SFB
823

Efficient designs for the estimation of mixed and self carryover effects

Joachim Kunert, Johanna Mielke

Discussion Paper

Nr. 8/2018

Efficient designs for the estimation of mixed and self carryover effects

Joachim Kunert
Department of Statistics, University of Dortmund, Germany
Johanna Mielke
Novartis Pharma AG, Basel, Switzerland

Abstract

Biosimilars are copies of biological medicines that are developed by a competitor after the patent for the originator drug has expired. Extensive clinical trials are required to show therapeutic equivalence between the biosimilar and its reference product before a biosimilar can be sold on the market. However, even after more than 10 years of experience with biosimilars in Europe, there is still some uncertainty if the patients who are already taking the reference product can switch between the biosimilar and its reference product. One convenient way to assess the impact of switches is the analysis of mixed and self carryover effects: if the products are switchable, there should not be any difference in the carryover effects. This paper determines a series of simple designs which are highly efficient for the comparison of the mixed and self carryover effects of two treatments. The proof of efficiency is not straightforward because the information matrix of the efficient designs is not completely symmetric.

Keywords: A-optimality, biosimilars, crossover designs, linear model, switchability

1 Introduction

After the patent of a pharmaceutical product expires, it is possible for competing companies to produce and sell a copy of the originator product. In the context of small molecule drugs, this is already well established and the copied products are known as generics. For large molecule drugs (so-called biologics), however, this is still a fairly new concept with the first product approved in Europe in 2006 and in the United States in 2015. A copy of a biologic is called a biosimilar. The development of biosimilars is currently in the focus of attention because several important patents on biological drugs expired in recent years or will expire soon (Generics and Biosimilar Initiative 2014).

Due to limited experience with biosimilars in practice, the higher complexity and the fact that, from a chemical point of view, a biosimilar is only required to be similar (Schellekens 2004), and not, as for generics, identical, to the originator product, there is still uncertainty among patients, physicians and health care providers if a patient who is already taking the originator product should be switched to the biosimilar product or if even substituting of the treatment at the pharmacy level is acceptable (e.g., Ebbers et al. 2012) which could lead to multiple switches between the biosimilar and the originator product. For providing assurance that alternating the treatments, for example, at pharmacy level is justifiable, it is necessary to confirm that the treatment response during multiple switches between the biosimilar and its reference product is comparable.

This calls for a crossover study where the units are observed over several periods, with the possibility of changing the treatment between periods. For gaining market authorisation as a biosimilar, it would be sufficient to observe the first period only. If in the first period there is no difference between the treatments, this means that the direct effects are similar. In later periods, however, there may additionally be carryover effects. One way of confirming that switching does not influence the efficacy of the treatment is to analyse the so-called self and mixed carryover effects. The model with mixed and self carryover effects was introduced by Afsarinejad and Hedayat (2002). It assumes that each treatment has two different carryover effects, one that is present if a subject stays on the treatment from the previous period (self carryover effect, continuous treatment) and the other one that is present if subjects change the treatment (mixed carryover effect, switching). If switching has no impact, there should be no differences between the carryover effects of the treatments.

Kunert and Stufken (2002, 2008) determined optimal crossover designs for estimating the direct treatment effects in the model with self and mixed carryover effects. The second article, Kunert and Stufken (2008), dealt with the case of two treatments which is relevant for our application (biosimilar and originator). However, for showing the switchability of a biosimilar and its reference product, the direct effects are not of primary interest. When the question of switchability is addressed, biosimilarity is already confirmed and there is a strong indication that there is no relevant difference between the treatment effects in the first period, i.e. between the direct effects. For confirming switchability, we therefore need to focus on estimating the carryover effects. In this paper, we propose efficient designs for estimating mixed and self carryover effects.

2 Description of the model

We consider the model that was assumed by Kunert and Stufken (2008). The response $y_{u,r}$ of subject u in period r is modelled as

$$y_{u,r} = \begin{cases} \alpha_u + \beta_r + \tau_{d(u,r)} + \rho_{d(u,r-1)} + e_{u,r}, & \text{if } d(u,r) \neq d(u,r-1), \\ \alpha_u + \beta_r + \tau_{d(u,r)} + \chi_{d(u,r-1)} + e_{u,r}, & \text{if } d(u,r) = d(u,r-1). \end{cases}$$

Here, $d(u,r)$ is the treatment assigned to subject u in period r $(1 \leq u \leq n, 1 \leq r \leq p)$ by the design d, α_u is the effect of subject u, β_r is the effect of period r, τ_i is the direct effect, ρ_i the mixed carryover effect and χ_i is the self carryover effect of treatment i. No carryover effect is present in the first period, i.e., $\rho_{d(u,0)} = \chi_{d(u,0)} = 0$. The errors $e_{u,r}, 1 \leq u \leq n, 1 \leq r \leq p$ are assumed to be independent identically distributed with expectation 0 and variance $\sigma^2 > 0$.

The set of all designs with t treatments, n subjects and p periods is denoted by $\Omega_{t,n,p}$. We focus on the case of two treatments (Reference treatment R, Test treatment T), i.e. $t = 2$.

For a given design $d \in \Omega_{2,n,p}$ we define $\mathbf{T}_d$ as the design matrix of direct effects, while $\mathbf{S}_d$ is the design matrix of the self carryover effects and $\mathbf{M}_d$ is the design matrix of the mixed carryover effects. We also consider the matrices $\mathbf{U} = \mathbf{I}_n \otimes \mathbf{1}_p$ and $\mathbf{P} = \mathbf{1}_n \otimes \mathbf{I}_p$, where $\otimes$ denotes the Kronecker-product of matrices, $\mathbf{I}_s$ is the $(s \times s)$-identity matrix and $\mathbf{1}_s$ is the s-vector with all entries 1. Then $\mathbf{U}$ and $\mathbf{P}$ are the design matrices for subject and period effects, respectively, and the model in vector notation can be written as

$$\mathbf{y} = \mathbf{T}_d\tau + \mathbf{S}_d\chi + \mathbf{M}_d\rho + \mathbf{U}\alpha + \mathbf{P}\beta + \mathbf{e},$$

where τ is the vector of direct effects, χ is the vector of self carryover and ρ is the vector of mixed carryover effects. Further, α, β and $\mathbf{e}$ are the vectors of subject effects, period effects and of the residual errors, respectively. We are interested in the simultaneous estimation of the 4-dimensional vector of all carryover effects

$$\delta = \begin{bmatrix} \chi \\ \rho \end{bmatrix}.$$

For a matrix $\mathbf{A}$, we define $\omega(\mathbf{A}) = \mathbf{A}(\mathbf{A}^T\mathbf{A})^+\mathbf{A}^T$, where $\mathbf{A}^T$ is the transpose and $(\mathbf{A}^T\mathbf{A})^+$ is the Moore-Penrose generalized inverse. Setting $\omega^\perp(\mathbf{A}) = \mathbf{I} - \omega(\mathbf{A})$, the information matrix for the estimation of δ then is

$$\mathbf{C}_d = [\mathbf{S}_d, \mathbf{M}_d]^T \omega^\perp([\mathbf{P}, \mathbf{U}, \mathbf{T}_d])[\mathbf{S}_d, \mathbf{M}_d].$$

Note that $[\mathbf{S}_d, \mathbf{M}_d]\mathbf{1}_4 = \mathbf{P}[0, 1, \ldots, 1]^T$, because each subject in all periods but the first experiences one of the four carryover effects. Therefore, the information matrix $\mathbf{C}_d$ has row and column sums zero and only contrasts of the carryover effects are estimable.

To compare the performances of the designs, we consider the A-criterion, see e.g. Pukelsheim (1993, p. 210). For a design $d \in \Omega_{2,n,p}$, we define $\lambda_1(\mathbf{C}_d) \geq \lambda_2(\mathbf{C}_d) \geq \lambda_3(\mathbf{C}_d) \geq \lambda_4(\mathbf{C}_d)$ as the ordered eigenvalues of $\mathbf{C}_d$. Note that $\lambda_4(\mathbf{C}_d) = 0$, since $\mathbf{C}_d$ has row- and column-sums zero. We then define the A-criterion φ_A as

$$\varphi_A(d) = \frac{1}{\left(\frac{1}{\lambda_1(\mathbf{C}_d)} + \frac{1}{\lambda_2(\mathbf{C}_d)} + \frac{1}{\lambda_3(\mathbf{C}_d)}\right)}.$$

An A-optimal design d^* maximizes $\varphi_A(d)$.

3 Determination of A-optimal designs

Since the information matrix $\mathbf{C}_d$ has row- and column-sums zero, we can rewrite

$$\mathbf{C}_d = \mathbf{B}_4\mathbf{C}_d\mathbf{B}_4 = \mathbf{B}_4[\mathbf{S}_d, \mathbf{M}_d]^T\omega^\perp([\mathbf{P}, \mathbf{U}, \mathbf{T}_d])[\mathbf{S}_d, \mathbf{M}_d]\mathbf{B}_4,$$

where $\mathbf{B}_4 = \omega^\perp(\mathbf{1}_4)$.

For two square matrices $\mathbf{G} \in \mathbb{R}^{k\times k}$ and $\mathbf{D} \in \mathbb{R}^{k\times k}$, we write $\mathbf{G} \leq \mathbf{D}$ if $\mathbf{D} - \mathbf{G}$ is nonnegative definite. With this notation, we get an immediate upper bound for $\mathbf{C}_d$, namely

$$\mathbf{C}_d \leq \tilde{\mathbf{C}}_d = \mathbf{B}_4[\mathbf{S}_d, \mathbf{M}_d]^T\omega^\perp([\mathbf{U}, \mathbf{T}_d])[\mathbf{S}_d, \mathbf{M}_d]\mathbf{B}_4,$$

see Kunert (1983). Equality holds iff

$$\mathbf{B}_4[\mathbf{S}_d, \mathbf{M}_d]^T\omega^\perp([\mathbf{U}, \mathbf{T}_d])\mathbf{P} = 0. \tag{1}$$

Since $\omega^\perp[\mathbf{U}, \mathbf{T}_d] = \omega^\perp(\mathbf{U}) - \omega^\perp(\mathbf{U})\mathbf{T}_d(\mathbf{T}_d^T\omega^\perp(\mathbf{U})\mathbf{T}_d)^+\mathbf{T}_d^T\omega^\perp(\mathbf{U})$, the matrix $\tilde{\mathbf{C}}_d$ can be split up into

$$\tilde{\mathbf{C}}_d = \mathbf{C}_{d11} - \mathbf{C}_{d12}\mathbf{C}_{d22}^+\mathbf{C}_{d12}^T,$$

where

$$\mathbf{C}_{d11} = \mathbf{B}_4[\mathbf{S}_d, \mathbf{M}_d]^T\omega^\perp(\mathbf{U})[\mathbf{S}_d, \mathbf{M}_d]\mathbf{B}_4,$$
$$\mathbf{C}_{d12} = \mathbf{B}_4[\mathbf{S}_d, \mathbf{M}_d]^T\omega^\perp(\mathbf{U})\mathbf{T}_d,$$

and

$$\mathbf{C}_{d22} = \mathbf{T}_d^T\omega^\perp(\mathbf{U})\mathbf{T}_d.$$

This implies that there is an upper bound for the A-criterion

$$\varphi_A(d) \leq \tilde{\varphi}_A(d),$$

where

$$\tilde{\varphi}_A(d) = \frac{1}{\left(\frac{1}{\lambda_1(\tilde{\mathbf{C}}_d)} + \frac{1}{\lambda_2(\tilde{\mathbf{C}}_d)} + \frac{1}{\lambda_3(\tilde{\mathbf{C}}_d)}\right)}.$$

This bound is easier to treat.

Each subject receives a sequence of treatments. Consider an arbitrary sequence $z \in \{R, T\}^p = Z_p$, say. For this sequence, we define

- $\mathbf{T}_z$ is the design matrix for direct effects for this sequence, i.e., the design matrix for direct effects we would get from a design consisting of one subject only, receiving sequence z,
- $\mathbf{S}_z$ is the design matrix for self carryover effects for this sequence,
- $\mathbf{M}_z$ is the design matrix for mixed carryover effects for this sequence.

For a design $d \in \Omega_{2,n,p}$, define $\pi_d(z)$ as the proportion of subjects receiving sequence z, $z \in Z_p$. It is clear that all $\pi_d(z) \geq 0$ and that $\sum_{z\in Z_p} \pi_d(z) = 1$. Since the number of subjects receiving sequence $z \in Z_p$ is a natural number or 0, we conclude that $\pi_d(z)$ must be an integral multiple of $1/n$, where n is the number of subjects. It then is often convenient to allow for *approximate designs*, for which this restriction is removed.

It is easy to see that the $\mathbf{C}_{dij}$ are linear in the sequences, $\mathbf{C}_{dij} = n\sum_{z\in Z} \pi_d(z)\mathbf{C}_{ij}(z)$, where

$$\begin{aligned}
\mathbf{C}_{11}(z) &= \mathbf{B}_4[\mathbf{S}_z, \mathbf{M}_z]^T \omega^{\perp}(\mathbf{1}_p)[\mathbf{S}_z, \mathbf{M}_z]\mathbf{B}_4, \\
\mathbf{C}_{12}(z) &= \mathbf{B}_4[\mathbf{S}_z, \mathbf{M}_z]^T \omega^{\perp}(\mathbf{1}_p)\mathbf{T}_z, \\
\mathbf{C}_{22}(z) &= \mathbf{T}_z^T \omega^{\perp}(\mathbf{1}_p)\mathbf{T}_z.
\end{aligned}$$

Making use of the linearity of the $\mathbf{C}_{dij}$, Kushner (1997) introduced a general method for deriving optimal crossover designs. However, Kushner's (1997) original method works only if all $\mathbf{C}_{dij}$ are square matrices, i.e., matrices with the same number of rows and columns. Since for our problem, $\mathbf{C}_{d12}$ is a (4×2)-matrix, we have to adapt Kushner's (1997) method to our situation.

Proposition 1 *Assume $\mathbf{X} \in \mathbb{R}^{2\times 4}$ is an arbitrary matrix. Then*

$$\tilde{\mathbf{C}}_d \leq \mathbf{C}_{d11} - \mathbf{C}_{d12}\mathbf{X} - \mathbf{X}^T\mathbf{C}_{d12}^T + \mathbf{X}^T\mathbf{C}_{d22}\mathbf{X}.$$

Equality holds if $\mathbf{X} = \mathbf{X}_d$, where

$$\mathbf{X}_d = \mathbf{C}_{d22}^{+}\mathbf{C}_{d12}^T.$$

Proof: Define $\mathbf{\Delta} = \mathbf{X} - \mathbf{X}_d$. Then

$$\mathbf{C}_{d11} - \mathbf{C}_{d12}\mathbf{X} - \mathbf{X}^T\mathbf{C}_{d12}^T + \mathbf{X}^T\mathbf{C}_{d22}\mathbf{X}$$

$$
\begin{aligned}
&= \mathbf{C}_{d11} - \mathbf{C}_{d12}(\mathbf{\Delta} + \mathbf{X}_d) - (\mathbf{\Delta} + \mathbf{X}_d)^T\mathbf{C}_{d12}^T + (\mathbf{\Delta} - \mathbf{X}_d)^T\mathbf{C}_{d22}(\mathbf{\Delta} - \mathbf{X}_d) \\
&= \mathbf{C}_{d11} - \mathbf{C}_{d12}\mathbf{X}_d - \mathbf{X}_d^T\mathbf{C}_{d12}^T + \mathbf{X}_d^T\mathbf{C}_{d22}\mathbf{X}_d \\
&\quad - \mathbf{C}_{d12}\mathbf{\Delta} - \mathbf{\Delta}^T\mathbf{C}_{d12}^T + \mathbf{\Delta}^T\mathbf{C}_{d22}\mathbf{\Delta} + \mathbf{X}_d^T\mathbf{C}_{d22}\mathbf{\Delta} + \mathbf{\Delta}^T\mathbf{C}_{d22}\mathbf{X}_d.
\end{aligned}
$$

Now, by setting $\mathbf{Q} = \omega^{\perp}(\mathbf{U})\mathbf{T}_d$ and using that $\mathbf{Q}^T\omega(\mathbf{Q}) = \mathbf{Q}^T$, we observe that

$$
\begin{aligned}
\mathbf{C}_{d22}\mathbf{X}_d &= \mathbf{T}_d^T\omega^{\perp}(\mathbf{U})\mathbf{T}_d\,(\mathbf{T}_d^T\omega^{\perp}(\mathbf{U})\mathbf{T}_d)^{+}\mathbf{T}_d^T\omega^{\perp}(\mathbf{U})[\mathbf{S}_d, \mathbf{M}_d]\mathbf{B}_4 \\
&= \mathbf{T}_d^T\omega^{\perp}(\mathbf{U})[\mathbf{S}_d, \mathbf{M}_d]\mathbf{B}_4 \\
&= \mathbf{C}_{d12}^T.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\mathbf{C}_{d11} - \mathbf{C}_{d12}\mathbf{X} - \mathbf{X}^T\mathbf{C}_{d12}^T + \mathbf{X}^T\mathbf{C}_{d22}\mathbf{X} \\
&= \mathbf{C}_{d11} - \mathbf{C}_{d12}\mathbf{X}_d + \mathbf{\Delta}^T\mathbf{C}_{d22}\mathbf{\Delta} \\
&\geq \mathbf{C}_{d11} - \mathbf{C}_{d12}\mathbf{X}_d = \tilde{\mathbf{C}}_d.
\end{aligned}
$$

This completes the proof.

Note that the right-hand side of the inequality in Proposition 1 is linear in the sequences,

$$
\begin{aligned}
&\mathbf{C}_{d11} - \mathbf{C}_{d12}\mathbf{X} - \mathbf{X}^T\mathbf{C}_{d12}^T + \mathbf{X}^T\mathbf{C}_{d22}\mathbf{X} \\
&= n\sum_{z\in Z}\pi_d(z)\big(\mathbf{C}_{11}(z) - \mathbf{C}_{12}(z)\mathbf{X} - \mathbf{X}^T\mathbf{C}_{12}^T(z) + \mathbf{X}^T\mathbf{C}_{22}(z)\mathbf{X}\big).
\end{aligned} \tag{2}
$$

As a first step, we can use this proposition to derive an upper bound for the smallest non-zero eigenvalue. Define

$$
\mathbf{b}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}.
$$

Then $\mathbf{b}_2\mathbf{b}_2^T = \mathbf{B}_2$. With this notation we get an immediate consequence of Proposition 1.

Proposition 2 *Assume* $\mathbf{k}$ *is any 4-dimensional vector, such that* $\mathbf{k}^T\mathbf{1}_4 = 0$ *and* $x \in \mathbb{R}$. *Then*

$$
\mathbf{k}^T\tilde{C}_d\mathbf{k} \leq \mathbf{k}^T\mathbf{C}_{d11}\mathbf{k} - 2\mathbf{k}^T\mathbf{C}_{d12}\mathbf{b}_2 x + \mathbf{b}_2^T\mathbf{C}_{d22}\mathbf{b}_2 x^2.
$$

Equality holds for $x = \mathbf{k}^T\mathbf{C}_{d12}\mathbf{C}_{d22}^{+}\mathbf{b}_2 = x_d$, *say.*

Proof: For an arbitrary $x \in \mathbb{R}$, define $J(x) = \mathbf{k}^T\mathbf{C}_{d11}\mathbf{k} - 2\mathbf{k}^T\mathbf{C}_{d12}\mathbf{b}_2 x + \mathbf{b}_2^T\mathbf{C}_{d22}\mathbf{b}_2 x^2$.

Case 1: $\mathbf{k}^T\mathbf{C}_{d12}\mathbf{b}_2 \neq 0$.

Consider the matrix

$$\mathbf{X} = \mathbf{C}_{d12}^T \frac{x}{\mathbf{k}^T\mathbf{C}_{d12}^T\mathbf{b}_2} \in \mathbb{R}^{2\times 4}.$$

Then $\mathbf{k}^T\mathbf{X}^T\mathbf{b}_2 = x$.

It follows from Proposition 1 that

$$\mathbf{k}^T\tilde{C}_d\mathbf{k} \leq \mathbf{k}^T\mathbf{C}_{d11}\mathbf{k} - \mathbf{k}^T\mathbf{C}_{d12}\mathbf{X}\mathbf{k} - \mathbf{k}^T\mathbf{X}^T\mathbf{C}_{d12}^T\mathbf{k} + \mathbf{k}^T\mathbf{X}^T\mathbf{C}_{d22}\mathbf{X}\mathbf{k}.$$

Since $\mathbf{C}_{d12}$ has row-sums 0, we have $\mathbf{C}_{d12}\mathbf{b}_2\mathbf{b}_2^T = \mathbf{C}_{d12}$. Since $\mathbf{C}_{d22}$ has both row- and column-sums 0, we even have $\mathbf{b}_2\mathbf{b}_2^T\mathbf{C}_{d22}\mathbf{b}_2\mathbf{b}_2^T = \mathbf{C}_{d22}$. Hence

$$\begin{aligned}\mathbf{k}^T\tilde{C}_d\mathbf{k} &\leq \mathbf{k}^T\mathbf{C}_{d11}\mathbf{k} - \mathbf{k}^T\mathbf{C}_{d12}\mathbf{b}_2\mathbf{b}_2^T\mathbf{X}\mathbf{k} \\ &\quad -\mathbf{k}^T\mathbf{X}^T\mathbf{b}_2\mathbf{b}_2^T\mathbf{C}_{d12}^T\mathbf{k} + \mathbf{k}^T\mathbf{X}^T\mathbf{b}_2\mathbf{b}_2^T\mathbf{C}_{d22}\mathbf{b}_2\mathbf{b}_2^T\mathbf{X}\mathbf{k} \\ &= \mathbf{k}^T\mathbf{C}_{d11}\mathbf{k} - \mathbf{k}^T\mathbf{C}_{d12}\mathbf{b}_2 x - x\mathbf{b}_2^T\mathbf{C}_{d12}^T\mathbf{k} + x\mathbf{b}_2^T\mathbf{C}_{d22}\mathbf{b}_2 x \\ &= J(x).\end{aligned}$$

Because of Proposition 1, equality holds iff $\mathbf{X} = \mathbf{X}_d = \mathbf{C}_{d22}^+\mathbf{C}_{d12}^T$, i.e., iff

$$x = \mathbf{b}_2^T\mathbf{C}_{d22}^+\mathbf{C}_{d12}^T\mathbf{k} = x_d.$$

Case 2: $\mathbf{k}^T\mathbf{C}_{d12}\mathbf{b}_2 = 0$.

Then $J(x) = \mathbf{k}^T\mathbf{C}_{d11}\mathbf{k} + \mathbf{b}_2^T\mathbf{C}_{d22}\mathbf{b}_2 x^2 \geq \mathbf{k}^T\mathbf{C}_{d11}\mathbf{k}$. Equality holds if $x = 0$.

On the other hand, it follows from $\mathbf{k}^T\mathbf{C}_{d12}\mathbf{b}_2 = 0$ that $\mathbf{k}^T\mathbf{C}_{d12}\mathbf{b}_2\mathbf{b}_2^T = 0$, and, therefore, that $\mathbf{k}^T\mathbf{C}_{d12} = 0$. Hence, $\mathbf{k}^T\mathbf{C}_{d11}\mathbf{k} = \mathbf{k}^T\tilde{C}_d\mathbf{k}$. Furthermore, $x_d = \mathbf{k}^T\mathbf{C}_{d12}\mathbf{C}_{d22}^+\mathbf{b}_2 = \mathbf{k}^T\mathbf{C}_{d12}\mathbf{b}_2\mathbf{b}_2^T\mathbf{C}_{d22}^+\mathbf{b}_2 = 0$. This completes the proof.

Proposition 3 *Consider an arbitrary design $d \in \Omega_{2,n,p}$. Assume $\mathbf{k} \in \mathbb{R}^4$ and $x \in \mathbb{R}$ are arbitrarily chosen, with the restriction that $\mathbf{k}^T\mathbf{1}_4 = 0$. We then get for the smallest nonzero eigenvalue of $\mathbf{C}_d$ that*

$$\lambda_3(\mathbf{C}_d) \leq n\frac{1}{\mathbf{k}^T\mathbf{k}}\max_{z\in Z_p}\{\mathbf{k}^T\mathbf{C}_{11}(z)\mathbf{k} - 2\mathbf{k}^T\mathbf{C}_{12}(z)\mathbf{b}_2 x + \mathbf{b}_2^T\mathbf{C}_{22}(z)\mathbf{b}_2 x^2\}.$$

Proof: It follows from Proposition 2 and Equation 2 that

$$\begin{aligned}\frac{1}{n}\mathbf{k}^T\tilde{\mathbf{C}}_d\mathbf{k} &\leq \sum_{z\in Z_p}\pi_d(z)\{\mathbf{k}^T\mathbf{C}_{11}(z)\mathbf{k}-2\mathbf{k}^T\mathbf{C}_{12}(z)\mathbf{b}_2x+\mathbf{b}_2^T\mathbf{C}_{22}(z)\mathbf{b}_2x^2\}\\ &\leq \max_{z\in Z_p}\{\mathbf{k}^T\mathbf{C}_{11}(z)\mathbf{k}-2\mathbf{k}^T\mathbf{C}_{12}(z)\mathbf{b}_2x+\mathbf{b}_2^T\mathbf{C}_{22}(z)\mathbf{b}_2x^2\}.\end{aligned}$$

Since

$$\lambda_3(\tilde{\mathbf{C}}_d)=\min_{\ell:\ell^T\mathbf{1}_4=0}\frac{1}{\ell^T\ell}\ell^T\tilde{\mathbf{C}}_d\ell$$

and since

$$\lambda_3(\mathbf{C}_d)\leq\lambda_3(\tilde{\mathbf{C}}_d),$$

the desired inequality follows.

We use another consequence of Propostion 1 to derive a bound for the trace of $\tilde{\mathbf{C}}_d$, where the trace is the sum of the elements on the main diagonal of the matrix.

Proposition 4 *Consider any matrix $\mathbf{X}\in\mathbf{R}^{2\times4}$ and define*

$$L=\max_{z\in Z_p} tr\left(\mathbf{C}_{11}(z)-2\mathbf{C}_{12}(z)\mathbf{X}+\mathbf{X}^T\mathbf{C}_{22}(z)\mathbf{X}\right),$$

where $tr(A)$ denotes the trace of a matrix A. Then for any design $d\in\Omega_{2,n,p}$ we have $tr(\mathbf{C}_d)\leq nL$.

Proof: Since $tr(\tilde{\mathbf{C}}_d)\geq tr(\mathbf{C}_d)$, it follows directly from Proposition 1 and Equation 2 that

$$\begin{aligned}tr(\mathbf{C}_d)/n &\leq tr\left(\sum_{z\in Z_p}\pi_d(z)\big(\mathbf{C}_{11}(z)-\mathbf{C}_{12}(z)\mathbf{X}-\mathbf{X}^T\mathbf{C}_{12}^T(z)+\mathbf{X}^T\mathbf{C}_{22}(z)\mathbf{X}\big)\right)\\ &\leq \max_{z\in Z_p}\Big(tr(\mathbf{C}_{11}(z))-2tr(\mathbf{C}_{12}(z)\mathbf{X})+tr(\mathbf{X}^T\mathbf{C}_{22}(z)\mathbf{X})\Big)=L.\end{aligned}$$

This completes the proof.

How do we choose $\mathbf{X}$ in Proposition 4? Assume that there is a design $d^*\in\Omega_{2,n,p}$, for which we hope that d^* maximizes the trace of $\mathbf{C}_d$. It follows from Proposition 1 that

$$\begin{aligned}tr(\mathbf{C}_{d^*}) &\leq tr(\tilde{\mathbf{C}}_{d^*})\\ &= tr\left(\mathbf{C}_{d^*11}-\mathbf{C}_{d^*12}\mathbf{X}_{d^*}-\mathbf{X}_{d^*}^T\mathbf{C}_{d^*12}^T+\mathbf{X}_{d^*}^T\mathbf{C}_{d^*22}\mathbf{X}_{d^*}\right).\end{aligned}$$

As an immediate consequence of Proposition 4, we therefore get a sufficient condition that $tr(\mathbf{C}_{d^*})=\max_{d\in\Omega_{2,n,p}} tr(\mathbf{C}_d)$.

Proposition 5 *Assume* $d^* \in \Omega_{2,n,p}$ *is such that for every sequence* $z \in Z_p$ *we have*

$$tr\left(\mathbf{C}_{11}(z) - 2\mathbf{C}_{12}(z)\mathbf{X}_{d^*} + \mathbf{X}_{d^*}^T\mathbf{C}_{22}(z)\mathbf{X}_{d^*}\right) \le tr\left(\mathbf{C}_{d^*}\right)/n,$$

where $\mathbf{X}_{d^*} = \mathbf{C}_{d^*22}^{+}\mathbf{C}_{d^*12}^{T}$, *as in Proposition 1. Then* $tr(\mathbf{C}_{d^*}) = \max_{d\in\Omega_{2,n,p}} tr(\mathbf{C}_d)$.

Proof: In Proposition 4 choose $\mathbf{X} = \mathbf{X}_{d^*}$. Then the conditions of Proposition 5 imply that $L \le tr(\mathbf{C}_{d^*})/n$. For any design $d \in \Omega_{2,n,p}$ it then follows from Proposition 4 that $tr(\mathbf{C}_d) \le tr(\mathbf{C}_{d^*})$.

For any sequence $z \in Z_p$, there is a dual sequence $\bar{z} \in Z_p$, where each T in z is replaced by an R in $\bar{z}$ and vice versa. A design $d \in \Omega_{2,n,p}$ is called *dual balanced* if $\pi_d(z) = \pi_d(\bar{z})$ for each pair of dual sequences z and $\bar{z}$ in Z_p.

Proposition 6 *If we allow for approximate designs, then for each design* $d \in \Omega_{2,n,p}$ *there is a dual balanced design* $f \in \Omega_{2,n,p}$, *such that*

$$\tilde{\varphi}_A(f) \ge \tilde{\varphi}_A(d).$$

Proof: The design d has weights $\pi_d(z)$, $z \in Z_p$. Consider the dual design $\bar{d} \in \Omega_{2,n,p}$ with weights $\pi_{\bar{d}}(z)$, $z \in Z_p$, where for each $z \in Z_p$ the dual design $\bar{d}$ allots the weight that d has allotted to the dual sequence $\bar{z}$, i.e. $\pi_{\bar{d}}(z) = \pi_d(\bar{z})$. If we define

$$\mathbf{H}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

then $\mathbf{S}_{\bar{d}} = \mathbf{S}_d\mathbf{H}_2$, $\mathbf{M}_{\bar{d}} = \mathbf{M}_d\mathbf{H}_2$ and $\mathbf{T}_{\bar{d}} = \mathbf{T}_d\mathbf{H}_2$. Therefore,

$$\mathbf{C}_{\bar{d}11} = \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix} \mathbf{C}_{d11} \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix}, \mathbf{C}_{\bar{d}12} = \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix} \mathbf{C}_{d12}\mathbf{H}_2$$

and

$$\mathbf{C}_{\bar{d}22} = \mathbf{H}_2\mathbf{C}_{d22}\mathbf{H}_2.$$

This implies that

$$\tilde{\mathbf{C}}_{\bar{d}} = \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix} \mathbf{C}_{d11} \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix}$$

$$
- \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix} \mathbf{C}_{d12}\mathbf{H}_2(\mathbf{H}_2\mathbf{C}_{d22}^{+}\mathbf{H}_2)\mathbf{H}_2\mathbf{C}_{d12}^{T} \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix}
$$

$$
= \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix} \tilde{\mathbf{C}}_d \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix}.
$$

It follows that $\tilde{C}_{\bar{d}}$ has the same eigenvalues as $\tilde{C}_d$ and, consequently, that $\tilde{\varphi}_A(\bar{d}) = \tilde{\varphi}_A(d)$.

Now consider the dual balanced design f which allots to each sequence z the weight $\pi_f(z) = \frac{1}{2}\pi_d(z) + \frac{1}{2}\pi_{\bar{d}}(z)$. It then follows from Proposition 1 of Kunert and Martin (2000) that

$$
\tilde{\mathbf{C}}_f \geq \frac{1}{2}\tilde{\mathbf{C}}_d + \frac{1}{2}\tilde{\mathbf{C}}_{\bar{d}},
$$

which implies that

$$
\tilde{\varphi}_A(f) \geq \frac{1}{2}\tilde{\varphi}_A(d) + \frac{1}{2}\tilde{\varphi}_A(\bar{d}) = \tilde{\varphi}_A(d),
$$

since the A-criterion is concave and increasing.

4 Some efficient designs

For a given sequence $z \in Z_p$, it is possible to explicitly give the entries of $\mathbf{C}_{ij}(z)$. We define n_R and n_T as the number of appearances of treatment R and T in z. Let m_{RT}, m_{TR} be the number of appearances of the mixed carryover effects of R and T, respectively, and s_{RR}, s_{TT} the number of appearances of the self carryover effects of R and T in z. Then

$$
\begin{aligned}
\mathbf{S}_z^T\omega^{\perp}(\mathbf{1}_p)\mathbf{S}_z &= \begin{pmatrix} s_{RR} & 0 \\ 0 & s_{TT} \end{pmatrix} - \frac{1}{p}\begin{pmatrix} s_{RR}^2 & s_{RR}s_{TT} \\ s_{RR}s_{TT} & s_{TT}^2 \end{pmatrix}, \\
\mathbf{S}_z^T\omega^{\perp}(\mathbf{1}_p)\mathbf{M}_z &= -\frac{1}{p}\begin{pmatrix} m_{RT}s_{RR} & m_{RT}s_{TT} \\ m_{TR}s_{RR} & m_{TR}s_{TT} \end{pmatrix}, \\
\mathbf{S}_z^T\omega^{\perp}(\mathbf{1}_p)\mathbf{T}_z &= \begin{pmatrix} s_{RR} & 0 \\ 0 & s_{TT} \end{pmatrix} - \frac{1}{p}\begin{pmatrix} s_{RR}n_R & s_{RR}n_T \\ s_{TT}n_R & s_{TT}n_T \end{pmatrix} \\
&= \frac{1}{p}\begin{pmatrix} s_{RR}n_T & -s_{RR}n_T \\ -s_{TT}n_R & s_{TT}n_R \end{pmatrix},
\end{aligned}
$$

where we have used that $n_R + n_T = p$. Similarly,

$$\begin{aligned}
\mathbf{M}_z^T \omega^{\perp}(\mathbf{1}_p)\mathbf{M}_z &= \begin{pmatrix} m_{RT} & 0 \\ 0 & m_{TR} \end{pmatrix} - \frac{1}{p}\begin{pmatrix} m_{RT}^2 & m_{RT}m_{TR} \\ m_{RT}m_{TR} & m_{TR}^2 \end{pmatrix}, \\
\mathbf{M}_z^T \omega^{\perp}(\mathbf{1}_p)\mathbf{T}_z &= \begin{pmatrix} 0 & m_{RT} \\ m_{TR} & 0 \end{pmatrix} - \frac{1}{p}\begin{pmatrix} m_{RT}n_R & m_{RT}n_T \\ m_{TR}n_R & m_{TR}n_T \end{pmatrix} \\
&= \frac{1}{p}\begin{pmatrix} -m_{RT}n_R & m_{RT}n_R \\ m_{TR}n_T & -m_{TR}n_T \end{pmatrix}.
\end{aligned}$$

These (2×2)- matrices then can be used to determine the (4×4)-matrix $\mathbf{C}_{11}(z)$ and the (4×2)-matrix $\mathbf{C}_{12}(z)$. The matrix $\mathbf{C}_{22}(z)$ is given by:

$$\begin{aligned}
\mathbf{C}_{22}(z) &= \mathbf{T}_z^T \omega^{\perp}(\mathbf{1}_p)\mathbf{T}_z \\
&= \begin{pmatrix} n_R & 0 \\ 0 & n_T \end{pmatrix} - \frac{1}{p}\begin{pmatrix} n_R^2 & n_T n_R \\ n_T n_R & n_T^2 \end{pmatrix} \\
&= \begin{pmatrix} n_R(1 - \frac{1}{p}n_R) & -\frac{1}{p}n_T n_R \\ -\frac{1}{p}n_T n_R & n_T(1 - \frac{1}{p}n_T) \end{pmatrix} \\
&= \frac{1}{p}n_T n_R \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \\
&= \frac{2}{p}n_T n_R \mathbf{B}_2.
\end{aligned}$$

The fact that $\mathbf{C}_{22}(z)$ is proportional to $\mathbf{B}_2$ for any z implies that for any design d there is a c such that $\mathbf{C}_{d22} = c\mathbf{B}_2$. This is important because we need to determine a g-inverse of $\mathbf{C}_{d22}$. Because of the proportionality the g-inverse is given by

$$\mathbf{C}_{d22}^{+} = \frac{1}{c}\mathbf{B}_2.$$

4.1 Designs for p=3

If $p = 3$, there are 8 possible sequences (see Table 1). We try to find an approximate design d which maximizes $\tilde{\varphi}_A$. Note that $\tilde{\varphi}_A(d)$ is uniquely determined by the 8 proportions

Table 1: Possible sequences with three periods ($p = 3$). It should be noted that z_1 and z_2, z_3 and z_4, z_5 and z_6, z_7 and z_8 are pairs of dual sequences.

	Sequence	m_{TR}	m_{RT}	s_{RR}	s_{TT}	n_R	n_T
z_1	TTT	0	0	0	2	0	3
z_2	RRR	0	0	2	0	3	0
z_3	RTT	0	1	0	1	1	2
z_4	TRR	1	0	1	0	2	1
z_5	RTR	1	1	0	0	2	1
z_6	TRT	1	1	0	0	1	2
z_7	RRT	0	1	1	0	2	1
z_8	TTR	1	0	0	1	1	2

$\pi_d(z)$, $z \in Z_3$. We conclude from Proposition 6 that a best design is among the dual-balanced designs, i.e., $\pi_d(z_1) = \pi_d(z_2) = p_1$, say, $\pi_d(z_3) = \pi_d(z_4) = p_3$, $\pi_d(z_5) = \pi_d(z_6) = p_5$ and $\pi_d(z_7) = \pi_d(z_8) = p_7$. With this restriction, we get

$$\mathbf{S}_d^T \omega^\perp(\mathbf{U})\mathbf{S}_d/n = \sum_z \pi_d(z)\mathbf{S}_z^T \omega^\perp(\mathbf{1}_3)\mathbf{S}_z$$

$$= \begin{pmatrix} \frac{2}{3}(p_1 + p_3 + p_7) & 0 \\ 0 & \frac{2}{3}(p_1 + p_3 + p_7) \end{pmatrix} = \frac{2}{3}(p_1 + p_3 + p_7)\mathbf{I}_2,$$

$$\mathbf{S}_d^T \omega^\perp(\mathbf{U})\mathbf{M}_d/n = \sum_z \pi_d(z)\mathbf{S}_z^T \omega^\perp(\mathbf{1}_3)\mathbf{M}_z$$

$$= \begin{pmatrix} -1/3p_7 & -1/3p_3 \\ -1/3p_3 & -1/3p_7 \end{pmatrix},$$

$$\mathbf{S}_d^T \omega^\perp(\mathbf{U})\mathbf{T}_d/n = \sum_z \pi_d(z)\mathbf{S}_z^T \omega^\perp(\mathbf{1}_3)\mathbf{T}_z$$

$$= \begin{pmatrix} \frac{1}{3}p_3 + \frac{1}{3}p_7 & -\frac{1}{3}p_3 - \frac{1}{3}p_7 \\ -\frac{1}{3}p_3 - \frac{1}{3}p_7 & \frac{1}{3}p_3 + \frac{1}{3}p_7 \end{pmatrix} = (\frac{2}{3}p_3 + \frac{2}{3}p_7)\mathbf{B}_2,$$

$$\mathbf{M}_d^T \omega^\perp(\mathbf{U})\mathbf{M}_d/n = \sum_z \pi_d(z)\mathbf{M}_z^T \omega^\perp(\mathbf{1}_3)\mathbf{M}_z$$

$$= \begin{pmatrix} \frac{2}{3}(p_3+p_7)+\frac{4}{3}p_5 & -\frac{2}{3}p_5 \\ -\frac{2}{3}p_5 & \frac{2}{3}(p_3+p_7)+\frac{4}{3}p_5 \end{pmatrix},$$

$$\mathbf{M}_d^T \omega^\perp(\mathbf{U})\mathbf{T}_d/n = \sum_z \pi_d(z)\mathbf{M}_z^T \omega^\perp(\mathbf{1}_3)\mathbf{T}_z$$

$$= -(\frac{2}{3}p_3 + \frac{4}{3}p_7 + 2p_5)\mathbf{B}_2,$$

and

$$\mathbf{T}_d^T \omega^\perp(\mathbf{U})\mathbf{T}_d/n = \sum_z \pi_d(z)\mathbf{T}_z^T \omega^\perp(\mathbf{1}_3)\mathbf{T}_z$$

$$= \frac{8}{3}\left(p_1 \cdot 0 + p_3 + p_5 + p_7\right)\mathbf{B}_2.$$

Combining these results, we have

$$\frac{1}{n}\tilde{C}_d = \mathbf{B}_4 \begin{pmatrix} a & b & e & f \\ b & a & f & e \\ e & f & c & d \\ f & e & d & c \end{pmatrix} \mathbf{B}_4,$$

where

$$a = \frac{2}{3}(p_1+p_3+p_7) - \frac{p_3^2+p_7^2+2p_3p_7}{12(p_3+p_5+p_7)},$$

$$b = \frac{p_3^2+p_7^2+2p_3p_7}{12(p_3+p_5+p_7)},$$

$$c = \frac{2}{3}(p_3+p_7) + \frac{4}{3}p_5 - \frac{p_3^2+4p_7^2+9p_5^2+4p_3p_7+6p_3p_5+12p_7p_5}{12(p_3+p_5+p_7)},$$

$$d = -\frac{2}{3}p_5 + \frac{p_3^2+4p_7^2+9p_5^2+4p_3p_7+6p_3p_5+12p_7p_5}{12(p_3+p_5+p_7)},$$

$$e = -1/3p_7 + \frac{p_3^2+3p_7p_3+3p_3p_5+2p_7^2+3p_7p_5}{12(p_3+p_5+p_7)},$$

$$f = -\frac{1}{3}p_3 - \frac{p_3^2+3p_7p_3+3p_3p_5+2p_7^2+3p_7p_5}{12(p_3+p_5+p_7)}.$$

This matrix has eigenvalues

$$\mu_1 = \frac{a-b+c-d}{2} + \sqrt{(e-f)^2 + \left(\frac{c-d-a+b}{2}\right)^2},$$

$$\mu_2 = \frac{a-b+c-d}{2} - \sqrt{(e-f)^2 + \left(\frac{c-d-a+b}{2}\right)^2},$$

$$\mu_3 = \frac{a+b+c+d}{2} - e - f$$

and $\mu_4 = 0$. With a numerical search, we found that $\tilde{\varphi}_A(d) = 1/(1/\mu_1 + 1/\mu_2 + 1/\mu_3)$ is maximized, if

$$p_1 = 0.0951,\ p_3 = 0.1033,\ p_5 = 0.1684 \text{ and } p_7 = 0.1332.$$

The $\tilde{\varphi}_A$-criterion for a design $\tilde{d}$ with these proportions is $\tilde{\varphi}_A(\tilde{d}) = 0.0636n$. However, there are two problems with $\tilde{d}$. Firstly, it takes a very large number of experimental subjects to construct an exact design with these proportions. Secondly, the true A-criterion of $\tilde{d}$ is less than the bound: $\varphi_A(\tilde{d}) < \tilde{\varphi}_A(\tilde{d})$. This is because $\tilde{d}$ does not fulfill Equation 1.

A sufficient condition to fulfill Equation 1 is as follows.

Assume the design d is such that in all periods both direct effects appear in exactly half of the subjects, and that in each of the periods $2, \ldots, p$ each of the four carryover effects appears in exactly one quarter of the subjects. This implies that

$$\sum_z \pi_d(z)\mathbf{T}_z^T = \frac{1}{2}\mathbf{1}_2\mathbf{1}_p^T,$$

and

$$\sum_z \pi_d(z)[\mathbf{S}_z,\ \ \mathbf{M}_z]^T = \mathbf{1}_4 \left[\begin{array}{ccccc} 0 & \frac{1}{4} & \frac{1}{4} & \ldots & \frac{1}{4} \end{array}\right].$$

Now note that

$$\begin{aligned} & \mathbf{B}_4[\mathbf{S}_d,\ \mathbf{M}_d]^T\omega^\perp([\mathbf{U}, \mathbf{T}_d])\mathbf{P} = \Big(\mathbf{B}_4[\mathbf{S}_d,\ \mathbf{M}_d]^T - \mathbf{C}_{d12}\mathbf{C}_{d22}^+\mathbf{T}_d^T\Big)\omega^\perp(\mathbf{U})\mathbf{P} \\ = & \Big(\mathbf{B}_4[\mathbf{S}_d,\ \mathbf{M}_d]^T - \mathbf{C}_{d12}\mathbf{C}_{d22}^+\mathbf{T}_d^T\Big)(\mathbf{1}_n \otimes \omega^\perp(\mathbf{1}_p)) \\ = & \ n\left(\mathbf{B}_4 \sum_z \pi_d(z)[\mathbf{S}_z,\ \ \mathbf{M}_z]^T - \mathbf{C}_{d12}\mathbf{C}_{d22}^+ \sum_z \pi_d(z)\mathbf{T}_z^T\right)\omega^\perp(\mathbf{1}_p). \end{aligned}$$

This implies for our d that

$$\mathbf{B}_4[\mathbf{S}_d,\ \mathbf{M}_d]^T\omega^\perp([\mathbf{U}, \mathbf{T}_d])\mathbf{P} = 0$$

and, therefore, $\mathbf{C}_d = \tilde{\mathbf{C}}_d$.

It is clear that the design $\tilde{d}$ does not fulfill the sufficient condition: In each period, the number of subjects receiving a mixed carryover is larger than the number of subjects receiving a self carryover. The difference is larger in period 3 than in period 2.

If, instead of the design $\tilde{d}$, we use an exact design d^* which allots $\pi_{d^*}(z) = \frac{1}{8}$ to all sequences in Z_3, we get $\tilde{\varphi}_A(d^*) = 0.0628n$. It is easy to verify that this design d^* fulfills the sufficient conditions for Equation 1. Hence, $\varphi_A(d^*) = 0.0628n$ and it comes very close to the numerically derived upper bound for the A-criterion ($0.0636n$).

4.2 Designs for p=1 mod 4

We consider the case that $p = 4\ell+1$, where ℓ is a natural number, and that n is divisible by four. For this case, consider the exact design $d^* \in \Omega_{2,n,p}$, where one quarter of the subjects receive the sequence

$$z_1 = [\ R \quad T \quad T \quad R \quad R \quad \ldots \quad T \quad T \quad R \quad R\ ],$$

one quarter of the subjects receive the dual sequence

$$\bar{z}_1 = [\ T \quad R \quad R \quad T \quad T \quad \ldots \quad R \quad R \quad T \quad T\ ],$$

the third quarter of the subjects receive the sequence

$$z_2 = [\ R \quad R \quad T \quad T \quad R \quad R \quad \ldots \quad T \quad T \quad R\ ],$$

and the final quarter of the subjects receive the dual sequence

$$\bar{z}_2 = [\ T \quad T \quad R \quad R \quad T \quad T \quad \ldots \quad R \quad R \quad T\ ].$$

These designs d^* have some appeal for practice: while the design is not too complicated from an organizational viewpoint, it is not too obvious for subjects when they receive the same treatment as in the period before and when a different treatment.

For z_1 we get $n_R = 2\ell + 1$ and $n_T = 2\ell$, while $m_{RT} = m_{TR} = s_{RR} = s_{TT} = \ell$. This implies that

$$\begin{aligned} \mathbf{S}_{z_1}^T \omega^{\perp}(\mathbf{1}_p)\mathbf{S}_{z_1} &= \begin{pmatrix} \ell & 0 \\ 0 & \ell \end{pmatrix} - \frac{1}{p}\begin{pmatrix} \ell^2 & \ell^2 \\ \ell^2 & \ell^2 \end{pmatrix}, \\ \mathbf{S}_{z_1}^T \omega^{\perp}(\mathbf{1}_p)\mathbf{M}_{z_1} &= -\frac{1}{p}\begin{pmatrix} \ell^2 & \ell^2 \\ \ell^2 & \ell^2 \end{pmatrix}, \end{aligned}$$

$$
\begin{aligned}
\mathbf{S}_{z_1}^T \omega^{\perp}(\mathbf{1}_p)\mathbf{T}_{z_1} &= \frac{1}{p}\begin{pmatrix} (2\ell+1)\ell & -(2\ell+1)\ell \\ -2\ell^2 & 2\ell^2 \end{pmatrix}, \\
\mathbf{M}_{z_1}^T \omega^{\perp}(\mathbf{1}_p)\mathbf{M}_{z_1} &= \begin{pmatrix} \ell & 0 \\ 0 & \ell \end{pmatrix} - \frac{1}{p}\begin{pmatrix} \ell^2 & \ell^2 \\ \ell^2 & \ell^2 \end{pmatrix},
\end{aligned}
$$

and

$$
\mathbf{M}_{z_1}^T \omega^{\perp}(\mathbf{1}_p)\mathbf{T}_{z_1} = \frac{1}{p}\begin{pmatrix} -2\ell^2 & 2\ell^2 \\ (2\ell+1)\ell & -(2\ell+1)\ell \end{pmatrix}.
$$

With straightforward algebra we get

$$
\mathbf{C}_{11}(z_1) = \ell\mathbf{B}_4,\ \mathbf{C}_{12}(z_1) = \frac{1}{2}\begin{pmatrix} \ell & -\ell \\ -\ell & \ell \\ -\ell & \ell \\ \ell & -\ell \end{pmatrix}, \text{ and } \mathbf{C}_{22}(z_1) = \frac{4\ell(2\ell+1)}{p}\mathbf{B}_2.
$$

Since the sequence z_2 has the same parameters, $n_R = 2\ell + 1$, $n_T = 2\ell$ and $m_{RT} = m_{TR} = s_{RR} = s_{TT} = \ell$, we find that all $\mathbf{C}_{ij}(z_2) = \mathbf{C}_{ij}(z_1)$.

For the dual sequences $\bar{z}_1$ and $\bar{z}_2$, the roles of R and T are interchanged. Hence $n_T = 2\ell + 1$ and $n_T = 2\ell$, but we also have $m_{RT} = m_{TR} = s_{RR} = s_{TT} = \ell$. It then is easy to see that, again, $\mathbf{C}_{11}(\bar{z}_i) = \mathbf{C}_{11}(z_1)$, $\mathbf{C}_{12}(\bar{z}_i) = \mathbf{C}_{12}(z_1)$ and $\mathbf{C}_{22}(\bar{z}_i) = \mathbf{C}_{22}(z_1)$, for $i = 1, 2$. This implies for the design d^* that

$$
\mathbf{C}_{d^*11} = n\ell\mathbf{B}_4,\ \mathbf{C}_{d^*12} = \frac{n}{2}\begin{pmatrix} \ell & -\ell \\ -\ell & \ell \\ -\ell & \ell \\ \ell & -\ell \end{pmatrix} \text{ and } \mathbf{C}_{d^*22} = \frac{4\ell(2\ell+1)n}{p}\mathbf{B}_2.
$$

We therefore have

$$
\tilde{\mathbf{C}}_{d^*} = n\ell\mathbf{B}_4 - \frac{n\ell p}{8(2\ell+1)}\begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}
$$

$$= \frac{n(p-1)}{16(p+1)} \begin{pmatrix} 2p+3 & -1 & -1 & -2p-1 \\ -1 & 2p+3 & -2p-1 & -1 \\ -1 & -2p-1 & 2p+3 & -1 \\ -2p-1 & -1 & -1 & 2p+3 \end{pmatrix}.$$

To show that $\mathbf{C}_{d^*} = \tilde{\mathbf{C}}_{d^*}$, we verify that in each period the direct effect of each treatment appears in exactly two of the sequences, while in each of the periods $2, \ldots, p$ each of the four carryover effects appears in exactly one of the four sequences.

This implies that Equation 1 holds, and, therefore, $\mathbf{C}_{d^*} = \tilde{\mathbf{C}}_{d^*}$. The eigenvalues of $\mathbf{C}_{d^*}$ are $\mu_1(d^*) = \mu_2(d^*) = n\frac{p-1}{4}$, $\mu_3(d^*) = n\frac{p-1}{4(p+1)}$ and 0. The eigenvector corresponding to $\mu_3(d^*)$ is $\mathbf{k}_3 = \frac{1}{2}[\ 1 \quad -1 \quad -1 \quad 1\ ]^T$.

The A-criterion of the design d^* therefore is

$$\varphi_A(d^*) = n\frac{p-1}{4+4+4(p+1)} = n\frac{p-1}{4(p+3)}.$$

Note that this cannot be larger than $\frac{n}{4}$. Even if the number of periods p goes to ∞, we have $\varphi_A(d^*) \to \frac{n}{4}$. We will show in the next section that no other design can perform much better. This is similar to what happens for the estimation of the direct effects in our model (see Kunert and Stufken 2008): a large number of periods is only of limited use.

5 An upper bound for the A-criterion

In this section we derive an upper bound for the A-criterion for an arbitrary p. We will find that the class of designs d^* derived in Section 4.2 is highly efficient. We begin with a technical lemma.

Proposition 7 *Consider an arbitrary sequence $z \in Z_p$, starting with R. Then the design-matrices $\mathbf{T}_z$, $\mathbf{S}_z$ and $\mathbf{M}_z$ fulfill the equality*

$$\mathbf{S}_z + \mathbf{M}_z\mathbf{H}_2 - \mathbf{T}_z = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix},$$

where H_2 is as in the proof of Proposition 6.

Proof The first row of $\mathbf{T}_z$ equals $[1\,0]$, since the sequence starts with R. The first rows of both $\mathbf{S}_z$ and $\mathbf{M}_z$ are $[0\,0]$. Therefore, the first two of $\mathbf{S}_z + \mathbf{M}_z\mathbf{H}_2 - \mathbf{T}_z$ must equal $[-1\,0]$.

Now consider the i-th row, for $i \geq 2$.

Case 1: The preceding treatment was R, the current treatment is R. Then

the i-th row of $\mathbf{S}_z$ is $[1\,0]$,

the i-th row of $\mathbf{M}_z$ is $[0\,0]$,

and

the i-th two of $\mathbf{T}_z$ is $[1\,0]$.

Hence, the i-th row of $\mathbf{S}_z + \mathbf{M}_z\mathbf{H}_2 - \mathbf{T}_z$ equals $[0\,0]$.

Case 2: The preceding treatment was R, the current treatment is T. Then

the i-th row of $\mathbf{S}_z$ is $[0\,0]$,

the i-th row of $\mathbf{M}_z$ is $[1\,0]$,

and

the i-th two of $\mathbf{T}_z$ is $[0\,1]$.

Hence, the i-th row of $\mathbf{M}_z\mathbf{H}_2$ is $[0\,1]$ and the i-th row of $\mathbf{S}_z + \mathbf{M}_z\mathbf{H}_2 - \mathbf{T}_z$ equals $[0\,0]$.

Case 3: The preceding treatment was T, the current treatment is R. Then

the i-th row of $\mathbf{S}_z$ is $[0\,0]$,

the i-th row of $\mathbf{M}_z$ is $[0\,1]$,

and

the i-th two of $\mathbf{T}_z$ is $[1\,0]$.

Hence, the i-th row of $\mathbf{M}_z\mathbf{H}_2$ is $[1\,0]$ and the i-th row of $\mathbf{S}_z + \mathbf{M}_z\mathbf{H}_2 - \mathbf{T}_z$ equals $[0\,0]$.

Case 4: The preceding treatment was T, the current treatment is T. Then

the i-th row of $\mathbf{S}_z$ is $[0\,1]$,

the i-th row of $\mathbf{M}_z$ is $[0\,0]$,

and

the i-th two of $\mathbf{T}_z$ is $[0\,1]$.

Hence, the i-th row of $\mathbf{S}_z + \mathbf{M}_z\mathbf{H}_2 - \mathbf{T}_z$ equals $[0\,0]$.

This completes the proof.

Proposition 8 *Consider an arbitrary sequence* $z \in Z_p$*, choose* $\mathbf{k} = \frac{1}{2}[1, -1, -1, 1]^T$ *and* *choose* $x = \frac{1}{\sqrt{2}}$. *Then*

$$\mathbf{k}^T\mathbf{C}_{11}(z)\mathbf{k} - 2\mathbf{k}^T\mathbf{C}_{12}(z)\mathbf{b}_2 x + \mathbf{b}_2^T\mathbf{C}_{22}(z)\mathbf{b}_2 x^2 \le \frac{p-1}{4p}$$

Proof: Observing that

$$\mathbf{k} = \frac{1}{\sqrt{2}}\begin{bmatrix} \mathbf{b}_2 \\ -\mathbf{b}_2 \end{bmatrix}$$

we get

$$\begin{aligned} \mathbf{k}^T\mathbf{C}_{11}(z)\mathbf{k} &= \left[\mathbf{b}_2^T/\sqrt(2), -\mathbf{b}_2^T/\sqrt{2}\right]\mathbf{B}_4[\mathbf{S}_z, \mathbf{M}_z]^T\omega^\perp(\mathbf{1}_p)[\mathbf{S}_z, \mathbf{M}_z]\mathbf{B}_4\begin{bmatrix} \mathbf{b}_2/\sqrt{2} \\ -\mathbf{b}_2/\sqrt{2} \end{bmatrix} \\ &= \frac{1}{2}(\mathbf{b}_2^T\mathbf{S}_z^T - \mathbf{b}_2^T\mathbf{M}_z^T)\omega^\perp(\mathbf{1}_p)(\mathbf{S}_z\mathbf{b}_2 - \mathbf{M}_z\mathbf{b}_2). \end{aligned}$$

and

$$\mathbf{k}^T\mathbf{C}_{12}(z)\mathbf{b}_2 = \frac{1}{\sqrt{2}}(\mathbf{b}_2^T\mathbf{S}_z^T - \mathbf{b}_2^T\mathbf{M}_z^T)\omega^\perp(\mathbf{1}_p)\mathbf{T}_z\mathbf{b}_2.$$

Therefore,

$$\begin{aligned} &\mathbf{k}^T\mathbf{C}_{11}(z)\mathbf{k} - 2\mathbf{k}^T\mathbf{C}_{12}(z)\mathbf{b}_2 x + \mathbf{b}_2^T\mathbf{C}_{22}(z)\mathbf{b}_2 x^2 \\ = &\frac{1}{2}(\mathbf{b}_2^T\mathbf{S}_z^T - \mathbf{b}_2^T\mathbf{M}_z^T)\omega^\perp(\mathbf{1}_p)(\mathbf{S}_z\mathbf{b}_2 - \mathbf{M}_z\mathbf{b}_2) - 2\frac{1}{\sqrt{2}}(\mathbf{b}_2^T\mathbf{S}_z^T - \mathbf{b}_2^T\mathbf{M}_z^T)\omega^\perp(\mathbf{1}_p)\mathbf{T}_z\mathbf{b}_2 x \\ &+\mathbf{b}_2^T\mathbf{T}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{T}_z\mathbf{1}_2 x^2 \\ = &\frac{1}{2}\Big((\mathbf{b}_2^T\mathbf{S}_z^T - \mathbf{b}_2^T\mathbf{M}_z^T - \mathbf{b}_2\mathbf{T}_z^T)\omega^\perp(\mathbf{1}_p)(\mathbf{S}_z\mathbf{b}_2 - \mathbf{M}_z\mathbf{b}_2) \\ &-(\mathbf{b}_2^T\mathbf{S}_z^T - \mathbf{b}_2^T\mathbf{M}_z^T - \mathbf{b}_2\mathbf{T}_z^T)\omega^\perp(\mathbf{1}_p)\mathbf{T}_z\mathbf{b}_2\Big) \\ = &\frac{1}{2}\Big(\mathbf{b}_2^T(\mathbf{S}_z^T - \mathbf{M}_z^T - \mathbf{T}_z^T)\omega^\perp(\mathbf{1}_p)(\mathbf{S}_z - \mathbf{M}_z - \mathbf{T}_z)\mathbf{b}_2\Big). \end{aligned}$$

If $\mathbf{H}_2$ is as in Proposition 7, then $\mathbf{H}_2\mathbf{b}_2 = -\mathbf{b}_2$. Hence,

$$(\mathbf{S}_z - \mathbf{M}_z - \mathbf{T}_z)\mathbf{b}_2 = (\mathbf{S}_z + \mathbf{M}_z\mathbf{H}_2 - \mathbf{T}_z)\mathbf{b}_2.$$

Therefore, if the sequence z starts with R, it follows from Proposition 7 that $(\mathbf{S}_z - \mathbf{M}_z - \mathbf{T}_z)\mathbf{b}_2 = [\frac{1}{\sqrt{2}}\, 0 \ldots 0]^T$. If z starts with T, we get with the same methods that $(\mathbf{S}_z - \mathbf{M}_z - \mathbf{T}_z)\mathbf{b}_2 = [-\frac{1}{\sqrt{2}}\, 0 \ldots 0]^T$.

This implies for any sequence z that

$$\mathbf{k}^T\mathbf{C}_{11}(z)\mathbf{k} - 2\mathbf{k}^T\mathbf{C}_{12}(z)\mathbf{b}_2 x + \mathbf{b}_2^T\mathbf{C}_{22}(z)\mathbf{b}_2 x^2$$

$$= \frac{1}{2} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \dots & 0 \end{bmatrix} \omega^{\perp}(\mathbf{1}_p) \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \frac{p-1}{4p}.$$

This completes the proof.

Now consider an arbitrary design $d \in \Omega_{2,n,p}$. Since $\mathbf{k}^T \mathbf{1}_4 = 0$ and $\mathbf{k}^T \mathbf{k} = 1$, it follows from Proposition 2 and Proposition 8 that

$$\lambda_3(d) \le n \max_{z \in Z_p} \{ k^T \mathbf{C}_{11}(z) k - 2k^T \mathbf{C}_{12}(z) \mathbf{b}_2 x + \mathbf{b}_2^T \mathbf{C}_{22}(z) \mathbf{b}_2 x^2 \} \le n \frac{p-1}{4p}. \quad (3)$$

Remember that $\mu_3(d^*) = n\frac{p-1}{4(p+1)}$., see Section 4.2. So the smallest nonzero eigenvalue of d^* is slightly less than this bound.

We now show that d^* maximizes the trace of the information matrix.

Proposition 9 *Choose* $\mathbf{X} = c \begin{bmatrix} \mathbf{B}_2 & -\mathbf{B}_2 \end{bmatrix}$, *where*

$$c = \frac{p}{2(p+1)}.$$

Consider an arbitrary sequence $z \in Z_p$ which starts with R. Then

$$tr \left(\mathbf{C}_{11}(z) - 2\mathbf{C}_{12}(z)\mathbf{X} + \mathbf{X}^T \mathbf{C}_{22}(z) \mathbf{X} \right) \le \frac{(2p+3)(p-1)}{4(p+1)}.$$

Proof: see Appendix.

As an immediate consequence of Proposition 9 we find that a sequence $z^* \in Z_p$ attains the bound for

$$G_z = tr(\mathbf{C}_{11}(z) - 2\mathbf{C}_{12}(z)\mathbf{X} + \mathbf{X}^T \mathbf{C}_{22}(z)\mathbf{X})$$

if and only if $m_{RT} = m_{TR}$ and $s_{TT} = s_{RR} = \frac{p-1}{4}$. Note that this is only possible if there is an $\ell \in \mathbb{N}$ such that $p = 4\ell + 1$. The two sequences z_1 and z_2 from Section 4.2 fulfill this condition.

For the dual $\bar{z}$ of a sequence z, observe as in the proof of Proposition 6 that

$$\mathbf{C}_{11}(\bar{z}) = \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix} \mathbf{C}_{11}(z) \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix}, \mathbf{C}_{12}(\bar{z}) = \begin{pmatrix} \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 \end{pmatrix} \mathbf{C}_{12}(z)\mathbf{H}_2,$$

and

$$\mathbf{C}_{22}(\bar{z}) = \mathbf{H}_2\mathbf{C}_{22}(z)\mathbf{H}_2.$$

The matrix $\mathbf{X}$ from Proposition 9 fulfills that

$$\mathbf{X}\begin{pmatrix}\mathbf{H}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_2\end{pmatrix} = \mathbf{H}_2\mathbf{X} = -\mathbf{X}.$$

So we get for a sequence z starting with T that

$$\begin{aligned}
& tr(\mathbf{C}_{11}(z)) - 2tr(\mathbf{C}_{12}(z)\mathbf{X}) + tr(\mathbf{X}^T\mathbf{C}_{22}(z)\mathbf{X}) \\
= \; & tr\left(\left(\begin{pmatrix}\mathbf{H}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_2\end{pmatrix}\mathbf{C}_{11}(\bar{z})\begin{pmatrix}\mathbf{H}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_2\end{pmatrix}\right)\right) - 2\,tr\left(\left(\begin{pmatrix}\mathbf{H}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_2\end{pmatrix}\mathbf{C}_{12}(\bar{z})\mathbf{H}_2\mathbf{X}\right)\right. \\
& + tr(\mathbf{X}^T\mathbf{H}_2\mathbf{C}_{22}(\bar{z})\mathbf{H}_2\mathbf{X}) \\
= \; & tr\left(\mathbf{C}_{11}(\bar{z})\begin{pmatrix}\mathbf{H}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_2\end{pmatrix}\begin{pmatrix}\mathbf{H}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_2\end{pmatrix}\right) - 2tr\left(\mathbf{C}_{12}(\bar{z})(-\mathbf{X})\begin{pmatrix}\mathbf{H}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_2\end{pmatrix}\right) \\
& + tr\left((-\mathbf{X}^T)\mathbf{C}_{22}(\bar{z})(-\mathbf{X})\right) \\
= \; & tr(\mathbf{C}_{11}(\bar{z})) - 2tr(\mathbf{C}_{12}(\bar{z})\mathbf{X}) + tr(\mathbf{X}^T\mathbf{C}_{22}(\bar{z})\mathbf{X}) \leq \frac{(2p+3)(p-1)}{4(p+1)},
\end{aligned}$$

because the sequence $\bar{z}$ starts with R and Proposition 9 can be applied. This implies that the bound from Proposition 9 also holds for sequences z starting with T.

We therefore have a bound for the trace of the information matrix of any design $d \in \Omega_{2,n,p}$, namely

$$tr(\mathbf{C}_d) \leq tr(\tilde{\mathbf{C}}_d) \leq n\sum_{z\in Z_p}\pi_d(z)tr\Big(\mathbf{C}_{11}(z) - 2\mathbf{C}_{12}(z)\mathbf{X} + \mathbf{X}^T\mathbf{C}_{22}(z)\mathbf{X}\Big) \leq n\frac{(2p+3)(p-1)}{4(p+1)}.$$

This helps to derive a bound for the A-criterion of any design.

Proposition 10 *Assume the design $d \in \Omega_{2,n,p}$ has an information matrix with eigenvectors $\mu_1 \geq \mu_2 \geq \mu_3$, with the side-conditions that*

$$\mu_1 + \mu_2 + \mu_3 \leq A, \text{ and } \mu_3 \leq q,$$

where $q \leq A/3$. Then we have for the A-criterion of the design that

$$\varphi_A(d) \leq \frac{q(A-q)}{A+3q}.$$

Proof: For given μ_2 and μ_3, we have that $\mu_1 \leq A - \mu_2 - \mu_3$. Hence,

$$\varphi_A(d) = \frac{1}{\frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3}} \leq \frac{1}{\frac{1}{A-\mu_2-\mu_3} + \frac{1}{\mu_2} + \frac{1}{\mu_3}}.$$

Holding μ_3 fixed, this bound is maximal if $\mu_2 = A - \mu_2 - \mu_3$, i.e. if $\mu_2 = (A - \mu_3)/2$. This implies that

$$\varphi_A(d) \leq \frac{1}{\frac{2}{A-\mu_3} + \frac{2}{A-\mu_3} + \frac{1}{\mu_3}}.$$

This bound, however, gets maximal if μ_3 gets as near to $(A - \mu_3)/2$ as possible, which means that $\mu_3 = q$. This gives

$$\varphi_A(d) \leq \frac{1}{\frac{2}{A-q} + \frac{2}{A-q} + \frac{1}{q}},$$

which completes the proof.

For any $d \in \Omega_{2,n,p}$ we have concluded from Proposition 9 that the sum of the three eigenvalues, i.e. the trace of $\mathbf{C}_d$, cannot be more than $A = n\frac{(2p+3)(p-1)}{4(p+1)}$, while we get from inequality (3) that $\mu_3(d) \leq n\frac{p-1}{4p}$. We therefore conclude from Proposition 10 that the A-criterion of the design fulfills

$$\varphi_A(d) \leq n\frac{(p-1)(2p^2+2p-1)}{4p(2p^2+6p+3)} = \varphi_A^*,$$

say.

Remember that the A-criterion of the design d^* from Section 4.2 is

$$\varphi_A(d^*) = n\frac{p-1}{4(p+3)}.$$

This means that the efficiency of the design d^* is at least

$$\frac{\varphi_A(d^*)}{\varphi_A^*} = \frac{2p^3+6p^2+3p}{2p^3+8p^2+5p-3},$$

which equals 0.88 for $p = 5$ and 0.92 for $p = 10$. If $p \to \infty$, the efficiency goes to 1.

Appendix: Proof of Proposition 9

Define

$$G_z = tr(\mathbf{C}_{11}(z)) - 2tr(\mathbf{C}_{12}(z)\mathbf{X}) + tr(\mathbf{X}^T\mathbf{C}_{22}(z)\mathbf{X}).$$

Then

$$\begin{aligned}
G_z &= tr\left(\mathbf{B}_4 \begin{bmatrix} \mathbf{S}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{S}_z & \mathbf{S}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{M}_z \\ \mathbf{M}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{S}_z & \mathbf{M}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{M}_z \end{bmatrix} \mathbf{B}_4\right) \\
&\quad -2c\, tr\left(\mathbf{B}_4 \begin{bmatrix} \mathbf{S}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{T}_z \\ \mathbf{M}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{T}_z \end{bmatrix} \begin{bmatrix} \mathbf{B}_2 & -\mathbf{B}_2 \end{bmatrix}\right) \\
&\quad +c^2 tr\left(\begin{bmatrix} \mathbf{B}_2 \\ -\mathbf{B}_2 \end{bmatrix} \mathbf{T}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{T}_z \begin{bmatrix} \mathbf{B}_2 & -\mathbf{B}_2 \end{bmatrix}\right) \\
&= tr(\mathbf{S}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{S}_z) + tr(\mathbf{M}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{M}_z) \\
&\quad -\frac{1}{4}\mathbf{1}_2^T\left(\mathbf{S}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{S}_z + \mathbf{S}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{M}_z + \mathbf{M}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{S}_z + \mathbf{M}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{M}_z\right)\mathbf{1}_2 \\
&\quad -2c\, tr(\mathbf{S}_2^T\omega^\perp(\mathbf{1}_p)\mathbf{T}_z\mathbf{B}_2) + 2c\, tr(\mathbf{M}_z^T\omega^\perp(\mathbf{1}_p))\mathbf{T}_z\mathbf{B}_2) \\
&\quad +2c^2\, tr(\mathbf{B}_2\mathbf{T}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{T}_z\mathbf{B}_2),
\end{aligned}$$

where we have used that $tr(\mathbf{A}_1\mathbf{A}_2) = tr(\mathbf{A}_2\mathbf{A}_1)$ and that $\begin{bmatrix} \mathbf{B}_2 & -\mathbf{B}_2 \end{bmatrix}\mathbf{B}_4 = \begin{bmatrix} \mathbf{B}_2 & -\mathbf{B}_2 \end{bmatrix}$.

We split G_z up into several parts. First of all,

$$\begin{aligned}
&\frac{1}{4}\mathbf{1}_2^T\left(\mathbf{S}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{S}_z + \mathbf{S}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{M}_z + \mathbf{M}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{S}_z + \mathbf{M}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{M}_z\right)\mathbf{1}_2 \\
=\ &\frac{1}{4}\mathbf{1}_2^T(\mathbf{S}_z^T + \mathbf{M}_z^T)\omega^\perp(\mathbf{1}_p)(\mathbf{S}_z + \mathbf{M}_z)\mathbf{1}_2 \\
=\ &\frac{1}{4}[\ 0 \quad 1 \quad \ldots \quad 1\ ]\omega^\perp(\mathbf{1}_p)\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \frac{p-1}{4p}.
\end{aligned}$$

Further, we have

$$\begin{aligned}
&tr(\mathbf{S}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{S}_z) - 2c\, tr(\mathbf{S}_2^T\omega^\perp(\mathbf{1}_p)\mathbf{T}_z\mathbf{B}_2) + c^2\, tr(\mathbf{B}_2\mathbf{T}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{T}_z\mathbf{B}_2) \\
=\ &tr\big((\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)^T\omega^\perp(\mathbf{1}_p)(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)\big) \\
=\ &tr\big((\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)^T(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)\big) - \frac{1}{p}\mathbf{1}_p^T(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)^T\mathbf{1}_p
\end{aligned}$$

and

$$\begin{aligned}
&tr(\mathbf{M}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{M}_z) + 2c\, tr(\mathbf{M}_z^T\omega^\perp(\mathbf{1}_p))\mathbf{T}_z\mathbf{B}_2) + c^2\, tr(\mathbf{B}_2\mathbf{T}_z^T\omega^\perp(\mathbf{1}_p)\mathbf{T}_z\mathbf{B}_2) \\
=\ &tr\big((\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)^T\omega^\perp(\mathbf{1}_p)(\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)\big)
\end{aligned}$$

$$= \; tr\big((\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)^T(\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)\big) - \frac{1}{p}\mathbf{1}_p^T(\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)(\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)^T\mathbf{1}_p.$$

Since z starts with R, the first row of $\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2$ equals $[\; -c/2 \quad c/2 \;]$. The i-th row $(i \geq 2)$ of $\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2$ equals

- $[\; 1 - c/2, \quad c/2 \;]$, if $z(i-1) = R$, $z(i) = R$,
- $[\; c/2, \quad -c/2 \;]$, if $z(i-1) = R$, $z(i) = T$,
- $[\; -c/2, \quad c/2 \;]$, if $z(i-1) = T$, $z(i) = R$,
- $[\; c/2, \quad 1 - c/2 \;]$, if $z(i-1) = T$, $z(i) = T$.

On the other hand, the first row of $\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2$ is $[\; c/2 \quad -c/2 \;]$. For $i \geq 2$ the i-th row of $\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2$ equals

- $[\; c/2, \quad -c/2 \;]$, if $z(i-1) = R$, $z(i) = R$,
- $[\; 1 - c/2, \quad c/2 \;]$, if $z(i-1) = R$, $z(i) = T$,
- $[\; c/2, \quad 1 - c/2 \;]$, if $z(i-1) = T$, $z(i) = R$,
- $[\; -c/2, \quad c/2 \;]$, if $z(i-1) = T$, $z(i) = T$.

This implies that we can rewrite

$$\begin{aligned}(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)^T(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2) \;&=\; \begin{bmatrix} \frac{c^2}{4} & -\frac{c^2}{4} \\ -\frac{c^2}{4} & \frac{c^2}{4} \end{bmatrix} + s_{RR}\begin{bmatrix} (1-\frac{c}{2})^2 & \frac{c}{2}(1-\frac{c}{2}) \\ \frac{c}{2}(1-\frac{c}{2}) & \frac{c^2}{4} \end{bmatrix} \\ &+ m_{RT}\begin{bmatrix} \frac{c^2}{4} & -\frac{c^2}{4} \\ -\frac{c^2}{4} & \frac{c^2}{4} \end{bmatrix} + m_{TR}\begin{bmatrix} \frac{c^2}{4} & -\frac{c^2}{4} \\ -\frac{c^2}{4} & \frac{c^2}{4} \end{bmatrix} \\ &+ s_{TT}\begin{bmatrix} \frac{c^2}{4} & \frac{c}{2}(1-\frac{c}{2}) \\ \frac{c}{2}(1-\frac{c}{2}) & (1-\frac{c}{2})^2 \end{bmatrix}\end{aligned}$$

and

$$tr\big((\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)^T(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)\big) = (1 + m_{RT} + m_{TR})\frac{c^2}{2} + \big(s_{RR} + s_{TT}\big)\Big(\frac{c^2}{4} + (1 - \frac{c}{2})^2\Big).$$

Similarly,

$$tr\big((\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)^T(\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)\big) = (1 + s_{RR} + s_{TT})\frac{c^2}{2} + \big(m_{RT} + m_{TR}\big)(\frac{c^2}{4} + (1 - \frac{c}{2})^2).$$

Noting that $s_{RR} + s_{TT} + m_{RT} + m_{TR} = p - 1$, we get

$$\begin{aligned}
& tr\big((\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)^T(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)\big) + tr\big((\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)^T(\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)\big) \\
= & (p+1)\frac{c^2}{2} + (p-1)(\frac{c^2}{4} + (1 - \frac{c}{2})^2) \\
= & \frac{3p^3 + 4p^2 - 2p - 4}{4(p+1)^2}.
\end{aligned}$$

Note that all terms considered so far turn out to be the same for any sequence $z \in Z_p$.

On the other hand

$$\begin{aligned}
& (\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)^T\mathbf{1}_p \\
= & \begin{bmatrix} -\frac{c}{2} \\ \frac{c}{2} \end{bmatrix} + s_{RR}\begin{bmatrix} 1 - \frac{c}{2} \\ \frac{c}{2} \end{bmatrix} + m_{RT}\begin{bmatrix} \frac{c}{2} \\ -\frac{c}{2} \end{bmatrix} + m_{TR}\begin{bmatrix} -\frac{c}{2} \\ \frac{c}{2} \end{bmatrix} + s_{TT}\begin{bmatrix} \frac{c}{2} \\ 1 - \frac{c}{2} \end{bmatrix} \\
= & \begin{bmatrix} s_{RR} - \frac{c}{2}\left(1 + s_{RR} - s_{TT} - (m_{RT} - m_{TR})\right) \\ s_{TT} + \frac{c}{2}\left(1 + s_{RR} - s_{TT} - (m_{RT} - m_{TR})\right) \end{bmatrix}
\end{aligned}$$

and, similarly,

$$\begin{aligned}
& (\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)^T\mathbf{1}_p \\
= & \begin{bmatrix} \frac{c}{2} \\ -\frac{c}{2} \end{bmatrix} + s_{RR}\begin{bmatrix} \frac{c}{2} \\ -\frac{c}{2} \end{bmatrix} + m_{RT}\begin{bmatrix} 1 - \frac{c}{2} \\ \frac{c}{2} \end{bmatrix} + m_{TR}\begin{bmatrix} \frac{c}{2} \\ 1 - \frac{c}{2} \end{bmatrix} + s_{TT}\begin{bmatrix} -\frac{c}{2} \\ \frac{c}{2} \end{bmatrix} \\
= & \begin{bmatrix} m_{RT} + \frac{c}{2}\left(1 + s_{RR} - s_{TT} - (m_{RT} - m_{TR})\right) \\ m_{TR} - \frac{c}{2}\left(1 + s_{RR} - s_{TT} - (m_{RT} - m_{TR})\right) \end{bmatrix}.
\end{aligned}$$

These formulas get slightly clearer if we reparametrize. Define $s = s_{TT}$, $d_s = s_{RR} - s_{TT}, m = m_{TR}$ and $d_m = m_{RT} - m_{TR}$. With this notation, we get

$$\begin{aligned}
& \mathbf{1}_p^T(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)(\mathbf{S}_z - c\mathbf{T}_z\mathbf{B}_2)^T\mathbf{1}_p + \mathbf{1}_p^T(\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)(\mathbf{M}_z + c\mathbf{T}_z\mathbf{B}_2)^T\mathbf{1}_p \\
= & (s + d_s - \frac{c}{2}(1 + d_s - d_m))^2 + (s + \frac{c}{2}(1 + d_s - d_m))^2 \\
& +(m - \frac{c}{2}(1 + d_s - d_m))^2 + (m + d_m + \frac{c}{2}(1 + d_s - d_m))^2 \\
= & F(s, d_s, d_m),
\end{aligned}$$

say. Note that m is not an independent parameter. Due to $s_{RR}+s_{TT}+m_{RT}+m_{TR} = p-1$, we have

$$m = \frac{p-1-2s-d_s-d_m}{2}.$$

We minimize F by choosing s, d_s and d_m. As a first step, we leave d_s and d_m constant and determine an s^* minimizing $F(., d_s, d_m)$.

The partial derivative of F with respect to s equals

$$\frac{\partial F}{\partial s} = 2(s + d_s + s - m - m - d_m) = 2(4s + 2d_s - (p-1)),$$

where we have used that the partial derivative of m with respect to s is -1. Hence, the derivative of F is zero if

$$s = \frac{p-1-2d_s}{4} = s^*(d_s),$$

say. It is negative if $s < s^*(d_s)$ and positive if $s > s^*(d_s)$. This implies that (for given d_s and d_m) the function F is minimal if $s = s^*(d_s)$.

Note that there are only two possible values for d_m. The sequence starts with R. If it also ends with an R, then $m_{RT} = m_{TR}$ and, hence, $d_m = 0$. If, however, the sequence ends with a T, then $d_m = 1$. No other values of d_m are possible.

To derive the minimum of F, we consider the two possibilities for d_m separately.

Case 1: $d_m = 1$. The derivative of $F(s^*(d_s), d_s, 1)$ with respect to d_s equals

$$\frac{d_s(p^2+2p+2)+p(p+1)}{2(p+1)^2}.$$

Hence the derivative is zero for

$$d_s = -\frac{p(p+1)}{p^2+2p+2} \in (-1, 0).$$

It is negative for smaller d_s and positive for larger d_s. Since d_s must be an integer, the minimum of F for Case 1 must therefore be either $F(s^*(-1), -1, 1)$ or $F(s^*(0), 0, 1)$. Since

$$F(s^*(0), 0, 1) = \frac{p^2-2p+3}{4} > \frac{p^4-p^2+4p+5}{4(p+1)^2} = F(s^*(-1), -1, 1),$$

the minimum for Case 1 is $F(s^*(-1), -1, 1)$.

Case 2: $d_m = 0$. The derivative of $F(s^*(d_s), d_s, 0)$ with respect to d_s equals

$$\frac{d_s(p^2+2p+2)-p}{2(p+1)^2}.$$

Hence, the derivative is zero if

$$d_s = \frac{p}{p^2 + 2p + 2} \in (0, 1).$$

Again, it is negative for smaller d_s and positive for larger d_s. Hence, the minimum for Case 2 must be either $F(s^*(0), 0, 0)$ or $F(s^*(1), 1, 0)$. Since,

$$F(s^*(0), 0, 0) = \frac{p^4 - p^2 + 1}{4(p+1)^2} < \frac{p^4 + 3}{4(p+1)^2} = F(s^*(1), 1, 0),$$

the minimum for Case 2 is $F(s^*(0), 0, 0)$.

Comparing the two cases, we find that $F(s^*(0), 0, 0) < F(s^*(-1), -1, 1)$. Therefore we have shown that

$$\min_{s, d_s, d_m} F(s, d_s, d_m) = \frac{p^4 - p^2 + 1}{4(p+1)^2}.$$

Combining the results for the three terms, we conclude for any $z \in Z_p$ that

$$\begin{aligned} G_z &\leq -\frac{p-1}{4p} + \frac{3p^3 + 4p^2 - 2p - 4}{4(p+1)^2} - \frac{F(s^*(0), 0, 0)}{p} \\ &= -\frac{(p-1)(p+1)^2}{4p(p+1)^2} + \frac{3p^4 + 4p^3 - 2p^2 - 4p}{4p(p+1)^2} - \frac{p^4 - p^2 + 1}{4p(p+1)^2} \\ &= \frac{(2p+3)(p-1)}{4(p+1)}. \end{aligned}$$

This completes the proof.

Acknowledgment The work of the first author was supported by the Deutsche Forschungsgemeinschaft (SFB 823). The paper was written while the first author was visiting the Free University of Bozen-Bolzano. The work of the second author was supported by funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 999754557. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Swiss Government. This work is part of the IDEAS European training network (http://www.ideas-itn.eu) from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 633567.

References

Afsarinejad, K. and Hedayat, A. (2002), "Repeated measurements designs for a model with self and mixed carryover effects," *Journal of Statistical Planning and Inference*, 106, 449 - 459.

Ebbers, H. C., Crow, S. A., Vulto, A. G., and Schellekens, H. (2012), "Interchangeability, immunogenicity and biosimilars", *Nature Biotechnology*, 30, 1186-1190.

Generics and Biosimilar Initiative (2014), "US$ 67 billion worth of biosimilar patents expiring before 2020", Available at *http://www.gabionline.net/Biosimilars/General/US-67-billion-worth-of-biosimilar-patents-expiring-before-2020* (last accessed 2017-09-20).

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

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

Kunert, J. and Martin, R. J. (2000), "On the determination of optimal designs for an interference model," *Annals of Statistics*, 28, 1728 - 1742.

Kunert, J. and 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, J. and 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, H. B. (1997), "Optimal repeated measurements designs: the linear optimality equations," *Annals of Statistics*, 25, 2328 - 2344. Corr. (1998), 26, 2081.

Schellekens, H. (2004), "How similar do 'biosimilars' need to be?", *Nature Biotechnology*, 22, 1357-1359.