The Annals of Statistics
1995, Vol. 23, No. 6, 2081–2101

ON *G*-EFFICIENCY CALCULATION FOR POLYNOMIAL MODELS

BY HOLGER DETTE[1] AND WENG KEE WONG

Technische Universität Dresden and University of California, Los Angeles

We study properties of the variance function of the least squares estimator for the response surface. For polynomial models, we identify a class of approximate designs for which their variance functions are maximized at the extreme points of the design space. As an application, we examine robustness properties of D-optimal designs and D_{n-r}-optimal designs under various polynomial model assumptions. Analytic formulas for the G-efficiencies of these designs are derived, along with their D-efficiencies.

1. Introduction. This work examines a practical issue that sometimes arises in designing an experiment: what types of designs have their variance function maximized at the extreme points of the design space? Many allusions to this question have been raised informally in the literature, often expressed in statements like "the model is most strained near the extreme points of the design space." The implication is that model-based inference on the relationship between the covariates and the response variable becomes less reliable near the extreme points of the design space. One of our goals here is to identify a large class of designs for which their variance functions are maximized at the extreme points of the design space, and we show that many of the commonly used designs have this property. We do this for the case when we have a polynomial model with a single covariate and the design space Ω is assumed to be a given compact space. Generalizations to the case when there are several covariates are straightforward, especially if one considers product models.

The statistical model of interest is

$$y = f_j^T(x)\beta + e, \qquad x \in \Omega,$$

where y is the response, the regression function is $f_j^T(x) = (1, x, x^2, \dots, x^j)$, $\beta^T = (\beta_0, \beta_1, \dots, \beta_j)$ is the vector of model parameters and e is a random error with mean zero and constant variance, independent of x. Following Kiefer and Wolfowitz (1960), all designs considered in this paper are approximate or continuous, and so they are probability measures defined on Ω. This

Received March 1994; revised March 1995.
[1]Research partially supported by the Deutsche Forschungsgemeinschaft.
AMS 1991 *subject classifications*. Primary 62K05; secondary 65D30.
Key words and phrases. Approximate designs, canonical moments, D- and G-optimal designs, D_{n-r}-optimal designs, homoscedasticity, information matrix, orthogonal polynomials.

means if a given number n of uncorrelated observations are to be taken from the experiment, and a design ξ with mass m_i at $x_i \in \Omega, 1, 2, \dots, t$, is used, then approximately nm_i observations are taken at x_i, $i = 1, 2, \dots, t$. The set of all approximate designs on Ω is denoted by Ξ. For a given $f_j(x)$ and a given $\xi \in \Xi$, the information contained in ξ is measured by its information matrix:

$$M_j(\xi) = \int_\Omega f_j(x) f_j^T(x)\, d\xi(x).$$

Here and throughout, we focus attention only on designs whose information matrices are nonsingular. Such designs are called nonsingular.

For estimating model parameters, a popular criterion is D-optimality. Given $f_j(x)$, this criterion seeks a design ξ_j so that the determinant of the information matrix is maximized over Ξ, that is,

$$\xi_j = \arg\max_{\xi \in \Xi} |M_j(\xi)|.$$

Under the assumptions of homoscedasticity, the D-optimal design ξ_j is also G-optimal [Kiefer and Wolfowitz (1960), theorem]. This means ξ_j minimizes the maximum variance of the estimated response surface across Ω. Since the variance of the estimated response at the point x using design ξ is proportional to $d_j(x, \xi) = f_j^T(x) M_j(\xi)^{-1} f_j(x)$, this is equivalent to the assertion

$$\xi_j = \arg\min_{\xi \in \Xi} \max_{x \in \Omega} d_j(x, \xi).$$

G-optimality is particularly appealing when it is desired to estimate the entire response surface, as it provides global protection against unreliable estimates at points in Ω after the experiment is run.

Following standard convention, we compare the worth of a nonsingular design ξ by its efficiency. If ξ is an arbitrary nonsingular design ξ, the G- and D-efficiency of ξ are, respectively, given by

$$G_j(\xi) = \frac{j+1}{\max_{x \in \Omega} d_j(x, \xi)} \quad \text{and} \quad D_j(\xi) = \left\{ \frac{|M_j(\xi)|}{|M_j(\xi_j)|} \right\}^{1/j+1}.$$

All subsequent comparisons of design are based on either one of these measures.

There is a vast amount of work on D-optimal designs; the analytical formula of the D-optimal design ξ_j for the homoscedastic model is known, and properties of these designs are well studied [see Fedorov (1972), Kiefer (1985) and the references therein]. In particular, it is known that ξ_j has a minimal number of $j + 1$ support points so that the design ξ_j cannot be used to test if there is a lack of fit in the model. This drawback, however, may be overcome by using the optimal design for an expanded model if the resulting loss in efficiencies is not severe [Kendall and Stuart (1968), Kussmaul (1969) and Atkinson and Fedorov (1975a, b)]. The work here examines this issue under the G-optimality criterion for a class of designs.

Numerical work [see, e.g., Thibodeau (1977) and Wong (1994)] suggests that many popular designs have the property that their variance functions are maximized at the extreme points of the design space. The implication is that the G-efficiencies of these (nonsingular) designs can be readily determined. In this work, we formalize a method for identifying such designs and prove that these designs include the frequently used D-optimal designs and the D_{n-r}-optimal designs. The latter class of designs is introduced by Studden (1980, 1982) and is useful for estimating a subset of the parameters in a polynomial model. As an application, we examine how G- (and D-) efficiencies of these designs change with the degree of the polynomial. This is an important consideration since in practice the true model is often unknown and polynomial approximations are often used. In the process, we generalize Kussmaul's results [Kussmaul (1969)] and also prove Thibodeau's conjecture [Thibodeau (1977)] concerning the G-efficiency of ξ_n when the regression function is $f_j(x)$, $n > j$.

Our analysis relies heavily on the theory of canonical moments, which is a common tool for studying D-optimal designs [Lau and Studden (1985) and Studden (1980, 1982, 1989)]. Because canonical moments do not change when the designs are linearly transformed, we may, without loss of generality, assume the design space Ω to be $[-1, 1]$. Consequently, the D- and G-efficiency results here remain the same when Ω is any other compact interval.

The rest of the paper is organized as follows. Section 2 contains our main results. In Theorem 2.6, we illustrate how our results could be useful for heteroscedastic models as well. In Section 3, we apply our results to answer some of the issues raised earlier. This is followed by a brief discussion on applications to other fields in Section 4 and a summary in Section 5. Auxiliary results on canonical moments and proofs of the main results are given in the Appendix.

2. Main results. To our knowledge, the question of when a design has its variance function maximized at the extreme points of the design space has not been adequately addressed in the literature. The sufficient conditions stated in Theorems 2.1 and 2.2 provide partial answers that enable us to show that many popular designs have this property. Consequently, their G-efficiencies under various polynomial assumptions can be easily assessed.

Let ξ be a design defined on $\Omega = [-1, 1]$, and let $c_j = \int_{-1}^{1} x^j \, d\xi(x)$, $j = 1, 2, \dots,$ denote the jth (ordinary) moment of ξ. Define c_i^+ to be the maximum value of the ith moment for fixed $c_0, c_1, c_2, \dots, c_{i-1}$, and similarly define c_i^- to be the corresponding minimum. The canonical moments of ξ are defined by

$$p_i = \frac{c_i - c_i^+}{c_i^+ - c_i^-}, \qquad i = 1, 2, \dots .$$

Note that $0 \le p_i \le 1$. Whenever $c_i^+ = c_i^-$, the canonical moments are left undefined for $j > i$, and the sequence is terminated. It is well known [Skibinsky (1986)] that every probability measure on the interval $[-1, 1]$ is uniquely determined by its corresponding sequence of canonical moments. In what follows, it is helpful to define $q_i = 1 - p_i$, $i = 1, 2, \dots$.

THEOREM 2.1. *Let ξ be a given design on $\Omega = [-1, 1]$.*

(a) *If ξ is symmetric with canonical moments*

$$0 < p_{2j} \le \tfrac{1}{2}, \qquad j = 1, 2, \dots, k, \tag{2.0}$$

then

$$\max_{x \in \Omega} d_j(x, \xi) = d_j(1, \xi) = d_j(-1, \xi), \qquad j = 1, 2, \dots, k.$$

(b) *If ξ is not symmetric but satisfies* (2.0) *and*

$$0 < p_{2j-1} \le \tfrac{1}{2}, \qquad j = 1, 2, \dots, k,$$

then

$$\max_{x \in \Omega} d_j(x, \xi) = d_j(1, \xi), \qquad j = 1, 2, \dots, k.$$

(c) *If ξ is not symmetric but satisfies* (2.0) *and*

$$1 > p_{2j-1} \ge \tfrac{1}{2}, \qquad j = 1, 2, \dots, k,$$

then

$$\max_{x \in \Omega} d_j(x, \xi) = d_j(-1, \xi), \qquad j = 1, 2, \dots, k.$$

THEOREM 2.2. *Let ξ be a given design on $\Omega = [-1, 1]$ with canonical moments $p_j \in (0, 1)$, $j = 1, 2, \dots, 2k - 1$, and $p_{2k} \in (0, 1]$. If $k = 1$, define $p_0 = 0$.*

(a) *If ξ is symmetric and its canonical moments satisfy*

$$0 < \frac{1 - p_{2j}}{3 - 4p_{2j}} \le p_{2j+2} \qquad j = 1, 2, \dots, k - 2, \tag{2.1}$$

and

$$\frac{2p_{2k-2} - 1}{1 - p_{2k-2}} \le p_{2k}, \tag{2.2}$$

then

$$\max_{x \in \Omega} d_j(x, \xi) = d_j(-1, \xi) = d_j(1, \xi), \qquad j = 1, 2, \dots, k.$$

(b) *If the canonical moments of ξ satisfy*

$$\frac{2p_{2j} - 1}{1 - p_{2j}} \frac{1 - p_{2j+1}}{p_{2j+1}} \le \frac{2p_{2j+2} - 1}{p_{2j+2}}, \qquad j = 1, 2, \dots, k - 2, \tag{2.3}$$

$$\frac{2p_{2k-2} - 1}{1 - p_{2k-2}} \frac{1 - p_{2k-1}}{p_{2k-1}} \le p_{2k} \tag{2.4}$$

and

$$p_{2j-1} \le \tfrac{1}{2}, \qquad j = 1, 2, \dots, k \tag{2.5}$$

then

$$\max_{x \in \Omega} d_j(x, \xi) = d_j(1, x), \qquad j = 1, 2, \dots, k.$$

(c) *If the canonical moments of* ξ *satisfy*

$$\frac{2p_{2j} - 1}{1 - p_{2j}} \frac{p_{2j+1}}{1 - p_{2j+1}} \le \frac{2p_{2j+2} - 1}{p_{2j+2}}, \qquad j = 1, 2, \dots, k-2, \tag{2.6}$$

$$\frac{2p_{2k-2} - 1}{1 - p_{2k-2}} \frac{p_{2k-1}}{1 - p_{2k-1}} \le p_{2k} \tag{2.7}$$

and

$$p_{2j-1} \ge \tfrac{1}{2}, \qquad j = 1, 2, \dots, k, \tag{2.8}$$

then

$$\max_{x \in \Omega} d_j(x, \xi) = d_j(-1, \xi), \qquad j = 1, 2, \dots, k.$$

REMARK 2.0. It is worth mentioning that in general the bounds in Theorems 2.1 and 2.2 cannot be improved in the following sense. For every positive integer k, there exists a design with canonical moments satisfying all conditions of Theorem 2.1 (or Theorem 2.2) except one condition such that the variance function is not maximized at the extreme points of the interval $[-1, 1]$. As an illustration, consider the case $k = 3$ and a symmetric design with canonical moments $p_2 = \frac{4}{7}$, $p_4 = \frac{3}{5}$, and $p_6 \in (0, \frac{1}{2})$ which satisfies condition (2.1) but not (2.2). Straightforward calculation shows the quantities in (A.1) in the Appendix are $a_1 = a_2 = \frac{1}{4}$, $b_3 = 2/p_6$, $c_3 = q_6/2$, $f_0 = f_1 = f_2 = 0$ and the orthonormal polynomials with respect to the measures $d\xi(x)$ and $(1 - x^2)\, d\xi(x)$ are given by

$$P_1(x) = \frac{x}{\sqrt{p_2}}, \qquad P_2(x) = \frac{x^2 - p_2}{\sqrt{p_2 q_2 p_4}} \quad \text{and} \quad Q_2(x) = \frac{x^2 - p_2 q_4}{\sqrt{q_2 p_2 q_4 p_4 q_6}}$$

[see Lau (1983)]. Thus we obtain, from (A.0) and Lemma A.3,

$$\begin{aligned} d_3(x, \xi) = {} & \frac{5}{2p_6} + \left(1 - \frac{1}{2p_6}\right)\left\{1 + \frac{7}{4}x^2 + \frac{245}{36}\left(x^2 - \frac{4}{7}\right)^2\right\} \\ & - \frac{49}{72}\frac{25}{p_6}(1 - x^2)\left(x^2 - \frac{8}{35}\right)^2. \end{aligned}$$

It is now straightforward to verify that this function attains its maximum at an interior point of the interval $[-1, 1]$ whenever $p_6 < \frac{1}{2}$.

Note that, in practice, conditions in Theorems 2.1 and 2.2 are verified by first calculating the ordinary moments of the design and subsequently canonical moments are found in a standard way [see Karlin and Shapely (1953), page 59]. We now apply the above results to establish results for G-efficiencies of D-optimal designs and D_{n-r}-optimal designs under various polynomial model assumptions.

THEOREM 2.3. *Let ξ denote a design such that $p_j \in (0,1)$, $j = 1,\dots,2k-1$, and $p_{2k} \in (0,1]$. Then, for $j = 1,2,\dots,k$,*

$$d_j(1,\xi) = 1 + \sum_{i=1}^{j} \left(\prod_{m=1}^{i} \frac{q_{2m-1}}{p_{2m-1}} \prod_{m=1}^{i-1} \frac{q_{2m}}{p_{2m}} \right) \frac{1}{p_{2i}}$$

and

$$d_j(-1,\xi) = 1 + \sum_{i=1}^{j} \left(\prod_{m=1}^{i} \frac{p_{2m-1}}{q_{2m-1}} \prod_{m=1}^{i-1} \frac{q_{2m}}{p_{2m}} \right) \frac{1}{p_{2i}}.$$

If ξ is symmetric and $[r]$ denotes the largest integer less than or equal to r,

$$d_j(0,\xi) = 1 + \sum_{i=1}^{[j/2]} \left(\prod_{m=1}^{i} \frac{p_{4m-2}}{q_{4m-2}} \prod_{m=1}^{i-1} \frac{q_{4m}}{p_{4m}} \frac{1}{p_{4i}} \right), \qquad j = 1,2,\dots,k.$$

REMARK 2.1. If the regression function is $f_j(x)$, Theorem 2.3 provides an upper bound for the G-efficiency of a nonsingular design ξ in terms of its canonical moments:

$$G_j(\xi) \le \frac{j+1}{\max\bigl(d_j(1,\xi), d_j(-1,\xi)\bigr)}.$$

THEOREM 2.4. *Assume the regression function $f_n(x)$ is a polynomial of degree n.*

(a) *Let ξ_n denote the D-optimal design for $f_n(x)$. The G-efficiency of ξ_n for $f_j(x)$ is*

$$G_j(\xi_n) = \frac{n(j+1)}{n + 2nj - j^2}, \qquad j = 1,2,\dots,n.$$

(b) *Let ξ_{n,D_1} denote the optimal design for estimating the coefficient of x^n for $f_n(x)$. The G-efficiency of ξ_{n,D_1} for $f_j(x)$ is*

$$G_j(\xi_{n,D_1}) = \begin{cases} \dfrac{j+1}{2j+1}, & \text{if } 1 \le j \le n-1, \\[2ex] \dfrac{n+1}{2n}, & \text{if } j = n. \end{cases}$$

REMARK 2.2. Note that, from theorem 2.4(a), $G_1(\xi_n) = 2n/(3n-1)$ if $n \ge 1$, and $G_2(\xi_n) = 3n/(5n-4)$ if $n \ge 2$, which coincide with the results of Kussmaul (1969).

REMARK 2.3. In a numerical study of properties of robust designs, Thibodeau (1977) conjectured that

$$\max_{x\in\Omega} d_j(x,\xi_n) = d_j(1,\xi_n) = d_j(-1,\xi_n) = n+1-\frac{(n-j)^2}{n}.$$

Our proof of Theorem 2.4(a) in the Appendix will prove Thibodeau's conjecture as a by-product.

REMARK 2.4. It is interesting to note that both the D-optimal designs and the D_1-optimal designs are extreme cases in that their canonical moments satisfy the inequalities of Theorems 2.1 and 2.2 with equalities.

Next, we consider the class of D_{n-r}-optimal designs proposed by Studden (1980). For the regression function $f_n(x)$, he defined a D_{n-r}-optimal design as one which minimizes the determinant of the covariance matrix of the least squares estimates of the "highest" $n-r$ parameters, $\beta_{r+1}, \beta_{r+2}, \dots, \beta_n$. Note that (i) when $r=0$, the design $\xi_{n,D_{n-r}}$ becomes ξ_{n,D_n}, which coincides with the D-optimal design ξ_n for $f_n(x)$ in Theorem 2.4(a) [Studden (1980)] and (ii) when $r=n-1$, $\xi_{n,D_{n-r}}$ reduces to ξ_{n,D_1} in Theorem 2.4(b). The next result generalizes the case to any values of r between 0 and $n-1$.

THEOREM 2.5. *Let $0\le r\le n-1$ and let $\xi_{n,D_{n-r}}$ denote the D_{n-r}-optimal design for the polynomial regression function $f_n(x)$ of degree n. The G-efficiency of ξ_{n,D_n-r} for $f_j(x)$ is*

$$G_j(\xi_{n,D_n-r}) = \begin{cases} \dfrac{j+1}{2j+1}, & \text{if } 1\le j\le r,\\[2ex] \dfrac{j+1}{n+1+r-(n-j)^2/(n-r)}, & \text{if } r+1\le j\le n. \end{cases}$$

From the proof of this theorem in the Appendix, it will be apparent that the conclusions in Theorems 2.1 and 2.2 hold as long as each of the first $2k$ canonical moments of the design satisfies either one of the conditions in Theorem 2.1 or Theorem 2.2. Thus, the conditions in these two theorems are not as stringent as they appear to be.

Our next result may be used to assess the loss in G-efficiency when we erroneously assume heteroscedasticity is present in the model. Following Fedorov [(1972), page 39], we represent the heteroscedasticity by an efficiency function $\lambda(x)$. This function is positive and its value at the point x is inversely proportional to the variance of the response at the point x. The interest here is the loss in efficiency if we determine the D-optimal design assuming the efficiency function is $\lambda(x)=(1+x)^{\alpha+1}(1-x)^{\beta+1}$, $\alpha>-1$ and $\beta>-1$, when in reality the efficiency function is constant across $\Omega=[-1,1]$. The G-efficiency of the (heteroscedastic) D-optimal design $\xi_n^{(\alpha,\beta)}$ under a homoscedastic model is now given. A numerical example is worked out in Section 3.

THEOREM 2.6. *Suppose* $\Omega = [-1, 1]$, *the regression function* $f_j(x)$ *is a polynomial of degree* j *and the efficiency function* $\lambda(x)$ *is constant across* Ω. *Let* $\xi_n^{(\alpha,\beta)}$ *denote the (heteroscedastic) D-optimal design for* $f_n(x)$ *assuming* $\lambda(x) = (1+x)^{\alpha+1}(1-x)^{\beta+1}$, $\alpha > -1$ *and* $\beta > -1$. *Then the variance function for* $\xi_n^{(\alpha,\beta)}$ *satisfies*

$$\max_{x\in\Omega} d_j\big(x, \xi_n^{(\alpha,\beta)}\big) = d_j\big(-1, \xi_n^{(\alpha,\beta)}\big) \quad \text{if } \alpha \le \beta,$$

and

$$\max_{x\in\Omega} d_j\big(x, \xi_n^{(\alpha,\beta)}\big) = d_j\big(-1, \xi_n^{(\alpha,\beta)}\big) \quad \text{if } \alpha \ge \beta.$$

Furthermore, for $j = 1, 2, \dots, n$,

$$d_j\big(1, \xi_n^{(\alpha,\beta)}\big) = 1 + \sum_{i=1}^{j} \frac{(\beta+n+2-i)_i}{(\alpha+n+2-i)_i} \frac{(\alpha+\beta+3+n-i)_{i-1}}{(n-i+1)_i} \times\{\alpha+\beta+3+2n-2i\}$$

and

$$d_j\big(-1, \xi_n^{(\alpha,\beta)}\big) = 1 - \sum_{i=1}^{j} \frac{(\alpha+n+2-i)_i}{(\beta+n+2-i)_i} \frac{(\alpha+\beta+3+n-i)_{i-1}}{(n-i+1)_i} \times\{\alpha+\beta+3+2n-2i\},$$

where we have used the notation $(a)_0 = 1$ *and* $(a)_k = a(a+1)\cdots(a+k-1)$.

REMARK 2.5. The heteroscedastic D-optimal design $\xi_n^{(\alpha,\beta)}$ is well known [Fedorov (1972), page 89].

For the sake of comparison, we now state the D-efficiencies of ξ_n and $\xi_{n,D_{n-r}}$ for the regression function $f_j(x)$, $n \ge j \ge 1$. The proof of these results are omitted since they can be deduced from Studden (1980) or Lau (1983).

THEOREM 2.7. *Let* $f_j(x)$ *denote the polynomial regression function of degree* j. *For* $n \ge j \ge 1$, *we have the following*:

(a) *The D-efficiency of* ξ_n *for* $f_j(x)$ *is given by*

$$D_j(\xi_n)^{j+1} = \prod_{i=2}^{j} \left\{ \left(\frac{n-i+1}{j-i+1}\right)^2 \frac{(2j-2i+1)}{(2n-2i+1)} \frac{(2j-2i+3)}{(2n-2i+3)} \right\}^{j-i+1} \times \left(\frac{n}{j}\frac{2j-1}{2n-1}\right)^j.$$

(b) *The D-efficiency of* $\xi_{n,D_{n-r}}$ *for* $f_j(x)$ *is given by*

$$D_j\big(\xi_{n,D_{n-r}}\big)^{j+1} = 2^{-j^2}\left(\frac{2j-1}{j}\right)^j \prod_{i=2}^{j} \left\{ \frac{(2j-2i+1)(2j-2i+3)}{(j-i+1)^2} \right\}^{j-i+1}$$

if $j \leq r$; *otherwise*, *the right-hand side expression is replaced by*

$$2^{-2jr+r^2}\left(\frac{n-r}{2n-2r-1}\right)^{j-r}\left(\frac{2j-1}{j}\right)^{j}$$
$$\times \prod_{i=2}^{r+1}\left\{\frac{(2j-2i+1)(2j-2i+3)}{(j-i+1)^2}\right\}^{j-i+1}$$
$$\times \prod_{i=r+2}^{j}\left\{\left(\frac{n-i+1}{j-i+1}\right)^2\frac{(2j-2i+1)(2j-2i+3)}{(2n-2i+1)(2n-2i+3)}\right\}^{j-i+1}.$$

Theorem 2.7(a) yields, for example, for $j = 1$, 2 and 3,

$$D_1(\xi_n) = \left\{\frac{n}{2n-1}\right\}^{1/2}, \qquad D_2(\xi_n) = \frac{3}{2n-1}\left\{\frac{n^2(n-1)^2}{8n-12}\right\}^{1/3} \tag{2.9}$$

and

$$D_3(\xi_n) = 2.5(n-1)(n-2)^{1/2}\left\{\frac{5n^3}{(2n-5)(2n-3)^3(2n-1)^5}\right\}^{1/4}.$$

Letting ξ_∞ denote the limiting design of ξ_n [which exists by Kiefer and Studden (1976)], we have

$$G_j(\xi_\infty) = \frac{j+1}{1+2j} \tag{2.10}$$

and

$$D_j(\xi_\infty)^{j+1} = \frac{1}{j!}2^{-j^2}j^j\prod_{i=1}^{j}\left(\frac{2j-2i+1}{j-i+1}\right)^{2j-2i+1}. \tag{2.11}$$

The G-efficiency follows directly from Theorem 2.4, and its D-efficiency follows from Theorem 2.7(a) after some algebra. Theorem 2.3 of Kiefer and Studden (1976), which is expressed in terms of partial sums of the zeta function, is a complicated version of (2.11). The expression in (2.11) has the advantage that it is more compactly written and is numerically more efficient to compute.

3. Examples and applications. We discuss some practical implications of the results in the previous section in designing an experiment. Suppose the relationship between the true expected response and a covariate x is a polynomial. Since the degree of the polynomial is often not known, it is prudent to choose a design which is robust to polynomial assumptions. Ideally, we would like to have a design that remains efficient for moderate changes in the assumed degree of the polynomial.

For G-efficiency, it is clear from Theorem 2.4(a) that, for fixed n $G_j(\xi_n)$ is a monotonic increasing function of j, provided $j^2 + 2j \geq n \geq j$, which is true for most practical cases. Furthermore, if the true regression function is a polynomial of degree j, the G-efficiency of using ξ_{j+h} (instead of ξ_j) is at least g provided

$$0 \le h \le \frac{(1-g)(j^2+j)}{(2j+1)g-1-j}, \qquad j \ge 1, \tfrac{2}{3} \le g \le 1. \tag{3.0}$$

The constraint $g \ge \frac{2}{3}$ ensures the denominator is positive for all $j \ge 1$; this restriction is reasonable since designs with high efficiency are sought. When $g = 1$ in (3.0), h is 0, confirming the uniqueness of the G-optimal design. By Atwood's inequality [Atwood (1969)], ξ_{j+h} has a D-efficiency of at least g when the regression function is $f_j(x)$ and h satisfies (3.0).

Further implications in terms of the loss in G- and D-efficiencies of using ξ_n when the regression function is $f_j(x)$, $n > j$, can be evaluated by applying Theorems 2.4 and 2.7. Since the practical cases of interest are typically when $n = j + 1$, $j + 2$ and possibly $j + 3$ (moderate changes in the assumed degree of the polynomial model), we evaluate these cases by substituting n for one of these values in Theorem 2.7. The resulting expressions are all monotonic functions of j: $D_j(\xi_{j+1}) \ge 0.8165$, $D_j(\xi_{j+2}) \ge 0.7746$ and $D_j(\xi_{j+3}) \ge 0.7559$ for all $j \ge 1$ with equality at $j = 1$. When $j \ge 20$, $D_j(\xi_{j+k}) \ge 0.9666$ for $1 \le k \le 3$. The practical implication here is that the D-optimal design ξ_n remains relatively efficient for the model $f_j(x)$ as long as $1 \le j \le n \le j + 3$. Similar conclusions are obtained for the G-efficiency: $G_j(\xi_{j+1}) \ge 0.8100$, $G_j(\xi_{j+2}) \ge 0.7500$ and $G_j(\xi_{j+3}) \ge 0.7273$ for all $j \ge 1$ with equality at $j = 1$. When $j \ge 20$, $G_j(\xi_{j+k}) \ge 0.8895$ for $1 \le k \le 3$. As in the case of D-efficiency, these calculations suggest that correct model specification becomes increasingly less important if one uses any ξ_n's, as long as $n \ge j$ and j is sufficiently large.

Atwood(1969), Thibodeau (1977) and, recently, Wong (1994) tabulated the D- and G-efficiencies for selected cases studied here. Theorems 2.4 and 2.7 generalize their numerical results and also may be combined to express the D-efficiency of ξ_n for $f_j(x)$ in terms of its G-efficiency and vice versa. For instance, if $n \ge j = 1$, (2.9) yields

$$D_1(\xi_n) = \frac{3n-1}{2n}\left\{\frac{n}{2n-1}\right\}^{1/2} G_1(\xi_n).$$

For the design ξ_∞, it can be verified that both $-D_j(\xi_\infty)$ and $G_j(\xi_\infty)$ are monotonically decreasing functions in j. A direct calculation shows $G_j(\xi_\infty) =$ 0.67, 0.60 and 0.57 for $j = 1$, 2 and 3, respectively, and decreases in the limit to 0.5. Also, if $n \ge j \to \infty$ in such a way that $0 < j/n = r \le 1$, the limiting value of $G_j(\xi_n)$ is $1/(2 - r)$. In contrast, (2.10) yields $D_j(\xi_\infty) = 0.71$, 0.75 and 0.79 for $j = 1$, 2 and 3, respectively, and equals 1 in the limit. Thus, if the true regression function is a polynomial of degree 3, say, then the $D(G)$-efficiency of ξ_n is at least 0.79 (0.50) as long as $n \ge 3$.

Similar deductions can be made for the D_{n-r}-optimal designs but, for space consideration, their D- and G-efficiencies for the regression function $f_j(x)$ are displayed in Table 1 for selected values of j, r and n. Since ξ_n and ξ_{D_n} are the same design [Studden (1980)], this table includes results for the D-optimal design for all the parameters as well. It is clear from Table 1 that, for fixed n and j, the G- and D-efficiencies of the design $\xi_{D_{n-r}}$ increase as r decreases. However, for fixed j and r $(\leq n)$, both the G- and D-efficiencies of $\xi_{D_{n-r}}$ do not change if n is sufficiently large. Other properties of these efficiencies can be deduced from the table. Again, by Atwood's inequality [Atwood (1969)], note that the D-efficiencies always exceed the G-efficiencies.

We now give an example to illustrate the use of Theorem 2.6 in practice. Consider, for example, the case when $\Omega = [-1, 1]$, the regression function is $f_n(x)$ and the D-optimal design for the efficiency $\lambda(x) = (1+x)^{\alpha+1}(1-x)^{\beta+1}$, $\alpha > -1$, $\beta > -1$, is used in the homoscedastic model. To see how the G-efficiency is affected, we discuss two special cases: (i) $\alpha = \beta = -\frac{1}{2}$, with $\lambda(x) = (1-x^2)^{1/2}$; and (ii) $\alpha = \beta = \frac{1}{2}$, with $\lambda(x) = (1-x^2)^{3/2}$. Other situations can be treated similarly. A straightforward calculation shows, for case (i),

$$G_j\big(\xi_n^{(-1/2,-1/2)}\big) = \frac{j+1}{2j+1}, \qquad j = 1, 2, \ldots, n,$$

TABLE 1

$G(D)$-efficiencies of the optimal designs, $\xi_{n,D_{n-r}}$ for $f_j(x)$, $2 \leq j \leq 6$, $1 \leq r \leq j$, $2 \leq n \leq 7$

		n					
j	$n-r$	2	3	4	5	6	7
2	1	0.750 (0.750)	0.600 (0.750)	0.600 (0.750)	0.600 (0.750)	0.600 (0.750)	0.600 (0.750)
	2	1.000 (1.000)	0.667 (0.826)	0.600 (0.750)	0.600 (0.750)	0.600 (0.750)	0.600 (0.750)
3	1	0	0.667 (0.786)	0.571 (0.786)	0.571 (0.786)	0.571 (0.786)	0.571 (0.786)
	2	0	0.800 (0.975)	0.615 (0.845)	0.571 (0.786)	0.571 (0.786)	0.571 (0.786)
	3	0	1.000 (1.000)	0.706 (0.875)	0.600 (0.823)	0.571 (0.786)	0.571 (0.786)
4	1	0	0	0.625 (0.934)	0.556 (0.813)	0.556 (0.813)	0.556 (0.813)
	2	0	0	0.714 (0.966)	0.588 (0.861)	0.556 (0.813)	0.556 (0.813)
	3	0	0	0.833 (0.986)	0.652 (0.886)	0.517 (0.843)	0.556 (0.813)
	4	0	0	1.000 (1.000)	0.741 (0.902)	0.625 (0.862)	0.571 (0.835)
5	1	0	0	0	0.600 (0.936)	0.546 (0.834)	0.546 (0.834)
	2	0	0	0	0.667 (0.963)	0.571 (0.875)	0.546 (0.834)
	3	0	0	0	0.750 (0.979)	0.621 (0.896)	0.563 (0.834)
	4	0	0	0	0.857 (0.991)	0.706 (0.909)	0.600 (0.876)
	5	0	0	0	1.000 (1.000)	0.857 (0.920)	0.652 (0.888)
6	1	0	0	0	0	0.583 (0.939)	0.539 (0.850)
	2	0	0	0	0	0.636 (0.962)	0.560 (0.886)
	3	0	0	0	0	0.700 (0.976)	0.600 (0.904)
	4	0	0	0	0	0.778 (0.986)	0.651 (0.916)
	5	0	0	0	0	0.875 (0.994)	0.714 (0.925)
	6	0	0	0	0	1.000 (1.000)	0.793 (0.932)

which is the same as the D_1-optimal design except when $j = n$. For case (ii),

$$G_j\big(\xi_n^{(1/2,1/2)}\big) = \frac{(n-j+1)(n-j+2)(j+1)}{2(n-j)(j-1)(n+1)+(3n+5)n+2}, \qquad j = 1,2,\dots,n.$$

If we specialize to the cases when $j = 1$ and n,

$$G_1\big(\xi_n^{(1/2,1/2)}\big) = \frac{2n}{3n+2} \quad \text{and} \quad G_n\big(\xi_n^{(1/2,1/2)}\big) = \frac{2}{3n+2}.$$

Observe now the high cost in terms of G-efficiency of the erroneous assumption of heteroscedasticity. If $n = 2$, $G_2(\xi_2^{(1/2,1/2)}) = 0.25$ and if $n = 3$, $G_3(\xi_3^{(1/2,1/2)}) = 0.182$. The implication is that one should be very careful about the heteroscedastic assumption since use of the heteroscedastic optimal design for the homoscedastic model can result in very severe loss in G-efficiency.

4. Further applications. The results stated in Section 2 are closely related to some problems associated with the Gauss–Jacobi quadrature and we will indicate some of these relations very briefly here. The interested reader is referred to the paper by Nevai (1986), which provides an excellent overview on this topic. See also Freud (1972) and Nevai (1986) for important applications of the Gauss–Jacobi quadrature in numerical integration and approximation theory.

On the compact interval $[-1,1]$, the Christoffel function of order n, with respect to a given measure $d\xi(x)$, is defined by

$$\lambda_n(d\xi, x) = \min\left\{ \int_{-1}^{1} |\pi(t)|^2 \, d\xi(t) \,\middle|\, \pi(t) \text{ is a polynomial of degree less than or equal to } n-1 \text{ and } \pi(x) = 1 \right\}.$$

If $P_n(x)$ denotes the nth orthonormal polynomial with respect to the measure $d\xi(x)$ and $x_1, x_2, \dots, x_n$ are the zeros of $P_n(x)$, then

$$\sum_{i=1}^{n} \lambda_n(d\xi, x_i)\pi(x_i) = \int_{-1}^{1} \pi(x)\, d\xi(x)$$

for all polynomials of degree $2n-1$ (i.e., the Gauss—Jacobi quadrature with knot at x_i and weight $\lambda_n(d\xi_i, x_i)$, $i = 1,2,\dots,n$, integrates these polynomials exactly). It is well known that

$$\{\lambda_n(d\xi, x)\}^{-1} = \sum_{i=0}^{n-1} P_i^2(x) = d_{n-1}(x, \xi),$$

and, consequently, the results of Section 2 state sufficient conditions under which the Christoffel function is minimized at the extreme points of the interval $[-1, 1]$. There are several results in the literature addressing this issue, but all of them are motivated primarily from the monotonic properties of the Christoffel functions. The approach of this paper in addressing this question is new and has the advantages that it avoids the common assumption that $d\xi(x)$ has to be an absolute continuous measure. Consequently, our results are applicable to discrete measures as well.

5. Summary. We gave sufficient conditions where the variance function of a continuous design is maximized at the extreme points of the design space. These results are applied to study the D- and G-efficiencies of D- and D_{n-r}-optimal designs when there is uncertainty in the degree of the polynomial model. Applications of our results to numerical analysis and approximation theory are also briefly noted.

In this paper, our attention has been confined to D- and G-efficiencies. Other measures of efficiencies, such as A- and E-efficiencies, could also be studied. However, it appears difficult to obtain analogous analytical results for the A- and E-efficiencies. A reason for this is that A- and E-optimal designs cannot be described in a nice closed form like those of D- and G-optimal designs. See Wong (1994) for numerical results for A- and E-efficiencies in selected cases under the setting considered here.

APPENDIX

Auxiliary results and proofs. Here we state several auxiliary results on canonical moments. Let $|M^*(\xi)|$ denote the determinant of an $(m+1) \times (m+1)$ information matrix

$$M^*(\xi) = \int_{-1}^{1} \lambda(x) f_n(x) f_n^T(x)\, d\xi(x)$$

in a weighted polynomial regression with efficiency function $\lambda(x)$. Note that the choice $\lambda(x) = 1$ gives the homoscedastic case considered in Sections 1–4. Recalling that c_i denotes the ith moment of a design ξ, it is easy to see that

$$\underline{D}_{2m}(\xi) = |c_{i+j}|_{0 \le i, j \le m}, \qquad \overline{D}_{2m}(\xi) = |c_{i+j} - c_{i+j+2}|_{0 \le i, j \le m-1},$$

$$\underline{D}_{2m+1}(\xi) = |c_{i+j} + c_{i+j+1}|_{0 \le i, j \le m}, \quad \overline{D}_{2m+1}(\xi) = |c_{i+j} - c_{i+j+1}|_{0 \le i, j \le m}$$

are the determinants of the information matrices for weighted polynomial regression with efficiency functions $\lambda(x) = 1$, $\lambda(x) = 1 - x^2$, $\lambda(x) = 1 + x$ and $\lambda(x) = 1 - x$, respectively. In terms of the canonical moments, they are given by the following lemma.

LEMMA A.1 [Lau and Studden (1985)]. *We have*

$$\underline{D}_{2m}(\xi) = 2^{m(m+1)} \prod_{i=1}^{m} (\zeta_{2i-1}\zeta_{2i})^{m+1-i},$$

$$\overline{D}_{2m}(\xi) = 2^{m(m+1)} \prod_{i=1}^{m} (\gamma_{2i-1}\gamma_{2i})^{m+1-i},$$

$$\underline{D}_{2m+1}(\xi) = 2^{(m+1)^2} \prod_{i=0}^{m} (\zeta_{2i}\zeta_{2i+1})^{m+1-i},$$

$$\overline{D}_{2m+1}(\xi) = 2^{(m+1)^2} \prod_{i=0}^{m} (\gamma_{2i}\gamma_{2i+1})^{m+1-i},$$

where $\zeta_0 = 1$, $\zeta_1 = p_1$, $\gamma_0 = 1$, $\gamma_1 = q_1$, $\zeta_j = q_{j-1}p_j$, $\gamma_j = p_{j-1}q_j$, $j \geq 2$, $q_j = 1 - p_j$ *and* $\{p_j\}_{j\geq 1}$ *are the canonical moments of* ξ.

There are two results that will be used repeatedly.

1. If ξ has canonical moments $p_j \in (0,1)$, $j = 1, 2, \dots, 2k$, then ξ has at least $k+1$ support points [see Karlin and Shapely (1953) or Karlin and Studden (1966)]. Consequently, $M_j(\xi)$ is nonsingular for $j = 1, 2, \dots, k$.
2. Let $P_0(x), \dots, P_k(x)$ denote the orthonormal polynomials with respect to the measure $d\xi(x)$, let $\hat{P}_k(x)^T = (P_0(x), \dots, P_k(x))$ and let A be a nonsingular matrix such that $f_k(x) = A\hat{P}_k(x)$. Then the variance function of ξ is given by

$$(\text{A.0}) \qquad \begin{aligned} d_k(x,\xi) &= f_k^T(x)\left\{A\int_{-1}^{1} \hat{P}_k(x)\hat{P}_k(x)^T \, d\xi(x)\, A^T\right\}^{-1} f_K(x) \\ &= \hat{P}_k(x)^T\hat{P}_k(x) = \sum_{i=0}^{k} P_i^2(x). \end{aligned}$$

Therefore the discussion of variance functions is intimately related to the properties of the orthonormal polynomials $P_i(x)$'s with respect to the measure $d\xi(x)$. The proof of the next lemma is straightforward and therefore omitted.

LEMMA A.2. *Let* $\{s_j\}_{1\leq j\leq k}$ *be real numbers, and let* $\{P_j(x)\}_{1\leq j\leq k}$ *denote the orthornormal polynomials with respect to the probability measure* $d\xi(x)$. *Then the variance functions* $d_1(x,\xi), \dots, d_k(x,\xi)$ *satisfy*

$$\sum_{j=1}^{m-1} s_j P_j^2(x) = -s_1 + \sum_{j=1}^{m-2} (s_j - s_{j+1}) d_j(x,\xi) + s_{m-1} d_{m-1}(x,\xi), \qquad 2 \leq m \leq k+1.$$

One of the key steps for obtaining conditions for which the maximum of the variance function is attained at the extreme points of the design space is described in the next lemma.

LEMMA A.3. *Let ξ denote a probability measure on the interval $[-1,1]$ with canonical moments $p_j \in (0,1)$, $j = 1,2,\dots,2k-1$, $p_{2k} \in (0,1]$. Then the variance functions $d_1(x,\xi), d_2(x,\xi),\dots,d_k(x,\xi)$ satisfy the recursive relation*

$$\begin{aligned} d_{m+1}(x,\xi) = (1+a_1)b_{m+1} &+ (1-a_m b_{m+1})d_m(x,\xi) \\ &-(1-x^2)b_{m+1}c_{m+1}Q_m^2(x) \\ &+ b_{m+1}\sum_{j=1}^{m-1}(a_{j+1}-a_j)d_j(x,\xi) - (1-x)b_{m+1}\sum_{j=0}^{m} f_j S_j^2(x), \\ &\qquad m = 1,2,\dots,k-1, \end{aligned}$$

where $\{S_j(x)\}_{0\le j\le k-1}$ and $\{Q_j(x)\}_{1\le j\le k-1}$ are the orthonormal polynomials with respect to the measures $(1-x)\,d\xi(x)$ and $(1-x^2)\,d\xi(x)$, respectively. Here,

$$a_m = \frac{p_{2m-1}}{q_{2m-1}}\prod_{j=1}^{m-1}\left\{\frac{p_{2j-1}}{q_{2j-1}}\frac{q_{2j}}{p_{2j}}\right\}\left\{1-\frac{q_{2m}}{p_{2m}}\right\}, \qquad m = 1,2,\dots,k-1,$$

$$\text{(A.1)}\quad b_m = \frac{1}{p_{2m}}\frac{q_{2m-1}}{p_{2m-1}}\prod_{j=1}^{m-1}\left\{\frac{q_{2j-1}}{p_{2j-1}}\frac{p_{2j}}{q_{2j}}\right\}, \qquad m = 1,2,\dots,k,$$

$$c_m = q_{2m}\frac{p_{2m-1}}{q_{2m-1}}\prod_{j=1}^{m-1}\left\{\frac{p_{2j-1}}{q_{2j-1}}\frac{q_{2j}}{p_{2j}}\right\}, \qquad m = 1,2,\dots,k,$$

and

$$f_m = \prod_{j=1}^{m}\frac{p_{2j-1}}{q_{2j-1}}\frac{q_{2j}}{p_{2j}}\left\{1-\frac{p_{2m+1}}{q_{2m+1}}\right\}, \qquad m = 0,1,\dots,k-1.$$

PROOF. Using Theorem 4.1(a) in Dette (1993), it follows that the orthonormal polynomials $P_j(x)$, $Q_j(x)$ and $S_j(x)$ with respect to the measures $d\xi(x)$, $(1-x^2)\,d\xi(x)$ and $(1-x)\,d\xi(x)$ satisfy the following identities, $m = 1,2,\dots,k$:

$$\text{(A.2)}\quad \begin{aligned} &\sum_{j=1}^{m-1}\frac{\underline{D}_{2j-1}(\xi)}{\overline{D}_{2j-1}(\xi)}\left\{\frac{\overline{D}_{2j-2}(\xi)}{\underline{D}_{2j-2}(\xi)} - \frac{\overline{D}_{2j}(\xi)}{\underline{D}_{2j}(\xi)}\right\}P_j^2(x) \\ &\quad + \frac{\overline{D}_{2m-2}(\xi)}{\underline{D}_{2m-2}(\xi)}\frac{\underline{D}_{2m}(\xi)}{\underline{D}_{2m}^{+}(\xi)}\frac{\underline{D}_{2m-1}(\xi)}{\overline{D}_{2m-1}(\xi)}P_m^2(x) \\ &\quad + (1-x)\sum_{j=0}^{m-1}\frac{\overline{D}_{2j}(\xi)}{\underline{D}_{2j}(\xi)}\left\{\frac{\underline{D}_{2j-1}(\xi)}{\overline{D}_{2j-1}(\xi)} - \frac{\underline{D}_{2j+1}(\xi)}{\overline{D}_{2j+1}(\xi)}\right\}S_j^2(x) \\ &= 1 + (x^2-1)\frac{\underline{D}_{2m-1}(\xi)}{\overline{D}_{2m-1}(\xi)}\frac{\overline{D}_{2m}(\xi)}{\underline{D}_{2m}^{+}(\xi)}Q_{m-1}^2(x). \end{aligned}$$

where $\underline{D}^{+}_{2m}(\xi) = \underline{D}_{2m}(\xi) + [\underline{D}_{2m-2}(\xi)/\overline{D}_{2m-2}(\xi)]\overline{D}_{2m}(\xi)$ [Karlin and Shapely (1953), page 59]. By Lemmas A.1 and A.2, (A.2) can be rewritten as

$$\begin{aligned} P_m^2(x) &= b_m\Bigg\{1 - (1-x^2)c_m Q_{m-1}^2(x) \\ &\qquad - \sum_{j=1}^{m-1} a_j P_j^2(x) - (1-x)\sum_{j=0}^{m-1} f_j S_j^2(x)\Bigg\} \\ &= b_m\Bigg\{1 + a_1 + \sum_{j=1}^{m-2}(a_{j+1}-a_j)d_j(x,\xi) - a_{m-1}d_{m-1}(x,\xi) \\ &\qquad - (1-x^2)c_m Q_{m-1}^2(x) - (1-x)\sum_{j=0}^{m-1} f_j S_j^2(x)\Bigg\}, \end{aligned}$$

where the quantities a_j, b_j, c_j and f_j are defined in (A.1). Using Lemma A.2 and (A.0), we obtain the following relationship for the variance functions:

$$\begin{aligned} d_{m+1}(x,\xi) &= \sum_{j=0}^{m+1} P_j^2(x) = d_m(x,\xi) + P_{m+1}^2(x) \\ &= (1 - a_m b_{m+1}) d_m(x,\xi) \\ &\quad + b_{m+1}\sum_{j=1}^{m-1}(a_{j+1}-a_j)d_j(x,\xi) + (1+a_1)b_{m+1} \\ &\quad - b_{m+1}c_{m+1}(1-x^2)Q_m^2(x) - b_{m+1}(1-x)\sum_{j=0}^{m} f_j S_j^2(x), \\ &\qquad\qquad m = 1,2,\dots,k-1. \end{aligned}$$

This proves Lemma A.3. □

Proof of Theorem 2.1. Consider case (b) and let $\{S_j(x)\}_{0\le j\le k-1}$, $\{Q_j(x)\}_{0\le j\le k-1}$ and $\{P_j(x)\}_{0\le j\le k}$ denote the orthonormal polynomials with respect to the measures $(1-x)\,d\xi(x)$, $(1-x^2)\,d\xi(x)$, and $d\xi(x)$, respectively. By Theorem 4.1(c) of Dette (1993), these polynomials satisfy for $j = 0,1,\dots,k-1$, the identity

$$\begin{aligned} &(1-x^2)\sum_{m=0}^{j-1}\frac{\underline{D}_{2m+1}(\xi)}{\overline{D}_{2m+1}(x)}\left\{\frac{\underline{D}_{2m}(\xi)}{\overline{D}_{2m}(\xi)} - \frac{\underline{D}_{2m+2}(\xi)}{\overline{D}_{2m+2}(\xi)}\right\}Q_m^2(x) \\ &\quad + (1-x^2)\frac{\underline{D}_{2j}(\xi)}{\overline{D}_{2j}(\xi)}\frac{\underline{D}_{2j+1}(\xi)}{\overline{D}_{2j+1}(\xi)}\frac{\overline{D}_{2j+2}(\xi)}{\overline{D}^{-}_{2j+2}(\xi)}\frac{\overline{D}_{2j+2}(\xi)}{\overline{D}^{-}_{2j+2}(\xi)}Q_j^2(x) \\ &\quad + (1-x)\sum_{m=0}^{j}\frac{\underline{D}_{2m}(\xi)}{\overline{D}_{2m}(\xi)}\left\{\frac{\underline{D}_{2m-1}(\xi)}{\overline{D}_{2m-1}(\xi)} - \frac{\underline{D}_{2m+1}(\xi)}{\overline{D}_{2m+1}(\xi)}\right\}S_m^2(x) \\ &= 1 - \frac{\underline{D}_{2j+1}(\xi)}{\overline{D}_{2j+1}(\xi)}\frac{\underline{D}_{2j+2}(\xi)}{\overline{D}^{-}_{2j+2}(\xi)}P_{j+1}^2(x) \end{aligned} \tag{A.3}$$

and

$$\overline{D}^{-}_{2j+2}(\xi) = \overline{D}_{2j+2}(\xi) + \frac{\underline{D}_{2j+2}(\xi)}{\underline{D}_{2j}(\xi)}\overline{D}_{2j}(\xi) \tag{A.4}$$

[see Karlin and Shapely (1953), page 59]. By an application of Lemma A.1 and assumption (2.0), we obtain, for $m = 1, 2, \dots, k-1$,

$$\frac{\underline{D}_{2m+1}(\xi)}{\overline{D}_{2m+1}(\xi)}\left\{\frac{\underline{D}_{2m}(\xi)}{\overline{D}_{2m}(\xi)} - \frac{\underline{D}_{2m+2}(\xi)}{\overline{D}_{2m+2}(\xi)}\right\}$$
$$= \frac{p_{2m+1}}{q_{2m+1}}\prod_{j=1}^{m}\frac{p_{2j-1}}{q_{2j-1}}\frac{p_{2j}}{q_{2j}}\left(1 + \frac{p_{2m+2}}{q_{2m+2}}\right) \geq 0.$$

Similarly, using the condition on p_{2j-1} in Theorem 2.1(b), it can be shown that all the "coefficients" of the polynomials $S_j^2(x)$ are nonnegative. Thus we have from (A.3), for all $x \in [-1, 1]$,

$$0 \leq 1 - \frac{\underline{D}_{2j+1}(\xi)}{\overline{D}_{2j+1}(\xi)}\frac{\underline{D}_{2j+2}(\xi)}{\overline{D}^{-}_{2j+2}(\xi)}P^2_{j+1}(x), \qquad j = 0, \dots, k-1. \tag{A.5}$$

It follows that, for any $x \in [-1, 1]$, we have

$$P^2_{j+1}(x) \leq \frac{\overline{D}_{2j+1}(\xi)}{\underline{D}_{2j+1}(\xi)}\frac{\overline{D}^{-}_{2j+2}(\xi)}{\underline{D}_{2j+2}(\xi)} = P^2_{j+1}(1), \tag{A.6}$$

where the last equality is obtained from (A.3) for $x = 1$. The assertion (b) of the theorem now follows from the representation of the variance function in (A.0). To prove (c), let ξ^* denote the reflection of ξ at the origin so that $d_j(x, \xi^*) = d_j(-x, \xi)$, $j = 1, \dots, k$. Since the canonical moments of ξ^* and ξ are related by

$$p^*_{2j} = p_{2j} \quad \text{and} \quad p^*_{2j-1} = q_{2j-1}, \qquad j = 1, \dots, k$$

[Lau and Studden (1985)], it follows that ξ^* satisfies the assumptions of Theorem 2.1(b). Thus we obtain, for $j = 1, 2, \dots, k$,

$$\max_{x\in\Omega} d_j(x, \xi) = \max_{x\in\Omega} d_j(-x, \xi) = \max_{x\in\Omega} d_j(x, \xi^*) = d_j(1, \xi^*) = d_j(-1, \xi),$$

proving (c). Finally, part (a) follows from (b) or (c) since the symmetry of ξ implies all canonical moments of odd order are $\frac{1}{2}$[Lau (1983)]. □

Proof of Theorem 2.2. First, we prove part (b) of the theorem by induction. For $j = 1$, we calculate the first orthonormal polynomial with respect to the measure $d\xi(x)$ as in Lau (1983) to get $P_1(x) = (x + 1 - 2\zeta_1)/(4\zeta_1\zeta_2)^{1/2}$ and, consequently, the variance function is $d_1(x, \xi) = 1 + (x + 1 - 2p_1)^2/(4p_1q_1p_2)$. Because $p_1 \leq \frac{1}{2}$ by assumption, this function attains its maximum in the interval $[-1, 1]$ at the point 1, which proves the assertion for $j = 1$. For the step j to $j + 1 \leq k$, we assume that the variance functions $d_i(x,\xi)$, $i=1,2,\dots,j$, all attain their maximum in the interval $[-1,1]$ at the point 1. By Lemma A.3, we have

$$
\begin{aligned}
d_{j+1}(x,\xi) &= (1+a_1)b_{j+1} + (1-a_j b_{j+1})d_j(x,\xi)\\
&\quad -(1-x^2)b_{j+1}c_{j+1}Q_j^2(x)\\
&\quad +b_{j+1}\sum_{i=1}^{j-1}(a_{i+1}-a_i)d_i(x,\xi)\\
&\quad -(1-x)b_{j+1}\sum_{i=0}^{j} f_i S_i^2(x), \qquad j=1,\dots,k-1,
\end{aligned} \tag{A.7}
$$

where the quantities a_i, b_i, c_i and f_i are defined in (A.1). In terms of the canonical moments, the difference $a_{i+1}-a_i$, can be written as

$$
a_{i+1}-a_i = \prod_{j=1}^{i}\frac{p_{2j-1}}{q_{2j-1}}\prod_{j=1}^{i-1}\frac{q_{2j}}{p_{2j}}\left\{-1+\frac{q_{2i}}{p_{2i}}+\frac{p_{2i+1}}{q_{2i+1}}\frac{q_{2i}}{p_{2i}}\left(1-\frac{q_{2i+2}}{p_{2i+2}}\right)\right\},
$$

and it is straightforward to show that the nonnegativity of this term is equivalent to (2.3). Similarly, it follows that (2.5) is equivalent to the assertion that $f_i \ge 0$, for $i=0,\dots,k-1$. From (2.3), we obtain

$$
\frac{2p_{2j}-1}{1-p_{2j}}\,\frac{1-p_{2j+1}}{p_{2j+1}} \le p_{2j+2}, \qquad j=1,\dots,k-2,
$$

which is equivalent to the inequality $1-a_j b_{j+1}\ge 0$, $j=1,2,\dots,k-2$. In the remaining case, $j=k-1$, this inequality follows directly from assumption (2.4) and, consequently, the terms $1-a_j b_{j+1}$, $b_{j+1}c_{j+1}$, b_{j+1}, f_i and $a_{i+1}-a_i$, $i=1,\dots,j-1$, in (A.7) are all nonnegative. By the induction hypotheses and Lemma A.3, we have

$$
\begin{aligned}
d_{j+1}(x,\xi) &\le (1+a_1)b_{j+1} + (1-a_j b_{j+1})d_j(1,\xi)\\
&\quad + b_{j+1}\sum_{i=1}^{j-1}(a_{i+1}-a_i)d_i(1,\xi)\\
&= d_{j+1}(1,\xi) \quad \text{for all } x\in[-1,1].
\end{aligned}
$$

This is the assertion for $j+1\le k$ and hence proves part (b) of the theorem. Part (c) is obtained by the same "reflection" argument as in the proof of Theorem 2.1(c). Finally, part (a) follows from part (b) because the symmetry of the design yields $p_{2j-1}=\frac{1}{2}$ for all j, [Lau (1983)]. □

Proof of Theorem 2.3. From (A.3), we have

$$
P_j^2(1) = \frac{\overline{D}_{2j-1}(\xi)}{\underline{D}_{2j-1}(\xi)}\,\frac{\overline{D}_{2j}^{-}(\xi)}{\underline{D}_{2j}(\xi)}, \qquad j=1,2,\dots,k,
$$

where

$$\frac{\overline{D}_{2j-1}(\xi)}{\underline{D}_{2j-1}(\xi)} = \prod_{i=1}^{j} \frac{q_{2i-1}}{p_{2i-1}} \quad \text{(by Lemma A.1)}$$

and

$$\frac{\overline{D}_{2j}^{-}(\xi)}{\underline{D}_{2j}(\xi)} = \prod_{m=1}^{j-1} \frac{q_{2m}}{p_{2m}} \left(\frac{q_{2j}}{p_{2j}} + 1 \right) = \prod_{m=1}^{j-1} \frac{q_{2m}}{p_{2m}} \frac{1}{p_{2j}} \quad \text{[by (A.4) and Lemma A.1]}.$$

The assertion for $d_j(1, \xi)$ can now be obtained from (A.0). The representation for $d_j(-1, \xi)$ follows by similar arguments as in the second part of the proof of Theorem 2.1. Finally, if ξ is symmetric, then the monic orthogonal polynomials with respect to $d\xi(x)$ satisfy the recursive relation $P_0^*(x) = 1$, $P_{-1}^*(x) = 0$, and

$$P_{j+1}^*(x) = xP_j^*(x) - q_{2j-2}\, p_{2j} P_{j-1}^*(x) \quad \text{for } j \geq 0,$$

with L_2-norm given by $\delta_j^2 = \int_{-1}^{1} P_j^*(x)^2 \, d\xi(x) = \prod_{i=1}^{j} q_{2i-2}\, p_{2i}$ [Lau (1983)]. A straightforward calculation now yields for the orthonormal polynomials $P_j(x) = P_j^*(x)/\delta_j$, with respect to the measure $d\xi(x)$,

$$P_{2i}(0)^2 = \prod_{m=1}^{i} \frac{p_{4m-2}}{q_{4m-2}} \prod_{m=1}^{i-1} \frac{q_{4m}}{p_{4m}} \frac{1}{p_{4i}}$$

and $P_{2i-1}(0) = 0$. The assertion now follows from (A.0). □

Proof of Theorem 2.4. (a) From Studden (1980), the canonical moments of ξ_n are given by

$$\text{(A.8)} \qquad p_{2j} = \frac{n-j+1}{2n-2j+1} \quad \text{and} \quad p_{2j-1} = \frac{1}{2}, \qquad j = 1, 2, \dots, n.$$

It is easy to see that, for $j = k - 1$, (2.1) implies (2.2) and, consequently, the first part of Theorem 2.2 is applicable, where the special choice of canonical moments in (A.8) yields equality in (2.1), for all $j = 1, 2, \dots, k-1$. Together with Theorems 2.2 and 2.3, this yields

$$\begin{aligned} \max_{x \in \Omega} d_j(x, \xi_n) &= d_j(1, \xi_n) = 1 + \sum_{i=1}^{j} \frac{2(n-i)+1}{n} \\ &= n + 1 - \frac{(n-j)^2}{n}, \qquad j = 1, 2, \dots, n. \end{aligned}$$

This proves part (a) of the theorem by the definition of G-efficiency.

(b) This is proved similarly by an application of Theorem 2.1 and Lemma A.3 and by noting that the canonical moments of the D_1-optimal design for $f_n(x)$ are given by $p_j = \frac{1}{2}$, $j = 1, 2, \dots, 2n-1$, and $p_{2n} = 1$ [Studden (1982)]. □

PROOF OF THEOREM 2.5. We assume that $r \geq 1$ [the case $r = 0$ is treated in Theorem 2.4(a)]. By Theorem 3.1 of Studden (1980), $\xi_{n,D_{n-r}}$ has all odd canonical moments equal to $\frac{1}{2}$ and even canonical moments given by

$$p_{2j} = \begin{cases} \dfrac{1}{2}, & \text{if } 1 \leq j \leq r, \\ \dfrac{n-j+1}{2n-2j+1}, & \text{if } r+1 \leq j \leq n. \end{cases} \tag{A.9}$$

For $j = 1, 2, \dots, r$, the assertion follows directly from Theorem 2.1. For $j \geq r+1$, we apply Lemma A.3 and obtain the recursive relation, for $m = r, r+1, \dots, n-1$,

$$\begin{aligned} d_{m+1}(x, \xi_{n,D_{n-r}}) = b_{m+1} &+ (1 - a_m b_{m+1}) d_m(x, \xi_{n,D_{n-r}}) \\ &- (1 - x^2) b_{m+1} c_{m+1} Q_m^2(x) \\ &+ b_{m+1} \sum_{j=r}^{m-1} (a_{j+1} - a_j) d_j(x, \xi_{n,D_{n-r}}). \end{aligned} \tag{A.10}$$

Note that $a_1 = a_2 = \cdots = a_r = 0$ because $p_{2m} = \frac{1}{2}$, $m = 1, 2, \dots, r$. By the definition of a_m in (A.1) and the representation (A.9) of the canonical moments of the D_{n-r}-optimal designs, we obtain $a_{r+1} = a_{r+2} = \cdots = a_n = 1/(n-r)$ and, consequently, (A.10) simplifies to

$$\begin{aligned} d_{m+1}(x, \xi_{n,D_{n-r}}) = b_{m+1} &+ (1 - a_m b_{m+1}) d_m(x, \xi_{n,D_{n-r}}) \\ &+ \frac{b_{m+1}}{n-r} d_r(x, \xi_{n,D_{n-r}}) \\ &- (1 - x^2) b_{m+1} c_{m+1} Q_m^2(x), \\ & \qquad m = r, r+1, \dots, n-1. \end{aligned} \tag{A.11}$$

It is easy to see that, for the canonical moments in (A.9), $1 - a_m b_{m+1} \geq 0$ and the assertion for $j = r, \dots, n$ follows from (A.11) by a similar induction argument as in the proof of Theorem 2.2. □

PROOF OF THEOREM 2.6. From Studden (1982), we have, for the canonical moments of $\xi_n^{(\alpha,\beta)}$,

$$p_{2i+1} = \frac{\alpha + n + 1 - i}{\alpha + \beta + 2n + 2 - 2i}, \qquad i = 0, 1, \dots, n,$$

and

$$p_{2i} = \frac{n + 1 - i}{\alpha + \beta + 3 + 2n - 2i}, \qquad i = 1, 2, \dots, n+1.$$

If $\alpha + \beta \geq -1$, the assertion follows from Theorems 2.1 and 2.3 and if $-2 < \alpha + \beta < -1$, the result follows from Theorems 2.2 and 2.3. □

Acknowledgments. The research on this topic was initiated while H. Dette was visiting UCLA in the summer of 1993. This author would like to thank the Department of Biostatistics for its hospitality and the Deutsche

Forschungsgemeinschaft for the financial support that made the visit possible. The authors would also like to thank the referees and an Associate Editor for their valuable comments.

REFERENCES

ATKINSON, A. C. and FEDOROV, V. V. (1975a). The design of experiments for discriminating between two rival models. *Biometrika* **62** 57–70.

ATKINSON, A. C. and FEDOROV, V. V. (1975b). Optimal design: experiments for discriminating between several models. *Biometrika* **62** 289–303.

ATWOOD, C. L. (1969). Optimal and efficient designs of experiments. *Ann. Math. Statist.* **40** 1570–1602.

DETTE, H. (1993). New identities for orthogonal polynomials on a compact interval. *J. Math. Anal. Appl.* **179** 547–573.

FEDOROV, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York. (Translated and edited by W. J. Studden and E. M. Klimko.)

FREUD, G. (1972). On Hermite–Fejér interpolation processes. *Studia Sci. Math. Hungar.* **7** 337–342.

KARLIN, S and SHAPELY, L. S. (1953). Geometry of moment spaces. *Mem. Amer. Math. Soc.* **12**.

KARLIN, S. and STUDDEN, W. J. (1966). Tchebycheff systems: with applications in analysis and statistics. Interscience, New York.

KENDALL, M. G. and STUART, A. (1968). *The Advanced Theory of Statistics* **3**. Hafner, New York.

KIEFER, J. (1985). *Jack Carl Kiefer Collected Papers III: Design of Experiments*. Springer, New York.

KIEFER, J. and STUDDEN, W. J. (1976). Optimal designs for large degree polynomial regression. *Ann. Statist.* **4** 1113–1123.

KIEFER, J. and WOLFOWITZ, J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* **12** 363–366.

KUSSMAUL, K. (1969). Protection against assuming the wrong degree in polynomial regression. *Technometrics* **11** 677–682.

LAU, T. S. (1983). Theory of canonical moments and its applications, I and II. Technical Reports 83-23, 83-24. Dept. Statistics, Purdue Univ.

LAU, T. S. and STUDDEN, W. J. (1985). Optimal designs for trigonometric and polynomial regression using canonical moments. *Ann. Statist.* **13** 383–394.

NEVAI, P. (1986). Géza Freud, orthogonal polynomials and Christoffel functions. A case study. *J. Approx. Theory* **48** 3–167.

SKIBINSKY, M. (1986). Principal representations and canonical moment sequences on an interval. *J. Math. Anal. Appl.* **120** 95–120.

STUDDEN, W. J. (1980). D_s-optimal designs for polynomial regression using continued fractions. *Ann. Statist.* **8** 1132–1141.

STUDDEN, W. J. (1982). Optimal designs for weighted polynomial regression using canonical moments. In *Statistical Decision Theory and Related Topics III* (S. S. Gupta and J. O. Berger, eds.) **2** 335–350. Academic Press, New York.

STUDDEN, W. J. (1989). Note on some ϕ_p-optimal designs for polynomial regression. *Ann. Statist.* **17** 618–623.

THIBODEAU, L. A. (1977). Robust design for regression problems. Ph.D. dissertation, Univ. Minnesota.

WONG, W. K. (1994). Comparing robust properties of A, D, E and G-optimal designs. *Comput. Statist. Data Anal.* **18** 441–448.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
TECHNISCHE UNIVERSITÄT DRESDEN
01062 DRESDEN
GERMANY

DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024-1772