AISM (2006) 58: 407–426
DOI 10.1007/s10463-006-0031-2

Holger Dette · Viatcheslav B. Melas
Andrey Pepelyshev

Local c- and E-optimal designs for exponential regression models

Received: 1 December 2003 / Revised: 6 May 2005 / Published online: 11 May 2006

Abstract In this paper we investigate local E- and c-optimal designs for exponential regression models of the form $\sum_{i=1}^{k} a_i \exp(-\mu_i x)$. We establish a numerical method for the construction of efficient and local optimal designs, which is based on two results. First, we consider for fixed k the limit $\mu_i \to \gamma$ $(i = 1, \dots, k)$ 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 local optimal designs in the exponential regression model are analytic functions of the nonlinear parameters $\mu_1, \dots, \mu_k$. This result is used for the numerical calculation of the local E-optimal designs by means of a Taylor expansion for any vector $(\mu_1, \dots, \mu_k)$. 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.

Keywords E-Optimal design · c-Optimal design · Exponential models · Local optimal designs · Chebyshev systems · Heteroscedastic polynomial regression

AMS Subject classification 62K05 · 41A50

H. Dette (✉)
Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany
E-mail: holger.dette@ruhr-uni-bochum.de

V. B. Melas · A. Pepelyshev
St. Petersburg State University, Department of Mathematics, St. Petersburg, Russia
E-mail: v.melas@pobox.spbu.ru
E-mail: andrey@ap7236.spb.edu

1 Introduction

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

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

(see, for example, Alvarez, Virto, Raso and Condon, 2003; Becka and Urfer, 1996; Becka, Bolt and Urfer, 1993; Landaw and DiStefano, 1984 among many others). 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; Ford, Torsney and Wu, 1992; Melas, 1978). Local optimal designs depend on an initial guess for the unknown parameter, but are the basis for all advanced design strategies, (see Chaloner and Verdinelli, 1995; Ford & Silvey, 1980; Han and Chaloner, 2003; Pronzato and Walter, 1985; 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; Dette and Wong, 1999; among others, who gave some results for models with two parameters).

It is the purpose of the present paper to study local c-optimal and E-optimal designs for the nonlinear regression model (1). For this purpose we prove two main results. First, we show that in the case $\mu_i \to \gamma$ $(i = 1, \ldots, k)$, where k is fixed, the local optimal designs for the model (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). The main reason for consideration of the limiting model is that the optimal designs in the limiting model can easily be calculated. Moreover, the resulting optimal designs are usually very efficient in the general model (1), which will be confirmed by our numerical results in Sect. 4. Secondly, it is proved that the support points and weights of the local optimal designs in the exponential regression model are analytic functions of the nonlinear parameters $\mu_1, \ldots, \mu_k$. 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 Sect. 2 we introduce the necessary notation, while the main results are stated in Sect. 3. In Sect. 4

we illustrate our method considering several examples and compare local D- and E-optimal designs. Finally all technical details are deferred to an Appendix (see Sect. 5).

2 Preliminaries

Consider the common exponential regression model with homoscedastic error

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

where the explanatory variable x varies in the experimental domain $\mathcal{X} = [b, +\infty)$ with $b \in \mathbb{R}$, $\beta^{\mathrm{T}} = (a_1, \mu_1, a_2, \ldots, \mu_k)$ 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 $0 < \mu_1 < \mu_2 < \cdots < \mu_k$. An approximate design ξ is a probability measure

$$\xi = \begin{pmatrix} x_1 & \ldots & x_n \\ w_1 & \ldots & w_n \end{pmatrix} \tag{3}$$

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 the total number of 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, 1992). If $n \geq 2k$, $w_i > 0$, $i = 1, \ldots, n$, it is well known that under some additional assumptions of regularity (see e.g. Jennrich, 1969) the nonlinear least squares estimator $\hat{\beta}$ for the parameter β in model (2) is asymptotically unbiased with covariance matrix satisfying

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

where

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

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

$$f(x) = \frac{\partial \eta(x, \beta)}{\partial \beta} = (\mathrm{e}^{-\mu_1 x}, -a_1 x \mathrm{e}^{-\mu_1 x}, \ldots, \mathrm{e}^{-\mu_k x}, -a_k x \mathrm{e}^{-\mu_k x})^{\mathrm{T}} \tag{4}$$

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

$$M(\xi) = \sum_{i=1}^{n} f(x_i) f^{\mathrm{T}}(x_i) w_i. \tag{5}$$

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. Pukelsheim, 1993; Silvey, 1980). In this paper we restrict ourselves to three well known optimality criteria. Following Chernoff (1953) we call a design ξ local D-optimal in the exponential regression model (2) 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. Local D-optimal designs in various non-linear regression models have been discussed by numerous authors (see e.g. Dette et al., 1999; He et al., 1996; Melas, 1978 among many others). D-optimal designs for the regression model (1) have been studied by Melas (1978, 2001), while some D- and c-optimal designs for other exponential regression models can be found in Han and Chaloner (2003), Fang and Wiens (2004).

Note that an ellipsoid is usually not used as a confidence region for the vector of parameters, because it is difficult to handle for practitioners, which usually prefer rectangular regions obtained from Bonferroni's method. Good designs for these confidence regions minimize the variance of the parameter estimates, which can be reflected by the c-optimality criterion. To be precise we call for a given vector $c \in \mathbb{R}^{2k}$ a design ξ local c-optimal if $c \in \text{Range}(M(\xi))$ and ξ minimizes $c^T M^-(\xi)c$. This 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 local e_i-optimal designs are also called optimal designs for estimating the ith coefficient ($i = 1, \dots, 2k$). Local c-optimal designs for nonlinear models with two parameters have been studied by Ford et al. (1992) among others. Finally we consider the E-optimality criterion, which determines the design such that the minimal eigenvalue, say $\lambda_{\min}(M(\xi))$, of the information matrix $M(\xi)$ is maximal. This corresponds to 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 $p^T p = 1$. In this paper we will demonstrate theoretically (see Theorem 2) and empirically (see the examples in Sect. 4) that in the exponential regression model (1) E-optimal designs usually behave substantially more reliably with respect to minimzation of the variances of the parameter estimates than do D-optimal designs. However, the problem of determining E-optimal designs is substantially harder than the D-optimal design problem.

Some local 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 local E- and c-optimal designs for the exponential regression model (1). We also compare these designs with the corresponding local D-optimal designs for the exponential model, which have been studied by Melas (1978).

Note that in contrast to local D-optimal designs the local E-optimal designs in model (2) depend on all parameters $a_1, \dots a_k, \mu_1, \dots, \mu_k$. The information matrix in the nonlinear regression model (1) is given by $K_a^{-1} M(\xi, e, \mu) K_a^{-1}$, where $e = (1, \dots, 1)^T$ and the matrix $K_a \in \mathbb{R}^{k \times k}$ is defined by

$$K_a = \text{diag}\Big(1, \frac{1}{a_1}, 1, \dots, 1, \frac{1}{a_k}\Big). \tag{6}$$

Consequently, a local E-optimal design problem in the model (1) with respect to the parameter $(a_1, \dots, a_k, \mu_1, \dots, \mu_k)^T$ corresponds to the problem of maximizing

$\lambda_{\min}(K_a^{-1} M(\xi, e, \mu) K_a^{-1})$. For the sake of transparency we will mainly concentrate on the case $a = (a_1, \dots, a_k) = e = (1, \dots 1)^{\mathrm{T}}$. The general case is treated exactly in the same way. In what follows we will only reflect the dependence of the information matrix in our notation, if it is not clear from the context. In particular, we do not reflect the dependence of the local E-optimal design on the parameter $a = e$. 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 1 *A design ξ^* is E-optimal if and only if there exists a nonnegative definite matrix A^* such that* $\operatorname{tr} A^* = 1$ *and*

$$\max_{x \in \mathcal{X}} f^{\mathrm{T}}(x) A^* f(x) \leq \lambda_{\min}(M(\xi^*)). \tag{7}$$

Moreover, we have equality in (7) for any support point of ξ^*, and the matrix A^* can be represented as

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

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 local c- and E-optimal designs in the exponential regression model (2). In order to indicate the dependence of the optimal designs on the nonlinear parameters in model (2) we denote the local c- and E-optimal design by $\xi_c^*(\mu)$ and $\xi_E^*(\mu)$, respectively. We begin with an investigation of the behaviour of the local optimal designs if the vector of nonlinear parameters $\mu = (\mu_1, \dots, \mu_k)^{\mathrm{T}}$ is contained in a neighbourhood of a point $\gamma(1, \dots, 1)^{\mathrm{T}}$, where $\gamma > 0$ is an arbitrary parameter. The information matrix (5) of any design becomes singular as $\mu \to \gamma(1, \dots, 1)^{\mathrm{T}}$. However, we will show that the corresponding local optimal designs are still weakly convergent, where the limiting measure has $2k$ support points.

To be precise let

$$\mu_i = \gamma - r_i \delta, \quad i = 1, \dots, k \tag{8}$$

where $\delta > 0$ and $r_1, \dots, r_k \in \mathbb{R} \setminus \{0\}$ are arbitrary fixed numbers such that $r_k > r_{k-1} > \dots > r_1$. If δ is small, local c- and E-optimal designs in the exponential regression model (2) are closely related to optimal designs in the heteroscedastic polynomial regression model

$$\mathbf{E}(Y(x)) = \sum_{i=1}^{2k} a_i x^{i-1}, \quad \mathbf{V}(Y(x)) = \exp(2\gamma x), \quad x \in [b, \infty) \tag{9}$$

where $\gamma > 0$ is assumed to be known. Note that for a design of the form (3) the information matrix in this model is given by

$$\bar{M}(\xi) = \sum_{i=1}^{n} \mathrm{e}^{-2\gamma x_i} \bar{f}(x_i) \bar{f}^{\mathrm{T}}(x_i) w_i, \tag{10}$$

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

$$\bar{f}(x) = (1, x, \ldots, x^{2k-1})^{\mathrm{T}}. \tag{11}$$

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, i.e. the design which minimizes $e_{2k}^{\mathrm{T}} \bar{M}^{-1}(\xi) e_{2k}$ for $e_{2k} = (0, \ldots 0, 1)^{\mathrm{T}} \in \mathbb{R}^{2k}$, appears as a weak limit of the local c- and E-optimal design $\xi_c^*(\mu)$ and $\xi_E^*(\mu)$ in the model (2). The proof is complicated and therefore deferred to the Appendix.

Theorem 2 *(1) For any design with at least* $2k$ *support points and* $\gamma > 0$ *there exists a neighbourhood* Ω_γ *of the point* $\gamma(1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^k$ *such that for any vector* $\mu = (\mu_1, \ldots, \mu_k)^{\mathrm{T}} \in \Omega_\gamma$ *the minimal eigenvalue of information matrix* $M(\xi)$ *in (5) is simple.*

(2) If condition (8) is satisfied and $\delta \to 0$*, then the local* E*-optimal design* $\xi_E^*(\mu)$ *in the exponential regression model (2) converges weakly to the* e_{2k}*-optimal design* $\bar{\xi}_{e_{2k}}$ *in the heteroscedastic polynomial regression model (9).*

(3) Assume that condition (8) is satisfied and define a vector $l = (l_1, \ldots, l_{2k})^{\mathrm{T}}$ *with* $l_{2i} = 0$, $(i = 1, \ldots, k)$,

$$l_{2i-1} = -\prod_{j \neq i} (r_i - r_j)^2 \sum_{j \neq i} \frac{2}{r_i - r_j}, \quad i = 1, \ldots, k. \tag{12}$$

If $l^T c \neq 0$ *and* $\delta \to 0$ *then the local* c*-optimal design* $\xi_c^*(\mu)$ *in the exponential regression model (2) converges weakly to the* e_{2k}*-optimal design* $\bar{\xi}_{e_{2k}}$ *in the heteroscedastic polynomial regression model (9).*

Remark 1 (a) Note that in the case, where all parameters μ_i share a common value, say γ, the parameters a_i in the model (1) are not identifiable and this model reduces to the model $\beta \exp(-\gamma x)$, for which a two-point design is optimal (see also our Example 1). Theorem 2 shows that the calculation of the design and the consideration of the limit in (8) with $\delta \to 0$ cannot be interchanged. Note that for a fixed design ξ the information matrix $M(\xi)$ becomes singular if $\delta \to 0$ in (8). On the other hand the local optimal design ξ_E^* depends on the parameter μ, say $\xi_E^*(\mu)$, and this measure converges weakly to a $2k$-point design.

(b) Note that Theorem 2 shows that as $\delta \to 0$ the local E-optimal and e_i-optimal designs have the same limiting design, namely the design $\bar{\xi}_{e_{2k}}$ is optimal for estimating the highest coefficient in model (9). Roughly speaking this indicates that E-optimal designs behave substantially better with respect to the estimation of the individual parameters than D-optimal designs. This is also confirmed by our numerical results in Sect. 4.

(c) E-optimal designs are easy to find if the minimum eigenvalue of the E-optimal information matrix is simple (see Dette and Studden, 1993; Heiligers, 1994 among others). In general this property cannot be proved for the model (1). However, if all components of the vector of nonlinear parameters $\mu_1, \dots, \mu_k$ converge to the same value, then, by Theorem 2, the information matrix of any design with $2k$ support points converges to a matrix with a simple minimum eigenvalue. In other words, if the components of the vector of nonlinear parameters are close, the minimum eigenvalue of the information matrix of the E-optimal design has multiplicity 1. As a consequence the E-optimal design can be found easily in such cases. It is also interesting to note that the limiting design is usually rather efficient for a broad range of values of the nonlinear parameters.

Remark 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 (9) has $2k$ support points, say

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

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

$$\sup_{x \in [b,\infty)} |p^*(x)| = \min_{\alpha_1,\dots,\alpha_{2k-1}} \sup_{x \in [b,\infty)} \exp(-\gamma x) \left| 1 + \sum_{i=1}^{2k-1} \alpha_i x^i \right|. \tag{13}$$

Moreover, also the weights $w_1^*(\gamma), \dots, w_{2k}^*(\gamma)$ of the e_{2k}-optimal design $\bar{\xi}_{e_{2k}}(\gamma)$ in model (9) can be obtained explicitly, i.e.

$$w^*(\gamma) = (w_1^*(\gamma), \dots, w_1^*(\gamma))^{\mathrm{T}} = \frac{J \bar{F}^{-1} e_{2k}}{\mathbf{1}_{2k} J \bar{F}^{-1} e_{2k}}, \tag{14}$$

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

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

$J = \mathrm{diag}(1, -1, 1, \dots, 1, -1)$, respectively, $\mathbf{1}_{2k} = (1, \dots, 1)^{\mathrm{T}} \in \mathbb{R}^{2k}$ and the vector $\bar{f}(x)$ is defined in (11) (see Pukelsheim and Torsney, 1991).

Remark 3 Let Ω denote the set of all vectors $\mu = (\mu_1, \dots, \mu_k)^{\mathrm{T}} \in \mathbb{R}^k$ with $\mu_i \neq \mu_j$, $i \neq j$, $\mu_i > 0$, $i = 1, \dots, k$, such that the minimum eigenvalue of the information matrix of the local E-optimal design (with respect to the vector μ) is simple. The following properties of local 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 local E-optimal designs substantially.

1. For any $\mu \in \Omega$ the local E-optimal design for the exponential regression model (2) (with respect to the parameter μ) is unique.
2. For any $\mu \in \Omega$ the support points of the local E-optimal design for the exponential regression model (2) (with respect to the parameter μ) do not depend on the parameters $a_1, \dots, a_k$.

3. For any $\mu \in \Omega$ the local E-optimal design for the exponential regression model (2) (with respect to the parameter μ) has $2k$ support points; moreover the point b is always a support point of the local E-optimal design. The support points of the E-optimal design are the extremal points of the Chebyshev function $p^{\mathrm{T}} f(x)$, where p is an eigenvector corresponding to the minimal eigenvalue of the information matrix $M(\xi_E^*(\mu))$.
4. For any $\mu \in \Omega$ the weights of the local E-optimal design for the exponential regression model (2) (with respect to the parameter μ) are given by

$$w^* = \frac{JF^{-1}c}{c^{\mathrm{T}}c}, \qquad (15)$$

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

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

and $x_1^*, \ldots, x_{2k}^*$ denote the support points of the local E-optimal design.
5. If $\mu \in \Omega$, let $x_{1;b}^*(\mu), \ldots, x_{2k;b}^*(\mu)$ denote the support points of the local E-optimal design for the exponential regression model (2) with design space $\mathcal{X} = [b, +\infty)$. Then $x_{1;0}^*(\mu) \equiv 0$,

$$x_{i;b}^*(\mu) = x_{i;0}^*(\mu) + b, \quad i = 2, \ldots, 2k.$$
$$x_{i;0}^*(\nu\mu) = x_{i;0}^*(\mu)/\nu, \quad i = 2, \ldots, 2k$$

for any $\nu > 0$.

We now study some analytical properties of local E-optimal designs for the exponential regression model (2). Theorem 2 indicates that the structure of the local 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 $p^{\mathrm{T}} f(x)$, where p is the eigenvector corresponding to the minimal eigenvalue of the information matrix $M(\xi_E^*(\mu))$. If the multiplicity is greater than 1 then the problem of constructing E-optimal designs is more complex. Observing Remark 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 $p = (p_1, \ldots, p_{2k})^{\mathrm{T}}$ is an eigenvector corresponding to the minimal eigenvalue of $M(\xi)$ we define a vector

$$\Theta = (\theta_1, \ldots, \theta_{6k-3})^{\mathrm{T}} = (q_2, \ldots, q_{2k}, x_2, \ldots, x_{2k}, w_2, \ldots, w_{2k})^{\mathrm{T}}, \qquad (16)$$

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 $x_1 = 0$, $w_1 = 1 - w_2 - \cdots - w_{2k}$) and $q = (1, q_2, \ldots, q_{2k})^{\mathrm{T}} = p/p_1$ 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 (16). Recall the definition of the set Ω in Remark 3. For each vector $\mu \in \Omega$ the minimum eigenvalue of the information matrix of a local E-optimal design $\xi_E^*(\mu)$ (for the parameter μ) has multiplicity 1 and for $\mu \in \Omega$ let

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

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

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

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

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

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

$$\frac{\partial \Lambda}{\partial \theta_i}(\Theta, \mu) = 0, \quad i = 1, \dots, 6k - 3, \tag{17}$$

and a straightforward differentiation shows that this system is equivalent to

$$\begin{cases} (M(\xi)q)_- - \Lambda(\Theta, \mu) q_- = 0, \\ 2q^T f(x_i) q^{\mathrm{T}} f'(x_i) w_i = 0, \quad i = 2, \dots, 2k, \\ (q^T f(x_i))^2 - (q^{\mathrm{T}} f(0))^2 = 0, \quad i = 2, \dots, 2k, \end{cases} \tag{18}$$

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

$$\begin{cases} M(\xi) p = \Lambda(\Theta, \mu) p, \\ p^{\mathrm{T}} f'(x_i) = 0, \quad i = 2, \dots, 2k, \\ (p^{\mathrm{T}} f(x_i))^2 = (p^{\mathrm{T}} f(0))^2, \quad i = 2, \dots, 2k, \end{cases} \tag{19}$$

and by the first part of Theorem 2 there exists a neighbourhood, say Ω_1, of the point $(1, \dots, 1)^{\mathrm{T}}$ such that for any $\mu \in \Omega_1$ the vector $\Theta^*(\mu)$ and the local E-optimal design $\xi_E^*(\mu)$ and its corresponding eigenvector p^* satisfy (17) and (19), respectively.

Theorem 3 *For any $\mu \in \Omega$ the system of equations (17) has a unique solution*

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

The local E-optimal design for the exponential regression model (2) is given by

$$\xi_E^*(\mu) = \begin{pmatrix} 0 & x_2^*(\mu) & \dots & x_{2k}^*(\mu) \\ w_1^*(\mu) & w_2^*(\mu) & \dots & w_{2k}^*(\mu) \end{pmatrix},$$

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

It follows from Theorem 3 that for any $\mu_0 \in \mathbb{R}^k$ such that the minimal eigenvalue of the information matrix corresponding to the local E-optimal design $\xi_E^*(\mu)$ has multiplicity 1 there exists a neighbourhood, say $\mathcal{U}$, of μ_0 such that for all $\mu \in \mathcal{U}$ the function $\Theta^*(\mu)$ can be expanded in convergent Taylor series of the form

$$\Theta^*(\mu) = \Theta^*(\mu_0) + \sum_{j=1}^{\infty} \Theta^*(j, \mu_0)(\mu - \mu_0)^j. \tag{20}$$

It was shown in Dette et al. (2004) that the coefficients $\Theta^*(j, \mu_0)$ in this expansion can be calculated recursively and therefore this expansion provides a numerical method for the determination of the local E-optimal designs using the analytic properties of the support points and weights as function of μ. From a theoretical point of view it is possible that several expansions have to be performed in order to cover the whole range of Ω of all values μ such that the minimum eigenvalue of the information matrix of the local 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 4 Note that the procedure described in the previous paragraph would not give the local 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 1. In all cases considered in our numerical study the equivalence theorem confirmed our designs to be local 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) was larger than 1. Some illustrative examples are presented in the following section.

4 Examples

Example 1 Consider the exponential model $\mathbf{E}(Y(x)) = a_1 \mathrm{e}^{-\mu_1 x}$ on the interval $[0, \infty)$ corresponding to the case $k = 1$. It is easy to verify that for this model the information matrix $M(\xi)$ has a simple minimum eigenvalue for all values of $\mu_1 > 0$ (see also Han and Chaloner, 2003). In this case the Chebyshev function $\phi(x) = (1 + q_2^* x)\mathrm{e}^{-\mu_1 x}$ minimizing

$$\sup_{x \in [0,\infty)} |(1 + \alpha x)\mathrm{e}^{-\mu_1 x}|$$

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

$$\mathrm{e}^{-\mu_1 x_2} - \mu_1 x_2 + 1 = 0, \quad \alpha \mathrm{e}^{-\mu_1 x_2} + \mu_1 = 0.$$

Therefore, the second point of the local E-optimal design is given by $x_2^* = t^*/\mu_1$, where t^* is the unique solution of the equation $\mathrm{e}^{-t} = t - 1$ (the other support point is 0) and the local 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, that is

$$w_1^* = \frac{x_2^* \mathrm{e}^{-\mu_1 x_2^*} + \mu_1}{x_2^* \mathrm{e}^{-\mu_1 x_2^*} + \mu_1 + \mu_1 \mathrm{e}^{\mu_1 x_2^*}},$$
$$w_2^* = \frac{\mu_1 \mathrm{e}^{\mu_1 x_2^*}}{x_2^* \mathrm{e}^{-\mu_1 x_2^*} + \mu_1 + \mu_1 \mathrm{e}^{\mu_1 x_2^*}}.$$

Example 2 For the exponential regression model

$$\mathbf{E}(Y(x)) = a_1 \mathrm{e}^{-\mu_1 x} + a_2 \mathrm{e}^{-\mu_2 x} \tag{21}$$

on the interval $[0, \infty)$ corresponding to the case $k = 2$ the situation is more complicated and the solution of the local E-optimal design problem can not be determined directly. In this case we used the Taylor expansion (20) for the construction of the local E-optimal design, where the point μ_0 in this expansion was given by the vector $\mu_0 = (1.5, 0.5)^{\mathrm{T}}$. By Remark 3(5) we can restrict ourselves to the case $\mu_1 + \mu_2 = 2$. Local E-optimal designs for arbitrary values of $\mu_1 + \mu_2$ can be easily obtained by rescaling the support points of the local E-optimal design found under the restriction $\mu_1 + \mu_2 = 2$, while the weights have to be recalculated using Remark 3(4). We consider the parameterization $\mu_1 = 1 + z$, $\mu_2 = 1 - z$ and study the dependence of the optimal design on the parameter z. Because $\mu_1 > \mu_2 > 0$, 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 local E-optimal design, that is

$$x_i^* = x_i(z) = \sum_{j=0}^{\infty} x_{i(j)} (z - 0.5)^j,\ w_i^* = w_i(z) = \sum_{j=0}^{\infty} w_{i(j)} (z - 0.5)^j, \tag{22}$$

(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 Fig. 1. We observe for a broad range of the interval $(0, 1)$ only a weak dependence of the local E-optimal design on the parameter z. Consequently, it is of some interest to investigate the robustness of the local E-optimal design for the parameter value $z = 0$, which corresponds to the vector $\mu = (1, 1)$. This vector yields the limiting model (9) and by Theorem 2 the local E-optimal designs converge weakly to the design $\bar{\xi}_{e_{2k}}^*$, which will be denoted

Table 1 The coefficients of the Taylor expansion (22) for the support points and weights of the local E-optimal design in the exponential regression model (21)

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

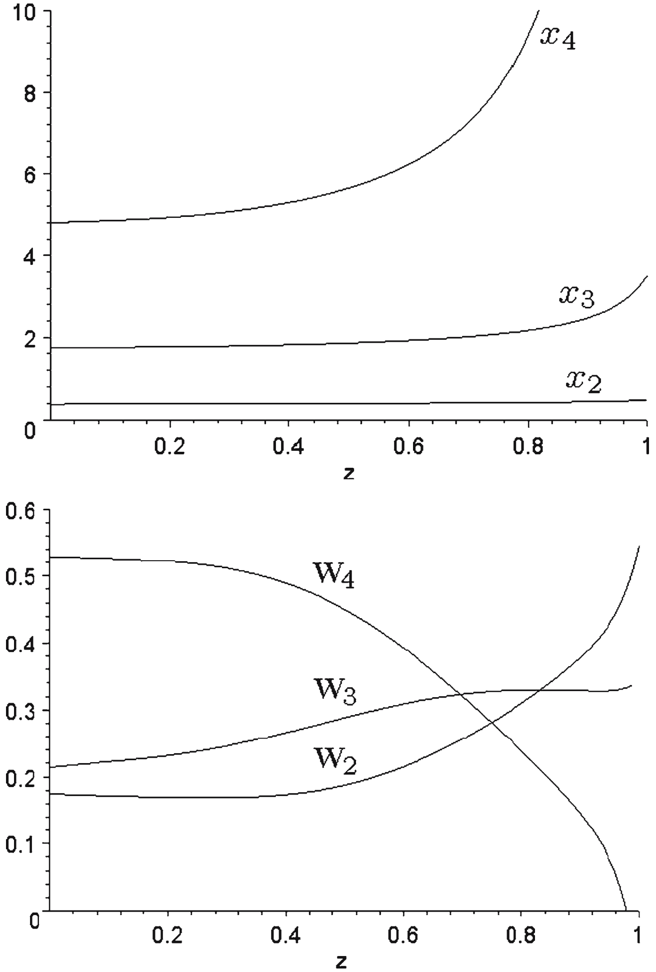

Fig. 1 Support points and weights of the local E-optimal design $\xi_E^*(\mu)$ in the exponential regression model (2), where $k = 2$ and $\mu = (1+z, 1-z)^T$

by $\bar{\xi}_E^*$ throughout this section. The support points of this design can be 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

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

and the weights of design $\bar{\xi}_E^*$ defined in Theorem 2 are calculated using formula (14), that is

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

Some E-efficiencies

$$I_E(\xi, \mu) = \frac{\lambda_{\min}(M(\xi))}{\lambda_{\min}(M(\xi_E^*(\mu)))} \tag{23}$$

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 local D-optimal design $\xi_D^*(\mu)$, the D-efficiencies

$$I_D(\xi, \mu) \;=\; \left(\frac{\det M(\xi)}{\sup_\eta \det M(\eta)}\right)^{\frac{1}{2k}} \tag{24}$$

of the local E-optimal design $\xi_E^*(\mu)$ and the corresponding efficiencies of the weak limit of the local D-optimal designs $\bar{\xi}_D^*$. We observe that the design $\bar{\xi}_D^*$ is very robust with respect to the D-optimality criterion. On the other hand the D-efficiencies of the E-optimal designs $\xi_E^*(\mu)$ and its corresponding limit $\bar{\xi}_E^*$ are substantially higher than the E-efficiencies of the designs $\xi_D^*(\mu)$ and $\bar{\xi}_D^*$.

We finally investigate the efficiencies

$$I_i(\xi, \mu) = \frac{\inf_\eta e_i^{\mathrm{T}} M^{-1}(\eta) e_i}{e_i^{\mathrm{T}} M^{-1}(\xi) e_i}, \quad i = 1, \ldots, 2k, \tag{25}$$

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 3 For the exponential model

$$\mathbf{E}(Y(x)) = a_1 \mathrm{e}^{-\mu_1 x} + a_2 \mathrm{e}^{-\mu_2 x} + a_3 \mathrm{e}^{-\mu_3 x} \tag{26}$$

Table 2 Efficiencies of local D- and E-optimal designs in the exponential regression model (21) ($\mu_1 = 1 + z, \mu_2 = 1 - z$)

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
$I_D(\xi_E^*(\mu))$	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
$I_E(\xi_D^*(\mu))$	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

The local D- and E-optimal designs are denoted by $\xi_D^*(\mu)$ and $\xi_E^*(\mu)$, respectively, while $\bar{\xi}_D^*$ and $\bar{\xi}_E^*$ denote the weak limit of the local D- and E-optimal design as $\mu \to (1, 1)$, respectively

Table 3 Efficiencies (25) of the designs $\bar{\xi}_D^*$ and $\bar{\xi}_E^*$ [obtained as the weak limit of the corresponding local optimal designs as $\mu \to (1, 1)$] for estimating the individual coefficients in the exponential regression model (21) ($\mu_1 = 1 + z$, $\mu_2 = 1 - z$)

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

corresponding to the case $k = 3$ the local E-optimal designs can be calculated by similar methods. For the sake of brevity we present only the limiting designs [obtained from the local D- and E-optimal designs if $\mu \to (1, 1, 1)$] 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 (9) (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

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

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

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

For the determination of this design we note that by Remark 3(5) we can restrict ourselves to the case $\mu_1 + \mu_2 + \mu_3 = 3$. The support points in the general case are obtained by a rescaling, while the weights have to be recalculated using Remark 3(4). For the sake of brevity we do not present the local E-optimal designs, but restrict ourselves to some efficiency considerations. For this we introduce the parameterization $\mu_1 = 1 + u + v$, $\mu_2 = 1 - u$, $\mu_3 = 1 - v$, where the restriction $\mu_1 > \mu_2 > \mu_3 > 0$ yields

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

In Table 4 we show the E-efficiency defined in (23) of the design $\bar{\xi}_E^*$, which is the weak limit of the local E-optimal design $\xi_E^*(\mu)$ as $\mu \to (1, 1, 1)$ (see Theorem 2). Two conclusions can be drawn from our numerical results. On the one hand we observe that the optimal design $\bar{\xi}_E^*$ is robust in a neighbourhood of the point (1, 1, 1). On the other hand we see that the local E-optimal design $\xi_E^*(\mu)$ is also robust if the nonlinear parameters μ_1, μ_2, μ_3 do not differ 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 (24)

Table 4 Efficiencies of local D-, E-optimal designs and of the corresponding limits $\bar{\xi}_D^*$ and $\bar{\xi}_E^*$ (obtained as the weak limit of the corresponding local optimal designs as $\mu \to (1, 1, 1)$) in the exponential regression model (26) ($\mu_1 = 1 + u + v$, $\mu_2 = 1 - u$, $\mu = 1 - v$)

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
$I_D(\xi_E^*(\mu))$	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
$I_E(\xi_D^*(\mu))$	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 5 Efficiencies (25) of the designs $\bar{\xi}_D^*$ and $\bar{\xi}_E^*$ [obtained as the weak limit of the corresponding local optimal designs as $\mu \to (1, 1, 1)$] for estimating the individual coefficients in the exponential regression model (26) ($\mu_1 = 1 + u + v$, $\mu_2 = 1 - u$, $\mu = 1 - v$)

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

and the E-efficiencies of the local D-optimal design $\xi_D^*(\mu)$ and its corresponding weak limit as $\mu \to (1, 1, 1)$. 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 local E- and D-optimal designs if $\mu \to (1, 1, 1)$ with respect to the criterion of estimating the individual coefficients in the exponential regression model (26). 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 (26). We observe that in most cases the limit of the local E-optimal designs $\bar{\xi}_E^*$ yields substantially larger efficiencies than the corresponding limit of the local D-optimal design $\bar{\xi}_D^*$. Moreover, this design is robust for many values of the parameter (u, v).

We conclude this section with some general recommendations regarding the design of experiments in the exponential regression model (1). The limiting designs have reasonable efficiencies for moderate values of $\max_{i \neq j} |\mu_i - \mu_j|$. Moreover, Theorem 2 shows that local E- and optimal designs for estimating individual coefficients in the regression model (1) behave similarly if the quantity $\max_{i \neq j} |\mu_i - \mu_j|$ is small. The numerical results presented in this section indicate that this observation can be transferred to general values of the parameters μ_i. In all cases considered in our study we found that local E-optimal designs yield substantially smaller variances for the estimates of the individual parameters than the local D-optimal designs. Therefore, if a confidence ellipsoid for the vector of parameters is not the primary interest of the experimenter, the application of E-optimal designs for the model (1) is strongly recommended.

5 Appendix

5.1 Proof of Theorem 2

Using the notation $\delta_j = r_j\delta$ and observing the approximation in (8) we obtain from the Taylor expansion

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

the representation

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

where the vectors f and $\bar{f}$ are defined in (4) and (11), respectively, the remainder term is of order

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

and the matrix L is given by

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

Consequently, the information matrix in the general exponential regression model (2) 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 (9) defined by (10). It can be shown by a straightforward but tedious calculation (see the technical report of Dette, Melas, & Peplysev, 2002) that for small δ

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

where $\mathbf{0}$ is $2k \times (2k-1)$ matrix with all entries equals 0 and the vector l is defined by (12) in Theorem 2. This yields for the information matrix of the design ξ

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

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 2 we note the local E-optimal design $\xi_E^*(\mu)$ is defined by

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

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

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

Finally, the proof of the third part of Theorem 2 can be obtained by similar arguments and is left to the reader. □

5.2 Proof of Theorem 3

In Sect. 3 we have already shown that the function $\Theta^*(\mu)$ as solution of (17) is uniquely determined. In this paragraph we prove that the Jacobi matrix

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

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

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

where the matrices B_1 and B_2 are defined by

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

respectively, and $w = (w_2, \dots, w_{2k})^{\mathrm{T}}$ and $x = (x_2, \dots, x_{2k})^{\mathrm{T}}$. Consequently, the Jacobi matrix of the system (17) has the structure

$$G = \frac{1}{q^{\mathrm{T}}q}\begin{pmatrix} D & B_1^{\mathrm{T}} & B_2^{\mathrm{T}} \\ B_1 & E & 0 \\ B_2 & 0 & 0 \end{pmatrix}. \tag{27}$$

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

$$B_1^{\mathrm{T}} = 2\tilde{c}\left(w_2 f_-'(x_2^*) \vdots - w_3 f_-'(x_3^*) \vdots w_4 f_-'(x_4^*) \vdots \dots \vdots - w_{2k-1} f_-'(x_{2k-1}^*) \vdots w_{2k} f_-'(x_{2k}^*)\right),$$

$$B_2^{\mathrm{T}} = 2\tilde{c}\left(f_-(0) + f_-(x_2^*) \vdots f_-(0) - f_-(x_3^*) \vdots \dots \vdots f_-(0) - f_-(x_{2k-1}^*) \vdots f_-(0) + f_-(x_{2k}^*)\right).$$

In the following we study some properties of the blocks of the matrix G defined in (27). Note that the matrix D in the upper left block of G is nonnegative definite. This follows from

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

and the inequality

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

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

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

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 equivalence Theorem 1, which shows that the function $(q^{*\mathrm{T}} f(x))^2$ 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 the block matrix that

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

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. The authors are also grateful to two unknown referees for their constructive comments, which led to a substantial improvement of an earlier version of this paper.

References

Alvarez, I., Virto, R., Raso, J., Condon, S. (2003). Comparing predicting models for the *Escherichia coli* inactivation by pulsed electric fields. *Innovative Food Science & Emerging Technologies*, *4*(2), 195–202.

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

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

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

Chernoff, H. (1953). Local optimal designs for estimating parameters. *Annals of Mathematical Statistics*, *24*, 586–602.

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

Dette, H., Studden, W.J. (1993). Geometry of E-optimality. *Annals of Statistics*, *21*, 416–433.

Dette, H., Haines, L., Imhof, L.A. (1999). Optimal designs for rational models and weighted polynomial regression. *Annals of Statistics*, *27*, 1272–1293.

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

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

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

Fang, Z., Wiens, D. (2004). Bayesian minimally supported D-optimal designs for an exponential regression model. *Communications in Statistics – Theory and Methods*, *33*, 1187–1204.

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

Ford, I., Torsney, B., Wu, C.F.J. (1992). The use of a canonical form in the construction of local optimal designs for non-linear problems. *Journal of the Royal Statistical Society, Series B*, *54*, 569–583.

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

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

Han, C., Chaloner, K. (2003). D- and c-optimal designs for exponential regression models used in pharmacokinetics and viral dynamics. *Journal of Statistical Planning and Inference*, *115*, 585–601.

He, Z., Studden, W.J., Sun, D. (1996). Optimal designs for rational models. *Annals of Statistics*, *24*, 2128–2142.

Heiligers, B. (1994). E-optimal designs in weighted polynomial regression. *Annals of Statistics*, *22*, 917–929.

Jennrich, R.I. (1969). Asymptotic properties of non-linear least squares estimators. *Annals of Mathematical Statistics*, *40*, 633–643.

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

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

Landaw, E.W., DiStefano, J.J. III. (1984). Multiexponential, multicompartmental, and noncompartmental modeling. II. Data analysis and statistical considerations. *American Journal of Physiology*, *246*, 665–677.

Melas, V.B. (1978). Optimal designs for exponential regression. *Mathematische Operationsforschung Statistik, Series Statistics*, *9*, 45–59.

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

Melas, V.B. (2001). Analytical properties of local D-optimal designs for rational models. In: MODA 6 – advances in model-oriented design and analysis (pp. 201–210). A.C. Atkinson, P. Hackel, W. G. Müller (Eds.) Heidelberg: Physica Verlag.

Pronzato, L., Walter, E. (1985). Robust experimental design via stochastic approximation. *Mathematical Biosciences*, *75*, 103–120.
Pukelsheim, F., Rieder, S. (1992). Efficient rounding of approximate designs. *Biometrika*, *79*, 763–770.
Pukelsheim, F., Torsney, B. (1991). Optimal weights for experimental designs on linearly independent support points. *Annals of Statistics*, *19*, 1614–1625.
Pukelsheim, F. (1993). Optimal design of experiments. New York: Wiley.
Ratkowsky, D.A. (1983). Nonlinear regression. New York: Dekker.
Ratkowsky, D.A. (1990). Handbook of nonlinear regression models. New York: Dekker
Seber, G.A.J., Wild, C.J. (1989). Nonlinear regression. New York: Wiley.
Silvey, S.D. (1980). Optimum design. London: Chapman and Hall.
Studden, W.J., Tsay, J.Y. (1976). Remez's procedure for finding optimal designs. *Annals of Statistics*, *4*, 1271–1279.
Wu, C.F.J. (1985). Efficient sequential designs with binary data. *Journal of the American Statistical Association*, *80*, 974–984.