Optimal Designs for Multi-Exponential Models with Covariance Structure

BY

JIEWEI ZENG
Bachelor of Science, Fudan University, Shanghai, China, 2007
Master of Science, University of Illinois at Chicago, Chicago, Illinois, 2009

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Chicago, 2013

Chicago, Illinois

Defense Committee:

Samad Hedayat, Chair and Advisor
Min Yang
Jie Yang
Jing Wang
Sally Freels, School of Public Health

This thesis is dedicated to my parents, my parents in law,

my husband Yue Yu, and my daughter Sophia S. Yu

ACKNOWLEDGMENTS

I wish to express my grateful appreciation to my advisor Professor Samad Hedayat, whose extraordinary support and enlightening guidance have profoundly influenced my study, research, and future career.

I would like to give my most sincere thanks to Professor Min Yang for his tremendous help in my research work, which has been accomplished with his continuous, illuminating and patient instructions.

I would like to thank Professor Jing Wang, Professor Jie Yang and Professor Sally Freels for their great help in my course work and research work.

I would like to thank all my committee members, Professor Min Yang, Professor Jing Wang, Professor Jie Yang, and Professor Sally Freels, for their kindness and help.

I want to give my special thanks to Professor Bikas Sinha for his helpful guidance in my research work.

Last, but not the least, I want to thank my husband, my parents, my parents in law, for their continuous understanding and support.

The research in this PhD dissertation is sponsored under Grant DMS-0904125 from the US National Science Foundation with Professor S. Hedayat as its PI.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Nonlinear models have been widely applied to explore the characteristics of data in biological, medical, environmental, and agricultural areas. Among them, the class of exponential regression models is of particular interest attributable to their capabilities of handling data in dynamic population, drugs in the bloodstream, temperature, and some biomedical problems. A majority of work has been accomplished to identify the optimal designs for nonlinear models. Recently Yang developed the strategy to determine the upper bounds of the cardinality of design support for nonlinear models. Unlike the traditional approximate design, where measurements are taken at a single design point, and the usual clinical design, where all the subjects start from the same time point, a new design setup which allows different starting points for each subject is proposed in this dissertation. Both autocorrelated covariance structure and intraclass correlation structure are adopted to address the correlations among the consecutive observations for a single subject. Multi-exponential models, in particular, one-compartment PK model and general bi-exponential model, are studied in this dissertation. Theoretical results for the design supports under equally spaced designs are provided for different models. It is found that the optimal design under one-compartment PK model is a saturated design, while the support size of the optimal design under the bi-exponential model with 4 parameters is no more than 5. Numerical results for unequally spaced designs are also obtained. Comparisons of design efficiencies are presented at the end of each model.

CHAPTER 1

INTRODUCTION

1.1 Introduction to Optimal Design

Experimentation is widely used in evaluating physical objects, chemical formulations, structures, components, and materials. Before any experiment is carried out, a well planned design is necessary to ensure the right type of data is collected and the desired precision of the results is achieved. For example, we may want to evaluate the efficacy of a certain drug in a clinical trial. An experimental design in this case typically includes what type of design should be adopted, how many subjects will be recruited, how to setup the inclusion or exclusion criteria in the screening process, which time points should samples be collected, etc. These are all within the domain of experimental designs.

Optimal design is a class of experimental design that is optimal with respect to some judiciously selected statistical criteria, a majority of which are constructed based on the Fisher information matrix. (We omit the word "Fisher" in the rest of this dissertation for simplicity.) It is well known that the information matrix is approximately proportional to the inverse of the variance covariance matrix of the parameter estimates. Roughly speaking, a design maximizing the information matrix will equivalently provide the parameter estimates with minimum variances. Different information matrix based optimality criteria have been proposed, such as A-, D-, E-, Ψ_p-, etc. The details can be found in Hedayat (1), Fedorov (2), and Silvey (3).

Linear models and nonlinear models are two major categories of classical models which have aroused great interest to statisticians. Due to the appealing and relatively simple structures of linear models in which parameters and explanatory variables are multiplicative, quite a few theorems have been developed to obtain the optimal designs for the linear models. These theorems and approaches serve as the fundamental tools in the optimal design area and will be discussed in detail.

1.1.1 Approximate Design Theory

The approximate design theory was proposed by Kiefer (4) (5) and was originally developed for the linear regression models.

Let $\mathfrak{B}$ be any Borel field of subsets of design space $\mathfrak{X}$ which includes $\mathfrak{X}$ and all sets which consist of a finite number of points, and let $\Xi = \{\xi\}$ be any class of probability measures on $\mathfrak{B}$ which includes all probability measures with finite support.

For the general linear model:

$$y = f(x)'\theta + \varepsilon,$$

where y is a $t \times 1$ vector of observations, x can be either a scalar or a vector which represents the explanatory variable(s) dependent on $\mathfrak{X}$. $f(x) = (f_1(x), \ldots, f_p(x))$, where $f_1(x), \ldots, f_p(x)$ are linearly independent real functions on $\mathfrak{X}$, and the range R of $f(x)$ is a compact set in the p-dimensional Euclidean space. θ is a $p \times 1$ vector of unknown parameters, and ε is a $t \times 1$ random error vector.

Then the information matrix is defined as

$$M(\xi) = \int_{\mathfrak{X}} f(x)'f(x)\xi(dx) = E_\xi f(x)'f(x).$$

Suppose we take a total of n observations, and the design space is $(x_1, x_2, \ldots, x_t)$. The exact design states that n_1 observations are allocated at x_1, n_2 observations at x_2, …, and n_t observations at x_t, such that $\sum_{i=1}^{t} n_i = n$. The information matrix is thus

$$M(\xi) = \sum_{i=1}^{t} n_i f'(x_i) f(x_i).$$

Alternatively, let $\omega_i = n_i/n$ be the proportion of observations assigned at x_i, the information matrix is thus

$$M(\xi) = n \sum_{i=1}^{t} \omega_i f'(x_i) f(x_i),$$

here n is a constant pre-determined before the design, thus $M(\xi)$ is only dependent on $\sum_{i=1}^{t} \omega_i f'(x_i) f(x_i)$.

To summarize, under the approximate design theory, we may consider the discrete design measure ξ as:

$$\xi = \begin{pmatrix} x_1 & x_2 & \cdots & x_t \\ \omega_1 & \omega_2 & \cdots & \omega_t \end{pmatrix}.$$

The optimal design problem is that given a total of n observations, how to choose $x_1, \ldots, x_t$ and $\omega_1, \ldots, \omega_t$ in order to maximize $M(\xi)$ under a certain criterion.

There are both advantages and disadvantages of the approximate design theory over the exact design theory. One major advantage is the simplicity in optimization over probabilities compared to that over integers. Also the approximate design theory is feasible for an arbitrary number of n. However, there are cases when the optimal designs are not discrete. Under such circumstances, we need to show that for any design ξ, there exists a discrete design ξ^+ such that $M(\xi) = M(\xi^+)$. Even when ξ is discrete, by converting the probabilities p_i to observations n_i, i.e, $np_i = n_i$, n_i is not necessarily an integer. In such cases, we may resort to a "nearby" design and do some approximations.

1.1.2 Design Criteria

Some of the optimality criteria are constructed based on the information matrix, while others are based on the variances of predictions. The major criteria are listed below:

(a) L-optimality

Definition 1. A design $\xi^* \in \Xi$ is L-optimal if and only if $M(\xi^*)$ is non-singular and

$$\min_{\xi \in \Xi} L(M(\xi)^{-1}) = L(M(\xi^*)^{-1}),$$

where L is a nonnegative linear function.

(b) A-optimality

A-optimality is a special form of L-optimality if we let

$$L(M(\xi)^{-1}) = \mathrm{tr}(M(\xi)^{-1}).$$

Definition 2. A design $\xi^* \in \Xi$ is A-optimal if and only if $M(\xi^*)$ is non-singular and

$$\min_{\xi \in \Xi} \mathrm{tr}(M(\xi)^{-1}) = \mathrm{tr}(M(\xi^*)^{-1}).$$

An A-optimal design minimizes the average variance of the parameter estimates.

(c) G-optimality

Definition 3. A design $\xi^* \in \Xi$ is G-optimal if and only if

$$\min_{\xi \in \Xi} \max_{x \in \mathfrak{X}} var_{\xi} \hat{EY_x} = \max_{x \in \mathfrak{X}} var_{\xi^*} \hat{EY_x},$$

where $\hat{EY_x}$ is the best linear unbiased estimator of EY_x.

(d) D-optimality

Definition 4. A design $\xi^* \in \Xi$ is D-optimal if and only if $M(\xi^*)$ is non-singular and

$$\min_{\xi \in \Xi} det(M(\xi)^{-1}) = det(M(\xi^*)^{-1}).$$

This criterion has the following implications:

(i) When normality is assumed, D-optimal design will minimize the volume of the confidence region for the parameters estimates, which is equivalent to minimizing the variance covariance matrix of the parameter estimates.

(ii) If we are interested in the full set of the parameters, D-optimality and G-optimality are equivalent.

(iii) D-optimality is scale invariant. If there exists a linear transformation between two sets of parameters, D-optimal design for one set of the parameters is also feasible for the other set.

(e) E-optimality

Definition 5. Let $\lambda_{\xi_1} \geq \lambda_{\xi_2} \geq \ldots \geq \lambda_{\xi_p}$ be the eigenvalues of $M(\xi)$, a design $\xi^* \in \Xi$ is E-optimal if and only if

$$\max_{\xi \in \Xi} \lambda_{\xi_p} = \lambda_{\xi_p^*}.$$

In other words, an E-optimal design maximizes the minimum eigenvalue of the information matrix. There are some statistical interpretations for this optimality criterion.

(i) In hypothesis testing of $\theta = 0$, where θ is a vector of parameters, under the normality assumption, we have the F statistics as follows:

$$F = \frac{(\hat{\theta})' M(\xi) \hat{\theta} / v}{\hat{\sigma}^2},$$

where $\hat{\sigma}^2$ is the estimate of the mean square error.

The power of the F test is thus $\frac{(\theta' M(\xi)\theta)/p}{\hat{\sigma}^2}$. We may want to choose a design $\xi^* \in \Xi$ which is able to maximize the power of the test. Since $\hat{\sigma}^2$ and p are known, maximization of the power is the same as maximizing $(\theta' M(\xi)\theta)$.

The value of $(\theta' M(\xi)\theta)$ varies with different sets of θ, subject to $\theta'\theta = 1$. One strategy of maximizing $(\theta' M(\xi)\theta)$ is to maximize its minimum attained with a certain set of θ. By matrix theory, we have

$$\min_{\theta : \theta'\theta = 1} (\theta' M(\xi)\theta) = \lambda_{\xi_p}. \tag{1.1}$$

So maximizing the minimum eigenvalue of the information matrix λ_{ξ_p} which is attained by an E-optimal design will simultaneously maximize the power of F test.

(ii) Suppose we have a linear transformation $c'\theta$, where $c = (c_1, \ldots, c_p)'$. A design $\xi^* \in \Xi$ which minimizes the maximum variance of the best linear unbiased estimator (b.l.u.e.) of $c'\theta$ under the constraint $c'c = 1$, i.e,

$$\begin{aligned} \min_{\xi \in \Xi} \max_{c:c'c=1} var(c'\hat{\theta}) &= \sigma^2 \min_{\xi \in \Xi} \max_{c:c'c=1} c'M(\xi)^{-1}c \\ &= \sigma^2 \min_{\xi \in \Xi} \frac{1}{\lambda_{\xi_p}} \\ &= \sigma^2 \lambda_{\xi_p^*}, \end{aligned}$$

is E-optimal.

(iii) From the geometric point of view, the confidence region for θ is defined as follows:

$$(\theta - \hat{\theta})' M(\xi)^{-1} (\theta - \hat{\theta}). \tag{1.2}$$

Therefore the maximum expected semi axis of this ellipsoid is proportional to $1/\lambda_{\xi_p}$, which is the largest eigenvalue of $M(\xi)^{-1}$, and the minimum expected semi axis is proportional to $1/\lambda_{\xi_1}$, which is the smallest eigenvalue of $M(\xi)^{-1}$. Thus the idea of E-optimality is to minimize the largest expected semi axis of the the ellipsoid.

1.1.3 Equivalence Theorem

Kiefer and Wolfowitz (4) (6), Kiefer (7) established the well known equivalence theorem which has widespread applications in identifying the optimal design or verifying the optimality of a given certain design.

Under the approximate design theory, we have each element of $M(\xi)$ as

$$m_{ij}(\xi) = \int_{\mathfrak{X}} f_i(x) f_j(x) \xi(dx), \quad i, j = 1, \ldots, p.$$

When $M(\xi)$ is non-singular, we can define

$$d(x, \xi) = f(x)' M(\xi)^{-1} f(x).$$

Theorem 6. *(Kiefer and Wolfowitz) Suppose $\mathfrak{X}$ is compact, the following three statements are equivalent:*

(I) ξ^* *maximizes* $\det M(\xi)$,

(II) ξ^* *minimizes* $\max_{x \in \mathfrak{X}} d(x, \xi)$,

(III) $\max_{x \in \mathfrak{X}} d(x, \xi^*) = p$.

White (8) extended the general equivalence theorem to nonlinear models. For the nonlinear model $f_{Y_x}(Y, x, \theta_1, \ldots, \theta_p)$, let

$$I_{ij}(x) = E_{Y_x}\{(\frac{\partial}{\partial \theta_i} \log f_{Y_x})(\frac{\partial}{\partial \theta_j} \log f_{Y_x})\}.$$

Define

$$m_{ij}(\xi) = \int_{\mathfrak{X}} I_{ij}(x)\xi(dx).$$

We can further define

$$d(x,\xi) = \mathrm{tr}\{I(x)M(\xi)^{-1}\}.$$

Theorem 7. *(White) The following conditions for a design measure ξ^* are equivalent:*

(I) ξ^* maximizes det $M(\xi)$,

(II) ξ^* minimizes $\max_{x\in\mathfrak{X}} d(x,\xi)$,

(III) $\max_{x\in\mathfrak{X}} d(x,\xi^*) = p$.

1.1.4 Carathéodory's Theorem

Definition 8. (Convex Hull) Let S be any set, then the convex hull of S is

$$con(s) = \{x : x = \sum_{i=1}^{k} \alpha_i x_i, \quad x_i \in S, \quad 0 \le \alpha_i \le 1, \quad \sum_{i=1}^{k} \alpha_i = 1, \quad k = 1, 2, \ldots\}.$$

Definition 9. (Dirac measure) ξ_x is Dirac if $P(X = x) = 1$ and $P(X \neq x) = 0$.

Theorem 10. *(Carathéodory's Theorem) Let $S \subset R^n$, then any point x in $con(s)$ can be written as $x = \sum_{i=1}^{n+1} \alpha_i x_i$, where $x_i \in S, \alpha_i \ge 0, \sum_{i=1}^{n+1} \alpha_i = 1$. Moreover if x is an extreme point of $con(s)$ then $x = \sum_{i=1}^{n} \alpha_i x_i$, where $x_i \in S, \alpha_i \ge 0, \sum_{i=1}^{n} \alpha_i = 1$.*

An upper bound for the support size of a design measure can be given in Theorem 10. Specifically, let $\mathfrak{M} = con\{M(\xi_x) : \xi_x \textit{ is Dirac}\}$. If $M(\xi) \in \mathfrak{M}$, and $M(\xi)$ is a $p \times p$ matrix, it can be written as a linear combination of at most $p(p+1)/2+1$ information matrices of Dirac designs. Thus the optimal design is based on at most $p(p+1)/2+1$ points.

1.2 On the de la Garza Phenomenon

Early works such as general equivalence theorem by Kiefer (4) (6) (7), geometric method by Elfving (9) and the discussions of efficiency by Hoel (10) contain substantial contributions for the optimal designs under linear models.

On the other hand, with the developing algorithms and numerical capabilities, one may be more interested in determining the number of support points in the optimal design class instead of theoretically obtaining the optimal design points. With some knowledge of the support size, it is relatively easy to identify the optimal design using existing algorithms.

Given a linear model with p parameters, in order to have all the parameters estimable, the number of support points is at least p. On the contrary, according to Carathéodory's Theorem, the upper bound is $p(p+1)/2+1$.

The theorem proposed by de la Garza (11) and later discussed in detail by Pukelsheim (12) and named the de la Garza Phenomenon by Khuri et al (13) stated that given a $(p-1)$th-degree polynomial regression model (p parameters in total) with i.i.d random errors, for any n point design where $n > p$, there exists a design with exactly p support points such that the information matrix of the latter one is not inferior to that of the former one under the Loewner ordering.

For nonlinear models with two parameters, Yang and Stufken (14) proposed an approach to identify the subclass of design: for any design ξ which does not belong to this class, there is a design in the class with an information matrix that dominates ξ under the Loewner ordering. It was then generalized to nonlinear models with an arbitrary number of parameters (Yang (15)).

The strategy is given as follows: suppose we have a nonlinear regression model with response variables belonging to the exponential family. In the framework of the approximate design theory, we can have $\xi = \{(c_i, \omega_i), i = 1, \ldots, n\}$, where c_i is a one-to-one transformation of the design point x_i. Then the information matrix for θ can be written as

$$M(\xi) = P(\theta)(\sum_{i=1}^{n} \omega_i C(\theta, c_i))(P(\theta))^T, \tag{1.3}$$

where $P(\theta)$ is a $p \times p$ nonsingular matrix that only depends on θ and

$$C(\theta, c_i) = \begin{pmatrix} \Psi_{11}(c_i) & \Psi_{12}(c_i) & \cdots & \Psi_{1p}(c_i) \\ \Psi_{12}(c_i) & \Psi_{22}(c_i) & \cdots & \Psi_{2p}(c_i) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_{1p}(c_i) & \Psi_{2p}(c_i) & \cdots & \Psi_{pp}(c_i) \end{pmatrix}.$$

Suppose we have two designs $\xi = \{(c_i, \omega_i), i = 1, \ldots, n\}$ and $\xi^* = \{(\tilde{c}_j, \tilde{\omega}_j), j = 1, \ldots, \tilde{n}\}$. Since $P(\theta)$ does not depend on ξ, $M(\xi) \le M(\xi^*)$ is equivalent to $\sum_{i=1}^{n} \omega_i C(\theta, c_i) \le \sum_{j=1}^{\tilde{n}} \tilde{\omega}_j C(\theta, \tilde{c}_j)$ under the Loewner ordering.

One sufficient condition to yield $M(\xi) \leq M(\xi^*)$ is that all the off-diagonal elements of the two matrices are equal, while at least one diagonal element of one matrix is not greater than the corresponding one of the other matrix. Specifically it is illustrated by the following equations:

$$\sum_{i=1}^{n} \omega_i \Psi_{lt}(c_i) = \sum_{j=1}^{\tilde{n}} \tilde{\omega}_j \Psi_{lt}(\tilde{c}_j),$$

for $1 \leq l \leq t \leq p$ except for some $l = t$

$$\sum_{i=1}^{n} \omega_i \Psi_{ll}(c_i) \leq \sum_{j=1}^{\tilde{n}} \tilde{\omega}_j \Psi_{ll}(\tilde{c}_j).$$

The detailed approach is initiated by the definition of $f_{l,t}$ such that

$$f_{l,t}(c) = \begin{cases} \Psi_l'(c) & t = 1, l = 1, \ldots, k, \\ (\frac{f_{l,t-1}(c)}{f_{t-1,t-1}(c)})' & 2 \leq t \leq k; t \leq l \leq k. \end{cases}$$

where $\Psi_1, \ldots, \Psi_k$ are functions assumed to be infinitely differentiable defined on $[A, B]$ and all $f_{l,l}$ have no zero value on $[A, B]$.

The structure of $f_{l,t}$ can be expressed as follows:

$$\begin{pmatrix} f_{1,1} = \Psi_1' & & & \\ f_{2,1} = \Psi_2' & f_{2,2} = \left(\frac{f_{2,1}}{f_{1,1}}\right)' & & \\ f_{3,1} = \Psi_3' & f_{3,2} = \left(\frac{f_{3,1}}{f_{1,1}}\right)' & f_{3,3} = \left(\frac{f_{3,2}}{f_{2,2}}\right)' & \\ f_{4,1} = \Psi_4' & f_{4,2} = \left(\frac{f_{4,1}}{f_{1,1}}\right)' & f_{3,3} = \left(\frac{f_{4,2}}{f_{2,2}}\right)' & f_{4,4} = \left(\frac{f_{4,3}}{f_{3,3}}\right)' \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Theorem 11. *(Yang) For a nonlinear regression model, suppose the information matrix can be written as* (1.3) *and* $c_i \in [A, B]$. *Rename all distinct* $\Psi_{lt}, 1 \leq l \leq t \leq p$, *to* $\Psi_1, \ldots, \Psi_k$ *such that (i)* Ψ_k *is one of* $\Psi_{ll}, 1 \leq l \leq p$ *and (ii) there is no* $\Psi_{lt} = \Psi_k$ *for* $l < t$. *Let* $F(c) = \prod_{l=1}^{k} f_{l,l}(c), c \in [A, B]$. *For any given design* ξ, *there exists a design* ξ^* *such that* $M(\xi) \leq M(\xi^*)$. *The support size of* ξ^* *varies with different values of* k *and* $F(c)$ *and is discussed in details here:*

(a) when k *is odd and* $F(c) < 0$, ξ^* *is based on at most* $(k+1)/2$ *points including point* A.

(b) when k *is odd and* $F(c) > 0$, ξ^* *is based on at most* $(k+1)/2$ *points including point* B.

(c) when k *is even and* $F(c) > 0$, ξ^* *is based on at most* $k/2+1$ *points including points* A *and* B.

(d) when k *is even and* $F(c) < 0$, ξ^* *is based on at most* $k/2$ *points.*

We can illustrate this approach by a simple example. Consider the Emax model,

$$\eta(x,\theta) = \theta_0 + \frac{\theta_1 x}{x+\theta_2}.$$

The information matrix can be decomposed as (1.3) with

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{\theta_2} & -\frac{1}{\theta_2} & 0 \\ \frac{1}{\theta_2^2} & -\frac{1}{\theta_2^2} & \frac{1}{\theta_1\theta_2} \end{pmatrix},$$

and

$$C(\theta, c_i) = \begin{pmatrix} 1 & c_i & c_i^2 \\ c_i & c_i^2 & c_i^3 \\ c_i^2 & c_i^3 & c_i^4 \end{pmatrix},$$

where $c_i = 1/(x_i+\theta_2)$. Let $\Psi_1 = c, \Psi_2 = c^2, \Psi_3 = c^3, \Psi_4 = c^4$. We can find that $f_{1,1} = 1, f_{2,2} = 2, f_{3,3} = 3, f_{4,4} = 4$, thus $F(c) = 24$. By situation (c), we have at most 3 points in the optimal design including the boundary points.

1.3 Literatures on Optimal Designs for Nonlinear Models

Nonlinear models have been extensively applied to explore the characteristics of data in biological, medical, environmental, and agricultural areas, see Seber and Wild (16) for different kinds of nonlinear models and their statistical inferences. In particular, a number of nonlinear models have been acknowledged in pharmacokinetics (PK) and pharmacodynamics (PD) analysis to model the time-concentration relationship after the administration of a drug, in order to characterize the pharmacological processes within the body, typically, the absorption, distribution, metabolism, and excretion (ADME) of the drug. Comprehensive introductions to PK/PD concepts and models can be found in a variety of books, see Gabrielsson and Weiner (17), Bonate (18), and Rosembaum (19). Among them, the class of exponential regression models is of particular interest attributable to their capabilities of handling data in compartmental systems.

Determinations of optimal designs for nonlinear models are not as well-established as those for linear models due to the dependence of the information matrix on the unknown parameters. As global optimization can hardly be obtained for nonlinear models, locally optimal designs have been proposed which are optimal in the neighborhood of the given values of the parameters (Chernoff (20)). Ford, Torsney and Wu (21) reduced the problem to a canonical form and constructed the locally optimal designs for nonlinear models within the exponential family.

In addition to Theorem 6, 7, 10, 11, there are a lot of literatures which focus on derivations of optimal designs for specific nonlinear models, see Hedayat, Zhong and Nie (22) for nonlinear models with two parameters. Among all kinds of nonlinear models, PK/PD models are of great interest due to their widespread applications in clinical studies. Li and Majumdar (23) worked on the one-compartment PK model. Fang and Hedayat (24) identified the D-optimal designs for the blending of Emax and one-compartment PK model. Ogungbenro et al (25) considered the unbalanced designs under PK/PD models. Zocchi and Atkinson (26) explored the multinomial logistic models in dose level experiments. Gueorguieva et al (27) studied the multivariate response PK models. Some other models have also been addressed. Li and

Majumdar (28) investigated the logistic models with three and four parameters. Konstantinou et al (29) analyzed the two parameter survival models. Dette et al (30) discussed the D-optimal designs for exponential regression models.

CHAPTER 2

SPECIAL DESIGN SETUP FOR NONLINEAR MODELS

2.1 Design Setup with Different Starting Points for Each Subject

The approximate design in which all subjects are measured once at a chosen design point might not be suitable for some clinical studies. On the other hand, the appropriateness of the usual clinical designs where measurements are taken at the same time points throughout the study for each subject is also questionable. In such cases when some subjects are only available at the early stages of the study, while others are available at the late stages, the data collected under the usual clinical design setup will most likely end up with many missing values.

To address this problem, we propose a special design which allows different starting points for each subject. In other words, after drug administration, the first sample of each subject can be taken at an arbitrary time point. This design setup shares some extent of similarities with the design used in growth curve models, see Abt et al (31) (32). But differently, we assume that subsequent samples for each subject are drawn at the following k time points. Here k is a constant chosen by the researcher which is assumed to be the same for all the subjects. Observations from different subjects are assumed to be independent.

Here we consider the equally spaced design with the support

$$0, 1, 2, \cdots, n,$$

which refers to the starting points of the subjects.

A simple example is presented here to illustrate the idea. Let $k = 5$, 50% of the subjects will start to be measured from time 5, and then they are continuing to be measured at time 6, 7, 8, 9, 10. 40% of the subjects will start from time 8, thus they are measured at time 8, 9, 10, 11, 12, 13. The remaining 10% will start from time 15, thus they are measured at time 15, 16, 17, 18, 19, 20. Overlapping is allowed under this design setup.

We assume that no distinctions is present among the subjects. After we come up with the design with several starting points and the associated weights, the allocations of subjects to different groups are entirely due to the preferences of the subjects or the researchers, which have no statistical interpretations.

We now present the design in a statistical way. Suppose the total number of subjects is N, and n_i is the number of subjects that start at time point t_i, $0 \leq t_i \leq n$. Let $\omega_i = n_i/N$, we represent the design ξ by the measure

$$\xi = \begin{pmatrix} 0 & 1 & \cdots & n \\ \omega_0 & \omega_1 & \cdots & \omega_n \end{pmatrix},$$

where $\sum_{i=0}^{n} \omega_i = 1$.

2.2 Covariance Structure

The autocorrelated covariance structure or the intraclass correlation structure is relevant here since each subject is measured at $k+1$ consecutive time points. We will assume that for the observations taken from time t_i through time $t_i + k$, the covariance structure is given by

$$\Sigma_i = \sigma^2((\rho^{|r-s|}))_{t_i \leq r,s \leq t_i+k}$$

for the autocorrelated covariance structure, and

$$\Sigma_i = \sigma^2[(1-\rho)I + \rho J]$$

for the intraclass correlation structure, where I is the identity matrix, and J is a matrix of ones.

Since Σ_i is assumed to be the same for all the subjects, we omit i in the later discussions.

The autocorrelated covariance structure assumes higher correlations between closer observations and lower correlations between farther observations, while the intraclass correlation structure assumes that the same correlation is present between every two observations within the same subject. We will study both in the later discussions.

2.3 Locally Optimal Design for General Nonlinear Model

Consider the general nonlinear model:

$$y = f(t, \theta) + \varepsilon.$$

Here y is a vector of all observations. Take a design with 2 starting points as an example, suppose n_i subjects start at time t_i and n_j subjects start at time t_j, then

$$y = (y_{1,t_i}, y_{1,t_i+1}, \cdots, y_{1,t_i+k}, y_{2,t_i}, \cdots, y_{n_i,t_i+k}, y_{1,t_j}, \cdots, y_{1,t_j+k}, \cdots, y_{n_j,t_j+k})'.$$

$t = (0, 1, \cdots, n+k)$ are the time points, $\theta = (\theta_1, \cdots, \theta_p)$ is a vector of p unknown parameters, and ε is the corresponding random error vector.

Based on the locally optimal design theory, we can derive the information matrix for the given nominal values of θ. To reduce the possibilities of misspecifications of θ, we may consult historical data in order to obtain more accurate guesses. Alternatively, a pilot study can be conducted to estimate the unknown parameters and the results will serve as the initial guesses.

2.4 Information Matrix for the Nonlinear Model with Covariance Structure

Let $\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_p)$ and $\tilde{\theta}_j$ be the initial guess of θ_j, $j = 1, \ldots, p$, and Σ be the covariance structure among observations for one single subject. Then the information matrix for one subject from t_i through $t_i + k$ is

$$M_i = G_i^T \Sigma^{-1} G_i,$$

where

$$G_i = \begin{pmatrix} \frac{\partial f(t_i, \tilde{\theta}_1)}{\partial \theta_1} & \cdots & \frac{\partial f(t_i, \tilde{\theta}_p)}{\partial \theta_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(t_i + k, \tilde{\theta}_1)}{\partial \theta_1} & \cdots & \frac{\partial f(t_i + k, \tilde{\theta}_p)}{\partial \theta_p} \end{pmatrix}. \tag{2.1}$$

For an exact design ξ with nonnegative integers n_i, $M(\xi)$ is given by

$$M(\xi) = \sum_{i=0}^{n} n_i M_i.$$

After dropping N, $M(\xi)$ can be written as

$$M(\xi) = \sum_{i=0}^{n} \omega_i M_i. \tag{2.2}$$

In the following chapters, we will implement the design setup on mono-exponential model and bi-exponential model.

CHAPTER 3

OPTIMAL DESIGNS FOR MONO-EXPONENTIAL MODEL

3.1 Introduction and Application

There are various situations in which explanatory variables are nonlinearly related. Among them, the class of exponential regression models is of particular interest due to their capabilities of handling data in dynamic population, drugs in the bloodstream, temperature, and some biomedical problems.

In biological science, the amount of a protein or substance usually follows an exponential model. If there are more than one process in the lifetime, it will be modeled by a multi-exponential model. For example, a single radioactive decay mode of a nuclide is described by a mono-exponential model. If two decay modes exist, for the second decay mode, another exponential term will be added, thus resulting in a bi-exponential model.

The simplest mono-exponential model is the one-compartment PK model with single dose and first order elimination rate. In a PK study, researchers are often interested in the ADME of the drug. ADME are often used to describe the disposition of a compound. In particular, all of the four aspects affect the kinetics of the drug as well as the drug levels. As a result, they will influence the performance and pharmacological activities of the drug.

Before the certain compound is taken up by the target cells, it has to reach a tissue through the bloodstream - usually via the digestive tract. Absoption critically determines the bioavailability of the compound. Oral administration of the drug is most preferable, but in case drugs are absorbed poorly when taken orally, some less desirable ways will be adopted, such as intravenously or by inhalation. After the absoption, the compound has to be carried to its effector site via the bloodstream. Distribution is defined as the reversible transfer of a drug between two compartments. Right after the drug enters the body, the compound starts to break down to metabolites, and this process is called metabolism. Excretion is the final step when the compound and its metabolites are removed from the body, usually through the kidneys or in the feces.

Understanding of the ADME of a drug is critical in the early phase of a clinical trial, when the toxicity of the compound is mainly studied. Therefore compartment models are important in describing the pharmacokinetics of a drug, which depict and predict the concentration-time curve.

One-compartment PK model makes the assumptions that the drug enters the central compartment (or compartment one) from somewhere outside of the body, then leaves the central compartment at a constant rate. It regards the entire body as one entity, assuming that the drug achieves instantaneous distribution throughout the body and equilibrates instantaneously between tissues. Thus the drug concentration-time profile shows a monophasic response as in Figure 1.

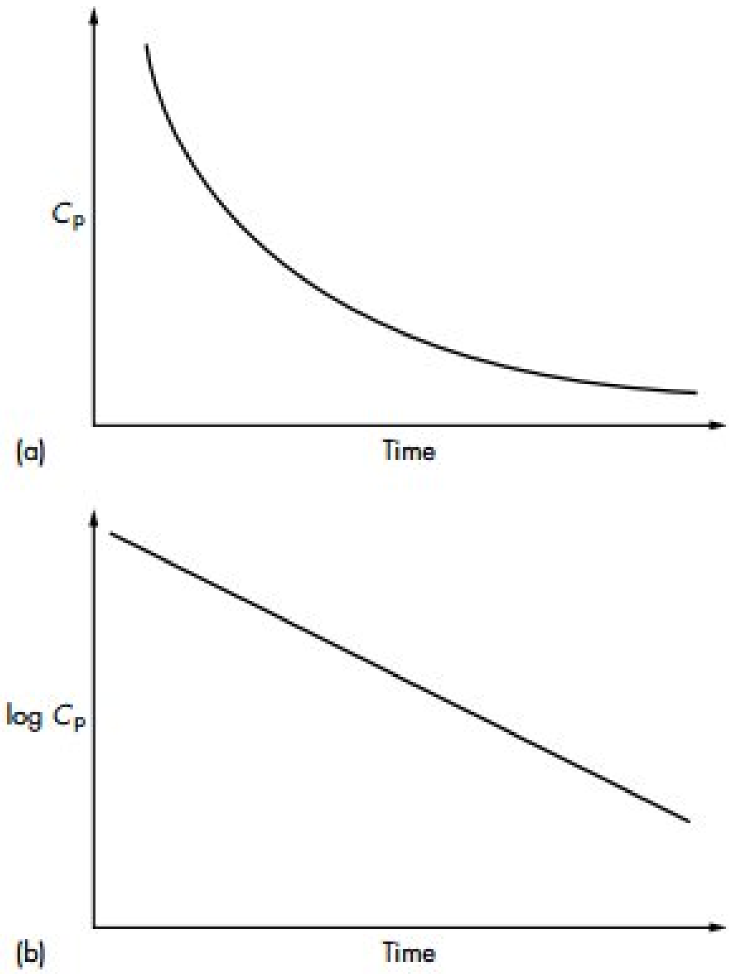

Figure 1: Plasma concentration (Cp) versus time profile of a drug showing a one-compartment model. (b) Time profile of a one-compartment model showing log Cp versus time.

Suppose the amount of drug A is decreasing at a rate that is proportional to the amount of drug A remaining in the body, denoted as $A(t)$, where t refers to the time. Then the elimination rate of drug A with respect to time t can be described by the following differential equation:

$$\frac{dA(t)}{dt} = -\lambda A(t).$$

Thus the solution of $A(t)$ is

$$A(t) = \frac{D}{V} e^{-\lambda t}, \tag{3.1}$$

where D is the dose of the drug administered, V is the volume of distribution, which is the theoretical volume that a drug would have to occupy to provide the same concentration as it currently is in blood plasma, and λ is the elimination rate.

3.2 Model Setup

As D represents the drug dose and thus can be considered as a constant, only V and λ are of interest. It is more convenient to replace $\frac{D}{V}$ by a single parameter a. Thus the model (3.1) can be rewritten as

$$y = ae^{-\lambda t} + \varepsilon$$

where $t = (0, 1, \cdots, n + k)$.

Theorem 12. *Under model* (3.1)*, a locally optimal design ξ has at most 2 starting points, and 0 is one of the starting points.*

We start proving the theorem for two measurements ($k = 1$) and then extend the proof to the case of $k + 1$ measurements.

3.3 Case of Two Measurements

We first consider the case with only two subsequent measurements, i.e, $k = 1$. Therefore, for each subject, samples will be taken at time t_i and $t_i + 1$.

In this case, both intraclass correlation structure and the autocorrelated covariance structure are the same as follows:

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

According to (2.2), we have shown that $M(\xi)$ is the weighted sum of M_i, where $M_i = G_i^T \Sigma^{-1} G_i, i = 0, \ldots, n$. From (2.1) we have

$$G_i^T = \begin{pmatrix} e^{-\lambda t_i} & e^{-\lambda(t_i+1)} \\ -at_i e^{-\lambda t_i} & -a(t_i+1)e^{-\lambda(t_i+1)} \end{pmatrix}.$$

By substituting $-\lambda t_i$ by a single parameter c_i, $M(\xi)$ can be decomposed as

$$M(\xi) = P(\theta)A\left(\sum_{i=0}^{n} \omega_i L_1(c_i) B \Sigma^{-1} B^T L_1(c_i)^T\right)(P(\theta)A)^T,$$

with

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}, L_1(c_i) = \begin{pmatrix} 0 & e^{c_i} \\ e^{c_i} & c_i e^{c_i} \end{pmatrix}, P(\theta) = \begin{pmatrix} 1 & 0 \\ 1 & \frac{1}{a} \end{pmatrix}^{-1}, B = \begin{pmatrix} 1 & 0 \\ \frac{1}{\lambda} & \frac{1}{\lambda e^{\lambda}} \end{pmatrix}.$$

Let $C(\theta, c_i) = L_1(c_i)SL_1(c_i)^T$, where $S = B\Sigma^{-1}B^T$. As S does not depend on ξ, it can be regarded as a constant matrix and we can simply denote it as $S = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$.

Further we denote $E(\xi) = \sum_{i=1}^{n} \omega_i C(\theta, c_i)$. Accordingly, the elements of $E(\xi)$ are represented by some functions of ω_i and c_i as shown below:

$$\begin{aligned}
E_{11} &= \sigma_{22} \sum_{i=0}^{n} \omega_i e^{2c_i}, \\
E_{12} &= \sigma_{12} \sum_{i=0}^{n} \omega_i e^{2c_i} + \sigma_{22} \sum_{i=0}^{n} \omega_i c_i e^{2c_i}, \\
E_{22} &= \sigma_{11} \sum_{i=0}^{n} \omega_i e^{2c_i} + 2\sigma_{12} \sum_{i=0}^{n} \omega_i c_i e^{2c_i} + \sigma_{22} \sum_{i=0}^{n} \omega_i c_i^2 e^{2c_i},
\end{aligned}$$

here, E_{11}, E_{12}, E_{22} are elements of $E(\xi)$.

We need to find a design $\xi^* = (\tilde{\omega}_i, \tilde{c}_i)$ such that $M(\xi) \leq M(\xi^*)$, which is equivalent to $E(\xi) \leq E(\xi^*)$. By the Loewner ordering, the sufficient conditions to prove the inequality are:

$$E_{11} = E_{11}^*, E_{12} = E_{12}^*, E_{22} \leq E_{22}^*, \tag{3.2}$$

where E_{11}^*, E_{12}^* and E_{22}^* are the elements of $E(\xi^*)$.

As long as $\sigma_{22} > 0$ which is guaranteed by the nonnegative definite property of S, (3.2) is equivalent to

$$\sum_{i=1}^{n} \omega_i e^{2c_i} = \sum_{j=1}^{n} \tilde{\omega}_i e^{2\tilde{c}_i},$$

$$\sum_{i=1}^{n} \omega_i c_i e^{2c_i} = \sum_{j=1}^{n} \tilde{\omega}_i \tilde{c}_i e^{2\tilde{c}_i},$$

$$\sum_{i=1}^{n} \omega_i c_i^2 e^{2c_i} \leq \sum_{j=1}^{n} \tilde{\omega}_i \tilde{c}_i^2 e^{2\tilde{c}_i}. \tag{3.3}$$

Let $\Psi_1 = e^{2c}, \Psi_2 = ce^{2c}, \Psi_3 = c^2 e^{2c}$, thus $F(c) = 4e^{2c}$. Clearly $F(c)$ is positive. According to situation (b) in Yang (15), the locally optimal design has at most 2 starting points, while 0 is one of the starting points.

3.4 Case of $k+1$ Measurements

We now consider the case of $k+1$ consecutive measurements in which subjects starting at time t_i will be measured at each time point through time $t_i + k$.

Generally we can denote the covariance structure as Σ. We have

$$G_i^T = \begin{pmatrix} e^{-\lambda t_i} & e^{-\lambda(t_i+1)} & \cdots & e^{-\lambda(t_i+k)} \\ -at_i e^{-\lambda t_i} & -a(t_i+1)e^{-\lambda(t_i+1)} & \cdots & -a(t_i+k)e^{-\lambda(t_i+k)} \end{pmatrix}.$$

Again by substituting $-\lambda t_i$ by c_i, $M(\xi)$ can be decomposed as

$$M(\xi) = P(\theta)A\Big(\sum_{i=0}^{n} \omega_i L_k(c_i) DB\Sigma^{-1}B^T D^T L_k(c_i)^T\Big)A^T P(\theta)^T,$$

with

$$P(\theta) = \begin{pmatrix} 1 & 0 \\ 1 & \frac{1}{a} \end{pmatrix}^{-1}, D = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 1 \end{pmatrix}, A = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}, L_k(c_i) = \begin{pmatrix} 0 & e^{c_i} \\ e^{c_i} & c_i e^{c_i} \end{pmatrix},$$

$$B = \begin{pmatrix} 1 & 0 & -e^{-2\lambda} & \cdots & (1-k)e^{-k\lambda} \\ \frac{1}{\lambda} & \frac{1}{\lambda e^{\lambda}} & \frac{1}{\lambda e^{2\lambda}} & \cdots & \frac{1}{\lambda e^{k\lambda}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{\lambda} & \frac{1}{\lambda e^{\lambda}} & \frac{1}{\lambda e^{2\lambda}} & \cdots & \frac{1}{\lambda e^{k\lambda}} \end{pmatrix},$$

Let $C(\theta, c_i) = L_k(c_i)DB\Sigma^{-1}B^TD^TL_k(c_i)^T$ and $E(\xi) = \sum_{i=0}^{n}\omega_i C(\theta, c_i)$, despite some constant coefficients, we still need to find a design $\xi^* = (\tilde{\omega}_i, \tilde{c}_i)$ which satisfies (3.2) and thus the three equations and inequalities in (3.3). Note that we have the same Ψ_l's as those in the case of $k = 1$, thus the result remains the same, i.e, the locally optimal design for case of $k+1$ measurements has at most two starting points, and 0 is one of the starting points.

Remark. *The correlation structure is treated as a constant matrix in our proof. Therefore the specific structure and the choice of ρ will not impact our conclusions about the number of support points. However, it will affect the actual design points and the associated weights, which will be presented later.*

3.5 Generalization to Equally Spaced Designs with Arbitrary Increments

We have discussed the equally spaced design with increment 1. However, when the design space is relatively large compared to the number of measurements for each subject, the equally spaced design with increment 1 is not capable of capturing all the information. For example, given the design space $[0, 240]$ and $k = 5$, we may come up with a design such that 30% of the subjects are measured at 30, 31, 32, 33, 34, 35, while 70% of the subjects are measured at 180, 181, 182, 183, 184, 185. Obviously a great deal of information will be lost under this design.

In this case, it is desirable to generalize the preceding result to equally spaced designs with arbitrary increments.

Suppose the increment is denoted as d, here d is also treated as a constant and assumed to be the same for each subject. Suppose the total number of measurements is still fixed at $k+1$, then for subjects starting at time t_i, samples will be collected at time $t_i, t_i+d, t_i+2d, \cdots, t_i+kd$. We have the following theorem.

Theorem 13. *Under the equally spaced design with arbitrary increments, the locally optimal design for model* (3.1) *has at most 2 starting points, and 0 is one of the starting points.*

Proof. Under the design setup with increment d and measurements $k+1$, we have

$$G_i^T = \begin{pmatrix} e^{-\lambda t_i} & e^{-\lambda(t_i+d)} & \cdots & e^{-\lambda(t_i+kd)} \\ -at_i e^{-\lambda t_i} & -a(t_i+d)e^{-\lambda(t_i+d)} & \cdots & -a(t_i+kd)e^{-\lambda(t_i+kd)} \end{pmatrix}.$$

By substituting $-\lambda t_i$ by c_i, $M(\xi)$ can be decomposed as

$$M(\xi) = P(\theta)\Big(\sum_{i=0}^{n} \omega_i L_{kd}(c_i)\Sigma^{-1}L_{kd}(c_i)^T\Big)P(\theta)^T,$$

with

$$P(\theta) = \begin{pmatrix} 1 & 0 \\ 1 & \frac{1}{a} \end{pmatrix}^{-1}$$

and

$$L_{kd}(c_i) = \begin{pmatrix} e_i^c & e^{c_i-\lambda d} & \cdots & e^{c_i-\lambda kd} \\ \frac{c_i}{\lambda}e^{c_i} & \frac{c_i}{\lambda}e^{c_i-\lambda d} - de^{c_i-\lambda d} & \cdots & \frac{c_i}{\lambda}e^{c_i-\lambda kd} - kde^{c_i-\lambda kd} \end{pmatrix}_{2\times(k+1)}.$$

Let $C(\theta, c_i) = L_{kd}(c_i)\Sigma^{-1}L_k(c_i)^T$ and $E(\xi) = \sum_{i=0}^{n} \omega_i C(\theta, c_i)$, despite some constant coefficients, we need to find a design $\xi^* = (\tilde{\omega}_i, \tilde{c}_i)$ which satisfies (3.2) and thus the three equations and inequalities in (3.3). Note that we have the same Ψ_l's and the result remains the same, i.e, we have at most two starting points, and 0 is one of the starting points. □

3.6 Numerical Results

3.6.1 Equally Spaced Designs

In real cases, studies are often designed with regard to the actual situations. It should be taken into account the characteristics of the treatments, the recruiting of the study subjects, the regulations, the side effects or adverse events that may occur, etc. Optimality is not always a first priority in the real studies.

However, because of the superiority of optimal designs in providing the parameter estimates with minimum variances, it is beneficial to have some knowledge of them. In effect, an optimal design can serve as a gold standard for all other designs. In cases where optimal designs are not conducted, we may want to have an idea of the performance of the selected design.

There is a widely accepted criterion in comparisons of different designs, which is the so called "efficiency". In the design theory, it is a measure of the performance of the design.

$D-$ and $A-$optimality criteria are among the most popular optimality criteria. Under D-optimality, efficiency is defined as $(\frac{|M(\xi)|}{|M(\xi^*)|})^{1/p}$, where ξ^* is the optimal design, and p is the number of parameters to be estimated and studied. Under A-optimality, since it leads to the design with minimum trace of the inverse of the information matrix, the efficiency is thus defined as $(\frac{\mathrm{tr}(M(\xi^*))}{\mathrm{tr}(M(\xi))})^{1/p}$. In a geometric way, no matter which criterion we choose, essentially the optimal design is to minimize the confidence region of the parameter estimates. All the other designs are unable to reduce the confidence region to the extent that the optimal design does. As a matter of fact, efficiency is a measure of the capability of the chosen design in reducing the confidence region as a percentage of that for the optimal design. It is a value which is always less than 1, and higher efficiency leads to a better design.

Under the one-compartment PK model, theoretically we have found that the number of starting points in the optimal design class should not exceed 2.

The initial guesses of the parameters were pre-determined at $\tilde{a} = 1$ and $\tilde{\lambda} = 0.2$. The design space was $[0, 180]$. We set increment d at 15 and the total number of measurements at 13 (k=12).

3.6.1.1 Autocorrelated Covariance Structure

We first adopt the autocorrelated covariance structure to address the correlations among observations. According to the Remark, the values of ρ will not affect the support size but will have an impact on the actual design points and the associated weights. In an attempt to give a rough picture of the optimal design under a specific value of ρ, we obtain the following design when $\rho = 0.5$ by numerical searching.

The referenced equally spaced design is arbitrarily chosen to be: 2, 17, 32, 47, 62, 77, 92, 107, 122, 137, 152, 167, 182, and the D-optimal design when $\rho = 0.5$ has two starting points, with 59% of the subjects starting at time 0 and 41% starting at time 5. Specifically, 59% of the subjects will be measured at 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180, and 41% of the subjects will be measured at 5, 20, 35, 50, 65, 80, 95, 110, 125, 140, 155, 170, 185.

As a result of the influence of different ρ's on the design points and their weights, the efficiencies of the optimal designs with respect to the referenced design will also vary with different values of ρ. The summary of the efficiencies under different values of ρ is given in Table I.

TABLE I: D-Efficiency of an Equally Spaced Design for One-Compartment PK Model with Autocorrelated Covariance Structure

ρ	D_{opt}	D_{ref}	**Efficiency**$((D_{ref}/D_{opt})^{\frac{1}{2}})$
0.9	24.38	2.62	32.8%
0.8	6.93	0.75	32.9%
0.7	3.53	0.38	32.8%
0.6	2.29	0.25	33.0%
0.5	1.70	0.18	32.5%
0.4	1.38	0.15	33.0%
0.3	1.20	0.13	32.9%
0.2	1.10	0.12	33.0%
0.1	1.06	0.11	32.2%

On average the referenced design is shown 33% efficiency of the optimal design. As we know, the equally spaced designs with different starting points but with the same d and k have different efficiencies compared with the optimal designs. If the starting point is by chance chosen at 0 or 5, which is one of the starting points in the optimal design, the efficiency will be much higher than the design in which the starting point is neither 0 nor 5.

The A-optimal design when $\rho = 0.5$ is chosen to be: all of the subjects will be measured at 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180. The efficiencies under different values of ρ are given in Table II.

TABLE II: A-Efficiency of an Equally Spaced Design for One-Compartment PK Model with Autocorrelated Covariance Structure

ρ	A_{opt}	A_{ref}	**Efficiency**$((A_{opt}/A_{ref})^{\frac{1}{2}})$
0.9	0.44	2.75	40%
0.8	0.82	5.51	38.6%
0.7	1.14	8.23	37.2%
0.6	1.41	10.89	36.0%
0.5	1.63	13.43	34.8%
0.4	1.81	15.80	33.8%
0.3	1.94	17.96	32.9%
0.2	2.04	19.86	32.0%
0.1	2.10	21.43	31.3%

3.6.1.2 Intraclass Correlation Structure

We now adopt the intraclass correlation structure and study the efficiency. The D-optimal design when $\rho = 0.5$ is chosen to be: 59% of the subjects will be measured at 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180, and 41% of the subjects will be measured at 5, 20, 35, 50, 65, 80, 95, 110, 125, 140, 155, 170, 185. The efficiencies under different values of ρ are given in Table III.

TABLE III: D-Efficiency of an Equally Spaced Design for One-Compartment PK Model with Intraclass Correlation Structure

ρ	D_{opt}	D_{ref}	**Efficiency**$((D_{ref}/D_{opt})^{\frac{1}{2}})$
0.9	88.49	9.46	32.7%
0.8	22.17	2.37	32.7%
0.7	9.88	1.06	32.8%
0.6	5.57	0.60	32.8%
0.5	3.58	0.38	32.6%
0.4	2.50	0.27	32.9%
0.3	1.86	0.20	32.8%
0.2	1.44	0.15	32.3%
0.1	1.18	0.13	33.2%

The A-optimal design when $\rho = 0.5$ is chosen to be: 75% of the subjects will be measured at 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180, and 25% of the subjects will be measured at 5, 20, 35, 50, 65, 80, 95, 110, 125, 140, 155, 170, 185. The efficiencies under different values of ρ are given in Table IV.

TABLE IV: A-Efficiency of an Equally Spaced Design for One-Compartment PK Model with Intraclass Correlation Structure

ρ	A_{opt}	A_{ref}	**Efficiency**$((A_{opt}/A_{ref})^{\frac{1}{2}})$
0.9	0.23	2.37	31.2%
0.8	0.46	4.73	31.2%
0.7	0.69	7.10	31.2%
0.6	0.91	9.45	31.0%
0.5	1.14	11.80	31.1%
0.4	1.37	14.14	31.1%
0.3	1.59	16.46	31.1%
0.2	1.80	18.73	31.0%
0.1	2.00	20.89	30.9%

3.6.2 Unequally Spaced Designs

In real studies, researchers may have more interest in the early stages of the study since usually they will provide more information. Consequently more measurements are taken at the early phases than the late phases. In this sense, it is prudent to apply the unequally spaced designs rather than the equally spaced designs.

Although we don't have theoretical results, we can use numerical methods to find unequally spaced designs in which samples can be taken with different time increments.

The referenced design is chosen to be: 0, 2, 5, 10, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, which comes from a real PK study (33). For simplicity, we reduce our search for the optimal designs within the design space of $[0, 19]$ so that we only identify the former 4 points between 0 and 19 while keeping the latter 10 points, i.e, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, the same as the referenced design.

3.6.2.1 Autocorrelated Covariance Structure

We first adopt the autocorrelated covariance structure and study the efficiency. The D-optimal design when $\rho = 0.5$ is chosen to be: 94% of the subjects are measured at 0, 4, 5, 6, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, and 6% of the subjects are measured at 0, 5, 6, 7, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240. The summary of the efficiencies under different ρ is given in Table V.

TABLE V: D-Efficiency of an Unequally Spaced Design for One-Compartment PK Model with Autocorrelated Covariance Structure

ρ	D_{opt}	D_{ref}	**Efficiency**$((D_{ref}/D_{opt})^{\frac{1}{2}})$
0.9	64.66	27.08	64.7%
0.8	18.82	9.67	71.7%
0.7	10.88	6.19	75.4%
0.6	8.34	5.05	77.8%
0.5	7.36	4.71	80.0%
0.4	7.10	4.78	82.1%
0.3	7.29	5.17	84.2%
0.2	7.84	5.84	86.3%
0.1	8.88	6.87	88.0%

The A-optimal design when $\rho = 0.3$ is chosen to be: 96% of the subjects will be measured at 0, 1, 6, 7, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, and 4% of the subjects will be measured at 0, 6, 7, 8, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240. The efficiencies under different values of ρ are given in Table VI.

TABLE VI: A-Efficiency of an Unequally Spaced Design for One-Compartment PK Model with Autocorrelated Covariance Structure

ρ	A_{opt}	A_{ref}	**Efficiency**$((A_{opt}/A_{ref})^{\frac{1}{2}})$
0.9	0.37	0.67	74.3%
0.8	0.67	0.98	82.7%
0.7	0.88	1.12	88.6%
0.6	0.99	1.18	91.6%
0.5	1.04	1.20	93.1%
0.4	1.06	1.20	94.0%
0.3	1.05	1.19	93.9%
0.2	1.00	1.16	92.8%
0.1	0.95	1.11	85.6%

3.6.2.2 Intraclass Correlation Structure

We now adopt the intraclass correlation structure and study its efficiency. The D-optimal design when $\rho = 0.5$ is chosen to be: 31% of the subjects will be measured at 0, 4, 5, 6, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, and 69% of the subjects will be measured at 0, 1, 5, 6, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240. The efficiencies under different values of ρ are given in Table VII.

TABLE VII: D-Efficiency of an Unequally Spaced Design for One-Compartment PK Model with Intraclass Correlation Structure

ρ	D_{opt}	D_{ref}	**Efficiency**$((D_{ref}/D_{opt})^{\frac{1}{2}})$
0.9	735.04	582.62	89.0%
0.8	184.57	146.27	89.0%
0.7	82.48	65.35	89.0%
0.6	46.73	37.01	89.0%
0.5	30.19	23.90	89.0%
0.4	21.24	16.81	89.0%
0.3	15.92	12.59	89.0%
0.2	12.61	9.96	88.9%
0.1	10.66	8.40	88.8%

The A-optimal design when $\rho = 0.5$ is chosen to be: all of the subjects will be measured at 0, 1, 6, 7, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240. The efficiencies under different values of ρ are given in Table VIII.

TABLE VIII: A-Efficiency of an Unequally Spaced Design for One-Compartment PK Model with Intraclass Correlation Structure

ρ	A_{opt}	A_{ref}	**Efficiency**$((A_{opt}/A_{ref})^{\frac{1}{2}})$
0.9	0.10	0.11	95.3%
0.8	0.20	0.23	93.3%
0.7	0.29	0.34	92.4%
0.6	0.39	0.45	93.1%
0.5	0.48	0.57	91.8%
0.4	0.58	0.68	92.4%
0.3	0.67	0.79	92.1%
0.2	0.76	0.89	92.4%
0.1	0.84	0.99	92.1%

CHAPTER 4

OPTIMAL DESIGNS FOR BI-EXPONENTIAL MODEL

4.1 Introduction and Application

We now consider the general bi-exponential model. It is assumed that the two processes are independent, or the two compartments have no exchange rate. Thus the two-compartment PK model is not considered here. Figure 2 shows the graph of a general bi-exponential model.

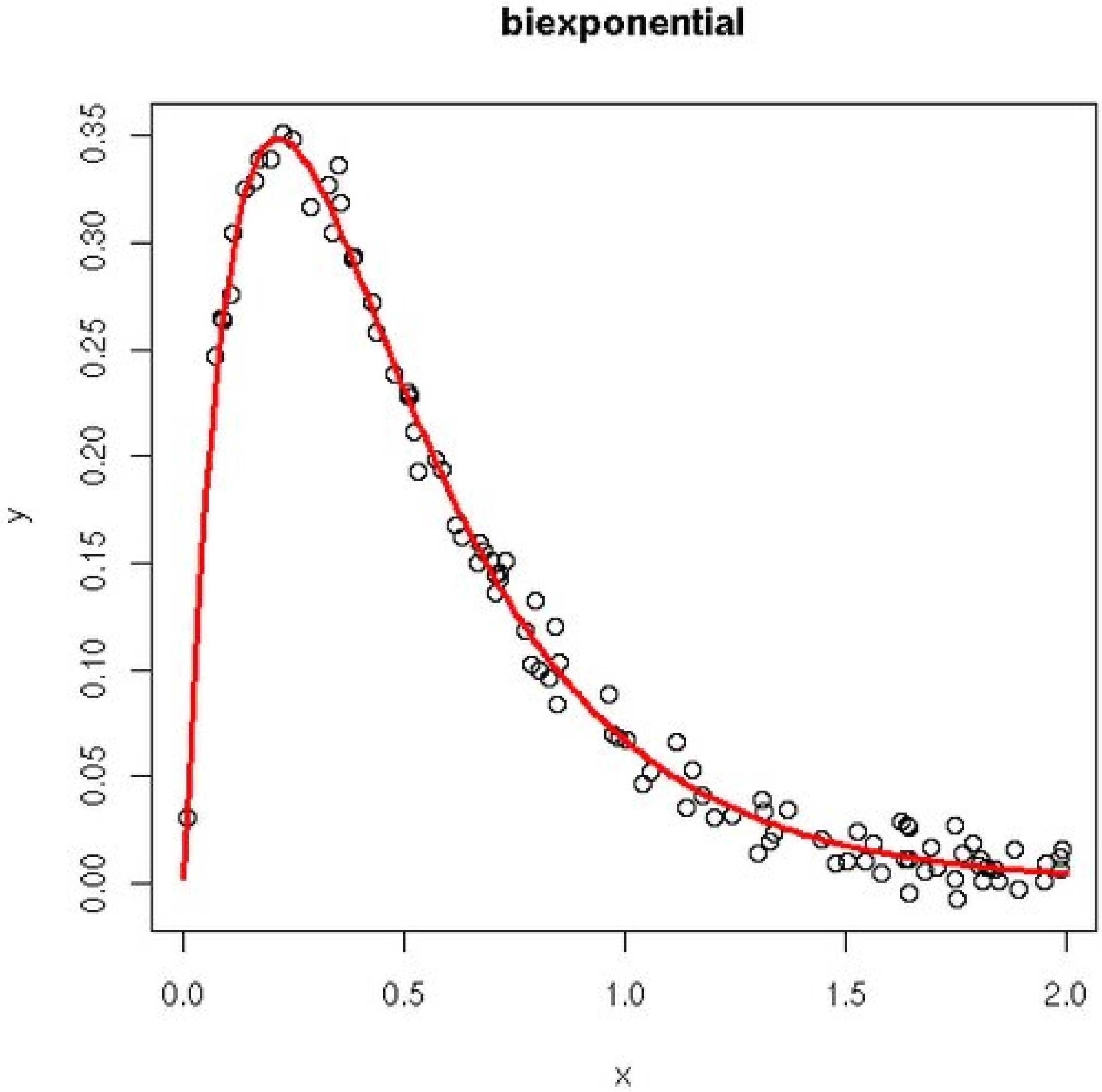

Figure 2: Response variable y versus explanatory variable x showing a bi-exponential model.

A typical example of bi-exponential model with 4 parameters is provided by Wells et al (34). The intention of their experiment is to measure bi-exponential transverse relaxation of the arterial spin labeling (ASL) signal so as to estimate arterial oxygen saturation and the time of exchange of labeled blood water into cortical brain tissue.

ASL is a non-invasive magnetic resonance imaging (MRI) technique in which blood water is harnessed as an endogenous tracer to map quantitative cerebral blood flow. After labeling, some of the labeled blood water enters the imaging volume and exchanges into the tissue (extravascular (EV) space) and some remains in the vessels (intravascular (IV) space).

The mean ASL signal as a function of measurement time is fitted by a bi-exponential model:

$$\triangle M(t) = \triangle M_A e^{-\lambda_1 t} + \triangle M_B e^{-\lambda_2 t},$$

where $\triangle M_A$ is the proportion of the signal in compartment EV with transverse decay rate λ_1, and $\triangle M_B$ is the proportion of the signal in compartment IV with transverse decay rate λ_2, and $\triangle M$ is the mean cortical ASL signal. In this case, the two compartments (EV and IV) are regarded as independent, and in each compartment, the amount of ASL signal is decaying exponentially.

4.2 Model Setup

Consider the general bi-exponential model

$$y = a_1 e^{-\lambda_1 t} + a_2 e^{-\lambda_2 t} + \varepsilon \tag{4.1}$$

where a_1, a_2, λ_1 and λ_2 are the parameters of interest, and $t = (0, 1, \cdots, n+k)$.

Theorem 14. *Under model* (4.1)*, a locally optimal design* ξ *has at most 5 starting points, and 0 is one of the starting points.*

Again we start proving the theorem for two measurements ($k = 1$) and then extend the proof to the case of $k+1$ measurements.

4.3 Case of Two Measurements

According to (2.2), we have $M_i = G_i^T \Sigma^{-1} G_i$, where

$$G_i^T = \begin{pmatrix} e^{-\lambda_1 t_i} & e^{-\lambda_1 (t_i+1)} \\ -a_1 t_i e^{-\lambda_1 t_i} & -a_1 (t_i+1) e^{-\lambda_1 (t_i+1)} \\ e^{-\lambda_2 t_i} & e^{-\lambda_2 (t_i+1)} \\ -a_2 t_i e^{-\lambda_2 t_i} & -a_2 (t_i+1) e^{-\lambda_2 (t_i+1)} \end{pmatrix}.$$

After substituting $-\lambda_1 t_i$ by c_i, $M(\xi)$ can be decomposed as

$$M(\xi) = P(\theta) \Big(\sum_{i=0}^{n} \omega_i L_1(c_i) \Sigma^{-1} L_1(c_i)^T \Big) P(\theta)^T,$$

with

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & \frac{1}{a_1} & 1 & \frac{1}{a_2} \\ 1 & 0 & 1 & 0 \\ 1 & \frac{1}{a_1} & 1 & \frac{1}{a_2} \end{pmatrix}^{-1}$$

and

$$L_1(c_i) = \begin{pmatrix} e^{c_i} & e^{c_i-\lambda_1} \\ 2e^{c_i} + \frac{c_i}{\lambda_1}e^{c_i} + \frac{c_i}{\lambda_1}e^{\frac{\lambda_2}{\lambda_1}c_i} + e^{\frac{\lambda_2}{\lambda_1}c_i} & \frac{c_i}{\lambda_1}e^{c_i-\lambda_1} + e^{c_i-\lambda_1} + \frac{c_i}{\lambda_1}e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2} \\ e^{c_i} + e^{\frac{\lambda_2}{\lambda_1}c_i} & e^{c_i-\lambda_1} + e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2} \\ e^{c_i} + \frac{c_i}{\lambda_1}e^{c_i} + e^{\frac{\lambda_2}{\lambda_1}c_i} + \frac{c_i}{\lambda_1}e^{\frac{\lambda_2}{\lambda_1}c_i} & \frac{c_i}{\lambda_1}e^{c_i-\lambda_1} + \frac{c_i}{\lambda_1}e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2} \end{pmatrix},$$

Let $C(\theta, c_i) = L_1(c_i)\Sigma^{-1}L_1(c_i)^T$. Also Σ^{-1} can be regarded as a constant matrix.

Denote $E(\xi) = \sum_{i=0}^{n} \omega_i C(\theta, c_i)$, we need to find a design $\xi^* = (\tilde{\omega}_i, \tilde{c}_i)$ such that $M(\xi) \leq M(\xi^*)$, which is equivalent to $E(\xi) \leq E(\xi^*)$. By the Loewner ordering, it is sufficient to show that

$$E_{11} = E_{11}^*, E_{12} = E_{12}^*, E_{13} = E_{13}^*, E_{14} = E_{14}^*,$$

$$E_{23} = E_{23}^*, E_{24} = E_{24}^*, E_{33} = E_{33}^*, E_{34} = E_{34}^*,$$

$$E_{22} \leq E_{22}^*, E_{44} \leq E_{44}^*. \tag{4.2}$$

where $E_{ij}, 1 \leq i \leq j \leq 4$, are elements of $E(\xi)$, and $E_{ij}^*, 1 \leq i \leq j \leq 4$, are elements of $E(\xi^*)$.

The sufficient conditions for (4.2) are as follows:

$$\sum_{i=0}^{n} \omega_i e^{2c_i} = \sum_{i=0}^{n} \tilde{\omega}_i e^{2\tilde{c}_i}$$

$$\sum_{i=0}^{n} \omega_i e^{(\frac{\lambda_2}{\lambda_1}+1)c_i} = \sum_{i=0}^{n} \tilde{\omega}_i e^{(\frac{\lambda_2}{\lambda_1}+1)\tilde{c}_i}$$

$$\sum_{i=0}^{n} \omega_i e^{\frac{2\lambda_2}{\lambda_1}c_i} = \sum_{i=0}^{n} \tilde{\omega}_i e^{\frac{2\lambda_2}{\lambda_1}\tilde{c}_i}$$

$$\sum_{i=0}^{n} \omega_i c_i e^{2c_i} = \sum_{i=0}^{n} \tilde{\omega}_i \tilde{c}_i e^{2\tilde{c}_i}$$

$$\sum_{i=0}^{n} \omega_i c_i e^{(\frac{\lambda_2}{\lambda_1}+1)c_i} = \sum_{i=0}^{n} \tilde{\omega}_i \tilde{c}_i e^{(\frac{\lambda_2}{\lambda_1}+1)\tilde{c}_i}$$

$$\sum_{i=0}^{n} \omega_i c_i e^{\frac{2\lambda_2}{\lambda_1}c_i} = \sum_{i=0}^{n} \tilde{\omega}_i \tilde{c}_i e^{\frac{2\lambda_2}{\lambda_1}\tilde{c}_i}$$

$$\sum_{i=0}^{n} \omega_i c_i^2 e^{2c_i} \leq \sum_{i=0}^{n} \tilde{\omega}_i \tilde{c}_i^2 e^{2\tilde{c}_i}$$

$$\sum_{i=0}^{n} \omega_i c_i^2 e^{(\frac{\lambda_2}{\lambda_1}+1)\tilde{c}_i} \leq \sum_{i=0}^{n} \tilde{\omega}_i \tilde{c}_i^2 e^{(\frac{\lambda_2}{\lambda_1}+1)\tilde{c}_i}$$

$$\sum_{i=0}^{n} \omega_i c_i^2 e^{\frac{2\lambda_2}{\lambda_1}\tilde{c}_i} \leq \sum_{i=0}^{n} \tilde{\omega}_i \tilde{c}_i^2 e^{\frac{2\lambda_2}{\lambda_1}\tilde{c}_i} \tag{4.3}$$

Denote $\Psi_1 = e^{2c}, \Psi_2 = ce^{2c}, \Psi_3 = c^2e^{2c}, \Psi_4 = e^{(\frac{\lambda_2}{\lambda_1}+1)c}, \Psi_5 = ce^{(\frac{\lambda_2}{\lambda_1}+1)c}, \Psi_6 = c^2e^{(\frac{\lambda_2}{\lambda_1}+1)c}, \Psi_7 = e^{\frac{2\lambda_2}{\lambda_1}c}, \Psi_8 = ce^{\frac{2\lambda_2}{\lambda_1}c}, \Psi_9 = c^2e^{\frac{2\lambda_2}{\lambda_1}c}$, we can have

$$F(c) = \frac{32e^{2c}(\lambda_1 - \lambda_2)^6(e^{\frac{c(\lambda_1-\lambda_2)}{\lambda_2}})^2\lambda_1}{\lambda_2^7}.$$

Clearly $F(c)$ is positive. According to situation (b) in Yang (15), the locally optimal design has at most 5 starting points, while 0 is one of the starting points.

4.4 Case of $k+1$ Measurements

We now consider the case of $k+1$ measurements. We have

$$G_i^T = \begin{pmatrix} e^{-\lambda_1 t_i} & e^{-\lambda_1(t_i+1)} & \cdots & e^{-\lambda_1(t_i+k)} \\ -a_1 t_i e^{-\lambda_1 t_i} & -a_1(t_i+1)e^{-\lambda_1(t_i+1)} & \cdots & -a_1(t_i+k)e^{-\lambda_1(t_i+k)} \\ e^{-\lambda_2 t_i} & e^{-\lambda_2(t_i+1)} & \cdots & e^{-\lambda_2(t_i+k)} \\ -a_2 t_i e^{-\lambda_2 t_i} & -a_2(t_i+1)e^{-\lambda_2(t_i+1)} & \cdots & -a_2(t_i+k)e^{-\lambda_2(t_i+k)} \end{pmatrix}.$$

Again by substituting $-\lambda_1 t_i$ by c_i, $M(\xi)$ can be decomposed as

$$M(\xi) = P(\theta)\Big(\sum_{i=0}^{n} \omega_i L_k(c_i)\Sigma^{-1}L_k(c_i)^T\Big)P(\theta)^T,$$

with

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & \frac{1}{a_1} & 1 & \frac{1}{a_2} \\ 1 & 0 & 1 & 0 \\ 1 & \frac{1}{a_1} & 1 & \frac{1}{a_2} \end{pmatrix}^{-1}$$

and

$$L_k(c_i) = \begin{pmatrix} e^{c_i} & e^{c_i-\lambda_1 d} & \cdots & e^{c_i-\lambda_1 kd} \\ (2+\frac{c_i}{\lambda_1})e^{c_i} + (1+\frac{c_i}{\lambda_1})e^{\frac{\lambda_2}{\lambda_1}c_i} & (1+\frac{c_i}{\lambda_1})e^{c_i-\lambda_1 d} + \frac{c_i}{\lambda_1}e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_1 d} & \cdots & (2-k+\frac{c_i}{\lambda_1})e^{c_i-k\lambda_1} + (1-k+\frac{c_i}{\lambda_1})e^{\frac{\lambda_2}{\lambda_1}c_i-k\lambda_2} \\ e^{c_i} + e^{\frac{\lambda_2}{\lambda_1}c_i} & e^{c_i-\lambda_1} + e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2} & \cdots & e^{c_i-k\lambda_1} + e^{\frac{\lambda_2}{\lambda_1}c_i-k\lambda_2} \\ (1+\frac{c_i}{\lambda_1})e^{c_i} + (1+\frac{c_i}{\lambda_1})e^{\frac{\lambda_2}{\lambda_1}c_i} & \frac{c_i}{\lambda_1}e^{c_i-\lambda_1} + \frac{c_i}{\lambda_1}e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2} & \cdots & (1-k+\frac{c_i}{\lambda_1})e^{c_i-k\lambda_1} + (1-k+\frac{c_i}{\lambda_1})e^{\frac{\lambda_2}{\lambda_1}c_i-k\lambda_2} \end{pmatrix}$$

Let $C(\theta, c_i) = L_k(c_i)\Sigma^{-1}L_k(c_i)^T$ and $E(\xi) = \sum_{i=0}^{n} \omega_i C(\theta, c_i)$, and we also need to find a design $\xi^* = (\tilde{\omega}_i, \tilde{c}_i)$ which satisfies (4.2) and thus the nine equations and inequalities in (4.3). Note that we have the same Ψ_l's as in the case of $k = 1$ and the result remains the same.

4.5 Generalization to Equally Spaced Designs with Arbitrary Increments

We can generalize the preceding results to the equally spaced designs with arbitrary increments, thus we have the following theorem.

Theorem 15. *Under the equally spaced designs with arbitrary increments, the locally optimal design for model* (4.1) *has at most 5 starting points, and 0 is one of the starting points.*

Proof. Under the design setup with increment d and measurements $k + 1$, we have

$$G_i^T = \begin{pmatrix} e^{-\lambda_1 t_i} & e^{-\lambda_1 (t_i+d)} & \cdots & e^{-\lambda_1 (t_i+kd)} \\ -a_1 t_i e^{-\lambda_1 t_i} & -a_1(t_i + d)e^{-\lambda_1 (t_i+d)} & \cdots & -a_1(t_i + kd)e^{-\lambda_1 (t_i+kd)} \\ e^{-\lambda_2 t_i} & e^{-\lambda_2 (t_i+d)} & \cdots & e^{-\lambda_2 (t_i+kd)} \\ -a_2 t_i e^{-\lambda_2 t_i} & -a_2(t_i + d)e^{-\lambda_2 (t_i+d)} & \cdots & -a_2(t_i + kd)e^{-\lambda_2 (t_i+kd)} \end{pmatrix}.$$

By substituting $-\lambda_1 t_i$ by c_i, $M(\xi)$ can be decomposed as

$$M(\xi) = P(\theta)\left(\sum_{i=0}^{n} \omega_i L_{kd}(c_i)\Sigma^{-1}L_{kd}(c_i)^T\right)P(\theta)^T,$$

with

$$P(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & \frac{1}{a_1} & 1 & \frac{1}{a_2} \\ 1 & 0 & 1 & 0 \\ 1 & \frac{1}{a_1} & 1 & \frac{1}{a_2} \end{pmatrix}^{-1}$$

and

$$L_{kd}(c_i) = \begin{pmatrix} e^{c_i} & e^{c_i-\lambda_1 d} & \cdots & e^{c_i-\lambda_1 kd} \\ (2+\frac{c_i}{\lambda_1})e^{c_i} + (1+\frac{c_i}{\lambda_1})e^{\frac{\lambda_2}{\lambda_1}c_i} & (1+\frac{c_i}{\lambda_1})e^{c_i-\lambda_1 d} + \frac{c_i}{\lambda_1}e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_1 d} & \cdots & (2-kd+\frac{c_i}{\lambda_1})e^{c_i-\lambda_1 kd} + (1-kd+\frac{c_i}{\lambda_1})e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2 kd} \\ e^{c_i} + e^{\frac{\lambda_2}{\lambda_1}c_i} & e^{c_i-\lambda_1 d} + e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2 d} & \cdots & e^{c_i-\lambda_1 kd} + e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2 kd} \\ (1+\frac{c_i}{\lambda_1})e^{c_i} + (1+\frac{c_i}{\lambda_1})e^{\frac{\lambda_2}{\lambda_1}c_i} & \frac{c_i}{\lambda_1}e^{c_i-\lambda_1 d} + \frac{c_i}{\lambda_1}e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2 d} & \cdots & (1-kd+\frac{c_i}{\lambda_1})e^{c_i-\lambda_1 kd} + (1-kd+\frac{c_i}{\lambda_1})e^{\frac{\lambda_2}{\lambda_1}c_i-\lambda_2 kd} \end{pmatrix}.$$

Let $C(\theta, c_i) = L_{kd}(c_i)\Sigma^{-1}L_{kd}(c_i)^T$ and $E(\xi) = \sum_{i=0}^{n} \omega_i C(\theta, c_i)$, and we also need to find a design $\xi^* = (\tilde{\omega}_i, \tilde{c}_i)$ which satisfies (4.2) and thus the nine equations and inequalities in (4.3). Note that we have the same Ψ_l's as the previous cases and the result remains the same. $\square$

4.6 Numerical Results

4.6.1 Equally Spaced Designs

Under the bi-exponential model, theoretically we have found that the number of starting points in the optimal design class should not exceed 5.

The initial guesses of the parameters were pre-determined at $\tilde{a}_1 = 1.38$, $\tilde{a}_2 = 0.42$, $\tilde{\lambda}_1 = 0.123$, $\tilde{\lambda}_2 = 0.00673$. The design space was $[0, 180]$ and we set $d = 15$ and $k = 12$.

4.6.1.1 Autocorrelated Covariance Structure

The referenced equally spaced design is arbitrarily chosen to be: 2, 17, 32, 47, 62, 77, 92, 107, 122, 137, 152, 167, 182. By numerical searching, we can find that the D-optimal design when $\rho = 0.5$ is: for 88% of the subjects, measurements are taken at 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180, and for 12% of the subjects, measurements are taken at 6, 21, 36, 51, 66, 81, 96, 111, 126, 141, 156, 171, 186. The summary of the efficiencies under different values of ρ is given in Table IX.

TABLE IX: D-Efficiency of an Equally Spaced Design for Bi-Exponential Model with Autocorrelated Covariance Structure

ρ	D_{opt}	D_{ref}	**Efficiency**$((D_{ref}/D_{opt})^{\frac{1}{4}})$
0.9	45498.0	16021.1	77.0%
0.8	12332.5	4323.7	77.0%
0.7	7559.5	2639.3	76.9%
0.6	6363.4	2213.5	76.8%
0.5	6339.1	2198.0	76.7%
0.4	7025.1	2429.3	76.8%
0.3	8406.6	2900.1	76.7%
0.2	10699.6	3683.7	76.6%
0.1	14382.1	4942.7	76.6%

The A-optimal design when $\rho = 0.5$ is chosen to be: 38% of the subjects will be measured at 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180, while 62% of the subjects will be measured at 4, 19, 34, 49, 64, 79, 94, 109, 124, 139, 154, 169, 184. The efficiencies under different values of ρ are given in Table X.

TABLE X: A-Efficiency of an Equally Spaced Design for Bi-Exponential Model with Autocorrelated Covariance Structure

ρ	A_{opt}	A_{ref}	**Efficiency**$((A_{opt}/A_{ref})^{\frac{1}{4}})$
0.9	3.38	3.74	97.5%
0.8	5.12	5.70	97.3%
0.7	6.17	6.82	97.5%
0.6	6.72	7.37	97.7%
0.5	6.91	7.55	97.8%
0.4	6.86	7.50	97.8%
0.3	6.64	7.32	97.6%
0.2	6.28	7.07	97.1%
0.1	5.83	6.78	96.3%

4.6.1.2 Intraclass Correlation Structure

By numerical searching, the D-optimal design when $\rho = 0.9$ is chosen to be: for 93% of the subjects, measurements are taken at 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180, and for 7% of the subjects, measurements are taken at 6, 21, 36, 51, 66, 81, 96, 111, 126, 141, 156, 171, 186. The summary of the efficiencies under different values of ρ is given in Table XI.

TABLE XI: D-Efficiency of an Equally Spaced Design for Bi-Exponential Model with Intraclass Correlation Structure

ρ	D_{opt}	D_{ref}	**Efficiency**$((D_{ref}/D_{opt})^{\frac{1}{4}})$
0.9	2459444	851730.75	76.7%
0.8	318226.94	110728.44	76.8%
0.7	103675.5	36129.06	76.8%
0.6	49504.38	17262.67	76.8%
0.5	29467.12	10278.15	76.9%
0.4	20480.23	7143.47	76.8%
0.3	16209.11	5652.10	76.8%
0.2	14663.21	5109.31	76.8%
0.1	15912	5534.31	76.8%

The A-optimal design when $\rho = 0.5$ is chosen to be: 64% of the subjects will be measured at 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, 180, 15% of the subjects will be measured at 4, 19, 34, 49, 64, 79, 94, 109, 124, 139, 154, 169, 184, and 21% of the subjects will be measured at 3, 18, 33, 48, 63, 78, 93, 108, 123, 138, 153, 168, 183. The efficiencies under different values of ρ are given in Table XII.

TABLE XII: A-Efficiency of an Equally Spaced Design for Bi-Exponential Model with Intraclass Correlation Structure

ρ	A_{opt}	A_{ref}	**Efficiency**$((A_{opt}/A_{ref})^{\frac{1}{4}})$
0.9	0.93	1.01	98.0%
0.8	1.43	1.60	97.2%
0.7	1.89	2.17	96.6%
0.6	2.34	2.74	96.1%
0.5	2.78	3.30	95.8%
0.4	3.23	3.86	95.6%
0.3	3.67	4.42	95.5%
0.2	4.12	4.98	95.4%
0.1	4.56	5.55	95.2%

4.6.2 Unequally Spaced Designs

We also consider the unequally spaced designs for bi-exponential model. The referenced design is 0, 2, 5, 10, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240.

4.6.2.1 Autocorrelated Covariance Structure

The D-optimal design when $\rho = 0.3$ is chosen to be: 67% of the subjects will be measured at 0, 6, 7, 8, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, and 33% of the subjects will be measured at 0, 7, 8, 9, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240. Efficiencies with respect to different values of ρ are given in Table XIII.

TABLE XIII: D-Efficiency of an Unequally Spaced Design for Bi-Exponential Model with Autocorrelated Covariance Structure

ρ	D_{opt}	D_{ref}	**Efficiency**$((D_{ref}/D_{opt})^{\frac{1}{4}})$
0.9	76424.3	34405.8	81.9%
0.8	22630.8	11252.7	84.0%
0.7	16013.6	8750.0	86.0%
0.6	15974.3	9501.4	87.8%
0.5	19030.3	12273.6	89.6%
0.4	25213.0	17613.0	91.4%
0.3	35906.7	27137.7	93.2%
0.2	54888.1	44101.5	94.7%
0.1	90312.2	74934.9	95.4%

The A-optimal design when $\rho = 0.5$ is chosen to be: 67% of the subjects are measured at 0, 3, 4, 19, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, 23% of the subjects are measured at 0, 3, 18, 19, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, and 10% of the subjects are measured at 0, 2, 3, 19, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240. Efficiencies with respect to different values of ρ are given in Table XIV.

TABLE XIV: A-Efficiency of an Unequally Spaced Design for Bi-Exponential Model with Autocorrelated Covariance Structure

ρ	A_{opt}	A_{ref}	**Efficiency**$((A_{opt}/A_{ref})^{\frac{1}{4}})$
0.9	2.31	3.01	93.6%
0.8	3.49	4.70	92.9%
0.7	4.26	5.63	93.3%
0.6	4.67	5.93	94.2%
0.5	4.79	5.81	95.3%
0.4	4.66	5.44	96.2%
0.3	4.38	4.95	97.0%
0.2	4.02	4.41	97.7%
0.1	3.60	3.87	98.2%

4.6.2.2 Intraclass Correlation Structure

The D-optimal design when $\rho = 0.5$ is chosen to be: all of the subjects are measured at 0, 1, 7, 8, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240. Efficiencies with respect to different values of ρ are given in Table XV.

TABLE XV: D-Efficiency of an Unequally Spaced Design for Bi-Exponential Model with Intraclass Correlation Structure

ρ	D_{opt}	D_{ref}	**Efficiency**$((D_{ref}/D_{opt})^{\frac{1}{4}})$
0.9	1875981	1583664	95.9%
0.8	2120688.6	1788842.3	95.8%
0.7	654415.87	551820.12	95.8%
0.6	303596.25	255949.59	95.8%
0.5	177601.21	149709.32	95.8%
0.4	122096.49	102912.39	95.8%
0.3	96017.28	80925.54	95.8%
0.2	86683.92	73055.02	95.8%
0.1	94566.34	79680.28	95.8%

The A-optimal design when $\rho = 0.5$ is chosen to be: 39% of the subjects are measured at 0, 3, 4, 19, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, 38% of the subjects are measured at 0, 1, 4, 5, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240, and 23% of the subjects are measured at 0, 1, 4, 19, 20, 30, 45, 60, 90, 120, 150, 180, 210, 240. Efficiencies with respect to different values of ρ are given in Table XVI.

TABLE XVI: A-Efficiency of an Unequally Spaced Design for Bi-Exponential Model with Intraclass Correlation Structure

ρ	A_{opt}	A_{ref}	**Efficiency**($(A_{opt}/A_{ref})^{\frac{1}{4}}$)
0.9	0.62	0.64	99.2%
0.8	0.97	1.01	99.0%
0.7	1.27	1.32	99.0%
0.6	1.55	1.62	98.9%
0.5	1.83	1.92	98.8%
0.4	2.10	2.21	98.7%
0.3	2.37	2.50	98.7%
0.2	2.64	2.78	98.7%
0.1	2.90	3.07	98.6%

CITED LITERATURE

1. Hedayat, A. S.: Study of optimality criteria in design of experiments. Statistics and Related Topics, Csorgo, M., Dawson, D. A., Rao. J. N. K and Salehh, A. K. Md. E., eds., North Holland Publishing Co., 39–56, 1981.

2. Fedorov, V. V.: Theory of Optimal Experiments. Academic Press, New York, 1972.

3. Silvey, S. D.: Optimal Design: An Introduction to the Theory for Parameter Estimation. Chapman and Hall, London, 1980.

4. Kiefer, J. and Wolfowitz, J.: Optimum designs in regression problems. The Annals of Mathematical Statistics, 30(2):271–294, 1959.

5. Kiefer, J.: General equivalence theory for optimum designs (approximate theory). The Annals of Statistics, 2(5):849–879, 1974.

6. Kiefer, J. and Wolfowitz, J.: The equivalence of two extremum problems. Canadian Journal of Mathematics, 12:363–366, 1960.

7. Kiefer, J.: Optimum designs in regression problems, II. The Annals of Mathematical Statistics, 32(1):298–325, 1961.

8. White, V. L.: An extension of the general equivalence theorem to nonlinear models. Biometrika, 60(2):345–348, 1973.

9. Elfving, G.: Optimum allocation in linear regression theory. The Annals of Mathematical Statistics, 23(2):255–262, 1952.

10. Hoel, G. P.: Efficiency problems in polynomial estimation. The Annals of Mathematical Statistics, 29(4):1134–1145, 1958.

11. De la Garza, A.: Spacing of information in polynomial regression. The Annals of Mathematical Statistics, 25(1):123–130, 1954.

12. Pukelsheim, F.: Optimal Design of Experiments. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006.

13. Khuri, I. A., Mukherjee, B., Sinha, K. B., and Ghosh, M.: Design issues for generalized linear models: a review. Statistical Science, 21(3):376–399, 2006.

14. Yang, M. and Stufken, J.: Support points of locally optimal designs for nonlinear models with two parameters. The Annals of Statistics, 37(1):518–541, 2009.

15. Yang, M.: On the de la garza phenomenon. The Annals of Statistics, 38(4):2499–2524, 2010.

16. Seber, G. A. F. and Wild, C. J.: Nonlinear Regression. Wiley, Hoboken, New Jersey, 2003.

17. Gabrielsson, J. and Weiner, D.: Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications. Swedish Pharmaceutical Press, Stockholm, 2007.

18. Bonate, L. P.: Pharmacokinetic-Pharmacodynamic Modeling and Simulation. Springer, New York, 2011.

19. Rosenbaum, E. S.: Basic Pharmacokinetics and Pharmacodynamics: An Integrated Textbook and Computer Simulations. Wiley, Hoboken, New Jersey, 2011.

20. Chernoff, H.: Locally optimal designs for estimating parameters. The Annals of Mathematical Statistics, 24(4):586–602, 1953.

21. Ford, I., Torsney, B., and Wu, C. F. J.: The use of a canonical form in the construction of locally optimal designs for non-linear problems. Journal of the Royal Statistical Society. Series B (Methodological), 54(2):569–583, 1992.

22. Hedayat, A. S., Zhong, J., and Nie, L.: Optimal and efficient designs for 2-parameter nonlinear models. Journal of Statistical Planning and Inference, 124(1):205–217, 2004.

23. Li, G. and Majumdar, D.: Some results on D-optimal designs for nonlinear models with applications. Biometrika, 96(2):487–493, 2009.

24. Fang, X. and Hedayat, A. S.: Locally D-optimal designs based on a class of composed models resulted from blending Emax and one-compartment models. The Annals of Statistics, 36(1):428–444, 2008.

25. Ogungbenro, K., Gueorguieva, I., Majid, O., Graham, G., and Aarons, L.: Optimal design for multiresponse pharmacokinetic–pharmacodynamic models–dealing with un-

balanced designs. Journal of Pharmacokinetics and Pharmacodynamics, 34(3):313–331, 2007.

26. Zocchi, S. S. and Atkinson, A. C.: Optimum experimental designs for multinomial logistic models. Biometrics, 55(2):437–444, 1999.

27. Gueorguieva, I., Aarons, L., Ogungbenro, K., Jorga, M. K., Rodgers, T., and Rowland, M.: Optimal design for multivariate response pharmacokinetic models. Journal of Pharmacokinetics and Pharmacodynamics, 33(2):97–124, 2006.

28. Li, G. and Majumdar, D.: D-optimal designs for logistic models with three and four parameters. Journal of Statistical Planning and Inference, 138(7):1950–1959, 2008.

29. Konstantinou, M., Biedermann, S., and Kimber, A.: Optimal designs for two-parameter nonlinear models with application to survival models. Statistica Sinica, Preprint.

30. Dette, H., Melas, B. V., and Wong, K. W.: Locally D-optimal designs for exponential regression models. Statistica Sinica, 16(3):789–803, 2006.

31. Abt, M., Liski, P. E., Mandal, K. N., and Sinha, K. B.: Optimal designs in growth curve models: part I correlated model for linear growth: optimal designs for slope parameter estimation and growth prediction. Journal of Statistical Planning and Inference, 64(1):141–150, 1997.

32. Abt, M., Liski, P. E., Mandal, K. N., and Sinha, K. B.: Optimal designs in growth curve models-II correlated model for quadratic growth: optimal designs for parameter estimation and growth prediction. Journal of Statistical Planning and Inference, 67(2):287–296, 1998.

33. Rowland, M., Thomson, D. P., Guichard, A., and Melmon, L. K.: Disposition kinetics of lidocaine in normal subjects. Annals of the New York Academy of Sciences, 179:383–398, 1971.

34. Wells, A. J., Siow, B., Lythgoe, F. M., and Thomas, L. D.: Measuring biexponential transverse relaxation of the ASL signal at 9.4 T to estimate arterial oxygen saturation and the time of exchange of labeled blood water into cortical brain tissue. Journal of Cerebral Blood Flow and Metabolism, 33(2):215–224, 2013.

VITA

Education

- Ph.D. in Statistics, University of Illinois at Chicago, 2013.
- M.S. in Statistics, University of Illinois at Chicago, 2009.
- B.S. in Mathematics and Applied Mathematics, Fudan University, 2007.

Working Experience

- Intern, Astellas Pharma Global Development, Deerfield, Illinois, 2009–2010.

Poster Prestation

- Topic on "Optimal Designs for Simultaneous Measurements in Compartment Models".

Membership

- American Statistical Association (ASA)
- Institute of Mathematical Statistics (IMS)