The Annals of Statistics
1998, Vol. 26, No. 4, 1496–1521

OPTIMAL DESIGNS FOR THE IDENTIFICATION OF THE ORDER OF A FOURIER REGRESSION

BY HOLGER DETTE AND GERD HALLER

Ruhr-Universität Bochum

For the Fourier regression model, we determine optimal designs for identifying the order of periodicity. It is shown that the optimal design problem for trigonometric regression models is equivalent to the problem of optimal design for discriminating between certain homo- and heteroscedastic polynomial regression models. These optimization problems are then solved using the theory of canonical moments, and the optimal discriminating designs for the Fourier regression model can be found explicitly. In contrast to many other optimality criteria for the trigonometric regression model, the optimal discriminating designs are not uniformly distributed on equidistant points.

1. Introduction. Consider the standard Fourier regression model

$$g_{2d}(x) = a_0 + \sum_{j=1}^{d} a_j \sin(jx) + \sum_{j=1}^{d} b_j \cos(jx), \qquad x \in [-\pi, \pi], \tag{1.1}$$

where $(a_0, a_1, \dots, a_d, b_1, \dots, b_d)^T$ denotes the $(2d+1)$-dimensional vector of unknown parameters. Note that one of the boundary points of the design space $[-\pi, \pi]$ could be omitted because of the periodicity of the regression functions. Applications of lower order trigonometric polynomials are given in Mardia (1972). The problem of the optimal design of experiments for model (1.1) has been discussed by several authors [see, e.g., Hoel (1965), Karlin and Studden (1966), page 347, Federov (1972), page 94, Hill (1978), Lau and Studden (1985), Riccomagno, Schwabe and Wynn (1997)]. It is well known that the equally spaced design points on an equidistant grid with at least $2d+1$ points are ϕ_p-optimal for estimating the parameters in the model (1.1) in the sense of Kiefer (1974) [see Pukelsheim (1993), page 241].

While most of this work considers a fixed given regression model, much less attention has been paid to the problem of constructing optimal designs for the identification of the relevant parameters in the trigonometric regression. However, discrimination designs have been discussed in the context of other linear and nonlinear models [see, e.g., Atkinson (1972), Atkinson and Cox (1974) for some early references and Spruill (1990) for some recent work]. In this paper we consider the optimal design problem, when Anderson's procedure [Anderson (1962)] is applied in order to reduce the degree of the

Received April 1997; revised April 1998.
AMS 1991 *subject classifications*. 62K05, 62J05.
Key words and phrases. Fourier regression model, optimal design, model discrimination, heteroscedastic polynomial regression, canonical moments.

Fourier regression model. More precisely, suppose that $n \geq 2d + 1$ independent responses $Y_1, \dots, Y_n$ are observed by the experimenter, where $Y_j \sim N(g(x_j), \sigma^2)$, $x_j \in [-\pi, \pi]$ $(j = 1, \dots, n)$, $\sigma^2 > 0$. We assume that the regression function g is unknown but belongs to the class of models

$$\mathscr{F}_{2d} := \{g_0, g_1, \dots, g_{2d}\}$$

where g_{2k} is defined by (1.1) ($k = 0, \dots, d$; $g_0 \equiv a_0$) and

$$g_{2k-1}(x) = a_0 + \sum_{j=1}^{k} a_j \sin(jx) + \sum_{j=1}^{k-1} b_j \cos(jx)$$

denotes the Fourier regression (1.1) without the "highest" term $\cos(kx)$. A reasonable procedure to identify how many trigonometric regression functions should be used for fitting the data is the following. Starting with the given regression $g_{2d}(x)$ in (1.1), one tests successively the hypotheses

$$(1.2) \quad H_0^{(2d)}: b_d = 0, \qquad H_0^{(2d-1)}: a_d = 0, \dots, H^{(2)}: b_1 = 0, \qquad H_0^{(1)}: a_1 = 0$$

and decides for the model g_{k_0} where k_0 is the first index for which $H_0^{(k_0)}$ is rejected. The statistical properties of this procedure in the context of multiple testing are carefully explained in the monograph of Anderson [(1994), pages 34–46]. Roughly speaking the investigator specifies error probabilities, say $\alpha_1, \dots, \alpha_{2d}$, of using the model g_k when in fact a model g_l with $l < k$ is needed (for all $k = 1, \dots, 2d$). Under these restrictions it is shown that this procedure satisfies several optimality properties; especially, it minimizes the error probability of choosing a Fourier regression model with too many parameters [see also Spruill (1990)].

It is easy to see that the statistical properties of the corresponding F-tests for the hypothesis $H_0^{(l)}$ depend on the design of experiment only through the noncentrality parameters, say δ_l, $l = 1, \dots, 2d$. The error probabilities $\alpha_1, \dots, \alpha_{2d}$ are independent of the design (because the corresponding noncentrality parameters vanish), while the power function of the test for the hypothesis $H_0^{(l)}$ is an increasing function of δ_l, $l = 1, \dots, 2d$. Therefore it is desirable to maximize $\delta_1, \dots, \delta_{2d}$ with respect to the choice of the underlying design. Unfortunately a simultaneous maximization is only possible in very special cases and it is common practice to maximize a concave function of the noncentrality parameters $\delta_1, \dots, \delta_{2d}$, which is called optimality criterion.

In this paper we determine optimal approximate discriminating designs which maximize a weighted p-mean of the noncentrality parameters. In Section 2 the optimality criterion is defined and some basic properties are derived. It is also shown that the problem of optimal design for identifying the appropriate model $g_{k_0} \in \mathscr{F}_{2d}$ can be reduced to the maximization of a composite optimality criterion for weighted polynomial regression models. This problem is then solved in Sections 3 and 4 by combining the general equivalence theory of optimal design [see Pukelsheim (1993), Chapter 7] with the theory of canonical moments which was introduced by Skibinsky (1967) and applied by Studden (1980, 1982a, b, 1989) in the context of optimal design.

The results provide some new insight into the theory of optimal design for the Fourier regression model. It is demonstrated that the optimal design problems for Fourier regression models are equivalent to design problems for linear models with certain weighted and unweighted polynomials as regression functions. In particular, the optimal discriminating designs derived in this paper are not necessarily uniformly distributed on equidistant points (in contrast to the classical ϕ_p-optimal designs for estimating the parameters in the trigonometric regression models). In contrast to polynomial regression models [see Dette (1995)] the designs for identifying the degree of the Fourier regression are not unique and cannot be characterized by their first $2d$ (trigonometric) moments. Moreover, by putting a special prior on the class of models $\mathscr{F}_{2d}$, recent results of Dette (1994, 1995) for discrimination designs in homoscedastic polynomial regression are obtained as special cases. Additionally, this paper extends these findings to some heteroscedastic polynomial models as well and highlights the particular role of uniform designs for Fourier regression models.

2. Approximate designs, preliminary results. In the context of approximate design theory, a design is treated as a probability measure σ with finite support on the interval $[-\pi, \pi]$ with the interpretation that observations are taken in proportion to the corresponding masses. The analogue of the matrix $X^T X$ in the Fourier regression model $g_k(x)$ is the information matrix

$$M_k(\sigma) = \int_{-\pi}^{\pi} f_k(x) f_k^T(x)\, d\sigma(x), \tag{2.1}$$

where

$$f_k(x) = \begin{cases} (1, \sin(x), \cos(x), \dots, \sin(jx), \cos(jx))^T, & \text{if } k = 2j, \\ (1, \sin(x), \cos(x), \dots, \sin((j-1)x), \cos((j-1)x), \sin(jx))^T, & \\ & \text{if } k = 2j-1 \end{cases}$$

($k = 1, \dots, 2d$). The quantities corresponding to the noncentrality parameter of the F-test for the hypothesis $H_0^{(k)}$ are given by

$$\delta_k(\sigma) = \left(e_k^T M_k^{-1}(\sigma) e_k\right)^{-1}, \qquad k = 1, \dots, 2d, \tag{2.2}$$

where e_k denotes the $(k+1)$th unit vector in $\mathbb{R}^{k+1}$ and the design σ is assumed to have at least $(2d+1)$ support points [see Pukelsheim (1993), page 70]. Throughout this paper, let σ_k^* denote the D_1-optimal design maximizing δ_k and define the D_1-efficiency of a design σ in the Fourier regression model $g_k(x)$ by

$$\text{eff}_k(\sigma) := \frac{\delta_k(\sigma)}{\delta_k(\sigma_k^*)}, \qquad k = 1, \dots, 2d. \tag{2.3}$$

The optimal discriminating designs are now defined in a similar manner as in Dette (1994, 1995). More precisely, a design is called a $\Phi_{p,\pi}$-optimal discriminating design for the class $\mathscr{F}_{2d}$ with respect to the prior $\pi = (\pi_1, \dots, \pi_{2d})$ if and only if it maximizes the function

$$\Phi_{p,\pi}(\sigma) = \left[\sum_{k=1}^{2d} \pi_k \left(\text{eff}_k(\sigma)\right)^p\right]^{1/p}. \tag{2.4}$$

Here $-\infty \le p < 1$ and the cases $p = 0$ and $p = -\infty$ are understood as the corresponding limits, that is,

$$\Phi_{0,\pi}(\sigma) = \prod_{k=1}^{2d} \left(\text{eff}_k(\sigma)\right)^{\pi_k},$$

$$\Phi_{-\infty,\pi}(\sigma) = \min\{\text{eff}_k(\sigma) | \pi_k > 0\}.$$

The prior π reflects the experimenter's belief about the adequacy of the different models, and a higher weight π_k gives more power to the F-test for the hypothesis $H_0^{(k)}$. Note that the multiple level of significance of Anderson's procedure does not depend on the prior used in the optimality criterion (2.4). Moreover, in practical applications, the order of the model usually exceeds a minimal number, say $m > 1$ (in other words, there are at least $m + 1$ terms in the regression) and this situation can be reflected in the criterion (2.4) by simply putting $\pi_1 = \cdots = \pi_m = 0$. The following result describes the relation between $\Phi_{p,\pi}$-optimal discriminating designs with respect to different values of p.

LEMMA 2.1. *A design σ^* is $\Phi_{p,\pi}$-optimal (for some given $p > -\infty$) for the class $\mathscr{F}_{2d}$ with respect to the prior π if it is $\Phi_{0,\hat{\pi}}$-optimal with respect to the prior $\hat{\pi} = (\hat{\pi}_1, \dots, \hat{\pi}_{2d})$ where*

$$\hat{\pi}_k := \frac{\pi_k \left(\text{eff}_k(\sigma^*)\right)^p}{\sum_{l=1}^{2d} \pi_l \left(\text{eff}_l(\sigma^*)\right)^p}, \qquad k = 1, \dots, 2d. \tag{2.5}$$

Let $\mathscr{N}(\sigma^*) := \{1 \le j \le 2d | \pi_j > 0,\ \Phi_{-\infty,\pi}(\sigma^*) = \text{eff}_j(\sigma^*)\}$, then σ^* is a $\Phi_{-\infty,\pi}$-optimal discriminating design for the class $\mathscr{F}_{2d}$ if and only if $|M_k(\sigma)| \ne 0$ whenever $\pi_k > 0$ and there exists a prior $\hat{\pi} = (\hat{\pi}_1, \dots, \hat{\pi}_{2d})$ with $\hat{\pi}_l = 0$ for all $l \notin \mathscr{N}(\sigma^*)$ such that σ^* is a $\Phi_{0,\hat{\pi}}$-optimal discriminating design with respect to the prior $\hat{\pi}$.

PROOF. The general equivalence theorem for mixtures of optimality criteria [see Pukelsheim (1993), Chapter 11] shows that for $p > -\infty$ the design σ^* is $\Phi_{p,\pi}$-optimal with respect to the prior $\pi = (\pi_1, \dots, \pi_{2d})$ if and only if the inequality

$$\sum_{k=1}^{2d} \pi_k \left[\text{eff}_k(\sigma)\right]^{p+1} \left[e_k^T M_k^{-1}(\sigma^*) f_k(x)\right]^2 \le \sum_{k=1}^{2d} \pi_k \left[\text{eff}_k(\sigma)\right]^p$$

holds for all $x \in [-\pi, \pi]$. Similarly, σ^* is $\Phi_{-\infty,\pi}$-optimal if and only if there exists a prior $\hat{\pi} = (\hat{\pi}_1, \dots, \hat{\pi}_{2d})$ such that $\hat{\pi}_k = 0$ whenever $k \notin \mathscr{N}(\sigma^*)$ and such that

$$\sum_{k \in \mathscr{N}(\sigma^*)} \hat{\pi}_k \operatorname{eff}_k(\sigma) \left[e_k^T M_k^{-1}(\sigma^*) f_k(x) \right]^2 \le 1$$

for all $x \in [-\pi, \pi]$. The assertion of Lemma 2.1 is now obvious from these characterizations. □

It follows by standard arguments [see Pukelsheim (1993), Chapters 4, 5] that $\Phi_{p,\pi}$ is a concave function on the set of designs on the interval $[-\pi, \pi]$ and invariant with respect to a reflection of the design σ at the origin. Consequently, there exists a $\Phi_{p,\pi}$-optimal discriminating design in the set Σ of all symmetric designs on the interval $[-\pi, \pi]$. In the following we will make extensive use of the fact that the set Σ of symmetric designs on the circle can be mapped onto the set of designs on the interval $[-1, 1]$, say $\Sigma_{[-1,1]}$, in a one-to-one manner. More precisely, define for a symmetric design σ on the interval $[-\pi, \pi]$ its projection ξ_σ onto $[-1, 1]$ by

$$\xi_\sigma(\cos x) = \begin{cases} 2\sigma(x) = 2\sigma(-x), & \text{if } 0 < x \le \pi, \\ \sigma(0), & \text{if } x = 0. \end{cases} \tag{2.6}$$

Now, consider a symmetric design $\sigma \in \Sigma$ on the interval $[-\pi, \pi]$ and its projection ξ_σ onto the interval $[-1, 1]$ and let

$$T_k(z) = \cos(k \arccos(z)), \qquad U_k(z) = \frac{\sin((k+1)\arccos(z))}{\sin(\arccos(z))}$$

denote the kth Chebyshev polynomial of the first and second kind, respectively [see Rivlin (1990)]. If $T^{(k)}(z) = (T_0(z), \dots, T_k(z))^T$ and $U^{(k)}(z) = (U_0(z), \dots, U_k(z))^T$ denote the vector of Chebyshev polynomials up to degree k; then a straightforward calculation shows that for $\sigma \in \Sigma$,

$$\begin{aligned} |M_{2k}(\sigma)| &= |M_k^c(\sigma)|\,|M_k^s(\sigma)|, \\ |M_{2k-1}(\sigma)| &= |M_{k-1}^c(\sigma)|\,|M_k^s(\sigma)|. \end{aligned} \tag{2.7}$$

The matrices $M_k^s(\sigma)$ and $M_k^c(\sigma)$ in (2.7) are defined as follows:

$$\begin{aligned} M_k^c(\sigma) &= \left(\int_{-\pi}^{\pi} \cos(ix)\cos(jx)\, d\sigma(x) \right)_{i,j=0}^{k} \\ &= \left(\int_{-1}^{1} T_i(z) T_j(z)\, d\xi_\sigma(z) \right)_{i,j=0}^{k} \\ &= \int_{-1}^{1} T^{(k)}(z) T^{(k)}(z)^T\, d\xi_\sigma(z) = T A_k(\xi_\sigma) T^T, \end{aligned} \tag{2.8}$$

$$
\begin{aligned}
M_k^s(\sigma) &= \left(\int_{-\pi}^{\pi} \sin(ix)\sin(jx)\, d\sigma(x) \right)_{i,j=1}^{k} \\
&= \left(\int_{-1}^{1} (1 - z^2) U_i(z) U_j(z)\, d\xi_\sigma(z) \right)_{i,j=0}^{k-1} \qquad (2.9)\\
&= \int_{-1}^{1} (1 - z^2) U^{(k-1)}(z) U^{(k-1)}(z)^T\, d\xi_\sigma(z) \\
&= U B_k(\xi_\sigma) U^T,
\end{aligned}
$$

where

$$
A_k(\xi_\sigma) = \left(\int_{-1}^{1} z^{i+j}\, d\xi_\sigma(z) \right)_{i,j=0}^{k}, \tag{2.10}
$$

$$
B_k(\xi_\sigma) = \left(\int_{-1}^{1} (1 - z^2) z^{i+j}\, d\xi_\sigma(z) \right)_{i,j=0}^{k-1} \tag{2.11}
$$

denote the information matrices of the design ξ_σ on the interval $[-1, 1]$ for a homoscedastic and heteroscedastic polynomial regression with efficiency function $\lambda(z) = (1 - z^2)$ [see Fedorov (1972), page 39], $T \in \mathbb{R}^{(k+1)\times(k+1)}$ and $U \in \mathbb{R}^{k\times k}$ are lower triangular matrices with diagonal elements $(1, 1, 2, \dots, 2^{k-1})$ and $(1, 2, \dots, 2^{k-1})$, respectively. Observing the definition of $\delta_k(\sigma)$ in (2.2), (2.7)–(2.11), we therefore obtain for a symmetric design σ,

$$
\delta_{2k}(\sigma) = \frac{|M_k^c(\sigma)|}{|M_{k-1}^c(\sigma)|} = \frac{|A_k(\xi_\sigma)|}{|A_{k-1}(\xi_\sigma)|} 2^{2(k-1)}, \qquad k = 1, \dots, d, \tag{2.12}
$$

$$
\delta_{2k-1}(\sigma) = \frac{|M_k^s(\sigma)|}{|M_{k-1}^s(\sigma)|} = \frac{|B_k(\xi_\sigma)|}{|B_{k-1}(\xi_\sigma)|} 2^{2(k-1)}, \qquad k = 1, \dots, d, \tag{2.13}
$$

where $B_0(\xi_\sigma) = A_0(\xi_\sigma) = 1$, $A_k(\xi_\sigma)$, $B_k(\xi_\sigma)$ are defined by (2.10) and (2.11), respectively, and ξ_σ is the projection of σ via the transformation (2.6). Consequently, the problem of determining $\Phi_{p,\pi}$-optimal discriminating designs for the Fourier regression models in $\mathscr{F}_{2d}$ can be solved by maximizing a certain function over the set of probability measures on the interval $[-1, 1]$ and transforming the maximizing measure back via (2.6). Note that the problem of maximizing the right-hand side of (2.12) and (2.13) over the set $\Sigma_{[-1,1]}$ is in fact a D_1-optimal design problem. More precisely, these problems arise in the determination of the optimal design for the estimation of the highest coefficient in a homoscedastic polynomial regression of degree k and a heteroscedastic polynomial regression of degree $k - 1$ with variance function $\sigma^2(x) = \sigma^2/(1 - x^2)$, $x \in (-1, 1)$, respectively. The solutions of these problems and the optimal values in (2.12) and (2.13) are well known [see Studden (1968, 1982b)] as

$$
\delta_k(\sigma_k^*) = \max_\sigma \delta_k(\sigma) = 1, \qquad k = 1, \dots, 2d
$$

and by (2.12) and (2.13) the efficiencies in (2.3) can be rewritten as

$$\text{(2.14)} \qquad \text{eff}_k(\sigma) = \begin{cases} 2^{2(j-1)} \dfrac{|A_j(\xi_\sigma)|}{|A_{j-1}(\xi_\sigma)|}, & \text{if } k = 2j, \\[2ex] 2^{2(j-1)} \dfrac{|B_j(\xi_\sigma)|}{|B_{j-1}(\xi_\sigma)|}, & \text{if } k = 2j-1. \end{cases}$$

This gives for the $\Phi_{p,\pi}$-optimality criterion,

$$\text{(2.15)} \qquad \begin{aligned} \Phi_{p,\pi}(\sigma) = \Bigg[\sum_{k=1}^{d} \pi_{2k-1} & \left(2^{2(k-1)} \frac{|B_k(\xi_\sigma)|}{|B_{k-1}(\xi_\sigma)|} \right)^p \\ & + \pi_{2k} \left(2^{2(k-1)} \frac{|A_k(\xi_\sigma)|}{|A_{k-1}(\xi_\sigma)|} \right)^p \Bigg]^{1/p}, \end{aligned}$$

which corresponds to a composite optimality criterion [in the sense of Atkinson and Donev (1992)] for the class of polynomial models

$$\left\{ \sum_{j=0}^{k} a_j x^j, \sqrt{1-x^2} \sum_{j=0}^{k-1} b_j x^j \,\Big|\, k = 1, \dots, d \right\}$$

on the interval $[-1, 1]$.

3. $\Phi_{p,\pi}$-optimal discriminating designs for $p > -\infty$. Note that for the prior $\pi = (0, \pi_2, 0, \dots, 0, \pi_{2d})$ the optimality criterion (2.15) reduces to a function which was already considered by Dette (1994, 1995) for determining optimal discriminating designs for the class of homoscedastic polynomial models up to degree d. Similarly, the prior $(\pi_1, 0, \pi_3, 0, \dots, \pi_{2d-1}, 0)$ corresponds to the problem of optimal discriminating design for heteroscedastic polynomial models with variance function $\sigma^2(x) = \sigma^2/(1-x^2)$, which has not been discussed so far. An important tool used in optimal design for polynomials is the theory of canonical moments which was introduced by Studden (1980, 1982a, b) in this context. We will only give a very brief heuristical introduction of this concept, which should be sufficient for the purpose of this paper. For more details which are needed in the Appendix, we refer to the work of Lau (1983, 1988), Skibinsky (1986) and the recent monograph of Dette and Studden (1997). It is well known that a probability measure on the interval $[-1, 1]$, say ξ, is determined by its sequence of moments $(c_1, c_2, \dots)$. Skibinsky (1967) defined a one-to-one mapping from the sequences of ordinary moments onto sequences $(p_1, p_2, \dots)$ whose elements vary independently in the interval $[0, 1]$. For a given probability measure on the interval $[-1, 1]$ the element p_j of the corresponding sequence is called the jth canonical moment of ξ. In order to indicate the dependence on ξ we use at some places the notation $p_j(\xi)$. The dependence on the design is omitted whenever it is clear from the context. If j is the first index for which

$p_j \in \{0, 1\}$, then the sequence of canonical moments terminates at p_j and the measure is supported at a finite number of points. The support points and corresponding masses can be found explicitly by evaluating certain orthogonal polynomials [see Skibinsky (1986) and Lau (1988)]. The set of probability measures on the interval $[-1, 1]$ with first k canonical moments equal to $(p_1, \dots, p_k) \in (0,1)^{k-1} \times [0, 1]$ is a singleton if and only if $p_k \in \{0, 1\}$. Otherwise there exists an uncountable number of probability measures corresponding to $(p_1, \dots, p_k)$ [see Skibinsky (1986)].

It turns out that the canonical moments of the $\Phi_{p,\pi}$-optimal discriminating designs can be found analytically, which provides a complete solution of the design problem. To this end we remark that the determinants of the matrices $A_k(\xi)$ and $B_k(\xi)$ can be easily expressed in terms of the canonical moments of the probability measure ξ [see Studden (1982b)], that is,

$$
\begin{aligned}
|A_k(\xi)| &= 2^{k(k+1)} \prod_{l=1}^{k} (q_{2l-2} p_{2l-1} q_{2l-1} p_{2l})^{k-l+1}, \\
|B_k(\xi)| &= 2^{k(k+1)} \prod_{l=1}^{k} (p_{2l-2} q_{2l-1} p_{2l-1} q_{2l})^{k-l+1},
\end{aligned} \tag{3.1}
$$

where $p_1, p_2, \dots$ denote the canonical moments of ξ ($p_0 = 1$) and $q_j = 1 - p_j$ ($j \geq 1$), $q_0 = 1$. Observing (2.15) and (3.1), we see that $\Phi_{p,\pi}$ is an increasing function of $p_{2j-1} q_{2j-1}$ $(j = 1, \dots, d)$ and consequently the canonical moments of the projection ξ_{σ^*} of the $\Phi_{p,\pi}$-optimal discriminating design for the class $\mathscr{F}_{2d}$ must satisfy

$$
p_{2l-1} = \tfrac{1}{2}, \qquad l = 1, \dots, d \tag{3.2}
$$

if $p > -\infty$. Similarly, there exists at least one $\Phi_{-\infty,\pi}$-optimal discriminating design for the class $\mathscr{F}_{2d}$ such that the canonical moments of the corresponding projection satisfy (3.2). Therefore, we can restrict ourselves to designs with this property and (2.14) and (2.15) reduce to

$$
\operatorname{eff}_k(\sigma) = \begin{cases} 2^{2j-2} p_{2j} \displaystyle\prod_{l=1}^{j-1} q_{2l} p_{2l}, & \text{if } k = 2j, \\ 2^{2j-2} q_{2j} \displaystyle\prod_{l=1}^{j-1} q_{2l} p_{2l}, & \text{if } k = 2j-1, \end{cases} \tag{3.3}
$$

$$
\begin{aligned}
\Phi_{p,\pi}(\sigma) = \Bigg[\sum_{k=1}^{d} \pi_{2k-1} \Bigg(2^{2k-2} q_{2k} \prod_{l=1}^{k-1} q_{2l} p_{2l} \Bigg)^p \\
+ \pi_{2k} \Bigg(2^{2k-2} p_{2k} \prod_{l=1}^{k-1} q_{2l} p_{2l} \Bigg)^p \Bigg]^{1/p},
\end{aligned} \tag{3.4}
$$

where $p_2, p_4, \dots$ denote the canonical moments of even order of the design ξ_σ on the interval $[-1, 1]$ satisfying (3.2) and corresponding to σ via (2.6). We first start with the $\Phi_{0,\pi}$-optimality criterion for which the solution is now relatively obvious. Throughout this paper we assume without loss of generality that either $\pi_{2d-1} > 0$ or $\pi_{2d} > 0$ and define $\sum_{k=i}^{j} a_k = 0$ if $i > j$.

LEMMA 3.1. *A symmetric design* σ^* *is a* $\Phi_{0,\pi}$*-optimal discriminating design for the class* $\mathscr{F}_{2d}$ *with respect to the prior* $\pi = (\pi_1, \dots, \pi_{2d})$ *if and only if the canonical moments of its projection* ξ_{σ^*} *via* (2.6) *satisfy*

$$(3.5)\qquad \begin{aligned} p_{2j-1} &= \frac{1}{2}, \\ p_{2j} &= \frac{\pi_{2j} + \sum_{i=j+1}^{d}(\pi_{2i-1} + \pi_{2i})}{\pi_{2j} + \pi_{2j-1} + 2\sum_{i=j+1}^{d}(\pi_{2i-1} + \pi_{2i})}, \qquad j = 1, \dots, d. \end{aligned}$$

Moreover, σ^* *is unique if and only if* $\pi_{2d} = 0$ *or* $\pi_{2d-1} = 0$.

PROOF. By the previous discussion, the canonical moments of ξ_{σ^*} satisfy (3.2). For $p = 0$ the optimality criterion therefore reduces to

$$\Phi_{0,\pi}(\sigma) = C \prod_{l=1}^{d} \left(p_{2l} \prod_{i=1}^{l-1} q_{2i} p_{2i} \right)^{\pi_{2l}} \left(q_{2l} \prod_{i=1}^{l-1} q_{2i} p_{2i} \right)^{\pi_{2l-1}},$$

which is uniquely maximized for the canonical moments in (3.5). Consequently, every design σ whose canonical moments of ξ_σ up to the order $2d$ satisfy (3.5) is $\Phi_{0,\pi}$-optimal discriminating with respect to the prior π.

If $\pi_{2d} = 0$ or $\pi_{2d-1} = 0$ we obtain $p_{2d} \in \{0, 1\}$ and there is exactly one design corresponding to $(p_1, \dots, p_{2d})$ [see Skibinsky (1986)]. □

It is worthwhile to mention that in contrast to the ordinary polynomial case the mapping from the set of priors onto the set of $\Phi_{0,\pi}$-optimal discriminating designs for the class $\mathscr{F}_{2d}$ is not one-to-one. On the one hand, Lemma 3.1 only specifies the first $2d$ canonical moments of the projection ξ_{σ^*} of a $\Phi_{0,\pi}$-optimal discriminating design σ^*. Thus every design with these canonical moments is $\Phi_{0,\pi}$-optimal discriminating for the class $\mathscr{F}_{2d}$ with respect to the prior π. Uniqueness only occurs in the cases $p_{2d} = 1$ or $p_{2d} = 0$, which are equivalent to $\pi_{2d-1} = 0$ or $\pi_{2d} = 0$, respectively. On the other hand, there are infinitely many priors corresponding to a given set of first d canonical moments $(p_2, \dots, p_{2d}) \in (0, 1)^{d-1} \times [0, 1]$ of even order. This is demonstrated by the following result, which provides a partial converse of Lemma 3.1. Roughly speaking, it shows that every design is in fact $\Phi_{0,\hat{\pi}}$-optimal with respect to an appropriately defined prior $\hat{\pi}$ for the class $\mathscr{F}_{2d}$.

THEOREM 3.2. *Let* $\sigma \in \Sigma$ *denote a design such that its projection* ξ_σ *via* (2.6) *has at least canonical moments of order* $2d$ *and satisfies* (3.2). *If* $d \geq 2$, *there exists an uncountable number of priors* $\hat{\pi}$ *such that* σ *is a* $\Phi_{0,\hat{\pi}}$*-optimal discriminating design for the class* $\mathscr{F}_{2d}$ *with respect to the prior* $\hat{\pi}$. *If* $d = 1$, *there exists exactly one prior* $\hat{\pi}$ *such that* σ *is a* $\Phi_{0,\hat{\pi}}$*-optimal discriminating design for the class* $\mathscr{F}_2$ *with respect to the prior* $\hat{\pi}$. *Moreover, if* $p_{2d} > 0$, *all* *such priors* $\hat{\pi} = (\hat{\pi}_1, \dots, \hat{\pi}_{2d})$ *are characterized by*

$$\hat{\pi}_{2l-1} = \frac{q_{2l}}{p_{2l}} \hat{\pi}_{2l} + \frac{1 - 2p_{2l}}{p_{2l}} \sum_{i=l+1}^{d} \left(\hat{\pi}_{2i} + \hat{\pi}_{2i-1}\right) \tag{3.6}$$

for $l = 1, \dots, d$, *where* $p_2, \dots, p_{2d}$ *denote the canonical moments of even order of* ξ_σ. *If* $p_{2d} = 0$, *all such priors satisfy* (3.6) *for* $l = 1, \dots, d-1$ *and additionally* $\hat{\pi}_{2d} = 0$.

PROOF. By Lemma 3.1, every prior $\hat{\pi}$ such that σ is a $\Phi_{0,\hat{\pi}}$-optimal discriminating design for the class $\mathscr{F}_{2d}$ with respect to the prior $\hat{\pi}$ satisfies the equations in (3.5). If $p_{2d} > 0$, this is equivalent to (3.6) for $l = 1, \dots, d$. If $p_{2d} = 0$, this is equivalent to (3.6) for $l = 1, \dots, d-1$ and additionally $\hat{\pi}_{2d} = 0$. It remains to show that there exists an uncountable number of nonnegative solutions of (3.5) if $d \geq 2$ and a unique solution if $d = 1$. To see this, consider the case $p_{2d} > 0$, put $\alpha_d = 1$, define $\beta_d = (q_{2d}/p_{2d})\alpha_d$ [see (3.6)] and successively for $l = d-1, \dots, 1$,

$$\alpha_l \geq \max\left\{0, \frac{2p_{2l} - 1}{q_{2l}} \sum_{i=l+1}^{d} (\alpha_i + \beta_i)\right\}$$

arbitrarily, and

$$\beta_l = \frac{q_{2l}}{p_{2l}} \alpha_l + \frac{1 - 2p_{2l}}{p_{2l}} \sum_{i=l+1}^{d} (\alpha_i + \beta_i).$$

By construction $\alpha_1, \dots, \alpha_d, \beta_1, \dots, \beta_d$ are nonnegative numbers satisfying (3.6) or equivalently (3.5). Consequently, by Lemma 3.1, the design σ is $\Phi_{p,\hat{\pi}}$-optimal with respect to the prior $\hat{\pi} = (\hat{\pi}_1, \dots, \hat{\pi}_{2d})$ where

$$\hat{\pi}_{2l-1} = \frac{\beta_l}{\sum_{i=1}^{d}(\alpha_i + \beta_i)}, \qquad \hat{\pi}_{2l} = \frac{\alpha_l}{\sum_{i=1}^{d}(\alpha_i + \beta_i)}, \qquad l = 1, \dots, d.$$

The assertion regarding $d = 1$ is obvious from this discussion. The case $p_{2d} = 0$ follows by a similar argument, starting with $\beta_d = 1$, $\alpha_d = 0$, and is omitted for the sake of brevity. □

REMARK 3.3. It is well known [see, e.g., Pukelshem (1993), Section 9.16] that every design σ_D with equal masses at at least $2d + 1$ points is ϕ_p-optimal for estimating the coefficients in the Fourier regression model g_{2d}, where ϕ_p denotes the ϕ_p-criterion of Kiefer (1974), $-\infty \leq p \leq 1$. Lau and Studden (1985) showed that the corresponding projection ξ_{σ_D} via (2.6) has canonical moments $p_j = \frac{1}{2}$ $(j = 1, \dots, 2d)$. Therefore, by Theorem 3.2, every ϕ_p-optimal design σ_D is also a $\Phi_{0,\hat{\pi}}$-optimal discriminating design with respect to any prior $\hat{\pi} = (\hat{\pi}_1, \dots, \hat{\pi}_{2d})$ satisfying $\hat{\pi}_{2l-1} = \hat{\pi}_{2l}$ $(l = 1, \dots, d)$. This observation is particularly important from a practical point of view because it provides a strong argument for the use of a uniform design in a Fourier regression. These designs are not only useful for parameter estimation but also efficient for model discrimination in the class $\mathscr{F}_{2d}$.

EXAMPLE 3.4. Consider the prior $\pi(2/3) = (1/(3d), 2/(3d), 1/(3d), \dots, 2/(3d))$ which puts double weight on the models $g_2, \dots, g_{2d}$. By Lemma 3.1, the $\Phi_{0,\pi}$-optimal discriminating designs σ^* are characterized by the first $2d$ canonical moments of its projection ξ_{σ^*},

$$p_{2j-1}^* = \frac{1}{2}, \qquad p_{2j}^* = \frac{2 + 3(d-j)}{3 + 6(d-j)}, \qquad j = 1, \dots, d$$

and there are infinitely many measures with these first $2d$ canonical moments [see Skibinsky (1986)]. For illustration, consider the case $d = 2$ which gives

$$p_1^* = p_3^* = \tfrac{1}{2}, \qquad p_2^* = \tfrac{5}{9}, \qquad p_4^* = \tfrac{2}{3}$$

and every design ξ on $[-1, 1]$ with these first four canonical moments corresponds to a $\Phi_{0,\pi}$-optimal discriminating design for the class $\mathscr{F}_4$ with respect to the prior $\pi(2/3)$ using the transformation (2.6). For example, if we terminate $(p_1^*, p_2^*, p_3^*, p_4^*)$ with $p_5^* = 0$ we obtain a design ξ_σ on $[-1, 1]$ which has masses $1/4, 3/8, 3/8$ at the points -1, -0.211, 0.878 [see Lau (1988)]. Transforming this design back onto $[-\pi, \pi]$ via (2.6) gives, for a $\Phi_{0,\pi}$-optimal discriminating design for the class $\mathscr{F}_4$ with respect to the prior $\pi(2/3) = (1/6, 1/3, 1/6, 1/3)$, the measure σ^- with masses $1/8, 3/16, 3/16, 3/16, 3/16, 1/8$ at the points $-\pi_{11}$, -1.783, -0.499, 0.499, 1.783 and π. If the sequence is terminated at $p_5^* = 1$, we obtain by a similar analysis the measure σ^+ with masses $3/16, 3/16, 1/4, 3/16, 3/16$ at the points -2.642, -1.358, 0, 1.358 and 2.642 as a further $\Phi_{0,\pi}$-optimal discriminating design with respect to the prior $\pi(2/3) = (1/6, 1/3, 1/6, 1/3)$.

The preceding example represents an interesting particular case, where the weights for the models g_{2l-1} are all equal and the weights for the models g_{2l} are all equal. It turns out that in this case a $\Phi_{0,\pi}$-optimal discriminating design can be found explicitly, which is "nearly" uniformly distributed on not necessarily equidistant points. Throughout this paper $(a)_n$ denotes the Pochhammer symbols, that is, $(a)_n = a(a+1)\cdots(1+n-1)$, $(a)_0 = 1$. The proof of this result is complicated and therefore deferred to the Appendix.

THEOREM 3.5. *Let $\pi(a) = (\pi_1, \dots, \pi_{2d})$ denote a prior such that $\pi_{2j} = a/d$, $\pi_{2j-1} = (1-a)/d$ $(j = 1, \dots, d)$ for some $a \in (0, 1]$. Let σ^- denote the design with equal masses $1/(2(d+a))$ at the $2d$ zeros of the polynomial*

$$P_d^-(\theta) = \sum_{j=0}^{d} \binom{d}{j} \Gamma(d+1+j)(j+1+a)_{d-j}(-1)^j \left(\cos\left(\frac{\theta}{2}\right)\right)^{2j}$$

in the interval $(-\pi, \pi)$ and masses $a/(2(d+a))$ at the points $-\pi$ and π; then σ^- is a $\Phi_{0,\pi(a)}$-optimal discriminating design with respect to the prior $\pi(a)$ supported at $2d+2$ points. Let σ^+ denote the design with equal *masses* $1/(2(d+a))$ *at the* $2d$ *zeros of the polynomial*

$$P_d^+(\theta) = \sum_{j=0}^{d} \binom{d}{j} \Gamma(d+1+j)(j+1-a)_{d-j}(-1)^j \left(\cos\left(\frac{\theta}{2}\right)\right)^{2j}$$

in the interval $(-\pi, \pi)$ *and mass* $a/(d+a)$ *at the point* 0; *then* σ^+ *is a* $\Phi_{0,\pi(a)}$-*optimal discriminating design with respect to the prior* $\pi(a)$ *supported at* $2d+1$ *points.*

THEOREM 3.6. *Let* $p \in (-\infty, 1)$, $\sigma \in \Sigma$ *denote a symmetric design on* $[-\pi, \pi]$ *and* ξ_σ *its projection onto* $\Sigma_{[-1,1]}$ *via* (2.6). *The design* σ *is a* $\Phi_{p,\pi}$-*optimal discriminating design for the class of Fourier regression models* $\mathscr{F}_{2d}$ *with respect to the prior* π *if and only if the canonical moments of its projection* ξ_σ *satisfy*

$$p_{2l-1} = \frac{1}{2} \quad (l = 1, \dots, d), \qquad p_{2d} = \frac{\pi_{2d}^{1/(1-p)}}{\pi_{2d-1}^{1/(1-p)} + \pi_{2d}^{1/(1-p)}} \tag{3.7}$$

and for $l = 1, \dots, d-1$,

$$\begin{aligned} &\pi_{2l-1} p_{2l}^{1-p} \\ &\quad + (2p_{2l} - 1) \sum_{i=l+1}^{d} \left\{ [\pi_{2i} p_{2i}^p + \pi_{2i-1} q_{2i}^p] \left(4^{i-l} \prod_{j=l+1}^{i-1} q_{2j} p_{2j}\right)^p \right\} \\ &\quad = \pi_{2l} q_{2l}^{1-p}. \end{aligned} \tag{3.8}$$

Moreover, there exists a unique $\Phi_{p,\pi}$-*optimal discriminating design for the class* $\mathscr{F}_{2d}$ *with respect to the prior* π *if and only if* $\pi_{2d} = 0$ *or* $\pi_{2d-1} = 0$.

PROOF. Note that in the case $p = 0$, the equations in (3.7) and (3.8) reduce to (3.5) and it remains to consider the case $p \neq 0$. It was already pointed out at the beginning of this section that for a $\Phi_{p,\pi}$-optimal discriminating design for the class $\mathscr{F}_{2d}$, the canonical moments of odd order less than or equal to $2d$ of the corresponding projection must be $1/2$. The representation of p_{2d} in (3.7) follows by straightforward algebra differentiating (3.4) with respect to p_{2d}.

If σ is a $\Phi_{p,\pi}$-optimal discriminating design for the class $\mathscr{F}_{2d}$ with respect to the prior $\pi = (\pi_1, \dots, \pi_{2d})$, then Lemma 2.1 shows that σ is also $\Phi_{0,\hat{\pi}}$-optimal with respect to the prior $\hat{\pi} = (\hat{\pi}_1, \dots, \hat{\pi}_{2d})$ specified by (2.5). Observing (3.3) we obtain for this prior,

$$\hat{\pi}_l = \begin{cases} s\pi_{2k}\left(2^{2k-2} p_{2k} \prod_{j=1}^{k-1} q_{2j} p_{2j}\right)^p, & \text{if } l = 2k, \\ s\pi_{2k-1}\left(2^{2k-2} q_{2k} \prod_{j=1}^{k-1} q_{2j} p_{2j}\right)^p, & \text{if } l = 2k-1, \end{cases} \tag{3.9}$$

$$l = 1, \dots, d,$$

where s is a normalizing constant such that $\sum_{l=1}^{2d} \hat{\pi}_l = 1$. Lemma 3.1 shows that the first d even canonical moments of the projection ξ_σ of a $\Phi_{0,\hat{\pi}}$-optimal discriminating design with respect to the prior $\hat{\pi}$ are given by

$$p_{2l} = \frac{\hat{\pi}_{2l} + \sum_{i=l+1}^{d}(\hat{\pi}_{2i-1} + \hat{\pi}_{2i})}{\hat{\pi}_{2l} + \hat{\pi}_{2l-1} + 2\sum_{i=l+1}^{d}(\hat{\pi}_{2i-1} + \hat{\pi}_{2i})}, \qquad l = 1, \dots, d-1.$$

Inserting in these equations (3.9) yields the system

$$\begin{aligned} p_{2l} &= \frac{\pi_{2l} p_{2l}^{p} + \sum_{i=l+1}^{d}\left[(\pi_{2i} p_{2i}^{p} + \pi_{2i-1} q_{2i}^{p})\left(4^{i-l}\prod_{j=l}^{i-1} q_{2j} p_{2j}\right)^{p}\right]}{\pi_{2l} p_{2l}^{p} + \pi_{2l-1} q_{2l}^{p} + 2\sum_{i=l+1}^{d}\left[(\pi_{2i} p_{2i}^{p} + \pi_{2i-1} q_{2i}^{p})\left(4^{i-l}\prod_{j=l}^{i-1} q_{2j} p_{2j}\right)^{p}\right]} \\ &= \frac{\pi_{2l} q_{2l}^{-p} + \sum_{i=l+1}^{d}\left[(\pi_{2i} p_{2i}^{p} + \pi_{2i-1} q_{2i}^{p})\left(4^{i-l}\prod_{j=l+1}^{i-1} q_{2j} p_{2j}\right)^{p}\right]}{\pi_{2l} q_{2l}^{-p} + \pi_{2l-1} p_{2l}^{-p} + 2\sum_{i=l+1}^{d}\left[(\pi_{2i} p_{2i}^{p} + \pi_{2i-1} q_{2i}^{p})\left(4^{i-l}\prod_{j=l+1}^{i-1} q_{2j} p_{2j}\right)^{p}\right]} \end{aligned}$$

($l = 1, \dots, d-1$), which is equivalent to (3.8). Consequently, if $\sigma \in \Sigma$ is a $\Phi_{p,\pi}$-optimal discriminating design with respect to the prior π, then the canonical moments of its projection ξ_σ must satisfy the system of equations in (3.7) and (3.8). On the other hand, it is straightforward to show that these equations have a unique solution $(p_1, \dots, p_{2d}) \in (0,1)^{d-1} \times [0,1]$. By standard arguments of optimum design theory [see Pukelsheim (1993), Section 7.13], a $\Phi_{p,\pi}$-optimal discriminating design $\sigma \in \Sigma$ exists and the assertion of the theorem follows. □

EXAMPLE 3.7. If the prior π^* satisfies $\pi_{2l-1}^* = \pi_{2l}^*$ ($l = 1, \dots, d$), then the unique solution of (3.7) and (3.8) is given by $p_l = 1/2$ ($l = 1, \dots, 2d$). Thus every D-optimal design for the Fourier regression model g_{2d} is also a Φ_{p,π^*}-optimal discriminating design for the class $\mathscr{F}_{2d}$ with respect to the prior π^* [independently of the value p].

As a "nontrivial" example, consider the case $d = 3$ and a prior of the form $\pi(a) = \frac{1}{3}((1-a), a, (1-a), a, (1-a), a)$ where $0 \le a \le 1$. In this case the system of equations in Theorem 3.6 gives $p_1 = p_3 = p_5 = 1/2$,

$$p_6 = \frac{1}{1 + (1/a - 1)^{1/(1-p)}}, \tag{3.10}$$

$$a q_4^{1-p} = (1-a) p_4^{1-p} + (2p_4 - 1)(a p_6^{p} + (1-a) q_6^{p}) 4^{p}, \tag{3.11}$$

$$\begin{aligned} a q_2^{1-p} &= (1-a) p_2^{1-p} + (2p_2 - 1) \\ &\quad \times \left[4^{p}(a p_4^{p} + (1-a) q_4^{p}) \right. \\ &\quad \left. + 4^{2p}(q_4 p_4)^{p}(a p_6^{p} + (1-a) q_6^{p})\right], \end{aligned} \tag{3.12}$$

which can easily be solved by standard software (e.g., Mathematica). If $a = 0$ or $a = 1$, we obtain $p_6 = 0$ or $p_6 = 1$ and there exists exactly one measure on $[-1, 1]$ with these canonical moments. The corresponding measure σ on the interval $[-\pi, \pi]$ obtained via the projection (2.6) is the $\Phi_{p,\pi(a)}$-optimal discriminating design with respect to the prior $\pi(a)$. In the remaining cases

TABLE 1
Positive support points and corresponding weights of "a lower principal representation" σ^- of the $\Phi_{p,\pi(\alpha)}$-optimal discriminating design for the class $\mathscr{F}_6$

p	**π(2/3)**				**π(3/4)**			
0	0.356	1.269	2.175	3.141	0.304	1.230	2.140	3.141
	0.136	0.136	0.136	0.092	0.133	0.133	0.133	0.101
−1	0.391	1.311	2.209	3.141	0.358	1.291	2.188	3.141
	0.142	0.137	0.135	0.086	0.141	0.134	0.131	0.094
−2	0.391	1.324	2.222	3.141	0.380	1.311	2.208	3.141
	0.143	0.138	0.136	0.083	0.144	0.135	0.132	0.089

$0 < a < 1$, we have $p_j \in (0,1)$, $j = 1, \dots, 6$ and there are infinitely many designs on $[-1,1]$ with the first six canonical moments equal to $p_1, \dots, p_6$. This implies the existence of infinitely many $\Phi_{p,\pi(a)}$-optimal discriminating designs for the class $\mathscr{F}_{2d}$ with respect to the prior $\pi(a)$. A solution with a minimal number of support points is obtained by using the measure with canonical moments $(p_1, \dots, p_6, p_7)$ where $p_1, \dots, p_6$ are determined by (3.10)–(3.12) and $p_7 = 0$ or $p_7 = 1$. This corresponds to a lower ($p_7 = 0$) or upper ($p_7 = 1$) principal representation of the point $(p_1, \dots, p_6)$ [see Skibinsky (1986)] and the corresponding measure on the interval $[-1,1]$ has four $(= d+1)$ support points including the point -1 (if $p_7 = 0$) or $+1$ (if $p_7 = 1$). The resulting $\Phi_{p,\pi}$-optimal design $\sigma^{\mp} \in \Sigma$ is obtained via (2.6) and has $8(= 2d+2)$ support points if $p_7(= p_{2d+1}) = 0$ and $7(= 2d+1)$ support points if $p_7(= p_{2d+1}) = 1$. We have calculated both cases for $a = \frac{2}{3}$, $a = \frac{3}{4}$ and $p = 0, -1, -2$. The results are listed in Tables 1 and 2. Table 1 shows positive support points and corresponding weights of the $\Phi_{p,\pi(a)}$-optimal discriminating design $\sigma^- \in \Sigma$ for the class of Fourier regression models $\mathscr{F}_6$ with respect to the prior $\pi(a) = \frac{1}{3}((1-a), a, (1-a), a, (1-a), a)$ for various values of p and a. The sequence of canonical moments of the projection ξ_{σ^-} is terminated with $p_7 = 0$. The $\Phi_{p,\pi(a)}$-optimal discriminating design has eight support points where the negative support points and corresponding masses are obtained by a reflection at the origin. Table 2 shows nonnegative support

TABLE 2
Nonnegative support points and corresponding weights of "an upper principal representation" σ^+ of the $\Phi_{p,\pi(\alpha)}$-optimal discriminating design for the class $\mathscr{F}_6$

p	**π(2/3)**				**π(3/4)**			
0	0.000	0.967	1.872	2.885	0.000	1.002	1.912	2.838
	0.184	0.136	0.136	0.136	0.202	0.133	0.133	0.133
−1	0.000	0.933	1.831	2.751	0.000	0.954	1.851	2.783
	0.172	0.135	0.137	0.142	0.188	0.131	0.134	0.141
−2	0.000	0.920	1.818	2.736	0.000	0.934	1.830	2.762
	0.166	0.136	0.138	0.143	0.178	0.132	0.135	0.144

points and corresponding weights of the $\Phi_{p,\pi(a)}$-optimal discriminating design $\sigma^+ \in \Sigma$ for the class of Fourier regression models $\mathscr{F}_6$ with respect to the prior $\pi(a) = \frac{1}{3}((1-a), a, (1-a), a, (1-a), a)$ for various values of p and a. The sequence of canonical moments of ξ_{σ^+} is terminated with $p_7 = 1$. The $\Phi_{p,\pi(a)}$-optimal discriminating design has seven support points where the negative support points and corresponding masses are obtained by a reflection at the origin.

4. Discriminating designs with respect to the maximin criterion. The determination of a $\Phi_{-\infty,\pi}$-optimal discriminating design turns out to be more complicated than in the differentiable case $-\infty < p < 1$. The previous discussion shows that, if $-\infty < p < 1$, the canonical moments of a $\Phi_{p,\pi}$-optimal discriminating design σ^* (more precisely of its projection ξ_{σ^*}) are unique up to the order $2d$. We will demonstrate now that for the maximin criterion $\Phi_{-\infty,\pi}$, this uniqueness statement is not necessarily correct. Note that the $\Phi_{-\infty,\pi}$-optimality criterion does not depend on the size of the elements in the prior, that is,

$$\Phi_{-\infty,\pi}(\sigma) = \min\{\text{eff}_j(\sigma) | \pi_j > 0\}.$$

We begin our investigations by considering priors of the form $(0, \pi_2, 0, \dots, 0, \pi_{2d})$ with $\pi_{2d} > 0$. In this case, the problem is equivalent to a problem of optimum design in homoscedastic polynomial regression and the solution is more transparent.

THEOREM 4.1. *Let $\pi^e = (0, \pi_2, 0, \dots, 0, \pi_{2d})$ denote a prior for the class of Fourier regression models $\mathscr{F}_{2d}$ with vanishing odd components ($\pi_{2d} > 0$) and define*

$$i_l = \#\{j \in \{l, \dots, d\} | \pi_{2j} > 0\}, \qquad l = 1, \dots, d \tag{4.1}$$

as the number of nonvanishing weights with an index $\geq 2l$.

(a) *A symmetric design $\sigma^* \in \Sigma$ is a $\Phi_{-\infty,\pi}$-optimal discriminating design for the class of Fourier regression models $\mathscr{F}_{2d}$ with respect to the prior π^e if and only if the canonical moments of its projection ξ_{σ^*} satisfy*

$$p^*_{2l} = \begin{cases} \dfrac{i_l + 1}{2 i_l}, & \text{if } \pi_{2l} > 0, \\[2mm] \dfrac{1}{2}, & \text{if } \pi_{2l} = 0, \end{cases} \qquad l = 1, \dots, d, \tag{4.2}$$

$$p^*_{2l-1} = \tfrac{1}{2}, \qquad l = 1, \dots, d. \tag{4.3}$$

Moreover, the $\Phi_{-\infty,\pi}$-optimal discriminating design for the class $\mathscr{F}_{2d}$ with respect to the prior π^e is unique.

(b) *Let $\hat{p}_1, \dots, \hat{p}_{2d_1} \in (0,1)$ be fixed ($1 \leq d_1 < d$) and $\hat{\sigma}$ denote the design whose projection $\xi_{\hat{\sigma}}$ has canonical moments $(\hat{p}_1, \dots, \hat{p}_{2d_1}, p^*_{2d_1+1}, \dots, p^*_{2d})$ where p^*_l is defined by (4.2) and (4.3) ($l = 2d_1 + 1, \dots, 2d$). Here $\hat{\sigma}$ is the* *unique $\Phi_{-\infty,\pi}$-optimal discriminating design for the class of Fourier regression models $\mathscr{F}_{2d}$ with respect to the prior*

$$\pi^{e,d_1} = \left(0,\dots,0,\pi_{2d_1+2},0,\pi_{2d_1+4},0,\dots,0,\pi_{2d}\right)$$

in the set

$$\hat{\Sigma} = \left\{\sigma \in \Sigma \mid p_l(\xi_\sigma) = \hat{p}_l \qquad \text{for } l = 1,\dots,2d_1\right\}$$

of all designs with first $2d_1$ canonical moments equal to $\hat{p}_1,\dots,\hat{p}_{2d_1}$.

Proof. (a) Observing (2.10), (2.15) and the definition of the prior π^e, it follows that the $\Phi_{-\infty,\pi}$-optimal design problem is equivalent to the problem of determining the $\Phi_{-\infty,\pi}$-optimal discriminating design for a homoscedastic polynomial regression on the interval $[-1,1]$. This has been solved by Dette (1995), and we obtain for the canonical moments of the projection ξ_{σ^*},

$$p^*_{2(d-l)} = \begin{cases} 1 - \left(\frac{1}{2}\right)^{2l} \prod_{i=d-l+1}^{d-1} \left(q^*_{2i} p^*_{2i}\right)^{-1}, & \text{if } \pi_{2(d-l)} > 0, \\ \dfrac{1}{2}, & \text{if } \pi_{2(d-l)} = 0, \end{cases} \tag{4.4}$$

$$l = 1,\dots,d-1,$$

$$p^*_{2l-1} = \tfrac{1}{2}, \qquad l = 1,\dots,d; \qquad p^*_{2d} = 1. \tag{4.5}$$

Because $p^*_{2d} = 1$, the $\Phi_{-\infty,\pi}$-optimal discriminating design for the class $\mathscr{F}_{2d}$ with respect to the prior π^e must be unique. In order to show that (4.2) and (4.4) coincide, we use an induction argument noting that only cases with $\pi_{2l} > 0$ are of interest. For $l = d$ we have $i_d = 1$ and both representations give $p^*_{2d} = 1$. Now assume that (4.2) and (4.3) is valid for $l = k,\dots,d$. If $\pi_{2j} = 0$ for all $1 \le j \le k-1$, there is nothing to show. In the other case, let k_1 denote the maximum index in $\{1,\dots,k-1\}$ for which $\pi_{2k_1} > 0$. Similarly, let k_2 be the minimum index in $\{k,\dots,d\}$ for which $\pi_{2k_2} > 0$. Observing the definition (4.1), we have $i_k = i_{k_2} = i_{k_1} - 1$ and $p^*_{2j} = \frac{1}{2}$ for $j = k_1+1,\dots,k_2-1$. Now the recursion (4.4) yields

$$\begin{aligned} p^*_{2k_1} &= 1 - \left(\frac{1}{2}\right)^{2(d-k_1)} \prod_{j=k_1+1}^{d-1} \left(q^*_{2j} p^*_{2j}\right)^{-1} \\ &= 1 - \frac{1}{4} \prod_{l=2}^{i_{k_1}-1} \frac{l^2}{l^2-1} = 1 - \frac{i_{k_1}-1}{2i_{k_1}} = \frac{i_{k_1}+1}{2i_{k_1}}, \end{aligned}$$

where the second equality follows from the fact that there are $i_{k_2} = i_{k_1} - 1$ nonvanishing elements among $\pi_{2k_1+2},\dots,\pi_{2d}$ corresponding to the canonical moments $(l+1)/2l$ $(l = 1,\dots,i_{k_1}-1)$. Repeating this argument proves part (a) of the theorem.

(b) By part (a) it follows that the measure $\xi_\sigma \in \Sigma_{[-1,1]}$ corresponding to the design $\sigma \in \Sigma$ which maximizes

$$\min\left\{\mathrm{eff}_{2l}(\sigma) \mid d_1 + 1 \le l \le d,\ \pi_{2l} > 0\right\} \tag{4.6}$$

has canonical moments $p_{2l-1} = 1/2$ $(l = 1,\dots,d)$, $p_{2l} = 1/2$ $(l = 1,\dots,d_1)$,

$$p_{2l} = \begin{cases} \dfrac{i_l + 1}{2i_l}, & \text{if } \pi_{2l} > 0, \\ \dfrac{1}{2}, & \text{if } \pi_{2l} = 0, \end{cases} \qquad l = d_1 + 1,\dots,d.$$

Because all efficiencies in (4.6) depend on $p_1,\dots,p_{2d_1}$, only through

$$q_{2d_1} \prod_{j=1}^{d_1} (q_{2j-2} p_{2j-1} q_{2j-1} p_{2j})$$

it follows that $\hat{\sigma}$ maximizes (4.6) within the class $\hat{\Sigma}$. The statement regarding the uniqueness follows from $p_{2d} = 1$. □

For the general case we need the following lemma. The proof is obvious, observing (2.15) and (3.1).

LEMMA 4.2. *Let $\pi = (\pi_1,\dots,\pi_{2d})$ denote a prior for the class of Fourier regression models $\mathscr{F}_{2d}$, $l_0 \in \{1,\dots,d\}$ and $\hat{\pi}$ be obtained from π by interchanging the $(2l_0 - 1)$th and $2l_0$th component, that is,*

$$\hat{\pi}_k = \begin{cases} \pi_k, & \text{if } k \neq 2l_0 - 1, 2l_0, \\ \pi_{2l_0 - 1}, & \text{if } k = 2l_0, \\ \pi_{2l_0}, & \text{if } k = 2l_0 - 1, \end{cases} \qquad k = 1,\dots,2d. \tag{4.7}$$

If σ^* denotes a $\Phi_{p,\pi}$-optimal discriminating design for the class of Fourier regression models $\mathscr{F}_{2d}$ with respect to the prior π, and $\hat{\sigma}$ denotes a design such that the canonical moments of the corresponding projections ξ_{σ^*} and $\xi_{\hat{\sigma}}$ are related by

$$\hat{p}_l = \begin{cases} p_l^*, & \text{if } l \neq 2l_0, \\ q_l^*, & \text{if } l = 2l_0, \end{cases} \qquad l = 1,\dots,2d,$$

then $\hat{\sigma}$ is a $\Phi_{p,\hat{\pi}}$-optimal design for the class $\mathscr{F}_{2d}$ with respect to the prior $\hat{\pi}$.

THEOREM 4.3. *Let π denote a prior for the class of Fourier regression models $\mathscr{F}_{2d}$,*

$$d_0^\pi := \max\big\{\{0\} \cup \{l \mid \pi_{2l-1}\pi_{2l} > 0\}\big\} \tag{4.8}$$

be the maximum index j for which the prior π assigns positive weight to both models g_{2j-1} and g_{2j} and let

$$i_l^\pi := \#\big\{j \in \{l,\dots,d\} \mid \pi_{2j-1} + \pi_{2j} > 0\big\} \tag{4.9}$$

be the number of pairs of models (g_{2j-1}, g_{2j}) $(j = l,\dots,d)$ with at least one positive corresponding weight.

(a) *If $d_0^\pi = 0$, then there exists a unique $\Phi_{-\infty,\pi}$-optimal discriminating design σ^* for the class of Fourier regression models $\mathscr{F}_{2d}$ and the canonical* *moments of the corresponding projection* $\xi_{\sigma^*} \in \Sigma_{[-1,1]}$ *are given by* $p_{2l-1} = \frac{1}{2}$, $p_{2d} = 1 (= 0)$ *if* $\pi_{2d} > 0$ ($\pi_{2d-1} > 0$) *and*

$$(4.10)\qquad p_{2l} = \begin{cases} 1 - \left(\dfrac{1}{2}\right)^{2(d-l)} \displaystyle\prod_{j=l+1}^{d-1} (q_{2j} p_{2j})^{-1} = \dfrac{i_l^\pi + 1}{2 i_l^\pi}, & \text{if } \pi_{2l} > 0, \\ \left(\dfrac{1}{2}\right)^{2(d-l)} \displaystyle\prod_{j=l+1}^{d-1} (q_{2j} p_{2j})^{-1} = \dfrac{i_l^\pi - 1}{2 i_l^\pi}, & \text{if } \pi_{2l-1} > 0, \\ \dfrac{1}{2}, & \text{if } \pi_{2l-1} = \pi_{2l} = 0, \end{cases}$$

$l = 1, \dots, d-1$.

(b) *If* $d_0^\pi = d$, *then there exists an uncountable number of* $\Phi_{-\infty,\pi}$*-optimal discriminating designs for the class of Fourier regression models* $\mathscr{F}_{2d}$. *The first* $2d$ *canonical moments of the corresponding projections onto* $\Sigma_{[-1,1]}$ *are uniquely determined by*

$$(4.11)\qquad p_l = \tfrac{1}{2} \quad (l = 1, \dots, 2d).$$

(c) *If* $1 \le d_0^\pi \le d-1$, *every projection* ξ_σ *with canonical moments satisfying*

$$(4.12)\qquad \begin{aligned} p_l &= \tfrac{1}{2}, & l &= 1, \dots, 2d_0^\pi, \\ p_{2l-1} &= \tfrac{1}{2}, & l &= d_0^\pi + 1, \dots, d, \end{aligned}$$

$$(4.13)\qquad p_{2l} \prod_{j=1}^{l-1} q_{2j} p_{2j} \ge 2^{1-2l} \quad \text{if } \pi_{2l} > 0, \qquad l = d_0^\pi + 1, \dots, d,$$

$$(4.14)\qquad q_{2l} \prod_{j=1}^{l-1} q_{2j} p_{2j} \ge 2^{1-2l} \quad \text{if } \pi_{2l-1} > 0, \qquad l = d_0^\pi + 1, \dots, d$$

corresponds via (2.6) *to a* $\Phi_{-\infty,\pi}$*-optimal discriminating design* σ *for the class of Fourier regression models* $\mathscr{F}_{2d}$. *A first solution* σ^* *of* (4.12)–(4.14) *is obtained if the canonical moments of the projection* ξ_{σ^*} *are given by* (4.11). *A second solution* $\xi_{\sigma^{**}}$ *is obtained by using* (4.12) *and* (4.10) *for* $l = d_0^\pi + 1, \dots, d$. *This sequence can be characterized by the fact that the corresponding* σ^{**} *is additionally the unique* $\Phi_{-\infty,\pi}$*-optimal discriminating design for the class* $\mathscr{F}_{2d}$ *with respect to the prior* $\hat{\pi} = (0, \dots, 0, \pi_{2d_0^\pi+1}, \pi_{2d_0^\pi+2}, \dots, \pi_{2d-1}, \pi_{2d})$. *In particular, there exists an uncountable number of* $\Phi_{-\infty,\pi}$*-optimal discriminating designs for the class of Fourier regression models* $\mathscr{F}_{2d}$.

Proof. Part (a) is an immediate consequence of Lemma 4.2 and Theorem 4.1 and its proof. In order to prove part (b) we note that for a design satisfying (3.2) the $\Phi_{-\infty,\pi}$ optimality criterion gives

$$(4.15)\qquad \begin{aligned} \Phi_{-\infty,\pi}(\sigma) = \min\Bigg\{ &\left\{ 2^{2k-2} q_{2k} \prod_{l=1}^{k-1} q_{2l} p_{2l} \,\middle|\, \pi_{2k-1} > 0 \right\} \\ &\cup \left\{ 2^{2k-2} p_{2k} \prod_{l=1}^{k-1} q_{2l} p_{2l} \,\middle|\, \pi_{2k} > 0 \right\} \Bigg\} \end{aligned}$$

and the sequence of canonical moments in (4.11) yields the criterion value $\Phi_{-\infty,\pi}(\sigma^*) = 1/2$. On the other hand, we obtain for every $\sigma \in \Sigma$,

$$(4.16)\quad \min\{\mathrm{eff}_{2d_0^\pi-1}(\sigma), \mathrm{eff}_{2d_0^\pi}(\sigma)\} \le \min\{p_{2d_0^\pi}, q_{2d_0^\pi}\} \le \tfrac{1}{2} \quad \text{if } d_0^\pi > 0$$

with equality if and only if the canonical moments of the corresponding projection satisfy the first part of (4.12). This shows that a design σ^* is $\Phi_{-\infty,\pi}$-optimal discriminating for the class $\mathscr{F}_{2d}$ if and only if the first $2d$ canonical moments of its projection ξ_{σ^*} satisfy (4.11), which proves part (b) of the theorem.

The remaining part (c) follows from these arguments and Theorem 4.1(b). From (4.16) we obtain that the first $2d_0^\pi$ canonical moments of the projection ξ_σ of a $\Phi_{-\infty,\pi}$-optimal discriminating design must satisfy (4.12) and that the remaining efficiencies must satisfy

$$(4.17)\qquad \mathrm{eff}_k(\sigma) \ge \tfrac{1}{2} \quad \text{whenever } \pi_k > 0$$

$(k = 2d_0^\pi + 1, \dots, 2d)$. Observing the representation (3.3) of the efficiencies in terms of canonical moments it follows that for a design satisfying (3.2) this is equivalent to (4.13) and (4.14). The design $\sigma^* \in \Sigma$ corresponding to (4.11) obviously satisfies (4.12)–(4.14). The second solution σ^{**} described in part (c) obviously satisfies (4.12) and the canonical moments $p_{2d_0^\pi+2}, \dots, p_{2d}$ of $\xi_{\sigma^{**}}$ are given by (4.10), by definition of σ^{**}. Let $l > d_0^\pi$ and assume that $\pi_{2l} > 0$; then, by (4.10),

$$\begin{aligned} p_{2l}\prod_{j=1}^{l-1} q_{2j}p_{2j} &= \frac{1}{q_{2l}}\prod_{j=1}^{d-1}(q_{2j}p_{2j})\prod_{j=l+1}^{d-1}(q_{2j}p_{2j})^{-1} \\ &= \left(\frac{1}{2}\right)^{2(l-d+d_0^\pi)}\prod_{j=d_0^\pi+1}^{d-1}(q_{2j}p_{2j}) \\ &= \begin{cases} \left(\frac{1}{2}\right)^{2(l-1)} q_{2d_1^\pi}, & \text{if } \pi_{2d_1^\pi-1} > 0, \\ \left(\frac{1}{2}\right)^{2(l-1)} p_{2d_1^\pi}, & \text{if } \pi_{2d_1^\pi} > 0, \end{cases} \end{aligned}$$

where $d_1^\pi = \min\{j \ge d_0^\pi + 1 | \pi_{2j-1} + \pi_{2j} > 0\}$. In both cases we obtain from (4.10),

$$p_{2l}\prod_{j=1}^{l-1}(q_{2j}p_{2j}) \ge 2^{1-2l},$$

which proves (4.13). The corresponding inequality (4.14) for the case $\pi_{2l-1} > 0$ is proved exactly in the same way. The characterization of σ^{**} as the unique $\Phi_{-\infty,\pi}$-optimal discriminating design with respect to the prior $\hat{\pi} = (0, \dots, 0, \pi_{2d_0^\pi+1}, \dots, \pi_{2d})$ now follows from Theorem 4.1(b) and part (a). □

REMARK 4.4. It is worthwhile to mention that a careful inspection of the proof of Theorem 4.3 shows that every projection ξ_σ with canonical moments satisfying

$$p_k = \tfrac{1}{2}, \qquad k = 1, \dots, 2d_0^\pi,$$
$$\mathrm{eff}_k(\sigma) \ge \tfrac{1}{2}, \qquad \pi_k > 0,\ k = 2d_0^\pi + 1, \dots, 2d$$

corresponds via (2.6) to a $\Phi_{-\infty,\pi}$-optimal discriminating design σ for the class $\mathscr{F}_{2d}$ with respect to the prior π. In other words, it is not necessary to require the canonical moments $p_{2d_0^\pi+1}, p_{2d_0^\pi+3}, \dots, p_{2d-1}$ to be 1/2. However this choice yields the largest efficiencies for testing the hypotheses $H_0^{2d}, \dots, H_0^{2d_0^\pi+1}$ in (1.2) with respect to choosing the odd canonical moments $p_{2d_0^\pi+1}, p_{2d_0^\pi+3}, \dots, p_{2d-1}$.

We will conclude this section by considering the minimax criterion in more detail, where the minimum of the efficiencies is not taken over all models of the class $\mathscr{F}_{2d}$. More precisely, we consider the four criteria

$$\min\{\mathrm{eff}_{2k}(\sigma) | k = 1, \dots, d\}, \tag{4.18}$$

$$\min\{\mathrm{eff}_{2k-1}(\sigma) | k = 1, \dots, d\}, \tag{4.19}$$

$$\begin{aligned}\min\Bigg\{&\left\{\mathrm{eff}_{2d-4j}(\sigma) | j = 0, \dots, \left[\frac{d-1}{2}\right]\right\} \\ &\cup \left\{\mathrm{eff}_{2d-3-4j}(\sigma) | j = 0, \dots, \left[\frac{d}{2} - \frac{3}{4}\right]\right\}\Bigg\},\end{aligned} \tag{4.20}$$

$$\begin{aligned}\min\Bigg\{&\left\{\mathrm{eff}_{2d-1-4j}(\sigma) | j = 0, \dots, \left[\frac{d}{2} - \frac{1}{4}\right]\right\} \\ &\cup \left\{\mathrm{eff}_{2d-2-4j}(\sigma) | j = 0, \dots, \left[\frac{d-2}{2}\right]\right\}\Bigg\}.\end{aligned} \tag{4.21}$$

Here (4.18) and (4.19) correspond to the $\Phi_{-\infty,\pi}$-criterion for the "priors," which put exactly positive weights at the models $\{g_{2k} | k = 1, \dots, d\}$ or at the models $\{g_{2k-1} | k = 1, \dots, d\}$, while in (4.20) and (4.21) alternating g_{2k-1} or g_{2k} $(k = 1, \dots, d)$ has positive weight starting either with g_{2d} [see (4.20)] or g_{2d-1} [see (4.21)]. For example, the criterion (4.19) could be used, if the experimenter is sure that the terms of highest periodicity do not contain a cosinus part. Similarly, (4.20) can be used in the construction of optimal designs for discriminating between the models $g_{2d}, g_{2d-3}, g_{2d-4}, g_{2d-7}, \dots$. It turns out that in these cases the optimal discriminating designs are unique, "nearly" uniform on not necessarily equidistant points, given by the zeros of certain trigonometric functions. The proof of the following result is deferred to the Appendix.

THEOREM 4.5. (a) *The $\Phi_{-\infty,\pi}$-optimal discriminating design maximizing* (4.18) *is unique and supported at the points* $-\pi, \pi, 0$ *and at the* $2d-2$

zeros of the function

$$\sum_{m=0}^{d-1} (m+1)(d-m)\cos[(d-2m-1)x] \tag{4.22}$$

in the interval $(-\pi, \pi)$. *The corresponding masses at the zeros of* (4.22) *are all equal to* $1/(2(d+2))$ *while the masses at* $\mp\pi$ *and* 0 *are given by* $3/(4(d+2))$ *and* $3/(2(d+2))$, *respectively.*

(b) *The* $\Phi_{-\infty,\pi}$*-optimal discriminating design maximizing* (4.19) *is unique and has equal masses at the points*

$$\left(\mp \frac{\pi k}{d+1} \middle| k = 1, \dots, d\right).$$

(c) *If* $d = 2k+1$, *the* $\Phi_{-\infty,\pi}$*-optimal design maximizing* (4.20) *is unique and supported at the* $2d+1$ *zeros of the function*

$$\sum_{j=0}^{k} (-1)^j \sin[(2k+1-2j)x] \tag{4.23}$$

in the interval $[-\pi, \pi]$. *The masses at the points minus or plus* π *and* 0 *are* $1/(4d)$ *and* $1/(2d)$, *respectively, while the masses at the remaining zeros of* (4.23) *are all equal to* $1/(2d)$.

If $d = 2k$, *the* $\Phi_{-\infty,\pi}$*-optimal discriminating design maximizing* (4.20) *is unique and supported at the* $2d+1$ *zeros of the function*

$$\sum_{j=0}^{k-1} (-1)^j (k-j)\sin[(2k-2j)x]$$

in the interval $[-\pi, \pi]$. *The masses at the points* 0, $\mp\pi$ *and* $\mp\pi/2$ *are* $1/(2(d+2))$, $1/(4(d+2))$ *and* $3/(2(d+2))$, *respectively, while the masses at all remaining* $2d-4$ *support points are all equal to* $1/(2(d+2))$.

(d) *If* $d = 2k+1$, *the* $\Phi_{-\infty,\pi}$*-optimal discriminating design maximizing* (4.21) *is unique and supported at the* $2d$ *zeros of the function*

$$\sum_{j=0}^{k} (-1)^j (2k+1-2j)\cos[(2k+1-2j)x]$$

in the interval $(-\pi, \pi)$. *The masses at the points minus or plus* $\pi/2$ *are* $3/(2(d+2))$ *while the masses at the remaining* $2d-2$ *support points are all equal to* $1/(2(d+2))$.

If $d = 2k$, *the* $\Phi_{-\infty,\pi}$*-optimal discriminating design maximizing* (4.21) *is unique and has equal masses at the* $2d = 4k$ *points*

$$\left\{\mp \frac{2l-1}{2k+1}\frac{\pi}{2} \middle| l = 1, \dots, k, k+2, \dots, 2k+1\right\}. \tag{4.24}$$

APPENDIX

Proof of Theorem 3.5. Observing Lemma 3.1 and the definition of the prior $\pi(a)$, we obtain from (3.5) for the first $2d$ canonical moments of the projection ξ_σ of a $\Phi_{0,\pi}$-optimal discriminating design σ,

$$p^a_{2j-1} = \frac{1}{2}, \qquad p^a_{2j} = \frac{a + (d-j)}{1 + 2(d-j)}, \qquad j = 1, \dots, d. \tag{A.1}$$

Terminating this sequence with $p_{2d+1} = 0$ yields a projection ξ_{σ^-} with $d+1$ support points [see Skibinsky (1986)]. These can be calculated as the roots of the polynomial $(1+x)Q_d(x)$, where $Q_d(x)$ is the dth monic orthogonal polynomial with respect to $(1+x)\,d\xi^R_\sigma(x)$ and the measure ξ^R_σ corresponds to the "reversed" sequence $(\tilde{p}_1, \dots, \tilde{p}_{2d}, 0)$,

$$\tilde{p}_j = p^a_{2d+1-j} = \begin{cases} \dfrac{a + l - 1}{2l - 1}, & \text{if } j = 2l - 1, \\ \dfrac{1}{2}, & \text{if } j = 2l, \end{cases} \tag{A.2}$$

$(j = 1, \dots, 2d)$ [see Studden (1982b) or Lau (1983)]. By a result of Skibinsky (1969), the measure corresponding to the sequence $(\tilde{p}_d)_{j\in\mathbb{N}}$ is the beta-distribution with density proportional to $(1+x)^{a-1}(1-x)^{-a}$, and $Q_d(x)$ must be proportional to the dth Jacobi polynomial $P_d^{(-a,a)}(x)$, that is,

$$\begin{aligned} Q_d(x) &= 2^d \binom{2d}{d}^{-1} P_d^{(-a,a)}(x) \\ &= \frac{\Gamma(d+1-a)2^d}{(2d)!} \sum_{j=0}^{d} \binom{d}{j} \frac{\Gamma(d+1+j)}{\Gamma(j+1-a)} \left(\frac{x-1}{2}\right)^j \end{aligned} \tag{A.3}$$

[see Van Assche (1987), page 2]. The assertion regarding the support points now follows from (2.6), putting $x = \cos\theta$ and observing that $P_d^{(\alpha,\beta)}(-x) = (-1)^d P_d^{(\beta,\alpha)}(x)$. For the calculation of the weights, we note that the Stieltjes transform of ξ_{σ^-} is given by

$$h(z) = \int_{-1}^{1} \frac{d\xi_{\sigma^-}(x)}{z - x} = \frac{P_d(z)}{(1+z)Q_d(z)}, \tag{A.4}$$

where $P_d(z)$ is the monic polynomial of degree d whose d zeros give the d interior support points of the measure with canonical moments $(q^a_1, \dots, q^a_{2d}, 1)$. This follows from formula (2.6), (2.14), and Lemma 2.1 in Skibinsky (1986) (where the results have to be transformed onto the interval $[-1,1]$, $n = 2d$) and from Lemma 2.9 in Studden (1982b). By Lemma 2.10 in Studden (1982b), the support is equal to the support of the measure $\tilde{\xi}$ corresponding to the sequence $(p^a_{2d}, \dots, p^a_1, 1)$ which is given by the dth orthogonal polynomial with respect to the measure $(1-x)\,d\tilde{\xi}(x)$. Now (A.1) and the previous discussion shows that $p^a_{2d}, \dots, p^a_1$ are the first $2d$ canonical moments of the beta-distribution with parameters $(-a, a-1)$. Consequently we obtain $P_d(x) = \binom{2d}{d}^{-1} 2^d P_d^{(-a+1,a-1)}(x)$ and, by (A.3), the identity in (A.4) reduces to

$$h(z) = \frac{P_d^{(-a+1,a-1)}(z)}{(1+z)P_d^{(-a,a)}(z)}.$$

The weights of ξ_{σ^-} at the corresponding support points can be obtained from

$$\xi_{\sigma^-}(x) = h(z)(z-x)|_{z=x} = \frac{P_d^{(-a+1,a-1)}(x)}{P_d^{(-a,a)}(x) + (1+x)(d/dz)P_d^{(-a,a)}(z)|_{z=x}}$$
$$= \frac{P_d^{(-a+1,a-1)}(x)}{(1-a)P_d^{(-a,a)}(x) + (d+a)P_d^{(-a+1,a-1)}(x)},$$

where the last line follows from

$$(d/dz)P_d^{(\alpha,\beta)}(z) = \tfrac{1}{2}(d+\alpha+\beta+1)P_{d-1}^{(\alpha+1,\beta+1)}(z)$$

and formulas (22.7.16), (22.7.18), (22.7.19) in Abramowitz and Stegun (1964). Now if x_0 is an interior support point, we have $P_d^{(-a,a)}(x_0) = 0$ and obtain $\xi_{\sigma-}(x_0) = 1/(d+a)$. The transformation (2.6) then yields $\sigma^-(x) = 1/(2(d+a))$ for all $2d$ interior support points of the corresponding measure σ^-. The assertion regarding the weight at the points minus or plus π follows from the symmetry of σ^-.

The second part of the theorem can be proved by similar arguments, which are omitted for the sake of brevity. □

Proof of Theorem 4.5. All cases are very similar and we restrict ourselves to part (d) and the case $d = 2k$. Observing the definition of d_0^π in (4.8) we have $d_0^\pi = 0$ and part (a) of Theorem 4.3, $(i_l = d - l + 1)$ shows that the $\Phi_{-\infty,\pi}$-optimal discriminating design σ^* must be unique and the corresponding projection ξ_{σ^*} has canonical moments

$$p_{2l-1}^* = \tfrac{1}{2}, \qquad l = 1,\dots,2k, \tag{A.5}$$

$$p_{4k-2j}^* = \begin{cases} \dfrac{j}{2+2j}, & \text{if } j \text{ is even,} \\[2ex] \dfrac{j+2}{2+2j}, & \text{if } j \text{ is odd} \end{cases} \tag{A.6}$$

$(j = 0,\dots,2k-1)$. By results of Studden (1982b), the support of ξ_{σ^*} is given by the zeros of the polynomial $Q_{2k}(x)$ whose zeros give the support points of the measure corresponding to $(p_{4k-1}^*, p_{4k-2}^*, \dots, p_2^*, p_1^*, 0)$. This polynomial is recursively defined by $Q_0(x) = 1$, $Q_1(x) = x$, $Q_2(x) = x^2 - p_{4k-2}^* = x^2 - 3/4$,

$$\begin{aligned} Q_{l+1}(x) &= xQ_l(x) + p_{4k-2l}^* q_{4k-2l+2}^* Q_{l-1}(x) \\ &= \begin{cases} xQ_l(x) - \dfrac{l-1}{4(l+1)}Q_{l-1}(x), & \text{if } l \text{ is even,} \\[2ex] xQ_l(x) - \dfrac{l+2}{4l}Q_{l-1}(x), & \text{if } l \text{ is odd} \end{cases} \end{aligned} \tag{A.7}$$

$(l = 1,\dots,2k-1)$. For the polynomials of even order we thus obtain the recursion $Q_0(x) = 1$, $Q_2(x) = x^2 - 3/4$,

$$Q_{2l}(x) = \left(x^2 - \tfrac{1}{2}\right)Q_{2l-2}(x) - \tfrac{1}{16}Q_{2l-4}(x), \qquad l = 2,\dots,k. \tag{A.8}$$

Recall the definition of the Chebyshev polynomials of the first kind, $T_j(x) = \cos(j \arccos x)$ ($j = 0, 1, 2, \dots$) and the recursion $T_0(x) = 1$, $T_1(x) = x$,

$$T_{j+1}(x) = 2xT_j(x) - T_{j-1}(x), \qquad j \geq 1$$

[see Rivlin (1990)]. Because the leading coefficient of $T_j(x)$ is 2^{j-1}, it follows by a simple induction that

$$Q_{2k}(x) = \frac{T_{2k+1}(x)}{2^{2k}x} = \frac{\cos[(2k+1)\arccos(x)]}{2^{2k}\cos[\arccos(x)]}, \tag{A.9}$$

which gives for the support of ξ_{σ^*}

$$\operatorname{supp}(\xi_{\sigma^*}) = \left\{\cos\left(\frac{2l-1}{2k+1}\frac{\pi}{2}\right)\middle| l = 1, \dots, k, k+2, \dots, 2k+1\right\}. \tag{A.10}$$

For the derivation of the corresponding weights, we use an alternative representation of $Q_{2k}(x)$, namely,

$$Q_{2k}(x) = \frac{1}{2^{2k-1}}\left[\sum_{j=0}^{k-1}(-1)^j T_{2k-2j}(x) + \frac{(-1)^k}{2}\right], \tag{A.11}$$

which follows by induction, using (A.8), (A.9) and the recursive relation for the Chebyshev polynomials $T_{2l}(x)$ of even order. In order to derive the corresponding polynomial $P_{2k-1}(x)$ in the numerator of the Stieltjes transform of ξ_{σ^*},

$$h(z) = \int_{-1}^{1}\frac{d\xi_{\sigma^*}(x)}{z-x} = \frac{P_{2k-1}(z)}{Q_{2k}(z)}, \tag{A.12}$$

we note that by formula (2.6) and (2.14) in Skibinsky (1986) the polynomial $P_{2k-1}(x)$ is determined by the property that its interior zeros give the support of the measure corresponding to the sequence of canonical moments $(q_1^*, \dots, q_{4k-1}^*, 1)$. This support coincides with the support of the measure corresponding to the sequence $(p_{4k-1}^*, \dots, p_1^*, 1)$ [see Studden (1982b)]. By a result of Studden (1982b), $P_{2k-1}(x)$ can be calculated recursively, by $P_0(x) = 1$, $P_1(x) = x$,

$$\begin{aligned} P_{j+1}(x) &= P_j(x) - p_{4k-2j}^* q_{4k-2j-2}^* P_{j-1}(x) \\ &= \begin{cases} xP_j(x) - \dfrac{j}{4(j+2)}P_{j-1}(x), & \text{if } j \text{ is even}, \\ xP_j(x) - \dfrac{j+3}{4(j+1)}P_{j-1}(x), & \text{if } j \text{ is odd}, \end{cases} \end{aligned} \tag{A.13}$$

which gives for the polynomials of even order $P_0(x) = 1$, $P_2(x) = x^2 - 1/2$ and

$$P_{2l}(x) = \left(x^2 - \tfrac{1}{2}\right)P_{2l-2}(x) - \tfrac{1}{16}P_{2l-4}(x), \qquad l = 2, \dots, k-1.$$

A simple induction now shows that

$$P_{2l}(x) = \frac{1}{2^{2l}} \sum_{j=0}^{l} (-1)^j U_{2l-2j}(x), \qquad l = 1, \dots, k-1, \tag{A.14}$$

where we used the recursive relation for the Chebyshev polynomials of the second kind, $U_0(x) = 1$, $U_1(x) = 2x$, $U_2(x) = 4x^2 - 1$,

$$\begin{aligned} U_{l+1}(x) &= 2xU_l(x) - U_{l-1}(x), \\ U_{2l}(x) &= 2(2x^2 - 1)U_{2l-2}(x) - U_{2l-4}(x) \end{aligned} \tag{A.15}$$

($l \geq 1$). A further induction using (A.13), (A.14) and (A.15) gives

$$P_{2l+1}(x) = \frac{2^{-2l-1}}{l+1} \sum_{j=0}^{l} (-1)^j (l+1-j) U_{2l+1-2j}(x), \qquad l = 0, \dots, k-1. \tag{A.16}$$

Observing (A.11), (A.12), (A.16) and the well-known fact $T_d'(x) = dU_{d-1}(x)$ (which readily follows from the trigonometric representation of the Chebyshev polynomials) yields for the weights of ξ_{σ^*},

$$\begin{aligned} \xi_{\sigma^*}(x) &= h(z)(z-x)|_{z=x} = \frac{P_{2k-1}(x)}{Q'_{2k}(x)} \\ &= \frac{(1/k)\sum_{j=0}^{k-1}(-1)^j(k-j)U_{2k-1-2j}(x)}{\sum_{j=0}^{k-1}(-1)^j(2k-2j)U_{2k-1-2j}(x)} = \frac{1}{2k} = \frac{1}{d} \end{aligned}$$

for all $x \in \operatorname{supp}(\xi_\sigma^*)$. Thus ξ_{σ^*} is the uniform distribution at the zeros of the polynomial in (A.9) and the assertion follows by the transformation (2.6). □

Acknowledgments. This work contains part of the second author's Ph.D. thesis, written at the Ruhr-Universität Bochum, Germany. The authors thank I. Gottschlich for her excellent typing of this manuscipt and two referees for very helpful comments which improved the representation of the results.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A. (1964). *Handbook of Mathematical Functions*. Dover, New York.

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

ANDERSON, T. W. (1994). *The Statistical Analysis of Time Series*. Wiley, New York.

ATKINSON, A. C. (1972). Planning experiments to detect inadequate regression models. *Biometrika* **59** 275–293.

ATKINSON, A. C. and COX, D. R. (1974). Planning experiments for discriminating between models (with discussion). *J. Roy. Statist. Soc. Ser. B* **36** 321–348.

ATKINSON, A. C. and DONEV, A. N. (1992). *Optimum Experimental Designs*. Oxford Univ. Press.

DETTE, H. (1994). Discrimination designs for polynomial regression on a compact interval. *Ann. Statist.* **22** 890–904.

DETTE, H. (1995). Optimal designs for identifying the degree of a polynomial regression. *Ann. Statist.* **23** 1248–1267.

DETTE, H. and STUDDEN, W. J. (1997). *The Theory of Canonical Moments with Applications in Statistics, Probability and Analysis*. Wiley, New York.

FEDOROV, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.

HILL, P. D. H. (1978). A note on the equivalence of D-optimal design measures for three rival linear models. *Biometrika* **65** 666–667.

HOEL, P. G. (1965). Minimax designs in two-dimensional regression. *Ann. Math. Statist.* **36** 1097–1106.

KARLIN, S. and STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. Interscience, New York.

KIEFER, J. C. (1974). General equivalence theory for optimum designs. *Ann. Statist.* **2** 849–879.

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

LAU, T. S. (1988). D-optimal designs on the unit q-ball. *J. Statist. Plann. Inference* **19** 299–315.

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

MARDIA, K. (1972). *The Statistics of Directional Data*. Academic Press, New York.

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

RICCOMAGNO, E., SCHWABE, R., and WYNN, H. P. (1997). Lattice-based D-optimum design for Fourier regression. *Ann. Statist.* **25** 2313–2327

RIVLIN, T. J. (1990). *Chebyshev Polynomials*. Wiley, New York.

SKIBINSKY, M. (1967). The range of the $(n+1)$th moment for distributions on $[0,1]$. *J. Appl. Probab.* **4** 543–552.

SKIBINSKY, M. (1969). Some striking properties of binomial and beta moments. *Ann. Math. Stast.* **40** 1753–1764.

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

SPRUILL, M. G. (1990). Good designs for testing the degree of a polynomial mean. *Sankhyā Ser. B* **52** 67–74.

STUDDEN, W. J. (1968). Optimal designs on Chebshev points. *Ann. Math. Statist.* **39** 1435–1447.

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

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

STUDDEN, W. J. (1982b). Optimal designs for weighted polynomial regression using canonical moments. In *Statistical Decision Theory and Related Topics 3* (S. S. Gupta and J. O. Berger, eds.) 335–350. Academic, New York.

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

VAN ASSCHE, W. (1987). *Asymptotics for Orthogonal Polynomials*. *Lecture Notes in Math*. **1265**. Springer, New York.

RUHR-UNIVERSITÄT BOCHUM
FAKULTÄT FÜR MATHEMATIK
44780 BOCHUM
GERMANY
E-MAIL: holger.dette@ruhr-uni-bochum.de