

# Optimal designs for enzyme inhibition kinetic models

Kirsten Schorning, Holger Dette, Katrin Kettelhake, Tilman Möller  
Fakultät für Mathematik  
Ruhr-Universität Bochum  
44780 Bochum  
Germany

## Abstract

In this paper we present a new method for determining optimal designs for enzyme inhibition kinetic models, which are used to model the influence of the concentration of a substrate and an inhibition on the velocity of a reaction. The approach uses a nonlinear transformation of the vector of predictors such that the model in the new coordinates is given by an incomplete response surface model. Although there exist no explicit solutions of the optimal design problem for incomplete response surface models so far, the corresponding design problem in the new coordinates is substantially more transparent, such that explicit or numerical solutions can be determined more easily. The designs for the original problem can finally be found by an inverse transformation of the optimal designs determined for the response surface model. We illustrate the method determining explicit solutions for the  $D$ -optimal design and for the optimal design problem for estimating the individual coefficients in a non-competitive enzyme inhibition kinetic model.

Keywords and Phrases: modelling enzyme kinetics reactions, incomplete response surface models,  $D$ -optimal designs, optimal designs for individual coefficients

AMS Subject classification: Primary 62K05 ; Secondary: 62K20

## 1 Introduction

The Michaelis-Menten model

$$\eta = \eta(S) = \frac{V \cdot S}{K_m + S} \quad (1.1)$$

was introduced by Michaelis and Menten (1913) and is widely used to represent an enzyme kinetics reaction, where enzymes bind substrates and turn them into products. Here  $V$  and  $S$  denotes the maximum velocity of the reaction and the concentration of the substrate, respectively, while  $K_m$  is the value of  $S$  at which half of the maximum velocity  $V$  is reached, i.e., the

Michaelis-Menten constant. Because of its importance the problem of designing experiments for statistical analysis based on the Michaelis-Menten model has found considerable interest in the literature and we refer to Rasch (1990); Song and Wong (1998); López-Fidalgo and Wong (2002); Dette and Biedermann (2003); Matthews and Allcock (2004) and Dette and Kunert (2014) among many others, who investigate optimal and efficient designs for the model (1.1) from different perspectives.

The Michaelis-Menten model is very well justified in the absence of enzyme inhibition which are molecules that decrease the activity of enzymes. However, many diseases require co-administration of several drugs and for this reason new drugs are usually also screened for their inhibitory potential. Adequate modeling has to reflect this fact and therefore the Michaelis-Menten model has been extended to include the effect of inhibitor concentration, say  $I$ . In these extended models  $\eta$  is a function of the concentration of the substrate  $S$  and the inhibition  $I$ , that is  $\eta = \eta(S, I)$  and one usually distinguishes between competitive and non-competitive inhibition [see Segel (1993), for example]. There does not exist much literature on optimal designs for these type of models and we refer to Youdim et al. (2010); Bogacka et al. (2011); Atkinson and Bogacka (2013) who mainly studied designs optimal with respect to determinant criteria, i.e.,  $D$ - or  $D_s$ -optimal designs. In most cases the optimal design problem is an extremely difficult optimisation problem and optimal designs have to be found numerically as the commonly used models for competitive and non-competitive inhibition are highly nonlinear and therefore difficult to analyse mathematically. As a consequence not much - except of empirical findings - is known about the structure of optimal designs.

In this paper we present a new approach to derive optimal designs for regression models for enzyme kinetics reactions involving inhibition which simplifies the analytical and numerical construction in this context substantially. The main idea is to use a nonlinear transformation of the explanatory variables  $\psi : (S, I) \rightarrow (x, y)$  such that the model in the new variables  $x$  and  $y$  is a simple (incomplete) multivariate polynomial regression model (also called response surface model). The design problem can then be analysed in a linear response surface model and in a second step the optimal designs from this model can be transformed back into the  $(S, I)$  coordinates to obtain optimal designs for enzyme inhibition kinetic models. We illustrate this approach determining optimal designs for a specific non-competitive enzyme inhibition kinetic model, but the general idea is also applicable in other enzyme inhibition kinetic models as well. In particular, we provide a complete proof that a design found in Bogacka et al. (2011) is in fact  $D$ -optimal. Moreover, we derive new optimal designs for estimating the individual coefficients in this model. This optimisation problem is a substantially harder than the  $D$ -optimal design problem and might be intractable in the classical parametrisation. However, with the transformation introduced here complete solution of the optimal design problem is possible.

Optimal designs for multivariate quadratic regression models have been widely studied in the literature and we refer to the seminal papers of Kono (1962); Farrel et al. (1967) and to Draper et al. (2000); Draper and Pukelsheim (2003) and Dette and Grigoriev (2014) for some more

recent work among many others. In these and many other references optimal designs for the “complete” model are considered, which means the model under consideration includes all multivariate polynomials up to a given degree. In the simplest case of a cubic polynomial in two variables, say  $x$  and  $y$ , this means that the corresponding linear model contains the functions

$$1, x, y, xy, x^2, y^2, xy^2, x^2y^2, x^3, y^3 . \quad (1.2)$$

However, the mapping  $\psi$  introduced in this paper to transform design problems for enzyme inhibition kinetic models to designs problems for linear response surface models usually results in “incomplete” multivariate polynomial regression models, for example a cubic model involving only the functions

$$xy, xy^2, yx^2 . \quad (1.3)$$

For such models the determination of optimal designs is substantially harder and there does not exist any solution so far. To our best knowledge - we are only aware of two references, where product designs for “incomplete” multivariate polynomials are determined [see Dette and Roeder (1996, 1997)]. Although the determination of optimal designs (not restricted to a product structure) is extremely hard we are able to find solutions for the case (1.3), which is the important one for the non-competitive enzyme inhibition kinetic model under consideration.

The remaining part of this paper is organised as follows. In Section 2 we introduce the regression model and some basic optimal design terminology. Section 3 is devoted to the nonlinear transformation and its basic properties. Additionally we illustrate our approach determining  $D$ -optimal designs. In Section 4 we use the new methodology to explicitly determine optimal designs for estimating the individual coefficients in a non-competitive enzyme inhibition kinetic model. Finally, the technical arguments are deferred to an appendix in Section 5.

## 2 A non-competitive model

Let  $S$  and  $I$  denote the concentration of the substrate and inhibition, respectively, and let  $\theta = (V, K_m, K_{ic})^T$  denote a vector of parameters. The non-competitive regression model considered in this paper is defined by

$$\eta(S, I, \theta) = \frac{V \cdot S}{(K_m + S)(1 + \frac{I}{K_{ic}})} , \quad (2.1)$$

where the explanatory variable  $(S, I)$  varies in the design space  $\mathcal{S} = [S_{\min}, S_{\max}] \times [I_{\min}, I_{\max}]$  defined by the constants  $0 \leq S_{\min} < S_{\max}$ ,  $0 \leq I_{\min} < I_{\max}$ . We assume that observations

$$Y_i = \eta(S_i, I_i, \theta) + \varepsilon_i , \quad i = 1, \dots, n \quad (2.2)$$

are available at experimental conditions  $(S_1, I_1), \dots, (S_n, I_n) \in \mathcal{S}$ , where  $\varepsilon_1, \dots, \varepsilon_n$  denote independent normally distributed random variables with mean 0 and variance  $\sigma^2 > 0$ . Following

Kiefer (1974) we define an approximate design as probability measure, say  $\xi$ , with finite support in  $\mathcal{S}$ . If the design  $\xi$  has masses  $\omega_i > 0$  at the different points  $(S_i, I_i)$  ( $i = 1, \dots, k$ ) and  $n$  observations can be made by the experimenter, the quantities  $\omega_i n$  are rounded to integers, say  $n_i$ , satisfying  $\sum_{i=1}^k n_i = n$ , and the experimenter takes  $n_i$  observations at each location  $(S_i, I_i)$  ( $i = 1, \dots, k$ ). Now, under regularity assumptions, standard asymptotic theory shows that the maximum likelihood estimator for the parameter  $\theta$  is asymptotically unbiased with variance  $\frac{\sigma^2}{n} M^{-1}(\xi, \theta)$ , where

$$M(\xi, \theta) = \int_{\mathcal{S}} \frac{\partial \eta(S, I, \theta)}{\partial \theta} \left( \frac{\partial \eta(S, I, \theta)}{\partial \theta} \right)^T d\xi(S, I). \quad (2.3)$$

denotes the information matrix of a design  $\xi$  in model (2.1) on the design space  $\mathcal{S}$  and

$$\frac{\partial \eta(S, I, \theta)}{\partial \theta} = \frac{S}{(K_m + S)} \frac{1}{(1 + I/K_{ic})} \left( 1, -\frac{V}{K_m + S}, \frac{V \cdot I/K_{ic}^2}{1 + I/K_{ic}} \right)^T, \quad (2.4)$$

is the gradient of the regression function with respect to the parameter  $\theta$ . An optimal design maximises an appropriate functional, say  $\Phi$ , of the matrix  $M(\xi, \theta)$  in the class of all designs on the design space  $\mathcal{S}$ , which is called optimality criterion. Several criteria have been proposed in the literature to discriminate between competing designs [see for example Pukelsheim (2006)] and in this paper we are interested in the widely used  $D$ -optimality criterion, which determines the design such that

$$\Phi_D\{M(\xi, \theta)\} = \det\{M(\xi, \theta)\} \quad (2.5)$$

is maximal, and in optimal designs for estimating the individual coefficients, which maximise

$$\Phi_{e_j}\{M(\xi, \theta)\} = (e_j M^{-1}(\xi, \theta) e_j)^{-1}, \quad j = 1, 2, 3. \quad (2.6)$$

Here  $e_j$  denotes the  $j$ th unit vector in  $\mathbb{R}^3$ ,  $M^{-1}(\xi, \theta)$  is a generalised inverse of the matrix  $M(\xi, \theta)$  and optimisation is only performed over the set of designs, such that the information matrices satisfy the range inclusion  $e_j \in \text{Range}(M(\xi, \theta))$ . In this case the right-hand side of (2.6) does not depend on the specific choice of the generalised inverse [see Pukelsheim (2006)]. A design maximising (2.5) is called  $D$ -optimal design, while a design maximising (2.6) is called  $e_j$ -optimal or optimal for estimating the  $j$ th coefficient of the vector  $\theta = (\theta_1, \theta_2, \theta_3)^T = (V, K_m, K_{ic})^T$  ( $j = 1, 2, 3$ ). For example, the  $e_2$ -optimal design is the optimal design for estimating the Michaelis-Menten constant  $\theta_2 = K_m$ .

Note that we discuss locally optimal designs, which require a-priori information about the unknown model parameters as they appear in the model in a nonlinear way [see Chernoff (1953)]. When preliminary knowledge regarding the unknown parameters of a nonlinear model is available, the application of locally optimal designs is well justified [see for example Dette et al. (2008)]. Locally optimal designs are also typically used as benchmarks for commonly

used designs and often serve as basis for constructing optimal designs with respect to more sophisticated optimality criteria, which are robust against a misspecification of the unknown parameters [see Pronzato and Walter (1985) or Chaloner and Verdinelli (1995), Dette (1997) among others]

### 3 A nonlinear transformation and $D$ -optimal designs

#### 3.1 Transformation of the design space

In this section we will provide a transformation which makes the structure of the resulting design problem more visible. The transformation will be used to solve the  $D$ -optimal design problem for the model (2.1) in Section 3.2 and for the solution of the  $e_j$ -optimal design problems in Section 4. Recalling the definition of the regression function in (2.1) we define a transformation by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \psi(S, I) = \begin{pmatrix} \frac{S}{K_m + S} \\ \frac{1}{1 + I/K_{ic}} \end{pmatrix}. \quad (3.1)$$

Note that (3.1) defines a one-to-one mapping from the original design space  $\mathcal{S} = [S_{\min}, S_{\max}] \times [I_{\min}, I_{\max}]$  onto the rectangle

$$\mathcal{X} = [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}], \quad (3.2)$$

where the boundary points of the two intervals are defined by

$$x_{\min} = \frac{S_{\min}}{K_m + S_{\min}}; \quad x_{\max} = \frac{S_{\max}}{K_m + S_{\max}}; \quad y_{\min} = \frac{1}{1 + I_{\max}/K_{ic}}; \quad y_{\max} = \frac{1}{1 + I_{\min}/K_{ic}}. \quad (3.3)$$

Moreover, the inverse transformation is given by

$$\begin{pmatrix} S \\ I \end{pmatrix} = \psi^{-1}(x, y) = \begin{pmatrix} \frac{xK_m}{1-x} \\ \frac{K_{ic}(1-y)}{y} \end{pmatrix}. \quad (3.4)$$

A straightforward calculation shows that the gradient in (2.4) can be represented by

$$\frac{\partial \eta(S, I, \theta)}{\partial \theta} = A(\theta) f(x, y),$$

where the non-singular matrix  $A(\theta)$  and the vector  $f$  are given by

$$A(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{V}{K_m} & \frac{V}{K_m} & 0 \\ \frac{V}{K_{ic}} & 0 & -\frac{V}{K_{ic}} \end{pmatrix},$$

$$f(x, y) = xy(1, x, y)^T, \quad (3.5)$$

respectively. If  $\tilde{\xi}$  denotes the design on the design space  $\mathcal{X}$  induced from the design  $\xi$  by the transformation (3.1), then the information matrix  $M(\xi, \theta)$  in (2.3) can be written as

$$M(\xi, \theta) = A(\theta)\tilde{M}(\tilde{\xi})A^T(\theta), \quad (3.6)$$

where the matrix  $\tilde{M}(\tilde{\xi})$  is defined by

$$\tilde{M}(\tilde{\xi}) = \int_{\mathcal{X}} f(x, y)f^T(x, y)d\tilde{\xi}(x, y). \quad (3.7)$$

As the transformation (3.1) is one-to-one, any design  $\xi$  on the design space  $\mathcal{S}$  induces a unique design  $\tilde{\xi}$  on  $\mathcal{X}$  and vice versa. Consequently, it follows from (3.6) that optimal designs maximising the functional  $\Phi(M(\xi, \theta))$  can be obtained by maximising the functional

$$\Phi(A^T(\theta)\tilde{M}(\tilde{\xi})A(\theta)).$$

### 3.2 $D$ -optimal designs

$D$ -optimal designs for the model (2.1) have been determined by Bogacka et al. (2011). In this section we explain how these designs can be obtained in a more transparent way from the regression model (3.5) using the transformation introduced in Section 3.1. We will use the same method to derive new optimal designs for estimating the individual coefficients in the subsequent Section 4.

In the case of  $D$ -optimality we have

$$\Phi_D(M(\xi, \theta)) = \Phi_D(A^T(\theta)\tilde{M}(\tilde{\xi})A(\theta)) = (\det A(\theta))^2(\det \tilde{M}(\tilde{\xi}))$$

where the matrix  $\tilde{M}(\tilde{\xi})$  is given by (3.7). Consequently, the  $D$ -optimal designs can be determined by first maximising the determinant  $\det \tilde{M}(\tilde{\xi})$  in the class of all designs  $\tilde{\xi}$  on the design space  $\mathcal{X}$  and then applying the inverse transformation (3.4). Our first result provides the solution of this problem.

**Theorem 3.1** *The  $D$ -optimal design maximising  $\det \tilde{M}(\tilde{\xi})$  in the class of all designs  $\tilde{\xi}$  on the set  $\mathcal{X}$  defined in (3.2) is given by*

$$\tilde{\xi}^* = \left( \begin{array}{ccc} (\max\{x_{\min}, \frac{x_{\max}}{2}\}, y_{\max}) & (x_{\max}, \max\{\frac{y_{\max}}{2}, y_{\min}\}) & (x_{\max}, y_{\max}) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right).$$

We can now use the inverse transformation (3.4) to obtain the  $D$ -optimal design on the original design space.

**Corollary 3.1** *The locally  $D$ -optimal design for the non-competitive model (2.1) in the class of all designs  $\xi$  on the set  $\mathcal{S}$  has equal masses at the three points*

$$\left(\max\left\{S_{\min}, \frac{S_{\max}K_m}{S_{\max} + 2K_m}\right\}, I_{\min}\right); \left(S_{\max}, \min\{K_{ic} + 2I_{\min}, I_{\max}\}\right); \left(S_{\max}, I_{\min}\right).$$

**Proof.** We use the inverse transformation (3.4) and Theorem 3.1 to obtain the  $D$ -optimal design on the original design space. For this purpose note that the first coordinate in (3.4) defines an increasing function of  $x$ , while the second coordinate is a decreasing function of  $y$ . Observing the relations in (3.3) we obtain that

$$\psi^{-1}\left(\max\left\{x_{\min}, \frac{x_{\max}}{2}\right\}, y_{\max}\right) = \left(\max\left\{S_{\min}, \frac{S_{\max}K_m}{S_{\max} + 2K_m}\right\}, I_{\min}\right).$$

□

## 4 Optimal designs for estimating individual coefficients

In this section we use the transformation (3.1) to derive optimal designs for estimating the individual coefficients  $V, K_m$  and  $K_{ic}$  in the non-competitive regression model (2.1). To our best knowledge all results presented here are new.

For this purpose consider the vector of parameters  $\theta = (\theta_1, \theta_2, \theta_3)^T = (V, K_m, K_{ic})^T$  and recall that an optimal design for estimating the coefficient  $\theta_j$  maximises the expression (2.6) in the class of all designs  $\xi$  satisfying the range inclusion

$$e_j \in \text{Range}(M(\xi, \theta)). \tag{4.1}$$

(here  $e_j \in \mathbb{R}^3$  denotes the  $j$ th unit vector ( $j = 1, 2, 3$ )). Now recall the definition of the transformation in (3.6). As the matrix  $A(\theta)$  is non-singular, it follows that any generalised inverse  $G$  of the matrix  $M(\xi, \theta)$  induces a generalised inverse of the matrix  $\tilde{M}(\tilde{\xi})$  by the transformation  $G \rightarrow A^T(\theta)GA(\theta)$  and vice versa. Similarly, if  $\xi$  satisfies (4.1), then  $\tilde{\xi}$  satisfies the range inclusion

$$c_j = B(\theta)e_j \in \text{Range}(\tilde{M}(\tilde{\xi})), \tag{4.2}$$

where

$$B(\theta) = (A(\theta))^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \frac{K_m}{V} & 0 \\ 1 & 0 & -\frac{K_{ic}}{V} \end{pmatrix}. \tag{4.3}$$

Consequently, an optimal design  $\xi^*$  maximising the functional (2.6) in the class of all designs  $\xi$  satisfying (4.1) can be determined by maximising the functional

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

in the class of all designs  $\tilde{\xi}$  satisfying (4.2) and transforming the design  $\tilde{\xi}$  back onto the design space  $\mathcal{S}$  using the inverse transformation (3.4).

In the following two sections we derive the optimal designs for estimating the individual coefficients using this approach. Section 4.1 contains the optimal designs for the Michaelis-Menten constant  $K_m$  and the dissociation constant  $K_{ic}$ , that is  $j = 2, 3$ . It will turn out that these two cases are almost the same in the transformed model (3.5). Section 4.2 contains the optimal design for estimating the maximum velocity  $V$  (that is  $j = 3$ ), where the determination of the optimal designs is much harder.

#### 4.1 Optimal designs for estimating the Michaelis-Menten constant and the dissociation constant

The optimal design problem for estimating the Michaelis-Menten constant  $K_m$  in the non-competitive model (2.1) corresponds to the choice  $j = 2$  in (2.6). By the discussion at the beginning of this section we have to determine the design maximising the expression

$$(c_2^T \tilde{M}^{-}(\tilde{\xi}) c_2)^{-1} = ((B(\theta) e_2)^T \tilde{M}^{-}(\tilde{\xi}) (B(\theta) e_2))^{-1} = \left( \frac{V}{K_m} \right)^2 (e_2^T \tilde{M}^{-}(\tilde{\xi}) e_2)^{-1} \quad (4.4)$$

in the class of all designs satisfying the range inclusion (4.2) for  $j = 2$ . If an optimal design for estimating the parameter  $K_{ic}$  is of particular interest, we use the criterion (2.6) with  $j = 3$ . Consequently, we have to maximise

$$(c_3^T \tilde{M}^{-}(\tilde{\xi}) c_3)^{-1} = ((B(\theta) e_3)^T \tilde{M}^{-}(\tilde{\xi}) (B(\theta) e_3))^{-1} = \left( \frac{V}{K_{ic}} \right)^2 (e_3^T \tilde{M}^{-}(\tilde{\xi}) e_3)^{-1} \quad (4.5)$$

in the class of all designs satisfying the range inclusion (4.2) for  $j = 3$ .

**Theorem 4.1** (1) *The optimal design maximising (4.4) is of the form*

$$\tilde{\xi} = \begin{pmatrix} (x_{\max}, y_{\max}) & (\bar{x}, y_{\max}) \\ \frac{\bar{x}}{1+\bar{x}} & \frac{1}{1+\bar{x}} \end{pmatrix} \quad (4.6)$$

where  $\bar{x} = \max \{x_{\min}, (\sqrt{2} - 1)x_{\max}\}$ .



(2) The optimal design maximising (4.5) is of the form

$$\tilde{\xi} = \begin{pmatrix} (x_{\max}, y_{\max}) & (x_{\max}, \bar{y}) \\ \frac{\bar{y}}{1+\bar{y}} & \frac{1}{1+\bar{y}} \end{pmatrix} \quad (4.7)$$

where  $\bar{y} = \max \{y_{\min}, (\sqrt{2} - 1)y_{\max}\}$ .

#### Corollary 4.1

(1) The optimal design for estimating the Michaelis-Menten constant in model (2.1) is given by

$$\xi = \begin{pmatrix} (S_{\max}, I_{\min}) & (\bar{S}, I_{\min}) \\ 1 - \omega & \omega \end{pmatrix}. \quad (4.8)$$

where

$$\bar{S} = \max \left\{ S_{\min}, \frac{K_m S_{\max} (\sqrt{2} - 1)}{K_m + (2 - \sqrt{2}) S_{\max}} \right\} \quad (4.9)$$

$$\omega = \left( 1 + \max \left\{ \frac{S_{\min}}{K_m + S_{\min}}, \frac{(\sqrt{2} - 1) S_{\max}}{K_m + S_{\max}} \right\} \right)^{-1}. \quad (4.10)$$

(2) The optimal design for estimating the parameter  $K_{ic}$  in model (2.1) is given by

$$\xi = \begin{pmatrix} (S_{\max}, I_{\min}) & (S_{\max}, \bar{I}) \\ 1 - \omega & \omega \end{pmatrix} \quad (4.11)$$

where

$$\bar{I} = \min \left\{ I_{\max}, I_{\min} (\sqrt{2} + 1) + K_{ic} \sqrt{2} \right\}$$

$$\omega = \left( 1 + K_{ic} \max \left\{ \frac{1}{K_{ic} + I_{\max}}, \frac{(\sqrt{2} - 1)}{K_{ic} + I_{\min}} \right\} \right)^{-1}.$$

**Proof.** We need to transform the design given by (4.6) and (4.7) to the actual design space  $\mathcal{S}$ . For the design (4.6) we note that (3.1) and (3.3) yield

$$\bar{x} = \max \left\{ \frac{S_{\min}}{S_{\min} + K_m}, (\sqrt{2} - 1) \frac{S_{\max}}{S_{\max} + K_m} \right\},$$

which gives for  $\frac{1}{1+\bar{x}}$  the weight in (4.10). Due to the fact that the first component of the inverse  $\psi^{-1}$  defined in (3.4) is increasing with respect to  $x$  we obtain

$$(\psi^{-1}(\bar{x}, y_{\max}))_1 = \bar{S},$$

where  $\bar{S}$  is defined in (4.9). Similar calculations for the other support point result in the design  $\xi$  in (4.8), which completes the proof of the first part of Corollary 4.1. The second part follows by similar arguments and the details are omitted for the sake of brevity.  $\square$

## 4.2 Optimal designs for estimating the maximal velocity

If the maximal velocity  $V$  is the parameter of main interest, one can use the optimality criterion (2.6) with  $j = 1$  to find the optimal design for a most precise estimation of  $V$ . Observing (4.3) we obtain from (4.2) that  $c_1 = (1, 1, 1)^T$  and the optimal design for estimating the maximum velocity can be found by maximising the criterion

$$(c_1^T \tilde{M}^{-1}(\tilde{\xi}) c_1)^{-1} \quad (4.12)$$

in the class of all designs such that the range inclusion (4.2) holds for  $j = 1$ . The solution of this problem is substantially harder and is based on the solution of an auxiliary optimal design problem for the weighted polynomial regression model

$$\mathbb{E}[Y|x] = xg(x, q)(\theta_1 + \theta_2 x), \quad (4.13)$$

where  $q \in [0, 1]$  is a fixed but arbitrary constant,

$$g(x, q) = qx + 1 - q \quad (4.14)$$

and the design space is given by  $[x_{\min}, x_{\max}]$ . Note that  $x_{\max} < 1$ , which follows from the definition (3.3), and without loss of generality we assume  $x_{\min} > 0$  (see Section 5.3 for further explanations). As the function  $g(x, q)$  is positive on the interval  $[x_{\min}, x_{\max}]$  it is easy to show that the set  $\{xg(x, q), x^2g(x, q)\}$  is a Chebychev system on the interval  $[x_{\min}, x_{\max}]$  [see Karlin and Studden (1966) for a definition of Chebyshev systems]. It then follows that there exists a unique function (depending on the value of  $q$ ) of the form

$$\Psi(x, q) = xg(x, q)(\psi_1 + \psi_2 x) \quad (4.15)$$

(with constants  $\psi_1, \psi_2 \in \mathbb{R}$ ) such that

1.  $|\Psi(x, q)| \leq 1$  for all  $x \in [x_{\min}, x_{\max}]$ .
2.  $\Psi(x_{\max}, q) = 1$  and there exists exactly one additional point  $\bar{x}(q) \in [x_{\min}, x_{\max}]$  such that  $\Psi(\bar{x}(q), q) = -1$ .

This function  $\Psi$  is called equi-oscillating polynomial in the literature and the points  $x_{\max}$  and  $\bar{x}(q)$  are called extremal points of  $\Psi$ . Some equi-oscillating polynomials with their corresponding extremal points for different values of  $q$  are depicted in Figure 1. We observe that  $x_{\max}$  is always

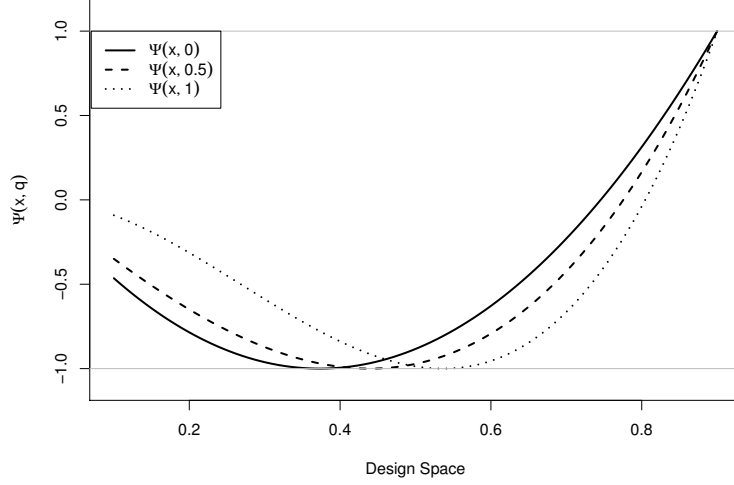


Figure 1: The equi-oscillating polynomials  $\Psi(x, q)$  for  $q = 0$  (solid line),  $q = 0.5$  (dashed line) and  $q = 1$  (dotted line) where the design space is given by  $[0.1, 0.9]$ .

an extremal point and that the other extremal point increases with the vale of  $q$ . We are now in a position to describe the design maximising (4.12) explicitly.

**Theorem 4.2**

a) If  $x_{\max} \leq y_{\max}$ , the optimal design maximising (4.12) is of the form

$$\tilde{\xi}(q^*) = \begin{pmatrix} (\bar{x}, g(\bar{x}, q^*)) & (x_{\max}, y_{\max}) \\ \omega(q^*) & 1 - \omega(q^*) \end{pmatrix} \quad (4.16)$$

where

$$\omega(q) = \frac{x_{\max}g(x_{\max}, q)(1 - x_{\max})}{x_{\max}g(x_{\max}, q)(1 - x_{\max}) + \bar{x}(q)g(\bar{x}(q), q)(1 - \bar{x}(q))} \quad (4.17)$$

and

$$q^* = \frac{1 - y_{\max}}{1 - x_{\max}}. \quad (4.18)$$

(b) If  $x_{\max} \geq y_{\max}$ , the optimal design maximising (4.12) can be obtained from part (a) by interchanging the roles of  $x$  and  $y$ .

In Figure 2 we display the support point  $\bar{x}(q)$  and the corresponding weight  $\omega(q)$  of the  $c$ -optimal design in (4.16) in dependence of the value  $q$  for  $x_{\max} = 0.9 < y_{\max}$ . We observe that the support point and corresponding weight are increasing with the value of  $q$ .

Using similar arguments as given in Section 3 the optimal design for estimating the maximum velocity can be obtained by the inverse transformation (3.4) from the design in Theorem 4.2.

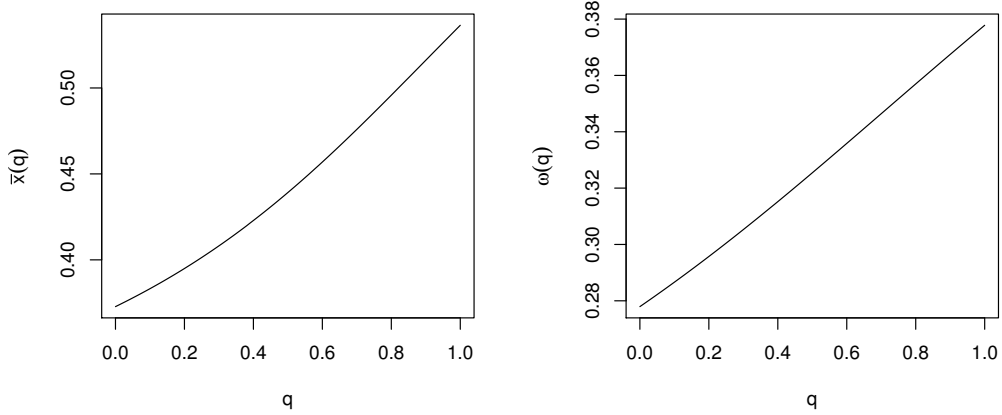


Figure 2: The support point  $\bar{x}(q)$  (left panel) and the corresponding weight  $\omega(q)$  (right panel) of the  $c$ -optimal design given by (4.16) in dependence of the value  $q$  for  $x_{\max} = 0.9 < y_{\max}$ .

The explicit formulas in the general case are omitted for the sake of brevity and we restrict ourselves to the case  $I_{\min} = 0$ , which is most important from a practical point of view.

**Corollary 4.2** *If  $I_{\min} = 0$  the optimal design for estimating the maximum velocity in the non-competitive regression model (2.1) is given by*

$$\xi^* = \begin{pmatrix} (\bar{S}, I_{\min}) & (S_{\max}, I_{\min}) \\ \omega & 1 - \omega \end{pmatrix},$$

where

$$\bar{S} = \max \left\{ S_{\min}, \frac{K_m S_{\max} (\sqrt{2} - 1)}{K_m + (2 - \sqrt{2}) S_{\max}} \right\}$$

$$\omega = \frac{S_{\max} (K_m + \bar{S})^2}{S_{\max} (K_m + \bar{S})^2 + \bar{S} (K_m + S_{\max})^2}.$$

**Proof.** If  $I_{\min} = 0$  it follows from (3.1) that  $y_{\max} = 1$ , which gives  $q = 0$  and for the function  $g$  in (4.14)  $g(x, 0) \equiv 1$ . The equi-oscillating polynomial corresponding to the Chebyshev system  $\{x, x^2\}$  on the interval  $[x_{\min}, x_{\max}]$  is given by

$$\underline{\Psi}(x, 0) = \begin{cases} -\left(\frac{x}{\bar{x}}\right)^2 + 2\frac{x}{\bar{x}} & \text{if } x_{\min} \leq x_{\max}(\sqrt{2} - 1) \\ \frac{(x_{\min} + x_{\max})}{x_{\min} x_{\max} (x_{\max} - x_{\min})} x^2 - \frac{(x_{\min}^2 + x_{\max}^2)}{x_{\min} x_{\max} (x_{\max} - x_{\min})} x & \text{if } x_{\min} > x_{\max}(\sqrt{2} - 1) \end{cases}$$

where  $\bar{x} = \bar{x}(0) = x_{\max}(\sqrt{2} - 1)$ . Here the equi-oscillating polynomial in the case  $x_{\min} \leq x_{\max}(\sqrt{2} - 1)$  is determined in the proof of Theorem 4.2 (see Section 5.3) and the remaining case

follows by similar arguments. Consequently, Theorem 4.2 shows that the the design maximising (4.5) is given by

$$\tilde{\xi} = \left( \begin{array}{cc} (\max(x_{\min}, \bar{x}), y_{\max}) & (x_{\max}, y_{\max}) \\ \frac{x_{\max}(1-x_{\max})}{x_{\max}(1-x_{\max})+\bar{x}(1-\bar{x})} & \frac{\bar{x}(1-\bar{x})}{x_{\max}(1-x_{\max})+\bar{x}(1-\bar{x})} \end{array} \right), \quad (4.19)$$

and the assertion now follows by the inverse transformation (3.4).  $\square$

**Acknowledgements** The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise. This work has also been supported in part by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823, Teilprojekt C2) of the German Research Foundation (DFG).

## References

- Atkinson, A. C. and Bogacka, B. (2013). Robust experimental design for choosing between models of enzyme inhibition. In D. Ucinski, A. C. Atkinson, M. P., editor, *mODa 10 - Advances in Model-Oriented Design and Analysis. Contributions to Statistics*. Springer.
- Bogacka, B., Patan, M., Johnson, P. J., Youdim, K., and Atkinson, A. (2011). Optimum design of experiments for enzyme inhibition kinetic models. *Journal of Biopharmaceutical Statistics*, 21(3):555–572.
- 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.
- 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 michaelismenten model. *Journal of the American Statistical Association*, 98(463):679–686.
- Dette, H., Bretz, F., Pepelyshev, A., and Pinheiro, J. C. (2008). Optimal designs for dose finding studies. *Journal of the American Statistical Association*, 103(483):1225–1237.
- Dette, H. and Grigoriev, Y. (2014). E-optimal designs for second-order response surface models. *Ann. Statist.*, 42(4):1635–1656.
- Dette, H. and Kunert, J. (2014). Optimal designs for the Michaelis-Menten model with correlated observations. *Statistics: A Journal of Theoretical and Applied Statistics*, 48(6):1254–1267.
- Dette, H. and Roeder, I. (1996). Optimal product designs for multivariate regression with missing terms. *Scan. J. Statist.*, 23(2):195208.
- Dette, H. and Roeder, I. (1997). Optimal discrimination designs for multifactor experiments. *Ann. Statist.*, 25(3):1161–1175.
- Draper, N. R., Heiligers, B., and Pukelsheim, F. (2000). Kiefer ordering of simplex designs for second-degree mixture models with four or more ingredients. *Annals of Statistics*, 28(2):578–590.

- Draper, N. R. and Pukelsheim, F. (2003). Canonical reduction of second-order fitted models subject to linear restrictions. *Statistics & Probability Letters*, 63(4):401–410.
- Elfving, G. (1952). Optimal allocation in linear regression theory. *The Annals of Mathematical Statistics*, 23:255–262.
- Farrel, R. H., Kiefer, J., and Walbran, A. (1967). Optimum multivariate designs. *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, 1:113–138.
- Harville, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. Springer, New York.
- Karlin, S. and Studden, W. J. (1966). *Tchebysheff Systems: With Application in Analysis and Statistics*. Wiley, New York.
- Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *Annals of Statistics*, 2:849–879.
- Kono, K. (1962). Optimum design for quadratic regression on  $k$ -cube. *Memoirs of the Faculty of Science. Kyushu University. Series A. Mathematics*, 16:114–122.
- López-Fidalgo, J. and Wong, W. K. (2002). Design issues for the Michaelis-Menten model. *Journal of Theoretical Biology*, 215(1):1–11.
- Matthews, J. N. S. and Allcock, G. C. (2004). Optimal designs for michaelismenten kinetic studies. *Statistics in Medicine*, 23(3):477–491.
- Michaelis, L. and Menten, M. (1913). Die Kinetik der Invertinwirkung. *Biochemische Zeitung*, 49:333–369.
- Pronzato, L. and Walter, E. (1985). Robust experimental design via stochastic approximation. *Mathematical Biosciences*, 75:103–120.
- Pukelsheim, F. (2006). *Optimal Design of Experiments*. SIAM, Philadelphia.
- Rasch, D. (1990). Optimum experimental design in nonlinear regression. *Communications in Statistics: Theory Methods.*, 19(12):4789–4806.
- Segel, I. H. (1993). *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*, 2nd edn. Wiley, New York.
- 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(6):1503–1516.
- Youdim, K. A., son, A. C., Patan, M., Bogacka, B., and Johnson, P. J. (2010). Potential application of d-optimal designs in the efficient investigation of cytochrome p450 inhibition kinetic models. *Drug Metabolism and Disposition*, 38(7):1019–1023.

## 5 Appendix: Proofs

### 5.1 Proof of Theorem 3.1

A further transformation shows that

$$f(x, y) = xy(1, x, y)^T = Cf\left(\frac{x}{x_{\max}}, \frac{y}{y_{\max}}\right),$$

where the matrix  $C$  is given by the diagonal  $3 \times 3$  matrix  $C = x_{\max}y_{\max} \cdot \text{diag}(1, x_{\max}, y_{\max})$ . Consequently, we may assume without loss of generality

$$0 \leq x_{\min} < x_{\max} = 1; \quad 0 < y_{\min} < y_{\max} = 1.$$

For the sake of brevity we now restrict the discussion to the case that  $0 \leq x_{\min}, y_{\min} < \frac{1}{2}$ , the other cases can be treated similarly. In this case the resulting design

$$\tilde{\xi}^* = \begin{pmatrix} (\frac{1}{2}, 1) & (1, \frac{1}{2}) & (1, 1) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad (4.20)$$

and, by the equivalence theorem for  $D$ -optimality this design is  $D$ -optimal if and only if the inequality

$$\begin{aligned} \kappa(x, y) &= x^2y^2(1, x, y)\tilde{M}^{-1}(\tilde{\xi}^*)(1, x, y)^T - 3 \\ &= 3x^2y^2(20x^2 - 44x + 8xy + 20y^2 - 44y + 41) - 3 \leq 0 \end{aligned} \quad (4.21)$$

holds for all  $(x, y) \in \mathcal{X} = [x_{\min}, 1] \times [y_{\min}, 1]$ . Note that the function  $\kappa$  is symmetric in  $x$  and  $y$ . We now calculate the points of  $\kappa$  in  $\mathcal{X}$  where the maxima are attained and show that the inequality (4.21) holds for these points.

The stationary points of  $\kappa$  in  $(x_{\min}, 1) \times (y_{\min}, 1)$  are given by

$$(x_1, y_1) = (\frac{1}{172}(55 - \sqrt{73}), \frac{1}{172}(55 - \sqrt{73})), \quad (x_2, y_2) = (\frac{1}{172}(55 + \sqrt{73}), \frac{1}{172}(55 + \sqrt{73})).$$

Considering the Hessian matrix for these points it follows that  $(x_1, y_1)$  is a saddle point of  $\kappa$  whereas  $(x_2, y_2)$  is a minimum. Consequently, the maxima of the function  $\kappa$  are attained at the boundary of  $\mathcal{X}$ .

Because of the symmetry of  $\kappa$  it is sufficient to calculate the maxima of the functions

$$\begin{aligned} \kappa_1(x) &= \kappa(x, y_{\min}) = 60y_{\min}^2 x^4 + (24y_{\min}^3 - 132y_{\min}^2) x^3 \\ &\quad + (60y_{\min}^4 - 132y_{\min}^3 + 123y_{\min}^2) x^2 - 3. \end{aligned} \quad (4.22)$$

$$\kappa_2(x) = \kappa(x, 1) = 60x^4 - 36x^3 + 17x^2 - 3 \quad (4.23)$$

for  $x \in [x_{\min}, 1]$ . We start with the maximisation of  $\kappa_1(x)$ . For  $0 < y_{\min} \leq \frac{1}{2}$  the only local extremum of  $\kappa_1$  is a minimum attained in  $x^* = 0$  with value  $\kappa_1(0) = -3$ . Consequently,  $\kappa_1$  is increasing for  $x > 0$  and its maximum on  $[x_{\min}, 1]$  is attained in  $x^* = 1$ .

The stationary points of the function  $\kappa_2$  in (4.23) are given by  $x = 0$ ,  $x = \frac{1}{2}$ ,  $x = \frac{17}{20}$ . Since  $\kappa_2$  is a polynomial of degree 4 with positive leading coefficient the locally minima of  $\kappa_2$  are attained at  $x = 0$  and  $x = \frac{17}{20}$ . Consequently, the maxima of  $\kappa_2$  on  $[x_{\min}, 1]$  are attained in  $x^* = \frac{1}{2}$  and

$x^* = 1$ . Using the symmetry the maxima of the function  $\kappa$  given in (4.21) on  $\mathcal{X}$  are given by

$$(x_1^*, y_1^*) = (1, y_{\min}), \quad (x_2^*, y_2^*) = (\tfrac{1}{2}, 1), \quad (x_3^*, y_3^*) = (1, \tfrac{1}{2}), \quad (x_4^*, y_4^*) = (1, 1).$$

Obviously,  $\kappa(x_2^*, y_2^*) = \kappa(x_3^*, y_3^*) = \kappa(x_4^*, y_4^*) = 0$  and additionally  $\kappa(x_1^*, y_1^*) \leq 0$  for  $y_{\min} \leq \frac{1}{2}$ , since the function

$$h(y_{\min}) = \kappa_1(1) = 60y_{\min}^4 - 108y_{\min}^3 + 51y_{\min}^2 - 3$$

is increasing on  $[0, \frac{1}{2}]$  with  $h(\frac{1}{2}) = 0$ . Consequently, the inequality (4.21) holds for all  $(x, y) \in \mathcal{X}$  and the design given by (4.20) is  $D$ -optimal.  $\square$

## 5.2 Proof of Theorem 4.1

(1) We have to solve the  $e_2$ -optimal design problem in a linear regression model with a vector of regression functions given by (3.5). For its solution we use Elfving's theorem [see Elfving (1952)], which states that a design  $\tilde{\xi}$  with masses  $p_1, \dots, p_k$  at the points  $(x_1, y_1), \dots, (x_k, y_k)$  is  $e_2$ -optimal if and only if there exists a constant  $\gamma > 0$  and constants  $\varepsilon_1, \dots, \varepsilon_k \in \{-1, 1\}$ , such that the vector  $\gamma e_2$  has the representation

$$\gamma e_2 = \sum_{j=1}^k \varepsilon_j p_j f(x_j, y_j), \quad (4.24)$$

and is a boundary point of the *Elfving set*

$$\mathcal{R} = \text{conv}(\{\varepsilon f(x, y) \mid (x, y) \in \mathcal{X}; \varepsilon \in \{-1, 1\}\}) \quad (4.25)$$

(here  $\text{conv}(\mathcal{A})$  denotes the convex hull of a set  $\mathcal{A} \subset \mathbb{R}^3$ ). We now consider the design  $\tilde{\xi}$  given by (4.6). A similar reasoning as in Section 3.2 shows that in the case of  $j = 2$  we may assume without loss of generality that  $x_{\max} = y_{\max} = 1$ . Then it is easy to see that the design  $\tilde{\xi}$  defined in (4.6) satisfies (4.24) with  $k = 2$ ,  $\varepsilon_1 = -\varepsilon_2 = 1$ ,  $(x_1, y_1) = (1, 1)$ ,  $(x_2, y_2) = (\bar{x}, 1)$  and

$$p_1 = \frac{\bar{x}}{1 + \bar{x}}; \quad p_2 = \frac{1}{1 + \bar{x}}; \quad \gamma = \frac{\bar{x}(1 - \bar{x})}{1 + \bar{x}}.$$

Therefore, it remains to prove that the corresponding vector  $\gamma e_2 = (0, \gamma, 0)^T$  is a boundary point of the Elfving set  $\mathcal{R}$  defined in (4.25). For this purpose we consider the vector

$$n = (n_1, n_2, n_3)^T = \left(1 - \frac{1 + \bar{x}}{\bar{x}(1 - \bar{x})}, \frac{1 + \bar{x}}{\bar{x}(1 - \bar{x})}, 0\right)^T$$



and show that this vector defines a supporting hyperplane to the set  $\mathcal{R}$  at  $\gamma e_2 = (0, \gamma, 0)^T$ , i.e.,

$$(\gamma e_2)^T n = 1, \quad (4.26)$$

$$z^T n \leq 1 \quad \text{for all } z \in \mathcal{R}. \quad (4.27)$$

Condition (4.26) is obviously satisfied. In order to prove (4.27) we establish the sufficient condition  $|f^T(x, y)n| \leq 1$  for all  $(x, y) \in \mathcal{X}$ , which can be rewritten as

$$-\bar{x}(1 - \bar{x}) \leq xy(-(\bar{x}^2 + 1) + (1 + \bar{x})x) \leq \bar{x}(1 - \bar{x}). \quad (4.28)$$

We now distinguish the cases

$$\begin{aligned} (i) \quad & 1 \geq x \geq \frac{1+\bar{x}^2}{1+\bar{x}} \\ (ii) \quad & \frac{1+\bar{x}^2}{1+\bar{x}} > x > x_{\min}. \end{aligned}$$

- (i) Here the first inequality in (4.28) is obviously satisfied and the second (upper) inequality holds if it can be established for  $y = 1$ , that is

$$x^2(1 + \bar{x}) - x(\bar{x}^2 + 1) - \bar{x}(1 - \bar{x}) \leq 0.$$

This inequality holds for all  $x \in [\frac{-1+\bar{x}}{1+\bar{x}}, 1]$  which proves (4.28) in the case (i).

- (ii) Here the second inequality in (4.28) is obviously satisfied and the first (lower) inequality holds if it can be established for  $y = 1$ , that is

$$h(x) = x^2(1 + \bar{x}) - x(\bar{x}^2 + 1) + \bar{x}(1 - \bar{x}) \geq 0.$$

The minimum of this function is attained at the point  $x^* = \frac{1+\bar{x}^2}{2(1+\bar{x})}$ . For  $\bar{x} = \sqrt{2} - 1$  the corresponding minimum value is given by  $h(x^*) = 0$  following that  $h(x) \geq 0$ . For  $\bar{x} = x_{\min} > \sqrt{2} - 1$  the point  $x^*$  is not contained in the interval  $[x_{\min}, 1]$ . For  $x > x^*$  the function  $h$  is increasing with  $h(x_{\min}) = 0$  such that (4.28) also holds in this case.

Consequently, by Elfving's theorem the design (4.6) is  $e_2$ -optimal in the linear regression model defined by the vector (3.5).

**(2)** In order to prove that the design given by (4.7) is  $e_3$ -optimal in the regression model defined by (3.5), we define the permutation matrix

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

and the information matrix

$$\tilde{M}_P(\tilde{\xi}) = P\tilde{M}(\tilde{\xi})P = \int_{\mathcal{X}} xy(1, y, x)^T xy(1, y, x) d\tilde{\xi}(x, y) = \int_{\mathcal{X}} f(y, x)f^T(y, x)d\tilde{\xi}(x, y).$$

Note that the permutation matrix  $P$  causes the exchange of the second and third component of the vector (3.5). Moreover, by using Lemma 9.2.4 of Harville (1997) we get

$$c_3^T \tilde{M}^{-1}(\tilde{\xi}) c_3 = \left(\frac{K_{ie}}{V}\right)^2 e_3^T \tilde{M}^{-1}(\tilde{\xi}) e_3 = \left(\frac{K_{ie}}{V}\right)^2 e_2^T P^T \tilde{M}^{-1}(\tilde{\xi}) P e_2 = \left(\frac{K_{ie}}{V}\right)^2 e_2^T \tilde{M}_P^{-1}(\tilde{\xi}) e_2.$$

By the first part of Theorem 4.1 it follows that  $(e_2^T \tilde{M}_P^{-1}(\tilde{\xi}) e_2)^{-1}$  is maximised by the design  $\tilde{\xi}$  given by (4.7).  $\square$

### 5.3 Proof of Theorem 4.2

Recall the definition of the weighted polynomial regression model (4.13) and note that in this model the vector of regression functions is given by

$$f(x, q) = xg(x, q) \begin{pmatrix} 1 \\ x \end{pmatrix}, \quad (5.1)$$

where the function  $g$  is defined in (4.14). The information matrix of a design  $\hat{\xi}$  in this model is given by

$$\hat{M}(\hat{\xi}) = \int_{x_{\min}}^{x_{\max}} f(x, q)f^T(x, q)d\hat{\xi}(x) = \int_{x_{\min}}^{x_{\max}} x^2 g^2(x, q)(1, x)(1, x)^T d\hat{\xi}(x). \quad (5.2)$$

If the lower bound of the design space is given by  $x_{\min} = 0$ , an optimal design does not contain 0 as support point. In order to prove this fact we consider the design

$$\hat{\xi} = \begin{pmatrix} 0 & x_2 & \dots & x_k \\ \omega_1 & \omega_2 & \dots & \omega_k \end{pmatrix},$$

with mass  $\omega_1 > 0$  at the point 0. Now consider the design

$$\hat{\mu} = \begin{pmatrix} x_2 & \dots & x_k \\ \frac{\omega_2}{1-\omega_1} & \dots & \frac{\omega_k}{1-\omega_1} \end{pmatrix}, \quad (5.3)$$

i.e., we remove the support point 0 and rescale the weights for the remaining support points appropriately. Then we obtain the following inequality with respect to the Loewner ordering

$$\hat{M}(\hat{\xi}) = \omega_1 \hat{M}(\delta_0) + (1 - \omega_1) \hat{M}(\hat{\mu}) = (1 - \omega_1) \hat{M}(\hat{\mu}) \leq M(\hat{\mu}),$$

where  $\delta_t$  is the Dirac measure in  $t$ . The last inequality is strict, whenever  $M(\hat{\mu})$  is non-singular. Consequently, every design  $\hat{\xi}$  which contains 0 as support point can be improved by another design  $\hat{\mu}$  of the form (5.3). Since all optimality criteria  $\Phi$  considered in this paper (in particular the  $c$ -optimality criterion) are monotonically increasing with respect to the Loewner ordering it follows that  $\Phi(\hat{M}(\hat{\xi})) \leq \Phi(\hat{M}(\hat{\mu}))$  and we always find an optimal design which does not contain the point 0 in its support. Therefore we can restrict ourselves to the consideration of  $x_{\min} > 0$  throughout this section.

The following auxiliary result gives a solution of the  $c$ -optimal design problem for the vector  $c = (1, 1)^T \in \mathbb{R}^2$ .

**Lemma 5.1** *The  $c$ -optimal design for the vector  $c = (1, 1) \in \mathbb{R}^2$  maximising*

$$((1, 1)\hat{M}^{-1}(\hat{\xi})(1, 1)^T)^{-1}$$

*in the class of all designs on the interval  $[x_{\min}, x_{\max}] \subset (0, 1)$  satisfying  $(1, 1)^T \in \text{Range}(\hat{M}(\hat{\xi}))$  is given by*

$$\hat{\xi}^*(q) = \begin{pmatrix} \bar{x}(q) & x_{\max} \\ \omega(q) & 1 - \omega(q) \end{pmatrix} \quad (5.4)$$

*where  $\omega(q)$  is defined by (4.17) and  $\bar{x}(q)$  and  $x_{\max}$  are the extremal points of the equi-oscillating polynomial  $\Psi(x, q)$  introduced at the beginning of Section 4.2.*

**Proof.** Recall that  $\{xg(x, q), x^2g(x, q)\}$  is a Chebychev system and note that the definition of the vector  $f(x, q)$  in (5.1) yields  $f(1, q) = c = (1, 1)^T$ . As  $x_0 = 1 \notin [x_{\min}, x_{\max}]$  the  $c$ -optimal design problem for the vector  $c = (1, 1)^T$  is in fact an optimal extrapolation design problem for the point  $x_0 = 1$ . Therefore it follows from Theorem X.7.7 in Karlin and Studden (1966) that the support points of  $\hat{\xi}^*(q)$  are given by the extremal points of the equi-oscillating polynomial  $\Psi(x, q)$  and the corresponding weights  $\omega_1(q), \omega_2(q)$  are given by

$$\omega_1(q) = 1 - \omega_2(q) = \frac{|L_1(1, q)|}{|L_1(1, q)| + |L_2(1, q)|},$$

where  $L_i(x, q) = xg(x, q)(\ell_{i1} + \ell_{i2}x)$  is the  $i$ th Lagrange interpolation polynomial with knots  $x_1 = \bar{x}(q)$  and  $x_2 = x_{\max}$  defined by the property  $L_i(x_j) = \delta_{ij}$ ,  $i = 1, 2$ ,  $j = 1, 2$  (here  $\delta_{ij}$  is the Kronecker symbol). A straightforward calculation now shows  $\omega(q)$  is given by (4.17), completing the proof of Lemma 5.1.  $\square$

It is also well known that the equi-oscillating polynomial  $\Psi(x, q)$  can be expressed in terms of the optimal design  $\hat{\xi}^*(q)$ , that is

$$\Psi(x, q) = \frac{(1, 1)\hat{M}^{-1}(\hat{\xi}^*(q))(1, x)^T xg(x, q)}{((1, 1)\hat{M}^{-1}(\hat{\xi}^*(q))(1, 1)^T)^{1/2}} \quad (5.5)$$

and the condition  $|\underline{\Psi}(x, q)| \leq 1$  for all  $x \in [x_{\min}, x_{\max}]$  is just an equivalent formulation of the equivalence theorem for  $c = (1, 1)^T$ -optimality in the regression model (4.13) [see Karlin and Studden (1966)].

We are now able to prove Theorem 4.2 using the equivalence theorem for  $c$ -optimality [see Pukelsheim (2006)], which states that the design  $\tilde{\xi}(q^*)$  (here  $q^*$  is defined by (4.18)) is  $c_1$ -optimal in the linear regression model defined by the vector (3.5) if and only if there exists a generalised inverse  $G$  of the matrix  $\tilde{M}(\tilde{\xi}(q^*))$  such that the inequality

$$(c_1^T G f(x, y))^2 \leq \kappa := c_1^T \tilde{M}^{-1}(\tilde{\xi}(q^*)) c_1 \quad (5.6)$$

holds for all  $(x, y) \in \mathcal{X}$ . For a proof of this statement we use the notation

$$\tilde{\xi} = \tilde{\xi}(q^*), \quad \hat{\xi}^* = \hat{\xi}^*(q^*), \quad g(x) = g(x, q^*), \quad \bar{x} = \bar{x}(q^*),$$

and first construct an appropriate generalised inverse  $G$  of the matrix  $\tilde{M}(\tilde{\xi})$  which can be used in (5.6). For this purpose define

$$G = P^{-T} H P^{-1}, \quad (5.7)$$

where the  $(3 \times 3)$ -matrices  $P$  and  $H$  are given by

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 - q^* & q^* & 1 \end{pmatrix}, \quad H = \begin{pmatrix} \hat{M}^{-1}(\hat{\xi}^*) & (\frac{\sqrt{\kappa}}{\bar{x}g^2(\bar{x})}, 0)^T \\ 0_2^T & 0 \end{pmatrix}, \quad (5.8)$$

and  $\hat{M}(\hat{\xi})$  and  $\kappa$  are defined in (5.2) and (5.6), respectively. Note that  $\tilde{M}(\tilde{\xi})$  can be rewritten in terms of  $P$  and  $\hat{M}(\hat{\xi})$ , that is

$$\begin{aligned} \tilde{M}(\tilde{\xi}) &= \int_{\mathcal{X}} \tilde{f}(x, y) \tilde{f}^T(x, y) d\tilde{\xi}(x, y) = \int_{\mathcal{X}} xy(1, x, y) xy(1, x, y)^T d\tilde{\xi}(x, y) \\ &= \int_{\hat{\mathcal{X}}} xg(x)(1, x, g(x)) xg(x)(1, x, g(x))^T d\hat{\xi}^*(x) \\ &= P \int_{\hat{\mathcal{X}}} xg(x)(1, x, 0) xg(x)(1, x, 0)^T d\hat{\xi}^*(x) P^T = P \begin{pmatrix} \hat{M}(\hat{\xi}^*) & 0 \\ 0_2^T & 0 \end{pmatrix} P^T. \end{aligned}$$

By a straightforward calculation we have

$$\begin{aligned}
\tilde{M}(\tilde{\xi})G\tilde{M}(\tilde{\xi}) &= P \begin{pmatrix} \hat{M}(\hat{\xi}^*) & 0 \\ 0_2^T & 0 \end{pmatrix} P^T P^{-T} H P^{-1} P \begin{pmatrix} \hat{M}(\hat{\xi}^*) & 0 \\ 0_2^T & 0 \end{pmatrix} P^T \\
&= P \begin{pmatrix} \hat{M}(\hat{\xi}^*) & 0 \\ 0_2^T & 0 \end{pmatrix} \begin{pmatrix} \hat{M}^{-1}(\hat{\xi}) & (\frac{\sqrt{\kappa}}{\bar{x}g^2(\bar{x})}, 0)^T \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{M}(\hat{\xi}^*) & 0 \\ 0_2^T & 0 \end{pmatrix} P^T \\
&= P \begin{pmatrix} \hat{M}(\hat{\xi}^*) & 0 \\ 0_2^T & 0 \end{pmatrix} P^T = \tilde{M}(\tilde{\xi}),
\end{aligned}$$

and it follows that  $G$  defined in (5.7) is a generalised inverse of the matrix  $\tilde{M}(\tilde{\xi})$ .

A simple calculation now shows  $\kappa = c_1^T \tilde{M}^{-1}(\tilde{\xi}) c_1 = (1, 1) \hat{M}^{-1}(\hat{\xi}^*) (1, 1)^T$  and using (5.5) we obtain for the left-hand side of (5.6)

$$\begin{aligned}
(c_1^T G f(x, y))^2 &= \left( (1, 1) \hat{M}^{-1}(\hat{\xi}^*) (1, x)^T xy + \frac{\sqrt{\kappa}}{\bar{x}g^2(\bar{x})} (y - g(x)) xy \right)^2 \\
&= \left( \frac{\sqrt{\kappa}}{\sqrt{\kappa}} (1, 1) \hat{M}^{-1}(\hat{\xi}^*) (1, x)^T xg(x) \frac{y}{g(x)} + \sqrt{\kappa} \frac{g^2(x)}{g^2(\bar{x})} \left( \frac{y}{g(x)} - 1 \right) \frac{x}{\bar{x}} \frac{y}{g(x)} \right)^2 \\
&= \kappa \left( \Psi(x, q^*) \frac{y}{g(x)} + \frac{g^2(x)}{g^2(\bar{x})} \left( \frac{y}{g(x)} - 1 \right) \frac{x}{\bar{x}} \frac{y}{g(x)} \right)^2 \\
&:= \kappa \tau^2(x, y),
\end{aligned}$$

where the last inequality defines the function  $\tau$  in an obvious manner. Consequently, proving that (5.6) is satisfied for all  $(x, y) \in \mathcal{X}$  is equivalent to proving that  $|\tau(x, y)| \leq 1$  for all  $(x, y) \in \mathcal{X}$ .

For the sake of transparency we therefore restrict ourselves to the case  $y_{\max} = 1$  and  $x_{\min} \leq x_{\max}(\sqrt{2} - 1)$ . The other cases can be proved in a similar way but with substantially more involved calculations. In this case it follows that  $q^* = 0$  and the function  $g$  reduces to  $g(x, 0) \equiv 1$ . Then the design  $\hat{\xi}^*$  is given by

$$\hat{\xi}^* = \hat{\xi}^*(0) = \begin{pmatrix} \bar{x} & x_{\max} \\ \frac{x_{\max}(1-x_{\max})}{x_{\max}(1-x_{\max})+\bar{x}(1-\bar{x})} & \frac{\bar{x}(1-\bar{x})}{x_{\max}(1-x_{\max})+\bar{x}(1-\bar{x})} \end{pmatrix}, \quad \bar{x} = x_{\max}(\sqrt{2} - 1) \quad (5.9)$$

with the corresponding oscillating polynomial  $\Psi(x, 0) = -\left(\frac{x}{\bar{x}}\right)^2 + 2\frac{x}{\bar{x}}$ . Consequently,

$$\tau(x, y) = y \frac{x}{\bar{x}} \left( y + \frac{x}{\bar{x}} - 3 \right). \quad (5.10)$$

The non-zero stationary point of  $\tau$  is given by  $(x^*, y^*) = (\bar{x}, 1)$  which is a minimum of  $\tau$  with  $\tau(x^*, y^*) = -1$ .

Now we consider  $\tau$  on the boundary of  $\mathcal{X} = [x_{\min}, x_{\max}] \times [y_{\min}, 1]$ . First, we set  $y = y_{\max} = 1$  so that  $\tau(x, y)$  reduces to

$$\tau_1(x) = \tau(x, y_{\max}) = \frac{x}{\bar{x}} \left( \frac{x}{\bar{x}} - 2 \right),$$

which is a parabola with minimum in  $x^* = \bar{x}$  and  $|\tau_1(x^*)| = 1 \leq 1$ . Moreover,  $\tau_1(x_{\max}) = (\sqrt{2} - 1)(\sqrt{2} + 1) = 1$  and for  $\tau_1(x_{\min}) \in (\tau_1(\bar{x}), \tau_1(0)) = (-1, 0)$ , which yields  $|\tau_1(x)| \leq 1$  for all  $x \in [x_{\min}, x_{\max}]$ . Next, we set  $y = y_{\min}$  so that  $\tau(x, y)$  reduces to

$$\tau_2(x) = \tau(x, y_{\min}) = y_{\min} \frac{x}{\bar{x}} \left( y_{\min} + \frac{x}{\bar{x}} - 3 \right)$$

which is again a parabola with minimum in  $x^* = \frac{3-y_{\min}}{2} \bar{x}$  and  $|\tau_x(x^*)| = \left| \left( \frac{3-y_{\min}}{2} \right)^2 y_{\min} \right|$ . Maximising the later term with respect to  $y_{\min} \in (0, y_{\max})$  it follows that  $|\tau_2(x^*)| \leq \frac{49}{54} \leq 1$ . By a similar argument it also follows that  $|\tau_2(x_{\max})| \leq 1$  and  $|\tau_2(x_{\min})| \leq 1$ . Setting  $x = x_{\min}$  and  $x = x_{\max}$  and using similar arguments we get  $|\tau(x_{\min}, y)| \leq 1$  and  $|\tau(x_{\max}, y)| \leq 1$  for all  $y \in [y_{\min}, 1]$ , respectively.

Consequently, the inequality  $|\tau(x, y)| \leq 1$  holds for all  $(x, y) \in \mathcal{X}$  and therefore the inequality (5.6) is satisfied  $(x, y) \in \mathcal{X}$ , which proves the  $c_1$ -optimality of  $\tilde{\xi} = \tilde{\xi}(0)$ .  $\square$