Optimal Designs for Identifying the Degree of a Polynomial Regression

by

Holger Dette
Technische Universität Dresden
and Purdue University

Technical Report #93-33

Department of Statistics
Purdue University

July 1993

△

OPTIMAL DESIGNS FOR IDENTIFYING THE DEGREE OF A POLYNOMIAL REGRESSION

by

Holger Dette*
Institut für Mathematische Stochastik
Abteilung Mathematik
Technische Universität Dresden
Mommsenstr. 13
01062 Dresden
GERMANY

ABSTRACT

If an experimenter wants to determine the degree of a polynomial regression on the basis of a sample of observations, Anderson (1962) showed that the following method is optimal. Starting with the highest (specified) degree the procedure is to test in sequence whether the coefficients are 0. In this paper optimal designs for Anderson's procedure are determined explicitly. The optimal design maximizes the minimum power of a given set of alternatives.

1. Introduction. A frequent problem in regression analysis is to determine how many independent variables have to be included in the fitted regression function. In many cases (e.g. polynomial regression) the underlying models are nested

$$h_\ell(x) = \sum_{i=0}^{\ell} \vartheta_i f_i(x) \qquad \ell = 0,\dots,n \tag{1.1}$$

and based on a sample of observations the experimenter has to identify the appropriate model $h_\ell(x)$. Anderson (1962) studied the following decision rule. For a given set of levels $(\alpha_1,\dots,\alpha_n)$ the procedure chooses the largest integer in $\{1,\dots,n\}$, for which the F-test in the model $h_j(x)$ rejects the hypotheses $H_0\colon \vartheta_j = 0$ $(j = 1,\dots,n)$. It is well known that Anderson's method has several optimality properties (see Anderson (1962) or Spruill (1990)). In the following $Y_1,\dots,Y_m$ denote m independent normally distributed observations with common variance $\sigma^2 > 0$ and mean given by one of the models in (1.1), that is

$$Y = Z_\ell \theta_\ell + \varepsilon \qquad \text{for some} \quad \ell = 0,\dots,n$$

AMS Subject Classification: 62K05
Keywords and Phrases: testing the degree of a polynomial regression, minimax designs, Chebyshev polynomials, canonical moments.
*Research supported in part by the Deutsche Forschungsgemeinschaft

where $Y = (Y_1, \ldots, Y_m)', \varepsilon \sim \mathcal{N}(0, \sigma^2 I_m), \theta_\ell = (\vartheta_0, \ldots, \vartheta_\ell)'$ and $Z_\ell = (f_j(x_i))_{i=0,\ldots,m}^{j=0,\ldots,\ell}$ $(l = 0, \ldots, n)$. For Anderson's procedure the probability of the error of choosing too many functions is independent of the matrix Z provided that Z has full rank (see Anderson (1962)). If the model is $h_\ell(x) = h_{\ell-1}(x) + \vartheta_\ell f_\ell(x)$ the distribution of the test statistic of the F-test for the hypothesis $H_0: \vartheta_\ell = 0$ has the noncentrality parameter

$$\delta_\ell^2 = \frac{\vartheta_\ell^2}{\sigma^2}(e_\ell'(Z_\ell' Z_\ell)^{-1} e_\ell)^{-1} \tag{1.2}$$

where $e_\ell = (0, \ldots, 0, 1)' \in \mathbb{R}^{\ell+1}$. Consequently, the probability of deciding in favor of $h_{\ell-1}(x)$ (when the model is in fact $h_\ell(x)$) is a decreasing function of δ_ℓ^2 and a good choice of a design $x_0, \ldots, x_m$ will make the quantities in (1.2) as large as possible.

In a recent paper Spruill (1990) considered the case of polynomial regression on the interval $[-1, 1]$ (which is the region where observations can be taken), the alternatives $\vartheta_1 = \ldots = \vartheta_n = 1$ and determined the optimal approximate design with respect to a maximin criterion. It is the aim of this paper to extend these results to arbitrary alternatives ϑ_l and intervals $[a, b]$. We obtain the surprising fact that the structure of the optimal design changes completely when the observations can be taken in a different design space or different alternatives are assumed. In Section 2 we describe some general aspects of (approximate) design theory and introduce the optimality criterion. Section 3 gives a short review of the theory of canonical moments which were introduced by Studden (1980, 1982a, b) in the context of design theory. These results are applied in Section 3 and 4 to obtain a complete solution of the optimal design problem which contains the result of Spruill (1990) as a special case (i.e. $[a, b] = [-1, 1]$, $\vartheta_\ell = 1$, $\ell = 1, \ldots, n$). Finally some asymptotic considerations and examples are presented in Section 5.

2. The optimality criterion.

Let $\mathcal{X}$ denote a compact space with a sigma field containing all one point sets and at least $n+1$ points. In the following we consider $n+1$ linearly independent, real valued and continuous regression functions $f_0(x), \ldots, f_n(x)$ (defined on $\mathcal{X}$) and collect the first $\ell+1$ functions in a vector $g_\ell(x) = (f_0(x), \ldots, f_\ell(x))'(\ell = 0, \ldots, n)$. The model in (1.1) can now be written as $h_\ell(x) = \theta_\ell' g_\ell(x)$. A (approximate) design ξ is a probability measure on $\mathcal{X}$ and the matrix

$$M_\ell(\xi) = \int_{\mathcal{X}} g_\ell(x) g_\ell'(x) d\xi(x)$$

is called the information matrix for the model $h_\ell(x)$ ($\ell = 0,\ldots,n$). If ξ puts masses $\frac{m_i}{m}$ ($i = 1,\ldots,s$) at the points $x_1,\ldots,x_s$ the experimenter takes m uncorrelated observations, m_i at each x_i ($i = 1,\ldots,s$), and the inverse of the information matrix $M_\ell(\xi)$ (in the model h_ℓ) is proportional to the covariance matrix of the least squares estimator for θ_ℓ, that is $\sigma^2(Z_\ell Z_\ell')^{-1}$. According to the discussion following (1.2) we call an approximate design ξ (maximin) optimal discriminating design if ξ maximizes the function

$$\Phi(\xi) = \min\{\vartheta_\ell^2(e_\ell' M_\ell^{-1}(\xi)e_\ell)^{-1} \mid \ell = 1,\ldots,n\} \tag{2.1}$$

(see also Spruill (1990)). Note that this optimality criterion is a local criterion in the sense that it depends on the alternatives ϑ_ℓ of the corresponding hypotheses $H_0\colon \vartheta_\ell = 0$ ($\ell = 1,\ldots,n$). Spruill (1990) used $\vartheta_1 = \ldots = \vartheta_n = 1$ and we will discuss a generalization of this special choice in the following sections. An extremely useful tool for determining optimal (approximate) designs are equivalence theorems which provide necessary and sufficient conditions for a design to be optimal (see e.g. Pukelsheim (1993)). In order to derive such a result for the optimality criterion (2.1) we define for a design ξ with nonsingular information matrix $M_n(\xi)$

$$\widetilde{M}(\xi) = \begin{pmatrix} M_1(\xi) & & \\ & \ddots & \\ & & M_n(\xi) \end{pmatrix} \in \mathbb{R}^{n(n+3)/2 \times n(n+3)/2},$$

$$K = \begin{pmatrix} \vartheta_1^{-1} e_1 & & \\ & \ddots & \\ & & \vartheta_n^{-1} e_n \end{pmatrix} \in \mathbb{R}^{n(n+3)/2 \times n}$$

and

$$C_K(\widetilde{M}) = (K'\widetilde{M}(\xi)^{-1}K)^{-1} = \begin{pmatrix} \vartheta_1^2(e_1'M_1^{-1}(\xi)e_1)^{-1} & & \\ & \ddots & \\ & & \vartheta_n^2(e_n'M_n^{-1}(\xi)e_n)^{-1} \end{pmatrix} \in \mathbb{R}^{n\times n}$$

(note that all other entries in these matrices are zero and that an optimal discriminating design must have nonsingular matrices $M_1(\xi),\ldots,M_n(\xi)$). If an information function j is defined on the nonnegative definite $n \times n$ matrices by $j(A) = \lambda_{\min}(A)$ (here $\lambda_{\min}(A)$ denotes the minimum eigenvalue of A) then it follows for the criterion (2.1) that

$$\Phi(\xi) = j(C_K(\widetilde{M}(\xi)) \qquad (\det M_n(\xi) > 0).$$

By an application of the general equivalence theorem in Pukelsheim (1993, p. 175) it is now straightforward to show the following equivalence theorem for the optimality criterion defined in (2.1).

Theorem 2.1. Let ξ denote a design such that $M_n(\xi)$ is nonsingular and let $\mathcal{N}(\xi):= \{j|\vartheta_j^2 e_j' M_j^{-1}(\xi)e_j = \max_{\ell=1}^m \vartheta_\ell^2 e_\ell' M_\ell^{-1}(\xi)e_\ell\}$. The design ξ is an optimal discriminating design if and only if there exist nonnegative numbers α_ℓ ($\ell \in \mathcal{N}(\xi)$) with $\sum_{\ell\in\mathcal{N}(\xi)} \alpha_\ell = 1$ such that

$$\sum_{\ell\in\mathcal{N}(\xi)} \alpha_\ell \frac{(e_\ell' M_\ell^{-1}(\xi) g_\ell(x))^2}{e_\ell' M_\ell^{-1}(\xi)e_\ell} \leq 1 \tag{2.2}$$

for all $x \in \mathcal{X}$. Moreover, in (2.2) equality holds for all support points of every optimal discriminating design.

Remark 2.2. Theorem 2.1 can easily be generalized by changing the e_l to arbitrary vectors $c_\ell \in \mathrm{R}^{\ell+1}$ and not necessarily nested regression functions. In this cases the design ξ has to satisfy $c_\ell \in$ range $(M_\ell(\xi))$ ($\ell = 1, \ldots, n$) and the inverses in Theorem 2.1 have to be replaced by general inverses (see Pukelsheim (1993, p. 283)).

3. Polynomial regression models. Let $g_\ell(x) = (1, x, \ldots, x^\ell)'$ denote the vector of monomials up to the order ℓ and $\mathcal{X} = [a, b]$. Thus the problem of Section 1 and 2 is to determine the optimal design when Anderson's procedure is applied for testing the degree of a polynomial regression. For this purpose we need some basic facts about canonical moments which were introduced in the context of design theory by Studden (1980, 1982a, b) (see also Lau (1983) for more details). Let ξ denote a probability measure on $[a, b]$ with moments $c_i = \int_a^b x^i d\xi(x)$. For a given set of moments $c_0, \ldots, c_{i-1}$ let c_i^+ denote the maximum of the ith moment $\int_a^b x^i d\xi(x)$ over the set of all probability measures μ on $[a, b]$ having the given moments $c_0, c_1, \ldots, c_{i-1}$. Similarly let c_i^- denote the corresponding minimum. The ith canonical moment is defined by

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

Note that $0 \leq p_i \leq 1$ and that the canonical moments are left undefined whenever $c_i^+ = c_i^-$. If i is the first index for which this equality holds, then $0 < p_k < 1, k = 1, \ldots, i-2, p_{i-1}$

must have the value 0 or 1 and the design ξ is supported at a finite number of points. In this case ξ is the "lower" or "upper principal representation" of its corresponding moment point $(c_0, \ldots, c_{i-1})$ (see Skibinsky (1986), Section 1). The optimality criterion (2.1) can easily be expressed in terms of canonical moments (see e.g. Lau and Studden (1985)).

$$
\begin{aligned}
\Phi(\xi) &= \min\left\{\vartheta_\ell^2 \frac{\det M_\ell(\xi)}{\det M_{\ell-1}(\xi)} | \ell = 1, \ldots, n\right\} \\
&= \min\left\{\vartheta_\ell^2 (b-a)^{2\ell} \prod_{j=1}^{\ell} q_{2j-2} p_{2j-1} q_{2j-1} p_{2j} \mid \ell = 1, \ldots, n\right\}
\end{aligned}
\tag{3.1}
$$

(here $q_j = 1 - p_j, j \geq 1, q_0 = 1$). A probability measure ξ is symmetric about the point $(a+b)/2$ if and only if all canonical moments of odd order satisfy $p_{2i-1} = 1/2$.

The following result shows that every symmetric design maximizes a (weighted) geometric mean of the ratios $\det M_\ell(\xi)/\det M_{\ell-1}(\xi)$ and can be proved by similar arguments as in Dette (1991).

Theorem 3.1. Let ξ^* denote a design with canonical moments $(\frac{1}{2}, p_2^*, \frac{1}{2}, p_4^*, \ldots, \frac{1}{2}, p_{2n-2}^*, \frac{1}{2}, 1)$ and define

$$
\beta_\ell^* = \left(1 - \frac{q_{2\ell}^*}{p_{2\ell}^*}\right) \prod_{j=1}^{\ell-1} \frac{q_{2j}^*}{p_{2j}^*} \quad \ell = 1, \ldots, n.
\tag{3.2}
$$

Then the design ξ maximizes the weighted geometric mean

$$
\prod_{j=1}^{n} \left(\frac{\det M_j(\xi)}{\det M_{j-1}(\xi)}\right)^{\beta_j^*}
$$

among all design on the interval $[a, b]$.

If ξ has canonical moments $(p_1, p_2, p_3, p_4, p_5, \ldots)$ then its reflection ξ' about the point $(a+b)/2$ has canonical moments $(q_1, p_2, q_3, p_4, q_5, \ldots)$ (see e.g. Lau and Studden (1985), p. 387). Consequently we obtain from (3.1) $\Phi(\xi) = \Phi(\xi')$ and it follows by standard arguments of optimal design theory that the optimal discriminating design is symmetric about the point $(a+b)/2$.

Theorem 3.2. The optimal discriminating design ξ^* has canonical moments $\frac{1}{2}, p_2^*, \frac{1}{2}, p_4^*, \frac{1}{2}, \ldots, p_{2n-2}^*, \frac{1}{2}, 1$ where the canonical moments of even order are defined for $j = 1, \ldots, n-1$

recursively by ($p^*_{2n}=1$)

$$(3.3)\qquad p^*_{2(n-j)}=\max\left\{1-\frac{\vartheta^2_{n-j}}{\vartheta^2_n}\left(\frac{b-a}{2}\right)^{-2j}\prod_{i=n-j+1}^{n-1}(q^*_{2i}p^*_{2i})^{-1},\frac{1}{2}\right\}.$$

Proof: Consider the design ξ^* with canonical moments $(\frac{1}{2},p^*_2,\frac{1}{2},\ldots,\frac{1}{2},p^*_{2n-2},\frac{1}{2},1)$ and let

$$\gamma_j:=\ 1-\frac{\vartheta^2_{n-j}}{\vartheta^2_n}\left(\frac{b-a}{2}\right)^{-2j}\left(\prod_{i=n-j+1}^{n-1}q^*_{2i}p^*_{2i}\right)^{-1}\qquad(j=1,\ldots,n-1).$$

If $\gamma_j\geq\frac{1}{2}$, then it is easy to see (observing (3.1)) that

$$\vartheta^2_{n-j}\frac{\det M_{n-j}(\xi^*)}{\det M_{n-j-1}(\xi^*)}=\vartheta^2_n\frac{\det M_n(\xi^*)}{\det M_{n-1}(\xi^*)}.$$

On the other hand, if $\gamma_j<\frac{1}{2}$, then $p^*_{2(n-j)}=q^*_{2(n-j)}=\frac{1}{2}$ and it follows that $(1-\gamma_j)^{-1}<2$ which implies

$$\frac{\vartheta^2_n}{\vartheta^2_{n-j}}\frac{\det M_n(\xi^*)}{\det M_{n-1}(\xi^*)}\frac{\det M_{n-j-1}(\xi^*)}{\det M_{n-j}(\xi^*)}=\frac{q^*_{2(n-j)}}{1-\gamma_j}<1.$$

Consequently the set $\mathcal{N}(\xi^*)$ defined in Theorem 2.1 is given by ($\gamma_0:=1$)

$$(3.4)\qquad \mathcal{N}(\xi^*)=\left\{j\in\{1,\ldots,n\}\mid\gamma_{n-j}\geq\frac{1}{2}\right\}.$$

By Theorem 3.1 the design ξ^* maximizes the weighted geometric mean

$$\prod_{\ell=1}^{n}\left(\frac{\det M_\ell(\xi)}{\det M_{\ell-1}(\xi)}\right)^{\beta^*_\ell}$$

where the weights β^*_ℓ are defined in (3.2). If all weights satisfy $\beta^*_l\geq 0$ $(l=1,\ldots,n)$ then Theorem 2.1 in Dette (1993) shows that this property is equivalent to the condition

$$(3.5)\qquad \sum_{\ell=1}^{n}\beta^*_\ell\frac{(e'_\ell M_\ell^{-1}(\xi^*)g_\ell(x))^2}{e'_\ell M_\ell^{-1}(\xi^*)e_\ell}\leq 1$$

for all $x\in[a,b]$. If $\ell\notin\mathcal{N}(\xi^*)$, then we have $\gamma_{n-\ell}<\frac{1}{2}$, by (3.3) $p^*_{2\ell}=\frac{1}{2}$ and by (3.2) this implies $\beta^*_\ell=0$. On the other hand, if $\gamma_{n-l}\geq\frac{1}{2}$, it follows from $p^*_{2l}\geq\frac{1}{2}$ that $\beta^*_l\geq 0$. Therefore Theorem 2.1 in Dette (1993) can be applied and (3.5) may be rewritten as

$$\sum_{\ell\in\mathcal{N}(\xi^*)}\beta^*_\ell\frac{(e'_\ell M_\ell^{-1}(\xi^*)g_\ell(x))^2}{e'_\ell M_\ell^{-1}(\xi^*)e_\ell}\leq 1$$

for all $x \in [a,b]$. The assertion follows now from Theorem 2.1. ∎

The weights and the support points of the optimal discriminating design ξ^* corresponding to the terminating sequence of canonical moments can be calculated by standard techniques (see Lau (1988)). Define $\sigma = (b-a)/2, \tau = (a+b)/2$,

$$K\begin{pmatrix} & a_1 & & \dots & a_n \\ b_0 & & b_1 & \dots & & b_n \end{pmatrix} = \det \begin{pmatrix} b_0 & -1 & & & \\ a_1 & b_1 & -1 & & \\ & & \ddots & \ddots & \\ & & \ddots & b_{n-1} & -1 \\ & & & a_n & b_n \end{pmatrix}$$

(all other entries in the matrix are 0) and

$$P_n(x) = K\begin{pmatrix} & -\sigma^2 q^*_{2n-2}p^*_{2n} & & \dots & & -\sigma^2 q^*_2 p^*_4 & \\ x-\tau & & x-\tau & \dots & x-\tau & & x-\tau \end{pmatrix}, \tag{3.6}$$

$$Q_{n-1}(x) = K\begin{pmatrix} & -\sigma^2 p^*_{2n-4}q^*_{2n-2} & & & & -\sigma^2 p^*_2 q^*_4 & \\ x-\tau & & x-\tau & \dots & x-\tau & & x-\tau \end{pmatrix} \tag{3.7}$$

$(P_0(x) = Q_0(x) = 1, P_1(x) = Q_1(x) = x - \tau)$, then the following result holds (see e.g. Lau (1988)).

Proposition 3.3. The optimal discrimination design ξ^* is supported at the zeros $x_0, \dots, x_n$ of the polynomial $(x-a)(x-b)Q_{n-1}(x)$ with masses

$$\xi(\{x_j\}) = \frac{P_n(x_j)}{\frac{d}{dx}(x-a)(x-b)Q_{n-1}(x)|_{x=x_j}} \qquad j = 0, \dots, n.$$

Example 3.4. Let $a = -b = -1$, and $n = 2$ (quadratic regression), by Theorem 3.2 the optimal discriminating design has canonical moments (of even order) $p_4 = 1$ and $p_2 = \max\{1 - \frac{\vartheta_1^2}{\vartheta_2^2}, \frac{1}{2}\}$. Therefore, if $|\vartheta_1|$ is small in proportion to $|\vartheta_2|$ (i.e. $|\vartheta_1| \le |\vartheta_2|/\sqrt{2}$), the optimal design ξ_1^* puts masses $1 - \frac{\vartheta_1^2}{2\vartheta_2^2}, \frac{\vartheta_1^2}{\vartheta_2^2}, 1 - \frac{\vartheta_1^2}{2\vartheta_2^2}$ at the points $-1, 0, 1$. This corresponds to the intuitive fact, that one only needs a few observations at zero in order to distinguish between a linear or quadratic regression on $[-1,1]$ when it is known that the extremum of the quadratic function is attained in a neighbourhood of 0. On the other hand, if $|\vartheta_2|$ is small compared to $|\vartheta_1|$ (i.e. $|\vartheta_1| > |\vartheta_2|/\sqrt{2}$), the optimal discriminating design ξ_2^* has $p_2^* = \frac{1}{2}$ and masses $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ at $-1, 0, 1$. In this case the minimum of the quadratic polynomial is attained outside of the interval $[-1,1]$ and the linear and quadratic function have a similar form inside of $[-1,1]$. In order to distinguish between these models one has to take the best design for testing the highest coefficient of the quadratic regression which is given by ξ_2^* (see also Kiefer and Wolfowitz (1959)).

4. Explicit solutions. Throughout this and the following section we will assume that the unknown parameters in the criterion (2.1) satisfy $\vartheta_\ell = \vartheta^\ell$ $(\ell = 1, \ldots, n)$ for some $\vartheta > 0$. A similar assumption was made by Spruill (1990) who considered the problem on the interval $[-1,1]$ and the case $\vartheta = 1$. For this case the criterion (2.1) reduces to (observing (3.1) and the fact that the optimal discriminating design is symmetric, which means $p_{2j-1} = \frac{1}{2}$, $j = 1, \ldots, n$)

$$\Phi(\xi) = \min\left\{ (\vartheta(b-a))^{2\ell} \frac{\det M_l(\xi)}{\det M_{l-1}(\xi)} \mid \ell = 1, \ldots, n \right\}$$

(4.1)

$$= \min\left\{ \rho^{2\ell} \prod_{j=1}^{\ell} q_{2j-2} p_{2j} \mid \ell = 1, \ldots, n \right\}$$

where $\rho = \vartheta(b-a)/2$. An optimal design with respect to the criterion (4.1) maximizes the noncentrality parameters of the corresponding F-distributions assuming that the alternatives ϑ_ℓ of the hypotheses $H_\ell \colon \vartheta_\ell = 0$ are of the form $\vartheta_\ell = \vartheta^\ell$ $(\ell = 1, \ldots, n)$ for some $\vartheta > 0$. The following result gives the canonical moments of the optimal design with respect to this criterion. Here and throughout this paper $U_n(x)$ will denote the Chebyshev polynomial of the second kind (see e.g. Rivlin (1990), p. 7).

Theorem 4.1. The optimal discriminating design (with respect to the criterion (4.1)) has canonical moments $(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2}, p_{2k}^*, \frac{1}{2}, \ldots, \frac{1}{2}, p_{2n-2}^*, \frac{1}{2}, 1)$ where

(4.2)
$$p_{2j}^* = \frac{U_{n-j+1}(\frac{\rho}{2})}{\rho U_{n-j}(\frac{\rho}{2})} \qquad j = n, n-1, \ldots, k$$

and

(4.3)
$$k = \min\left\{ j \leq n \mid \frac{U_{n-i+1}(\frac{\rho}{2})}{\rho U_{n-i}(\frac{\rho}{2})} > \frac{1}{2} \quad \text{for } i = j, \ldots, n \right\}$$

Proof. Let $\varepsilon_n = 1$ and consider the sequence

$$\varepsilon_j = 1 - \frac{\rho^{-2}}{\varepsilon_{j+1}} \quad (j = n-1, \dots, 1). \tag{4.4}$$

It is straightforward to show that $\varepsilon_j \leq \varepsilon_{j+1}$ whenever $\varepsilon_n, \dots, \varepsilon_{j+1} > 0$ and obviously $\varepsilon_n = 1 = U_1(\frac{\rho}{2})/\rho U_0(\frac{\rho}{2})$ (see Rivlin (1990), p. 39). In the following we show that ε_j is given by (4.2) for all $j = 1, \dots, n$. To this end we use the recursive relation for the Chebyshev polynomials of the second kind

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x) \qquad (U_0(x) = 1, U_{-1}(x) = 0)$$

and obtain by induction $(j+1 \to j)$

$$\varepsilon_j = 1 - \frac{\rho^{-2}\rho U_{n-j-1}(\frac{\rho}{2})}{U_{n-j}(\frac{\rho}{2})} = \frac{U_{n-j+1}(\frac{\rho}{2})}{\rho U_{n-j}(\frac{\rho}{2})} \qquad j = 1, \dots, n. \tag{4.5}$$

By Theorem 3.2 the canonical moments of even order of the optimal discriminating design (with respect to the criterion (4.1)) are given by

$$p_{2j}^* = \max\left\{1 - \rho^{-2(n-j)} \prod_{i=j+1}^{n-1} (q_{2i}^* p_{2i}^*)^{-1}, \frac{1}{2}\right\}. \tag{4.6}$$

If $\rho \leq 1$, then we have $p_{2j}^* = \frac{1}{2}$ $(j = 1, \dots, n-1)$ and $k = n$ because it follows from (4.3), (4.4) and (4.5) that $\varepsilon_{n-1} = U_2(\frac{\rho}{2})/\rho U_1(\frac{\rho}{2}) = 1 - \rho^{-2} \leq 0$. In this case the assertion of Theorem 4.1 is obviously correct. On the other hand, if $\rho \geq 2$, then we obtain from (4.6) that $p_{2n-2}^* = 1 - \rho^{-2} \geq \frac{3}{4}$ and by an induction argument $p_{2j}^* > \frac{1}{2}$ $(j = 1, \dots, n)$. In this case we have from (4.6)

$$p_{2j}^* = 1 - \rho^{-2(n-j)} \left(\prod_{i=j+1}^{n-1} q_{2i}^* p_{2i}^*\right)^{-1} \qquad j = n-1, \dots, 1 \tag{4.7}$$

and it can easily be seen that (4.7) and (4.4) define the same sequences. The assertion now follows from the representation (4.5) (note that we obtain $k = 1$ in (4.3)). Finally, if $1 < \rho < 2$ and there exists an index k such that

$$1 - \rho^{-2(n-k)} \left(\prod_{i=k+1}^{n-1} q_{2i}^* p_{2i}^*\right)^{-1} < \frac{1}{2}$$

then $p_{2k}^* = \frac{1}{2}$ and we obtain

$$1-\rho^{-2(n-k+1)}\left(\prod_{i=k}^{n-1} q_{2i}^* p_{2i}^*\right)^{-1} = 1-\rho^{-2(n-k)}\left(\prod_{i=k+1}^{n-1} q_{2i}^* p_{2i}^*\right)^{-1} 4\rho^{-2} < \frac{1}{2}$$

which also implies $p_{2k-2}^* = \frac{1}{2}$. Consequently the sequence of canonical moments of the optimal discriminating design is of the form $(\frac{1}{2}, \dots, \frac{1}{2}, p_{2k}^*, \frac{1}{2}, \dots, \frac{1}{2}, p_{2n-2}^*, \frac{1}{2}, 1)$ where $p_{2k}^*, \dots, p_{2n-2}^* > \frac{1}{2}$ can be calculated recursively by (4.4) (or equivalently by (4.7)). The assertion of the theorem now follows from (4.5) which shows that k is defined by (4.3). ■

Remark 4.2. The proof of Theorem 4.1 shows that there are three different cases for the sequence of the optimal discriminating design ξ^*.

a) $\rho \le 1$: ξ^* has canonical moments $(\frac{1}{2}, \dots, \frac{1}{2}, 1)$. This case was originally considered by Spruill (1990) ($b = -a = 1, \vartheta = 1$) and the set $\mathcal{N}(\xi^*)$ of Theorem 3.1 is given by $\mathcal{N}(\xi^*) = \{n\}$.

b) $1 < \rho < 2$: In this case there exists an index $1 \le k \le n$ (depending on n and ρ) such that ξ^* has canonical moments $(\frac{1}{2}, \dots, \frac{1}{2}, p_{2k}^*, \frac{1}{2}, \dots, \frac{1}{2}, p_{2n-2}^*, \frac{1}{2}, 1)$ where $p_{2j}^* > \frac{1}{2}$ $(j = k, \dots, n-1)$, is defined recursively by $(p_{2n}^* = 1)$

$$p_{2j}^* = 1 - \frac{\rho^{-2}}{p_{2j+2}^*} \qquad j = n-1, \dots, k \tag{4.8}$$

and explicitly given by (4.2). It is straightforward to show that for sufficiently large n this index always satisfies $k > 1$ (this is a consequence of the fact that for $\rho < 2$ the equation $1-\rho^{-2}/z = z$ has no fixpoint and that the case $k = 1$ implies $\frac{1}{2} \le p_{2j}^* < p_{2j+2}^*$ for all $j = 1, \dots, n-1$). For this choice we have

$$\mathcal{N}(\xi^*) = \begin{cases} \{k, \dots, n\} & \text{if } 2U_{n-k+2}(\frac{\rho}{2}) < \rho U_{n-k+1}(\frac{\rho}{2}) \\ \{k-1, k, \dots, n\} & \text{if } 2U_{n-k+2}(\frac{\rho}{2}) = \rho U_{n-k+1}(\frac{\rho}{2}) \end{cases}$$

c) $\rho \ge 2$: ξ^* has canonical moments $(\frac{1}{2}, p_2^*, \frac{1}{2}, \dots, \frac{1}{2}, p_{2n-2}^*, \frac{1}{2}, 1)$ where the canonical moments of even order $p_{2j}^* > \frac{1}{2}$ are defined recursively by (4.8) for all $j = 1, \dots, n-1$ (or equivalently by (4.2)). In this case we have $\mathcal{N}(\xi^*) = \{1, \dots, n\}$.

In the following we will identify the support points and the weights of the optimal discriminating design ξ^* in Theorem 4.1. At first we will consider the case $\rho \ge 2$ which turn out to be essential for the general case.

Theorem 4.3. Let $\rho \geq 2$, then the support points $x_0, \dots, x_n$ of the optimal discriminating design ξ^* are given by $x_j = z_j + \frac{a+b}{2}$ $(j = 0, \dots, n)$ where $z_0, \dots, z_n$ are the zeros of the polynomial

$$U_{n+1}\left(\frac{x\vartheta}{2}\right) U_{n-1}\left(\frac{\rho}{2}\right) - U_{n-1}\left(\frac{x\vartheta}{2}\right) U_{n+1}\left(\frac{\rho}{2}\right)$$

and the weights at these points are

$$\xi(\{x_j\}) = \frac{2U_n(\frac{z_j\vartheta}{2})U_{n-1}(\frac{\rho}{2})}{U'_{n+1}(\frac{z_j\vartheta}{2})U_{n-1}(\frac{\rho}{2}) - U'_{n-1}(\frac{z_j\vartheta}{2})U_{n+1}(\frac{\rho}{2})} \qquad j = 0, \dots, n.$$

Proof: The support points and weights of the optimal discriminating design ξ^* are given in Proposition 3.3. It is easy to see that the polynomials $Q_\ell(x)$ and $P_\ell(x)$ are functions of $x - \frac{a+b}{2}$ and we may assume without loss of generality that the design space is the interval $[-\frac{b-a}{2}, \frac{b-a}{2}]$, that is $\tau = \frac{a+b}{2} = 0$. By Lemma 2.10 in Studden (1982b) the design $\tilde{\xi}$ with canonical moments $(\frac{1}{2}, \tilde{p}_2, \frac{1}{2}, \dots, \frac{1}{2}, \tilde{p}_{2n-2}, \frac{1}{2}, 1)$ and

$$\tilde{p}_{2j} = 1 - p^*_{2(n-j)} = 1 - \frac{U_{j+1}(\frac{\rho}{2})}{\rho U_j(\frac{\rho}{2})} = \frac{U_{j-1}(\frac{\rho}{2})}{\rho U_j(\frac{\rho}{2})} \qquad (j = 1, \dots, n-1) \tag{4.9}$$

has the same support points as ξ^* (here we used (4.3) and the recursive relation for the Chebyshev polynomials of the second kind (see e.g. Rivlin (1990), p. 40)). An expansion of the determinent in (3.7) yields that the support points of $\tilde{\xi}$ are given by zeros of the polynomial $(x^2 - \sigma^2)\tilde{Q}_{n-1}(x)$ where $\sigma = (b-a)/2$ and $\tilde{Q}_\ell(x)$ is defined recursively by

$$\begin{aligned} \tilde{Q}_{\ell+1}(x) &= x\tilde{Q}_\ell(x) - \sigma^2 \tilde{p}_{2\ell}\tilde{q}_{2\ell+2}\tilde{Q}_{\ell-1}(x) \\ &= x\tilde{Q}_\ell(x) - \vartheta^{-2}\frac{U_{\ell-1}(\frac{\rho}{2})U_{\ell+2}(\frac{\rho}{2})}{U_\ell(\frac{\rho}{2})U_{\ell+1}(\frac{\rho}{2})}\tilde{Q}_{\ell-1}(x) \end{aligned} \tag{4.10}$$

$(\ell = 0, \dots, n-2), \tilde{Q}_0(x) = 1, \tilde{Q}_1(x) = x$. Using the recursive relation for the polynomials $U_\ell(x)$ it now follows by an induction argument and formula (22.7.27) in Abramowitz and Stegun (1964) that

$$\tilde{Q}_\ell(x) = \frac{U_{\ell+2}(\frac{x\vartheta}{2})U_\ell(\frac{\rho}{2}) - U_\ell(\frac{x\vartheta}{2})U_{\ell+2}(\frac{\rho}{2})}{\vartheta^\ell((x\vartheta)^2 - \rho^2)U_\ell(\frac{\rho}{2})} = \frac{U_{\ell+2}(\frac{x\vartheta}{2})U_\ell(\frac{\rho}{2}) - U_\ell(\frac{x\vartheta}{2})U_{\ell+2}(\frac{\rho}{2})}{\vartheta^{\ell+2}(x^2 - \sigma^2)U_\ell(\frac{\rho}{2})} \tag{4.11}$$

and the assertion about the support points is an immediate consequence of Proposition 3.3 (note that $\rho = \vartheta\sigma$). For the second part we remark that the polynomial $Q_{n-1}(x)$ in (3.7) (corresponding to the canonical moments of ξ^*) satisfies $Q_{n-1}(x) = \tilde{Q}_{n-1}(x)$. Because $\rho \geq 2$ we have $\mathcal{N}(\xi^*) = \{1, \ldots, n\}$ and obtain from (4.1) and the discussion in the proof of Theorem 4.1

$$q^*_{2\ell} p^*_{2\ell+2} \;=\; \rho^{-2} \;=\; \sigma^{-2}\vartheta^{-2} \qquad (\ell = 1, \ldots, n-1). \tag{4.12}$$

This implies, for the polynomials $P_\ell(x)$ defined in (3.6), that

$$P_{\ell+1}(x) = xP_\ell(x) - \vartheta^{-2}P_{\ell-1}(x) \quad (\ell = 0, 1, \ldots, n-1)$$

$(P_0(x) = 1, P_1(x) = x)$ and it is easy to see that these polynomials are given by

$$P_\ell(x) = \vartheta^{-\ell} U_\ell(\frac{x\vartheta}{2}).$$

The assertion about the weights now follows directly from Proposition 3.3 and (4.11). ∎

Theorem 4.4. Let $\rho > 0$ and k be defined by (4.3), then the support points $x_0, \ldots, x_n$ of the optimal discriminating design are given by $x_j = z_j + \frac{a+b}{2}$ $(j = 0, \ldots, n)$ where $z_0, \ldots, z_n$ are the zeros of the polynomial

$$\begin{aligned} H_{n+1}(x) = \; & U_{k-1}\left(\frac{x\vartheta}{\rho}\right)\left\{U_{n-k+2}\left(\frac{x\vartheta}{2}\right)U_{n-k}\left(\frac{\rho}{2}\right) - U_{n-k}\left(\frac{x\vartheta}{2}\right)U_{n-k+2}\left(\frac{\rho}{2}\right)\right\} \\ & - U_{k-2}\left(\frac{x\vartheta}{\rho}\right)\left\{U_{n-k+1}\left(\frac{x\vartheta}{2}\right)U_{n-k-1}\left(\frac{\rho}{2}\right) - U_{n-k-1}\left(\frac{x\vartheta}{2}\right)U_{n-k+1}\left(\frac{\rho}{2}\right)\right\} \end{aligned}$$

while the weights are given by

$$\xi(\{x_j\}) = \frac{U_{n-k+1}(\frac{z_j\vartheta}{2})U_{k-1}(\frac{z_j\vartheta}{\rho})U_{n-k}(\frac{\rho}{2}) - U_{n-k}(\frac{z_j\vartheta}{2})U_{k-2}(\frac{z_j\vartheta}{\rho})U_{n-k+1}(\frac{\rho}{2})}{\vartheta^{-1}\cdot\frac{d}{dx}H_{n+1}(x)|_{x=z_j}}$$

$(j = 0, \ldots, n)$.

Proof. As in the proof of Theorem we assume without loss of generality that the design space is given by the interval $[-\frac{b-a}{2}, \frac{b-a}{2}]$ and consider the reversed sequence $(\frac{1}{2}, \tilde{p}_2, \frac{1}{2}, \ldots, \frac{1}{2}, \tilde{p}_{2n-2}, \frac{1}{2}, 1)$ where $\tilde{p}_{2j} = 1 - p^*_{2(n-j)}$ $(j = 1, \ldots, n-1)$. The corresponding design $\tilde{\xi}$ has the same support points as ξ^* and from the definition of k in (4.3) and Remark 4.2 we have $\tilde{p}_{2j} = \frac{1}{2}$ $(j = n-k+1, \dots, n-1)$. Thus it follows from (3.7) that $(\sigma = (b-a)/2)$

$$\tilde{Q}_{n-1}(x) = K\begin{pmatrix} \overbrace{-\left(\frac{\sigma}{2}\right)^2 \cdots -\left(\frac{\sigma}{2}\right)^2}^{k-2} & & -\sigma^2\tilde{p}_{2(n-k)}\tilde{q}_{2(n-k+1)} & \cdots & -\sigma^2\tilde{p}_2\tilde{q}_4 & \\ x \quad\quad x \;\; \dots \;\; x & & x & \dots & x & x \end{pmatrix}$$

$$= K\begin{pmatrix} \overbrace{-\left(\frac{\sigma}{2}\right)^2 \cdots -\left(\frac{\sigma}{2}\right)^2}^{k-2} \\ x \quad\quad x \;\; \dots \;\; x \end{pmatrix} K\begin{pmatrix} -\sigma^2\tilde{p}_{2(n-k-1)}\tilde{q}_{2(n-k)} & \cdots & -\sigma^2\tilde{p}_2\tilde{q}_4 & \\ x & x \;\; \dots & x & x \end{pmatrix}$$

$$- \;\sigma^2\tilde{p}_{2(n-k)}\tilde{q}_{2(n-k+1)} K\begin{pmatrix} & \overbrace{-\left(\frac{\sigma}{2}\right)^2 \cdots -\left(\frac{\sigma}{2}\right)^2}^{k-3} & \\ x & x \;\; \dots\dots \;\; x & x \end{pmatrix}$$

$$\times K\begin{pmatrix} -\sigma^2\tilde{p}_{2(n-k-2)}\tilde{q}_{2(n-k-1)} & & \sigma^2\tilde{p}_2\tilde{q}_4 & \\ x & x \;\; \dots & x & x \end{pmatrix}$$

where the last line follows from Sylvesters identity. Observing the reasoning (4.10) and (4.11) in the proof of Theorem 4.3 we see that the second factors in both terms are given by (4.11) for $\ell = n-k$ and $\ell = n-k-1$, respectively. By the same reasoning as at the end of the proof of Theorem 4.3 the first factors satisfy the recurrence relation

$$\tilde{P}_{l+1}(x) \;=\; x\tilde{P}_l(x) - \left(\frac{\sigma}{2}\right)^2\tilde{P}_{l-1}(x)$$

$(\tilde{P}_{-1}(x) = 0,\ \tilde{P}_0(x) = 1)$ and are given by $(\frac{\sigma}{2})^{k-1}U_{k-1}(\frac{x}{\sigma})$ and $(\frac{\sigma}{2})^{k-2}U_{k-2}(\frac{x}{\sigma})$ respectively. Thus it follows from (4.2) and $\tilde{p}_{2j} = 1 - p^*_{2(n-j)}$ that

$$\tilde{Q}_{n-1}(x) = \left(\frac{\sigma}{2}\right)^{k-1}\left[U_{k-1}\left(\frac{x\vartheta}{\rho}\right)Q^*_{n-k}(x) - \frac{\sigma}{\rho}\frac{U_{n-k-1}(\frac{\rho}{2})}{U_{n-k}(\frac{\rho}{2})}U_{k-2}\left(\frac{x\vartheta}{\rho}\right)Q^*_{n-k-1}(x)\right]$$

where $Q^*_\ell(x)$ is defined by (4.11). This yields

$$(x^2 - \sigma^2)\tilde{Q}_{n-1}(x) = \left(\frac{\sigma}{2}\right)^{k-1}\vartheta^{k-n-2}\cdot\frac{H_{n+1}(x)}{U_{n-k}\left(\frac{\rho}{2}\right)} \tag{4.13}$$

and proves the assertion for the support points. For the calculation of the weights of the design ξ^* we obtain for the polynomial $P_n(x)$ in (3.6) by a similar reasoning

$$\begin{aligned} P_n(x) = \frac{\left(\frac{\sigma}{2}\right)^{k-1}\vartheta^{k-n-1}}{U_{n-k}\left(\frac{\rho}{2}\right)}\Bigg\{ & U_{n-k+1}\left(\frac{x\vartheta}{2}\right)U_{k-1}\left(\frac{x\vartheta}{\rho}\right)U_{n-k}\left(\frac{\rho}{2}\right) - \\ & U_{n-k}\left(\frac{x\vartheta}{2}\right)U_{k-2}\left(\frac{x\vartheta}{\rho}\right)U_{n-k+1}\left(\frac{\rho}{2}\right)\Bigg\} \end{aligned} \tag{4.14}$$

and Theorem 4.4 follows by a further application of Proposition 3.3. ■

Remark 4.5. Spruill (1990) considered the case $[a,b] = [-1,1], \vartheta = 1$ which yields $\rho = 1, k = n$. For this case the polynomial $H_{n+1}(x)$ in Theorem 4.4 is given by

$$U_{n-1}(x)(x^2-1)$$

and it is not too hard to show that the weights at the support points $x_j = -\cos\left(\frac{\pi}{n}j\right)$ $(j = 0,\ldots,n)$ are proportional to $1{:}2{:}\ldots{:}2{:}1$ (see Spruill (1990)). Another case of interest is $\rho = 2$ because at this point the sequence in (4.4) changes from a divergent sequence into a convergent sequence $(j \to -\infty)$. Observing (4.2) and $U_n(1) = n+1$ we obtain

$$p^*_{2j} = \frac{n-j+2}{2(n-j)+2} \qquad (j = 1,\ldots,n)$$

$(p^*_{2j-1} = \frac{1}{2})$ and Theorem 2.5 of Dette (1992) yields that in this case the optimal discriminating design ξ^* has masses proportional to $2{:}3{:}\ldots{:}3{:}2$ at the zeros of the polynomial $(x^2-1)U_n'(x)$.

5. Asymptotic distributions. In this section we consider the same setting as in Section 4 when the degree n of the polynomial is large. By Theorem 4.1 and (4.4) the canonical moments (of even order) of the optimal discriminating design ξ^* are given by

$$(5.1) \qquad \begin{aligned} p^*_{2j} &= \frac{1}{2} & 1 \le j \le k-1 \\ p^*_{2j} &= 1 - \frac{\rho^{-2}}{p^*_{2j+2}} & k \le j \le n-1 \end{aligned}$$

where k is defined in (4.3). If $\rho < 2$, then it follows from the discussion in Remark 4.2 that for sufficiently large n there exists an index j_0 such that $p^*_{2i} = \frac{1}{2}$ for all $i \le n - j_0$. Consequently the optimal discriminating design converges (weakly) to the probability measure with canonical moments $p_i = \frac{1}{2}$ for all $i \in \mathbb{N}$. If $\rho \ge 2$, then Remark 4.2 shows that $p^*_{2j} > \frac{1}{2}$ for all $j \in \mathbb{N}$. The sequence (5.1) is decreasing (i.e. $p^*_{2j-2} < p^*_{2j}$) and consequently, as n tends to infinity, the canonical moments of the optimal design ξ^* converge to the distribution for which

$$\begin{aligned} p_{2j} &= z \\ p_{2j-1} &= \frac{1}{2} \end{aligned}$$

for all $j \in \mathbb{N}$ where $z \geq \frac{1}{2}$ is the fixpoint of the equation $z = 1 - \frac{\rho^{-2}}{z}$, that is

$$z = \frac{1}{2} + \frac{1}{2}\sqrt{1 - 4\rho^{-2}}. \tag{5.2}$$

Theorem 5.1.

a) If $\rho < 2$, then the optimal discriminating design ξ^* converges weakly to the arc-sin distribution with density

$$f(x) = \begin{cases} \frac{1}{\pi} \frac{1}{\sqrt{(x-a)(b-x)}} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

b) If $\rho \geq 2$, then the optimal discriminating design ξ^* converges weakly to the distribution with density

$$f(x) = \begin{cases} \frac{2(b-a)}{\pi} \frac{\sqrt{|16\vartheta^{-2} - (2x - (a+b)|^2}}{[(b-a) + \sqrt{(b-a)^2 - 16\vartheta^{-2}}][(b-a)^2 - (2x - (a+b))^2]} & \text{if } |x - (a+b)/2| < 2|\vartheta|^{-1} \\ 0 & \text{otherwise} \end{cases}$$

Proof. By the discussion at the beginning of this Section we have to identify the (unique) probability measure η corresponding to the sequence of canonical moments $z, z, \ldots$, where $z \geq \frac{1}{2}$. Let $\sigma = (b-a)/2$, $\tau = (a+b)/2$ then it follows (see e.g. Dette (1992) p. 245) that the Stieltjes transform of η is given by the continued fraction expansion

$$\Phi(w) = \int_a^b \frac{d\eta(t)}{w-t} = \frac{1}{w - \tau - \sigma^2 z G(w)}$$

where

$$G(w) = \frac{1 \mid}{|w - \tau} - \frac{\sigma^2 z(1-z)|}{| \ w - \tau} - \ \cdots \ = \frac{1}{w - \tau - \sigma^2 z(1-z)G(w)}$$

Solving with respect to $G(w)$ yields

$$G(w) = \frac{w - \tau - \sqrt{(w-\tau)^2 - 4\sigma^2(1-z)z}}{2\sigma^2 z(1-z)}$$

and consequently we obtain by straightforward calculation

$$\Phi(w) = -\frac{1}{2z}\frac{(1-2z)(w-\tau) - \sqrt{(w-\tau)^2 - 4\sigma^2 z(1-z)}}{(w-\tau)^2-\sigma^2}.$$

¿From the inversion formula for Stieltjes transforms (see e.g. Perron (1954)) it follows that η has the density

$$-\frac{1}{\pi}\operatorname{Im}(\Phi(x)) = \frac{1}{2\pi z}\frac{\sqrt{4\sigma^2 z(1-z) - (x-\tau)^2}}{\sigma^2 - (x-\tau)^2}$$

where $|x-\tau|^2 \le 4\sigma^2 z(1-z)$. If $\rho \ge 2$ then $4\sigma^2 z(1-z) = 4\vartheta^{-2}$ (which follows from 4.12)), $\tau = (a+b)/2$ and using (5.2) the density of η can be written as

$$f(x) = \frac{2(b-a)}{\pi}\frac{\sqrt{16\vartheta^{-2} - (2x-(a+b))^2}}{[(b-a)+\sqrt{(b-a)^2 - 16\vartheta^{-2}}][(b-a)^2 - (2x-(a+b))^2]}$$

($|x-(a+b)/2| \le 2|\vartheta|^{-1}$). Similarly we have for $\rho < 2$ that $z = \frac{1}{2}$ and the limiting density is given by

$$f(x) = \frac{1}{\pi}\frac{1}{\sqrt{(b-x)(x-a)}} \qquad (a < x < b)\ .$$

■

Example 5.2. It might be of interest how the designs behave in a situation where they are not optimal (e.g. Spruill's design on the interval $[-b,b]$). To this end consider the case $a=-b, \vartheta = 1, \rho = b$ and $n=3$ (cubic regression). If $b \le \sqrt{2}$, then the design determined by Spruill (1990) is the optimal discriminating design, that is

$$\xi_1 = \begin{pmatrix} -b & -\frac{b}{2} & \frac{b}{2} & b \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$$

and the canonical moments (of even order) of ξ_1 are $p_2 = p_4 = \frac{1}{2}$ and $p_6 = 1$ (this follows from Theorem 4.1 and 4.3). If $\sqrt{2} \le b \le \sqrt{3}$ we obtain from the results of Section 4 that the optimal discriminating design is given by

$$\xi_2 = \begin{pmatrix} -b & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & -b \\ \frac{1}{2}\frac{b^2-1}{2b^2-1} & \frac{1}{2}\frac{b^2}{2b^2-1} & \frac{1}{2}\frac{b^2}{b^2-1} & \frac{1}{2}\frac{b^2-1}{2b^2-1} \end{pmatrix}$$

with canonical moments (of even order) $p_2 = \frac{1}{2}, p_4 = 1-b^{-2}, p_6 = 1$. Note that the interior support points are independent of $b \in [\sqrt{2},\sqrt{3}]$. Finally if $b \ge \sqrt{3}$ the optimal design is

$$\xi_3 = \begin{pmatrix} -b & -\sqrt{\frac{b^2-2}{b^2-1}} & \sqrt{\frac{b^2-2}{b^2-1}} & b \\ \frac{1}{2}\left(1 - \frac{b^2}{b^4-2b^2+2}\right) & \frac{1}{2}\frac{b^2}{b^4-2b^2+2} & \frac{1}{2}\frac{b^2}{b^4-2b^2+2} & \frac{1}{2}\left(1-\frac{b^2}{b^4-2b^2+2}\right) \end{pmatrix}$$

with canonical moment (of even order) $p_2 = \frac{b^2-2}{b^2-1}, p_4 = 1 - b^{-2}, p_6 = 1$. (this follows from Theorem 4.3 or 4.4). For the values of the optimality criterion (4.1) at these point we obtain

$$\Phi(\xi_1) = \begin{cases} \frac{b^6}{16} & \text{if } b \le \sqrt{2} \\ \frac{b^4}{8} & \text{if } \sqrt{2} \le b \le 2 \\ \frac{b^2}{2} & \text{if } 2 \le b \end{cases}$$

$$\Phi(\xi_2) = \begin{cases} \frac{b^2}{4}(b^2-1) & \text{if } 1 < b \le \sqrt{3} \\ \frac{b^2}{2} & \text{if } b > \sqrt{3} \end{cases}$$

$$\Phi(\xi_3) = b^2 \frac{b^2-2}{b^2-1} \qquad \text{for all } b \ge \sqrt{2}.$$

Table 5.3 gives some efficiencies for different values of the parameter b. Note that the designs ξ_2 and ξ_3 are not defined if b is too small (i.e. $b \le 1$ for ξ_2 and $b \le \sqrt{2}$ for ξ_3). This is caused by the fact that the designs have to satisfy $\mathcal{N}(\xi_2) = \{2,3\}$ and $\mathcal{N}(\xi_3) = \{1,2,3\}$ which is impossible if b is too small. We conclude with the statement that the performance of the different designs will depend heavily on the length of the interval.

design/b	$\sqrt{2}$	1.5	1.6	1.7	$\sqrt{3}$	2	3	10	∞
ξ_1	1	0.9	0.8205	0.7646	0.75	0.75	0.5714	0.5051	0.5
ξ_2	1	1	1	1	1	0.75	0.5714	0.5051	0.5
ξ_3	0	0.64	0.9204	0.9966	1	1	1	1	1

Table 5.1: Efficiencies of the designs ξ_1, ξ_2 and ξ_3 for different intervals $[-b, b]$

Acknowledgements. This paper was written while the author was visiting Purdue University. The author would like to thank the Department of Statistics for its hospitality and the Deutsche Forschungsgemeinschaft for the support that made this visit possible. I am also indebted to Bill Studden for many helpful discussions during the preparation of the paper and to Mary Epperson for her excellent typing of the manuscript..

References

M. Abramowitz, I.S. Stegun (1964). Handbook of Mathematical functions. Dover, New York.

T.W. Anderson (1962). The choice of the degree of a polynomial regression as a multiple decision problem. *Ann. Math. Statist.*, **33**, 255–265.

H. Dette (1991). A note on model robust designs for polynomial regression. *Journal of Statistical Planning and Inference*, **28**, 223–232.

H. Dette (1992). Optimal designs for a class of polynomials of odd and even degree. *Ann. Statist.*, **20**, 238–259.

H. Dette (1993). Elfving's theorem for D-optimality. *Ann. Statist.*, **21**, June.

J. Kiefer, J. Wolfowitz (1959). Optimum designs in regression problems. *Ann. Math. Statist.*, **30**, 271–294.

T.S. Lau (1983). Theory of canonical moments and its application in polynomial regression I and II. Technical Reports 83–23, 83–24, Department of Statistics, Purdue University.

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

T.S. Lau (1988). D-optimal designs on the unit q-ball. *Journal of Statistical Planning and Inference*, **19**, 299–315.

O. Perron (1954). *Die Lehre von den Kettenbrüchen (Band I, II).* B.G. Teubner, Stuttgart.

F. Pukelsheim (1993). *Optimal design of experiments.* Wiley: New York.

T.J. Rivlin (190). *Chebyshev polynomials.* Wiley: New York.

M. Skibinsky (1986). Principal representations and canonical moment sequences for distributions on an interval. *J. Math. Anal. and Applications*, **120**, 95–120.

M.G. Spruill (1990). Good designs for testing the degree of a polynomial mean. *Sankhyā: The Indian Journal of Statistics, Series B*, **52**, 67–74.

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

W.J. Studden (1982a). Some robust-type D-optimal designs in polynomial regression. *J. American Statist. Assoc.*, **77**, 916–921.

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