambda)\omega^\perp((I_n \otimes V_\lambda)[U, M_d, S_d])(I_n \otimes V_\lambda)P = 0, \tag{3}$$

see [Kunert(1983)].

To continue, we define the matrix

$$W_\lambda = \Lambda^{-1} - (1_p^T \Lambda^{-1} 1_p)^{-1} \Lambda^{-1} 1_p 1_p^T \Lambda^{-1}, \tag{4}$$

where

$$\Lambda^{-1} = \begin{pmatrix} 1 & -\lambda & 0 & \dots & \dots & 0 \\ -\lambda & 1+\lambda^2 & -\lambda & \ddots & & \vdots \\ 0 & -\lambda & 1+\lambda^2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1+\lambda^2 & -\lambda \\ 0 & \dots & \dots & 0 & -\lambda & 1 \end{pmatrix}$$

is the inverse of Λ. We then can write

$$\begin{aligned} & T_d^T (I_n \otimes V_\lambda) \omega^\perp ((I_n \otimes V_\lambda)[U, M_d, S_d])(I_n \otimes V_\lambda) T_d \\ &= C_{d11} - C_{d12} C_{d22}^- C_{d12}^T - (C_{d13} - C_{d12} C_{d22}^- C_{d23}) \\ &\qquad \times (C_{d33} - C_{d23}^T C_{d22}^- C_{d23})^- (C_{d13} - C_{d12} C_{d22}^- C_{d23})^T, \end{aligned}$$

where $C_{d11} = T_d^T (I_n \otimes W_\lambda) T_d$, $C_{d12} = T_d^T (I_n \otimes W_\lambda) M_d$, $C_{d13} = T_d^T (I_n \otimes W_\lambda) S_d$, $C_{d22} = M_d^T (I_n \otimes W_\lambda) M_d$, $C_{d23} = M_d^T (I_n \otimes W_\lambda) S_d$, and $C_{d33} = S_d^T (I_n \otimes W_\lambda) S_d$, see [Kunert and Martin(2000)].

Now define $c_{dij} = tr(B_t C_{dij})$, the trace of $B_t C_{dij}$, with $B_t = I_t - \frac{1}{t} 1_t 1_t^T$. From [Kunert and Martin(2000)], we have

$$tr(T_d^T (I_n \otimes V_\lambda) \omega^\perp ((I_n \otimes V_\lambda)[U, M_d, S_d])(I_n \otimes V_\lambda) T_d) \le q_d^*, \tag{5}$$

where q_d^* is defined as follows.

1. If $c_{d22} c_{d33} - c_{d23}^2 > 0$, then $q_d^* = c_{d11} - \dfrac{c_{d12}^2 c_{d33} - 2 c_{d12} c_{d13} c_{d23} + c_{d13}^2 c_{d22}}{c_{d22} c_{d33} - c_{d23}^2}$.
2. If $c_{d22} c_{d33} - c_{d23}^2 = 0$ and $c_{d22} > 0$, then $q_d^* = c_{d11} - \dfrac{c_{d12}^2}{c_{d22}}$.
3. If $c_{d22} = 0$ and $c_{d33} > 0$, then $q_d^* = c_{d11} - \dfrac{c_{d13}^2}{c_{d33}}$.
4. If $c_{d22} = c_{d33} = 0$, then $q_d^* = c_{d11}$.

In (5) we have equality, if all matrices C_{dij} are completely symmetric.

Combining what we have seen so far, we get

$$tr(C_d) \leq q_d^* \tag{6}$$

with equality holding iff (3) holds and all matrices C_{dij} are completely symmetric.

We want to find designs with a maximum value of q_d^* where equality holds in (6). It then is useful to split up the design matrices into the contributions from each unit, that is, we write

$$T_d = \begin{pmatrix} T_{d1} \\ T_{d2} \\ \vdots \\ T_{dn} \end{pmatrix}, M_d = \begin{pmatrix} M_{d1} \\ M_{d2} \\ \vdots \\ M_{dn} \end{pmatrix} \quad \text{and} \quad S_d = \begin{pmatrix} S_{d1} \\ S_{d2} \\ \vdots \\ S_{dn} \end{pmatrix},$$

where the $(p \times t)$-matrices T_{du}, M_{du} and S_{du} correspond to unit u, $1 \leq u \leq n$. Further define

$$\begin{aligned} c_{d11}^{(u)} &= tr[B_t(T_{du}^T W_\lambda T_{du})], \\ c_{d12}^{(u)} &= tr[B_t(T_{du}^T W_\lambda M_{du})], \\ c_{d13}^{(u)} &= tr[B_t(T_{du}^T W_\lambda S_{du})], \\ c_{d22}^{(u)} &= tr[B_t(M_{du}^T W_\lambda M_{du})], \\ c_{d23}^{(u)} &= tr[B_t(M_{du}^T W_\lambda S_{du})], \\ c_{d33}^{(u)} &= tr[B_t(S_{du}^T W_\lambda S_{du})]. \end{aligned}$$

Then we get $c_{dij} = \sum_{u=1}^n c_{dij}^{(u)}$, $1 \leq i \leq j \leq 3$. Observe that $c_{dij}^{(u_1)} = c_{dij}^{(u_2)}$ if unit u_1 receives the same sequence of treatments as u_2. This equality remains true if the sequence of unit u_1 can be transformed to the one of unit u_2 by relabeling the treatments. We thus can merge the sequences into K equivalence classes. Each equivalence class ℓ can be identified by a representative sequence. For example, the equivalence class with the representative sequence [1112] contains, among others, the sequences [1112], [2223], [4441]. For class ℓ, we define

$$c_{ij}(\ell) = c_{dij}^{(u_\ell)}$$

where u_ℓ is an unit receiving the representative sequence for class ℓ. With this notation we have

$$c_{dij} = n(\sum_{\ell=1}^{K} \pi_{d\ell} c_{ij}(\ell))$$

where $\pi_{d\ell}$ is the proportion of units in the design receiving sequences from class ℓ. This implies that q_d^* depends on the design only through the $\pi_{d\ell}$. So if we want to maximize q_d^* we have to find appropriate values for $\pi_{d\ell}$. The determination of the appropriate $\pi_{d\ell}$ is made a lot easier by Kushner's method.

We define the bivariate function

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

Then $q_d(x, y) \geq q_d^*$ for all x, y and there is at least one point (x^*, y^*) such that $q_d(x^*, y^*) = q_d^*$, see [Kunert and Martin(2000)]. Defining

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

for $1 \leq \ell \leq K$, we have $q_d(x, y) = n \sum_{\ell=1}^{K} \pi_{d\ell} h_\ell(x, y)$. Hence $q_d(x, y)$ can be derived from the weighted mean of the sequence classes. In particular, we get

$$q_d^* \leq n \min_{x,y} \max_{\ell} \ h_\ell(x, y) \tag{7}$$

for all designs $d \in \Omega_{t,n,p}$.

3 Results

In this section we derive the form of the optimal designs for the cases that $t \geq p \geq 4$. For that we consider all K equivalence classes ℓ with their representative sequences $[s_1^\ell, s_2^\ell, \ldots, s_p^\ell]$. Of particular interest is the class 1 with sequence $[s_1^1, s_2^1, \ldots, s_p^1] = [1, 2, \ldots, p]$. For λ not too small, we show that only class 1 is needed for the optimal designs. In particular, this implies that each treatment should appear at most once within each unit. Our main proposition identifies the upper bound in (7).

Proposition 3.1. *For $t \geq p \geq 4$ define*

$$\lambda^*(p) = (p - 2 - \sqrt{p^2 - 8})/(2(p - 3)).$$

If we assume that $\lambda \in [\lambda^*(p), 1)$ *then*

$$h_1(x^*, 0) = \min_{x,y} \; \max_{\ell} \; h_\ell(x, y). \tag{8}$$

Here $x^* = \arg\min_x h_1(x, 0) = -c_{12}(1)/c_{22}(1)$.

Beweis. If unit u receives a sequence from the equivalence class 1, then $S_{du} = 0$. Therefore, $c_{13}(1) = c_{23}(1) = c_{33}(1) = 0$. Hence, the minimum of $h_1(x, y)$ does not depend on y. We conclude that $h_1(x, y)$ attains its minimum if

$$x = -\frac{c_{12}(1)}{c_{22}(1)},$$

that is, if $x = x^*$, see Lemma A.5. Therefore, it holds for all x and y that

$$\max_{\ell} \; h_\ell(x, y) \geq h_1(x, y) \geq h_1(x^*, 0). \tag{9}$$

On the other hand, we show in the appendix for any ℓ and $x \in [0, 1]$ that

$$h_1(x, 0) \geq h_\ell(x, 0),$$

see Lemma A.6. Since $x^* \in [0, 1]$, see Lemma A.5, this implies that

$$h_1(x^*, 0) \geq h_\ell(x^*, 0).$$

Thus

$$h_1(x^*, 0) = \max_{\ell} \; h_\ell(x^*, 0) \geq \min_{x,y} \; \max_{\ell} \; h_\ell(x, y). \tag{10}$$

Combining (9) and (10) we get equation (8). □

An example of a design d^* such that C_{d^*} is completely symmetric and $tr(C_d) = n\, h_1(x^*, 0)$ is given by any $OA_I(n, p, t, 2)$, that is by any type I orthogonal array of strength 2. An $OA_I(n, p, t, 2)$ is a $p \times n$ matrix with entries from the set $\{1, \ldots, t\}$, such that the columns of any $2 \times n$ submatrices contain all $t(t-1)$ ordered pairs of treatments i and j with $i \neq j$ equally often. For more details, see e.g. [Hedayat and Yan(2008)].

4 Examples

Now we present some examples with a focus on A-optimality which is included in universal optimality. The A-criterion of a design d is given by

$$A(d) = \frac{tr(C_d^+)}{t-1},$$

where C_d^+ denotes the Moore-Penrose inverse of the information matrix C_d. In our situation, it describes the average variance of all pairwise comparisons of direct treatment effects.

Either the following designs are universally optimal or we give lower bounds for their A-efficiencies which depend on λ. The A-efficiency of a design d can be defined by the quotient

$$e_A(d) = \frac{A(d^*)}{A(d)},$$

where d^* is the A-optimal design. If we do not know the A-optimal design, we use the lower bound

$$\frac{(t-1)^2}{nh_1(x^*, 0)}$$

for $tr(C_{d^*}^+)$, see [Kunert and Martin(2000)]. We get a lower bound for the A-efficiency of a design d which is given by

$$\widetilde{e}_A(d) = \frac{(t-1)^2}{tr(C_d^+)nh_1(x^*, 0)}. \quad (11)$$

Example 1: $t = 4$, $p = 4$, $n = 12$

$$d_1 = \begin{bmatrix} 1 & 2 & 3 & 4 & 1 & 2 & 3 & 4 & 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 & 3 & 4 & 1 & 2 & 4 & 3 & 2 & 1 \\ 3 & 4 & 1 & 2 & 4 & 3 & 2 & 1 & 2 & 1 & 4 & 3 \\ 4 & 3 & 2 & 1 & 2 & 1 & 4 & 3 & 3 & 4 & 1 & 2 \end{bmatrix}$$

Design d_1 is an $OA_I(12, 4, 4, 2)$. It is universally optimal for $\lambda \in [1-\sqrt{2}, 1)$.

Example 2: $t = 4$, $p = 4$, $n = 4$

$$d_2 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \\ 3 & 1 & 4 & 2 \\ 4 & 3 & 2 & 1 \end{bmatrix}$$

Design d_2 is universally optimal for $\lambda = 0$, see [Kunert and Stufken(2002)]. Taking account of (11), the A-efficiency of d_2 is at least 0.90508 for any $\lambda \in [0, 1)$. If we restrict attention to designs where each treatment appears at most once within each unit, then a numerical evaluation shows that d_2 is A-optimal within this class of designs for all $\lambda \in (0, 1)$.

Example 3: $t = 5$, $p = 5$, $n = 10$

$$d_3 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 4 & 5 & 1 & 2 & 3 \\ 2 & 3 & 4 & 5 & 1 & 3 & 4 & 5 & 1 & 2 \\ 5 & 1 & 2 & 3 & 4 & 5 & 1 & 2 & 3 & 4 \\ 3 & 4 & 5 & 1 & 2 & 2 & 3 & 4 & 5 & 1 \\ 4 & 5 & 1 & 2 & 3 & 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

Design d_3 is a Williams-Design. It is optimal for $\lambda = 0$. Taking account of (11) again, its A-efficiency is at least 0.85590 for all $\lambda \in (0, 1)$.

A Proofs

Lemma A.1. *Consider the matrix W_λ, defined in (4). Assume $p \geq 4$ and $\lambda \in [\lambda^*(p), 1)$ with $\lambda^*(p)$ as in Proposition 3.1. We then get for the entries w_{ij} of W_λ that $w_{ii} > 0$ for $1 \leq i \leq p$ and $w_{ij} \leq 0$ for $1 \leq i \neq j \leq p$.*

Beweis. It was shown by [Kunert(1985)] for $\lambda \in [\lambda^*(p), 1)$ that $w_{ij} \leq 0$ for $i \neq j$ and $w_{ii} \geq 0$.

To see that the diagonal elements are in fact positive, define $(1-\lambda)^k =: L_k$. We observe $1_p^T \Lambda^{-1} = [L_1, L_2, \ldots, L_2, L_1]$ and $1_p^T \Lambda^{-1} 1_p = L_1(p-\lambda(p-2)) = z_p$, say. For $2 \leq i \leq p-1$ the substitution $p = v + 4$ yields

$$w_{ii} = 1 + \lambda^2 - \frac{L_4}{z_p} = \frac{(1-\lambda)(1+\lambda^2)v - \lambda^3 + \lambda^2 + \lambda + 3}{(1-\lambda)v - 2\lambda + 4}$$

which is positive for $v \geq 0$ and therefore also for $p \geq 4$.

The entries w_{11} and w_{pp} are positive because, using the substitution $p = v + 4$ again, we have

$$w_{11} = w_{pp} = 1 - \frac{L_2}{z_p} = \frac{(1-\lambda)v - \lambda + 3}{(1-\lambda)v - 2\lambda + 4}$$

which is positive. □

Lemma A.2. *For $t \geq p \geq 4$ and $\lambda \in [\lambda^*(p), 1)$ consider an arbitrary sequence class ℓ. Then $c_{11}(1) \geq c_{11}(\ell)$ and $c_{11}(1) + 2c_{12}(1) \geq c_{11}(\ell) + 2c_{12}(\ell)$.*

Beweis. Assume u is a unit receiving a sequence $s = [s_1, ..., s_p]$ from class ℓ. Then

$$c_{11}(\ell) = tr(B_t T_{du}^T W_\lambda T_{du})$$

and

$$c_{12}(\ell) = tr(B_t T_{du}^T W_\lambda M_{du}).$$

Since W_λ has column-sums zero and since $T_{du} 1_t = 1_p$, it follows that

$$c_{11}(\ell) = tr(T_{du}^T W_\lambda T_{du}) \quad \text{and} \quad c_{12}(\ell) = tr(T_{du}^T W_\lambda M_{du}).$$

For $1 \leq i \leq t$ denote the i-th column of T_{du} by $t_{du}^{(i)}$ and the i-th column of M_{du} by $m_{du}^{(i)}$. Observe that

- the j-th entry of $t_{du}^{(i)}$ is 1 if $s_j = i$ and 0 otherwise,
- the j-th entry of $m_{du}^{(i)}$ is 1 if $j \geq 2, s_{j-1} = i$ and $s_{j-1} \neq s_j$. In all other cases it is 0.

For any given i it follows that

$$t_{du}^{(i)T} W_\lambda t_{du}^{(i)} = \sum_{j=1}^{p} I(s_j = i) \sum_{r=1}^{p} w_{jr} I(s_r = i),$$

where $I(statement)$ is 1 if $statement$ is true and 0 otherwise. Hence,

$$\begin{aligned} c_{11}(\ell) &= \sum_{i=1}^{t} t_{du}^{(i)T} W_\lambda t_{du}^{(i)} \\ &= \sum_{i=1}^{t} \left(\sum_{j=1}^{p} \sum_{r=1}^{p} w_{jr} I(s_j = s_r = i) \right) = \sum_{j=1}^{p} \sum_{r=1}^{p} w_{jr} I(s_j = s_r) \\ &= \sum_{j=1}^{p} w_{jj} + 2 \sum_{j=1}^{p-1} \sum_{r=j+1}^{p} w_{jr} I(s_j = s_r) \qquad (12) \\ &\leq \sum_{j=1}^{p} w_{jj}, \end{aligned}$$

because all $w_{jr} \leq 0$ if $j \neq r$. It is easy to see that $c_{11}(1) = \sum_{j=1}^{p} w_{jj}$ and, therefore, we have proved that

$$c_{11}(\ell) \leq c_{11}(1).$$

We can also use equation (12) to derive the slightly sharper bound

$$c_{11}(\ell) \leq \sum_{j=1}^{p} w_{jj} + 2 \sum_{j=1}^{p-1} w_{j,j+1} I(s_j = s_{j+1}). \tag{13}$$

On the other hand,

$$t_{du}^{(i)T} W_\lambda m_{du}^{(i)} = \sum_{j=1}^{p} I(s_j = i) \sum_{r=2}^{p} w_{jr} I(s_r \neq s_{r-1} = i).$$

Therefore,

$$\begin{aligned} c_{12}(\ell) &= \sum_{i=1}^{t} \left(\sum_{j=1}^{p} \sum_{r=1}^{p} w_{jr} I(s_j = i, s_r \neq s_{r-1} = i) \right) \\ &= \sum_{j=1}^{p} \sum_{r=1}^{p} w_{jr} I(s_j = s_{r-1} \neq s_r) \\ &= \sum_{j=1}^{p} \sum_{r \neq j} w_{jr} I(s_j = s_{r-1} \neq s_r), \end{aligned}$$

because $s_j = s_{r-1} \neq s_r$ can never hold for $r = j$. On the other hand, if $r = j+1$, then $s_j = s_{r-1} \neq s_r$ becomes $s_j \neq s_{j+1}$. Making use of the fact that all $w_{jr} \leq 0$ for all $r \neq j$, we conclude that

$$c_{12}(\ell) \leq \sum_{j=1}^{p-1} w_{j,j+1} I(s_j \neq s_{j+1}).$$

Combining this with equation (13), we conclude that

$$c_{11}(\ell) + 2c_{12}(\ell) \leq \sum_{j=1}^{p} w_{jj} + 2 \sum_{j=1}^{p-1} w_{j,j+1} I(s_j = s_{j+1}) + 2 \sum_{j=1}^{p-1} w_{j,j+1} I(s_j \neq s_{j+1}).$$

Since it is easy to see that

$$c_{12}(1) = \sum_{j=1}^{p-1} w_{j,j+1}, \tag{14}$$

we have proved that $c_{11}(\ell) + 2c_{12}(\ell) \leq c_{11}(1) + 2c_{12}(1)$. □

Lemma A.3. *Under the conditions of Lemma A.2 we have* $c_{22}(1) \geq c_{22}(\ell)$.

Beweis. If M_{du} is as in the proof of Lemma A.2, then

$$c_{22}(\ell) = tr(B_t M_{du}^T W_\lambda M_{du}) = tr(M_{du}^T W_\lambda M_{du}) - \frac{1}{t} 1_t^T M_{du}^T W_\lambda M_{du} 1_t.$$

Observe that $M_{du}1_t$ is a p-dimensional vector with entries 1 or 0. The first element of $M_{du}1_t$ is always 0, the j-th element, for $j \geq 2$, is 0, if $s_{j-1} = s_j$, and 1, otherwise. Hence,

$$\begin{aligned} 1_t^T M_{du}^T W_\lambda M_{du} 1_t &= \sum_{j=2}^{p} \sum_{r=2}^{p} w_{jr} I(s_{j-1} \neq s_j) I(s_{r-1} \neq s_r) \\ &= \sum_{j=2}^{p} w_{jj} I(s_{j-1} \neq s_j) \\ &\quad + 2 \sum_{j=2}^{p-1} \sum_{r=j+1}^{p} w_{jr} I(s_{j-1} \neq s_j) I(s_{r-1} \neq s_r). \end{aligned}$$

Defining $m_{du}^{(i)}$ as in the proof of Lemma A.2, we get

$$\begin{aligned} tr(M_{du}^T W_\lambda M_{du}) &= \sum_{i=1}^{t} m_{du}^{(i)T} W_\lambda m_{du}^{(i)} \\ &= \sum_{i=1}^{t} \left(\sum_{j=2}^{p} \sum_{r=2}^{p} w_{jr} I(s_j \neq s_{j-1} = i = s_{r-1} \neq s_r) \right) \\ &= \sum_{j=2}^{p} \sum_{r=2}^{p} w_{jr} I(s_j \neq s_{j-1} = s_{r-1} \neq s_r) \\ &= \sum_{j=2}^{p} w_{jj} I(s_j \neq s_{j-1}) \\ &\quad + 2 \sum_{j=2}^{p-1} \sum_{r=j+1}^{p} w_{jr} I(s_j \neq s_{j-1} = s_{r-1} \neq s_r) \\ &\leq \sum_{j=2}^{p} w_{jj} I(s_j \neq s_{j-1}). \end{aligned}$$

Combining the two parts, we get

$$\begin{aligned} c_{22}(\ell) &\leq \sum_{j=2}^{p} w_{jj} I(s_j \neq s_{j-1}) \\ &\quad - \frac{1}{t} \Big(\sum_{j=2}^{p} w_{jj} I(s_j \neq s_{j-1}) \\ &\quad + 2 \sum_{j=2}^{p-1} \sum_{r=j+1}^{p} w_{jr} I(s_{j-1} \neq s_j) I(s_{r-1} \neq s_r) \Big) \\ &= \frac{t-1}{t} \sum_{j=2}^{p} w_{jj} I(s_j \neq s_{j-1}) \\ &\quad - \frac{2}{t} \sum_{j=2}^{p-1} \sum_{r=j+1}^{p} w_{jr} I(s_{j-1} \neq s_j) I(s_{r-1} \neq s_r) \\ &\leq \frac{t-1}{t} \sum_{j=2}^{p} w_{jj} - \frac{2}{t} \sum_{j=2}^{p-1} \sum_{r=j+1}^{p} w_{jr}, \end{aligned}$$

where the last inequality made use of the fact that all $w_{jj} > 0$ and all $w_{jr} \leq 0$, for $j \neq r$. It is easy to see that this bound for $c_{22}(\ell)$ equals $c_{22}(1)$. $\square$

Lemma A.4. *Under the conditions of Lemma A.2 consider an $x \in [0,1]$. We than have that*

$$c_{11}(1) + 2xc_{12}(1) \geq c_{11}(\ell) + 2xc_{12}(\ell).$$

Beweis. Define $f_{(\ell)}(x) := c_{11}(\ell)+2xc_{12}(\ell)$. Due to the linearity of the function $f_{(\ell)}(x)$, it is sufficient to show that $f_{(1)}(0) \geq f_{(\ell)}(0)$ and $f_{(1)}(1) \geq f_{(\ell)}(1)$, $1 \leq \ell \leq K$. The inequalities are valid because of Lemma A.2. $\square$

Lemma A.5. *Under the conditions of Lemma A.2, we have that $c_{22}(1) > 0$ and the function $h_1(x, 0)$ has an unique minimum at*

$$x^* = -\frac{c_{12}(1)}{c_{22}(1)} \in [0, 1].$$

Beweis. It follows from the proof of Lemma A.3 that

$$c_{22}(1) = \frac{t-1}{t} \sum_{j=1}^{p} w_{jj} - \frac{2}{t} \sum_{j=2}^{p-1} \sum_{r=j+1}^{p} w_{jr}.$$

Making use of Lemma A.1, we see that $c_{22}(1) > 0$. This implies that $h_1(x, 0) = c_{11}(1) + 2xc_{12}(x) + x^2 c_{22}(x)$ has a unique minimum at the point x^*. It only remains to show that $0 \leq x^* \leq 1$.

For each j, it holds that $\sum_{r=1}^{p} w_{jr} = 0$. This implies that

$$\begin{aligned} \frac{t-1}{t} \sum_{j=2}^{p} w_{jj} &= -\frac{t-1}{t} \sum_{j=2}^{p} \left(\sum_{r=1}^{j-1} w_{jr} + \sum_{r=j+1}^{p} w_{jr} \right) \\ &= -\frac{t-1}{t} \sum_{j=2}^{p} \sum_{r=1}^{j-1} w_{jr} - \frac{t-1}{t} \sum_{j=2}^{p-1} \sum_{r=j+1}^{p} w_{jr}. \end{aligned}$$

We therefore can rewrite $c_{22}(1)$ as

$$c_{22}(1) = -\frac{t-1}{t} \sum_{j=2}^{p} \sum_{r=1}^{j-1} w_{jr} - \frac{t+1}{t} \sum_{j=2}^{p-1} \sum_{r=j+1}^{p} w_{jr}.$$

Since all w_{ij} in this sum are non-positive, we get

$$\begin{aligned} c_{22}(1) &\geq -\frac{t-1}{t}w_{21} - \frac{t+1}{t}\sum_{j=2}^{p-1} w_{j,j+1} \\ &\geq -\frac{t-1}{t}w_{21} - \sum_{j=2}^{p-1} w_{j,j+1} - \frac{1}{t}w_{p-1,p}. \end{aligned}$$

Observing that $w_{p-1,p} = w_{21} = w_{12}$, we conclude that

$$c_{22}(1) \geq -\sum_{j=1}^{p-1} w_{j,j+1}.$$

Equation (14) then shows that $0 \leq -c_{12}(1) \leq c_{22}(1)$. This completes the proof. □

Lemma A.6. *Under the conditions of Lemma A.2 we get for any* $x \in [0,1]$ *that*

$$h_1(x,0) \geq h_\ell(x,0).$$

Beweis. Making use of Lemma A.3 and A.4 we observe that

$$\begin{aligned} h_1(x,0) &= c_{11}(1) + 2xc_{12}(1) + x^2c_{22}(1) \\ &\geq c_{11}(\ell) + 2xc_{12}(\ell) + x^2c_{22}(\ell) \\ &= h_\ell(x,0). \end{aligned}$$

□

Acknowledgements The financial support by the Deutsche Forschungsgemeinschaft (SFB 823, Statistik nichtlinearer dynamischer Prozesse) is gratefully acknowledged.

Literatur

[Afsarinejad and Hedayat(2002)] Afsarinejad, K., Hedayat, A. S., 2002. Repeated measurements designs for a model with self and simple mixed carryover effects. Journal of Statistical Planning and Inference 106, 449–459.

[Hedayat and Yan(2008)] Hedayat, A. S., Yan, Z., 2008. Crossover designs based on type I orthogonal arrays for a self and simple mixed carryover effects model with correlated errors. Journal of Statistical Planning and Inference 138, 2201–2213.

[Kiefer(1975)] Kiefer, J., 1975. Construction and optimality of generalized Youden designs. In: Srivastava, J. N. (Ed.), A Survey of Statistical Design and Linear Models. North-Holland, Amsterdam, pp. 333–353.

[Kunert(1983)] Kunert, J., 1983. Optimal design and refinement of the linear model with applications to repeated measurements designs. The Annals of Statistics 11, 247–257.

[Kunert(1985)] Kunert, J., 1985. Optimal repeated measurements designs for correlated observations and analysis by weighted least squares. Biometrika 72, 375–389.

[Kunert and Martin(2000)] Kunert, J., Martin, R., 2000. On the determination of optimal designs for an interference model. The Annals of Statistics 28, 1728–1742.

[Kunert and Stufken(2002)] Kunert, J., Stufken, J., 2002. Optimal crossover designs in a model with self and mixed carryover effects. Journal of the American Statistical Association 97, 898–906.

[Kunert and Stufken(2008)] Kunert, J., Stufken, J., 2008. Optimal crossover designs for two treatments in the presence of mixed and self carryover effects. Journal of the American Statistical Association 103, 1641–1647.

[Kushner(1997)] Kushner, H. B., 1997. Optimal repeated measurements designs: the linear optimality equations. The Annals of Statistics 25, 2328–2344.