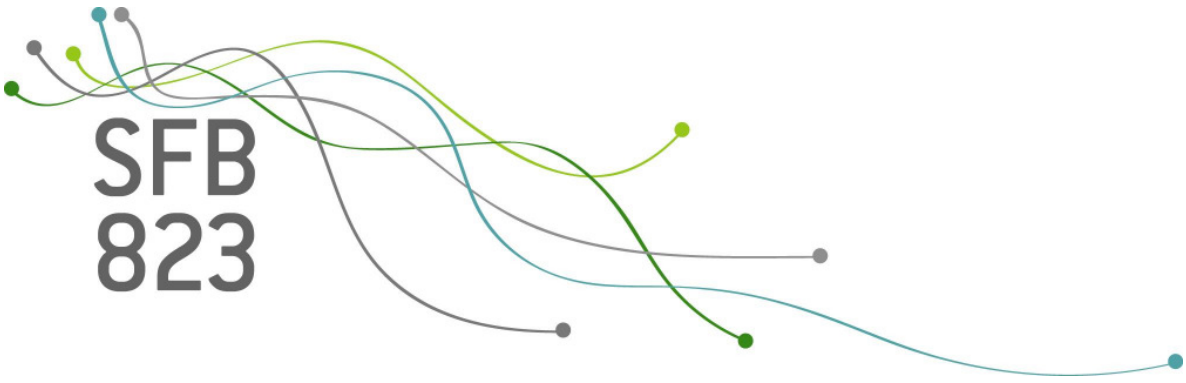
SFB 823

Optimal designs for random effect models with correlated errors with applications in population pharmacokinetics

Holger Dette, Andrey Pepelyshev,
Tim Holland-Letz

Nr. 17/2009

Discussion Paper



Optimal designs for random effect models with correlated errors with applications in population pharmacokinetics

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

Andrey Pepelyshev
St. Petersburg State University
Department of Mathematics
St. Petersburg, Russia
email: andrey@ap7236.spb.edu

Tim Holland-Letz
Ruhr-Universität Bochum
Medizinische Fakultät
44780 Bochum, Germany
email: tim.holland-letz@rub.de

16 July 2009

Abstract

We consider the problem of constructing optimal designs for population pharmacokinetics which use random effect models. It is common practice in the design of experiments in such studies to assume uncorrelated errors for each subject. In the present paper a new approach is introduced to determine efficient designs for nonlinear least squares estimation which addresses the problem of correlation between observations corresponding to the same subject. We use asymptotic arguments to derive optimal design densities, and the designs for finite sample size are constructed from the quantiles of the corresponding optimal distribution function. It is demonstrated that compared to the optimal exact designs, whose determination is a hard numerical problem, these designs are very efficient. Alternatively, the designs derived from asymptotic theory could be used as starting designs for the numerical computation of exact optimal designs. Several examples of linear and nonlinear models are presented in order to illustrate the methodology.

Keyword and Phrases: random effect models, nonlinear least squares estimate, correlated observations, compartmental models, asymptotic optimal design density

AMS Subject Classification: 62K05

1 Introduction

The work presented in this paper is motivated by some problems encountered in the optimal design of a clinical trial to establish the pharmacokinetics of Uzara®, a digoxin related herbal diarrhea medication [based on Thürmann et al. (2004)]. These kinds of trials pose methodological design challenges because they combine both the estimation of global population parameters and correlated measurement errors. The trial in question included a number of patients each given an oral application of Uzara®, with the resulting serum concentration of digitoxin being measured repeatedly during the next 36 hours.

Situations of this kind are rather common in the evaluation of the pharmacokinetics and the pharmacodynamics of drugs (see Buelga et al. (2005), Colombo et al. (2006) among others) and are usually modeled by linear or nonlinear random effects models, which allow the estimation of population parameters, that is, the mean and the inter-individual variability of the parameters. Under the additional assumption of a normal distribution, the population characteristics are usually estimated by maximum likelihood methods. In many cases the likelihood cannot be evaluated explicitly and approximations are used for the calculation of the estimate. Efficient algorithms are available for this purpose [see Aarons (1999)]. Loosely speaking, under a Gaussian assumption this approach corresponds to weighted nonlinear least squares estimation. It was pointed out by several authors that the application of an appropriate design in these studies can increase the efficiency of the population approach substantially. Usually the choice of an appropriate design is based on the Fisher information matrix which cannot be derived explicitly in pharmacokinetic models with random effects. For this reason many authors propose an approximation of the likelihood [see for example Retout et al. (2002), Mentré et al. (1997) or Retout and Mentré (2003) or Schmelter (2007a) oder Schmelter (2007b) among others], which is then used to derive an approximation for the Fisher information matrix. This matrix is the basis of various optimality criteria, which have been proposed in the literature for the construction of optimal designs for random effect regression models.

The relation between time and concentration in the analysis of the Uzara® trial can be described using the theory of one compartment models with oral application (Atkinson et al. (1993), Shargel (1993)). More precisely, the digitoxin concentration was modeled using the 3 parameter Bateman function

$$\eta(t, b) = b_3(e^{-b_1 t} - e^{-b_2 t}) \tag{1.1}$$

where $\eta(t,b)$ denotes the measured concentration, t is the time (in hours) and $b=(b_1,b_2,b_3)$ is the vector of parameters (Garrett (1994)). The parameters are assumed to vary between patients and it is the aim to estimate their global means (and sometimes variances) over all patients. Measurements within the same patient are correlated, and we assume this correlation to be proportional to the time lag between measurements. More precisely, if $t_1 < \ldots < t_n$ denote the time points of measurements, then the $(n \times n)$ covariance matrix of the observations corresponding to a patient is given by

$$V_\varepsilon = \sigma^2(\gamma \exp(-\lambda|t_i - t_j|) + (1-\gamma)\delta_{i,j})_{i,j=1,\ldots,n} \tag{1.2}$$

where $\delta_{i,j}$ denotes Kronecker's symbol and $\gamma \in [0,1]$ is a constant. In a follow-up trial Thrmann considered $n=15$ measurements each on $K=18$ patients. Measurements for different patients are assumed to be independent and were taken at non-optimized time points $0, 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8, 10, 12, 15, 24, 36$ (hours). If $\beta=(\beta_1,\beta_2,\beta_3)^T$ denotes the vector of the population mean of the parameters b, V_p the corresponding population covariance, an approximation of the covariance of a single patient can be expressed as

$$\Sigma_{pop} = \frac{\partial \eta(\mathbf{t},\beta)}{\partial \beta}^T V_p \frac{\partial \eta(\mathbf{t},\beta)}{\partial \beta} + V_\varepsilon, \tag{1.3}$$

where $\eta(\mathbf{t},\beta) = (\eta(t_1,\beta),...,\eta(t_n,\beta))^T$ denotes the vector of expected responses at $t_1,...,t_n$. We have received the estimates $\hat{\beta} = (0.2, 0.135, 28)^T$,

$$\hat{V}_p = \begin{pmatrix} 0.0025 & 0.0019 & 0 \\ 0.0019 & 0.0016 & 0 \\ 0 & 0 & 144 \end{pmatrix}$$

for the population parameters, while the estimates of the parameters in the covariance matrix (1.3) are given by $\hat{\gamma}=0.8$, $\hat{\lambda}=0.01$, $\hat{\sigma}^2=0.2$.

While we were considering optimal design problems for trials of this type, several questions appeared which motivated the research presented in this paper. First, the estimation of the population mean and the construction of corresponding optimal designs for population pharmacokinetics depends sensitively on the Gaussian assumption, which is usually made for computational convenience. The maximum likelihood estimates may be inconsistent if basic distributional assumptions are violated. As a consequence, the derived optimal designs might be inefficient. Second, most authors derive the approximation for the Fisher information matrix under the additional assumption that the random errors corresponding to the measurements of each individual are uncorrelated [see e.g. Retout et al. (2001), Retout et al. (2002) or Retout and Mentré (2003) among many others]. However, this is not a realistic assumption for many applications in population pharmacokinetics and a general concept for constructing optimal designs in the general context is still missing. Third, even if the Gaussian assumption and the assumption of

uncorrelated errors for each subject can be justified, the numerical construction of the estimate and the corresponding optimal design is extremely hard.
In the present paper we try to address these problems, in particular the problem of correlation caused by measurements at the same individual. We consider nonlinear least squares estimation in random effect regression models which does not require a specification of the underlying distributions. For this estimate we introduce a new methodology which can be, say, used to derive efficient or optimal designs in very general situations. More precisely, we embed the discrete optimal design problem in a continuous optimal design problem, where a nonlinear functional of the design density has to be minimized or maximized. This approach allows us to address the problem of correlation and yields to an asymptotic optimal design density which has to be determined numerically in all cases of practical interest. We propose an efficient algorithm based on approximations by polynomials. For a fixed sample size, say n, for each individual an exact design can finally be obtained from the quantiles of the corresponding optimal distribution function. It will be demonstrated in concrete examples that these designs are extremely efficient. Moreover, if the experimenter really wants to find the optimal exact design, the designs derived from the asymptotic optimal design density are very good starting designs for any numerical procedure. To our knowledge this is the first systematic approach to determine optimal designs for linear and nonlinear mixed effect models with correlated errors.

The remaining part of this paper is organized as follows. In Section 2 we consider the case of a linear random effect model and explain the basic concepts in this context. In Section 3 we also introduce the idea of asymptotic optimal designs for linear regression models with correlated observations, which was first considered by Bickel and Herzberg (1979). Some examples for linear and quadratic regression models are presented in Section 4. Finally, Section 5 deals with the case of nonlinear random effect models. In particular, we consider a compartmental model with correlated random errors and derive D-optimal designs and optimal designs for estimating the area under the curve. Finally, the Uzara® example is re-analyzed and optimal designs for model (1.1) are determined.

2 Preliminaries

In this Section we consider the common random-effect linear regression model

$$Y_{ij} = b_i^T f(t_{ij}) + \varepsilon_{ij} \ , \ i = 1, \dots, K; \ j = 1, \dots n_i \ ; \tag{2.1}$$

where Y_{ij} denotes the jth observation of the ith subject at the experimental condition t_{ij}, $\varepsilon_{11}, \dots, \varepsilon_{K,n_K}$ are centered random variables, $f(t) = (f_1(t), \dots, f_p(t))^T$ is a given vector of linearly independent regression functions, b_i is a p-dimensional random vector representing the individual parameters of the ith subject, $i = 1, \dots, K$. The explanatory variables variables t_{ij} can be cho-

sen by the experimenter from a compact interval, say T, which will be specified in the concrete examples. We assume that errors $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{i,n_i})$ for different subjects are independent but the errors for the same subject are correlated, that is

$$\mathrm{Cov}(\varepsilon_{ij}, \varepsilon_{is}) = \sigma^2(\gamma r(t_{ij} - t_{is}) + (1-\gamma)\delta_{j,s}), \tag{2.2}$$

where $\gamma \in [0,1]$ is a constant, $r(t)$ a given correlation function such that $r(0) = 1$ and $\delta_{j,s}$ denotes Kronecker's symbol. The corresponding covariance matrix is denoted by V_ε. Moreover, we also assume that random variables b_i representing the individual parameters b_i have mean β, covariance matrix V_p and are independent of the random variables ε_i. This means that the covariance between two observations at time t_{ij} and t_{is} $(j \neq s)$ is

$$\mathrm{Cov}(Y_{ij}, Y_{is}) = f^T(t_{ij}) V_p f(t_{is}) + \sigma^2 \gamma r(t_{ij} - t_{is}) \, ,$$

while the variance of Y_{ij} is given by $f^T(t_{ij}) V_p f(t_{ij}) + \sigma^2$. It was shown by Schmelter (2007b) that an optimal design necessarily advices the experimenter to investigate all subjects at the same experimental settings, i.e. $t_{ij} = t_j$ $(i = 1, \ldots, K,\ j = 1, \ldots, n)$. Consequently in the situation under consideration an exact design $\xi = \{t_1, \ldots, t_n\}$ is an n-dimensional vector which describes the experimental conditions for each subject. Without loss of generality we assume that the components are ordered, i.e. $t_1 < \ldots < t_n$.

Suppose that n observations are taken according to the design ξ. Then the model (2.1) for the ith subject can be written as

$$Y_i = X b_i + \varepsilon_i \ ; \ i = 1, \ldots, K, \tag{2.3}$$

where $Y_i = (Y_{i1}, \ldots, Y_{in})^T$ and the matrix X is given by $X = (f(t_1), \ldots, f(t_n))^T$. This model is a special case of the random-effect models discussed in Harville (1976), which are called generalized MANOVA . The ordinary least squares estimate of the parameter β (i.e. the mean of the random variables b_i) is given by

$$\hat{\beta}_{\mathrm{OLS}} = \frac{1}{K} \sum_{i=1}^{K} (X^T X)^{-1} X^T Y_i, \tag{2.4}$$

where each summand is the least squares estimate corresponding to the individual effect b_i. The covariance matrix of $\hat{\beta}_{\mathrm{OLS}}$ is given by

$$\begin{aligned} \mathbf{D}(\hat{\beta}_{\mathrm{OLS}}) &= \frac{1}{K} (X^T X)^{-1} X^T (V_\varepsilon + X V_p X^T) X (X^T X)^{-1} \\ &= \frac{1}{K} \left((X^T X)^{-1} X^T V_\varepsilon X (X^T X)^{-1} + V_p \right) . \end{aligned} \tag{2.5}$$

If the covariance matrix V_ε of the errors and the covariance matrix of the random effects V_p were known (or can be well estimated) the weighted least squares statistic

$$\hat{\beta}_{\mathrm{WLS}} = \frac{1}{K} \sum_{i=1}^{K} \left(X^T (V_\varepsilon + X V_p X^T)^{-1} X \right)^{-1} X^T (V_\varepsilon + X V_p X^T)^{-1} Y_i, \tag{2.6}$$

could alternatively be used to estimate the parameter β. The covariance matrix of the estimate $\hat{\beta}_{\mathrm{WLS}}$ is given by

$$\mathbf{D}(\hat{\beta}_{\mathrm{WLS}}) = \frac{1}{K}(X^T(V_\varepsilon + XV_pX^T)^{-1}X)^{-1}.$$

Optimal designs minimize appropriate functionals of the covariance matrix of the ordinary or weighted least squares estimate. For this reason we assume throughout this paper without loss of generality $K = 1$ for the design problem for the random effect model (2.3). We consider the problem of constructing optimal designs for ordinary least squares estimation, because on the one hand it is our experience that in applications it is usually difficult to specify the matrices V_p and V_ε with sufficient precision, such that the improvement in the quality of the estimates is significant. On the other hand, we will also investigate the efficiency of the constructed optimal designs for ordinary squares estimation if these are used for weighted least squares estimation (see Section 3 and 4). In particular, it is demonstrated that these designs are also rather efficient for weighted least squares.

3 Asymptotic optimal designs

Although the theory of optimal design has been discussed intensively for uncorrelated observations [see for example Fedorov (1972), Pázman (1986) and Atkinson and Donev (1992)] less results can be found for dependent observations. For linear and nonlinear random effect models several authors have investigated optimal design problems under the additional assumption of uncorrelated errors [see e.g. Schmelter (2007a), Schmelter (2007b), Mentré et al. (1997) or Retout and Mentré (2003) among others]. For fixed effect regression models with correlated errors numerous authors suggest to derive optimal designs by asymptotic considerations. Sacks and Ylvisaker (1966, 1968), considered a fixed design space, where the number of design points in this set tends to infinity. As a result of this assumption the asymptotic optimal designs depend only on the behavior of the correlation function in a neighborhood of the point 0. In the present paper we use an approach of Bickel and Herzberg (1979) and Bickel et al. (1981), who considered a design interval expanding proportionally to the number of observation points. This case is equivalent to the consideration of a fixed interval with correlation function depending on the sample size. To be precise, we assume that the design space is given by an interval, say T, and that the design points are are generated by a sequence of designs $\xi_n = \{t_{11}, \ldots, t_{nn}\}$, where

$$t_{jn} = a\left((j-1)/(n-1)\right),\ j = 1, \ldots, n; \tag{3.1}$$

and $a : [0,1] \to T$ denotes the inverse of a distribution function. Note that the function a is obtained from the density of the weak limit of the sequence ξ_n as $n \to \infty$. For example,

if $T = [-1, 1]$, the function $a(u) = (2u - 1)$ corresponds to the equally-spaced design with distribution function $a^{-1}(x) = \frac{x+1}{2}$ and density $(a^{-1})'(x) = \frac{1}{2} I_{[-1,1]}(x)$. Furthermore, we assume that the correlation function $r(t)$ of the errors ε_i in (2.2) depends on n in the form

$$r_n(t) = \rho(nt), \tag{3.2}$$

such that $\rho(t) = o(t)$ as $t \to \infty$. For the numerical construction of asymptotic optimal designs we derive in the following Lemma an asymptotic representation for the covariance matrix of the ordinary least squares estimate. For this purpose we make the following regularity assumptions.

(C1) The regression functions $f_1(t), \ldots, f_p(t)$ are linearly independent and bounded on the interval T and satisfy a first order Lipschitz condition,i.e.

$$|f_i(t) - f_i(s)| \leq M|t - s| \quad \text{and } |f_i(t)| \leq M \quad \text{for all } t, s \in T \ , i = 1, \ldots, p.$$

(C2) The function a is twice differentiable and there exists a positive constant $M < \infty$ such that for all $u \in (0, 1)$

$$\frac{1}{M} \leq a'(u) \leq M, \quad |a''(u)| \leq M. \tag{3.3}$$

(C3) The correlation function ρ is differentiable with bounded derivative and satisfies $\rho'(t) \leq 0$ for sufficiently large t.

The following result is obtained by similar arguments as given in Bickel and Herzberg (1979) and its proof therefore omitted.

Lemma. *Assume that conditions (C1), (C2) and (C3) are satisfied, then the covariance matrix of the ordinary least squares estimate defined in* (2.4) *satisfies*

$$\mathbf{D}(\hat{\beta}_{\mathrm{OLS}}) = \frac{\sigma^2}{n} \left(W^{-1}(a) + 2\gamma W^{-1}(a) R(a) W^{-1}(a) \right) + V_p + o(\frac{1}{n}), \tag{3.4}$$

and the matrices W and R are defined by

$$\begin{aligned} W(a) &= \left(\int_0^1 f_i(a(u)) f_j(a(u))\, du \right)_{i,j=1}^{p}, \\ R(a) &= \left(\int_0^1 f_i(a(u)) f_j(a(u)) Q(a'(u))\, du \right)_{i,j=1}^{p}, \end{aligned}$$

and the function Q is given by

$$Q(t) = \sum_{j=1}^{\infty} \rho(jt).$$

Note that only the first term in (2.5) (and (3.4)) depends on the underlying (asymptotic) design, and this term will be the basis for the construction of optimal designs in the following discussion. If the function a is the inverse of a continuous distribution with density, say φ, then $a'(t) = 1/\varphi(t)$ and for large n the first term of the covariance matrix of the ordinary least squares estimate can be approximated by the matrix $V(\varphi)/n$, where the $p \times p$ matrix V is given by

$$V(\varphi) := \sigma^2 \left(W^{-1}(\varphi) + 2\gamma W^{-1}(\varphi) R(\varphi) W^{-1}(\varphi) \right). \tag{3.5}$$

Here the matrices W and R are defined by

$$\begin{aligned} W(\varphi) &= \left(\int_T f_i(t) f_j(t) \varphi(t)\, dt \right)_{i,j=1}^{p}, \\ R(\varphi) &= \left(\int_T f_i(t) f_j(t) Q(1/\varphi(t)) \varphi(t)\, dt \right)_{i,j=1}^{p}, \end{aligned}$$

respectively. An asymptotic optimal density density φ^* minimizes an appropriate functional of the matrix $V(\varphi)$. For this purpose numerous criteria have been proposed in the literature [see Silvey (1980), Atkinson and Donev (1992), Pukelsheim (1993)] and exemplarily we consider in the following section the D- and c-optimality criterion which minimize $\det V(\varphi)$ and $c^T V(\varphi) c$ for a given vector $c \in \mathbb{R}^p$, respectively. The application of our methodology for other optimality criteria will be obvious from these examples. The general procedure for constructing an efficient design minimizing a given functional of the covariance matrix of the ordinary least squares estimate is as follows:

(1) The correlation structure (i.e. the function ρ in (3.2)) has to be specified.

(2) The exact design problem is embedded in an asymptotic design problem and an appropriate functional of the matrix $V(\varphi)$ in (3.5) is minimized with respect to φ. This yields the asymptotic optimal design density φ^*.

(3) If φ^* denotes the density minimizing the functional specified in step (2) the exact designs for a fixed sample size n are derived from the quantiles of the corresponding distribution function, say Φ^*. This gives:

$$t_{i,n} = (\Phi^*)^{-1} \left(\frac{i-1}{n-1} \right) \; ; \quad i = 1, \ldots, n. \tag{3.6}$$

The optimal density φ^* in step (2) of this procedure is determined numerically as follows. We use a parametric representation of the density by a polynomial of the form

$$\varphi(t) = \frac{(p_0 + p_1 t + \ldots + p_r t^r)_+}{\int_T (p_0 + p_1 t + \ldots + p_r t^r)_+ dt}$$

and apply the Nelder-Mead algorithm to find the optimal density minimizing the specified functional of the matrix $V(\varphi)$ in this parametric class. We run the algorithm for different degrees of the polynomial and different initial values and chose as result the density corresponding to the minimal value of the optimality criterion. All integrals are calculated by the Simpson quadrature formula. Usually no substantial improvements are obtained for polynomials of degree larger than $r = 6$. We also investigated the performance of similar algorithms if the density is represented in terms of rational or exponential functions. The results were very similar and on the basis of our numerical experiments we conclude that the optimal density φ^* can be very well approximated of polynomials of degree 6.

The derived designs from asymptotic theory can, on the one hand, be used to construct efficient designs for a given sample size as specified in step (3) of the algorithm. On the other hand, these designs can also be used to determine exact optimal designs by using them as initial values in a further (discrete) optimization procedure. More precisely for the determination of an exact optimal design we propose to add a fourth step in the above algorithm:

(4) The Nelder-Mead algorithm is used for the determination of an exact optimal design [minimizing a functional of the covariance matrix in (2.5)], where the n-point design obtained in step (3) is used as an initial design.

In the following section we will illustrate this procedure in the case of a linear and quadratic regression model. In Section 5 we extend the methodology to nonlinear models and investigate its performance in a compartmental model. In particular, it is demonstrated that the designs derived from the asymptotic approach are very efficient compared to exact optimal designs.

4 Some numerical results for linear models

In this section we consider the linear and quadratic regression model. Efficient and optimal exact designs are derived by the algorithm proposed in Section 3, where the D-optimality criterion is used to compare competing designs.

4.1 Linear regression

We begin our investigations with the classical linear regression model, where $p = 2$, $f_1(t) = 1$, $f_2(t) = t$ and $T = [-1, 1]$. The correlation function in (3.2) is given by $\rho(t) = e^{-\lambda t}$, i.e.

$$r_n(t) = e^{-\lambda n t}, \lambda > 0. \tag{4.1}$$

The asymptotic D-optimal design densities have been calculated by the procedure described in Section 2 and are given in Figure 1 for different choices of the parameters λ and γ. Note

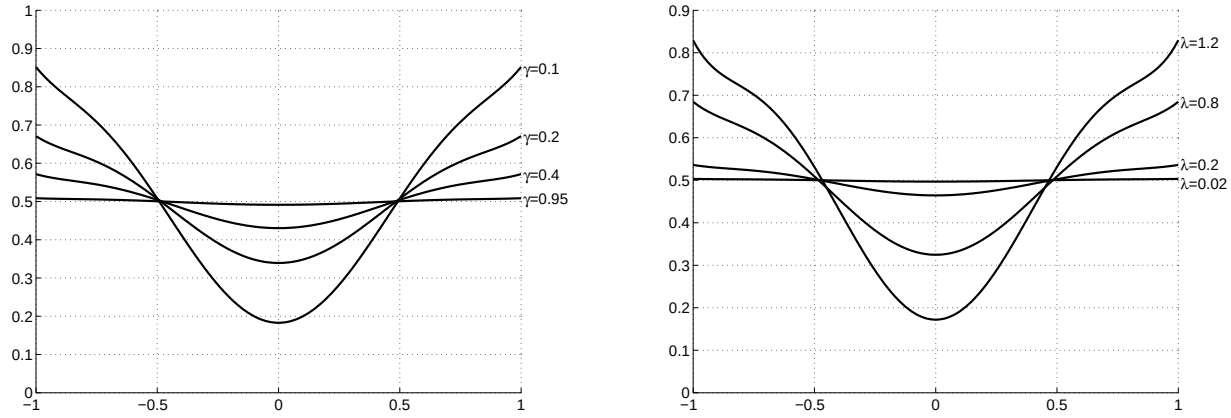


Figure 1: *Asymptotic D-optimal design densities for ordinary least squares estimation in a random effect linear regression model for different choices of parameters in the covariance function (2.2) with $r(t)$ defined by (4.1). Right part: $\gamma = 0.6$; Left Part: $\lambda = 0.2$*

that the numerically calculated optimal design densities are symmetric, but we were not able to prove the symmetry of the asymptotic optimal density. It can be seen from Figure 1 that the optimal density approximates the density of the uniform distribution if $\gamma \to 1$ or $\lambda \to 0$. Note that the case $\lambda \approx 0$ corresponds to a more slowly decreasing correlation between the errors. If $\gamma \to 0$ or $\lambda \to \infty$ the situation of uncorrelated errors is approximated and it can be observed from Figure 1 that in this case the optimal design density puts more mass at the boundary of the design space. Note that for uncorrelated observations the D-optimal design is concentrated at the boundary of the design space [see Hohmann and Jung (1975)].

In the following we will investigate the efficiency of an exact design derived from the asymptotic theory for ordinary and weighted least squares estimation, where the parameters in the correlation function (2.2) are given by

$$\gamma = 0.6, \sigma^2 = 0.5,\ V_p = \mathrm{diag}(\sigma^2_{\beta_1}, \sigma^2_{\beta_2}) = \mathrm{diag}(0.3^2, 0.3^2). \tag{4.2}$$

For this purpose let ξ_n^u be an n-point equidistant design and ξ_n^a be an n-point design obtained by the transformation (3.6) from the asymptotic optimal density. The points of the design ξ_n^a are displayed in the left part of Figure 2 for $\lambda = 1.2$, which also shows the exact optimal designs for ordinary and weighted least squares estimation. These designs are calculated as described in step (4) of our procedure using a discrete optimization routine with the design ξ_n^a as starting design. In the right part of Figure 2 we present the efficiencies

$$\begin{aligned}
\mathrm{eff}_{\mathrm{OLS}}(\xi) &= \left(\frac{\det[(X^TX)^{-1}X^TV_\varepsilon X(X^TX)^{-1} + V_p]}{\det[(X_{\mathrm{OLS}}^TX_{\mathrm{OLS}})^{-1}X_{\mathrm{OLS}}^TV_\varepsilon X_{\mathrm{OLS}}(X_{\mathrm{OLS}}^TX_{\mathrm{OLS}})^{-1} + V_p]} \right)^{1/p} \\
\mathrm{eff}_{\mathrm{WLS}}(\xi) &= \left(\frac{\det[X^T(V_\varepsilon + XV_pX^T)^{-1}X]}{\det[X_{\mathrm{WLS}}^T(V_\varepsilon + X_{\mathrm{WLS}}V_pX_{\mathrm{WLS}}^T)^{-1}X_{\mathrm{WLS}}]} \right)^{1/p}
\end{aligned}$$

of the design ξ_n^a derived from asymptotic theory and the uniform design ξ_n^u if they are used for ordinary or weighted least squares estimation. Here X denotes the design matrix obtained from the design ξ under consideration, while X_{OLS} and X_{WLS} correspond to the optimal exact design for ordinary and weighted least squares estimation. We observe that - although the design points of the optimal designs may be different - the D-efficiency of design ξ_n^a for ordinary or weighted least squares estimation is very large. Note also that the D-efficiency of the uniform design ξ_n^u is also large since the asymptotic optimal density is close to uniform density for the given choice of parameters.

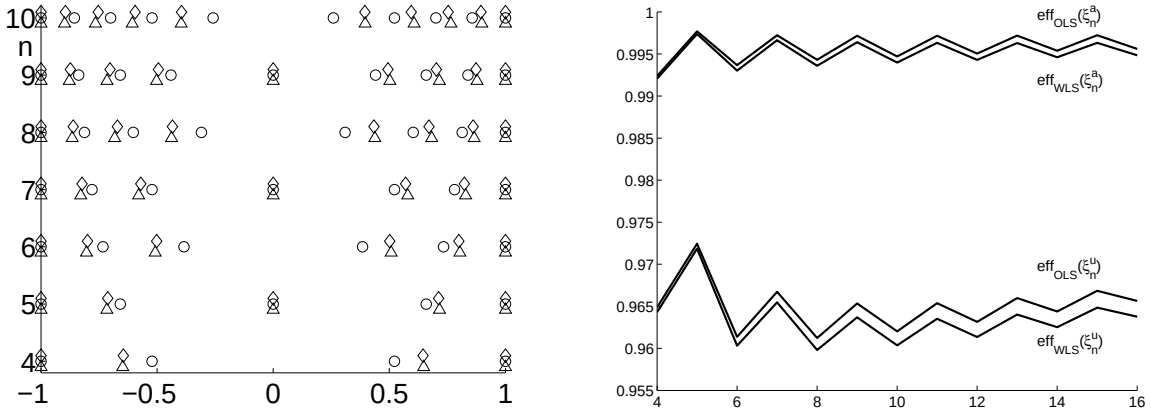


Figure 2: *Left part: Various designs for ordinary and weighted least squares estimation in the random effect linear regression model. Exact D-optimal designs derived from asymptotic theory: ball; exact D-optimal for ordinary least squares estimation: diamond; exact D-optimal designs for weighted least squares estimation: triangle. Right part: Efficiency of the designs ξ_n^a and the equidistant design ξ_n^u for ordinary and weighted least squares estimation. The parameters are given by (4.2), where $\lambda = 1.2$.*

4.2 Quadratic model

As second example of a random effect linear model we consider the quadratic regression model, that is $p = 3$, $f_1(t) = 1$, $f_2(t) = t$, $f_2(t) = t^2$ and assume again that the design space is given by the interval $T = [-1, 1]$. The asymptotic D-optimal densities for different choices of the parameters λ and γ and the parameters specified in (4.2) are shown in Figure 3. We observe again that the D-optimal density converges to the density of the uniform design, if $\gamma \to 1$ or $\lambda \to 0$. On the other hand, if $\gamma \to 0$ or $\lambda \to \infty$ it can be seen that the asymptotic D-optimal design density is more concentrated at the points -1, 0 and 1, which are the points of the exact D-optimal design for a quadratic fixed effect model with uncorrelated observations [see Gaffke and Krafft (1982)]. This corresponds to intuition because in this case the errors are less

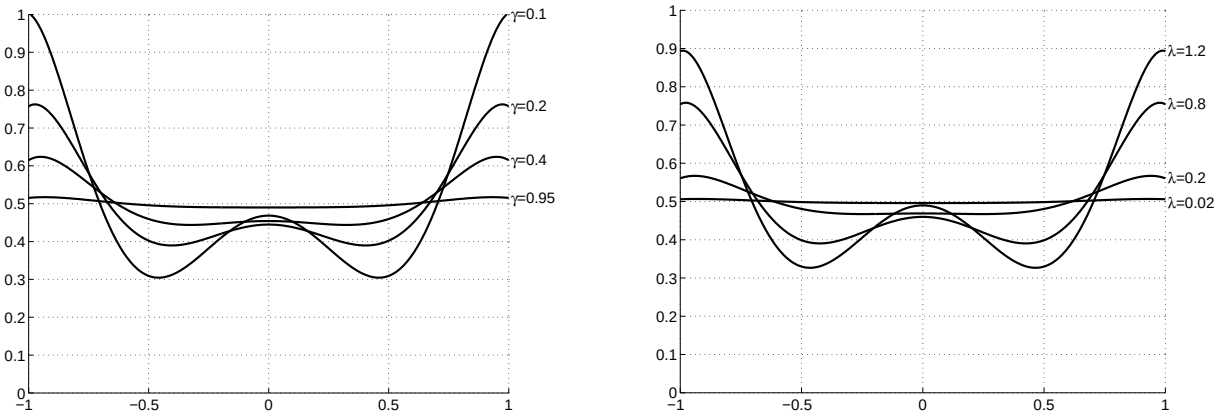


Figure 3: *Asymptotic D-optimal design densities for ordinary least squares estimation in a random effect quadratic model for different choices of parameters in the covariance function (2.2) with $r(t)$ defined by (4.1). Right panel $\gamma = 0.6$; Left panel $\lambda = 0.2$.*

correlated. The points of the design ξ_n^a derived from the asymptotic D-optimal design density and the exact D-optimal designs for ordinary and weighted least squares estimation are depicted in the left part of Figure 4. The right part of this Figure shows the D-efficiencies for ordinary and weighted least squares estimation in the quadratic regression model, where the parameters are given by (4.2) with $\lambda = 1.2$. We observe that the design points concentrate in two regions located at the points of the exact D-optimal design for a quadratic regression with uncorrelated errors. It is also noteworthy that the designs derived from the asymptotic theory are very efficient for ordinary and weighted least squares estimation.

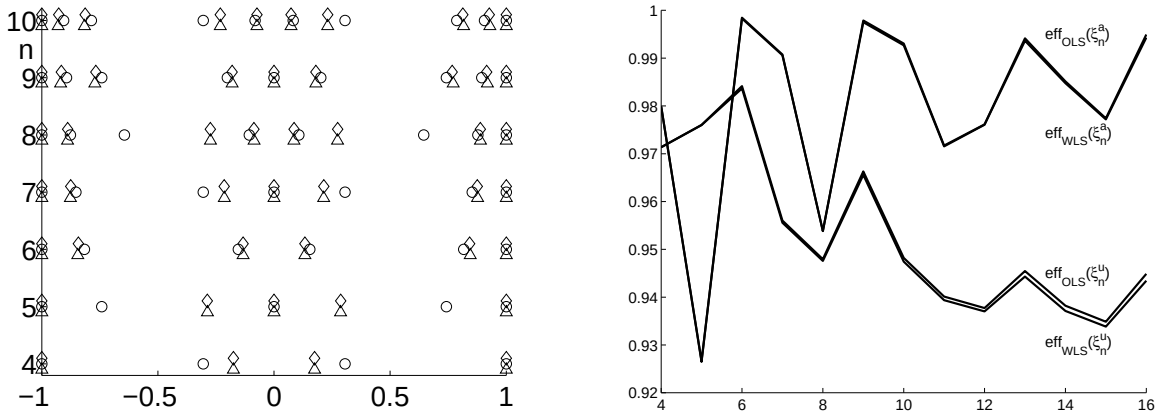


Figure 4: *Left part: Various designs for ordinary and weighted least squares estimation in the random effect quadratic regression model. Exact D-optimal designs derived from asymptotic theory: ball; exact D-optimal for ordinary least squares estimation: diamond; exact D-optimal designs for weighted least squares estimation: triangle. Right part: Efficiency of the designs* ξ_n^a *and equidistant design* ξ_n^u *for ordinary and weighted least squares estimation. The parameters are given by (4.2), where* $\lambda = 1.2$.

5 Nonlinear random effect models

In this section we illustrate how the theory can be extended to nonlinear random effect models, which have found considerable interest in the literature on pharmacokinetics. In this case the model under investigation is given by

$$Y_{ij} = \eta(t_j, b_i) + \varepsilon_{ij} \ , \ i = 1, \ldots, K; \ j = 1, \ldots n. \tag{5.1}$$

Since the model (5.1) is nonlinear with respect to the variables b_i there exists no analytical expression for the likelihood function and various approximations have been considered in the literature [see Mentré et al. (1997), Retout et al. (2002), Retout and Mentré (2003) among others]. These approximations are used for the calculation of the maximum likelihood estimate and a corresponding Fisher information. Alternatively, an estimate of the population mean β could easily be obtained as average of the nonlinear least squares estimates $\hat{b}_i$ for the different individuals, but due to the nonlinearity of the model an explicit representation of the corresponding covariance matrix cannot be derived. Following Retout and Mentré (2003) we propose to use a first-order Taylor expansion to derive an approximation of this covariance matrix. To be precise we use (under suitable assumptions of differentiability of the regression function) the approximation

$$\eta(t, b) \approx \eta(t, \beta) + f(t, \beta)(b - \beta)^T, \tag{5.2}$$

where

$$f(t, b) = \frac{\partial \eta(t, b)}{\partial b}$$

denotes the gradient of the regression function with respect to b. This means that the nonlinear model (5.1) is approximated by a the linear model (5.2), where for the construction of the design we assume that knowledge about the parameter β is available from previous or similar experiments. This corresponds to the concept of locally optimal designs as introduced by Chernoff (1953) in the context of fixed effect nonlinear regression models. Usually locally optimal designs serve as benchmarks for commonly used designs and are the basis for the construction of optimal designs with respect to more sophisticated optimality criteria using a Bayesian or Minimax approach (see Chaloner and Verdinelli (1995) or Dette (1995)). As a consequence, the covariance matrix of the nonlinear least squares estimate in the model (5.1) is approximated by replacing the matrix X in model (2.1) with $f(t) = f(t, b)|_{b=\beta}$, and the methodology described in Section 2 and 3 can be applied to determine efficient designs for ordinary and weighted nonlinear least squares estimation. In the following we illustrate this concept in several examples and derive D- and c-optimal designs for inference in a random effect model.

5.1 D-optimal design for a random effect compartmental model

We consider a random effect compartmental model with first-order absorption, i.e.

$$\eta(t,b) = \frac{b_1}{b_1 - b_2}(e^{-b_2 t} - e^{-b_1 t}). \tag{5.3}$$

The model (5.3) is a special case of the Bateman function mentioned in the introduction [see Garrett (1994)] and has found considerable attention in chemical sciences, toxicology or pharmacokinetics [see for example Gibaldi and Perrier (1982)]. The optimal design problem in the compartmental fixed effect model has also been studied by numerous authors [see for example Box and Lucas (1959), Atkinson et al. (1993), Dette and O'Brien (1999), Biedermann et al. (2004) among others], but much less results are available under the assumption of random effects. Recently optimal approximate designs for the random-effect compartmental model (5.3) have been determined by Atkinson (2008), but we did not find references dealing with exact designs for models with correlated errors. We will now derive such designs from the asymptotic optimal design density and compare these with the exact optimal designs for ordinary and weighted least squares estimation.
Note that the gradient of the function η with respect to b is given by

$$f(t,b) = \left(\frac{b_2(e^{-b_1 x} - e^{-b_2 x}) + ({b_1}^2 x - b_1 b_2 x)e^{-b_1 x}}{(b_1 - b_2)^2}, \frac{b_1(e^{-b_1 x} - e^{-b_2 x}) + ({b_1}^2 x - b_1 b_2 x)e^{-b_2 x}}{(b_1 - b_2)^2} \right)^T . \tag{5.4}$$

In order to illustrate the methodology we assume that the parameters of the population distribution and the error distribution are given by $\beta^{(0)} = (1, 0.5)^T$,

$$\gamma = 0.6, \sigma^2 = 0.01,\ V_p = \mathrm{diag}(\sigma^2_{\beta_1}, \sigma^2_{\beta_2}) = \mathrm{diag}(0.1^2, 0.05^2). \tag{5.5}$$

[see Atkinson (2008)] and that the design space is given by the interval $T = [0, 10]$. We assume again that the function $r(t)$ in (2.2) is given by (4.1). The asymptotic D-optimal design densities for different choices of the parameters are shown in Figure 5. We observe again that for $\gamma \to 1$ or $\lambda \to 0$ the D-optimal design densities approximate the uniform design, while for larger values of λ or smaller values of γ the asymptotic D-optimal designs put more weight at two specific regions of the design space.

Again this corresponds to intuition, because the (approximate) D-optimal design for the model (5.3) with uncorrelated observations is a two-point design [see e.g. Box and Lucas (1959)]. In the following discussion we will investigate the performance of the uniform and an exact design derived from asymptotic theory. For this purpose we define ξ_n^u as an n-point equidistant design $\{10/n, 20/n, \ldots, 10\}$ and ξ_n^a as the n-point design obtained by the transformation

$$t_j = (\Phi^*)^{-1}\left(j/(n+1)\right),\ j = 1, \ldots, n,$$

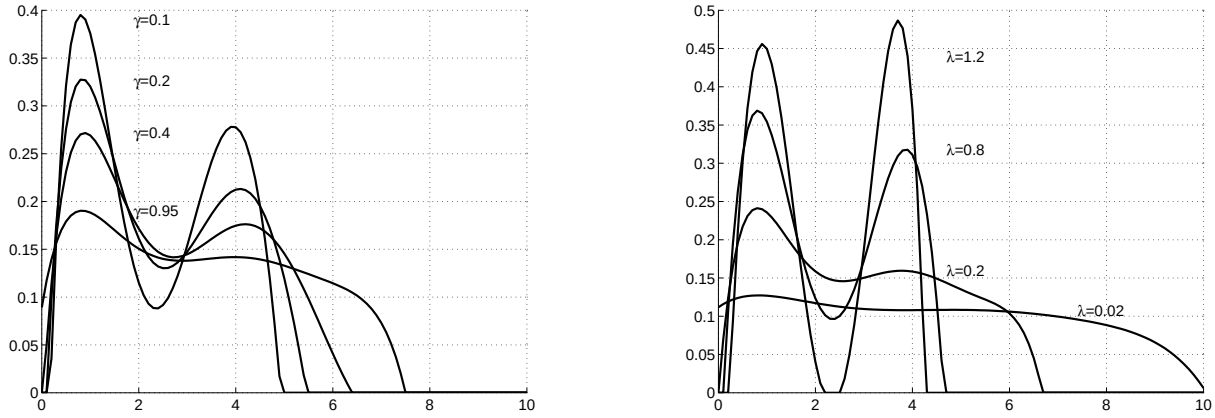


Figure 5: *Asymptotic D-optimal design densities for nonlinear least squares estimation in the compartmental model* (5.3) *for different choices of the parameters in the covariance function (2.2) with $r(t)$ defined by (4.1). Left part: $\lambda = 0.2$; right part: $\gamma = 0.6$.*

where Φ^* denotes the distribution function corresponding to the asymptotic D-optimal design density. Note this transformation is slightly different from the transformation (3.1) in order to exclude the point 0 from the design points. Obviously, it is not reasonable to take observations $t = 0$ in model (5.3), because in the model (5.3) it is assumed that the drug is administered at time $t = 0$. The corresponding points of the exact designs are depicted in the left part of Figure 6 while the right part of the figure shows the D-efficiencies of the different designs. We observe that the designs derived from the asymptotic theory have a substantially larger D-efficiency compared to the uniform design. For example, if $n = 6$ the D-efficiency of the uniform design is approximately 50% for ordinary and weighted nonlinear least squares estimation, while the D-efficiency of the design ξ_n^a is close to 90%.

It is worthwhile to mention that in nonlinear random effect models the optimal designs depend additionally on the mean β of the distribution of the population parameters b_i. Therefore it is also of interest to investigate the sensitivity of the designs with respect to a misspecification of this parameter. For a study of the impact of such a misspecification on the efficiency of the resulting designs we consider the case $n = 4$ and $n = 6$ and the corresponding designs $\xi_4^a = \{1.04, 2.01, 3.16, 4.33\}$ and $\xi_6^a = \{0.83, 1.47, 2.32, 3.30, 4.20, 5.20\}$, respectively. In Figure 7 we display the efficiencies

$$\left(\frac{\det[(X^TX)^{-1}X^TV_\varepsilon X(X^TX)^{-1} + V_p]}{\det[(X_{\mathrm{OLS},\beta}^TX_{\mathrm{OLS},\beta})^{-1}X_{\mathrm{OLS},\beta}^TV_\varepsilon X_{\mathrm{OLS},\beta}(X_{\mathrm{OLS},\beta}^TX_{\mathrm{OLS},\beta})^{-1} + V_p]} \right)^{-1/p} \tag{5.6}$$

for different values of β. Here X denotes the design matrix obtained from the design ξ_n^a under the assumption that $\beta^{(0)} = (1, 0, 5)^T$, while the matrix $X_{\mathrm{OLS},\beta}$ corresponds to the exact D-optimal design for ordinary nonlinear least squares estimation for a specific β. The efficiencies

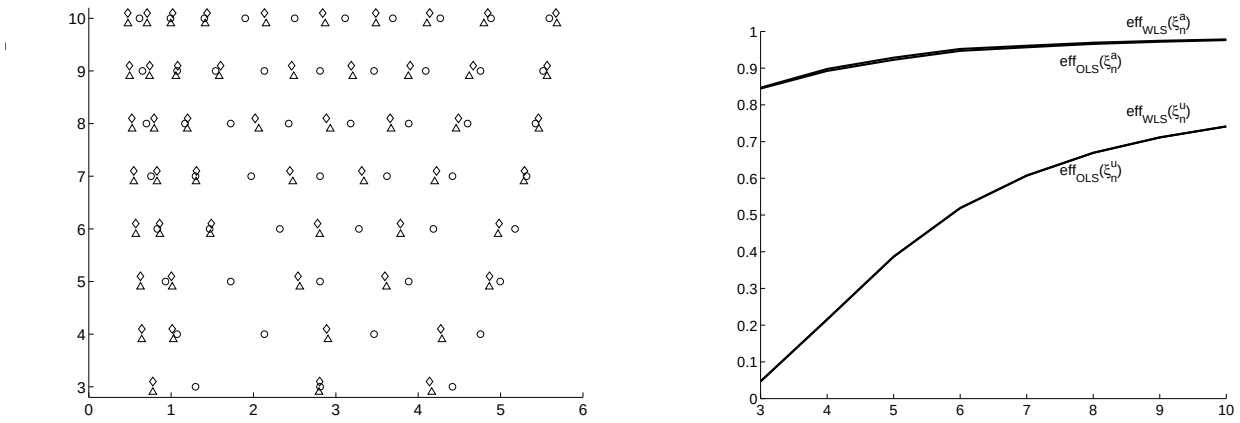


Figure 6: *Left part: Various designs for ordinary and weighted nonlinear least squares estimation in the compartmental model* (5.3)*. Exact D-optimal designs derived from asymptotic theory: ball; exact D-optimal for ordinary least squares estimation: diamond; exact D-optimal designs for weighted least squares estimation: triangle. Right part: Efficiency of the designs* ξ_n^a *and* ξ_n^u *for ordinary and weighted least squares estimation. The parameters are given by (5.5), where* $\lambda = 0.2$*.*

are plotted in Figure 7 for the rectangle $[\beta_i^{(0)} - 3\sigma_{\beta_i^{(0)}}, \beta_i^{(0)} + 3\sigma_{\beta_i^{(0)}}] = [0.7, 1.3] \times [0.35, 0.65]$. It can be seen that the exact optimal designs derived from the asymptotic theory yield also very good D-efficiencies over a broad range of the population mean.

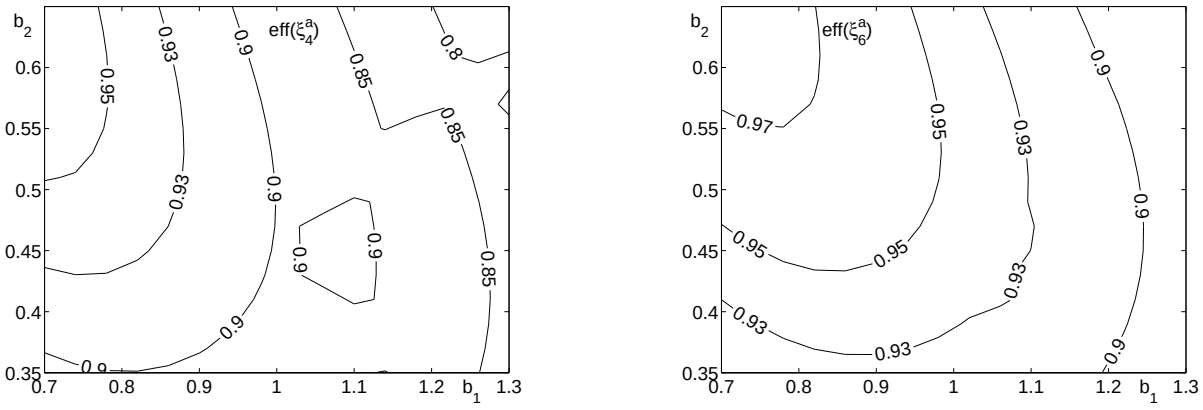


Figure 7: *D-efficiencies of the designs* $\xi_4^a = \{1.04, 2.01, 3.16, 4.33\}$ *(left part) and the design* $\xi_6^a = \{0.83, 1.47, 2.32, 3.30, 4.20, 5.20\}$ *(right part) (these designs are obtained from asymptotic density) for ordinary nonlinear least squares estimation in the random-effect compartmental model* (5.3)*, if the mean of the population distribution has been misspecified. The parameters of the population distribution are given by* (5.5) *with* $\lambda = 0.2$*.*

5.2 Optimal designs for estimating the AUC

In bioavailability studies the experimenter is often interested in the estimation of the area under curve (AUC), which is for the compartmental model (5.3) defined by

$$\text{AUC} = \int_0^\infty \eta(t, \beta)\, dt = \frac{1}{b_2}.$$

Thus the locally AUC-optimal design minimizes the variance of the nonlinear least squares estimate for the parameter β_2, which can be approximated by

$$(0,1)\left((X^TX)^{-1}X^TV_\varepsilon X(X^TX)^{-1} + V_p\right)(0,1)^T.$$

This correspond to a c-optimality criterion, which has been discussed extensively in the literature for fixed effect models with uncorrelated errors [see for example Ford et al. (1992), Fan and Chaloner (2003) or Dette et al. (2008) among others]. The asymptotic optimal design densities for estimating the area under the curve are shown in Figure 8. We observe again that the uniform design density is approximated if $\lambda \to 0$ or $\gamma \to 1$. On the other hand, if λ is large or $\gamma \to 0$, the AUC-optimal design density is more concentrated, which reflects the fact that the optimal design for estimating the area under the curve in the fixed effect compartmental model with uncorrelated observations is a one-point design. In Figure 9 we show the designs derived from the asymptotic optimal design densities and the exact optimal designs for estimating the area under the curve in the compartmental model. We observe that the designs derived from the asymptotic optimal design density are very close to the exact optimal designs for ordinary least squares estimation of the area under the curve. Moreover, the design ξ_n^a yields a substantial improvement in efficiency compared to the uniform design [see the right part of Figure 9].

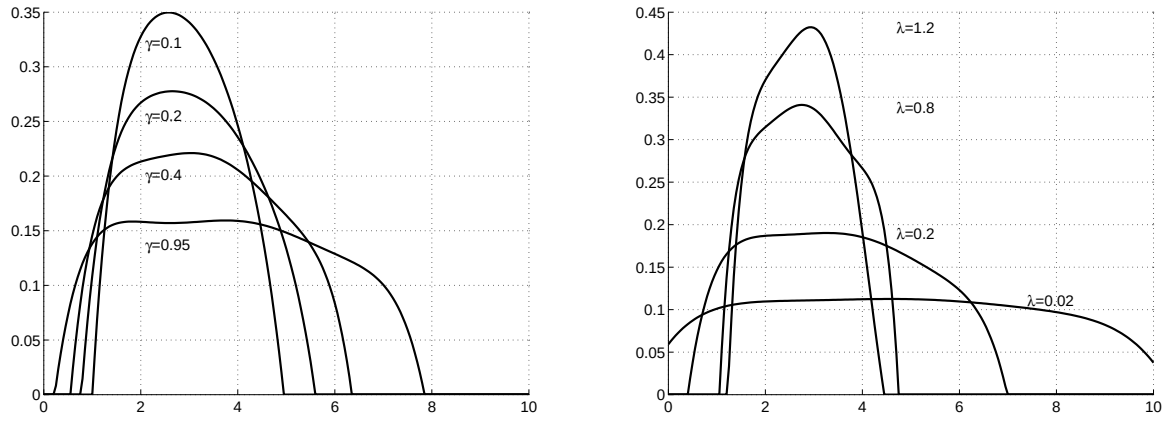


Figure 8: *Asymptotic optimal densities for estimating the area under the curve in the compartmental model* (5.3) *for different choices of the parameters in the covariance function (2.2) with* $r(t)$ *defined by (4.1). Left part:* $\lambda = 0.2$*, right part:* $\gamma = 0.6$*.*

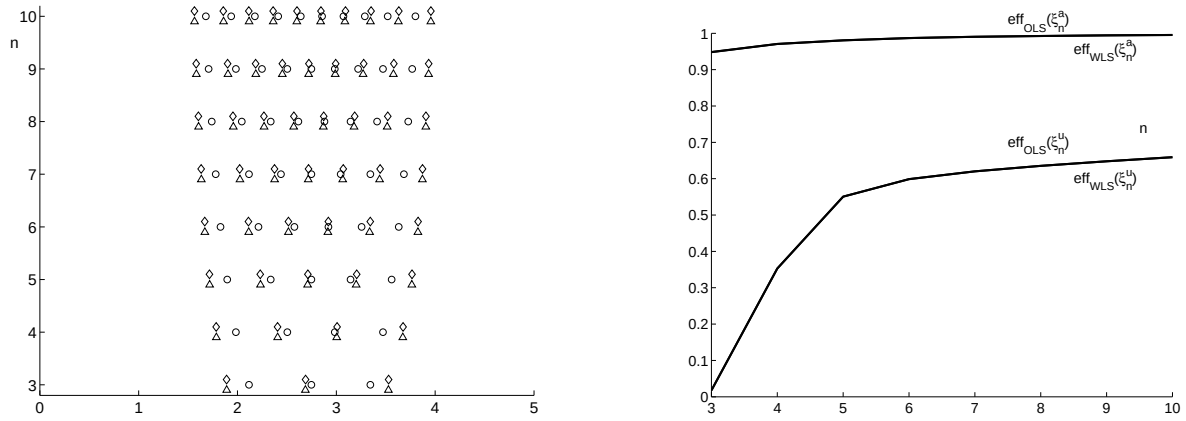


Figure 9: *Left part: Various designs for estimating the area under the curve in the compartmental model* (5.3). *Designs derived from asymptotic theory: ball; exact optimal designs for least squares estimation of the area under the curve: diamond; exact optimal designs for weighted least squares estimation: triangle. Right part: Efficiency of the designs* ξ_n^a *and* ξ_n^u *for ordinary and weighted least squares estimation. The parameters are given by* (5.5), *where* $\lambda = 1.2$.

5.3 Optimal designs for estimating the AUC in the Uzara example

We finally consider the optimal design problem for estimating the AUC in the example presented in the introduction. Original measurements were taken at non-optimized time points

$$0,\ 0.5,\ 1, 1.5,\ 2,\ 3,\ 4,\ 5,\ 6,\ 8,\ 10,\ 12,\ 15,\ 24,\ 36.$$

For the parameter estimates given in the introduction we have derived the asymptotic optimal design density, which is depicted in Figure 10 for the design interval $T = [0, 36]$. The resulting design from this density is given by

$$2.09,\ 4.55,\ 7.49,\ 10.8,\ 13.9,\ 16.8,\ 19.2,\ 21.5,\ 23.6,\ 25.7,\ 27.7,\ 29.8,\ 31.9,\ 34.0,\ 36$$

We observe that compared to original design the optimal design derived from the asymptotic theory is closer to an equidistant allocation of the observations.

Acknowledgements. The support of the Deutsche Forschungsgemeinschaft (SFB 823, "Statistik nichtlinearer dynamischer Prozesse ") is gratefully acknowledged. The work of H. Dette was also supported in part by an NIH grant award IR01GM072876:01A1 and by the BMBF project SKAVOE. The authors would like to thank Prof. Petra Thrmann, who provided the Uzara dataset and Martina Stein, who typed parts of the paper with considerable technical expertise.

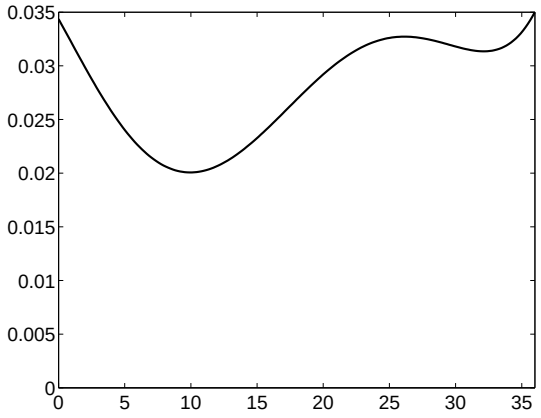


Figure 10: *Asymptotic optimal density for estimating the area under the curve in the 3-parameter Bateman model with covariance function (1.2).*

References

Aarons, L. (1999). Software for population pharmacokinetics and pharmacodynamics. *Clinical Pharmacokinetics*, 36:255–264.

Atkinson, A. C. (2008). Examples of the use of an equivalence theorem in constructing optimum experimental designs for random-effects nonlinear regression models. *Journal of Statistical Planning and Inference*, 138(9):2595–2606.

Atkinson, A. C., Chaloner, K., Herzberg, A. M., and Juritz, J. (1993). Optimum experimental designs for properties of a compartmental model. *Biometrics*, 49:325–337.

Atkinson, A. C. and Donev, A. (1992). *Optimum Experimental Designs.* Clarendon Press, Oxford.

Bickel, P. J. and Herzberg, A. M. (1979). Robustness of design against autocorrelation in time I: Asymptotic theory, optimality for location and linear regression. *Ann. Statist.*, 7(1):77–95.

Bickel, P. J., Herzberg, A. M., and Schilling, M. F. (1981). Robustness of design against autocorrelation in time II: Optimality, theoretical and numerical results for the first-order autoregressive process. *J. Amer. Statist. Assoc.*, 76(376):870–877.

Biedermann, S., Dette, H., and Pepelyshev, A. (2004). Maximin optimal designs for a compartmental model. *MODA 7, Advances in Model-oriented Design and Analysis*, pages 41–48.

Box, G. E. P. and Lucas, H. L. (1959). Design of experiments in non-linear situations. *Biometrika*, 46:77–90.

Buelga, D., de Gatta M., D. M. F., Herrera, E., Dominguez-Gil, A., and Garcia, M. (2005). The Bateman function revisited: a critical reevaluation of the quantitative expressions to characterize

concentrations in the one compartment body model as a function of time with first-order invasion and first-order elimination. *Antimicrobial agents and chemotherapy*, 49.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science.*, 10.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Statist.*, 24:586–602.

Colombo, S., Buclin, T., Cavassini, M., Decosterd, L., Telenti, A., Biollaz, J., and Csajka, C. (2006). Population pharmacokinetics of atazanavir in patients with human immunodeficiency virus infection. *Antimicrobial agents and chemotherapy*, 50.

Dette, H. (1995). Designing of experiments with respect to "standardized" optimality criteria. *Journal of the Royal Statistical Society, Ser. B*, 59:97–110.

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 O'Brien, T. (1999). Optimality criteria for regression models based on predicted variance. *Biometrika*, 86:93–106.

Fan, S. K. and Chaloner, K. (2003). A geometric method for singular c-optimal designs. *J. Statist. Plann. Inference*, 113:249–257.

Fedorov, V. V. (1972). *Theory of Optimal Experiments.* Academic Press, New York.

Ford, I., Torsney, B., and Wu, C. F. J. (1992). The use of canonical form in the construction of locally optimum designs for nonlinear problems. *Journal of the Royal Statistical Society, Ser. B*, 54:569–583.

Gaffke, N. and Krafft, O. (1982). Exact D-optimum designs for quadratic regression. *J. Roy. Statist. Soc. Ser. B*, 44(3):394–397.

Garrett, E. R. (1994). The Bateman function revisited: a critical reevaluation of the quantitative expressions to characterize concentrations in the one compartment body model as a function of time with first-order invasion and first-order elimination. *Journal of Pharmacokinetics and Biopharmaceutics*, 22:103–128.

Gibaldi, M. and Perrier, D. (1982). *Parmacokinetics. Second edition.* Dekker, New York.

Harville, D. (1976). Extension of the Gauss-Markov theorem to include the estimation of random effects. *Ann. Statist.*, 4(2):384–395.

Hohmann, G. and Jung, W. (1975). On sequential and nonsequential D-optimal experimental design. *Biometrische Zeitschrift*, 17(5):329–336.

Mentré, F., Mallet, A., and Baccar, D. (1997). Optimal design in random-effects regression models. *Biometrika*, 84(2):429–442.

Pázman, A. (1986). *Foundations of Optimum Experimental Design*. D. Reidel Publishing Company, Dordrecht.

Pukelsheim, F. (1993). *Optimal Design of Experiments*. John Wiley & Sons, New York.

Retout, S., Duffull, S., and Mentre, F. (2001). Development and implementation of the Fisher information matrix for the evaluation of population pharmakokinetic designs. *Computer Methods and Programs in Biomedicine*, 65:141–151.

Retout, S. and Mentré, F. (2003). Further developments of the Fisher information matrix in nonlinear mixed-effects models with evaluation in population pharmacokinetics. *Journal of Biopharmaceutical Statistics*, 13:209–227.

Retout, S., Mentré, F., and Bruno, R. (2002). Fisher information matrix for nonlinear mixed-effects models: evaluation and application for optimal design of enoxaparin population pharmacokinetics. *Statistics in Medicine*, 21:2623–2639.

Sacks, J. and Ylvisaker, N. D. (1966). Designs for regression problems with correlated errors. *Ann. Math. Statist.*, 37:66–89.

Sacks, J. and Ylvisaker, N. D. (1968). Designs for regression problems with correlated errors; many parameters. *Ann. Math. Statist.*, 39:49–69.

Schmelter, T. (2007a). Considerations on group-wise identical designs for linear mixed models. *Journal of Statistical Planning and Inference*, 137:4003–4010.

Schmelter, T. (2007b). The optimality of single-group designs for certain mixed models. *Metrika*, 65(2):183–193.

Shargel, L. (1993). *Applied Biopharmaceutics and Pharmacokinetics*. Chapman and Hall, London.

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

Thürmann, P., Neff, A., and Fleisch, J. (2004). Interference of Uzara glycosides in assays of digitalis glycosides. *International Journal of Clinical Pharmacology and Therapeutics*, 42(5):281–284.