Efficient design of experiments in the Monod model

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 70 94 559

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

Nikolai Strigul
Stevens Institute of Technology,
Department of Civil Environmental
and Ocean Engineering
USA
email: nstrigul@stevens-tech.edu

Abstract

In this paper the estimation problem and the problem of designing experiments in a nonlinear regression model, used in microbiology, are studied. The model is called Monod model, defined imlicitly by a differential equation for the regression function and has numerous applications in microbial growth kinetics, water research, pharmacokinetics and plant physiology. It is proved that least squares parameter estimates are asymptotically unbiased and normally distributed. The asymptotic covariance matrix of the least squares estimator is the basis for construction of efficient designs of experiments. In particular locally D-, E-and c-optimal designs are determined and their properties are studied. Moreover the performance of the designs (determined by the asymptotic theory) is confirmed in simulation experiments for realistic sample sizes. If certain intervals for the nonlinear parameters can be specified based on microbiological background, locally optimal designs can be constructed, which are robust with respect to misspecification of the initial parameters and which allow efficient estimation of the parameters in the Monod model. The results indicate that parameter variances can be decreased by a factor two by simply sampling at optimal times during the experiment.

Keywords and phrases: Monod model, microbial growth kinetics, least squares estimate, locally optimal designs, D-optimality, E-optimality, c-optimality.

1 Introduction

The Monod model is widely applied for modelling biodegradation rates. It is used for describing microbial growth and substrate degradations in all kinds of applications, e.g. batch and continuous fermentation, activated sludge wastewater treatment, pharmacokinetics, plant physiology etc. [see e.g. Pirt (1975), Holmberg, (1982)]. Much of the versatility of the Monod model is due to the fact that it can describe biodegradation rates following zero-one first order kinetics with respect to the target substrate concentration [see Holmberg (1982)]. Roughly speaking the model consists of a first-order differential equation, that is

$$\eta'(t) = \mu(t)\eta(t), \tag{1.1}$$

where the function μ is defined by

$$\mu(t) = \vartheta_1 \frac{s(t)}{s(t) + \vartheta_2}, \tag{1.2}$$

and the function s is given by the expression

$$s(t) - s_0 = (\eta_0 - \eta(t))/\vartheta_3, \tag{1.3}$$

(here $s_0 = s(0)$, $\eta_0 = \eta(0)$ are given initial conditions). In microbiology a traditional notation is used for the unknown parameters of the Monod model [see e.g. Pirt (1975)]. The parameter ϑ_1 denotes the maximum growth rate and is usually denoted by $\mu_{\max}$ or $V_{\max}$, ϑ_2 is the saturation of affinity constant and is often denoted by K_s. The parameter ϑ_3 is the yield coefficient (often denoted by Y) and $\eta(t)$ and $s(t)$ denote the concentration of microorganisms and the concentration of the substrate, respectively. The explanatory variable t usually denotes the time, which varies in a compact interval $[0, T]$, where the maximal time T can be of quite different size. The minimum is several hours for optimal microbiological media, while the maximum is one year or more for specialized groups of microorganisms. Due to natural biological conditions we can assume $\vartheta_i > 0$, $i = 1, 2, 3$, the initial conditions s_0 and η_0 are usually known and positive and the explanatory variable t varies in an interval, say $[0, T]$.

In a number of recent papers the problem of parameter estimation and the problem of designing experiments for this model was discussed by an extensive empirical study [see Vanrolleghem, Van Daele and Dochaine (1995); Merkel, Schwarz, Fritz, Reuss and Krauth (1996); Ossenbruggen (1996)] . The results of these authors indicate that the information quality of the experiments is highly dependent on the design and major improvements can be obtained by choosing the observations at appropriate allocations [see Vanrolleghem et. al. (1995)].

It it the purpose of the present paper to provide some more theoretical background for statistical inference in the Monod model. To be precise we assume that at experimental conditions $t_1, \dots, t_n$ independent observations $y_{1(1)}, \dots, y_{1(r_1)}, \dots, y_{n(1)}, \dots, y_{n(r_n)}$ are available, which are given by

$$y_{j(i)} = \eta(t_j, \theta) + \varepsilon_{j(i)},\ i = 1, \dots, r_j,\ j = 1, \dots, n, \tag{1.4}$$

where $\eta = \eta(\cdot, \theta)$ is a solution of the Monod equation defined by (1.1). In other words r_j denotes the number of observations at time point t_j ($j = 1, \dots, n$), $\varepsilon_{j(i)}$ are independent identically distributed random values with zero mean and constant variance $\sigma^2 > 0$ and $\theta = (\vartheta_1, \vartheta_2, \vartheta_3)^T$ denotes the vector of parameters. Throughout this paper the "true" value for θ in the model (1.1) will be denoted by $\theta^* = (\vartheta_1^*, \vartheta_2^*, \vartheta_3^*)^T$.

In Section 2 we demonstrate that statistical inference in this model is closely related to analysis in an equivalent linear regression model and we establish consistency and asymptotic normality of the least squares estimator (LSE) $\hat{\theta} = (\hat{\vartheta}_1, \hat{\vartheta}_2, \hat{\vartheta}_3)^T$, which minimizes the expression

$$\sum_{j=1}^{n} \sum_{i=1}^{r_j} (y_{j(i)} - \eta(t_j, \theta))^2 \ . \tag{1.5}$$

For sufficiently large sample size we show that the covariance matrix of the least squares estimator can be approximated by the inverse of the Fisher information matrix

$$\frac{N}{\sigma^2} \left(\sum_{j=1}^{n} \frac{r_j}{N} \frac{\partial}{\partial \vartheta_i} \eta(t_k, \theta)\Big|_{\theta=\theta^*} \frac{\partial}{\partial \vartheta_j} \eta(t_k, \theta)\Big|_{\theta=\theta^*} \right)_{i,j=1}^{3} , \tag{1.6}$$

where $N = \sum_{j=1}^{n} r_j$ denotes the total number of observations. In the second part of this paper we will study locally optimal designs [see Chernoff (1953)] for estimating the parameters in the non-linear regression function obtained as a solution of the Monod equation. These designs minimize an appropriate functional of the Fisher information matrix defined in (1.6) [see Fedorov (1972), Silvey (1980), Atkinson and Donev (1992) or Pukelsheim (1993)]. Although the regression function in the Monod model is only given implicitly, we are able to obtain an explicit representation of the information matrix (1.6), which can be used for the construction of locally optimal designs. Exemplarily we determine optimal designs with respect to the D-, E- and c- optimality criterion. Section 3 deals with the D-optimality criterion. We find the best three point designs and show that these designs are D-optimal within the class of all designs if the design region $[0, T]$ is sufficiently large and the initial value η_0 is sufficiently close to 0. These results are used for the construction of efficient designs (with respect to the D-criterion) on arbitrary design spaces. In Section 4 we present some numerical results, compare the D-optimal designs with the uniform design (which is commonly used for this type of problems) and study the sensitivity of the locally D-optimal designs with respect to changes of the initial values for the parameters. Locally E-optimal designs and optimal designs for estimating the individual parameters are investigated in Section 5, which also contains some numerical results. Some conclusions and recommendations are given in Section 6, while all technical details are deferred to an appendix in Section 7. Finally it must be stressed at this point that locally optimal designs are influenced by a preliminary "guess" for the parameter values. Our results demonstrate that without any prior information locally optimal designs for the Monod model are not robust with respect to misspecification of the initial parameters and therefore give some arguments in favour of more robust optimality criteria [see e.g. Pronzato and Walter (1985), Chaloner and Larntz (1989) or Haines (1995) among many others]. The results of the present paper provide a first step in the construction of optimal designs for the Monod model with respect to these more sophisticated criteria and also indicate that the determination of Bayesian- or robust optimal designs in the Monod model seems to be an extremely hard and challenging problem (even numerically). On the other hand, if certain intervals for the nonlinear parameters can be specified based on microbiological background, we are able to construct locally optimal designs, which are robust with respect to misspecification of the initial parameters and allow efficient estimation of the parameters in the Monod model. In many applications of the Monod model the statistical inference is made in two steps and some information regarding the parameters in the model is available from the first step, which can be used for the construction of efficient designs in the second step [see e.g. Merkel et. al. (1996)]. In such cases the application of the optimal designs determined in this paper is well justified and yields a substantial improvement with respect to the accuracy of the parameter estimates.

2 Approximate designs, equivalent regression models and asymptotic properties of the LSE.

Following Kiefer (1974) we introduce the concept of approximate design, which means that an experimental design is determined by a discrete probability measure

$$\xi = \begin{pmatrix} t_1 & \dots & t_n \\ w_1 & \dots & w_n \end{pmatrix}, \tag{2.1}$$

where $t_1, \dots, t_n \in [0, T]$ are the distinct experimental conditions at which observations have to be taken and $w_1, \dots, w_n > 0$ ($\sum_{j=1}^n w_j = 1$) are positive weights representing the proportions of total observations taken at the corresponding points [see Fedorov (1972), Silvey (1980) or Pukelsheim (1993)]. Throughout this paper we will call any (approximate) experimental design of the form (2.1) simply a design for the sake of brevity. If N observations can be taken a rounding procedure is applied to obtain integers r_j from the not necessarily integer valued quantities $w_j N$ ($j = 1, \dots n$) [see Pukelsheim and Rieder (1992)]. The analogue (up to the constant N/σ^2) of the Fisher information matrix (1.6) is the matrix

$$M(\xi, \theta^*) = \left(\sum_{k=1}^n w_k \frac{\partial}{\partial \vartheta_i} \eta(t_k, \theta) \Big|_{\theta=\theta^*} \frac{\partial}{\partial \vartheta_j} \eta(t_k, \theta) \Big|_{\theta=\theta^*} \right)_{i,j=1}^3, \tag{2.2}$$

which is called the information matrix of the design ξ. Our first result shows that least squares estimates based on an approximate design and a simple rounding procedure are consistent and asymptotically normal distributed. The proof can be found in the Appendix.

Theorem 1 *Let ξ denote an arbitrary design of the form (2.1) on the interval $[0, T]$, $n \geq 3$ and assume that r_j observations are taken at the points t_j, where the values r_j are obtained by rounding the values $w_j N$ to integers such that $\sum_{j=1}^n r_j = N$. If $\eta_0, s_0 > 0$,*

$$\theta^* \in \Omega = \{\theta = (\vartheta_1, \vartheta_2, \vartheta_3)^T : \vartheta_i > 0, \ i = 1, 2, 3\}, \tag{2.3}$$

then nonlinear least squares estimate $\hat{\theta}$ of the parameter θ^* minimizing (1.5) satisfies

$$\sqrt{N}(\hat{\theta} - \theta^*) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2 M^{-1}(\xi, \theta^*)) \quad N \to \infty \ .$$

In other words the vector $\sqrt{N}(\hat{\theta} - \theta^*)$ has asymptotically a normal distribution with mean zero and covariance matrix $\sigma^2 M^{-1}(\xi, \theta^*)$.

Note that Theorem 1 provides asymptotic unbiasedness and normality of the least squares estimator $\hat{\theta}$ in the Monod model (1.1). Moreover, asymptotically the covariance matrix of the vector $\sqrt{N}\hat{\theta}_N$ is given by

$$\sigma^2 M^{-1}(\xi, \theta^*),$$

where the information matrix is defined in (2.2). Following Chernoff (1953) we assume that θ^0 is a prior guess of the "true" parameter θ^* and call a design $\xi^*_{\theta^0}$ locally D-optimal design, if

$$\det M(\xi^*_{\theta^0}, \theta^0) = \max_{\xi} \det M(\xi, \theta^0),$$

where the maximum is taken over all designs on the interval $[0, T]$. The concept of local optimality for nonlinear regression models requires two assumptions:

- The number of observations is sufficiently large such that the asymptotic theory is applicable.
- If the design is optimal for prior guess θ^0, then it is also efficient for the "true" (but unknown) parameter θ^*, i.e.

$$I_{\theta^*}(\xi^*_{\theta^0}) \;=\; \left(\frac{\det M(\xi^*_{\theta^0}, \theta^*)}{\det M(\xi^*_{\theta^*}, \theta^*)}\right)^{1/m} > 1 - \delta,$$

  where m is the number of parameters ($m = 3$ in this case), δ is a small positive constant.

The quantity $I_{\theta^*}(\xi^*_{\theta^0})$ is called "relative efficiency" of the design $\xi^*_{\theta^0}$ with respect to the (unknown) design $\xi^*_{\theta^*}$ and $I^{-1}_{\theta^*}(\xi^*_{\theta^0})$ represents the (relative) additional amount of observations which is necessary to obtain the same accuracy with the design $\xi^*_{\theta^0}$ compared to the "ideal" (but unknown) design $\xi^*_{\theta^*}$. The first assumption can be verified by simulations, while the second assumption can be verified by a robustness study, if the locally D-optimal designs are known [see Section 4 for an example]. If these assumptions can be justified the problem of searching a locally D-optimal design for the Monod model (1.1) is equivalent to the problem of searching a D-optimal design for the linear regression model

$$\theta^T f(t) \;=\; \theta^T \frac{\partial}{\partial\theta} \eta(t, \theta)\Big|_{\theta=\theta^0},$$

and all results from the classical linear theory can be transferred to local optimality criteria. For example the local D-optimality criterion guarantees (asymptotically) a minimum volume of the confidence ellipsoid if the random errors in the model (1.4) follow a normal distribution [see Karlin, Studden (1966), Ch. X] and the local D-optimality can be characterized by the celebrated equivalence theorem of Kiefer, Wolfowitz (1960), which shows, that a design $\xi^*_{\theta^0}$ is locally D-optimal if and only if

$$f^T(t) M^{-1}(\xi^*_{\theta^0}, \theta^0) f(t) \le m, \quad \forall t \in [0, T], \tag{2.4}$$

where m is the number of parameters in the model. Moreover, the D-efficiency of a given design with respect to a locally D-optimal design can be evaluated by Kiefer's inequality without an explicit construction of a locally D-optimal design. This inequality yields for the D-efficiency

$$\left(\frac{\det M(\xi, \theta^0)}{\max_\xi \det M(\xi, \theta^0)}\right)^{1/m} \;\ge\; e^{1-v/m}, \tag{2.5}$$

where the constant v is defined by

$$v = \max_{t\in[0,T]} f^T(t) M^{-1}(\xi, \theta^0) f(t)$$

[see Pukelsheim (1993)]. Two further optimality criteria will be discussed in Section 5, which have been proposed as alternative for the Monod model [see for example Vanrolleghem et. al. (1995)

or Versyck, Bernaerts Geeraerd and Van Impe (1999)]. A design is called locally E-optimal if it maximizes the minimum eigenvalue

$$\lambda_{\min}\left(M(\xi,\theta^0)\right) \tag{2.6}$$

of the information matrix in the class of all (approximate) designs. An E-optimal design mimini-mizes the worst variance

$$\max_{p^Tp=1} \operatorname{Var}(p^T\hat{\theta}),$$

taken over the variances of all (normalized) linear combinations of the parameter estimates $p^T\hat{\theta} = \sum_{i=1}^m p_i\hat{\vartheta}_i$ for the specific value θ^0. If only one linear combination is of interest, say $c^T\theta$, the locally c-optimality criterion might be useful which determines a design minimizing the quantitiy

$$c^TM^-(\xi,\theta^0)c, \tag{2.7}$$

where A^- denotes the generalized inverse of the matrix A and the minimum is taken over the class of all designs ξ for which the linear combination $c^T\theta$ is estimable, that is $c \in \operatorname{range}(M(\xi,\theta^0))$ [see Pukelsheim (1993)]. Note that for the special choice of a unit vector $c = e_k = (0,\dots,0,1,0,\dots,0)^T$ the c-optimal design minimizes the variance of the least squares estimate for the parameter ϑ_k, $k = 1,\dots,m$, for the specific value θ^0.

Locally optimal designs have been constructed for several nonlinear regression models [see Box and Lucas (1959), Melas (1978), Rasch (1990), Ford, Torsney and Wu (1992), Haines (1992,1993), Sitter and Torsney (1995), He, Studden and Sun (1996), Dette and Wong (1999) among many others]. The problem of determining optimal designs for the Monod model under consideration is substantially more complex because the regression function in the model (1.1) is only defined implicitly and the model has two non-linear parameters. The main step in the solution of this design problem consists in a derivation of an alternative representation of the information matrix defined in (2.2). For this purpose we introduce the notation

$$\begin{aligned} c &= c(\theta) = s_0\vartheta_3 + \eta_0, \\ b &= b(\theta) = \vartheta_2\vartheta_3/c, \end{aligned} \tag{2.8}$$

Combining formulas (1.2) and (1.3) we obtain from (1.1) the differential equation

$$\eta'(t) = \vartheta_1 \frac{s_0\vartheta_3 + \eta_0 - \eta(t)}{s_0\vartheta_3 + \eta_0 - \eta(t) + \vartheta_2\vartheta_3}\eta(t) \;=\; \vartheta_1 \frac{c - \eta(t)}{c(1+b) - \eta(t)}\eta(t)\ . \tag{2.9}$$

From the initial conditions $s_0 > 0$, $\eta_0 > 0$, $\vartheta_i > 0$, $i = 1, 2, 3$ it follows that

$$\eta'(0) = \vartheta_1 \frac{s_0\vartheta_3}{s_0\vartheta_3 + \vartheta_2\vartheta_3} > 0.$$

and the following lemma describes some general properties of the regression function η.

Lemma 1 *Let η denote a non-constant solution of the Monod equation (1.1) and $T > 0$, then*

$$\begin{aligned} \eta(t) \;>\; 0\ ,\quad \eta'(t) \;>0\ ,&\quad t \in [0,T], \\ \eta_0 \;\le\; \eta(t) \;<\; c\ ,&\quad t \in [0,T], \\ \lim_{t\to\infty} \eta(t) \;=\; c& \end{aligned}$$

Note that the function η is strictly increasing on the interval $[0, T]$ and consequently for a fixed vector θ the inverse

$$t(x) = t(x, \theta) = \eta^{-1}(x, \theta) \tag{2.10}$$

exists on the interval $[\eta_0, \eta(T)]$ and satsifies $t(\eta_0) = 0$. From (2.9) it follows by a straightforward calculation

$$\begin{aligned} \frac{d\eta}{dt} &= \left.\frac{d\eta}{dt}\right|_{t=\eta^{-1}(u)} = \vartheta_1 \left(\frac{1+b}{u} + \frac{b}{c-u}\right)^{-1} \\ \frac{dt}{du} &= \frac{d\eta^{-1}}{du} = \frac{1}{\vartheta_1}\left(\frac{1+b}{u} + \frac{b}{c-u}\right) , \end{aligned}$$

and integrating the last formula we obtain

$$t(x) = \int_{\eta_0}^{x} \frac{1}{\vartheta_1}\left(\frac{1+b}{u} + \frac{b}{c-u}\right) du = \frac{1}{\vartheta_1}\left((1+b)\ln\frac{x}{\eta_0} + b\ln\frac{c-x}{c-\eta_0}\right) . \tag{2.11}$$

Throughout this paper let

$$\mathcal{X} = \{\eta(t, \theta^0) \mid t \in [0, T] \} \tag{2.12}$$

denote the induced design space, then any design ξ of the form (2.1) on the interval $[0, T]$ with $0 \le t_1 < t_2 < \ldots < t_n \le T$ induces a design ζ on $\mathcal{X}$ by the transformation

$$\eta_i = \eta(t_i, \theta^0),\ i = 1, \ldots, n. \tag{2.13}$$

In the following we will prove that the matrix $M(\xi, \theta^0)$ can be represented as a function of the points $\eta_1, \ldots, \eta_n$. Note that such a relation is obvious for linear models [see Pukelsheim (1993), p.3], but in general not clear for the Monod model under consideration, because the information matrix for this model contains the partial derivatives of the regression function η evaluated at the points t_i. In order to prove this dependency we differentiate the identity

$$\eta(t(x, \theta), \theta) \equiv x\ , \quad \theta \in \Omega\ ,\ x \in [\eta_0, c] \tag{2.14}$$

with respect to the parameters ϑ_i $(i = 1, 2, 3)$ and obtain

$$\frac{\partial\eta(t, \theta)}{\partial t}\frac{\partial t(x, \theta)}{\partial\vartheta_i} + \frac{\partial\eta(t, \theta)}{\partial\vartheta_i} = 0\ , \quad i = 1, 2, 3 \tag{2.15}$$

with $t = t(x, \theta)$, $x = \eta(t, \theta)$. Now observing (2.11) and (2.15) it follows by a direct computation that for any $t \ge 0$

$$\frac{\partial\eta(t, \theta)}{\partial\theta} = K\varphi(x), \tag{2.16}$$

where $x = \eta(t, \theta)$ and the matrix $K \in \mathbb{R}^{3\times 3}$ is defined by

$$K = \begin{pmatrix} \frac{1+b}{\vartheta_1} & \frac{b}{\vartheta_1} & 0 \\ -\frac{b}{\vartheta_2} & -\frac{b}{\vartheta_2} & 0 \\ -\frac{b\eta_0}{c\vartheta_3} & -\frac{b\eta_0}{c\vartheta_3} & -\frac{b}{\vartheta_3} \end{pmatrix} . \tag{2.17}$$

Here

$$\varphi(x) = (\varphi_1(x), \varphi_2(x), \varphi_3(x))^T$$

denotes a vector of regression functions with components

$$(2.18)\qquad \begin{aligned} \varphi_1(x) &= \varphi_1(x,\theta) = v(x)\ln\tfrac{x}{\eta_0},\\ \varphi_2(x) &= \varphi_2(x,\theta) = v(x)\ln\tfrac{c-x}{c-\eta_0},\\ \varphi_3(x) &= \varphi_3(x,\theta) = v(x)\tfrac{x-\eta_0}{c-x}, \end{aligned}$$

and

$$(2.19)\qquad v(x) = v(x,\theta) = \frac{x(c-x)}{(1+b)c-x}\ .$$

Note that for each θ the function $\eta(t) = \eta(t,\theta)$ is strictly increasing with limit $\lim_{t\to\infty}\eta(t) = c$ [see Lemma 1]. Thus it is possible to extend this function by the definition $\eta(\infty) = c$. By the above discussion we can now transfer the original (locally) optimal design problem for the Monod model to a design problem for a linear model on the induced design space $\mathcal{X}$. To this end let

$$(2.20)\qquad \zeta = \begin{pmatrix} x_1 & \dots & x_n \\ w_1 & \dots & w_n \end{pmatrix},\ \eta_0 \le x_1 < x_2 < \ldots < x_n \le \bar{c}$$

denote an arbitrary design on the interval $[\eta_0, \bar{c}]$, where $\bar{c} \le c$, and define

$$(2.21)\qquad \xi_\zeta = \begin{pmatrix} t_1 & \dots & t_n \\ w_1 & \dots & w_n \end{pmatrix},$$

with $t_i = t(x_i, \theta^0)$, $i = 1, \ldots, n$ as the corresponding design on the original design space $[0, T]$ with $T = t(\bar{c}, \theta^0)$. We define

$$\bar{M}(\zeta) = \bar{M}(\zeta, \theta^0) = \sum_{j=1}^{n} w_j \varphi(x_j)\varphi(x_j)^T,$$

as the information matrix of the design ζ in the (homoscedastic) linear regression model

$$(2.22)\qquad \beta^T\varphi(x)\ ,$$

where $\beta = (\beta_1, \beta_2, \beta_3)$ denotes the vector of parameters. It follows from (2.2) and formula (2.16), that

$$(2.23)\qquad \begin{aligned} M(\xi_\zeta, \theta^0) &= K\bar{M}(\zeta, \theta^0)K^T,\\ \det M(\xi_\zeta, \theta^0) &= \tfrac{b^4}{(\vartheta_1^0\vartheta_2^0\vartheta_3^0)^2}\det \bar{M}(\zeta, \theta^0). \end{aligned}$$

The following results are now obvious from these considerations.

Theorem 2 *A design ξ_ζ is a locally D-optimal design for the Monod model (1.1) on the interval $[0, T]$ if and only if the induced design ζ is D-optimal for the regression model (2.22) on the interval $[\eta_0, \bar{c}]$, $\bar{c} = \eta(T)$, under the standard assumptions about the measurement errors.*

Theorem 3 *A design ξ_ζ is a locally $E-(e_k-)$optimal design for the Monod model (1.1) on the interval $[0,T]$ if and only if the design ζ is $E-(e_k-)$ optimal for regression function $\beta^T K\varphi(x)$ on the interval $[\eta_0,\bar{c}]$, $\bar{c}=\eta(T)$, under the standard assumptions on the measurement errors.*

Consequently, it is sufficient to construct locally D-optimal designs for the regression model (2.22) and locally E- and e_k-optimal designs for the regression model $\beta^T K\varphi(x)$. The locally optimal designs for the Monod model (1.1) are simply obtained by tranforming the design η in (2.20) to the design ξ_ζ in (2.21). We will illustrate this method in the following sections discussing the different optimality criteria seperately.

3 Locally D-optimal designs

Due to Theorem 2 it will be sufficient to study D-optimal designs for the linear regression model (2.22) under the standard assumptions about measurements errors. Recall the notation $c=\eta(\infty,\theta^0)$, $\bar{c}=\eta(T,\theta^0)\le c$, then it is easy to see that the definition of the vector of regression functions φ in the model (2.22) can be continuously extended by putting

$$\begin{aligned}&\varphi_i(c)=0,\ i=1,2\\ &\varphi_3(c)=\tfrac{c-\eta_0}{b}\ =\ \tfrac{s_0c}{\vartheta_2}\ .\end{aligned}$$

In other words the vector of regression functions in the model (2.22) is well defined on intervals $[\eta_0,\bar{c}]$, where $\bar{c}=\eta(T,\theta^0)$, $0<T\le\infty$, $\bar{c}\le c$. Let us denote designs that are D-optimal in the class of all designs supported at k support points as D-optimal k-point designs (E- and e_k-optimal k-point designs are defined similary). It is well known [see Karlin, Studden (1966), Ch. X] that if the number of design points is less than the number of estimated parameters in the regression model, then the information matrix is singular. Thus the D-optimal design has at least three support points. The following result determines the best 3-point design for the regression model (2.22). The proof is deferred to the Appendix.

Lemma 2 *Assume that $\theta^0\in\Omega$ and that x_1^*,x_2^* are determined by the relation*

$$\Phi(x_1^*,x_2^*,\theta^0)\ =\ \max\Big\{\Phi(x_1,x_2,\theta^0)\ \Big|\ \eta_0\le x_1<x_2\le\bar{c}\ \Big\}\ , \tag{3.1}$$

where the function Φ is defined by

$$\Phi(x_1,x_2,\theta^0)=\det(\varphi_i(x_j))_{i,j=1}^3, \tag{3.2}$$

with $x_3=\bar{c}$. The design

$$\zeta_{\bar{c}}^*=\begin{pmatrix}x_1^* & x_2^* & \bar{c}\\ 1/3 & 1/3 & 1/3\end{pmatrix} \tag{3.3}$$

is a D-optimal 3-point design for the regression model (2.22) on the interval $[\eta_0,\bar{c}]$ for any $\bar{c}$ with $\eta_0<\bar{c}\le c$.

In general it is not clear if there exist better designs (with respect to the D-criterion) with more than three support points. In our numerical study we did not find D-optimal designs for the regression model (2.22) with more than three support points, but a general proof of this property for arbitrary design regions seems to be difficult. However, it is possible to obtain theoretical results in this direction if the right endpoint $\bar{c}$ of the design space is sufficiently large and the initial condition η_0 is sufficiently small. For this purpose we consider at first the design problem for the regresion model (2.22) on the interval $[0, c]$. In this case the vector of the regression functions can be rewritten in a more convenient form by the substitution $x \to cx$

$$\hat{\varphi}(x) = cv(x)\left(\ln\frac{x}{\tilde{\eta}_0}, \ln\frac{1-x}{1-\tilde{\eta}_0}, \frac{x-\tilde{\eta}_0}{1-x}\right)^T,$$

where $\tilde{\eta}_0 = \eta_0/c$. Thus we can assume without loss of generality that $c = 1$ and the D-optimal designs on the general interval $[0, c]$ are obtained by a rescaling from the D-optimal designs on the interval $[0, 1]$ (calculated for $c = 1$ and initial condition $\tilde{\eta}_0 = \eta_0/c$). Because D-optimal designs are not changed under nonsingular transformations of the regression functions, the problem is reduced to the investigation of D-optimal designs for the regression model $\beta^T\tilde{\varphi}(x, \tilde{\eta}_0)$, where the vector $\tilde{\varphi}$ is defined by

$$\tilde{\varphi}(x, \tilde{\eta}_0) \;=\; (\tilde{\varphi}_1(x), \tilde{\varphi}_2(x), \tilde{\varphi}_3(x))^T \;=\; v(x)\left(\frac{\ln(x/\tilde{\eta}_0)}{-\ln\tilde{\eta}_0}, \ln\frac{1-x}{1-\tilde{\eta}_0}, \frac{x-\tilde{\eta}_0}{1-x}\right)^T.$$

A direct computation shows that for $\tilde{\eta}_0 \to 0$

$$\begin{aligned}&\tilde{\varphi}(x, \tilde{\eta}_0) \to \psi(x)\\ &\tilde{\Phi}(x_1, x_2, \theta^0) = \det(\tilde{\varphi}_i(x_j))_{i,j=1}^3 \;\to\; \Psi(x_1, x_2).\end{aligned}$$

(in the last equation we put $x_3 = 1$), where the functions ψ and Ψ are defined by

$$\begin{aligned}\psi(x) \;&=\; v(x)\left(1, \ln(1-x), \frac{x}{1-x}\right)^T,\\ \Psi(x_1, x_2) \;&=\; \frac{1}{b}v(x_1)v(x_2)[\ln(1-x_1) - \ln(1-x_2)]\\ &=\; \frac{1}{b}\frac{\ln(1-x_1) - \ln(1-x_2)}{(1-x_1+b)(1-x_2+b)}x_1x_2(1-x_1)(1-x_2).\end{aligned} \tag{3.4}$$

Lemma 3

(1) For for any $b = \vartheta_2\vartheta_3 > 0$ the unique D-optimal design for the regression model $\beta^T\psi(x)$ on the interval $[0, 1]$ is given by

$$\tilde{\zeta} = \begin{pmatrix}\tilde{x}_1 & \tilde{x}_2 & 1\\ 1/3 & 1/3 & 1/3\end{pmatrix},$$

where the points $\tilde{x}_1 = \tilde{x}_1(b)$ and $\tilde{x}_2 = \tilde{x}_2(b)$ are determined by the relation

$$\Psi(\tilde{x}_1, \tilde{x}_2) = \max\left\{\; \Psi(x_1, x_2) \;\middle|\; 0 \le \tilde{x}_1 < \tilde{x}_2 < 1 \;\right\}, \tag{3.5}$$

and the function Ψ is defined in (3.4).

(2) *There exist positive numbers $\varepsilon > 0$ and $\delta > 0$ such that for any $\eta_0 < \varepsilon$ and any $\bar{c} > c - \delta$ the design given in Lemma 2 by formula (3.3) is the unique D-optimal design for the regression model (2.22) with design region $[\eta_0, \bar{c}]$. Moreover, if $\eta_0 \to 0$ and $\bar{c} \to 1$ we have*

$$V^T \bar{M}^{-1}(\zeta^*_{\bar{c}}) V \to \tilde{M}^{-1}(\tilde{\zeta}),$$

where $\tilde{M}(\zeta)$ is the information matrix of design ζ in the regression model $\beta^T \psi(x)$, $\bar{M}(\zeta)$ is the information matrix of design ζ in the regression model $\beta^T \varphi(x)$ and $V = \mathrm{diag}(-1/\ln \eta_0, 1, 1)$. In particular it follows that

$$e_3^T \bar{M}^{-1}(\zeta^*_{\bar{c}}) e_3 \to 3b^2.$$

The following theorem is now obtained by the combination of Lemma 2, 3 and Theorem 2.

Theorem 4 *Consider the Monod model defined by (1.1) - (1.3) on the interval $[0, T]$.*

(1) *For any $\theta^0 \in \Omega$ and $0 < T \leq \infty$ a locally D-optimal 3-point design is given by*

$$\xi^*_T = \begin{pmatrix} t_1^* & t_2^* & T \\ 1/3 & 1/3 & 1/3 \end{pmatrix}, \tag{3.6}$$

where

$$t_i^* = \frac{1}{\vartheta_1}\left[(1+b)\ln(x_i^*/\eta_0) + b \ln \frac{c-\eta_0}{c - x_i^*}\right], \quad i = 1, 2$$

and the points x_i^* are determined by the relation (3.1).

(2) *The design ξ^*_T determined by (3.6) is the unique locally D-optimal design for sufficiently large T and sufficiently small η_0.*

(3) *For $N \to \infty$, $T \to \infty$ $\eta_0 \to 0$ the following relation holds:*

$$\frac{N}{\sigma^2} D(\hat{\vartheta}_3) \;\to\; 3\vartheta_1^{-2}.$$

4 A numerical study of locally D-optimal designs

4.1 Some locally D-optimal designs for the Monod model

We begin with a numerical construction of the design ζ^*_c defined by (3.3), which is D-optimal for the regression model (2.22) on the interval $[0, c]$ according to the second part of Lemma 3. The function $\Phi(x_1, x_2, \theta^0)$ defined in (3.2) for $x_3 = c$ is of the from

$$\Phi(x_1, x_2, \theta^0) = v(x_1)v(x_2)\left[\ln\frac{x_1}{\eta_0}\ln\frac{c - x_2}{c-\eta_0} - \ln\frac{x_2}{\eta_0}\ln\frac{c-x_1}{c-\eta_0}\right]\cdot\frac{c-\eta_0}{b}, \tag{4.1}$$

where the function v is defined by (2.19). The maximum of Φ on the set $\{\ (x_1, x_2) \mid \eta_0 \leq x_1 < x_2 \leq c\ \}$ can be calculated by a standard gradient method. Define $\tilde{\eta}_0 = \eta_0/c$ then it is again sufficient to consider only the case $c = 1$ [see the discussion in Section 3]. The designs ζ^*_c on the interval $[\eta_0, c]$ can be simply obtained from the designs on the interval $[\tilde{\eta}_0, 1]$ multiplying the support points with c. Table 4.1 shows locally D-optimal 3-point designs on the interval $[\tilde{\eta}_0, 1]$ and the diagonal elements of the corresponding covariance matrix for various values of the parameters $b = \vartheta_2\vartheta_3$ and $\tilde{\eta}_0$. According to Lemma 2 these designs are locally D-optimal among all designs with three support points. The optimality in the class of all designs was checked by an application of Kiefer's equivalence theorem [see Kiefer and Wolfowitz (1960) or equation (2.4)]. We observed in all considered cases that the checking condition was satisfied and our numerical study shows that the D-optimal 3-point designs are in fact D-optimal within the class of all approximate designs on interval $[\eta_0, c]$.

Table 4.1: Locally D-optimal 3-point designs ζ_1^* for the equivalent regression model (2.22) on the interval $[\tilde{\eta}_0, 1]$ for various values of $\tilde{\eta}_0$ and $b = \vartheta_2\vartheta_3$. These designs are of the form (3.3) and determined by Lemma 2. The table also shows the diagonal elements $\bar{m}^{ii}$ $(i = 1, 2)$ of the matrix $\bar{M}^{-1}(\zeta_1^*)$ rounded to integers. Optimal designs on the interval $[\tilde{\eta}_0 c, c]$ can be obtained by rescaling the design ζ_1^* with the factor c. The value of $\bar{m}^{33} = (\bar{M}^{-1}(\zeta_1^*))_{33}$ is given by $3b^2/(1-\tilde{\eta}_0)^2$.

b	x_1^*	x_2^*	$\bar{m}^{11}$	$\bar{m}^{22}$	$\bar{m}^{33}$	x_1^*	x_2^*	$\bar{m}^{11}$	$\bar{m}^{22}$	$\bar{m}^{33}$
	$\tilde{\eta}_0 = 0.2$					$\tilde{\eta}_0 = 0.1$				
0.1	0.70	0.95	31	23	0.05	0.67	0.95	13	19	0.04
0.25	0.65	0.93	71	70	0.29	0.61	0.92	28	57	0.23
0.75	0.59	0.91	279	367	2.64	0.55	0.89	101	284	2.08
1	0.58	0.90	428	593	0.54	4.69	0.89	153	456	3.70
2	0.56	0.89	1321	2013	18.7	0.51	0.88	458	1521	14.8
	$\tilde{\eta}_0 = 0.05$					$\tilde{\eta}_0 = 0.01$				
0.1	0.65	0.94	7	17	0.03	0.62	0.94	2	15	0.03
0.25	0.59	0.92	14	50	0.21	0.56	0.91	5	43	0.19
0.75	0.52	0.89	49	244	1.87	0.49	0.88	16	205	1.72
1	0.51	0.88	74	389	3.32	0.48	0.87	24	325	3.06
2	0.48	0.87	217	1286	13.3	0.45	0.86	69	1065	12.24

We will now discuss the corresponding designs for the Monod model defined by the differential equation (1.1), which are related to the D-optimal designs for the linear regression model (2.22) by the relations (2.20) and (2.21). Because $\eta(\infty) = c$ the corresponding designs for the Monod model have a support point at $T = \infty$ and cannot be realized in practice. This fact was also observed empirically by Vanrolleghem et. al. (1995), who showed that for the commonly used optimality criteria (including the D-, A-, and E-criterion) the optimal strategies yield to prohibitively long experiments. However, even if the D-optimal designs are not directly implementable, they can be used as a basis for evaluating the efficiency of other designs, which are used in practice. For example consider the following design

$$\hat{\xi} = \begin{pmatrix} t_1^* & t_2^* & t_3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}, \qquad (4.2)$$

where $t_i^* = t_i^*(\theta^0)$, $i = 1, 2$ are points of the design

$$\xi_\infty^* = \begin{pmatrix} t_1^* & t_2^* & \infty \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

on the infinite design space $[0, \infty]$, which was determined in Theorem 4. We will now choose the point t_3 such that the efficiency of the design $\hat{\xi}$ with respect to the design ξ^*_∞

$$I = \left(\frac{\det M(\hat{\xi})}{\det M(\xi^*_\infty)} \right)^{1/3} = \left(\frac{\det \bar{M}(\zeta_{\hat{\xi}})}{\det \bar{M}(\zeta_{\xi^*_\infty})} \right)^{1/3} \tag{4.3}$$

is equal to a given value $1-\delta$ (note that $\zeta_{\xi^*_\infty} = \zeta^*_c$). The corresponding values of the point t_3 are represented in Table 4.2 for various values of δ, b and η_0. From this table we can conclude that for $t_3 = 2t^*_2$ the efficiency of the design $\hat{\xi}$ is at least 0.98 for all considered parameter combinations. Thus if $t^*_2 < T/2$, the designs

$$\xi^*_D = \begin{pmatrix} t^*_1 & t^*_2 & 2t^*_2 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}. \tag{4.4}$$

are close to the locally D-optimal designs on the interval $[0,\infty]$ and they can be realized in practice. Note that the true efficiency of these designs on the interval $[0,T]$ is usually larger than $1-\delta$, because our comparison is based on a design for an infinite design space and the locally D-optimal design on the interval $[0,T]$ has always a smaller determinant than the locally D-optimal design ξ^*_∞ on the infinite design space $[0,\infty]$. In other words the designs given in (4.4) have at least D-efficiency $1-\delta$ for the concrete interval $[0,T]$, whenever $t^*_2 \le T/2$.

Table 4.2: The third support point t_3 of the design $\hat{\xi}$ defined in (4.2), such that the relative D- efficiency (4.3) is at least $1-\delta$. The last two columns show the ratio $k(1-\delta) = t_3/t^*_2$.

b	η_0	t^*_1	t^*_2	$1-\delta$				$k(0.95)$	$k(0.99)$
				0.9	0.95	0.98	0.99		
				t_3					
0.25	0.1	2.5	3.4	4.2	4.4	4.7	4.9	1.3	1.4
0.5	0.1	3.0	4.4	6.0	6.5	7.1	7.4	1.5	1.7
0.75	0.1	3.5	5.4	7.8	8.5	9.3	9.9	1.6	1.8
1	0.1	4.0	6.5	9.7	10.5	11.6	12.4	1.6	1.9
1.5	0.1	5.1	8.5	13.3	14.6	16.2	17.4	1.7	2.1
2	0.1	6.1	10.5	17.0	18.7	20.8	22.5	1.8	2.1
0.25	0.2	1.7	2.5	3.4	3.6	3.9	4.0	1.4	1.6
0.25	0.05	3.3	4.2	5.1	5.3	5.6	5.8	1.2	1.4
0.25	0.01	5.2	6.2	7.1	7.3	7.6	7.8	1.2	1.2

We now compare the design ξ^*_D defined in (4.4) with the uniform design

$$\xi_u = \begin{pmatrix} T/n & 2T/n & \dots & T \\ 1/n & 1/n & \dots & 1/n \end{pmatrix}. \tag{4.5}$$

A numerical study indicates that the information matrix of the uniform design for $n \ge 10$, $T > 1.5t^*_2$ does not depend sensitively on the parameters n and T (these results are not displayed for the sake of brevity). Exemplarily we consider the case $n = 10, T = 2t^*_2$, other situations give similar results. A comparison of the uniform design ξ_u and the design ξ^*_D given in (4.4) with respect to the D-criterion shows that the D-efficiencies

$$I(\xi_u) \;=\; \left(\frac{\det M(\xi_u)}{\det M(\xi^*_D)} \right)^{\frac{1}{3}} \tag{4.6}$$

of the uniform design vary between 150% and 200% [see Table 4.3]. This indicates a rather poor performance of the uniform design ξ_u. A more refined comparison is obtained by looking at the the asymptotic variances of the estimators for the parameters ϑ_1 , ϑ_2 and ϑ_3. Note that the ratio of these variances is given by

$$d_i = \frac{(M^{-1}(\xi_u))_{ii}}{(M^{-1}(\xi_D^*))_{ii}} \ , \quad i = 1, 2, 3. \tag{4.7}$$

We observe from Table 4.3 that the uniform design produces a smaller (asymptotic) variance of the estimator for the parameter ϑ_3 compared to the design (4.4). On the other hand the asymptotic variances obtained for the estimators for the parameters ϑ_1 and ϑ_2 are substantially smaller than the corresponding variances obtained from the uniform design (4.5). However in realistic situations [see Pirt (1975) and Blok (1994) and the simulations of the following paragraph] an efficient estimation of ϑ_1 and ϑ_2 is more important, because the parameter ϑ_3 is usualy estimated with much higher precision than the parameters ϑ_1 and ϑ_2. Moreover, the loss of efficiency using the unifom design for estimating ϑ_1 and ϑ_2 is substantially larger than the loss of efficiency using the design ξ_D^* for estimating the parameter ϑ_3. Thus, if we consider estimation of the parameters ϑ_1 and ϑ_2 as more important, the design ξ_D^* yields a reduction of approximately 50% of the variance compared to the uniform design ξ_u [see the results in Table 4.3] provided that the sample size N is sufficiently large.

Table 4.3: The efficiency of the design ξ_D^* defined in (4.4) in relation to the uniform design ξ_u defined in (4.5) in the Monod model (1.1). Here the efficiency d_i defined in (4.7) ($i = 1, 2, 3$) corresponds to the estimation of the individual parameters $\vartheta_1, \vartheta_2, \vartheta_3$ and depends only on the parameters $\tilde{\eta}_0 = \eta_0/c$ and b. $I(\xi_u)$ corresponds to the D-criterion and is defined in (4.6).

$\tilde{\eta}_0$	0.2					0.1				
b	0.1	0.25	0.75	1	2	0.1	0.25	0.75	1	2
d_1	2.0	2.0	2.2	2.3	2.5	1.9	2.1	2.2	2.3	2.3
d_2	2.2	2.0	2.2	2.3	2.4	2.1	2.2	2.2	2.3	2.3
d_3	1.0	0.7	0.6	0.6	0.6	1.2	0.8	0.6	0.6	0.6
$I(\xi_u)$	1.6	1.4	1.4	1.5	1.5	1.7	1.6	1.5	1.5	1.5
$\tilde{\eta}_0$	0.05					0.01				
b	0.1	0.25	0.75	1	2	0.1	0.25	0.75	1	2
d_1	3.0	2.1	2.3	2.3	2.4	2.8	2.5	2.4	2.5	2.4
d_2	3.4	2.2	2.2	2.3	2.4	3.0	2.6	2.4	2.5	2.4
d_3	1.4	0.9	0.7	0.7	0.6	1.7	1.2	0.8	0.7	0.7
$I(\xi_u)$	2.1	1.6	1.6	1.5	1.6	2.4	2.0	1.7	1.7	1.7

In order to investigate how these asymptotic observations can be transferred to realistic sample sizes a small simulation study was conducted. We simulated observations according to the model (1.4) where the erors are normally distributed with variance $\sigma^2 = 0.01^2$ and $\sigma^2 = 0.02^2$. The parameter $\theta = (\vartheta_1, \vartheta_2, \vartheta_3)^T$ was estimated by the least squares technique using the Nelder-Mead simplex method. For the parameter θ^* we fixed the value $(0.25, 0.5, 0.25)^T$ which corresponds to parameters observed in studies of microbial growth [see Pirt (1975) or Blok (1994)]. The simulation was repeated 400 times for the sample sizes $N = 20, 30, 40, 60$ and the designs ξ_D^* and ξ_u defined in (4.4) and (4.5), respectively. The simulated variances of the estimators for the parameters ϑ_1, ϑ_2, ϑ_3 are represented in Table 4.4 and 4.5 corresponding to the choices $\sigma = 0.01$ and $\sigma = 0.02$, respectively. For the sake of transparency these values are multiplied with N/σ^2 (and rounded to integers). The results confirm our asymptotic findings for realistic sample sizes. The design ξ_D^* defined in (4.4) allows a substantially more precise estimation of the parameters ϑ_1 and ϑ_2 compared to the design ξ_u. Note that the variances for the estimation of the parameter ϑ_3 are substantially smaller (independent of the design) than the corresponding variances of the estimators for ϑ_1 and ϑ_2. Finally, it is worthwhile to mention that the asymptotic considerations of the previous paragraph are applicable if $N \geq 30$. Thus our simulation study confirmed that the design ξ_D^* performs much better than the uniform design ξ_u.

Table 4.4: Simulated and asymptotic variances of the estimates for the parameters in the Monod model (1.1) for different sample sizes (N). The variances are multiplied with N/σ^2 ($\sigma = 0.01$) and are presented for the uniform design ξ_u and the design ξ_D^* in (4.4) obtained from the D-optimal design on an infinite design space.

N	20	30	40	60	∞
	uniform design				
θ_1	771	606	575	529	551
θ_2	17693	13637	12336	11944	12500
θ_3	2.1	2.1	2.3	2.1	2.3
	optimal design				
θ_1	432	295	195	180	269
θ_2	10101	6789	4357	4070	6055
θ_3	3.3	2.9	2.9	2.6	3.0

Table 4.5: Simulated and asymptotic variances of the estimates for the parameters in the Monod model (1.1) for different sample sizes (N). The variances are multiplied with N/σ^2 ($\sigma = 0.02$) and are presented for the uniform design ξ_u and the design ξ_D^* in (4.4) obtained from the D-optimal design on an infinite design space.

N	20	30	40	100	∞
	uniform design				
θ_1	2936	2080	1791	564	551
θ_2	72541	49874	43325	13494	12500
θ_3	4.0	3.8	3.7	3.6	2.3
	optimal design				
θ_1	1056	867	798	278	269
θ_2	23024	20390	14880	6359	6055
θ_3	2.6	2.9	2.9	3.0	3.0

4.2 Robustness with respect to the inital parameter choice

For practical purposes it is important to study the sensitivity of the locally D-optimal design with respect to misspecification of the initial parameters. Up to now the optimal designs were determined under the assumption that $\theta^0 = \theta^*$ but the "true" vector of parameters θ^* is not known in practice and the vector θ^0 is a preliminary guess, for example obtained from the least squares estimates in a similar experiment or from a preliminary investigation as recommended by Merkel et. al. (1996). In order to obtain some feeling about this sensitivity we choosed $\vartheta_2^* = 0.5$, $\vartheta_3^* = 0.25$ for the nonlinear parameters in the Monod model and calculated the design ξ_D^* defined in (4.4) for $\theta = \theta^*$. In Table 4.6 we show the efficiency for estimating the individual parameters ϑ_1, ϑ_2 defined by

$$d_i = \left(\frac{(M^{-1}(\xi_0^*))_{ii}}{(M^{-1}(\xi_D^*))_{ii}} \right)^{-1}, \quad i = 1, 2, \tag{4.8}$$

where ξ_0^* denotes the locally optimal design (4.4) calculated under the preliminary assumption $\vartheta_2^0 = \vartheta_2^* + u$, $\vartheta_3^0 = \vartheta_3^* + v$. Note that we did not present the quantity d_3 in this table, because it is close to one in all cases under considerations. In other words the design given in (4.4) is extremely robust for the estimation of the parameter ϑ_3 with respect to misspecification of the initial parameters. However, we observe from Table 4.6 that misspecification of ϑ_2^* and ϑ_3^* may yield a substantial loss of efficiency for the estimation of the parameters ϑ_1 and ϑ_2. Because estimation of ϑ_1 and ϑ_2 is usually considered as more difficult we conclude from our numerical study that the design (4.4) derived from the locally D-optimal design is rather sensitive with respect to misspecification of the initial parameter θ^0.

Table 4.6: Sensitivity of the design (4.4) drived from the locally D-optimal design with respect to misspecification of the inital values. The table shows the loss of efficiency defined in (4.8) for estimating the individual parameters and the loss of D-efficiency (I) if the initial parameters ϑ_2^* and ϑ_3^* have been misspecified by an amount of u and v, respectively.

u	0.05	0.05	-0.05	-0.05	0.1	0.1	-0.1	-0.1	0.15	0.15	-0.15	-0.15
v	0.05	-0.05	0.05	-0.05	0.1	-0.1	0.1	-0.1	0.15	-0.15	0.15	-0.15
I	1.12	1.02	1.01	1.23	1.45	1.18	1.02	3.01	1.92	1.92	1.04	50.0
d_1	1.42	1.14	0.98	2.09	2.70	1.87	0.98	35.70	5.30	8.08	0.96	18055
d_2	1.43	1.15	0.97	2.14	2.83	1.84	0.97	31.18	5.79	7.22	0.94	1017

Remark 4.1. The numerical results given in Table 4.6 can be used to give a recommendation for practical applications. Usually intervals for the unknown parameters ϑ_2^* and ϑ_3^* can be specified by the experimenter, based on microbiological background. Assume that these intervals are given by $[\alpha_1, \alpha_2]$, and $[\beta_1, \beta_2]$, respectively. In this case we propose as the design for the experiment with the Monod model the locally D-optimal design for the parameter $(\vartheta_2, \vartheta_3) = (\alpha_2, \beta_1)$, because in this case the true parameter $(\vartheta_2^*, \vartheta_3^*) \in [\alpha_1, \alpha_2] \times [\beta_1, \beta_2]$ satsifies $\vartheta_2^* = \vartheta_2 + u$ and $\vartheta_3^* = \vartheta_3 + v$ with negative u and positive v. The results of Table 4.6 show that the designs constructed by this procedure are extremely robust with respect to misspecification of the initial parameter values.

5 Locally E- and e_k-optimal designs

In this section we present a theoretical and numerical study of the locally E- and e_k-optimal designs in the Monod model. The optimality criteria are defined in (2.6) and (2.7), respectively. Due to

Theorem 3 it will be sufficient to study the optimal designs for the regression model $\beta^T K\varphi(x)$ on the interval $[\eta_0, \bar{c}]$, where the matrix K is defined by (2.17). It will be shown in Proposition 1 of the Appendix that the functions $\varphi_1(x)$, $\varphi_2(x)$ and $\varphi_3(x)$ defined in (2.18) generate a Chebyshev-system on the interval $[\eta_0, \bar{c}]$. Then it is well known [see Karlin, Studden (1966), Ch. 1] that there exists a function $g(x)$,

$$g(x) = r_1\varphi_1(x) + r_2\varphi_2(x) + r_3\varphi_3(x), \tag{5.1}$$

with real coefficients r_1, r_2 and r_3, and (unique) points $\tilde{x}_1$, $\tilde{x}_2$ and $\tilde{x}_3$ satisfying $\eta_0 \leq \tilde{x}_1 < \tilde{x}_2 < \tilde{x}_3 \leq \bar{c}$ such that

$$g(\tilde{x}_i) = (-1)^{i+1}, \ \ i = 1, 2, 3 \tag{5.2}$$

and such that the inequality

$$|g(x)| \leq 1, \tag{5.3}$$

holds for all $x \in [\eta_0, \bar{c}]$. Counting the number of possible zeros of the function $g'(x)$ shows that at least one of the points $\tilde{x}_i$ has to be a boundary point of the interval $[\eta_0, \bar{c}]$. Since $g(\eta_0) = 0$ it follows that $\tilde{x}_3 = \bar{c}$. The matrix

$$\tilde{F} = \left(e_i^T K\varphi(\tilde{x}_j)\right)_{i,j=1}^3 .$$

is nonsingular, because the functions φ_1, φ_2 and φ_3 generate a Chebyshev-system on the interval $[\eta_o, \bar{c}]$ and the matrix K is nonsingular. The following lemma gives the locally e_k- and E-optimal designs and will be proved in the Appendix.

Lemma 4

(1) *If $\eta_0 > 0$ is sufficiently small, then the design*

$$\zeta_E = \begin{pmatrix} \tilde{x}_1 & \tilde{x}_2 & \bar{c} \\ \tilde{\omega}_1 & \tilde{\omega}_2 & \tilde{\omega}_3 \end{pmatrix} \tag{5.4}$$

is a locally E-optimal design in the regression model $\beta^T K\varphi(x)$ on the interval $[\eta_0, \bar{c}]$. Here the support points are defined by the condition (5.2) and (5.3) and the weights $\tilde{\omega}_i$ are given by the formula

$$\tilde{\omega}_i = |\tilde{A}_i| \Big/ \sum_{j=1}^3 |\tilde{A}_j| \ , \tag{5.5}$$

where $\tilde{A}_i = e_i^T \tilde{F}^{-1} r$, $i = 1, 2, 3$, and $r = (r_1, r_2, r_3)^T$ denotes the vector of coefficients of the function g defined in (5.1) - (5.3).

(2) *For arbitrary $\bar{c} \leq c$, $\eta_0 > 0$ the design ζ_E defined by (5.4) is a locally E-optimal design for the regression model $\beta^T K\varphi(x)$ on the interval $[\eta_0, \bar{c}]$ if and only if the condition*

$$r^T \bar{M}(\zeta_E) r = (r^T (K^T K)^{-1} r)\lambda_{\min}(K\bar{M}(\zeta_E)K^T)$$

is satisfied, where $r = (r_1, r_2, r_3)^T$ denotes the vector of coefficients of the function g defined in (5.1) - (5.3). In this case the design ζ_E is the unique (locally) E-optimal design.

(3) *If* $\eta_0 > 0$ *is sufficiently small, then the design*

$$\zeta_{e_k} \;=\; \begin{pmatrix} \tilde{x}_1 & \tilde{x}_2 & \bar{c} \\ \tilde{\omega}_1(k) & \tilde{\omega}_2(k) & \tilde{\omega}_3(k) \end{pmatrix}.$$

is a locally e_k*-optimal design in the regression model* $\beta^T K\varphi(x)$ *on the interval* $[\eta_0, \bar{c}]$*. Here the support points are defined by the condition (5.2) and (5.3) and the weights* $\tilde{\omega}_i$ *are given by the formula*

$$\tilde{\omega}_i(k) \;=\; |\tilde{A}_{ik}| \Big/ \sum_{j=1}^{3} |\tilde{A}_{jk}|, \;\; k = 1, 2, 3,$$

with $\tilde{A}_{ij} = e_i^T \tilde{F}^{-1} e_j$ $(i, j = 1, 2, 3)$.

(4) *For arbitrary* $\eta_0 > 0$ *the design* ζ_{e_k} *is a locally* e_k*-optimal design for the regression model* $\beta^T K\varphi(x)$ *on the interval* $[\eta_0, \bar{c}]$ *if and only if*

$$\tilde{A}_{ik} \geq 0 \;, \quad i = 1, 2, 3 \;.$$

In this case the design ζ_{e_k} *is the unique (locally)* e_k*-optimal design.*

Note that it follows from Theorem 3 that a design ζ is locally $E-$ (e_k-) optimal for the regression model $\beta^T K\varphi(x)$ on the interval $[\eta_0, \bar{c}]$ if and only if the design ξ_ζ induced by the transformation (2.21) is locally $E-$ (e_k-) optimal for Monod model (1.1) on the interval $[0, T]$, where $T = t(\bar{c})$. Therefore, if no constraints are imposed on the desired real time operation, the locally e_k- and $E-$ optimal designs yield designs with prohibitively long experiments. This fact was also observed empirically for the E-criterion by Vanrolleghem et. al. (1995).

For the numerical construction of the E-optimal design ζ_E it is sufficient to maximize the function

$$Q(x_1, x_2) \;=\; \lambda_{min}(K^T \bar{M}(\zeta) K)$$

on the set

$$U = \{(x_1, x_2)^T; \eta_0 \leq x_1 \leq x_2 \leq \bar{c}\} \;,$$

where ζ is defined in (5.4) and (5.5). The design ζ_{e_k} can be constructed in a similar way and the optimality can be checked by the necessary and sufficient characterizations (2) and (4) in Lemma 4. We conclude this section with a comparison of locally D, E- and e_2-optimal designs. The characteristics of the locally optimal designs can be found in Table 5.1, which shows the design (4.4) derived from the D-optimal design and the E- and e_2- optimal design

$$\xi_E^* \;=\; \begin{pmatrix} t_{1E}^* & t_{2E}^* & 2t_{2D}^* \\ w_1 & w_2 & w_3 \end{pmatrix}$$

$$\xi_{e_2}^* \;=\; \begin{pmatrix} t_{1E}^* & t_{2E}^* & 2t_{2D}^* \\ m_1 & m_2 & ,m_3 \end{pmatrix}$$

on the interval $[0, 2t_{2D}^*]$. Note that ξ_E^* and $\xi_{e_2}^*$ have the same support points. In all cases considered in our numerical study the necessary and sufficient conditions from part (2) and (4) of Lemma 4

were fulfilled which leads to the conjecture that the designs defined by formula (5.4) and (5.6) are in fact locally E- and e_k-optimal on any interval $[0, T]$ and for any initial condition $\eta_0 > 0$.
The designs were firstly compared by their D-efficiencies

$$I(\xi) := \left(\frac{\det M(\xi)}{\det M(\xi_D^*)} \right)^{1/3} \tag{5.6}$$

and we did not observe substantial differences with respect to this criterion [see the last two columns in Table 5.1]. For a more refined comparison we calculated the asymptotic efficiencies

$$\begin{aligned} d_i &= \frac{(M^{-1}(\xi_E^*))_{ii}}{(M^{-1}(\xi_D^*))_{ii}} \quad , \ i = 1, 2, 3 \\ \hat{d}_i &= \frac{(M^{-1}(\xi_{e_2}^*))_{ii}}{(M^{-1}(\xi_D^*))_{ii}} \quad , \ i = 1, 2, 3 \end{aligned} \tag{5.7}$$

for estimating the individual coefficients in the Monod model. We observe a very similar behaviour of the E- and e_2-optimal design, which provide more efficient estimates for the parameters ϑ_1 and ϑ_2 as the design ξ_D^* derived from the D-optimal. On the other hand this design is more efficient for the estimation of the parameter ϑ_3. However, if improvement of accuracy in the estimation of the parameters ϑ_1 and ϑ_2 is considered as more important the E- and e_2-optimal design have some advantages.

6 Conclusions

In this paper we discussed the problem of estimation and design of experiments for the Monod model, which is widely used for describing microbial growth and substrate degradations in activated sludge wastewater treatment, pharmacokinetics and plant physiology. Although the model is only defined implicitly by a differential equation we demonstrate that asymptotic inference is possible. It is shown that the least squares method provides consistent estimates of the parameters in this model. The asymptotic covariance matrix derived from these estimators can be used for the construction and evaluation of designs provided that the sample size is larger than $N = 30$ and locally D-, e_2- and E-optimal designs are determined. A numerical study shows that uniform designs are a good choice for an initial study. The performance of these designs is rather robust with respect to misspecification of the initial values for the parameters and its efficiency with respect to the locally D-optimal designs is approximately 50% for the estimation of the important parameters in the Monod model. Moreover, locally E- (and also e_2-) optimal designs have similar properties as the designs determined by the D-criterion. They are slightly better than locally D-optimal designs but require different numbers of experiments under different experimental conditions.
Locally optimal designs are rather sensitive with respect to a specification of initial values for the parameter in the model if no prior information is available. However, if certain intervals for the nonlinear parameters can be specified based on microbiological background, we are able to construct locally optimal designs, which are robust with respect to misspecification of the initial parameters and allow efficient estimation of the parameters in the Monod model [see Remark 4.1

Table 5.1: Comparison of D-, E-, e_2-optimal designs for various values of $b = \vartheta_2\vartheta_3$. The E- and $\xi^*_{e_2}$-optimal design have the same suport $t^*_{1E}, t^*_{2E}, 2t^*_{2D}$ but different weights w_1, w_2, w_3 and m_1, m_2, m_3, respectively. The D-optimal design has equal masses at the points $t^*_{1D}, t^*_{2D}, 2t^*_{2D}$ and the efficiencies d_i and $\hat{d}_i$ are defined in (5.7), while $I(\xi)$ is defined by (5.6).

b	0.1	0.25	0.5	0.75	1	1.5	2
t^*_{1D}	2.92	3.29	3.95	4.62	5.30	6.67	8.04
t^*_{2D}	3.51	4.25	5.45	6.64	7.82	10.18	12.54
$(M^{-1})_{11}$	42.3	110.5	269.8	486.4	760.4	1480.2	2428.5
$(M^{-1})_{22}$	24.9	103.0	423.9	1136.5	2460.0	8055.4	19914.0
$(M^{-1})_{33}$	3.00	3.00	3.00	3.00	3.00	3.00	3.00
t^*_{1E}	2.65	2.94	3.47	4.03	4.59	5.73	6.88
t^*_{2E}	3.52	4.26	5.48	6.68	7.88	10.26	12.65
w_1	0.45	0.42	0.41	0.40	0.40	0.40	0.40
w_2	0.35	0.36	0.37	0.37	0.37	0.37	0.37
w_3	0.20	0.21	0.22	0.22	0.23	0.23	0.23
m_1	0.40	0.39	0.39	0.39	0.39	0.39	0.40
m_2	0.39	0.39	0.38	0.38	0.38	0.38	0.37
m_3	0.21	0.22	0.23	0.23	0.23	0.23	0.23
d_1	0.80	0.82	0.83	0.84	0.84	0.84	0.84
d_2	0.84	0.85	0.85	0.85	0.85	0.85	0.85
d_3	1.66	1.58	1.52	1.49	1.48	1.47	1.47
$\hat{d}_1$	0.82	0.83	0.84	0.84	0.84	0.84	0.84
$\hat{d}_2$	0.83	0.84	0.85	0.85	0.85	0.85	0.85
$\hat{d}_3$	1.60	1.50	1.47	1.46	1.46	1.46	1.46
$I(\xi^*_E)$	1.12	1.11	1.10	1.10	1.10	1.10	1.10
$I(\xi^*_{e_2})$	1.11	1.10	1.10	1.10	1.10	1.10	1.10

for details]. Morover, in many applications of the Monod model the statistical inference is made in two steps and some information regarding the parameters in the model is available from the first step, which can be used for the construction of efficient designs in the second step [see e.g. Merkel et. al. (1996)]. If prior information for the values of the parameters is available from the data obtained in the first step, the efficiency of the estimates can be substantially improved by using locally D-, e_2- or E-optimal designs. Additionally the results of the present paper provide further insight in the optimal design problem for the Monod model, which will be useful for the construction of Bayesian- or robust optimal designs.

7 Appendix

We will begin proving a Chebyshev property for a systems of functions, which will be crucial for a proof of the statements in Section 2 - 5. Recall that a system of functions $\psi_1(t), \ldots, \psi_m(t)$ is called a Chebyshev system (T-system) on an interval $[\alpha, \beta]$ if

$$\det\left(\psi_i(t_j)\right)_{i,j=1}^m > 0$$

for any $\alpha \leq t_1 < \ldots < t_m \leq \beta$ [see Karlin and Studden (1966), Ch.1] . The main property of a Chebyshev-system is that any non trivial linear combination $\sum_{i=1}^m \alpha_i \psi_i(t)$ of the functions $\psi_1, \ldots, \psi_m$ has at most $m-1$ distinct roots in the interval $[\alpha, \beta]$.

Proposition 1

a) *The system of functions*

$$\left\{ \frac{1}{u}, \frac{1}{b_1 - u}, \frac{1}{(b_1 - u)(b_2 - u)} \right\}$$

is a Chebyshev system on any interval $[a, b] \subset (0, \infty]$, *whenever* $b_1, b_2 \geq b$;

b) *If a system of functions* $\psi_1(x), \ldots, \psi_m(x)$ *is a Chebyshev system on the interval* $[\eta_0, b]$, *then the system of functions*

$$\int_{\eta_0}^u \psi_1(x)dx, \ldots, \int_{\eta_0}^u \psi_m(x)dx$$

is also a Chebyshev system on any interval $[a, b] \subset (\eta_0, b]$.

Proof.

a) Consider the determinant

$$J = \det\left(\psi_i(u_j)\right)_{i,j=1}^3, \ a \leq u_1 < u_2 < u_3 \leq b,$$

where $\psi_1(u) = 1/u$, $\psi_2(u) = 1/(b_1 - u)$, $\psi_3(u) = 1/[(b_1 - u)(b_2 - u)]$. Let us multiply i-th row with $u_i(b_1 - u_i)(b_2 - u_i)$, $i = 1, 2, 3$. By a linear transformation of the columns the resulting determinant can be reduced to the Vandermonde determinant

$$\det(u_j^{i-1})_{i,j=1}^3 = \prod_{j<i}(u_i - u_j) > 0$$

Hence $J > 0$ and the system $\psi_1(u), \psi_2(x), \psi_3(x)$ is a Chebyshev system on the interval $[a, b] \subset (0, \infty]$.

b) Let us suppose that assertion b) of Proposition 1 is not true. Then there exist real numbers $\alpha_1, \ldots, \alpha_m$, not all equal to zero such that the linear combination

$$\sum_{i=1}^m \alpha_i \int_{\eta_0}^u \psi_i(x)dx =: q(u)$$

has at least m distinct zeros in the interval $[a, b] \subset (\eta_0, b]$. Moreover, we have $q(\eta_0) = 0$, and therefore the derivative

$$q'(u) = \sum_{i=1}^m \alpha_i \psi_i(u)$$

has m different roots in the interval $[\eta_0, b]$, which contradicts to the Chebyshev property of the system $\{\psi_i(u)\}$. ∎

7.1 Proof of Lemma 1:

The derivative of η is positive at the point $t = 0$ and by continuity it must also be positive in a neighbourhood of the origin. Moreover, the function $\bar{\eta}(t) \equiv c$ is obviously a solution of the differential equation (1.1). By the implicit function theorem [see Gunning and Rossi (1965)] it follows that the differential equation (1.1) can not have two differentiable solutions which coincide at a point. Thus the function η should be less than c for any $t > 0$ and must be increasing. Consequently there exists the limit $\lim_{t\to\infty} \eta(t)$. Now equation (1.1) implies that the limit of the derivative of η also exists. Because η is a bounded function it follows from (2.9) that

$$\begin{aligned} \lim_{t\to\infty} \eta'(t) &= 0 \\ \lim_{t\to\infty} \eta(t) &= c, \end{aligned}$$

which completes the proof of Lemma 1. ∎

7.2 Proof of Theorem 1:

Recall the definiton of the set $\Omega = \{\theta = (\vartheta_1, \vartheta_2, \vartheta_3)^T : \vartheta_i > 0,\ i = 1, 2, 3\}$ in Theorem 1, then the assertion of the theorem follows from results of Jennrich (1969) and the following Lemma.

Lemma 5 *Let $\eta(t, \theta)$ denote the regression function determined by (1.1)-(1.3).*

(a) For any fixed vector $\theta \in \Omega$ there exist the derivatives

$$\frac{\partial \eta}{\partial \vartheta_i}(t, \theta),\ \frac{\partial^2 \eta}{\partial \vartheta_i \partial \vartheta_j}(t, \theta),\ i, j = 1, 2, 3.$$

(b) For any fixed $\theta^0 \in \Omega$ the function

$$g(\theta) = \sum_{j=1}^{n} (\eta(t_j, \theta) - \eta(t_j, \theta^0))^2$$

with $0 \le t_1 < \ldots < t_n < \infty$ and $n \ge 3$ attains its minimum value (equal to zero) in the set Ω if and only if $\theta = \theta^0$.

(c) For any $\theta^0 \in \Omega$ and for any design ξ with more than 3 support points it follows that

$$\det M(\xi, \theta^0) \neq 0.$$

Proof. Statement (a) is an immediate consequence of the identities (2.11) and (2.15). Let us suppose that condition (b) is not valid. Then as $n \ge 3$ there exist two vectors $\theta_{(1)}$ and $\theta_{(2)}$, such that

$$\eta(t_i, \theta_{(1)}) = \eta(t_i, \theta_{(2)}),\ i = 1, 2, 3. \tag{7.1}$$

Define $\eta_i = \eta(t_i, \theta_{(1)})$ $(i = 1, 2, 3)$ and consider the two functions

$$t_{(j)}(x) = t(x, \theta_{(j)}) = \eta^{-1}(x, \theta_{(j)}),\ \ j = 1, 2.$$

Due to (7.1) we have $t_{(1)}(\eta_i) = t_{(2)}(\eta_i) = t_i$ $(i = 1, 2, 3)$, and observing (2.11) we obtain

$$t_i = a_1(\theta_{(j)}) \int_{\eta_0}^{\eta_i} \frac{1}{u} du + a_2(\theta_{(j)}) \int_{\eta_0}^{\eta_i} \frac{1}{a_3(\theta_{(j)}) - u} du,$$

$i = 1, 2, 3$, $j = 1, 2$, where the constants $a_i(\theta)$ are given by

$$\begin{aligned} a_1(\theta) &= \frac{1+b}{\vartheta_1} &= \frac{1}{\vartheta_1} \frac{s_0\vartheta_3 + \eta_0 + \vartheta_2\vartheta_3}{s_0\vartheta_3 + \eta_0}, \\ a_2(\theta) &= \frac{b}{\vartheta_1} &= \frac{1}{\vartheta_1} \frac{\vartheta_2\vartheta_3}{s_0\vartheta_3 + \eta_0}, \\ a_3(\theta) &= c &= s_0\vartheta_3 + \eta_0. \end{aligned}$$

It is easy to verify that the conditions $a_i(\theta_{(1)}) = a_i(\theta_{(2)})$, $i = 1, 2, 3$ imply $\theta_{(1)} = \theta_{(2)}$. Consequently, there exist two different vectors

$$a_{(j)} = (a_{1(j)}, a_{2(j)}, a_{3(j)}) = (a_1(\theta_{(j)}), a_2(\theta_{(j)}), a_3(\theta_{(j)}))^T,\ j = 1, 2,$$

such that the equations

$$a_{1(j)} \int_{\eta_0}^{\eta_i} \frac{1}{u} du + a_{2(j)} \int_{\eta_0}^{\eta_i} \frac{1}{a_{3(j)} - u} du = t_i,$$

are satisfied for all $i = 1, 2, 3$, $j = 1, 2$. Subtracting the equalities for $j = 2$ from the equalities for $j = 1$ we obtain

$$\begin{aligned} \Delta_1 \int_{\eta_0}^{\eta_i} \tfrac{1}{u} du &- a_{2(1)} \int_{\eta_0}^{\eta_i} \tfrac{\Delta_3}{(a_{3(1)}-u)(a_{3(2)}-u)} du + \\ &+ \quad \Delta_2 \int_{\eta_0}^{\eta_i} \tfrac{1}{a_{3(1)}-u} du = 0, \end{aligned} \tag{7.2}$$

for some constants $\Delta_i = a_{i(1)} - a_{i(2)}$, $i = 1, 2, 3$. From Proposition 1 it follows that the functions

$$\psi_1(x) = \int_{\eta_0}^{x} \frac{du}{u},\ \psi_2(x) = \int_{\eta_0}^{x} \frac{du}{b_1 - u},\ \psi_3(x) = \int_{\eta_0}^{x} \frac{du}{(b_1 - u)(b_2 - u)},$$

form a Chebyshev-system on an interval $(\eta_0, d]$ whenever $b_1, b_2 \geq d$. Inserting $d = \min\{a_{3(1)}, a_{3(2)}\}$, $b_1 = a_{3(1)}$, $b_2 = a_{3(2)}$, $\alpha_1 = \Delta_1$, $\alpha_2 = -a_{2(1)}\Delta_3$, $\alpha_3 = \Delta_2$ we obtain from (7.2) that

$$\sum_{i=1}^{3} \alpha_i \psi_i(\eta_j) = 0,\ j = 1, 2, 3,$$

where the coefficients α_i are not all equal to zero and we have from Lemma 1 $\eta_i \leq \min\{a_{3(1)}, a_{3(2)}\}$ $(i = 1, 2, 3)$. This equality contradicts to the main property of Chebyshev-systems and the proof of the assertion (b) is completed.

Let us finally prove that assertion (c) is valid. From Proposition 1 it follows that the functions $\varphi_1(x)$, $\varphi_2(x)$, $\varphi_3(x)$ defined in (2.18) generate a Chebyshev-system on the interval $(0, \infty]$. If $n = 3$ a direct calculation shows shows that

$$M(\xi, \theta^0) = F^T W F, \tag{7.3}$$

where $W = \text{diag}\{w_1, w_2, w_3\}$,

$$F = \left(\frac{\partial \eta(t_i, \theta^0)}{\partial \vartheta_j} \right)_{i,j=1}^{3}.$$

Consequently it follows that

$$\det M(\xi, \theta^0) = w_1 w_2 w_3 (\det F)^2, \tag{7.4}$$

while due to equality (2.16) we have

$$\det F = \det K \cdot \det \left(\varphi_j(\eta_i)\right)_{i,j=1}^{3} \neq 0,$$

where we used the Chebyshev property of the system $\{\varphi_i(x)\}_{i=1}^3$. In the general case $n > 3$ let

$$\alpha = (i_1, i_2, i_3),\ \xi_\alpha = \begin{pmatrix} t_{i_1} & t_{i_2} & t_{i_3} \\ w_{i_1} & w_{i_2} & w_{i_3} \end{pmatrix},\ 1 \le i_1 < i_2 < i_3 \le n,$$

and denote by τ the set of all different multiindices α, then the Cauchy-Binet formula shows

$$\det M(\xi, \theta^0) = \sum_{\alpha \in \tau} \det M(\xi_\alpha, \theta^0), \tag{7.5}$$

and all terms in the sum on the right-hand side are positive. ∎

7.3 Proof of Lemma 2

It follows by a standard argument that the optimal weights (with respect to the D-criterion) are equal for any design with three support points. Consequently it is enough to verify that

$$\frac{\partial}{\partial x_3} \det F(x_1, x_2, x_3) > 0$$

for $\eta_0 \le x_1 < x_2 < x_3$, where $F(x_1, x_2, x_3) = \det(\varphi_i(x_j))_{i,j=1}^3$. To this end we introduce the notation

$$L(x) = \frac{\partial}{\partial x} \det F(x_1, x_2, x),$$

$$G(x) = \det \begin{pmatrix} f_1(x_1) & f_2(x_1) & f_3(x_1) \\ f_1(x_2) & f_2(x_2) & f_3(x_2) \\ f_1(x) & f_2(x) & f_3(x) \end{pmatrix}. \tag{7.6}$$

with $f_i(x) = \varphi_i(x)/v(x)$, then a direct computation shows that

$$0 \le \det F(x_1, x_2, x) = v(x_1)v(x_2)v(x)G(x)$$

and

$$L(x) = v(x_1)v(x_2)v'(x)G(x) + v(x_1)v(x_2)v(x)G'(x). \tag{7.7}$$

Consequently, we obtain

$$L(x) = v(x_1)v(x_2)v(x)G'(x)$$

for $x = x_1, x_2$. Observing (2.18) and (7.6) it follows that the function $G(x)$ has zeros at $x = \eta_0, x_1, x_2$. Hence there exist points u_1, u_2 with $\eta_0 < u_1 < x_1 < u_2 < x_2$ and $G'(u_i) = 0$, $i = 1, 2$. Moreover, the functions $\{f_i'(x)\}$ form a Chebyshev-system on the interval $[\eta_0, \bar{c}]$ and the functions $\{f_i(x)\}$ have the same property on the interval $(\eta_0, \bar{c}]$ by Proposition 1. Therefore the functions G' and G have at most two zeros in the interval $(\eta_0, \bar{c}]$ and we obtain from $G(x) > 0$, $G'(x) > 0$ for $x > x_2$ that both terms on the right-hand side of the equation (7.7) are positive for $x > x_2$. This implies $L(x) > 0$ for any $x > x_2$ and consequently the largest support point of the D-optimal 3-point design must be attained at the boundary, i.e. $x_3 = \bar{c}$. ∎

7.4 Proof of Lemma 3.

Proof of the assertion (1). Let $b > 0$ be an arbitrary fixed number, ζ be an arbitrary fixed design with at least $n \geq 3$ support points . Consider a transformation of the function of the equivalence theorem for D-optimality in the linear regression model $\beta^T \psi(x)$ defined by (3.4), that is

$$d(x) = d(x, \zeta) = d_1(x) - d_2(x),$$

where

$$\begin{aligned} d_1(x) &= d_1(x, \zeta) = \frac{1}{v^2(x)} \psi^T(x) \tilde{M}^{-1}(\zeta) \psi(x) - \frac{3b^2}{(1-x)^2}, \\ d_2(x) &= \frac{3}{v^2(x)} - \frac{3b^2}{(1-x)^2} \end{aligned}$$

and $v(x)$ is given by (2.19) for $c = 1$. We define $g(x) = g_1(x) - g_2(x)$, with

$$\begin{aligned} g_1(x) &= [x(1-x)^2 d_1'(x)]''', \\ g_2(x) &= [x(1-x)^2 d_2'(x)]'''. \end{aligned}$$

and introduce

$$\begin{pmatrix} A & B & C \\ B & G & E \\ C & E & F \end{pmatrix} = \tilde{M}^{-1}(\zeta),$$

where $\tilde{M}(\zeta)$ is the information matrix of the design ζ in the linear regression model $\beta^T \psi(x)$. Assume for a moment that $F = 3b^2$. In this case a direct computation gives

$$\begin{aligned} (1-x)^3 g_1(x) &= [-4Gx^2 + (10G + 2E)x - 6(E + G)], \\ x^5 g_2(x) &= [144(1+b)^2 - 36(2+b)(1+b)x]. \end{aligned} \tag{7.8}$$

It is not difficult to see that all elements of the matrix $\tilde{M}^{-1}(\zeta)$ are positive and, in particular, $G > 0$, $E > 0$. Moreover, for $x \in (0, 1)$ we have

$$g_2(x) > 0, \; g_1(x) < 0.$$

The first of these inequalities is obvious. The second inequality will be verified by considerating the roots $x_{(1)}$ and $x_{(2)}$ of the equation

$$-4Gx^2 + (10G + 2E)x - 6(E + G) = 0 \; ,$$

which are given by

$$x_{(1),(2)} = \frac{E + 5G \pm \sqrt{E^2 - 14EG + G^2}}{4G} .$$

Since E, G are positive it follows that $x_{(1)}, x_{(2)} > 1$ and we have from (7.8) $g_1(x) < 0$ for $x \in (0,1)$. Therefore the function $g(x)$ is strictly negative for $x \in (0,1)$ if the condition $F = 3b^2$ is valid.
Now let

$$\zeta = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ w_1 & w_2 & \dots & w_n \end{pmatrix}, \ 0 \le x_1 < \ldots < x_n \le 1.$$

be a D-optimal design. From the proof of Lemma 2 and the Cauchy-Binet formula it follows that $x_n = 1$. Notice that for $x \to 0$ we have $\psi(x) \to (0,0,0)^T$ and consequently the left boundary $x = 0$ of the design space is not a support point of a D-optimal design, i.e. $x_1 > 0$. By the equivalence theorem for D-optimality we have

$$\begin{aligned} \tilde{d}(x_i) &:= v^2(x_i)d(x_i) = \psi^T(x_i)\tilde{M}^{-1}(\zeta)\psi(x_i) - 3 = 0 \ , \quad i = 1, \ldots, n, \\ \tilde{d}'(x_i) &= 0 \ , \quad i = 1, \ldots, n-1. \end{aligned} \tag{7.9}$$

Let us assume that the design ζ contains $n > 3$ points. From (3.4) we obtain

$$\lim_{x \to 1} v^2(x)d(x) \ = \ 0$$

and it follows that $F = 3b^2$. Consequently, the arguments of the previous paragraph are applicable and we have from (7.9)

$$d'(x) = 0$$

for $x = x_1, x_2, x_3$ and for some points $\tilde{x}_1 \in (x_1, x_2)$, $\tilde{x}_2 \in (x_1, x_2)$. Thus the function $d'(x)$ and also the function $x(1-x)^2 d'(x)$ has at least five roots in the interval $(0,1)$. Hence $g(x)$ being the third derivative of the function $x(1-x)^2 d'(x)$ has at least $5 - 3 = 2$ zeros in the interval $(0,1)$. But we proved above that this is impossible. Thus the assumption $n \ge 4$ yields a contradiction and any D-optimal design for the regression model $\beta^T\psi(x)$ on the interval $[0,1]$ is supported at exactly three points.
Let us finally show that there exists a unique D-optimal design. The existence of a D-optimal design follows from continuity of the function $\psi(x)$ and compactness of the interval $[0,1]$. Assume that $\zeta_{(1)}$ and $\zeta_{(2)}$ are two different D-optimal designs. Then it follows by a standard concavity argument [see Fedorov (1972)] that the design $\tilde{\zeta} := \zeta_{(1)}/2 + \zeta_{(2)}/2$ is also D-optimal. But this is impossible because the design $\tilde{\zeta}$ contains more than three distinct points.

Proof of statement (2). From the proof of statement (1) it follows that the function

$$v^2(x)d(x)$$

has at most three maxima on the interval $[0,1]$ and statement (2) follows by a continuity argument.
■

7.5 Proof of Lemma 4.

We will only present a proof of assertion (1) and (2). The remaining statements regarding the e_k-optimal designs are proved similary. Let $h(x) = K\varphi(x)$, $s = (K^T)^{-1}r$ and define

$$(7.10) \qquad N(\zeta_E) \;=\; \int h(x)h^T(x)d\zeta_E(x) \;=\; K\bar{M}(\zeta_E)K^T \ ,$$

then it is easy to see that the assunption in (2) can be rewritten as

$$(7.11) \qquad r^T\bar{M}(\zeta_E)r \;=\; s^TN(\zeta_E)s = (s^Ts)\lambda_{\min}(N(\zeta_E)) \ .$$

From (5.1) - (5.3) we obtain

$$(7.12) \qquad s^Th(x_i) \;=\; (-1)^{i+1} \ , \quad i = 1,2,3$$

$$(7.13) \qquad s^Th(x) \;\leq\; 1 \ \ \forall x$$

and Elfvings theorem [Elfving (1952)] and (7.12) imply that the vector s is an eigenvector of the matrix $N(\zeta_E)$, that is $N(\zeta_E)s = \lambda s$. Observing (7.12) and (7.11) it follows that

$$1 \;=\; s^TN(\zeta_E)s \;=\; \lambda\cdot(s^Ts) \;=\; \frac{\lambda}{\lambda_{\min}(N(\zeta_E))}$$

and we obtain from (7.13) the inequality

$$h^T(x)\Big(\frac{ss^T}{s^Ts}\Big)h(x) \;\leq\; \frac{1}{s^Ts} \;=\; \lambda_{\min}(N(\zeta_E)) \ \ \forall x \ .$$

Part (2) of Lemma 4 now follows immediately from the corresponding equivalence theorems for E-criterion [see Pukelsheim (1993)]. The design is unique, because the function $g(x)$ defined in (5.1) is unique [see e.g. Karlin, Studden (1966), Ch. 1].
To prove (1) let us note that for $\eta_0 \to 0$ we have

$$\tilde{\varphi}^T(x) \;=\; v(x)\left(-\frac{\ln\frac{x}{\eta_0}}{\ln\eta_0}, \ln\frac{c-x}{c-\eta_0}, \frac{x-\eta_0}{c-x}\right) \;\sim\; v(x)\left(1, \ln\frac{c-x}{c-\eta_0}, \frac{x-\eta_0}{c-x}\right) \;=:\; \tilde{\psi}(x)$$

and that any subset with two functions of

$$\left\{v(x), v(x)\ln\frac{c-x}{c-\eta_0}, v(x)\frac{x-\eta_0}{c-x}\right\}$$

generates a Chebyshev-system on the interval $[\eta_0, \bar{c}]$. Define

$$\tilde{K} \;=\; \begin{pmatrix} \frac{1+b}{\vartheta_1}\ln\eta_0 & \frac{b}{\vartheta_1} & 0 \\ -\frac{b}{\vartheta_2}\ln\eta_0 & -\frac{b}{\vartheta_2} & 0 \\ -\frac{b\eta_0}{c\vartheta_3}\ln\eta_0 & -\frac{b\eta_0}{c\vartheta_3} & -\frac{b}{\vartheta_3} \end{pmatrix} .$$

and

$$\tilde{M}(\xi) \;:=\; \int \tilde{\psi}(x)\tilde{\psi}^T(x)d\xi(x) \ .$$

A straightforward but tedious calculation shows that for sufficiently small η_0 the sign pattern of the matrix

$$(\tilde{K}^T)^{-1}(\tilde{M}(\xi))^{-1}\tilde{K}^{-1}$$

is of the form

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

and consequently the matrix D has a simple eigenvalue for sufficently small η_0 [see Gantmacher (1959)]. It follows from general results on E-optimality [see Dette and Studden (1993)] that the E-optimal design for the vector of regression functions $\tilde{\psi}(x)$ is supported on the Chebyshev points. The assertion (1) of the Lemma is now obtained by a continuity argument. ∎

Acknowledgements. The work of V.B. Melas was partially supported by the Russian Foundation of Basic Research (grant No. 00-01-00495). The work of H. Dette and V.B. Melas was supported by the Deutsche Forschungsgemeinschaft (SFB 475: Komplexitätsreduktion in multivariaten Datenstrukturen).

References

[1] Atkinson, A.C. and Donev, (1992). Optimum Experimental Designs. Oxford: Clarendon Press.

[2] Atkinson, A.C., Demetrio, C.G.B., Zocchi, S.S. (1995). Optimum dose levels when males and females differ in response. *J. R. Stat. Soc.*, Ser. C 44, 213-226.

[3] Blok J. (1994). Classification of biodegradability by growth kinetic parameters. *Ecotoxicology and Environmental Safety* 27, 294–305.

[4] Box G.E.P., Lucas H.L. (1959). Designs of experiments in nonlinear situations. *Biometrika* 46, 77–90.

[5] Chaloner, K. and Larntz, K. (1989). Optimal Bayesian designs applied to logistic regression experiments. *J. Statist. Plann. Inference* 21, 191-208.

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

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

[8] Dette H., Studden, W.J. (1993). Geometry of E-optimality. *Ann. Stat.* 21, 416-433.

[9] Elfving (1952). Optimum allocation in linear regression theory. *Ann. Math. Stat.* 23, 255-262.

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

[11] Ford, I., Torsney, B., Wu, C.F.J. (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.

[12] Gantmacher (1959). Matrizenrechnung I, II. VEB Deutscher Verlag der Wissenschaften, Berlin.

[13] Gunning R.C., Rossi H. (1965). *Analitical functions of several complex variables.* Prentice-Hall, NewYork.

[14] Haines L.M. (1992). Optimal design for inverse quadratic polynomials. *South African Statistical Journal* 26, 25-41.

[15] Haines, L. M. (1993). Optimal design for nonlinear regression models. *Commun. Statist. A*, 22, 1613-1627.

[16] Haines, L. M. (1995). A geometric approach to optimal design for one-parameter nonlinear models. *J. R. Stat. Soc.,* Ser. B 57, 575-598.

[17] He Z., Studden W.J., Sun D. (1996). Optimal designs for rational models. *Ann. Statist.* 24, 2128 -2147.

[18] Holmberg, A. (1982). On the practical identifiability of microbial gowth models incorporating Michaelis-Menten type nonlinearities. *Math. Biosci.* 62, 23–43.

[19] Jennrich, R.I. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Stat.* 40, 633-643.

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

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

[22] Kiefer J., Wolfowitz J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* 14, 363–366.

[23] Melas V.B. (1978). Optimal designs for exponential regression // *Math. Operations forsh. Statist.* Ser. Statistics, 9., 45–59.

[24] Merkel W., Schwarz A., Fritz S., Reuss M. and Krauth K. (1996). New strategies for estimating kinetic parameters in anaerobic wastewater treatment plants. *Water Science and Technology* 34, 393–401.

[25] Ossenbruggen P. J., Spanjers H. and Klapwik A. (1996). Assessment of a two-step nitrification model for activated sludge. *Water Research* 30, 939–953.

[26] Pirt, S. J. (1975). Principles of microbe and cell cultivation. John Wiley & Sons, New York.

[27] Pukelsheim F. (1993). *Optimal designs of experiments.* Wiley, NewYork.

[28] Pukelsheim F. and Rieder S. (1992). Efficient rounding of approximate designs. *Biometrika* 79, 763-770.

[29] Pronzato, L. and Walter, E. (1985). Robust experimental design via stochastic approximation. *Math. Biosci.* 75, 103-120.

[30] Rasch D. (1990). Optimum experimental design in nonlinear regression.*Commun. Statist., Theor. Meth.* 19, 4789-4806.

[31] Silvey, S.D. (1980). Optimal Design. Chapman and Hall, London.

[32] Sitter, R., Torsney, B. (1995). Optimal designs for binary response experiments with two design variables. *Statist. Sinica* 5, 405-419.

[33] Vanrolleghem P.A., Van Daele M. and Dochain D. (1995). Practical identifiability of a biokinetic model of activated sludge respiration. *Water Research* 29, 2561–2570.

[34] Versyck, K.J., Bernaerts, K., Geeraerd, A.H., Van Impe, J.F. (1999). Introducing optimal experimental design in predictive modeling: A motivating example. *International Journal of Food Microbiology* 51, 39–51.