The Annals of Statistics
1997, Vol. 25, No. 6, 2328–2344

OPTIMAL REPEATED MEASUREMENTS DESIGNS: THE LINEAR OPTIMALITY EQUATIONS[1]

BY H. B. KUSHNER

Nathan S. Kline Institute for Psychiatric Research

In approximate design theory, necessary and sufficient conditions that a repeated measurements design be universally optimal are given as linear equations whose unknowns are the proportions of subjects on the treatment sequences. Both the number of periods and the number of treatments in the designs are arbitrary, as is the covariance matrix of the normal response model. The existence of universally optimal "symmetric" designs is proved; the single linear equation which the proportions satisfy is given. A formula for the information matrix of a universally optimal design is derived.

1. Introduction. In the class RMD(t, N, p) of repeated measurements designs—where t is the number of treatments, N the number of subjects and p the number of periods—a subject is repeatedly exposed to various treatments. Many authors [Cheng and Wu (1980), Kunert (1983), Hedayat and Zhao (1990)] assume the response model

$$y_{ij} = \tau_{d(i,j)} + \rho_{d(i,j-1)} + \pi_j + \beta_i + \varepsilon_{ij}, \qquad 1 \le i \le N, \quad 1 \le j \le p, \tag{1.1}$$

in which the design allocates treatment $d(i, j)$ to the ith subject in the jth period. $\tau_{d(i,j)}$ and $\rho_{d(i,j)}$ are the treatment and carryover effects, π_j is the period effect and β_i is the subject effect. The N independent vectors $\varepsilon_i = (\varepsilon_{ij})$, $1 \le j \le p$, are multivariate normal with mean 0, and $p \times p$ covariance matrix, C.

Our chief interest is to determine all designs which minimize functions which depend on the proportions of subjects assigned to each treatment sequence. We use "proportions" in the sense customary in approximate design theory, the approach of this paper. The functions considered are the standard information functions.

Previous optimality results contain restrictions on p, t, C, the competing class of designs and the class of designs proved to be optimal. A necessary and sufficient condition for universal optimality was given for the case $p = 2$ in Hedayat and Zhao (1990), who also pointed out that most other optimality

Received October 1995; revised July 1997.
[1]Research supported in part by National Institute of Mental Health Grants MH 42959 and P50MH51359.
AMS 1991 *subject classifications*. Primary 62K05; secondary 62K10.
Key words and phrases. Optimal repeated measurements designs, universal optimality, treatment effect, carryover effect, treatment sequence.

results relate to the situation $t \le p$. The conditions $p = t$ and $p \equiv 1 \pmod t$ appear in Cheng and Wu (1980) and $p = t$ in Kunert (1985). The case $t = 2$ was treated in Laska and Meisner (1985) and in Matthews (1987).

Pioneering work [Hedayat and Afsarinejad (1978); Cheng and Wu (1980)] concentrated on the case $C = I_p$. However, a subject's p responses are not likely to be independent. Moreover, the optimality of a design usually depends on C (an exception is the case $p = 2$). In almost all subsequent work [Kunert (1985); Matthews (1987)], C is assumed to be a first-order autoregressive matrix, but since this assumption is highly specialized, there remains the open problem of obtaining optimal designs for a general covariance matrix.

Also, many authors place limitations on the class of competing designs. For example, in Cheng and Wu (1980), some optimality results are proved when the class of competing designs excludes those in which a treatment is preceded by itself. Even when the class of competing designs is the full class RMD(t, N, p), optimality results are often only proved for interesting subclasses of designs [Cheng and Wu (1980), Section 3; Kunert (1985)].

None of the above limitations, except for $p = 2$, appear in Hedayat and Zhao (1990) (their assumption that C is a correlation matrix is easy to remove). The present paper provides a comparable treatment of the general optimal design problem in approximate theory.

The main results of this paper are the following. For any p, t and C, a necessary and sufficient condition that the proportions of subjects on treatment sequences determine a universally optimal (resp. strictly Φ-optimal) design is that they satisfy a system of linear equations (Theorems 5.3 and 5.4). Surprisingly, "strict Φ-optimality"—the optimality of a design with respect to any strictly concave information function Φ—is equivalent to the universal optimality of the design. A formula (4.13) for the information matrix of a universally optimal (resp. strictly Φ-optimal) design is derived. For any Φ, the related formula (4.12) can be used both to check a design's optimality and to compute the Φ-efficiency of a nonoptimal design. For any p, t or C, there exist "symmetric" designs which are universally optimal. In a symmetric design (Section 3), equal proportions of subjects are assigned to all treatment sequences within a "symmetry block" (Section 5). These proportions are determined by just one linear equation (Theorem 5.5), which is also a necessary condition for any optimal design. A symmetric design consisting of at most two symmetry blocks always exists (Corollary 5.6).

In Kushner (1996, 1997), we study solutions of the optimality equations. We derive the designs of Cheng and Wu (1980), Kunert (1983) and Matthews (1987, 1990) and give new universally optimal designs.

2. The information matrix of treatment and carryover effects. In this section, we give a formula (Lemma 2.4) for $C_d(\tau, \rho)$, the information matrix of treatment and carryover effects. The formula concretely exhibits the treatment sequence proportions whose determination is the aim of optimal design theory. The matrix, $C^{\sim}$, which also appears in the formula, significantly affects a design's optimality.

2.1. *Notation.* I_n denotes the $n \times n$ identity matrix, J_n the $n \times n$ matrix of 1's and $J_{m,n} = J^1_{m,n}$ the $m \times n$ matrix of 1's. $J^2_{m,n}$ is the $m \times n$ matrix whose first row is zero and final $m-1$ rows are 1's. $1_m = 1^1_m = J^1_{m,1}$ and $1^2_m = J^2_{m,1}$. T^{μ}, $1 \le \mu \le 2$, are, respectively, the $Np \times t$ design matrices of treatment and carryover effects; T^1_i and T^2_i are the $p \times t$ period-treatment and period-carryover incidence matrices for subject i, so that $T^{\mu} = [(T^{\mu}_1)', (T^{\mu}_2)', \dots, (T^{\mu}_N)']'$. The first row of T^2_i is zero; for $2 \le j \le p$, the jth row of T^2_i is the $(j-1)$st row of T^1_i. Note that

$$1^{\mu}_p = T^{\mu}_i 1_t, \qquad 1 \le i \le N, \quad 1 \le \mu \le 2. \tag{2.1}$$

Write $T^{\mu}_{d,i}$ to designate the matrix T^{μ}_i when it is necessary to indicate the design d. In Kunert (1991), $T^1 = T_d$, $T^2 = F_d$, $T^1_i = T_{d,i}$, $T^2_i = F_{d,i}$. Let

$$M^{\mu} = N^{-1} \sum_{i=1}^{N} T^{\mu}_i, \qquad 1 \le \mu \le 2. \tag{2.2}$$

2.2. *Formulas for $C_d(\tau, \rho)$. The matrix $C^{\sim}$.* Define the $p \times p$ matrix $C^{\sim}$ by

$$C^{\sim} = C^{-1} - C^{-1} J_p C^{-1} / 1'_p C^{-1} 1_p. \tag{2.3}$$

Note that if x is a p-dimensional vector, then

$$C^{\sim} x = 0 \quad \text{iff} \quad x = \alpha 1_p, \ \alpha \text{ a scalar}. \tag{2.4}$$

LEMMA 2.1. *The information matrix of treatment and carryover effects is given by*

$$C_d(\tau, \rho) = \begin{pmatrix} C_{d11} & C_{d12} \\ C_{d21} & C_{d22} \end{pmatrix}, \tag{2.5}$$

where $C_{d\mu\nu} = (T^{\mu})' S T^{\nu}$, $1 \le \mu, \nu \le 2$, *are* $t \times t$ *matrices and* $S = (I_N - J_N/N) \otimes C^{\sim}$.

See Cheng and Wu (1980) for the case $C = I_p$. An application of Lemma 2.1 is: when $p = 2$, optimality is independent of the covariance matrix. For any C, $C^{\sim} = \beta\left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)$, β a scalar, and scalar multiples of information matrices do not affect optimality. See Hedayat and Zhao (1990) for the case C a correlation matrix.

COROLLARY 2.2. *Any design which is optimal when $C = C_0$ is also optimal if*

$$C = \alpha C_0 + \gamma 1'_p + 1_p \gamma', \qquad \alpha \text{ a scalar}, \ \gamma \text{ a } p\text{-dimensional vector}. \tag{2.6}$$

PROOF. The C in (2.6) is the most general solution of the matrix equation, $C^{\sim} = \alpha^{-1}C_0^{\sim}$. Hence, $C_d(\tau, \rho)$ is the same for the covariance matrices (2.6) and $C = \alpha C_0$. Consequently, the same situation holds for the information matrix of treatment effects. □

An application of Corollary 2.2 is: designs optimal with $C = I_p$ are also optimal with $C = \alpha I_p + \gamma 1_p' + 1_p \gamma'$, a "type-H" matrix [Huynh and Feldt (1970)]. In fact, for any C of this type,

$$C^{\sim} = \alpha^{-1}(I_p - J_p/p).$$

LEMMA 2.3. (i) *The submatrices of* $C_d(\tau, \rho)$ *are given by*

$$C_{d\mu\nu} = \sum_{i=1}^{N} (T_i^{\mu} - M^{\mu})' C^{\sim} (T_i^{\nu} - M^{\nu}), \qquad 1 \le \mu, \nu \le 2. \tag{2.7}$$

(ii) *The row and column sums of* $C_{d\mu\nu}$, $1 \le \mu, \nu \le 2$, *are zero.*

2.3. *Treatment sequences*: *proportions and treatment/carryover incidence matrices.* Let τ denote a treatment sequence characterized by a p-tuple of integers, $\tau = (t_1, t_2, \dots, t_p)$, $1 \le t_j \le t$, where t_j is the treatment given in the jth period, $1 \le j \le p$. The matrices T_i^{μ} depend only on the treatment sequence to which a subject is assigned: $T_{\tau}^{\mu} = T_i^{\mu}$, where T_{τ}^1 and T_{τ}^2 are, respectively, the common treatment and carryover incidence matrices for all subjects assigned to τ. If n_{τ} is the number of such subjects, then $N = \sum_{\tau} n_{\tau}$, where τ—and in subsequent sums omitting the domain of τ—runs over all t^p treatment sequences. With $p_{\tau} = n_{\tau}/N$ the treatment sequence proportion,

$$1 = \sum_{\tau} p_{\tau}, \qquad p_{\tau} \ge 0. \tag{2.8}$$

2.4. $C_d(\tau, \rho)$ *as a sum over treatment sequences and their proportions. The matrices* $T_{\tau}, T_{\tau}^{\mu}, \hat{T}_{\tau}, \hat{T}_{\tau}^{\mu}$ *and their convex means,* $M_d, M_d^{\mu}, \hat{M}_d, \hat{M}_d^{\mu}$. *The matrices* $L_d, \hat{L}_d, \hat{L}_{d\mu\nu}$. Let T_{τ} be the $p \times 2t$ matrix $T_{\tau} = [T_{\tau}^1, T_{\tau}^2]$. Consistent with (2.2), define the convex means, M_d^{μ}, M_d, of the matrices T_{τ}^{μ}, T_{τ} by $M_d^{\mu} = \sum_{\tau} p_{\tau} T_{\tau}^{\mu}$, $1 \le \mu \le 2$, and $M_d = \sum_{\tau} p_{\tau} T_{\tau}$. Let $L_d = \sum_{\tau} p_{\tau} T_{\tau}' C^{\sim} T_{\tau}$.

LEMMA 2.4. *The information matrix of treatment and carryover effects is given by the sum*

$$C_d(\tau, \rho) = N \sum_{\tau} p_{\tau} (T_{\tau} - M_d)' C^{\sim} (T_{\tau} - M_d). \tag{2.9}$$

From (2.9), $C_d(\tau, \rho)$ may be thought of as the weighted variance of the T_{τ} matrices occurring with multiplicity Np_{τ}. A matrix version of a familiar identity is $C_d(\tau, \rho) = NL_d - NM_d' C^{\sim} M_d$. More generally, for any $p \times 2t$ matrix, T,

$$C_d(\tau, \rho) = N \sum_{\tau} p_{\tau} (T_{\tau} - T)' C^{\sim} (T_{\tau} - T) - N(T - M_d)' C^{\sim} (T - M_d). \tag{2.10}$$

Set $T = [J^1_{p,t}, J^2_{p,t}]/t$ and define $\hat{T}_\tau$, $\hat{T}^1_\tau$, $\hat{T}^2_\tau$, $\hat{M}_d$, $\hat{M}^1_d$, $\hat{M}^2_d$ and $\hat{L}_d$, the centered counterparts of the matrices without the "hats," by

$$\hat{T}_\tau = T_\tau - T = \left[\hat{T}^1_\tau, \hat{T}^2_\tau\right] = \left[T^1_\tau - J^1_{p,t}/t, T^2_\tau - J^2_{p,t}/t\right], \tag{2.11}$$

$$\hat{M}_d = M_d - T = \left[\hat{M}^1_d, \hat{M}^2_d\right] = \left[M^1_d - J^1_{p,t}/t, M^2_d - J^2_{p,t}/t\right], \tag{2.12}$$

$$\hat{L}_d = \sum_\tau p_\tau \hat{T}'_\tau C^{\sim} \hat{T}_\tau. \tag{2.13}$$

From (2.10),

$$C_d(\tau, \rho) = N\hat{L}_d - N\left(\hat{M}_d\right)' C^{\sim} \hat{M}_d. \tag{2.14}$$

Finally, define the $t \times t$ submatrices of $\hat{L}_d$,

$$\hat{L}_{d\mu\nu} = \sum_\tau p_\tau \left(\hat{T}^\mu_\tau\right)' C^{\sim} \hat{T}^\nu_\tau, \qquad 1 \le \mu, \nu \le 2. \tag{2.15}$$

3. Permuted and symmetric designs, convex combinations of designs. In this section, we discuss permuted designs, the convex combination of designs and symmetric designs. The exploitation of designs of these types is the theme of the method of this paper.

3.1. *P, the array of proportions. Permuted designs and their information matrices.* A repeated measurements design d is specified by $P = (p_\tau)$, the t^p proportions on the treatment sequences. Sometimes we refer to the design d by P and indicate the relation between d and P by $d \leftrightarrow P$. Let σ denote a permutation of $\{1, 2, \dots, t\}$, that is, σ is an element of the symmetric group, S_t. Let $\sigma\tau$ be the treatment sequence $\sigma\tau = (\sigma(t_1), \sigma(t_2), \dots, \sigma(t_p))$. For each $\sigma \in S_t$, define a new design, the permuted design, P_σ, by $P_\sigma = (p_{\sigma^{-1}\tau})$, designated also by d_σ. Let H_σ be the $t \times t$ matrix representing the permutation σ. Since d_σ results from a relabelling of the treatments, the d and d_σ incidence matrices of the ith subject and their means are related by

$$T^\mu_{d_\sigma, i} = T^\mu_{d,i} H_\sigma, \qquad M^\mu_{d_\sigma} = M^\mu_d H_\sigma, \qquad \sigma \in S_t, \quad 1 \le i \le N, \quad 1 \le \mu \le 2. \tag{3.1}$$

It follows from (2.7) and (3.1) that

$$C_{d_\sigma\mu\nu} = H'_\sigma C_{d\mu\nu} H_\sigma, \qquad \sigma \in S_t, \quad 1 \le \mu, \nu \le 2,$$

$$C_{d_\sigma}(\tau, \rho) = (I_2 \otimes H'_\sigma) C_d(\tau, \rho)(I_2 \otimes H_\sigma), \qquad \sigma \in S_t, \tag{3.2}$$

$$T^\mu_{\sigma\tau} = T^\mu_\tau H_\sigma, \qquad \sigma \in S_t, \quad 1 \le \mu \le 2, \tag{3.3}$$

and, with $C_d(\tau)$ the information matrix of treatment effects [see Shah and Sinha (1989), page 4],

$$C_{d_\sigma}(\tau) = H'_\sigma C_d(\tau) H_\sigma, \qquad \sigma \in S_t. \tag{3.4}$$

3.2. *The convex combination of designs.* For $1 = \sum_{i=1}^I \alpha_i$, $\alpha_i \ge 0$, $1 \le i \le I$, the convex combination of the I designs $d_i \leftrightarrow P^i = (p^i_\tau)$, $1 \le i \le I$, is $d_\alpha \leftrightarrow$

$P^{\alpha} = (p_{\tau}^{\alpha}) = \sum_{i=1}^{I} \alpha_i P^i$, where $p_{\tau}^{\alpha} = \sum_{i=1}^{I} \alpha_i p_{\tau}^i$. We write $d_{\alpha} = \sum_{i=1}^{I} \alpha_i d_i$. The d_{α} play an important part in the sequel. However, they may not exist in exact design theory. For example, $(1,2)$ and $(2,1)$ are two sequences of integers whose convex sum, $(1,2)/2 + (2,1)/2 = (3/2, 3/2)$, is not a sequence of integers.

3.3. *Concavity of $C_d(\tau)$ as a function of d.*

LEMMA 3.1 [Pukelsheim (1993), pages 74–77]. *A^s, the Schur complement of A, is a concave, nonincreasing function of $A \geq 0$.*

Let ξ denote all the fixed effects in (1.1) and let $C_d(\xi)$ denote their information matrix.

THEOREM 3.2. *$C_d(\tau)$ is a concave function of d:*

$$C_{d_{\alpha}}(\tau) \geq \sum_{i=1}^{I} \alpha_i C_{d_i}(\tau). \tag{3.5}$$

PROOF. From $C_d(\xi) = N \sum_{\tau} p_{\tau} B_{\tau}$, where B_{τ} are fixed matrices,

$$C_{d_{\alpha}}(\xi) = \sum_{i=1}^{I} \alpha_i C_{d_i}(\xi). \tag{3.6}$$

From (3.6), Lemma 3.1 and $C_d(\tau) = C_d^s(\xi)$ for a suitable Schur complement, we get (3.5). □

3.4. *Optimality of convex combinations of designs.* Let Φ denote a concave information function, defined on the set of symmetric nonnegative matrices, of the type standard in optimality theory: d is Φ-optimal if it maximizes $\Phi(C_d(\tau))$. If Φ is strictly concave, d is strictly Φ-optimal.

LEMMA 3.3. (i) *The set of Φ-optimal designs is convex: if d_i, $1 \leq i \leq I$, are Φ-optimal designs, then so is d_{α}. In particular, if d is Φ-optimal, then so is $d_{\alpha} = \sum_{\sigma \in S_t} \alpha_{\sigma} d_{\sigma}$, $1 = \sum_{\sigma \in S_t} \alpha_{\sigma}$, $\alpha_{\sigma} \geq 0$, $\sigma \in S_t$.*

(ii) *$C_d(\tau) = \beta(tI_t - J_t)$, β a scalar, for every strictly Φ-optimal d.*

PROOF. (i) Equation (3.5) and the concavity of Φ.

(ii) Equation (3.4) and Kiefer's (1975), Proposition 1 Remark. □

NOTES. (a) In a different problem, Silvey [(1980), page 17] notes the convexity of the set of optimal designs.

(b) Part (ii) also holds in exact design theory.

3.5. *Symmetric optimal designs.* A symmetric design generalizes the notion of a dual balanced design, introduced for $t = 2$ by Laska and Meisner

(1985). Design d is symmetric if

$$d = d_\sigma, \qquad \sigma \in S_t, \tag{3.7}$$

that is, if the proportions of d satisfy

$$p_{\sigma\tau} = p_\tau, \qquad \sigma \in S_t, \quad \text{all } \tau. \tag{3.8}$$

For a model without fixed subject effects, Laska and Meisner (1985) proved that an optimal dual balanced design exists; for model (1.1), see Matthews (1987).

THEOREM 3.4. *A symmetric Φ-optimal design exists.*

PROOF. Use Lemma 3.3(i) with $\alpha_\sigma = 1/t!$, $\sigma \in S_t$. The proportions in the Φ-optimal design

$$P_S = \left(\sum_{\sigma \in S_t} P_\sigma \right) \Big/ t! \tag{3.9}$$

satisfy (3.8). □

We call $d_S \leftrightarrow P_S$ the symmetrized design of $d \leftrightarrow P$.

4. The information matrix of strictly Φ-optimal and universally optimal designs. Here we derive a formula for the treatment effects information matrix of a design which is either strictly Φ-optimal or universally optimal. Important roles in the analysis are played by the quadratic function associated to a treatment sequence, the quadratic function of the design and by symmetric and "symmetrized" designs. We briefly discuss the efficiency of nonoptimal designs.

4.1. *The quadratic function associated to a treatment sequence.* For each treatment sequence, τ, define a nonnegative quadratic, $q_\tau(s)$, by

$$\begin{aligned} q_\tau(s) &= \operatorname{tr}\left[\left(\hat{T}_\tau^1 + s\hat{T}_\tau^2\right)' C^{\sim} \left(\hat{T}_\tau^1 + s\hat{T}_\tau^2\right)\right] \\ &= q_{11}^\tau + 2q_{12}^\tau s + q_{22}^\tau s^2, \qquad -\infty < s < \infty, \end{aligned} \tag{4.1}$$

where

$$q_{\mu\nu}^\tau = \operatorname{tr}\left[\left(\hat{T}_\tau^\mu\right)' C^{\sim} \hat{T}_\tau^\nu\right], \qquad 1 \le \mu, \nu \le 2, \quad \left(q_{12}^\tau = q_{21}^\tau\right). \tag{4.2}$$

NOTE. Equation (4.1) defines a genuine quadratic. For,

$$q_{22}^\tau = \operatorname{tr}\left[\left(\hat{T}_\tau^2\right)' C^{\sim} \hat{T}_\tau^2\right] = 0$$

implies $0 = (\hat{T}_\tau^2)' C^{\sim} \hat{T}_\tau^2$ and then, from Pukelsheim [(1993), page 15] and (2.4), $0 = C^{\sim} \hat{T}_\tau^2$, and $1_p x' = \hat{T}_\tau^2 = T_\tau^2 - J_{p,t}^2/t$, where x, a t-dimensional column vector, must vanish since both T_τ^2 and $J_{p,t}^2$ have zero first rows. This gives the contradictory $0 = tT_\tau^2 - J_{p,t}^2$. Both $q_{11}^\tau = 0$ and $q_{11}^\tau q_{22}^\tau - (q_{12}^\tau)^2 = 0$ can and do occur.

LEMMA 4.1. *The quadratics corresponding to treatment sequences τ and $\sigma\tau$ are identical*:

$$q_{\sigma\tau}(s) = q_\tau(s), \qquad \sigma \in S_t.$$

PROOF. This follows from (3.3) and (2.11). □

Let $C^{-1} = (c^{jk})$, $1 \le j, k \le p$, so that

$$C^{\sim} = (\tilde{c}^{jk}) = \left(c^{jk} - c^j c^k / c^{\text{tot}}\right), \qquad 1 \le j, k \le p, \tag{4.3}$$

where c^j are the row sums, and $c^{\text{tot}} = \sum_{j=1}^{p} \sum_{k=1}^{p} c^{jk}$ is the total sum, of C^{-1}. Let

$$\delta_{\mu\nu}(\tau) = \left(\delta_{t_{j+1-\mu}}^{t_{k+1-\nu}}\right), \qquad 1 \le j, k \le p, \quad 1 \le \mu, \nu \le 2$$

(where the symbol t_0 is left undefined but $0 \equiv \delta_{t_0}^{t_0} = \delta_m^{t_0}$, $1 \le m \le t$). A convenient computational formula for $q_{\mu\nu}^{\tau}$ is

$$q_{\mu\nu}^{\tau} = \operatorname{tr}\left[\delta_{\mu\nu}(\tau) C^{\sim}\right] - \tilde{c}^{11} \delta_{\mu\nu}^{22} / t, \qquad 1 \le \mu, \nu \le 2.$$

4.2. *$Q(s, P)$, the quadratic of a design. Information matrix of a symmetric design.* For any array of proportions, P, define the quadratic of a design, $Q(s, P)$, and its coefficients, $q_{\mu\nu}(P)$ $[q_{12}(P) = q_{21}(P)]$, by

$$Q(s, P) = \sum_{\tau} p_\tau q_\tau(s) = q_{11}(P) + 2q_{12}(P)s + q_{22}(P)s^2,$$

so that, from (2.15) and (4.2),

$$q_{\mu\nu}(P) = \operatorname{tr} \hat{L}_{d\mu\nu}, \qquad 1 \le \mu, \nu \le 2. \tag{4.4}$$

LEMMA 4.2. *For any design, P, and its symmetrized design, P_S, $Q(s, P) = Q(s, P_S)$.*

PROOF. Use Lemma 4.1 and (3.9). □

LEMMA 4.3. *Suppose that $d \leftrightarrow P$ is a symmetric design. Then*

$$C_d(\tau, \rho) = N \begin{pmatrix} q_{11}(P) & q_{12}(P) \\ q_{21}(P) & q_{22}(P) \end{pmatrix} \otimes (tI_t - J_t)/t(t-1), \tag{4.5}$$

$$\begin{aligned} C_d(\tau) &= N\left(q_{11}(P) - q_{12}^2(P)/q_{22}(P)\right)(tI_t - J_t)/t(t-1) \\ &= N \min_{-\infty < s < \infty} Q(s, P)(tI_t - J_t)/t(t-1), \end{aligned} \tag{4.6}$$

$$\hat{M}_d^{\mu} = 0, \qquad 1 \le \mu \le 2. \tag{4.7}$$

PROOF. If d is symmetric, then (3.1) and (3.7) imply that

$$M_d = X \otimes 1_t' \quad \text{where } X \text{ is a } p \times 2 \text{ matrix.} \tag{4.8}$$

But from (2.1), $M_d 1_t = [1_p, 1_p^2]$. Hence, $X = [1_p, 1_p^2]/t$, which, with (4.8), implies (4.7). Similarly, Lemma 2.3(ii), (3.2) and (3.7) imply

$$(4.9)\quad C_d(\tau,\rho) = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \otimes (tI_t - J_t), \quad d \text{ symmetric}, \ \alpha_{\mu\nu} \text{ scalars}.$$

But from (2.14) and (4.7),

$$(4.10)\qquad C_d(\tau,\rho) = N\hat{L}_d, \quad d \text{ symmetric}.$$

Hence, (4.5) follows from (4.4), (4.9)–(4.10) and

$$t(t-1)\alpha_{\mu\nu} = \operatorname{tr} C_{d\mu\nu} = N \operatorname{tr} \hat{L}_{d\mu\nu}, \qquad 1 \le \mu, \nu \le 2, \quad d \text{ symmetric}.$$

Equation (4.6) is obtained by taking Schur complements of (4.5). □

4.3. *The maximum of* $\Phi(C_d(\tau))$, *for any* Φ. *The information matrix of a strictly* Φ-*optimal* (*resp. universally optimal*) *design.*

THEOREM 4.4. *Let*

$$(4.11)\qquad b = \max_P \min_{-\infty < s < \infty} Q(s,P).$$

Then

(i) *for any* Φ,

$$(4.12)\qquad \max_d \Phi(C_d(\tau)) = Nb\Phi(tI_t - J_t)/t(t-1).$$

(ii) d *is* Φ-*optimal for* Φ *strictly concave* (*resp. universally optimal*) *if and only if its treatment effects information matrix is*

$$(4.13)\qquad C_d(\tau) = Nb(tI_t - J_t)/t(t-1).$$

PROOF. From Theorem 3.4, Lemma 4.2 and (4.6), for any Φ,

$$(4.14)\qquad \begin{aligned} \max_d \Phi(C_d(\tau)) &= \max_{d \text{ symmetric}} \Phi(C_d(\tau)) \\ &= N\Phi(tI_t - J_t) \max_P \min_{-\infty<s<\infty} Q(s,P)/t(t-1), \end{aligned}$$

which is (4.12).

Necessity of (4.13): If Φ is strictly concave, (4.14) also shows that, in Lemma 3.3(ii), $\beta = Nb/t(t-1)$. If d is universally optimal, then, of course, it is optimal for the strictly concave Φ's, so (4.13) is also true in this case.

Sufficiency of (4.13): this follows from (4.12). □

4.4. *Minimax and quadratics.* The study of the b in (4.11) is the subject of this section. We present an analysis for any quadratics, rather than just for the quadratics (4.1). One can, for example, apply these results to the quadratics obtained from (4.1) with C^{-1} in place of $C^{\sim}$, which occur in Laska and Meisner (1985). Suppose that $B_i(x)$, $1 \le i \le I$, are nonnegative, nonconstant quadratics in x: $B_i(x) = a_i + b_i x + c_i x^2 \ge 0$, $-\infty < x < \infty$, $c_i > 0$, $1 \le i \le I$.

Define the scalars a, b and the function $B(x)$ by

$$B(x) = \max_{1 \le i \le I} \{B_i(x)\}, \qquad -\infty < x < \infty,$$

$$b = \min_{-\infty < x < \infty} B(x), \tag{4.15}$$

$$B(a) = \min_{-\infty < x < \infty} \max_{1 \le i \le I} \{B_i(x)\} = b. \tag{4.16}$$

The quantity a is uniquely defined by (4.16), since $B_i(x)$ is a convex function and $B(x)$, the maximum of convex functions, is convex and is not constant on any interval. [For simplicity, we use the same symbol "b" in (4.11) and (4.15). When the quadratics are (4.1), Theorem 4.5 shows that the two b's are identical.] Let $\mathscr{N} \subset \{1, 2, \dots, I\}$ be the set of indices for which

$$b = B(a) = B_i(a), \qquad i \in \mathscr{N}. \tag{4.17}$$

In a neighborhood of $x = a$, $B(x)$ depends only on the quadratics whose indices are in $\mathscr{N}$: for $\delta > 0$ sufficiently small, $B(x) = \max_{i \in \mathscr{N}} \{B_i(x)\}$, $a - \delta < x < a + \delta$. With S_I the simplex in Section 3.2, define

$$B(x, \alpha) = \sum_{i=1}^{I} \alpha_i B_i(x), \qquad -\infty < x < \infty, \quad \alpha = (\alpha_1, \alpha_2, \dots, \alpha_I) \in S_I.$$

THEOREM 4.5. (i) *$B(x, \alpha)$ has the minimax property*

$$\max_{\alpha \in S_I} \min_{-\infty < x < \infty} B(x, \alpha) = \min_{-\infty < x < \infty} \max_{\alpha \in S_I} B(x, \alpha) = b. \tag{4.18}$$

(ii) *The two equations*

$$B(x, \alpha) = b \tag{4.19a}$$

and

$$B'(x, \alpha) = \sum_{i=1}^{I} \alpha_i B_i'(x) = 0 \tag{4.19b}$$

are simultaneously satisfied if and only if

$$x = a, \tag{4.20}$$

$$\alpha_i = 0, \qquad i \notin \mathscr{N}, \tag{4.21a}$$

$$1 = \sum_{i \in \mathscr{N}} \alpha_i, \qquad \alpha_i \ge 0, \quad i \in \mathscr{N}, \tag{4.21b}$$

$$B'(a, \alpha) = \sum_{i \in \mathscr{N}} \alpha_i B_i'(a) = 0. \tag{4.22}$$

Moreover, a solution of (4.21)–(4.22) *exists with either one or two nonzero α_i's.*

PROOF. (i) It is obvious that

$$\max_{\alpha \in S_I} \min_{-\infty < x < \infty} B(x, \alpha) \le \min_{-\infty < x < \infty} \max_{\alpha \in S_I} B(x, \alpha)$$
$$= \min_{-\infty < x < \infty} \max_{1 \le i \le I} B_i(x) = \min_{-\infty < x < \infty} B(x) = b.$$

To prove (4.18), we show that

$$\min_{-\infty < x < \infty} B(x, \alpha) = b \quad \text{for some } \alpha \in S_I. \tag{4.23}$$

Case 1. There is an $i^* \in \mathcal{N}$ such that $B'_{i^*}(a) = 0$. Then $\min_{-\infty < x < \infty} B_{i^*}(x) = B_{i^*}(a) = b$. Hence, (4.23) holds for $\alpha_i^* = \delta_i^{i^*}$.

Case 2. There are $i_1, i_2 \in \mathcal{N}$ such that $B'_{i_1}(a) > 0$ and $B'_{i_2}(a) < 0$. Then for

$$\begin{aligned} \alpha_i^* &= -B'_{i_2}(a) / \big(B'_{i_1}(a) - B'_{i_2}(a)\big), \quad \text{if } i = i_1, \\ &= B'_{i_1}(a) / \big(B'_{i_1}(a) - B'_{i_2}(a)\big), \quad \text{if } i = i_2, \\ &= 0, \quad \text{otherwise,} \end{aligned} \tag{4.24}$$

we get $B'(a, \alpha^*) = 0$ and therefore (4.23) holds for the α^* in (4.24). The solutions of (4.22) given in cases 1 and 2 have at most two components of α nonzero.

Case 3. The derivatives $B'_i(a)$, $i \in \mathcal{N}$, have the same sign. This case cannot occur because $B(x)$ would be strictly increasing or decreasing in a neighborhood of $x = a$ and could not attain the minimum (4.23) there.

(ii) For an x and a satisfying (4.19),

$$b = \sum_{i=1}^{I} \alpha_i B_i(x) \le \sum_{i=1}^{I} \alpha_i B_i(a) = \sum_{i \in \mathcal{N}} \alpha_i B_i(a) + \sum_{i \notin \mathcal{N}} \alpha_i B_i(a). \tag{4.25}$$

But

$$b > B_i(a), \qquad i \notin \mathcal{N}, \tag{4.26}$$

and if (4.21a) were not true, (4.17), (4.25) and (4.26) would give a contradiction. If (4.20) were not true, then (4.17) and (4.25) with (4.21) inserted gives the contradictory

$$b = \sum_{i \in \mathcal{N}} \alpha_i B_i(x) < \sum_{i \in \mathcal{N}} \alpha_i B_i(a) = b.$$

Hence (4.20) is true and (4.22) results from (4.19b). □

The concluding result of this section is an algorithm for calculating a, b and the set $\mathcal{N}$. Let $1 \le i, j \le I$. An intersection point of $B_i(x)$ and $B_j(x)$, that is, a point (x_{ij}, y_{ij}) such that

$$y_{ij} = B_i(x_{ij}) = B_j(x_{ij}),$$

is called *admissible* if $B'_i(x_{ij}) \le 0$ and $B'_j(x_{ij}) \ge 0$. In this definition, we do not exclude $i = j$, in which case $(x_{ii}, y_{ii}) = (-b_i/2c_i, a_i - b_i^2/4c_i)$ is still called an admissible "intersection" point.

LEMMA 4.6. (i) *$B_i(x)$ and $B_j(x)$ have at most one admissible intersection point.*

(ii)

$$b = \max_{\substack{1 \le i \le j \le I \\ (x_{ij}, y_{ij}) \\ \text{admissible}}} y_{ij}$$

(iii) *$a = x_{mn}$, where (m, n) is any pair such that $b = y_{mn}$.*

INDICATION OF PROOF. Use $\max_{\alpha \in S_I} B(x, \alpha) = \max_{1 \le i \le I} B_i(x)$ and (4.18).

For the set of quadratics $\{q_\tau(s)\}$, define a, b and $q(s)$ by

$$q(s) = \max_{\tau} \{q_\tau(s)\}, \qquad -\infty < s < \infty, \tag{4.27}$$

$$b = \min_{-\infty < s < \infty} q(s), \tag{4.28}$$

$$q(a) = \min_{-\infty < s < \infty} \max_{\tau} \{q_\tau(s)\} = b. \tag{4.29}$$

Let $\mathscr{T}$ be the set of treatment sequences at which minimax (4.29) is achieved:

$$b = q(a) = q_\tau(a), \qquad \tau \in \mathscr{T}. \tag{4.30}$$

When C is known, a, b and $\mathscr{T}$ can be rapidly determined by Lemma 4.6. From (4.21a), in an optimal design,

$$p_\tau = 0, \qquad \tau \in \mathscr{T}. \tag{4.31}$$

Hence the support of any optimal design is a subset, possibly proper, of $\mathscr{T}$.

4.5. *Design efficiency.* The Φ-efficiency of design d is defined as

$$\text{eff}_\Phi(d) = \Phi(C_d(\tau))/\Phi(C_{d^*}(\tau)), \tag{4.32}$$

where d is any design and d^* is Φ-optimal. The denominator in (4.32) is given by (4.12). If $d \leftrightarrow P$ is symmetric, the numerator in (4.32) is obtained from (4.6), yielding

$$\text{eff}(d) = b_d/b, \quad d \text{ symmetric}, \tag{4.33}$$

where

$$b_d = \min_{-\infty < s < \infty} Q(s, P) = q_{11}(P) - q_{12}^2(P)/q_{22}(P).$$

We omit Φ in (4.33) because, for symmetric d, Φ-efficiency is independent of Φ.

5. The optimality equations. The matrix equation (4.13) can be considered to be a very complicated system of equations for the proportions of a universally optimal (resp. strictly Φ-optimal) design. However, it is possible to reduce (4.13) to a system of linear equations featuring the matrices $\hat{L}_{d\mu\nu}$ and $\hat{M}_d^\mu$. These matrices are much simpler than $C_d(\tau)$. The entries in $C_d(\tau, \rho)$ are quadratic functions of the proportions and those of $\hat{L}_d$ and $\hat{M}_d$ are linear functions of the proportions. In this section, if R is a $2t \times 2t$ matrix, then the $t \times t$ submatrices R_{ij}, $1 \le i, j \le 2$, are defined as in (2.5). R^+ denotes the Moore–Penrose inverse of R. If w is a scalar, then $w^+ = 1/w$, $w \ne 0$ and $0^+ = 0$.

5.1. *The matrix* $\hat{L}^s_d$.

LEMMA 5.1. *If* $R \geq 0$ *is a* $2t \times 2t$ *matrix, then*

$$\operatorname{tr} R^s \leq \operatorname{tr} R_{11} - (\operatorname{tr} R_{12})^2 (\operatorname{tr} R_{22})^+. \tag{5.1}$$

PROOF. From Pukelsheim [(1993), page 75], for $T = (I_t, X_t)'$, X_t a $t \times t$ matrix,

$$T'RT = R^s + (X_t + R_{22}^+ R_{21})' R_{22} (X_t + R_{22}^+ R_{21}). \tag{5.2}$$

Set $X_t = xI_t$, x a scalar. Then $R^s \leq T'RT = R_{11} + xR_{12} + xR_{21} + x^2 R_{22}$ and

$$\operatorname{tr} R^s \leq \operatorname{tr} R_{11} + 2x \operatorname{tr} R_{12} + x^2 \operatorname{tr} R_{22}, \quad \text{all } x. \tag{5.3}$$

In (5.3), set $x = -(\operatorname{tr} R_{12})(\operatorname{tr} R_{22})^+$ to obtain (5.1). □

LEMMA 5.2. (i) *For any* d,

$$C_d(\tau, \rho) \leq N\hat{L}_d. \tag{5.4}$$

(ii) *If* d *is optimal, then*

$$C_d(\tau) = N\hat{L}^s_d. \tag{5.5}$$

PROOF. (i) This follows from (2.14).
(ii) From (5.4),

$$C_d(\tau) = C^s_d(\tau, \rho) \leq N\hat{L}^s_d. \tag{5.6}$$

From (4.4), Lemma 5.1 and (4.13),

$$N \operatorname{tr} \hat{L}^s_d \leq N\big(q_{11}(P) - q_{12}^2(P)/q_{22}(P)\big) = \operatorname{tr} C_d(\tau). \tag{5.7}$$

Equation (5.5) follows from (5.6) and (5.7). □

5.2. *First form of the optimality equations.*

THEOREM 5.3. *With* a, b *and* $\mathscr{T}$ *given in* (4.28)–(4.30), $\hat{T}^\mu_\tau, \hat{M}^\mu_d$ *in* (2.11)–(2.12) *and* $C^\sim$ *in* (2.3), d *is a universally optimal* (*resp. strictly* Φ-*optimal*) *design in approximate design theory iff the treatment sequence proportions*, p_τ, *satisfy the linear equations*

$$\sum_{\tau \in \mathscr{T}} p_\tau \Big(\big(\hat{T}^1_\tau\big)' C^\sim \hat{T}^1_\tau + a \big(\hat{T}^1_\tau\big)' C^\sim \hat{T}^2_\tau \Big) = b(tI_t - J_t)/t(t-1), \tag{5.8}$$

$$\sum_{\tau \in \mathscr{T}} p_\tau \Big(\big(\hat{T}^2_\tau\big)' C^\sim \hat{T}^1_\tau + a \big(\hat{T}^2_\tau\big)' C^\sim \hat{T}^2_\tau \Big) = 0, \tag{5.9}$$

$$C^\sim \bigg(\sum_{\tau \in \mathscr{T}} p_\tau \Big(\hat{T}^1_\tau + a\hat{T}^2_\tau \Big) \bigg) = 0, \tag{5.10}$$

$$p_\tau = 0, \qquad \textit{if } \tau \notin \mathscr{T}. \tag{5.11}$$

PROOF. With $\hat{L}_{d\mu\nu}$ given in (2.15), we will prove (5.8)–(5.10) in the form

$$\hat{L}_{d11} + a\hat{L}_{d12} = b(tI_t - J_t)/t(t-1), \tag{5.8$'$}$$

$$\hat{L}_{d21} + a\hat{L}_{d22} = 0, \tag{5.9$'$}$$

$$C^{\sim}\left(\hat{M}_d^1 + a\hat{M}_d^2\right) = 0. \tag{5.10$'$}$$

Let f be a symmetric optimal design and $g = d/2 + f/2$, an optimal design from Lemma 3.3(i). Set $A = \hat{L}_f$, $B = \hat{L}_d$ and $D = \hat{L}_g = A/2 + B/2$. From (5.2) for $R = A, B$ and D, (4.13) and (5.5),

$$\begin{aligned}(X_t + A_{22}^+ A_{21})' A_{22}(X_t + A_{22}^+ A_{21})\\ +(X_t + B_{22}^+ B_{21})' B_{22}(X_t + B_{22}^+ B_{21})\\ = (X_t + D_{22}^+ D_{21})' D_{22}(X_t + D_{22}^+ D_{21}), \quad \text{for all } X_t.\end{aligned} \tag{5.12}$$

Setting $X_t = -D_{22}^+ D_{21}$ in (5.12), we obtain $0 = Y'A_{22}Y + W'B_{22}W$, where $Y = -D_{22}^+ D_{21} + A_{22}^+ A_{21}$ and $W = -D_{22}^+ D_{21} + B_{22}^+ B_{21}$, implying $0 = Y'A_{22}Y = W'B_{22}W$, and from Pukelsheim [(1993), page 15],

$$0 = A_{22}Y = B_{22}W. \tag{5.13}$$

From (2.13), the row and column sums of $A_{\mu\nu}$, $B_{\mu\nu}$ and $D_{\mu\nu}$, $1 \le \mu, \nu \le 2$, vanish and, therefore, so do those of Y. From (4.10) and Lemma 4.3, $A_{\mu\nu} = q_{\mu\nu}(P)(tI_t - J_t)/t(t-1)$, $1 \le \mu, \nu \le 2$, where $f \leftrightarrow P$, and from Section 4.1, $q_{22}(P) \ne 0$. Hence, the first equality in (5.13) implies $0 = Y = -D_{22}^+ D_{21} - a(I_t - J_t/t)$ and then the second equality in (5.13) implies

$$W = a(I_t - J_t/t) + B_{22}^+ B_{21}. \tag{5.14}$$

Equations (5.13)–(5.14) give (5.9)$'$: $0 = B_{21} + aB_{22} = \hat{L}_{d21} + a\hat{L}_{d22}$. Equation (5.8)$'$ follows from (4.13), (5.5) and (5.9)$'$:

$$\begin{aligned} b(tI_t - J_t)/t(t-1) &= N^{-1}C_d(\tau) = \hat{L}_{d11} - \hat{L}_{d12}\hat{L}_{d22}^+\hat{L}_{d21}\\ &= \hat{L}_{d11} + a\hat{L}_{d12}\hat{L}_{d22}^+\hat{L}_{d22}\\ &= \hat{L}_{d11} + a\hat{L}_{d12}.\end{aligned}$$

Finally, from (4.13), (5.3) with $R = C_d(\tau, \rho)$, (2.14) and (4.4),

$$\begin{aligned} Nb = \operatorname{tr} C_d(\tau) &\le \operatorname{tr}\left[C_{d11} + 2xC_{d12} + x^2 C_{d22}\right]\\ &= NQ(x, P) - N\operatorname{tr}\left[\left(\hat{M}_d^1 + x\hat{M}_d^2\right)' C^{\sim}\left(\hat{M}_d^1 + x\hat{M}_d^2\right)\right],\end{aligned} \tag{5.15}$$

and setting $x = a$ in (5.15), gives $b \le b - \operatorname{tr}[(\hat{M}_d^1 + a\hat{M}_d^2)' C^{\sim}(\hat{M}_d^1 + a\hat{M}_d^2)]$. Hence, $0 = \operatorname{tr}[(\hat{M}_d^1 + a\hat{M}_d^2)' C^{\sim}(\hat{M}_d^1 + a\hat{M}_d^2)]$, equivalent to (5.10)$'$. □

5.3. *Second form of the optimality equations.* Define treatment-period proportion sums by $P_m^j = \sum_\tau p_\tau \delta_{t_j}^m = \sum_{t_j = m} p_\tau$, $P_{mn}^{jk} = \sum_\tau p_\tau \delta_{t_j}^m \delta_{t_k}^n = \sum_{(t_j, t_k) = (m, n)} p_\tau$, $1 \le j, k \le p$, $1 \le m, n \le t$,and treatment proportion sums by $P_m = \sum_{j=1}^p P_m^j$, $1 \le m \le t$. Set $P_{mn}^{j0} = P_{mn}^{0j} = P_{mn}^{00} = P_m^0 = 0$, $1 \le j \le p$, $1 \le m, n \le t$.

THEOREM 5.4. *The optimality equations in Theorem* 5.3 *are equivalent to* (5.11) *and the following linear system*:

$$b(t\delta_{mn}-1)/t^2(t-1)+(a/t)w_m u = \sum_{j,k=1}^{p}\left(P_{mn}^{j,k}+aP_{mn}^{j,k-1}\right)\tilde{c}^{jk}, \qquad 1\le m,n\le t,$$

$$-w_m u/t + a\tilde{c}^{11}/t^2 = \sum_{j,k=1}^{p}\left(P_{mn}^{j-1,k}+aP_{mn}^{j-1,k-1}\right)\tilde{c}^{jk}, \qquad 1\le m,n\le t,$$

$$P_m^j = w_m s_j + 1/t, \qquad 1\le m\le t,\ 1\le j\le p,$$

where a, b *and* $\mathscr{T}$ *are as given in Theorem* 5.3 *and* $\tilde{c}^{jk}$ *in* (4.3); $s_j=(1-(-a)^j)/(1+a)$, $1\le j\le p$; $w_m=w(P_m-p/t)$, $1\le m\le t$; $w=(1+a)^2/(p+a(1+p-(-a)^p))$; $u=(1+a)^{-1}\sum_{j=1}^{p}(-a)^j\tilde{c}^{j1}$. (*If* $a=-1$, *formulas for* s_j *and* u *are obtained by taking limits.*)

INDICATION OF PROOF. Sum (5.8)–(5.10) using $T_\tau^\mu=(\delta^m_{t_{j+1-\mu}})$, $1\le j\le p$, $1\le m\le t$ for $1\le\mu\le 2$. □

5.4. *Symmetry blocks, inequivalent treatment sequences, symmetric and symmetrized designs.* The "symmetry block," $\langle\kappa\rangle$, a set of treatment sequences, is defined by $\langle\kappa\rangle=\{\tau:\ \tau=\sigma\kappa,\ \sigma\in S_t\}=$ the orbit of κ under S_t. $\mathrm{card}\langle\kappa\rangle=(t)_k=t(t-1)\cdots(t-k+1)$, where k = number of distinct integers in $\kappa=(k_1,k_2,\dots,k_p)$. Call κ_1 equivalent to κ_2 iff $\kappa_1=\sigma\kappa_2$, $\sigma\in S_t$. Let κ run over a set of inequivalent treatment sequences. Then

$$\text{all } t^p \text{ treatment sequences} = \bigcup_{\kappa}\langle\kappa\rangle$$

is a disjoint union. In a symmetric design, $P=(p_\tau)$, proportions are the same for every treatment sequence in a symmetry block:

(5.16a) $$p_\tau=p_\kappa, \quad \text{if } \tau\in\langle\kappa\rangle,$$

(5.16b) $$1=\sum_{\kappa}\mathrm{card}\langle\kappa\rangle p_\kappa, \qquad p_\kappa\ge 0,$$

where the κ in (5.16b)—and in subsequent sums in this section—runs over a set of inequivalent treatment sequences. In any design, P, let

(5.17a) $$P_\kappa=\sum_{\tau\in\langle\kappa\rangle}p_\tau$$

denote the sum of proportions over symmetry block $\langle\kappa\rangle$. Thus

(5.17b) $$1=\sum_{\kappa}P_\kappa.$$

The proportions of the symmetrized design (3.9) are given by

(5.18) $$p_\tau^S=P_\kappa/\mathrm{card}\langle\kappa\rangle, \quad \text{if } \tau\in\langle\kappa\rangle.$$

5.5. *A condition necessary for an optimal design and necessary and sufficient for a symmetric optimal design.*

THEOREM 5.5. (i) *The equation*

$$0 = \sum_{\kappa \in \mathscr{T}} P_\kappa q'_\kappa(a) \tag{5.19}$$

is a necessary condition that d be universally optimal, where P_κ is given in (5.17a) *and the κ in* (5.19) *runs over inequivalent treatment sequences in $\mathscr{T}$.*
(ii) *A symmetric design is optimal iff* (5.11) *and* (5.19) *hold.*

PROOF. Lemma 4.2, Theorem 4.4(i) and Theorem 4.5 imply (5.19). If d is symmetric and (5.19) holds, then (4.6) and Theorem 4.5 imply that d is optimal. □

NOTE. We write simply "optimal" in Theorem 5.5 because it is true for any Φ.

COROLLARY 5.6. *If $\mathscr{T}$ contains two or more inequivalent treatment sequences, then there exists a universally optimal symmetric design, consisting of just two symmetry blocks, $\langle \kappa_1 \rangle$ and $\langle \kappa_2 \rangle$, with sum of proportions P_{κ_i}, $1 \le i \le 2$, given by*

$$\begin{aligned} P_{\kappa_1} &= -q'_{\kappa_2}(a)/\big(q'_{\kappa_1}(a) - q'_{\kappa_2}(a)\big), \\ P_{\kappa_2} &= q'_{\kappa_1}(a)/\big(q'_{\kappa_1}(a) - q'_{\kappa_2}(a)\big) \end{aligned} \tag{5.20}$$

and proportions $p_{\kappa_i} = P_{\kappa_i}/(t)_{\kappa_i}$. If all treatment sequences in $\mathscr{T}$ are equivalent, that is, if $\mathscr{T} = \langle \kappa \rangle$, then there exists a universally optimal symmetric design with common proportion $p_\kappa = 1/(t)_k$.

PROOF. Equation (5.20) results from (5.19) when just two P_κ's are nonzero. If $\mathscr{T}$ consists of two symmetry blocks, then the derivatives in (5.20) are automatically of opposite sign. Otherwise, choose κ_1 and κ_2 to be any treatment sequences in $\mathscr{T}$ such that $q'_{\kappa_i}(a)$ have opposite signs. □

Acknowledgments. I am deeply grateful to Morris Meisner for many valuable discussions and to Eugene Laska, who suggested the topic of this research. In a number of places in this paper, my debt to their previous research in this field is also recorded. I am especially indebted to a referee, whose long labor resulted in a huge shortening of the paper, substantial gains in clarity and many contributions. Among these are his short proof of Theorem 4.5(i) and the use of the matrix inequality (5.4). He has also shown that Theorem 4.4 can be derived from this inequality and Kiefer's Proposition 1.

REFERENCES

CHENG, C. S. and WU, C. F. (1980). Balanced repeated measurements designs. *Ann. Statist.* **8** 1272–1283. (Correction (1983). *Ann. Statist.* **11** 349.)

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

HEDAYAT, A. and ZHAO, W. (1990). Optimal two-period repeated measurements designs. *Ann. Statist.* **18** 1805–1816.

HUYNH, H. and FELDT, L. S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact *F*-distributions. *J. Amer. Statist. Assoc.* **65** 1582–1589.

KIEFER, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (J. N. Srivistava, ed.). North-Holland, Amsterdam.

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

KUNERT, J. (1985). Optimal repeated measurements designs for correlated observations and analysis by weighted least squares. *Biometrika* **72** 375–389.

KUNERT, J. (1991). Cross-over designs for two treatments and correlated errors. *Biometrika* **78** 315–324.

KUSHNER, H. B. (1996). Optimal and efficient repeated measurements designs for uncorrelated observations. Submitted for publication.

KUSHNER, H. B. (1997). Optimality and efficiency of two-treatment repeated measurements designs. *Biometrika* **84** 455–468.

LASKA, E. and MEISNER, M. (1985). A variational approach to optimal two-treatment crossover designs: application to carryover-effect models. *J. Amer. Statist. Assoc.* **80** 704–710.

MATTHEWS, J. N. S. (1987). Optimal crossover designs for the comparison of two treatments in the presence of carryover effects and auto-correlated errors. *Biometrika* **74** 311–320.

MATTHEWS, J. N. S. (1990). Optimal dual-balanced two treatment crossover designs. *Sankhyā* **52** 332–337.

PUKELSHEIM, F. (1993). *Optimal Design of Experiments*. Wiley, New York.

SHAH, K. R. and SINHA, B. K. (1989). *Theory of Optimal Designs. Lecture Notes in Statist.* **54**. Springer, New York.

SILVEY, S. D. (1980). *Optimal Design. An Introduction to the Theory for Parameter Estimation.* Chapman and Hall, London.

NATHAN S. KLINE INSTITUTE FOR
PSYCHIATRIC RESEARCH
STATISTICAL SCIENCES AND
EPIDEMIOLOGY DIVISION
140 OLD ORANGEBURG ROAD
ORANGEBURG, NEW YORK 10962
E-MAIL: kushner@rfmh.org