T-optimal designs for discrimination between two polynomial models

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: holger.dette@rub.de
FAX: +49 234 3214 559

Viatcheslav B. Melas, Petr Shpilev
St. Petersburg State University
Department of Mathematics
St. Petersburg , Russia
email: v.melas@pobox.spbu.ru
email: pitshp@hotmail.com

Abstract

The paper is devoted to the explicit construction of optimal designs for discrimination between two polynomial regression models of degree $n-2$ and n. In a fundamental paper Atkinson and Fedorov (1975a) proposed the T-optimality criterion for this purpose. Recently Atkinson (2010) determined T-optimal designs for polynomials up to degree 6 numerically and based on these results he conjectured that the support points of the optimal design are cosines of the angles that divide a half of the circle into equal parts if the coefficient of x^{n-1} in the polynomial of larger degree vanishes. In the present paper we give a strong justification of the conjecture and determine all T-optimal designs explicitly for any degree $n \in \mathbb{N}$. In particular, we show that there exists a one-dimensional class of T-optimal designs. Moreover, we also present a generalization to the case when the ratio between the coefficients of x^{n-1} and x^n is smaller than a certain critical value. Because of the complexity of the optimization problem T-optimal designs have only been determined numerically so far and this paper provides the first explicit solution of the T-optimal design problem since its introduction by Atkinson and Fedorov (1975a). Finally, for the remaining cases (where the ratio of coefficients is larger than the critical value) we propose a numerical procedure to calculate the T-optimal designs. The results are also illustrated in an example.

AMS Subject Classification: 62K05

Keywords and Phrases: T-optimum design; discrimination designs; uniform approximation; Chebyshev polynomials; model uncertainty; goodness-of-fit test

1 Introduction

The problem of identifying an appropriate model in a class of competing regression models is of fundamental importance in regression analysis and occurs often in real experimental studies. It is nowadays widely accepted that good experimental designs can improve the performance of discrimination, and several authors have addressed the problem of constructing optimal designs for this purpose [see Hunter and Reiner (1965), Stigler (1971), Atkinson and Fedorov (1975a,b), Hill (1978), Studden (1982), Spruill (1990), Dette (1994, 1995), Dette and Haller (1998), Song and Wong (1999), Ucinski and Bogacka (2005), Wiens (2009, 2010) among many others]. In a fundamental paper Atkinson and Fedorov (1975a) introduced the T-optimality criterion for discriminating between two competing regression models. As an example, these authors constructed T-optimal designs for a constant and a quadratic model. Since its introduction the problem of determining T-optimal designs has been considered by numerous authors [see Atkinson and Fedorov (1975b), Ucinski and Bogacka (2005), Wiens (2009), Tommasi and López-Fidalgo (2010) among others]. The T-optimal design problem is essentially a minimax problem and except for very simple models the corresponding optimal designs are not easy to find and have to be determined numerically. In a recent paper Dette and Titoff (2009) discussed the T-optimal design problem from a general point of view and related it to a nonlinear problem in approximation theory. As an illustration, designs for discriminating between a linear model and a cubic model without quadratic term were presented and it was shown that T-optimal designs are in general not unique. Atkinson (2010) considered a similar problem of this type and studied the problem of discriminating between two competing polynomial regression models which differ in the degree by two. This author determined T-optimal designs for polynomials up to degree 6 numerically where the coefficient of x^{n-1} in the polynomial of larger degree (say n) vanishes. Based on these results he conjectured that the support points of the T-optimal design are cosines of angles dividing a half of circle into equal parts.

The present paper has two purposes. In particular, we prove the conjecture raised in Atkinson (2010) and derive explicit solutions of the T-optimal design problem for discriminating between polynomial regression models of degree $n-2$ and n for any $n \in \mathbb{N}$. Moreover, we also determine the T-optimal designs analytically in the case when the ratio of the coefficients of the terms x^{n-1} and x^n is sufficiently small. The situation considered in Atkinson (2010) corresponds to the case where this ratio vanishes, and in this case we show that there exists a one-dimensional class of T-optimal designs. To our best knowledge these results provide the first explicit solution of the T-optimal design problem in a non-trivial situation. Our results provide further insight into the complicated structure of the T-optimal design problem. Finally, in the case where the coefficient exceeds the critical value we suggest a procedure to determine the T-optimal design numerically.

2 The T-optimal design problem revisited

Consider the classical regression model

$$y = \eta(x) + \varepsilon, \tag{2.1}$$

where the explanatory variable x varies in the design space $\mathcal{X}$ and observations at different locations, say x and x' are assumed to be uncorrelated with the same variance. In (2.1) the quantity ε denotes a random variable with mean 0 and variance σ^2 and η is a function, which is called regression function in the literature. We assume that the experimenter has two parametric models for this function in mind, that is

$$\eta_1(x, \theta_1) \quad \text{and} \quad \eta_2(x, \theta_2) \tag{2.2}$$

and the first goal of the experiment is to discriminate between these two models. In (2.2) the quantities θ_1 and θ_2 denote unknown parameters which vary in compact parameter spaces, say $\Theta_1 \subset \mathbb{R}^{m_1}$ and $\Theta_2 \subset \mathbb{R}^{m_2}$, and have to be estimated from the data. In order to find "good" designs for discriminating between the models η_1 and η_2 we consider approximate designs in the sense of Kiefer (1974), which are defined as probability measures on the design space $\mathcal{X}$ with finite support. The support points of an (approximate) design ξ give the locations where observations are taken, while the weights give the corresponding relative proportions of total observations to be taken at these points. If the design ξ has masses $\omega_i > 0$ at the different points x_i $(i = 1, \ldots, k)$ and N observations can be made by the experimenter, the quantities $\omega_i N$ are rounded to integers, say n_i, satisfying $\sum_{i=1}^k n_i = N$, and the experimenter takes n_i observations at each location x_i $(i = 1, \ldots, k)$.
To determine a good design for discriminating between the models η_1 and η_2 Atkinson and Fedorov (1975a) proposed in a fundamental paper to fix one model, say η_1 (more precisely its corresponding parameter θ_1) and to determine the design which maximizes the minimal deviation between the model η_1 and the class of models defined by η_2, that is

$$\xi^* = \arg\max_{\xi} \int_{\mathcal{X}} \left(\eta_1(x, \theta_1) - \eta_2(x, \theta_2^*)\right)^2 \xi(dx),$$

where the parameter θ_2^* minimizes the expression

$$\theta_2^* = \arg\min_{\theta_2 \in \Theta_2} \int_{\mathcal{X}} \left(\eta_1(x, \theta_1) - \eta_2(x, \theta_2)\right)^2 \xi(dx).$$

Note that θ_2^* is not an estimate but corresponds to best approximation of the "given" model $\eta_1(\cdot, \theta_1)$ by models of the form $\{\eta_2(\cdot, \theta_2) \mid \theta_2 \in \Theta_2\}$ with respect to a weighted L_2-norm. Since its introduction the T-optimal design problem has found considerable interest in the literature and we refer the interested reader to the work of Ucinski and Bogacka (2005) or Dette and Titoff (2009) among others. In general, the determination of T-optimal designs is a very difficult problem and explicit solutions are – to our best knowledge – not available except for very simple models with a few parameters. In this paper we present analytical results for T-optimal designs, if the interest is in the discrimination between two polynomial models which differ in the degree by two. To be precise, we consider the case where the regression functions $\eta_1(x,\theta_1)$ and $\eta_2(x,\theta_2)$ are given by

$$\eta_1(x,\theta_1) = \theta_{10} + \theta_{11}x + \ldots + \theta_{1n-2}x^{n-2} + \theta_{1n-1}x^{n-1} + \theta_{2n}x^n, \tag{2.3}$$

and

$$\eta_2(x,\theta_2) = \theta_{20} + \theta_{21}x + \ldots + \theta_{2n-2}x^{n-2}, \tag{2.4}$$

respectively, and the design space is given by $\mathcal{X} = [-1,1]$. In model (2.3) the parameter θ_1 is given by $\theta_1 = (\theta_{10},\theta_{11},\ldots,\theta_{1n-2},b\theta_{1n},\theta_{1n})^T$, where the ratio of the coefficients corresponding to the highest powers $b = \theta_{1n-1}/\theta_{1n}$ and the parameter θ_{1n} specify the deviation from a polynomial of degree $n-2$.
In the following discussion we define

$$\overline{\eta}(x,\alpha,b,\theta_{1n}) = \eta_1(x,\theta_1) - \eta_2(x,\theta_2) = \alpha_0 + \alpha_1 x + \ldots + \alpha_{n-2}x^{n-2} + \theta_{1n}(bx^{n-1} + x^n), \tag{2.5}$$

where we use the notation $\alpha_i = \theta_{1i} - \theta_{2i}$ $(i = 0,\ldots,n-2)$, then the problem of finding the T-optimal design for the models η_1 and η_2 can be reduced to

$$\xi^* = \arg\max_{\xi} \int_{\mathcal{X}} \left(\alpha_0^* + \alpha_1^* x + \ldots + \alpha_{n-2}^* x^{n-2} + \theta_{1n}(bx^{n-1} + x^n)\right)^2 \xi(dx)$$

where $\alpha^* = (\alpha_1^*,\ldots,\alpha_{n-2}^*)^T$ is a vector minimizing the expression

$$\alpha^* = \arg\min_{\alpha} \int_{\mathcal{X}} \left(\overline{\eta}(x,\alpha,b,\theta_{1n})\right)^2 \xi(dx).$$

It is now easy to see that for a fixed value of $b = \theta_{1n-1}/\theta_{1n}$ the T-optimal design does not depend on the parameter θ_{1n}. In the next section we give the complete solution of the T-optimal design problem if the absolute value of the parameter $b = \theta_{1n-1}/\theta_{1n}$ less or equal than some critical value.

3 T-optimal designs for small values of $|b| = |\theta_{1n-1}/\theta_{1n}|$

Throughout this section we assume that the parameter b satisfies

$$|b| = |\theta_{1n-1}/\theta_{1n}| \le n(1-\cos\left(\frac{\pi}{n}\right))/(1+\cos\left(\frac{\pi}{n}\right)) = n\tan^2\left(\frac{\pi}{2n}\right), \tag{3.1}$$

then it is easy to see that all points

$$t_i^*(b) = -\left(1 + \frac{|b|}{n}\right)\cos\left(\frac{i\pi}{n}\right) - \frac{|b|}{n}, \quad i = 1,\ldots,n \tag{3.2}$$

are located in the interval $[-1,1]$. Our first result gives an explicit solution of the T-optimal design problem in the case $b=\theta_{1n-1}=0$ and – as a by-product – proves the conjecture raised in Atkinson (2010).

Theorem 3.1 *A design ξ is T-optimal for discriminating between the models* (2.3) *and* (2.4) *with $\theta_{1n-1}=0$ on the interval $[-1,1]$ if and only if it can be represented in the form $\xi=(1-\alpha)\xi_1+\alpha\xi_2$, where $\alpha\in[0,1]$, the measures ξ_1 and ξ_2 are defined by*

$$\xi_1=\begin{pmatrix} t_1^*(0) & \dots & t_n^*(0) \\ \omega_1^* & \dots & \omega_n^* \end{pmatrix}, \qquad \xi_2=\begin{pmatrix} -t_n^*(0) & \dots & -t_1^*(0) \\ \omega_n^* & \dots & \omega_1^* \end{pmatrix}, \tag{3.3}$$

and the weights and support points are given by

$$\omega_i^*=\frac{2}{n}\sin^2\left(\frac{i\pi}{2n}\right),\quad \omega_{n-i}^*=\frac{2}{n}\cos^2\left(\frac{i\pi}{2n}\right),\quad i=1,\dots,\left\lfloor\frac{n}{2}\right\rfloor,\quad \omega_n^*=\frac{1}{n}, \tag{3.4}$$

and (3.2) *for $b=0$, respectively.*

Proof of Theorem 3.1. It was proved by Dette and Titoff (2009) [see Theorem 2.1] that any T-optimal design on the interval $[-1,1]$ for discriminating between the polynomials $\sum_{j=0}^{n-2}\theta_{2j}x^j$ and

$$\eta_1(x,\theta_1)=\sum_{j=0}^{n-2}\theta_{1j}x^{j^2}+\theta_{1n}x^n$$

(note that $\theta_{1n-1}=0$) is supported at the set of the extremal points

$$\mathcal{A}=\left\{x\in[-1,1]\;\middle|\;\psi^*(x)=\sup_{t\in[-1,1]}|\psi^*(t)|\right\}$$

where $\psi^*(x)=\eta_1(x,\theta_1)-\sum_{j=0}^{n-2}\overline{\theta}_{2j}x^j$ and

$$\overline{\theta}_2=(\overline{\theta}_{20},\dots,\overline{\theta}_{2n-2})^T=\arg\min_{\theta_2\in\mathbb{R}^{n-1}}\sup_{x\in[-1,1]}|\;\eta_1(x,\theta_1)-\sum_{j=0}^{n-2}\theta_{2j}x^j\;| \tag{3.5}$$

is the parameter corresponding to the best approximation of $\eta_1(x,\theta_1)$ with respect to the sup-norm. By a standard result in approximation theory [see Achiezer (1956), Section 35 and 43] it follows that the solution of the problem (3.5) is unique and given by $\psi^*(x)=\theta_{1n}2^{-(n-1)}T_n(x)$, where $T_n(x)=\cos(n\arccos x)$ is the nth Chebyshev polynomial of the first kind. Note that $T_n(x)$ is an even or odd polynomial of degree n with leading coefficient 2^{n-1} [see Szegö (1975)]. The corresponding extremal points are given by $x_0=t_1^*(0)=-1$, $x_i=t_i^*(0)=-\cos\frac{i\pi}{n}$, $i=1,\dots,n-1$, $x_n=t_n^*(0)=1$.

Now it follows from Theorem 2.2 in Dette and Titoff (2009) that a design ξ^* is T-optimal if and only if it satisfies the system of linear equations

$$\int_{\mathcal{A}} \psi^*(x) x^k d\xi^*(x) = 0 \qquad k = 0, \ldots, n-2 \tag{3.6}$$

(note that in the case of linear models the necessary condition in Theorem 2.2 in Dette and Titoff (2009) is also sufficient). Therefore for proving that $\xi_1^* = \xi_1$ is a T-optimal design it is sufficient to verify the identities

$$\int \psi^*(x) d\xi_1^*(x) = \theta_{1n} 2^{-(n-1)} (-1)^n \sum_{i=1}^{n} (-1)^i x_i^k \omega_i^* = 0, \;\; k = 0, 1, \ldots, n-2, \tag{3.7}$$

which will be done in the Appendix. In a similar way we can check that the design ξ_2^* in (3.3) is a T-optimal design. Note that

$$\text{supp } (\xi_1^*) \cup \text{ supp } (\xi_2^*) = \Big\{ x_i = -\cos\Big(\frac{\pi}{n} i\Big) \mid i = 0, \ldots, n \Big\} = \mathcal{A}$$

because $t_{n-i}^*(0) = -t_i^*(0)$. Moreover, (3.6) defines a system of linear equations of the form $F\omega = 0$ for the vector $\omega = (\omega_0, \ldots, \omega_n)^T$ of the T-optimal design ξ^*, where the matrix F is given by $F = ((-1)^i x_i^k)_{i=0,\ldots,n}^{k=0,\ldots,n-2} \in \mathbb{R}^{n-1 \times n+1}$ and has rank $n-1$. Additionally, the components of the vector ω satisfy $\sum_{i=0}^{n} \omega_i = 1$. Therefore the set of solutions has dimension 1. Because the vectors of weights corresponding to the designs ξ_1^* and ξ_2^* are given by $\omega^{(1)} = (0, \omega_1^*, \ldots, \omega_n^*)^T$ and $\omega^{(2)} = (\omega_n^*, \ldots, \omega_1^*, 0)^T$ and are therefore linearly independent (note that $\omega_i^* > 0, \; i = 1, \ldots, n$), any vector of weights corresponding to a T-optimal design must be a convex combination of $\omega^{(1)}$ and $\omega^{(2)}$. Consequently, any T-optimal design can be represented in the form $\xi = (1-\alpha)\xi_1^* + \alpha\xi_2^*$, which proves the assertion of Theorem 3.1. □

Note that the T-optimal design is not unique in the case $b = 0$. On the other hand, the T-optimal designs are unique, whenever $\theta_{1n-1} \neq 0$, and, if the ratio $|\theta_{1n-1}/\theta_{1n}|$ is not too large, the T-optimal designs can also be found explicitly as demonstrated in our following result.

Theorem 3.2 *If the parameter $b = \theta_{1n-1}/\theta_{1n}$ satisfies* (3.1)*, then there exists a unique T-optimal design on the interval $[-1, 1]$ for discriminating between the models* (2.3) *and* (2.4)*. For positive b this design has the form*

$$\xi^* = \begin{pmatrix} t_1^*(b) & \ldots & t_n^*(b) \\ \omega_1^* & \ldots & \omega_n^* \end{pmatrix}, \tag{3.8}$$

where the points $t_i^*(b)$ and weights $w_i^*(b)$ are defined in (3.2) *and* (3.4)*, respectively (note that $t_1^*(b) \geq -1, t_n^*(b) = 1$). The T-optimal design for negative b has the form*

$$\xi^* = \begin{pmatrix} -t_n^*(b) & \ldots & -t_1^*(b) \\ \omega_n^* & \ldots & \omega_1^* \end{pmatrix}$$

(note that $-t_n^*(b) = -1, -t_1^*(b) \leq 1$).

Proof of Theorem 3.2. We consider the case $0 < b \le n(1-\cos\left(\frac{\pi}{n}\right))/(1+\cos\left(\frac{\pi}{n}\right))$ where direct calculations show that the points $t_i^*(b), i = 1,\dots,n$ are contained in the interval $[-1,1]$. Moreover, these points are the extremal points of the polynomial

$$c_n T_n\left(\frac{-x-\frac{b}{n}}{1+\frac{b}{n}}\right), c_n = (-1)^n\left(\frac{1}{2}\right)^{n-1}\left(1+\frac{b}{n}\right)^n \tag{3.9}$$

where T_n is the Chebyshev polynomial of the first kind. For later purposes we note that the coefficient of x^{n-1} in this polynomial is equal to

$$\sum_{i=1}^{n}\left[\left(1+\frac{b}{n}\right)u_i + \frac{b}{n}\right] = b, \tag{3.10}$$

where $u_1,\dots,u_n$ are the roots of the polynomial $T_n(x)$, that is $u_i = \cos(\frac{2i-1}{2n}\pi)$ $(i = 1,\dots,n)$, $\sum_{i=1}^n u_i = 0$. It can be shown by a standard argument in approximation theory [see Achiezer (1956), Section 35 and 43] that $\theta_{1n}\psi^*(x)$ with

$$\psi^*(x) = c_n T_n\left(\frac{-x-\frac{b}{n}}{1+\frac{b}{n}}\right)$$

is the unique solution of the extremal problem

$$\min_{\theta_2\in\mathbb{R}^{n-1}} \sup_{x\in[-1,1]} \mid \eta_1(x,\theta_1) - \sum_{j=0}^{n-2}\theta_{2j}x^j \mid,$$

where $\eta_1(x,\theta_1) = \sum_{j=0}^n \theta_{1j}x^j$. Therefore by Theorem 2.1 and 2.2 in Dette and Titoff (2009) a T-optimal design is supported at the n extremal points $t_1^*(b),\dots,t_n^*(b)$ (note that we use $b \le n\tan^2(\frac{\pi}{2n})$ at this point, which implies $|t_j^*(b)| \le 1; j = 1,\dots,n$) and the weights are determined by (3.6). Because the set of extremal points is given by $\mathcal{A} = \{t_1^*(b),\dots,t_n^*(b)\}$ this system reduces to

$$\sum_{i=1}^{n} t_i^{*k}(b)(-1)^i\omega_i^* = 0,\ k = 0,1,\dots,n-2, \tag{3.11}$$

and we will prove in the appendix that the weights given in (3.4) define a solution of (3.11). Therefore the design ξ^* specified in (3.8) is a T-optimal design for $0 < b \le n(1-\cos\pi/n)/(1+\cos\pi/n)$. Since the function $\psi^*(x)$ is unique, any T-optimal design is supported at the points $t_1^*(b),\dots,t_n^*(b)$ [see Theorem 2.1 in Dette and Titoff (2009)]. By Theorem 2.2 in the same reference it follows that the weights of any T-optimal design satisfy the system of linear equations (3.11) with $\omega_i^* = \omega_i$ and $\sum_{i=1}^n \omega_i = 1$. Since $\psi^*(t_i^*(b)) = (-1)^i$ $(i = 1,\dots,n)$ we can rewrite this system as

$$F\omega = e_n, \tag{3.12}$$

Table 1: *The critical values* $b_n^* = n\tan^2\left(\frac{\pi}{2n}\right)$ *for various values* $n \in \mathbb{N}$.

n	3	4	5	6	7	8	9	10
b_n^*	1	0.6864	0.5280	0.4306	0.3646	0.3168	0.2801	0.2509

where $\omega = (\omega_1, \dots, \omega_n)^T$ is the vector of weights, the last row of the matrix F is given by $(1, \dots, 1)$ and corresponds to the condition $\sum_{i=1}^n \omega_i = 1$, $e_n = (0, \dots, 0, 1)^T \in \mathbb{R}^n$ denotes the nth unit vector and the columns of the matrix F are given by

$$a_i = (-1)^i (1, t_i^*(b), \dots, (t_i^*(b))^{n-2}, \psi^*(t_i^*(b)))^T, \; i = 1, 2, \dots, n.$$

The remaining assertion of Theorem 3.2 follows if we prove that $\det F \neq 0$, which implies that the solution of (3.12) and therefore the T-optimal design is unique. For this purpose assume that the opposite holds. In this case the rows of the matrix F would be linearly dependent and there exists a vector $h = (h_1, \dots, h_{n-1}, 1)^T$ such that $a_i^T h = 0, \; i = 1, 2, \dots, n$. But the function $k(x) = (1, x, \dots, x^{n-2}, \psi^*(x))^T h$ is a polynomial of degree n with coefficient of x^{n-1} given by b. Since $a_i h = k(t_i^*(b)) = 0$ this polynomial has roots at the points $t_i^*(b)$, moreover

$$\sum_{i=1}^n t_i^*(b) = -b - \sum_{i=1}^n \left(1 + \frac{b}{n}\right) \cos\left(\frac{i\pi}{n}\right) = -b + 1 + \frac{b}{n}.$$

However, by (3.10) the sum of the roots must equal $-b$ by Vieta's formula. This contradiction proves that $\det F \neq 0$. Therefore the system of equations in (3.12) has a unique solution, which means that the T-optimal design is unique.
The case of negative b is considered in a similar way and the details are omitted for the sake of brevity. □

The critical values $b_n^* = n\tan^2\left(\frac{\pi}{2n}\right)$ for various values of $n \in \mathbb{N}$ are displayed in Table 1. Theorem 3.1 and 3.2 give an explicit solution of the T-optimal design problem for discriminating between a polynomial regression of degree $n-2$ and n, whenever $|b| = |\theta_{1n-1}|/|\theta_{1n}| \leq b_n$. In the opposite case the solution is not so transparent and will be discussed in the following section.

4 T-optimal designs for large values of $|b|$

In this section we consider the case $|b| \geq n\tan^2\left(\frac{\pi}{2n}\right)$ for which the T-optimal design cannot be found explicitly. Therefore we present a numerical method to determine the optimal designs. The method was described by Dette et al. (2004) in the context of determining optimal designs for estimating individual coefficients in a polynomial regression model [see also Melas (2006)] and for the sake of brevity we only explain the basic principle. For this purpose we rewrite the function $\overline{\eta}$ in (2.5) as

$$\overline{\eta}(x,\alpha,\bar{b}) = \alpha_0 + \alpha_1 x + \ldots + \alpha_{n-2}x^{n-2} + \theta_{1n-1}(x^{n-1} + \bar{b}x^n), \tag{4.1}$$

where $\bar{b} = 1/b = \theta_{1n}/\theta_{1n-1}$. Note that for fixed $\bar{b}$ the T-optimal design is independent of the parameter θ_{1n-1} and that the choice

$$\bar{b} \in \left[-\frac{1}{n}\cot^2\left(\frac{\pi}{2n}\right), \frac{1}{n}\cot^2\left(\frac{\pi}{2n}\right) \right]$$

corresponds to the case $|b| \geq n\tan^2\left(\frac{\pi}{2n}\right)$ considered in this section. In order to express the dependence on the parameter $\bar{b}$ we use the notation $t_i^*(\bar{b})$ for the support points and $\omega_i^*(\bar{b})$ for the weights of the T-optimal design in this section.

The main idea of the algorithm is a representation of $t_i^*(\bar{b})$ and $\omega_i^*(\bar{b})$ in terms of a Taylor series, where the coefficients can be determined explicitly as soon as the series is known for a particular point $\bar{b}$. In the present situation this point is given by $\bar{b} = 0$, which corresponds to the situation of discriminating between a polynomial of degree $n-2$ and $n-1$. For this case it follows from Dette and Titoff (2009) that the T-optimal design coincides with the D_1-optimal design. This design has been determined explicitly by Studden (1980) and puts masses $\omega_i(0) = \frac{1}{n-1}$ at the points $t_i(0) = \cos\left(\frac{(i-1)\pi}{n-1}\right)$ $(i = 2, \ldots, n-1)$ and masses $\omega_1(0) = \omega_n(0) = \frac{1}{2(n-1)}$ at the points $t_1(0) = -1$ and $t_n(0) = 1$.

For the constructions of the Taylor expansion we now associate to each vector

$$\tau \in \mathcal{U} = \left\{ (t_2,\ldots,t_{n-1},\omega_1,\ldots,\omega_{n-1})^T \Big| -1 < t_2 < \ldots < t_{n-1} < 1; \omega_i > 0, \sum_{j=1}^{n-1}\omega_j < 1 \right\},$$

a design with n support points defined by

$$\xi_\tau = \begin{pmatrix} -1 & t_2 & \ldots & t_{n-1} & 1 \\ \omega_1 & \omega_2 & \ldots & \omega_{n-1} & \omega_n \end{pmatrix}.$$

As pointed out in the previous discussion there exists a corresponding extremal problem defined by

$$\inf_{q\in\mathbb{R}^{n-1}} \sup_{x\in[-1,1]} \left| \bar{b}x^n + x^{n-1} - \bar{f}^T(x)q \right| \tag{4.2}$$

with a unique solution corresponding to the T-optimal design problem under consideration, where we use the notation $\bar{f}^T(x) = (1, x, \ldots, x^{n-2})$. For each vector q in (4.2) define vectors $d_q = (q^T, 1, \bar{b})^T, \Theta = (q,\tau)$ and a quadratic form

$$H(\Theta,\bar{b}) = H(q,\tau,\bar{b}) = d_q^T M(\xi_\tau) d_q,$$

where $M(\xi_\tau)$ is the information matrix of the design ξ_τ for the regression model (4.1). It then follows by similar results as in Dette et al. (2004) that the design ξ_{τ^*} is a T-optimal design for discriminating between the polynomials of degree n and $n-2$ and the vector q^* is a solution of an extremal problem (4.2) if the points $\Theta^* = (q^*, \tau^*) \in \mathbb{R}^{n-1} \times \mathcal{U}$ is the unique solution of the system

$$\frac{\partial}{\partial \Theta} H(\Theta, \bar{b})\Big|_{\Theta=\Theta^*} = 0,$$

such that the inequality $\left|d_{q^*}^T f(x)\right|^2 \le d_{q^*}^T M(\xi_{\tau^*}) d_{q^*}$ holds for all $x \in [-1,1]$. Additionally, the function

$$\Theta^* : \begin{cases} I \longrightarrow \mathbb{R}^{3n-4} \\ \bar{b} \longrightarrow \Theta^*(\bar{b}) = (\Theta_1^*(\bar{b}), \ldots, \Theta_{3n-4}^*(\bar{b})) = (q^*(\bar{b})^T, \tau^*(\bar{b})^T). \end{cases}$$

which maps the parameter $\bar{b} \in I = [-\frac{1}{n}\cot^2\left(\frac{\pi}{2n}\right), \frac{1}{n}\cot^2\left(\frac{\pi}{2n}\right)]$ to the coordinates of the best approximation $q^*(\bar{b})$ and the support points $t_i^*(\bar{b})$ and weights $\omega^*(\bar{b})$ of the T-optimal design, is a real analytical function. The coefficients in the corresponding Taylor expansion

$$\Theta^*(\bar{b}) = \Theta^*(\bar{b}_0) + \sum_{j=1}^{\infty} \Theta^*(j, \bar{b}_0)(\bar{b} - \bar{b}_0)^j$$

in a neighborhood of any point $\bar{b}_0 \in I$ can be calculated by the recursive formulas

$$\Theta^*(s+1, \bar{b}_0) = -\frac{1}{(s+1)!} J^{-1}(\bar{b}_0) \left(\frac{d}{db}\right)^{s+1} g(\Theta_{(s)}^*(\bar{b}), \bar{b})\Big|_{\bar{b}=\bar{b}_0}, s = 0, 1, 2, \ldots,$$

where

$$\begin{aligned} \Theta_{(s)}^*(\bar{b}) &= \Theta_{(s)}^*(\bar{b}_0) + \sum_{j=1}^{s} \Theta^*(j, \bar{b}_0)(\bar{b} - \bar{b}_0)^j, \\ g(\Theta, \bar{b}) &= \frac{\partial}{\partial \Theta} H(\Theta, \bar{b}) \\ J(\bar{b}_0) &= \left(\frac{\partial^2}{\partial \Theta_i \partial \Theta_j} H(\Theta, \bar{b})\right)\Big|_{\Theta=\Theta^*(\bar{b}_0)}. \end{aligned}$$

We can use this result to calculate the T-optimal design for discriminating between polynomials of degree n and $n-2$ in the cases which are not covered by Theorem 3.1 and 3.2. We illustrate the methodology in the following example.

Example 4.1 Consider the T-optimal design problem for a model of degree 5 and a cubic polynomial model. Note that for $n=5$ we have $n\tan^2(\frac{\pi}{2n}) \simeq 0.528$. Therefore if $b \in [0, 0.528]$ a T-optimal design is given by Theorem 3.1, that is

$$\xi_T^* = \begin{pmatrix} t_1(b) & t_2(b) & t_3(b) & t_4(b) & 1 \\ 0.038 & 0.138 & 0.262 & 0.362 & \frac{1}{5} \end{pmatrix},$$

$$t_i^*(b) = -\left(1 + \frac{b}{5}\right) \cos\left(\frac{i\pi}{5}\right) - \frac{b}{5}, \quad i = 1, \ldots, 5.$$

In order to construct the T-optimal design on the interval $[0.528, \infty]$ we introduce the notation $\bar{b} = 1/b \in [0, 1.894]$. With the results of the previous paragraph we obtain a Taylor expansion for the interior support points $t_2^*(\bar{b}), t_3^*(\bar{b}), t_4^*(\bar{b})$ and weights $\omega_1^*(\bar{b}), \omega_2^*(\bar{b}), \omega_3^*(\bar{b}), \omega_4^*(\bar{b})$ of the T-optimal design for discriminating between a cubic and a polynomial of degree 5 where $\bar{b} = \theta_{1n}/\theta_{1n-1}$. By the results of Studden (1980) the vector of support points and weights corresponding to the center of the expansion at the point $\bar{b}_0 = 0$ is explicitly known, that is

$$(t_2^*(0), t_3^*(0), t_4^*(0), \omega_1^*(0), \ldots, \omega_4^*(0)) = (-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, \frac{1}{8}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}).$$

At the first step we use a Taylor expansion at the point $\bar{b}_0 = 0$ to determine the T-optimal design for $\bar{b} \in [0, 0.4]$. When we have found the vector $\Theta^*(0.4)$ we construct a further Taylor expansion at the point $\bar{b}_0 = 0.4$ and this process is continued in order to determine the vector $\Theta^*(\bar{b})$ for any value $\bar{b} \in [0, 1.894]$. The support points and weights are depicted in Figure 1 as a function of the parameter $\bar{b} = 1/b = \theta_{1n}/\theta_{1n-1}$. Note that in all cases $b \neq 0$ the T-optimal design for discriminating between a polynomial of degree 5 and 3 is supported at 5 points.

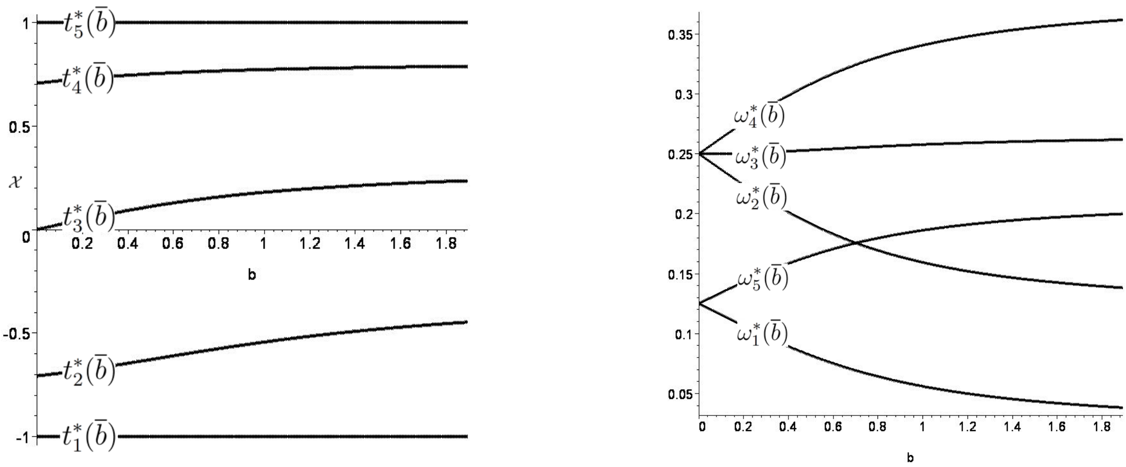

Figure 1: *The support points (left panel) and weights (right panel) of the T-optimal design for discriminating between a polynomial of degree* 3 *and* 5 *for various values of* $\bar{b} = 1/b \in [0, 1.894]$.

Acknowledgements. This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Teilprojekt C2) of the German Research Foundation (DFG). The work by V. B. Melas and Petr Shpilev was partly supported by Russian Foundation of Basic Research (project 09-01-00508). The authors would like to thank Martina Stein, who typed numerous versions of this manuscript with considerable technical expertise.

5 Appendix. Proof of the identities (3.7) and (3.11)

Note that the identities in (3.7) and (3.11) can be written in the form

$$\sum_{i=1}^{n} t_i^{*k}(b)(-1)^i \omega_i^* = 0, \ \ k = 0, 1, \ldots, n-2, \tag{5.1}$$

where $t_i^*(0) = \cos(\frac{i\pi}{n}) = x_i$. We will prove that these equalities hold for any real number b. Since

$$t_i^{*k}(b) = \sum_{j=0}^{k} a_j \cos\left(\frac{ji\pi}{n}\right), i = 0, 1, \ldots, n, k = 0, 1, \ldots, n-2 \tag{5.2}$$

for some coefficients $a_j = a_j(b)$ $(j = 0, 1, \ldots, k)$ the identities (5.1) follow from

$$\sum_{i=1}^{n} (-1)^i \cos\left(\frac{ki\pi}{n}\right) \omega_i^* = 0, \ \ k = 0, 1, \ldots, n-2. \tag{5.3}$$

In order to prove (5.3) consider first the case $k = 0, n = 2s$ for some s, where the left hand side of (5.3) reduces to

$$\begin{aligned}
\sum_{i=1}^{n} \omega_i^*(-1)^i &= \frac{1}{n}\Big[\sum_{i=1}^{s-1}\Big[(1-\cos\Big(\frac{i\pi}{n}\Big))(-1)^i + (1+\cos\Big(\frac{i\pi}{n}\Big))(-1)^i\Big] + (-1)^s + 1\Big] \\
&= \frac{1}{n}\Big[\sum_{i=1}^{s-1} 2(-1)^i + (-1)^s + 1\Big] = 0,
\end{aligned}$$

which proves (5.3). If $k = 0, n = 2s+1$ we get

$$\begin{aligned}
\sum_{i=1}^{n} \omega_i^*(-1)^i &= \frac{1}{n}\Big[\sum_{i=1}^{s}\Big[(1-\cos\Big(\frac{i\pi}{n}\Big))(-1)^i - (1+\cos\Big(\frac{i\pi}{n}\Big))(-1)^i\Big] + (-1)\Big] \\
&= \frac{1}{n}\Big[2\sum_{i=1}^{s} \cos\left(\frac{i\pi}{n}\right)(-1)^{i+1} - 1\Big] \\
&= \frac{1}{n}\Big[1 - \frac{\cos\Big[\frac{\pi(1+2(n+1)s)}{2n}\Big]}{\cos\Big(\frac{\pi}{2n}\Big)} - 1\Big] = -\frac{1}{n}\frac{\cos\Big(\frac{(2s+1)\pi}{2}\Big)}{\cos\Big(\frac{\pi}{2n}\Big)} = 0
\end{aligned}$$

where the third identity follows by standard results for trigonometrical summation [see e.g. Jolley (1961), formula (428)]. This proves (5.3) for the case $k = 0, \ n = 2s+1$. Now consider the case of even $n, n = 2s$ for some odd $s, s = 2l-1$ and k of the form $k = 2(2r-1)$. In this case the left hand side of (5.3) reduces to

$$
\begin{aligned}
&\frac{1}{n}\Big[\sum_{i=1}^{s-1}\Big[(1-\cos\Big(\frac{i\pi}{n}\Big)) + (1+\cos\Big(\frac{i\pi}{n}\Big))\Big](-1)^i \cos\left(\frac{ki\pi}{n}\right) + (-1)^s \cos\left(\frac{k\pi}{2}\right) + \cos(k\pi)\Big] \\
&= \frac{1}{n}\Big[2\sum_{i=1}^{s-1}(-1)^i \cos\left(\frac{ki\pi}{n}\right) + (-1)^s \cos\left(\frac{k\pi}{2}\right) + \cos(k\pi)\Big] \\
&= \frac{1}{n}\Big\{\Big(\cos\Big(\frac{k\pi}{4s}\Big)\Big)^{-1}\Big[\cos\Big(\frac{\pi k}{4s}-\pi\Big) + \cos\Big(\frac{\pi k}{4s}+\frac{\pi}{2}(k+2s-2)\Big)\Big] + 2\Big\} \\
&= \frac{1}{n}\Big\{(-1) + (-1)^{2s-1} + 2\Big\} = 0
\end{aligned}
$$

where we have again used well known results on trigonometric summation [see Jolley (1961), formula (428)]. Therefore we obtain the equality (5.3) in the case $n = 2s, s = 2l - 1$ and $k = 2(2r - 1)$. The other cases can be proved in a similar way, and the details are omitted for the sake of brevity.

References

Achiezer, N. I. (1956). *Theory of Approximation*. Ungar, New York.

Atkinson, A. C. (2010). The non-uniqueness of some designs for discriminating between two polynomial models in one variablel. *MODA 9, Advances in Model-Oriented Design and Analysis*, pages 9–16.

Atkinson, A. C. and Fedorov, V. V. (1975a). The designs of experiments for discriminating between two rival models. *Biometrika*, 62:57–70.

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

Dette, H. (1994). Discrimination designs for polynomial regression on a compact interval. *Annals of Statistics*, 22:890–904.

Dette, H. (1995). Optimal designs for identifying the degree of a polynomial regression. *Annals of Statistics*, 23:1248–1267.

Dette, H. and 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., and Pepelyshev, A. (2004). Optimal designs for estimating individual coefficient in polynomial regression - a functional approach. *Journal of Statistical Planning and Inference*, 118:201–219.

Dette, H. and Titoff, S. (2009). Optimal discrimination designs. *Annals of Statistics*, 37(4):2056–2082.

Hill, P. D. (1978). A review of experimental design procedures for regression model discrimination. *Technometrics*, 20(1):15–21.

Hunter, W. G. and Reiner, A. M. (1965). Designs for discriminating between two rival models. *Technometrics*, 7(3):307–323.

Jolley, L. B. W. (1961). *Summation of Series. 2nd revised ed.* Dover Publications, New York.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *Annals of Statistics*, 2:849–879.

Melas, V. B. (2006). *Functional Approach to Optimal Experimental Design (Lecture Notes in Statistics 184).* Springer, New York.

Song, D. and Wong, W. K. (1999). On the construction of g_{rm}-optimal designs. *Statistica Sinica*, 9:263–272.

Spruill, M. C. (1990). Good designs for testing the degree of a polynomial mean. *Sankhya, Ser. B*, 52(1):67–74.

Stigler, S. (1971). Optimal experimental design for polynomial regression. *Journal of the American Statistical Association*, 66:311–318.

Studden, W. J. (1980). D_s-optimal designs for polynomial regression using continued fractions. *Annals of Statistics*, 8(5):1132–1141.

Studden, W. J. (1982). Some robust-type D-optimal designs in polynomial regression. *Journal of the American Statistical Association*, 77(380):916–921.

Szegö, G. (1975). *Orthogonal Polynomials.* American Mathematical Society, Providence, R.I.

Tommasi, C. and López-Fidalgo, J. (2010). Bayesian optimum designs for discriminating between models with any distribution. *Computational Statistics & Data Analysis*, 54(1):143–150.

Ucinski, D. and Bogacka, B. (2005). T-optimum designs for discrimination between two multiresponse dynamic models. *Journal of the Royal Statistical Society, Ser. B*, 67:3–18.

Wiens, D. P. (2009). Robust discrimination designs, with Matlab code. *Journal of the Royal Statistical Society, Ser. B*, 71:805–829.

Wiens, D. P. (2010). Robustness of design for the testing of lack of fit and for estimation in binary response models. *Computational Statistics and Data Analysis*, 54:3371–3378.