Robust and Efficient Designs for the Michaelis-Menten Model

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

Abstract

For the Michaelis-Menten model we determine designs which maximize the minimum of the D-efficiencies over a certain interval for the non-linear parameter. The best two point designs can be found explicitly and a characterization is given, when these designs are optimal within the class of all designs. In most cases of practical interest the determined designs are highly efficient and robust with respect to misspecification of the non-linear parameter. The results are illustrated and applied in an example of hormone-receptor assay.

---

*Professor Holger Dette is Professor of Mathematics (Stochastics) at the Department of Mathematics at Ruhr-Universität Bochum, 44780 Bochum, Germany, e-mail: holger.dette@ruhr-uni-bochum.de. Stefanie Biedermann works at the same Department as Assistant Professor, e-mail: stefanie.biedermann@ruhr-uni-bochum.de. The support of the Deutsche Forschungsgemeinschaft (DE 502/14-1 and SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen, Teilprojekt A2) is gratefully acknowledged. Parts of this paper were written during a visit of the first author at Purdue University and this author would like to thank the Department of Statistics for their hospitality. The authors would like to thank four referees and the associate editor for their very constructive comments on earlier versions of this paper, which improved the presentation of the results substantially. The authors are also grateful to A. Pepelyshev for computational assistance and to I. Gottschlich, who typed most parts of this paper with considerable technical expertise.

KEY WORDS: Michaelis-Menten model, robust optimal design, local D-optimality, receptor assay

1. INTRODUCTION

The rectangular hyperbola has been used to describe saturation functions for numerous physical and biological phenomena [see Cressie and Keightley (1979), Johansen (1984), Beverton and Holt (1957), Cornish-Bowden (1995), Hay, Meznarich, DiGiacomo, Hirst and Zerbe (1988)]. The form encountered most often in biology is the familiar Michaelis-Menten enzyme kinetic function

$$y = \frac{a\,x}{b+x}\,; \qquad x \in [0, x_0]\,, \tag{1}$$

where y is the reaction velocity, a the maximum velocity of this reaction, x the concentration of a substrate and b the half-saturation constant, i.e. the value of x, where y is half-maximal. Several methods have been proposed in the literature to estimate the parameters a, b [see e.g. Bliss and James (1966), Glick, Landman and Roufogalis (1979), Zivin and Waud (1982), Currie (1982) or Raaijmakers (1987)]. Most of these are based on transformations of the equation (1) to a linear plot and an application of ordinary linear regression techniques. Recently Raaijmaakers (1987) gave some arguments in favor of the application of maximum likelihood techniques for the estimation of the parameters in the Michaelis-Menten model.

The problem of designing experiments for the Michaelis-Menten model has also found considerable attention in the literature [see e.g. Duggleby (1979), Dunn (1988), Boer, Rasch and Hendrix (2000), Dette and Wong (1999) or Lopez-Fidalgo and Wong (2002)]. However, optimal designs are difficult to implement in practice, because in non-linear models the (locally) optimal designs depend on the unknown parameters [see Chernoff (1953) or Silvey (1980)] and usually are not robust with respect to their misspecification. For this reason, Song and Wong (1998) proposed Bayesian optimal designs for the Michaelis-Menten model and studied some of their properties. In this paper we study a different approach, which is based on the minimax concept and therefore does not require the specification of a prior distribution [see Müller (1995), Haines (1995), Dette (1997), Imhof and Wong (2000) or Imhof (2001)]. For the sake of transparency, we concentrate on maximum likelihood estimation and the D-optimality criterion, which means that an (locally) optimal design maximizes the determinant of the Fisher information matrix. Following Dette (1997) we propose as an optimality criterion the maximization of the minimum D-efficiency (with respect to the locally D-optimal design) taken over a certain range for the parameters. This criterion is motivated by the fact that in many cases the experimenter is able to specify a certain range for the non-linear parameters [see e.g. Cressie and Keightley (1981)], but has no preference for specific values, which would allow the specification of a prior distribution for a Bayesian approach. The optimality criterion is carefully defined in Section 2, which also discusses a useful reparametrization of the model. In Section 3 we determine the standardized maximin D-optimal two point designs and give an explicit characterization, when these designs are optimal within the class of all designs. It is demonstrated that in most cases the optimal designs are supported on two points and have equal masses at these points. Moreover, in the cases where the standardized maximin D-optimal designs are supported at more than two points it is indicated that the best two point designs are rather efficient. This section also contains a brief discussion of designing experiments for the Michaelis-Menten model in the case of heteroscedasticity. The finite sample performance of the standardized maximin D-optimal two point designs is assessed in Section 4 by comparing several designs in an extensive simulation study based on a real data example, which was originally analyzed by Cressie and Keightley (1981). Finally, some of the more technical arguments are deferred to an appendix.

2. STANDARDIZED MAXIMIN D-OPTIMAL DESIGNS

Recalling the definition of the model (1) we introduce a slightly different parametrization, that is

$$a = \alpha\, x_0, \quad b = \beta\, x_0, \tag{2}$$

and assume that Y follows a distribution from an exponential family with mean

$$E[\,Y \mid x] = \frac{\alpha\, x_0\, x}{\beta\, x_0 + x}, \quad x \in [0, x_0] \tag{3}$$

and constant variance. The Fisher information for the parameters (α, β) at a point x is then given by

$$I(x, \alpha, \beta) = x_0^2\, \frac{(x/x_0)^2}{(\beta + x/x_0)^2} \begin{pmatrix} 1 & -\frac{\alpha}{\beta + x/x_0} \\ -\frac{\alpha}{\beta + x/x_0} & \frac{\alpha^2}{(\beta + x/x_0)^2} \end{pmatrix}. \tag{4}$$

A design is a probability measure ξ on the interval $[0, x_0]$ with finite support, where the support points correspond to the location of the observations and the weights give the relative proportions of total observations to be taken at the corresponding points. The Fisher information matrix of a design ξ is defined as

$$M(\xi, \alpha, \beta) = \int_0^{x_0} I(x, \alpha, \beta)\, d\xi(x), \tag{5}$$

and a locally D-optimal design $\xi^*_{\alpha,\beta}$ maximizes the determinant of the matrix $M(\xi, \alpha, \beta)$. Observing the structure of the Fisher information $I(x, \alpha, \beta)$ defined in (4) and the fact that we are discussing the determinant criterion, it is easy to see that the factor α has no effect on the maximization of the determinant $|M(\xi, \alpha, \beta)|$, and we may assume without loss of generality that $\alpha = 1$. For the sake of a transparent notation we introduce

$$M(\xi, \beta) = M(\xi, 1, \beta).$$

Similarly, it is easy to see that we can restrict ourselves to the case $x_0 = 1$ and that designs on the general interval $[0, x_0]$ are simply obtained by rescaling the corresponding measures on $[0, 1]$ with the factor x_0. For example the locally D-optimal design on the interval $[0, 1]$ is given by the measure

$$\xi^*_\beta = \begin{pmatrix} \frac{\beta}{2\beta+1} & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{6}$$

(here the first row represents the support points and the second row the corresponding weights) and the locally D-optimal design on the interval $[0, x_0]$ has equal masses at the points $\beta x_0/(2\beta+1)$ and x_0 [see Rasch (1990)]. The value of the determinant of the locally D-optimal design can easily be calculated as

$$|M(\xi^*_\beta, \beta)| = \frac{1}{64\, \beta^2\, (1+\beta)^6}\,. \tag{7}$$

Note that the optimal design depends heavily on the unknown parameter β and some prior knowledge regarding this parameter is necessary to implement this design in practice. However, precise knowledge of this parameter is rarely available in practical problems and for this reason locally optimal designs are not widely used in applications. Nevertheless, in many situations a certain range for the parameter β can be specified by the experimenter [see for example Cressie and Keightley (1981), p.237, where a specific range for the dissociation constant β for the receptor-estradiol interaction is given] and a maximin approach might be appropriate, which maximizes the minimum of the determinants $|M(\xi, \beta)|$ taken over a certain interval, say $[\beta_0, \beta_1]$, for the parameter β with respect to the choice of the design ξ. It was pointed out by Dette (1997) in a more general context that the values of the determinants $|M(\xi,\beta)|$ at different points β might be of rather different size and for this reason not comparable. In fact it can be shown by an application of Hölder's inequality that $|M(\xi,\beta)|$ is decreasing with increasing value of β (independently of the design ξ) and consequently the design maximizing the minimum of the determinants is in fact locally D-optimal for the largest possible parameter β_1, for which the minimum is computed. Obviously this design is inefficient for small values in the interval $[\beta_0, \beta_1]$. A more reasonable criterion, which takes different sizes of the determinants into account, is the standardized maximin D-optimality criterion [see Dette (1997)]

$$\begin{aligned} \Phi(\xi) &= \min\Big\{ \frac{|M(\xi,\beta)|}{\sup_\eta |M(\eta,\beta)|} \;\Big|\; \beta \in [\beta_0,\beta_1] \Big\} \\ &= \min\Big\{ \frac{|M(\xi,\beta)|}{|M(\xi^*_\beta,\beta)|} \;\Big|\; \beta \in [\beta_0,\beta_1] \Big\}. \end{aligned} \tag{8}$$

This criterion seeks for a design maximizing the worst D-efficiency

$$\mathrm{eff}_\beta(\xi) = \Big(\frac{|M(\xi,\beta)|}{|M(\xi^*_\beta,\beta)|} \Big)^{1/2} \tag{9}$$

over the interval $[\beta_0, \beta_1]$. Following Dette (1997) a design ξ^* maximizing the criterion Φ in (8) is called standardized maximin D-optimal design. The following result characterizes the design ξ^* and can be proved by standard methods [see e.g. Wong (1992) or Pukelsheim (1993)].

Lemma 1 *A design ξ^* is standardized maximin D-optimal if and only if there exists a prior π supported on the set*

$$\mathcal{N}(\xi^*) = \{\beta \in [\beta_0,\beta_1] \mid \Phi(\xi^*) = [\mathit{eff}_\beta(\xi^*)]^2\} \tag{10}$$

such that the inequality

$$\int_{[\beta_0,\beta_1]} f^T(x,\beta) M^{-1}(\xi^*,\beta) f(x,\beta)\, \pi(d\beta) \le 2 \tag{11}$$

holds for all $x \in [0,1]$, where

$$f^T(x,\beta) = \frac{x}{\beta+x}(1, -\frac{1}{\beta+x}).$$

Moreover, there is equality in (11) for all support points of the design ξ^*.

Although Lemma 1 is attractive from a theoretical point of view, its practical impact is limited, because it does not answer the question how to choose the prior π in (11). For this reason, standardized maximin optimal designs have only been found in rare circumstances [see Müller (1995), or Imhof (2001)]. In all of these cases the authors concentrate on the optimization in the subclass of minimum supported designs. To be precise, we call a design with k support points standardized maximin D-optimal k-point design, if it maximizes the function Φ defined in (8) in the class of all designs with at most k support points. Our main result of this section gives the standardized maximin D-optimal two point design for the Michaelis-Menten model. The proof is deferred to the appendix.

Theorem 2 *The standardized maximin D-optimal two point design ξ_{z^*} for the Michaelis-Menten model on the interval $[0,1]$ has equal masses at the points 1 and*

$$z^* = z^*(\beta_0, \beta_1) = \frac{\beta_1\sqrt{\beta_0(1+\beta_0)} - \beta_0\sqrt{\beta_1(1+\beta_1)}}{\sqrt{\beta_1(1+\beta_1)} - \sqrt{\beta_0(1+\beta_0)}}. \qquad (12)$$

Moreover, for this design the set $\mathcal{N}(\xi_{z^*})$ defined in (10) is given by

$$\mathcal{N}(\xi_{z^*}) = \{\beta_0, \beta_1\}.$$

The question if and when the design ξ_{z^*} coincides with the standardized maximin D-optimal design within the class of all designs will be deferred to the following section. Roughly speaking this statement is correct, whenever

$$\frac{\beta_0}{\beta_1} \geq \frac{7}{36},$$

but the lower bound can be improved (decreased) with increasing value of β_1 [see Section 3]. We will conclude this section with an example illustrating the transformation introduced in (2) and the application of Theorem 2 for finding the best two point designs.

Example 3 For the analysis of data from hormone-receptor assays the relation between the response and explanatory variable is usually modelled by the Michaelis-Menten equation [see Cressie and Keightley (1981)]. In the specific example considered by these authors the range for the variable x is the interval

$[0, x_0] = [0, 2000]$. We consider three possible ranges for the interval for the parameter b in the model (1), namely

$$\begin{aligned} B_A &= [100, 2000] \\ B_B &= [100, 1000] \\ B_C &= [100, 500]. \end{aligned} \tag{13}$$

Using the transformation (2) this gives for the β-scale the intervals $[0.05, 1], [0.05, 0.5]$ and $[0.05, 0.25]$, respectively. The standardized maximin D-optimal two point designs are now obtained from Theorem 2 and the transformation $x \to 2000\,x$, which gives

$$\xi_A = \begin{pmatrix} 267.35 & 2000 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{14}$$

$$\xi_B = \begin{pmatrix} 223.78 & 2000 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{15}$$

$$\xi_C = \begin{pmatrix} 177.83 & 2000 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{16}$$

It follows from the proof of Theorem 2 in the Appendix that the worst D-efficiency of the standardized maximin D-optimal two point design is given by

$$\min_{\beta \in [\beta_0, \beta_1]} \text{eff}(\xi_{z^*}) = \frac{4\,\beta_0\,(1+\beta_0)\,z^*\,(1-z^*)}{(z^*+\beta_0)^2} = [u(z^*, \beta_0)]^{1/2}.$$

Figure 1 shows the D-efficiency (with respect to the locally D-optimal design ξ_β^*) of the designs ξ_B given in (15). The pictures for the designs ξ_A and ξ_C are similar and therefore omitted. We observe that for a broad range of the intervals B_A, B_B and B_C the standardized maximin D-optimal two point design has at least local D-efficiency 80%. For example, if we used the design ξ_B and the true parameter in the model (1) would be $b = 600$, then the design ξ_B has efficiency 90% compared to the locally D-optimal design, which assumes precise knowledge of the parameter b.

3. GLOBAL OPTIMALITY AND EFFICIENCY CONSIDERATIONS

The optimality of the determined best two point designs can be checked with the equivalence theorem given in Lemma 1. The following result determines the prior π, which is the main ingredient in Lemma 1. The proof is deferred to the appendix.

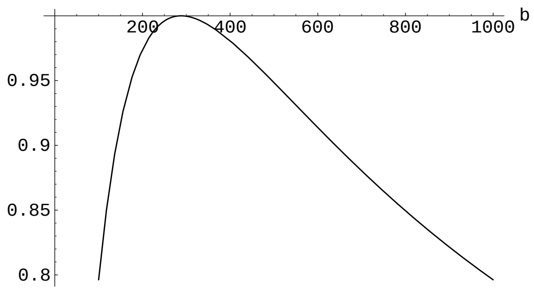

Figure 1: D-efficiencies of the standardized maximin D-optimal two point design ξ_B, where the minimum is taken over the range $[100, 1000]$

Lemma 4 *Let ξ_{z^*} denote the standardized maximin D-optimal two point design given in Theorem 2, then $\mathcal{N}(\xi_{z^*}) = \{\beta_0, \beta_1\}$ and the only prior on the set $\mathcal{N}(\xi_{z^*})$, for which the checking condition (11) is possibly satisfied, is the uniform distribution on $\mathcal{N}(\xi_{z^*})$.*

Note that it is not guaranteed that the design ξ_{z^*} of Theorem 2 and the uniform prior on the set $\{\beta_0, \beta_1\}$ satisfy the inequality (11), but the only prior for which this could be possible is the uniform prior. Consequently, we can use Lemma 1 with the uniform prior in order to check the optimality of the best two point design ξ_{z^*} given in Theorem 2. A straightforward calculation shows that the inequality (11) holds for all $x \in [0, 1]$, if and only if the function

$$L(\xi_{z^*}) = \frac{2}{\max\limits_{x \in [0,1]} \int_{\beta_0}^{\beta_1} f^T(x, \beta) M^{-1}(\xi_{z^*}, \beta) f(x, \beta)\, \pi(d\beta)} \qquad (17)$$

satisfies the equation

$$L(\xi_{z^*}) = 1. \qquad (18)$$

Moreover, it was shown by Dette (1996) that the function L is a lower boundary for the Φ-efficiency of a design with respect to the standardized maximin optimality criterion defined in (8), that is

$$\text{eff}_\Phi(\xi) = \frac{\Phi(\xi)}{\max\limits_{\eta} \Phi(\eta)} \geq L(\xi) \qquad (19)$$

for all designs ξ. For the best two point design ξ_{z^*} and the uniform prior we obtain by a straightforward calculation that the inequality (11) is equivalent to

$$g(x,\beta_0,\beta_1) \leq 2 \quad \forall\, x \in [0,1]\,, \tag{20}$$

where the function g is defined by

$$\begin{aligned} g(x,\beta_0,\beta_1) \;=\; & \sum_{i=0}^{1} \frac{(1+\beta_i)^4\,(z^*+\beta_i)^4\,x^2}{(z^*)^2\,(1-z^*)^2\,(x+\beta_i)^2}\,\Big\{\frac{(z^*)^2}{(z^*+\beta_i)^4} \\ & + \frac{1}{(1+\beta_i)^4} - \frac{2}{x+\beta_i}\Big(\frac{(z^*)^2}{(z^*+\beta_i)^3}+\frac{1}{(1+\beta_i)^3}\Big) \\ & + \frac{1}{(x+\beta_i)^2}\Big(\frac{(z^*)^2}{(z^*+\beta_i)^2}+\frac{1}{(1+\beta_i)^2}\Big)\Big\}, \end{aligned} \tag{21}$$

and the function L is given by

$$L(\xi_{z^*}) \;=\; \frac{2}{\max\limits_{x\in[0,1]} g(x,\beta_0,\beta_1)}\,. \tag{22}$$

Example 5 Recall the situation of Example 3. The condition (20) for the best two point designs in (14) - (16) is illustrated in Figure 2, where the designs have been rescaled to the interval $[0,1]$. From the inequality (20) it follows that the design ξ_C is the standardized maximin D-optimal design in the class of all designs, and that this conclusion is not correct for the designs ξ_A and ξ_B (see Figure 2). However, we obtain from (19) and (22) the lower bounds 80.24% and 93.8% for the Φ-efficiencies of the designs ξ_A and ξ_B, respectively. Note that $L(\xi_{z^*})$ provides only a lower bound for the Φ-efficiency and the true Φ-efficiency could be substantially larger. In fact it can be shown by numerical calculations that the standardized maximin D-optimal designs for the intervals B_A and B_B are three point designs and given by

$$\xi_A^* \;=\; \begin{pmatrix} 109.62 & 635.81 & 2000 \\ 0.2352 & 0.3207 & 0.4441 \end{pmatrix}$$

$$\xi_B^* \;=\; \begin{pmatrix} 116.89 & 457.44 & 2000 \\ 0.2483 & 0.2872 & 0.4645 \end{pmatrix}.$$

These designs have been determined by the Nelder-Mead method using Mathlab in the class of all three point designs. The global optimality was then verified using Lemma 1. This gives for the efficiencies of the two point designs

$$\mathrm{eff}_\Phi(\xi_A) = \frac{\Phi(\xi_A)}{\Phi(\xi_A^*)} \approx 0.9096, \quad \mathrm{eff}_\Phi(\xi_B) = \frac{\Phi(\xi_B)}{\Phi(\xi_B^*)} \approx 0.9787.$$

Thus one cannot expect a substantial improvement by determining the standardized maximin D-optimal design in the class of all designs. In a numerical study we considered numerous other situations for the interval $[\beta_0, \beta_1]$ and in all cases the optimal design for the Michaelis-Menten model has at most three support points.

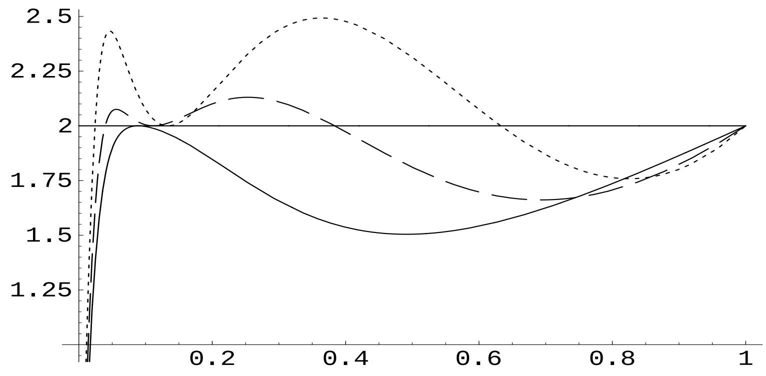

Figure 2: The function g defined in (21) for the designs ξ_A (dotted line), ξ_B (dashed line) and ξ_C (solid line). The design is standardized maximin D-optimal if and only if the curve stays below the line $x = 2$. Twice the inverse of the maximal excess gives a lower bound for the Φ-efficiency [see inequality (19)]

For the following discussion we use a slightly different parametrization of the interval $[\beta_0, \beta_1]$. Let $c \in (0, 1]$ be defined by

$$\beta_0 = c\beta_1$$

and consider the function

$$h(c, \beta_1) = \max_{x \in [0,1]} g(x, c\beta_1, \beta_1),$$

where the function g is given in (21). By Lemma 1 and the discussion preceding Example 5 it follows that the best two point design is in fact standardized maximin D-optimal within the class of all measures if and only if $h(c, \beta_1) = 2$. For example, for the choice $c = 1$, the interval $[\beta_0, \beta_1]$ collapses to the point β_1 and the standardized maximin D-optimality criterion reduces to the local D-optimality criterion. We trivially obtain $h(1, \beta_1) = 2$, because the locally D-optimal design is in fact supported on two points. We are now interested in the bound

$$k(\beta_1) = \inf\{\, c \in (0, 1] \mid h(u, \beta_1) = 2 \ \forall\ u \geq c\}.$$

In other words $k(\beta_1)$ is the smallest number in the interval $(0,1]$, such that for any value $c > k(\beta_1)$ the best two point design of Theorem 2 is actually the standardized maximin D-optimal design in the class of all models (where the minimum is taken over the interval $[c\beta_1, \beta_1]$). Although an analytical derivation of this function seems to be intractable, we are able to derive this function numerically by a careful investigation of the function g defined in (21). If the ratio $c = \beta_0/\beta_1$ is close to one, the function g has only one local maximum (see the solid line in Figure 2). If c approaches a critical value, say $\tilde{k}(\beta_1)$ from above the function g has two local maxima but is still smaller than 2 [see the solid and dotted line in Figure 3]. If c is slightly smaller than $\tilde{k}(\beta_1)$ one of the maxima exceeds the critical bound 2 [see the dashed line in Figure 3], and if c is further decreased the other maximum turns to a local minimum [see the dotted and dashed line in Figure 2]. Based on this information we determined the critical bound $\tilde{k}(\beta_1)$ numerically and our calculations show that it is in fact inverse linear. More precisely, our numerical results indicate that $k(\beta_1)$ is determined by the relation

$$\tilde{k}(\beta_1) = \frac{1}{4.14307231\,\beta_1 + 5.14307231}\,. \tag{23}$$

Moreover, our numerical study shows that for all values smaller than $\tilde{k}(\beta_1)$ the best two point design is not optimal within the class of all designs.

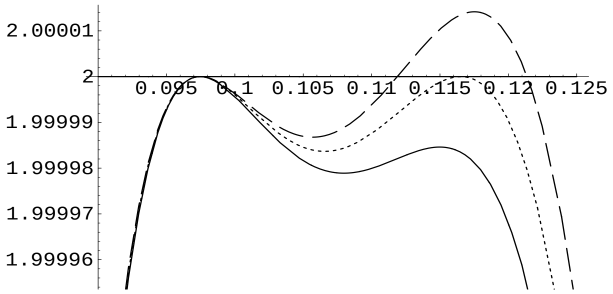

Figure 3: The form of the function g defined in (21) if c is close to $k(\beta_1)$. Solid line: c is larger than $k(\beta_1)$, dotted line: $c = k(\beta_1)$, dashed line: c is smaller than $k(\beta_1)$.

Numerical fact 6 *The standardized maximin D-optimal two point design given in Theorem 2 is standardized maximin D-optimal within the class of all designs, if and only if*

$$c = \frac{\beta_0}{\beta_1} \geq \tilde{k}(\beta_1),$$

where the function $\tilde{k}$ is defined in (23).

We finally note that in the situation of Example 5 and Example 3 the constant c is given by 0.05, 0.1 and 0.2, corresponding to the intervals B_A, B_B and B_C in (13) and the values for the critical point $\tilde{k}(\beta_1)$ are $\tilde{k}(1) = 0.10768$, $\tilde{k}(0.5) = 0.13860$ and $\tilde{k}(0.25) = 0.16184$, respectively, which confirms the results given in Example 5.

Remark 7 Many biological phenomena, where the Michaelis-Menten model is used, exhibit heteroscedasticity [see e.g. Currie (1982) or Dunn (1988).] Usually it is assumed that the conditional variance of Y in (3) is a function of the mean, i.e. $\mathrm{Var}\,(Y \mid x) = v(g(x,a,b))$, where $g(x,a,b) = a\,x/(b+x)$ denotes the conditional expectation of Y at x. Using results of Atkinson and Cook (1995) it can be shown that the Fisher information at a point x in the Michaelis-Menten model (1) with heteroscedastic structure is given by

$$I_v(x,a,b) = \frac{g^2(x,a,b)}{a^2} \begin{pmatrix} 1 & -\frac{a}{b+x} \\ -\frac{a}{b+x} & \frac{a^2}{(b+x)^2} \end{pmatrix} \left\{ \frac{1}{v(g(x,a,b))} + \frac{1}{2}\Big(\frac{v'(g(x,a,b))}{v(g(x,a,b))}\Big)^2 \right\} \qquad (24)$$

and an analogous analysis as described in the homoscedastic case ($v \equiv 1$) can be performed, where the structure of the standardized maximin D-optimal design depends on the assumed link function v.

4. EXAMPLE

In this section, we reanalyze a design used by Cressie and Keightley (1981) for an application of the Michaelis-Menten model to estrogen receptor assays, which are used in the study of human breast cancer. Here the response Y represents the amount of estrogen bound to receptor and x is the amount of hormone not bound to receptor. For more details we refer to Cressie and Keightley (1981). In our notation the design space is the interval $[0, 2000]$ and the design used by these authors is given in Table 1 of their paper. This means that $n = 10$ observations are available, where the incubating concentrations of the hormone (estradiol - 17 β) are nearly carried out in duplicate. The maximum likelihood estimates for a and b from these data are given by $\hat{a} = 43.95$ and $\hat{b} = 236.53$, respectively, where the variance is estimated as $\hat{\sigma}^2 = 1.39$.

In order to study the finite sample properties of the standardized maximin D-optimal designs we used a simulation study and generated data according to the model

$$Y_i = \frac{\hat{a}\, x_i}{\hat{b} + x_i} + \hat{\sigma}\, \varepsilon_i \quad i = 1, \ldots, n, \tag{25}$$

where the explanatory variables x_i vary in the interval $[0, 2000]$ and the ε_i are independent standard normal random variables. In this study we compared 8 designs:

(1) The standardized maximin D-optimal two point designs ξ_A, ξ_B and ξ_C defined by (14), (15) and (16) and corresponding to the parameter spaces $[b_0, b_1] = [100, 2000]$, $[100, 1000]$ and $[100, 500]$.

(2) The locally D-optimal designs η_A, η_B and η_C for the parameter $b^* = (b_0 + b_1)/2$ for the three cases considered in (1). This design was proposed by a referee.

(3) The uniform design on the interval $[0, 2000]$ with two replications at each point. This design takes two observations at the points $4000\, i/n$ $(i = 1, \ldots, n/2)$.

(4) The design given in Table 1 of Cressie and Keightley (1981), where we put $n/10$ observations at the points 83.9, 84.6, 147.8, 148.2, 463.8, 463.9, 964.1, 967.6, 1900, 1925 (the case $n = 10$ corresponds precisely to the situation considered by these authors).

Table 1 shows the simulated mean squared error of the maximum likelihood estimates based on data generated from the model (25), where the sample sizes are $n =$ 10, 20, 30, 40, 50 and 10000 simulation runs were carried out for each scenario. We observe substantial differences between the different designs. The standardized maximin D-optimal two point designs ξ_A, ξ_B and ξ_C yield rather similar results, where the performance is improved by using smaller regions in the maximin criterion. This corresponds to intuition and means that the design ξ_C (corresponding to the interval $[b_0, b_1] = [100, 500]$) is the best. All other designs yield a substantially larger MSE than the standardized maximin D-optimal two point designs. While this observation is not surprising for the uniform design and the design used by Cressie and Keightley (1981), the results for the locally D-optimal design (with respect to the midpoint of the assumed parameter space) are notable. In the case of an unprecise specification of the parameter space these designs show a substantially

Table 1: Simulated mean squared error of the maximum likelihood estimates in the Michaelis-Menten model with $a = 43.95$, $b = 236.53$, $\sigma^2 = 1.39$. The table shows the MSE for the standardized maximin D-optimal two point designs ξ_A, ξ_B and ξ_C defined in (14), (15) and (16), respectively. The designs η_A, η_B and η_C are the locally D-optimal designs for the parameter $b = (b_0 + b_1)/2 = 1050$, 550, 300, respectively, while the last two rows show the results for the design used by Cressie and Keightley (1981) and the uniform design on the interval $[0, 2000]$ with two replications at each design point.

	$n = 10$		$n = 20$		$n = 30$		$n = 40$		$n = 50$	
	$\hat{b}$	$\hat{a}$	$\hat{b}$	$\hat{a}$	$\hat{b}$	$\hat{a}$	$\hat{b}$	$\hat{a}$	$\hat{b}$	$\hat{a}$
ξ_A	295.45	0.661	145.81	0.323	100.78	0.221	73.76	0.168	59.84	0.133
η_A	634.98	0.941	308.03	0.467	204.39	0.308	155.63	0.234	125.40	0.188
ξ_B	267.48	0.639	132.99	0.320	87.79	0.210	65.90	0.158	53.03	0.124
η_B	371.62	0.720	188.43	0.366	128.17	0.251	95.69	0.186	75.12	0.151
ξ_C	252.53	0.607	121.45	0.299	84.07	0.199	61.81	0.147	48.41	0.117
η_C	267.50	0.625	137.10	0.315	89.40	0.210	67.08	0.158	54.25	0.124
C & K	360.76	1.029	182.83	0.540	118.81	0.343	88.20	0.256	73.16	0.215
uniform	711.22	1.069	322.99	0.529	216.05	0.363	159.72	0.268	129.62	0.219

worse performance than the standardized maximin D-optimal two point designs, but yield a smaller MSE than the uniform design and the design used by Cressie and Keightley (1981). For example, the design η_C, which requires a rather precise specification for the parameter b, i.e. $b \in [100, 500]$, is comparable with the standardized maximin D-optimal two point design ξ_A, which only requires the information $b \in [100, 2000]$. Summarizing this study we conclude that the inference performed by Cressie and Keightley (1981) could have been improved by using one of the standardized maximin D-optimal designs proposed in this paper and that the approach (2) does not yield robust and efficient designs for the Michaelis-Menten model.

It might also be of interest how the designs perform if the value b used in the simulation is not in the interval $[100, 500]$, which means that the interval has been misspecified by the experimenter. We investigated the case $b = 750$ and the uniform design, the standardized maximin D-optimal design ξ_C for the interval $[100, 500]$ and the design of Cressie and Keightley (1981)

Table 2: Simulated mean squared error of the maximum likelihood estimator in the Michaelis-Menten model with $a = 43.95$, $b = 750$, $\sigma^2 = 1.39$. The table shows the MSE for the standardized maximin D-optimal design ξ_C defined in (16), the design used by Cressie and Keightley (1981) and the uniform design on the interval $[0, 2000]$ with two replications at each point.

	$n = 10$		$n = 20$		$n = 30$		$n = 40$		$n = 50$	
	$\hat{b}$	$\hat{a}$	$\hat{b}$	$\hat{a}$	$\hat{b}$	$\hat{a}$	$\hat{b}$	$\hat{a}$	$\hat{b}$	$\hat{a}$
ξ_C	8104.5	3.554	3837.1	1.596	2775.1	1.049	2134.4	0.812	1640.5	0.626
C & K	8137.8	4.441	4803.0	3.037	3075.1	1.893	2291.1	1.428	1821.7	1.137
uniform	8907.3	4.196	5688.7	3.103	3829.1	2.028	2837.3	1.527	2273.8	1.235

in this situation. The results are depicted in Table 2. We observe again substantial improvements by the standardized maximin D-optimal design. We note that we could even improve the parameter estimates by using standardized maximin D-optimal designs for a larger parameter space (e.g. the design ξ_A).
Another important issue in this example is robustness with respect to the model assumptions. The optimal two point designs should only be used if there is strong evidence that the enzyme exhibits Michaelis-Menten kinetics and that this aspect has not to be checked. In other circumstances the optimal designs are often used as benchmarks to assess the efficiencies of designs, which might have more widespread acceptance among practitioners. Our results show that the uniform design and the design of Cressie and Keightley (1981) are rather inefficient. The standardized maximin D-optimal designs can be used to modify these commonly applied designs such that they become more efficient for estimation. For example a simple modification of the uniform design would be to take 10% of the observations at the points 550, 900, 1250, 1600 and 30% of the observations at 177 and 1950 (motivated by the design ξ_C). Compared to the uniform design this modification reduces the mean squared error of the estimates substantially (for example if $n = 50$ we have $\mathrm{MSE}(\hat{b}) \approx 77.01$, $\mathrm{MSE}(\hat{a}) \approx 0.149$) and the modified design can also be used for model checking.

APPENDIX: PROOFS

Proof of Theorem 2: It is easy to check that the optimal weights of a two point design must be equal [see e.g. Silvey (1980)]. If the support points are denoted by $z_1, z_2 \in [0, 1]$ we obtain for the Fisher information matrix of a design ξ_{z_1,z_2} with equal weights at the points z_1 and z_2

$$M(\xi,\beta)=\frac{1}{2}\sum_{i=1}^{2}\begin{pmatrix}\frac{z_i^2}{(z_i+\beta)^2} & -\frac{z_i^2}{(z_i+\beta)^3}\\ -\frac{z_i^2}{(z_i+\beta)^3} & \frac{z_i^2}{(z_i+\beta)^4}\end{pmatrix}, \tag{A.1}$$

and a straightforward calculation gives for the determinant

$$|M(\xi,\beta)|=\frac{1}{4}\frac{z_1^2\,z_2^2\,(z_2-z_1)^2}{(z_1+\beta)^4\,(z_1+\beta)^4}.$$

Therefore the standardized maximin D-optimality criterion reduces to

$$\begin{aligned}\Phi(\xi_{z_1,z_2}) \;=\; \min\Big\{16\,\beta^2\,(1+\beta)^6\,\frac{z_1^2\,z_2^2\,(z_2-z_1)^2}{(z_1+\beta)^4\,(z_2+\beta)^4}\;\Big|\\ \beta\in[\beta_0,\beta_1]\Big\},\end{aligned} \tag{A.2}$$

where we have used the representation for the determinant of the locally D-optimal design in (7). Now it is easy to see that $\Phi(\xi_{z_1,z_2})\le\Phi(\xi_{z_1,1})$ for all $z_2\in(z_1,1]$ and consequently 1 is a support point of the best two point design. For $\xi_z=\xi_{z,1}$ the criterion in (A.2) reduces to

$$\Phi(\xi_z)=\min\{u(z,\beta)\mid\beta\in[\beta_0,\beta_1]\}, \tag{A.3}$$

where the function u is defined by

$$u(z,\beta)=16\,\beta^2\,(1+\beta)^2\,\frac{z^2\,(1-z)^2}{(z+\beta)^4}\,. \tag{A.4}$$

Recalling the definition of the point z^* in (12) we will prove below the following two assertions.

- For any fixed $z\in(0,1)$ the function $\beta\to\log u(z,\beta)$ is unimodal.
- If ξ_{z^*} maximizes the function Φ in the class of all two point designs then

$$u(z^*,\beta_0)=u(z^*,\beta_1). \tag{A.5}$$

If these assertions have been established we obtain the representation

$$\Phi(\xi_z)=\min\{u(z,\beta_0),u(z,\beta_1)\} \tag{A.6}$$

for the criterion (A.3) from the unimodality of the function $\beta \to \log u(z, \beta)$, and (A.5) gives a condition on the optimal design ξ_{z^*}, which determines z^* uniquely. Solving equation (A.5) with respect to z^* it follows by a straightforward calculation

$$z^* = z^*(\beta_0, \beta_1) = \frac{\beta_1\sqrt{\beta_0(1+\beta_0)} - \beta_0\sqrt{\beta_1(1+\beta_1)}}{\sqrt{\beta_1(1+\beta_1)} - \sqrt{\beta_0(1+\beta_0)}}, \quad \text{(A.7)}$$

which proves the assertion of Theorem 2. We complete the proof by showing the unimodality of the function $\beta \to \log u(z, \beta)$ and the identity (A.5).

Proof of unimodality: This follows by a straightforward investigation of

$$\frac{\partial}{\partial \beta} \log u(z, \beta) = \frac{1}{\beta} + \frac{1}{1+\beta} - \frac{2}{z+\beta}$$

which has at most one zero for fixed z.

Proof of (A.5): Since the function $\beta \to \log u(z, \beta)$ has at most one maximum it follows that the standardized maximin D-optimal two point design ξ_{z^*} can be found by maximizing the function $\Phi(\xi_z)$ defined in (A.6) over the interval $(0, 1)$. This maximization can be divided into three separate steps, namely the maximization over the sets

$$\begin{aligned} \mathcal{M}_< &:= \{z \in (0,1) \mid u(z, \beta_0) < u(z, \beta_1)\} \\ \mathcal{M}_> &:= \{z \in (0,1) \mid u(z, \beta_0) > u(z, \beta_1)\} \\ \mathcal{M}_= &:= \{z \in (0,1) \mid u(z, \beta_0) = u(z, \beta_1)\}. \end{aligned} \quad \text{(A.8)}$$

We will now demonstrate that the maximization over the first two sets yields a maximum at the boundary, which proves the assertion of (A.5). To this end consider the first set $\mathcal{M}_<$, the second case is treated exactly in the same way. A straightforward calculation shows

$$\mathcal{M}_< = (0, z^*),$$

where the point z^* is defined in (A.7). Because for $z \in \mathcal{M}_<$ we have $\Phi(\xi_z) = u(z, \beta_0)$ a straightforward maximization of $\Phi(\xi_z)$ over the set $\mathcal{M}_<$ yields

$$\tilde{z} = \frac{\beta_0}{1 + 2\beta_0} \quad \text{and} \quad u(\tilde{z}, \beta_0) = 1.$$

Now assume that $\tilde{z}$ is in the set $\mathcal{M}_<$ then $u(\tilde{z}, \beta_0) < u(\tilde{z}, \beta_1)$ must hold and a tedious calculation shows that this inequality is equivalent to $(\beta_0 - \beta_1)^2 < 0$. This contradiction shows that the maximum is not attained in the interior of the set $\mathcal{M}_<$ and a similar argument for the set $\mathcal{M}_>$ establishes the assertion (A.5) completing the proof of Theorem 2. $\Box$

Proof of Lemma 4: It follows from Theorem 2 that $\mathcal{N}(\xi_{z^*}) = \{\beta_0, \beta_1\}$. If a prior π puts weights α_0 and $\alpha_1 = (1 - \alpha_0)$ at the points β_0 and β_1, respectively, then the inequality (11) can be rewritten as (note that $\alpha_0 = 1 - \alpha_1$)

$$d(x, \alpha_0) = \sum_{i=0}^{1} \alpha_i f^T(x, \beta_i) M^{-1}(\xi^*, \beta_i) f(x, \beta_i) \leq 2.$$

If ξ_{z^*} is in fact the (global) standardized maximin D-optimal design we must have equality for the support points of ξ_{z^*}. The equalities $d(1, \alpha_0) = d(z^*, \alpha_0) = 2$ hold for all $\alpha_0 \in (0, 1)$ and therefore give no information regarding the weight. However at the point $x = z^*$ the function d must have a local maximum, which yields the condition

$$\frac{\partial}{\partial x} d(x, \alpha_0) \mid_{x=z^*} = 0. \qquad \text{(A.9)}$$

Now a straightforward but very tedious calculation shows that this equation implies $\alpha_0 = 1/2$, which establishes the assertion of the Lemma. $\Box$

References

[1] Atkinson, A. C., and Cook, R. D. (1995), "D-optimum Designs for Heteroscedastic Linear Models," *Journal of the American Statistical Association*, 90, No. 429, 204-212.

[2] Beverton, R. J. H., and Holt, S. J. (1957), "On the Dynamics of Exploited Fish Populations," *London: Her Majesty's Stationary Office.*

[3] Bliss, C. I., and James, A. T. (1966), "Fitting the Rectangular Hyperbola," *Biometrics*, 22, 573-602.

[4] Boer, E. P. J., Rasch, D., and Hendrix, E. M. T. (2000), "Locally Optimal Designs in Non-Linear Regression: A Case Study for the Michaelis-Menten Function," Balakrishnan, N. (ed.) et al., *Advances in Stochastic Simulation Methods, Selected Papers at the 3rd St. Petersburg Workshop on Simulation, St. Petersburg, Russia, June 28 - July 3, 1998*, Boston: Birkhäuser *Statistics for Industry and Technology*, 177-188.

[5] Chernoff, H. (1953), "Locally Optimal Designs for Estimating Parameters," *Annals of Mathematical Statistics*, 24, 586-602.

[6] Cornish-Bowden, A. (1995), *Fundamentals of Enzyme Kinetics*, (Rev. Ed.), Portland Press, London.

[7] Cressie, N. A. C., and Keightley, D. D. (1979), "The Underlying Structure of a Direct Linear Plot with Applications to the Analysis of Hormone-Receptor Interactions," *Journal of Steroid Biochemistry*, 11, 1173-1180.

[8] Cressie, N. A. C., and Keightley, D. D. (1981), "Analysing Data from Hormone-Receptor Assays," *Biometrics*, 37, 235-249.

[9] Currie, D. J. (1982), "Estimating Michaelis-Menten Parameters: Bias, Variance and Experimental Design," *Biometrics*, 38, 907-919.

[10] Dette, H. (1996), "Lower Bounds for Efficiencies with Applications," Research Developments in Probability and Statistics: Festschrift in Honor of M. L. Puri's 65th Birthday (eds. E. Brunner, M. Denker), 111-124, VSP Utrecht.

[11] Dette, H. (1997), "Designing Experiments with Respect to Standardized Optimality Criteria," *Journal of the Royal Statistical Society, Ser. B*, 59, No.1, 97-110.

[12] Dette, H., and Wong, W. K. (1999), "E-Optimal Designs for the Michaelis-Menten Model," *Statistics & Probability Letters*, 44, No.4, 405-408.

[13] Duggleby, R. G. (1979), "Experimental Designs for Estimating the Kinetic Parameters for Enzyme-Catalysed Reactions," *Journal of Theoretical Biology*, 81, 671-684.

[14] Dunn, G. (1988), "Optimal Designs for Drug, Neurotransmitter and Hormone-Receptor Assays," *Statistics in Medicine*, 7, 805-815.

[15] Glick, N., Landman, A. D., and Roufogalis, B.D. (1979), "Correcting Lineweaver-Burk Calculations of V and K_m," *Trends in Biochemical Sciences*, 4, N82-N83.

[16] Haines, L. M. (1995), "A Geometric Approach to Optimal Design for One-Parameter Nonlinear Models," *Journal of the Royal Statistical Society, Ser. B*, 57, 575-598.

[17] Hay, W. W., Meznarich, H. K., DiGiacomo, J. E., Hirst, K., and Zerbe G. (1988), "Effects of Insulin and Glucose Concentration on Glucose Utilization in Fetal Sheep," *Pediatric Research*, 23, 381-387.

[18] Imhof, L. A. (2001), "Maximin Designs for Exponential Growth Models and Heteroscedastic Polynomial Models," *Annals of Statistics*, 29, No.2, 561-576.

[19] Imhof, L. A., and Wong, W. K. (2000), "A Graphical Method for Finding Maximin Designs," *Biometrics*, 56, 113-117.

[20] Johansen, S. (1984), "Functional Relations, Random Coefficients and Nonlinear Regression, with Application to Kinetic Data," *Lecture Notes in Statistics*, 22, Springer, N.Y.

[21] Lopez-Fidalgo, J., and Wong, W. K. (2002), "Optimal Designs for the Michaelis-Menten Model," *Journal of Theoretical Biology*, 215, No.1, 1-11.

[22] Müller, C. H. (1995), "Maximin Efficient Designs for Estimating Nonlinear Aspects in Linear Models," *Journal of Statistical Planning and Inference*, 44, No.1, 117-132.

[23] Pukelsheim, F. (1993), *Optimal Design of Experiments*, Wiley, N.Y.

[24] Raaijmaakers, J. G. W. (1987), "Statistical Analysis of the Michaelis-Menten Equation," *Biometrics*, 43, No.4, 793-803.

[25] Rasch, D. (1990), "Optimum Experimental Design in Nonlinear Regression," *Communications in Statistics, Theory and Methods*, 19, No.12, 4789-4806.

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

[27] Song, D., and Wong, W. K. (1998), "Optimal Two Point Designs for the Michaelis-Menten Model with Heteroscedastic Errors," *Communications in Statistics, Theory and Methods*, 27, No.6, 1503-1516.

[28] Wong, W. K. (1992), "A Unified Approach to the Construction of Minimax Designs," *Biometrika*, 79, No.3, 611-619.

[29] Zivin, J. A., and Waud, D. R. (1982), "How to Analyze Binding, Enzyme and Uptake Data: the Simplest Case, a Single Phase," *Life Sciences*, 30, 1407-1422.