

Crossover designs based on type I orthogonal arrays for a self and simple mixed carryover effects model with correlated errors[☆]

A.S. Hedayat, Zhiwu Yan*

Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607-7045, USA

Received 15 November 2006; received in revised form 24 April 2007; accepted 25 September 2007

Available online 12 October 2007

Abstract

We investigate the performance of crossover designs based on type I orthogonal arrays for a self and simple mixed carryover effects model in the presence of correlated errors. Assuming that between-subject errors are independent while within-subject errors behave according to the stationary first-order autoregressive and moving average processes, analytical optimality results for 3-period designs are established and, as an illustration, numerical details for a number of 4-period cases are tabulated.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Crossover designs; Self carryover effects; Simple mixed carryover effects; Universal optimality; Type I orthogonal arrays

1. Introduction

In crossover designs, experimental subjects (units) are repeatedly exposed to a sequence of different or identical treatments. An observation on a subject may not only be affected by the treatment assigned most recently to this subject, but could also be affected by lingering effects of treatments that the subject received in earlier periods. Such lingering effects are known as carryover (or residual) effects.

The traditional model for crossover designs assumes that each treatment has a carryover effect that does not interact with the direct effect of the treatment applied in the next period. From many points of view, practical and theoretical, this seems to be the most plausible model and therefore has received tremendous attention in literature. Selected references include Hedayat and Afsarinejad (1978), Cheng and Wu (1980), Kunert (1983, 1984), Matthews (1987, 1990), Hedayat and Zhao (1990), Stufken (1991, 1996), Kushner (1997, 1998), and Hedayat and Yang (2003, 2004, 2005). But in certain situations good arguments can also be made for other models. For example in sensory experiments of medical interventions, it is more convincing to assume that the carryover effects depend on the treatment in the following period. To deal with such situations, Afsarinejad and Hedayat (2002) suggested an alternative model that contains two types of carryover effects, one is called a self carryover effect if a treatment is followed by itself and the other one

[☆] Research supported by National Science Foundation (NSF) Grant DMS-0603761, and National Institutes of Health (NIH) Grant P50-AT00155 (jointly supported by National Center for Complementary and Alternative Medicine, the Office of Dietary Supplements, the Office of Research on Women's Health, and National Institute of General Medicine). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF and the NIH.

* Corresponding author.

E-mail address: zwyan@math.uic.edu (Z. Yan).

is called a simple mixed carryover effect if a treatment is followed by any other treatment. Under this latter model, Kunert and Stufken (2002) proved that a totally balanced design is universally optimal for all direct treatment effects over the class of all designs with t treatments, n subjects and p periods, as long as $t \geq 3$ and $3 \leq p \leq 2t$. They assumed, like most other researchers have done in literature, that all observations on the n subjects are uncorrelated. In many situations, such a simple assumption is technically desirable but practically questionable, in the context of crossover designs it is often reasonable to argue that the error terms are actually correlated if they correspond to measurements on the same subject. One way to look at such a correlation structure is to consider the p observations for a subject a short time series and hence the entire set of observations as n time series.

The main purpose of the present paper is to identify a class of optimal and efficient designs-type I orthogonal arrays for the self and simple mixed carryover effects model with correlated errors. We shall be particularly concerned with the cases where the errors within each subject follow the stationary first-order autoregressive (AR(1)) and moving average (MA(1)) processes, but our method can also be used in conjunction with correlation structures of other forms. Due to certain technical hurdles we will only consider 3- and 4-period cases in this article, but fortunately designs with small number of periods are attractive to practitioners and thus tend to be used more often in practice. For both AR(1) and MA(1) structures with any correlation coefficient for which stationarity holds, we prove that type I orthogonal arrays are universally optimal for direct effects over all designs with $p = 3$ and $t \geq 3$. For the case with $p = 4$, numerical results based on $t = 4, 5$ and 7 show that type I orthogonal arrays are highly efficient and are very likely to be optimal. The important case of 2-period designs is not addressed here because the within-subject error structure has no impact on the optimality of designs. The interested reader is referred to Afsarinejad and Hedayat (2002) for more details about the 2-period studies under the homoscedastic model.

The statistical tool we employ in this paper was developed by Kunert and Martin (2000a) who generalized the results of Kushner (1997). The tool relates the problem of identifying designs with maximal traces of information matrices to a minimax problem of a sequence of bivariate quadratic functions, which are solely determined by design parameters and the error variance–covariance structure.

2. Notations and preliminary results

We start with the following notations.

By 1_a and 0_a we will mean the $a \times 1$ vectors of 1s and 0s, respectively. By I_a , $0_{a \times b}$ and $J_{a \times b}$ we will mean the $a \times a$ identity matrix, the $a \times b$ matrix with all entries equal to 0 and the $a \times b$ matrix with all entries equal to 1, respectively. For simplicity, we may drop the subscripts if doing so will not cause any confusion. Finally, we will denote the class of all crossover designs with t treatments, n subjects and p periods by $\Omega_{t,n,p}$.

We assume that the response y_{ij} obtained from subject j at period i , where $j \in \{1, \dots, n\}$ and $i \in \{1, \dots, p\}$, can be written as

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

Here:

- (i) $d(i,j) \in \{1, \dots, t\}$ denotes the treatment assigned to subject j in period i ;
- (ii) α_i is the effect due to the i th period;
- (iii) β_j is the effect due to the j th subject;
- (iv) $\tau_{d(i,j)}$ is the direct effect due to treatment $d(i,j)$;
- (v) $\rho_{d(i-1,j)}$ is the simple mixed carryover effect due to treatment $d(i-1,j)$ (with $\rho_{d(0,j)} = 0$);
- (vi) $\gamma_{d(i-1,j)}$ is the self carryover effect due to treatment $d(i-1,j)$ (with $\gamma_{d(0,j)} = 0$) and
- (vii) ε_{ij} is the non-observable random error term.

All effects (except for ε_{ij}) considered in model (1) are assumed to be fixed but unknown. We also assume that the errors between different subjects are uncorrelated, while the errors associated with the p observations within each subject have mean zero and a common variance–covariance matrix $\sigma^2 A$, where σ^2 is an unknown scalar and A is a known $p \times p$ positive definite matrix, to be specified later.

In matrix notation, model (1) can be written as

$$Y = P\alpha + U\beta + T_d\tau + M_d\rho + S_d\gamma + \varepsilon, \quad (2)$$

where $Y = (y_{11}, y_{21}, \dots, y_{pn})^T$, $\alpha = (\alpha_1, \dots, \alpha_p)^T$, $\beta = (\beta_1, \dots, \beta_n)^T$, $\tau = (\tau_1, \dots, \tau_t)^T$, $\rho = (\rho_1, \dots, \rho_t)^T$, $\gamma = (\gamma_1, \dots, \gamma_t)^T$, $\varepsilon = (\varepsilon_{11}, \varepsilon_{21}, \dots, \varepsilon_{pn})^T$; the matrices $P = 1_n \otimes I_p$, $U = I_n \otimes 1_p$, $T_d = (T_{d1}^T, \dots, T_{dn}^T)^T$, $M_d = (M_{d1}^T, \dots, M_{dn}^T)^T$ and $S_d = (S_{d1}^T, \dots, S_{dn}^T)^T$ are, respectively, the design matrices of the period, subject, direct treatment, simple mixed carryover and self carryover effects; while T_{du} , M_{du} and S_{du} are the corresponding design matrices for a single subject u , $1 \leq u \leq n$; and $\otimes$ denotes the Kronecker product. Our assumptions on the errors imply that $\text{Cov}(\varepsilon) = \sigma^2(I_n \otimes A)$.

For any $a \times b$ matrix X , we define $\omega^\perp(X) = I_a - X(X^T X)^- X^T$, where $(X^T X)^-$ is a generalized inverse of $X^T X$, thus $\omega^\perp(X)$ stands for the projection onto the space spanned by the vectors that are orthogonal to the column space of X . Then the information matrix $C_d(\tau)$ for the direct treatment effects τ can be written as

$$C_d(\tau) = T_d^T (I_n \otimes A^{-1/2}) \omega^\perp((I_n \otimes A^{-1/2})[P, U, M_d, S_d])(I_n \otimes A^{-1/2}) T_d,$$

where $A^{-1/2}$ is the $p \times p$ matrix with the property $A^{-1/2} A^{-1/2} = A^{-1}$. Note that because $T_d 1_t$ and $U 1_n$ are equal, we have that the row and column sums of $C_d(\tau)$ are all equal to zero for any design $d \in \Omega_{t,n,p}$.

We are interested in optimal designs for the direct treatment effects τ (and suppose that all elementary treatment contrasts are of equal interest). It follows from Kiefer's (1975) Proposition 1 that a design d^* for which the information matrix $C_{d^*}(\tau)$ is completely symmetric and maximizes $\text{Tr}[C_d(\tau)]$, the trace of $C_d(\tau)$, over all $d \in \Omega_{t,n,p}$ is universally optimal for τ . Complete symmetry of a matrix X means that it can be written as $X = a_1 I + a_2 J$, where a_1 and a_2 are real numbers.

Following Kunert and Stufken (2002), we shall determine an upper bound for the information matrix $C_d(\tau)$ and show that this upper bound can be attained by designs with certain properties. Let $\tilde{C}_d(\tau)$ be the information matrix for τ under the model obtained by deleting $P\alpha$ in (2), i.e.,

$$\tilde{C}_d(\tau) = T_d^T (I_n \otimes A^{-1/2}) \omega^\perp((I_n \otimes A^{-1/2})[U, M_d, S_d])(I_n \otimes A^{-1/2}) T_d.$$

Then as in earlier work by Kunert (1983), we have $C_d(\tau) \leq \tilde{C}_d(\tau)$ in the Loewner sense, i.e., $\tilde{C}_d(\tau) - C_d(\tau)$ is non-negative definite, with equality if and only if the following orthogonality condition is satisfied

$$T_d^T (I_n \otimes A^{-1/2}) \omega^\perp((I_n \otimes A^{-1/2})[U, M_d, S_d])(I_n \otimes A^{-1/2}) P = 0. \quad (3)$$

Before giving sufficient conditions on a design under which (3) is fulfilled, we need the following terminology.

A design $d \in \Omega_{t,n,p}$ is said to be

- (a) *Invariant on the periods for the direct effects* if each treatment in d appears equally often in each of the p periods.
- (b) *Invariant on the periods for the self carryover effects* if for each treatment in d the self carryover effect due to the treatment appears equally often in each of the last $p - 1$ periods.
- (c) *Invariant on the periods for the simple mixed carryover effects* if for each treatment in d the simple mixed carryover effect due to the treatment appears equally often in each of the last $p - 1$ periods.

Now we are ready to present

Lemma 1. *If a design $d \in \Omega_{t,n,p}$ is invariant on the periods for the direct effects, the self and simple mixed carryover effects, then Eq. (3) holds for any $p \times p$ positive definite matrix A .*

Proof. Since A is positive definite, so is A^{-1} . Let λ_{ij} be the (i, j) th entry of A^{-1} , then it can be verified that the 2×2 matrix $D = \begin{pmatrix} \lambda_{11} & \sum_{k=2}^p \lambda_{1k} \\ \sum_{k=2}^p \lambda_{k1} & \sum_{k=2}^p \sum_{l=2}^p \lambda_{kl} \end{pmatrix}$ is also positive definite. In fact, $D = Q^T A^{-1} Q$, where $Q^T = \begin{pmatrix} 1 & 0_{1 \times (p-1)} \\ 0 & 1_{1 \times (p-1)} \end{pmatrix}$. Consequently, there is a set of x_i and y_i , for $i = 1, \dots, p$, such that they are solutions to the following system of linear equations:

$$\begin{cases} \lambda_{11}x_i + (\sum_{k=2}^p \lambda_{1k})y_i = \lambda_{1i}, \\ (\sum_{k=2}^p \lambda_{k1})x_i + (\sum_{k=2}^p \sum_{l=2}^p \lambda_{kl})y_i = \sum_{k=2}^p \lambda_{ki}. \end{cases}$$

Define

$$G = \begin{pmatrix} x_1 & x_2 & \dots & x_p \\ y_1 J_{(p-1) \times 1} & y_2 J_{(p-1) \times 1} & \dots & y_p J_{(p-1) \times 1} \end{pmatrix},$$

then it can be verified that

- (i) the first row of the matrix $A^{-1}(I_p - G)$ is $0_{1 \times p}$; and
- (ii) all column sums of $A^{-1}(I_p - G)$ are zeros.

Let $A = 1_n \otimes G$ and $B = P - A = 1_n \otimes (I_p - G)$.

Since design d is invariant on the periods for the direct effects, we have

$$T_d^T (I_n \otimes A^{-1}) B = T_d^T \{1_n \otimes (A^{-1}(I_p - G))\} = \left(\sum_{j=1}^n T_{dj}^T \right) \{A^{-1}(I_p - G)\} = 0_{t \times p}.$$

Similarly, due to the invariance properties of the self and simple mixed carryover effects, we have

$$\begin{aligned} ((I_n \otimes A^{-1/2})[U, M_d, S_d])^T (I_n \otimes A^{-1/2}) B &= [U, M_d, S_d]^T \{1_n \otimes (A^{-1}(I_p - G))\} \\ &= 0_{(n+2t) \times p}. \end{aligned}$$

Let $H = \begin{pmatrix} x_1 J_{n \times 1} & x_2 J_{n \times 1} & \dots & x_p J_{n \times 1} \\ (y_1 - x_1) J_{2t \times 1} & (y_2 - x_2) J_{2t \times 1} & \dots & (y_p - x_p) J_{2t \times 1} \end{pmatrix}$, then we obtain

$$[U, M_d, S_d]H = A,$$

and thus

$$(I_n \otimes A^{-1/2})A = (I_n \otimes A^{-1/2})[U, M_d, S_d]H,$$

which implies that the column span of $(I_n \otimes A^{-1/2})A$ is in that of $(I_n \otimes A^{-1/2})[U, M_d, S_d]$ and therefore

$$\omega^\perp((I_n \otimes A^{-1/2})[U, M_d, S_d])(I_n \otimes A^{-1/2})A = 0.$$

Accordingly we have

$$\begin{aligned} T_d^T (I_n \otimes A^{-1/2}) \omega^\perp((I_n \otimes A^{-1/2})[U, M_d, S_d])(I_n \otimes A^{-1/2})P \\ &= T_d^T (I_n \otimes A^{-1/2}) \omega^\perp((I_n \otimes A^{-1/2})[U, M_d, S_d])(I_n \otimes A^{-1/2})(A + B) \\ &= T_d^T (I_n \otimes A^{-1/2}) \omega^\perp((I_n \otimes A^{-1/2})[U, M_d, S_d])(I_n \otimes A^{-1/2})B \\ &= T_d^T (I_n \otimes A^{-1})B \\ &= 0. \quad \square \end{aligned}$$

Following [Kunert and Martin \(2000a\)](#) and [Kunert and Stufken \(2002\)](#), we can decompose $\tilde{C}_d(\tau)$ as follows:

$$\begin{aligned} \tilde{C}_d(\tau) &= T_d^T (I_n \otimes A^{-1/2}) \omega^\perp((I_n \otimes A^{-1/2})[U, M_d, S_d])(I_n \otimes A^{-1/2})T_d \\ &= C_{d11} - C_{d12} C_{d22}^- C_{d12}^T - (C_{d13} - C_{d12} C_{d22}^- C_{d23}) \\ &\quad \times (C_{d33} - C_{d23}^T C_{d22}^- C_{d23})^- (C_{d13} - C_{d12} C_{d22}^- C_{d23})^T, \end{aligned}$$

where

$$\begin{aligned} C_{d11} &= T_d^T (I_n \otimes W) T_d, & C_{d12} &= T_d^T (I_n \otimes W) M_d, & C_{d22} &= M_d^T (I_n \otimes W) M_d, \\ C_{d13} &= T_d^T (I_n \otimes W) S_d, & C_{d23} &= M_d^T (I_n \otimes W) S_d, & C_{d33} &= S_d^T (I_n \otimes W) S_d, \end{aligned}$$

and

$$W = A^{-1} - (1_p^T A^{-1} 1_p)^{-1} A^{-1} 1_p 1_p^T A^{-1}.$$

Let $B_t = \omega^\perp(1_t)$, and for any design $d \in \Omega_{t,n,p}$ we define $c_{dij} = \text{Tr}[B_t C_{dij} B_t] = \text{Tr}[B_t C_{dij}]$ for $1 \leq i \leq j \leq 3$. Following the proof of Proposition 2 in [Kunert and Martin \(2000a\)](#), with minor modifications, we obtain $\text{Tr}[\tilde{C}_d(\tau)] \leq q_d^*$ for any design $d \in \Omega_{t,n,p}$, with equality if all matrices C_{dij} , $1 \leq i \leq j \leq 3$, are completely symmetric, where q_d^* is defined by the following four cases:

1. if $c_{d22}c_{d33} - c_{d23}^2 > 0$, then $q_d^* = c_{d11} - (c_{d12}^2 c_{d33} - 2c_{d12}c_{d13}c_{d23} + c_{d13}^2 c_{d22}) / (c_{d22}c_{d33} - c_{d23}^2)$;
2. if $c_{d22}c_{d33} - c_{d23}^2 = 0$ and $c_{d22} > 0$, then $q_d^* = c_{d11} - c_{d12}^2 / c_{d22}$;
3. if $c_{d22} = 0$ and $c_{d33} > 0$, then $q_d^* = c_{d11} - c_{d13}^2 / c_{d33}$; and
4. if $c_{d22} = c_{d33} = 0$, then $q_d^* = c_{d11}$.

In all, it follows that the inequalities

$$\text{Tr}[C_d(\tau)] \leq \text{Tr}[\tilde{C}_d(\tau)] \leq q_d^*$$

hold for every design $d \in \Omega_{t,n,p}$, and that the equalities

$$\text{Tr}[C_d(\tau)] = \text{Tr}[\tilde{C}_d(\tau)] = q_d^*$$

hold if (3) holds and if all matrices C_{dij} , $1 \leq i \leq j \leq 3$, are completely symmetric. We shall soon see that a class of designs with these properties is the class of type I orthogonal arrays of strength 2, denoted by $\text{OA}_I(n, p, t, 2)$. An $\text{OA}_I(n, p, t, 2)$ is a $p \times n$ array with entries from $\{1, \dots, t\}$ such that any $2 \times n$ sub-array contains all $t(t-1)$ ordered pairs without repetition equally often. Clearly, such an array exists only if $p \leq t$ and if n is an integer multiple of $t(t-1)$. It is also well known that an $\text{OA}_I(t(t-1), p, t, 2)$ exists for any $p \leq t$ if t is a prime power (see [Rao, 1961](#)).

Type I orthogonal arrays were introduced by [Rao \(1961\)](#). Such arrays have been recently revealed to possess many desirable properties, including balance, high efficiency and optimality (see [Majumdar and Martin, 2004](#)). In the presence of correlated errors (especially in crossover designs), type I orthogonal arrays seem to be quite promising to produce efficient designs. Unfortunately, very little research has been done to exploit these arrays in applications beyond the traditional model.

If a crossover design d^* uses an $\text{OA}_I(n, p, t, 2)$ with columns corresponding to subjects and rows corresponding to periods, then obviously it is invariant on the periods for the direct effects, the self and simple mixed carryover effects and thus by Lemma 1 Eq. (3) holds. Since such a design has no pairs of consecutive identical treatments on the same subject, thus $S_{d^*} = 0$, then C_{d^*13} , C_{d^*23} and C_{d^*33} are all matrices of zeros, therefore, $C_{d^*}(\tau) = \tilde{C}_{d^*}(\tau) = C_{d^*11} - C_{d^*12} C_{d^*22}^{-1} C_{d^*12}^T$. Note also that $M_{d^*} = (I_n \otimes L) T_{d^*}$, where L is the $p \times p$ matrix with all $(i+1, i)$ th entries equal to 1 and 0 otherwise, then $C_{d^*11} = T_{d^*}^T (I_n \otimes W) T_{d^*}$, $C_{d^*12} = T_{d^*}^T (I_n \otimes (WL)) T_{d^*}$, and $C_{d^*22} = T_{d^*}^T (I_n \otimes (L^T W L)) T_{d^*}$, thus by [Martin and Eccleston \(1998\)](#), we conclude that C_{d^*11} , C_{d^*12} , C_{d^*22} and $C_{d^*}(\tau)$ are all completely symmetric. If we can further show that $q_{d^*}^*$ is maximized by d^* , then d^* would accordingly be universally optimal.

In what follows, we proceed along the lines of [Kunert and Martin \(2000a\)](#) to deduce an upper bound for q_d^* that is independent of any design d . By writing

$$c_{d11}^{(u)} = \text{Tr}[B_t (T_{du}^T W T_{du})], \quad c_{d12}^{(u)} = \text{Tr}[B_t (T_{du}^T W M_{du})],$$

$$c_{d13}^{(u)} = \text{Tr}[B_t (T_{du}^T W S_{du})], \quad c_{d22}^{(u)} = \text{Tr}[B_t (M_{du}^T W M_{du})],$$

$$c_{d23}^{(u)} = \text{Tr}[B_t (M_{du}^T W S_{du})], \quad c_{d33}^{(u)} = \text{Tr}[B_t (S_{du}^T W S_{du})],$$

for $u \in \{1, \dots, n\}$, we then obtain $c_{dij} = \sum_{u=1}^n c_{dij}^{(u)}$, for $1 \leq i \leq j \leq 3$.

Note that the $c_{dij}^{(u)}$'s are determined by the sequence of treatments assigned to subject u . We say that two treatment sequences are equivalent, if one can be transformed to the other by relabelling the treatments. It is obvious that two equivalent treatment sequences give the same $c_{dij}^{(u)}$'s. Therefore, for given t and p , we can partition the set of all

possible treatment sequences into K equivalence classes $s_1, \dots, s_K$, such that $c_{dij}^{(u)}$'s are the same for all subject u receiving a sequence from a given class. We denote by $\pi_{d\ell}$ the proportion of subjects receiving sequences from the class s_ℓ , $1 \leq \ell \leq K$, in a given design $d \in \Omega_{t,n,p}$. We also define $c_{ij}(\ell) = c_{dij}^{(u)}$, where u is any subject receiving a sequence from the ℓ th class s_ℓ . Then we obtain

$$c_{dij} = n \sum_{\ell=1}^K \pi_{d\ell} c_{ij}(\ell) \quad (4)$$

for $1 \leq i \leq j \leq 3$. This implies that the bound q_d^* of any design $d \in \Omega_{d,n,p}$ is determined by the proportions $\pi_{d\ell}$. However, q_d^* is a non-linear function of the $\pi_{d\ell}$, which makes the maximization of q_d^* difficult. The problem is linearized by introducing the function

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

By the Proposition 3 in [Kunert and Martin \(2000a\)](#), we have that $q_d^* \leq q_d(x, y)$ for every x and y and that there is at least one point, (x^*, y^*) , say, such that $q_d^* = q_d(x^*, y^*)$. We will see in Section 4 that the determination of (x^*, y^*) is a critical step in the identification of optimal designs.

For the ℓ th equivalent class s_ℓ , $1 \leq \ell \leq K$, we define

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

and obtain

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

which is a linear combination of the $h_\ell(x, y)$. We immediately have

Lemma 2. For any design $d \in \Omega_{t,n,p}$, we have the following inequality:

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

Proof.

$$\begin{aligned} q_d^* &\leq \min_{x,y} q_d(x, y) = \min_{x,y} \left\{ n \sum_{\ell=1}^K \pi_{d\ell} h_\ell(x, y) \right\} \\ &\leq n \min_{x,y} \left\{ \sum_{\ell=1}^K \pi_{d\ell} \left[\max_{\ell} h_\ell(x, y) \right] \right\} \\ &= n \min_{x,y} \max_{\ell} h_\ell(x, y). \quad \square \end{aligned}$$

3. Main results for AR(1) correlation structure

To determine $\min_{x,y} \max_{\ell} h_\ell(x, y)$, we need to calculate, for $\ell \in \{1, \dots, K\}$, $c_{ij}(\ell)$'s which by definition depend, through the matrix W , on the covariance matrix A . It is therefore necessary to specify the within-subject correlation structure. Throughout this section we assume that the errors within each subject follow a stationary first-order autoregressive (AR(1)) process. Then the (i, j) th entry of the matrix A can be written as $\lambda^{|i-j|}/(1 - \lambda^2)$, $1 \leq i, j \leq p$, with $-1 < \lambda < 1$ being the correlation coefficient. We also assume that λ is known, which allows us to calculate the weighted least-square estimate for direct treatment effects.

Table 1
Equivalence classes and $c_{ij}(\ell)$'s for $p = 3$ and $t \geq 3$ (AR(1))

ℓ	Sequence	$c_{11}(\ell)$	$c_{12}(\ell)$	$c_{13}(\ell)$	$c_{22}(\ell)$	$c_{23}(\ell)$	$c_{33}(\ell)$
1	[111]	0	0	0	0	0	$\frac{2(t-1)}{t(3-\lambda)}$
2	[112]	$\frac{4}{3-\lambda}$	$-\frac{2}{3-\lambda}$	$\frac{1+\lambda}{3-\lambda}$	$\frac{2(t-1)}{t(3-\lambda)}$	$-\frac{(t-1)(1+\lambda)}{t(3-\lambda)}$	$\frac{2(t-1)(1+\lambda)}{t(3-\lambda)}$
3	[121]	$\frac{4(1+\lambda)}{3-\lambda}$	$-\frac{3(1+\lambda)}{3-\lambda}$	0	$\frac{2t(2+\lambda)-2}{t(3-\lambda)}$	0	0
4	[122]	$\frac{4}{3-\lambda}$	$-\frac{1+\lambda}{3-\lambda}$	$\frac{1-\lambda}{3-\lambda}$	$\frac{2(t-1)(1+\lambda)}{t(3-\lambda)}$	$\frac{1+\lambda}{t(3-\lambda)}$	$\frac{2(t-1)}{t(3-\lambda)}$
5	[123]	$\frac{2(3+\lambda)}{3-\lambda}$	$-\frac{2(1+\lambda)}{3-\lambda}$	0	$\frac{2t(2+\lambda)-2}{t(3-\lambda)}$	0	0

Note that

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

For any fixed p and t , the computation of $c_{ij}(\ell)$'s for all equivalence classes can be done with the help of MAPLE. However, with large p the number K of the equivalence classes increases rapidly and $h_\ell(x, y)$'s become more complicated to deal with. For example, when $p = t = 5$ there are 52 classes of equivalent sequences, which makes it almost impossible to find an analytical solution for $\min_{x,y} \max_{\ell} h_\ell(x, y)$. In fact, we will see in the sequel that even for $p = 4$ an analytical solution for $\min_{x,y} \max_{\ell} h_\ell(x, y)$ is unlikely to be available and has to be obtained by computer search. In the present paper we will only consider the cases with $p = 3$ and 4, which are very popular in practical applications.

For the case $p = 3$ and $t \geq 3$, there are five different classes of equivalent sequences. The representative sequences and the corresponding $c_{ij}(\ell)$'s are given in Table 1.

The following important lemma builds on the results in Table 1.

Lemma 3. Let $x^* = t(1+\lambda)/(2t+t\lambda-1)$ with $\lambda \in (-1, 1)$, then for $p = 3$ and any $t \geq 3$,

$$h_5(x^*, -1) = \min_{x,y} \max_{\ell \in \{1, \dots, 5\}} h_\ell(x, y).$$

Proof. Straightforward calculations and calculus arguments yield

$$\begin{aligned} & h_5(x^*, -1) - h_1(x^*, -1) \\ &= \frac{2(3t^2 + 2t^2\lambda - 1)}{t(3-\lambda)(2t+t\lambda-1)} \geq 0 \quad \text{if } -1 < \lambda < 1. \\ & h_5(x^*, -1) - h_2(x^*, -1) \\ &= \frac{2t(1+\lambda)[t^2(3-\lambda-\lambda^2) + t(3+8\lambda+2\lambda^2) - (4+3\lambda)] + 2}{t(3-\lambda)(2t+t\lambda-1)^2} \\ &\geq \begin{cases} \frac{2t(1+\lambda)(t^2+3t-7)+2}{t(3-\lambda)(2t+t\lambda-1)^2} \geq 0 & \text{if } 0 \leq \lambda < 1; \\ \frac{2t(1+\lambda)(3t^2-3t-4)+2}{t(3-\lambda)(2t+t\lambda-1)^2} \geq 0 & \text{if } -1 < \lambda < 0. \end{cases} \end{aligned}$$

$$\begin{aligned}
& h_5(x^*, -1) - h_3(x^*, -1) \\
&= \frac{2(3t + t\lambda + \lambda - 1)}{(3 - \lambda)(2t + t\lambda - 1)} \geq 0 \quad \text{if } -1 < \lambda < 1. \\
& h_5(x^*, -1) - h_4(x^*, -1) \\
&= \frac{2t[t^2(3 + \lambda - 2\lambda^2 - \lambda^3) + t(3 + 10\lambda + 8\lambda^2 + 2\lambda^3) - (4 + 4\lambda + \lambda^2)] + 2}{t(3 - \lambda)(2t + t\lambda - 1)^2} \\
&\geq \begin{cases} \frac{2t(t^2 + 3t - 9) + 2}{t(3 - \lambda)(2t + t\lambda - 1)^2} \geq 0 & \text{if } 0 \leq \lambda < 1; \\ \frac{2t(t^2 - t - 4) + 2}{t(3 - \lambda)(2t + t\lambda - 1)^2} \geq 0 & \text{if } -1 < \lambda < 0. \end{cases}
\end{aligned}$$

Thus we obtain

$$h_5(x^*, -1) = \max_{\ell \in \{1, \dots, 5\}} h_\ell(x^*, -1) \geq \min_{x, y} \max_{\ell \in \{1, \dots, 5\}} h_\ell(x, y).$$

On the other hand, it can be verified that $h_5(x, y)$ is minimized at $(x^*, -1)$, therefore

$$\max_{\ell \in \{1, \dots, 5\}} h_\ell(x, y) \geq h_5(x, y) \geq h_5(x^*, -1),$$

which implies that

$$\min_{x, y} \max_{\ell \in \{1, \dots, 5\}} h_\ell(x, y) \geq h_5(x^*, -1),$$

hence we have

$$h_5(x^*, -1) = \min_{x, y} \max_{\ell \in \{1, \dots, 5\}} h_\ell(x, y). \quad \square$$

We are now ready to formulate the main result for the case $p = 3$ and $t \geq 3$.

Theorem 1. Under model (1) with errors within each subject following a stationary first-order autoregressive process, for $p = 3$, $t \geq 3$ and any $\lambda \in (-1, 1)$, a cross-over design d^* based on an $\text{OA}_I(n, 3, t, 2)$ with columns corresponding to subjects and rows corresponding to periods is universally optimal in $\Omega_{t, n, 3}$ for the direct treatment effects.

Proof. The complete symmetry of $C_{d^*}(\tau)$ has already been established in the previous section, it remains to show that d^* maximizes the trace of $C_d(\tau)$. Since design d^* is based on an $\text{OA}_I(n, 3, t, 2)$, thus all n subjects receive treatment sequences equivalent to the same sequence [123], i.e., the 5th class in Table 1. Therefore by Eq. (4), we have

$$c_{d^*ij} = n \sum_{\ell=1}^5 \pi_{d^*\ell} c_{ij}(\ell) = n c_{ij}(5).$$

Observe that $c_{d^*22} c_{d^*33} - c_{d^*23}^2 = 0$ and $c_{d^*22} > 0$, thus by definition, we obtain

$$q_{d^*}^* = c_{d^*11} - c_{d^*12}^2 / c_{d^*22} = n \{c_{11}(5) - c_{12}^2(5) / c_{22}(5)\}.$$

In addition, $h_5(x^*, -1) = \min_{x, y} h_5(x, y) = c_{11}(5) - c_{12}^2(5) / c_{22}(5)$, thus by Lemma 3

$$q_{d^*}^* = n h_5(x^*, -1) = n \min_{x, y} \max_{\ell \in \{1, \dots, 5\}} h_\ell(x, y).$$

Hence by Lemma 2, we obtain $q_{d^*}^* \geq q_d^*$, for any $d \in \Omega_{t, n, 3}$.

We then have, for any $d \in \Omega_{t, n, 3}$,

$$\text{Tr}[C_{d^*}(\tau)] = q_{d^*}^* \geq q_d^* \geq \text{Tr}[C_d(\tau)]. \quad \square$$

Table 2
Numerical values of $n \min_{x,y} \max_{\ell} h_{\ell}(x, y)$ for $p = 4$ (AR(1))

λ	−0.99	−0.9	−0.8	−0.7	−0.6	−0.5	−0.4
$t = 4, n = 12$	31.202	30.689	30.366	30.266	30.345	30.560	30.871
$t = 5, n = 20$	52.024	51.164	50.619	50.447	50.576	50.934	51.460
$t = 7, n = 42$	109.299	107.481	106.322	105.946	106.209	106.965	108.085
λ	−0.3	−0.2	−0.1	0	0.1	0.2	0.3
$t = 4, n = 12$	31.251	31.684	32.173	32.727	33.371	34.134	35.050
$t = 5, n = 20$	52.106	52.850	53.689	54.643	55.749	57.053	58.612
$t = 7, n = 42$	109.474	111.081	112.902	114.975	117.372	120.190	123.542
λ	0.4	0.5	0.6	0.7	0.8	0.9	0.99
$t = 4, n = 12$	36.154	37.486	39.088	41.005	43.292	46.015	48.912
$t = 5, n = 20$	60.485	62.737	65.435	68.658	72.496	77.061	81.912
$t = 7, n = 42$	127.551	132.351	138.087	144.922	153.045	162.693	172.935

For $p = 4$ and $t \geq 4$, there are 15 equivalence classes. The representative sequences and $c_{ij}(\ell)$'s for the 15 classes are given in Table 4 in the Appendix. While the procedure described in the proof of Lemma 3 works nicely for the case $p = 3$, it becomes quite cumbersome for the present situation. However, we are able to obtain numerical solutions of $\min_{x,y} \max_{\ell} h_{\ell}(x, y)$ for fixed λ values with the help of MATLAB. Table 2 shows numerical results (rounded to the third decimal position) of $n \min_{x,y} \max_{\ell} h_{\ell}(x, y)$ for $p = 4$, $t = 4, 5, 7$ and a broad range of selected λ values respectively. Since design efficiencies of type I orthogonal arrays do not depend on n , the number of subjects, thus without loss of generality we simply took $n = t(t - 1)$ in Table 2. Straightforward computations showed that, for all λ considered here, the trace (also rounded to the third decimal position) of the information matrix of the corresponding type I orthogonal array was equal to $n \min_{x,y} \max_{\ell} h_{\ell}(x, y)$. This implies an efficiency of more than 0.999. It could even be hopeful that these designs are in deed optimal or highly efficient for any $\lambda \in (-1, 1)$.

4. Main results for MA(1) correlation structure

In a first-order moving average (MA(1)) correlation structure, only errors from adjacent periods are correlated. The within-subject dispersion matrix can be written as $A = I_p + \lambda N_p$, where λ is the correlation coefficient and lies in the interval $(-\frac{1}{2}, \frac{1}{2})$ for stationarity, and N_p is a $p \times p$ matrix with all $(i, i + 1)$ th and $(i + 1, i)$ th entries equal to 1 and 0 otherwise.

If we take $x^* = t/(2t - t\lambda + \lambda - 1)$ with $\lambda \in (-\frac{1}{2}, \frac{1}{2})$ in Lemma 3, then we can show that the corresponding result also holds for MA(1) correlation structure and consequently optimality of an $OA_I(n, 3, t, 2)$ will follow for $p = 3$ and $t \geq 3$. To avoid heavy algebra as shown in Lemma 3, here we use a more succinct argument as follows. First it can be shown that when $p = 3$

$$W = \frac{1}{4v - 1} \begin{pmatrix} 2v & -1 & 1 - 2v \\ -1 & 2 & -1 \\ 1 - 2v & -1 & 2v \end{pmatrix} \quad (5)$$

holds for both AR(1) and MA(1) correlation structures with $v = 1/(1 + \lambda)$ for AR(1) and $v = 1 - \lambda$ for MA(1) (see Martin and Pooladsaz, 2005). We also note that the information matrix $C_d(\tau)$ depends on the covariance matrix A only through W , thus a type I orthogonal array that is universally optimal under AR(1) would also be optimal under MA(1). We then immediately have

Theorem 2. Under model (1) with errors within each subject following a stationary first-order moving average process, for $p = 3$, $t \geq 3$ and any $\lambda \in (-\frac{1}{2}, \frac{1}{2})$, a crossover design d^* based on an $OA_I(n, 3, t, 2)$ with columns corresponding to subjects and rows corresponding to periods is universally optimal in $\Omega_{t,n,3}$ for the direct treatment effects.

Table 3

Numerical values of $n \min_{x,y} \max_{\ell} h_{\ell}(x, y)$ for $p = 4$ (MA(1))

λ	−0.49	−0.4	−0.3	−0.2	−0.1	−0
$t = 4, n = 12$	33.254	32.670	32.283	32.157	32.296	32.727
$t = 5, n = 20$	55.433	54.469	53.837	53.644	53.896	54.643
$t = 7, n = 42$	116.434	114.430	113.133	112.764	113.344	114.975
λ	0.1	0.2	0.3	0.4	0.49	
$t = 4, n = 12$	33.514	34.785	36.813	40.324	46.984	
$t = 5, n = 20$	55.991	58.160	61.622	67.618	79.018	
$t = 7, n = 42$	117.892	122.567	130.024	142.956	167.602	

For $p = 4$ and $t \geq 4$, the $c_{ij}(\ell)$'s for the 15 equivalence classes are given in Table 5 in the Appendix. Table 3 shows numerical results of $n \min_{x,y} \max_{\ell} h_{\ell}(x, y)$ for $p = 4$ and $t = 4, 5, 7$, respectively.

A pattern similar to that in Section 3 was observed, that is, for all λ considered here, the trace of the information matrix of the corresponding type I orthogonal array was equal to $n \min_{x,y} \max_{\ell} h_{\ell}(x, y)$. Therefore type I orthogonal arrays are at least highly efficient for $p = 4$ and $t = 4, 5, 7$ for these λ values, and again these designs are very likely to be optimal or highly efficient for any $\lambda \in (-\frac{1}{2}, \frac{1}{2})$.

5. Discussion

We have shown that in the self and simple mixed carryover effects model, when $t \geq p = 3$, type I orthogonal arrays are universally optimal for AR(1) with any $\lambda \in (-1, 1)$, and for MA(1) with $\lambda \in (-\frac{1}{2}, \frac{1}{2})$; when $p = 4$ and $t = 4, 5, 7$, the efficiencies of type I orthogonal arrays are at least 0.999 for both AR(1) and MA(1) with selected λ values running through the entire intervals that apply. In the case of the traditional model, however, the optimality and efficiency of type I orthogonal arrays depend on the value of λ . For example (see Kunert and Martin, 2000b), when $t \geq p = 3$, type I orthogonal arrays are optimal only if $\lambda > 1/t$ for AR(1), and $\lambda > 1/(t + 1)$ for MA(1); when $p = t = 4$ and under AR(1), the efficiency of type I orthogonal arrays is more than 0.9933 if $\lambda \geq -(\sqrt{2} - 1)$, and is 1 if $\lambda > 0.1193$ and it drops rapidly for more negative λ , approaching 0.7917 for $\lambda \rightarrow -1$.

Numerical results for $p = 4$ and any $t > 7$ can be easily obtained for both correlation structures in a similar manner and high efficiency (at least 0.999) of type I orthogonal arrays can also be established as expected (details are omitted here). So it is fairly safe to conjecture that a design based on an $OA_I(n, 4, t, 2)$ is indeed optimal for both AR(1) and MA(1) as long as $t \geq 4$.

The special case $\lambda = 0$ for both correlation structures corresponds to the homoscedastic model, for which our results can also be justified by Kunert and Stufken (2002). This is so because the class of type I orthogonal arrays is a subclass of totally balanced designs.

Observe that an $OA_I(n, p', t, 2)$ can be obtained by deleting the last $p - p'$ rows of an $OA_I(n, p, t, 2)$ ($p' < p$), thus if one initially embarks on a 4-period (or 3-period) design based on an $OA_I(n, 4, t, 2)$ (or $OA_I(n, 3, t, 2)$), but has to stop after three or two (or two) periods, one would still be using a highly efficient (indeed optimal) design (see Afsarinejad and Hedayat, 2002 for the optimality of an $OA_I(n, 2, t, 2)$).

An obvious drawback of designs using type I orthogonal arrays is that such designs do not allow the assessment of the self carryover effects and thus the appropriateness of the model assumptions about the carryover effects. If this is a primary concern in practice, one should consider other appropriate designs. However, in many applications the restriction on the number of subjects would make such an assessment difficult in any case and the use of these designs needs to be based, at least in part, on prior judgment and an act of faith.

Type I orthogonal arrays exist only if n , the number of subjects, is an integer multiple of $t(t - 1)$, which may be unacceptably large in practice unless t is small. However, as pointed out by Majumdar and Martin (2004), designs found still have two useful aspects. Firstly, it gives a lower bound for the criterion values for designs with fewer subjects, and secondly it suggests the structure that an efficient design with fewer subjects is likely to need. This leads the experimenter to designs that may be utilized in real applications.

Table 4
Equivalence classes and $c_{ij}(\ell)$'s for $p = 4$ and $t \geq 4$ (AR(1)).

ℓ	sequence	$c_{11}(\ell)$	$c_{12}(\ell)$	$c_{13}(\ell)$
1	[1111]	0	0	0
2	[1112]	$\frac{3-\lambda}{2-\lambda}$	$-\frac{3-\lambda}{4-2\lambda}$	$\frac{1}{2-\lambda}$
3	[1121]	$\frac{3+\lambda+\lambda^2-\lambda^3}{2-\lambda}$	$-\frac{4+3\lambda-\lambda^3}{4-2\lambda}$	$\frac{1+\lambda+\lambda^2-\lambda^3}{4-2\lambda}$
4	[1122]	$\frac{4-4\lambda+3\lambda^2-\lambda^3}{2-\lambda}$	$\frac{-2+\lambda-2\lambda^2+\lambda^3}{4-2\lambda}$	$\frac{4-4\lambda+3\lambda^2-\lambda^3}{4-2\lambda}$
5	[1123]	$\frac{5-2\lambda+2\lambda^2-\lambda^3}{2-\lambda}$	$\frac{-3-\lambda-\lambda^2+\lambda^3}{4-2\lambda}$	$\frac{2-\lambda+2\lambda^2-\lambda^3}{4-2\lambda}$
6	[1211]	$\frac{3+\lambda+\lambda^2-\lambda^3}{2-\lambda}$	$\frac{-2-\lambda-\lambda^2+\lambda^3}{2-\lambda}$	$\frac{1-2\lambda+\lambda^2}{4-2\lambda}$
7	[1212]	$\frac{4+4\lambda-\lambda^2-\lambda^3}{2-\lambda}$	$\frac{-6-7\lambda+\lambda^2+2\lambda^3}{4-2\lambda}$	0
8	[1213]	$\frac{5+2\lambda-\lambda^3}{2-\lambda}$	$\frac{-5-5\lambda+2\lambda^3}{4-2\lambda}$	0
9	[1221]	$\frac{4}{2-\lambda}$	$\frac{-2}{2-\lambda}$	$\frac{1}{2-\lambda}$
10	[1222]	$\frac{3-\lambda}{2-\lambda}$	$\frac{-1-2\lambda+\lambda^2}{4-2\lambda}$	$\frac{2-3\lambda+\lambda^2}{4-2\lambda}$
11	[1223]	$\frac{5-\lambda}{2-\lambda}$	$\frac{-3-2\lambda+\lambda^2}{4-2\lambda}$	$\frac{1}{2-\lambda}$
12	[1231]	$\frac{5+\lambda+\lambda^2-\lambda^3}{2-\lambda}$	$\frac{-4-3\lambda+\lambda^3}{4-2\lambda}$	0
13	[1232]	$\frac{5+2\lambda-\lambda^3}{2-\lambda}$	$\frac{-4-7\lambda+2\lambda^2+\lambda^3}{4-2\lambda}$	0
14	[1233]	$\frac{5-2\lambda+2\lambda^2-\lambda^3}{2-\lambda}$	$\frac{-2-3\lambda+\lambda^3}{4-2\lambda}$	$\frac{2-3\lambda+\lambda^2}{4-2\lambda}$
15	[1234]	$\frac{6+\lambda^2-\lambda^3}{2-\lambda}$	$\frac{-3-5\lambda+\lambda^2+\lambda^3}{4-2\lambda}$	0
		$c_{22}(\ell)$	$c_{23}(\ell)$	$c_{33}(\ell)$
1	[1111]	0	0	$\frac{(t-1)(3-\lambda)}{t(4-2\lambda)}$
2	[1112]	$\frac{(t-1)(3-\lambda)}{t(4-2\lambda)}$	$-\frac{t-1}{t(2-\lambda)}$	$\frac{2(t-1)}{t(2-\lambda)}$
3	[1121]	$\frac{t(6+\lambda^2-\lambda^3)+(-4+4\lambda-3\lambda^2+\lambda^3)}{t(4-2\lambda)}$	$\frac{t(-1-\lambda-\lambda^2+\lambda^3)+(2-\lambda+2\lambda^2-\lambda^3)}{t(4-2\lambda)}$	$\frac{(t-1)(3+\lambda+\lambda^2-\lambda^3)}{t(4-2\lambda)}$
4	[1122]	$\frac{(t-1)(3+\lambda+\lambda^2-\lambda^3)}{t(4-2\lambda)}$	$\frac{t(-1-\lambda-\lambda^2+\lambda^3)+(2+3\lambda-\lambda^3)}{t(4-2\lambda)}$	$\frac{t(6+\lambda^2-\lambda^3)+(-4-4\lambda+\lambda^2+\lambda^3)}{t(4-2\lambda)}$
5	[1123]	$\frac{t(6+\lambda^2-\lambda^3)+(-4+4\lambda-3\lambda^2+\lambda^3)}{t(4-2\lambda)}$	$\frac{t(-1-\lambda-\lambda^2+\lambda^3)+(2-\lambda+2\lambda^2-\lambda^3)}{t(4-2\lambda)}$	$\frac{(t-1)(3+\lambda+\lambda^2-\lambda^3)}{t(4-2\lambda)}$
6	[1211]	$\frac{t(3+\lambda+\lambda^2-\lambda^3)-2}{t(2-\lambda)}$	$\frac{t(-1+2\lambda-\lambda^2)+2}{t(4-2\lambda)}$	$\frac{(t-1)(3-\lambda)}{t(4-2\lambda)}$
7	[1212]	$\frac{t(7+5\lambda-2\lambda^3)+(-3+\lambda)}{t(4-2\lambda)}$	0	0
8	[1213]	$\frac{t(7+5\lambda-2\lambda^3)+(-3+\lambda)}{t(4-2\lambda)}$	0	0
9	[1221]	$\frac{t(6+\lambda^2-\lambda^3)+(-4-4\lambda+\lambda^2+\lambda^3)}{t(4-2\lambda)}$	$\frac{t(-1-2\lambda+\lambda^2)+(2+3\lambda-\lambda^3)}{t(4-2\lambda)}$	$\frac{(t-1)(3+\lambda+\lambda^2-\lambda^3)}{t(4-2\lambda)}$
10	[1222]	$\frac{(t-1)(3+\lambda+\lambda^2-\lambda^3)}{t(4-2\lambda)}$	$\frac{2-\lambda+2\lambda^2-\lambda^3}{t(4-2\lambda)}$	$\frac{(t-1)(4-4\lambda+3\lambda^2-\lambda^3)}{t(4-2\lambda)}$
11	[1223]	$\frac{t(6+\lambda^2-\lambda^3)+(-4-4\lambda+\lambda^2+\lambda^3)}{t(4-2\lambda)}$	$\frac{t(-1-2\lambda+\lambda^2)+(2+3\lambda-\lambda^3)}{t(4-2\lambda)}$	$\frac{(t-1)(3+\lambda+\lambda^2-\lambda^3)}{t(4-2\lambda)}$
12	[1231]	$\frac{t(9+\lambda+2\lambda^2-2\lambda^3)+(-3+\lambda)}{t(4-2\lambda)}$	0	0
13	[1232]	$\frac{t(9+\lambda+2\lambda^2-2\lambda^3)+(-3+\lambda)}{t(4-2\lambda)}$	0	0
14	[1233]	$\frac{t(3+\lambda+\lambda^2-\lambda^3)-2}{t(2-\lambda)}$	$\frac{1}{t(2-\lambda)}$	$\frac{(t-1)(3-\lambda)}{t(4-2\lambda)}$
15	[1234]	$\frac{t(9+\lambda+2\lambda^2-2\lambda^3)+(-3+\lambda)}{t(4-2\lambda)}$	0	0

Table 5

Equivalence classes and $c_{ij}(\ell)$'s for $p = 4$ and $t \geq 4(\text{MA}(1))$.

ℓ	Sequence	$c_{11}(\ell)$	$c_{12}(\ell)$	$c_{13}(\ell)$
1	[1111]	0	0	0
2	[1112]	$\frac{3-4\lambda}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-3+4\lambda}{2(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{1}{2-\lambda}$
3	[1121]	$\frac{3-2\lambda-\lambda^2}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-4+\lambda+3\lambda^2}{2(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{1+\lambda^2}{2(2-\lambda)(1-\lambda-\lambda^2)}$
4	[1122]	$\frac{2-3\lambda}{1-\lambda-\lambda^2}$	$\frac{-1+\lambda}{2(1-\lambda-\lambda^2)}$	$\frac{2-3\lambda}{2(1-\lambda-\lambda^2)}$
5	[1123]	$\frac{5-7\lambda+\lambda^2}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-3+2\lambda+\lambda^2}{2(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{1-\lambda}{2(1-\lambda-\lambda^2)}$
6	[1211]	$\frac{3-2\lambda-2\lambda^2}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-1}{1-\lambda-\lambda^2}$	$\frac{1-3\lambda}{2(2-\lambda)(1-\lambda-\lambda^2)}$
7	[1212]	$\frac{2+\lambda}{1-\lambda-\lambda^2}$	$\frac{-3-2\lambda}{2(1-\lambda-\lambda^2)}$	0
8	[1213]	$\frac{5-3\lambda-\lambda^2}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-5}{2(2-\lambda)(1-\lambda-\lambda^2)}$	0
9	[1221]	$\frac{4}{2-\lambda}$	$\frac{-2}{2-\lambda}$	$\frac{1}{2-\lambda}$
10	[1222]	$\frac{3-4\lambda}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-1-\lambda+2\lambda^2}{2(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{1-2\lambda}{2(1-\lambda-\lambda^2)}$
11	[1223]	$\frac{5-6\lambda-2\lambda^2}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-3+\lambda+4\lambda^2}{2(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{1}{2-\lambda}$
12	[1231]	$\frac{5-4\lambda-3\lambda^2}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-4+\lambda+3\lambda^2}{2(2-\lambda)(1-\lambda-\lambda^2)}$	0
13	[1232]	$\frac{5-3\lambda-\lambda^2}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-4-3\lambda+5\lambda^2}{2(2-\lambda)(1-\lambda-\lambda^2)}$	0
14	[1233]	$\frac{5-7\lambda+\lambda^2}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-1-\lambda}{2(1-\lambda-\lambda^2)}$	$\frac{1-2\lambda}{2(1-\lambda-\lambda^2)}$
15	[1234]	$\frac{6-6\lambda-\lambda^2}{(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{-3-2\lambda+3\lambda^2}{2(2-\lambda)(1-\lambda-\lambda^2)}$	0
		$c_{22}(\ell)$	$c_{23}(\ell)$	$c_{33}(\ell)$
1	[1111]	0	0	$\frac{(t-1)(3-4\lambda)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$
2	[1112]	$\frac{(t-1)(3-4\lambda)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{1-t}{t(2-\lambda)}$	$\frac{2(t-1)}{t(2-\lambda)}$
3	[1121]	$\frac{t(6-6\lambda-\lambda^2)-4+8\lambda-3\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{t(-1-\lambda^2)+2-3\lambda+\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{(t-1)(3-2\lambda-\lambda^2)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$
4	[1122]	$\frac{(t-1)(3-2\lambda-\lambda^2)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{t(-1-\lambda^2)+2+\lambda-\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{t(6-6\lambda-\lambda^2)-4+\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$
5	[1123]	$\frac{t(6-6\lambda-\lambda^2)-4+8\lambda-3\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{t(-1-\lambda^2)+2-3\lambda+\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{(t-1)(3-2\lambda-\lambda^2)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$
6	[1211]	$\frac{t(3-2\lambda-\lambda^2)+2+2\lambda+2\lambda^2}{t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{t(-1+3\lambda)+2-2\lambda-2\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{(t-1)(3-4\lambda)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$
7	[1212]	$\frac{t(7-2\lambda-2\lambda^2)-3+4\lambda}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	0	0
8	[1213]	$\frac{t(7-2\lambda-2\lambda^2)-3+4\lambda}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	0	0
9	[1221]	$\frac{t(6-6\lambda-\lambda^2)-4+\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{t(-1-\lambda+2\lambda^2)+2+\lambda-\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{(t-1)(3-2\lambda-\lambda^2)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$
10	[1222]	$\frac{(t-1)(3-2\lambda-\lambda^2)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{1-\lambda}{2t(1-\lambda-\lambda^2)}$	$\frac{(t-1)(2-3\lambda)}{2t(1-\lambda-\lambda^2)}$
11	[1223]	$\frac{t(6-6\lambda-\lambda^2)-4+\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{t(-1-\lambda+2\lambda^2)+2+\lambda-\lambda^2}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{(t-1)(3-2\lambda-\lambda^2)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$
12	[1231]	$\frac{t(9-8\lambda-2\lambda^2)-3+4\lambda}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	0	0
13	[1232]	$\frac{t(9-8\lambda-2\lambda^2)-3+4\lambda}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	0	0
14	[1233]	$\frac{t(3-2\lambda-\lambda^2)+2+2\lambda+2\lambda^2}{t(2-\lambda)(1-\lambda-\lambda^2)}$	$\frac{1}{t(2-\lambda)}$	$\frac{(t-1)(3-4\lambda)}{2t(2-\lambda)(1-\lambda-\lambda^2)}$
15	[1234]	$\frac{t(9-8\lambda-2\lambda^2)-3+4\lambda}{2t(2-\lambda)(1-\lambda-\lambda^2)}$	0	0

Finally, only two correlation structures, AR(1) and MA(1), are considered in this paper. It is not only intuitively appealing, but also just impractical to make too complicated assumptions to address optimal design problems. Since the AR(1) and MA(1) structures are not introduced until Sections 3 and 4, so the method of this paper may be used for other correlation structures as well. In the worst case scenario when we have no any prior information about the correlation structure in advance, type I orthogonal arrays are extremely appealing since they are optimal for any correlation structure in the class of all binary designs (the results of Kunert and Martin, 2000b can be easily extended to the model considered in this paper) and are very likely to be optimal or highly efficient in the entire class for a variety of correlation structures.

Acknowledgments

The authors would like to thank the editor, associate editor and two referees for their valuable comments and suggestions. Our thanks are also due to Professor R.J. Martin for pointing out the simple argument to reach the conclusion of Theorem 2.

Appendix

Equivalence classes and $c_{ij}(\ell)$'s for $p = 4$ and $t \geq 4$ are shown in Tables 4 (AR(1)) and 5 (MA(1)).

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.
- Cheng, C.-S., Wu, C.-F., 1980. Balanced repeated measurements designs. *Ann. Statist.* 8, 1272–1283.
- Hedayat, A.S., Afsarinejad, K., 1978. Repeated measurements designs II. *Ann. Statist.* 6, 619–628.
- Hedayat, A.S., Yang, M., 2003. Universal optimality of balanced uniform crossover designs. *Ann. Statist.* 31, 978–983.
- Hedayat, A.S., Yang, M., 2004. Universal optimality for selected crossover designs. *J. Amer. Statist. Assoc.* 99, 461–466.
- Hedayat, A.S., Yang, M., 2005. Optimal and efficient crossover designs for comparing test treatments with a control treatment. *Ann. Statist.* 33, 915–943.
- Hedayat, A.S., Zhao, W., 1990. Optimal two-period repeated measurements designs. *Ann. Statist.* 18, 1817–1828.
- Kiefer, J., 1975. Construction and optimality of generalized Youden designs. In: Srivastava, J.N. (Ed.), *A Survey of Statistical Design and Linear Models*, North-Holland, Amsterdam, pp. 333–353.
- Kunert, J., 1983. Optimal designs and refinement of linear model with applications to repeated measurements designs. *Ann. Statist.* 11, 247–257.
- Kunert, J., 1984. Optimality of balanced uniform repeated measurements designs. *Ann. Statist.* 12, 1006–1017.
- Kunert, J., Martin, R.J., 2000a. On the determination of optimal designs for an interference model. *Ann. Statist.* 28, 1728–1742.
- Kunert, J., Martin, R.J., 2000b. Optimality of type I orthogonal arrays for cross-over models with correlated errors. *J. Statist. Plann. Inference* 87, 119–124.
- 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., 1997. Optimal repeated measurements designs: the linear optimality equations. *Ann. Statist.* 25, 2328–2344.
- Kushner, H., 1998. Optimal and efficient repeated-measurements designs for uncorrelated observations. *J. Amer. Statist. Assoc.* 93, 1176–1187.
- Majumdar, D., Martin, R.J., 2004. Efficient designs based on orthogonal arrays of type I and type II for experiments using units ordered over time or space. *Statist. Method.* 1, 19–35.
- Martin, R.J., Eccleston, J.A., 1998. Variance-balanced change-over designs for dependent observations. *Biometrika* 85, 883–892.
- Martin, R.J., Pooladsaz, S., 2005. Optimal bordered complete block designs under interference and dependent observations. *J. Statist. Plann. Inference*, Preprint.
- Matthews, J.N.S., 1987. Optimal crossover designs for the comparison of two treatments in the presence of carryover effects and autocorrelated errors. *Biometrika* 74, 311–320.
- Matthews, J.N.S., 1990. Optimal dual-balanced two-treatment crossover designs. *Sankhya B* 52, 332–337.
- Rao, C.R., 1961. Combinatorial arrangements analogous to orthogonal arrays. *Sankhya A* 23, 283–286.
- Stufken, J., 1991. Some families of optimal and efficient repeated measurements designs. *J. Statist. Plann. Inference* 27, 75–83.
- Stufken, J., 1996. Optimal crossover designs. In: Ghosh, S., Rao, C.R., (Eds.), *Handbook of Statistics*, vol. 13, North-Holland, Amsterdam, pp. 63–90.