Optimal cross-over designs for total effects under a model with self and mixed carryover effects

Pierre Druilhet* and Walter Tinsson **

*Laboratoire de Mathématiques, UMR CNRS 6620, Clermont Université, Université Blaise Pascal, 63177 AUBIERE Cedex, France
**Laboratoire de Mathématiques Appliquées, UMR CNRS 5142, Université de Pau, BP 1155, 64013 PAU Cedex, France

Abstract

We consider cross-over designs for a model that includes specific carryover effects when a treatment is preceded by itself. When the parameters of interest are total effects, i.e. the sum of direct effects of treatment and self carryover effects, we show that optimal designs are a compromise between designs balanced on subjects such as balanced binary block designs and designs with subjects having a single treatment. We also propose universally optimal designs with a reduced number of subjects.

Keywords: Approximate design; Neighbour design; Optimal design; Total effect; Universal optimality.

1. Introduction

In cross-over designs, it is often assumed that the response on a given period depends on both the treatment applied to that period (direct treatment effect) and the treatment applied to the previous period (carryover effect). The optimal designs depend on the way the interference between these two effects is modelized (see Bose and Dey, 2009, for a recent review of optimal cross-over designs). The simpler way to modelize this interference is to assume that carryover and direct treatment effects are additive, which means that the carryover effect of a treatment is the same no matter the treatment applied to the following period is. For example Kunert (1984), Kushner (1998), Bailey and Druilhet (2004) obtained optimal or efficient designs for this model, Zheng (2013) consider optimal designs in the presence of drop out subjects. The additive model is often too coarse. To enrich the model, Kempton et al. (2001) proposed a model where carryover effects are proportional to direct effects and Bailey and Kunert (2006) obtained optimal designs for that model. Sen and Mukerjee (1987) proposed a model with interaction between carryover and direct treatment effect. Park et al. (2011) obtained efficient cross-over designs under that model. As a compromise between additive and full interaction models, Afsarinejad and Hedayat (2002) proposed a model with two different kinds of carryover effect for a treatment: a self carryover when the following treatment is the same one and a mixed carryover effect when the following treatment is a different one. This is equivalent to assume a partial interaction between treatment and carryover effects. Kunert and Stufken (2002) obtained optimal designs for the estimation of direct treatment effects under this model. For designs with pre-periods and circularity conditions, Druilhet and Tinsson (2009) obtained efficient designs when the parameters of interest are total effects, i.e. the effects of treatments preceded by themselves.

In this paper we consider designs without pre-period and therefore without circularity condition for the model with self and carryover effects. We obtain optimal designs for total effects based on the construction of optimal sequences initially proposed by in Kushner (1997). Then, we propose a method to derive universally optimal designs with a limited number of subjects.

2. Models with self and mixed carryover effects.

Let b be the number of subjects, k the number of periods, t the number of treatments and $n = bk$ the total number of observations. For $1 \leqslant u \leqslant b$ and $1 \leqslant j \leqslant k$, denote by $d(u, j)$ the treatment assigned to subject u in period j. As in Afsarinejad & Hedayat (2002) , we assume that the response y_{uj} is

$$y_{uj} = \beta_u + \tau_{d(u,j)} + \lambda_{d(u,j-1)} + \chi_{d(u,j-1)d(u,j)} + \varepsilon_{uj}, \tag{1}$$

where β_u is the effect of subject u, τ_i is the effect of treatment i, λ_i is the general carryover effect of treatment i, $\chi_{ii'}$ is the additional specific carryover effect when treatment i is followed by itself, with $\chi_{ii'} = 0$ if $i \neq i'$, and ε_{uj} are independent identically distributed errors with expectation 0 and variance σ^2. In vector notation, we have

$$Y = B\,\beta + T_d\,\tau + L_d\,\lambda + S_d\,\chi + \varepsilon \tag{2}$$

where Y is the n-vector of responses, β the b-vector of subject effects, τ the t-vector of treatment effects, λ the t-vector of carryover effects and χ the t-vector of self-carryover effects whose entries are χ_{ii}, $1 \leqslant i \leqslant t$. The matrices B, T_d, L_d and S_d are the design matrices of subjects, direct treatments, carryover and specific self-carryover effects. Note that $\mathrm{var}(\varepsilon) = \sigma^2 I_{bk}$. We define the vector ϕ of total effects by $\phi = \tau + \lambda + \chi$, which corresponds to the direct effect of a treatment in addition to that treatment's carryover effect when preceded by itself. If $\theta' = (\tau', \lambda', \chi')$ and $K' = (I_t|I_t|I_t)$, then

$$\phi = K'\theta.$$

The model we have described does not include period effects. However, it will be seen in §3.3 that the optimal designs obtained for this model are also optimal when period effects are present. We denote by $\Omega_{t,b,k}$ the set of all cross-over designs with t treatments, b subjects and k periods. We also denote respectively by I_n, J_n and $\mathbb{I}_n$ the $n \times n$ identity matrix, the $n \times n$ matrix of ones and the n-vector of ones.

3. Linearization of the problem

3.1. Information matrices

There are two equivalent ways to define the information matrix for the parameter ϕ (see Pukelsheim, 1993, chapter 3). The first one is to consider a linear reparameterize of the model by $\theta \mapsto \eta = (\phi', \psi')'$, then calculate the partitioned information matrix $C_d(\eta)$ of η and derive the information matrix $C_d(\phi)$ for ϕ by taking the Schur-complement in $C_d(\eta)$. This approach allows one to compute the information matrix for a given design, but may lead to untractable formulae to derive optimal designs. In order to adapt Kushner's (1997) methods to our case, it is preferable to use a definition of $C_d(\phi)$ through an extremal representation which allows linearization techniques. This approach is presented below.

The information matrix for the whole parameter $\theta' = (\tau', \lambda', \chi')$ is given by:

$$C_d(\theta) = (T_d|L_d|S_d)'\,\omega_B^{\perp}\,(T_d|L_d|S_d)$$

denoting $\omega_B = B\,(B'B)^{-1}\,B'$ the projection matrix onto the column span of B and $\omega_B^{\perp} = I_n - \omega_B$. So:

$$C_d\,(\theta) = \begin{pmatrix} T_d'\omega_B^{\perp}T_d & T_d'\omega_B^{\perp}L_d & T_d'\omega_B^{\perp}S_d \\ L_d'\omega_B^{\perp}T_d & L_d'\omega_B^{\perp}L_d & L_d'\omega_B^{\perp}S_d \\ S_d'\omega_B^{\perp}T_d & S_d'\omega_B^{\perp}L_d & S_d'\omega_B^{\perp}S_d \end{pmatrix} = \begin{pmatrix} C_{d11} & C_{d12} & C_{d13} \\ C_{d12}' & C_{d22} & C_{d23} \\ C_{d13}' & C_{d23}' & C_{d33} \end{pmatrix}. \tag{3}$$

The information matrix for the total effects ϕ may be obtained from $C_d\,(\theta)$ by the following extremal representation proposed by Gaffke (1987):

$$C_d(\phi) = \min_{L \in \mathbb{R}^{3t \times t} : L'K = I_t} L'C_d(\theta)L, \tag{4}$$

where the minimum, which exists and is unique, is taken relative to the Loewner ordering. We recall that, for two $t \times t$ symmetric matrices M and N, $M \leqslant N$ relative to the Loewner ordering means that $u'Mu \leqslant u'Nu$ for any t-vector u.

Lemma 1. *The row and column sums of $C_d(\phi)$ are zero, i.e. $C_d(\phi)\mathbb{I}_t = 0$.*

Proof. It is equivalent to prove that $\mathbb{I}_t' C_d(\phi)\mathbb{I}_t = 0$. Since $T_d\mathbb{I}_t = \mathbb{I}_n$ and $\omega_B^\perp \mathbb{I}_n = 0$, we have $\mathbb{I}_t' C_{d11}\mathbb{I}_t = 0$. Consider $L_1 = (I_t|0_t|0_t)'$ where 0_t is the $t \times t$ zero matrix. L_1 satisfies the constraint $L_1'K = I_t$. Therefore, from (4), $\mathbb{I}_t'\, C_d(\phi)\, \mathbb{I}_t \leqslant \mathbb{I}_t'\, L_1'\, C_d(\theta)\, L_1\, \mathbb{I}_t = \mathbb{I}_t'\, C_{d11}\, \mathbb{I}_t = 0$. □

The definition of $C_d(\phi)$ given by (4) does not provide an explicit expression. Especially, the matrix L^* that achieves the minimum has usually an untractable form (see Pukelsheim, 1993, chapter 3). A design d is said to be *symmetric* if all its matrices C_{dij} are completely symmetric, i.e. if $C_{dij} = (a_{ij}\, I_t + b_{ij}\, J_t)$ for some scalars a_{ij} and b_{ij}. For such designs, the matrix L^* has the simple parametric form given by the following lemma.

Lemma 2 (Druilhet and Tinsson, 2009, Prop. 1)**.** *If a design d is symmetric, then $C_d(\phi)$ is completely symmetric. Moreover, let L^* be a matrix that achieves the minimum in (4), i.e. such that*

$$C_d(\phi) = L^{*\prime} C_d(\theta) L^*.$$

Then, L^* may be chosen among the matrices L having the following form

$$L = \left((1 - x_1 - x_2)\, I_t - (y_1 + y_2)\, J_t \right| x_1 I_t + y_1 J_t \,|\, x_2 I_t + y_2 J_t\left.\right), \tag{5}$$

where x_1, x_2, y_1 and y_2 are scalars.

Note that the parametric form of L in (5) includes the constraint $L^{*\prime}K = I_t$. We denote x_1^*, x_2^*, y_1^* and y_2^* the scalars corresponding to L^*. A completely symmetric matrix with zero row and column sums is entirely determined by its trace. Therefore, for a symmetric design, we can obtained scalars x_1^*, x_2^*, y_1^* and y_2^* the by minimizing $\mathrm{tr}\,(L^{*\prime} C_d(\theta) L^*)$. For any design d, we denote $c_{dij} = \mathrm{tr}\,(C_{dij})$ and $\tilde{c}_{dij} = \mathrm{tr}\,(C_{dij} J_t)$, $i, j = 1, 2, 3$.

Lemma 3. *If a design d is symmetric,*

$$\mathrm{tr}(C_d(\phi)) = \min_{x_1, x_2, y_1, y_2} q_d\,(x_1, x_2, y_1, y_2) = q_d\,(x_1^*, x_2^*, y_1^*, y_2^*)\,,$$

where x_1^*, x_2^*, y_1^* and y_2^* are the scalars defined in Lemma 2 and

$$\begin{aligned} q_d\,(x_1, x_2, y_1, y_2) = \quad & (1 - x_1 - x_2)^2\, c_{d11} + 2x_1\,(1 - x_1 - x_2)\, c_{d12} + 2x_2\,(1 - x_1 - x_2)\, c_{d13} \\ & + x_1^2 c_{d22} + \left(2x_1 y_1 + t y_1^2\right) \tilde{c}_{d22} + x_2^2 c_{d33} + \left(2x_2 y_2 + t y_2^2\right) \tilde{c}_{d33} \\ & + 2x_1 x_2 c_{d23} + 2\,(x_1 y_2 + x_2 y_1 + t y_1 y_2)\, \tilde{c}_{d23}. \end{aligned}$$

Proof. Since $T_d\mathbb{I}_t = \mathbb{I}_n$ and $\omega_B^\perp \mathbb{I}_n = 0$, $C_{d1j}\mathbb{I}_t = \mathbb{I}_t' C_{d1j} = 0$ and $C_{d1j} J_t = J_t C_{d1j} = 0$ for $j = 1, 2, 3$. Thus,

$$\begin{aligned} L^{*\prime} C_d\,(\theta)\, L^* = \quad & (1 - x_1^* - x_2^*)^2\, C_{d11} + 2x_1^*\,(1 - x_1^* - x_2^*)\, C_{d12} + 2x_2^*\,(1 - x_1^* - x_2^*)\, C_{d13} \\ & + (x_1^* I_t + y_1^* J_t)\, C_{d22}\, (x_1^* I_t + y_1^* J_t) + (x_2^* I_t + y_2^* J_t)\, C_{d33}\, (x_2^* I_t + y_2^* J_t) \\ & + 2\,(x_1^* I_t + y_1^* J_t)\, C_{d23}\, (x_2^* I_t + y_2^* J_t). \end{aligned}$$

Remark that $\mathrm{tr}\,(J_t C_{dij} J_t) = \mathrm{tr}\left(C_{dij} J_t^2\right) = t\,\mathrm{tr}\,(C_{dij} J_t) = t\tilde{c}_{dij}$. Therefore, from lemma 2, $\mathrm{tr}(C_d(\phi)) = \mathrm{tr}\,(L^{*\prime} C_d\,(\theta)\, L^*) = q_d\,(x_1^*, x_2^*, y_1^*, y_2^*)$. Since the minimum in (4) exists and is unique, we have

$$\mathrm{tr}(C_d(\phi)) = \min_{L \in \mathbb{R}^{3t \times 3} : L'K = I_t} \mathrm{tr}(L' C_d(\theta) L) = \min_{x_1, x_2, y_1, y_2} q_d\,(x_1, x_2, y_1, y_2)\,.$$

□

Lemma 4. *Let d be a symmetric design, from lemmas 1, 2 and 3, we have*

$$C_d(\phi) = \frac{q_d\,(x_1^*, x_2^*, y_1^*, y_2^*)}{t - 1} \left(I_t - \frac{1}{t} J_t\right).$$

Proof. From Lemma 2, the matrix $C_d(\phi)$ is completely symmetric. From Lemma 1, $C_d(\phi)\mathbb{I}_t = 0$. So $C_d(\phi) = \alpha\left(I_t - \frac{1}{t} J_t\right)$ for some scalar α. Then $\mathrm{tr}(C_d(\phi)) = \alpha(t - 1)$ and, from Lemma 3, $\alpha(t - 1) = q_d\,(x_1^*, x_2^*, y_1^*, y_2^*)$. □

Lemma 5. *For any design d, we have:*

$$\operatorname{tr}(C_d(\phi)) \leqslant \min_{x_1,x_2,y_1,y_2} q_d\left(x_1,x_2,y_1,y_2\right)$$

Proof. The proof of this lemma is similar to the proof of Proposition 2 by Druilhet and Tinsson (2009) and will be therefore omitted. □

3.2. Decomposition over the subjects

We denote by T_{du}, L_{du} and S_{du} the incidence matrices restricted to subject u, thus $T'_d = (T'_{d1}|\ldots|T'_{db})$, $L'_d = (L'_{d1}|\ldots|L'_{db})$ and $S'_d = (S'_{d1}|\ldots|S'_{db})$. Since $\omega_B^{\perp} = I_b \otimes Q_k$ with $Q_k = I_k - k^{-1}J_k$, each coefficients c_{dij} and $\tilde{c}_{dij}$ can be decomposed using the contribution $c_{dij}^{(u)}$ and $\tilde{c}_{dij}^{(u)}$ of each subject u. For example when $i=j=1$:

$$c_{d11} = \operatorname{tr}\left(C_{d11}\right) = \operatorname{tr}\left(T'_d\omega_B^{\perp}T_d\right) = \sum_{u=1}^{b} c_{d11}^{(u)} \text{ with } c_{d11}^{(u)} = \operatorname{tr}\left(T'_{du}Q_kT_{du}\right).$$

In order to calculate all the c_{dij} and $\tilde{c}_{dij}$, we introduce the following notations for each subject u: n_{ui} is the number of occurrences of treatment i, m_{ui} is the number of occurrences of treatment i followed by itself and t_{ui} is 1 if treatment i is in the last period, 0 otherwise. We denote also by n_u, m_u and t_u the vectors constituted by the t values n_{ui}, m_{ui} and t_{ui} and $\langle .,. \rangle$ is the usual scalar product on $\mathbb{R}^t$. With these notations, $c_{dij}^{(u)}$ and $\tilde{c}_{dij}^{(u)}$ can be written:

$$\begin{aligned}
c_{d11}^{(u)} &= k - \frac{1}{k}\left\|n_u\right\|^2, & c_{d12}^{(u)} &= \langle m_u, \mathbb{I}_t\rangle - \frac{1}{k}\left\langle n_u, n_u - t_u\right\rangle, \\
c_{d13}^{(u)} &= \langle m_u, \mathbb{I}_t\rangle - \frac{1}{k}\left\langle n_u, m_u\right\rangle, & c_{d22}^{(u)} &= (k-1) - \frac{1}{k}\left\|n_u - t_u\right\|^2, \\
c_{d23}^{(u)} &= \langle m_u, \mathbb{I}_t\rangle - \frac{1}{k}\left\langle n_u - t_u, m_u\right\rangle, & c_{d33}^{(u)} &= \langle m_u, \mathbb{I}_t\rangle - \frac{1}{k}\left\|m_u\right\|^2, \\
\tilde{c}_{d22}^{(u)} &= \frac{1}{k}\left(k-1\right), & \tilde{c}_{d23}^{(u)} &= \frac{1}{k}\langle m_u, \mathbb{I}_t\rangle, \\
\tilde{c}_{d33}^{(u)} &= \langle m_u, \mathbb{I}_t\rangle\left(1 - \frac{\langle m_u, \mathbb{I}_t\rangle}{k}\right).
\end{aligned}$$

Note that it is not necessary to compute all these 9 values for each subject because $\tilde{c}_{d22}^{(u)}$ depends only on the size k and $\tilde{c}_{d33}^{(u)} = k\tilde{c}_{d23}^{(u)}\left(1 - \tilde{c}_{d23}^{(u)}\right)$. It follows that:

$$q_d\left(x_1,x_2,y_1,y_2\right) = \sum_{u=1}^{b} h_d^{(u)}\left(x_1,x_2,y_1,y_2\right) \text{ with:}$$

$$\begin{aligned}
h_d^{(u)}\left(x_1,x_2,y_1,y_2\right) = \;& c_{d11}^{(u)} + 2\left(c_{d12}^{(u)} - c_{d11}^{(u)}\right)x_1 + 2\left(c_{d13}^{(u)} - c_{d11}^{(u)}\right)x_2 \\
&+ \left(2c_{d11}^{(u)} - 2c_{d12}^{(u)} - 2c_{d13}^{(u)} + 2c_{d23}^{(u)}\right)x_1x_2 + \left(2\tilde{c}_{d22}^{(u)}\right)x_1y_1 + \left(2\tilde{c}_{d33}^{(u)}\right)x_2y_2 \\
&+ \left(2\tilde{c}_{d23}^{(u)}\right)x_1y_2 + \left(2\tilde{c}_{d23}^{(u)}\right)x_2y_1 + \left(2t\tilde{c}_{d23}^{(u)}\right)y_1y_2 + t\tilde{c}_{d22}^{(u)}y_1^2 + t\tilde{c}_{d33}^{(u)}y_2^2 \\
&+ \left(c_{d11}^{(u)} - 2c_{d12}^{(u)} + c_{d22}^{(u)}\right)x_1^2 + \left(c_{d11}^{(u)} - 2c_{d13}^{(u)} + c_{d33}^{(u)}\right)x_2^2.
\end{aligned}$$

Two sequences of treatments in two subjects u_1 and u_2 are said to be equivalent if $h_d^{(u_1)} = h_d^{(u_2)}$ (that is if they have the same values for the c_{dij} and $\tilde{c}_{dij}$ in h_d). Therefore, for given k and t, we can divide the set of all possible treatment sequences into $\mathcal{L}$ equivalence classes. By abuse of notation ℓ refers now to an equivalence class as well as its index, so we can write $1 \leqslant \ell \leqslant \mathcal{L}$ and:

$$q_d\left(x_1,x_2,y_1,y_2\right) = b\sum_{\ell=1}^{\mathcal{L}} \pi_{d\ell}\, h_\ell\left(x_1,x_2,y_1,y_2\right) \tag{6}$$

with $\pi_{d\ell}$ the proportion of subjects assigned to the class ℓ and h_ℓ common value of $h_d^{(u)}$ for all the subjects u of the class ℓ.

3.3. Model with period effects

Consider the model (1) with the addition of period effects:

$$Y = P\,\gamma + B\,\beta + T_d\,\tau + L_d\,\lambda + S_d\,\chi + \varepsilon \tag{7}$$

where γ is the k-vector of periods effects and $P = \mathbb{I}_b \otimes I_k$ is the plot×period incidence matrix. Write $\tilde{\theta} = (\tau|\lambda|\phi|\gamma)$. The information matrix for $\tilde{\theta}$ is:

$$\tilde{C}_d(\tilde{\theta}) = \begin{pmatrix} C_d(\theta) & \tilde{C}_{d12} \\ \tilde{C}_{d21} & \tilde{C}_{d22} \end{pmatrix},$$

where $C_d(\theta)$ is the information matrix given by (3) corresponding to Model (2), $\tilde{C}_{d21} = (P'\omega_B^{\perp}T_d|P'\omega_B^{\perp}L_d|P^{'}\omega_B^{\perp}S_d)$, $\tilde{C}_{d12} = \tilde{C}'_{d21}$ and $\tilde{C}_{d22} = P'\omega_B^{\perp}P = b(I_k - J_k/k)$. Note that $(T'_d\omega_B^{\perp}\,P)_{ij} = p_{dij} - r_{di}/k$, where p_{dij} is the number of occurrences of treatment i on period j and r_{di} is the number of occurrences of treatment i over the whole design, $(L'_d\omega_B^{\perp}\,P)_{ij} = p_{di(j-1)} - \tilde{r}_{di}/k$ where $p_{di0} = 0$ and $\tilde{r}_{di}$ is the number of occurrences of treatment i over the whole design except the first period and $(S'_d\omega_B^{\perp}\,P)_{ij} = s_{dij} - s_{di}/k$ where s_{dij} is the number of times treatment i is preceded by itself on period j and s_{dj} is the number of times treatment i is preceded by itself over the whole design.

Denote by $\tilde{C}_d(\phi)$ the information matrix for ϕ under Model (7):

$$\tilde{C}_d(\phi) = \min_{\tilde{L}\in\mathbb{R}^{(3t+k)\times t}:\tilde{L}'K=I_t} \tilde{L}'\tilde{C}_d(\tilde{\theta})\tilde{L}, \tag{8}$$

where $K' = (I_t|I_t|I_t|0_{t\times k})$.

Lemma 6. *The row and column sums of $\tilde{C}_d(\phi)$ are zero.*

Proof. The proof is similar to the proof of Lemma 1 with $L_1 = (I_t|0_t|0_t|0_{t\times k})'$.
□

A design is said to be balanced on the periods if each treatment appears equally often in each period and if the number of times a treatment is preceded by itself on a given period, except the first one, does not depend on the treatment label.

Proposition 1. *If d is a symmetric design balanced on the periods, then the information matrix for total effects is the same under Model (2) and Model (7).*

Proof.
Step 1: if d is a design balanced on the periods, then $p_{dij} = b/t$, $r_{di} = bk/t$ and $\tilde{r}_{di} = b(k-1)/t$ for $i = 1, ..., t$ and $j = 1, ...t$. Moreover $a_{dj} = s_{dij} - s_{di}/k$ do not depend on i and $\tilde{C}'_{d12} = (0_{k\times t}|E'_d|F'_d)$ with E_d the $(t \times k)$ matrix $\frac{b}{kt}\,(-(k-1)\mathbb{I}_t|\mathbb{I}_t|...|\mathbb{I}_t)$ and F_d the $(t \times k)$ matrix $(a_{d1}\mathbb{I}_t|...|a_{dk}\mathbb{I}_t)$. Since $Q_t\mathbb{I}_t = 0$, we have $Q_tE_d = Q_tF_d = 0$ and then $(I_3 \otimes Q_t)\,\tilde{C}_{d12} = 0$.
Step 2: write $\tilde{L}^{*'} = (M'|N')$ a matrix L that achieves the minimum in (8) where $M' = (A'|B'|C')$ with A, B and C of size $t\times t$ such that $A+B+C = I$ and N is a $k\times t$ matrix. By the same argument that used in Proposition 1 by Druilhet and Tinsson (2009), the matrices A, B and C can be chosen completely symmetric. Therefore, A, resp. B and C, commutes with Q_t and $Q_tM' = M'(I_3\otimes Q_t)$.
Step 3: from Lemma 6, $\tilde{C}_d(\phi)\mathbb{I}_t = 0$. Thus, $\tilde{C}_d(\phi) = Q_t\tilde{C}_d(\phi)Q_t = Q_tM'C_d(\theta)MQ_t + Q_tM'\tilde{C}_{d12}NQ_t + Q_tN'\tilde{C}_{d21}MQ_t + Q_tN'\tilde{C}_{d22}NQ_t$. But, from step 1 and 2, $QM'\tilde{C}_{d12} = M'(I_3 \otimes Q_t)\tilde{C}_{d12} = 0$ and $(\tilde{C}_{d21}MQ_t)' = QM'\tilde{C}_{d12} = 0$. Therefore, $\tilde{C}_d(\phi) = Q_tM'C_d(\theta)MQ_t + Q_tN'\tilde{C}_{d22}NQ_t$. Since M and N minimize this expression and N is allowed to vary freely, N can be chosen equal to 0 and then $\tilde{C}_d(\phi) = C_d(\phi)$.

□

Corollary 1. *A symmetric period-balanced design which is universally optimal under model (2) is also universally optimal under model (7).*

4. Optimal designs

Our goal is now to obtain universally optimal designs. Since, for any design d, $C_d(\phi)$ have row and column sums equal to zero, we know from Kiefer (1975) that a design d^* for which the information matrix $C_{d^*}(\phi)$ is completely symmetric and maximizes the trace over all the designs d in $\Omega_{t,b,k}$ is universally optimal for the estimation of total effects.

Proposition 2. *Let d^* be a symmetric design in $\Omega_{t,b,k}$ with proportions $\pi^* = (\pi_{d^*1}, ..., \pi_{d^*\mathcal{L}})$ that achieve the maximum in*

$$\max_{\pi} \min_{(x_1,x_2,y_1,y_2)} b \sum_{\ell=1}^{\mathcal{L}} \pi_\ell \, h_\ell \left(x_1, x_2, y_1, y_2\right), \tag{9}$$

then d^* is universally optimal for the estimation of total effects over $\Omega_{t,b,k}$.

Proof. From Lemma 2, C_{d^*} is completely symmetric. Therefore, it is sufficient to prove that d^* maximises the trace of the information matrix. By Lemma 3 and Lemma 5, for any design d in $\Omega_{t,b,k}$, we have

$$\begin{aligned} \operatorname{tr}(C_d(\phi)) &\leqslant \min_{x_1,x_2,y_1,y_2} q_d\left(x_1, x_2, y_1, y_2\right) = \min_{x_1,x_2,y_1,y_2} b \sum_{\ell=1}^{\mathcal{L}} \pi_{d\ell} \, h_\ell \left(x_1, x_2, y_1, y_2\right) \\ &\leqslant \max_{\pi} \min_{(x_1,x_2,y_1,y_2)} b \sum_{\ell=1}^{\mathcal{L}} \pi_{d\ell} \, h_\ell \left(x_1, x_2, y_1, y_2\right) = \operatorname{tr}(C_{d^*}(\phi)). \end{aligned}$$

□

The following proposition by Kunert and Martin (2000) provides a simple way to check if a symmetric design is universally optimal.

Proposition 3. *Consider a symmetric design $d^* \in \Omega_{t,b,k}$ and a point $(x_1^*, x_2^*, y_1^*, y_2^*)$ such that the first derivatives of q_{d^*} are equal to zero. If we have:*

$$\forall \; \ell = 1, ..., \mathcal{L} \;, \; bh_\ell \left(x_1^*, x_2^*, y_1^*, y_2^*\right) \leqslant q_{d^*} \left(x_1^*, x_2^*, y_1^*, y_2^*\right)$$

then d^* is universally optimal over $\Omega_{t,b,k}$.

Proof. Denote q_d^* the minimum of the function q_d. For every design $d \in \Omega_{t,b,k}$ it is clear that $q_d^* \leqslant q_d\left(x_1^*, x_2^*, y_1^*, y_2^*\right)$ and:

$$q_d\left(x_1^*, x_2^*, y_1^*, y_2^*\right) = b\sum_{\ell=1}^{\mathcal{L}} \pi_{d\ell} h_\ell \left(x_1^*, x_2^*, y_1^*, y_2^*\right) \leqslant b\sum_{\ell=1}^{\mathcal{L}} \pi_{d\ell} \frac{q_{d^*}\left(x_1^*, x_2^*, y_1^*, y_2^*\right)}{b}.$$

Since $q_{d^*}\left(x_1^*, x_2^*, y_1^*, y_2^*\right) = q_{d^*}^*$, d^* maximizes the trace of the information matrix over $\Omega_{t,b,k}$.

□

In order to prove that the optimal design d^* is generated by only one treatment sequence ℓ_1 (*i.e.* $q_{d^*}\left(x_1, x_2, y_1, y_2\right) = bh_{\ell_1}\left(x_1, x_2, y_1, y_2\right)$), we have to find $(x_1^*, x_2^*, y_1^*, y_2^*)$ that minimize h_{ℓ_1}, find the minimum $q_{d^*}^*$ of q_{d^*} (that is only find the minimum of h_{ℓ_1}) and check that for $1 \leqslant \ell \leqslant \mathcal{L}$, $bh_\ell\left(x_1^*, x_2^*, y_1^*, y_2^*\right) \leqslant q_{d^*}^*$. Otherwise an optimal design can be generated by two or more sequences and only sequences ℓ such that $h_\ell\left(x_1^*, x_2^*, y_1^*, y_2^*\right) = \max_{1 \leqslant \ell \leqslant \mathcal{L}} h_\ell\left(x_1^*, x_2^*, y_1^*, y_2^*\right)$ have to be considered. Therefore $(x_1^*, x_2^*, y_1^*, y_2^*)$ must be at the intersection of two or more of the h_ℓ and the proportions $\pi_{d\ell}$ must be chosen such that the partial derivatives of $q_d\left(x_1^*, x_2^*, y_1^*, y_2^*\right)$ are both 0.

5. Some examples

We give here the way to obtain optimal or efficient designs from optimal sequences. First, the case $k = 4$ is treated with all the details. Then, we consider all the cases for $k \leqslant 10$. We can see that the optimal designs obtained here are most often generated by a single sequence whereas optimal circular designs obtained by Druilhet and Tinsson (2009) are generated by a mixture of sequences. However, it is worth noting that if we use the optimal sequences obtained here to generate circular designs, we obtain very efficient designs.

Example 1. Consider the case $k = 4$. If $t \geqslant 4$ then the following table lists the equivalence classes and the corresponding values $c_{d11}^{(\ell)}$, $c_{d12}^{(\ell)}$, $c_{d13}^{(\ell)}$, $c_{d22}^{(\ell)}$, $c_{d23}^{(\ell)}$, $c_{d33}^{(\ell)}$ and $\tilde{c}_{d23}^{(\ell)}$.

ℓ	Sequence	$c_{d11}^{(\ell)}$	$c_{d12}^{(\ell)}$	$c_{d13}^{(\ell)}$	$c_{d22}^{(\ell)}$	$c_{d23}^{(\ell)}$	$c_{d33}^{(\ell)}$	$\tilde{c}_{d23}^{(\ell)}$
01	[1 1 1 1]	0	0	0	3/4	3/4	3/4	3/4
02	[1 1 1 2]	3/2	−1/4	1/2	3/4	1/2	1	1/2
03	[1 1 2 1]	3/2	−3/4	1/4	7/4	1/2	3/4	1/4
04	[1 1 2 2]	2	1/2	1	7/4	5/4	3/2	1/2
05	[1 1 2 3]	5/2	−1/4	1/2	7/4	1/2	3/4	1/4
06	[1 2 1 2]	2	−3/2	0	7/4	0	0	0
07	[1 2 1 3]	5/2	−5/4	0	7/4	0	0	0
08	[1 2 2 1]	2	−1/2	1/2	7/4	1/2	3/4	1/4
09	[1 2 2 2]	3/2	1/4	1/2	7/4	1	1	1/2
10	[1 2 3 1]	5/2	−1	0	9/4	0	0	0
11	[1 2 3 3]	5/2	0	1/2	9/4	3/4	3/4	1/4
12	[1 2 3 4]	3	−3/4	0	9/4	0	0	0

For the sequence [1 1 2 2] the corresponding function h_4 is then:

$$\begin{aligned} h_4(x_1,x_2,y_1,y_2) = \;& 2 - 3x_1 - 2x_2 + \frac{7}{2}x_1x_2 + \frac{3}{2}x_1y_1 + 2x_2y_2 \\ & + x_1y_2 + x_2y_1 + ty_1y_2 + \frac{3t}{4}y_1^2 + ty_2^2 + \frac{11}{4}x_1^2 + \frac{3}{2}x_2^2 \end{aligned}$$

The partial derivatives of h_4 are zero if and only if:

$$x_1^* = \frac{8t(t-2)}{8-34t+17t^2},\; x_2^* = \frac{2t^2}{8-34t+17t^2},\; y_1^* = \frac{-8(t-2)}{8-34t+17t^2},\; y_2^* = \frac{-2t}{8-34t+17t^2}.$$

and then:

$$h_4(x_1^*,x_2^*,y_1^*,y_2^*) = \frac{4\left(5t^2-11t+4\right)}{8-34t+17t^2}$$

Finally we check that:

$$\forall\, \ell = 1,...,12 \;\; , h_\ell(x_1^*,x_2^*,y_1^*,y_2^*) \leqslant h_4(x_1^*,x_2^*,y_1^*,y_2^*)$$

and then the sequence [1 1 2 2] is optimal.

Example 2. This example is devoted to designs generated by a single optimal sequences. For $k > 4$ it is a tedious task to obtain explicit results, so optimal designs are computed numerically. The tables given below list all the results for $k = 4, 5, 6, 7, 8, 9, 10$ and $t \geqslant k$:

k	Single optimal sequence
4	[1 1 2 2]
5	[1 1 2 2 2]
6	*none*
7	[1 1 2 2 3 3 3]

k	Single optimal sequence
8	[1 1 1 2 2 3 3 3]
9	[1 1 1 2 2 2 3 3 3]
10	[1 1 1 2 2 2 3 3 3 3]

A simple way to construct a symmetric design balanced on the periods which is universally optimal is to consider all the possible treatment permutations from an optimal sequence. For example, for $t = 4$ and $k = 7$, the following design is universally optimal:

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

When the number of different treatments that appear in the optimal sequence is 3, this method required $t(t-1)(t-2)$ subjects. We propose now a method which requires, when it is feasible, only $t(t-1)$ subjects. We first construct a binary block design with 3 periods such that each ordered pair of distinct treatments appear equally often on each pair of periods. The resulting design is a balanced incomplete block design in the usual sense which is also neighbour balanced at distance 1 and 2 and balanced on the period. From this design, we construct the final design by replicating the treatment according to the optimal sequence pattern. It is straightforward to check that the resulting design is a symmetric design balanced on the periods and therefore is universally optimal according to Proposition 2 and Corollary 1. For example, again for $t = 4$ and $k = 7$, we consider first the following design with 3 periods and 12 subjects:

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

It satisfies the conditions above. From this design and the optimal sequence, we obtain the design:

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

which is universally optimal among all possible designs with 12 subjects, 4 treatments and 7 periods.

Example 3. Consider the case $k = 6$. The optimal design is then a mixture of the two sequences [1 1 1 2 2 2] and [1 1 2 2 3 3]. The optimal proportions of these sequences depend on t: some values are given below:

Optimal sequences	$t = 6$	$t = 7$	$t = 8$	$t = 10$	$t = 20$
[1 1 1 2 2 2]	0.482	0.428	0.388	0.330	0.210
[1 1 2 2 3 3]	0.518	0.572	0.612	0.670	0.790

Note that the sequence [1 1 2 2 3 3] is always predominant in the mixture. In practice it can be used alone because of its high efficiency which is given below:

Sequence	$t = 6$	$t = 7$	$t = 8$	$t = 10$	$t = 20$
[1 1 2 2 3 3]	0.989	0.992	0.994	0.996	0.998

Example 4. Now consider a situation where t is less than the number of treatments that appear in an optimal sequence. For example $k = 10$ and $t = 2$. Such cases are easier to compute because the number of equivalence classes to be considered is lower than in the general case. In our example there are only 46 equivalence classes (instead of 973 for $t \geqslant 10$) and we find that an optimal design is generated by the following single sequence :

[1 1 1 1 1 2 2 2 2 2]

References

[1] Afsarinejad, K. & Hedayat, A. S. (2002). Repeated measurements designs for a model with self and simple mixed carryover effects. *J. Statist. Plan. Infer.* **106**, 449–59.

[2] Bailey, R. A. & Druilhet, P. (2004). Optimality of neighbor-balanced designs for total effects. *Ann. Statist.* **32**, 1650–61.

[3] Bailey, R. A. & Kunert, J. (2006). On optimal crossover designs when carryover effects are proportionnal to direct effects. *Biometrika* **93**, 613–25.

[4] Bose, M. & Dey, A (2009) *Optimal crossover designs.* World Scientific Publishing

[5] Druilhet, P. & Tinsson, W. (2009). Optimal repeated measurement designs for a model with partial interactions. *Biometrika* **96**, 677–690.

[6] Druilhet, P. & Tinsson, W. (2012). Efficient circular neighbour designs for spatial interference model *J. Statist. Plan. Infer.* **142**, 1161–1169.

[7] Gaffke, N. (1987). Further characterizations of design optimality and admissibility for partial parameter estimation in linear regression. *Ann. Statist.* **15**, 942–57.

[8] Kempton, R. A., Ferris, S. J. & David, O. (2001). Optimal change-over designs when carryover effects are proportional to direct effects of treatments. *Biometrika* **88**, 391–9.

[9] Kiefer, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models*, Ed. J. N. Srivastava, pp. 333–53, Amsterdam: North-Holland.

[10] Kunert, J. (1984). Optimality of balanced uniform repeated measurements designs. *Ann. Statist.* **12**, 1006-1017.

[11] Kunert, J. & Martin, R. J. (2000). On the determination of optimal designs for an interference model. *Ann. Statist.* **28**, 1728–42.

[12] Kunert, J. & Stufken, J. (2002). Optimal crossover designs in a model with self and mixed carryover effects. *J. Am. Statist. Assoc.* **97**, 898–906.

[13] Kushner, H. B. (1997). Optimal repeated measurements designs: the linear optimality equations. *Ann. Statist.* **25**, 2328–44.

[14] Kushner, H.B. (1998). Optimal and Efficient Repeated-Measurements Designs for Uncorrelated Observations. *J. Am. Statist. Assoc.* **93**, 1176–1187.

[15] Park, D. K., Bose, M., Notz, W. I.& Dean, A. M. (2011). Efficient crossover designs in the presence of interactions between direct and carry-over treatment effects. *J. Statist. Plan. Infer.* **141**, 846–860.

[16] Pukelsheim, F. (1993). *Optimal Design of Experiments.* New York: John Wiley & Sons Inc.

[17] Sen, M. & Mukerjee, R. (1987). Optimal repeated measurements designs under interaction. *J. Statist. Plan. Infer.* **17**, 81–91.

[18] Zheng, W. (2013). Universally Optimal Crossover Designs under Subject Dropout. *Ann. Statist.* To appear.