Locally E-optimal designs for exponential regression models

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de

Viatcheslav B. Melas
St. Petersburg State University
Department of Mathematics
St. Petersburg
Russia
email: v.melas@pobox.spbu.ru

Andrey Pepelyshev
St. Petersburg State University
Department of Mathematics
St. Petersburg
Russia
email: andrey@ap7236.spb.edu

November 21, 2003

Abstract

In this paper we investigate locally E- and c-optimal designs for exponential regression models of the form $\sum_{i=1}^{k} a_i \exp(-\lambda_i x)$. We establish a numerical method for the construction of efficient and locally optimal designs, which is based on two results. First we consider the limit $\lambda_i \to \gamma$ and show that the optimal designs converge weakly to the optimal designs in a heteroscedastic polynomial regression model. It is then demonstrated that in this model the optimal designs can be easily determined by standard numerical software. Secondly, it is proved that the support points and weights of the locally optimal designs in the exponential regression model are analytic functions of the nonlinear parameters $\lambda_1, \dots, \lambda_k$. This result is used for the numerical calculation of the locally E-optimal designs by means of a Taylor expansion for any vector $(\lambda_1, \dots, \lambda_k)$. It is also demonstrated that in the models under consideration E-optimal designs are usually more efficient for estimating individual parameters than D-optimal designs.

AMS Subject Classification: 62K05, 41A50
Keywords and phrases: E-optimal design, c-optimal design, exponential models, locally optimal designs, Chebyshev systems, heteroscedastic polynomial regression.

1 Introduction

Nonlinear regression models are widely used to describe the dependencies between a response and an explanatory variable [see e.g. Seber and Wild (1989), Ratkowsky (1983) or Ratkowsky (1990)]. An important class of models in environmental and ecological statistics consists of exponential regression models defined by

$$\sum_{i=1}^{k} a_i e^{-\lambda_i x} \ , \quad x \geq 0, \tag{1.1}$$

[see, for example, Becka and Urfer (1996) or Becka, Bolt and Urfer (1993)]. An appropriate choice of the experimental conditions can improve the quality of statistical inference substantially and therefore many authors have discussed the problem of designing experiments for nonlinear regression models [see for example Chernoff (1953), Melas (1978) and Ford, Torsney and Wu (1992)]. Locally optimal designs depend on an initial guess for the unknown parameter, but are the basis for all advanced design strategies, [see Pronzato and Walter (1985), Chaloner and Verdinelli (1995), Ford and Silvey (1980) or Wu (1985)]. Most of the literature concentrates on D-optimal designs (independent of the particular approach), which maximize the determinant of the Fisher information matrix for the parameters in the model, but much less attention has been paid to E-optimal designs in nonlinear regression models, which maximize the minimum eigenvalue of the Fisher information matrix [see Dette and Haines (1994) or Dette and Wong (1999), who gave some results for models with two parameters].
It is the purpose of the present paper to study locally c-optimal and E-optimal designs for the nonlinear regression model (1.1). For this purpose we prove two main results. First we show that in the case $\lambda_i \to \gamma$ $(i = 1, \ldots, k)$ the locally optimal designs for the model (1.1) converge weakly to the optimal designs in a heteroscedastic polynomial regression model of degree $2k$ with variance proportional to $\exp(2\gamma x)$. It is then demonstrated that in most cases the E- and c-optimal designs are supported at the Chebyshev points, which are the local extrema of the equi-oscillating best approximation of the function $f_0 \equiv 0$ by a normalized linear combination of the form $\sum_{i=0}^{2k-1} a_i \exp(-\gamma x) x^i$. These points can be easily determined by standard numerical software [see for examples Studden and Tsay (1976)]. Secondly it is proved that the support points and weights of the locally optimal designs in the exponential regression model are analytic functions of the nonlinear parameters $\lambda_1, \ldots, \lambda_k$. This result is used to provide a Taylor expansion for the weights and support points as functions of the parameters, which can easily be used for the numerical calculation of the optimal designs. It is also demonstrated that in the models under consideration E-optimal designs are usually more efficient for estimating individual parameters than D-optimal designs.
The remaining part of the paper is organized as follows. In Section 2 we introduce the necessary notation, while the main results are stated in Section 3. In Section 4 we illustrate our method considering several examples and compare locally D- and E-optimal designs. Finally all technical details are deferred to an Appendix (see Section 5).

2 Preliminaries

Consider the common exponential regression model with homoscedastic error

$$\mathbf{E}(Y(x)) = \eta(x,\beta) = \sum_{i=1}^{k} a_i e^{-\lambda_i x}, \;\; \mathbf{V}(Y(x)) = \sigma^2 > 0, \tag{2.1}$$

where the explanatory variable x varies in the experimental domain $\mathcal{X} = [b, +\infty)$ with $b \in \mathbb{R}$, $\beta^T = (a_1, \lambda_1, a_2, \ldots, \lambda_k)$ denotes the vector of unknown parameters and different measurements are assumed to be uncorrelated. Without loss of generality we assume $a_i \neq 0$, $i = 1, \ldots, k$ and $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_k$. An approximate design ξ is a probability measure

$$\xi \;=\; \begin{pmatrix} x_1 & \ldots & x_n \\ w_1 & \ldots & w_n \end{pmatrix} \tag{2.2}$$

with finite support on $[b, \infty)$, where $x_1, \ldots, x_n$ give the locations, where observations are taken and $w_1, \ldots, w_n$ denote the relative proportions of total observations taken at these points [see Kiefer (1974)]. In practice a rounding procedure is applied to obtain the samples sizes $N_i \approx w_i N$ at the experimental conditions x_i, $i = 1, 2, \ldots, n$ [see e.g. Pukelsheim and Rieder (1993)]. If $n \geq 2k$, $w_i > 0$, $i = 1, \ldots, n$, it is well known that the least squares estimator $\hat{\beta}$ for the parameter β in model (2.1) is asymptotically unbiased with covariance matrix satisfying

$$\lim_{N\to\infty} \mathrm{Cov}(\sqrt{N}\hat{\beta}) = \sigma^2 M^{-1}(\xi, a, \lambda),$$

where

$$M(\xi) \;=\; M(\xi, a, \lambda) \;=\; \left(\sum_{s=1}^{n} \frac{\partial \eta(x_s, \beta)}{\partial \beta_i} \frac{\partial \eta(x_s, \beta)}{\partial \beta_j} w_s \right)_{i,j=1}^{2k}$$

denotes the information matrix of the design ξ. Throughout this paper we will use the notation

$$f(x) = \frac{\partial \eta(x,\beta)}{\partial \beta} = (e^{-\lambda_1 x}, -a_1 x e^{-\lambda_1 x}, \ldots, e^{-\lambda_k x}, -a_k x e^{-\lambda_k x})^T \tag{2.3}$$

for the gradient of the mean response function $\eta(x, \beta)$. With this notation the information matrix can be conveniently written as

$$M(\xi) = \sum_{i=1}^{n} f(x_i) f^T(x_i) w_i. \tag{2.4}$$

Note that for nonlinear models the information matrix depends on values of the unknown parameters, but for the sake of brevity we only reflect this dependence in our notation, if it is not clear from the context. An optimal design maximizes a concave real valued function of the information matrix and there are numerous optimality criteria proposed in the literature to discriminate between competing designs [see e.g. Silvey (1980) or Pukelsheim (1993)]. In this paper we restrict ourselves to three well known optimality criteria. Following Chernoff (1953) we call a design ξ locally D-optimal in the exponential regression model (2.1) if it

maximizes $\det M(\xi)$. The optimal design with respect to the determinant criterion minimizes the content of a confidence ellipsoid for the parameter β, based on the asymptotic covariance matrix. Locally D-optimal designs in various non-linear regression models have been discussed by numerous authors [see e.g. Melas (1978), He, Studden and Sun (1996) or Dette, Haines and Imhof (1999) among many others]. For a given vector $c \in \mathbb{R}^{2k}$ a design ξ is called locally c-optimal if $c \in \text{Range}(M(\xi))$ and ξ minimizes $c^T M^-(\xi)c$, which corresponds to the minimization of the asymptotic variance of the least squares estimator for the linear combination $c^T\beta$. If $c = e_i = (0, \dots, 0, 1, 0, \dots, 0)^T \in \mathbb{R}^{2k}$ is the ith unit vector, the locally e_i-optimal designs are also called optimal designs for estimating the ith coefficient $(i = 1, \dots, 2k)$. Locally c-optimal designs for nonlinear models with two parameters have been studied by Ford, Torsney and Wu (1992) among others. Finally we consider the E-optimality criterion, which determines the design such that the minimal eigenvalue of information matrix $M(\xi)$ is maximal. This corresponds to the the minimization of the worst variance of the least squares estimator for the linear combination $p^T\beta$ taken over all vectors p such that $p^Tp = 1$. Some locally E-optimal designs for models with two parameters have been found by Dette and Haines (1994) and Dette and Wong (1999).
In the following sections we determine and investigate locally E- and c-optimal designs for the exponential regression model (1.1). Secondly, we compare these designs with the corresponding locally D-optimal designs for the exponential model, which have been studied by Melas (1978). We begin our investigations with an important tool for analyzing E-optimal designs. A proof can be found in Pukelsheim (1993) or Melas (1982).

Theorem 2.1. *A design ξ^* is E-optimal if and only if there exists a nonnegative definite matrix A^* such that* $\text{tr}A^* = 1$ *and*

$$\max_{x \in \mathcal{X}} f^T(x) A^* f(x) \leq \lambda_{min}(M(\xi^*)). \tag{2.5}$$

Moreover, we have equality in (2.5) for any support point of ξ^*, and the matrix A^* can be represented as

$$A^* = \sum_{i=1}^{s} \alpha_i p_{(i)} p_{(i)}^T,$$

where s is the multiplicity of the minimal eigenvalue, $\alpha_i \geq 0$, $\sum_{i=1}^{s} \alpha_i = 1$, $\{p_{(i)}\}_{i=1,\dots,s}$ is a system of orthonormal eigenvectors corresponding to the minimal eigenvalue.

3 Main results

In this section we study some important properties of locally c- and E-optimal designs in the exponential regression model (2.1). In order to indicate the dependence of the optimal designs on the nonlinear parameters of the model we denote the locally c- and E-optimal design by $\xi_c^*(\lambda)$ and $\xi_E^*(\lambda)$, respectively. We begin with an investigation of the behaviour of the locally optimal designs if the vector of nonlinear parameters $\lambda = (\lambda_1, \dots, \lambda_k)$ is contained in a neighbourhood of a point $\gamma(1, \dots, 1)^T$, where $\gamma > 0$ is an arbitrary parameter. The information matrix (2.4) of any design becomes singular as $\lambda \to \gamma(1, \dots, 1)^T$. However we will show that the corresponding

locally optimal designs are still weakly convergent, where the limiting measure has $2k$ support points.
To be precise let

$$\lambda_i = \gamma - r_i\delta \ , \quad i = 1, \ldots, k \tag{3.1}$$

where $\delta > 0$ and $r_1, \ldots, r_k \in \mathbb{R}\backslash\{0\}$ are arbitrary fixed numbers such that $r_i \neq r_j$, $i, j = 1, \ldots, k$. If δ is small, locally c- and E-optimal designs in the exponential regression model (2.1) are closely related to optimal designs in the heteroscedastic polynomial regression model

$$\mathbf{E}(Y(x)) = \sum_{i=1}^{2k} a_i x^{i-1}, \ \ \mathbf{V}(Y(x)) = \exp(2\gamma x) \ , \quad x \in [b, \infty) \tag{3.2}$$

where $\gamma > 0$ is assumed to be known. Note that for a design of the form (2.2) the information matrix in this model is given by

$$\bar{M}(\xi) = \sum_{i=1}^{n} e^{-2\gamma x_i} \bar{f}(x_i) \bar{f}^T(x_i) w_i \ , \tag{3.3}$$

[see Fedorov (1972)], where the vector of regression functions defined by

$$\bar{f}(x) = (1, x, \ldots, x^{2k-1})^T. \tag{3.4}$$

The corresponding c-optimal designs are denoted by $\bar{\xi}_c^*$, where the dependence on the constant γ is not reflected in our notation, because it will be clear from the context. The next theorem shows that the e_{2k}-optimal design $\bar{\xi}_{2k}^* = \bar{\xi}_{e_{2k}}^*$ in the heteroscedastic polynomial regression appears as a weak limit of the locally c- and E-optimal design $\xi_c^*(\lambda)$ and $\xi_E^*(\lambda)$ in the model (2.1). The proof is complicated and therefore deferred to the Appendix.

Theorem 3.1.

1) *For any design with at least $2k$ support points and $\gamma > 0$ there exists a neighbourhood Ω_γ of the point $\gamma(1, \ldots, 1)^T \in \mathbb{R}^k$ such that for any vector $\lambda = (\lambda_1, \ldots, \lambda_k)^T \in \Omega_\gamma$ the minimal eigenvalue of information matrix $M(\xi)$ in (2.4) is simple.*

2) *If condition (3.1) is satisfied and $\delta \to 0$, then the locally E-optimal design $\xi_E^*(\lambda)$ in the exponential regression model (2.1) converges weakly to the e_{2k}-optimal design $\bar{\xi}_{e_{2k}}$ in the heteroscedastic polynomial regression model (3.2).*

3) *Assume that condition (3.1) is satisfied and define a vector $l = (l_1, \ldots, l_{2k})^T$ with $l_{2i} = 0$, $(i = 1, \ldots, k)$,*

$$l_{2i-1} = -\prod_{j \neq i} (r_i - r_j)^2 \sum_{j \neq i} \frac{2}{r_i - r_j} \ , \ i = 1, \ldots, k. \tag{3.5}$$

If $l^T c \neq 0$ and $\delta \to 0$ then the locally c-optimal design $\xi_c^*(\lambda)$ in the exponential regression model (2.1) converges weakly to the e_{2k}-optimal design $\bar{\xi}_{e_{2k}}$ in the heteroscedastic polynomial regression model (3.2).

Remark 3.2. It is well known [see e.g. Karlin and Studden (1966)] that the e_{2k}-optimal design $\bar{\xi}_{e_{2k}}$ in the heteroscedastic polynomial regression model (3.2) has $2k$ support points, say

$$x_1^*(\gamma) < \ldots < x_{2k}^*(\gamma).$$

These points are given by the extremal points of the Chebyshev function $p^*(x) = q^{*T} \bar{f}(x) e^{-\gamma x}$, which is the solution of the problem

$$\text{(3.6)} \qquad \sup_{x \in [b,\infty)} |p^*(x)| = \min_{\alpha_1,\ldots,\alpha_{2k-1}} \sup_{x \in [b,\infty)} \exp(-\gamma x) \Big| 1 + \sum_{i=1}^{2k-1} \alpha_i x^i \Big|.$$

Moreover, also the weights $w_1^*(\gamma), \ldots, w_{2k}^*(\gamma)$ of the e_{2k}-optimal design $\bar{\xi}_{e_{2k}}(\gamma)$ in model (3.2) can be obtained explicitly, i.e.

$$\text{(3.7)} \qquad w^*(\gamma) = (w_1^*(\gamma), \ldots, w_1^*(\gamma))^T = \frac{J \bar{F}^{-1} e_{2k}}{\mathbf{1}_{2k} J \bar{F}^{-1} e_{2k}},$$

where the matrixes $\bar{F}$ and J are defined by

$$\bar{F} = (\bar{f}(x_1^*(\gamma)) e^{-\gamma x_1^*(\gamma)}, \ldots, \bar{f}(x_{2k}^*(\gamma)) e^{-\gamma x_{2k}^*(\gamma)}) \in \mathbb{R}^{2k \times 2k},$$

$J = \operatorname{diag}(1, -1, 1, \ldots, 1, -1)$, respectively, $\mathbf{1}_{2k} = (1, \ldots, 1)^T \in \mathbb{R}^{2k}$ and the vector $\bar{f}(x)$ is defined in (3.4) [see Pukelsheim and Torsney (1991)].

Remark 3.3. Let Ω denote the set of all vectors $\lambda = (\lambda_1, \ldots, \lambda_k)^T \in \mathbb{R}^k$ with $\lambda_i \neq \lambda_j$, $i \neq j$, $\lambda_i > 0$, $i = 1, \ldots, k$, such that the minimum eigenvalue of the information matrix of the locally E-optimal design (with respect to the vector λ) is simple. The following properties of locally E-optimal designs follow by standard arguments from general results on E-optimal designs [see Dette and Studden (1993), Pukelsheim (1993)] and simplify the construction of locally E-optimal designs substantially.

1. For any $\lambda \in \Omega$ the locally E-optimal design for the exponential regression model (2.1) (with respect to the parameter λ) is unique.

2. For any $\lambda \in \Omega$ the support points of the locally E-optimal design for the exponential regression model (2.1) (with respect to the parameter λ) do not depend on the parameters $a_1, \ldots, a_k$.

3. For any $\lambda \in \Omega$ the locally E-optimal design for the exponential regression model (2.1) (with respect to the parameter λ) has $2k$ support points, moreover the point b is always a support point of the locally E-optimal design. The support points of the E-optimal design are the extremal points of the Chebyshev function $p^T f(x)$, where p is an eigenvector corresponding to the minimal eigenvalue of the information matrix $M(\xi_E^*(\lambda))$.

4. For any $\lambda \in \Omega$ the weights of the locally E-optimal design for the exponential regression model (2.1) (with respect to the parameter λ) are given by

$$w^* = \frac{JF^{-1}c}{c^Tc}, \tag{3.8}$$

where $c^T = \mathbf{1}_{2k}^T JF^{-1}$, $J = \operatorname{diag}(1, -1, 1, \ldots, 1, -1)$,

$$F = (f(x_1^*), \ldots, f(x_m^*)) \in \mathbb{R}^{2k \times 2k}$$

and $x_1^*, \ldots, x_{2k}^*$ denote the support points of the locally E-optimal design.

5. If $\lambda \in \Omega$, let $x_{1;b}^*(\lambda), \ldots, x_{2k;b}^*(\lambda)$ denote the support points of the locally E-optimal design for the exponential regression model (2.1) with design space $\mathcal{X} = [b, +\infty)$. Then $x_{1;0}^*(\lambda) \equiv 0$,

$$x_{i;b}^*(\lambda) \;=\; x_{i;0}^*(\lambda) + b, \qquad i = 2, \ldots, 2k.$$

$$x_{i;0}^*(\nu\lambda) \;=\; x_{i;0}^*(\lambda)/\nu, \qquad i = 2, \ldots, 2k$$

for any $\nu > 0$.

We now study some analytical properties of locally E-optimal designs for the exponential regression model (2.1). Theorem 3.1 indicates that the structure of the locally E-optimal design depends on the multiplicity of the minimal eigenvalue of its corresponding information matrix. If the multiplicity is equal to 1 then the support of an E-optimal design consists of the extremal points of the Chebyshev function $p^T f(x)$, where p is the eigenvector corresponding to the minimal eigenvalue of the information matrix $M(\xi_E^*(\lambda))$. If the multiplicity is greater than 1 then the problem of constructing E-optimal designs is more complex. Observing Remark 3.3(5) we assume that $b = 0$ and consider a design

$$\xi \;=\; \begin{pmatrix} x_1 & \ldots & x_{2k} \\ w_1 & \ldots & w_{2k} \end{pmatrix}$$

with $2k$ support points, $x_1 = 0$, such that the minimal eigenvalue of the information matrix $M(\xi)$ has multiplicity 1. If $p = (p_1, \ldots, p_{2k})^T$ is an eigenvector corresponding to the minimal eigenvalue of $M(\xi)$ we define a vector

$$\Theta = (\theta_1, \ldots, \theta_{6k-3})^T = (q_2, \ldots, q_{2k}, x_2, \ldots, x_{2k}, w_2, \ldots, w_{2k})^T, \tag{3.9}$$

where the points w_i and x_i ($i = 2, \ldots, 2k$) are the non-trivial weights and support points of the design ξ (note that $x_1 = 0$, $w_1 = 1 - w_2 - \ldots - w_{2k}$) and $q = (1, q_2, \ldots, q_{2k})^T = p/p_1$ is the normalized eigenvector of the information matrix $M(\xi)$. Note that there is a one-to-one correspondence between the pairs (q, ξ) and the vectors of the form (3.9). Recall the definition of the set Ω in Remark 3.3. For each vector $\lambda \in \Omega$ the maximum eigenvalue of the information matrix of a locally E-optimal design $\xi_E^*(\lambda)$ (for the parameter λ) has multiplicity 1 and for $\lambda \in \Omega$ let

$$\Theta^* = \Theta^*(\lambda) = (q_2^*, \ldots, q_{2k}^*, x_2^*, \ldots, x_{2k}^*, w_2^*, \ldots, w_{2k}^*)^T$$

denote the vector corresponding to the locally E-optimal design with respect to the above transformation. We consider the function

$$\Lambda(\Theta, \lambda) = \frac{\sum_{i=1}^{2k} (q^T f(x_i))^2 w_i}{q^T q}$$

(note that $x_1 = 0$, $w_1 = 1 - w_2 - \ldots - w_{2k}$), then it is easy to see that

$$\Lambda(\Theta^*(\lambda), \lambda) = \frac{q^{*T} M(\xi_E^*(\lambda)) q^*}{q^{*T} q^*} = \lambda_{min}(M(\xi_E^*(\lambda))),$$

where $\lambda_{min}(M)$ denotes the minimal eigenvalue of the matrix M. Consequently $\Theta^* = \Theta^*(\lambda)$ is an extremal point of the function $\Lambda(\Theta, \lambda)$. A necessary condition for the extremum is given by the system of equations

$$\frac{\partial \Lambda}{\partial \theta_i}(\Theta, \lambda) = 0, \; i = 1, \ldots, 6k-3, \tag{3.10}$$

and a straightforward differentiation shows that this system is equivalent to

$$\begin{cases} (M(\xi)q)_- - \Lambda(\Theta, \lambda) q_- = 0, \\ 2q^T f(x_i) q^T f'(x_i) w_i = 0, \quad i = 2, \ldots, 2k, \\ (q^T f(x_i))^2 - (q^T f(0))^2 = 0, \quad i = 2, \ldots, 2k, \end{cases} \tag{3.11}$$

where the vector $p_- \in \mathbb{R}^{2k-1}$ is obtained from the vector the $p \in \mathbb{R}^{2k-1}$ by deleting the first coordinate. This system is equivalent to the following system of equations

$$\begin{cases} M(\xi) p = \Lambda(\Theta, \lambda) p, \\ p^T f'(x_i) = 0, \quad i = 2, \ldots, 2k, \\ (p^T f(x_i))^2 = (p^T f(0))^2, \quad i = 2, \ldots, 2k, \end{cases} \tag{3.12}$$

and by the first part of Theorem 3.1 there exists a neighbourhood Ω_1 of the point $(1, \ldots, 1)^T$ such that for any $\lambda \in \Omega_1$ the vector $\Theta^*(\lambda)$ and the locally E-optimal design $\xi_E^*(\lambda)$ and its corresponding eigenvector p^* satisfy (3.10) and (3.12), respectively.

Theorem 3.4. *For any $\lambda \in \Omega$ the system of equations (3.10) has a unique solution*

$$\Theta^*(\lambda) = (q_2^*(\lambda), \ldots, q_{2k}^*(\lambda), x_2^*(\lambda), \ldots, x_{2k}^*(\lambda), w_2^*(\lambda), \ldots, w_{2k}^*(\lambda))^T.$$

The locally E-optimal design for the exponential regression model (2.1) is given by

$$\xi_E^*(\lambda) \;=\; \begin{pmatrix} 0 & x_2^*(\lambda) & \ldots & x_{2k}^*(\lambda) \\ w_1^*(\lambda) & w_2^*(\lambda) & \ldots & w_{2k}^*(\lambda) \end{pmatrix}$$

where $w_1^*(\lambda) = 1 - w_2^*(\lambda) - \ldots - w_{2k}^*(\lambda)$ and $q^*(\lambda) = (1, q_2^*(\lambda), \ldots, q_{2k}^*(\lambda))^T$ is an (normalized) eigenvector of the information matrix $M(\xi_E^*(\lambda))$. Moreover, the vector $\Theta^*(\lambda)$ is a real analytic function of λ.

It follows from Theorem 3.4 that for any $\lambda_0 \in \mathbb{R}^k$ such that the minimal eigenvalue of the information matrix corresponding to the locally E-optimal design $\xi_E^*(\lambda)$ has multiplicity 1 there exists a neighbourhood, say $\mathcal{U}$ of λ_0 such that for all $\lambda \in \mathcal{U}$ the function $\Theta^*(\lambda)$ can be expanded in convergent Taylor series of the form

$$\Theta^*(\lambda) = \Theta^*(\lambda_0) + \sum_{j=1}^{\infty} \Theta^*(j, \lambda_0)(\lambda - \lambda_0)^j. \tag{3.13}$$

It was shown in Dette, Melas and Pepelyshev (2004) that the coefficients $\Theta^*(j, \lambda_0)$ in this expansion can be calculated recursively and therefore this expansion provides a numerical method for the determination of the locally E-optimal designs using the analytic properties of the support points and weights as function of λ. From a theoretical point it is possible that several expansions have to be performed in order to cover the whole range of Ω of all values λ such that the minimum eigenvalue of the information matrix of the locally E-optimal design has multiplicity 1. However, in all our numerical examples only one expansion was sufficient (although we can not prove this in general).

Remark 3.5. Note that the procedure described in the previous paragraph would not give the locally E-optimal design for the exponential regression model in the case, where the minimum eigenvalue of the corresponding information matrix has multiplicity larger than 1. For this reason all designs obtained by the Taylor expansion were checked for optimality by means of Theorem 2.1. In all cases considered in our numerical study the equivalence theorem confirmed our designs to be locally E-optimal and we did not find cases where the multiplicity of the minimum eigenvalue of the information matrix in the exponential regression model (2.1) was larger than 1. Some illustrative examples are presented in the following section.

4 Examples

Example 4.1. Consider the exponential model $\mathbf{E}(Y(x)) = a_1 e^{-\lambda_1 x}$ corresponding to the case $k = 1$. In this case the Chebyshev function $\phi(x) = (1 + q_2^* x)e^{-\lambda_1 x}$ minimizing

$$\sup_{x \in [0,\infty)} |(1 + \alpha x)e^{-\lambda_1 x}|$$

with respect to the parameter $\alpha \in \mathbb{R}$ and the corresponding extremal point x_2^* are determined by the equations $\phi(x_2^*) = -\phi(0)$ and $\phi'(x_2^*) = 0$, which are equivalent to

$$e^{-\lambda_1 x_2} - \lambda_1 x_2 + 1 = 0\ , \quad \alpha e^{-\lambda_1 x_2} + \lambda_1 = 0.$$

Therefore, the second point of the locally E-optimal design is given by $x_2^* = t^*/\lambda_1$, where t^* is the unique solution of the equation $e^{-t} = t - 1$ (the other support point is 0) and the locally

E-optimal design is given by $\{0, x_2^*; w_1^*, w_2^*\}$, where the weights are calculated by the formula given in Remark 3.3, that is

$$w_1^* = \frac{x_2^* e^{-\lambda_1 x_2^*} + \lambda_1}{x_2^* e^{-\lambda_1 x_2^*} + \lambda_1 + \lambda_1 e^{\lambda_1 x_2^*}}\ , \quad w_2^* = \frac{\lambda_1 e^{\lambda_1 x_2^*}}{x_2^* e^{-\lambda_1 x_2^*} + \lambda_1 + \lambda_1 e^{\lambda_1 x_2^*}}.$$

j	0	1	2	3	4	5	6
$x_{2(j)}$	0.4151	0.0409	0.0689	0.0810	0.1258	0.1865	0.2769
$x_{3(j)}$	1.8605	0.5172	0.9338	1.2577	2.1534	3.6369	6.3069
$x_{4(j)}$	5.6560	4.4313	10.505	20.854	44.306	90.604	181.67
$w_{2(j)}$	0.1875	0.2050	0.6893	0.3742	-1.7292	-1.2719	7.0452
$w_{3(j)}$	0.2882	0.2243	-0.0827	-0.8709	-0.1155	2.7750	1.8101
$w_{4(j)}$	0.4501	-0.4871	-0.9587	0.2323	2.9239	-0.2510	-12.503

Table 1: *The coefficients of the Taylor expansion (4.2) for the support points and weights of the locally E-optimal design in the exponential regression model (4.1).*

Example 4.2. For the exponential regression model

$$\mathbf{E}(Y(x)) = a_1 e^{-\lambda_1 x} + a_2 e^{-\lambda_2 x} \tag{4.1}$$

corresponding to the case $k = 2$ the situation is more complicated and the solution of the locally E-optimal design problem can not be determined directly. In this case we used the Taylor expansion (3.13) for the construction of the locally E-optimal design, where the point λ_0 in this expansion was given by the vector $\lambda_0 = (1.5, 0.5)^T$. By Remark 3.3 (5) we can restrict ourselves to the case $\lambda_1 + \lambda_2 = 2$. Locally E-optimal designs for arbitrary values of $\lambda_1 + \lambda_2$ can be easily obtained by rescaling the support points of the locally E-optimal design found under the restriction $\lambda_1 + \lambda_2 = 2$, while the weights have to be recalculated using Remark 3.3(4). We consider the parameterization $\lambda_1 = 1 + z, \lambda_2 = 1 - z$ and study the dependence of the optimal design on the parameter z. Because $\lambda_1 > \lambda_2 > 0$, an admissible set of values z is the interval $(0, 1)$. We choose the center of this interval as the origin for the Taylor expansion. Table 1 contains the coefficients in the Taylor expansion for the points and weights of the locally E-optimal design, that is

$$x_i^* = x_i(z) = \sum_{j=0}^{\infty} x_{i(j)}(z - 0.5)^j,\ w_i^* = w_i(z) = \sum_{j=0}^{\infty} w_{i(j)}(z - 0.5)^j, \tag{4.2}$$

(note that $x_1^* = 0$ and $w_1^* = 1 - w_2^* - w_3^* - w_4^*$). The points and weights are depicted as a function of the parameter z in Figure 1. We observe for a broad range of the interval $(0, 1)$ only a weak dependence of the locally E-optimal design on the parameter z. Consequently, it is of some interest to investigate the robustness of the locally E-optimal design for the parameter $z = 0$, which corresponds to the vector $\lambda = (1, 1)$. This vector yields to the limiting model (3.2)

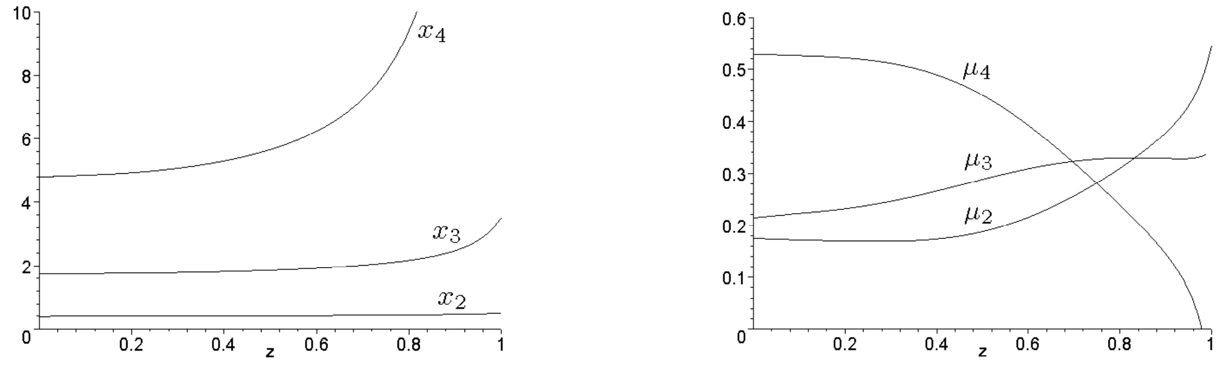

Figure 1: *Support points and weights of the locally E-optimal design $\xi_E^*(\lambda)$ in the exponential regression model (2.1), where $k=2$ and $\lambda=(1+z,1-z)^T$.*

and by Theorem 3.1 the locally E-optimal designs converge weakly to the design $\bar{\xi}^*_{e_{2k}}$, which will be denoted by $\bar{\xi}^*_E$ throughout this section. The support points of this design can obtained from the corresponding Chebyshev problem

$$\inf_{\alpha_1,\alpha_2,\alpha_3} \sup_{x\in[0,\infty)} |(1+\alpha_1 x+\alpha_2 x^2+\alpha_3 x^3)e^{-x}|$$

The solution of this problem can be found numerically using the Remez algorithm [see Studden and Tsay (1976)], i.e.

$$P_3(x) = (x^3 - 3.9855x^2 + 3.15955x - 0.27701)e^{-x}.$$

The extremal points of this polynomial are given by

$$x_1^* = 0,\ x_2^* = 0.40635,\ x_3^* = 1.75198,\ x_4^* = 4.82719.$$

and the weights of design $\bar{\xi}^*_E$ defined in Theorem 3.1 are calculated using formula (3.7), that is

$$w_1^* = 0.0767,\ w_2^* = 0.1650,\ w_3^* = 0.2164,\ w_4^* = 0.5419.$$

Some E-efficiencies

$$I_E(\xi,\lambda) = \frac{\lambda_{min}(M(\xi))}{\lambda_{min}(M(\xi_E^*(\lambda)))} \tag{4.3}$$

of the limiting design $\bar{\xi}^*_E$ are given in Table 2 and we observe that this design yields rather high efficiencies, whenever $z\in(0,0.6)$. In this table we also display the E-efficiencies of the locally D-optimal design $\xi_D^*(\lambda)$, the D-efficiencies

$$I_D(\xi,\lambda) \ = \left(\frac{\det M(\xi)}{\sup_\eta \det M(\eta)}\right)^{1/2k} \tag{4.4}$$

of the locally E-optimal design $\xi_E^*(\lambda)$ and the corresponding efficiencies of the weak limit of the locally D-optimal designs $\bar{\xi}^*_D$. We observe that the design $\bar{\xi}^*_D$ is very robust with respect to the

z	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$I_D(\bar{\xi}_D^*)$	1.00	1.00	1.00	0.99	0.98	0.95	0.90	0.80	0.61
$I_D(\xi_E^*(\lambda))$	0.75	0.74	0.75	0.75	0.78	0.82	0.87	0.90	0.89
$I_D(\bar{\xi}_E^*)$	0.74	0.74	0.76	0.77	0.78	0.79	0.78	0.72	0.58
$I_E(\bar{\xi}_E^*)$	1.00	1.00	0.98	0.94	0.85	0.72	0.58	0.45	0.33
$I_E(\xi_D^*(\lambda))$	0.66	0.66	0.66	0.67	0.70	0.74	0.79	0.82	0.80
$I_E(\bar{\xi}_D^*)$	0.65	0.64	0.62	0.59	0.56	0.52	0.47	0.41	0.33

Table 2: *Efficiencies of locally D- and E-optimal design in the exponential regression model (4.1) ($\lambda_1 = 1 + z, \lambda_2 = 1 - z$). The locally D- and E-optimal design are denoted by $\xi_D^*(\lambda)$ and $\xi_E^*(\lambda)$, respectively, while $\bar{\xi}_D^*$ and $\bar{\xi}_E^*$ denote the weak limit of the locally D- and E-optimal design as $\lambda \to (1, 1)$, respectively.*

z	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$I_1(\bar{\xi}_E^*, \lambda)$	1.00	1.00	0.98	0.93	0.84	0.69	0.53	0.40	0.27
$I_1(\bar{\xi}_D^*, \lambda)$	0.65	0.64	0.61	0.57	0.50	0.41	0.32	0.26	0.19
$I_2(\bar{\xi}_E^*, \lambda)$	0.99	0.97	0.92	0.85	0.76	0.65	0.55	0.44	0.34
$I_2(\bar{\xi}_D^*, \lambda)$	0.68	0.70	0.70	0.68	0.65	0.60	0.54	0.46	0.37
$I_3(\bar{\xi}_E^*, \lambda)$	1.00	1.00	0.98	0.93	0.85	0.73	0.56	0.38	0.20
$I_3(\bar{\xi}_D^*, \lambda)$	0.65	0.64	0.62	0.58	0.52	0.45	0.35	0.24	0.13
$I_4(\bar{\xi}_E^*, \lambda)$	1.00	0.99	0.97	0.94	0.88	0.76	0.57	0.33	0.10
$I_4(\bar{\xi}_D^*, \lambda)$	0.63	0.59	0.54	0.49	0.42	0.34	0.24	0.13	0.04

Table 3: *Efficiencies (4.5) of the designs $\bar{\xi}_D^*$ and $\bar{\xi}_E^*$ [obtained as the weak limit of the corresponding locally optimal designs as $\lambda \to (1, 1)$] for estimating the individual coefficients in the exponential regression model (4.1) ($\lambda_1 = 1 + z, \lambda_2 = 1 - z$).*

D-optimality criterion. On the other hand the D-efficiencies of the E-optimal designs $\xi_E^*(\lambda)$ and its corresponding limit $\bar{\xi}_E^*$ are substantially higher than the E-efficiencies of the designs $\xi_D^*(\lambda)$ and ξ_D^*.
We finally investigate the efficiencies

$$(4.5) \qquad I_i(\xi,\lambda) \;=\; \frac{\inf_\eta e_i^T M^{-1}(\eta) e_i}{e_i^T M^{-1}(\xi) e_i}, \quad i = 1,\dots,2k,$$

of the optimal designs $\bar{\xi}_D^*$ and $\bar{\xi}_E^*$ for the estimation of the individual parameters. These efficiencies are shown in Table 3. Note that in most cases the design $\bar{\xi}_E^*$ is substantially more efficient for estimating the individual parameters than the design $\bar{\xi}_D^*$. The design $\bar{\xi}_E^*$ can be recommended for a large range of possible values of z.

Example 4.3. For the exponential model

$$(4.6) \qquad \mathbf{E}(Y(x)) = a_1 e^{-\lambda_1 x} + a_2 e^{-\lambda_2 x} + a_3 e^{-\lambda_3 x}$$

corresponding to the case $k = 3$ the locally E-optimal designs can be calculated by similar methods. For the sake of brevity we present only the limiting designs [obtained from the locally D- and E-optimal designs if $\lambda \to (1,1,1)$] and investigate the robustness with respect to the D- and E-optimality criterion. The support points of the e_6-optimal designs in the heteroscedastic polynomial regression model (3.2) (with $\gamma = 1$) can be found as the extremal points of the Chebyshev function

$$P_5(x) = (x^5 - 11.7538x^4 + 42.8513x^3 - 55.6461x^2 + 21.6271x - 1.1184)e^{-x}$$

which are given by

$$x_1^* = 0,\ x_2^* = 0.2446,\ x_3^* = 1.0031,\ x_4^* = 2.3663,\ x_5^* = 4.5744,\ x_6^* = 8.5654.$$

For the weights of the limiting design $\bar{\xi}_E^* := \bar{\xi}_{e_6}^*$ we obtain from the results of Section 3

$$w_1^* = 0.0492,\ w_2^* = 0.1007,\ w_3^* = 0.1089,\ w_4^* = 0.1272,\ w_5^* = 0.1740,\ w_6^* = 0.4401.$$

For the investigation of the robustness properties of this design we note that by Remark 3.3(5) we can restrict ourselves to the case $\lambda_1 + \lambda_2 + \lambda_3 = 3$. The support points in the general case are obtained by a rescaling, while the weights have to be recalculated using Remark 3.3(4). For the sake of brevity we do not present the locally E-optimal designs, but restrict ourselves to some efficiency considerations. For this we introduce the parameterization $\lambda_1 = 1+u+v, \lambda_2 = 1-u, \lambda_3 = 1-v$, where the restriction $\lambda_1 > \lambda_2 > \lambda_3 > 0$ yields

$$u < v, v < 1\ , \quad u > -v/2.$$

In Table 4 we show the E-efficiency defined in (4.3) of the design $\bar{\xi}_E^*$, which is the weak limit of the locally E-optimal design $\xi_E^*(\lambda)$ as $\lambda \to (1,1,1)$ (see Theorem 3.1). Two conclusions can be drawn from our numerical results. On the one hand we observe that the optimal design

u	0	0	0	-0.2	-0.2	0.2	0.2	0.4	0.4	0.7
v	0.2	0.5	0.8	0.6	0.8	0.3	0.8	0.5	0.8	0.8
$I_D(\bar{\xi}_D^*)$	1.00	0.98	0.83	0.97	0.86	0.99	0.79	0.92	0.70	0.48
$I_D(\xi_E^*(\lambda))$	0.78	0.85	0.90	0.86	0.90	0.66	0.90	0.61	0.86	0.50
$I_D(\bar{\xi}_E^*)$	0.75	0.78	0.74	0.78	0.75	0.77	0.71	0.78	0.65	0.47
$I_E(\bar{\xi}_E^*)$	0.98	0.76	0.43	0.71	0.48	0.93	0.36	0.53	0.19	0.02
$I_E(\xi_D^*(\lambda))$	0.65	0.73	0.79	0.74	0.79	0.55	0.79	0.52	0.74	0.48
$I_E(\bar{\xi}_D^*)$	0.63	0.57	0.37	0.53	0.40	0.46	0.31	0.23	0.09	0.01

Table 4: *Efficiencies of locally D-, E-optimal designs and of the corresponding limits $\bar{\xi}_D^*$ and $\bar{\xi}_E^*$ (obtained as the weak limit of the corresponding locally optimal designs as $\lambda \to (1,1,1)$) in the exponential regression model (4.6) ($\lambda_1 = 1 + u + v, \lambda_2 = 1 - u, \lambda = 1 - v$).*

$\bar{\xi}_E^*$ is robust in a neighbourhood of the point $(1,1,1)$. On the other hand we see that the locally E-optimal design $\xi_E^*(\lambda)$ is also robust if the nonlinear parameters $\lambda_1, \lambda_2, \lambda_3$ differ not too substantially (i.e. the "true" parameter is contained in a moderate neighbourhood of the point $(1,1,1)$). The table also contains the D-efficiencies of the E-optimal designs defined in (4.4) and the E-efficiencies of the locally D-optimal design $\xi_D^*(\lambda)$ and its corresponding weak limit as $\lambda \to (1,1,1)$. Again the D-efficiencies of the E-optimal designs are higher than the E-efficiencies of the D-optimal designs.

We finally compare briefly the limits of the locally E- and D-optimal designs if $\lambda \to (1,1,1)$ with respect to the criterion of estimating the individual coefficients in the exponential regression model (4.6). In Table 5 we show the efficiencies of these designs for estimating the parameters in $a_1, b_1, a_2, b_2, a_3, b_3$ in the model (4.6). We observe that in most cases the limit of the locally E-optimal designs $\bar{\xi}_E^*$ yields substantially larger efficiencies than the corresponding limit of the locally D-optimal design $\bar{\xi}_D^*$. Moreover this design is robust for many values of the parameter (u, v).

5 Appendix

5.1 Proof of Theorem 3.1.

Using the notation $\delta_j = r_j \delta$ and observing the approximation in (3.1) we obtain from the Taylor expansion

$$e^{-(\gamma - r_j\delta)x} = e^{-\gamma x}(1 + \sum_{i=1}^{2k-1} \delta_j^i x^i / i!) + o(\delta^{2k-1}) \ , \quad (j = 1, \ldots, k)$$

the representation

$$f(x) = L\bar{f}(x)e^{-\gamma x} + H(\delta),$$

u	0	0	0	-0.2	-0.2	0.2	0.2	0.4	0.4	0.7
v	0.2	0.5	0.8	0.6	0.8	0.3	0.8	0.5	0.8	0.8
$I_1(\bar{\xi}_E^*)$	0.98	0.77	0.43	0.71	0.48	0.86	0.35	0.52	0.26	0.11
$I_1(\bar{\xi}_D^*)$	0.63	0.56	0.36	0.53	0.40	0.59	0.30	0.41	0.22	0.10
$I_2(\bar{\xi}_E^*)$	0.97	0.74	0.43	0.70	0.48	0.80	0.37	0.49	0.29	0.19
$I_2(\bar{\xi}_D^*)$	0.65	0.59	0.42	0.55	0.43	0.63	0.38	0.48	0.33	0.23
$I_3(\bar{\xi}_E^*)$	0.90	0.73	0.43	0.71	0.48	0.93	0.38	0.53	0.16	0.02
$I_3(\bar{\xi}_D^*)$	0.71	0.59	0.38	0.53	0.40	0.46	0.47	0.23	0.04	0.01
$I_4(\bar{\xi}_E^*)$	0.99	0.82	0.41	0.73	0.47	0.93	0.31	0.53	0.17	0.02
$I_4(\bar{\xi}_D^*)$	0.60	0.50	0.29	0.51	0.36	0.48	0.20	0.25	0.10	0.01
$I_5(\bar{\xi}_E^*)$	0.99	0.85	0.30	0.76	0.35	0.93	0.21	0.53	0.11	0.02
$I_5(\bar{\xi}_D^*)$	0.55	0.39	0.12	0.33	0.14	0.46	0.09	0.23	0.05	0.01
$I_6(\bar{\xi}_E^*)$	0.99	0.84	0.26	0.75	0.31	0.93	0.18	0.53	0.09	0.02
$I_6(\bar{\xi}_D^*)$	0.53	0.34	0.08	0.27	0.10	0.45	0.06	0.22	0.03	0.01

Table 5: *Efficiencies (4.5) of the designs $\bar{\xi}_D^*$ and $\bar{\xi}_E^*$ (obtained as the weak limit of the corresponding locally optimal designs as $\lambda \to (1,1,1)$) for estimating the individual coefficients in the exponential regression model (4.6) ($\lambda_1 = 1+u+v, \lambda_2 = 1-u, \lambda = 1-v$).*

where the vectors f and $\bar{f}$ are defined in (2.3) and (3.4), respectively, the remainder term is of order

$$H(\delta) = \big(o(\delta^{2k-1}), o(\delta^{2k-2}), \ldots, o(\delta^{2k-1}), o(\delta^{2k-2})\big)^T$$

and the matrix L is given by

$$L = \begin{pmatrix} 1 & \delta_1 & \frac{\delta_1^2}{2!} & \frac{\delta_1^3}{3!} & \ldots & \frac{\delta_1^{2k-1}}{(2k-1)!} \\ 0 & 1 & \delta_1 & \frac{\delta_1^2}{2!} & \ldots & \frac{\delta_1^{2k-2}}{(2k-2)!} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \delta_k & \frac{\delta_k^2}{2!} & \frac{\delta_k^3}{3!} & \ldots & \frac{\delta_k^{2k-1}}{(2k-1)!} \\ 0 & 1 & \delta_k & \frac{\delta_k^2}{2!} & \ldots & \frac{\delta_k^{2k-2}}{(2k-2)!} \end{pmatrix}.$$

Consequently the information matrix in the general exponential regression model (2.1) satisfies

$$M^{-1}(\xi) = L^{-1^T} \bar{M}^{-1}(\xi) L^{-1} + o(\delta^{4k-2}) ,$$

where $\bar{M}(\xi)$ is the information matrix in the heteroscedastic polynomial regression model (3.2) defined by (3.3). It can be shown by a straightforward but tedious calculation [see the technical report of Dette, Melas and Pepelyshev (2002)] that for small δ

$$\delta^{2k-1} L^{-1} = (2k-1)!(\mathbf{0} \vdots l)^T + o(1),$$

where $\mathbf{0}$ is $2k \times (2k-1)$ matrix with all entries equals 0 and the vector l is defined by (3.5) in Theorem 3.1. This yields for the information matrix of the design ξ

$$\delta^{4k-2} M^{-1}(\xi) = ((2k-1)!)^2 \left(\bar{M}^{-1}(\xi)\right)_{2k,2k} l\, l^T + o(1).$$

Therefore, if δ is sufficiently small, it follows that maximal eigenvalue of the matrix $M^{-1}(\xi)$ is simple.
For a proof of the second part of Theorem 3.1 we note the locally E-optimal design $\xi_E^*(\lambda)$ is defined by

$$\xi_E^*(\lambda) = \arg\min_{\xi} \max_{c, c^T c=1} c^T M^{-1}(\xi) c\ .$$

If $\delta \to 0$ it therefore follows from the arguments of previous paragraph that this design converges weakly to the design $\bar{\xi}^*_{e_{2k}}$, which minimizes a function

$$\max_{c,\ c^T c=1} (c^T l)^2 e_{2k}^T \bar{M}^{-1}(\xi) e_{2k},$$

Finally, the proof of the third part of Theorem 3.1 can be obtained by similar arguments and is left to the reader. □

5.2 Proof of Theorem 3.3.

In Section 3 we have already shown that the function $\Theta^*(\lambda)$ as solution of (3.10) is uniquely determined. In this paragraph we prove that the Jacobi matrix

$$G = G(\lambda) = \left(\frac{\partial^2 \Lambda}{\partial\theta_i \partial\theta_i}(\Theta^*(\lambda), \lambda)\right)_{i,j=1}^{3m-3}$$

is nonsingular. It then follows from the Implicit Function Theorem (see Gunning, Rossi, 1965) that the function $\Lambda(\Theta, \lambda)$ is real analytic. For this purpose we note that a direct calculation shows

$$\begin{array}{l}
q^T q \frac{\partial^2 \Lambda}{\partial w \partial w}(\Theta^*(\lambda), \lambda) = 0, \\
q^T q \frac{\partial^2 \Lambda}{\partial x \partial w}(\Theta^*(\lambda), \lambda) = 0, \\
q^T q \frac{\partial^2 \Lambda}{\partial x \partial x}(\Theta^*(\lambda), \lambda) = E = \operatorname{diag}\{(q^T f(x_i^*))^2 w_i\}_{i=2,\dots,2k}, \\
q^T q \frac{\partial^2 \Lambda}{\partial q_- \partial q_-}(\Theta^*(\lambda), \lambda) = (M(\xi^*) - \Lambda I_{2k})_-, \\
q^T q \frac{\partial^2 \Lambda}{\partial q_- \partial x}(\Theta^*(\lambda), \lambda) = B_1^T, \\
q^T q \frac{\partial^2 \Lambda}{\partial q_- \partial w}, (\Theta^*(\lambda), \lambda) = B_2^T\ ,
\end{array}$$

where the matrices B_1 and B_2 are defined by

$$\begin{aligned}
B_1^T &= 2\left(q^T f(x_2^*) w_2 f_-'(x_2^*) \vdots \dots \vdots q^T f(x_{2k}^*) w_{2k} f_-'(x_{2k}^*)\right), \\
B_2^T &= 2\left(q^T f(x_2^*)\, f_-(x_2^*) - q^T f(0)\, f_-(0) \vdots \dots \vdots q^T f(x_{2k}^*)\, f_-(x_{2k}^*) - q^T f(0)\, f_-(0)\right),
\end{aligned}$$

respectively, and $w = (w_2, \dots, w_{2k})^T$ and $x = (x_2, \dots, x_{2k})^T$. Consequently the Jacobi matrix of the system (3.10) has the structure

$$G = \frac{1}{q^T q} \begin{pmatrix} D & B_1^T & B_2^T \\ B_1 & E & 0 \\ B_2 & 0 & 0 \end{pmatrix}. \tag{A.1}$$

Because $(p^{*T} f(x_i^*))^2 = \lambda_{min}$ we obtain $q^{*T} f(x_i^*) = (-1)^i \tilde{c}$ $(i = 1, \dots, 2k)$ for some constant $\tilde{c}$, and the matrices B_1 and B_2 can be rewritten as

$$\begin{aligned} B_1^T &= 2\tilde{c} \left(w_2 f_-'(x_2^*) \vdots - w_3 f_-'(x_3^*) \vdots w_4 f_-'(x_4^*) \vdots \dots \vdots - w_{2k-1} f_-'(x_{2k-1}^*) \vdots w_{2k} f_-'(x_{2k}^*) \right), \\ B_2^T &= 2\tilde{c} \left(f_-(0) + f_-(x_2^*) \vdots f_-(0) - f_-(x_3^*) \vdots \dots \vdots f_-(0) - f_-(x_{2k-1}^*) \vdots f_-(0) + f_-(x_{2k}^*) \right). \end{aligned}$$

In the following we study some properties of the blocks of the matrix G defined in (A.1). Note that the matrix D in the upper left block of G is nonnegative definite. This follows from

$$\min_v \frac{v^T M(\xi^*) v}{v^T v} \le \min_u \frac{u^T M_-(\xi^*) u}{u^T u}$$

and the inequality

$$\lambda_{min}(M_-(\xi^*)) \ge \lambda_{min}(M(\xi^*)) = \Lambda(\Theta^*(\lambda), \lambda) ,$$

where M_- denotes the matrix obtained from M deleting the first row and column. Thus we obtain for any vector $u \in \mathbb{R}^{2k-1}$

$$u^T D u = u^T M_-(\xi^*) u - \Lambda u^T u \ge u^T u (\lambda_{min}(M_-(\xi^*)) - \Lambda) \ge 0,$$

which shows that the matrix D is nonnegative definite. The diagonal matrix E is negative definite because all it's diagonal elements are negative. This property follows from the the equivalence Theorem 2.1, which shows that the function $(q^{*T} f(x))^2$ is concave in a neighbourhood of every point x_i^*, $i = 2, \dots, 2k$. Moreover, the matrices B_1 and B_2 are of full rank and we obtain from the formula for the determinant of block matrix that

$$\det G = -\det E \, \det(D - B_1^T E^{-1} B_1) \, \det(B_2^T (D - B_1^T E^{-1} B_1)^{-1} B_2).$$

Since each determinant is nonzero, the matrix G is nonsingular, which completes the proof of the theorem. □

Acknowledgements: The support of the Deutsche Forschungsgemeinschaft (SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen, Teilprojekt A2, Sachbeihilfe De 502/18-1, 436 RUS 113/712/0-1) is gratefully acknowledged.

References

M. Becka, W. Urfer (1996). Statistical aspects of inhalation toxicokinetics. Environmental and Ecological Statistics 3, 51-64.

M. Becka, H.M. Bolt, W. Urfer (1993). Statistical evaluation of toxicokinetic data. Environmetrics 4, 311-322.

K. Chaloner, I. Verdinelli (1995). Bayesian experimental design: a review. Statistical Science 10, 273-304.

H. Chernoff (1953). Locally optimal designs for estimating parameters. Ann. Math. Statist. 24, 586-602.

H. Dette, W.J. Studden (1993). Geometry of E-optimality. Ann. Statist. 21, 416-433.

H. Dette, L. Haines (1994). E-optimal designs for linear and nonlinear models with two parameters. Biometrika 81, 739–754.

H. Dette, L. Haines and L.A. Imhof (1999). Optimal designs for rational models and weighted polynomial regression. Ann. Statist. 27, 1272-1293.

H. Dette, V.B. Melas, A. Pepelyshev (2004). Optimal designs for estimating individual coefficients in polynomial regression – a functional approach. Journal of Stat. Plann. and Inf. 118, 201-219.

H. Dette, V.B. Melas, A. Pepelyshev (2002). Optimal designs for a class of nonlinear regression models. Preprint, Ruhr-Universität Bochum.
http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm

H. Dette, W.K. Wong (1999). E-optimal designs for the Michaelis Menten model. Statistics & Probability Letters 44, 405-408.

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

I. Ford, B. Torsney, C.F.J. Wu (1992). The use of a canonical form in the construction of locally optimal designs for non-linear problems. J.R. Stat. Soc., Ser. B 54, 569-583.

I. Ford, S.D. Silvey (1980). A sequentially constructed design for estimating a nonlinear parametric function. Biometrika 67, 381-388.

R.C. Gunning , H. Rossi (1965). Analytical functions of several complex variables. Prentice-Hall, Inc., NewYork.

Z. He, W.J. Studden and D. Sun (1996). Optimal designs for rational models. Ann. Statist. 24, 2128-2142.

B. Heiligers (1994). E-optimal designs in weighted polynomial regression. Ann. Statist. 22, 917-929.

S. Karlin, W.J. Studden (1966). Tchebycheff Systems: with Applications in Analysis and Statistics. Interscience, New York.

V.B. Melas (1978). Optimal designs for exponential regression. Math. Operationsforsch. Statist., Ser. Statistics 9, 45-59.

V.B. Melas (1982) A duality theorem and E-optimality. Industrial Laboratory, 48, 275-296 (translated from Russian).

V.B. Melas (2001). Analytical properties of locally D-optimal designs for rational models. In: MODA 6 – Advances in Model-Oriented Design and Analysis (eds. A.C. Atkinson, P. Hackel, W.G. Müller). Physica Verlag, Heidelberg, 201-210.

L. Pronzato, E. Walter (1985). Robust experimental design via stochastic approximation. Math. Biosciences 75, 103-120.

F. Pukelsheim, B. Torsney (1991). Optimal weights for experimental designs on linearly independent support points. Ann. Statist. 19, 1614-1625.

F. Pukelsheim, S. Rieder (1992). Efficient rounding of approximate designs. Biometrika 79, 763-770.

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

D.A. Ratkowsky (1990). Handbook of Nonlinear Regression Models. Dekker, New York.

D.A. Ratkowsky (1983). Nonlinear Regression. Dekker, New York.

G.A.J. Seber, C.J. Wild (1989). Nonlinear Regression. Wiley, New York.

S.D. Silvey (1980). Optimum design. Chapman and Hall, London.

W.J. Studden, J.Y. Tsay (1976). Remez's procedure for finding optimal designs. Ann. Statist. 4, 1271-1279.

C.F.J. Wu (1985). Efficient sequential designs with binary data. J. Amer. Stat. Assoc. 80, 974-984.