*The Annals of Statistics*
2015, Vol. 43, No. 2, 501–518
DOI: 10.1214/14-AOS1287


# UNIVERSALLY OPTIMAL DESIGNS FOR TWO INTERFERENCE MODELS

BY WEI ZHENG

*Indiana University–Purdue University Indianapolis*

A systematic study is carried out regarding universally optimal designs under the interference model, previously investigated by Kunert and Martin [*Ann. Statist.* **28** (2000) 1728–1742] and Kunert and Mersmann [*J. Statist. Plann. Inference* **141** (2011) 1623–1632]. Parallel results are also provided for the undirectional interference model, where the left and right neighbor effects are equal. It is further shown that the efficiency of any design under the latter model is at least its efficiency under the former model. Designs universally optimal for both models are also identified. Most importantly, this paper provides Kushner's type linear equations system as a necessary and sufficient condition for a design to be universally optimal. This result is novel for models with at least two sets of treatment-related nuisance parameters, which are left and right neighbor effects here. It sheds light on other models in deriving asymmetric optimal or efficient designs.

**1. Introduction.** One issue with the application of block designs in agricultural field trials is that a treatment assigned to a particular plot typically has effects on the neighboring plots besides the effect on its own plot. See Rees (1967), Draper and Guttman (1980), Kempton (1982), Besag and Kempton (1986), Langton (1990), Gill (1993), Goldringer, Brabant and Kempton (1994), Clarke, Baker and DePauw (2000), David et al. (2001) and Connolly et al. (2008) for examples in various backgrounds. Interference models have been suggested for the analysis of data in order to avoid systematic bias caused by these neighbor effects. Various designs have been proposed by Gill (1993), Druilhet (1999), Filipiak and Markiewicz (2003, 2005, 2007), Bailey and Druilhet (2004), Ai, Ge and Chan (2007), Ai, Yu and He (2009), Druilhet and Tinssonb (2012) and Filipiak (2012) among others. All of them considered circular designs, where each block has a guard plot at each end so that each plot within the block has two neighbors.

To study noncircular designs, Kunert and Martin (2000) investigated the case when the block size, say $k$, is 3 or 4, which is extended by Kunert and Mersmann (2011) to $t \geq k \geq 5$, where $t$ is the number of treatments. Both of them restricted to the subclass of pseudo symmetric designs and the assumption that the within-block covariance matrix is proportional to the identity matrix. This paper

Received May 2014; revised September 2014.
*MSC2010 subject classifications.* Primary 62K05; secondary 62J05.
*Key words and phrases.* Approximate design theory, interference model, linear equations system, pseudo symmetric designs, universally optimal designs.

provides a unified framework for deriving optimal pseudo symmetric designs for an arbitrary covariance matrix as well as the general setup of $k \geq 3$ and $t \geq 2$. Most importantly, the Kushner's type linear equations system is developed as a necessary and sufficient condition for any design to be universally optimal, which is a powerful device for deriving asymmetric designs. Moreover, a new approach of finding the optimal sequences are proposed. These results are novel for models with at least two sets of treatment-related nuisance parameters, which are left and right neighbor effects here. They shed light on other similar or more complicated models such as the one in Afsarinejad and Hedayat (2002) and Kunert and Stufken (2002) for the study of crossover designs. Here, parallel results are also provided for the undirectional interference model where the left and right neighbor effects are equal. It is further established that the efficiency of any given design under the latter model is not less than the one under the former model, for the purpose of estimating the direct treatment effects.

Throughout the paper, we consider designs in $\Omega_{n,k,t}$, the set of all possible block designs with $n$ blocks of size $k$ and $t$ treatments. The response, denoted as $y_{dij}$, observed from the $j$th plot of block $i$ is modeled as

$$Y_{dij} = \mu + \beta_i + \tau_{d(i,j)} + \lambda_{d(i,j-1)} + \rho_{d(i,j+1)} + \varepsilon_{ij}, \tag{1}$$

where $\mathbb{E}\varepsilon_{ij} = 0$. The subscript $d(i, j)$ denotes the treatment assigned in the $j$th plot of block $i$ by the design $d : \{1, 2, \dots, n\} \times \{1, 2, \dots, k\} \to \{1, 2, \dots, t\}$. Furthermore, $\mu$ is the general mean, $\beta_i$ is the $i$th block effect, $\tau_{d(i,j)}$ is the direct treatment effect of treatment $d(i, j)$, $\lambda_{d(i,j-1)}$ is the neighbor effect of treatment $d(i, j-1)$ from the left neighbor, and $\rho_{d(i,j+1)}$ is the neighbor effect of treatment $d(i, j+1)$ from the right neighbor. One major objective of design theorists is to find optimal or efficient designs for estimating the direct treatment effects in the model.

If $Y_d$ is the vector of responses organized block by block, model (1) is written in a matrix form of

$$Y_d = 1_{nk}\mu + U\beta + T_d\tau + L_d\lambda + R_d\rho + \varepsilon, \tag{2}$$

where $\beta = (\beta_1, \dots, \beta_n)'$, $\tau = (\tau_1, \dots, \tau_t)'$, $\lambda = (\lambda_1, \dots, \lambda_t)'$ and $\rho = (\rho_1, \dots, \rho_t)'$. The notation $'$ means the transpose of a vector or a matrix. Here, we have $U = I_n \otimes 1_k$ with $\otimes$ as the Kronecker product, and $1_k$ represents a vector of ones with length $k$. Also, $T_d$, $L_d$ and $R_d$ represent the design matrices for the direct, left neighbor and right neighbor effects, respectively. We assume there is no guard plots, that is, $\lambda_{d(i,0)} = \rho_{d(i,k+1)} = 0$. Then we have $L_d = (I_n \otimes H)T_d$ and $R_d = (I_n \otimes H')T_d$, where $H(i, j) = \mathbb{I}_{i=j+1}$ with the indicator function $\mathbb{I}$.

Here, we merely assume $\mathrm{Var}(\varepsilon) = I_n \otimes \Sigma$, with $\Sigma$ being an arbitrary $k \times k$ positive definite symmetric matrix. Given a matrix, say $G$, we define the projection $\mathrm{pr}^{\perp} G = I - G(G'G)^{-}G'$. The information matrix for the direct treatment effect $\tau$ is

$$C_d = T_d' V' \mathrm{pr}^{\perp}(VU|VL_d|VR_d)VT_d, \tag{3}$$

where $V$ is the matrix such that $V^2 = I_n \otimes \Sigma^{-1}$. By direct calculations, we have

$$\begin{aligned}
C_d &= E_{d00} - E_{d01} E_{d11}^- E_{d10},\\
E_{d00} &= C_{d00},\\
E'_{d10} &= E_{d01} = \begin{pmatrix} C_{d01} & C_{d02} \end{pmatrix},\\
E_{d11} &= \begin{pmatrix} C_{d11} & C_{d12} \\ C_{d21} & C_{d22} \end{pmatrix},
\end{aligned}$$

where $C_{dij} = G'_i(I_n \otimes \tilde{B})G_j$, $0 \le i, j \le 2$ with $G_0 = T_d$, $G_1 = L_d$, $G_2 = R_d$ and $\tilde{B} = \Sigma^{-1} - \Sigma^{-1} J_k \Sigma^{-1} / 1'_k \Sigma^{-1} 1_k$ with $J_k = 1_k 1'_k$. It is obvious that $C_{dij} = C'_{dji}$. For the special case of $\Sigma = I_k$, we have the simplification of $\tilde{B} = I_k - k^{-1} J_k = \mathrm{pr}^{\perp}(1_k)$, and the latter is denoted by $B_k$. Kushner (1997) pointed out that when $\Sigma$ is of type-$H$, that is, $aI_k + b1'_k + 1_k b'$ with $a \in \mathbb{R}$ and $b \in \mathbb{R}^k$, we have

$$\tilde{B} = B_k / a. \tag{4}$$

Hence, the choices of designs agree with that for $\Sigma = I_k$. This special case will be particularly dealt with in Section 5. We allow $\Sigma$ to be an arbitrary covariance matrix throughout the rest of the paper.

Note that a design in $\Omega_{n,k,t}$ could be considered as a result of selecting $n$ elements from the set, $\mathcal{S}$, of all possible $t^k$ block sequences with replacement. For sequence $s \in \mathcal{S}$, define the sequence proportion $p_s = n_s/n$, where $n_s$ is the number of replications of $s$ in the design. A design is determined by $n_s$, $s \in \mathcal{S}$, which is in turn determined by the *measure* $\xi = (p_s, s \in \mathcal{S})$ for any fixed $n$.

For $0 \le i, j \le 2$, define $C_{sij}$ to be $C_{dij}$ when the design consists of the single sequence $s$, and let $C_{\xi ij} = \sum_{s \in \mathcal{S}} p_s C_{sij}$. Then we have $C_{dij} = nC_{\xi ij}$, $0 \le i, j \le 2$. Similarly, $E_{dij} = n \sum_{s \in \mathcal{S}} p_s E_{sij} = nE_{\xi ij}$, $0 \le i, j \le 1$. Note that $C_d$ is a Schur's complement of $A_d = (E_{dij})_{0 \le i,j \le 1}$, for which we also have $A_d = n \sum_{s \in \mathcal{S}} p_s A_s = nA_\xi$. It is obvious that $C_d = nC_\xi$, where $C_\xi = E_{\xi 00} - E_{\xi 01} E_{\xi 11}^- E_{\xi 10}$. In approximate design theory, we try to find the optimal measure $\xi$ among the set $\mathcal{P} = \{(p_s, s \in \mathcal{S}) | \sum_{s \in \mathcal{S}} p_s = 1, p_s \ge 0\}$ to maximize $\Phi(C_\xi)$ for a given function $\Phi$ satisfying the following three conditions [Kiefer (1975)]:

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

A measure $\xi$ which achieves the maximum of $\Phi(C_\xi)$ among $\mathcal{P}$ for any $\Phi$ satisfying (C.1)–(C.3) is said to be *universally optimal*. Such measure is optimal under criteria of $A$, $D$, $E$, $T$, etc.

The rest of the paper is organized as follows. Section 2 provides some preliminary results as well as a necessary and sufficient condition for a pseudo symmetric measure to be universally optimal among $\mathcal{P}$. The latter is critical for deriving the optimal sequence proportions through an algorithm. Section 3 provides a linear equations system of $p_s, s \in \mathcal{S}$, as a necessary and sufficient condition for a measure to be universally optimal. Section 4 provides similar results for the model with $\lambda = \rho$. Further, it is shown that the efficiency of any design under the latter model would be at least its efficiency under model (2). Also, an alternative approach is given to derive the optimal sequence proportions. Section 5 derives theoretical results regarding feasible sequences when $\Sigma$ is of type-$H$. Section 6 provides some examples of optimal or efficient designs for various combinations of $k, t, n$ and $\Sigma$.

**2. Pseudo symmetric measure.** Let $\mathcal{G}$ be the set of all $t!$ permutations on symbols $\{1, 2, \dots, t\}$. For permutation $\sigma \in \mathcal{G}$ and sequence $s = (t_1 \cdots t_k)$ with $1 \le t_i \le t$ and $1 \le i \le k$, we define $\sigma s = (\sigma(t_1) \cdots \sigma(t_k))$. For measure $\xi = (p_s, s \in \mathcal{S})$, we define $\sigma\xi = (p_{\sigma^{-1}s}, s \in \mathcal{S})$. A measure is said to be *symmetric* if $\sigma\xi = \xi$ for all $\sigma \in \mathcal{G}$. For sequence $s$, denote by $\langle s \rangle = \{\sigma s : \sigma \in \mathcal{G}\}$ the *symmetric block* generated by $s$. Such symmetric blocks are also called equivalent classes by Kushner (1997), due to the fact that symmetric blocks generated by two different sequences are either identical or mutually disjoint. Now let $m$ be the total number of distinct symmetric blocks which partition $\mathcal{S}$. Without loss of generality, suppose these $m$ symmetric blocks are generated by sequences $s_i$, $1 \le i \le m$. Then we have $\mathcal{S} = \bigcup_{i=1}^m \langle s_i \rangle$. For a symmetric measure, we have

$$p_s = p_{\langle s_i \rangle} / \left|\langle s_i \rangle\right| \qquad \text{for } s \in \langle s_i \rangle, 1 \le i \le m, \tag{5}$$

where $p_{\langle s_i \rangle} = \sum_{s \in \langle s_i \rangle} p_s$ and $|\langle s_i \rangle|$ is the cardinality of $\langle s_i \rangle$. The linearity of $A_d$, conditions (C.1)–(C.3) and properties of Schur's complement together yield the following lemma.

LEMMA 1. *For any measure, say $\xi$, there exists a symmetric measure, say $\xi^*$, such that $\Phi(C_\xi) \le \Phi(C_{\xi^*})$ for any $\Phi$ satisfying* (C.1)–(C.3).

Define a measure to be *pseudo symmetric* if $C_{\xi ij}$, $0 \le i, j \le 2$ are all completely symmetric. It is easy to verify that a symmetric measure is also pseudo symmetric. The difference is that (5) does not has to hold for a general pseudo symmetric measure. Lemma 1 indicates that an optimal measure in the subclass of (pseudo) symmetric measures is automatically optimal among $\mathcal{P}$. For a pseudo symmetric measure, we have $C_{\xi ij} = c_{\xi ij} B_t/(t-1) + (1_t' C_{\xi ij} 1_t) J_t/t^2$, $0 \le i, j \le 2$, where $c_{\xi ij} = \operatorname{tr}(B_t C_{\xi ij} B_t)$. Hence $E_{\xi 11} = Q_\xi \otimes B_t/(t-1) + \tilde{Q}_\xi \otimes J_t/t^2$, where $Q_\xi = (c_{\xi ij})_{1 \le i, j \le 2}$ and $\tilde{Q}_\xi = (1_t' C_{\xi ij} 1_t)_{1 \le i, j \le 2}$. Now we show that both $Q_\xi$ and $\tilde{Q}_\xi$ are positive definite for any measure, and hence $E_{\xi 11}$ is positive definite for any pseudo symmetric measure. The latter is the key to prove Theorem 3.

LEMMA 2. *$Q_\xi$ is positive definite for any measure $\xi$.*

PROOF. It is sufficient to show the nonsingularity of $Q_s$ for all $s \in \mathcal{S}$. Suppose $Q_s$ is singular, there exists a nonzero vector $x = (x_1, x_2)'$ such that

$$\begin{aligned} 0 = x' Q_s x &= \sum_{i=1}^{2}\sum_{j=1}^{2} x_i x_j c_{sij} \\ &= \operatorname{tr}\left(\sum_{i=1}^{2}\sum_{j=1}^{2} x_i x_j B_t C_{sij} B_t\right). \end{aligned}$$

Since $\sum_{i=1}^{2}\sum_{j=1}^{2} x_i x_j B_t C_{sij} B_t$ is a nonnegative definite matrix, we have

$$\begin{aligned} 0 &= \sum_{i=1}^{2}\sum_{j=1}^{2} x_i x_j B_t C_{sij} B_t \\ &= B_t (x_1 L_s + x_2 R_s)' \tilde{B} (x_1 L_s + x_2 R_s) B_t, \end{aligned}$$

which in turn yields

$$0 = \tilde{B}(x_1 L_s + x_2 R_s) B_t. \tag{6}$$

Equation (6) is only possible when each column of $M = (x_1 L_s + x_2 R_s) B_t$ consists of identical entries, that is, the rows of $M$ are identical. In the sequel, we investigate the possibility of (6) for sequence $s = (t_1 \cdots t_k)$. Define $e_i$ to be a zero–one vector of length $t$ with only its $i$th entry as one, then the first, second and last rows of $M$ are given by $x_2(e_{t_2} - 1_t/t)'$, $x_1(e_{t_1} - 1_t/t)' + x_2(e_{t_3} - 1_t/t)'$ and $x_1(e_{t_{k-1}} - 1_t/t)'$, respectively. Now we continue the discussion in the following four cases. (i) If $x_1 = x_2$, the equality of the first two rows of $M$ indicates $e_{t_1} + e_{t_3} - e_{t_2} = 1_t/t$, which is impossible since the left-hand side is a vector of integers and the right-hand side is a vector of fractional numbers. (ii) If $x_1 \neq x_2$ and $t_2 = t_{k-1}$, the first and the last rows of $M$ cannot be the same. (iii) If $x_1 \neq x_2$, $t_2 \neq t_{k-1}$ and $t = 2$, the equality of the first and the last rows of $M$ necessities $x_1 + x_2 = 0$, which together with the equality of the first two rows of $M$ indicates $e_{t_1} + e_{t_2} - e_{t_3} = 1_t/t$, which is again impossible. (iv) If $x_1 \neq x_2$, $t_2 \neq t_{k-1}$ and $t \geq 3$, by looking at the $t_2$th and $t_{k-1}$th entries of the first and last rows of $M$, (6) necessities $x_2(1 - 1/t) = -x_1/t$ and $x_1(1 - 1/t) = -x_2/t$ which is impossible by simple algebra. □

LEMMA 3. *$\tilde{Q}_\xi$ is positive definite for any measure $\xi$.*

PROOF. Since $\tilde{B}$ has column and row sums as zero. We have

$$\tilde{Q}_\xi = \begin{pmatrix} \tilde{B}(1,1) & \tilde{B}(1,k) \\ \tilde{B}(k,1) & \tilde{B}(k,k) \end{pmatrix}, \tag{7}$$

where $\tilde{B}(i,j)$ means the $(i,j)$th entry in $\tilde{B}$. For vector $x = (x_1, x_2)' \in \mathbb{R}^2$, define $w = (x_1, 0, \ldots, 0, x_2)' \in \mathbb{R}^k$. For any nonzero $x$, we have

$$x' \tilde{Q}_\xi x = w' \tilde{B} w > 0, \tag{8}$$

in view of the fact that $\tilde{B}1_k = 0$, $\tilde{B} \geq 0$ and the rank of $\tilde{B}$ is $k-1$. Hence, the lemma is concluded. $\square$

LEMMA 4. *For a pseudo symmetric measure, say $\xi$, we have $C_\xi = q_\xi^* B_t/(t-1)$, where*

$$q_\xi^* = c_{\xi 00} - \ell_\xi' Q_\xi^{-1} \ell_\xi, \tag{9}$$

*with $\ell_\xi = (c_{\xi 01}, c_{\xi 02})'$.*

REMARK 1. In proving Lemma 4, we used the equations $1_t' C_{\xi 0j} = 0$, $0 \leq j \leq 2$. Note that $nq_\xi^*$ is the $q_d^*$ as defined in Kunert and Martin (2000). Lemma 2 shows that only case (i) of the four cases proposed by them is possible. Hence the generalized inverse $Q_\xi^-$ in Kunert and Martin (2000) is now replaced by $Q_\xi^{-1}$ in (9).

By applying Lemmas 1 and 4, we derive the following proposition.

PROPOSITION 1. *Let $y^* = \max_{\xi \in \mathcal{P}} q_\xi^*$. A measure $\xi \in \mathcal{P}$ is universally optimal* (i) *if it is a pseudo symmetric measure with $q_\xi^* = y^*$*, (ii) *if and only if $C_\xi = y^* B_t/(t-1)$.*

Let $R_s = (c_{sij})_{0 \leq i,j \leq 2}$ and $R_\xi = \sum_{s \in \mathcal{S}} p_s R_s$. By Lemma 2 we have $q_\xi^* = \det(R_\xi)/\det(Q_\xi)$, where $\det(\cdot)$ means the determinant of a square matrix. For measure $\xi = (p_s, s \in \mathcal{S})$, we call the set $\mathcal{V}_\xi = \{s : p_s > 0, s \in \mathcal{S}\}$ the *support* of $\xi$. One can identify universally optimal pseudo symmetric measures based on the following theorem. See Zheng (2013b) for an algorithm based on a similar theorem.

THEOREM 1. *A pseudo symmetric measure, say $\xi$, is universally optimal if and only if* $\det(R_\xi) > 0$ *and*

$$\max_{s \in \mathcal{S}} \left[\operatorname{tr}\left(R_s R_\xi^{-1}\right) - \operatorname{tr}\left(Q_s Q_\xi^{-1}\right)\right] = 1. \tag{10}$$

*Moreover, each sequence in $\mathcal{V}_\xi$ reaches the maximum in* (10).

PROOF. If $\det(R_\xi) = 0$, we have $q_\xi^* = 0$, which means that such design has no information regarding $\tau$, and hence can be readily excluded from the consideration. In the sequel, we restrict the discussion to the case of $\det(R_\xi) > 0$.

By Lemmas 1, 2 and 4, a pseudo symmetric measure, say $\xi$, is universally optimal if and only if it achieves the maximum of $\varphi(\xi) = \log(\det(R_\xi)/\det(Q_\xi))$, which is equivalent to

$$\lim_{\delta \to 0} \frac{\varphi[(1-\delta)\xi + \delta\xi_0] - \varphi(\xi)}{\delta} \leq 0, \tag{11}$$

for any measure $\xi_0 \in \mathcal{P}$. It is well known that

$$\lim_{\delta\to 0} \frac{\log(\det(R_{(1-\delta)\xi+\delta\xi_0})) - \log(\det(R_\xi))}{\delta} = \operatorname{tr}(R_{\xi_0} R_\xi^{-1}) - 3. \tag{12}$$

The same result holds for $Q(\xi)$ except that 3 should be replaced by 2. By applying (12) to (11), we have

$$\operatorname{tr}(R_{\xi_0} R_\xi^{-1}) - \operatorname{tr}(Q_{\xi_0} Q_\xi^{-1}) \le 1. \tag{13}$$

In (13), by setting $\xi_0$ to be a degenerated measure which puts all its mass on a single sequence, we derive

$$\max_{s\in\mathcal{S}} (\operatorname{tr}(R_s R_\xi^{-1}) - \operatorname{tr}(Q_s Q_\xi^{-1})) \le 1.$$

By taking $\xi_0 = \xi$, we have the equal sign for (13). Also observe that conditioning on fixed $\xi$, the left-hand side of (13) is a linear function of the proportions in $\xi_0$. Thus, we have

$$\max_{s\in\mathcal{S}} (\operatorname{tr}(R_s R_\xi^{-1}) - \operatorname{tr}(Q_s Q_\xi^{-1})) \ge 1.$$

Hence, the theorem follows. □

**3. The linear equations system: A necessary and sufficient condition for universal optimality.** For sequence $s$ and vector $x \in \mathbb{R}^2$, define the quadratic function $q_s(x) = c_{s00} + 2\ell_s' x + x' Q_s x$. For measure $\xi = (p_s, s \in \mathcal{S})$, define $q_\xi(x) = \sum_{s\in\mathcal{S}} p_s q_s(x) = c_{\xi 00} + 2\ell_\xi' x + x' Q_\xi x$. One can verify that $q_\xi^* = \min_{x\in\mathbb{R}^2} q_\xi(x)$. Since $q_s(x)$ is strictly convex for all $s \in \mathcal{S}$ in view of Lemma 2, thus $r(x) := \max_{s\in\mathcal{S}} q_s(x)$ is also strictly convex. Let $x^*$ be the unique point in $\mathbb{R}^2$ which achieves minimum of $r(x)$ and define $\mathcal{T} = \{s : q_s(x^*) = r(x^*), s \in \mathcal{S}\}$. Recall $y^* = \max_{\xi\in\mathcal{P}} q_\xi^*$ and $\mathcal{V}_\xi = \{s : p_s > 0, s \in \mathcal{S}\}$, now we derive Theorem 2 below which is important for proving Theorem 3 and results in Section 4.

THEOREM 2. (i) $y^* = r(x^*)$. (ii) $q_\xi^* = y^*$ *implies* $x^* = -Q_\xi^{-1}\ell_\xi$. (iii) $q_\xi^* = y^*$ *implies* $\mathcal{V}_\xi \subset \mathcal{T}$.

PROOF. First, we have

$$y^* = \max_{\xi\in\mathcal{P}} \min_{x\in\mathbb{R}^2} q_\xi(x) \le \min_{x\in\mathbb{R}^2} \max_{\xi\in\mathcal{P}} q_\xi(x) = \min_{x\in\mathbb{R}^2} \max_{s\in\mathcal{S}} q_s(x) = r(x^*).$$

Then (i) is proved if we can show $y^* \ge r(x^*)$. To see the latter, define $\mathcal{T}_0 = \{s : q_s(x^*) = r(x^*), s \in \{s_1 \cdots s_m\}\}$. (1) If $\mathcal{T}_0$ contains a single sequence, say $s_1$, let $\xi_0$ be the measure with $p_{\langle s_1\rangle} = 1$, then we have $\min_{x\in\mathbb{R}^2} q_{\xi_0}(x) = r(x^*)$. Hence, $y^* \ge r(x^*)$. (2) If $\mathcal{T}_0$ contains more than one sequences, let $\nabla q_s(x^*)$ be the gradient of $q_s(x)$ evaluated at point $x = x^*$ and define $\Xi$ to be the convex hull of $\{\nabla q_s(x^*) : s \in \mathcal{T}_0\}$. We claim $0 \in \Xi$, since otherwise we could find a vector $z \in \mathbb{R}^2$ so that $z'\nabla q_s(x^*) < 0$ for all $s \in \{\nabla q_s(x^*) : s \in \mathcal{T}_0\}$, which would indicate that $x^*$

is not the minimum point of $r(x)$, and hence the contradiction is reached. Note that $0 \in \Xi$ indicates there exists a measure, say $\xi_0$, such that $q_{\xi_0}(x^*) = r(x^*)$ and $\nabla q_{\xi_0}(x^*) = 0$, which yields $\min_{x \in \mathbb{R}^2} q_{\xi_0}(x) = r(x^*)$ and hence $y^* \geq r(x^*)$. (i) is thus proved.

Observe that the minimum of $q_\xi(x)$ is achieved at the unique point $x = -Q_\xi^{-1}\ell_\xi := \tilde{x}$. If $\tilde{x} \neq x^*$, we have $y^* = r(x^*) \geq q_\xi(x^*) > q_\xi(\tilde{x}) = q_\xi^*$ and hence the contradiction is reached. (ii) is thus concluded.

For (iii), if there is a sequence, say $s$, with $s \in \mathcal{V}_\xi$ and $s \notin \mathcal{T}$, we have $y^* > q_\xi(x^*) \geq q_\xi^*$, and hence the contradiction is reached. $\square$

THEOREM 3. *A measure $\xi = (p_s, s \in \mathcal{S})$ is universally optimal among $\mathcal{P}$ if and only if*

$$\sum_{s \in \mathcal{T}} p_s\big[E_{s00} + E_{s01}\big(x^* \otimes B_t\big)\big] = y^* B_t/(t-1), \tag{14}$$

$$\sum_{s \in \mathcal{T}} p_s\big[E_{s10} + E_{s11}\big(x^* \otimes B_t\big)\big] = 0, \tag{15}$$

$$\sum_{s \notin \mathcal{T}} p_s = 0. \tag{16}$$

PROOF. Note that (14)–(16) is equivalent to

$$E_{\xi 00} + E_{\xi 01}\big(x^* \otimes B_t\big) = y^* B_t/(t-1), \tag{17}$$

$$E_{\xi 10} + E_{\xi 11}\big(x^* \otimes B_t\big) = 0, \tag{18}$$

$$\sum_{s \in \mathcal{T}} p_s = 1. \tag{19}$$

Necessity. By Proposition 1, there exists a symmetric measure, say $\xi_1$, which is universally optimal. Further, we have $C_\xi = C_{\xi_1} = y^* B_t/(t-1)$. Define $\xi_2 = (\xi + \xi_1)/2$. Then we have $A_{\xi_2} = (A_\xi + A_{\xi_1})/2$, which indicates $C_{\xi_2} \geq (C_\xi + C_{\xi_1})/2 = y^* B_t/(t-1)$. The latter combined with Proposition 1 yields $C_{\xi_2} = y^* B_t/(t-1)$. Hence, by similar arguments as in Kushner (1997), we have

$$E_{\xi 11}\big(E^+_{\xi 11} E_{\xi 10} - E^+_{\xi_2 11} E_{\xi_2 10}\big) = 0, \tag{20}$$

$$E_{\xi_1 11}\big(E^+_{\xi_1 11} E_{\xi_1 10} - E^+_{\xi_2 11} E_{\xi_2 10}\big) = 0, \tag{21}$$

where $^+$ means the Moore–Penrose generalized inverse. Since $\xi_1$ is a symmetric measure, we have $E_{\xi_1 11} = Q_{\xi_1} \otimes B_t/(t-1) + \tilde{Q}_{\xi_1} \otimes J_t/t^2$. By Lemmas 2, 3 and the orthogonality between $B_t$ and $J_t$, we obtain $\det(E_{\xi_1 11}) = \det(Q_\xi)^{t-1}\det(\tilde{Q}_\xi)/[(t-1)^{2t-2}t^3] > 0$. Applying the latter to (21) yields

$$\begin{aligned} E^+_{\xi_2 11} E_{\xi_2 10} &= E^+_{\xi_1 11} E_{\xi_1 10} \\ &= Q^{-1}_{\xi_1} \ell_{\xi_1} \otimes B_t \\ &= -x^* \otimes B_t, \end{aligned} \tag{22}$$

in view of Theorem 2(ii). Now (18) is derived from (20) and (22). By (18), we have

$$y^* B_t/(t-1) = C_\xi = E_{\xi 00} - E_{\xi 01} E_{\xi 11}^- E_{\xi 10} \tag{23}$$

$$= E_{\xi 00} + E_{\xi 01} E_{\xi 11}^- E_{\xi 11}(x^* \otimes B_t) \tag{24}$$

$$= E_{\xi 00} + E_{\xi 01}(x^* \otimes B_t), \tag{25}$$

which is essentially (17).

By (5.2) of Kushner (1997), we have $C_\xi \le H' A_\xi H$ for any $3t \times t$ matrix $H$. Set $H = (x_0, x_1, x_3)' \otimes B_t$ with $x_0 \equiv 1$, we have

$$C_\xi \le \sum_{i=0}^{2} \sum_{j=0}^{2} x_i x_j B_t C_{\xi ij} B_t. \tag{26}$$

By taking the trace of both sides of (26), we have

$$\begin{aligned} \operatorname{tr}(C_\xi) &\le \sum_{i=0}^{2} \sum_{j=0}^{2} x_i x_j c_{\xi ij} \\ &= q_\xi(x), \end{aligned}$$

for $x = (x_1, x_2)'$. Now set $x = -Q_\xi^{-1} \ell_\xi$, we have $\operatorname{tr}(C_\xi) \le q_\xi^* \le y^*$. Note that $\operatorname{tr}(C_\xi) = y^*$ in view of Proposition 1(ii). As a result, we have $q_\xi^* = y^*$ and thus (19) in view of Theorem 2(iii).

Sufficiency of (17)–(19) is trivial in view of (23)–(25). □

**4. Undirectional interference model.** In many occasions, it is reasonable to believe that the neighbor effects of each treatment from the left and the right should be the same, that is, $\lambda = \rho$. With this condition, model (2) reduces to

$$Y_d = 1_{nk}\mu + U\beta + T_d\tau + (L_d + R_d)\lambda + \varepsilon. \tag{27}$$

The information matrix, $\tilde{C}_d$, for $\tau$ under model (27) is given by

$$\begin{aligned} \tilde{C}_d &= C_{d00} - \tilde{C}_{d01} \tilde{C}_{d11}^- \tilde{C}_{d10}, \\ \tilde{C}_{d10}' &= \tilde{C}_{d01} = T_d'(I_n \otimes \tilde{B})(L_d + R_d), \\ \tilde{C}_{d11} &= (L_d + R_d)'(I_n \otimes \tilde{B})(L_d + R_d). \end{aligned}$$

It is obvious that $\tilde{C}_d/n$ only depends on the measure $\xi = (p_s, s \in \mathcal{S})$, and we denote such matrix by $\tilde{C}_\xi$. Let $\tilde{q}_s(z) = q_s((z, z)')$ and $\tilde{r}(z) = \max_{s \in \mathcal{S}} \tilde{q}_s(z)$ for $z \in \mathbb{R}$. Note that $\tilde{r}(z)$ is strictly convex due to the strict convexity of $r(x)$, hence there is an unique minimizer of $\tilde{r}(z)$ which is denoted by $z^*$ here. By following similar arguments as in Sections 2 and 3, one can derive the following theorem for universally optimal measures under model (27) in view of Lemma 5(ii).

THEOREM 4. *Let* $y_0 = \tilde{r}(z^*)$ *and* $\mathcal{T}_0 = \{s \in \mathcal{S} : \tilde{q}_s(z^*) = y_0\}$. *For measure* $\xi = (p_s, s \in \mathcal{S})$, *the following three sets of conditions are equivalent*. (i) $\xi$ *is universally optimal*. (ii) $\tilde{C}_\xi = y_0 B_t/(t-1)$. (iii)

$$\sum_{s \in \mathcal{T}_0} p_s\big[C_{s00} + z^* \tilde{C}_{s01} B_t\big] = y_0 B_t/(t-1), \tag{28}$$

$$\sum_{s \in \mathcal{T}_0} p_s\big[\tilde{C}_{s10} + z^* \tilde{C}_{s11} B_t\big] = 0, \tag{29}$$

$$\sum_{s \in \mathcal{T}_0} p_s = 1. \tag{30}$$

The following lemma is the key to build up the connections between the two models as given by Theorem 5.

LEMMA 5. *If* $\Sigma$ *is persymmetric*, *we have the following*. (i) $x^* = (z^*, z^*)'$. (ii) $y^* = y_0$. (iii) $\mathcal{T} = \mathcal{T}_0$.

PROOF. For sequence $s = (t_1 t_2 \cdots t_p)$, define its *dual* sequence as $s' = (t_p, t_{p-1} \cdots t_1)$. First we claim that

$$\ell_s = \Lambda_2 \ell_{s'}, \tag{31}$$

$$Q_s = \Lambda_2 Q_{s'} \Lambda_2, \tag{32}$$

where $\Delta_h = (\mathbb{I}_{i+j=h+1})_{1 \le i,j \le h}$. Then the function $r(x)$ is symmetric about the line $x_1 = x_2$, where $x = (x_1, x_2)'$. This indicates that the two components of $x^* \in \mathbb{R}^2$ are identical. From this, (i) and (ii) follows immediately. (iii) follows directly from (i) and (ii) by definitions of $\mathcal{T}$ and $\mathcal{T}_0$.

To prove (31) and (32), it is sufficient to show $L_s = \Delta_k R_{s'}$, $R_s = \Delta_k L_{s'}$ and $\Delta_k \tilde{B} \Delta_k = \tilde{B}$. The first two equations are trivial. To see the latter, note that the persymmetry (and hence the bisymmetry) of $\Sigma$ indicates the bisymmetry of $\Sigma^{-1}$ in view of Laplace's formula for calculating the matrix inverse. Hence, the sum of the $i$th column (or row) of $\Sigma^{-1}$ is equal to the sum of its $(k+1-i)$th column, which indicates the bisymmetry of $\Sigma^{-1} J_k \Sigma^{-1}$, and hence the bisymmetry of $\tilde{B}$. $\square$

REMARK 2. There is a wide range of covariance matrices which are persymmetric. Examples include the identity matrix, the completely symmetric matrix, the AR(1) type covariance matrix and the one used in Section 6. By Corollary 2.2 of Kushner (1997), Lemma 5 still holds if $\Sigma = \Sigma_0 + \gamma 1_k' + 1_k \gamma'$ with $\Sigma_0$ being persymmetric. In fact, the lemma holds as long as $\tilde{B}$ is persymmetric. When $\tilde{B}$ is not persymmetric, empirical evidence indicates that we typically have $x^* \neq (z^*, z^*)'$ and $y^* < y_0$. Even though we observe $\mathcal{T} = \mathcal{T}_0$ very often, however, the optimal proportions for sequences in the support would be different for the two models.

A measure $\xi = (p_s, s \in \mathcal{S})$ is said to be *dual* if $p_{\langle s\rangle} = p_{\langle s'\rangle}, s \in \mathcal{S}$, where $s'$ is the dual sequence of $s$ as defined in the proof of Lemma 5.

THEOREM 5. *If $\Sigma$ is persymmetric, we have the following.* (i) *For any measure, its universal optimality under model* (2) *implies its universal optimality under model* (27). (ii) *For a pseudo symmetric dual measure, its universal optimality under model* (27) *implies its universal optimality under model* (2). (iii) *Given any criterion function satisfying conditions* (C.1)–(C.3), *the efficiency of any measure under model* (27) *is at least its efficiency under model* (2).

PROOF. (i) is readily proved by the direct comparison between equations (14)–(16) and equations (28)–(30).

For a pseudo symmetric measure, say $\xi$, it is universally optimal for the two models as long as it maximizes the traces of the information matrices, that is, $\mathrm{tr}(C_\xi) = \min_{x\in\mathbb{R}^2} q_\xi(x)$ and $\mathrm{tr}(\tilde{C}_\xi) = \min_{z\in\mathbb{R}} \tilde{q}_\xi(z)$, respectively. If $\xi$ is also dual, $q_\xi(x)$ is a function symmetric about the line of $x_1 = x_2$ in view of (31) and (32). This indicates that $\min_{x\in\mathbb{R}^2} q_\xi(x) = \min_{z\in\mathbb{R}} \tilde{q}_\xi(z)$. Hence, the universal optimality under the two models will be equivalent for such measure, and thus (ii) follows.

Since the information matrices of universally optimal designs are the same for the two models in view of Proposition 1 and Theorem 4, hence (iii) is verified as long as we can show

$$C_d \le \tilde{C}_d, \tag{33}$$

for any design $d$. To see (33), note that the column space of $L_d + R_d$ is a subset of the column space of $[L_d|R_d]$, hence we have $\mathrm{pr}^\perp(VU|VL_d|VR_d) \le \mathrm{pr}^\perp(VU|V(L_d+R_d))$. Now (33) follows in view of (3) and $\tilde{C}_d = T_d'V'\mathrm{pr}^\perp(VU|V(L_d+R_d))VT_d$. $\square$

COROLLARY 1. (i) *A measure with $C_{d\xi 00}$, $\tilde{C}_{\xi 01}$ and $\tilde{C}_{\xi 11}$ being completely symmetric is universally optimal under model* (27) *if and only if*

$$\sum_{s\in\mathcal{T}} p_s \frac{\partial \tilde{q}_s(z)}{\partial z}\bigg|_{z=z^*} = 0, \tag{34}$$

$$\sum_{s\in\mathcal{T}} p_s = 1. \tag{35}$$

(ii) *When $\Sigma$ is persymmetric, a pseudo symmetric dual measure is universally optimal under model* (2) *if and only if* (34) *and* (35) *holds.*

REMARK 3. Since $\tilde{q}_s(z)$ is a univariate function, one can use the Kushner's (1997) method to find $z^*$ and $\mathcal{T}$ with the computational complexity of $O(m^2)$, where $m$ is the total number of symmetric blocks. If we have to deal with multivariate functions such as $q_s(x)$ (e.g., when $\Sigma$ is not persymmetric and the side effects are directional), the computation of $x^*$ and $\mathcal{T}$ is more involved but manageable. See Bailey and Druilhet (2014) for an example where $x$ is 5-dimensional. Alternatively, one can build an efficient algorithm (see the Appendix) based on (10) to derive the optimal measure, which further induces $x^*$ and $\mathcal{T}$.

**5. The set $\mathcal{T}$ for type-$H$ covariance matrix.** By restricting to the type-$H$ covariance matrix $\Sigma$, we derive theoretical results regarding $\mathcal{T}$ for $2 \le t < k$. Note that the cases of $3 \le k \le 4$ and $5 \le k \le t$ have been studied by Kunert and Martin (2000) and Kunert and Mersmann (2011). Two special cases of type-$H$ covariance matrix are the identity matrix and a completely symmetric matrix.

THEOREM 6. *Assume $\Sigma$ to be of type-$H$.* (i) *If $2 \le t \le k-2$, we have*

$$z^* = 0,$$
$$y^* = k(t-1)/t - v(t-v)/kt,$$
$$\mathcal{T} = \{s : f_{s,m} = u \text{ or } u+1,\ 1 \le m \le t\},$$

*where $u$ and $v$ are the integers satisfying $k = ut + v$ and $0 \le v < t$.*

(ii) *If $2 \le t = k-1$, we have*

$$z^* = \frac{1}{2[k(k-3)+1/t]}, \tag{36}$$

$$y^* = k - 1 - \frac{2}{k} - \frac{1}{2k[k(k-3)+1/t]}, \tag{37}$$

$$\mathcal{T} = \langle s_0 \rangle \cup \langle s_0' \rangle, \tag{38}$$

*where $s_0 = (1\,1\,2 \cdots t)$ and $s_0'$ is its dual sequence. Moreover, a measure maximizes $q_\xi^*$ if and only if $p_{\langle s_0 \rangle} = p_{\langle s_0' \rangle} = 1/2$.*

PROOF. Due to (4), here we assume $\Sigma = I_k$ throughout the proof without loss of generality. For sequence $s = (t_1 \cdots t_k)$, define the quantities $\phi_s = \sum_{i=1}^{k-1} \mathbb{I}_{t_i = t_{i+1}}$, $\varphi_s = \sum_{i=2}^{k-1} \mathbb{I}_{t_{i-1} = t_{i+1}}$, $f_{s,m} = \sum_{i=1}^{k} \mathbb{I}_{t_i = m}$, $\chi_s = \sum_{m=1}^{t} f_{s,m}^2$. By direct calculations, we have

$$\tilde{q}_s(z) = q_{s,0} + q_{s,1} z + q_{s,2} z^2, \tag{39}$$

$$q_{s,0} = c_{s00} = k - \chi_s / k, \tag{40}$$

$$q_{s,1} = c_{s01} + c_{s02} = 2(2k\phi_s + f_{s,t_1} + f_{s,t_k} - 2\chi_s)/k, \tag{41}$$

$$\begin{aligned} q_{s,2} &= c_{s11} + 2c_{s12} + c_{s22} \\ &= 2\big[\varphi_s + k - 1 - (k+t-2)/kt\big] \\ &\quad - 2(2\chi_s - 2f_{s,t_1} - 2f_{s,t_k} + \mathbb{I}_{t_1 = t_k})/k. \end{aligned} \tag{42}$$

(i) follows by the same approach as in Theorem 1.a of Kushner (1998) with only more tedious arguments based on (39)–(42).

Now we focus on $t = k - 1$. First, we have $\phi_{s_0} = 1$, $\varphi_{s_0} = 0$ and $\chi_{s_0} = k + 2$, and hence $q_{s_0,0} = k - 1 - 2/k$, $q_{s_0,1} = -2/k$ and $q_{s_0,2} = 2(k - 3) + 2/kt$. It can be verified that $\tilde{q}_{s_0}(z)$ reaches its minimum at $z = z^*$. Since $\tilde{q}_{s_0}(z) = \tilde{q}_{s_0'}(z)$, it is sufficient to show $\tilde{q}_{s_0}(z^*) = \max_{s\in\mathcal{S}} \tilde{q}_s(z^*)$ for the purpose of proving (ii).

We first restrict the consideration to the subset $\mathcal{S}_1 = \{s : t_1 \neq t_k, s \in \mathcal{S}\}$. If we only exchange the treatments in locations $\{2, \dots, k-1\}$, the values of $\chi_s$, $f_{s,t_1}$ and $f_{s,t_k}$ remain invariant. Note that $\tilde{q}_s(z^*)$ is increasing in the quantity $\phi_s + 2^{-1}z^*\varphi_s$. If for a certain location, say $i$, we have $t_{i-1} = t_{i+1} \neq t_i$. At least one of $i - 1$ and $i + 1$ would be in the set $\{2, \dots, k-1\}$. After switching this location with location $i$, $\phi_s$ will be increased by 1, and at the same time the amount of decrease for $\varphi_s$ will be at most 2. Note that $z^*/2 \leq 1/2$ for all $p \geq 3$ and $t \geq 2$, and hence a sequence, say $s$, which maximizes $\tilde{q}_s(z^*)$ should be of the format $s = (1'_{f_{s,1}}1|\cdots|1'_{f_{s,h}}h)$, without loss of generality. Here, $h := h(s)$ is the number of distinct treatments in sequence $s$ and $\sum_{i=1}^h f_{s,i} = p$. Among sequences of this particular format, the sequence which maximizes $\tilde{q}_s(z^*)$ should satisfy $\min(f_{s,1}, f_{s,h}) \geq \max_{2\leq i\leq h-1} f_{s,i}$, where we take the maximization over the empty set to be 0. Without loss of generality, we assume $t_1 = \max_{1\leq i\leq t} f_{s,i}$. Now we shall show $f_{s,1} \leq 2$ for maximizing sequences as follows. Suppose $f_{s,1} \geq 3$, this indicates $h < t$. By decreasing $f_{s,1}$ by one and changing $f_{s,h+1}$ from 0 to 1, the quantity $\tilde{q}_s(z^*)$ is increased by the amount of

$$\Delta_s = \frac{2}{k}\big[f_{s,1} - 1 + (4f_{s,1} - 5 - 2k)z^* + (4f_{s,1} - 8 - k)(z^*)^2\big].$$

If $k = 3$, we have $\Delta_s > 0$ in view of $z^* > 0$ and $f_{s,1} \geq 3$. Suppose $k \geq 4$, we have $0 < z^* \leq (2k)^{-1}$, hence we have

$$\begin{aligned} k\Delta_s/2 &= f_{s,1} - 1 - 2kz^* - p(z^*)^2 + (4f_{s,1} - 5)z^* + (4f_{s,1} - 8)(z^*)^2 \\ &> f_{s,1} - 2 - (4k)^{-1} > 0. \end{aligned}$$

At this point, we have shown $\tilde{q}_{s_0}(z^*) = \max_{s\in\mathcal{S}_1} \tilde{q}_s(z^*)$. By similar arguments, one can show that the sequence $s_1 = (1\ 2\cdots t1)$ maximizes $\tilde{q}_s(z^*)$ among $s \notin \mathcal{S}_1$. By direct calculations, we have

$$\begin{aligned} \tilde{q}_{s_0}(z^*) - \tilde{q}_{s_1}(z^*) &= (4 - 2/k)z^* - 4(z^*)^2/k \\ &\geq z^*(10/3 - 2/k^2) > 0. \end{aligned}$$

Hence, (36)–(38) are proved. For the rest of (ii), the sufficiency of $p_{\langle s_0\rangle} = p_{\langle s_0'\rangle} = 1/2$ is indicated by the proof of Theorem 5. For the necessity, it is enough to note that the two components of $\nabla q_\xi(x^*) = 2(\ell_\xi + Q_\xi x^*) = 2\sum_{s\in\mathcal{S}} p_s(\ell_s + Q_s x^*)$ will not be identical if $p_{\langle s_0\rangle} \neq p_{\langle s_0'\rangle}$. Hence, the lemma is concluded. $\square$

**6. Examples.** This section tries to illustrate the theorems of this paper through several examples for various combinations of $k, t, n$ and $\Sigma$. By Theorem 5(iii), the efficiency of a design is higher under model (27) than under model (2) for any criterion function $\Phi$ satisfying (C.1)–(C.3) under a mild condition, that is, $\Sigma$ is persymmetric. Hence, it is sufficient to propose optimal or efficient designs under model (2). The existence of the universally optimal measure in $\mathcal{P}$ is obvious in view of Lemmas 1 and 4. However, to derive an exact design, one has to restrict the consideration to the subset $\mathcal{P}_n = \{\xi \in \mathcal{P} : n\xi \text{ is a vector of integers}\}$. Universally, optimal measure does not necessarily exist in $\mathcal{P}_n$ except for certain combinations of $k, t, n$. In this case, one can convert $p_s$ in the equations of Theorem 3 into $n_s$ by multiplying both sides of the equations by $n$. Then one can define a distance between two sides of the equations and find the solution, say $\{n_s, s \in \mathcal{T}\}$, to minimize this distance. If there is universally optimal measure in $\mathcal{P}_n$, such approach automatically locates the universally optimal exact design; otherwise, the exact designs thus found are typically highly efficient under the different criteria. See Zheng (2013a) and Figure 1 for evidence.

Let $0 \le a_1 \le a_2 \le a_{t-1}$ be the $t$ eigenvalues of $C_d$ for an exact design $d$. If $d$ is universally optimal, we have $a_i = ny^*/(t-1)$, $1 \le i \le t-1$. Here, we define $A$-, $D$-, $E$- and $T$-efficiencies of design $d$ as follows:

$$\mathcal{E}_A(d) = \frac{t-1}{ny^*}\frac{t-1}{(\sum_{i=1}^{t-1} a_i^{-1})} = \frac{(t-1)^2}{ny^*(\sum_{i=1}^{t-1} a_i^{-1})},$$

$$\mathcal{E}_D(d) = \frac{t-1}{ny^*}\left(\prod_{i=1}^{t-1} a_i\right)^{1/(t-1)},$$

$$\mathcal{E}_E(d) = \frac{(t-1)a_1}{ny^*},$$

$$\mathcal{E}_T(d) = \frac{t-1}{ny^*}\left(\frac{1}{t-1}\sum_{i=1}^{t-1} a_i\right) = \frac{\sum_{i=1}^{t-1} a_i}{ny^*}.$$

It is well known that a universally optimal measure has unity efficiency under these four criteria.

We begin with the discussion on the case when $\Sigma$ is of type-$H$. For the latter, Kunert and Martin (2000) studied the conditions on $p_{\langle s\rangle}$ for a pseudo symmetric design to be universally optimal for $k = 3$ and 4, which was further extended by Kunert and Mersmann (2011) to $t \ge k \ge 5$. We would comment on these cases and then explore the case of $k \ge 5$ and $t < k$. Finally, irregular form of $\Sigma$ will be briefly discussed.

For $(k, t) = (4, 2)$, Corollary 1 indicates that the necessary and sufficient condition for a pseudo symmetric design to be universally optimal is $p_{\langle(1\ 1\ 2\ 2)\rangle} = 3p_{\langle(1\ 2\ 1\ 2)\rangle} + p_{\langle(1\ 2\ 2\ 1)\rangle}$. Theorem 2 of Kunert and Martin (2000) proposed

$p_{\langle(1\ 1\ 2\ 2)\rangle} = p_{\langle(1\ 2\ 2\ 1)\rangle} = 1/2$, which is sufficient but not necessary for universal optimality. For $k = 3$ and $(k, t) = (4, 3)$, Corollary 1 indicates that sufficient conditions regarding $p_{\langle s\rangle}$ given by Theorems 1 and 3 of Kunert and Martin (2000) are also necessary.

For $t \geq k = 4$, Kunert and Martin (2000) showed that the optimal values of $p_{\langle s\rangle}$ are given by irrational numbers, and hence an exact universally optimal design does not exist. In fact, based on Theorem 3 here, one can derive efficient exact designs for the majority values of $t$ and $n$. For example, $d_1$ below with $t = 4$ and $n = 10$ yields the efficiencies of $\mathcal{E}_A(d_1) = 0.9943$, $\mathcal{E}_D(d_1) = 0.9946$, $\mathcal{E}_E(d_1) = 0.9682$ and $\mathcal{E}_T(d_1) = 0.9949$. Note that the $E$-efficiency is relatively lower than other efficiencies due to the asymmetry of the design.

$$d_1 = \begin{bmatrix} 2 & 1 & 4 & 3 & 1 & 1 & 3 & 2 & 4 & 3 \\ 2 & 1 & 4 & 3 & 2 & 4 & 4 & 3 & 2 & 2 \\ 1 & 3 & 3 & 1 & 4 & 3 & 2 & 1 & 1 & 4 \\ 1 & 4 & 2 & 2 & 4 & 3 & 2 & 4 & 3 & 1 \end{bmatrix}.$$

For $t \geq k \geq 5$, Kunert and Mersmann (2011) showed that the set $\mathcal{T}$ should include sequences $(1\ 2 \cdots k)$, $(1\ 1\ 2 \cdots k-3\ k-2\ k-2)$, $s_0$ and its dual sequence $s_0'$ as defined in Theorem 6. The optimal proportion for them are again irrational numbers. Further, they proposed the use of type I orthogonal array ($OA_I$), that is, $p_{\langle(1\ 2 \cdots k)\rangle} = 1$, and proved that the $T$-efficiencies of such designs are at least 0.94. Note that $OA_I$ is pseudo symmetric, hence its efficiencies are identical under criteria $A$, $D$, $E$ and $T$.

When $t = k - 1$, Theorem 6(ii) indicates that a pseudo symmetric design with $p_{\langle s_0\rangle} = p_{\langle s_0'\rangle} = 1/2$ will be universally optimal. For example, when $t = 4$ and $k = 5$, $d_2$ below with $n = 24$ is universally optimal. Here, the first 12 sequences are equivalent to (1 1 2 3 4) while the rest are equivalent to (1 2 3 4 4).

$$d_2 = \left[\begin{matrix} 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 4 \\ 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 4 \\ 4 & 2 & 3 & 1 & 4 & 3 & 1 & 4 & 2 & 1 & 2 & 3 \\ 2 & 3 & 4 & 4 & 3 & 1 & 2 & 1 & 4 & 3 & 1 & 2 \\ 3 & 4 & 2 & 3 & 1 & 4 & 4 & 2 & 1 & 2 & 3 & 1 \end{matrix}\right.$$

$$\left.\begin{matrix} 3 & 4 & 2 & 3 & 1 & 4 & 4 & 2 & 1 & 2 & 3 & 1 \\ 2 & 3 & 4 & 4 & 3 & 1 & 2 & 1 & 4 & 3 & 1 & 2 \\ 4 & 2 & 3 & 1 & 4 & 3 & 1 & 4 & 2 & 1 & 2 & 3 \\ 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 4 \\ 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 4 \end{matrix}\right].$$

When $2 \leq t < k - 1$, there is a large variety of symmetric blocks in $\mathcal{T}$ and there will be infinity many solutions for optimal sequence proportions. Even for $t = 2$ and $k = 5$, we shall have $\mathcal{T} = \langle(1\ 1\ 1\ 2\ 2)\rangle \cup \langle(1\ 1\ 2\ 2\ 2)\rangle \cup \langle(1\ 1\ 2\ 1\ 2)\rangle \cup \langle(1\ 2\ 1\ 2\ 2)\rangle \cup \langle(1\ 1\ 2\ 2\ 1)\rangle \cup \langle(1\ 2\ 2\ 1\ 1)\rangle \cup \langle(1\ 2\ 1\ 1\ 2)\rangle \cup \langle(1\ 2\ 2\ 1\ 2)\rangle \cup \langle(1\ 2\ 1\ 2\ 1)\rangle \cup \langle(1\ 2\ 2\ 2\ 1)\rangle$. Let $p_1, \ldots, p_{10}$ be the proportions of these symmetric

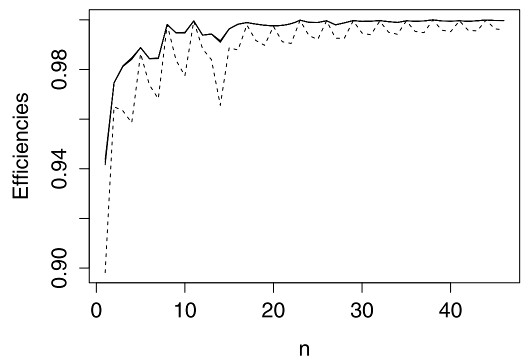


FIG. 1. *The efficiencies of exact designs for $5 \le n \le 50$ when $k = 4$, $t = 3$ and $\eta = 0.5$. The $E$-efficiency is plotted by the dashed line, while $A$-, $D$- and $T$-efficiencies are all plotted by the same solid line.*

blocks. A pseudo symmetric design with $p_1 = p_2$, $p_3 = p_4$, $p_5 = p_6$, $p_7 = p_8$, $1.8(p_1 + p_2) = 2.2(p_3 + p_4 + p_7 + p_8) + 4p_9 + 0.4p_{10}$, $\sum_{i=1}^{10} p_i = 1$ and $p_i \ge 0$ will be universally optimal. One simple solution is $p_5 = p_6 = 1/2$. Hence a design which assigns $1/4$ of its blocks to sequences (1 1 2 2 1), (2 2 1 1 2), (1 2 2 1 1) and (2 1 1 2 2) is universally optimal.

At last, we would like to convey the message that the deviation of $\Sigma$ from type-$H$ has large impact on the choice of designs. For simplicity of illustration, we consider the form $\Sigma = (\mathbb{I}_{i=j} + \eta\mathbb{I}_{|i-j|=1})_{1\le i,j\le k}$. When $k = t = 5$ and $\eta = 0.5$, the efficiency of $OA_I$ reduces to 0.8232. In fact, Corollary 1 indicates that $\langle(1\ 1\ 2\ 3\ 3)\rangle$, instead of $\langle(1\ 2\ 3\ 4\ 5)\rangle$ for $\eta = 0$, becomes the dominating symmetric block among the four. To be more specific, a pseudo symmetric design with sequences solely from $\langle(1\ 1\ 2\ 3\ 3)\rangle$ yields the efficiency of 0.9999 for all four criteria. When we tune $\eta$ to 0.9, the efficiency of $OA_I$ further reduces to 0.3395, while the symmetric design based on $\langle(1\ 1\ 2\ 3\ 3)\rangle$ becomes even more efficient. One the other hand, when $\eta$ takes negative values, the efficiency of $OA_I$ becomes even higher than 0.94. Similar phenomena are observed for other cases of $t \ge k$.

For $t < k$, we also observe that the value of $\eta$ influences the choice of design substantially. The details are omitted due to the limit of space. We end this section by Figure 1. It shows that the linear equations system in Theorem 3 is powerful in deriving efficient exact designs for arbitrary values of $n$.

## APPENDIX: THE ALGORITHM BASED ON THEOREM 1

Recall that $m$ is be total number of distinct symmetric blocks and $s_1, s_2, \dots, s_m$ are the $m$ representatives for each of the symmetric blocks. Note that two pseudo symmetric measures with the same vector of $P_\xi = (p_{\langle s_1\rangle}, p_{\langle s_2\rangle}, \dots, p_{\langle s_m\rangle})$ have the same information matrix and hence the same performance under all optimality criteria. For a measure $\xi$ and a sequence $s$, we define

$$\theta(P_\xi, s) = \operatorname{tr}(R_s R_\xi^{-1}) - \operatorname{tr}(Q_s Q_\xi^{-1}). \tag{43}$$

We also define $\theta^*(P_\xi) = \max_{1\le i\le m} \theta(P_{\langle d\rangle}, s_i)$ and $e_i$ to be vector of length $m$ with the $i$th entry as 1 and other entries as 0.

*Step* 0: Choose tuning parameters $\epsilon > 0$ and $\omega$ such that $\epsilon$ is in a small neighborhood of zero and $\omega$ is in a neighborhood of one.

*Step* 1: Choose initial measure $P^{(0)} = P_{\xi_0}$. Put $i_0 = \operatorname{argmin}_{1\le j\le m} \theta^*(e_i)$ and $n = 0$, then let $P^{(0)} = e_{i_0}$.

*Step* 2: Check optimality. If $\theta_n := \theta^*(P^{(n)}) > 1 + \epsilon$, go to step 3. Otherwise, output the optimal measure as $P^{(n)}$.

*Step* 3: Update the measure. Let $i_{n+1} = \operatorname{argmax}_{1\le i\le m} \theta(P^{(n)}, s_i)$ and the updated measure is $P^{(n+1)} = (\theta_n - 1)^\omega e_{i_{n+1}} + (1 - (\theta_n - 1)^\omega)P^{(n)}$. Increase $n$ by 1 and go back to step 2.

REMARK 4. There is a possibility of tie in choosing $i_0$ in step 1 and $i_{n+1}$ in step 3. The strategy in such case is quite arbitrary. Let $\Xi_n = \{i : \theta(P^{(n)}, s_i) = \theta^*(P^{(n)})\}$. If $|\Xi_n| > 1$, one can either choose an arbitrary $j_n \in \Xi_n$ and let $i_{n+1} = j_n$ or replace $e_{i_{n+1}}$ in step 3 by $|\Xi_n|^{-1} \sum_{i\in\Xi_n} e_i$. The same strategy applies to the choice of $i_0$.

REMARK 5. Note that the update algorithm in step 3 is essentially a steepest descent algorithm. The parameter $\omega$ is to adjust for the length of step for the *best* direction. By the concavity of the optimality criteria, the global optimum is guaranteed to be found. In the examples of this paper, $\omega = 1$ works well enough. The parameter $\epsilon$ is used to adjust for time of convergence. When the sequential algorithm converges very slow, one can increase $\epsilon$ to save time. In most examples of this paper, setting $\epsilon = 10^{-7}$ enable us to obtain the optimal design within 10 seconds.

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

## REFERENCES

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

AI, M.-Y., GE, G.-N. and CHAN, L.-Y. (2007). Circular neighbor-balanced designs universally optimal for total effects. *Sci. China Ser. A* **50** 821–828. MR2353065

AI, M., YU, Y. and HE, S. (2009). Optimality of circular neighbor-balanced designs for total effects with autoregressive correlated observations. *J. Statist. Plann. Inference* **139** 2293–2304. MR2507991

BAILEY, R. A. and DRUILHET, P. (2004). Optimality of neighbor-balanced designs for total effects. *Ann. Statist.* **32** 1650–1661. MR2089136

BAILEY, R. A. and DRUILHET, P. (2014). Optimal cross-over designs for full interaction models. *Ann. Statist.* **42** 2282–2300. MR3269980

BESAG, D. and KEMPTON, R. A. (1986). Statistical analysis of field experiments using neighboring plots. *Biometrics* **42** 231–251.

CLARKE, F. R., BAKER, R. J. and DEPAUW, R. M. (2000). Plot direction and spacing effects on interplot interference in spring wheat cultivar trials. *Crop Science* **40** 655–658.

CONNOLLY, T., MCNICOL, J. W., WASTIE, R. L. and STEWART, H. E. (2008). Evaluating between-plant interference in field trials for assessing potato genotype resistence to late blight. *Annals of Applied Biology* **127** 2732–282.

DAVID, O., MONOD, H., LORGEOU, J. and PHILIPPEAU, G. (2001). Control of interplot interference in grain maize: A multi-site comparison. *Crop Science* **41** 406–414.

DRAPER, N. R. and GUTTMAN, I. (1980). Incorporating overlap effects from neighbouring units into response surface models. *Appl. Statist.* **29** 128–134.

DRUILHET, P. (1999). Optimality of neighbour balanced designs. *J. Statist. Plann. Inference* **81** 141–152. MR1718460

DRUILHET, P. and TINSSON, W. (2012). Efficient circular neighbour designs for spatial interference model. *J. Statist. Plann. Inference* **142** 1161–1169. MR2879760

FILIPIAK, K. (2012). Universally optimal designs under an interference model with equal left- and right-neighbor effects. *Statist. Probab. Lett.* **82** 592–598. MR2887476

FILIPIAK, K. and MARKIEWICZ, A. (2003). Optimality of neighbor balanced designs under mixed effects model. *Statist. Probab. Lett.* **61** 225–234. MR1959129

FILIPIAK, K. and MARKIEWICZ, A. (2005). Optimality and efficiency of circular neighbor balanced designs for correlated observations. *Metrika* **61** 17–27. MR2126434

FILIPIAK, K. and MARKIEWICZ, A. (2007). Optimal designs for a mixed interference model. *Metrika* **65** 369–386. MR2299558

GILL, P. S. (1993). Design and analysis of field experiments incorporating local and remote effects of treatments. *Biom. J.* **35** 343–354. MR1222502

GOLDRINGER, I., BRABANT, P. and KEMPTON, R. A. (1994). Adjustment for competition between genotypes in single-row-plot trials of winter wheat (*Triticum awstivum*). *Plant Breeding* **112** 294–300.

KEMPTON, R. A. (1982). Adjustment for competition between varieties in plant breeding trials. *J. Agric. Sci., Camb.* **98** 599–611.

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. Srivistava, eds.) 333–353. North-Holland, Amsterdam. MR0395079

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 MERSMANN, S. (2011). Optimal designs for an interference model. *J. Statist. Plann. Inference* **141** 1623–1632. MR2747931

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.

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.

LANGTON, S. (1990). Avoiding edge effects in agroforestry experiments; the use of neighbour-balanced designs and guard areas. *Agroforestry Systems* **12** 173–185.

REES, D. H. (1967). Some designs of use in serology. *Biometrics* **23** 779–791.

ZHENG, W. (2013a). Universally optimal crossover designs under subject dropout. *Ann. Statist.* **41** 63–90. MR3059410

ZHENG, W. (2013b). Optimal crossover designs for the proportional model. *Ann. Statist.* **41** 2218–2235. MR3127864

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