Ann Inst Stat Math (2009) 61:143–157
DOI 10.1007/s10463-007-0133-5

Constrained optimal discrimination designs for Fourier regression models

Stefanie Biedermann · Holger Dette · Philipp Hoffmann

Received: 26 June 2006 / Revised: 9 March 2007 / Published online: 7 July 2007

Abstract In this article, the problem of constructing efficient discrimination designs in a Fourier regression model is considered. We propose designs which maximize the power of the F-test, which discriminates between the two highest order models, subject to the constraints that the tests that discriminate between lower order models have at least some given relative power. A complete solution is presented in terms of the canonical moments of the optimal designs, and for the special case of equal constraints even more specific formulae are available.

Keywords Constrained optimal designs · Trigonometric regression · D_1-optimal designs · Chebyshev polynomials · Canonical moments

1 Introduction

The Fourier regression or trigonometric regression model

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

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

S. Biedermann
University of Southampton, School of Mathematics, Highfield SO17 1BJ, UK

H. Dette (✉) · P. Hoffmann
Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany
e-mail: holger.dette@ruhr-uni-bochum.de; holger.dette@rub.de

where $d \in \mathbb{N}_0$, is used to describe periodic phenomena (see, e.g., Mardia (1972), Kitsos et al. (1988) or the collection of research papers in biology edited by Lestrel (1997)). Moreover, there are several applications of trigonometric regression models in two-dimensional shape analysis in biology. We refer to Younker and Ehrlich (1977) and Currie et al. (2000) for concrete examples. The value $2d$ in (1) or $2d - 1$ in (2) is usually denoted as the degree of the Fourier regression model. The coefficients $a_0, a_1, \ldots, a_d, b_1, \ldots, b_d$ denote unknown parameters, which have to be estimated from the data. The problem of designing experiments for models of the form (1) has been discussed by several authors; see, e.g., Karlin and Studden (1966), p. 347, Fedorov (1972), p. 94, Hill (1978), Lau and Studden (1985) for optimal designs on the full circle, as well as Dette et al. (2002a) and Dette et al. (2002b) for optimal designs on a partial circle. Most authors concentrate on the problem of determining optimal designs for the estimation of the full vector of unknown parameters, whereas the problem of constructing optimal designs for model discrimination has been considered by Dette and Haller (1998), Dette and Melas (2003) and Zen and Tsai (2004). The present paper is devoted to the problem of constructing optimal discrimination designs using constrained optimality criteria.

Constrained optimal designs have primarily been considered by Stigler (1971), Studden (1982b) and Lee (1988a,b), whereas Cook and Wong (1994), Dette (1995) and Clyde and Chaloner (1996) investigated the relation between this approach and compound optimality criteria.

Although these results are interesting from a theoretical point of view constrained optimal designs still had to be found numerically and explicit results could only be inferred in rare cases. In particular, it turns out that there is a one-to-one correspondence between compound and constrained optimal designs, which, however, can only be exploited in rare cases to find the constrained optimal design from the corresponding compound optimal design, which is usually much simpler to calculate. Dette and Franke (2000) characterized constrained optimal discriminating designs for polynomial regression models utilizing the theory of canonical moments, which was introduced by Skibinsky (1967) and applied by Studden (1980, 1982a,b, 1989) for determining optimal designs in polynomial regression models. The problem of finding constrained optimal discriminating designs for Fourier regression models, however, has not been considered yet.

The present paper is devoted to this problem. For the construction of constrained optimal designs we assume that the highest frequency of the model has been fixed and determine the design such that the coefficient corresponding to this frequency is estimated with maximal efficiency subject to the constraints that the coefficients corresponding to the highest frequencies in the models of lower degree can be estimated with some guaranteed efficiency. The optimality criterion is carefully described in Sect. 2. In Sect. 3 we briefly review some facts from the theory of canonical moments (see Dette and Studden 1997), which is the basic tool for the construction of optimal discrimination designs. A complete characterization of the constrained optimal discriminating designs is given in terms of their canonical moments, and in the special case of equal bounds we further specify the optimal designs in terms of their supporting polynomials and explicit formulae for the weights. Finally, in Sect. 4 we give some concluding remarks discussing our approach.

2 Constrained optimal designs in Fourier regression models

We assume that we can make $n \geq 2d+1$ independent observations $Y_1, \ldots, Y_n$ where $Y_i \sim \mathcal{N}(g(x_i), \sigma^2)$, and the regression function $g(x)$ belongs to the class of trigonometric models $\{g_0, g_1, \ldots, g_{2d}\}$ where g_{2l} and g_{2l-1} are defined in (1) and (2), respectively. For $k = 0, \ldots, 2d$ we define

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

and

$$\theta_k = \begin{cases} (a_0, a_1, b_1, \ldots, a_j, b_j)^T & \text{if } k = 2j \\ (a_0, a_1, b_1, \ldots, a_{j-1}, b_{j-1}, a_j)^T & \text{if } k = 2j-1. \end{cases}$$

Then the models $g_k(x)$ where $k = 0, \ldots, 2d$, can be written as $g_k(x) = f_k(x)^T \theta_k$.

An approximate design is a probability measure σ with finite support on the interval $[-\pi, \pi]$ with the interpretation that observations are taken at the support points in proportion to the corresponding masses. The information matrix in the Fourier regression model $g_k(x)$ is given by

$$M_k(\sigma) = \int_{-\pi}^{\pi} f_k(x) f_k^T(x) \mathrm{d}\sigma(x). \tag{3}$$

An optimal design maximizes an appropriate information function of the information matrix (see Pukelsheim 1993, p. 131). There are numerous criteria which can be used for the characterization of efficient designs. Most of these criteria focus on precise parameter estimation in a model of given degree. In many practical situations, however, it is not known before the experiment, up to which degree a Fourier regression model should be fitted. As sparse modeling is advisable we turn our attention to designs that allow successive testing of the higher order coefficients with high power, thus guaranteeing good discrimination properties of the testing procedure. Our optimality criterion for constructing discrimination designs is therefore based on a multiple F-test, where, starting with the given regression model $g_{2d}(x)$ in (1) one tests the hypotheses $H_0^{(2d)} : b_d = 0,\ H_0^{(2d-1)} : a_d = 0,\ H_0^{(2d-2)} : b_{d-1} = 0,\ H_0^{(2d-3)} : a_{d-1} = 0, \ldots, H_0^{(0)} : a_0 = 0$, in the models $g_{2d}, g_{2d-1}, \ldots, g_0$, successively, and decides for the model g_{k_0} where k_0 is the first index for which the hypothesis $H_0^{(k_0)} : \theta_{k_0} = 0$ is rejected. (Note that this sequence of tests can be stopped earlier if the minimal degree of the Fourier regression model is pre-specified.) The statistical properties of this testing procedure are elaborately presented in Anderson (1994). The quantities corresponding to the noncentrality parameter of the F-test for the hypothesis $H_0^{(k)}$ are given by

$$\delta_k(\sigma) = (e_k^T M_k^{-1}(\sigma) e_k)^{-1} \qquad k = 1, \ldots, 2d, \tag{4}$$

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, p. 70). A design σ_k^* is called D_1-optimal for the model g_k or D_1^k-optimal if it maximizes δ_k. This is in fact equivalent to maximizing the power of the corresponding test. Collecting data following a D_1^k-optimal sampling scheme therefore gives the best results for discriminating between models g_k and g_{k-1}. To discriminate between more than two different models, one has to construct an optimality criterion based on several functions $\delta_k(\sigma)$, $k = 2d, 2d-1, \ldots, 0$. As these quantities are of different scalings we standardize them by using the corresponding efficiencies to make them comparable in size. The expression

$$\mathrm{eff}_k(\sigma) := \frac{\delta_k(\sigma)}{\delta_k(\sigma_k^*)}, \quad k = 1, \ldots, 2d \tag{5}$$

is called the D_1^k-efficiency of the design σ in the Fourier regression model $g_k(x)$.

Dette and Haller (1998) proposed to maximize a weighted p-mean of the efficiencies $\mathrm{eff}_1, \ldots, \mathrm{eff}_{2d}$ for the construction of an optimal design for discriminating between the models $\{g_1, \ldots g_{2d}\}$. In the present paper, we consider an alternative optimality criterion to obtain efficient discriminating designs. This approach is attractive if the main interest of the experimenter is in discriminating between the two models of highest degree, while at the same time the optimal design should allow for an efficient discrimination between the models of lower degrees.

We consider two criteria for determining a constrained optimal discriminating design for the Fourier regression model. The first approach considers the highest cosine frequency as most important and a constrained optimal discriminating design σ^* maximizes

$$\mathrm{eff}_{2d}(\sigma) \quad \text{subject to } \ \mathrm{eff}_k(\sigma) \geq c_k, \quad k = 2d-1, 2d-2, \ldots, 2d-2j-1 \tag{6}$$

for some $j \in \{0, \ldots, d-1\}$. The second criterion, however, determines the design which maximizes

$$\mathrm{eff}_{2d-1}(\sigma) \text{ subject to the constraints } \ \mathrm{eff}_k(\sigma) \geq c_k, \quad k = 2d, 2d-2, \ldots, 2d-2j-1, \tag{7}$$

and $j \in \{1, \ldots, d-1\}$. For both criteria, the quantities $c_{2d-2j-1}, \ldots, c_{2d} \in (0, 1)$ are given by the experimenter reflecting the desired minimal relative power of testing $H_0^{(k)}$, $k = 2d-2j-1, \ldots, 2d$. A necessary condition for the existence of optimal designs is $c_{2d-2j} + c_{2d-2j-1} \leq 1$. Unlike criterion (6), criterion (7) corresponds to the situation where the highest sine frequency is regarded as most important, i.e. the power of testing $H_0^{(2d-1)}$ is maximized whereas the preceding test of $H_0^{(2d)}$ (and all the other hypotheses $H_0^{(k)}$) have some prespecified minimal relative power. The situation where the experimenter prefers to start with model (2) for some practical reason, and therefore start the testing procedure with $H_0^{(2d-1)}$, can be incorporated in criterion (7) by putting $c_{2d} = 0$. For the solution of the constrained optimization problems (6) and (7) we need several tools, which will be explained in what follows.

It follows by standard arguments (see Pukelsheim 1993, Chap. 4, 5) that eff_k 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, if there exists a constrained optimal discriminating design, then there also exists an optimal design in the set Σ of all symmetric designs on the interval $[-\pi, \pi]$. We note that these symmetric designs induce designs ξ_σ on the interval $[-1, 1]$ by the projection

$$\xi_\sigma(\cos x) = \begin{cases} 2\sigma(x) = 2\sigma(-x) & \text{if } \ 0 < x \le \pi \\ \sigma(0) & \text{if } \ x = 0 \end{cases} \tag{8}$$

for any symmetric design $\sigma \in \Sigma$. The corresponding set of the measures ξ_σ on $[-1, 1]$ will be denoted by $\Sigma_{[-1,1]}$. It was shown in Dette and Haller (1998) that for any $\sigma \in \Sigma$

$$\delta_k(\sigma) = \begin{cases} 2^{2(j-1)} \frac{|A_j(\xi_\sigma)|}{|A_{j-1}(\xi_\sigma)|} & \text{if } \ k = 2j \\[2ex] 2^{2(j-1)} \frac{|B_j(\xi_\sigma)|}{|B_{j-1}(\xi_\sigma)|} & \text{if } \ k = 2j-1 \end{cases} \tag{9}$$

where $B_0(\xi_\sigma) = A_0(\xi_\sigma) = 1$, and

$$A_k(\xi_\sigma) = \left(\int_{-1}^{1} z^{i+j} \mathrm{d}\xi_\sigma(z) \right)_{i,j=0}^{k} \tag{10}$$

and

$$B_k(\xi_\sigma) = \left(\int_{-1}^{1} (1 - z^2) z^{i+j} \mathrm{d}\xi_\sigma(z) \right)_{i,j=0}^{k-1} \tag{11}$$

denote the information matrices of the design ξ_σ on the interval $[-1, 1]$ for a homoscedastic and a heteroscedastic polynomial regression model with efficiency function $\lambda(z) = (1 - z^2)$ (see Karlin and Studden 1966), respectively. Consequently, the problem of determining constrained optimal discriminating designs for the Fourier regression model 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 (8).

The problem of maximizing the right hand side of (9) over the set $\Sigma_{[-1,1]}$ if $k = 2j$ is in fact the D_1-optimal design for the ordinary polynomial regression model, while for odd values of k the right hand side of (9) corresponds to the weighted polynomial regression with efficiency function $\sigma^2(x) = \sigma^2/(1 - x^2), x \in (-1, 1)$. The solutions of these problems are well known (see Studden 1968, 1982b and yield $\delta_k(\sigma_k^*) = \max_\sigma \delta_k(\sigma) = 1$ $(k = 1, \ldots, 2d)$, and therefore the efficiency of a symmetric design σ defined in (5) can be rewritten as

$$\mathrm{eff}_k(\sigma) = \begin{cases} 2^{2(j-1)} \frac{|A_j(\xi_\sigma)|}{|A_{j-1}(\xi_\sigma)|} & \text{if } k = 2j \\[2ex] 2^{2(j-1)} \frac{|B_j(\xi_\sigma)|}{|B_{j-1}(\xi_\sigma)|} & \text{if } k = 2j-1. \end{cases} \tag{12}$$

3 The solution of the constrained optimal design problem

For the characterization of the measure $\xi_{\sigma^*} \in \Sigma_{[-1,1]}$ corresponding to the constrained optimal discriminating design σ^* by the relation (8) we require some basic facts about the theory of canonical moments which has been introduced by Studden (1980, 1982a,b) in the context of optimal design. We will only give a very brief heuristical introduction and refer to the monograph of Dette and Studden (1997) for more details.

It is well known that a probability measure on the interval $[-1, 1]$, say ξ, is determined by its sequence of moments $(m_1, m_2, \ldots)$. Skibinsky (1967) defined a one to one mapping from the sequences of ordinary moments onto sequences $(p_1, p_2, \ldots)$ 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. 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 Dette and Studden 1997, Chap. 3). The set of probability measures on the interval $[-1, 1]$ with first k canonical moments equal to $(p_1, \ldots, p_k) \in (0, 1)^{k-1} \times [0, 1]$ is a singleton if and only if $p_k \in \{0, 1\}$. Otherwise there exists an uncountable number of probability measures corresponding to $(p_1, \ldots, p_k)$ (see Skibinsky 1986).

It turns out that the determinants in (12) can be described in terms of the canonical moments $p_1, p_2, \ldots$ of the measure ξ_σ (see Studden 1982b), that is

$$|A_k(\xi_\sigma)| = 2^{k(k+1)} \prod_{\ell=1}^{k} (q_{2\ell-2}p_{2\ell-1}q_{2\ell-1}p_{2\ell})^{k-\ell+1} \tag{13}$$

$$|B_k(\xi_\sigma)| = 2^{k(k+1)} \prod_{\ell=1}^{k} (p_{2\ell-2}q_{2\ell-1}p_{2\ell-1}q_{2\ell})^{k-\ell+1} \tag{14}$$

where $p_0 = 1,\ q_0 = 1$ and $q_j = 1 - p_j$ for $j \geq 1$. Observing (12), (13) and (14), we find that the efficiencies are increasing functions of the terms $p_{2l-1}q_{2l-1}$, and consequently the odd canonical moments of the optimal projection design ξ_{σ^*} satisfy

$$p_{2\ell-1} = \frac{1}{2} \quad \ell = 1, \ldots, d. \tag{15}$$

Therefore we can restrict ourselves to designs with this property, and (12) reduces to

$$\mathrm{eff}_k(\sigma) = \begin{cases} 2^{2j-2} p_{2j} \prod_{\ell=1}^{j-1} q_{2\ell} p_{2\ell} & \text{if } k = 2j \\ 2^{2j-2} q_{2j} \prod_{\ell=1}^{j-1} q_{2\ell} p_{2\ell} & \text{if } k = 2j-1 \end{cases} \tag{16}$$

where $p_2, p_4, \ldots$ denote the canonical moments of even order of the design $\xi_\sigma \in \sum_{[-1,1]}$ satisfying (15) and corresponding to the measure σ via (8). Our main result gives a characterization of the canonical moments of ξ_{σ^*}.

Theorem 1 *(a) If there exists a constrained optimal discriminating design for the vector* $(c_{2d-2j-1}, \dots, c_{2d-1})$ *in* (6)*, then there also exists a symmetric optimal discriminating design* σ^**. The canonical moments up to the order* $2d$ *of the corresponding projection* ξ_{σ^*} *are determined by the system of equations*

$$
\begin{aligned}
p_{2n-1} &= \frac{1}{2} \quad n = 1, \dots, d \\
p_{2n} &= \frac{1}{2} \quad n = 1, \dots, d-j-1 \\
p_{2d-2j+2n} &= \begin{cases} 1 - \max\left\{\frac{1}{2}, \frac{c_{2d-2j+2n-1}}{2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}}\right\}, & \text{if } c_{2d-2j+2n-1} > c_{2d-2j+2n} \\ \max\left\{\frac{1}{2}, \frac{c_{2d-2j+2n}}{2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}}\right\}, & \text{if } c_{2d-2j+2n} \geq c_{2d-2j+2n-1} \end{cases} \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad n = 0, \dots, j-1 \\
p_{2d} &= 1 - \frac{c_{2d-1}}{2^{2j} \prod_{l=d-j}^{d-1} p_{2l} q_{2l}}.
\end{aligned}
$$

(b) If there exists a constrained optimal discriminating design for the vector $(c_{2d-2j-1}, \dots, c_{2d-2}, c_{2d})$ *in* (7)*, then there also exists a symmetric constrained optimal discriminating design* σ^**. The canonical moments up to the order* $2d$ *of the corresponding projection* ξ_{σ^*} *are determined by the system of equations*

$$
\begin{aligned}
p_{2n-1} &= \frac{1}{2} \quad n = 1, \dots, d \\
p_{2n} &= \frac{1}{2} \quad n = 1, \dots, d-j-1 \\
p_{2d-2j+2n} &= \begin{cases} 1 - \max\left\{\frac{1}{2}, \frac{c_{2d-2j+2n-1}}{2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}}\right\}, & \text{if } c_{2d-2j+2n-1} > c_{2d-2j+2n} \\ \max\left\{\frac{1}{2}, \frac{c_{2d-2j+2n}}{2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}}\right\}, & \text{if } c_{2d-2j+2n} \geq c_{2d-2j+2n-1} \end{cases} \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad n = 0, \dots, j-1 \\
p_{2d} &= \frac{c_{2d}}{2^{2j} \prod_{l=d-j}^{d-1} p_{2l} q_{2l}}.
\end{aligned}
$$

A necessary condition for the existence of optimal designs with respect to either criterion ((6) *or* (7)) *is given by* $c_{2d-2j} + c_{2d-2j-1} \leq 1$.

Proof Because both parts are proven similarly, we restrict ourselves to a proof of part (a). By the previous discussion the canonical moments of odd order $1, 3, \ldots, 2d-1$ must be $1/2$. Note that

$$\mathrm{eff}_{2d-2j+2n}(\sigma) = p_{2d-2j+2n} 2^{2(d-j+n-1)} \prod_{l=1}^{d-j+n-1} p_{2l} q_{2l}\,, \quad n = 0, \ldots, j.$$

In order to maximize these efficiencies we have to choose the canonical moments such that the products $p_{2l}q_{2l}$ are as large as possible. This can be accomplished by choosing p_{2l} as close as possible to the value $1/2$ such that the constraints in (6) are satisfied. Since there are no restrictions on the efficiencies $\mathrm{eff}_1(\sigma), \ldots, \mathrm{eff}_{2d-2j-2}(\sigma)$ we obtain $p_2 = \cdots = p_{2d-2j-2} = \frac{1}{2}$. Substituting this choice into the formulae for the higher order efficiencies, (16) reduces to

$$\mathrm{eff}_{2d-2j+2n-1}(\sigma) = q_{2d-2j+2n} 2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}$$

$$\mathrm{eff}_{2d-2j+2n}(\sigma) = p_{2d-2j+2n} 2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l} q_{2l}.$$

We start with the case $n = 0$, for which the representations $\mathrm{eff}_{2d-2j-1}(\sigma) = q_{2d-2j}$, $\mathrm{eff}_{2d-2j}(\sigma) = p_{2d-2j}$ yield the constraints

$$p_{2d-2j} \geq c_{2d-2j}, \quad q_{2d-2j} \geq c_{2d-2j-1}.$$

Consequently any design ξ_σ for which $p_{2d-2j} \in [c_{2d-2j}, 1 - c_{2d-2j-1}]$ satisfies the constraints of order $2d-2j$ and $2d-2j-1$. We therefore assume that $c_{2d-2j} + c_{2d-2j-1} \leq 1$ in what follows to ensure the existence of such a design. If $\frac{1}{2} \in [c_{2d-2j}, 1 - c_{2d-2j-1}]$ one can choose $p_{2d-2j} = \frac{1}{2}$ to maximize $p_{2d-2j}q_{2d-2j}$. Else we have either $c_{2d-2j} \geq \frac{1}{2}$ or $1 - c_{2d-2j-1} \leq \frac{1}{2}$, and we choose $p_{2d-2j} = c_{2d-2j}$ or $p_{2d-2j} = 1 - c_{2d-2j-1}$, respectively.

If $n > 0$ we note that the constraints $\mathrm{eff}_{2d-2j+2n}(\sigma) \geq c_{2d-2j+2n}$ and $\mathrm{eff}_{2d-2j+2n-1}(\sigma) \geq c_{2d-2j+2n-1}$ reduce to

$$p_{2d-2j+2n} \geq \frac{c_{2d-2j+2n}}{2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l}q_{2l}} =: c'_{2d-2j+2n}$$

$$q_{2d-2j+2n} \geq \frac{c_{2d-2j+2n-1}}{2^{2n} \prod_{l=d-j}^{d-j+n-1} p_{2l}q_{2l}} =: c'_{2d-2j+2n-1}$$

Therefore the same arguments as presented for the case $n = 0$ yield the corresponding result for $p_{2d-2j+2n}$. Finally we consider the case $n = j$, where there is only one constraint $\mathrm{eff}_{2d-1}(\sigma) \geq c_{2d-1}$, which can be rewritten as

$$p_{2d} \leq 1 - \frac{c_{2d-1}}{2^{2d-2} \prod_{l=1}^{d-1} p_{2l}q_{2l}}.$$

In order to maximize this expression one has to choose p_{2d} such that there is equality. This proves the final assertion of part (a) in Theorem 1. □

Remark 1 Note that Theorem 1 characterizes the canonical moments up to the order $2d$ of the projection ξ_{σ^*} of the (symmetric) constrained optimal discriminating design σ^*. In general $p_{2d} \notin \{0, 1\}$ and in these cases there exists an infinite number of probability measures on the interval $[-1, 1]$ with the canonical moments $p_1, \dots p_{2d}$ (see Skibinsky 1986). Each of these measures corresponds to a constrained optimal discriminating design by the projection (8). One possible choice among these designs with a reasonable small support will be illustrated later in the proof of Theorem 2, Eq. (24).

In what follows, we present two further results, where a solution of the constrained optimal design problem can be found explicitly. For this purpose let $T_j(x)$ and $U_j(x)$ denote the jth Chebyshev polynomial of the first and second kind, respectively (see Rivlin 1974).

Theorem 2 *Consider the constrained optimal design problem in* (6) *where* $c_{2d-2j} = \cdots = c_{2d-2} = c \in (0, 1)$, $c_l < c$ $(l = 2d-2j-1, \dots, 2d-1)$. *If there exists a constrained optimal discriminating design, then there also exists a symmetric constrained optimal discriminating design* σ^*.

(a) If $c > 1/2$, *define*

$$\kappa = \kappa(c_{2d-1}, c) = \frac{1}{2} - \frac{c_{2d-1}}{4c} + \frac{2c-1}{2((2c-1)j-2c)},$$

$$\begin{aligned} P^*_{d+1}(x) &= (xU_j(x) - 2\kappa U_{j-1}(x))T_{d-j}(x) \\ &\quad - 2c(xU_{j-1}(x) - 2\kappa U_{j-2}(x))T_{d-j-1}(x), \end{aligned} \tag{17}$$

$$\begin{aligned} P^{(1)}_d(x) &= (xU_j(x) - 2\kappa U_{j-1}(x))U_{d-j-1}(x) \\ &\quad - 2c(xU_{j-1}(x) - 2\kappa U_{j-2}(x))U_{d-j-2}(x) \end{aligned} \tag{18}$$

The polynomial $P^*_{d+1}(x)$ *has* $d+1$ *distinct roots* $x_0, \dots x_d$ *in the interval* $(-1, 1)$, *and the design* ξ_{σ^*} *with masses*

$$\lambda_k = \frac{P^{(1)}_d(x_k)}{\frac{\mathrm{d}}{\mathrm{d}x} P^*_{d+1}(x)|_{x=x_k}}, \quad k = 0 \dots, d \tag{19}$$

at $x_0, \dots, x_d$ *corresponds to a constrained optimal discriminating design by the projection* (8) *for the optimization problem* (6).

(b) If $c < 1/2$, *define*

$$P^*_{d+1}(x) = T_{d+1}(x) - (1-2c)T_{d-1}(x),$$
$$P^{(1)}_d(x) = U_d(x) - (1-2c)U_{d-2}(x).$$

The polynomial $P^*_{d+1}(x)$ *has* $d+1$ *distinct roots* $x_0, \dots, x_d$ *in the interval* $(-1, 1)$, *and the design* ξ_{σ^*} *with masses* (19) *at* $x_0, \dots, x_d$ *corresponds to a constrained optimal discrimination design by the projection* (8) *for the optimization problem* (6)

Remark 2 A necessary condition for the existence of a symmetric constrained optimal discrimination design for the design problem (6) is $\kappa > 0$, which ensures that the value of the canonical moment p_{2d} will be within the interval (0, 1). If $\kappa \leq 0$ it is therefore recommended to modify the choices of c and c_{2d-1} accordingly so that κ attains a positive value, before starting to calculate the optimal design.

Proof We only prove part (a) of the Theorem. Part (b) follows by similar (and even simpler) arguments. Note that the canonical moments of the constrained optimal discriminating design can be obtained by Theorem 1. The canonical moments of odd order satisfy

$$p_{2n-1} = \frac{1}{2}, \quad n = 1, \dots, d, \tag{20}$$

while the canonical moments of even order less or equal than $2d - 2j - 2$ are given by

$$p_{2n} = \frac{1}{2}, \quad n = 1, \dots, d - j - 1. \tag{21}$$

For the next canonical moment of even order we have from Theorem 1

$$p_{2d-2j} = \max\left\{\frac{1}{2}, c\right\} = c,$$

and it can be shown by a straightforward induction that

$$p_{2d-2j+2t} = \frac{1}{2}\frac{(2c-1)t - 2c}{(2c-1)(t+1) - 2c}, \quad t = 0, \dots, j-1. \tag{22}$$

Note that this representation implies

$$\frac{1}{2} < c < \frac{j+1}{2j},$$

because the canonical moments vary in the interval (0, 1). The remaining canonical moment of order $2d$ is obtained by a direct calculation, that is

$$p_{2d} = 1 - \frac{1}{2}\frac{c_{2d-1}}{c}\frac{(2c-1)j - 2c}{(2c-1)(j+1) - 2c}. \tag{23}$$

It follows from a straightforward but tedious calculation that $p_{2d} \in (0, 1) \Leftrightarrow \kappa > 0$, which proves Remark 2. Note that (20)–(23) do not determine a design on the interval $[-1, 1]$ (except in the case $c_{2d-1} = 0$, which is excluded). In order to obtain a design with finite support we extend this sequence by

$$p_{2d+1} = \frac{1}{2}, \quad p_{2d+2} = 0. \tag{24}$$

The design ξ_σ^* on the interval $[-1, 1]$ with canonical moments (20)–(24) is uniquely determined and has $d + 1$ support points not including 1 or -1 (see Skibinsky 1986), entailing that the design σ^* will be supported on $2d+2$ points. For the calculation of the support points and corresponding weights we apply Theorem 3.6.1 in Dette and Studden (1997). By this result the design ξ_{σ^*} has weights

$$\lambda_k \;=\; \frac{P_d^{(1)}(x_k)}{\frac{\mathrm{d}}{\mathrm{d}x}P_{d+1}^*(x)|_{x=x_k}} \tag{25}$$

at the roots $x_0,\ldots,x_d$ of the polynomial $P_{d+1}^*(x)$, where $P_{d+1}^*(x)$ and $P_d^{(1)}(x)$ are obtained from the recursion

$$W_{k+1}(x) = xW_k(x) - q_{2k-2}p_{2k}W_{k-1}(x) \tag{26}$$

(note that $p_{2j-1} = \frac{1}{2}$ for $j = 1,\ldots,d+1$) with different initial conditions, that is

$$P_{d+1}^*(x) = W_{d+1}(x) \quad \text{for } W_{-1}(x) \equiv 0,\ W_0(x) \equiv 1 \tag{27}$$

$$P_d^{(1)}(x) = W_{d+1}(x) \quad \text{for } W_0(x) \equiv 0,\ W_1(x) \equiv 1. \tag{28}$$

We now calculate these polynomials using (21)–(24) and begin with $P_{d+1}^*(x)$. From the initial condition in (27) and (21) we obtain by a straightforward calculation

$$W_{d-j}(x) = \frac{1}{2^{d-j-1}}T_{d-j}(x)\ , \quad W_{d-j-1}(x) = \frac{1}{2^{d-j-2}}T_{d-j-1}(x) \tag{29}$$

Observing (22) and

$$\begin{aligned} q_{2l-2}p_{2l} &= \left(1 - \frac{1}{2}\frac{(2c-1)(l-1-d+j)-2c}{(2c-1)(l-d+j)-2c}\right)\frac{1}{2}\frac{(2c-1)(l-d+j)-2c}{(2c-1)(l-d+j+1)-2c} \\ &= \frac{1}{4}\frac{(2c-1)(l-d+j+1)-2c}{(2c-1)(l-d+j+1)-2c} = \frac{1}{4} \end{aligned}$$

$(d-j < l \le d-1)$, we obtain the recursion

$$\begin{aligned} W_{d-j+1} &= xW_{d-j}(x) - \frac{1}{2}cW_{d-j-1}(x) \\ W_{l+1}(x) &= xW_l(x) - \frac{1}{4}W_{l-1}(x) \quad \text{if } d-j < l \le d-1. \end{aligned}$$

Now a straightforward induction yields

$$W_{d-j+l}(x) = \frac{1}{2^{l+d-j-1}}(U_l(x)T_{d-j}(x) - 2cU_{l-1}(x)T_{d-j-1}(x))\ , \quad l = 1,\ldots,j.$$

We finally note that by (22) and (23) we have $q_{2d-2}p_{2d} = \kappa$, from which it follows that

$$P_{d+1}^*(x) = \frac{1}{2^{d-1}}\Big[(xU_j(x) - 2\kappa U_{j-1}(x))T_{d-j}(x) - 2c(xU_{j-1}(x) - 2\kappa U_{j-2}(x))T_{d-j-1}(x)\Big]$$

using (26) and (27). Observing the initial conditions in (28) it follows that the polynomial $P_d^{(1)}(x)$ can be calculated analogously, where (29) is replaced by

$$W_{d-j}(x) = \frac{1}{2^{d-j-1}}U_{d-j-1}(x)\ , \quad W_{d-j-1}(x) = \frac{1}{2^{d-j-2}}U_{d-j-2}(x).$$

Consequently, by a straightforward induction we obtain

$$P_d^{(1)}(x) = \frac{1}{2^{d-1}}\Big[(xU_j(x) - 2\kappa U_{j-1}(x))U_{d-j-1}(x) - 2c(xU_{j-1}(x) - 2\kappa U_{j-2}(x))U_{d-j-2}(x)\Big],$$

and the assertion (a) of Theorem 1 follows from Theorem 3.6.1 in Dette and Studden (1997). □

We conclude this section with an analogue for the optimization problem (7). The proof is similar and omitted for brevity.

Theorem 3 *Consider the constrained optimal design problem in* (7) *where* $c_{2d-2j} = \cdots = c_{2d-2} = c_{2d} = c \in (0, 1)$, $c_l < c$ $(l = 2d - 2j - 1, \ldots, 2d - 3)$. *If there exists a constrained optimal discriminating design then there also exists a symmetric constrained optimal discriminating design* σ^*.

(a) If $c > 1/2$, *define*

$$\kappa = \kappa(c_{2d}, c) = \frac{c_{2d}}{4c},$$

and consider for this κ *the polynomials* $P_{d+1}^*(x)$ *and* $P_d^{(1)}(x)$ *defined by* (17) *and* (18), *respectively. The polynomial* $P_{d+1}^*(x)$ *has* $d+1$ *distinct roots* $x_0, \ldots, x_d$ *in the interval* $(-1, 1)$, *and the design* ξ_{σ^*} *which has masses* (19) *at the points* $x_0, \ldots, x_d$ *corresponds to a constrained optimal discriminating design for the optimization problem* (7) *by the projection* (8).

(b) If $c < 1/2$, *define*

$$P_{d+1}^*(x) = T_{d+1}(x) + (1 - 2c)T_{d-1}(x),$$
$$P_d^{(1)}(x) = U_d(x) + (1 - 2c)U_{d-2}(x).$$

The polynomial $P^*_{d+1}(x)$ *has* $d+1$ *distinct roots* $x_0, \ldots, x_d$ *in the interval* $(-1, 1)$*, and the design* ξ_{σ^*} *with masses* (19) *corresponds to a constrained optimal discrimination design for the optimization problem* (7) *by the projection* (8)*.*

4 Concluding remarks and discussion

We have applied the theory of canonical moments to derive constrained optimal designs for discriminating between Fourier models of different degree. For general constraints, we have found explicit recursive relations (see Theorem 1) for the first $2d$ canonical moments of the optimal designs, from which optimal designs can be computed applying Theorem 3.6.1 in Dette and Studden (1997). For some special cases (with respect to the constraints) we present explicit formulae for the supporting polynomials and the weights, which allow the direct computation of one of the optimal designs (see Theorems 2 and 3).

The choice of the lower bounds c_l, $l = 2d - 2j - 1, \ldots, 2d$, for the efficiencies is up to the experimenter according to his interest in a specific testing problem. There is, however, no guarantee that there exists an optimal design satisfying these constraints. Apart from the necessary conditions that $\kappa > 0$ (see Remark 2) and $c'_{2d-2j+2n} + c'_{2d-2j+2n-1} \leq 1$, where $n = 0, \ldots, j-1$ and c'_k are defined in the proof of Theorem 1 there seems to be no simple check if a particular choice of values c_l will be admissible. Naturally, the experimenter would like to have the bounds as large as possible, which might, however, contradict the existence of an optimal design. Once the bounds have been chosen, we recommend to either compute the optimal canonical moments by the recursive relations given in Theorem 1 or the optimal supporting polynomials $P^*_{d+1}(x)$ via Theorems 2 or 3 (if appropriate). If any of the canonical moments is outside the open interval $(0, 1)$ or the roots of the polynomial $P^*_{d+1}(x)$ are either outside $(-1, 1)$ or not distinct then there exists no optimal design with respect to these constraints. In this situation, the choice of the constraints has to be modified by lowering the values of the less important bounds.

If the experimenter does not favour any testing problem over another, a natural approach to choose the bounds is $c_l = 1/2$, $l \in \{2d - 2j - 1, \ldots, 2d\}$ for some j. Applying Theorem 1 yields that in this situation all canonical moments up to the order $2d$ equal $1/2$. A somewhat related approach would be to maximize the minimal efficiency $\text{eff}_l(\sigma)$ for any l within some subset $\mathcal{L}$ of $\{1, 2, \ldots, 2d\}$. This criterion has already been considered in Dette and Haller (1998), Sect. 4, Theorem 4.3. From part (b) of this theorem, it follows for example that for the choice $\mathcal{L} = \{2d - 2j - 1, \ldots, 2d\}$, the optimal designs also have canonical moments $p_k = 1/2$, $k = 1, \ldots, 2d$. These designs therefore coincide with the constrained optimal designs where all values c_l, $l \in \mathcal{L}$, are equal to $1/2$. This corresponds to intuition since neither approach gives preference to any testing problem within $\mathcal{L}$ over another. For different choices of $\mathcal{L}$ the reader is referred to Sect. 4 in Dette and Haller (1998). In practice, however, there will be higher interest in testing the higher order coefficients in most situations.

A challenging problem for future considerations will be the generalization of our results and techniques to models in more than one dimension. In this situation, one approach may be to find optimal designs within the class of product designs [see, e.g., Dette and Studden (1997), Sect. 5.8, for some results on multivariate polynomials]. Another method worth trying will be a lattice design approach as described in Riccomagno et al. (1997), i.e. embed the higher dimensional problem into a one-dimensional structure and exploit results in one dimension, including those in the present paper.

Acknowledgments This work of the authors was supported by the SFB 475 (Komplexitätsreduktion in multivariaten Datenstrukturen). The work of H. Dette was supported in part by a NIH grant award IR01GM072876:01A1. The authors would also like to thank I. Gottschlich, who typed this paper with considerable technical expertise. We are also grateful to two referees for their constructive comments, which led to a substantial improvement of an earlier version of this paper.

References

Anderson, T. W. (1994). The statistical analysis of time series. Classics Library, New York: Wiley.

Currie, A. J., Ganeshanandam, S., Noiton, D. A., Garrick, D., Shelbourne, C. J. A., Oraguzie, N. (2000). Quantitative evaluation of apple (Malus × domestica Borkh.) fruit shape by principle component analysis of Fourier descriptors. *Euphytica, 111*, 219E27.

Clyde, M., Chaloner, K. (1996). The equivalence of constrained and weighted designs in multiple objective design problems. *Journal of the American Statistical Association, 91*, 1236–1244.

Cook, R. D., Wong, W. K. (1994). On the equivalence of constrained and compound optimal designs. *Journal of the American Statistical Association, 89*, 687–692.

Dette, H. (1995). Discussion of the paper "Constrained optimization of experimental design" by V. Fedorov and D. Cook. *Statistics, 26*, 153–161.

Dette, H., Franke, T. (2000). Constrained D- and D_1-optimal designs for polynomial regression. *Annals of Statistics, 28,* 1702–1727.

Dette, H., Haller, G. (1998). Optimal designs for the identification of the order of a Fourier regression. *Annals of Statistics, 26*, 1496–1521.

Dette, H., Melas, V. B. (2003). Optimal designs for estimating individual coefficients in Fourier regression models. *Annals of Statistics, 31*, 1669–1692.

Dette, H., Melas, V. B., Biedermann, S. (2002a). A functional-algebraic determination of D-optimal designs for trigonometric regression models on a partial circle. *Statistics & Probability Letters, 58(4)*, 389–397.

Dette, H., Melas, V. B., Pepelyshev, A. (2002b). D-optimal designs for trigonometric regression models on a partial circle. *Annals of the Institute of Statistical Mathematics, 54(4),* 945–959.

Dette, H., Studden, W. J. (1997). The theory of canonical moments with applications in statistics, probability and analysis. New York: Wiley.

Fedorov, V. V. (1972). *Theory of optimal experiments.* New York: Academic Press.

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

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

Kitsos, C. P., Titterington, D. M., Torsney, B. (1988). An optimal design problem in rhythmometry. *Biometrics, 44,* 657–671.

Lau, T. S., Studden, W. J. (1985). Optimal designs for trigonometric and polynomial regression. *Annals of Statistics, 13*, 383–394.

Lee, C. M. S. (1988a). Constrained optimal designs. *Journal of Statistical Planninng and Inference, 18*, 377–389.

Lee, C. M. S. (1988b). D-optimal designs for polynomial regression, when lower degree parameters are more important. *Utilitas Mathematica, 34,* 53–63.

Lestrel, P. E. (1997). *Fourier descriptors and their applications in biology.* Cambridge: Cambridge University Press.

Mardia, K. (1972). *The statistics of directional data.* New York: Academic Press.

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

Riccomagno, E., Schwabe, R., Wynn, H.P. (1997). Lattice-based D-optimum designs for Fourier regression models. *Annals of Statistics, 25(6)*, 2313–2327.

Rivlin, T. J. (1974). *Chebyshev polynomials.* New York: Wiley.

Skibinsky, M. (1967). The range of the $(n+1)$th moment for distributions on [0, 1]. *Journal of Applied Probability, 4*, 543–552.
Skibinsky, M. (1986). Principal representations and canonical moment sequences for distributions on an interval. *Journal of Mathematical Analysis and Applications, 120,* 95–120.
Stigler, S. (1971). Optimal experimental designs for polynomial regression. *Journal of the American Statistical Association, 66*, 311–318.
Studden, W. J. (1968). Optimal designs on Tchebycheff points. *Annals of Mathematical Statistics, 39,* 1435–1447.
Studden, W. J. (1980). D_s-optimal designs for polynomial regression using continued fractions. *Annals of Statistics, 8*, 1132–1141.
Studden, W. J. (1982a). Optimal designs for weighted polynomial regression using canonical moments. In S. S. Gupta, J. O. Berger (Eds.), Third Purdue symposium on decision theory and related topics (Vol. 2, pp. 335–350).
Studden, W. J. (1982b). Some robust type D-optimal designs in polynomial regression. *Journal of the American Statistical Association, 8*, 916–921.
Studden, W. J. (1989). Note on some ϕ_p-optimal designs for polynomial regression. *Annals of Statistics, 17*, 618–623.
Younker, J. L., Ehrlich, R. (1977). Fourier biometrics: Harmonic amplitudes as multivariate shape descriptors. *Systematic Zoology, 26*, 336E42.
Zen, M.-M., Tsai, M.-H. (2004). Criterion-robust optimal designs for model discrimination and parameter estimation in Fourier regression models. *Journal of Statistical Planninng and Inference, 124*, 475–487.