The Annals of Statistics
1987, Vol. 15, No. 4, 1604–1628

ON THE OPTIMALITY OF FINITE WILLIAMS II(a) DESIGNS

BY J. KUNERT AND R. J. MARTIN

Universität Trier and University of Sheffield

In this paper, we consider the type II(a) designs of Williams. It was shown, essentially, by Kiefer that the type II(a) designs are asymptotically universally optimum for a first order autoregression with parameter $\lambda > 0$. We concentrate on the stationary first order autoregression with $\lambda > 0$ and the extra plot version of the II(a) designs. Our main results are that the design is D- and A-optimal then, but is not necessarily E-optimal when λ is small.

1. Introduction. In a seminal paper, Williams (1952) considered experimental design when the plots are contiguous and laid out in a line and when the errors or plot effects are assumed correlated according to a first or second order autoregressive process. For a first order autoregression AR(1) with positive lag-one correlation coefficient λ, he looked at II(a) designs in which every treatment occurs equally often next to every other, but never to itself.

Kiefer (1961) showed that for an AR(1) with $0 < \lambda < 1$, the II(a) designs belong to the class of asymptotically optimum designs as n, the number of plots, tends to infinity. The optimality criterion he used was a broad one, including many of the commonly used criteria. Recently, Kiefer and Wynn (1984) have produced conditions for a design to be asymptotically optimum for the general pth order autoregression.

There are, however, no known optimality results for finite n. Cox (1952) and Atkinson (1969) both conjectured A-optimality of II(a) designs for $0 < \lambda < 1$. We restrict our investigation to the type II(a) designs with extra plot originally considered by Williams (1952) and the case of stationary AR(1) errors. We also restrict attention to the case of more than two treatments. Discussion of results on II(a) designs for other forms of AR(1) and the case of two treatments is contained in Kunert and Martin (1987).

In this paper, we show that Cox's and Atkinson's conjecture is true for all $0 < \lambda < 1$ and can be extended to include D-optimality. We also show that an extension to E-optimality is not in general possible for all $0 < \lambda < 1$, but is possible for λ not too small. For a restricted set of designs, we can show a more general optimality of the II(a) designs for all $0 < \lambda < 1$. The results are of importance in illustrating the problems that arise when autocorrelation is present and provide an interesting application of results in Cheng (1987).

Note that we do not consider the case $\lambda < 0$ since the II(a) design is clearly inefficient then. Further discussion of this case is in Kunert and Martin (1987).

Section 2 contains introductory material and formal statements of our results on the II(a) designs. In Section 3, we give some general optimality results that

Received July 1985; revised April 1987.
AMS 1980 *subject classifications*. Primary 62K05; secondary 62K10, 62P10.
Key words and phrases. Autoregression, correlated errors, experimental design, optimal design, φ_p-criteria.

are needed later. Section 4 contains the rather long proof of D-optimality and Section 5 gives the remaining proofs.

2. Definitions. Suppose we have $t \geq 3$ treatments labelled $1, 2, \ldots, t$ and that ϑ_i, $i = 1, \ldots, t$, is the additive parameter associated with the effect of treatment i. We consider contrasts of the ϑ_i of the form $l'\vartheta$, where l is a t-vector with $l'l = 1$ and $l'1_t = 0$, 1_t is the t-vector of 1's and $\vartheta = (\vartheta_i)$. Let $\tau_i = \vartheta_i - \Sigma\vartheta_j/t$, $\tau = (\tau_i)$, and $\hat{\tau}$ be the BLUE for τ. Then $l'\hat{\tau}$ is the BLUE for $l'\vartheta$.

An experimental design prescribes which treatment is to be applied to which plot. The set of all designs with t treatments and n plots is denoted $\Omega_{t,n}$. For data y_i and errors ε_i, $i = 1, \ldots, n$, following a stationary AR(1), we consider the linear model

$$Y = T_d\vartheta + \varepsilon,$$

where T_d is the treatment design matrix. The kth row of T_d has a 1 in the ith column if treatment i is applied to plot k, $1 \leq k \leq n$. The other entries of the kth row are 0. There are no nuisance parameters in the model. Let $\sigma^2 S_\lambda$ be the covariance matrix of ε. We define V_λ as an arbitrary matrix with $V_\lambda' V_\lambda = S_\lambda^{-1}$ and we get

$$V_\lambda Y = V_\lambda T_d \vartheta + e, \tag{1}$$

where the e_i are uncorrelated. Without changing model (1), we can write

$$V_\lambda Y = V_\lambda T_d \tau + V_\lambda 1_n \alpha + e,$$

where $\alpha = \Sigma\vartheta_i/t$ is the overall mean. For every $d \in \Omega_{t,n}$, define the information matrix

$$\begin{aligned}\mathscr{C}_d &= T_d' V_\lambda' \left(I_n - V_\lambda 1_n (1_n' V_\lambda' V_\lambda 1_n)^{-1} 1_n' V_\lambda'\right) V_\lambda T_d \\ &= T_d' \left(S_\lambda^{-1} - \left(1_n' S_\lambda^{-1} 1_n\right)^{-1} S_\lambda^{-1} 1_n 1_n' S_\lambda^{-1}\right) T_d\end{aligned}$$

[see Kiefer (1975)], where I_n is the unit matrix of order n. It is easily shown that $\mathscr{C}_d$ has row and column sums zero [see Kunert (1983)] and that τ is estimable if and only if $\mathscr{C}_d$ has rank $t-1$. The Moore–Penrose generalized inverse $\mathscr{C}_d^+$ of $\mathscr{C}_d$ then is the covariance matrix of the BLUE for τ.

To determine an optimal design, we consider the class of φ_p-criteria [Kiefer (1975)], where $-\infty \leq p \leq \infty$. Let $\mu_{d1} \geq \mu_{d2} \geq \cdots \geq \mu_{d,t-1} \geq 0$ denote the eigenvalues of $\mathscr{C}_d$. Then,

$$\varphi_p(\mathscr{C}_d) = \left((t-1)^{-1} \sum \mu_{di}^{-p}\right)^{1/p}, \qquad p \notin \{-\infty, 0, \infty\},$$

while

$$\varphi_{-\infty}(\mathscr{C}_d) = \lim_{p \to -\infty} \varphi_p(\mathscr{C}_d) = \mu_{d1}^{-1},$$

$$\varphi_0(\mathscr{C}_d) = \lim_{p \to 0} \varphi_p(\mathscr{C}_d) = (\Pi \mu_{di})^{-1/(t-1)}$$

and

$$\varphi_\infty(\mathscr{C}_d) = \lim_{p \to \infty} \varphi_p(\mathscr{C}_d) = \mu_{d,t-1}^{-1}.$$

Then φ_0, φ_1 and φ_∞ are, respectively, the well known D-, A- and E-criteria. A design $d^* \in \Omega_{t,n}$ is said to be φ_p-optimal over a set $\Delta \subset \Omega_{t,n}$ if d^* minimizes $\varphi_p(\mathscr{C}_d)$ over Δ.

For the stationary AR(1), S_λ has (i, j)th entry $\lambda^{|i-j|}/(1-\lambda^2)$ and

$$S_\lambda^{-1} = \begin{pmatrix} 1 & -\lambda & & & & & \\ -\lambda & 1+\lambda^2 & -\lambda & & & 0 & \\ & -\lambda & & \ddots & & & \\ & & \ddots & & \ddots & -\lambda & \\ & 0 & & & -\lambda & 1+\lambda^2 & -\lambda \\ & & & & & -\lambda & 1 \end{pmatrix}.$$

For a II(a) design with extra plot to exist, we must have $n = rt + 1$, where r is an integer multiple of $(t-1)/2$ if t is odd or of $t-1$ if t is even.

Take a fixed but arbitrary design of $d \in \Omega_{t,n}$. Denote by r_{di}, $1 \le i \le t$, the number of times treatment i appears in d and by N_{dij}, $1 \le i \ne j \le t$, the number of times treatment i appears adjacent to j. Let N_{dii}, $1 \le i \le t$, be the number of times treatment i appears next to itself. Further, denote by a_{di}, $1 \le i \le t$, the number of times treatment i appears at an end plot. Then $a_{di} \in \{0, 1, 2\}$. The ith diagonal element of $\mathscr{C}_d$ then equals

$$c_{dii} = r_{di}(1+\lambda^2) - a_{di}\lambda^2 - 2N_{dii}\lambda - \frac{1-\lambda}{n-(n-2)\lambda}\{r_{di}(1-\lambda) + a_{di}\lambda\}^2,$$

while the (i, j)th element equals

$$c_{dij} = -N_{dij}\lambda - \frac{1-\lambda}{n-(n-2)\lambda}\{r_{di}(1-\lambda) + a_{di}\lambda\}\{r_{dj}(1-\lambda) + a_{dj}\lambda\}.$$

A design d^* for which

1. $r_{d^*1} = r + 1$, $r_{d^*2} = \cdots = r_{d^*t} = r$,
2. $N_{d^*ij} = 2r/(t-1)$, $1 \le i \ne j \le t$,
3. $N_{d^*ii} = 0$, $1 \le i \le t$,

is called a Williams II(a) design. Note that this generalizes Williams' (1952) original definition since we do not insist on there being r complete blocks. Since for any design $d \in \Omega_{t,n}$, the φ_p-criterion does not change when the treatments are relabelled, we have defined treatment 1 to be the one which appears most often. Then $a_{d^*1} = 2$.

The eigenvalues of the information matrix of a Williams II(a) design $d^* \in \Omega_{t,n}$ are [see Kunert and Martin (1987)]

$$\mu_{d^*1} = r(1+\lambda^2) + \frac{2r\lambda}{t-1} + \frac{r(1-\lambda)^2(1+\lambda)(t-1)}{n-(n-2)\lambda}$$

with multiplicity 1 and

$$\mu_{d^*2} = r(1+\lambda^2) + \frac{2r\lambda}{t-1}$$

with multiplicity $t-2$. Note that $\mu_{d^*1} > \mu_{d^*2}$ for all $0 < \lambda < 1$.

We now state the following results, which will be proved later.

THEOREM 2.1. *For all* $0 < \lambda < 1$, *a Williams* II(a) *design* $d^* \in \Omega_{t,n}$ *is* φ_p*-optimal*, $0 \le p \le \infty$, *over all* $d \in \Omega_{t,n}$ *for which* $N_{d11} \le (t-1)\Sigma N_{dii}$, *where the summation is over all other treatments.*

For the D- and A-criteria, we can show optimality over all designs.

THEOREM 2.2. *For all* $0 < \lambda < 1$, *a Williams* II(a) *design* $d^* \in \Omega_{t,n}$ *is* D- *and* A*-optimal over* $\Omega_{t,n}$.

In the case of E-optimality (if we allow designs with $N_{d11} > 0$), we cannot in general show optimality for all $0 < \lambda < 1$. We only have

THEOREM 2.3. *A Williams* II(a) *design* $d^* \in \Omega_{t,n}$ *is* E*-optimal over* $\Omega_{t,n}$, *provided*

$$\frac{t-1}{t}\,\frac{(1-\lambda)^2(1+\lambda)(n-1)}{n-(n-2)\lambda} \le t(t-2)\lambda,$$

i.e., if λ *is not too near to* 0.

We can show that the bound in Theorem 2.3 is sharp (i.e., cannot be improved upon) by exhibiting designs that are E-better than the Williams II(a) design when the bound is exceeded. For example, take $t = 3$. For given n, we consider a series of designs $d(i)$, $i = 0, \dots, r-1$, for which $r_{d(i)1} = r + 1$, $r_{d(i)2} = \cdots = r_{d(i)t} = r$, $a_{d(i)1} = 2$ and for which $N_{d(i)1j} = r - i$, $j = 2, 3$, $N_{d(i)23} = r + i$ and $N_{d(i)11} = i$. Such designs exist for all $n = 3r + 1$, $r > 1$; see Kunert and Martin (1987). The eigenvalues of $\mathscr{C}_{d(i)}$ are easily found to equal

$$r(1+\lambda^2) + r\lambda - 3i\lambda + 2(1-\lambda)^2(1+\lambda)r/\{n-(n-2)\lambda\},$$
$$r(1+\lambda^2) + r\lambda + i\lambda \quad \text{and} \quad 0.$$

[See Kunert and Martin (1987).] When $\lambda = 0$, these become $r + 2r/n$ and r, so that there are small values of $\lambda > 0$ for which the first eigenvalue of $d(i)$ is bigger than the second. Note that $d(0)$ is the Williams II(a) design. An extension of the proof of Theorem 2.3 shows that there is a sequence $(\lambda_{3ni})_i$ with

$$0 = \lambda_{3nr} < \lambda_{3,n,r-1} < \cdots < \lambda_{3n1} < \lambda_{3n0} = 1$$

such that the design $d(i)$ is E-optimal for $\lambda_{3,n,i+1} \le \lambda \le \lambda_{3ni}$. It immediately follows that the Williams type II(a) design is not E-optimal for $\lambda < \lambda_{3n1}$. Note that λ_{3ni} for fixed i does not tend to 0 as $n \to \infty$. In fact, λ_{3ni} increases (slightly) with n. For instance $\lambda_{3,7,1} \approx 0.174$ and $\lambda_{3n1} \to 0.212$ as $n \to \infty$. Note also that this result does not contradict the asymptotic optimality of the type II(a) design shown by Kiefer (1961) since, for fixed i, $r^{-1}\mathscr{C}_{d(i)} \to (1 + \lambda + \lambda^2)(I_3 - 3^{-1}1_3 1_3')$ as $r \to \infty$. Taking $i = r - 1$ gives the following result, which it is of interest to compare with Kiefer's (1961) Theorem 3.1.1 and the remark following it.

PROPOSITION 2.4. *There exists a sequence of designs* $(d(r))_r$, *where* $d(r) \in \Omega_{3,3r+1}$, *such that* $d(r)$ *is E-optimal over* $\Omega_{3,3r+1}$ *for* $0 \leq \lambda \leq \lambda_r^*$ *and that* $\lim_{r\to\infty} N_{d(r)12}/r = \lim_{r\to\infty} N_{d(r)13}/r = 0$, *while* $\lim_{r\to\infty} N_{d(r)23}/r = 2$.

It should be noted that $\lambda_r^* = \lambda_{3,3r+1,r-1}$ decreases with r and becomes 0 in the limit.

If the number of treatments is more than three, similar results can be obtained. For further details see Kunert and Martin (1987). The example also shows that the condition in Theorem 2.1 is needed, since $\Sigma N_{dii} = 0$, but $N_{d11} > 0$ for the designs which perform better under the E-criterion than the Williams II(a) designs.

We end this section by noting why the optimality proofs are difficult. First, we know that d^* in general is not E-optimal. Second, $\mathscr{C}_{d^*}$ is not completely symmetric. Third, although $\mathscr{C}_{d^*}$ has only two different eigenvalues, it does not have maximal trace. Any design d with $r_{di} = r_{d^*i}$ and $N_{dii} = 0$ for all i, which has $a_{d1} < 2$, will have $\mathrm{tr}(\mathscr{C}_d) > \mathrm{tr}(\mathscr{C}_{d^*})$—see Kunert and Martin (1987) and the following example. This means that we cannot use the theorem of Cheng (1978).

EXAMPLE 2.5. Compare the Williams II(a) design

$$d^* = [1 \quad 2 \quad 1 \quad 3 \quad 2 \quad 3 \quad 1] \in \Omega_{3,7}$$

and the design

$$d = [2 \quad 1 \quad 3 \quad 1 \quad 2 \quad 1 \quad 3] \in \Omega_{3,7}.$$

Then $\mathrm{tr}\,\mathscr{C}_d > \mathrm{tr}\,\mathscr{C}_{d^*}$.

For designs with two different nonzero eigenvalues, there are, roughly speaking, two factors which determine their performance under the φ_p-criteria. One factor is the trace of the information matrix; the other factor is the size of the smaller of the two eigenvalues. If p is increasing from $0 \to \infty$, the size of the smaller eigenvalue becomes more and more important, while the trace can be neglected in the limit.

This is exemplified in the case considered here. The designs which perform better under the φ_∞-criterion than the Williams II(a) designs, all have smaller trace. On the other hand, it was difficult to show that the designs with larger trace were not better under the φ_0-criterion. The φ_1-criterion is intermediate.

3. General results on φ_p-criteria. In this section, we give general results which establish optimality of a design d^* over a subset $\Delta \subset \Omega_{t,n}$. These results concern optimal designs d^* for which $\mathscr{C}_{d^*}$ has at most two different nonzero eigenvalues. The main theorem, Theorem 3.3, is a corollary to results in the paper of Cheng (1987). The most interesting point of this result is that it does not need maximization of $\mathrm{tr}\,\mathscr{C}_d$. In what follows, we always assume that $\mathscr{C}_{d^*}$ has rank $t-1$.

THEOREM 3.1. *Assume* $d^* \in \Omega_{t,n}$ *has the following properties*:

(i) d^* *maximizes* $\operatorname{tr} \mathscr{C}_d$ *over* Δ.

(ii) $\mathscr{C}_{d^*}$ *has only two different nonzero eigenvalues, the greater of which has multiplicity* 1.

(iii) d^* *maximizes* $\operatorname{tr} \mathscr{C}_d - \mu_{d1}$ *over* Δ, *where* μ_{d1} *is the maximal eigenvalue of* $\mathscr{C}_d$, $d \in \Delta$.

Then d^* *is* φ_p*-optimal over* Δ *for all* $p \geq 0$.

PROOF. Assume $p \geq 0$. The φ_p-criterion becomes smaller if one of the eigenvalues is increased and all the others remain fixed. Consequently, a hypothetical design with eigenvalues

$$\tilde{\mu}_{d1} = \mu_{d1} + \operatorname{tr} \mathscr{C}_{d^*} - \operatorname{tr} \mathscr{C}_d \quad \text{and} \quad \mu_{d2}, \dots, \mu_{d,t-1}, 0$$

performs at least as well as d. The convexity of the φ_p-criterion implies that if the design also has

$$\tilde{\mu}_{d2} = \cdots = \tilde{\mu}_{d,t-1} = (\operatorname{tr} \mathscr{C}_d - \mu_{d1})/(t-2) = (\operatorname{tr} \mathscr{C}_{d^*} - \tilde{\mu}_{d1})/(t-2),$$

then the φ_p-value must be at least as small. Since $(\operatorname{tr} \mathscr{C}_{d^*} - \mu_{d1})/(t-2) \leq \tilde{\mu}_{d1}$ for all $d \in \Delta$ and since $\tilde{\mu}_{d1} + \mu_{d2} + \cdots + \mu_{d,t-1} = \operatorname{tr} \mathscr{C}_{d^*}$ for all $d \in \Delta$, the convexity of φ_p implies the optimality of d^*. □

The conditions of Theorem 3.1 are fulfilled if (iii) is replaced by

(iii′) d^* minimizes the maximal eigenvalue μ_{d1} over Δ.

This more restrictive version can be found in Whittinghill (1984).

The conditions of Theorem 3.1 are also fulfilled if (ii) and (iii) are kept but (i) is replaced by

(i′) d^* maximizes μ_{d1} over Δ.

This version will be used in the proof of Theorem 3.3.

Although conditions (i′) and (iii′) appear contradictory, it can easily be seen that both versions imply that the conditions of Theorem 3.1 hold.

THEOREM 3.2 [Jacroux (1985) and Kunert (1985)]. *Assume* $d^* \in \Omega_{t,n}$ *has the following properties*:

(i) d^* *maximizes* $\operatorname{tr} \mathscr{C}_d$ *over* Δ.

(ii) $\mathscr{C}_{d^*}$ *has at most two different nonzero eigenvalues, the greater of which has multiplicity* $t-2$.

(iii) d^* *maximizes* $\mu_{d,t-1}$ *over* Δ.

Then d^* *is* φ_p*-optimal over* Δ *for all* $p \geq 0$.

A proof for Theorem 3.2 can be found in Kunert (1985).

These two theorems need maximization of the trace; that is, we cannot show optimality of a Williams II(a) design over the whole set $\Omega_{t,n}$ with these two theorems only. We thus need results for designs with only two different eigenvalues, which need not have maximal trace. The following theorem is an immediate consequence of results in the paper of Cheng (1987).

THEOREM 3.3. *Consider a fixed* $b \neq -1$ *and a design* $d \in \Delta$. *Let the two nonnegative real numbers* $\mu_1(d)$ *and* $\mu_2(d)$, *where* $\mu_1(d) \geq \mu_2(d)$, *be such that* $\operatorname{tr} \mathscr{C}_d = \mu_1(d) + (t-2)\mu_2(d)$ *and* $\varphi_b(\mathscr{C}_d)$ *equals the* φ_b*-criterion of a hypothetical design with eigenvalues* $\mu_1(d), \mu_2(d), \dots, \mu_2(d)$ *and* 0. *If a design* d^* *exists with the properties*

(i) d^* *is* φ_b*-optimal over* Δ,

(ii) $\mathscr{C}_{d^*}$ *has at most two different nonzero eigenvalues, the greater of which has multiplicity* 1,

(iii) d^* *maximizes* $\mu_2(d)$ *over* Δ,

then d^* *is* φ_p*-optimal over* Δ *for all* $p \geq \max\{b, 0\}$.

PROOF. The case $b = -\infty$ was solved in Theorem 3.1 and the remarks following it. If $b = \infty$, nothing remains to show. Thus assume $|b| < \infty$. Also assume $p < \infty$.

First of all, parts (a) and (b) of Lemma 2.2 of Cheng (1987) show that $\mu_1(d)$ and $\mu_2(d)$ do exist and are unique. Theorem 2.1 of Cheng (1987) directly implies that for all $p \geq \max\{0, b\}$ the φ_p-criterion of a hypothetical design with eigenvalues $\mu_1(d), \mu_2(d), \dots, \mu_2(d)$ and 0 is a lower bound for the φ_p-criterion of the design d. To see this, one can apply the arguments in the proof of Corollary 3.3 of Cheng (1987). For this hypothetical design, the φ_b-criterion is a monotone function of

$$g_b(\mu_1(d)) + (t-2)g_b(\mu_2(d)),$$

where $g_s(x) = -\ln(x)$ if $s = 0$ and $g_s(x) = x^{-s}$ if $s \neq 0$. Similarly, the φ_p-criterion is a monotone function of

$$g_p(\mu_1(d)) + (t-2)g_p(\mu_2(d)).$$

Since $p \geq b$ and $p \geq 0$, there is a convex and monotone function h such that $g_p(x) = h(g_b(x))$ for every x. Consequently, the φ_p-criterion is a monotone function of

$$h(g_b(\mu_1(d))) + (t-2)h(g_b(\mu_2(d)))$$

and, if the φ_b-criterion remains fixed, it is minimal if $g_b(\mu_2(d))$ and $g_b(\mu_1(d))$ are as nearly equal as possible, that is, if $\mu_2(d)$ is maximal.

So we have shown that d^* is φ_p-optimal for all p such that $\max\{0, b\} < p < \infty$. The continuity of the φ_p-criteria in p implies that d^* also is φ_∞-optimal. □

REMARK. A problem in using Theorem 3.3 is that we appear to need knowledge of the φ_b-criterion for every $d \in \Delta$ to determine $\mu_1(d)$ and $\mu_2(d)$. However, very often it suffices to know upper bounds for $\mu_2(d)$. In the case considered here, these bounds will be derived with the help of Theorems 3.1 and 3.2.

4. *D*-optimality of Williams II(a) designs. We now apply our general theorems to the special case of a Williams II(a) design $d \in \Omega_{t,n}$. Note that d^* has the structure of eigenvalues needed in Theorem 3.3.

Now consider an arbitrary design $d \in \Omega_{t,n}$. As $n = rt + 1$, there is at least one treatment which appears $r + 1 + z$ times in the design, where $z \in \{0, 1, \dots, r(t-1)\}$. Let treatment 1 be one of those that appear most often in d. Define $b \in \{0, 1, 2\}$ as the number of end plots of d, where treatment 1 does not appear, i.e., $b = 2 - a_{d1}$. Abbreviate the number N_{d11} of direct adjacencies of treatment 1 to itself by m. These three values determine the first diagonal element of the information matrix of d, that is,

$$c_{d11} = (r + 1 + z)(1 + \lambda^2) - (2 - b)\lambda^2 - 2m\lambda$$
$$- \frac{1 - \lambda}{n - (n-2)\lambda}\{(r + 1 + z)(1 - \lambda) + (2 - b)\lambda\}^2.$$

In this section, we use Theorems 3.1 and 3.2 to construct a lower bound of the φ_p-criterion, $0 \le p \le \infty$, for every design $d \in \Omega_{t,n}$. This bound only depends on the three values m, z and b. Then we prove D-optimality of the II(a) design over a restricted subset Δ of $\Omega_{t,n}$.

PROPOSITION 4.1. *Consider any $d \in \Omega_{t,n}$ such that treatment 1 appears $r + 1 + z$ times, $N_{d11} = m$ and $2 - a_{d1} = b$. Then for every $0 \le p \le \infty$, we have*

$$\varphi_p(\mathscr{C}_d) \ge \bar{\varphi}_p(m, z, b),$$

where $\bar{\varphi}_p(m, z, b)$ is the φ_p-criterion of a hypothetical design with eigenvalues $\mu_1(m, z, b), \mu_2(m, z, b), \dots, \mu_2(m, z, b)$ and 0 and where

$$\mu_1(m, z, b) = r(1 + \lambda^2) + \frac{2r\lambda}{t-1} + \frac{(1-\lambda)^2(1+\lambda)r(t-1)}{n - (n-2)\lambda}$$
$$+ \frac{z}{t-1}\{(t-2)(1+\lambda^2) + 4\lambda\} + \frac{b}{t-1}\{(t-2)\lambda^2 + 2\lambda\} - \frac{2mt\lambda}{t-1}$$
$$+ \frac{1-\lambda}{n-(n-2)\lambda}\left[2(1+\lambda)\{b\lambda - z(1-\lambda)\} - \frac{t}{t-1}\{b\lambda - z(1-\lambda)\}^2\right]$$

and

$$\mu_2(m, z, b) = r(1 + \lambda^2) + \frac{2r\lambda}{t-1}$$
$$- \frac{z}{(t-1)(t-2)}\{(t-2)(1+\lambda^2) + 4\lambda\}$$
$$- \frac{b}{(t-1)(t-2)}\{(t-2)\lambda^2 + 2\lambda\} + \frac{2m\lambda}{(t-1)(t-2)}.$$

PROOF. The greatest eigenvalue of $\mathscr{C}_d$ equals the maximum of $x'\mathscr{C}_d x/x'x$ over $x \in \mathbb{R}^t$. The second smallest eigenvalue of $\mathscr{C}_d$ equals the minimum of

$x'\mathscr{C}_d x/x'x$ over $\mathbb{R}^t$, where $x'1_t = 0$. Considering the vector x with first entry $t-1$ and all other entries -1, we find that μ_{d1} is at least $tc_{d11}/(t-1) = \mu_1(m, z, b)$, while $\mu_{d,t-1}$ is at most $\mu_1(m, z, b)$. An upper bound for $\operatorname{tr}\mathscr{C}_d$ is reached if all c_{dii}, $i \geq 2$, equal

$$c_{d2} = \left(r - \frac{z}{t-1}\right)(1+\lambda^2) - \frac{b\lambda^2}{t-1} - \frac{1-\lambda}{n-(n-2)\lambda}\left\{\left(r - \frac{z}{t-1}\right)(1-\lambda) + \frac{b\lambda}{t-1}\right\}^2.$$

In that case, the sum of all eigenvalues equals

$$c_{d11} + (t-1)c_{d2} = \mu_1(m, z, b) + (t-2)\mu_2(m, z, b).$$

Now assume a hypothetical design $\tilde{d}$ with eigenvalues $\mu_1(m, z, b)$, $\mu_2(m, z, b)$, $\ldots$, $\mu_2(m, z, b)$ and 0. We distinguish among two cases:

(i) $\mu_1(m, z, b) \geq \mu_2(m, z, b)$. Then $\tilde{d}$ has the structure of eigenvalues needed in Theorem 3.1. It maximizes $\operatorname{tr}\mathscr{C}_d$ over all d with fixed m, z and b and it maximizes $\operatorname{tr}\mathscr{C}_d - \mu_{d1}$ over all such d. Consequently, $\varphi_p(\mathscr{C}_{\tilde{d}})$ is a lower bound for the φ_p-criterion of all such d, $0 \leq p \leq \infty$.

(ii) $\mu_1(m, z, b) < \mu_2(m, z, b)$. Then $\tilde{d}$ has the structure of eigenvalues needed in Theorem 3.2. It maximizes $\operatorname{tr}\mathscr{C}_d$ over all d with fixed m, z and b and it maximizes μ_{dt-1} over all such d. Consequently, $\varphi_p(\mathscr{C}_{\tilde{d}})$ is a lower bound for the φ_p-criterion of all such d, $0 \leq p \leq \infty$. □

Note that for the Williams II(a) design d^* and for every $0 \leq p \leq \infty$, we get $\varphi_p(\mathscr{C}_{d^*}) = \bar{\varphi}_p(0,0,0)$. Thus, every design $d \in \Omega_{t,n}$ with $\bar{\varphi}_p(m, z, b) > \bar{\varphi}_p(0,0,0)$ performs worse under the φ_p-criterion than the Williams II(a) design, $0 \leq p \leq \infty$.

We will now use this bound to prove the following theorem.

THEOREM 4.2. *For all* $0 < \lambda < 1$, *a Williams* II(a) *design* $d^* \in \Omega_{t,n}$ *is D-optimal* (*i.e.*, φ_0-*optimal*) *over all* $d \in \Omega_{t,n}$ *for which*

$$N_{d11} \leq (t-1)(N_{d22} + \cdots + N_{dtt}).$$

The proof of Theorem 4.2 is very long and, therefore, is presented through a series of propositions. The method of proof is to successively remove from consideration designs that cannot be D-better than the II(a) design and to show that finally no other designs are left. Thus, Proposition 4.11 removes all competing designs that do not have $N_{dii} = 0$ for all i and, hence, those with $m \neq 0$. Then, Proposition 4.12 and 4.13 show that we only need consider designs with $m = z = 0$ and $N_{dii} = 0$ for all i. Next, designs with $b \neq 0$ are excluded. Propositions 4.14 and 4.15 exclude designs with $b = 2$ and, finally, Propositions 4.16–4.18 exclude $b = 1$.

We will be able to exclude most competing designs by showing that the bounds $\bar{\varphi}_0(m, z, b)$ derived in Theorem 4.1 are at least $\bar{\varphi}_0(0,0,0)$. However, for λ close to 1, this is not always possible and we will sometimes use sharper bounds than $\bar{\varphi}_0(m, z, b)$. We think that this is of interest in view of Theorem 2.3. If p is large, we cannot show optimality for small λ, while if p is small, there are difficulties in the proof of optimality for large λ.

We begin with some preliminary results in Propositions 4.3–4.10.

PROPOSITION 4.3. *If a Williams* II(a) *design exists in* $\Omega_{t,n}$, *then*

$$n = rt + 1 \geq t(t-1)/2 + 1 \geq 2(t-1) \geq 4.$$

PROPOSITION 4.4.

$$n - (n-2)\lambda \geq n(1-\lambda) \geq (n-1)(1-\lambda) \quad \textit{for all } 0 < \lambda < 1.$$

PROPOSITION 4.5.

$$\mu_1(0,0,0) > \mu_2(0,0,0) \quad \textit{for all } 0 \leq \lambda < 1.$$

PROPOSITION 4.6. *Define* $\delta_1(m, z, b) = \mu_1(m, z, b) - \mu_1(0,0,0)$. *Then*

$$\delta_1(m, z, b) \leq \frac{z}{t-1}\{(t-2)(1+\lambda^2) + 4\lambda\} + \frac{b}{t-1}\{(t-2)\lambda^2 + 4\lambda\} - \frac{2mt}{t-1}\lambda.$$

PROOF. Using Proposition 4.4 and the fact that $1 + \lambda < 2$ yields

$$\frac{2(1-\lambda)(1+\lambda)b\lambda}{n-(n-2)\lambda} \leq \frac{4b\lambda}{n}.$$

Proposition 4.3 then gives the desired result. □

PROPOSITION 4.7. *For all m, z and b we have*

$$\mu_1(0,0,0) + (t-2)\mu_2(0,0,0) \geq \mu_1(m, z, b) + (t-2)\mu_2(m, z, b) - 2b\lambda/(t-1) + 2m\lambda.$$

PROOF.

$$\begin{aligned}\mu_1(m, z, b) &+ (t-2)\mu_2(m, z, b) - \mu_1(0,0,0) - (t-2)\mu_2(0,0,0)\\ &= -2m\lambda + \frac{1-\lambda}{n-(n-2)\lambda}\left[2(1+\lambda)\{b\lambda - z(1-\lambda)\} - \frac{t}{t-1}\{b\lambda - z(1-\lambda)\}^2\right]\\ &\leq -2m\lambda + \frac{2b\lambda}{t-1},\end{aligned}$$

where the inequality was derived as in Proposition 4.6. □

PROPOSITION 4.8. *If in the design $d \in \Omega_{t,n}$ a treatment appears at least once adjacent to itself, then* $\operatorname{tr} \mathscr{C}_d \le \operatorname{tr} \mathscr{C}_{d^*}$, *where d^* is the Williams* II(a) *design.*

PROOF. If one treatment i, $1 \le i \le t$, appears at least once adjacent to itself, this decreases the trace by 2λ. Thus, for the design d we get

$$\operatorname{tr} \mathscr{C}_d \le \mu_1(0, z, b) + (t-2)\mu_2(0, z, b) - 2\lambda.$$

Proposition 4.7 now gives the desired result, since $2b/(t-1) \le 2$. □

PROPOSITION 4.9. *Define* $\delta_2(m, z, b) = (t-2)\{\mu_2(0,0,0) - \mu_2(m, z, b)\}$. *Then*

$$\delta_2(m, z, b) \ge \delta_2(m, 0, b) \ge \delta_2(m, 0, 0)$$

for all m, z, b and all $0 < \lambda < 1$. Equality holds in the first inequality if and only if $z = 0$ and in the second inequality if and only if $b = 0$.

PROPOSITION 4.10. *For all m, z and b, we have*

$$\delta_1(m, z, b) - \delta_2(m, z, b) = -2m\lambda + \frac{1-\lambda}{n-(n-2)\lambda}\left[2(1+\lambda)\{b\lambda - z(1-\lambda)\} - \frac{t}{t-1}\{b\lambda - z(1-\lambda)\}^2\right].$$

PROPOSITION 4.11. *Assume at least one treatment appears adjacent to itself in d and $N_{d11} \le (t-1)\sum_{i=2}^{t} N_{dii}$. Then d performs worse under the D-criterion than d^*.*

PROOF. Proposition 4.8 shows that $\operatorname{tr} \mathscr{C}_d < \operatorname{tr} \mathscr{C}_{d^*}$. The Williams II(a) design d^* has the structure of the eigenvalues needed in Theorem 3.1. Further,

$$\begin{aligned}
\mu_2(d)(t-2) &\le \operatorname{tr} \mathscr{C}_d - \mu_1(d) \le \operatorname{tr} \mathscr{C}_d - \mu_1(m, z, b) \\
&\le (t-2)\mu_2(m, z, b) - 2\sum_{i=2}^{t} N_{dii}\lambda \\
&\le (t-2)\mu_2(0,0,0) + \frac{2m\lambda}{t-1} - 2\sum_{i=2}^{t} N_{dii}\lambda \\
&\le (t-2)\mu_2(0,0,0).
\end{aligned}$$

Here we have used Proposition 4.9 to show that

$$(t-2)\mu_2(m, z, b) \le (t-2)\mu_2(0,0,0) + 2m\lambda/(t-1).$$

Now define $\Delta = \{d, d^*\}$. Then Theorem 3.1 gives the desired result. □

For the rest of this section, we can restrict attention to designs for which no treatment ever appears adjacent to itself. It is clear that we only have to consider such z and b for which $\bar{\varphi}_0(m, z, b) < \infty$, that is, for which both $\mu_1(m, z, b)$ and $\mu_2(m, z, b)$ are nonzero.

PROPOSITION 4.12. *For all z and b, we have $\bar{\varphi}_0(0, z, b) \le \bar{\varphi}_0(0,0,0)$ only if*

$$\{\delta_1(0, z, b) - \delta_2(0, z, b)\}\mu_2(0,0,0) \ge \delta_2(0, z, b)\{\mu_1(0,0,0) - \mu_2(0, z, b)\}.$$

PROOF. $\bar{\varphi}_0(0, z, b) \le \bar{\varphi}_0(0,0,0)$ if and only if

$$\mu_1(0, z, b)\mu_2(0, z, b)^{t-2} \ge \mu_1(0,0,0)\mu_2(0,0,0)^{t-2}.$$

The concavity of the product implies that

$$\mu_2(0,0,0)^{t-2} \ge \mu_2(0, z, b)^{t-3}\{\mu_2(0, z, b) + \delta_2(0, z, b)\}.$$

Consequently,

$$\mu_1(0, z, b)\mu_2(0, z, b) \ge \mu_1(0,0,0)\{\mu_2(0, z, b) + \delta_2(0, z, b)\},$$
$$\delta_1(0, z, b)\mu_2(0, z, b) \ge \delta_2(0, z, b)\mu_1(0,0,0)$$

and

$$\{\delta_1(0, z, b) - \delta_2(0, z, b)\}\mu_2(0, z, b) \ge \delta_2(0, z, b)\{\mu_1(0,0,0) - \mu_2(0, z, b)\}.$$

Proposition 4.9 implies that the right-hand side is always nonnegative, and the inequality can hold only if $\delta_1(0, z, b) \ge \delta_2(0, z, b)$. The fact that $\mu_2(0, z, b) \le \mu_2(0, 0, 0)$ then gives the desired result. □

The next thing we do is to show, by a step by step consideration of the values z and b can take, that there is no design that performs better than the Williams II(a) design and for which no treatment is adjacent to itself.

PROPOSITION 4.13. *If $z \ge 1$, then $\bar{\varphi}_0(0, z, b) > \bar{\varphi}_0(0,0,0)$ for every $b \in \{0,1,2\}$.*

PROOF.

CASE (i), $b = 0$. The right-hand side of the inequality in Proposition 4.12 is positive for all $z > 0$. Consequently, $\bar{\varphi}_0'(0, z, 0)$ can be as small as $\bar{\varphi}_0(0,0,0)$ only if

$$\delta_1(0, z, 0) - \delta_2(0, z, 0) > 0.$$

According to Proposition 4.10, this would imply

$$z(1 - \lambda)\left\{-2(1 + \lambda) - \frac{t}{t-1}z(1 - \lambda)\right\} > 0,$$

and that can never be true.

CASE (ii), $b = 1$. According to Proposition 4.10, we have

$$\delta_1(0, z, 1) - \delta_2(0, z, 1) = \frac{1-\lambda}{n-(n-2)\lambda}\left[2(1+\lambda)\{\lambda - z(1-\lambda)\} - \frac{t}{t-1}\{\lambda - z(1-\lambda)\}^2\right].$$

Assume $\bar{\varphi}_0(0, z, 1)$ is as small as $\bar{\varphi}_0(0,0,0)$. The inequality in Proposition 4.12 now implies

$$\frac{1-\lambda}{n-(n-2)\lambda}\left[2(1+\lambda)\{\lambda - z(1-\lambda)\} - \frac{t}{t-1}\{\lambda - z(1-\lambda)\}^2\right] \times\left\{r(1+\lambda^2) + \frac{2r\lambda}{t-1}\right\}$$
$$\geq \left[\frac{z}{t-1}\{(t-2)(1+\lambda^2)+4\lambda\} + \frac{1}{t-1}\{(t-2)\lambda^2 + 2\lambda\}\right] \times\left\{\frac{(t-1)r(1-\lambda)^2(1+\lambda)}{n-(n-2)\lambda} + \frac{z(1+\lambda^2)+\lambda^2}{t-1}\right\}$$

Defining $f(z)$ as the left-hand side minus the right-hand side of the inequality, we find that the derivative of f with respect to z is negative for $z \geq 1$. Thus, the inequality can hold for some $z \geq 1$, only if it holds for $z = 1$. Proposition 4.4 implies that $\{n-(n-2)\lambda\}/\{r(1-\lambda)\} > t$ and, consequently,

$$\left\{2(2\lambda-1)(1+\lambda) - \frac{t}{t-1}(2\lambda-1)^2\right\}\left(1+\lambda^2+\frac{2\lambda}{t-1}\right)$$
$$\geq \frac{1}{t-1}\{(t-2)(1+2\lambda^2)+6\lambda\}\left\{(t-1)(1-\lambda^2) + \frac{(1+2\lambda^2)t}{t-1}\right\}.$$

Some algebra shows that this cannot be true for $t \geq 3$.

CASE (iii), $b = 2$. According to Proposition 4.10, we have

$$\delta_1(0, z, 2) - \delta_2(0, z, 2) = \frac{1-\lambda}{n-(n-2)\lambda}\left[2(1+\lambda)\{2\lambda - z(1-\lambda)\} - \frac{t}{t-1}\{2\lambda - z(1-\lambda)\}^2\right].$$

Assume $\bar{\varphi}_0(0, z, 2) \leq \bar{\varphi}_0(0,0,0)$. The inequality of Proposition 4.12 then implies, as in Case (ii), that

$$\tilde{f}(z) = \left[2(1+\lambda)\{2\lambda - z(1-\lambda)\} - \frac{t}{t-1}\{2\lambda - z(1-\lambda)\}^2\right]\left(1+\lambda^2+\frac{2\lambda}{t-1}\right)$$
$$-\left[\frac{z}{t-1}\{(t-2)(1+\lambda^2)+4\lambda\} + \frac{2}{t-1}\{(t-2)\lambda^2+2\lambda\}\right]$$
$$\times\left\{(t-1)(1-\lambda^2) + \frac{t}{t-1}(1+3\lambda^2)\right\} \geq 0.$$

We find that the derivative of $\tilde{f}(z)$ with respect to z is negative. That means that the inequality can hold for some $z \geq 1$ only if it holds for $z = 1$. Inserting $z = 1$, we get

$$\left\{2(1+\lambda)(3\lambda-1) - \frac{t}{t-1}(3\lambda-1)^2\right\}\left(1+\lambda^2+\frac{2\lambda}{t-1}\right)$$

$$\geq \frac{1}{t-1}\{(t-2)(1+3\lambda^2)+8\lambda\}\left\{(t-1)(1-\lambda^2)+\frac{t}{t-1}(1+3\lambda^2)\right\}.$$

Further manipulation shows that this inequality can never hold. □

Let us summarize what we have shown so far. In every competing design, treatment 1 is a treatment which appears most often. The competing designs are restricted to those for which $N_{d11} \leq (t-1)(N_{d22} + \cdots + N_{dtt})$. Assume in this subset of $\Omega_{t,n}$ there is a design d performing better under the D-criterion than the Williams II(a) design $d^* \in \Omega_{t,n}$.

Then in d, no treatment ever appears adjacent to itself (Proposition 4.11). Since no treatment appears more often than treatment 1, Proposition 4.13 implies that every treatment appears at most $r+1$ times in the design. If treatment 1 appears at both end plots, then we have $m = z = b = 0$ for d and the design d cannot be better than d^*, since then $\varphi_0(\mathscr{C}_d) \geq \bar{\varphi}_0(0,0,0)$ (Proposition 4.1).

If treatment 1 does not appear at an end plot, then $\varphi_0(\mathscr{C}_d) \geq \bar{\varphi}_0(0,0,2)$. Unfortunately, if $t \geq 7$ and $\lambda > (t-2)/t$, it can happen that $\bar{\varphi}_0(0,0,0) > \bar{\varphi}_0(0,0,2)$. Similarly, if treatment 1 appears at only one end plot, then $\varphi_0(\mathscr{C}_d) \geq \bar{\varphi}_0(0,0,1)$ and if $\lambda > \frac{1}{2}$, it can happen that $\bar{\varphi}_0(0,0,0) > \bar{\varphi}_0(0,0,1)$.

This means that in the remaining cases, where at least one end plot is not occupied by treatment 1, we have to do a more detailed analysis to show that d does not perform better than d^*.

PROPOSITION 4.14. *Assume $d \in \Omega_{t,n}$ is such that $r_{d1} = r+1$ and $a_{d1} = 0$, i.e., $b = 2$. Further, assume that $N_{dii} = 0$ for all $i \in \{1,\dots,t\}$ and that there is a treatment t, say, which appears not more than $r-1$ times in the design and which appears at an end plot. Then d performs worse under the D-criterion than the Williams* II(a) *design $d^* \in \Omega_{t,n}$.*

PROOF. Consider the tth diagonal element of $\mathscr{C}_d$. Then $a_{dt} \geq 1$ and

$$c_{dtt} = r_{dt}(1+\lambda^2) - a_{dt}\lambda^2 - \frac{1-\lambda}{n-(n-2)\lambda}\{r_{dt}(1-\lambda)+a_{dt}\lambda\}^2$$

$$\leq (r-1)(1+\lambda^2) - \frac{1-\lambda}{n-(n-2)\lambda}\{(r-1)(1-\lambda)+\lambda\}^2.$$

Note that

$$\begin{aligned}\{r(1-\lambda)-1+2\lambda\}^2 &= \frac{1}{t^2}\{rt-rt\lambda-t+2t\lambda\}^2\\ &= \frac{1}{t^2}\{rt+1-(rt-1)\lambda-(t+1)+(2t-1)\lambda\}^2\\ &\geq \frac{1}{t^2}\{n-(n-2)\lambda\}^2\\ &\quad -\frac{2}{t^2}\{n-(n-2)\lambda\}\{(t+1)-(2t-1)\lambda\}.\end{aligned}$$

Hence,

$$\begin{aligned}c_{dtt} &\leq (r-1)(1+\lambda^2)-\frac{1-\lambda}{t^2}\{n-(n-2)\lambda-2(t+1)+2(2t-1)\lambda\}\\ &= (r-1)(1+\lambda^2)-\frac{1-\lambda}{t^2}\{rt(1-\lambda)-(2t+1)+(4t-1)\lambda\}\\ &\leq (r-1)(1+\lambda^2)-\frac{1-\lambda}{t^2}\{(r-3)t(1-\lambda)\}.\end{aligned}$$

The eigenvalue $\mu_{d,t-1}$ of $\mathscr{C}_d$ satisfies

$$\mu_{d,t-1} \leq \frac{t}{t-1}c_{dtt}.$$

If we use the bound for c_{dtt} derived earlier, we get

$$\mu_{d,t-1} \leq \frac{rt}{t-1}(1+\lambda^2)-\frac{r}{t-1}(1-\lambda)^2-\frac{t}{t-1}(1+\lambda^2)+\frac{3}{t-1}(1-\lambda)^2.$$

It follows that

$$\begin{aligned}&\mu_2(0,0,0)-\delta_2(0,0,2)-\mu_{d,t-1}\\ &\quad\geq r(1+\lambda^2)+\frac{2r\lambda}{t-1}-\frac{2}{t-1}\{(t-2)\lambda^2+2\lambda\}\\ &\qquad -\frac{rt}{t-1}(1+\lambda^2)+\frac{r}{t-1}(1-\lambda)^2+\frac{t}{t-1}(1+\lambda^2)-\frac{3}{t-1}(1-\lambda)^2\\ &\quad=\frac{1}{t-1}\{(t-3)(1-\lambda^2)+2(\lambda-\lambda^2)\}>0.\end{aligned}$$

We, thus, know that

$$\begin{aligned}\sum\mu_{di} &\leq \mu_1(0,0,2)+(t-2)\mu_2(0,0,2)\\ &= \mu_1(0,0,2)+(t-2)\mu_2(0,0,0)-\delta_2(0,0,2)\end{aligned}$$

and that

$$\mu_{d1}\geq\mu_1(0,0,2),\qquad \mu_{d,t-1}\leq\mu_2(0,0,0)-\delta_2(0,0,2).$$

Since

$$\delta_1(0,0,b) - \delta_2(0,0,b) = \frac{1-\lambda}{n-(n-2)\lambda}\left\{2(1+\lambda)b\lambda - \frac{t}{t-1}b^2\lambda^2\right\} > 0$$

and since $\delta_2(0,0,2) > 0$ (Proposition 4.9), we have $\mu_1(0,0,2) > \mu_1(0,0,0) > \mu_2(0,0,0)$. The concavity of the product then implies that

$$\mu_{d1} \cdot \cdots \cdot \mu_{d,t-1} \le \mu_1(0,0,2)\mu_2(0,0,0)^{t-3}\{\mu_2(0,0,0) - \delta_2(0,0,2)\}.$$

Consequently, d can perform as well as the Williams II(a) design only if

$$\mu_1(0,0,2)\{\mu_2(0,0,0) - \delta_2(0,0,2)\} \ge \mu_1(0,0,0)\mu_2(0,0,0).$$

This is equivalent to

$$\{\delta_1(0,0,2) - \delta_2(0,0,2)\}\mu_2(0,0,0) \ge \delta_2(0,0,2)\{\mu_1(0,0,2) - \mu_2(0,0,0)\}.$$

Note that this inequality differs from the one in Proposition 4.12. In fact,

$$\begin{aligned}\mu_1(0,0,2) - \mu_2(0,0,0) &= \mu_1(0,0,0) - \mu_2(0,0,0) + \delta_1(0,0,2)\\ &> \mu_1(0,0,0) - \mu_2(0,0,0) + \delta_2(0,0,2)\\ &> \mu_1(0,0,0) - \mu_2(0,0,0) + \delta_2(0,0,2)/(t-2)\\ &= \mu_1(0,0,0) - \mu_2(0,0,2).\end{aligned}$$

Consequently, we would need

$$\begin{aligned}&\{\delta_1(0,0,2) - \delta_2(0,0,2)\}\mu_2(0,0,0)\\ &\quad\ge \delta_2(0,0,2)\left\{\frac{(t-1)(1-\lambda)^2(1+\lambda)r}{n-(n-2)\lambda} + \delta_2(0,0,2)\right\}.\end{aligned}$$

As in Proposition 4.13, it would then follow that

$$\begin{aligned}&\left\{4(1+\lambda)\lambda - \frac{4t\lambda^2}{t-1}\right\}\left(1+\lambda^2+\frac{2\lambda}{t-1}\right)\\ &\quad\ge \frac{2}{t-1}\{(t-2)\lambda^2+2\lambda\}\left\{(t-1)(1-\lambda^2) + \frac{2t(t-2)\lambda^2}{t-1} + \frac{4t\lambda}{t-1}\right\}\end{aligned}$$

and, consequently, that

$$\begin{aligned}&\{2(t-1) - 2\lambda\}\{(1+\lambda^2)(t-1) + 2\lambda\}\\ &\quad\ge \{(t-2)\lambda + 2\}(t-1)^2(1-\lambda^2) + t\lambda\{2t(t-2)\lambda^2 + 4t\lambda\}.\end{aligned}$$

This cannot be fulfilled. □

PROPOSITION 4.15. *Assume $d \in \Omega_{t,n}$ is such that $r_{d1} = r+1$, $a_{d1} = 0$ and $N_{dii} = 0$ for all $1 \le i \le t$. Further, assume that no treatment appearing at an end plot appears less than r times in the design. Then d performs worse under the D-criterion than the Williams* II(a) *design $d^* \in \Omega_{t,n}$.*

PROOF. We can show that $\operatorname{tr}\mathscr{C}_d$ is smaller than the bound which we assumed for the construction of $\mu_2(m, z, b)$.

There are at most two treatments $t-1$ and t, say, appearing at end plots. For both, we have $r_{di} \geq r$. Thus,

$$\begin{aligned}\operatorname{tr}\mathscr{C}_d &\leq n(1+\lambda^2) - 2\lambda^2 \\ &\quad - \frac{1-\lambda}{n-(n-2)\lambda}\left[\sum_{i=1}^{t-2}\{r_{di}(1-\lambda)\}^2 + \sum_{i=t-1}^{t}\{r_{di}(1-\lambda) + a_{di}\lambda\}^2\right] \\ &\leq n(1+\lambda^2) - 2\lambda^2 - \frac{1-\lambda}{n-(n-2)\lambda} \\ &\qquad \times\left\{tr^2(1-\lambda)^2 + 2r(1-\lambda)^2 + (1-\lambda)^2 + 4r\lambda(1-\lambda) + 2\lambda^2\right\}.\end{aligned}$$

The bound used in Proposition 4.1 for $b = 2$ and $z = 0$ equals

$$n(1+\lambda^2) - 2\lambda^2 - \frac{1-\lambda}{n-(n-2)\lambda}\left\{tr^2(1-\lambda)^2 + 2r(1-\lambda)^2 + (1-\lambda)^2 + 4r\lambda(1-\lambda) + \frac{4\lambda^2}{t-1}\right\}.$$

Thus, the difference between the bound for the trace and the true trace is at least

$$\rho = \frac{1-\lambda}{n-(n-2)\lambda}\,\frac{2(t-3)}{t-1}\lambda^2.$$

This implies that $\mu_2(d)$ is at most $\mu_2(0,0,2) - \rho/(t-2)$ and d can be D-better than d^* only if

$$\mu_1(0,0,2)\{\mu_2(0,0,2) - \rho/(t-2)\}^{t-2} \geq \mu_1(0,0,0)\mu_2(0,0,0)^{t-2}.$$

With the same reasoning as in Proposition 4.12, we conclude that this can be true only if

$$\begin{aligned}&\{\delta_1(0,0,2) - \delta_2(0,0,2) - \rho\}\mu_2(0,0,0) \\ &\quad\geq \delta_2(0,0,2)\left\{\frac{(t-1)(1-\lambda)^2(1+\lambda)r}{n-(n-2)\lambda} + \frac{\delta_2(0,0,2)}{(t-2)}\right\}\end{aligned}$$

and, consequently, if

$$\begin{aligned}&(2\lambda - \lambda^2)\left(1 + \lambda^2 + \frac{2\lambda}{t-1}\right) \\ &\quad\geq \frac{1}{t-1}\{(t-2)\lambda^2 + 2\lambda\}\left\{(t-1)(1-\lambda^2) + \frac{2t\lambda^2}{t-1} + \frac{4t\lambda}{(t-1)(t-2)}\right\}.\end{aligned}$$

Some algebra shows that this is not true. □

Propositions 4.14 and 4.15 show that designs for which $b = 2$ cannot perform as well as the Williams II(a) design d^*. This means that if there is a design d performing better than d^*, it must have $b = 1$, that is, treatment 1 appearing at exactly one end plot.

PROPOSITION 4.16. *Assume a design $d \in \Omega_{t,n}$ exists with $r_{d1} = r + 1$, $a_{d1} = 1$ and $N_{dii} = 0$ for all $1 \le i \le t$. The design d can perform as well under the D-criterion as the Williams* II(a) *design $d^* \in \Omega_{t,n}$ only if $t \ge 4$ and*

$$\operatorname{tr} \mathscr{C}_d \le \mu_1(0,0,1) + (t-2)\mu_2(0,0,1) - \frac{1-\lambda}{n-(n-2)\lambda}\frac{t-2}{t-1}\lambda^2.$$

PROOF. Since $a_{d1} = 1$, it follows that another treatment t, say, appears at an end plot of the design. Assume $r_{dt} \le r - 1$. Then there must be another treatment 2, say, such that $r_{d2} \ge r + 1$ and $a_{d2} = 0$. Relabelling the treatments to exchange 1 and 2 leads to the case $a_{d1} = 0$, which was already solved.

So treatment t appears at least r times and

$$\begin{aligned}\operatorname{tr} \mathscr{C}_d &= n(1+\lambda^2) - 2\lambda^2 - \frac{1-\lambda}{n-(n-2)\lambda}\\ &\quad\times\left[\{(r+1)(1-\lambda)+\lambda\}^2 + \sum_{i=2}^{t-1} r_{di}^2(1-\lambda)^2 + \{r_{dt}(1-\lambda)+\lambda\}^2\right]\\ &\le n(1+\lambda^2) - 2\lambda^2 - \frac{1-\lambda}{n-(n-2)\lambda}\\ &\quad\times\left[\{(r+1)(1-\lambda)+\lambda\}^2 + (t-1)\left\{r(1-\lambda)+\frac{\lambda}{(t-1)}\right\}^2\right]\\ &\quad-\frac{1-\lambda}{n-(n-2)\lambda}\frac{t-2}{t-1}\lambda^2.\end{aligned}$$

Thus, $\operatorname{tr} \mathscr{C}_d$ is smaller than we assumed for the determination of $\mu_2(0,0,1)$. The difference is at least

$$\tilde{\rho} = \frac{1-\lambda}{n-(n-2)\lambda}\frac{t-2}{t-1}\lambda^2.$$

In the case of three treatments, d can perform better than d^* only if

$$\mu_1(0,0,1)\{\mu_2(0,0,1) - \tilde{\rho}\} \ge \mu_1(0,0,0)\mu_2(0,0,0).$$

Since $t - 2 = 1$, it follows that

$$\delta_1(0,0,1)\{\mu_2(0,0,1) - \tilde{\rho}\} \ge \mu_1(0,0,0)\{\delta_2(0,0,1) + \tilde{\rho}\}$$

and thus

$$\begin{aligned}&\{\delta_1(0,0,1) - \delta_2(0,0,1) - \tilde{\rho}\}\mu_2(0,0,0)\\ &\quad\ge \delta_2(0,0,1)\left\{\frac{(t-1)(1-\lambda)^2(1+\lambda)r}{n-(n-2)\lambda} + \delta_2(0,0,1)\right\}.\end{aligned}$$

Application of Proposition 4.4 shows that

$$2\lambda\left(1+\lambda^2+\frac{2\lambda}{t-1}\right)\geq\frac{1}{t-1}\{(t-2)\lambda^2+2\lambda\}\left\{(t-1)(1-\lambda^2)+\frac{t\lambda^2+2t\lambda}{t-1}\right\},$$

where $t=3$. This cannot be true for $0<\lambda<1$. □

Proposition 4.17. *If a design $d\in\Omega_{t,n}$ is such that $r_{d1}=r+1$, $r_{d2}=\cdots=r_{dt}=r$, $a_{d1}=a_{dt}=1$ and $N_{dii}=0$ for all $1\leq i\leq t$, then it performs worse under the D-criterion than the Williams* II(a) *design $d^*\in\Omega_{t,n}$.*

Proof. Treatment 1 has exactly $2(r+1)-1$ neighbors. This implies that $2r+1$ of the $n-1=rt$ possible adjacencies in the design are occupied by adjacencies with treatment 1. So there are only $(t-2)r-1$ adjacencies possible among treatments $2,\dots,t$, that is, $\sum_{i=2}^{t-1}\sum_{j=i+1}^{t}N_{dij}=(t-2)r-1$. Consequently, at least one of the $(t-1)(t-2)/2$ different N_{dij} must be at most $2r/(t-1)-1$.

The eigenvalue $\mu_{d,t-1}$ fulfills

$$\mu_{d,t-1}\leq\frac{1}{2}(c_{dii}+c_{djj}-2c_{dij})\leq r(1+\lambda^2)+\frac{2r\lambda}{t-1}-\lambda.$$

As a consequence of Proposition 4.16, we only have to consider the case $t\geq4$. We want to show by contradiction that d performs worse than d^*. We know that

$$\mu_{d1}\geq\mu_1(0,0,1),\qquad \mu_{d,t-1}\leq\mu_2(0,0,0)-\lambda$$

and from Proposition 4.16 that

$$\mu_{d1}+\cdots+\mu_{d,t-1}\leq\mu_1(0,0,1)+(t-2)\mu_2(0,0,0)-\delta_2(0,0,1)-\tilde{\rho}.$$

Note that

$$\mu_1(0,0,1)\geq\mu_2(0,0,0)-\frac{\delta_2(0,0,1)+\tilde{\rho}-\lambda}{t-3}$$

and that

$$\mu_2(0,0,0)-\lambda\leq\mu_2(0,0,0)-\frac{\delta_2(0,0,1)+\tilde{\rho}-\lambda}{t-3}.$$

Consequently, the concavity of the product implies that

$$\mu_{d1}\times\cdots\times\mu_{d,t-1}$$
$$\leq\mu_1(0,0,1)\left\{\mu_2(0,0,0)-\frac{\delta_2(0,0,1)+\tilde{\rho}-\lambda}{t-3}\right\}^{t-3}\{\mu_2(0,0,0)-\lambda\}.$$

The concavity of the product also implies that

$$\mu_2(0,0,0)^{t-2}\geq\left\{\mu_2(0,0,0)-\frac{\delta_2(0,0,1)+\tilde{\rho}-\lambda}{t-3}\right\}^{t-3}$$
$$\times\{\mu_2(0,0,0)+\delta_2(0,0,1)+\tilde{\rho}-\lambda\}.$$

If d were to perform better than d^*, we would have

$$\mu_1(0,0,1)\{\mu_2(0,0,0) - \lambda\} \geq \mu_1(0,0,0)\{\mu_2(0,0,0) + \delta_2(0,0,1) + \tilde{\rho} - \lambda\}$$

or, equivalently,

$$\delta_1(0,0,1)\{\mu_2(0,0,0) - \lambda\} \geq \mu_1(0,0,0)\{\delta_2(0,0,1) + \tilde{\rho}\}.$$

Subtract $\{\delta_2(0,0,1) + \tilde{\rho}\}\{\mu_2(0,0,0) - \lambda\}$ from both sides. It then easily follows that

$$\{\delta_1(0,0,1) - \delta_2(0,0,1) - \tilde{\rho}\}\mu_2(0,0,0) > \delta_2(0,0,1)\{\mu_1(0,0,0) - \mu_2(0,0,0) + \lambda\}.$$

Hence, we would need

$$\frac{1-\lambda}{n-(n-2)\lambda} 2\lambda\left\{r(1+\lambda^2) + \frac{2r\lambda}{t-1}\right\} > \frac{1}{t-1}\{(t-2)\lambda^2 + 2\lambda\}\left\{\frac{(t-1)(1-\lambda)^2(1+\lambda)r}{n-(n-2)\lambda} + \lambda\right\}.$$

Multiplication of both sides by $\{n-(n-2)\lambda\}/\{r(1-\lambda)\} > t$ implies

$$2\lambda\left(1 + \lambda^2 + \frac{2\lambda}{t-1}\right) > \{(t-2)\lambda^2 + 2\lambda\}\left\{1 - \lambda^2 + \frac{t\lambda}{t-1}\right\}.$$

This inequality cannot be true for $0 < \lambda < 1$. □

PROPOSITION 4.18. *Assume the design $d \in \Omega_{t,n}$ is such that $r_{d1} = r + 1$, $a_{d1} = 1$ and $N_{dii} = 0$, $1 \leq i \leq t$. Assume there is one treatment t, say, such that $r_{dt} \leq r - 1$. Then d performs worse under the D-criterion than the Williams* II(a) *design $d^* \in \Omega_{t,n}$.*

PROOF. Since treatment t appears less than r times, it has at most $2r-2$ neighbors. This means that there is at least one treatment i (which appears not more than $r+1$ times in the design) such that $N_{dit} \leq 2r/(t-1) - 1$. Consequently,

$$\begin{aligned}\frac{1}{2}(c_{dii} + c_{dtt} - 2c_{dit}) \leq \frac{1}{2}\Bigg\{&(r+1)(1+\lambda^2) - \frac{1-\lambda}{n-(n-2)\lambda}(r+1)^2(1-\lambda)^2 \\ &+ (r-1)(1+\lambda^2) - \frac{1-\lambda}{n-(n-2)\lambda}(r-1)^2(1-\lambda)^2 \\ &+ \frac{4r\lambda}{t-1} - 2\lambda + 2\frac{1-\lambda}{n-(n-2)\lambda}(r^2-1)(1-\lambda)^2\Bigg\} \\ \leq r(1+\lambda^2) + \frac{2r\lambda}{t-1} - \lambda.&\end{aligned}$$

Now, proceed as in Proposition 4.17. □

This excludes the last possible candidate for D-optimality and our proof is complete.

5. Further optimality proofs. We are now in a position to give a proof of the theorems in Section 2. We will make use of the methods in Section 3 to exend Theorem 4.2 to more general criteria.

PROOF OF THEOREM 2.1. We apply Theorem 3.3. From Theorem 4.2, we know that d^* is φ_0-optimal over the competing designs. Proposition 4.5 shows that d^* has the required structure of the eigenvalues. We have to verify condition (iii) of Theorem 3.3. As $\mu_2(d)$ is difficult to compute, we use an upper bound.

For $\operatorname{tr} \mathscr{C}_d$ fixed, the φ_0-criterion decreases in $\mu_2(d)$. It is minimal if $\mu_2(d)$ is as large as possible but not larger than $\mu_1(d)$. Thus for $\operatorname{tr} \mathscr{C}_d$ fixed, the minimal value of φ_0 is attained if $\mu_1(d) = \mu_{d1}$ and $\mu_2(d) = (\operatorname{tr} \mathscr{C}_d - \mu_{d1})/(t-2)$. The true $\mu_2(d)$ is not greater than that.

We have for every competing design that

$$(t-2)\mu_2(d) \le \operatorname{tr} \mathscr{C}_d - \mu_{d1} \le (t-2)\mu_2(0,0,0).$$

The details of the derivation of the last inequality can be found in the proof of Proposition 4.11. □

It is clear from the examples given in Proposition 2.4 that there are designs which are φ_p-better than the Williams II(a) designs for large p, especially for $p = \infty$. Thus Theorem 2.1 cannot be extended if p is large and at the same time λ is small. Theorems 2.2 and 2.3 illustrate some possible generalizations. On the basis of examples, we conjecture φ_p-optimality of the type II(a) designs for all $0 < \lambda < 1$ whenever $p < \{\log(3)/\log(9/7)\} - 1 \approx 3.37$.

PROOF OF THEOREM 2.2. The only designs left to compete have $N_{d11} > (t-1)\sum_{i=2}^{t} N_{dii}$. Thus Proposition 4.8 implies that $\operatorname{tr} \mathscr{C}_d \le \operatorname{tr} \mathscr{C}_{d^*}$ for all these d. If in addition $\mu_2(m, z, b) \le \mu_2(0,0,0)$, then Theorem 3.1 implies φ_p-optimality of d^* for all $p \ge 0$. We thus only have to consider designs with $m > 0$ and z, b such that $\mu_2(m, z, b) > \mu_2(0,0,0)$. As $\operatorname{tr} \mathscr{C}_{d^*} \ge \operatorname{tr} \mathscr{C}_d$, it follows that $\mu_1(m, z, b) < \mu_1(0,0,0)$. We know that

$$\sum_{i=1}^{t-1} \mu_{di}^{-1} \ge \mu_1(m, z, b)^{-1} + (t-2)\mu_2(m, z, b)^{-1};$$

see Proposition 4.1. So under the A-criterion d cannot be as good as d^* unless

$$\mu_1(m, z, b)^{-1} + (t-2)\mu_2(m, z, b)^{-1} \le \mu_1(0,0,0)^{-1} + (t-2)\mu_2(0,0,0)^{-1}.$$

This implies

$$\{\mu_1(0,0,0) - \mu_1(m, z, b)\}\mu_2(m, z, b)\mu_2(0,0,0) \le (t-2)\{\mu_2(m, z, b) - \mu_2(0,0,0)\}\mu_1(m, z, b)\mu_1(0,0,0).$$

As $\mu_2(m, z, b) > \mu_2(0,0,0)$ and $\mu_1(m, z, b) < \mu_1(0,0,0)$, it follows that

$$\{\mu_1(0,0,0) - \mu_1(m, z, b)\}\mu_2(0,0,0)^2 \le (t-2)\{\mu_2(m, z, b) - \mu_2(0,0,0)\}\mu_1(0,0,0)^2.$$

Note that

$$\frac{\mu_1(0,0,0)}{\mu_2(0,0,0)} = 1 + \left[(t-1)\frac{(1-\lambda)^2(1+\lambda)r}{n-(n-2)\lambda}\Big/\left\{r(1+\lambda^2)+\frac{2r\lambda}{t-1}\right\}\right]$$

$$\le 1 + \left[(t-1)\frac{(1-\lambda^2)}{n}\Big/\left\{1+\lambda^2+\frac{2\lambda}{t-1}\right\}\right] < 1 + \frac{t-1}{n} < \sqrt{t},$$

where we have used Proposition 4.4. Consequently, d can perform as well as d^* only if

$$\mu_1(0,0,0) - \mu_1(m,z,b) < t(t-2)\{\mu_2(m,z,b) - \mu_2(0,0,0)\}.$$

Proposition 4.7 shows that then

$$(t-2)\{\mu_2(m,z,b) - \mu_2(0,0,0)\} - 2b\lambda/(t-1) + 2m\lambda$$
$$< t(t-2)\{\mu_2(m,z,b) - \mu_2(0,0,0)\}$$

and, consequently,

$$2m\lambda < (t-1)(t-2)\{\mu_2(m,z,b) - \mu_2(0,0,0)\} + 2b\lambda/(t-1)$$
$$= -z\{(t-2)(1+\lambda^2)+4\lambda\} - b\{(t-2)\lambda^2 + 2\lambda\} + 2b\lambda/(t-1) + 2m\lambda.$$

This can be true only if

$$0 < -z\{(t-2)(1+\lambda^2)+4\lambda\} - b\{(t-2)\lambda^2 + 2(t-2)\lambda/(t-1)\}.$$

This cannot hold and thus under the A-criterion d cannot perform as well as d^*. Can it under the D-criterion?

We have just seen that

$$\mu_1(0,0,0) - \mu_1(m,z,b) \ge t(t-2)\{\mu_2(m,z,b) - \mu_2(0,0,0)\}.$$

Defining $\rho = \mu_2(m,z,b) - \mu_2(0,0,0) > 0$, we get

$$\mu_1(m,z,b) \le \mu_1(0,0,0) - t(t-2)\rho.$$

The design d can perform as well as d^* under the D-criterion only if

$$\mu_1(m,z,b)\mu_2(m,z,b)^{t-2} \ge \mu_1(0,0,0)\mu_2(0,0,0)^{t-2}.$$

The left-hand side is at most

$$f(\rho) = \{\mu_1(0,0,0) - t(t-2)\rho\}\{\mu_2(0,0,0)+\rho\}^{t-2}.$$

The derivative of $f(\rho)$ with respect to ρ equals

$$-t(t-2)\{\mu_2(0,0,0)+\rho\}^{t-2} + (t-2)\{\mu_1(0,0,0) - t(t-2)\rho\}$$
$$\times\{\mu_2(0,0,0)+\rho\}^{t-3}$$
$$= -(t-2)\{\mu_2(0,0,0)+\rho\}^{t-3}\{t(t-1)\rho + t\mu_2(0,0,0) - \mu_1(0,0,0)\}$$
$$< 0,$$

since $t\mu_2(0,0,0) > \sqrt{t}\,\mu_2(0,0,0) > \mu_1(0,0,0)$. This means that for $\rho > 0$, we get $f(\rho) < f(0) = \mu_1(0,0,0)\mu_2(0,0,0)^{t-2}$. □

We conclude this section with the proof of E-optimality for sufficiently large λ.

PROOF OF THEOREM 2.3. If we define $x \in \mathbb{R}^t$ with ith entry $t-1$ and all other entries equal to -1 and $y \in \mathbb{R}^t$ with the ith entry equal to 1, the jth entry equal to -1 and all others equal to 0, we can see as in Proposition 4.1 that:

(i) $\mu_{d,t-1} \le t/(t-1)c_{dii}$ for all $1 \le i \le t$.
(ii) $\mu_{d,t-1} \le \frac{1}{2}(c_{dii} + c_{djj} - 2c_{dij})$ for all $1 \le i < j \le t$.

Now assume there is a design $d \in \Omega_{t,n}$ which is better under the E-criterion than the Williams II(a) design d^*, i.e.,

$$\mu_{d,t-1} > \mu_{d^*,t-1} = r(1+\lambda^2) + \frac{2r\lambda}{t-1}.$$

We will determine the properties of d.

STEP 1. How often does each treatment appear in d? Without loss of generality, we can assume that treatment t is one which appears the least frequently. Then

$$r_{dt} \le r \quad \text{and} \quad c_{dtt} \le r_{dt}(1+\lambda^2) - \frac{(1-\lambda)^3}{n-(n-2)\lambda} r_{dt}^2.$$

Note that for $r_{dt} < r$, this bound is increasing in r_{dt}. Now (i) implies

$$\begin{aligned}\mu_{d,t-1} &\le \frac{t}{t-1} c_{dtt} \\ &\le r_{dt}(1+\lambda^2) + 2r_{dt}\lambda/(t-1) \\ &\quad + \frac{\{r_{dt}(1+\lambda^2)/(t-1)\}\{n - r_{dt}t - (n - r_{dt}t - 2)\lambda\}}{n-(n-2)\lambda}.\end{aligned}$$

Assume $r_{dt} \le r-1$. Then some algebra shows that

$$\mu_{d,t-1} < r(1+\lambda^2) + 2r\lambda/(t-1).$$

Thus we must have $\min\{r_{di}\} = r$ and, without loss of generality,

$$r_{d1} = r+1, \qquad r_{d2} = \cdots = r_{dt} = r.$$

STEP 2. How often are treatments i and j adjacent to each other? Assume that 2 and 3 are those treatments in $\{2,3,\dots,t\}$ with the smallest number of adjacencies between them. Then

$$-c_{d23} = N_{d23}\lambda + \frac{1-\lambda}{n-(n-2)\lambda}\{r(1-\lambda) + a_{d2}\lambda\}\{r(1-\lambda) + a_{d3}\lambda\}.$$

Using (ii), we get

$$\begin{aligned}\mu_{d,t-1} &\le \frac{1}{2}(c_{d22} + c_{d33} - 2c_{d23}) \\ &\le r(1+\lambda^2) + N_{d23}\lambda - \frac{1}{2}(a_{d2} + a_{d3})\lambda^2 \\ &\quad - \frac{1}{2}\frac{1-\lambda}{n-(n-2)\lambda}(a_{d2} - a_{d3})^2\lambda^2 \\ &\le r(1+\lambda^2) + N_{d23}\lambda.\end{aligned}$$

This means that treatments 2 and 3 must have at least $2r/(t-1)+1$ adjacencies.

Remember that $N_{d23} = \min\{N_{dij}\colon 2 \le i < j \le t\}$. This means that

$$\sum_{i=2}^{t-1} \sum_{j=i+1}^{t} N_{dij} \ge (t-2)r + (t-1)(t-2)/2.$$

How often are treatments $2, \dots, t$ adjacent to treatment 1? It can be seen that

$$\sum_{i=1}^{t} \sum_{j=i}^{t} N_{dij} = n - 1$$

and that

$$\sum_{j=2}^{t} N_{dij} + 2N_{d11} = 2r_{d1} - a_{d1} = 2r + 2 - a_{d1} = 2r + b.$$

As

$$\begin{aligned}\sum_{i=1}^{t} \sum_{j=i}^{t} N_{dij} &\ge \sum_{i=1}^{t-1} \sum_{j=i+1}^{t} N_{dij} + N_{d11} \\ &= \sum_{i=2}^{t-1} \sum_{j=i+1}^{t} N_{dij} + \sum_{j=2}^{t} N_{d1j} + N_{d11} \\ &\ge (t-2)r + (t-1)(t-2)/2 + 2r + b - N_{d11},\end{aligned}$$

it follows that

$$\begin{aligned}N_{d11} &\ge rt + b + (t-1)(t-2)/2 - (n-1) \\ &= (t-1)(t-2)/2 + b.\end{aligned}$$

STEP 3. What else must d satisfy? From (i), we get that

$$\mu_{d,t-1} \le \frac{t}{t-1} c_{d11} = \mu_1(m, z, b),$$

where $m \ge (t-1)(t-2)/2 + b$ and $z = 0$. Proposition 4.6 implies that

$$\begin{aligned}\mu_1(m, z, b) &\le \mu_1(0,0,0) + \frac{b}{t-1}\{(t-2)\lambda^2 + 4\lambda\} - t(t-2)\lambda - \frac{2bt}{t-1}\lambda \\ &\le \mu_1(0,0,0) - t(t-2)\lambda \\ &= \mu_2(0,0,0) + \frac{(t-1)(1-\lambda)^2(1+\lambda)r}{n-(n-2)\lambda} - t(t-2)\lambda\end{aligned}$$

and d can perform better than d^* only if

$$\frac{(t-1)r(1-\lambda)^2(1+\lambda)}{n-(n-2)\lambda} > t(t-2)\lambda.$$

This completes the proof. □

Acknowledgments. The authors wish to thank the referees for their helpful comments. We especially want to thank one referee for pointing out an error in an earlier draft of this paper and C.-S. Cheng for prior notice of his results, the use of which produced a simplification in the proof of Theorem 3.3.

REFERENCES

ATKINSON, A. C. (1969). The use of residuals as a concomitant variable. *Biometrika* **56** 33–41.

CHENG, C.-S. (1978). Optimality of certain asymmetrical experimental designs. *Ann. Statist.* **6** 1239–1261.

CHENG, C.-S. (1987). An optimization problem with applications to optimal design theory. *Ann. Statist.* **15** 712–723.

COX, D. R. (1952). Some recent work on systematic experimental designs. *J. Roy. Statist. Soc. Ser. B* **14** 211–229.

JACROUX, M. (1985). Some sufficient conditions for the type 1 optimality of block designs. *J. Statist. Plann. Inference* **11** 385–398.

KIEFER, J. (1961). Optimum experimental designs V, with applications to systematic and rotatable designs. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **1** 381–405. Univ. California Press.

KIEFER, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastava, ed.) 333–353. North-Holland, Amsterdam.

KIEFER, J. and WYNN, H. P. (1984). Optimum and minimax exact treatment designs for one-dimensional autoregressive error processes. *Ann. Statist.* **12** 431–450.

KUNERT, J. (1983). Optimal design and refinement of the linear model with applications to repeated measurements designs. *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. and MARTIN, R. J. (1987). Some results on the optimality properties of finite Williams designs. *Comm. Statist. A—Theory Methods.* To appear.

WHITTINGHILL, D. C. (1984). Block designs: General optimality results with applications to situations where balanced designs do not exist. Technical Report 84-21, Dept. Statistics, Purdue Univ.

WILLIAMS, R. M. (1952). Experimental design for serially correlated observations. *Biometrika* **39** 151–167.

FACHBEREICH IV—MATHEMATIK / STATISTIK
UNIVERSITÄT TRIER
POSTFACH 3825
D-5500 TRIER
WEST GERMANY

DEPARTMENT OF PROBABILITY
AND STATISTICS
UNIVERSITY OF SHEFFIELD
SHEFFIELD S3 7RH
ENGLAND