Optimal and Efficient Crossover Designs for Test-Control Study When Subject Effects Are Random

A. S. Hedayat and Wei Zheng

We study crossover designs based on the criteria of A-optimality and MV-optimality under the model with random subject effects, for the purpose of comparing several test treatments with a standard control treatment. Optimal and efficient designs are proposed, and their efficiencies are also evaluated. A family of totally balanced test-control incomplete crossover designs based on a function of the ratio of the subject effect variance to the error variance are shown to be highly efficient and robust. The results have interesting connections with those in Hedayat and Yang (2005) and Hedayat, Stufken, and Yang (2006). The omitted proofs in the article are included in a supplemental material online.

KEY WORDS: A-optimality; Fisher information matrix; Mixed-effects model; MV-optimality.

1. INTRODUCTION

To compare the effects of various treatments we could, if feasible, conduct a crossover design where each subject in the study receives each treatment in succession. Besides the budget consideration, one primary advantage of adopting crossover designs is that the treatment effects are no longer confounded with the subject effects, and hence the systematic bias in estimating the treatment effects could be reduced substantially. However, the period effects and carryover effects could come into the picture at the same time, which makes the design problem more complex. Our interest here is to find optimal or efficient designs for the purpose of estimating direct treatment effects rather than carryover effects.

As in Hedayat, Stufken, and Yang (2006), we assume the subject effects to be random, which could be considered as in the middle status between the models with the fixed subject effects and no subject effects, in terms of the information matrix for the direct treatment effects. They studied the universal optimality, for which the totally balanced designs proposed therein exhibit the symmetry among all treatments. In this article, we focus on the A-optimality and MV-optimality, which aim to compare multiple test treatments to a control treatment. We investigate the high efficiency of totally balanced test-control incomplete (TBTCI) crossover designs, which exhibit the symmetry among test treatments. Interestingly, the class of TBTCI designs covers the totally balanced designs as a special case when the control treatment has the same replication as test treatments. The gain in generality for designs is due to our relaxation of the criterion from the universal optimality to one particular optimality either A-optimality or MV-optimality. Note that \( \theta \), the ratio between the variance of the subject effects and the variance of the error term, is still critical for the choice of designs.

The TBTCI designs were first proposed by Hedayat and Yang (2005) when they studied the A-optimality and MV-optimality for the model with fixed subject effects, that is, \( \theta = \infty \) in our setup. They proved that a certain type of TBTCI designs would be optimal in \( \Lambda \), a mildly restricted subclass of competing designs, and conjectured that \( \Lambda \) does not exclude too many good designs so that the optimal designs in \( \Lambda \) are still highly efficient in \( \Omega \), the unrestricted class of designs. As for the question of how efficient are these designs, no investigation was carried out. In this article, we generalize their results and find optimal designs in \( \Lambda \) for general models with random subject effects, that is, for general \( \theta \geq 0 \). We also characterize the optimal designs in \( \Omega \) when \( \theta = 0 \), which corresponds to the model with no subject effects. Moreover, we give an explicit way to evaluate the efficiency of any design in \( \Omega \) without finding the optimal designs for any \( \theta \).

Other selected contributions in this area include, but are not limited to, Hedayat and Afsarinejad (1975, 1978), Cheng and Wu (1980), Kunert (1984), Hedayat and Zhao (1990), Stufken (1991), Kushner (1997, 1998), Kunert and Martin (2000), Kunert and Stufken (2002), and Hedayat and Yang (2003, 2004) among many others. Stufken (1996) is an excellent survey article about crossover designs. However, the majority of the literature deal with universal optimality when subject effects are fixed.

The article is organized as follows. Section 2 introduces the model based on which the study of designs is carried out, and further provides some notations which will be used in the subsequent sections. Section 3 introduces the optimality criteria and the corresponding optimal and efficient designs. Section 4 describes the method of deriving the lower bound of the A-efficiency of any design, which will be used to evaluate the efficiencies of the designs proposed in Section 3. Section 5 summarizes the results of this article by comparing them to relevant results in existing literatures and proposes some future directions of research. Section 6 gives the proofs for Theorems 1 and 2 in Section 3 through several lemmas, while further supports for the derivation of these lemmas will be postponed to the supplemental material.

Sam A. Hedayat is Professor (E-mail: Hedayat@uic.edu) and Wei Zheng is Ph.D. Candidate, Department of Mathematics, Statistics & Computer Science, University of Illinois at Chicago, Chicago, IL 60607. Research was primarily sponsored by the National Science Foundation grants DMS-0603561 and DMS-0904125, and the NIH grant P50-AT00155 (jointly supported by the National Center for Complementary and Alternative Medicine, the Office of Dietary Supplements, the Office for Research on Women’s Health, and the National Institute of General Medicine). The contents are solely the responsibility of the authors and do not necessarily represent the official views of NIH. We are very grateful to Professor Min Yang, Department of Statistics, University of Missouri at Columbia, for his many valuable and insightful comments and suggestions while this article was under preparation. We also acknowledge with thanks the suggestions we received in improving the article by the Editor, the Associate Editor, and the two referees.
2. MODEL ASSUMPTIONS AND NOTATIONS

We denote by $\Omega_{t+1,n,p}$ the set of all of designs where $n$ subjects are used in $p \geq 3$ occasions, called periods, for the purpose of evaluating and studying one control treatment and $t \geq 2$ test treatments. Hereafter, we shall designate the $t$ test treatments by $1, 2, \ldots, t$ and the control treatment by 0. For a continuous response $Y$, a plausible and useful linear model with random subject effects can be written as

$$Y_{dku} = \mu + \pi_k + \sigma_u + \tau(d,k,u) + \rho(d(k-1),u) + \varepsilon_{ku},$$

(1)

where we assume the subject effect $\varepsilon_u \sim (0, \sigma_u^2)$, the error term $\varepsilon_{ku} \sim (0, \sigma^2)$, and $\{\varepsilon_u\} \perp \{\varepsilon_{ku}\}$. The identical condition in the notation iid is not essential here, and hence could be removed. Here, $Y_{dku}$ denotes the response from subject $u$ in period $k$ to which treatment $d(k,u) \in \{0, 1, 2, \ldots, t\}$ was assigned by design $d \in \Omega_{t+1,n,p}$, $k = 1, \ldots, p$, and $u = 1, \ldots, n$. Furthermore, $\mu$ is the general mean, $\pi_k$ is the $k$th period effect, $\tau(d,k,u)$ is the (direct) treatment effect of treatment $d(k,u)$, and $\rho(d(k-1),u)$ is the (first-order) carryover or residual effect of treatment $d(k-1,u)$ that subject $u$ received in the previous period (by convention $\rho(d(0),u) = 0$). Finally, $\sigma^2$ is the subject effects variance, and $\sigma_u^2$ is the error variance.

Writing the $np \times 1$ response vector as $Y_d = (Y_{d11}, Y_{d12}, \ldots, Y_{dp1}, Y_{d12}, \ldots, Y_{dpm})'$, we have

$$E(Y_d) = 1_{np} \mu + TP \pi + T_d \tau + F_d \rho,$$

$$\text{var}(Y_d) = \sigma^2 V,$$

where $V = I_p \otimes (I_p + \theta J_p)$ and $\theta = \sigma^2 / \sigma_u^2$. Here $\pi = (\pi_1, \ldots, \pi_p)'$, $\tau = (\tau_0, \ldots, \tau_t)'$, $\rho = (\rho_0, \ldots, \rho_t)'$, $P = I_n \otimes I_p$, and $T_d$ and $F_d$ denote the treatment/subject and carryover/subject incidence matrices. The notation $\otimes$ represents the transpose operator; $\otimes$ represents the Kronecker product; $I_p$ represents the column vector of length $s$ with all its entries as 1; $1_n$ represents the $s \times s$ identity matrix; and $J_p = 1_p J_p$. The information matrix $C_d$ for $\tau$ under model (1) can now be expressed as

$$C_d = T_d V^{-1/2} \text{pr}^{-1} (V^{-1/2} |1_{np}| |P| |F_d|) V^{-1/2} T_d,$$

(3)

where $\text{pr}^{-1}$ is a projection operator such that $\text{pr}^{-1} A = I - A(A'A)^{-1} A'$ for any matrix $A$.

Throughout the article, for each design $d$, we adopt the notation $n_{d\eta} = (\bar{n}_{d\eta}, \bar{l}_{d\eta}, \bar{m}_{d\eta}, \bar{r}_{d\eta})$, to denote the number of times that treatment $d$ is assigned to subject $\eta$, the number of times this happens in the first $p - 1$ periods associated with subject $\eta$, the number of times treatment $i$ is assigned to period $k$, the number of times that treatment $i$ is immediately preceded by treatment $j$, the total replication of treatment $i$ in the $n$ experimental subjects, and the total replication of treatment $i$ limited to the first $p - 1$ periods in the design. Also, we would like to define the subclass of designs $d_{t+1,n,p} = \{d \in \Omega_{t+1,n,p}| \bar{n}_{d\eta} = n_{d\eta}/p, k = 1, \ldots, p \}$ and $m_{d\eta} = 0, i = 0, 1, \ldots, t\}$.

Therefore, for any design in $\Lambda$, the control treatment appears equally often in all periods and no treatment is allowed to be preceded by itself. Additionally, we have the following convention: For any two square matrices (e.g., $A, B$) of the same size, the inequality $A \leq B$ represents the loewner’s ordering of the matrices; $\text{Tr}(A)$ represents the trace of the matrix $A$; $[x]$ represents the greatest integer that is not greater than $x$; $B_i = \text{pr}^{-1} (1_i - J_i/s, 0, \ldots, 0)$ is the set of all $t \times t$ permutation matrices and

$$\bar{S}_i = \left(\begin{array}{c|c} 1 & 0_{t \times 1} \\ \hline 0_{t \times 1} & S_i \end{array}\right),$$

for $i = 1, \ldots, t$.

3. OPTIMAL AND EFFICIENT DESIGNS

In comparing two or more test treatments with a control, the most frequently used optimality criteria are A-optimality and MV-optimality.

Definition 1. (i) In a class of competing designs, a design is said to be A-optimal if it minimizes $\sum_{i=1}^t \text{var}(\hat{\tau}_i - \tau_0)$, where $\hat{\tau} = (\hat{\tau}_0, \hat{\tau}_1, \ldots, \hat{\tau}_t)'$ is the generalized least square estimate of $\tau$.

(ii) A design is said to be MV-optimal if it minimizes $\max_{i=1,\ldots,t} \text{var}(\hat{\tau}_i - \tau_0)$.

Since the row and column sums of $C_d$ are both 0, we could express $C_d$ as

$$C_d = \left(1_n^T M_{d1} 1_t - 1_n^T M_{d1} J_t \right).$$

(4)

The $t \times t$ submatrix $M_{d1}$ is closely related to the A-optimality and MV-optimality. Let $H = (-1_n | I_t)$, then $\var(H \tilde{\tau}) = \sigma^2 H_d^T H = \sigma^2 M_{d1}^{-1}$, since one could choose diag(0, $M_{d1}^{-1}$) as a generalized inverse of $C_d$ in view of (4). For completeness and convenience, we cite a well-known result as Lemma 1 below, which is similarly done by Hedayat and Yang (2005).

Lemma 1. A design $d^*$ is A-optimal if it minimizes $\text{Tr}(M_{d1}^{-1})$. Further, if the $M_{d1}$ is completely symmetric, then $d^*$ is also MV-optimal.

Lemma 1 indicates that $M_{d1}$ contains all the information needed to evaluate the A-optimality and MV-optimality of a design. Since $M_{d1}$ in turn is a submatrix of $C_d$ by ignoring the first row and first column, designs with the same $C_d$ should be equivalent in the A-sense and MV-sense. On the other hand, the matrix $C_d$ is a function of the unknown variable $\theta$, hence the determination of optimal designs depends on the value of $\theta$. As pointed out by Hedayat, Stufken, and Yang (2006), two extreme cases are worth mentioning. The case where $\theta = 0$ corresponds to the situation of no subject effect. It is easily seen that

$$\lim_{\theta \to 0} C_d = T_d^T \text{pr}^{-1} \left(1_{np} |P| |F_d| \right) T_d,$$

which would indeed be precisely the information matrix for $\tau$ if we were to ignore the subject effect. Under this special case, Theorem 1 below gives the optimal designs in $\Omega$. The other extreme case corresponds to $\theta = \infty$, and we have

$$\lim_{\theta \to \infty} C_d = T_d^T \text{pr}^{-1} \left(1_{np} |P| |U| |F_d| \right) T_d,$$

where $U = I_n \otimes I_p$. This limit is precisely the information matrix that we would have obtained had we treated the subject effects as fixed. Under this special case, Hedayat and Yang (2005) gave the optimal designs in the subclass $\Lambda$, and they conjectured that optimal designs in $\Lambda$ is still highly efficient in $\Omega$. In this article, we derive optimal designs in $\Lambda$ for any value of $\theta$, which covers their result as a special case. We also give the explicit way of evaluating the efficiency of any design for any value of $\theta$.
Theorem 1. If \( \theta = 0 \), then for any \( n, t, p \), a design \( d \) is simultaneously A-optimal and MV-optimal in \( \Omega_{\theta, t, p} \) if it satisfies:

1. \( m_{dii} = r_d \tilde{d}_{dii}/pn, i = 0, 1, \ldots, t \)
2. \( m_{dij}, m_{d0i}, m_{d00} \) are constants across all \( 1 \leq i \neq j \leq t \)
3. \( \ell_{dik} = \ell_{dij}, i = 0, 1, \ldots, t, k_1, k_2 = 1, 2, \ldots, p \)
4. \( r_d = r_{dij}, i, j = 1, 2, \ldots, t \)
5. \( r_{d0} = \arg \min_{\ell \in \{1, 2, \ldots, np\}} (t/(np - h) + 1/h) \)

Condition 5 is equivalent to \( r_{d0} = np/(1 + \sqrt{t}) \) when the latter is an integer.

We would like to give an intuitive explanation for the conditions imposed in Theorem 1. Conditions 2–4 impose some structure of symmetry among test treatments and periods. Condition 5 directly determines the replication of each treatment in conjunction with Condition 4. Finally, Condition 1 requires the exact relationship between two type of variables, which is too strong in general. When \( n = 18, t = 4, p = 3 \), we need \( r_{d0} = np/(1 + \sqrt{t}) = 18 \). Under this situation, the design below constructed by Yang and Stuken (2008) is optimal in \( \Omega_{18,4,3} \) when \( \theta = 0 \):

\[
d_1: \begin{align*}
d_{11} & = 111122333444000000, \\
d_{12} & = 12323434400001012, \\
d_{13} & = 4000112212340330401.
\end{align*}
\]

For general \( \theta \), it is hard to compare all designs in \( \Lambda \). However, we succeeded in finding optimal designs in the subclass \( \Lambda \). In order to introduce the result, it is necessary to give the definition of a class of designs proposed by Hedayat and Yang (2005).

Definition 2. A design \( d \) is said to be a TBTCI design if:

1. \( m_{dii} = 0 \) for all \( 0 \leq i \leq t \)
2. \( m_{d0i}, m_{d00}, m_{d0i} \) are constants across all \( 1 \leq i \neq j \leq t \)
3. \( \ell_{dik} = \ell_{dij}, i = 0, 1, \ldots, t, k_1, k_2 = 1, 2, \ldots, p \)
4. \( r_d = r_{dij}, i, j = 1, 2, \ldots, t \)
5. \( n_{dij} = 0 \) or 1 for all \( i = 1, 2, \ldots, t, n = 1, 2, \ldots, n \)
6. \( |n_{d0u} - n_{d0v}| \leq 1 \) and \( |n_{d0u} - n_{d0v}| \leq 1 \) for all \( 1 \leq u, v \leq n \)
7. \( \sum_{i=1}^n 1 + \sum_{i=1}^n 1 + \sum_{i=1}^n 1 + \sum_{i=1}^n 1 + \sum_{i=1}^n 1 + \sum_{i=1}^n 1 + \sum_{i=1}^n 1 + \sum_{i=1}^n 1 \) are constants across all \( 1 \leq i \neq j \leq t \)

Theorem 2. (i) Suppose \( p \geq 3 \) and \( t \geq \max(p - 1, 3) \), then for any value of \( \theta \) a design \( d \) is simultaneously A-optimal and MV-optimal in \( \Lambda_{1+1,n,p} \) if \( d \) is a TBTCI design and

\[
r_{d0} = \arg \min_{h \in \{1, 2, \ldots, np - 1\}} f(h, \theta),
\]

where

\[
f(r_d, \theta) = t(t - 1)(\alpha_1 - \beta_1/\gamma_1)^{-1} + t(\alpha_2 - \beta_2/\gamma_2)^{-1},
\]

\[
\alpha_1 = t(1 - \lambda_p)(np - r_{d0}) - \eta(np - r_{d0})^2 - r_{d0} + \lambda_p \chi_2 + \eta^2 r_{d0},
\]

\[
\beta_1 = \lambda_p(n(p - 1) - \tilde{r}_{d0}) + (n(p - 1) - \tilde{r}_{d0})(np - r_{d0}) - \lambda_p \chi_2 - \eta r_{d0} \tilde{r}_{d0},
\]

\[
\gamma_1 = (t + 1 - 2/p - \lambda_p)(n(p - 1) - \tilde{r}_{d0}) - n(p - 1)^2/p - \eta(n(p - 1) - \tilde{r}_{d0})^2 + \eta^2 r_{d0}^2 + \lambda_p \chi_3,
\]

\[
\alpha_2 = r_{d0} - \lambda_p \chi_1 - \eta^2 r_{d0}.
\]

(ii) Suppose \( p = 3 \) and \( t = 2 \), then the conclusion in (i) is still valid if we change the class of competing designs from \( \Lambda_{1+1,n,p} \) to \( \{ d \in \Lambda_{1+1,n,p}: r_d^0/n \leq 0.6306 \} \).

Theorem 2 indicates that \( \theta \) influences the determination of optimal designs in \( \Lambda \) by deciding the value of \( r_{d0} \). Based on this \( r_{d0} \), we could try to find a TBTCI design. Also note that Equation (5) is analogous to Condition 5 of Theorem 1. Now let us compare the seven conditions in Definition 2 with Conditions 1–4 in Theorem 1. Conditions 2–4 in Definition 2 are identical to Conditions 2–4 in Theorem 1. Conditions 5–7 in Definition 2 are about symmetry among subjects, and their association with treatments. Theorem 1 does not impose these conditions since \( \theta = 0 \) corresponds to the model with no subject effects. Observe that Condition 1 in Definition 2 directly contradicts with Condition 1 in Theorem 1, which indicates that TBTCI designs can not be optimal when \( \theta = 0 \). However, TBTCI designs satisfying Equation (5) are robust and highly efficient for all values of \( \theta \), which will be illustrated in the next section. Moreover, TBTCI designs exist more commonly than the designs in Theorem 1. Hedayat and Zheng (2010) gave different methods of constructing TBTCI designs.

4. EVALUATION OF THE EFFICIENCIES

For any \( \theta \), let us denote by \( d(\theta) \) the corresponding optimal design, then we can define the \( A \)-efficiency of a design \( d \) at this \( \theta \) to be \( A(\theta, d) = \text{Tr}(\text{M}_1^{-1} d(\theta)/\text{Tr}(\text{M}_d^{-1})) \). We would be interested in deriving \( A(\theta, d) \) for any design \( d \), which should be resorted to deriving \( \text{Tr}(\text{M}_d^{-1}(\theta)) \). However, \( d(\theta) \) is generally not known except when \( \theta = 0 \). In this article, we instead derive a lower-bound curve \( \ell(\cdot) \leq \text{Tr}(\text{M}_d^{-1}(\cdot)) \), then \( \ell(\cdot) \) serves as the lower bound of \( A(\theta, d) \). In the following, we will simply call \( \ell(\cdot) \) the efficiency of design \( d \). Section 3 gives optimal designs in \( \Omega \) when \( \theta = 0 \), and optimal designs in \( \Lambda \) for general \( \theta \). We would like to evaluate the efficiencies of these designs. For each design, we have

\[
C_d = T_d^0 V^{-1/2} \text{pr}^\perp (V^{-1/2} [\text{I}_{np}] [\text{P} [\text{F}_d] V^{-1/2}]) T_d^{-1/2}
\]

\[
\leq T_d V^{-1/2} \text{pr}^\perp (V^{-1/2} [\text{I}_{np}] [\text{F}_d B_{r+1}]) V^{-1/2} T_d^{-1/2}
\]

\[
\leq T_d V^{-1/2} \text{pr}^\perp (V^{-1/2} [\text{I}_{np}] [\text{F}_d B_{r+1}]) V^{-1/2} T_d^{-1/2}
\]
since \( Q_d(x, y, z) \) is a linear function of \( M_d(i) \), \( i = 1, 2, 3 \). By direct calculation, we have

\[
Q_d(x, y, z) = \frac{\eta(z - n(p - 1)/(t + 1))}{(t - 1)^2} + \sqrt{(t - 1)^2 + \frac{\eta}{q_d}} .
\]
where $w_{dj}$ is the proportion of the number of sequences from class $j$ in design $d$, and hence $\sum_{j=1}^{J} w_{dj} = 1$. Now we have

$$\max_d q_d = \max_d \min_{x,y,z} Q_d(x,y,z) \leq \min_{x,y,z} \max_d Q_d(x,y,z)$$

$$= \min_{x,y,z} \max_{h_1,h_2} \max_{d,r_{d0}=h_1} \sum_{j=1}^{J} w_{dj} H_j(x,y,z)$$

$$- \frac{\eta}{t} \left( \left( r_{d0} - \frac{n(p-1)}{t+1} \right) z + r_{d0} \right)^2$$

$$= \min_{x,y,z} \max_{h_1,h_2} \max_{d,r_{d0}=h_1} \sum_{j=1}^{J} w_{dj} H_j(x,y,z)$$

$$- \frac{\eta}{t} \left( \left( h_2 - \frac{n(p-1)}{t+1} \right) z + h_1 \right)^2,$$  \hspace{1cm} (16)

where

$$\eta_{h_1,h_2} := \max_{0,h_1-h_2} \max_{0,h_1-h_2} \eta_{h_1,h_2}.$$

By (9) and (16), we have:

**Theorem 3.** For any $n, t, p$, and $\theta$, the reciprocal of the right-hand side of (16) is a lower bound of $\text{Tr}(M^{-1}_d)$ for any $d \in \Omega_{n,t+1,p}$.

To implement Theorem 3, there are three levels of maximization/minimization: (I) Maximization over the designs in $\{d \in \Omega | r_{d0} = h_1, r_{d0} = h_2\}$ for given $(h_1, h_2)$ and $(x, y, z)$. (II) Maximization over $(h_1, h_2)$ for given $(x, y, z)$. (III) Minimization over $(x, y, z)$.

For level (I) maximization, the problem could be reduced to the linear programming problem:

Maximize $\sum_{j=1}^{J} w_j H_j$

Subject to $n \sum w_j a_j = h_1; n \sum w_j b_j = h_2; \sum w_j = 1; w_j \geq 0, j = 1, 2, \ldots, J$.

Here $a_j$ (resp., $b_j$) is the number of 0’s in the sequence (resp., the first $p-1$ symbols of the sequence) from class $j$. For example, when $p = 3, t \geq 3$, we have $J = 13$ and

| 000010011111111 | 00101100211222 | 0100212100213 |
| 02222111111000 | 02221111110000 |

For level (II) maximization, we could reduce the amount of search for $(h_1, h_2)$ by the following procedure:

1. Find a rough lower bound $e_1(r_{d0}, r_{d0}) \leq \text{Tr}(M^{-1}_d), d \in \Omega$.
2. Calculate $\text{Tr}(M^{-1}_{d0})$ for a particular design $d_0$ (e.g., TBT-CI design).
3. If $e_1(h_1, h_2) \leq \text{Tr}(M^{-1}_{d0})$, continue with the linear programming in level (I) maximization; otherwise, jump to next $(h_1, h_2)$.

**Lemma 2.** A lower bound of $\text{Tr}(M^{-1}_d)$ is $e_1(r_{d0}, r_{d0}) = t(t-1)^2/\theta_1 + t/\theta_1$, where

$$x_1 = t(1-\lambda_2)(np-r_{d0}) - r_{d0} - \eta np (np-2r_{d0}) + \lambda_2 \chi_1,$$

$$y_1 = r_{d0} - \eta r_{d0} - \lambda_2 \chi_1 - \chi_4/n,$$

$$\chi_4 = r_{d0} + (2r_{d0} - p + 1)(r_{d0}/(p-1)) + (p-1)(r_{d0}/(p-1))^2 + (r_{d0} - r_{d0})^2/r_{d0}/p$$

and $\chi_4$ is defined as in Theorem 2 in Section 3.

**Proof.** By ignoring $F_d$ in $C_d$, we have

$$\begin{align*}
C_d & \leq T_d^0 V^{-1/2} \text{pr}^{-1}(1_{np} V^{-1/2} P V^{-1/2} T_d^0)
\quad - T_d^0 V^{-1/2} \text{pr}^{-1}(1_{np} V^{-1/2} P V^{-1/2} P^0 V^{-1/2} \text{pr}^{-1}(1_{np}) V^{-1/2}
\quad - T_d^0 V^{-1/2} \text{pr}^{-1}(1_{np} V^{-1/2} P V^{-1/2} P^0 V^{-1/2} \text{pr}^{-1}(1_{np}) V^{-1/2}
\quad - n^{-1} J_n \otimes B_p.
\end{align*}$$

By the $S_i$ argument as in Lemma 4 in Section 6, we can obtain

$$\text{Tr}(M^{-1}_d) \geq t(t-1)^2/\theta_1 + t/\theta_1,$$  \hspace{1cm} (17)

where

$$\begin{align*}
\hat{x}_1 & = t(1-\lambda_2)(np-r_{d0}) - r_{d0} \\
& - \eta np (np-2r_{d0}) + \lambda_2 \sum_{u=1}^{n} n_{u0a}^2,
\end{align*}$$

$$\hat{y}_1 = r_{d0} - \eta r_{d0} - \lambda_2 \sum_{u=1}^{n} n_{u0a}^2 - n^{-1} \sum_{k=1}^{p} (l_{d0k} - r_{d0k}/p)^2.$$

In deriving (17), we used the fact $\sum_{k=1}^{p} (l_{d0k} - r_{d0k}/p)^2 \leq \sum_{k=1}^{p} (l_{d0k} - r_{d0k}/p)^2$. The lemma is now concluded by noting $\sum_{a=1}^{n} n_{a0a}^2 \geq \chi_1$ and $\sum_{k=1}^{p} (l_{d0k} - r_{d0k}/p)^2 \geq \chi_4$.

For level (III), we could use Newton-type algorithm.

When $n = 36, p = 3, t = 4$, a TBT-SC design $d_2$ with $r_{d0} = 36$ would be constructed through the method in Hedayat and Zheng (2010), also $d_3 = (1, 1) \otimes d_1$ is the optimal design when $\theta = 0$. The corresponding lower bound $\ell(\cdot)$ as specified by Theorem 3 could be calculated by the methods mentioned above; Note that the screening procedure for level (II) maximization saved the calculating time by 77.89%. The total time needed to calculate $\ell(\cdot)$ at one value of $\theta$ is 386.50 seconds (CPU: Intel 2 Duo 1.80 GHz; Software: R), while the direct search for optimal designs would require $1.11 \times 10^{26}$ years.
Figure 1. LB represents $\ell(\cdot), \lambda = \theta/(1 + \theta)$ transforms the range of $\theta$ from $[0, \infty]$ to $[0, 1]$.

Figure 2. The TBTCI design $d_2$ is quite robust with the efficiency $\geq 0.9$.

Figure 3. Efficiency in $\Lambda$: $\min_h f(h, \theta)/\text{Tr}(M_d^{-1})$.

By Lemma 4 in Section 6, the A-criterion of a TBTCI design $d$ has the form $\text{Tr}(M_d^{-1}) = n^{-1}g_p,r(\theta_0)/n, \theta)$ for some function $g$ which depends on the values of $p$ and $t$. Hence the A-criterion of a TBTCI design is proportional to $n^{-1}$ as long as the ratio $r_0/n$ and other parameters are fixed. Actually, the lower bound $\ell(\cdot)$ has a similar property. To see this, let $\ell^*(\cdot)$ be the reciprocal of the right-hand side of (16) when we allow $h_1$ and $h_2$ to be real numbers instead of integers under the maximization. Then $\ell(\cdot)$ should be very close to $\ell^*(\cdot)$ especially when $n$ is large, and we have $\ell^*(\theta) \leq \ell(\theta)$ for any $\theta$ in general. Also, we have $\ell^*(\theta) = n^{-1}g_p,t(\theta)$, hence $\ell^*(\theta)/\text{Tr}(M_d^{-1})$ only depends on the ratio $r_0/n$. Figure 3 gives the more conservative efficiencies of TBTCI designs with the ratio being 1, 0.95, 0.9 (if they exist) in both $\Omega$ and $\Lambda$ when $p = 3$ and $t = 4$. This figure gives the guidance regarding what kind of TBTCI design should be chosen. Similar work could be carried out for other configurations of $p$ and $t$. We postpone the discussion of possible improvements on current results to Section 5.

To evaluate the MV-efficiencies of TBTCI designs and the type of designs proposed in Theorem 1 in Section 3, it is useful to note that the matrix $M_d$ is completely symmetric for them. For any of these designs, we have $MV_d = A_d/t$, where $A_d$ and $MV_d$ are the A-criteria and MV-criteria of the design $d$. Let $d^{\mu}$ and $d^{mv}$ be the A-optimal and MV-optimal designs, then we have

$$\frac{MV_{d^{mv}}}{MV_d} \geq \frac{A_{d^{mv}}}{A_d} \geq \frac{A_d}{A_d}.$$ 

Hence we have:

**Corollary 1.** For any design proposed by Theorems 1 and 2 in Section 3, its MV-efficiency is at least as large as its A-efficiency regardless of the competing class of designs.

5. DISCUSSION AND CLOSING REMARKS

In a mixed linear model with random subject effects, the ratio $(\theta)$ of the variance of the subject effect to the variance of the error plays a crucial role in deciding which design should be used. Note that in estimating the treatment effects, $\theta = 0$ corresponds to the model without the subject effects while $\theta = \infty$
corresponds to the model with fixed subject effects. For the latter model, Hedayat and Yang (2005) found optimal designs in the subclass \( \Lambda \) and conjectured that \( \Lambda \) will not exclude too many good designs so that optimal designs in this subclass is still highly efficient or even optimal in \( \Omega \). As for how efficient these designs will be, they didn’t carry out the investigation.

In this article, we dealt with general \( \theta \geq 0 \) and found optimal designs in \( \Lambda \), which naturally covered the result in Hedayat and Yang (2005) as a special case when \( \theta = \infty \). Moreover, we found optimal designs in \( \Omega \) when \( \theta = 0 \). Further, we gave the algorithm to calculate the lower bound of \( \text{Tr}(M_d^{-1}) \), which turns out to be very close to the value of optimal designs. Hence it is proper to use this lower bound to evaluate the efficiencies of the designs proposed. Specifically, we evaluated the behavior of two types of designs. One is the optimal designs when \( \theta = 0 \), whose efficiency decreases dramatically as \( \theta \) increases. The other is the TBTCI designs with properly chosen value of \( r_{d0} \), which are shown to be robust across different values of \( \theta \). Since we usually don’t know the value of \( \theta \) in application, this robustness is critical. Also, it is gratifying that TBTCI designs commonly exist (Hedayat and Zheng 2010), while the optimal designs for \( \theta = 0 \) is not so common to exist due to the first condition of Theorem 1 in Section 3.

In studying universal optimality, Hedayat, Stufken, and Yang (2006) proved the high efficiency of the totally balanced designs, which is essentially a special type of TBTCI designs with the property of \( r_{d0} = r_{d1} \). Particularly, they proved the universal optimality of the totally balanced designs in \( \Omega^2 = \{d \in \Omega | l_{d0} = n/(t+1), m_{dii} = 0, i = 0, 1, \ldots, t\} \), which is certainly true for \( \Lambda \)-optimality and MV-optimality. However, as pointed out by one of the referees, optimality frameworks based on functions of \( M_d \) result in placing overriding weight on the control and hence higher control replication, which is shown in Section 4. Actually, the optimality of the totally balanced designs in \( \Omega^2 \) established by Hedayat, Stufken, and Yang (2006) could be explained by the fact that the class \( \Omega^2 \) rules out any designs with unequal replications of treatments. Also, the high efficiency of the totally balanced designs in \( \Omega \) established therein based on the trace approach could not be carried over to \( \Lambda \)-criteria and MV-criteria.

At the same time, we observe from Figure 3 that the gap between optimal designs in \( \Lambda \) and \( \Omega \) is more obvious for small values of \( \theta \). We conjecture that this is due to the restriction of \( m_{dii} = 0, i = 0, 1, \ldots, t \), in view of Condition 1 in Theorem 1 in Section 3. When \( \theta = \infty \), Hedayat and Yang (2006) and Yang and Park (2007) extended the class of competing designs from \( \Lambda \) to \( \Omega^1 = \{d \in \Omega | l_{d0} = r_{d0}/p, k = 1, 2, \ldots, p\} \) when either of the following conditions satisfies (i) \( p = 3 \) and \( 3 \leq t \leq 20 \); or (ii) \( p \geq 4, (p-3)(p-2)+2 \leq t \leq (p-2)(p-1)+1 \) and \( n \geq p(p-1)/2 \). They proposed designs allowing part of the subjects in the study to have identical treatments in the last two periods, which are \( \Lambda \)-better than TBTCI designs. However, we note that Conditions (i) and (ii) here imposed on \( n, t, p \) are quite restrictive, therefore it would be essential if we could find optimal designs in \( \Omega^1 \) or even \( \Omega \) for wider ranges of the parameters \( n, t, p \). Another possible direction for future research is to study the case when \( p > t+1 \), for which we have not seen any work carried out yet. We believe some designs with similar structures as TBTCI designs would be highly efficient in this case.

6. MAIN PROOFS

Lemma 1 in Section 3 gives an explicit relationship between the optimality criteria and the information matrix \( C_d \). In order to find optimal designs, we first need to find a function \( \ell_0(\theta) \) such that

\[
\text{Tr}(M_d^{-1}) \geq \ell_0(\theta) \tag{18}
\]

for any \((d, \theta)\). If at the same time we have \( \text{Tr}(M_{d*}^{-1}) = \ell_0(\theta^*) \), then the design \( d^* \) would be optimal when \( \theta = \theta^* \). To establish (18), we can start with maximizing \( C_d \) in the Loewner’s sense.

Lemma 3. For any design \( d \), we have

\[
C_d \leq T_d' V^{-1/2} \text{pr}^\perp (1_{\text{np}} V^{-1/2} F_d) V^{-1/2} T_d. \tag{19}
\]

The equality in (19) holds for any design \( d \) in which \( l_{dik} = r_{d0}/p, i = 0, \ldots, t \).

Proof. It is sufficient to prove \( T_d' V^{-1/2} \text{pr}^\perp (1_{\text{np}} V^{-1/2} F_d) \times V^{-1/2} P = 0 \) when \( l_{dik} = r_{d0}/p, i = 0, \ldots, t \). Let

\[
A = 1_n \otimes \left( \begin{array}{cc}
0 & 0 \\
0 & B_{p-1}
\end{array} \right)
\]

Then, we have \( \text{pr}^\perp (1_{\text{np}} V^{-1/2} F_d) V^{-1/2} (P - A) = 0 \) and \( 1_{\text{p}}' \times V^{-1/2} A = 0 \). Plus, \( l_{dik} = r_{d0}/p, i = 0, \ldots, t \), implies \( T_d' V^{-1/2} A = 0 \) and \( T_d' V^{-1} A = 0 \). The lemma is established.

Lemma 4. (i) For any design \( d \in \Omega_{i+1,n,p} \), we have

\[
\text{Tr}(M_d^{-1}) \geq \frac{(t-1)^2}{x_0} + \frac{t}{y_0}, \tag{20}
\]

where

\[
x_0 = a_1 - \frac{\beta^2}{\gamma_1},
\]

and

\[
y_0 = r_{d0} - \frac{\theta}{1 + \theta p} \sum_{u=1}^{n} n_{d0u}^2 - \frac{r_{d0}^2}{(1 + \theta p)n} - \left\{ \frac{n(p-1) - \tilde{r}_{d0}}{\tilde{r}_{d0}} \right\} \left( m_{d00} - \frac{\theta}{1 + \theta p} \sum_{u=1}^{n} n_{d0u} \tilde{n}_{d0u} \right. \\
\left. - \frac{1}{(1 + \theta p)n} r_{d0} \tilde{r}_{d0} \right)^2 \\
+ \tilde{r}_{d0} \left( \frac{r_{d0}}{p} - l_{d01} - m_{d00} + \frac{\theta}{1 + \theta p} \sum_{u=1}^{n} n_{d0u} \tilde{n}_{d0u} \right. \\
+ \frac{1}{(1 + \theta p)n} r_{d0} \tilde{r}_{d0} \right) \left\} \times \left\{ n(p-1) \left( \tilde{r}_{d0} - \frac{\theta}{1 + \theta p} \sum_{u=1}^{n} \tilde{n}_{d0u}^2 \right. \\
\left. - \frac{\tilde{r}_{d0}^2}{(1 + \theta p)n} \right)^{-1} - \frac{\tilde{r}_{d0}^2}{p} \right\}^{-1}
\]
with
\[
\alpha_1 = t\left(\frac{1 - \frac{\theta}{1+\theta}}{1+\theta p}\right)(np - r_{d0}) - \frac{t}{(1+\theta p)pn} \sum_{i=1}^{t} r_{di}^2 - r_{d0},
\]
\[
+ \frac{\theta}{1+\theta p} \sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} + \frac{r_{d0}^2}{(1+\theta p)pn}.
\]
\[
\beta_1 = \frac{\theta t}{1+\theta p} \sum_{i=1}^{t} \sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} + \frac{t}{(1+\theta p)pn} \sum_{i=1}^{t} r_{di}\tilde{r}_{di}
\]
\[
- \frac{\theta}{1+\theta p} \sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} - \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn},
\]
\[
\gamma_1 = \left(t + 1 - \frac{2}{p}\right)\frac{\theta}{1+\theta p} \sum_{i=1}^{t} \sum_{u=1}^{n} \frac{\gamma_{u}^2}{p_{n}^2} + \frac{t}{(1+\theta p)pn} \sum_{i=1}^{t} \frac{n_{duu}}{p_{n}^2}.
\]

Further, the equality in (20) holds when the following three conditions hold:

1. \(n_{duu}\) is either 0 or 1, \(1 \leq i \leq t, 0 \leq u \leq n\)
2. \(l_{dik} = r_{di}/p, i=0, \ldots, t\)
3. \(T_{d}^{*}V^{1/2}pr^{1/2}(1_{np})V^{-1/2}T_{d}, T_{d}^{*}V^{1/2}pr^{1/2}(1_{np})V^{-1/2}F_{d}, \) and \(F_{d}^{*}V^{1/2}pr^{1/2}(1_{np})V^{-1/2}F_{d}\) are invariant under any permutation of test treatments.

(ii) For any design \(d \in \Omega_{i+1,n,p}\) in which \(r_{d0} - l_{d01} = m_{dii} = 0, 0 \leq i \leq t,\) we have (20) with \(\alpha_1\) and \(\gamma_1\) therein keep unchanged, however \(\beta_1\) and \(\gamma_0\) have the following simpler forms
\[
\beta_1 = \frac{\theta t}{1+\theta p} \sum_{i=1}^{t} \sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} + \frac{t}{(1+\theta p)pn} \sum_{i=1}^{t} r_{di}\tilde{r}_{di}
\]
\[
- \frac{\theta}{1+\theta p} \sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} - \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn},
\]
\[
\gamma_0 = \frac{\alpha_2}{\gamma_2},
\]
\[
\alpha_2 = r_{d0} - \frac{\theta}{1+\theta p} \sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} - \frac{r_{d0}^2}{(1+\theta p)pn},
\]
\[
\beta_2 = \frac{\theta}{1+\theta p} \sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} + \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn},
\]
\[
\gamma_2 = \tilde{r}_{d0} - \frac{r_{d0}^2}{(p-1)pn} - \frac{\theta}{1+\theta p} \sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} - \frac{r_{d0}^2}{(1+\theta p)pn}.
\]

The equality in (20) still holds under the same conditions.

The proof of Lemma 4 is similar to lemma 4 of Hedayat and Yang (2005), and we postpone it to the Supplemental Materials. The merit of Lemma 4 is that we settle down with a very special subclass of designs such that the \(A\)-criterion of designs therein are easy to be calculated directly and at the same time the optimal design is guaranteed to be included in this subclass. For the special case of \(\theta = 0\), we are ready to prove Theorem 1 in Section 3.

**Proof of Theorem 1.** By Lemma 4(i), we have \(\text{Tr}(M_{d}^{-1}) \geq t(t-1)/(np - r_{d0}) + tnp/(r_{d0}(np - r_{d0})) = t/r_{d0} + r_{d0}^2/(np - r_{d0})\) for any design with the equality holds when the design satisfies Conditions 1–4. Hence the \(A\)-optimality is established. Furthermore, these conditions are sufficient for \(M_{d}\) to be completely symmetric.

For general value of \(\theta\), we first need the following four preliminary lemmas:

**Lemma 5.** For any \(d \in \Lambda_{i+1,n,p}\), we have \(r_{d0} \leq p[\frac{\theta}{4}, \tilde{r}_{d0} \leq (p-1)[\frac{\theta}{4}], l_{d0k} \leq \lfloor \frac{\theta}{4}\rfloor\).

**Proof.** If \(l_{d0k} > \lfloor \frac{\theta}{4}\rfloor\), it will conflict with the condition that \(m_{d00} = 0\). The other two inequalities follow immediately by noting that \(l_{d0k} = l_{d0k}'\) for any \(1 \leq k \neq k' \leq p\).

**Lemma 6.** For any design \(d \in \Lambda_{i+1,n,p}\), we have
\[
\sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} \leq \frac{1}{4}np(p-1).
\]

Consequently, we have \(\sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} \leq t(n(p-1) - \tilde{r}_{d0})\) as long as \(2t \geq p\).

**Proof.** To maximize \(\sum_{u=1}^{n} n_{duu}\tilde{n}_{duu}\), it is necessary for \(r_{d0}\) to attain its maximum, which is \(p\lfloor \frac{\theta}{4}\rfloor\] by Lemma 5. (i) When \(n\) is an even number, then \(\tilde{r}_{d0} = np/2\) and the distribution of control treatment in the design fall into one type: \(n/2\) of the subjects take the control treatment at all even periods; the remaining half of the subjects take the control treatment at all odd periods. Then the inequality trivially holds. (ii) When \(n\) is an odd number, the maximum of \(\sum_{u=1}^{n} n_{duu}\tilde{n}_{duu}\) will be attained when one of the subjects does not take the control treatment while the remaining \(n - 1\) subjects take the control treatment in the same way as in (i). To see this, suppose Subject 1 has the smallest value of \(\tilde{n}_{duu}\). If \(\tilde{n}_{duu} > 0\), without loss of generality let us suppose Subject 1 takes the control treatment at the second period. Then always exist a subject (say, 2) who takes test treatments in the second period as well as in the neighboring periods, that is, periods 1 and 3, since \(n\) is odd. Then we can exchange the treatments between these two subjects at the second period so that \(m_{d00}\) is still 0. By this exchange, the decrement of \(\tilde{n}_{duu}\) at \(d01\) is at most \(2\tilde{n}_{duu}\), while the corresponding increment of \(n_{duu}\) is at least \(2\tilde{n}_{duu} + 1\). Since \(\tilde{n}_{duu} > \tilde{n}_{duu}\), \(\sum_{u=1}^{n} n_{duu}\tilde{n}_{duu}\) is increased. Hence we have, by the argument in (i), that
\[
\sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} \leq \frac{1}{4}(n-1)p(p-1) \leq \frac{1}{4}np(p-1).
\]

**Lemma 7.** For any design \(d \in \Omega_{i+1,n,p}\), we have
\[
\sum_{i=1}^{t} \sum_{u=1}^{n} n_{duu}\tilde{n}_{duu} \geq n(p-1) - \tilde{r}_{d0}.
\]
Proof.

\[
\sum_{i=1}^{t} n_{diu} \tilde{n}_{diu} \geq \sum_{i=1}^{t} \sum_{u=1}^{n} \tilde{n}_{diu} = \sum_{i=1}^{t} \tilde{r}_{di} = n(p-1) - \tilde{r}_{d0}.
\]

Lemma 8. For any design \(d \in \Lambda_{i+1,n,p}\), we have \(\beta_1 + \gamma_1 \geq 0\) for any value of \(\theta\).

Proof. If we consider \(\beta_1 + \gamma_1\) as a function of \(\theta\), then by the simple derivative equations

\[
\frac{d}{d\theta} \left( \frac{\theta}{1 + \theta} \right) = \frac{1}{(1 + \theta)^2},
\]

we have

\[
\frac{d(\beta_1 + \gamma_1)}{d\theta} = \frac{K_d}{(1 + \theta)^2},
\]

where \(K_d\) is a constant which depends on the design \(d\) only. That means \(\beta_1 + \gamma_1\) is either nondecreasing or nondecreasing with respect to \(\theta\) when the design \(d\) is fixed. Hence, it is enough to prove the inequality for \(\theta = 0\) and \(\theta = \infty\). For \(\theta = 0\), we have, in view of Lemma 5, that

\[
\beta_1 + \gamma_1 = \left( t + 1 - \frac{2}{p} \right)(n(p-1) - \tilde{r}_{d0})
\]

\[
- \frac{n}{p}(p-1)^2 + \frac{t}{pn} \sum_{i=1}^{t} \tilde{r}_{di} \tilde{l}_{dip} - \frac{\tilde{r}_{d0} \tilde{l}_{d0p}}{pn}
\]

\[
\geq \left( t - \frac{1}{p} \right) n(p-1)
\]

\[
- \left( t + 1 - \frac{2}{p} + \frac{\tilde{l}_{d0p}}{pn} \right) \tilde{r}_{d0}
\]

\[
\geq \frac{tp}{p}(p-1) - \left( t + 1 - \frac{1}{p} \right) n(p-1) - \frac{n(p-1)}{2}(p(t-1) - \frac{1}{2}) \geq 0.
\]

For \(\theta = \infty\), we have, in view of Lemmas 5 and 7, that

\[
\beta_1 + \gamma_1 \geq \left( t + 1 - \frac{2}{p} \right)(n(p-1) - \tilde{r}_{d0}) - \frac{n}{p}(p-1)^2 - \frac{1}{p} \tilde{r}_{d0}
\]

\[
\geq \left( t - \frac{1}{p} \right) n(p-1) - \left( t + 1 - \frac{1}{p} \right) \tilde{r}_{d0}
\]

\[
\geq \frac{n(p-1)}{2p}(p(t-1) - 1) \geq 0.
\]

In the following, we shall search for designs which minimizes the right-hand side of (20), where the components \(\sum_{i=1}^{t} n_{diu} \tilde{n}_{diu}, \sum_{i=1}^{t} \tilde{n}_{diu}^2, \sum_{i=1}^{t} n_{diu} \tilde{n}_{diu}^2, \sum_{i=1}^{t} n_{diu} \tilde{n}_{diu}^2, \sum_{i=1}^{t} \tilde{n}_{diu} \tilde{n}_{diu}^2, \sum_{i=1}^{t} \tilde{n}_{dip}^2, \) and \(\sum_{i=1}^{t} \tilde{r}_{di} \tilde{r}_{dip}\) are related to each other. The latter four components are investigated by Lemmas 9 and 10 while the remaining three are investigated by Lemma 11. Since \(\sum_{i=1}^{t} n_{diu} \tilde{n}_{diu}^2, \sum_{i=1}^{t} \tilde{n}_{diu}^2, \sum_{i=1}^{t} \tilde{n}_{dip}^2, \) and \(\sum_{i=1}^{t} \tilde{r}_{di} \tilde{r}_{dip}\) cannot be minimized simultaneously, we need the following decomposition:

\[
\sum_{i=1}^{t} \tilde{r}_{di} \tilde{r}_{dip} = \sum_{i=1}^{t} \tilde{r}_{di}^2 + \sum_{i=1}^{t} \tilde{r}_{di} \tilde{l}_{dip},
\]

\[
\sum_{i=1}^{t} \tilde{r}_{di}^2 = \sum_{i=1}^{t} \tilde{r}_{di}^2 + 2 \sum_{i=1}^{t} \tilde{r}_{di} \tilde{l}_{dip} + \sum_{i=1}^{t} \tilde{l}_{dip}^2.
\]

By (24), we transform \(\sum_{i=1}^{t} \tilde{r}_{di} \tilde{l}_{dip}\) and \(\sum_{i=1}^{t} \tilde{l}_{dip}^2\) into \(\sum_{i=1}^{t} \tilde{r}_{di} \tilde{l}_{dip}\) and \(\sum_{i=1}^{t} \tilde{l}_{dip}^2\), respectively, we have

\[
\sum_{i=1}^{t} \tilde{r}_{di} \tilde{l}_{dip}
\]

\[
\geq \frac{1}{t}(n(p-1) - \tilde{r}_{d0})(n - \tilde{l}_{d0p})
\]

\[
- \sqrt{\left( a - \frac{(n(p-1) - \tilde{r}_{d0})^2}{t} \right) \left( b - \frac{(n - \tilde{l}_{d0p})^2}{t} \right)}.
\]

with the equality holds when \(\tilde{r}_{di} = \tilde{r}_{dij}\) and \(\tilde{l}_{dip} = \tilde{l}_{dip}\) for any \(1 \leq i \neq j \leq t\).

Lemma 9. For any design \(d \in \Omega_{i+1,n,p}\), when \(a\) and \(b\) are the values of \(\sum_{i=1}^{t} \tilde{r}_{di}^2\) and \(\sum_{i=1}^{t} \tilde{l}_{dip}^2\), respectively, we have

\[
\sum_{i=1}^{t} \tilde{r}_{di} \tilde{l}_{dip}
\]

\[
\geq \frac{1}{t}(n(p-1) - \tilde{r}_{d0})(n - \tilde{l}_{d0p})
\]

\[
- \sqrt{\left( a - \frac{(n(p-1) - \tilde{r}_{d0})^2}{t} \right) \left( b - \frac{(n - \tilde{l}_{d0p})^2}{t} \right)}.
\]

with the equality holds if and only if \(x_i = -c_0 y_i, i = 1, \ldots, t\), with

\[
c_0 = \sqrt{\frac{ia - (n(p-1) - \tilde{r}_{d0})^2}{tb - (n - \tilde{l}_{d0p})^2}}.
\]

Hence the lemma is established.

Lemma 10. If \(p \leq t + 1\) and \(\theta \geq 0\), we have, for any design \(d \in \Lambda_{i+1,n,p}\), that

\[
\frac{t}{x_0} + \frac{t}{y_0} \geq \frac{t(t-1)^2}{x_0} + \frac{t}{y_0},
\]

where

\[
x_0 = \bar{a}_1 - \frac{\tilde{r}_{d0}^2}{\tilde{r}_{d0}}
\]

\[
y_1 = \bar{y}_1.
\]
Thus, we have
\[ \tilde{\alpha}_1 = t \left( 1 - \frac{\theta}{1 + \theta p} \right) (np - r_{d0}) - \frac{(np - r_{d0})^2}{(1 + \theta p)pn} - r_{d0} + \frac{\theta}{1 + \theta p} \sum_{u=1}^{n} n_{d0u}^2 \right), \]
\[ \tilde{\beta}_1 = \frac{\theta t}{1 + \theta p} (n(p - 1) - \tilde{r}_{d0}) + \frac{(n(p - 1) - \tilde{r}_{d0})(np - r_{d0})}{(1 + \theta p)pn} - \frac{\theta}{1 + \theta p} \sum_{u=1}^{n} n_{d0u}^2 \tilde{n}_{d0u} - \frac{\theta}{1 + \theta p} \sum_{u=1}^{n} n_{d0u}^2 \tilde{n}_{d0u}, \]
\[ \tilde{\gamma}_1 = \left( t + 1 - \frac{2}{p} - \frac{\theta t}{1 + \theta p} \right) (n(p - 1) - \tilde{r}_{d0}) - \frac{n}{p} (p - 1)^2 - \frac{(n(p - 1) - \tilde{r}_{d0})^2}{(1 + \theta p)pn} + \frac{\tilde{r}_{d0}^2}{(1 + \theta p)pn} + \frac{\theta}{1 + \theta p} \sum_{u=1}^{n} n_{d0u}^2. \]

The equality in (26) holds when all \( n_{d0u} \)'s are binary and \( \tilde{r}_{d0} = \tilde{r}_{d}, l_{dip} = l_{dip} \) for any \( 1 \leq i \neq j \leq t \).

**Proof.** Let \( a, b, \) and \( c \) be the values of \( \sum_{i=1}^{t} \tilde{r}_{d0i}^2, \sum_{i=1}^{t} \tilde{r}_{d0i}l_{dip}, \) and \( \sum_{i=1}^{t} \tilde{r}_{d0i}l_{dip} \), respectively. In view of the decomposition (24), and Lemma 8, we have
\[ \frac{\partial x_0}{\partial c} = -2 (\tilde{\beta}_1 + \gamma_1) \leq 0. \]

Thus, we have \( x_0 \leq x_0' \), where \( x_0' \) is obtained from \( x_0 \) by replacing \( \sum_{i=1}^{t} \tilde{r}_{d0i}l_{dip} \) by its lower bound in the right-hand side of (25). Similarly, denote by \( \tilde{\beta}_1' \) the new term obtained from \( \beta_1 \) by the same replacement. Obviously, we have \( \tilde{\beta}_1 \geq \tilde{\beta}_1' \). In the following, the notation \( u \propto v \) stands for \( u = f(\theta, t, n, p, b, l_{dip}, \gamma_1) v \) with the function \( f > 0 \) for any \( d \in \Omega_{t+1, n, p} \) and \( \theta \geq 0 \), and the explicit form of \( f \) may vary from line to line. By direct calculation, we have
\[ \frac{\partial x_0'}{\partial b} \propto 1 + \frac{\beta_1'}{\gamma_1'} \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}} - 1 \]
\[ \propto (\beta_1' + \gamma_1') \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}} - \gamma_1' \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}} \]
\[ = \Psi_1 \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}} - \Psi_2 \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}}, \]
where
\[ \Psi_1 = \left( t + 1 - \frac{2}{p} - \frac{\theta t}{1 + \theta p} \right) (n(p - 1) - \tilde{r}_{d0}) + \frac{\theta t}{1 + \theta p} \sum_{i=1}^{t} \sum_{u=1}^{n} n_{d0u}n_{d0iu} - \frac{\theta}{1 + \theta p} \sum_{i=1}^{t} \sum_{u=1}^{n} n_{d0u}n_{d0iu} - (np - r_{d0})^2 \]
\[ + \frac{\theta}{1 + \theta p} \sum_{i=1}^{t} \sum_{u=1}^{n} n_{d0u} \tilde{n}_{d0u} - \frac{\theta}{1 + \theta p} \sum_{u=1}^{n} \tilde{n}_{d0u}^2, \]
\[ \Psi_2 = \left( t + 1 - \frac{2}{p} - \frac{\theta t}{1 + \theta p} \right) (n(p - 1) - \tilde{r}_{d0}) + \frac{\theta}{1 + \theta p} \sum_{i=1}^{n} n_{d0u} \tilde{n}_{d0u} + \frac{\theta}{1 + \theta p} \sum_{i=1}^{n} \tilde{n}_{d0u}^2. \]

Let \( r = \Psi_1 / \Psi_2 \), then (28) implies
\[ \frac{\partial x_0'}{\partial b} = 0 \quad \text{iff} \quad \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}} = \frac{\Psi_1}{\Psi_2} \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}}, \]
\[ \frac{\partial x_0'}{\partial b} > 0 \quad \text{iff} \quad \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}} < \frac{\Psi_1}{\Psi_2} \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}}, \]
\[ \frac{\partial x_0'}{\partial b} < 0 \quad \text{iff} \quad \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}} > \frac{\Psi_1}{\Psi_2} \sqrt{\frac{ta - (n(p - 1) - \tilde{r}_{d0})}{tb - (n - l_{d0p})^2}}. \]

Note that \( r \) is a positive constant with respect to the values of \( a \) and \( b \). (29)–(31) imply that the optimal vector \( (a, b) \in \mathbb{R}^2 \) maximizing \( x_0' \) has to satisfy the right-hand side of (29). By replacing \( b \) as a function of \( a \) using the equality in (29), we denote the updated expression of \( \beta_1' \) and \( x_0' \) by \( \beta_1'' \) and \( x_0'' \), respectively. Note that the latter two could also be obtained directly by plugging the following equations into \( \beta_1 \) and \( x_0 \), respectively:
\[ \sum_{i=1}^{t} r_{d0i} \tilde{r}_{d0i} = (1 - r) \left( a - \frac{(n(p - 1) - \tilde{r}_{d0})^2}{t} \right) + \frac{(np - r_{d0})(n(p - 1) - \tilde{r}_{d0})}{t}, \]
\[ \sum_{i=1}^{t} \tilde{r}_{d0i}^2 = (1 - r)^2 \left( a - \frac{(n(p - 1) - \tilde{r}_{d0})^2}{t} \right) + \frac{(np - r_{d0})^2}{t} \]
Notice that \( \beta_1'', \gamma_1, \) and \( x_0'' \) are functions of \( a \) without \( b \) involved. By direct calculation, we have
\[ \frac{\partial x_0''}{\partial a} \propto -\left( 1 - r + \frac{\beta_1''}{\gamma_1} \right) \leq 0. \]

Hence, \( x_0'' \) is maximized when \( a \) equals to its minimum, \( (n(p - 1) - \tilde{r}_{d0})^2 / t \). Then, by (25) and (29), the point of \( (a, b, c) \in \mathbb{R}^3 \) given by the following equations will attain the maximum of \( x_0' \):
\[ a = (n(p - 1) - \tilde{r}_{d0})^2 / t, \]
\[ b = (n - l_{d0p})^2 / t, \]
\[ c = (n(p - 1) - \tilde{r}_{d0})(n - l_{d0p}) / t. \]
Note that the preceding equations will be satisfied if
\[ \hat{r}_d = \hat{r}_{dj} \quad \text{and} \quad l_{dp} = l_{djp} \quad \text{for any} \ 1 \leq i \neq j \leq t. \quad (34) \]

By plugging the equations of (33) into \( \beta_1 \) and \( x_0 \), we denote the new terms derived by \( \beta_1'' \) and \( x_0'' \) respectively. Obviously, \( x_0 \leq x_0'' \) with equality holds when (34) holds. By Lemmas 6 and 7 in the supplemental material, we have
\[
\beta_1'' = \frac{\theta}{1 + \theta} \left( \frac{(t - 1)^2}{x_0 (t(x_1, x_2, x_3, \theta))} + \frac{t}{y_0 (x_1, x_2, x_3, \theta)} \right) \geq 0.
\]
That indicates that \( x_0'' \) increases when \( \sum_{i=1}^{t} \sum_{u=0}^{n} n_{du} \hat{n}_{du} - \sum_{u=1}^{n} n_{du0} \hat{n}_{du0} \) decreases. Notice that \( x_0 \) is obtained from \( x_0'' \) by replacing \( \sum_{i=1}^{t} \sum_{u=0}^{n} n_{du} \hat{n}_{du} \) by its minimum, \( (n(p - 1) - \hat{r}_{d0}) \), in view of Lemma 7. The conclusion is obtained.

Remark 1. In maximizing the \( x_0 \) in (20), the conditions \( \hat{r}_d = \hat{r}_{dj}, l_{dp} = l_{djp} \) minimize \( \sum_{i=1}^{t} \sum_{u=0}^{n} n_{du} \hat{n}_{du} \) and \( \sum_{i=1}^{t} \sum_{u=0}^{n} n_{du} \hat{n}_{du} \), however not \( \sum_{i=1}^{t} \sum_{u=0}^{n} n_{du} \hat{n}_{du} \). Suppose (n(p - 1) - \( \hat{r}_{d0} \))/(t - 1) is an integer, let \( r_{d1} = 0, l_{d1} = n - l_{d00}, \) and \( \hat{r}_d = (n(p - 1) - \hat{r}_{d0})/(t - 1) \) for \( i = 2, \ldots, t \), then we have
\[
\frac{t(t - 1)^2}{x_0} + \frac{t}{y_0} \geq \frac{t(t - 1)^2}{x_0} + \frac{t}{y_0}, \quad (37)
\]
for \( d \in \Lambda \) whenever \( p < t \).

To work on the right-hand side of (26) we now only need to investigate how it is influenced by \( \sum_{u=0}^{n} n_{du0} \hat{n}_{du0}, \sum_{u=0}^{n} n_{du0} \hat{n}_{du0}, \) and \( \sum_{u=0}^{n} \hat{n}_{du0}^2 \).

Lemma 11. (i) When \( p \geq 3 \) and \( t \geq \max(p - 1, 3) \), for any design \( d \in \Lambda_{t+1, n, p} \), we have
\[
t(t - 1)^2 + \frac{t}{y_0} \geq \frac{t(t - 1)^2}{x_0} + \frac{t}{y_0}, \quad (37)
\]
where \( x_0' \) and \( y_0' \) are derived from \( x_0 \) and \( y_0 \) respectively by replacing \( \sum_{u=0}^{n} n_{du0} \hat{n}_{du0}, \sum_{u=0}^{n} n_{du0} \hat{n}_{du0}, \) and \( \sum_{u=0}^{n} \hat{n}_{du0}^2 \) therein by their minimum with given \( r_{d0} \). Automatically, the equality in (37) holds for a design \( d^* \) which minimizes those three terms.

(ii) When \( p = 3 \) and \( t = 2 \), the above conclusion is still valid if \( r_{d0}/n \geq 0.6306 \).

Proof. By Lemma 6, we have
\[
\sum_{u=0}^{n} n_{du0} \hat{n}_{du0} \leq (t(n(p - 1) - \hat{r}_{d0}) \quad (38)
\]
which implies \( \hat{b}_1 \geq 0 \) for any \( \theta \geq 0 \). Denote by \( \xi_1, \xi_2, \) and \( \xi_3 \) the values of \( \sum_{u=0}^{n} n_{du0} \hat{n}_{du0}, \sum_{u=0}^{n} n_{du0} \hat{n}_{du0}, \) and \( \sum_{u=0}^{n} \hat{n}_{du0}^2 \). Among the restrictions by the nature of the design, we have
\[
t(n(p - 1) - \hat{r}_{d0}) \geq \xi_2 \geq \hat{r}_{d0} \geq \xi_3 \geq \hat{r}_{d0}.
\]
We can write \( \hat{b}_1 \) as a function \( \hat{b}_1(\xi_1, \xi_2, \xi_3, \theta) \) with the value \( r_{d0} \) fixed. The same argument could be applied to \( \hat{b}_1, \hat{z}_1, \hat{z}_2, \) and \( \hat{z}_3 \) and \( \hat{r}_{d0}, \) and we define
\[
H(\xi_1, \xi_2, \xi_3, \theta) = \frac{t(t - 1)^2}{x_0(\hat{r}_{d0}, \xi_2, \xi_3, \theta)} + \frac{t}{y_0(\hat{r}_{d0}, \xi_2, \xi_3, \theta)}. \quad (39)
\]
For notational convenience, we also define
\[
\Theta_1(\xi_1, \xi_2, \xi_3, \theta) = \frac{\hat{b}_1(\xi_1, \xi_2, \xi_3, \theta)}{\hat{y}_1(\xi_1, \xi_2, \xi_3, \theta)}, \quad \Theta_2(\xi_1, \xi_2, \xi_3, \theta) = \frac{\hat{b}_2(\xi_1, \xi_2, \xi_3, \theta)}{\hat{y}_2(\xi_1, \xi_2, \xi_3, \theta)}.
\]
In the following, we will omit the variables \( \hat{z}_1, \hat{z}_2, \xi_3 \) for the functions defined when there is no ambiguity. We have
\[
\frac{\partial H}{\partial z_1} \geq 0, \quad i = 1, 2, 3, \quad (41)
\]
for any value of \( \hat{z}_i, i = 1, 2, 3, \) and \( \theta \), it is enough to show
\[
\frac{\partial H}{\partial \xi_i} \geq 0, \quad i = 1, 2, 3, \quad (42)
\]
or
\[
\frac{\partial H(r_{d0}, \xi_i, \xi_3, \theta)}{\partial \hat{r}_{d0}} \geq 0, \quad i = 1, 2, 3, \quad (43)
\]
for any value of \( \theta \). Propositions 1–5 in the Supplemental Materials finish the proof.

Proof of Theorem 2. Combining Lemmas 1, 4(ii), 10, and 11, it is enough to prove that a totally balanced test-control incomplete crossover design satisfies:

- \( n_{du0} \) is either 0 or 1, \( 1 \leq i \leq t, 0 \leq u \leq n \)
- \( l_{dij} = r_{dij}/p, i = 0, \ldots, t \)
- \( T_{dV} V^{-1/2} \pr_{\perp} (\{1_{np}\}) V^{-1/2} T_{dV} \parallel T_{dV} V^{-1/2} \pr_{\perp} (\{1_{np}\}) V^{-1/2} F_{dV} \) are invariant under any permutation of test treatments
- \( \hat{r}_d = r_{d}, l_{dp} = l_{djp} \) for any \( 1 \leq i \neq j \leq t \)
- \( n_{du0} \) is minimized with respect to fixed \( r_{d0} \).

By comparing these conditions to the seven conditions in Definition 2, we conclude the theorem.

SUPPLEMENTAL MATERIALS

Supplemental Proofs: All the details of the proofs that were omitted in this article are given in this material. (Hedayat and Zheng- Supplementary.pdf)
REFERENCES
Cheng, C. S., and Wu, C. F. (1980), “Balanced Repeated Measurements Designs,” The Annals of Statistics, 8, 1272–1283. [1581]
Hedayat, A. S., and Afarinejad, K. (1975), “Repeated Measurements Designs, I,” A Survey of Statistical Design and Linear Models, ed. J. N. Srivastava, Amsterdam: North-Holland, pp. 229–240. [1581]
Hedayat, A. S., and Yang, M. (2003), “Universal Optimality of Balanced Uniform Crossover Designs,” The Annals of Statistics, 31, 978–983. [1581]
Hedayat, A. S., and Yang, M. (2004), “Universal Optimality for Selected Crossover Designs,” Journal of the American Statistical Association, 99, 461–466. [1581]
Hedayat, A. S., and Zheng, W. (2010), “Totally Balanced Test-Control Incomplete Crossover Designs and Their Statistical Applications,” Contemporary Mathematics, to appear. [1583,1585,1587]
Hedayat, A. S., Stufken, J., and Yang, M. (2006), “Optimal and Efficient Crossover Designs When Subject Effects Are Random,” Journal of the American Statistical Association, 101, 1031–1038. [1581,1582,1587]
Kunert, J. (1984), “Optimality of Balanced Uniform Repeated Measurements Designs,” The Annals of Statistics, 12, 1006–1017. [1581]
Kunert, J., and Martin, R. J. (2000), “On the Determination of Optimal Designs for an Interference Model,” The Annals of Statistics, 28, 1728–1742. [1581]
Kunert, J., and Stufken, J. (2002), “Optimal Crossover Designs in a Model With Self and Mixed Carryover Effects,” Journal of the American Statistical Association, 97, 898–906. [1581]
Kushner, H. B. (1997), “Optimal Repeated Measurements Designs: The Linear Optimality Equations,” The Annals of Statistics, 25, 2328–2344. [1581]
Kushner, H. B. (1998), “Optimal and Efficient Repeated-Measurements Designs for Uncorrelated Observations,” Journal of the American Statistical Association, 93, 1176–1187. [1581]
Stufken, J. (1991), “Some Families of Optimal and Efficient Repeated Measurements Designs,” Journal of Statistical Planning and Inference, 27, 75–83. [1581]
Stufken, J. (1996), “Optimal Crossover Designs,” in Design and Analysis of Experiments. Handbook of Statistics, Vol. 13, eds. S. Ghosh and C. R. Rao, Amsterdam: North-Holland, pp. 63–90. [1581]
Yang, M., and Park, M. (2007), “Efficient Crossover Designs for Comparing Test Treatments With a Control Treatment When $p = 3$,” Journal of Statistical Planning and Inference, 137, 2056–2067. [1587]
Yang, M., and Stufken, J. (2008), “Optimal and Efficient Crossover Designs for Comparing Test Treatments to a Control Treatment Under Various Models,” Journal of Statistical Planning and Inference, 138, 278–285. [1583]