

Optimal designs for the Michaelis Menten model with correlated observations

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: holger.dette@rub.de

Joachim Kunert
TU Dortmund
Fakultät Statistik
44221 Dortmund, Germany
email: joachim.kunert@udo.edu

January 19, 2012

Abstract

In this paper we investigate the problem of designing experiments for weighted least squares analysis in the Michaelis Menten model. We study the structure of exact D -optimal designs in a model with an autoregressive error structure. Explicit results for locally D -optimal are derived for the case where 2 observations can be taken per subject. Additionally standardized maximin D -optimal designs are obtained in this case. The results illustrate the enormous difficulties to find exact optimal designs explicitly for nonlinear regression models with correlated observations.

Keywords and Phrases: Autoregressive errors; Michaelis Menten model; exact designs, locally D -optimal designs; standardized maximin optimal design.

AMS Subject Classification: 62K05

1 Introduction

The Michaelis-Menten model

$$\eta(x, a, b) = \frac{ax}{b+x} \tag{1.1}$$

is widely used as a first approximation to describing complex biological systems, as in the study of saturable phenomena in enzyme kinetics. Its applications include such important

areas as agriculture, biochemistry, biology, microbiology, toxicology, environmental science, nutrition science, bio-pharmaceutical studies. In many cases η represents the velocity of a chemical reaction, the parameter $a \geq 0$ denotes the maximum velocity, the predictor $x \geq 0$ reflects the concentration of a substrate and the parameter $b \geq 0$ is the half-saturated constant, the concentration x where the velocity is half-maximal.

Some exemplary applications of the Michaelis-Menten model in different disciplines can be found in Clench (1979) (conservation biology research), Butler and Wolkowicz (1985) (nutrient uptake study), Rong and Rappaport (1996) (environmental research), Park et al. (2005) (pharmacokinetics) or Yu and Gu (2007) (agriculture) among many others. Because of its importance the problem of designing experiments for statistical analysis with the Michaelis-Menten model has found considerable attention in the literature. Early work dates back to Dunn (1988) and Rasch (1990), who determined locally D -optimal designs for maximum likelihood or least squares estimation in the Michaelis-Menten model for independent observations, respectively. Since this pioneering work several authors have continued to construct optimal designs in this model [see Dette and Wong (1999); Lopez-Fidalgo and Wong (2002) or Dette and Biedermann (2003) among many others], and a recent review about the state of the art in design of experiments for independent (and uncorrelated) data can be found in Dette et al. (2010). On the other hand the Michaelis-Menten is very often used for statistical analysis of dependent data. Typically, the explanatory variable x here represents time and dependencies appear naturally because observations are taken at the same subject [see López et al. (2000) for a typical example]. Despite of the wide spread use of the Michaelis-Menten model for analyzing correlated data - to the best knowledge of the authors - design issues in this context have not been addressed so far.

On the other hand, the problem of designing experiments for correlated data has found considerable interest in the literature, where mainly asymptotic optimal designs for location models have been studied. Early work dates back to Sacks and Ylvisaker (1966, 1968), which was extended and further developed by Bickel and Herzberg (1979); Bickel et al. (1981). More recent references are Dette et al. (2008, 2009); Zhigljavsky et al. (2010) who considered the design problem for correlated data in constant, linear and quadratic regression models under various dependence structures. However, less results are available for nonlinear regression models [see Pepelyshev (2007) or López-Fidalgo et al. (2008) for some results in this direction], in particular we are not aware of optimum designs for the Michaelis-Menten model in the case of correlated data.

It is the purpose of the present paper to provide insight into this design problem for the autoregressive error structure. We study D -optimal design problems for weighted least squares analysis of correlated data with the Michaelis-Menten model. It turns out that the design problem for the Michaelis-Menten model in the case of dependencies is substantially more complicated. In Section 2 we introduce the basic terminology and present some explicit results on locally D -optimal designs for the Michaelis-Menten model with two (correlated)

measurements per subject. These designs require prior information about the non-linear parameters of the model [see Chernoff (1953)] and usually serve as benchmark for the commonly used designs. They also form the basis for more robust design strategies as Bayesian or standardized maxmin optimal designs [see Chaloner and Verdinelli (1995) or Dette (1997)]. In the case of uncorrelated data, one of the runs of an optimal design always has to be taken at the largest point of the design space. One important result of this paper is that in the case of dependencies this statement is not necessarily correct. We derive a sufficient condition which guarantees that one of the observations of the locally D -optimal design has to be taken at the largest point of the design space. We demonstrate that in general this is not the case. In particular we give examples of locally D -optimal designs, where both observations have to be taken at two interior points of the design space. In Section 3 we present some numerical results of locally D -optimal designs in the case, where 3 or 4 observations are taken for each subject and make a similar observation. Section 4 is devoted to the problem of constructing standardized maximin optimal designs [see Dette (1997)], which require less precise knowledge about the unknown parameters. Finally some technical details are given in the Appendix.

2 Locally D -optimal designs

Consider the non-linear regression model (1.1) and assume that n subjects are investigated, where for each subject m replications are allowed, that is

$$Y_{i,j} = \eta(x_{i,j}, a, b) + \varepsilon_{i,j} = \frac{ax_{i,j}}{b + x_{i,j}} + \varepsilon_{i,j}, \quad j = 1, \dots, m; \quad i = 1, \dots, n \quad (2.1)$$

where $\varepsilon_{1,1}, \dots, \varepsilon_{n,n}$ are normally distributed random variables. We assume that observations from different subjects are independent, but dependencies may appear for the observations from the same subject that is

$$\text{Cov}(\varepsilon_{i,j}, \varepsilon_{i',k}) = \delta_{i,i'} \sigma^2 \lambda^{|x_{i,j} - x_{i',k}|}, \quad j, k = 1, \dots, m; \quad i, i' = 1, \dots, n \quad (2.2)$$

where $\delta_{i,i'}$ denotes Kronecker's delta. Throughout this paper we are interested in optimal designs for maximum likelihood estimation of the parameter (a, b) in the Michaelis Menten model (2.1), where we consider the parameters σ^2 and $\lambda \in (0, 1)$ as nuisance parameters. A standard calculation shows that for a large sample size the covariance matrix of the maximum likelihood estimator can be approximated (up to a constant, which does not depend on the design) by the matrix M^{-1} , where

$$M = \sum_{i=1}^n z_i^T \Sigma_i^{-1} z_i \quad (2.3)$$

where the matrices $z_i \in \mathbb{R}^{2 \times m}$ $\Sigma_i \in \mathbb{R}^{m \times m}$ are defined by

$$\begin{aligned} z_i &= \left(\left(\frac{x_{i,j}}{b+x_{i,j}}, -\frac{ax_{i,j}}{(b+x_{i,j})^2} \right)_{j=1,\dots,m}^T \right) \\ \Sigma_i &= \sigma^2 \left(\lambda^{|x_{i,j}-x_{i,k}|} \right)_{j,k=1,\dots,m} \end{aligned}$$

An optimal design maximizes a concave and monotone function of the matrix M defined (2.3) and there are numerous criteria which could be used to discriminate between competing designs [Pukelsheim (2006)]. Throughout this paper we consider the D -optimality criterion, which determines the design such that the determinant of the matrix (2.3) is maximal. It is easy to see that the D -optimal design does not depend on the parameters σ^2 and a and for this reason we put without loss of generality $\sigma^2 = a = 1$ throughout this paper. We also assume without loss of generality that the design space is given by the interval $[0, 1]$, D -optimal designs on other intervals can easily be obtained by a linear transformation. Note that a design maximizing the determinant of the matrix M still depends on the parameter b and λ . Therefore these designs require prior information about the non-linear parameters of the model and are called locally D -optimal designs [see Chernoff (1953)]. They usually serve as benchmark for the commonly used designs and locally optimal designs also form the basis for more robust design strategies as Bayesian or standardized maxmin optimal designs [see Chaloner and Verdinelli (1995) or Dette (1997)]. Robust designs will be discussed in Section 4.

We assume that all subjects are treated under the same experimental conditions and the observations of each subject have to be done at the same time, that is $x_{1,j} = x_{2,j} = \dots = x_{n,j}$. It is then easy to see that we may assume $n = 1$ for the determination of a D -optimal design without loss of generality. This simplifies the optimization problem substantially. Nevertheless a closed form of the determinant of the matrix M as a function of the experimental conditions $x_{1,1}, \dots, x_{1,m}$ is very complicated. In the following we will therefore discuss the optimization problem with two observations per subject.

2.1 Locally D -optimal designs with 2 observations per subject

In this section we consider the case $m = 2$ and use for the sake of simplicity the notations $u = x_{1,1}$ and $x = x_{1,2}$ and assume that $u < x$. A straightforward calculation yields for the determinant of the matrix M in (2.3)

$$C(x, u) = \frac{(D(x, u))^2}{1 - \lambda^{2(x-u)}}, \quad (2.4)$$

where the function D is given by

$$D(x, u) = \frac{xu(x-u)}{(b+x)^2(b+u)^2}$$

The following result provides the general structure of the locally D -optimal design in the case where the parameter b is not too small and the parameter λ in the correlation structure (2.2) is arbitrary.

Theorem. 2.1 *If $b \geq 1/3$, then the locally D -optimal for the Michaelis Menten model (1.1) with two observations per subject and correlation structure (2.2) is given by $x^* = 1$ and a second point $u^*(b, \lambda)$ defined as the unique solution of the equation*

$$\frac{b - (2b + 1)u}{u(1 - u)(b + u)} = \frac{\log(\lambda)\lambda^{2(1-u)}}{1 - \lambda^{2(1-u)}}. \quad (2.5)$$

In particular we have

$$u^*(b, 0) \leq u^*(b, \lambda) \leq u^*(b, 1),$$

where $u^*(b, 0)$ and $u^*(b, 1)$ do not depend on λ and are given by

$$u^*(b, 0) = \frac{b}{2b + 1} \quad \text{and} \quad u^*(b, 1) = \frac{1}{2} \left(2 + 3b - \sqrt{4 + 4b + 9b^2} \right), \quad (2.6)$$

respectively.

Proof. For fixed u the derivative of the function C with respect to x is given by

$$C'_1(x, u) = \frac{\partial}{\partial x} C(x, u) = \frac{2D(x, u)}{(1 - \lambda^{2(x-u)})^2} (D'_1(x, u)(1 - \lambda^{2(x-u)}) + D(x, u) \log(\lambda) \lambda^{2(x-u)}) \quad (2.7)$$

and $D'_1(x, u)$ the derivative of the function D with respect to x , that is

$$D'_1(x, u) = \frac{u}{(b + u)^2(b + x)^3} (x(2b + u) - ub).$$

An elementary calculation shows that for $r, s, t \geq 0$ the function

$$f(\lambda) = r(1 - \lambda^s) + t \log(\lambda) \lambda^s,$$

is nonnegative if $rs - t \geq 0$. Considering (2.7) this yields

$$rs - t = D_1(x, u)2(x - u) - D(x, u) = \frac{u(x - u)}{(b + u)^2(b + x)^3} (b(3x - 2u) - x(x - 2u))$$

which is nonnegative, whenever

$$b \geq \frac{x(x - 2u)}{3x - 2u}.$$

Because the maximum of the right hand side is $1/3$ (note that $u < x \leq 1$), the function defined in (2.7) is nonnegative, whenever $b \geq 1/3$, which proves the first assertion of Theorem 2.1. For the remaining part we introduce the notation

$$C(u) = C(1, u) = \frac{(D(u))^2}{1 - \lambda^{2(1-u)}}, \quad D(u) = D(1, u) = \frac{u(1 - u)}{(b + 1)^2(b + u)^2}$$

Checking the roots of the logarithmic derivative of the function $C(u)$ yields the equation

$$\frac{D'(u)}{D(u)} = \frac{\log(\lambda)\lambda^{2(1-u)}}{1 - \lambda^{2(1-u)}}.$$

which is the same as (2.5). The right hand side is a decreasing function of $\lambda \in (0, 1)$ which gives

$$\frac{\log(\lambda)\lambda^{2(1-u)}}{1 - \lambda^{2(1-u)}} > \lim_{\lambda \rightarrow 1} \frac{\log(\lambda)\lambda^{2(1-u)}}{1 - \lambda^{2(1-u)}} = -\frac{1}{2(1-u)}.$$

The solution of the equation

$$\frac{D'(u)}{D(u)} = -\frac{1}{2(1-u)}$$

is given by the point $u^*(b, 1)$ defined in (2.6). It is easy to see that $u^*(b, 0) \leq u^*(b, 1)$. Since, by definition, $D'(u) \geq 0$ if and only if $u \leq u^*(b, 0)$, it follows that

$$\frac{D'(u)}{D(u)} \geq 0 > \frac{\log(\lambda)\lambda^{2(1-u)}}{1 - \lambda^{2(u)}} > -\frac{1}{2(1-u)} \quad \text{if } u \leq u^*(b, 0), \quad (2.8)$$

while

$$\frac{D'(u)}{D(u)} < -\frac{1}{2(1-u)} < \frac{\log(\lambda)\lambda^{2(1-u)}}{1 - \lambda^{2(u)}} \quad \text{if } u^*(b, 1) < u \leq 1. \quad (2.9)$$

By continuity we therefore obtain from (2.9) and (2.8) the existence of a point $u^*(b, \lambda)$, which solves (2.5) and satisfies $u^*(b, 0) = \frac{b}{2b+1} \leq u^*(b, \lambda) \leq u^*(b, 1)$. The proof of Theorem 2.1 will now be concluded by showing that (2.5) has at most one solution. For this purpose we will show in Appendix 5.2 that the the derivative of the function

$$h(u) = \frac{1}{1-u} + \frac{\log(\lambda)\lambda^{2(1-u)}}{1 - \lambda^{2(1-u)}},$$

with respect to u is nonnegative for all $u \in [0, 1]$. Therefore the function h is strictly increasing. On the other hand a straightforward calculation yields that the function

$$\ell(u) = \frac{D'(u)}{D(u)} + \frac{1}{1-u} = \frac{1}{u} - \frac{2}{b+u}$$

is strictly decreasing in $u \in [0, u^*(b, 1)]$. Because equation (2.5) is equivalent to $h(u) = \ell(u)$ there is exactly one solution of (2.5) and the proof of Theorem 2.1 is completed. \square

In the following we are interested in the D -efficiency of designs using the largest point $x = 1$ and some u , that is

$$\text{eff}(u, b, \lambda) = \left(\frac{C(u)}{C(u^*(b, \lambda))} \right)^{1/2}.$$

The following result gives a lower bound for this efficiency. In particular we obtain a bound for the designs with experimental conditions $u^*(b, 1)$ and 1 which does not depend on the parameter λ .

b	0.33	0.67	1	2	5	10
$E(b)$	0.9854	0.9663	0.9519	0.9282	0.9069	0.8983

Table 1: Lower bounds for the D -efficiency of the design with experimental conditions $u^*(b, 1)$ and 1

Corollary. 2.2 *If $b \geq 1/3$ then we have for any $u < x \leq 1$*

$$\text{eff}(u, b, \lambda) \geq \frac{D(u)}{D(u^*(b, 0))} \left(\frac{1 - \lambda^{2(1-u^*(b, 1))}}{1 - \lambda^{2(1-u)}} \right)^{1/2}.$$

In particular

$$\text{eff}(u^*(b, 1), b, \lambda) \geq E(b) := \frac{D(u^*(b, 1))}{D(u^*(b, 0))}.$$

Proof. It is easy to see that the function $D(u) = D(1, u)$ is maximized for $u^*(b, 0)$. This implies for any $1 \geq x \geq u \geq 0$

$$C(x, u) = \frac{(D(x, u))^2}{1 - \lambda^{2(x-u)}} \leq \frac{(D(1, u^*(b, \lambda)))^2}{1 - \lambda^{2(1-u^*(b, \lambda))}} \leq \frac{(D(u^*(b, 0)))^2}{1 - \lambda^{2(1-u^*(b, \lambda))}} \leq \frac{(D(u^*(b, 0)))^2}{1 - \lambda^{2(1-u^*(b, 1))}},$$

and the assertion follows from the definition of the efficiency. \square

Table 3 gives the lower bound of the D -efficiency for the design which advises the experimenter to take observations at the points $u^*(b, 1)$ and 1 for some selected values of $b \geq 1/3$. Note that the "true" efficiency of this design is larger than the bound from the table.

For example, if $b = 5$ and $\lambda = 0.5$, numerical derivation gives $u^*(b, \lambda) = 0.5562$. Note that $u^*(5, 0) = 0.4545$ while $u^*(5, 1) = 0.6101$. We further calculate that $C(u^*(b, \lambda)) = 1.0736$, while $C(u^*(5, 1)) = 1.0556$. Therefore, the true efficiency of the design with experimental conditions $u^*(5, 1) = 0.6101$ and 1 is $\sqrt{1.0556/1.0736} = 0.9916$, while the lower bound is given by 0.9069. Nevertheless, the results from Table 3 show that the designs with experiments at $u^*(b, 1)$ and 1 perform rather well, even when the lower bound is used. We will use this bound in Section 4 to find standardized maximin optimal designs.

2.2 Improved bounds for b

In the previous section we have shown that in the case $b \geq \frac{1}{3}$ the locally D -optimal design for the Michaelis-Menten model with correlation structure (2.2) uses the largest point in the design space $x = 1$ as an experimental condition, independently of the parameter $\lambda \in [0, 1]$. In this section we show that the bound can be improved for a specific value of λ (the case $b \geq \frac{1}{3}$ actually corresponds to the case $\lambda \rightarrow 1$). We will demonstrate that there is a threshold

$b^*(\lambda)$ depending on $\lambda \in [0, 1]$, such that the locally D -optimal design is supported at the point $x = 1$ whenever $b > b^*(\lambda)$. To be precise, consider the logarithmic derivative of the function C defined in (2.4)

$$h_\lambda(x, u) = \frac{1}{2} \frac{\partial}{\partial x} \log C(x, u) = \frac{1}{x} + \frac{1}{x-u} - \frac{2}{b+x} + \frac{\lambda^{2(x-u)} \log(\lambda)}{1 - \lambda^{2(x-u)}}.$$

It then follows from the proof of Theorem 2.1 that it is only possible to have an optimal design (x^*, u^*) with $u^* < x^* < 1$ if $h_\lambda(x, u) < 0$ for some $u \leq x < 1$. We have seen in Theorem 2.1 that if $b \geq b^*(1) := \frac{1}{3}$, this is impossible for any $0 \leq \lambda < 1$. We now derive a smaller bound $b^*(\lambda)$ for a fixed λ .

For this purpose we show in Appendix 5.1 that for any $\lambda \in (0, 1)$ and for any $x \leq 1$, we have

$$h_\lambda(x, u) \geq h_\lambda(x, 0) = \frac{2b}{(b+x)x} + \frac{\lambda^{2x} \log \lambda}{1 - \lambda^{2x}}. \quad (2.10)$$

As a consequence of this inequality we obtain that the condition $h_\lambda(x, 0) > 0$ for all $x < 1$ implies that the locally D -optimal design (x^*, u^*) for the Michaelis-Menten model (2.1) with correlation structure (2.2) satisfies $x^* = 1$. Observing that the equation $h_\lambda(x, 0) = 0$ has exactly one solution with respect to b , the condition is equivalent to the inequality

$$b > b^*(\lambda) := \sup_{0 \leq x \leq 1} g_\lambda(x),$$

where the function g_λ is defined by

$$g_\lambda(x) = \frac{(-\log \lambda) \lambda^{2x} x^2}{2(1 - \lambda^{2x}) + (\log \lambda) \lambda^{2x} x}.$$

The following properties of g_λ are derived in Appendix 5.3

- (i) With the definition $g_\lambda(0) = 0$ the function g_λ is continuous on the interval $[0, 1]$ with $g_\lambda(x) > 0$ for all $x > 0$.
- (ii) For any fixed $x > 0$ the function $g_\lambda(x)$ is increasing with respect λ , where

$$\lim_{\lambda \rightarrow 0} g_\lambda(x) = 0.$$

- (iii) There exists a unique solution, say λ_0 , of the equation

$$-\log \lambda = \frac{4(1 - \lambda^2)}{4 + \lambda^2} \quad (2.11)$$

in the interval $(0, 1)$ with the following properties

λ	0.9	0.7	0.5	0.4	0.3	0.2	0.1	0.01
$b^*(\lambda)$	0.28963	0.20677	0.13061	0.09888	0.07526	0.05330	0.03935	0.01967
$g_\lambda(1)$	0.28963	0.20677	0.13061	0.09561	0.06331	0.03469	0.01177	0.00023

Table 2: *Some examples for the bound $b^*(\lambda)$ for selective vales $\lambda \in (0, 1)$.*

(iiia) For all $\lambda > \lambda_0$ it follows that

$$\max_{x \in (0,1)} g_\lambda(x) = g_\lambda(1) = \frac{(-\log \lambda)\lambda^2}{2(1 - \lambda^2) + (\log \lambda)\lambda^2},$$

(iiib) For all $\lambda \leq \lambda_0$ it follows that $\max_{x \in (0,1)} g_\lambda(x) = g_\lambda(x_0)$, where x_0 is the unique solution of the equation

$$(-\log \lambda)x = \frac{4(1 - \lambda^{2x})}{4 + \lambda^{2x}}.$$

in the interval $(0, 1)$.

A numerical calculation gives $\lambda_0 \approx 0.4863$. In Table 2 we show some representative values of $b^*(\lambda)$ and (for the sake of comparison) the values $g_\lambda(1)$ for some $\lambda \in (0, 1)$ (these differ, whenever $\lambda < \lambda_0$). We observe that $b^*(\lambda)$ and $g_\lambda(1)$ are increasing in λ , both with limits 0 and $1/3$ if λ approaches 0 or 1, respectively. If $b < b^*(\lambda)$, then the locally D -optimal design does not necessarily use $x^* = 1$ as experimental condition. Consider as an example the case $\lambda = 0.01$ and $b = 0.01$. We have numerically derived the locally D -optimal design which is given by $x^* = 0.05792$ and $u^* = 0.008186$ with corresponding value $C(x^*, u^*) = 650.0$. Figure 1 shows the function $C(x, u^*)$ for $x \in (u^*, 1]$. We clearly observe that the global maximum is attained at an interior point of the interval $(0, 1)$.

3 Designs with more observations per subject

In the case where more than 2 observations can be taken per subject, locally D -optimal designs have to be found numerically. Exemplarily, we present in Table 3 and 4 locally D -optimal designs for the Michaelis-Menten model with correlation structure (2.2) where 3 or 4 observations can be taken per subject. The results for 3 observations per subject are listed in Table 3 for various values of the parameter b . We observe that in all cases the largest point in the design space is an optimal experimental condition. Interestingly, in the case $\lambda = 0.5$ and $\lambda = 0.9$ the locally D -optimal designs advise the experimenter to take observations at the point 0. Despite the fact that at this point there is no information regarding the parameters in the mean of the response, information from the correlation between different observations can be used to improve the accuracy of the estimates for the parameters a and b . The

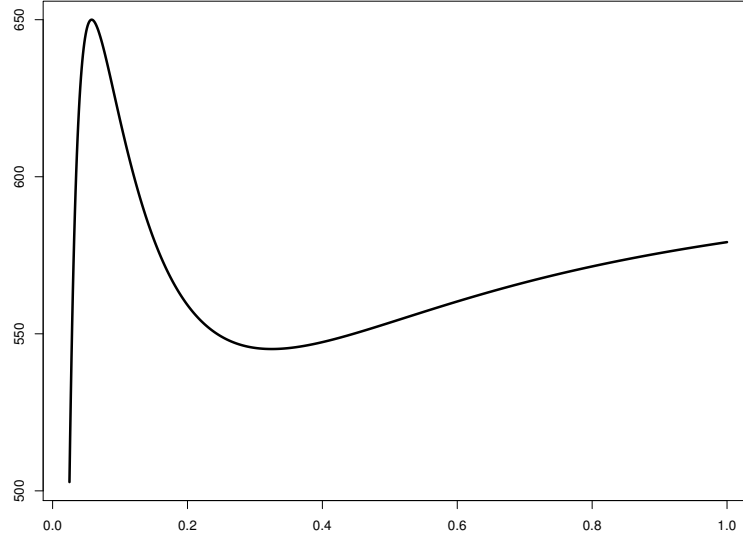


Figure 1: *The function $C(x, u^*)$ in the case $\lambda = 0.01$ and $b = 0.01$, where $u^* = 0.008186$*

corresponding results for 4 observations per subject are presented in Table 4 and show a similar behavior. It is interesting to note that for a large correlation ($\lambda = 0.5$ or $\lambda = 0.9$) the support points in the interior of the design space are increasing with the parameter b . On the other hand, if the correlation is smaller this is not necessarily the case (see the results for $\lambda = 0.1$).

b	$\lambda = 0.1$			$\lambda = 0.5$			$\lambda = 0.9$		
0.2	.135	.536	1.000	.000	.016	1.000	.000	.056	1.000
0.7	.244	.496	1.000	.000	.176	1.000	.000	.172	1.000
1.2	.292	.549	1.000	.000	.249	1.000	.000	.244	1.000
1.7	.321	.581	1.000	.000	.297	1.000	.000	.292	1.000
2.2	.339	.602	1.000	.000	.331	1.000	.000	.326	1.000
2.7	.352	.617	1.000	.000	.355	1.000	.000	.350	1.000

Table 3: *Locally D-optimal designs for the Michaelis-Menten model with 3 observations per subject.*

b	$\lambda = 0.1$				$\lambda = 0.5$				$\lambda = 0.9$			
0.2	.115	.229	.611	1.000	.000	.039	.129	1.000	.000	.052	.288	1.000
0.7	.219	.379	.648	1.000	.000	.101	.281	1.000	.000	.098	.272	1.000
1.2	.000	.179	.454	1.000	.000	.146	.384	1.000	.000	.143	.375	1.000
1.7	.000	.212	.501	1.000	.000	.178	.448	1.000	.000	.174	.440	1.000
2.2	.000	.234	.531	1.000	.000	.202	.489	1.000	.000	.197	.483	1.000
2.7	.000	.250	.550	1.000	.000	.219	.518	1.000	.000	.215	.513	1.000

Table 4: *Locally D-optimal designs for the Michaelis-Menten model with 4 observations per subject.*

4 Standardized maximin D -optimal designs

In order to obtain designs which are robust with respect to miss-specification of the unknown parameters we will construct standardized maximin optimal designs, which maximize a minimum efficiency calculated with respect to a given interval for the parameters in the optimality criterion [see Dette (1997)]. It is demonstrated by several authors that these designs are on the one hand robust and on the other hand have reasonable efficiency in regions of parameters over which the minimum is calculated. In the present context we concentrate on the case $b \geq 1/3$ and $m = 2$ and will work with the lower bound for the efficiency

$$E(u, b, \lambda) = \frac{D(u)}{D(u^*(b, 0))} \sqrt{\frac{1 - \lambda^{2(1-u^*(b,1))}}{1 - \lambda^{2(1-u)}}} \quad (4.1)$$

specified in Corollary 2.2, because an explicit solution of the locally D -optimal design problem is not available. To be precise, assume that $1/3 \leq p \leq q$, then we call a design with experimental conditions u and 1 standardized maximin optimal if it maximizes the function

$$\Phi(u) = \min_{p \leq b \leq q} \inf_{0 \leq \lambda < 1} E(u, b, \lambda). \quad (4.2)$$

The following result specifies the maximum of the function Φ .

Theorem. 4.1 *The maximum of the function Φ defined in (4.2) is given by the unique solution u^* of the equation*

$$E(u, p, 0) = E(u, q, 1) \quad (4.3)$$

in the interval $(0, 1)$, which is equivalent to

$$\frac{(q+u)^4}{(p+u)^4} = \left(\frac{q(q+1)}{p(p+1)} \right)^2 \frac{\sqrt{4+4q+9q^2-3q}}{2(1-u)}. \quad (4.4)$$

Proof. First note that the function

$$\lambda \rightarrow \frac{1 - \lambda^{2(1-\tilde{u})}}{1 - \lambda^{2(1-u)}}$$

corresponding to the second factor in (4.1) is decreasing and increasing if $u < \tilde{u}$ and $u > \tilde{u}$, respectively. This implies $\inf_{0 \leq \lambda < 1} E(u, b, \lambda) = \min\{E(u, b, 0), E(u, b, 1)\}$, where

$$E(u, b, 0) = \frac{4b(b+1)u(1-u)}{(b+u)^2},$$

$$E(u, b, 1) = \lim_{\lambda \rightarrow 1} E(u, b, \lambda) = \frac{4b(b+1)u(1-u)}{(b+u)^2} \sqrt{\frac{\sqrt{4+4b+9b^2}-3b}{2(1-u)}}.$$

Therefore the optimization problem reduces to finding a $u \in (0, 1)$ to maximize

$$\Phi(u) = \min_{p \leq b \leq q} \min\{E(u, b, 0), E(u, b, 1)\}.$$

This determination is now performed in two steps:

- (i) We show that the solution u^* of (4.3) is well-defined, unique and maximizes $\min\{E(u, p, 0), E(u, q, 1)\}$.
- (ii) We show that $\Phi(u^*) = \min\{E(u^*, p, 0), E(u^*, q, 1)\}$.

The existence and uniqueness of u^* is shown as follows. If we consider $E(u, p, 0)$ and $E(u, q, 0)$ as functions of u , then $E(u, p, 0) > E(u, q, 0)$ if $u < u^0(p, q)$, $E(u, p, 0) < E(u, q, 0)$ if $u > u^0(p, q)$, and there is equality for $u = u^0(p, q)$, where

$$u^0(p, q) = u^0(q, p) = \frac{q\sqrt{p(p+1)} - p\sqrt{q(q+1)}}{\sqrt{q(q+1)} - \sqrt{p(p+1)}},$$

[see Dette and Biedermann (2003)].

It is easy to see that $u^0(p, q)$ is increasing with respect to p and we therefore obtain

$$u^0(p, q) \leq \lim_{p \rightarrow q} u^0(p, q) = u^*(q, 0) = \frac{q}{2q+1} < u^*(q, 1),$$

where $u^*(q, 1)$ is defined in (2.6). Similarly,

$$u^0(p, q) \geq \lim_{q \rightarrow p} u^0(p, q) = u^*(p, 0) = \frac{p}{2p+1}.$$

In particular, this means that $E(u^*(q, 1), p, 0) < E(u^*(q, 1), q, 0)$.

Note that $E(u, q, 0) > E(u, q, 1)$ if $u < u^*(q, 1)$ and $E(u, q, 0) < E(u, q, 1)$ if $u > u^*(q, 1)$, where $u^*(q, 1)$ is the solution of $E(u, q, 0) = E(u, q, 1)$. Therefore, we have

$$E(u, q, 1) < E(u, q, 0) \leq E(u, p, 0), \text{ for all } u \leq u^0(p, q)$$

while

$$E(u, q, 1) \geq E(u, q, 0) > E(u, p, 0), \text{ for all } u \geq u^*(q, 1).$$

Hence, $E(u, q, 1)$ and $E(u, p, 0)$ must intersect at least once in the interval $(u^0(p, q), u^*(q, 1))$. However, note that for any fixed b , $E(u, b, 0)$ is strictly increasing in u for $u < u^*(b, 0)$ and strictly decreasing for $u > u^*(b, 0)$, while $E(u, b, 1)$ is increasing for $u < u^*(b, 1)$ and decreasing for larger u . This implies that in the whole interval $(u^*(p, 0), u^*(q, 1)) \supset (u^0(p, q), u^*(q, 1))$ the function $E(u, p, 0)$ is decreasing while $E(u, q, 1)$ is increasing. Hence, there can be only one intersection point.

We denote the solution by u^* and define $M^* = E(u^*, p, 0) = E(u^*, q, 1)$, where we do not reflect the dependence of u^* on p and q in our notation. From the monotonicity of $E(u, q, 1)$ on the left and of $E(u, p, 0)$ on the right of u^* , it then follows that $M^* = \max_u \min\{E(u, p, 0), E(u, q, 1)\}$. Note that this implies that $\Phi(u) < M$ for each $u \neq u^*$. A typical situation is depicted in Figure 2 for the case $p = 1/2$ and $q = 1$.

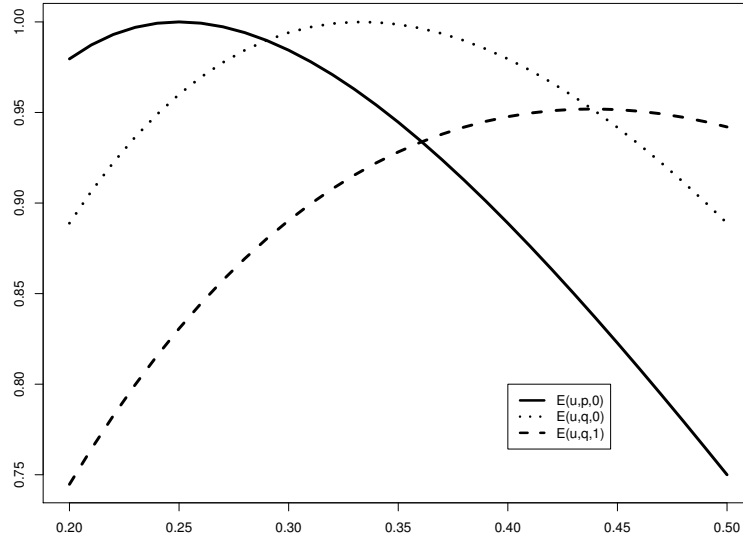


Figure 2: The functions $E(u, p, 0)$ (solid line), $E(u, q, 1)$ (dashed line) and $E(u, q, 0)$ (dotted line) in the case $p = 1/2$ and $q = 2$.

For Step (ii) it then only remains to show that $\Phi(u^*) \geq M$, that is

$$\min_{p \leq b \leq q} \min\{E(u^*, b, 0), E(u^*, b, 1)\} = \min\{E(u^*, p, 0), E(u^*, q, 1)\}.$$

It was already shown by Dette and Biedermann (2003) that for any fixed u ,

$$\min_{p \leq b \leq q} E(u^*, b, 0) = \min\{E(u^*, p, 0), E(u^*, q, 0)\}.$$

We now investigate the minimum of $E(u^*, b, 1)$ in the interval $[p, q]$. We note that the following arguments do not depend on the specific point u^* , but are correct for any $u \in [0, 1]$. A simple calculation shows that $E(u^*, b, 0) \leq E(u^*, b, 1)$ if and only if

$$b \leq \delta(u^*) := \frac{(2 - u^*)u^*}{2 - 3u^*}.$$

On the other hand, a straightforward but tedious calculation shows that $E(u, b, 1)$ is decreasing whenever $b > \delta(u^*)$. These arguments imply

$$\begin{aligned} \min_{p \leq b \leq q} \min\{E(u^*, b, 0), E(u^*, b, 1)\} &= \min\left\{\min_{p \leq b \leq \delta(u^*)} E(u^*, b, 0), \min_{\delta(u^*) \leq b \leq q} E(u^*, b, 1)\right\} \\ &= \min\{E(u^*, p, 0), E(u^*, q, 0), E(u^*, q, 1)\} \\ &= \min\{E(u^*, p, 0), E(u^*, q, 1)\}, \end{aligned}$$

which proves (ii) and completes the proof of Theorem (4.1). □

p	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1	1	1	$\frac{5}{3}$	$\frac{5}{3}$	$\frac{5}{3}$
q	$\frac{2}{3}$	$\frac{4}{3}$	2	$\frac{4}{3}$	2	$\frac{8}{3}$	2	$\frac{8}{3}$	$\frac{10}{3}$
u^*	.2889	.3389	.3628	.4523	.4738	.4867	.5025	.5307	.5378
$E(u^*, p, 0)$.9434	.8814	.8481	.9391	.9182	.9043	.9276	.9170	.9093

Table 5: *The smaller support point u^* of the standardized maximin optimal design with respect to the criterion (4.2). The value $E(u^*, p, 0)$ provides a lower bound of the D -efficiency of the design which uses the experimental conditions u^* and 1 for all $b \in [p, q]$ and $\lambda \in [0, 1]$.*

Example. 4.2 In Table 5 we show the solution of the equation (4.4) for selected values of p and q . The last row in the table gives the lower bound for the D -efficiency in the interval $[p, q]$. For example, if $p = 1/3, q = 2/3$, then we obtain from the first column in Table 5 $u^* = 0.2889$ and

$$E(u^*, 1/3, 0) = E(u^*, 2/3, 1) = 0.9434.$$

Therefore, a design using the experimental conditions 0.2889 and 1 has a D -efficiency of at least 94.3%, provided that the unknown parameter b is an element of the interval $[1/3, 2/3]$ independently of the value of the parameter $\lambda \in (0, 1)$. It might be of interest to investigate the efficiency of the design calculated under the assumption of uncorrelated errors, which uses the points $u^0(1/3, 2/3) = 0.2402$ and 1. The lower bound of the efficiency for this design provided by (4.2) is given by 90.0%.

5 Appendix: Technical details

5.1 Proof of (2.10)

The derivative of $h_\lambda(x, u)$ with respect to u is given by

$$\frac{\partial}{\partial u} h_\lambda(x, u) = \frac{1}{(x-u)^2} - \frac{2(\log \lambda)^2 \lambda^{2(x-u)}}{(1 - \lambda^{2(x-u)})^2}.$$

and the assertion follows if we are able to show that $\frac{\partial}{\partial u} h_\lambda(x, u) \geq 0$ for all $0 \leq u < x \leq 1$ and for all $0 < \lambda < 1$. In order to prove this inequality we consider for $\delta > 0$ the function

$$z_\delta(\lambda) = \frac{(-\log \lambda) \lambda^\delta}{1 - \lambda^{2\delta}}. \quad (5.1)$$

Then the derivative with respect to λ is

$$z'_\delta(\lambda) = \frac{\lambda^\delta}{\lambda(1 - \lambda^{2\delta})^2} \{ \lambda^{2\delta} - \delta(1 + \lambda^{2\delta}) \log \lambda - 1 \}.$$

The sign of $z'_\delta(\lambda)$ therefore is determined by the sign of $f_\delta(\lambda) = \lambda^{2\delta} - \delta(1 + \lambda^{2\delta}) \log \lambda - 1$. The derivative of $f_\delta(\lambda)$ equals

$$f'_\delta(\lambda) = \frac{\delta}{\lambda} \{ \lambda^{2\delta} (1 - 2\delta \log \lambda) - 1 \}$$

and is non-positive because $1 - 2\delta \log \lambda \leq \lambda^{-2\delta}$. Consequently, $f'_\delta(\lambda) \geq f'_\delta(1) = 0$ and, therefore, $z'_\delta(\lambda) \geq 0$ for all $\lambda \in (0, 1)$. Therefore we conclude that

$$z_\delta(\lambda) \leq \lim_{\lambda \rightarrow 1} z_\delta(\lambda) = \frac{1}{2\delta}.$$

Since $\frac{\partial}{\partial u} h_\lambda(x, u) = \frac{1}{(x-u)^2} - 2(z_{x-u}(\lambda))^2$, we conclude that

$$\frac{\partial}{\partial u} h_\lambda(x, u) \geq \frac{1}{(x-u)^2} - 2\left(\frac{1}{2(x-u)}\right)^2 > 0.$$

The desired inequality then easily follows. □

5.2 Completing the proof of Theorem 2.1

The function h defined in the proof of Theorem 2.1 has derivative

$$\frac{\partial}{\partial u} h(u) = \frac{1}{(1-u)^2} - \frac{2(\log(\lambda))^2 \lambda^{2(1-u)}}{(1 - \lambda^{2(1-u)})^2} = \frac{1}{(1-u)^2} - 2(z_{1-u}(\lambda))^2,$$

where $z_\delta(\lambda)$ is defined in (5.1). As in Appendix 5.1 it follows that $\frac{\partial}{\partial u} h(u) > 0$.

5.3 Properties of the function g_λ

The proof of (i) and (ii) is straightforward. To see (iii), consider the derivative of $g_\lambda(x)$ with respect to λ , which is

$$g'_\lambda(x) = \frac{(-\log \lambda)x}{(2(\lambda^{-2x} - 1) + x \log \lambda)^2} (4(\lambda^{-2x} - 1) + x(\log \lambda)(4\lambda^{-2x} + 1)).$$

For $x \in (0, 1]$, therefore $g'_\lambda(x) = 0$ if and only if

$$\ell_\lambda(x) := (-\log \lambda)x = r_\lambda(x) := \frac{4(1 - \lambda^{2x})}{4 + \lambda^{2x}}, \quad (5.2)$$

By calculating derivatives it follows that $r_\lambda(x)$ is concave and increasing with respect to x while $\ell_\lambda(x)$ is linear. Hence, there exist at most two solutions of (5.2), where one of them is $x = 0$. Moreover, $r'_\lambda(0) = \frac{40}{25} \log(\lambda)$ and is larger than the slope of ℓ_λ and therefore there is a second solution of (5.2) in the interval $(0, 1]$ if and only

$$\ell_\lambda(1) = -\log \lambda \geq r_\lambda(1) = \frac{4(1 - \lambda^2)}{4 + \lambda^2}. \quad (5.3)$$

It is now easy to see that equation (2.11) has exactly one solution, say λ_0 , such that (5.3) holds for all $\lambda \leq \lambda_0$ and the converse inequality holds for $\lambda > \lambda_0$. Hence, g_λ has exactly one local extremum in $[0, 1]$ if $\lambda \leq \lambda_0$ and there is no local extremum for $\lambda > \lambda_0$ which proves (iiia) and (iiib).

Acknowledgements

Support by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823, Teilprojekt C2) of the German Research Foundation (DFG) is gratefully acknowledged. The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise. This paper was written while both authors were at the Isaac Newton Institute, Cambridge, UK during the 2011 programme Design and Analysis of Experiments.

References

- Bickel, P. J. and Herzberg, A. M. (1979). Robustness of design against autocorrelation in time I: Asymptotic theory, optimality for location and linear regression. *Annals of Statistics*, 7(1):77–95.
- Bickel, P. J., Herzberg, A. M., and Schilling, M. F. (1981). Robustness of design against autocorrelation in time II: Optimality, theoretical and numerical results for the first-order autoregressive process. *Journal of the American Statistical Association*, 76(376):870–877.

- Butler, G. J. and Wolkowicz, G. S. K. (1985). A mathematical model of the chemostat with a general class of functions describing nutrient uptake. *SIAM Journal on Applied Mathematics*, 45:138–151.
- Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, 10(3):273–304.
- Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics*, 24:586–602.
- Clench, H. (1979). How to make regional lists of butterflies: some thoughts. *Journal of the Lepidopterists' Society*, 33:216–231.
- Dette, H. (1997). Designing experiments with respect to “standardized” optimality criteria. *Journal of the Royal Statistical Society, Ser. B*, 59:97–110.
- Dette, H. and Biedermann, S. (2003). Robust and efficient designs for the Michaelis-Menten model. *Journal of the American Statistical Association*, 98:679–686.
- Dette, H., Kiss, C., and Wong, W. K. (2010). A web-based tool for finding optimal designs for the Michaelis-Menten model and an overview. *Statistics in Biopharmaceutical Research*, 2(3):383–393.
- Dette, H., Kunert, J., and Pepelyshev, A. (2008). Exact optimal designs for weighted least squares analysis with correlated errors. *Statistica Sinica*, 18(1):135–154.
- Dette, H., Leonenko, N. N., Pepelyshev, A., and Zhigljavsky, A. (2009). Asymptotic optimal designs under long-range dependence error structure. *Bernoulli*, 15:1036–1056.
- Dette, H. and Wong, W. (1999). E -optimal designs for the Michaelis-Menten model. *Statistics and Probability Letters*, 44:405–408.
- Dunn, G. (1988). Optimal designs for drug, neurotransmitter and hormone receptor assays. *Statistics in Medicine*, 7:805–815.
- López, S., France, J., Gerrits, W. J., Dhanoa, M. S., Humphries, D. J., and Dijkstra, J. (2000). A generalized Michaelis-Menten equation for the analysis of growth. *Journal of Animal Science*, 78(7):1816–1828.
- López-Fidalgo, J., Martín-Martín, R., and Stehlik, M. (2008). Marginally restricted D -optimal designs for correlated observations. *Journal of Applied Statistics*, 35(6):617–632.
- Lopez-Fidalgo, J. and Wong, W. (2002). Optimal designs for the Michaelis-Menten model. *Journal of Theoretical Biology*, 215:1–11.

- Park, S. I., Felipe, C. R., Machado, P. G., Garcia, R., Skerjanec, A., Schmourer, R., Tedesco-Silva Jr, H., and Medina-Pestana, J. O. (2005). Pharmacokinetic/pharmacodynamic relationships of FTY720 in kidney transplant recipients. *Braz J Med Biol Res*, 38(5):683–694.
- Pepelyshev, A. (2007). Optimal designs for the exponential model with correlated observations. *MODA 8, Advances in Model-Oriented Design and Analysis*, pages 165–172.
- Pukelsheim, F. (2006). *Optimal Design of Experiments*. SIAM, Philadelphia.
- Rasch, D. (1990). Optimum experimental design in nonlinear regression. *Communications in Statistics*, 19:4789–4806.
- Rong, C. and Rappaport, S. M. (1996). Relation between pulmonary clearance and particle burden: A Michaelis-Menten-like kinetic model. *Occupational and Environmental Medicine*, 53:567–572.
- Sacks, J. and Ylvisaker, N. D. (1966). Designs for regression problems with correlated errors. *Annals of Mathematical Statistics*, 37:66–89.
- Sacks, J. and Ylvisaker, N. D. (1968). Designs for regression problems with correlated errors; many parameters. *Annals of Mathematical Statistics*, 39:49–69.
- Yu, X. Z. and Gu, J. D. (2007). Differences in Michaelis-Menten kinetics for different cultivars of maize during cyanide removal. *Ecotoxicology and Environmental Safety*, 67:254–259.
- Zhigljavsky, A., Dette, H., and Pepelyshev, A. (2010). A new approach to optimal design for linear models with correlated observations. *Journal of the American Statistical Association*, 105:1093–1103.