Optimal repeated measurement designs for a model with partial interactions

By P. DRUILHET

Laboratoire de mathématiques, UMR CNRS 6620

Université Blaise Pascal, 63 177 Aubière CEDEX , France

pierre.druilhet@math.univ-bpclermont.fr

AND W. TINSSON

Laboratoire de mathématiques appliquées, UMR CNRS 5142,

Université de Pau, BP 1155, 64013 PAU CEDEX, France

walter.tinsson@univ-pau.fr

Summary

We consider cross-over designs for a model with partial interactions (Afsarinejad and Hedayat, 2002). In this model the carryover effect is different depending on whether the treatment is preceded by itself or not. When the aim of an experiment is to select the most efficient treatment, the parameters of interest should be the total effects which correspond to the use of a single treatment. In this context, binary designs are inefficient. We obtain optimal designs by generalizing the method introduced by Kushner (1997) and Kunert and Martin (2000). This generalization places the method proposed by Bailey and Druilhet (2004) into Kushner's context. We also propose efficient designs generated by only one sequence.

Some key words: Approximate design, crossover design, optimal design, total effects, universal optimality.

1 Introduction

For chronic diseases, cross-over designs are the most common experimental devices for comparing treatments and selecting the best one. The models associated to these designs usually include carry-over effects. These effects may be additive or interact with treatment effects. Considering all possible interactions induces a model with too many parameters. Kempton et al. (2001) propose a model where carry-over effects are proportional to direct treatment effects: see Bailey and Kunert (2006) for optimal designs related to this model. Afsarinejad and Hedayat (2002) proposed another parsimonious model with partial interactions between treatment and carryover effects: the carryover effect of a treatment is different depending on whether the treatment is preceded by itself or not. For such models, Kunert and Stufken (2002) found that universally optimal designs for direct treatment effects have no consecutive pairs of identical treatments. However, most often the aim of an experiment is to select a single treatment which will be used alone and therefore will always be preceded by itself. In that case, the parameters of interest are the total effects which are the sum of direct treatment effects and self carryover effects. Bailey and Druilhet (2004) considered total effects for models without interactions and showed that binary designs are efficient for such effects. Unfortunately, total effects under the Afsarinejad and Hedayat (2002) model are not estimable for designs with no treatment preceded by itself and binary designs cannot be used in that context.

In this paper, we consider optimal cross-over designs for total effects under the model proposed by Afsarinejad and Hedayat (2002). In Section 2, we present the designs and the models. In Section 3, we show that the extremal representation of information matrices proposed by Gaffke (1987) and Pukelsheim (1993, chapter 3) may be used to generalize the techniques developed by Kushner (1997) and Kunert and Martin (2000). We also show how this generalization places the Bailey and Druilhet (2004) approach in Kushner's context. In Section 4, we obtain optimal approximate designs and also efficient designs generated by only one sequence. In Section 5, we show that the optimal sequences found in Section 4 also give optimal designs when period effects are included in the model.

2 The designs and the models

Let b be the number of subjects, k the number of periods and t the number of treatments. For $1 \leq u \leq b$ and $1 \leq j \leq k$, denote by $d(u,j)$ the treatment assigned to subject u in period j. We first consider a model without period effects. It will be seen in Section 5 that the optimal designs obtained for this model are also optimal when period effects are present. Following Hedayat and Afsarinejad (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 ($\chi_{ii'} = 0$ if $i \neq i'$), ε_{uj} are independent identically distributed errors with expectation 0 and variance σ^2. Note that this parametrization, although equivalent, is slightly different from that proposed by Hedayat and Afsarinejad (2002) or Kunert and Stufken (2002): the self carryover effect of Treatment v is here equal to $\lambda_v + \chi_{v,v}$. The vector χ corresponds to an interaction between direct treatment and carryover effects. In vector notation, we have:

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

where B, T_d, L_d and S_d are the incidence matrices of subjects, direct treatments, carryover and specific self-carryover effects. We define the vector ϕ of *total effects* by $\phi = \tau + \lambda + \chi$ which corresponds to the effect of a treatment when preceded by itself. Note that if $\theta' = (\tau', \lambda', \chi')$ and $K' = (I_t|I_t|I_t)$, then

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

Because the effect of having no treatment differs from the carryover effect of any treatment, we consider here only designs with pre-periods, i.e. designs with one period, called pre-period, added at the beginning. On this preperiod, each subject receives a treatment but the response is not used in the analysis. In this paper, we consider only circular designs, i.e. the treatment assigned to a subject in the pre-period is the same as the treatment assigned in the last period. The circularity condition may be written $d(i,0) = d(i,k)$. One interest of the circularity is that we may use the randomization proposed by Azaïs (1987) which preserve the neighbour structure of the treatment sequences. We denote by $\Omega_{t,b,k}$ the set of all circular designs with t treatments, b subjects and k periods.

3 Upper bound for the information matrix

In this section, we present a general method to derive optimal crossover designs. We generalize to more general effects the techniques developed by Kushner (1997), Kunert and Martin (2000) for direct treatment effects and by Bailey and Druilhet (2004) for total effects. The key point of this approach is the linearisation derived from the extremal representation of the information matrix.

Denote by $\mathbb{1}_k$, I_k and J_k respectively the vector of ones of length k, the (k,k) identity matrix and the (k,k) matrix of ones. For any matrix A, denote by A^+ the Moore-Penrose inverse of A and by $\mathrm{pr}_A = A(A'A)^+A'$ the projection matrix onto the column span of A. We also denote $\mathrm{pr}_A^\perp = I - \mathrm{pr}_A$ and $Q_k = \mathrm{pr}_{\mathbb{1}_k}^\perp = I_k - k^{-1}J_k$. For a square matrix A, $\mathrm{tr}(A)$ is the trace of A. For two symmetric matrices M and N, $M \leq N$ means that $N - M$ is a nonnegative definite matrix (Loewner ordering). A matrix M is *completely symmetric* if $M = aI + bJ$ for some scalars a and b.

3.1 Information matrices and its extremal representation

Consider a generic partitioned linear model:

$$Y = A\alpha + B\beta + \varepsilon \qquad \text{with} \qquad E(\varepsilon) = 0 \quad \text{and} \;\; \mathrm{var}(\varepsilon) = \sigma^2 I,$$

where α is a vector of length qt. From Kunert (1983), the information matrix $C[\alpha]$ of the parameter α is

$$C[\alpha] = A'\mathrm{pr}_B^\perp A. \tag{3}$$

Consider now a subsystem $K'\alpha$ where K is a (qt,s) matrix. The information matrix $C[K'\alpha]$ of $K'\alpha$ may be defined by the extremal representation (Gaffke, 1987 or Pukelsheim, 1993):

$$C[K'\alpha] = \min_{L \in \mathbb{R}^{qt\times s} : L'K = I_s} L'C[\alpha]L, \tag{4}$$

where the minimum is taken relative to the Loewner ordering. It is worth noting that the extremal representation (4) has a unique global minimum for any nonnegative symmetric matrix $C[\alpha]$. If L^* is a (qt,s) matrix that minimizes $L'C[\alpha]L$ under the constraint $L'K =$

I_s, then

$$C[K'\alpha] = L^{*\prime} C[\alpha] L^{*}. \tag{5}$$

The fact that, for a fixed L, $L'C[\alpha]L$ is linear in $C[\alpha]$ will be useful in Section 4.2, where $C[\alpha]$ will be decomposed in a sum of several matrices.

The main issue in constructing optimal designs will be to find L^* for the designs candidate to optimality. Because there exists a unique global minimum in (4), L^* minimizes $L'C[\alpha]L$ under the constraint $L'K = I_s$ iff L^* minimizes $\mathrm{tr}(L'C[\alpha]L)$. So, we have

$$\mathrm{tr}(C[K'\alpha]) = \mathrm{tr}(L^{*\prime} C[\alpha] L^{*}) = \min_{L \in \mathbb{R}^{qt \times s} : L'K = I_s} \mathrm{tr}(L'C[\alpha]L). \tag{6}$$

In our problem, the matrix $C[\alpha]$ will have a natural block structure with completely symmetric blocks. In that case, the following results show that the blocks of L^* are also completely symmetric.

Proposition 1 *Let $C[\alpha] = (\mathcal{C}_{ij})_{1\le i,j\le q}$ be a nonnegative symmetric block matrix, where the blocks $\mathcal{C}_{ij}$, $1 \le i, j \le q$, are (t,t) completely symmetric matrices. Let $K' = (K_1', ..., K_q')$ where for all i, K_i are (t,t) completely symmetric matrices. Then $C[K'\alpha]$ is completely symmetric. Moreover, if $L^{*\prime} = (L_1^{*\prime}, ..., L_q^{*\prime})$ satisfies both (5) and the constraint $L^{*\prime}K = I_t$, then L_i^*, $1 \le i \le q$, can be chosen to be completely symmetric.*

Proof : The proof is given in Appendix A. □

The interest of this result is that $L_i^* = a_i^* I_t + b_i^* J_t$ can be found out by minimizing $q(a_1, b_1, ..., a_q, b_q) = \mathrm{tr}(L'C[\alpha]L)$ under the constraint $L'K = I_t$, where $L' = (L_1', ..., L_q')$ and $L_i = a_i I_t + b_i J_t$. Note that q is a quadratic form in $a_1, b_1, ..., a_q, b_q$.

Corollary 2 *Under the notations and assumptions of Proposition 1, if moreover $\mathcal{C}_{ij}\mathbb{1}_t = 0$, $1 \le i, j \le q$, then L_i^* can be chosen to be equal to $a_i^* I_t$ for some scalar a_i^*.*

Proof : By Proposition 1, L_i^* can be chosen to be equal to $a_i^* I_t + b_i^* J_t$. Since $\mathcal{C}_{ij} J_t = J_t \mathcal{C}_{ij} = 0$, $L_i' \mathcal{C}_{ij} L_j = a_i^* a_j^* \mathcal{C}_{ij}$ does not depend on b_i^* and b_j^* and the result follows. □

The following lemma, although straightforward, is useful to explicitly calculate L^*.

Lemma 3 *Let C be a nonnegative (v, v) matrix, γ be a k-vector, and $\gamma \mapsto L_\gamma$ be a linear mapping with L_γ a (v, w) matrix. Then, $q(\gamma) = \mathrm{tr}\left(L_\gamma' C L_\gamma\right)$ is convex quadratic in $\gamma_1, ..., \gamma_k$. Moreover, $q(\gamma^*)$ is a minimum of $q(\gamma)$ if and only if $d\, q/d\, \gamma\ (\gamma^*) = 0$.*

3.2 Some examples

Under the notations and assumptions of Proposition 1 and Corollary 2, i.e. assuming that $\mathcal{C}_{ij}$ are completely symmetric and that $\mathcal{C}_{ij} \mathbb{1}_t = 0$, we show how in some cases the matrix L^* can be obtained. The three first examples give a new presentation of known results established by Kushner (1997), Kunert and Martin (2000) and Bailey and Druilhet (2004). The last example will be used in this paper. We denote $c_{ij} = \mathrm{tr}(\mathcal{C}_{ij})$.

Example 1 : $q = 2$ and $K' = (I_t|0)$. By Corollary 2 and because $L^{*\prime} K = I_t$, $L^{*\prime}$ can be chosen to be equal to $(I_t| x^*\, I_t)$ for some $x^* \in \mathbb{R}$. By (6), x^* can be found by minimizing

$$q(x) = \mathrm{tr}(L' C[\alpha] L) = c_{11} + 2x\, c_{12} + x^2 c_{22}.$$

This quadratic function was used by Kushner (1997, Eq. 4.1). When $c_{22} \neq 0$, the minimum is obtained for $x^* = -c_{12}/c_{22}$ and we found $C[K'\alpha] = \mathcal{C}_{11} - \mathcal{C}_{12} \mathcal{C}_{22}^{+} \mathcal{C}_{21}$, the Schur-complement of $\mathcal{C}_{22}$.

Example 2: $q = 3$ and $K' = (I_t|0|0)$. As in example 1, $L^{*\prime}$ can be chosen to be equal to

$(I_t|x^*I_t|y^*I_t)$, where x^* and y^* minimize the quadratic function

$$q(x,y) = \mathrm{tr}(L'C[\alpha]L) = c_{11} + x^2c_{22} + y^2c_{33} + 2xc_{12} + 2yc_{13} + 2xyc_{23}.$$

This quadratic function was used by Kunert and Martin (2000, Proposition 3).

Example 3: $q = 2$ and $K' = (I_t|I_t)$. We have $L^{*\prime} = (x^*\, I_t|(1-x^*)\, I_t)$, where x^* minimizes

$$q(x) = x^2c_{11} + (1-x)^2c_{22} + 2x(1-x)c_{12}.$$

If $\mathcal{C}_{11} = \mathcal{C}_{22}$, then $x^* = \frac{1}{2}$ and $L^* = \frac{1}{2}K$. Therefore, $C[K'\alpha] = \frac{1}{4}K'C[\alpha]K$. This equation was obtained in a different way by Bailey and Druilhet (2004) in order to construct optimal designs for total effects under models without interactions.

Example 4: $q = 3$ and $K' = (I_t|I_t|I_t)$. We assume that $\mathcal{C}_{11} = \mathcal{C}_{22}$ and that $\mathcal{C}_{13} = \mathcal{C}_{23}$ (see Section 4.1). We no longer assume that $\mathcal{C}_{33}\mathbb{1}_t = 0$. We have $L^{*\prime} = (x_1^*\, I_t|x_2^*I_t|(1-x_1^*-x_2^*)I_t$ $+y^*J_t)$, where x_1^*, x_2^* and y^* minimize

$$\begin{aligned} q(x_1,x_2,y) \quad &= \quad (x_1^2+x_2^2)\, c_{11} + 2\, x_1\, x_2\, c_{12} - (x_1+x_2)^2(2\, c_{13} - c_{33}) \\ &\quad +2(x_1+x_2)(c_{13} - c_{33} - y\, \widetilde{c}_{33}) \\ &\quad +2\, y\, \widetilde{c}_{33} + t\, y^2\, \widetilde{c}_{33} + c_{33} \qquad (7) \end{aligned}$$

with $\widetilde{c}_{33} = \mathrm{tr}(J_t\mathcal{C}_{33}) = t^{-1}\mathrm{tr}(J_t\mathcal{C}_{33}J_t)$. Note that L^* does not satisfy $L^{*\prime}K = I_t$. This is in fact a simplified form of

$$\widetilde{L}^{*\,\prime} = (x_1^*\, I_t + u^*J_t|x_2^*I_t - (u^*+y^*)J_t|(1-x_1^*-x_2^*)I_t + y^*J_t)\,,$$

with $\widetilde{L}^{*\,\prime}K = I_t$, noting that the terms u^*J_t and $(u^*+y^*)J_t$ vanish in the expression of $\widetilde{L}^{*\,\prime}C[\alpha]\widetilde{L}^*$. By Lemma 3 and by symmetry of $q(x_1,x_2,y)$ in x_1, x_2, it is easy to see that

x_1^* and x_2^* can be chosen to be equal. Denote $x^* = x_1^* = x_2^*$. From (7), x^* and y^* can be found by minimizing

$$\begin{aligned} q\,(x,y) = \;& c_{33} + 4\,(c_{13} - c_{33})\,x + 2\widetilde{c}_{33}y + 2\,(c_{11} + c_{12} - 4c_{13} + 2c_{33})\,x^2 \\ & + t\widetilde{c}_{33}y^2 - 4\widetilde{c}_{33}xy \end{aligned} \tag{8}$$

and the minimum, $q(x^*, y^*)$, is equal to $\mathrm{tr}(C[K'\alpha])$.

4 Optimal circular crossover designs for total effects

From Kiefer (1975), a design d^* for which the information matrix $C_{d^*}[\phi]$ is completely symmetric and that maximizes the trace of $C_d[\phi]$ over all the designs d in $\Omega_{t,b,k}$ is universally optimal. In this section, we propose a method to construct universally optimal circular designs for total effects.

4.1 Upper bound of $\mathrm{tr}\, C_d[\phi]$

For a design d, the information matrix for the whole parameter $\theta' = (\tau', \lambda', \chi')$ in Model (1) is:

$$C_d[\theta] = \begin{pmatrix} T_d' \,\mathrm{pr}_{(B)}^{\perp}\, T_d & T_d' \,\mathrm{pr}_{(B)}^{\perp}\, L_d & T_d' \,\mathrm{pr}_{(B)}^{\perp}\, S_d \\ L_d' \,\mathrm{pr}_{(B)}^{\perp}\, T_d & L_d' \,\mathrm{pr}_{(B)}^{\perp}\, L_d & L_d' \,\mathrm{pr}_{(B)}^{\perp}\, S_d \\ S_d' \,\mathrm{pr}_{(B)}^{\perp}\, T_d & S_d' \,\mathrm{pr}_{(B)}^{\perp}\, L_d & S_d' \,\mathrm{pr}_{(B)}^{\perp}\, S_d \end{pmatrix} = \begin{pmatrix} \mathcal{C}_{d11} & \mathcal{C}_{d12} & \mathcal{C}_{d13} \\ \mathcal{C}_{d12}' & \mathcal{C}_{d22} & \mathcal{C}_{d23} \\ \mathcal{C}_{d13}' & \mathcal{C}_{d23}' & \mathcal{C}_{d33} \end{pmatrix}.$$

By circularity of the designs, $B'T_d = B'L_d$, $T_d'T_d = L_d'L_d$ and $T_d'S_d = L_d'S_d = S_d'S_d = S_d'L_d = S_d'T_d$ and therefore $\mathcal{C}_{d22} = \mathcal{C}_{d11}$ and $\mathcal{C}_{d23} = \mathcal{C}_{d13}$. As in Section 3.2, we define $c_{dij} = \mathrm{tr}(\mathcal{C}_{dij})$ and $\widetilde{c}_{dij} = \mathrm{tr}(J_t\, \mathcal{C}_{dij})$.

A design is said to be *symmetric* if all the blocks C_{dij} are completely symmetric, or equivalently, if $C_d[\theta]$ is invariant by any permutations of the treatment labels.

Proposition 4 *For any design d, the information matrix for total effects ϕ satisfies:*

$$\operatorname{tr} C_d[\phi] \le \min_{x,y} q_d(x,y)$$

where $q_d(x,y)$ is equal to $q(x,y)$ defined by (8). Equality holds if $C_{dij}[\theta]$ are completely symmetric for $1 \le i \le 3$ and $1 \le j \le 3$.

Proof: For a design d, we denote $\overline{C}_d[\theta] = \frac{1}{t!}\sum_{\sigma\in\mathcal{S}_t}(I_3\otimes P_\sigma)\, C_d[\theta]\,(I_3\otimes P'_\sigma)$ the symmetrized information matrix for θ. By construction, the corresponding matrices $\overline{\mathcal{C}}_{dij}$ are completely symmetric for all i and j. Moreover, $c_{dij} = \operatorname{tr}(\mathcal{C}_{dij}) = \operatorname{tr}(\overline{\mathcal{C}_{dij}})$ and $\widetilde{c}_{dij} = \operatorname{tr}(J_t\,\mathcal{C}_{d33}) = \operatorname{tr}(J_t\,\overline{\mathcal{C}_{d33}})$. Denote by $\overline{C}_d[\phi]$ the matrix obtained from $\overline{C}_d[\theta]$ by (4). By concavity of $C_d[\phi]$ with respect to $C_d[\theta]$, (see Pukelsheim, 1993 p. 77), and by Lemma 10 in Appendix A,

$$\overline{C_d}[\phi] \ge \frac{1}{t!}\sum_{\sigma} P_\sigma C_d[\phi] P'_\sigma$$

and $\operatorname{tr} C_d[\phi] \le \operatorname{tr}\overline{C_d}[\phi] = \min_{x,y} q_d(x,y)$. □

Note that $q_d(x,y)$, resp. $\overline{C}_d[\theta]$, correspond to $q(x,y)$, resp. $C[\alpha]$, of Example 4 of Section 3.2 and Proposition 1.

A design is called degenerate if $\widetilde{c}_{d33} = \operatorname{tr}(J_t\,\mathcal{C}_{d33}[\theta]) = 0$. A design is degenerate if and only if its sequences either contains a single treatment or has no treatment preceded by itself. For such designs, total effects are not estimable. For example, binary designs are degenerate. Note that the information matrix of a degenerate design is null and therefore such a design is not considered. The following lemma shows that the minimization of $q_d(x,y)$ may be reduce to the minimization of a one variable quadratic function.

Lemma 5 *Let d be a non-degenerate design. The values x^* and y^* that minimize $q_d(x,y)$ satisfy $2x^* - ty^* - 1 = 0$. Moreover, x^* minimizes $q_d(x)$, where*

$$q_d(x) = c_{d33} - \frac{1}{t}\widetilde{c}_{d33} + 4(c_{d13} - c_{d33} + \frac{1}{t}\widetilde{c}_{d33})x + 2(c_{d11} + c_{d12} - 4c_{d13} + 2c_{d33} - \frac{2}{t}\widetilde{c}_{d33})x^2, \quad (9)$$

and $q_d(x^*, y^*) = q_d(x^*)$.

Proof : by Lemma 3, $\partial q_d / \partial y \ (x^*, y^*) = 0$ at the minimum. For a non-degenerate design, this is equivalent to $2x^* - ty^* - 1 = 0$. Replacing y by $(2x - 1)/t$ in $q_d(x, y)$, we find $q_d(x, (2\,x - 1)/t) = q_d(x)$. Therefore, x^* necessarily minimizes $q_d(x)$. □

4.2 Decomposition over the subjects

It is well known that $C_d[\theta]$ is the sum of the information matrices C_{du} corresponding to Subject u. Denote by T_{du}, L_{du} and S_{du} the incidence matrices for Subject u. Thus, $T'_d = (T'_{d1} \mid ... \mid T'_{db})$, $L'_d = (L'_{d1} \mid ... \mid L'_{db})$, $S'_d = (S'_{d1} \mid ... \mid S'_{db})$ and:

$$C_d[\theta] = \sum_{u=1}^{b} C_{du}[\theta] = \sum_{u=1}^{b} \begin{pmatrix} T'_{du} Q_k T_{du} & T'_{du} Q_k L_{du} & T'_{du} Q_k S_{du} \\ L'_{du} Q_k T_{du} & L'_{du} Q_k L_{du} & L'_{du} Q_k S_{du} \\ S'_{du} Q_k T_{du} & S'_{du} Q_k L_{du} & S'_{du} Q_k S_{du} \end{pmatrix} \tag{10}$$

We decompose in the same way $\mathrm{tr}\,(\mathcal{C}_{dij})$ and $\mathrm{tr}\,(\mathcal{C}_{dij} J_t)$:

$$c_{dij} = \mathrm{tr}\,(\mathcal{C}_{dij}) = \sum_{u=1}^{b} c_{dij}^{(u)} \text{ and } \widetilde{c}_{dij} = \mathrm{tr}\,(\mathcal{C}_{dij} J_t) = \sum_{u=1}^{b} \widetilde{c}_{dij}^{(u)},$$

denoting by $c_{dij}^{(u)}$ and by $\widetilde{c}_{dij}^{(u)}$ the contributions of Subject u. For example, for $i = j = 1$ we have:

$$c_{d11} = \mathrm{tr}\,(\mathcal{C}_{d11}) = \mathrm{tr}\left(\sum_{u=1}^{b} T'_{du} Q_k T_{du}\right) = \sum_{u=1}^{b} c_{d11}^{(u)} \text{ with } c_{d11}^{(u)} = \mathrm{tr}(T'_{du} Q_k T_{du}).$$

Simplifications of these forms give:

$$c_{d11}^{(u)} = k - \frac{n_u^s}{k},\ c_{d12}^{(u)} = m_u - \frac{n_u^s}{k},\ c_{d13}^{(u)} = m_u - \frac{l_u}{k},\ c_{d33}^{(u)} = m_u - \frac{m_u^s}{k},\ \widetilde{c}_{d33}^{(u)} = m_u - \frac{m_u^2}{k},$$

with $n_u^s = \sum_{i=1}^t n_{ui}^2$, $m_u = \sum_{i=1}^t m_{ui}$, $m_u^s = \sum_{i=1}^t m_{ui}^2$, $l_u = \sum_{i=1}^t n_{ui} m_{ui}$, denoting by n_{ui} the number of periods where Subject u receives treatment i and by m_{ui} the number of times Treatment i is preceded by itself for Subject u. It follows that:

$$q_d(x) = \sum_{u=1}^{b} h_d^{(u)}(x)$$

where

$$h_d^{(u)}(x) = c_{d33}^{(u)} - \frac{1}{t}\widetilde{c}_{d33}^{(u)} + 4(c_{d13}^{(u)} - c_{d33}^{(u)} + \frac{1}{t}\widetilde{c}_{d33}^{(u)})x + 2(c_{d11}^{(u)} + c_{d12}^{(u)} - 4c_{d13}^{(u)} + 2c_{d33}^{(u)} - \frac{2}{t}\widetilde{c}_{d33}^{(u)})x^2.$$

Two sequences of treatments in two subjects u_1 and u_2 are said to be *equivalent* if $\left(n_{u_1}^s, m_{u_1}, m_{u_1}^s, l_{u_1}\right) = \left(n_{u_2}^s, m_{u_2}, m_{u_2}^s, l_{u_2}\right)$, which is the case if one sequence is obtained from the other one by relabelling the treatments or by a circular permutation of the periods. So, for given k and t, we can divide the set of all possible treatment sequences into K equivalence classes of treatments. Since n_u^s, m_u, m_u^s, l_u and $c_{dij}^{(u)}$ are the same for all u from one equivalent class ℓ, we change the notation and write n_ℓ^s, m_ℓ, m_ℓ^s, l_ℓ and $c_{dij}(\ell)$ instead. We define:

$$\begin{aligned} h_\ell(x) = & \left(m_\ell - \frac{m_\ell^s}{k} - \frac{\delta_\ell}{t}\right) + \frac{4}{k}\left(m_\ell^s - l_\ell + \frac{k\delta_\ell}{t}\right)x \\ & + 2\left\{(k - m_\ell) + \frac{2}{k}(2l_\ell - n_\ell^s - m_\ell^s) - \frac{2\delta_\ell}{t}\right\}x^2 \end{aligned}$$

where $\delta_\ell = m_\ell(1 - m_\ell/k)$. For a design d, we denote by $\pi_{d\ell}$ the proportion of subjects assigned to the class ℓ $(1 \le \ell \le K)$. So, we have:

$$q_d(x) = b\sum_{\ell=1}^{K} \pi_{d\ell}\, h_\ell(x). \tag{11}$$

4.3 Approximate designs

A design, also called exact design, is characterized, up to a subject permutation, by the proportions of subjects receiving each sequence of treatments. These proportions are necessarily multiple of $1/b$. If we remove this restriction, we obtain approximate designs, sometimes called continuous block designs. The information matrix associated to an approximate design d is defined similarly to (10) by

$$C_d[\theta] = b\sum_s \pi^{(ds)} C_s[\theta]$$

where $\pi^{(ds)}$ is the proportion of sequences receiving the sequence s and $C_s[\theta]$ is the information matrix associated to the sequence s. The information matrix $C_d[\phi]$ for total effects is obtained from $C_{ds}[\theta]$ by (4). Note that (11) also hold for approximate designs with

$$\pi_{d\ell} = \sum_{s\in\ell} \pi^{(ds)}.$$

The notion of symmetric or universally optimal design can be directly extended to approximate designs. Let d be an exact or approximate design, then the symmetrized design, denoted by $\bar{d}$, is the symmetric design such that $\pi_{d\ell} = \pi_{\bar{d}\ell}$ for any equivalent class ℓ. Note that $\overline{C}_d[\theta]$ defined in the proof of Proposition 4 is equal to $C_{\bar{d}}[\theta]$.

4.4 Construction of optimal approximate design

Our goal, in order to find an optimal approximate designs d^*, is to obtain a value x^* and proportions $\pi'_{d^*} = (\pi_{d^*1}, ..., \pi_{d^*K})$ such that:

$$q_{d^*}(x^*) = \max_{\pi_d} q_d^* \text{ with } q_d^* = \min_x q_d(x).$$

The following propositions characterizes universally optimal approximate designs.

Proposition 6 (Kunert and Martin, 2000) *Consider a symmetric approximate design $d^* \in \Omega_{t,b,k}$ and a point x^* such that the first derivative of q_{d^*} is zero. If we have also for all $1 \leq \ell \leq K$:*

$$bh_\ell(x^*) \leq q_{d^*}^*,$$

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

An optimal design for this one-dimensional problem may be obtained using one class or a mixture of two different classes (see Kushner, 1997). The following method (for given k and t) can be used in order to prove that a design d^* is optimal:

- If the optimal symmetric design d^* is generated by one treatment sequence ℓ_1, *i.e.* $q_{d^*}(x) = bh_{\ell_1}(x)$:

  - find x^* that minimizes h_{ℓ_1} and then the minimum $q^*_{d^*}$ of q_{d^*},

  - check that for $1 \leq \ell \leq K$, $bh_\ell(x^*) \leq q^*_{d^*}$ (cf. Prop. 6).

- If the optimal symmetric design d^* is generated by two treatment sequences, ℓ_1 and ℓ_2, *i.e.* $q_{d^*}(x) = b\{\pi_{d^*\ell_1}h_{\ell_1}(x) + \pi_{d^*\ell_2}h_{\ell_2}(x)\}$:

  - find an admissible intersection point x^* according to the definition of Kushner (1997), *i.e.* $h_{\ell_1}(x^*) = h_{\ell_2}(x^*)$ and

$$\frac{\partial h_{\ell_1}}{\partial x}(x^*)\,\frac{\partial h_{\ell_2}}{\partial x}(x^*) \leq 0,$$

  - find the minimum $q^*_{d^*} = q_{d^*}(x^*)$ of q_{d^*},

  - check that for $1 \leq \ell \leq K$, $bh_\ell(x^*) \leq q^*_{d^*}$ (cf. Prop. 6).

Note that the optimal proportions can be found by the following method.

$$\frac{\partial q_{d^*}}{\partial x}(x^*) = 0 \Longleftrightarrow b\left\{\pi_{d^*\ell_1}\frac{\partial h_{\ell_1}}{\partial x}(x^*) + \pi_{d^*\ell_2}\frac{\partial h_{\ell_2}}{\partial x}(x^*)\right\} = 0.$$

Denote $a_i = (\partial h_{\ell_i}/\partial x)(x^*)$, $i = 1, 2$. The optimal proportions are then:

$$\pi^*_{d^*\ell_1} = \frac{a_2}{a_2 - a_1} \text{ and } \pi^*_{d^*\ell_2} = \frac{a_1}{a_1 - a_2}.$$

When $t \geq k$, the following proposition show that the number of equivalent classes that possibly appear in an optimal design may be reduce to $k - 1$. Note that for a sequence ℓ with a single treatment can be neglected since $h_\ell(x) \equiv 0$. We denote by $\lfloor x \rfloor$ the integer part of x.

Proposition 7 *For $t \geq k$, if a symmetric approximate design is universally optimal, then the treatment sequences present in the design necessarily satisfies:*

1. *all the periods receiving the same treatment are contiguous in the sequence.*

2. *each treatment present in the sequence occurs $\lfloor k/v \rfloor$ or $\lfloor k/v \rfloor + 1$ times, where v is the number of different treatments present in the sequence. The number of treatments that occur $\lfloor k/v \rfloor + 1$ times is $k - v\lfloor k/v \rfloor$ and the number of treatments that occur $\lfloor k/v \rfloor$ times is $v(\lfloor k/v \rfloor + 1) - k$.*

The proof is given in Appendix B. When $k \geq t$, it is worth noting that for each value of v there is only one equivalent class of treatment sequences that may appear in the optimal design.

For a sequence ℓ satisfying Conditions 1 and 2 of Proposition 7, we have $m_{\ell i} = n_{\ell i} - 1$ for any treatment present in ℓ. Moreover the number v_ℓ of treatments present in ℓ is equal to $k - m_\ell$, $n_\ell^s = m_\ell^s + m_\ell + k$ and $l_\ell = m_\ell^s + m_\ell$. Therefore, h_ℓ simplify to:

$$h_\ell\left(x\right) = \left(m_\ell - \frac{m_\ell^s}{k} - \frac{\delta_\ell}{t}\right) + \frac{4}{k}\left(\frac{k\delta_\ell}{t} - m_\ell\right)x + 2\left\{\left(k - m_\ell\right)\left(\frac{k-2}{k}\right) - \frac{2\delta_\ell}{t}\right\}x^2.$$

The restriction $t \geq k$ in Proposition 7 is purely technical. We conjecture that the proposition is still valid for $t < k$. This conjecture has been checked for $k \leq 10$ (see Section 4.5).

4.5 Examples of optimal approximate designs and efficient exact designs.

We give here optimal designs in the sense of approximate design theory for several values of k. For $k = 3, 4$ we give explicit formulae. For $k = 5, ..., 10$, we present numerical results. For each situation, we also propose efficient or optimal designs generated by one sequence, i.e. design obtained from one sequence by considering all the treatment permutations.

1) *The case* $k = 3$. An optimal design in the sense of approximate theory is generated by one or two sequence in the following set of sequences:

Sequence	n_ℓ^s	m_ℓ	m_ℓ^s	l_ℓ
[1 2 3]	3	0	0	0
[1 1 2]	5	1	1	2

So the functions h_ℓ are given by (we identify each class with its value of m_ℓ):

$$h_0(x) = 2x^2 \text{ and } h_1(x) = \frac{2}{3}\left(1 - \frac{1}{t}\right) - \frac{4}{3}\left(1 - \frac{2}{t}\right)x + \frac{4}{3}\left(1 - \frac{2}{t}\right)x^2.$$

It is impossible in that case to satisfy Proposition 6 using only one of these two sequences. So, we must find an admissible intersection point x^*. Some algebra shows that:

$$x^* = \frac{\left\{(t-2)^2 + (t-1)(t+4)\right\}^{1/2} - (t-2)}{t+4}.$$

The proportions in the optimal design d^* are then:

$$\pi^*_{d^*0} = 1 - \pi^*_{d^*1} \text{ and } \pi^*_{d^*1} = \frac{3t}{t+4}\left\{1 - \frac{(t-2)}{t^{1/2}(2t-1)^{1/2}}\right\}.$$

The following table gives the optimal proportions for several values of t.

t	3	4	5	10	20	∞
Prop. [1 2 3]	0.046	0.067	0.079	0.100	0.111	0.121
Prop. [1 1 2]	0.954	0.933	0.921	0.899	0.889	0.879

The sequence [1 1 2] is dominant in this mixture. So it can be interesting in practice to use designs generated by only this sequence. The quality of such designs can be quantified by the classical Φ_p criteria. We know (see e.g. Druilhet, 2004) that when the information matrix is completely symmetric Φ_p does not depend on p. Thus we can derive the efficiency factor of a design $d \in \Omega_{t,b,3}$ generated by one sequence ℓ:

$$\text{Eff}(d) = \frac{b\, h_\ell^*}{q_{d^*}(x^*)} \text{ with } h_\ell^* = \min_x h_\ell(x).$$

Note that the efficiency factors given in approximate design theory are lower bounds of efficiency factors obtained in exact design theory. Numerical applications are given in the following table:

t	3	4	5	10	20	∞
Eff. [1 1 2]	0.989	0.985	0.982	0.976	0.974	0.971
Eff. [1 2 3]	0	0	0	0	0	0

2) *The case* $k = 4$. An optimal approximate design is generated by the sequence [1 1 2 2]. As an example we can consider, for $t = 4$, the optimal design such that:

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

Note that, by circularity, the design obtained by taking away one out of every two columns is universally optimal over all the designs in $\Omega_{4,6,4}$.

3) *The case* $k = 5$. The optimal design is generated by the following mixtures:

t	3	4	5	10	20	∞
Prop. [1 1 2 2 3]	0.167	0.250	0.300	0.400	0.450	0.500
Prop. [1 1 1 2 2]	0.833	0.750	0.700	0.600	0.550	0.500

The efficiencies of designs generated by only one sequence are:

t	3	4	5	10	20	∞
Eff. [1 1 2 2 3]	0.844	0.895	0.918	0.949	0.961	0.970
Eff. [1 1 1 2 2]	0.984	0.977	0.972	0.963	0.959	0.955

It can be observed that the design generated by [1 1 2 2 3] is more efficient than the one generated by [1 1 1 2 2] for $t > 18$.

4) *The case* $k = 6$. The optimal design is generated by the following mixtures:

t	3	4	5	6	7	≥ 8
Prop. [1 1 2 2 3 3]	0.400	0.628	0.775	0.878	0.954	1.000
Prop. [1 1 1 2 2 2]	0.600	0.372	0.225	0.122	0.046	

The efficiencies of designs generated by only one sequence are:

t	3	4	5	6	7	≥ 8
Eff. [1 1 2 2 3 3]	0.962	0.989	0.997	0.999	0.999	1.000
Eff. [1 1 1 2 2 2]	0.962	0.942	0.930	0.922	0.917	

We note that the sequence [1 1 2 2 3 3] is always more efficient than the sequence [1 1 1 2 2 2]. It also generates an optimal design when $t \geq 8$.

5) *The case* $k = 7$. The optimal design is generated by the following mixtures:

t	3	≥ 4
Prop. [1 1 1 2 2 3 3]	0.682	1.000
Prop. [1 1 1 1 2 2 2]	0.318	

The efficiencies of designs generated by only one sequence are:

t	3	≥ 4
Eff. [1 1 1 2 2 3 3]	0.994	1.000
Eff. [1 1 1 1 2 2 2]	0.938	

We note that he sequence [1 1 1 2 2 3 3] is always better than the sequence [1 1 1 1 2 2 2] and generates the optimal design when $t \geq 4$.

6) *The case* $k = 8, 9, 10$. We find that the optimal design is generated, for every t, by only one sequence. These optimal sequences are given by:

t	8	9	10
Sequence	[1 1 1 2 2 2 3 3]	[1 1 1 2 2 2 3 3 3]	[1 1 1 1 2 2 2 3 3 3]

5 Models with period effects

We consider the model:

$$y_{ij} = \alpha_j + \beta_i + \tau_{d(i,j)} + \lambda_{d(i,j-1)} + \chi_{d(i,j-1)d(i,j)} + \varepsilon_{ij}, \tag{12}$$

where α_j is the effects of the period j. We denote by A the corresponding incidence matrix. A *strongly symmetric design* is a design such that all the sequences belonging to the same equivalent class appear equally often. This notion is slightly more restrictive than the notion of symmetric design defined in Section 4.1 but corresponds to the design generated by one sequence or a mixture of sequences.

Proposition 8 *A strongly symmetric design which is universally optimal for total effects under Model (1) is also universally optimal under Model (12).*

Proof : The proof is a direct consequence of the following lemma. □

Lemma 9 *For a strongly symmetric design, the information matrix for the total effects is the same under models (1) and (12)*

Proof : Let d be a strongly symmetric design. The difficulty is that interaction and period effects are not orthogonal. Write $\tilde{\theta}' = (\tau'|\lambda'|\chi'|\alpha')$ and $\tilde{C}_d[\tilde{\theta}]$ the corresponding information matrix. We want to prove that $\tilde{C}_d[\tilde{K}'\tilde{\theta}] = C_d[K'\theta]$ where $\tilde{K}' = (I_t|I_t|I_t|0_{t\times k})$. We write:

$$\tilde{C}_d[\tilde{\theta}] = \begin{pmatrix} C_d[\theta] & C_{d12} \\ C_{d21} & C_{d22} \end{pmatrix}$$

where $C'_{d12} = (0_{k\times t}|0_{k\times t}|D'_d)$ with D_d the $(t \times k)$ matrix $(p_{d1}\mathbb{1}_t|p_{d2}\mathbb{1}_t|...|p_{dk}\mathbb{1}_t)$. The scalars p_{dj} depend only on the number of times a treatment is preceded by itself on period j. Note that the usual orthogonality condition (see Kunert, 1983) between interaction and period effects is $D_d = 0$ which is not the case here. The key point of the proof is that

$$Q_t\, D_d = 0. \tag{13}$$

Denote by $\tilde{L}^*_d$ a matrix such that $\tilde{C}_d[\tilde{K}'\tilde{\theta}] = \tilde{L}^*_d\,'\tilde{C}_d[\tilde{\theta}]\tilde{L}^*_d$ and $\tilde{L}^*_d\,'\tilde{K} = I_t$. We write $\tilde{L}^*_d\,' = (M^*_d\,'|N^*_d\,')$ where M_d is a $(3t \times t)$ matrix and N^*_d is a $(k \times k)$ matrix. It is easy to see that $C_d[K'\theta]\mathbb{1}_t = 0$ and, since $\tilde{C}_d[\tilde{K}'\tilde{\theta}] \le C_d[K'\theta]$, we have $\tilde{C}_d[\tilde{K}'\tilde{\theta}]\mathbb{1}_t = 0$. So, $C_d[\tilde{K}'\tilde{\theta}] = Q_t\tilde{C}_d[\tilde{K}'\tilde{\theta}]Q_t$. Therefore, $\tilde{L}^*$ can be chosen to be equal to $(M^*_dQ_t|N^*_dQ_t)$. For any permutation σ, we have $(I_3 \otimes P_\sigma)C_{d12} = C_{d12}$ and therefore, similarly to Proposition 1, it can be shown that the three $(t \times t)$ blocks of M^*_d can be chosen to be completely symmetric. So, $\tilde{L}^*_d\,'$ can be chosen to be equal to $(x^*_1Q_t|x^*_2Q_t|(1-(x^*_1+x^*_2))Q_t|N^*)$. Put $\tilde{L}' = (M'|N')$ with $M' = (x_1Q_t|x_2Q_t|(1-(x_1+x_2)Q_t)$. By (13), we have $M'C_{d12} = 0$ and then

$$\tilde{C}_d[\tilde{K}'\tilde{\theta}] = \min_{x_1,x_1,N} \tilde{L}'\tilde{C}[\tilde{\theta}]\tilde{L} = \min_{x_1,x_1,N}(M'C_d[\theta]M + N'C_{d12}N).$$

Since x_1, x_2 and N vary freely, N can be chosen to be equal to 0 and therefore, from Example 4 of Section 3.2 and the constraint on y obtained in Lemma 5, $M^*\,'\tilde{C}[\tilde{\theta}]M^* = L^*\,'C_d[\theta]L^*$. The result follows. $\square$

6 Acknowledgement

The authors are deeply grateful to the referees whose numerous comments and suggestions resulted in many improvements of the paper.

Appendix A

We give here the proof of Proposition 1, using the same notations and assumptions. We first establish a general result where we do not assume that the block matrices $\mathcal{C}_{ij}$ are completely symmetric. This result will also be used in Proposition 4.

Lemma 10 *Consider a permutation σ on $\{1, ..., t\}$ and denote by P_σ the corresponding permutation matrix. Denote by $C[K'\alpha]$, resp. $C_\sigma[K'\alpha]$, the matrices defined from $C[\alpha]$, resp. from $(I_q \otimes P_\sigma)C[\alpha](I_q \otimes P'_\sigma)$, by (4). We have:*

$$C_\sigma[K'\alpha] = P_\sigma \; C[K'\alpha] \; P'_\sigma.$$

Proof :

$$\begin{aligned} C_\sigma[K'\alpha] &= \min_{L \in \mathbb{R}^{qt\times t}: L'K = I_t} L'(I_q \otimes P_\sigma)C[\alpha](I_q \otimes P'_\sigma)L \quad \text{(by (15))}, \\ &= P_\sigma \left\{ \min_{L \in \mathbb{R}^{qt\times t}: L'K = I_t} P'_\sigma L'(I_q \otimes P_\sigma)C[\alpha](I_q \otimes P'_\sigma)LP_\sigma \right\} P'_\sigma. \end{aligned}$$

Put $L_\sigma = (I_q \otimes P'_\sigma)LP_\sigma$. Since the blocks of K are completely symmetric, $(I_q \otimes P'_\sigma)K = KP_\sigma$ and then $L'_\sigma K = I_t$. Because $L \leftrightarrow L_\sigma$ is a one to one mapping:

$$C_\sigma[K'\alpha] = P_\sigma \left(\min_{L_\sigma \in \mathbb{R}^{qt\times t}: L'_\sigma K = I_t} L'_\sigma C[\alpha] L_\sigma \right) P'_\sigma = P_\sigma \; C[K'\alpha] \; P'_\sigma. \tag{14}$$

□

We now assume that the block matrices $\mathcal{C}_{ij}$ are completely symmetric.

Step 1: We show that the set E of matrices L that minimize $L'C[\alpha]L$ under the constraint $L'K = I_t$ is an affine subspace. Consider L_a and L_b in E with $L_a \neq L_b$. It is sufficient to prove that $L_\gamma = \gamma L_a + (1-\gamma)L_b$ belongs to E for any real γ. We have $L'_\gamma K = I_t$ and

$$L'_\gamma C[\alpha] L_\gamma = \gamma^2 L'_a C[\alpha] L_a + (1-\gamma)^2 L'_b C[\alpha] L_b + \gamma(1-\gamma)(L'_a C[\alpha] L_b + L'_b C[\alpha] L_a).$$

This quadratic function in γ, whose coefficients are symmetric matrices, admits a minimum relative to the Loewner ordering at two distinct values of γ: $\gamma = 0$ and $\gamma = 1$.

Therefore it is constant. So, L_γ minimizes $L'C[\alpha]L$ for any $\gamma \in \mathbb{R}$.

Step 2: For any permutation σ on $\{1,...,t\}$ and denote by P_σ the corresponding permutation matrix. We want to prove that if $L^* \in E$, so does $L^*_\sigma = (I_q \otimes P'_\sigma)\, L^* P_\sigma$. Since $\mathcal{C}_{ij}$ and K_i are completely symmetric:

$$(I_q \otimes P_\sigma)C[\alpha](I_q \otimes P'_\sigma) = C[\alpha]. \tag{15}$$

By Lemma 10,

$$C[K'\alpha] = C_\alpha[K'\alpha] = P_\sigma[K'\alpha]P'_\sigma$$

for any permutation σ and $C[K'\alpha]$ is completely symmetric. From (14),

$$C[K'\alpha] = \min_{L_\sigma \in \mathbb{R}^{qt\times t}: L'_\sigma K = I_t} L'_\sigma C[\alpha] L_\sigma$$

and therefore $C[K'\alpha] = L^{*\prime}_\sigma C[\alpha] L^*_\sigma$.

Step 3: If $L^* \in E$ then, by steps 1 and 2, $\bar{L}^* = \frac{1}{t!}\sum_\sigma L^*_\sigma$ also belongs to E. By construction, $\bar{L}^*_i$ is completely symmetric and the proof is complete.

Appendix B

We prove here Proposition 7. From (11), we have $h_\ell\,(x) = A_\ell + B_\ell\, x + C_\ell\, x^2$ where: $A_\ell = m_\ell - \dfrac{m^s_\ell}{k} - \dfrac{\delta_\ell}{t}$, $B_\ell = \dfrac{4}{k}\left(m^s_\ell - l_\ell + \dfrac{k\delta_\ell}{t}\right)$ and $C_\ell = 2\left\{(k - m_\ell) + \dfrac{2}{k}\,(2l_\ell - n^s_\ell - m^s_\ell) - \dfrac{2\delta_\ell}{t}\right\}$.
We denote by v_ℓ the number of treatments present in the sequence ℓ. For any treatment i in ℓ, we have:

$$n_{\ell i} \geq m_{\ell i} + 1 \tag{16}$$

with equality if and only if treatment i occurs in consecutive periods. In general, each treatment appears in the sequence ℓ in several groups of consecutive periods. The number of such groups is $\gamma_i = n_{\ell i} - m_{\ell i}$. From ℓ, we construct a new sequence $\widetilde{\ell}$ as follows: in each group of consecutive periods receiving the same treatment, we replace the treatment that occurs in this group by a new treatment, such that in the new sequence $\widetilde{\ell}$, each treatment have only one group. For example $\ell = (1, 1, 2, 2, 2, 1, 3, 2, 2)$ gives $\widetilde{\ell} = (1, 1, 2, 2, 2, 4, 3, 5, 5)$. We have $m_{\widetilde{\ell}} = m_\ell$ and $v_{\widetilde{\ell}} = \gamma = k - m_\ell$, where $\gamma = \sum_i \gamma_i$ is the number of different groups in ℓ. The existence of $\tilde{\ell}$ is assured by the fact that $t \geq k$.

Lemma 11 *If d is a non degenerate design, then x_d^* that minimizes $q_d(x)$ is positive.*

Proof. From Lemma 3 we know that $h_\ell(x)$ is convex, then $C_\ell \geq 0$. So it is sufficient to show that $B_\ell \leq 0$. If $C_\ell = 0$ then $B_\ell = 0$ because we know that $h_\ell(x)$ admits a minimum. In that case, any x is a minimum and can be chosen to be positive. We assume now that $C_\ell > 0$. By (16), we have $l_\ell = \sum_i m_{\ell i} n_{\ell i} \geq m_\ell + m_\ell^s$ and then

$$B_\ell \leq \frac{4}{kt} m_\ell (k - m_\ell - t)$$

with equality if and only if the treatments present in the sequence are contiguous. In that case a sequence $\widetilde{\ell}$ satisfies $k - m_{\widetilde{\ell}} = v_{\widetilde{\ell}}$ and, since $t \geq v_{\widetilde{\ell}}$, $B_{\widetilde{\ell}} \leq 0$. Consider now a general sequence ℓ and denote by $\widetilde{\ell}$ its associated sequence with contiguous treatments. Since $m_{\widetilde{\ell}} = m_\ell$ we have also $\delta_\ell = \delta_{\widetilde{\ell}}$ and then $B_\ell \leq B_{\widetilde{\ell}}$ if and only if $(l_\ell - m_\ell^s) \geq \left(l_{\widetilde{\ell}} - m_{\widetilde{\ell}}^s\right)$. But:

$$l_\ell - m_\ell^s = \sum_{i=1}^{v_\ell} m_{\ell i} (n_{\ell i} - m_{\ell i}) \geq \sum_{i=1}^{v_\ell} m_{\ell i} = m_\ell = \sum_{i=1}^{v_{\widetilde{\ell}}} m_{\widetilde{\ell} i} \left(n_{\widetilde{\ell} i} - m_{\widetilde{\ell} i}\right) = l_{\widetilde{\ell}} - m_{\widetilde{\ell}}^s$$

with equality if and only if $\ell = \widetilde{\ell}$. So we have for every sequence ℓ: $B_\ell \leq B_{\widetilde{\ell}} \leq 0$. Therefore, by (11), $x_d^* = -(\sum_\ell \pi_{dl} B_\ell)/(2 \sum_\ell \pi_{dl} C_\ell)$ is non-negative since the denominator is positive for a non degenerate design. □

The proof of the Proposition 7 is given below.

Step 1: Let ℓ be a sequence containing v_ℓ different treatments numbered $1, ..., v_\ell$ and $\widetilde{\ell}$ the associated sequence defined above. We want to prove that $h_{\widetilde{\ell}}(x) > h_\ell(x)$ for all $x \geq 0$ and $\ell \neq \widetilde{\ell}$. The idea is to show that $A_\ell < A_{\widetilde{\ell}}$, $B_\ell < B_{\widetilde{\ell}}$ and $C_\ell < C_{\widetilde{\ell}}$. Note that $m_\ell = m_{\widetilde{\ell}}$ so $\delta_\ell = \delta_{\widetilde{\ell}}$. Then:

• We show that $C_\ell < C_{\widetilde{\ell}}$. It is equivalent to show that $(n_\ell^s - 2l_\ell + m_\ell^s) > \left(n_{\widetilde{\ell}}^s - 2l_{\widetilde{\ell}} + m_{\widetilde{\ell}}^s\right)$. Since $n_{\widetilde{\ell}i} = m_{\widetilde{\ell}i} + 1$ we have:

$$n_\ell^s - 2l_\ell + m_\ell^s = \sum_{i=1}^{v_\ell} (n_{\ell i} - m_{\ell i})^2 = \sum_{i=1}^{v_\ell} \gamma_i^2 \geq \sum_{i=1}^{v_\ell} \gamma_i = k - m_\ell = \sum_{i=1}^{v_{\widetilde{\ell}}} \left(n_{\widetilde{\ell}i} - m_{\widetilde{\ell}i}\right)^2$$

with equality if and only if $\ell = \widetilde{\ell}$. The result follows.

• From Lemma 11, we have $B_\ell < B_{\widetilde{\ell}}$.

• We show that $A_\ell < A_{\widetilde{\ell}}$. It is equivalent to show that $m_\ell^s > m_{\widetilde{\ell}}^s$. This follows from the fact that $\sum_i m_{\ell i} = \sum_i m_{\widetilde{\ell}i}$ and, for all i, $m_{\ell i} \geq m_{\widetilde{\ell}i}$ with equality if and only if $\ell = \widetilde{\ell}$.

Step 2: we have shown that an optimal sequence l necessary satisfy $l = \tilde{l}$. Consider now the set $\mathcal{L}_v$ of sequences ℓ having the same number $v_\ell = v$ of distinct treatments and such that $\ell = \widetilde{\ell}$. For a sequence $\ell \in \mathcal{L}_v$, $n_{\ell i} = m_{\ell i} - 1$ and $m_\ell = k - v$. So we have for every $\ell_1, \ell_2 \in \mathcal{L}_v$ (see Section 4.2):

$$h_{\ell_1}(x) - h_{\ell_2}(x) = \frac{1}{k}\left(m_{\ell_2}^s - m_{\ell_1}^s\right).$$

If there exists a sequence $\ell^* \in \mathcal{L}_v$ such that, for any sequence ℓ in $\mathcal{L}_v$ which is not equivalent to ℓ^*, $h_{\ell^*}(x) - h_\ell(x) > 0$, then only the sequences in the equivalent class corresponding to ℓ^* can belong to the optimal design. Such a sequence minimizes $m_\ell^s = \sum_{i=1}^v m_{\ell i}^2$ under the constraint $\sum_{i=1}^v m_{\ell i} = m_\ell$. So, it is well known that the $m_{\ell^* i}$ or equivalently the $n_{\ell^* i}$ $(i = 1, ..., v)$ must be as equal as possible: if k/v is an integer, a treatment present in ℓ^* occurs k/v times; otherwise, a treatment present in ℓ^* occurs either $\lfloor k/v \rfloor$ or $\lfloor k/v \rfloor + 1$ times in ℓ^* $\square$

7 Acknowledgement

The authors are deeply grateful to the referees whose many comments and suggestions resulted in many improvements of the paper.

References

Afsarinejad, K. & Hedayat, A. S. (2002). Repeated measurements designs for a model with self and simple mixed carryover effects. *J. Statist. Plann. Inference* 106 449–459.

Azaïs, J.-M. (1987). Design of experiments for studying intergenotypic competition. *J. Roy. Statist. Soc. Ser. B* 49 334–345.

Bailey, R. A. & Druilhet, P. (2004). Optimality of neighbor-balanced designs for total effects. *Ann. Statist.* 32 1650–1661.

Bailey, R. A. & Kunert, J. (2006). On optimal crossover designs when carryover effects are proportionnal to direct effects. *Biometrika* 93 613–625.

Druilhet, P. (2004). Conditions for optimality in experimental designs. *Linear Algebra Appl.* 388 147–157.

Gaffke, N. (1987). Further characterizations of design optimality and admissibility for partial parameter estimation in linear regression. *Ann. Statist.* 15 942–957.

Kempton, R. A., Ferris, S. J. & David, O. (2001). Optimal change-over designs when carry-over effects are proportional to direct effects of treatments. *Biometrika* 88 391–399.

Kiefer, J. (1975). Construction and optimality of generalized Youden designs. In *A survey of statistical design and linear models (Proc. Internat. Sympos., Colorado State Univ., Ft. Collins, Colo., 1973).* Amsterdam: North-Holland, 333–353.

Kunert, J. (1983). Optimal design and refinement of the linear model with applications to repeated measurements designs. *Ann. Statist.* 11 247–257.

Kunert, J. & Martin, R. J. (2000). On the determination of optimal designs for an interference model. *Ann. Statist.* 28 1728–1742.

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

Kushner, H. B. (1997). Optimal repeated measurements designs: the linear optimality equations. *Ann. Statist.* 25 2328–2344.

Pukelsheim, F. (1993). *Optimal design of experiments.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons Inc. A Wiley-Interscience Publication.