The Annals of Statistics
2013, Vol. 41, No. 4, 2218–2235
DOI: 10.1214/13-AOS1148

OPTIMAL CROSSOVER DESIGNS FOR THE PROPORTIONAL MODEL

BY WEI ZHENG

Indiana University-Purdue University Indianapolis

In crossover design experiments, the proportional model, where the carryover effects are proportional to their direct treatment effects, has draw attentions in recent years. We discover that the universally optimal design under the traditional model is E-optimal design under the proportional model. Moreover, we establish equivalence theorems of Kiefer–Wolfowitz's type for four popular optimality criteria, namely A, D, E and T (trace).

1. Introduction. Let $\Omega_{p,t,n}$ be the collection of all crossover designs with p periods, t treatments, and n subjects. In an experiment based on design $d \in \Omega_{p,t,n}$, the response from subject $u \in \{1, 2, \dots, n\}$ in period $k \in \{1, 2, \dots, p\}$, to which treatment $d(k, u) \in \{1, 2, \dots, t\}$ was assigned by design d, is traditionally modeled as

$$Y_{dku} = \mu + \alpha_k + \beta_u + \tau_{d(k,u)} + \gamma_{d(k-1,u)} + \varepsilon_{ku}. \tag{1}$$

Here, μ is the general mean, α_k is the kth period effect, β_u is the uth subject effect, $\tau_{d(k,u)}$ is the (direct) effect of treatment $d(k, u)$, and $\gamma_{d(k-1,u)}$ is the carryover effect of treatment $d(k - 1, u)$ that subject u received in the previous period (by convention $\gamma_{d(0,u)} = 0$). A central problem in the area of crossover design is to find the best design among $\Omega_{p,t,n}$ for estimating the direct, and sometimes also carryover, treatment effects. Since Hedayat and Afsarinejad (1975, 1978) the optimal design problems have been mainly studied under model (1). Examples include Cheng and Wu (1980), Kunert (1984), Stufken (1991), Hedayat and Yang (2003, 2004) and Hedayat and Zheng (2010) among others. For approximate design solutions, see Kushner (1997, 1998), Kunert and Martin (2000), Kunert and Stufken (2002), and Zheng (2013a) among others.

Many variants of model (1) have been proposed in literature. The main focus is on different modelings of carryover effects, such as no carryover effects model, mixed carryover effects model [Kunert and Stufken (2002)] and the full interaction model [Park et al. (2011)]. The choice of model should be based on practical background and it is the responsibility of design theorists to provide recipes of optimal or efficient designs for each of these models. Here we consider model (2) below

Received April 2013; revised July 2013.
MSC2010 subject classifications. Primary 62K05; secondary 62J05.
Key words and phrases. Crossover designs, A-optimality, D-optimality, E-optimality, equivalence theorem, proportional model, pseudo symmetric designs.

because usually (i) it is essential to choose a parsimonious but reasonable model; (ii) The treatment having the larger direct effect in magnitude usually yields the larger carryover effect:

$$Y_{dku} = \mu + \alpha_k + \beta_u + \tau_{d(k,u)} + \lambda\tau_{d(k-1,u)} + \varepsilon_{ku}. \tag{2}$$

Throughout the paper, we call this model as the *proportional* model. Kempton, Ferris and David (2001) proposed this model and some theoretical results are later derived by Bailey and Kunert (2006) and Bose and Stufken (2007). The main difficulty is due to the nonlinear term $\lambda\tau_{d(k-1,u)}$ in the model. In this paper, we show that universally optimal designs for estimating treatment effects under the traditional model is E-optimal under the proportional model regardless the value of λ. Unlike the traditional model, the proportional model do not yield universally optimal designs in general. Instead, we derive equivalence theorems for four popular optimality criteria, namely A, D, E and T. Besides, we derive optimal designs for estimating λ.

The rest of this paper is organized as follows. Section 2 briefly introduces the universal optimality of Kushner's design under the traditional linear model as well as some necessary notation to be used for the rest of this paper. Section 3 studies the optimal design problem for the proportional model. Finally, Section 4 gives some examples of optimal designs under different values of p and t.

2. Some notation and Kushner's design. Let G be a temporary object whose meaning differs from context to context. For a square matrix G, we define G', G^- and $\mathrm{tr}(G)$ to represent the transpose, g-inverse and trace of G, respectively. The projection operator $\mathrm{pr}^\perp$ is defined as $\mathrm{pr}^\perp G = I - G(G'G)^- G'$. For two square matrices of equal size, G_1 and G_2, $G_1 \le G_2$ means that $G_2 - G_1$ is nonnegative definite. For a set G, the number of elements in the set is represented by $|G|$. Besides, I_k is the $k \times k$ identity matrix and 1_k is the vector of length k with all its entries as 1. We further define $J_k = 1_k 1_k'$ and $B_k = I_k - J_k/k$. Finally, $\otimes$ represents the Kronecker product of two matrices.

Let $Y_d = (Y_{d11}, Y_{d21}, \ldots, Y_{dp1}, Y_{d12}, \ldots, Y_{dpn})'$ be the $np \times 1$ response vector, then model (1) has the matrix form

$$Y_d = 1_{np}\mu + Z\alpha + U\beta + T_d\tau + F_d\gamma + \varepsilon, \tag{3}$$

where $\alpha = (\alpha_1, \ldots, \alpha_p)'$, $\beta = (\beta_1, \ldots, \beta_n)'$, $\tau = (\tau_1, \ldots, \tau_t)'$, $\gamma = (\rho_1, \ldots, \rho_t)'$, $Z = 1_n \otimes I_p$, $U = I_n \otimes 1_p$, and T_d and F_d denote the treatment/subject and carryover/subject incidence matrices. Here we assume $\mathbb{E}(\varepsilon) = 0$ and $\mathrm{Var}(\varepsilon) = I_n \otimes \Sigma$, where Σ is a nonsingular within subject covariance matrix. Define $\Sigma^{-1/2}$ to be the matrix such that $\Sigma^{-1} = \Sigma^{-1/2}\Sigma^{-1/2}$. Let $\tilde{T}_d = I_n \otimes \Sigma^{-1/2} T_d$, $\tilde{F}_d = I_n \otimes \Sigma^{-1/2} F_d$, $\tilde{Z} = I_n \otimes \Sigma^{-1/2} Z$ and $\tilde{U} = I_n \otimes \Sigma^{-1/2} U$. The information matrix for the direct treatment effect τ under model (3) is

$$\begin{aligned} C_d &= \tilde{T}_d' \,\mathrm{pr}^\perp(\tilde{Z}|\tilde{U}|\tilde{F}_d)\tilde{T}_d \\ &= C_{d11} - C_{d12} C_{d22}^- C_{d21}, \end{aligned}$$

where $C_{dij} = G_i' \operatorname{pr}^{\perp}(\tilde{Z}|\tilde{U})G_j, 1 \le i, j \le 2$ with $G_1 = \tilde{T}_d$ and $G_2 = \tilde{F}_d$. Define $\tilde{B} = \Sigma^{-1} - \Sigma^{-1}J_p\Sigma^{-1}/1_p'\Sigma^{-1}1_p$, and note that $\Sigma = I_p$ implies $\tilde{B} = B_p$. Straightforward calculations show that $C_{dij} = G_i'(B_n \otimes \tilde{B})G_j, 1 \le i, j \le 2$ with $G_1 = T_d$ and $G_2 = F_d$. A design is said to be universally optimal [Kiefer (1975)] if it maximizes $\Phi(C_d)$ for any Φ satisfying:

(C.1) Φ is concave.
(C.2) $\Phi(S'CS) = \Phi(C)$ for any permutation matrix S.
(C.3) $\Phi(bC)$ is nondecreasing in the scalar $b > 0$.

In approximate design theory, a design $d \in \Omega_{p,t,n}$ is considered as the result of selecting n elements with replacement from $\mathcal{S}$, the collection of all possible t^p treatment sequences. Now define the treatment sequence proportion $p_s = n_s/n$, where n_s is the number of replications of sequence s in the design. A design in approximate design theory is then identified by the vector $P_d = (p_s, s \in \mathcal{S})$ with the restrictions of $\sum_{s\in\mathcal{S}} p_s = 1$ and $p_s \ge 0$.

Let T_s (resp., F_s) be the $p \times t$ matrix T_d (resp., F_d) when d consists of a single sequence s. For sequence $s \in \mathcal{S}$ define $\hat{C}_{sij} = B_t G_i' \tilde{B} G_j B_t, 1 \le i, j \le 2$ with $G_1 = T_s$ and $G_2 = F_s$. By direct calculations, we have

$$C_{dij} = \hat{C}_{dij} - nG_i'\tilde{B}G_j, \qquad 1 \le i, j \le 2, \tag{4}$$

where $\hat{C}_{dij} = n\sum_{s\in\mathcal{S}} p_s\hat{C}_{sij}, 1 \le i, j \le 2$ with $G_1 = \sum_{s\in\mathcal{S}} p_s T_s B_t$ and $G_2 = \sum_{s\in\mathcal{S}} p_s F_s B_t$. Further, we define $c_{sij} = \operatorname{tr}(\hat{C}_{sij})$, $c_{dij} = \operatorname{tr}(\hat{C}_{dij}) = n\sum_{s\in\mathcal{S}} p_s c_{sij}$, the quadratic function $q_s(x) = c_{s11} + 2c_{s12}x + c_{s22}x^2$, $q(x) = \max_s q_s(x)$, $y^* = \min_{-\infty<x<\infty} q(x)$, x^* to be the unique solution of $q(x) = y^*$ and $\mathcal{Q} = \{s \in \mathcal{S} | q_s(x^*) = y^*\}$. Kushner (1997) derived the following theorem.

THEOREM 1 [Kushner (1997)]. *A design d is universally optimal under model (3) if and only if*

$$\sum_{s\in\mathcal{Q}} p_s[\hat{C}_{s11} + x^*\hat{C}_{s12}] = \frac{y^*}{t-1}B_t, \tag{5}$$

$$\sum_{s\in\mathcal{Q}} p_s[\hat{C}_{s21} + x^*\hat{C}_{s22}] = 0, \tag{6}$$

$$\sum_{s\in\mathcal{Q}} p_s\tilde{B}(T_s + x^*F_s)B_t = 0, \tag{7}$$

$$\sum_{s\in\mathcal{Q}} p_s = 1, \tag{8}$$

$$p_s = 0 \qquad \text{if } s \notin \mathcal{Q}. \tag{9}$$

Let σ be a permutation of the symbols $\{1, 2, \dots, t\}$. For a sequence $s = (t_1, \dots, t_p)$, we define $\sigma s = (\sigma(t_1), \dots, \sigma(t_p))$. Note that $q_s(x)$ is invariant to treatment permutations, that is,

$$q_s(x) = q_{\sigma s}(x), \qquad \sigma \in \mathcal{P}. \tag{10}$$

Define the design σd by $P_{\sigma d} = (p_{\sigma^{-1}s}, s \in \mathcal{S})$. A design d is said to be *symmetric* if $P_d = P_{\sigma d}$. Also we define symmetric blocks as $\langle s \rangle = \{\sigma s, \sigma \in \mathcal{P}\}$ where $\mathcal{P}$ is the collection of all possible $t!$ permutations. For a symmetric design, we have $p_{\tilde{s}} = p_{\langle s \rangle}/|\langle s \rangle|$ for any $\tilde{s} \in \langle s \rangle$, where $p_{\langle s \rangle} = \sum_{\tilde{s} \in \langle s \rangle} p_{\tilde{s}}$. Given p, t, n, a symmetric design d is uniquely determined by $(p_{\langle s \rangle}, s \langle \in \rangle \mathcal{S})$, where $s \langle \in \rangle \mathcal{S}$ means that s runs through all distinct symmetric blocks contained in $\mathcal{S}$. Equation (10) is essential for the following theorem.

THEOREM 2 [Kushner (1997)]. *A symmetric design is universally optimal under model* (3) *if*

$$\sum_{s \in \mathcal{Q}} p_s q_s'(x^*) = 0, \tag{11}$$

$$\sum_{s \in \mathcal{Q}} p_s = 1, \tag{12}$$

$$p_s = 0, \qquad \text{if } s \notin \mathcal{Q}, \tag{13}$$

where $q_s'(x)$ is the derivative of $q_s(x)$ with respective to x.

3. Proportional model.

3.1. *Problem formulation and literature review.* We are interested in model (2), which could be rewritten in the matrix form

$$Y_d = 1_{np}\mu + T_d\tau + \lambda F_d\tau + Z\alpha + U\beta + \varepsilon. \tag{14}$$

Here we assume $\varepsilon \sim N(0, I_n \otimes \Sigma)$. Fisher's information matrix for τ is

$$\begin{aligned} C_{d,\tau_0,\lambda_0}(\tau) &= (\tilde{T}_d + \lambda_0 \tilde{F}_d)' \operatorname{pr}^{\perp}(\tilde{Z}|\tilde{U}|\tilde{F}_d\tau_0)(\tilde{T}_d + \lambda_0 \tilde{F}_d) \\ &= C_{d11} + \lambda_0(C_{d12} + C_{d21}) + \lambda_0^2 C_{d22} \\ &\quad - (C_{d12} + \lambda_0 C_{d22})\tau_0\big(\tau_0' C_{d22}\tau_0\big)^{-1}\tau_0'(C_{d21} + \lambda_0 C_{d22}). \end{aligned} \tag{15}$$

Unlike model (3), model (14) is nonlinear, and therefore the choice of optimal designs depends on the unknown parameters λ_0 and τ_0; see (15). The nonlinearity of the model imposes the major difficulty on the problem. For this, Bose and Stufken (2007) assumes that λ_0 is a known parameter at the stage of data analysis, in which case the (Fisher's) information matrix does not depend on τ_0 and hence the same for the choice of optimal designs. But such strategy inevitably yields significant bias in the analysis stage when one do not have sufficient knowledge about λ_0.

Note that Kempton, Ferris and David (2001) and Bailey and Kunert (2006) also worked on $C_{d,\tau_0,\lambda_0}(\tau)$ even though they derived it from the aspect of model approximation [Fedorov and Hackl (1997), page 18] without normality assumption. For unknown τ_0 and λ_0, they adopted the following Bayesian type of criteria:

$$\phi_{g,\lambda_0}(d) = \int \Phi\big(C_{d,\tau_0,\lambda_0}(\tau)\big)g(\tau_0)\,d(\tau_0) = \mathbb{E}_g\big(\Phi\big(C_{d,\tau_0,\lambda_0}(\tau)\big)\big), \tag{16}$$

where g is the prior distribution of τ_0. Note that they only considered the special case of $\Sigma = I_p$ and Φ being the A-criterion function. Particularly, Kempton, Ferris and David (2001) gave a search algorithm for A-efficient designs when g is the density function of a special multivariate normal distribution. Bailey and Kunert (2006) proved the optimality of totally balanced design [Kunert and Stufken (2002)] when $\Sigma = I_p$, $3 \le p \le t$, the distribution g is exchangeable, and $-1 \le \lambda_0 < \lambda^*$ with

$$\lambda^* = \frac{1}{p-1} - \frac{pt-t-1}{(p-1)(t-2)(pt-t-1-t/p)^2}. \tag{17}$$

Note that $0 < \lambda^* < 1/(p-1)$. Hence the results of Bailey and Kunert (2006) will not be applicable when $p \ge t$ or the carryover effects is positively proportional to the direct treatment effects with a moderate or even larger magnitude. Here, we develop tools for finding optimal designs for any value of λ_0 and Σ and for four popular criteria, namely A, D, E and T. For E-criterion, the optimal design does not depend on the value of λ_0.

3.2. *Preliminary results.* Recall that the design σd is defined by $P_{\sigma d} = (p_{\sigma^{-1}s}, s \in \mathcal{S})$. Let S_σ be the unique permutation matrix such that $T_{\sigma d} = T_d S_\sigma$ and $F_{\sigma d} = F_d S_\sigma$ for any design d. Also define

$$\sigma\tau_0 = S_\sigma \tau_0.$$

Let δ_{τ_0} be the probability measure which puts equal mass to each element in $\{\sigma\tau_0 | \sigma \in \mathcal{P}\}$. We shall focus on the special case of $g = \delta_{\tau_0}$ and then extend the results to any arbitrary exchangeable distribution g. By definition we have

$$\phi_{\delta_{\tau_0},\lambda_0}(d) = \frac{1}{t!}\sum_{\sigma} \Phi\big(C_{d,\sigma\tau_0,\lambda_0}(\tau)\big),$$

where the summation runs through all $t!$ permutations. Now we have:

THEOREM 3. *In approximate design theory, given any values of the real number λ_0 and the vector τ_0, for any design d there exists a symmetric design, say d^*, such that*

$$\phi_{\delta_{\tau_0},\lambda_0}(d) \le \phi_{\delta_{\tau_0},\lambda_0}(d^*). \tag{18}$$

PROOF. First, we observe that

$$C_{\sigma d,\tau_0,\lambda_0}(\tau) = S'_\sigma C_{d,\sigma\tau_0,\lambda_0}(\tau) S_\sigma. \tag{19}$$

For any given permutation σ_0, by (19) we have

$$\begin{aligned}
\phi_{\delta_{\tau_0},\lambda_0}(\sigma_0 d) &= \frac{1}{t!}\sum_\sigma \Phi\big(C_{\sigma_0 d,\sigma\tau_0,\lambda_0}(\tau)\big) \\
&= \frac{1}{t!}\sum_\sigma \Phi\big(S'_{\sigma_0} C_{d,\sigma_0\sigma\tau_0,\lambda_0}(\tau) S_{\sigma_0}\big) \\
&= \frac{1}{t!}\sum_\sigma \Phi\big(C_{d,\sigma_0\sigma\tau_0,\lambda_0}(\tau)\big) \\
&= \phi_{\delta_{\tau_0},\lambda_0}(d).
\end{aligned} \tag{20}$$

By direct calculations, we have

$$\begin{aligned}
C_{d,\tau_0,\lambda_0}(\tau,\lambda,\alpha) &= \big(\tilde{T}_d + \lambda_0 \tilde{F}_d | \tilde{F}_d \tau_0 | \tilde{Z}\big)' \operatorname{pr}^\perp(\tilde{U})\big(\tilde{T}_d + \lambda_0 \tilde{F}_d | \tilde{F}_d \tau_0 | \tilde{Z}\big) \\
&= n \sum_s p_s \big(T_s + \lambda_0 F_s | F_s \tau_0 | I_p\big)' \tilde{B}\big(T_s + \lambda_0 F_s | F_s \tau_0 | I_p\big).
\end{aligned} \tag{21}$$

Define d^* to be the design such that

$$P_{d^*} = \frac{1}{t!}\sum_\sigma P_{\sigma d}.$$

It is easy to show that d^* is a symmetric design and

$$C_{d^*,\tau_0,\lambda_0}(\tau,\lambda,\alpha) = \frac{1}{t!}\sum_\sigma C_{\sigma d,\tau_0,\lambda_0}(\tau,\lambda,\alpha), \tag{22}$$

in view of (21). By Lemma 3.1 of Kushner (1997) and (22), we have

$$\frac{1}{t!}\sum_\sigma C_{\sigma d,\tau_0,\lambda_0}(\tau) \le C_{d^*,\tau_0,\lambda_0}(\tau). \tag{23}$$

By (20) and (23), we have

$$\begin{aligned}
\phi_{\delta_{\tau_0},\lambda_0}(d) &= \frac{1}{t!}\sum_\sigma \phi_{\delta_{\tau_0},\lambda_0}(\sigma d) \\
&= \frac{1}{(t!)^2}\sum_\sigma\sum_{\tilde{\sigma}} \Phi\big(C_{\sigma d,\tilde{\sigma}\tau_0,\lambda_0}(\tau)\big) \\
&= \frac{1}{(t!)^2}\sum_{\tilde{\sigma}}\sum_\sigma \Phi\big(C_{\sigma d,\tilde{\sigma}\tau_0,\lambda_0}(\tau)\big) \\
&\le \frac{1}{t!}\sum_{\tilde{\sigma}} \Phi\big(C_{d^*,\tilde{\sigma}\tau_0,\lambda_0}(\tau)\big) \\
&= \phi_{\delta_{\tau_0},\lambda_0}(d^*).
\end{aligned}$$

□

REMARK 1. In proving Theorem 3 we use the same approach in the proof of Theorem 3.2 of Kushner (1997) to derive (23). However, the proof of the latter theorem is not rigorous since (3.6) therein does not hold in general. Actually the gap can be filled by using (22) in replacement of (3.6) therein.

COROLLARY 1. *In approximate design theory, given any value the number* λ_0 *and the prior distribution* g *of* τ_0 *as long as the latter is exchangeable, for any design* d *there exists a symmetric design, say* d^*, *such that*

$$\phi_{g,\lambda_0}(d) \le \phi_{g,\lambda_0}(d^*).$$

PROOF. It is enough to notice that inequality (18) holds for any τ_0. □

By Corollary 1, there always exists a symmetric design which is optimal among $\Omega_{p,t,n}$. We define a design d to be *pseudo symmetric* if all treatments in d are equally replicated on each period and C_{dij}, $1 \le i, j \le 2$ are completely symmetric. A symmetric design is pseudo symmetric and thus an optimal design in the subclass of pseudo symmetric designs is automatically optimal among $\Omega_{p,t,n}$.

PROPOSITION 1. *Regardless the value of* τ_0, *Fisher's information matrix* $C_{d,\tau_0,\lambda_0}(\tau)$ *of a symmetric design* d *has eigenvalues of* $0, (t-1)^{-1}(c_{d11} - c_{d12}^2/c_{d22})$ *and* $(t-1)^{-1}(c_{d11} + 2\lambda_0 c_{d12} + \lambda_0^2 c_{d22})$ *with multiplicities* $1, 1$ *and* $t-2$, *respectively.*

PROOF. For a symmetric design d, we have $\sum_{s\in\mathcal{S}} p_s T_s B_t = 0 = \sum_{s\in\mathcal{S}} p_s F_s B_t$ and hence

$$C_{dij} = \hat{C}_{dij}, \qquad 1 \le i, j \le 2, \tag{24}$$

in view of (44). Moreover, these matrices are all completely symmetric and have row and column sum as zero, which together with (24) yields $C_{dij} = c_{dij} B_t/(t-1)$. Due to $1'\tau_0 = 0$ and hence $B_t \tau_0 = \tau_0$, we have

$$(t-1)C_{d,\tau_0,\lambda_0}(\tau) = \left(c_{d11} + 2\lambda_0 c_{d12} + \lambda_0^2 c_{d22}\right) B_t - \frac{(c_{d12} + \lambda_0 c_{d22})^2}{c_{d22}} \frac{\tau_0 \tau_0'}{\tau_0' \tau_0}.$$

Let $\{x_1, \dots, x_{t-2}\}$ be the orthogonal basis which is orthogonal to both 1 and τ_0. Then $\{x_1, \dots, x_{t-2}, \tau_0, 1\}$ forms the eigenvectors for the above matrix. Hence, the lemma is concluded. □

REMARK 2. Since c_{sij} is the same for sequences in the same symmetric block $\langle s \rangle$, we have $c_{dij} = \sum_{s\langle\in\rangle\mathcal{S}} p_{\langle s\rangle} c_{sij}$. In view of Corollary 1 and Proposition 1, one can derive an optimal design in two steps. First, we find the optimum value of $p_{\langle s\rangle}$ for all distinct symmetric blocks. Within each symmetric block with positive $p_{\langle s\rangle}$, we construct a pseudo symmetric design, and then assemble these designs according to the desired value of $p_{\langle s\rangle}$. For step one, see equivalence theorems in Section 3.4. For step two, one can utilize some combinatory structures such as type I orthogonal arrays, for the latter see Design 6 of Bailey and Kunert (2006), for example. For E-criterion, more general optimal designs could be derived. See Section 3.3 for details.

REMARK 3. The application of Corollary 1 and Proposition 1 for A-criterion leads to Proposition 1 of Bailey and Kunert (2006).

3.3. *E-optimality.* Let $\mathcal{E}_{g,\lambda_0}(d)$ be the criterion $\phi_{g,\lambda_0}(d)$ when Φ therein is evaluated by the second smallest eigenvalue of the information matrix. We call a design to be $\mathcal{E}_{g,\lambda_0}$-optimal if it maximizes $\mathcal{E}_{g,\lambda_0}(d)$.

PROPOSITION 2. *In approximate design theory, regardless the value of λ_0 and the prior distribution g as long as the latter is exchangeable, a design d is $\mathcal{E}_{g,\lambda_0}$-optimal if and only if $\mathcal{E}_{g,\lambda_0}(d) = ny^*/(t-1)$ with y^* as defined right before Theorem 1.*

PROOF. First, it is easy to verify that

$$c_{d11} - c_{d12}^2/c_{d22} \le c_{d11} + 2\lambda_0 c_{d12} + \lambda_0^2 c_{d22}$$

for any λ_0. By Theorem 4.5 of Kushner (1997), we have

$$\begin{aligned} y^* &= \min_{-\infty<x<\infty} \sum_{s\in\mathcal{S}} p_s q_s(x) \\ &= n^{-1} \max_d \bigl(c_{d11} - c_{d12}^2/c_{d22}\bigr). \end{aligned}$$

Hence, the proposition is proved in view of Corollary 1 and Proposition 1. □

THEOREM 4. *In approximate design theory, regardless of the value of λ_0 and the prior distribution g as long as the latter is exchangeable, a design is $\mathcal{E}_{g,\lambda_0}$-optimal if there exists a real number x such that*

$$\sum_{s\in\mathcal{Q}} p_s[\hat{C}_{s11} + x\hat{C}_{s12}] = \frac{y^*}{t-1} B_t, \tag{25}$$

$$\sum_{s\in\mathcal{Q}} p_s[\hat{C}_{s21} + x\hat{C}_{s22}] = 0, \tag{26}$$

$$\sum_{s\in\mathcal{Q}} p_s \tilde{B}(T_s + xF_s)B_t = 0, \tag{27}$$

$$\sum_{s\in\mathcal{Q}} p_s = 1, \tag{28}$$

$$p_s = 0 \qquad \textit{if } s \notin \mathcal{Q}. \tag{29}$$

PROOF. Since $C_{d,\tau_0,\lambda_0}(\tau)$ have column and row sums as zero, we have

$$\mathcal{E}_{g,\lambda_0}(d) = \mathbb{E}_g\Big[\min_{\ell' 1_t=0, \ell'\ell=1} \ell' C_{d,\tau_0,\lambda_0}(\tau)\ell\Big]. \tag{30}$$

For a design satisfying (25)–(29) we have

$$C_{d11} + xC_{d12} = \frac{ny^*}{t-1}B_t, \tag{31}$$

$$C_{d21} + xC_{d22} = 0, \tag{32}$$

in view of (44). Since C_{d22} is symmetric, (32) implies the symmetry of C_{d21} and hence $C_{12} = C_{21}$. Then by (15), (31) and (32), we have

$$C_{d,\tau_0,\lambda_0}(\tau) = \frac{ny^*}{t-1}B_t + (\lambda_0 - x)^2 C_{d22} - \frac{(\lambda_0 - x)^2}{\tau_0' C_{d22}\tau_0} C_{d22}\tau_0\tau_0' C_{d22}. \tag{33}$$

Let $\{0, a_1, \dots, a_{t-1}\}$ be the eigenvalues of C_{d22} with corresponding normalized eigenvectors $\{1_t, \ell_1, \dots, \ell_{t-1}\}$, then we have $C_{d22} = \sum_{i=1}^{t-1} a_i \ell_i \ell_i'$. Since $\tau_0' 1_t = 0$, we have the representation $\tau_0 = \sum_{i=1}^{t-1} c_i \ell_i$. For any vector ℓ with $\ell' 1_t = 0$ and $\ell'\ell = 1$, we have the expression of $\ell = \sum_{i=1}^{t-1} b_i \ell_i$ with the restriction $\sum_{i=1}^{t-1} b_i^2 = 1$, the equation $\ell' B_t \ell = 1$, and hence by (33)

$$\begin{aligned}
\ell' C_{d,\tau_0,\lambda_0}(\tau)\ell &= \frac{ny^*}{t-1} + (\lambda_0 - x)^2 \sum_{i=1}^{t-1} a_i b_i^2 - \frac{(\lambda_0 - x)^2}{\sum_{i=1}^{t-1} a_i c_i^2}\left(\sum_{i=1}^{t-1} a_i b_i c_i\right)^2 \\
&= \frac{ny^*}{t-1} + \frac{(\lambda_0 - x)^2}{\sum_{i=1}^{t-1} a_i c_i^2}\left[\left(\sum_{i=1}^{t-1} a_i b_i^2\right)\left(\sum_{i=1}^{t-1} a_i c_i^2\right) - \left(\sum_{i=1}^{t-1} a_i b_i c_i\right)^2\right] \qquad (34) \\
&\geq \frac{ny^*}{t-1},
\end{aligned}$$

the equality holds if and only if $\ell = \tau_0 / \|\tau_0\|$. The theorem is concluded in view of Proposition 2, (30) and (34). $\square$

REMARK 4. The advantage of Theorem 4 is that the design therein is optimal for any λ_0 while the A-optimality of totally balanced design [Bailey and Kunert (2006)] requires the condition of $-1 \leq \lambda_0 \leq \lambda^*$.

As a direct result of Theorems 1 and 4, we have the following corollary.

COROLLARY 2. *In approximate design theory, regardless the value of* λ_0 *and the prior distribution* g *as long as the latter is exchangeable, a universally optimal design for model* (3) *is also* $\mathcal{E}_{g,\lambda_0}$*-optimal for model* (14).

THEOREM 5. *The variable* x *in* (25)–(29) *takes the unique value of* x^*, *which is defined right above Theorem* 1.

PROOF. Given (25)–(29), we have by (31) and (32) that

$$\begin{aligned} C_{d11} - C_{d12}(C_{d22})^{-}C_{d21} &= C_{d11} + xC_{d12}(C_{d22})^{-}C_{d22} \\ &= C_{d11} + xC_{d12} \\ &= \frac{ny^*}{t-1}B_t, \end{aligned}$$

which indicates that d is universally optimal for model (3) in view of Theorem 1. Hence, we have $x = x^*$ by Theorem 1. □

As a direct result of Theorem 2, Corollary 2 and Remark 2, we have Corollary 3.

COROLLARY 3. *In approximate design theory, regardless of the value of* λ_0 *and the prior distribution* g *as long as the latter is exchangeable, a pseudo symmetric design is* $\mathcal{E}_{g,\lambda_0}$*-optimal if it satisfies* (11)–(13).

3.4. *Equivalence theorems.* In order to introduce the following results, we define $x_d = -c_{d12}/c_{d22}$ and $q_d(x) = \sum_{s\in\mathcal{S}} p_s q_s(x)$. Then we have

$$\begin{aligned} nq_d(x_d) &= c_{d11} - c_{d12}^2/c_{d22}, \\ nq_d(\lambda_0) &= c_{d11} + 2\lambda_0 c_{d12} + \lambda_0^2 c_{d22}. \end{aligned}$$

For a $t \times t$ matrix C with eigenvalues $0 = a_0 \le a_1 \le a_2 \le \cdots \le a_{t-1}$, define the criterion functions

$$\begin{aligned} \Phi_A(C) &= (t-1)\left(\sum_{i=1}^{t-1} a_i^{-1}\right)^{-1}, \\ \Phi_D(C) &= \left(\prod_{i=1}^{t-1} a_i\right)^{1/(t-1)}, \\ \Phi_T(C) &= (t-1)^{-1}\sum_{i=1}^{t-1} a_i. \end{aligned}$$

Let $\mathcal{A}_{g,\lambda_0}(d)$, $\mathcal{D}_{g,\lambda_0}(d)$ and $\mathcal{T}_{g,\lambda_0}(d)$ be the criterion $\phi_{g,\lambda_0}(d)$ when Φ therein is evaluated by Φ_A, Φ_D and Φ_T, respectively. We call a design to be $\mathcal{A}_{g,\lambda_0}$-optimal if it maximizes $\mathcal{A}_{g,\lambda_0}(d)$. Definitions for optimality of $\mathcal{D}_{g,\lambda_0}$ and $\mathcal{T}_{g,\lambda_0}$ are similar.

THEOREM 6. *In approximate design theory, regardless of the value of* λ_0 *and the prior distribution* g *as long as the latter is exchangeable, a pseudo symmetric design* d *is* $\mathcal{D}_{g,\lambda_0}$*-optimal if and only if*

$$\max_{s\in\mathcal{S}}\left(\frac{1}{t-1}\frac{q_s(x_d)}{q_d(x_d)} + \frac{t-2}{t-1}\frac{q_s(\lambda_0)}{q_d(\lambda_0)}\right) = 1. \tag{35}$$

Moreover, the sequences in design d *attain the maximum in* (35).

PROOF. For a real number x, let $\eta(\xi_1, \xi_2, x) = q_{\xi_2}(x)/q_{\xi_1}(x)$ with ξ_1 and ξ_2 being either a design or a sequence. Also define $\psi(P_d) = \log((c_{d11} - c_{d12}^2/c_{d22})) + (t-2)\log(c_{d11} + 2\lambda_0 c_{d12} + \lambda_0^2 c_{d22})$. By Theorem 2, Corollary 2, Remark 2 and the concavity of D-criterion, a pseudo symmetric design d^* is $\mathcal{D}_{g,\lambda_0}$-optimal if and only if for any other design d we have

$$\begin{aligned} 0 &\geq \lim_{\delta\to 0} \frac{\psi((1-\delta)P_{d^*} + \delta P_d) - \psi(P_{d^*})}{\delta} \\ &= \eta(d^*, d, x_{d^*}) + (t-2)\eta(d^*, d, \lambda_0) - (t-1). \end{aligned} \tag{36}$$

Take d in (37) to be a design consist of a single sequence s, we have

$$\max_{s\in\mathcal{S}}\left(\frac{1}{t-1}\eta(d^*, s, x_{d^*}) + \frac{t-2}{t-1}\eta(d^*, s, \lambda_0)\right) \leq 1. \tag{37}$$

Observe that

$$\eta(d, d, x_d) = 1 = \eta(d, d, \lambda_0), \tag{38}$$

then we have

$$\max_{s\in\mathcal{S}}\left(\frac{1}{t-1}\eta(d, s, x_d) + \frac{t-2}{t-1}\eta(d, s, \lambda_0)\right) \geq 1. \tag{39}$$

The theorem is completed in view of (37), (38) and (39). □

THEOREM 7. *In approximate design theory, regardless of the value of λ_0 and the prior distribution g as long as the latter is exchangeable, a pseudo symmetric design d is $\mathcal{A}_{g,\lambda_0}$-optimal if and only if*

$$\max_{s\in\mathcal{S}}\left(\pi_d \frac{q_s(x_d)}{q_d(x_d)} + (1-\pi_d)\frac{q_s(\lambda_0)}{q_d(\lambda_0)}\right) = 1, \tag{40}$$

where $\pi_d = q_d(\lambda_0)/(q_d(\lambda_0) + (t-2)q_d(x_d))$. Moreover, the sequences in design d attain the maximum in (40).

REMARK 5. Theorem 7 is essentially a generalization of the result of Bailey and Kunert (2006).

THEOREM 8. *In approximate design theory, regardless of the value of λ_0 and the prior distribution g as long as the latter is exchangeable, a pseudo symmetric design d is $\mathcal{E}_{g,\lambda_0}$-optimal if and only if*

$$\max_{s\in\mathcal{S}} \frac{q_s(x_d)}{q_d(x_d)} = 1. \tag{41}$$

Moreover, the sequences in design d attain the maximum in (41).

REMARK 6. In fact, (41) is equivalent to (11)–(13).

THEOREM 9. *In approximate design theory, regardless of the value of* λ_0 *and the prior distribution* g *as long as the latter is exchangeable, a pseudo symmetric design* d *is* $\mathcal{T}_{g,\lambda_0}$*-optimal if and only if*

$$\max_{s\in\mathcal{S}} \frac{q_s(x_d)+(t-2)q_s(\lambda_0)}{q_d(x_d)+(t-2)q_d(\lambda_0)} = 1. \tag{42}$$

Moreover, the sequences in design d *attain the maximum in* (42).

3.5. *Estimation of* λ_0. By the Cramér–Rao inequality, the variance of an unbiased estimator of λ_0 is bounded by the reciprocal of

$$C_{d,\tau_0,\lambda_0}(\lambda) = \tau_0'\tilde{F}_d' \operatorname{pr}^{\perp}(\tilde{T}_d+\lambda_0\tilde{F}_d|\tilde{Z}|\tilde{U})\tilde{F}_d\tau_0, \tag{43}$$

achievable by MLE asymptotically. Define $A_{d11} = F_d'(B_n\otimes\tilde{B})F_d$, $A_{d11} = F_d'(B_n\otimes\tilde{B})(T_d+\lambda_0F_d)$, $A_{d21} = A_{d12}'$ and $A_{d22} = (T_d+\lambda_0F_d)'(B_n\otimes\tilde{B})(T_d+\lambda_0F_d)$. Straightforward calculations show that $C_{d,\tau_0,\lambda_0}(\lambda) = \tau_0'A_d\tau_0$ where $A_d = A_{d11} - A_{d12}(A_{d22})^{-}A_{21}$.

As in (16) we define $\varphi_{g,\lambda_0}(d) = \mathbb{E}_g\tau_0'A_d\tau_0$, where the expectation is taken with respect to the prior distribution measure g. Then we have $\varphi_{\delta_{\tau_0},\lambda_0}(d) = \tau_0'\bar{A}_d\tau_0$ where $\bar{A}_d = \frac{1}{t!}\sum_\sigma S_\sigma' A_d S_\sigma$.

For each sequence s, define $\hat{A}_{sij} = B_tG_i'\tilde{B}G_jB_t$, $1\le i,j\le 2$ with $G_1 = F_s$ and $G_2 = T_s+\lambda_0F_s$. By direct calculations, we have

$$A_{dij} = \hat{A}_{dij} - nG_i'\tilde{B}G_j, \qquad 1\le i,j\le 2, \tag{44}$$

where $\hat{A}_{dij} = n\sum_{s\in\mathcal{S}} p_s\hat{A}_{sij}$, $1\le i,j\le 2$, $G_1 = \sum_{s\in\mathcal{S}} p_sF_sB_t$ and $G_2 = \sum_{s\in\mathcal{S}} p_s(T_s+\lambda_0F_s)B_t$.

Further define $h_{sij} = \operatorname{tr}(\hat{A}_{sij})$, $h_{dij} = \operatorname{tr}(\hat{A}_{dij}) = n\sum_{s\in\mathcal{S}} p_sh_{sij}$, the quadratic function $r_s(x) = h_{s11} + 2h_{s12}x + h_{s22}x^2$, $r(x) = \max_s r_s(x)$, $y_0 = \min_{-\infty<x<\infty} r(x)$, x_0 to be the unique solution of $r(x) = y_0$ and $\mathcal{R} = \{s\in\mathcal{S}|r_s(x_0) = y_0\}$. Now we have the following theorem.

THEOREM 10. *Given any* $-\infty<\lambda_0<\infty$, *a design maximizes* $\varphi_{\delta_{\tau_0},\lambda_0}(d)$ *for any* τ_0 *if*

$$\sum_{s\in\mathcal{R}} p_s[\hat{A}_{s11} + x_0\hat{A}_{s12}] = \frac{y_0}{t-1}B_t, \tag{45}$$

$$\sum_{s\in\mathcal{R}} p_s[\hat{A}_{s21} + x_0\hat{A}_{s22}] = 0, \tag{46}$$

$$\sum_{s\in\mathcal{R}} p_s\tilde{B}[\hat{F}_s + x_0(\hat{T}_s+\lambda_0\hat{F}_s)] = 0, \tag{47}$$

$$\sum_{s\in\mathcal{R}} p_s = 1, \tag{48}$$

$$p_s = 0 \qquad \textit{if } s\notin\mathcal{R}. \tag{49}$$

PROOF. Note that for any design d, there exists a symmetric design d^* with $\bar{A}_d \leq A_{d^*}$ by the same argument as for (23). For a symmetric design, we have $A_d = (t-1)^{-1}(h_{d11} - h_{d12}^2/h_{d22})B_t$. By similar arguments as in proof of Theorem 4.4 of Kushner (1997), we have $\max_d(h_{d11} - h_{d12}^2/h_{d22}) = ny_0$. By direct calculations, we know that (45)–(49) implies $A_d = ny_0 B_t/(t-1)$ and hence the theorem is proved. □

COROLLARY 4. *Given any value of the real number* λ_0, *a design maximizes* $\varphi_{g,\lambda_0}(d)$ *for any exchangeable prior distribution* g *if it satisfies* (45)–(49).

PROOF. The necessity is immediate. For sufficiency, it is enough to note that the design satisfying (45)–(49) does not depend on τ_0. □

COROLLARY 5. *Given any value of the real number* λ_0, *a design maximizes* $\varphi_{g,\lambda_0}(d)$ *for any exchangeable prior distribution* g *if it is a symmetric design with*

$$\sum_{s\langle\in\rangle\mathcal{R}} p_{\langle s\rangle} r_s'(x_0) = 0, \tag{50}$$

$$\sum_{s\langle\in\rangle\mathcal{R}} p_{\langle s\rangle} = 1, \tag{51}$$

$$p_s = 0 \qquad \text{if } s \notin \mathcal{R}, \tag{52}$$

where $r_s'(x)$ *is the derivative of* $r_s(x)$ *with respect to* x.

PROOF. It is enough to show that (50)–(52) implies (45)–(49). The proof of the latter is analogous to that of Theorem 2. □

A general necessary and sufficient optimality condition is given by the following.

THEOREM 11. *Given any value of the real number* λ_0, *a design maximizes* $\varphi_{g,\lambda_0}(d)$ *for any exchangeable prior distribution* g *if and only if* $\operatorname{tr}(A_d) = ny_0$.

PROOF. Note that $\bar{A}_d = \operatorname{tr}(A_d)B_t/(t-1)$, hence we have $\varphi_{\delta_{\tau_0},\lambda_0}(d) = \tau_0'\tau_0 \operatorname{tr}(A_d)/(t-1)$ and hence

$$\varphi_{g,\lambda_0}(d) = \frac{\mathbb{E}_g(\tau_0'\tau_0)}{t-1} \operatorname{tr}(A_d). \tag{53}$$

Through the proof of Theorem 10, we know that $\max_d \operatorname{tr}(A_d) = ny_0$, which together with (53) proves the theorem. □

4. Examples. In the spirit of Theorems 6–9 and Remark 2, we consider examples of optimal designs in the format of pseudo symmetric designs, even though a more general format could be proposed for E-optimality due to Theorem 4. Let m to be the total number of distinct symmetric blocks and suppose $s_1, s_2, \ldots, s_m$ are the representative sequences for each of the symmetric blocks. For a design d, define the vector $P_{\langle d \rangle} = (p_{\langle s_1 \rangle}, p_{\langle s_2 \rangle}, \ldots, p_{\langle s_m \rangle})$. Then two pseudo symmetric designs with the same $P_{\langle d \rangle}$ will have the same $\phi_{g,\lambda_0}(d)$ for any ϕ, g, λ_0 as long as g is exchangeable. In particular, they are equivalent in terms of $\mathcal{A}_{g,\lambda_0}$-, $\mathcal{D}_{g,\lambda_0}$-, $\mathcal{E}_{g,\lambda_0}$- and $\mathcal{T}_{g,\lambda_0}$-optimality. In the sequel, we will mainly focus on the determination of $P_{\langle d \rangle}$ based on the equivalence theorems 6–9. A general algorithm could be found in the supplemental article [Zheng (2013)].

For the following examples, we consider first order autocorrelation for within subject covariance matrix, namely $\Sigma = (\rho \mathbb{I}_{|i-j|=1} + \mathbb{I}_{i=j})_{1 \le i,j \le p}$, where $\mathbb{I}$ is the indicator function. Hence, $\rho = 0$ implies $\Sigma = I_p$. Following Kushner (1998), we define two special symmetric blocks. The symmetry block $\langle \mathrm{di} \rangle$ consists of all sequences having distinct treatments in the p periods. The symmetry block $\langle \mathrm{re} \rangle$ consists of all sequences having distinct treatments in the first $p-1$ periods, with the treatment in period $p-1$ repeating in period p. All examples given below are pseudo symmetric designs except otherwise specified. For ease of illustration by examples, we only consider $(\rho, \lambda_0) \in \{-1/2, 0, 1/2\} \times [-1, 1]$, even though other values of (ρ, λ_0) does not cause extra difficulty. Throughout this section, g is exchangeable unless otherwise specified.

Case of $(p, t) = (3, 3)$: Let d_1 be a design with $p_{\langle \mathrm{re} \rangle} = 1/6$ and $p_{\langle \mathrm{di} \rangle} = 5/6$. See, for instance, Example 1 of Kushner (1998) with $n = 36$ subjects. Define d_2 to be a design with $p_{\langle \mathrm{di} \rangle} = 1$, which requires n to be a multiple of 6 as an exact design. When $\rho = 0$, Theorem 8 shows the $\mathcal{E}_{g,\lambda_0}$-optimality of d_1 for any λ_0 and Bailey and Kunert (2006) shows the $\mathcal{A}_{g,\lambda_0}$-optimality of d_2 when $-1 \le \lambda_0 \le \lambda^* = 0.34375$. In fact, one can verify by Theorems 7, 6 and 9 that d_2 is even $\mathcal{A}_{g,\lambda_0}$-, $\mathcal{D}_{g,\lambda_0}$- and $\mathcal{T}_{g,\lambda_0}$-optimal when $-1 \le \lambda_0 \le 0.394$. At $\lambda_0 = 0.5$, d_1 is optimal under all four criteria. When we tune ρ to be $1/2$, d_2 is optimal under all four criteria for $-0.75 \le \lambda_0 \le 1$. When we tune ρ to be $-1/2$, d_2 is still $\mathcal{A}_{g,\lambda_0}$-, $\mathcal{D}_{g,\lambda_0}$- and $\mathcal{T}_{g,\lambda_0}$-optimal for small and negative values of λ_0, while the design with $p_{\langle \mathrm{re} \rangle} = 2/9$ and $p_{\langle \mathrm{di} \rangle} = 7/9$ is $\mathcal{E}_{g,\lambda_0}$-optimal. For moderate positive λ_0, designs for four criteria are all different, but they all consists of small portion of $\langle \mathrm{re} \rangle$ and large portion of $\langle \mathrm{di} \rangle$. All these designs are highly efficient for all criteria; see Table 1, for example.

Without surprise, ϕ_{g,λ_0}-optimal design for exchangeable g is not necessarily optimal when g is not exchangeable. We consider the prior distribution of $g = g_1$ which puts all its mass on the single point $\tau_0 = (0, 1, -1)'$. When $n = 36$, derive $d_{1'}$ from d_1 by replacing one sequence of 123 therein by 323, it turns out that $d_{1'}$ is 1.66% more $\mathcal{E}_{g_1,0}$-efficient when $\lambda_0 = \rho = 0$. However, in practice, one does not have accurate information of τ_0. Exchangeable prior distribution of τ_0 actually accounts for the case when nothing is known about τ_0. A further justification is that symmetry is usually a nice feature. If we search among pseudo symmetric

TABLE 1
Efficiency of d_1 and d_2 under $\mathcal{A}_{g,\lambda_0}$, $\mathcal{D}_{g,\lambda_0}$, $\mathcal{E}_{g,\lambda_0}$, and $\mathcal{T}_{g,\lambda_0}$-criteria for the case of $(p,t)=(3,3)$ when $\rho=\lambda_0=0$

Design	**A**	**D**	**E**	**T**
d_1	0.9782	0.9752	1	0.9722
d_2	1	1	0.9931	1

designs, the designs as proposed in this paper would be optimal under the corresponding criterion for any arbitrary g, which is not necessarily exchangeable. To see this, note that C_{dij}'s are all completely symmetric for a pseudo symmetric design. Hence, it is easily seen, by examining the proof of Proposition 1, that the value of $\phi_{g,\lambda_0}(d)$ is independent of the distribution g for both d_1 and d_2 regardless the value of λ_0 as well as the criterion function Φ.

Case of $(p,t)=(3,4)$: Let d_3 be a design with $p_{\langle\mathrm{re}\rangle}=1/8$ and $p_{\langle\mathrm{di}\rangle}=7/8$. Define d_4 to be a design with $p_{\langle\mathrm{di}\rangle}=1$, which requires n to be a multiple of 12 as an exact design. When $\rho=0$, Theorem 8 shows the $\mathcal{E}_{g,\lambda_0}$-optimality of d_3 for any λ_0 and Bailey and Kunert (2006) shows the $\mathcal{A}_{g,\lambda_0}$-optimality of d_4 when $-1\le\lambda_0\le\lambda^*=0.4455$. In fact, one can verify by Theorems 6, 7 and 9 that d_4 is even $\mathcal{A}_{g,\lambda_0}$-, $\mathcal{D}_{g,\lambda_0}$- and $\mathcal{T}_{g,\lambda_0}$-optimal when $-1\le\lambda_0\le0.463$. Similarly at $\lambda_0=0.5$, d_3 is optimal under all four criteria. When we tune ρ to be $1/2$, d_4 is optimal under all four criteria for $-0.35\le\lambda_0\le1$. It is still $\mathcal{A}_{g,\lambda_0}$-, $\mathcal{D}_{g,\lambda_0}$- and $\mathcal{E}_{g,\lambda_0}$-optimal and highly $\mathcal{T}_{g,\lambda_0}$-efficient for $-1\le\lambda_0<-0.35$. When $\rho=-0.5$, similar phonomania as for the case of $(p,t)=(3,4)$ is observed.

Case of $(p,t)=(3,5)$: We have similar observations as for case of $(p,t)=(3,4)$, except that the portion of $\langle\mathrm{re}\rangle$ becomes further smaller. This trend projects to larger values of t.

Case of $(p,t)=(4,3)$: When $\rho=0$, the design with $p_{\langle\mathrm{re}\rangle}=1$ is optimal under all four criteria for $0\le\lambda_0\le1$. For negative λ_0, designs are different for different criteria. However, they typically consist of symmetric blocks of $\langle1232\rangle$ and $\langle\mathrm{re}\rangle$. Table 2 shows the performance of these designs for $\lambda_0=-0.5$. Designs therein are identified by $p_{\langle\mathrm{re}\rangle}=1-p_{\langle1232\rangle}$. When ρ is nonzero, symmetric blocks of $\langle1123\rangle$, $\langle1231\rangle$,$\langle1232\rangle$ and $\langle\mathrm{re}\rangle$ will appear in different optimal designs. Note that Bailey and Kunert's (2006) result does not apply to this case since they deal with $3\le p\le t$.

Case of $(p,t)=(4,4)$: When $\rho=0$, the design with $p_{\langle\mathrm{re}\rangle}=1/12$ and $p_{\langle\mathrm{di}\rangle}=11/12$ is $\mathcal{E}_{g,\lambda_0}$-optimal for all λ_0, while the design with $p_{\langle\mathrm{di}\rangle}=1$ is $\mathcal{A}_{g,\lambda_0}$-, $\mathcal{D}_{g,\lambda_0}$- and $\mathcal{T}_{g,\lambda_0}$-optimal for any λ_0 between -1 and 0.318 ($>\lambda^*$). Interestingly, the design with $p_{\langle\mathrm{re}\rangle}=1$ is $\mathcal{A}_{g,\lambda_0}$-, $\mathcal{D}_{g,\lambda_0}$- and $\mathcal{T}_{g,\lambda_0}$-optimal for $0.625\le\lambda_0\le1$. For $0.318<\lambda<0.625$, the optimal designs consist of $\langle\mathrm{re}\rangle$ and $\langle\mathrm{di}\rangle$ with the proportion depending on different criteria. When $\rho=0.5$, the design with $p_{\langle\mathrm{di}\rangle}=1$ is optimal under all the four criteria for $0.368\le\lambda_0\le1$ and $\mathcal{E}_{g,\lambda_0}$-optimal for all λ_0. When

TABLE 2
Efficiency of designs under $\mathcal{A}_{g,\lambda_0}$, $\mathcal{D}_{g,\lambda_0}$, $\mathcal{E}_{g,\lambda_0}$, and $\mathcal{T}_{g,\lambda_0}$-criteria for the case of $(p,t)=(4,3)$ when $\rho=0$ and $\lambda_0=-0.5$

$p_{\langle 1232\rangle}$	**A**	**D**	**E**	**T**
0.4729	1	0.9964	0.9553	0.9768
0.6330	0.9952	1	0.9199	0.9888
0	0.9636	0.9475	1	0.9167
1	0.9422	0.9785	0.8000	1

$\rho=-0.5$, the design with $p_{\langle 1123\rangle}=1$ (resp., $p_{\langle \mathrm{re}\rangle}=1$) is $\mathcal{E}_{g,\lambda_0}$-optimal for all λ_0 and also optimal under the other three criteria for λ_0 close to zero (reps. 0.3). For moderate negative value of λ_0, the design with $p_{\langle \mathrm{di}\rangle}=1$ is optimal under these three criteria.

Case of $(p,t)=(4,5)$: Similar observation as the case of $(p,t)=(4,4)$ except that the symmetric metric $\langle 1122\rangle$ appears as small proportion in optimal designs when $\rho=-0.5$ and λ_0 takes a positive moderate value.

Case of $(p,t)=(5,3)$: When $\rho=0$, the design with $p_{\langle 12233\rangle}=2/5$ and $p_{\langle 12332\rangle}=3/5$ is $\mathcal{E}_{g,\lambda_0}$-optimal for all λ_0 and also optimal under the other three criteria when λ_0 is in a neighborhood of zero. For other values of ρ and λ_0, there is no specific symmetric block which will dominate, but we observe that all sequences in the optimal designs contain all three treatments.

Case of $(p,t)=(6,2)$: It is well known that $t=2$ indicates the equivalence of all optimality criteria for the classical model. For proportional model, this is also true. To see this, Proposition 1 shows that the information matrix $C_{d,\tau_0,\lambda_0}(\tau)$ only has one positive eigenvalue $(t-1)^{-1}(c_{d11}-c_{d12}^2/c_{d22})$ with multiplicity 1. Hence, the optimal design will be irrelevant of optimality criteria as well as the value of λ_0. When $\rho=0$, the design with $p_{\langle 111222\rangle}=5/8$ and $p_{\langle 121212\rangle}=3/8$ is optimal under all the four criteria for any λ_0. An exact design with 16 runs is given as

$$
\begin{array}{cccccccccccccccc}
1&1&1&1&1&2&2&2&2&2&1&1&1&2&2&2\\
1&1&1&1&1&2&2&2&2&2&2&2&2&1&1&1\\
1&1&1&1&1&2&2&2&2&2&1&1&1&2&2&2\\
2&2&2&2&2&1&1&1&1&1&2&2&2&1&1&1\\
2&2&2&2&2&1&1&1&1&1&1&1&1&2&2&2\\
2&2&2&2&2&1&1&1&1&1&2&2&2&1&1&1
\end{array}
$$

When $\rho=0.5$, the design with $p_{\langle 122121\rangle}=1$ is optimal. When $\rho=-0.5$, the design with $p_{\langle 111222\rangle}=2/11$ and $p_{\langle 122211\rangle}=9/11$ is optimal. Based on Corollary 6, these designs are also universally optimal for the classical model.

COROLLARY 6. *For any Σ, a pseudo symmetric design which is $\mathcal{E}_{g,\lambda_0}$-optimal for the proportional model is also universally optimal for the classical model.*

Acknowledgments. We are grateful to the Associate Editor and two referees for their constructive comments on earlier versions of this manuscript.

SUPPLEMENTARY MATERIAL

Appendix for optimal crossover designs for the proportional model (DOI: 10.1214/13-AOS1148SUPP; .pdf). This document is to provide a general algorithm to derive optimal $P_{\langle d \rangle}$ for arbitrary values of λ_0 and Σ based on the equivalence theorems.

REFERENCES

BAILEY, R. A. and KUNERT, J. (2006). On optimal crossover designs when carryover effects are proportional to direct effects. *Biometrika* **93** 613–625. MR2261446

BOSE, M. and STUFKEN, J. (2007). Optimal crossover designs when carryover effects are proportional to direct effects. *J. Statist. Plann. Inference* **137** 3291–3302. MR2363257

CHÊNG, C. S. and WU, C.-F. (1980). Balanced repeated measurements designs. *Ann. Statist.* **8** 1272–1283. MR0594644

FEDOROV, V. V. and HACKL, P. (1997). *Model-oriented Design of Experiments. Lecture Notes in Statistics* **125**. Springer, New York. MR1454123

HEDAYAT, A. S. and AFARINEJAD, K. (1975). Repeated measurements designs. In *A Survey of Statistical Design and Linear Models* (N. Srivastava, ed.) 229–240. North-Holland, Amsterdam.

HEDAYAT, A. and AFSARINEJAD, K. (1978). Repeated measurements designs. II. *Ann. Statist.* **6** 619–628. MR0488527

HEDAYAT, A. S. and YANG, M. (2003). Universal optimality of balanced uniform crossover designs. *Ann. Statist.* **31** 978–983. MR1994737

HEDAYAT, A. S. and YANG, M. (2004). Universal optimality for selected crossover designs. *J. Amer. Statist. Assoc.* **99** 461–466. MR2062831

HEDAYAT, A. S. and ZHENG, W. (2010). Optimal and efficient crossover designs for test-control study when subject effects are random. *J. Amer. Statist. Assoc.* **105** 1581–1592. MR2796573

KEMPTON, R. A., FERRIS, S. J. and DAVID, O. (2001). Optimal change-over designs when carryover effects are proportional to direct effects of treatments. *Biometrika* **88** 391–399. MR1844839

KIEFER, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (*Proc. Internat. Sympos., Colorado State Univ., Ft. Collins, Colo.*, 1973) (J. N. Srivastava, ed.) 333–353. North-Holland, Amsterdam. MR0395079

KUNERT, J. (1984). Optimality of balanced uniform repeated measurements designs. *Ann. Statist.* **12** 1006–1017. MR0751288

KUNERT, J. and MARTIN, R. J. (2000). On the determination of optimal designs for an interference model. *Ann. Statist.* **28** 1728–1742. MR1835039

KUNERT, J. and STUFKEN, J. (2002). Optimal crossover designs in a model with self and mixed carryover effects. *J. Amer. Statist. Assoc.* **97** 898–906. MR1941418

KUSHNER, H. B. (1997). Optimal repeated measurements designs: The linear optimality equations. *Ann. Statist.* **25** 2328–2344. MR1604457

KUSHNER, H. B. (1998). Optimal and efficient repeated-measurements designs for uncorrelated observations. *J. Amer. Statist. Assoc.* **93** 1176–1187. MR1649211

PARK, D. K., BOSE, M., NOTZ, W. I. and DEAN, A. M. (2011). Efficient crossover designs in the presence of interactions between direct and carry-over treatment effects. *J. Statist. Plann. Inference* **141** 846–860. MR2732954

STUFKEN, J. (1991). Some families of optimal and efficient repeated measurements designs. *J. Statist. Plann. Inference* **27** 75–83. MR1089354

ZHENG, W. (2013a). Universally optimal crossover designs under subject dropout. *Ann. Statist.* **41** 63–90.

ZHENG, W. (2013b). Supplement to "Optimal crossover designs for the proportional model." DOI:10.1214/13-AOS1148SUPP.

DEPARTMENT OF MATHEMATICAL SCIENCES
INDIANA UNIVERSITY-PURDUE UNIVERSITY INDIANAPOLIS
INDIANAPOLIS, INDIANA 46202-3216
USA
E-MAIL: weizheng@iupui.edu